The Interleukin 1 Gene Family in Systemic Juvenile Idiopathic Arthritis

Carmel Joanna Winstanley Stock

A thesis submitted for the degree of Doctor of Philosophy to University College London

2011

Centre for Paediatric and Adolescent Rheumatology
Department of Immunology and Molecular Pathology
Division of Infection and Immunity
Windeyer Institute of Medical Sciences
University College London
46 Cleveland Street
London
W1T 4JF

Abstract

Patients with systemic Juvenile Idiopathic Arthritis (sJIA) have elevated serum levels of inflammatory cytokines. Treatment with interleukin-1 (IL-1) receptor antagonist (Anakinra) shows remarkable improvement in some sJIA patients. The hypothesis of this thesis is that genetic variations in IL-1 family genes contribute to disease pathogenesis.

To investigate this, a two-stage case-control association study of 20 candidate genes was performed. Selected tagging SNPs were tested for association in 130 sJIA patients and 146 controls in stage-1 of the study. SNPs at significantly different frequencies in the cohorts were genotyped in an additional 105 sJIA patients and 184 controls, and stratified meta-analysis of the two-stage data performed. Analysis was also performed with 4,671 controls from the Wellcome Trust Case Control Consortium (WTCCC).

No associations were found with caspase-1, cryopyrin, or IL-18. Significant disease associations were identified with SNPs in the ligand *IL1A*, the receptor antagonist *IL1RN*, and a two-SNP haplotype in the IL-18 antagonist *IL18BP*. Associations were also identified in the decoy receptor *IL1R2,* and the co-receptor *IL1RAP*, although these were not confirmed when re-analysed with the WTCCC controls. Transient transfection assays with haplotype constructs, performed by Dr Wen, showed that the *IL18BP* haplotype affected gene transcription levels *in vitro*. This effect was not however reproduced using PBMCs from healthy individuals. Allele specific binding to one of the haplotype SNPs was predicted *in silico*, but no evidence for this was seen in EMSA experiments. Further functional studies are required to corroborate involvement of this haplotype in disease.

In summary, this study has identified genetic associations for susceptibility to sJIA with a number of *IL1* family members. These results indicate that there may be aberrant control of IL-1 activity in patients with sJIA. Further work is required to determine how these associated SNPs affect IL-1 activity, and thereby the inflammatory response in sJIA.

Acknowledgements

I would like to thank Professor Woo for giving me the opportunity to work in her group, and her support and supervision throughout my project.

I would also like to thank Dr Mark Fife for his continued supervision, and Professor Cathryn Lewis for advice and guidance on the statistical analysis.

I am very grateful to all of the members of the 'Woo crew' who I have worked with, especially Dr Jane Samuel, Miss Emma Ogilvie, Dr Ebun Omyonmi, Dr Dongling Zheng, and Dr Aiqing Wen who performed the transcription study.

I would like to thank Professor David Isenberg and the Oliver Bird Rheumatism Program for giving me the opportunity to do this PhD.

I would also like to thank all of the patients and healthy individuals who kindly donated DNA and blood samples, and the members of BSPAR who contributed to the DNA repository, without which this project would not have been possible.

Thank you to my parents for their support and understanding.

A huge thanks to my fellow Olli Birdies, who were always there to share problems and successes, as well as a few drinks.

I would also like to thank my partner Benit, whose support helped me to get to the end.

Table of Contents

Abstract 2

Acknowledgements 3

Table of Contents 4

Table of Tables 10

Table of Figures 12

Abbreviations 14

Glossary 17

1. Diseases 17
2. Disease features 19
3. Measures of disease activity 21
4. Medications 22
5. Cell lines 23
6. Somatic cells 23
7. Immunomodulatory molecules 25
8. Nomenclature of the genes investigated in this study 28

1. Introduction 30

1.1. Juvenile Idiopathic Arthritis 30
1.2. Systemic Juvenile Idiopathic Arthritis 30
1.2.1. Complications 31
1.2.2. Diagnosis 34
1.2.3. Treatment 34
1.3. Inflammation and Cytokines 40
1.4. Altered Cytokine Profile in sJIA 42
1.5. Genetic Associations 45
1.5.1. Major Histo-compatibility Complex 46
1.5.2. Non-MHC Genes 46
1.6. sJIA is distinct from other JIA subtypes 49
1.7. Interleukin 1 Gene Family 50
1.7.1. Activity 50
1.7.2. IL-1 Ligands 50
1.7.2.1. IL-1α 51
1.7.2.2. IL-1β 51
1.7.3. IL-1Ra 52
1.7.4. IL-1 Receptor Complex 52
1.7.4.1. IL-1 R1 52
1.7.4.2. IL-1R2 54
1.7.4.3. IL-1RAcP 54
1.7.5. IL-18 56
1.7.5.1. IL-18 Binding Protein 56
1.7.5.2. IL-18 Receptor Complex 57
1.7.6. Other IL-1 Family Members 58
1.7.6.1. Ligands 58
1.7.6.2. Receptors 59

1.7.7. IL-1 associated proteins 60
1.7.7.1. Caspase-1 61
1.7.7.2. Cryopyrin 61
1.7.8. IL-1 Family Gene Organisation 62
1.8. IL-1 in disease 62
1.9. IL-1 Gene Family Disease Associations 65
1.10. IL-1 in sJIA 67
1.11. Hypothesis of the project 68
1.12. Aims of the project 68
2. Materials 70
2.1. Patient and Control Samples 70
2.1.1. Healthy controls 70
2.1.2. Patients 70
2.2. Laboratory reagents 71
2.2.1. Chemicals 71
2.2.2. Buffers 73
2.2.3. Gels 74
2.2.4. DNA extraction 75
2.2.5. PCR 75
2.2.6. DNA purification 75
2.2.7. Cell culture 76
2.2.8. PBMC isolation 76
2.2.9. Bacterial culture 76
2.2.10. Cloning 77
2.2.11. RNA quantification 79
2.2.12. ELISA 79
2.2.13. EMSA 80
3. Methods 82
3.1. Association Study 82
3.1.1. Introduction 82
3.1.1.1. Tagging SNPs 82
3.1.1.2. Two-stage study design 85
3.1.2. Stage-1 86
3.1.2.1. tSNP Selection 86
3.1.2.2. Genotyping 97
3.1.2.2.1. Genotype calling 97
3.1.2.2.2. Golden Gate genotyping pilot study 101
3.1.2.3. Data Quality Controls 101
3.1.2.4. Association Analysis 104
3.1.3. Stage-2 105
3.1.3.1. Genotyping 105
3.1.3.2. Data Quality Controls 105
3.1.3.3. Association Analysis 106
3.1.4. WTCCC control cohort analysis 107
3.1.5. Identification of additional tagged SNPs 108
3.1.6. Validation of the tagging SNP selection results 108
3.2. Sequence feature prediction 109
3.2.1. Comparative Genomics 110
3.2.2. Transcription Factor Binding prediction 113

3.3. General experimental protocols 113
3.3.1. DNA extraction 113
3.3.1.1. From blood 113
3.3.1.2. From saliva 114
3.3.1.3. Nucleic acid Quantification 114
3.3.2. Polymerase Chain Reaction (PCR) 114
3.3.2.1. Primer Design 114
3.3.2.2. PCR reaction 115
3.3.2.3. PCR cycling 115
3.3.2.4. Optimisation 116
3.3.2.5. PCR product visualisation 116
3.3.2.6. PCR Product Purification 116
3.3.2.7. PCR Product Gel Extraction 116
3.3.3. Genotyping of Healthy Controls 118
3.3.4. Cloning 118
3.3.4.1. Generation of insert 118
3.3.4.2. PCR Product Ligation into Plasmid 118
3.3.4.3. Plasmid Transformation into Competent Cells 119
3.3.4.4. Plasmid amplification 119
3.3.4.5. Colony PCR 119
3.3.4.6. Small scale plasmid preparation 'Miniprep' 119
3.3.4.7. Sequencing 120
3.3.5. Cell culture 120
3.3.5.1. Peripheral Blood Mononuclear Cells (PBMCs) 120
3.3.5.2. THP-1 Cells 120
3.3.5.3. Cell counting 120
3.3.6. PBMC isolation 121
3.3.7. Enzyme Linked Immunosorbant Assay (ELISA) 121
3.3.7.1. Sandwich ELISA 123
3.3.7.2. ELISA Analysis 123
3.3.7.2.1. Standard Curve 123
3.3.7.2.2. Statistical analysis 123
3.3.8. RNA based protocols 124
3.3.8.1. Total RNA extraction 124
3.3.8.2. Reverse Transcription 124
3.3.8.3. Quantitative RT-PCR 125
3.3.8.3.1. Analysis 126
3.3.9. Electrophoretic Mobility Shift Assay (EMSA) 126
3.3.9.1. Nuclear extraction 126
3.3.9.2. Oligonucleotide annealing 128
3.3.9.3. Probe amplification 128
3.3.9.4. Probe labelling 128
3.3.9.5. Protein binding 129
3.3.9.6. Visualising DNA-protein complexes 130
3.4. IL18BP candidate region 130
3.4.1. Investigation of expression levels according to haplotype 130
3.4.1.1. Haplotype determination 130
3.4.1.1.1. Haplotype determination of uncertain phase 130
3.4.1.2. PBMC stimulation 131
3.4.1.3. Protein expression levels 133

3.4.1.3.1. Optimisation of cell concentration 133
3.4.1.3.2. Expression comparison between haplotypes 134
3.4.1.4. RNA expression levels 134
3.4.1.4.1. Validation of housekeeping genes 134
3.4.1.4.2. Expression comparison between haplotypes 135
3.4.2. Transcription factor binding 135
3.4.2.1. Probe design 135
3.4.2.2. Short probe EMSAs 135
3.4.2.3. Long probe EMSAs 137
3.5. Association analysis of the susceptibility alleles according to response to IL-1 blockade 137
4. Results 140
4.1. Association Study Quality Control 140
4.1.1. Power Calculation 140
4.1.2. Golden Gate genotyping pilot study 140
4.1.3. Stage-1 140
4.1.4. Stage-2 143
4.1.5. Validation of the tagging SNP selection results 149
4.1.5.3. HapMap3 compared to WTCCC2 158
4.2. IL18BP candidate region 163
4.2.1. Association Study 163
4.2.1.1. Tagging SNP selection 163
4.2.1.2. Stage-1 166
4.2.1.3. Stage-2 170
4.2.1.3.1. Meta-analysis 170
4.2.1.4. WTCCC control cohort analysis 170
4.2.1.5. Associated SNPs 173
4.2.1.5.1. Additional captured SNPs 173
4.2.1.5.2. Associated SNP positions 177
4.2.1.6. Summary 177
4.2.1.7. Discussion 180
4.2.2. Comparative Genomics 182
4.2.2.1. Discussion 182
4.2.3. Investigation of expression levels according to haplotype 187
4.2.3.1. Haplotype determination 187
4.2.3.1.1. Haplotype determination of uncertain phase 187
4.2.3.2. Protein expression levels 187
4.2.3.2.1. Optimisation of cell concentration 187
4.2.3.2.2. IL-18BP gene expression comparison according to haplotype 190
4.2.3.3. RNA expression levels 193
4.2.3.3.1. Validation of house keeping genes 193
4.2.3.3.2. IL-18BP gene expression comparison according to haplotype 197
4.2.3.4. Summary 201
4.2.3.5. Discussion 201
4.2.4. Transcription Factor Binding 203
4.2.4.1. Binding Prediction 203
4.2.4.2. EMSA 203
4.2.4.2.1. Short probe EMSAs 203
4.2.4.2.2. Long probe EMSAs 206
4.2.4.3. Summary 206

4.2.4.4. Discussion 206
4.3. IL1RAP candidate region 211
4.3.1. Association Study 211
4.3.1.1. tSNP selection 211
4.3.1.2. Stage-1 214
4.3.1.3. Stage- 2 231
4.3.1.3.1. Stratified analysis 235
4.3.1.4. WTCCC control cohort analysis 237
4.3.1.4. Associated SNPs 237
4.3.1.4.1. Additional captured SNPs 237
4.3.1.4.2. Associated SNP positions 237
4.3.1.5. Summary 240
4.3.1.6. Discussion 240
4.4. IL1 ligand cluster candidate region 244
4.4.1. Association study 244
4.4.1.1. tSNP selection 244
4.4.1.2. Stage-1 247
4.4.1.3. Stage-2 262
4.4.1.3.1. Stratified analysis 262
4.4.1.4. WTCCC control cohort analysis 268
4.4.1.5. Associated SNPs 270
4.4.1.5.1. Additional captured SNPs 270
4.4.1.5.2. Associated SNP positions 272
4.4.1.6. Summary 272
4.4.1.7. Discussion 275
4.5. IL1 receptor cluster candidate region 279
4.5.1. Association Study 279
4.5.1.1. Tagging SNP selection 279
4.5.1.2. Stage-1 281
4.5.1.3.1. Stratified analysis 302
4.5.1.4. WTCCC control cohort analysis 302
4.5.1.5. Associated SNPs 302
4.5.1.5.1. Additional captured SNPs 306
4.5.1.5.2. Associated SNP positions 306
4.5.1.6. Summary 306
4.5.1.7. Discussion 309
4.6. IL18 candidate region 312
4.6.1. Association Study 312
4.6.1.1. tSNP selection 312
4.6.1.2. Stage-1 315
4.6.2. Summary 315
4.6.3. Discussion 315
4.7. CASP1 candidate region 320
4.7.1. Association Study 320
4.7.1.1. Tagging SNP selection 320
4.7.1.2. Stage-1 322
4.7.2. Summary 325
4.7.3. Discussion 325
4.8. NALP3 candidate region 331
4.8.1. Association Study 331

4.8.1.1 Tagging SNP selection 331
4.8.1.1 Stage-1 334
4.8.2. Summary 334
4.8.3. Discussion 334
4.9. Association analysis of the susceptibility alleles according to response to IL-1 blockade 340
4.9.1. Discussion 341
5. Discussion 346
6. Future Work 352
Appendix 1. Golden Gate genotyping platform 398
A. Methodology 398
B. Genotyping Protocol 400
Appendix 2. Genotyping assay results of Stage-1 SNPs 402
Appendix 3. IL-18BP haplotype transient transfection assays 410
A. Methods 410
B. Results 412
Appendix 4. Publications ..... **Error! Bookmark not defined.**

Table of Tables

Table 1.1 Features of the JIA subtypes according to the ILAR classification 32
Table 1.2 Differential diagnosis of sJIA 35
Table 3.1 Flanking sequences scanned for PGA variation discovery..... 87
Table 3.2 CEPH samples in the population genotyping databases..... 92
Table 3.3 BSPAR repository patient samples used in GoldenGate Pilot Study..... 102
Table 4.1 BSPAR repository patient GoldenGate Pilot Study 141
Table 4.2 Genotyped SNPs classed as failed assays..... 142
Table 4.3 Genotyped SNPs significantly deviating from Hardy-Weinberg Equilibrium 147
Table 4.4 Genotyping discrepancies between the two genotyping platforms..... 148
Table 4.5 Comparison of LD (r2) values from different population genotyping data sets... 157
Table 4.6 Comparison of SNP frequencies in different population genotyping data sets 160
Table 4.7 *IL18BP* candidate region tSNPs..... 167
Table 4.8 *IL18BP* candidate region stage-1 analysis results 168
Table 4.9 *IL18BP* candidate region stage-2 analysis results 171
Table 4.10 *IL18BP* candidate region pooled cohorts analysis results..... 172
Table 4.11 *IL18BP* candidate region analysis results with WTCCC2 controls..... 174
Table 4.12 Additional SNPs tagged by *IL18BP* SNP1 176
Table 4.13 *IL18BP* SNP1 and 2 genotypes of WC healthy controls 188
Table 4.14 IL-18BP protein expression levels from IFNγ stimulated PBMCs..... 192
Table 4.15 *IL18BP* RNA expression levels from IFNγ stimulated PBMCs 199
Table 4.16 *IL1RAP* candidate region tSNPs..... 216
Table 4.17 Significant *IL1RAP* candidate region stage-1 analysis results..... 217
Table 4.18 Non-significant *IL1RAP* candidate region stage-1 analysis results..... 220
Table 4.19 Conditional analysis of associated *IL1RAP* region SNPs within LD clusters 230
Table 4.20 Conditional analysis of associated *IL1RAP* region SNPs between LD clusters . 233
Table 4.21 *IL1RAP* candidate region stage-2 analysis results..... 234
Table 4.22 *IL1RAP* candidate region stratified analysis results 236
Table 4.23 *IL1RAP* candidate region analysis results with WTCCC2 controls 238
Table 4.24 Additional SNPs tagged by *IL1RAP* SNP14..... 239
Table 4.25 *IL1* ligand cluster candidate region tSNPs..... 251
Table 4.26 *IL1* ligand cluster additional tagging of analysis excluded SNPs 252
Table 4.27 Significant *IL1* ligand cluster candidate region stage-1 analysis results 254
Table 4.28 Non-significant *IL1* ligand cluster candidate region stage-1 analysis results..... 258
Table 4.29 Conditional analysis of associated *IL1* ligand region SNPs within LD clusters. 264
Table 4.30 Conditional analysis of associated *IL1* ligand region SNPs between LD clusters 265
Table 4.31 *IL1* ligand candidate region stage-2 analysis results 266
Table 4.32 *IL1* ligand cluster candidate region stratified analysis results 267
Table 4.33 *IL1* ligand candidate region analysis results with WTCCC2 controls..... 269
Table 4.34 Conditional analysis of associated *IL1* ligand region SNPs with WTCCC2 controls 271
Table 4.35 Additional SNPs tagged by *IL1* ligand cluster SNPs 273
Table 4.36 *IL1* receptor cluster candidate region tSNPs..... 287
Table 4.37 *IL1* receptor cluster additional tagging of analysis excluded SNPs 288
Table 4.38 Significant *IL1* receptor cluster candidate region stage-1 analysis results 289
Table 4.39 Non-significant *IL1* receptor cluster candidate region stage-1 analysis results.. 294
Table 4.40 Conditional analysis of associated *IL1* receptor region SNPs 300

Table 4.41 Incompletely captured *IL1* receptor cluster candidate region SNPs 301
Table 4.42 *IL1* receptor candidate region stage-2 analysis results 303
Table 4.43 *IL1* receptor cluster candidate region stratified analysis results 304
Table 4.44 *IL1* receptor candidate region analysis results with WTCCC2 controls..... 305
Table 4.45 Additional SNPs tagged by *IL1* receptor cluster SNP8 307
Table 4.46 *IL18* candidate region tSNPs..... 316
Table 4.47 *IL18* candidate region stage-1 results 317
Table 4.48 *CASP1* candidate region tSNPs..... 326
Table 4.49 *CASP1* additional tagging of analysis excluded SNPs 327
Table 4.50 *CASP1* candidate region stage-1 results 328
Table 4.51 *NALP3* candidate region tSNPs..... 335
Table 4.52 *NALP3* candidate region stage-1 results 336
Table 4.53 IL-1 blockade treated patient information 342
Table 4.54 Allele frequencies according to IL-1 blockade response..... 343

Table of Figures

Figure 1.1 Treatment strategy for sJIA 36
Figure 1.2 IL-1 ligand-receptor interactions 53
Figure 1.3 IL-1 signal transduction 55
Figure 1.4 Organisation of the two IL-1 gene family clusters 63
Figure 2.1 Age of disease onset of patients 72
Figure 2.2 pGL3-Basic vector map 78
Figure 3.1 SeattleSNPs gene regions scanned for variation discovery 90
Figure 3.2 Graphical representation of LD patterns 94
Figure 3.3 Finalisation of tSNP set 96
Figure 3.4 Golden Gate Assay Work-flow 98
Figure 3.5 Genoplot of GoldenGate genotyping results 100
Figure 3.6 PCR condition optimisation 117
Figure 3.7 ELISA principle 122
Figure 3.8 Electroporetic Mobility Shift Assay principle 127
Figure 3.9 IL18BP cloning region for haplotype determination 132
Figure 3.10 *IL18BP* SNP1 EMSA probes 136
Figure 4.1 Genoplots of failed Stage-1 SNP assays 145
Figure 4.2 *IL18BP* candidate region 164
Figure 4.3 LD plot of the *IL18BP* region SNPs 165
Figure 4.4 Genoplots of the significant *IL18BP* SNPs 169
Figure 4.5 *IL18BP* flanking genes 178
Figure 4.6 Positions of the associated SNPs in the *IL18BP* candidate region 179
Figure 4.7 VISTA conservation homology plot of *IL18BP* 184
Figure 4.8 Sequence alignment of the *IL18BP* SNP nucleotides 186
Figure 4.9 Haplotype determination sequencing results 189
Figure 4.10 IL-18BP expression levels from stimulated PBMCs at different cell concentrations 191
Figure 4.11 Stimulated PBMC IL-18BP protein expression levels according to haplotype. 194
Figure 4.12 Melting curves of housekeeping qRT-PCR assay primers 195
Figure 4.13 Housekeeping gene C_T variation 196
Figure 4.14 Melting curve of IL-18BP qRT-PCR assay primers 198
Figure 4.15 Stimulated PBMC *IL18BP* RNA expression levels according to haplotype 200
Figure 4.16 *IL18BP* SNP1 transcription factor binding prediction 204
Figure 4.17 *IL18BP* short probe EMSA 205
Figure 4.18 *IL18BP* long probe poly dI-dC titration EMSA 209
Figure 4.19 *IL18BP* long probe EMSA 210
Figure 4.20 *IL1RAP* candidate region 212
Figure 4.21 LD plot of the SNPs in the *IL1RAP* gene region 213
Figure 4.22 Genoplots of the significant *IL1RAP* SNPs 226
Figure 4.23 D′ relationship of associated *IL1RAP* region SNPs 228
Figure 4.24 Positions of the associated SNPs in the *IL1RAP* candidate region 241
Figure 4.25 *IL1* ligand cluster candidate region 245
Figure 4.26 LD plot of the SNPs in the *IL1* ligand gene cluster region 246
Figure 4.27 Genoplots of the significant *IL1* ligand cluster region SNPs 260
Figure 4.28 D′ relationship of associated *IL1* ligand cluster region SNPs 263
Figure 4.29 Positions of the associated SNPs in the *IL1* ligand cluster candidate region 274
Figure 4.30 *IL1* receptor cluster candidate region 280

Figure 4.31 LD plot of the SNPs in the *IL1* receptor gene cluster region..... 282
Figure 4.32 Genoplots of the significant *IL1* receptor cluster SNPs 297
Figure 4.33 D′ relationship of associated *IL1* receptor cluster region SNPs..... 299
Figure 4.34 Positions of the associated SNPs in the *IL1* receptor cluster candidate region . 308
Figure 4.35 *IL18* candidate region..... 313
Figure 4.36 LD pattern over the *IL18* candidate region..... 314
Figure 4.37 *CASP1* candidate region 321
Figure 4.38 LD pattern over the *CASP1* candidate region..... 323
Figure 4.39 *NALP3* candidate region 332
Figure 4.40 LD plot of the SNPs in the *NALP3* gene region 333

Abbreviations

AAO - Age At Onset
Ab – Antibody
ACR - American College of Rheumatology
ANA - Anti-nuclear Antibodies
ANCA - Anti-neutrophil Cytoplasmic Antibodies
ANOVA - Analysis Of Variance
AOSD - Adult Onset Still's Disease
APS – Anti-phosphlopid Syndrome
AS - Ankylosing Spondylitis
ASCT - Autologous Stem Cell Transplant
ASO - Allele Specific Oligonucleotide
BD - Breslow-Day
BLAST - Basic Local Alignment Search Tool
BLAT - BLAST Like Alignment Tool
bp - Base Pair
BSA - Bovine Serum Albumin
BSPAR - British Society for Paediatric and Adolescent Rheumatology
CAPS - Cryopryin-associated Periodic Syndromes
CD – Crohn's Disease
cDNA – complementary Deoxyribonucleic Acid
CEPH – Centre d'Etude du Polymorphisme Humain
CEU - CEPH Utah residents
CI – Confidence Interval
CIA – Collagen Induced Arthritis
CIAP – Calf Intestinal Alkaline Phosphatase
CINCA - Chronic Infantile Neurological, Cutaneous, and Articular syndrome (same as NOMID)
CMH - Cochran-Mantel-Haenszel
CNS - Conserved Non-coding Sequence
cPCR – colony Polymerase Chain Reaction
CRP - C-Reactive Protein
df - Degrees of Freedom
DMSO - Dimethyl Sulfoxide
DNA - Deoxyribonucleic Acid
dNTP -Deoxyribonucleotide Triphosphate
DTT – Diothiothreitol
EDTA - Ethylenediaminetetraacetic Acid
ELISA - Enzyme-Linked Immunosorbent Assay
EM - Expectation-Maximisation
EMSA - Electro-Mobility Shift Assay
ESR - Erythrocyte Sedimentation Rate
FCAS - Familial Cold Autoinflamatory Syndrome

FCS - Fetal Calf Serum
FMF – Familial Mediterranian Fever
GC – Genotype Call
GOS - Great Ormond Street
GOSH – Great Ormond Street Hospital
GWAS - Genome-Wide Association Study
HapMap – International Human Haplotype Map
HLA - Human Leukocyte Antigen
HLH - Hemophagocytic lymphohistiocytosis
HRP - Horseradish Peroxidase
HWE - Hardy-Weinberg Equilibrium
i.v. – Intravenous
IFN – Interferon
IgG - Immunoglobulin G
IL – Interleukin
ILAR - International League of Associations for Rheumatology
JIA - Juvenile Idiopathic Arthritis
kb – Kilobase
KO – Knock Out
LB – Lysogeny Broth
LD - Linkage Disequilibrium
LOD – Logarithm of odds
LPS - Lipopolysaccaride
LSO - Locus Specific Oligonucleotide
MAF - Minor Allele Frequency
MAS - Macrophage Activation Syndrome
Mb – Megabase
MHC - Major Histocompatability Complex
mRNA - messenger Ribonucleic Acid
MTX – Methotrexate
MWS - Muckle-Wells Syndrome
NCBI - National Center for Biotechnology Information
NE - Nuclear Extract
NK - Natural Killer
NOMID - Neonatal Onset Multisystem Inflammatory Disease (same as CINCA)
NSAID - Non-Steroidal Anti-Inflammatory Drug
NTC – No Template Control
nts – Nucleotides
OA – Osteoarthritis
OD - Optical Density
OR - Odds Ratio
PBMC - Peripheral Blood Mononuclear Cells
PBS - Phosphate Buffered Saline
PBST - Phosphate Buffered Saline Tween-20

PCR - Polymerase Chain Reaction
PGA - Programmes for Genomic Applications
PMA - Phorbol 12-Ayristate 13-Acetate
qRT-PCR – quantitative Reverse Transcription Polymerase Chain Reaction
RA - Rheumatoid Arthritis
RE - Restriction Enzyme
RF - Rheumatoid Factor
RFLP - Restriction Fragment Length Polymorphism
RNA - Ribonucleic Acid
rs – RefSNP
RT – Reverse Transcriptase
RT-PCR - Reverse Transcription Polymerase Chain Reaction
SAA - Serum Amyloid A
SD – Standard deviation
sJIA - Systemic Juvenile Idiopathic Arthritis
SLE - Systemic Lupus Erythematosus
SNP - Single Nucleotide Polymorphism
TAE - Tris-acetate-EDTA
TBE - Tris-borate-EDTA
TDT – Transmission Disequilibrium Test
TE - Tris Ethylenediaminetetraacetic Acid
TF - Transcription Factor
TGF - Transforming Growth Factor
Tm – Melting Termperature
TMB – Tetramethylbenzidine
TMED - NNN'N'-Tetramethylethylenediamine
TNF - Tumor Necrosis Factor
tSNPs – Tagging Single Nucleotide Polymorphism
TSS – Transcription Start Site
UC – Ulcerative Colitis
UCSC - University of California Santa Cruz
UTR - Untranslated Region
UV – Ultraviolet
VNTR - Variable Number Tandem Repeat
WC - Windeyer Control (healthy controls working in the division)
WTCCC - Wellcome Trust Case Control Consortium

Glossary

1. Diseases

Adult-onset Still's disease (AOSD): An autoinflammatory disease characterised by recurrent fevers, joint pain, rash, and inflammation of multiple organs. It is the adult onset form of sJIA.

Alopecia areata: A disease resulting in sudden hair loss, usually on the scalp but can be any hair-producing part of the body. Ranges from patches of hair loss to total body hair loss. Thought to be the result of an autoimmune reaction against the hail follicles.

Amyloidosis: Aggregates of insoluble fibrous proteins (amyloid) deposited in an organ or tissue.

Ankylosing Spondylitis (AS): A chronic autoimmune arthritis of the spine and sacroiliac joints. 40% of patients also have inflammation of the eye.

Arthritis: A group of diseases involving joint damage.

Chronic Infantile Neurologic Cutaneous and Arthropathy Syndrome (CINCA): An autoinflammatory disease present from birth. It is characterised by a skin rash and articular and central nervous system involvement.

Coeliac disease: An autoimmune disease of the small intestine caused by a reaction to gluten.

Colitis: Inflammation of the colon, can be due to infection, autoimmune reactions, or a combination. Severe or chronic inflammation can cause damage to the lining resulting in ulceration.

Collagen Induced Arthritis: An animal model of rheumatoid arthritis, induced by immunisation with type II collagen.

Coronary artery disease: Thinning of the coronary arteries due to the formation of plaque. This reduces the supply of oxygenated blood to the heart increasing the risk of angina and heart attacks.

Crohn's disease: An inflammatory disease affecting any part of the gastrointestinal tract, most commonly the ileum.

Enthesitis related arthritis: A subtype of JIA in which patients have arthritis and inflammation at the point where a tendon or ligament inserts into bone.

Familial cold auto-inflammatory syndrome (FCAS): Alternative name for Familial cold urticaria.

Familial cold urticaria: An inflammatory disease in which symptoms develop within 1-2 hours of exposure to cold. Symptoms include conjunctivitis, chills, joint pain, fever, and rash.

Familial Mediterranean fever (FMF): An autoinflammatory disease characterised by recurring fever, skin rash, joint pain, and inflammation of organs in the abdomen and chest.

Glaucoma: An eye disorder in which the optic nerve is damaged.

Graves' disease: An autoimmune disease in which autoantibodies to thyroid-stimulating hormone receptor are produced, resulting in hyperthyroidism.

Hemophagocytic lymphohistiocytosis: A haematologic disorder characterised by fever, splenomegaly, and jaundice.

Inflammatory bowel disease: A group of inflammatory diseases of the colon and small intestine, includes Crohn's disease and ulcerative colitis.

Juvenile Idiopathic Arthritis: Arthritis of no known cause which begins before the age of sixteen and persists for at least six weeks. Divided into seven distinct subtypes according to the International League of Associations for Rheumatology criteria.

Kawasaki disease: An autoimmune necrotising vasculitis of small and medium sized blood vessels. Affected areas include the skin, mucous membranes, lymph nodes, extremities, and eyes. Patients have fevers often lasting for more than five days.

Macrophage Activation Syndrome (MAS): A potentially life-threatening complication of some childhood rheumatic diseases. Patients have uncontrolled activation and proliferation of macrophages and T-lymphocytes, and increased circulating cytokines causing multisystem inflammation and dysfunction. May be preceded by an infection.

Muckle-Wells syndrome (MWS): An autoinflammatory disease characterised by periodic episodes of fever, skin rash and joint pain. Progressive hearing loss, kidney damage, and amyloidosis can also occur.

Multiple Sclerosis: An inflammatory disease which results in demyelination of brain and spinal cord axons, leading to disability.

Neuroblastoma: A solid nuroendocrine tumor which forms on nervous tissue, usually beginning in the adrenal glands.

Oligoarthritis: A subtype of JIA in which no more than four joints are affected in the first six months. If more joints become effected after six months it is defined as extended oligoarthritis, and as persistent oligoarthritis if no more joints become affected throughout the course of the disease.

Osteoarthritis (OA): A degenerative joint disease in which the cartilage surface of joints is worn down leading to the development of bony growths, inflammation in the surrounding tissues and pain in moving the joint.

Osteoporosis: Reduced bone mass caused by an imbalance between bone resorption and formation. Increases the risk of fractures.

Periodontitis: An inflammatory disease affecting the tissues supporting the teeth. Caused by the presence of bacteria on the tooth surface and the immune response against them. Progressive loss of the surrounding bone can lead to tooth loss.

Polyarthritis: Encompasses two subtypes of JIA in which more than four joints are affected within the first six months of the disease. Patients are divided into rheumatoid factor (RF) positive and negative subtypes.

Psoratic arthritis: An inflammatory joint disease associated with psoriasis.

Psoriasis: An autoimmune chronic skin condition causing red scaly patches on the skin due to areas of inflammation and excess skin production.

Rheumatoid arthritis (RA): An autoimmune disease mainly affecting the synovial membrane in joints causing pain, swelling, and joint destruction and deformity. Majority of patients have rheumatoid factor autoantibody.

Sarcoidosis: Granulomas of clumped together inflammatory cells form in the body affecting organ function. Most often affects the lungs but can also occur in a number of other organs.

Scleroderma: An autoimmune disease characterised by fibrosis of the skin (limited scleroderma), or the skin and internal organs (diffuse scleroderma).

Systemic JIA: A subtype of JIA in which patients have arthritis associated with marked systemic features and a daily spiking fever.

Systemic Lupus Erythematosus (SLE): An autoimmune connective tissue disease affecting many organs including the kidneys, lungs, heart, liver, and nervous system. Patients may also have a characteristic butterfly rash on their face. Patients have apoptosis abnormalities, and circulating anti-nuclear (ANA) and double stranded DNA antibodies.

Systemic vasculitis: Blood vessel inflammation affecting many different organ systems.

Type 1 diabetes mellitus (T1DM): An insulin deficiency caused by autoimmune destruction of the beta-cells of the pancreas resulting in an inability to adsorb glucose.

Ulcerative colitis (UC): An inflammatory disease of the colon.

Undifferentiated JIA: A subtype of JIA in which the patient's symptoms either fulfils the criteria of more than one or none of the other subtypes.

Vitiligo: A disorder in which melanocytes do not function correctly, resuting in loss of skin pigmentation. May have an autoimmune component.

2. Disease features

Anaemia: Deficiency of red blood cells, haemoglobin, or total blood volume.

Coagulopathy: Disorder of blood coagulation.

Conjunctivitis: Inflammation of the conjunctiva, the outermost layer of the eye and inner surface of the eyelids.

Cytopenia: Reduction or deficiency in the number of blood cells.

Dactylitis: Inflammation of a finger or toe.

Encephalopathy: Disorder or disease of the brain.

Enthesitis: Inflammation at the point where a tendon or ligament inserts into bone.

Erythema nodosum: Inflammation of fat cells causing nodules on the skin, usually the shins.

Evanescent or erythematous rash: A transient or red rash.

Hepatomegaly: Abnormal enlargement of the liver.

Hilar lymphadenopathy: Inflammation of the thoracic lymph nodes.

Iridocyclitis: Inflammation of the iris and ciliary body behind the eye.

Leukocytosis: Abnormal increase in the number of white blood cells.

Lymphadenopathy: Abnormal swelling or enlargement of the lymph nodes.

Macular rash: A rash of flat, red spots.

Malar rash: A 'butterfly-like' facial rash.

Malaise: A feeling of being unwell.

Meningitis: Inflammation of the membranes covering the brain and spinal cord, normally caused by infection.

Mucocutaneous symptoms: Symptoms affecting the skin or mucus membranes.

Myalgia: Muscle pain.

Nephritis: Inflammation of the kidneys.

Neutropenia: An abnormally low number of neutrophils.

Onycholysis: Spontaneous separation of the nail from the nail bed.

Oophoritis: Inflammation of the ovaries.

Panniculitis: Inflammation of subcutaneous adipose tissue.

Pericarditis: Inflammation of the membrane enclosing the heart.

Pneumonitis: Inflammation of lung tissue.

Proteinuria: An excess of serum proteins in the urine.

Pruritus: Itching.

Pupura: Red or purple discolourations caused by bleeding underneath the skin.

Quotidian fever: A fever which recurs daily.

Sacroillitis: Inflammation of the sacroiliac joint.

Sepsis: An infection which has spread through the blood causing symptoms throughput the body.

Serositis: Inflammation of serous membranes which line body cavities and enclose organs.

Splenomegaly: Abnormal enlargement of the spleen.

Thrombocytosis: Overproduction of blood platelets.

Uticarial rash: Raised, itchy rash.

Uveitis: Inflammation of the uveal layer of the eye.

3. Measures of disease activity

Active arthritis: Defined as presence of swelling or limitation of motion with accompanying heat, pain, or tenderness in the joint.

American College of Rheumatology (ACR) score: This reflects the change in disease activity between time points. It is calculated using: the physician global assessment of disease activity, parent/patient assessment of overall well-being, functional ability, number of joints with active arthritis, number of joints with a limited range of motion, ESR, and CRP. ACR scores used are ACR30, 70, and 90 for which are required at least 30, 70, or 90 percent, respectively improvement from baseline in any three of the variables, with worsening by more than 30% in no more than one of the variables.

Anti-neutrophil cytoplasmic antibodies (ANCA): Auto-antibodies directed against antigens in the cytoplasm of neutrophil granulocytes and monocytes.

Anti-nuclear antibodies (ANA): Auto-antibodies directed against cell nucleus contents.

C-reactive Protein (CRP): Non-specific marker of inflammation. CRP is an acute-phase protein produced in the liver in response to adipocyte released factors. It is involved in complement binding and enhances macrophage phagocytosis.

Double stranded DNA antibodies (dsDNA): Auto-antibodies directed against the double stranded DNA normally only found within nuclei and not in circulation.

Erythrocyte count: The number of red blood cells per ml^3 of blood.

Erythrocyte sedimentation rate (ESR): Non-specific measure of inflammation. ESR is the distance in mm that red blood cells precipitate in one hour. When inflammation is present there is an increase in acute-phase proteins including fibrinogen. High fibrinogen levels cause red blood cells to stick together and form stacks called rouleaux which settle faster than separate cells. An increased ESR indicates the presence of inflammation.

Haemoglobin level: The level of the iron-containing metalloprotein haemoglobin found in red blood cells.

Parent/patient assessment of overall well-being: The assessor marks on a 10 centimetre visual analog scale (VAS) of well to not-well where they think the patient falls in terms of general well-being. The distance in centimetres of the mark from the 'well' end is used as the score.

Physician global assessment of disease activity: This score is calculated in the same way as the parent/patient assessment of overall well-being but is completed by the treating physician.

Rheumatoid Factor (RF): Auto-antibody directed against the Fc fraction of IgG antibodies. RF and IgG form immune complexes which contribute to the disease.

Serum Amyloid A (SAA): A family of acute-phase proteins. Levels in the blood increase in response to tissue damage and inflammation. They also act as cytokines increasing cell migration, adhesion, and proliferation.

4. Medications

Adalimumab: An anti-TNF alpha monoclonal antibody. Marketed as Humira (Abbott).

Anakinra: A recombinant IL-1 receptor antagonist. Marketed as Kineret (Amgen).

Anti-TNFα: Drugs targeting the TNF-α cytokine. Currently include Adalimumab, an anti-TNFα monoclonal antibody, Etanercept, a recombinant human TNF receptor fusion protein, and Infliximab, a chimeric anti-TNFα monoclonal antibody.

Autologous Stem Cell Transplant (ASCT): A treatment designed to completely re-start the immune system. The pluripotent stem cells are collected from patients' blood or bone marrow and stored. The immune system is then ablated and the patient's own stem cells transfused back to re-populate it.

Canakinumab: An anti-IL1β monoclonal antibody. Marketed as Ilaris (Novartis).

Corticosteroids: Synthetic steroid hormones. Reduce inflammation and suppress the immune system.

Etanercept: A recombinant TNFα receptor. Marketed as Enbrel (Amgen and Pfizer).

Humira: Brand name (Abbott) of Adalimumab.

IL-1 Trap: An IL-1 blocking fusion protein of the extracellular domain of the interleukin-1 receptor and the FC domain of IgG1 that binds and neutralizes IL-1 Marketed as Rilonacept (Regeneron).

Infliximab: A chimeric anti-TNFα monoclonal antibody. Marketed as Remicade (Centocor Ortho Biotech).

Methlyprednisolone: Synthetic corticosteroid.

Methotrexate (MTX): An anti-inflammatory drug.

Non Steroidal Anti-Inflammatory Drug (NSAID): Non steroid drugs with anti-inflammatory, analgesic, and anti-pyretic properties. Includes aspirin and ibuprofen.

Prednisolone: An anti-inflammatory corticosteroid drug.

Rilonacept: Brand name (Regeneron) of IL-1 Trap.

Tocilizumab: An anti-IL-6 humanised monoclonal antibody.

5. Cell lines

HeLa: Epithelial cells derived from a cervical adenocarcinoma.

HepG2: Hepatic cells derived from a liver carcinoma.

Jurkat: T lymphocytes derived from T cell leukemia cells recovered from peripheral blood.

THP-1: Monocytic cells derived from acute monocytic leukemia cells recovered from peripheral blood. Can be differentiated into macrophages by stimulation with phorbyl 12-myristate 13-acetate (PMA).

6. Somatic cells

Adipocytes: Fat cells.

B-cells/ B-lymphocytes: Antibody producing white blood cells.

Cytotoxic T-cells/ cytotoxic T-lymphocytes: Induce apoptosis in infected, or damaged cells by releasing perforin, granzymes, and graulysin.

Dendritic cells: Antigen-presenting cells.

Endothelial cells: Cells which line the inner surface of blood vessels.

Epithelial cells: Cells which line the inner and outer surfaces of the body.

Erythrocytes: Red blood cells.

Fibroblasts: Cells which synthesise extracellular matrix and collagen, important in wound healing.

Hematopoietic cells: A collective term for cells derived from the bone-marrow.

Hepatocytes: Liver cells.

Keratinocytes: The most common type of skin cells, produce keratin.

Leukocytes: White blood cells.

Lymphocytes: Immune cells. The main types are B-cells, T-cells, and NK cells.

Macrophage: Phagocytotic immune cells which present antigen to T helper cells and secrete enzymes, complement proteins, and cytokines.

Mast Cells: Connective tissue cells which release histamine in response to tissue damage triggering inflammation and attracting phagocytes.

Monocytes: White blood cells which migrate to sites of inflammation and differentiate into macrophages and dendritic cells.

Mononuclear cells: Collective name for haematopoetic cells which do not have lysosomal granules in their cytoplasm, and have round nuclei. Includes macrophages, monocytes, mast cells, plasma cells, and lymphocytes.

Natural killer (NK) cell: Cytotoxic lymphocytes which release perforin and granzyme causing target cells to apoptose. Also activate macrophages.

Neutrophils: Phagocytotic granulocytic white blood cells.

Osteoclasts: Bone cells responsible for bone resorption.

Peripheral blood mononuclear cells (PBMCs): Mononuclear immune cells found in the blood circulation. Includes lymphocytes, monocytes, and macrophages.

Phagocytes: White blood cells which engulf and ingest foreign pathogens and dead or dying cells. Includes monocytes, macrophages, dendritic cells, mast cells, and neutrophils.

Splenocytes: White blood cells found in the spleen.

Stem Cells: Pluripotent unspecialised cells capable of renewing themselves through cell division and developing into any cell type.

Synoviocytes: Fibroblasts, and fibroblast-like cells found the synovial membrane of the joints.

T-cells/T-lymphocytes: Lymphocytes which have T cell receptors (TCR) on their cell surface. TCR recognise antigen bound to major histocompatibility complex (MHC) molecules resulting in cell activation.

T helper cells: Lymphocytes with no phagocytic or cytotoxic activity. Activate and direct other immune cells.

Type 1 T helper (Th1) cells: Main cytokine produced is IFNγ, produces opsonising antibodies, and activates cytotoxic T-cells and macrophages.

Type 2 T helper (Th2) cells: Produces IL-4, 5, 6, 10 and 13. Stimulates B-cell proliferation, antibody class switching, and antibody production.

White blood cells: Immune system cells. Includes neutrophils, basophils, eosinophils, lymphocytes, and monocytes.

7. Immunomodulatory molecules

Acute-phase proteins: Proteins which change in plasma concentrations in response to inflammation. They include C-reactive protein, serum amyloid A and fibrinogen.

Angiotensin 1-converting enzyme: A pulmonary and renal endothelial cell secreted enzyme involved in vasoconstriction.

AP-1: A transcription factor which regulates gene expression in response to a number of factors including cytokines.

CCAAT/enhancer-binding protein β (C/EBPβ): A transcription factor which interacts with the CCAAT motif present in several gene promoters.

Caspases: A family of cysteine proteases which are the main effectors of apoptosis.

Chemokines: A family of inflammation induced cytokines which induce chemotaxis to the site of inflammation.

Complement: A cascade of enzymes which generate opsonins, chemotactic factors, anaphlatoxins, and the membrane attack complex.

Cytokines: Signalling molecules secreted by immune system cells.

Endotoxins: Toxins found in bacterial cell walls and released following destruction of the cell.

Exotoxins: Toxins secreted by bacterial cells.

Fibrinogen: A soluble plasma glycoprotein involved in blood coagulation.

Granulocyte macrophage colony-stimulating factor: Stimulates stem cells to differentiate into granulocytes and monocytes.

Human leucocyte antigens (HLA): See Major Histocompatibility Complex (MHC).

Hydrocortisone: A corticosteroid hormone released in response to stress, suppresses the immune system. Also known as cortisol.
Interferon gamma (IFNγ): Biases the immune response towards a cellular response by enhancing monocyte/macrophage function, inducing the production of inflammatory cytokines, increasing HLA expression, activation and proliferation of B cells, and T cell differentiation.

Interleukins (IL): A family of cell signalling cytokines regulating cell growth, differentiation, and motility.

IL-2: Produced by antigen bound T cells to stimulate a T cell response.

IL-4: Stimulates the proliferation of activated B and T cells, induces naïve T helper cells to differentiate into Th2 cells.

IL-6: A pleiotropic inflammatory cytokine which induces activated B cells to differentiate into plasma cells and secrete antibodies, stimulates haemopoiesis, and activates T cells to differentiate into cytotoxic T cells.

IL-8: Chemotactically attracts neutrophils to the site of inflammation.

IL-10: An anti-inflammatory cytokine mainly produced by monocytes. It down-regulates the expression of Th1 cytokines, enhances B cell proliferation and antibody production, and can block NF-κB activity.

IL-12: Directs the differentiation of T-cells into Th1 cells, induces production of IFNγ and TNFα from T and NK cells.

IL-20: Belongs to the IL-10 family, regulates the proliferation and differentiation of keratinocytes.

Interferon regulatory factor-1 (IRF1): Activates the transcription of interferon alpha and beta.

IL-1 receptor associated kinase (IRAK): A family of serine/threonine kinases required in the IL-1 signalling cascade.

Lipopolysacharride (LPS): The highly immunogenic major component of bacterial cell walls.

Macrophage migration inhibitory factor (MIF): Suppresses apoptosis and modulates the production of pro-inflammatory mediators by macrophages, also activates T cells.

Major Histocompatibility Complex (MHC): A large gene cluster encoding histocompatability antigens, split into three class regions. Class I and II encode human leucocyte antigens (HLA) which are displayed on the cell surface and define tissue type, class III encode components of the complement cascade. Class I HLA are present on most cell types and present antigens synthesised inside the cells. Class II HLA are only present on phagocytic immune cells and present antigens from digested particles. Expression of MHC

genes is controlled by cytokines with IFNγ and TNFα being potent inducers. The MHC region is approximately 100 times more polymorphic than the rest of the genome.

Phorbol esters: Direct antagonists of protein kinase C which naturally occur in plants. In cell culture they mimic the effects of various cytokines, includes phorbol 12-myristate 13-acetate (PMA).

Pyrogens: Fever causing agents, can be endogenous or exogenous.

Transforming growth factor (TGF)β: A secreted protein involved in regulation of cell cycle and apoptosis. Important in the regulation of regulatory T cells, and blocks the activation of lymphocytes and monocyte derived phagocytes.

Tumor necrosis factor (TNF)α: A proinflammatory cytokine mainly produced by activated monocytes and macrophages. Stimulates the acute phase reaction and induces apoptosis and inflammation.

8. Nomenclature of the genes investigated in this study

Name	Protein symbol	Gene symbol	Alternative symbols
Interleukin-1 alpha	IL-1α	*IL1A*	IL1F1
Interleukin-1 beta	IL-1β	*IL1B*	IL1F2
Interleukin-1 receptor antagonist	IL-1Ra	*IL1RN*	DIRA; IRAP; IL1F3; MVCD4; IL-1ra3; ICIL-1RA; MGC10430
Interleukin-1 receptor 1	IL-1R1	*IL1R1*	P80; IL1R; CD121A; D2S1473; IL-1R-α
Interleukin-1 receptor 2	IL-1R2	*IL1R2*	IL1RB; CD121b; MGC47725;
Interleukin-1 receptor accessory protein	IL-1RAcP	*IL1RAP*	IL1R3; C3orf13; FLJ37788
Interleukin-1 receptor-like 1	IL-1RL1	*IL1RL1*	T1; ST2; DER4; ST2L; ST2V; FIT-1; MGC32623;
Interleukin-1 receptor-like 2	IL-1RL2	*IL1RL2*	IL1RRP2; IL1R-rp2;
Interleukin-18	IL-18	*IL18*	IGIF; IL-1g; IL1F4; MGC12320
Interleukin-18 binding protein	IL-18BP	*IL18BP*	IL18BPa
Interleukin-18 receptor	IL-18R1	*IL18R1*	CD218a; IL18RA; IL1RRP; CDw218a; IL-1Rrp
Interleukin-18 receptor accessory protein	IL-18RAcP	*IL18RAP*	ACPL; CD218b; IL18RB; CDw218b; MGC120589; MGC120590
IL-1 family member 5	IL-1F5	*IL1F5*	IL-36Ra; FIL1; FIL1D; IL1L1; IL1HY1; IL1RP3; MGC29840; FIL1δ
IL-1 family member 6	IL-1F6	*IL1F6*	IL-36α; FIL1; FIL1E; MGC129552; MGC129553; IL1ε; FIL1ε;
IL-1 family member 7	IL-1F7	*IL1F7*	IL-37; FIL1; FIL1Z; IL1H4; IL-1H4; IL1RP1; IL-1RP1; FIL1ζ
IL-1 family member 8	IL-1F8	*IL1F8*	IL-36β; FIL1; FIL1H; IL1H2; IL-1H2; IL1-ETA; MGC126880; MGC126882; FIL1-η
IL-1 family member 9	IL-1F9	*IL1F9*	IL-36γ; IL1E; IL1H1; IL-1H1; IL-1RP2
IL-1 family member 10	IL-1F10	*IL1F10*	FKSG75; IL-1HY2; IL1-θ; MGC119831; MGC119832; MGC119833; FIL1- θ
Caspase-1	CASP1	*CASP1*	ICE; P45; IL1BC
Cryopyrin	NALP3	*NALP3*	AII; AVP; FCU; MWS; FCAS; CIAS1;C1orf7; CLR1.1; PYPAF1; AII/AVP; AGTAVPRL; FLJ95925; NLRP3

Chapter 1
Introduction

1. Introduction

1.1. Juvenile Idiopathic Arthritis

Juvenile Idiopathic Arthritis (JIA) is defined as arthritis of no known cause that begins before the age of sixteen and persists for at least six weeks (Petty et al., 2004). It is the most common of the rheumatic diseases that affect children. The only published data for the UK gave the annual incidence of JIA as 10 per 100,000 children, with a prevalence rate of 1 in 1,000 (Symmons et al., 1996).

JIA is a clinically heterogenic disease with seven subtypes defined by the International League Against Rheumatism (ILAR) classification (Petty, 1998; Petty, 2001; Petty et al., 2004). The diagnostic definition, prevalence, peak age at onset, and sex ratios of each subtype are outlined in Table 1.1.

1.2. Systemic Juvenile Idiopathic Arthritis

Systemic JIA (sJIA) can be the most severe subtype and is potentially fatal. Patients suffer from arthritis as well as marked systemic involvement. A characteristic feature of sJIA is a fever lasting for at least two weeks and following a quotidian pattern, with a once or twice daily spike and a return to baseline temperature in-between. In addition to arthritis and fever one or more other systemic features must be present to satisfy the diagnostic criteria. These include: an evanescent or erythematous rash, lymphadenopathy, hepatomegaly or splenomegaly, and serositis (Petty, 2001; Petty et al., 2004).

The evanescent or erythematous rash is also a characteristic feature of sJIA. It commonly appears over the trunk and proximal extremities of patients but can be generalised. The rash typically accompanies the fever, is salmon-pink and macular, although it is urticarial in 10% of patients (Schneider and Laxer, 1998). Pericarditis is the most common form of serositis seen in sJIA patients but is normally asymptomatic (Davidson J 2000). More severe myocardial involvement is potentially fatal but rare, although 10-20% of patients have subclinical myocarditis (Goldenberg et al., 1992a). During phases of active systemic involvement hepatomegaly is common, and splenomegaly is seen in over 50% of patients (Goldenberg et al., 1992b). On initial presentation arthritis is not always manifest. However, it does usually develop within the first few months of disease onset, but it may not develop for some years. The most commonly affected joints are the knees, wrists, and ankles. Cervical

spine and hip involvement, which is generally bilateral and can be rapidly destructive, are also seen (Schneider and Laxer, 1998).

sJIA is the only JIA subtype which does not predominantly affect females, showing very little discrimination between the sexes (Laxer and Schneider, 2004). There are three main courses which the disease follows. Patients may follow a monocyclic course with complete remission within a few years of onset, their systemic features may settle but be followed by persistent polyarthritis, or they may follow a polycyclic course with continued flares in systemic inflammation (Calabro et al., 1976; Lomater et al., 2000). The average duration of active disease is 5 to 6 years but some patients experience disease activity into adult life, with 40-50% of patients having active disease 10 years after onset (Wallace and Levinson, 1991).

1.2.1. Complications

There are a number of complications associated with sJIA. Some are the result of disease activity, while others are a consequence of the drugs used to control the inflammation. Many children experience joint abnormalities, osteoporosis, and growth abnormalities, which continue into adulthood. Of a cohort of sJIA patients still under the care of a rheumatologist five years after diagnosis, 18% showed a lack of linear growth, 75% had joint damage, and 30% suffered functional limitations. Eleven percent of the patients were unable to attend school full-time as a result of their disease, and 44% were unable to participate in a full school programme (Bowyer et al., 2003).

Impaired linear growth and delayed sexual maturation is often seen in patients with sJIA (Schneider and Laxer, 1998). These effects can be limited however, through the administration of growth hormone (Davies et al., 1994; Simon et al., 2003). Localised joint problems are also common if inflammation is not controlled early or sufficiently. Flexion deformities can develop rapidly at inflamed joints, persistent inflammation may lead to bony overgrowth at the affected joint, and if the affected joint is a knee, leg length discrepancy can develop. Developmental problems and muscle atrophy can also arise due to lack of or altered mechanical use, for example altered gait, as a result of pain avoidance (Davidson, 2000). Poor nutrition and anorexia, a result of hypercatabolism and lack of appetite due to malaise, are seen in sJIA patients and contribute to reduced growth. Osteoporosis, due to poor nutrition and reduced bone turnover is another condition commonly seen in sJIA patients which contributes to their reduced growth. Anaemia, due to poor nutritional intake, chronic disease, and gastrointestinal blood loss as a side-effect of some drug treatments, is also seen

Subtype name	Diagnostic definition	Incidence (per 100,000)	Peak age at onset (yrs)	Sex ratio F:M
1. Systemic arthritis	Arthritis, associated with or preceded, by a spiking quotidian fever for at least 3 days, and at least one of: evanescent erythematous rash, generalised lymphadenopathy, hepatomegaly and/or splenomegaly, and serositis	10	2-4	1:1
2. Oligoarthritis	Arthritis affecting between 1 and 4 joints in the first 6 months	60	<6	4:1
(a) Persistent	No more than 4 joints affected throughout the disease	42 to 48	n/a	n/a
(b) Extended	More than four joints affected after the first 6 months	12 to 18	n/a	n/a
3. Polyarthritis RF –ve	Arthritis affecting 5 or more joints in the first 6 months, patient negative for RF	40	6-7	3:1
4. Polyarthritis RF +ve	Arthritis affecting 5 or more joints in the first 6 months, patient positive for RF in 2 or more tests at least 3 months apart in the first 6 months of disease	10	9-12	9:1
5. Psoriatic arthritis	Arthritis with psoriasis or arthritis with at least 2 of: dactylitis, nail pitting or onycholysis, psoriasis in a first degree relative	15	7-10	2:1
6. Enthesitis related arthritis	Arthritis with enthesitis, or, either arthritis or enthesitis with at least two of: history or presence of sacroiliac joint tenderness and/or inflammatory lumbosacral pain, presence of HLA-B27, history of: AS, enthesitis related arthritis, sacroiliitis or acute anterior uveitis in a first degree relative, acute anterior uveitis, and onset of arthritis in a boy over 6 years	50	9-12	7:1
7. Undifferentiated arthritis	Arthritis that either fulfils the criteria of more than one, or none of the other subtypes	1	n/a	n/a

Table 1.1 Features of the JIA subtypes according to the ILAR classification

The main features of the JIA subtypes defined by the International League of Associations for Rheumatology (ILAR) are shown. The diagnostic criteria of each subtype (Petty et al., 2004) listed demonstrates the marked clinical heterogeneity between the subtypes. The incidence (Woo et al., 2007), peak age of onset, and sex distribution (Nistala et al., 2009) of the subtypes are also shown.

AS – Ankylosing Spondylitis, HLA- Human Leukocyte Antigen, RF- Rheumatoid Factor

in sJIA (Laxer and Schneider, 2004).

A rare but serious complication in sJIA is the development of macrophage activation syndrome (MAS). This is a set of clinical symptoms caused by the excessive activation of macrophages, which are found in the bone marrow and other reticular endothelial tissues. MAS can develop at any point during the course of sJIA; at the onset of disease, during a period of active disease, or even during an inactive period. The trigger for development is not known but it may be related to uncontrolled disease activity, infection, and the use of certain drugs including sulphasalazine, non-steroidal anti-inflammatory drugs (NSAIDs), and gold. Persistent fever, hepatosplenomegaly, lymphadenopathy and encephalopathy are typical features of MAS. Other features seen include pneumonitis, erythematous rash, panniculitis and bleeding. Patients can also have neurological involvement with changes in mental status, seizures, and coma (Laxer and Schneider, 2004). Blood tests show reduced white blood cell and platelet counts, elevation in hepatic transaminases, coagulopathy, and severe anaemia. In contrast to the high erythrocyte sedimentation rate (ESR) seen during a flare in systemic activity, ESR in patients with MAS is normal or reduced (Nistala et al., 2009), while C-reactive protein (CRP) is raised. MAS is a rare complication in sJIA with a reported prevalence of 7% in UK patients (Sawhney et al., 2001b). The prevalence may however be higher as it is possible that not all cases are diagnosed. This is because it has a superficially similar presentation to systemic flare and therefore may be not distinguished. Based on hematopathological evidence of MAS in bone marrow aspirates from sJIA patients it is suggested that 53% of sJIA patients develop MAS (Behrens et al., 2007). If not diagnosed and treated aggressively and promptly MAS can be fatal. Of the patients with MAS seen at Great Ormond Street Hospital between 1989 and 2000 there was a 22% mortality rate, mainly a consequence of multi organ failure (Sawhney et al., 2001a).

The other severe complication strongly associated with sJIA is the development of amyloidosis, where normally soluble proteins are deposited as insoluble fibrils. In sJIA cases in Europe is it reported to occur in approximately 9-10% of patients (Svantesson et al., 1983). The organs most commonly affected by this disorder are the liver, spleen, and kidneys. Late involvement includes the gastrointestinal tract, vasculature, and heart. Patients present with proteinuria, hypertension, hepatomegaly, diarrhoea and abdominal pain. Due to resulting kidney failure or infection, amyloidosis may be fatal (Schneider and Laxer, 1998).

In the past mortality from sJIA was reported to be as high as 14% (Stoeber, 1981; Hafner and Truckenbrodt, 1986). Mortality rates are believed to be much lower now, although no formal statistics are available. However, mainly due to the incidence of MAS and amyloidosis, sJIA is still perceived to be the most potentially fatal of the JIA subtypes (Woo et al., 2007).

1.2.2. Diagnosis

The diagnosis of systemic JIA may be difficult as the systemic features can precede the development of arthritis, required for a definite diagnosis, by several weeks or months. Patients may present non-specifically with a fever of unknown origin and a marked acute phase response on laboratory investigation. Laboratory features seen in patients include, very high CRP and ESR, normal to high RF and complement, leukocytosis with neutropenia, thrombocytosis, and anaemia. In severe cases liver enzymes, ferritin, and coagulation markers may also be abnormal (Nistala et al., 2009). There are currently no disease-specific markers for sJIA. In order to exclude a number of other conditions which can present in a similar manner to sJIA, an extensive investigative work-up is essential (Table 1.2).

1.2.3. Treatment

The treatment strategy for sJIA employed at Great Ormond Street Hospital is outlined in Figure 1.1, demonstrating the order in which the different therapies are administered. The majority of sJIA patients respond well to high dose corticosteroids. However, dose reductions, which are desirable due to the associated adverse side effects, often result in disease flare. Despite efficacy rates of less than 40%, the immunosuppressive drug Methotrexate (MTX) is currently recommended as a first line disease modifying drug. This is because MTX has fewer side effects, and more data regarding long term use, than other available drugs (personal communication, P.Woo, 2010).

For patients with only mild symptoms, NSAIDs administered to cover whole 24 hour periods may be sufficient to alleviate the disease. However, in patients with more severe disease who do not respond to NSAID therapy it may be necessary to use steroids, including prednisolone and methylprednisolone. Resolution of systemic features and a significant improvement in laboratory disease features in a mean of 2.1 months after treatment with alternate day high dose prednisolone has been shown in sJIA patients. These clinical improvements, and a decrease in the number of joints with active arthritis, were maintained over a one year

Condition	Differentiating features from sJIA
Infections	Positive cultures/antibodies/serology, clinical history, continuous fever and rash
Malignancy	Nonquotidian fevers, bone pain at night, low white blood cell and platelet counts, positive lymph node biopsy
Neuroblastoma	Nonquotidian fevers, systemically unwell constantly
CINCA	Fixed rash, undulating fevers, neurologic complications
Kawasaki disease	Fixed rash, mucocutaneous symptoms, conjunctivitis, extremity changes, coronary artery dilation
Systemic vasculitis	Undulating fevers, fixed and painful rashes or pupura, painful subcutaneous nodules, nephritis, positive ANCA
SLE	Constant fevers, malar rash, photosensitivity, low complement, positive ANA and dsDNA, cytopenias
Sarcoidosis	Increased calcium, ACE, erythema nodosum, hilar lymphadenopathy
Other periodic fever syndromes*	Characteristic fever patterns, family history, gene studies

Table 1.2 Differential diagnosis of sJIA

Adapted from (Woo, 2006; Woo et al., 2007).

A number of diseases can present in a similar manner to sJIA. In order to proceed with the appropriate treatment it is important for a correct diagnosis to be made using the clinical and laboratory features shown.

* Including Muckle-Wells syndrome, Familial Mediterranean fever, hyperimmunoglobulemia D with recurrent fever, TNF receptor associated periodic syndrome, and Familial cold uticaria.

ACE – angiotensin converting enzyme, ANCA- anti-neutrophil cytoplasmic antibodies, CINCA- Chronic Infantile Neurologic Cutaneous and Arthropathy Syndrome, SLE- Systemic Lupus Erythematosus

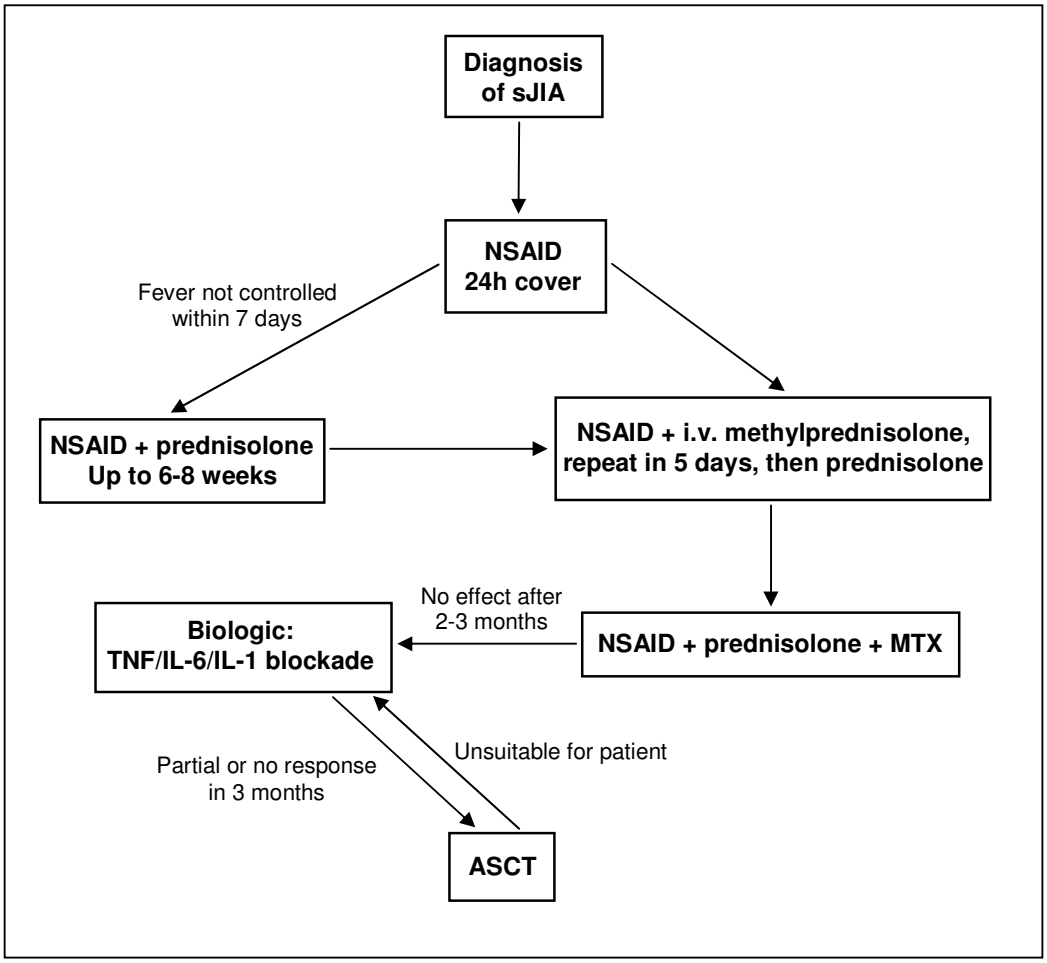


Figure 1.1 Treatment strategy for sJIA

Based on personal communication, P.Woo, 2010

The treatment strategy employed at Great Ormond Street Hospital for sJIA is illustrated. If a treatment regimen does not provide a satisfactory improvement in the patient, the algorithm is followed and the next regimen implemented.

ASCT-Autologous Stem Cell Transplant, i.v.–intravenous, MTX – Methotrexate, NSAID- Non Steroidal Anti-Inflammatory Drug, TNF-Tumour Necrosis Factor

follow-up period. However, a number of side-effects were also observed in these patients, including reduced longitudinal growth, mild cataracts, vertebral crush fracture, and a possible case of avascular necrosis (Kimura et al., 2000). The documented side-effects of steroid use include growth retardation, weight gain, osteoporosis, and immune suppression (Perez et al., 2000). This means that long term systemic treatment with steroids is undesirable, especially in young, growing patients. Therefore, for the control of arthritis, intra-articular steroid injections are, where possible, preferable. However, this shows the shortest period of effectiveness in patients with the systemic subtype of JIA (Breit et al., 2000).

Conflicting results of treatment with MTX have been reported in patients with sJIA. MTX acts through decreasing the expression levels of the pro-inflammatory cytokines IL-18, IL-1β, IL-12, and IFNγ, and increasing the expression of the anti-inflammatory cytokines IL-1Ra, IL-4, and IL-10 (Seitz et al., 1996; Constantin et al., 1998). sJIA patients with progressive disease, despite treatment with other medications, were treated in a double-blind crossover trial with MTX for four months, and with placebo for four months. A greater reduction in systemic features was seen with active treatment than placebo, but this was not statistically significant and there was no difference between the number of active joints, or the range of motion following either treatment. (Woo et al., 2000). In an earlier study however a third of a group of patients treated with low dose MTX, responded well, with four patients being able to reduce their steroid treatment. However, three patients had to have their steroids increased, and two experienced disease flares during treatment (Speckmaier et al., 1989). Although MTX has been shown to be less effective in sJIA than in other subtypes (Halle and Prieur, 1991), improvement in disease severity, enabling discontinuation of cortiosteroids, can be achieved, either alone or in combination with cyclophosphamide (Wallace and Sherry, 1997).

Anti-TNFα therapy has been shown to be effective in patients with oligoarticular and polyarticular JIA (Quartier et al., 2003a; Lovell et al., 2003), but is less efficacious in sJIA. Approximately 50% of a sJIA cohort, who had previously not tolerated or not responded to MTX, achieved a >30% improvement after three months treatment with the recombinant TNFα receptor Etanercept. However, four patients experienced severe side-effects and withdrew from treatment, three experienced systemic symptoms, and a further four had a disease flare during the first year of treatment. The oligoarthritis and polyarthritis patients also included in the trial achieved a significantly better response (Quartier et al., 2003b). Similar results were seen in another cohort, where only 30% of the sJIA patients showed a

response to Etanercept treatment, compared to 70% of patients with other subtypes (Horneff et al., 2004). Only 46% of a sJIA cohort showed a >50% reduction in disease severity after treatment with Etanercept. Although the patients who responded were able to discontinue use of steroids, 45% of the patients experienced a disease flare, and two developed MAS (Kimura et al., 2005). There is also some evidence that even patients who respond initially may not experience a sustained improvement. Only 24% of patients treated with anti-TNFα (Etanercept, or the anti-TNF antibodies Infliximab, or Adalimumab) initially satisfied the criteria for remission, but 45% of them later experienced a disease flare, with only 13% of the patients showing a sustained improvement (Russo and Katsicas, 2009).

Very good results have been achieved with the use of other biologics in blockade of IL-6 and IL-1. Tocilizumab is an anti-IL-6 receptor antibody which binds to both the membrane bound and soluble forms, effectively blocking cellular signalling through IL-6 (Nishimoto and Kishimoto, 2004). A single intravenous infusion of tocilizumab induced a marked clinical response within 48 hours in 18 Caucasian sJIA patients refractory to other treatments. One week after treatment 11 had achieved an ACR30 improvement score, accompanied by improvements in CRP and haemoglobin levels, ESR, and total white cell counts. The improvements were seen for up to eight weeks following administration. Two patients experienced a flare in disease activity but there were no other serious adverse events (Woo et al., 2005). Similar results were also seen in a cohort of 11 Japanese patients who received tocilizumab every two weeks, with an increase in the dose if active inflammation was present. Fever resolution and improvement in arthritis was seen in all patients within three days to four weeks. Two weeks after initiation of treatment seven patients achieved an ACR70 improvement score, and a further three achieved an ACR50 (Yokota et al., 2005). These significant improvements have been confirmed in the Japanese population in a phase III double-blind placebo control study in which 68% of patients achieved ACR70 during the open-label period and a significantly lower proportion of patients receiving active treatment than placebo withdrew from the trial during the double-blind section (Yokota et al., 2006). Promising initial results have also been reported from a double-blind placebo controlled five year global multicenter phase III study, with significantly more patients receiving active treatment achieving a positive response during the initial 12 weeks (de Benedetti et al., 2010).

Similarly dramatic responses have been achieved with IL-1 blockade therapy, of which Anakinra, a recombinant IL-1 receptor antagonist, is the most extensively investigated. The

first publication of the use of Anakinra in sJIA was in two patients who had severe disease refractory to all other available therapies including corticosteroids, MTX, etanercept, and infliximab. Both patients experienced resolution of fever, rash, myalgia, and joint pain within 24 hours and two days of treatment. One patient had active arthritis in 22 joints at initiation of treatment, all of which had fully resolved after three weeks. Laboratory measurements in both patients returned to normal, and both remained symptom free without any other medication for the follow-up periods of six months, and one year (Verbsky and White, 2004). In another trial of nine patients all seven with systemic symptoms at initiation of therapy, became afebrile within the first week and remained so throughout the follow-up period of two to twelve months. Of the eight patients with active arthritis at the initiation of therapy the arthritis score completely resolved in six, and improved in the remaining two. This was accompanied by improvement in the laboratory markers of inflammation; leukocytosis and thrombocytosis resolved, and haemoglobin levels, erythrocyte counts, and ESR improved in all patients (Pascual et al., 2005). Similar results have been achieved in other, smaller, studies, particularly with patients refractory to previous therapies (Henrickson, 2004; Irigoyen et al., 2004). Although these initial studies reported that nearly all patients responded well to Anakinra treatment, further studies suggest that there are at least two subpopulations within sJIA: Anakinra responsive and non-responsive (Quartier et al., 2006). In a trial of 20 sJIA patients approximately 50% showed a minimum ACR30 improvement, of which 10% showed an ACR70. Fifteen of the patients did however show improvement initially, and clinical symptoms resolved within three months in 14 patients (Lequerre et al., 2008). Similar results were also seen in an independent trial of 21 patients, 10 of whom were able to discontinue all other medications after four months and showed control of systemic and articular features over the mean follow-up period of 1.36 years. The systemic features, fever and rash, were however well controlled in the remaining 11 patients, who showed initial improvement but later disease recurrence despite dose increases (Gattorno et al., 2008). There have however been reports of patients experiencing serious adverse events due to Anakinra, including MAS (Gattorno et al., 2008). After initiation of treatment three patients developed acute hepatitis, with a rapid improvement in liver enzymes and disease symptoms on cessation of Anakinra, although two patients also experienced disease flares. The authors hypothesised that the hepatitis occurred as a result of Anakinra induced altered immune response to infection, or that it was an atypical presentation of MAS (Canna et al., 2009).

A major drawback of Anakinra is a short half life, making daily injections necessary. This

increases the risk of local reactions at the injection site, and is not ideal or practical, especially in younger patients. Other IL-1 blocking drugs are being developed to overcome this issue. At the time of writing these are still undergoing clinical trials, although the preliminary evidence is promising. Four weekly injections of the recombinant IL-1 type I receptor (IL-1 Trap) Rilonacept resolved fever and rash in all patients treated, 77.8% of whom achieved ACR50, and 44.4% ACR70, after four weeks (Lovell et al., 2006). Eleven of the 19 patients given the IL-1β antibody Canakinumab in a phase II trial, as a single injection repeated only following disease flare, fever recurrence, or elevated CRP, showed an improvement. All of the responders had an ACR50 improvement score after 15 days, and four were classified as being in remission. The time from treatment to evidence of relapse varied between 23 and 200 days, with only one adverse event of gastritis with ulcer bleeding occurring (Ruperto et al., 2008). Both of these drugs are currently undergoing larger long-term phase III trials to investigate efficacy and safety.

Autologous stem cell transplants (ASCT) have been carried out with positive results in patients with progressive disease refractory to all other therapies. Of 18 sJIA patients with progressive disease for a minimum of 5 years, 33% went into remission following ASCT. A further 27% did not show response to ASCT, but did go on to respond well to second-line treatment to which they had previously been refractory. However, there was a mortality rate in these patients of 22%. Two of the patients died shortly after ASCT, and another two, who had not responded to treatment, died as a result of infection while undergoing further intensive immunosuppression (Brinkman et al., 2007). Greater success rates have been achieved in other ASCT studies however. Of seven UK sJIA patients, 57% achieved treatment-free remission with catch-up growth over the 5-8 year follow up. Two of the patients experienced disease relapse within one year, and one patient (14%) died as a result of disseminated adenovirus reactivation (Abinun et al., 2009). 100% success has been reported in smaller studies of one, and three patients (Woolfrey et al., 2010; Wulffraat et al., 1999).

1.3. Inflammation and Cytokines

Inflammation, with hallmark symptoms of swelling, redness, heat, pain, and loss of function of the inflamed area, is an essential process in protection against external insults (Benjamini et al., 2000). The inflammatory process involves recognition of foreign stimulus, production of soluble mediators, acute phase response, cellular response, and resolution. The

inflammatory response is initiated by infection, trauma, allergy, or the release of inflammatory mediators, which may be derived from microbes, damaged tissue, mast cells, other leukocytes, and complement components, into tissues (Lydyard et al., 2000).

Following initiation of inflammation, leukocytes, including macrophages, lymphocytes, and neutrophils, are attracted to, and infiltrate the affected site. After phagocytosing microbes macrophages become stimulated and release cytokines. These increase vascular permeability and expression of adhesion molecules, attracting more cells to the site and further amplifying the inflammatory response (Benjamini et al., 2000). Pyrogenic cytokines, including interleukins (IL) -1 and -6, tumor necrosis factor (TNF), and interferons (IFN), act on the hypothalamus inducing fever to retard microbe growth, and on hepatocytes, inducing production of acute phase proteins (Benjamini et al., 2000). There are two main parts of immune defence, innate and adaptive immunity. Innate immunity is a non-specific, immediate, response to infection. The main components are cytokine and complement activation, and, through cytokine induced release of acute phase proteins, opsonisation of microbes enhancing phagocytosis. The adaptive immune response is specific and involves the activation of immune lymphocytes and other cellular changes through antigen presentation, and results in the development of immunological memory. Cytokines produced during inflammation trigger T-cell differentiation. IL-12 drives T-cells towards the Th1 pathway which is mainly a response to cellular infections leading to cell-mediated immunity. The Th2 pathway, in which IL-4 is one of the key cytokines, is usually in response to extra-cellular infections and results in antibody mediated immunity. The IL-17 producing Th17 cell type has recently been defined, and is thought to derive from Th1 cells (Nistala et al., 2010).

Once the injury or microorganism has been removed or controlled, inhibitors dampen inflammation, and tissue repair mechanisms become activated. Anti-inflammatory cytokines, including IL-10, IL-4, and TGFβ, and cytokine antagonists, including soluble receptors, are expressed. These neutralise the inflammatory response enabling repair of the damage through production of collagen by various cells including myofibroblasts and macrophages (Lydyard et al., 2000). If the cause of the inflammation is not removed, for example during chronic infection such as tuberculosis, then chronic inflammation occurs (Benjamini et al., 2000). Chronic inflammation can also result from persistence of a pro-inflammatory stimulus, for example the persistence of auto-antigen in the case of autoimmunity, an impairment of the mechanisms that neutralise inflammation, or a clonal disorder with pro-inflammatory

consequences (Davies and Woo, 2004). It is believed that the balance between the production of immuno-activating and inflammatory cytokines, and the release of immunosuppressive agents and anti-inflammatory cytokines, is key to the outcome of the inflammatory response. Disruption of this balance results in pathogenic chronic inflammation.

1.4. Altered Cytokine Profile in sJIA

A number of studies have shown that patients with sJIA have a different cytokine expression profile from both healthy controls and patients with other JIA subtypes (Pascual et al., 2005; Allantaz et al., 2007; Barnes et al., 2009; Ishikawa et al., 2009; Mangge et al., 1999). It has also been shown that the cytokine expression profile in sJIA patients varies over the disease course, with 286 genes up-regulated during active compared to inactive disease phases (Ogilvie et al., 2007). Some of the published results regarding cytokine levels have, however, been conflicting. There are a number of possible explanations for these discrepancies resulting from differences in study design. These include: measuring RNA or protein levels, use of peripheral blood mononuclear cells (PBMCs), plasma, serum, or synovial fluid samples, if the assay used detects only free cytokine or also cytokine/receptor complexes, the disease phase of the patients at the time of sampling, and also the time of day samples are collected as, for example, IL-6 follows a circadian pattern (Sothern et al., 1995a; Sothern et al., 1995b). The conflicting findings may also be a result of disease heterogeneity. Patients fall along a spectrum of disease severity, suffering from mild to severe disease, and respond differently to various treatments (section 1.2.3.). Differences in clinical phenotype may reflect differences in the underlying physiopathologies and thus the cytokine profiles observed.

Patients with sJIA have been shown to have up-regulation of IL-1 immune pathways (Pascual et al., 2005), although this was not replicated in another study (Ogilvie et al., 2007; Barnes et al., 2009). Levels of IL-1β and IL-1Ra have been shown to be significantly higher in sJIA than in healthy controls (Madson et al., 1994; de Benedetti et al., 1995; Muller et al., 1998; Lotito et al., 2007), but other publications have found opposing results (de Benedetti et al., 1995; Muller et al., 1998; Muzaffer et al., 2002). As levels of IL-1Ra have also been shown to parallel the fever curve in sJIA (Rooney et al., 1995; Prieur et al., 1996), and to correlate with disease activity (de Benedetti et al., 1995; Muzaffer et al., 2002), these conflicting results may be due to variation in disease activity at the time of sample collection in each

study. Pascual et al found that when PBMCs from healthy individuals were incubated with serum from active sJIA patients 46 genes showed a more than two fold increased expression compared to healthy PBMCs incubated without culture, and those incubated with autologous serum. Several members of the IL-1 cytokine and cytokine receptor family were included in the upregulated genes; *IL1β, IL1RN, IL1R1* and *IL1R2*. It was also shown that when incubated with sera from sJIA patients healthy PBMCs were induced to produce 63 ± 39 pg/ml IL-1β protein, where as when incubated with healthy sera they produced less than 10 pg/ml. In addition to this they showed that sera from febrile sJIA patients were even more efficient at inducing IL-1β secretion than sera from afebrile patients. They also showed that PBMCs from sJIA patients had a greater capacity to secrete IL-1β after stimulation with PMA-ionomycin than PBMCs from healthy individuals, 333 ± 204 pg/ml compared to only 21 ± 4 pg/ml, but there was no significant difference in the production of IL-6 and TNF between patients and controls in the same cultures (Pascual et al., 2005), although this finding was not replicated in other independent studies (Muller et al., 1998; Gattorno et al., 2008).

IL-6 is a pro-inflammatory, pyrogenic cytokine, which regulates the haematopoesis, immune response, the acute phase response, and inflammation (Taga and Kishimoto, 1997). It has been shown that compared to healthy controls, both RNA and protein levels of IL-6 are significantly higher in patients with sJIA during active, but not inactive, disease phases (de Benedetti et al., 1991a; Ogilvie et al., 2007; Barnes et al., 2009). Levels of IL-6 have also been shown to be higher in patients with sJIA than in patients with other subtypes (Rooney et al., 1995; Lotito et al., 2007). However, the converse results of IL-6 levels being lower in sJIA than the other subtypes, have also been shown (de Jager et al., 2007). Increased levels of IL-6 in sJIA are not accompanied by increased levels of the IL-6 antagonists soluble gp130 and IL-6 auto-antibodies, which are present at similar levels to healthy controls and patients with other JIA subtypes (Keul et al., 1998). This indicates that the balance of the IL-6 system in sJIA patients is towards a pro-inflammatory state. It has been demonstrated that the levels of IL-6 rise and fall paralleling temperature changes during the quotidian fever (Rooney et al., 1995; Prieur et al., 1996), and correlate with disease activity, including joint involvement and platelet count (de Benedetti et al., 1991b; Rooney et al., 1995), although no difference in IL-6 mRNA levels was seen in PBMCs between patients during active and inactive disease phases (Ogilvie et al., 2007). There are two forms of the IL-6 receptor, membrane bound and soluble, expressed through alternative splicing. Following binding to IL-6, both receptors can associate with the membrane bound co-receptor gp130, and initiate a signalling cascade

(Murakami et al., 1993). Although there is no significant difference between the levels of soluble IL-6 receptor (sIL-6R) in sJIA patients and healthy controls or individuals with oligo- or polyarthritis, a significantly higher level of circulating IL-6/sIL-6R complex is present. The authors hypothesised that while in a complex with the soluble receptor, IL-6 is protected from protease degradation, maintaining a higher level of active IL-6 in circulation (De Benedetti et al., 1994b). The soluble form of the IL-6 receptor has been shown to be negatively correlated to IL-6 levels, and that during febrile periods increased IL-6 is accompanied by decreased sIL-6R (De Benedetti et al., 1994a). This finding is unexpected, as decreased sIL-6R would result in decreased IL-6 signalling, and potentially resolution of the inflammatory response. It is possible however that there is an increased expression of the membrane bound IL-6 receptor which compensates and enables continuation of the IL-6 induced inflammation. Membrane bound proteins are however harder to measure than circulating proteins and no reports have been published.

Strikingly high levels of IL-18 have been shown in patients with sJIA compared to both healthy controls and all other JIA subtypes. Although levels of IL-18 are significantly lower during inactive than active sJIA disease phases it is still significantly higher, regardless of disease activity stage (de Jager et al., 2007; Jelusic et al., 2007; Lotito et al., 2007). The difference in IL-18 levels is very marked, with published levels in sJIA patients and healthy controls respectively of: 45,073pg/ml compared 305pg/ml (Maeno et al., 2002), and 130,000pg/ml compared to 140.5pg/ml (de Jager et al., 2007). Based only on levels of IL-18, patients with sJIA can be identified with 93% accuracy from patients with other JIA subtypes, demonstrating the subtype specificity of this pro-inflammatory cytokine (de Jager et al., 2007). This specificity also extends to comparisons with other inflammatory diseases; Kawakaki disease and Epstein-Barr Virus infection show significantly lower IL-18 levels than sJIA (Shimizu et al., 2010). IL-18 levels are also a marker of disease severity, correlating to CRP levels, the number of joints with active arthritis, and radiological scores of joint damage. Levels of IL-18 have also been reported to rapidly and dramatically rise during development of MAS, and to fall, although remaining significantly higher than in healthy controls, following resolution, (Lotito et al., 2007).

Conflicting results concerning expression levels of the anti-inflammatory cytokine IL-10 have been presented. It has been shown that compared to healthy controls and patients with other JIA subtypes, genes involved in IL-10 signalling are up-regulated in sJIA (Barnes et al.,

2009), and that expression levels are higher during active disease phases than inactive phases (Ogilvie et al., 2007). However, it has also been shown that following stimulation, cultured blood cells from sJIA patients express two to three times less IL-10 than cells from healthy control individuals flowing LPS stimulation (Muller et al., 1998). These conflicting results mean that the role of IL-10 in the pathogenesis of sJIA is uncertain.

Levels of TNFα have been shown to be higher in sJIA than in healthy controls, but not compared to other JIA subtypes (Madson et al., 1994). Levels of the TNFα receptors I and II however, are higher in sJIA than in both healthy controls and other JIA subtypes (Muller et al., 1998; Muzaffer et al., 2002). This would, despite the same ligand concentration, result in a greater TNFα inflammatory response in sJIA due to the higher number of ligand/receptor complexes able to induce a signalling cascade.

1.5. Genetic Associations

The first indication of a significant genetic component involved in susceptibility to a given disease, is that multiple affected members are found within families. In other rheumatic and non-rheumatic autoimmune diseases extensive kindreds with multiple affected family members in several generations are seen. This is not however not the case in JIA. Of over three thousand patients with JIA throughout Europe there were only 12 affected sibling pairs (Clemens et al., 1985), and in 1994 it was estimated that there were only 300 affected sibling pairs in the United States. Despite the rarity of familial cases there is evidence from family studies that JIA is a genetic disease. In a study of JIA affected sibling pairs, 76% were concordant for subtype. In the seven sets of twins (zygosity unknown) in the study there was marked concordance in the age of onset, with a mean difference of only 3.3 months, compared to two years in non-twins (Moroldo et al., 1997). A concordance rate of 25% has been estimated between monozygotic twins, which, given the population prevalence, equates to a relative risk of approximately 250 (Savolainen et al., 2000). Risk to siblings of an affected family member, λ_s, calculated as the prevalence of the disease in siblings of an affected child divided by the prevalence of the disease in the general population, has been estimated as fifteen, similar to that seen with insulin-dependant diabetes mellitus (Glass. D and Giannini EH 1999), which is widely acknowledged to have a genetic component. The recurrence risk of JIA in a first degree relative (parent, sibling, offspring) of a proband is significantly higher than for a relative of an unaffected individual (OR 30.4) (Prahalad et al.,

2004). There is also anecdotal evidence that families with a JIA proband have an increased risk of other rheumatic and autoimmune diseases. The prevalence of the auto-immune diseases alopecia areata, insulin-dependant diabetes mellitus, thyroid disease, vitiligo, and autoimmune oophoritis, are 16.1% and 10.6% higher than in the general population, in first and second degree relatives of a JIA proband respectively (Prahalad et al., 2002). This evidence suggests that in these families there may be a predisposition to autoimmunity in general, with a variation of diseases manifesting in different family members. Because of the non-Mendelian pattern of inheritance seen in JIA, and the risk to siblings, it is highly likely that JIA is a complex genetic trait, where many genes, each with a small effect, and environmental factors, interact to influence development of pathology. It is also postulated that there are likely to be some genetic associations common to all JIA subtypes but, due to disease heterogeneity, that the majority will be subtype specific.

1.5.1. Major Histo-compatibility Complex

The genetic associations with JIA as a whole which have been confirmed the most consistently are, as also seen in other autoimmune diseases, those in the Major Histocompatibility Complex (MHC), or Human Leukocyte Antigen (HLA) system (Glass and Giannini, 1999). A number of HLA associations have been shown with JIA (Date et al., 1999a; Thomson et al., 2002; Saila et al., 2004). DRB1*11, DQA1*05 (Thomson et al., 2002), DRB1*0405, and DQB1*0401 (Date et al., 1999b), show evidence of significant association with the systemic subtype.

The effects of HLA associations are, however, not sufficient to fully explain susceptibility to JIA; it has been estimated that the HLA associations account for only approximately 17% of the susceptibility risk for JIA (Prahalad et al., 2000). In addition, reported HLA associations with sJIA separately have not been replicated, and it is unlikely that they play as significant a role in this subtype as in others (Woo, 2006).

1.5.2. Non-MHC Genes

In addition to these MHC associations there have been several non-MHC genes found to be associated with sJIA, although these have been less reproducible in different patient populations.

Associations have been found with JIA as a whole and various genes. These include: *TNFα*, *CTLA4* (cytotoxic T lymphocyte-associated 4), *MIF* (macrophage migration inhibitory factor) (Miterski et al., 2004; Donn et al., 2002), *PTPN22* (protein tyrosine phosphatase N22) (Hinks et al., 2005), *ACE* (angiotensin 1-converting enzyme) (Alsaeid et al., 2003), and *IRF1* (interferon regulatory factor-1) (Donn et al., 2001a), although this association was not replicated in a larger cohort (Fife et al., 2007).

There are also a number of associations which are specific to the systemic subtype of JIA. Fishman et al found an association between sJIA and the G allele of a polymorphism at –174 of *IL6*. This association was even more pronounced in patients with disease onset before the age of five. Unstimulated HeLa cells transfected with a luciferase reporter construct containing the G allele showed a 0.62 fold higher transcription levels of *IL6* than cells transfected with the C allele construct. Following stimulation with either LPS or IL-1, cells transfected with the G allele construct showed an increased level of expression of *IL6*, while there was no increased expression with the C allele construct. The expression difference was also demonstrated at the protein level in healthy control individuals. Individuals homozygous for the G allele had IL-6 plasma levels approximately double that of individuals homozygous for the C allele (Fishman et al., 1998). These results suggest that due to genetic variation, patients with sJIA have higher, and inducible, expression levels of the pro-inflammatory cytokine IL-6. This association has been replicated in a multi-centre transmission disequilibrium test (TDT) family based study (Ogilvie et al., 2003), and the effect has also been shown to extend into a three marker haplotype with an increased level of association (Fife et al., 2005).

In a haplotype of three single nucleotide polymorphisms (SNPs) in the 5′ regulatory region of the anti-inflammatory cytokine *IL10*, the ATA haplotype was found to be significantly more common in JIA patients following a polyarticular, rather than an oligoarticular, disease course. In transient transfection assays the ATA haplotype was associated with lower transcriptional activity of *IL10*, and PBMCs from individuals carrying two copies if the ATA haplotype showed significantly lower IL-10 production following lipopolysaccharide (LPS) stimulation, than those from individuals carrying any other combination of haplotypes (Crawley E et al 1999). This significant reduction in IL-10 expression levels has, however, also been attributed to the A allele at -1082, which is the first SNP in the three SNP haplotype (Turner et al., 1997). A further study found that just the -1082 A allele, rather than the

haplotype, was significantly associated with the systemic subtype of JIA. This study also found a significant association of a SNP in the 5′ region of *IL-20* with sJIA (Fife et al., 2006).

The C allele of a polymorphism at position -173 in the 5' promoter region of the T-cell activator *MIF,* macrophage migration inhibitory factor, is at a significantly higher frequency in patients with sJIA than in healthy controls. This C allele creates a transcription factor binding site for activator protein 4 (Donn et al., 2001b). It was also found that patients carrying the C allele have significantly higher serum and synovial fluid levels of MIF than patients carrying two copies of the alternate G allele. Allele C carrying patients also required longer duration of treatment with glucocorticoids, showed a shorter clinical response to intra-articular injections, and had a significantly higher number of joints with active arthritis, and with limited range of motion (de Benedetti et al., 2003). These results indicate that not only is *MIF* -173 associated with susceptibility to sJIA, but that it is also a predictor of more severe disease.

A 32bp deletion in the chemokine receptor gene *CCR5* results in expression of a non-functional receptor. There have been conflicting reports of the significance of this Δ32 variant in the development and severity of RA (Garred et al., 1998; Cooke et al., 1998; Gomez-Reino et al., 1999; Zapico et al., 2000; Pokorny et al., 2005). A meta-analysis of all these studies showed a significant negative association, suggesting that CCR5Δ32 is protective against RA (Prahalad, 2006). It has also been shown in two independent studies to be associated with JIA as a whole, although one found it to be protective against disease (Prahalad et al., 2006) while the other found it to be a risk factor (Scheibel et al., 2008). In this study one of the most significant frequency differences was found with sJIA, with the Δ32 variant present in the patient cohort at a frequency of 0.25, compared to only 0.038 in healthy controls. This finding is contradictory to the results of the other studies, and is counterintuitive as CCR5Δ32 would result in a reduced inflammatory response. It is therefore unclear what role this variation plays in the development of sJIA.

In case-control and TDT tests in two independent cohorts, both association and linkage was shown between sJIA and *TPSN*, encoding the endoplasmic reticulum chaperone tapasin which is involved in antigen processing (Bukulmez et al., 2005). Diastrophic dysplasia, a disorder of bone and cartilage development, is caused by loss-of-function mutations in *SLC26A2*, solute carrier family 26. It was shown that homozygosity for six SNPs across the

gene locus is significantly associated with susceptibility to sJIA, but not with any of the other subtypes (Lamb et al., 2007).

As well as genetic associations with susceptibility to disease development, polymorphisms have also been identified as being associated with the development the severe complications seen with sJIA. The MAS phenotype is similar to familial hemophagocytic lymphohistiocytosis (HLH). HLH is associated with mutations in *MUNC13-4* which is involved in the cytolytic secretory pathway. Polymorphisms in this gene have been shown to be significantly more common in sJIA patients who have developed MAS than in healthy controls, or sJIA patients with no history of MAS (Zhang et al., 2008b). Further evidence for this association is that Munc13-4 has been shown to be differentially expressed in sJIA patients with sub-clinical MAS (Fall et al., 2007). A polymorphism in the 5' promoter region of the serum amyloid P component gene is associated with the development of systemic amyloidosis. A greater proportion of patients with amyolidosis were homozygous for this polymorphism than patients without amyloidosis. None of the patients with amyloidosis were homozygous for the alternate allele, whereas there was no difference in allele distribution between patients without amyloidosis and healthy controls (Woo P et al 1987).

All sJIA disease association studies published to date are candidate gene studies with small patient cohorts. Because of the small cohort sizes these studies have limited power and therefore some may be erroneous. Replication in independent populations is needed before any disease association can be established with certainty.

1.6. sJIA is distinct from other JIA subtypes

Although sJIA is classed as a subtype of JIA it is clinically distinct from the other subtypes. Features distinguishing sJIA from the other subtypes include differences in: sex ratio incidence, clinical presentation, pattern of joint disease, response to treatments, and the association of MAS, which is very rarely seen with the other subtypes (Woo and Colbert, 2009). Expression profiles identified in patients with sJIA are also significantly different from patients with other subtypes (Ramanan and Grom, 2005; Barnes et al., 2009). It is therefore important that sJIA is considered separately to the other subtypes as these marked distinctions indicate differences in the underlying pathogenesis (Adams and Lehman, 2005).

Due to sJIA sharing more clinical and gene expression profile similarities to a number of auto-inflammatory diseases (Allantaz et al., 2007; Ogilvie et al., 2007), including MWS, CINCA, FMF, many clinicians and researchers consider that, rather than with the other JIA subtypes, it is more appropriate to classify sJIA as an auto-inflammatory disease (Woo and Colbert, 2009).

1.7. Interleukin 1 Gene Family

A number of nomenclature assignments have been given to members of the IL-1 family. The protein and gene symbols given in section 8 of the glossary will be used throughout this thesis.

1.7.1. Activity

Interleukin 1 (IL-1) is a highly inflammatory cytokine and a potent pyrogen (Dinarello et al., 1986; Zheng et al., 1995). IL-1 elicits an inflammatory response through up-regulating expression of the transcription factor NFκB (Shirakawa and Mizel, 1989). NFκB upregulates the expression of a large number of immune response cytokines, including IL-1, thereby creating an autoregulatory feed back loop (Kopp and Ghosh, 1995).

Effects of IL-1 include osteoblast activation (Lin et al., 2010) and cartilage degradation (Isaev et al., 1992), and increased expression of adhesion molecules (Kaiserlian et al., 1991) resulting in infiltration of lymphocytes to the affected site (Miossec et al., 1986). IL-1 also results in activation of T-cells (Galy et al., 1990; Wesa and Galy, 2002; Ben Sasson et al., 2009), natural killer (NK) cells (Ben Aribia et al., 1987; Voiculescu et al., 1988), dendritic cells (Wesa and Galy, 2002; Guo et al., 2003), and the profileration of fibroblasts (Kohase et al., 1987). IL-1 is also a potent inducer of the acute phase response (Ramadori et al., 1985; Zheng et al., 1995; Josephs et al., 2000).

IL-1 affects nearly every cell type, including neutrophils (Forsyth and Levinsky, 1990), macrophages (Hanazawa et al., 1988), NK cells (Herman et al., 1985), dendritic cells (Jonuleit et al., 1996), T cells (Igarashi et al., 1990), endothelial cells (Ching et al., 2007), and hepatic cells (Bevan and Raynes, 1991).

1.7.2. IL-1 Ligands

There are two classical forms of IL-1, IL-1α and IL-1β, encoded by distinct genes. IL-1α acts

mainly as a regulator of intracellular events and a mediator of local inflammation, while IL-1β is released from cells and acts as a systemic mediator of inflammation (Dinarello, 1996).

1.7.2.1. IL-1α

IL-1α is synthesised as a precursor called proIL-1α, which is fully biologically active but remains intracellularly in the cytosol with no significant accumulation in any organelle. Pro-IL-1α is only found in the circulation during severe disease following release from apoptotic cells (Chen et al., 2007). In the cell proIL-1α is thought to act as an autocrine growth factor, particularly in epithelial and ectodermal cells, regulating cellular differentiation. proIL-1α is constitutively expressed in normal human skin (Hauser et al., 1986), and senescent skin cells contain high levels of IL-1α mRNA. Blockade with antisense oligonucleotides to the IL-1α transcript prevents cell senescence and extends the proliferative life span of skin cells (Maier et al., 1990).

Cytosolic proIL-1α can be myristoylated and transported to the cell surface, becoming membrane IL-1 (Stevenson et al., 1993). Membrane IL-1 is found on the surface of several cells, but particularly on stimulated monocytes and B lymphocytes (Kurt-Jones et al., 1985). Although membrane IL-1 is biologically active (Beuscher and Colten, 1988; Niki et al., 2004) it represents no more than five per cent of the total proIL-1α synthesised (Dinarello, 1996).

Following cleavage by activated calpains, calcium-dependant, membrane associated cysteine proteases, pro-IL-1α is converted into mature IL-1α and released from the cell, inducing an inflammatory response (Kobayashi et al., 1990; Watanabe and Kobayashi, 1994).

1.7.2.2. IL-1β

IL-1β is also synthesised as a precursor, proIL-1β, which remains in the cytosol. Unlike pro-IL-1α, proIL-1β but is only marginally active and has no membrane-bound form (Mosley et al., 1987a; Mosley et al., 1987b). Cleavage of proIL-1β by the IL-1β converting enzyme caspase 1 yields mature, active IL-1β (Cerretti et al., 1992; Thornberry et al., 1992; Miller et al., 1993) which is secreted from the cell. Although generation of mature IL-1β in response to endotoxin is dependent on proIL-β cleavage by caspase 1, during tissue necrosis proIL-1β can be cleaved by other proteases (Fantuzzi et al., 1997).

1.7.3. IL-1Ra

In addition to IL-1α and IL-1β there is a naturally occurring IL-1 receptor antagonist, IL-1Ra (Arend et al., 1985; Seckinger et al., 1987; Hannum et al., 1990; Carter et al., 1990). IL-1Ra binds to the IL-1 receptors, preventing the IL-1 ligands from binding to the receptor, but does not induce a signalling response (Dripps et al., 1991). No signal is induced because IL-1Ra contains only one of the two binding sites with which IL-1α and IL-1β bind to the receptor (Evans et al., 1995), so no conformational change is induced, and IL-1RAcP (section 1.7.4.3) is not recruited to the complex (Greenfeder et al., 1995). IL-1Ra binds to IL-1R1 (section 1.7.4.1) with approximately the same affinity as IL-1α and IL-1β (Hannum et al., 1990), but binds to the non-signalling IL-1R2 (section 1.7.4.2) with a 100-500 fold lower affinity than for IL-1R1 (Granowitz et al., 1991). An intracellular form of IL-1Ra, which lacks a leader sequence and remains intracellularly (Andersson et al., 1992), is constitutively produced in keratinocytes and epithelial cells, where it may block the binding of IL-1α to nuclear DNA (Haskill et al., 1991; Dinarello, 1997).

1.7.4. IL-1 Receptor Complex

The classical IL-1 receptors, which interact with IL-1α and IL-1β, include two IL-1 receptors and a receptor accessory protein. These three IL-1 receptors share significant extracellular domain homology, are members of the immunoglobulin superfamily, and are comprised of three IgG like domains (Sims et al., 1988; Savage et al., 1989; McMahan et al., 1991; Greenfeder et al., 1995).

The interactions between the IL-1 ligands and the classical IL-1 receptors is shown in Figure 1.2.

1.7.4.1. IL-1 R1

The type-1 IL-1 receptor (IL-1R1) has a single transmembrane segment, and a cytoplasmic domain with no apparent intrinsic tyrosine kinase activity. IL-1R1 is found predominantly on endothelial cells, smooth muscle cells, epithelial cells, hepatocytes, fibroblasts, keratinocytes, epidermal dendritic cells and T lymphocytes (Dinarello, 1997). The extracellular domain of IL-1R1 contains three Ig-like domains which bind to the two receptor binding domains of IL-1 ligands (Vigers et al., 1997). Binding of IL-1 to the first two binding domains of IL-1R1 induces a conformational change, enabling contact of the third IL-1R1 domain with the

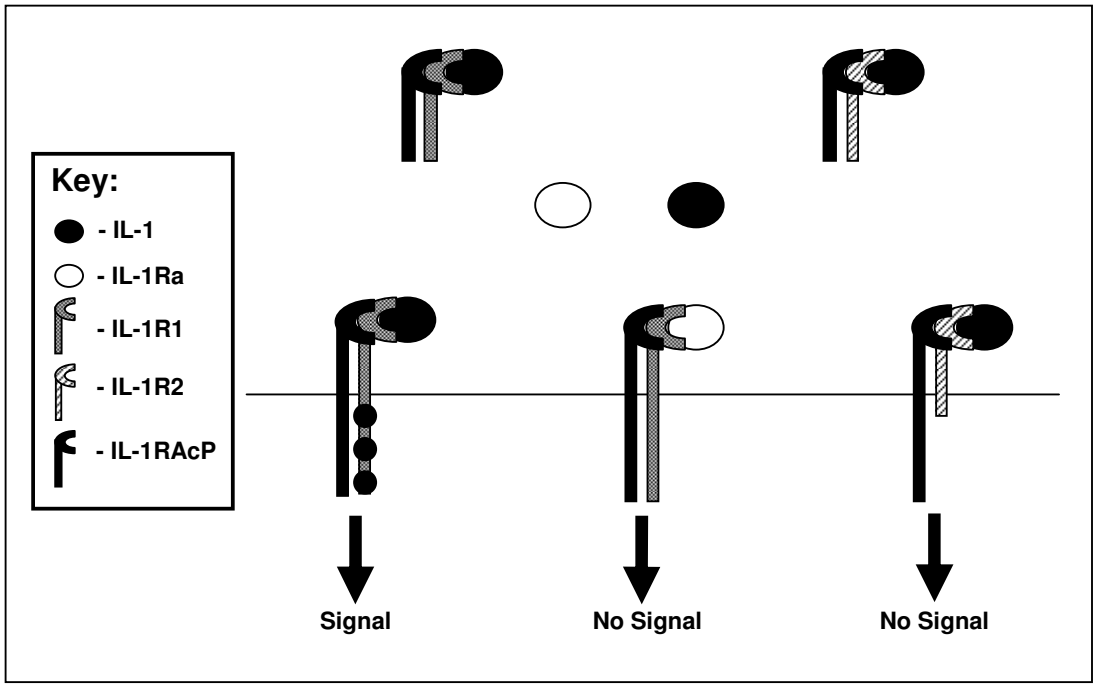


Figure 1.2 IL-1 ligand-receptor interactions

IL-1 ligands and the IL-1 receptor antagonist (IL-1Ra) are both able to bind to the IL-1 receptors. IL-1Ra occupies the binding site, preventing IL-1 ligands from binding, but does not induce signalling.Two different IL-1 receptors are present on the cell membrane. The type-1 IL-1 receptor (IL-1R1) is a functional, signal inducing receptor, while the type-2 IL-1 receptor (IL-1R2) has a short cytosolic domain and acts as a non-signalling, decoy receptor. A co-receptor (IL-1RAcP) does not bind to IL-1 directly, but is recruited to the IL-1 ligand/receptor complex and is required for signal induction. Isoforms of all three IL-1 receptors lacking the trans-membrane domain, generated either through alternative splicing or by enzymatic cleavage of the membrane bound protein, are secreted from cells. These secreted isoforms bind to circulating IL-1 ligands and remove them from circulation.

second binding site on IL-1. This structural change in IL-1R1 allows docking of the IL-1 receptor accessory protein (IL-1RAcP, section 1.7.4.3) with the IL-1R1/IL-1 complex, forming an IL-1RAcP/IL-1R1/IL-1 complex. It is this complex of ligand, receptor, and co-receptor which is required for signal transduction (Greenfeder et al., 1995). This signalling complex then recruits MyD88 and the IL-1 receptor activating kinase (IRAK), leading to phosphorylation of IRAK and, in turn, other kinases, including inhibitory-κ B, with subsequent translocation of NF-κB to the nucleus where it initiates gene expression of other cytokines (Figure 1.3).

Enzymatic cleavage of the membrane-bound protein yields a secreted form of IL-1R1 (sIL-1R1) which lacks the signal transducing transmembrane domain. sIL-1R1 binds to IL-1, removing it from the circulation and acting as a response antagonist (Symons et al., 1991).

1.7.4.2. IL-1R2

The type 2 IL-1 receptor IL-1R2 has a short cytosolic domain and is non-signal inducing (Colotta et al., 1993). By sequestering IL-1 ligands, preventing them from binding to the signalling IL-1R1, IL-1R2 acts as a decoy receptor (Stylianou et al., 1992; Sims et al., 1993; Sims et al., 1994). IL-1R2 is predominantly expressed on lymphoid and myeloid cells, including monocytes, neutrophils, bone marrow cells,and B cells (McMahan et al., 1991).

A secreted isoform of IL-1R2 (sIL-1R2) is expressed through alternative splicing. Like the membrane bound isoform, sIL-1R2 acts as an antagonist, binding both IL-1α and IL-1β. Unlike the memebrane bound form however, it also binds pro-IL1β, preventing it from being cleaved into the mature, active form by caspase-1. sIL-1R2 is also able to bind IL-1Ra, but with low affinity (Symons et al., 1995). Through interaction with sIL-1RAcP (section 1.7.4.3), the affinity with which sIL-1R2 binds IL-1 is enhanced, forming a high affinity IL-1 scavenger (Smith et al., 2003).

1.7.4.3. IL-1RAcP

IL-1RAcP does not bind IL-1 (Greenfeder et al., 1995; Wesche et al., 1998) but is a co-receptor essential for IL-1R1 signalling (Wesche et al., 1997b; Korherr et al., 1997; Huang et al., 1997; Cullinan et al., 1998) (section 1.7.4.1, and Figure 1.3).

A secreted form of IL-1RAcP (sIL-1RAcP), which lacks the transmembrane and intracellular

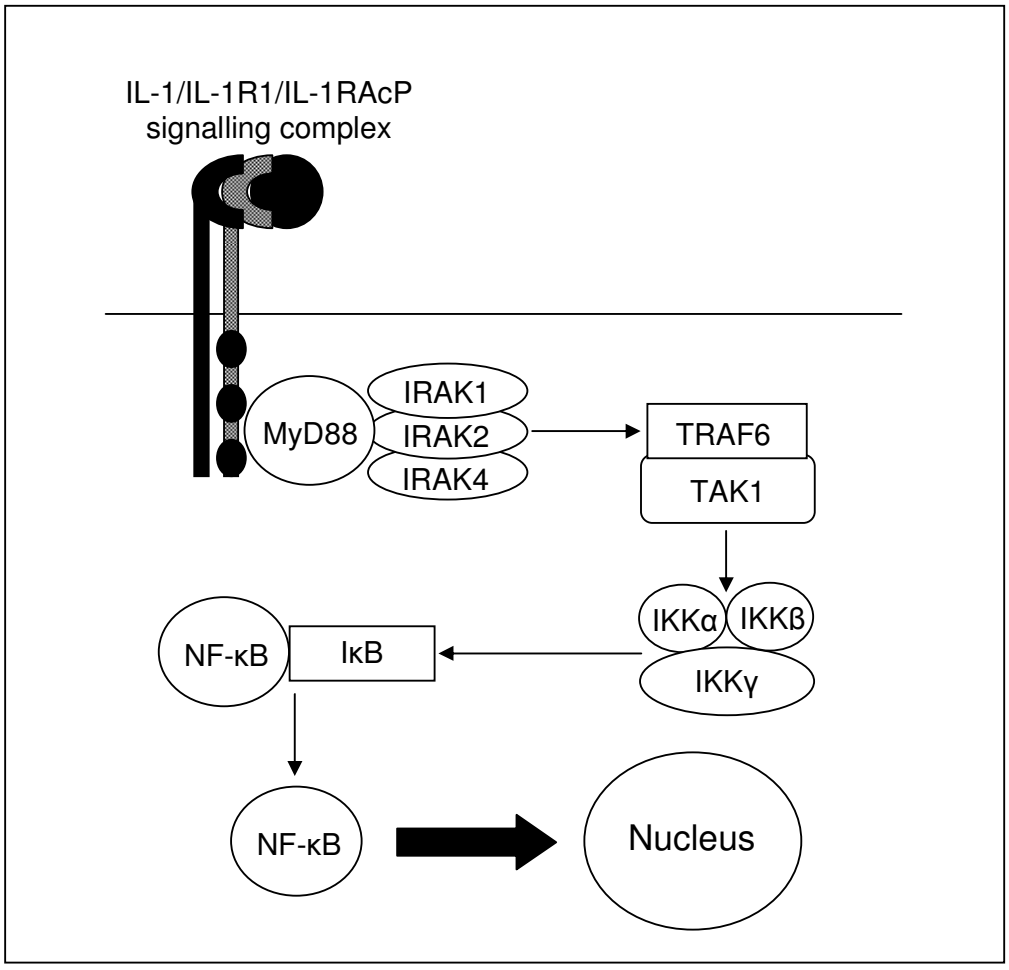


Figure 1.3 IL-1 signal transduction

The cytoplasmic domains of both IL-1R1 and IL-1RAcP share significant homology with the Drosophila receptor Toll, which is involved in development and innate immune responses. These domains are called Toll-IL1 Receptor-like (TIR) domains (Auron, 1998). In the IL-1/IL-1R1/IL-1RAcP signalling complex the TIR domains of the two receptors are brought together, enabling recruitment of Myeloid differentiation factor 88 (MyD88) by homophilic interaction with its TIR domain (Wesche et al., 1997a; Burns et al., 1998). IL-1 receptor-associated Kinase (IRAK) 1, IRAK2, and IRAK4 then associate with MyD88 via a death domain interaction (Muzio et al., 1997; Ringwood and Li, 2008). IRAK1 and IRAK2 are phosphorylated by IRAK4, and then mediate oligomerisation of the adapter molecule Tumor necrosis factor receptor associated factor 6 (TRAF6) (Cao et al., 1996), which in turn recruits transforming growth factor activated kinase 1 (TAK1) to the complex (Wang et al., 2001). Following recruitment to the signalling cascade, TAK1 phosphorylates the IKK complex of IKKα, IKKβ, and IKKγ (Yu et al., 2008). These in turn then phosphorylate IκB, resulting in proteolysis of IκB and the release of NFκB, which translocates to the nucleus and triggers gene expression (Scheidereit, 2006).

(adapted from (Loiarro et al., 2010))

domains of the membrane-bound form (Jensen and Whitehead, 2003a), is generated through alternative splicing. sIL-1RAcP is biologically active as an inhibitor of IL-1 *in vitro* by markedly reducing NF-κB activation induced by IL-1, but not by TNFα. Over-expression of sIL-1RAcP markedly ameliorates collagen-induced arthritis (CIA) in mice (Smeets et al., 2003). This inhibitory ability of sIL-1RAcP is through interaction with sIL-1R2 (section 1.7.4.2), forming a high affinity IL-1 scavenger. The sIL-1RAcP/sIL-1R2 complex has an approximately 100-fold increased affinity for IL-1α and IL-1β compared to sIL-1R2 alone, but no increased affinity for IL-1Ra (Smith et al., 2003).

As the membrane-bound and secreted isoforms of IL-1RAcP have opposing effects on the IL-1 response, the balance of expression of the two isoforms will impact on overall IL-1 activity levels. It has been shown that in HepG2 cells, during an inflammatory response to phorbol esters, and under various stress conditions, mRNA expression of sIL-1RAcP increases relative to that of mIL-1RAcP (Jensen et al., 2000b; Jensen and Whitehead, 2003b).

1.7.5. IL-18

IL-18, formerly known as IFN-γ-inducing factor, is a member of the IL-1 gene family with a similar structure to IL-1β. Like IL-1β, IL-18 is transcribed as a precursor, proIL-18, which requires cleavage by caspase-1 to become its mature, active form (Ghayur et al., 1997). IL-18 is constitutively expressed by PBMCs, with the main source being monocytes/macrophages and keratinocytes. IL-18 production is stimulated by macrophage stimulators including LPS, exotoxins, and various microbial products (Dinarello, 1999). The major function of IL-18, through activation of NF-κB, is to induce IFNγ production (Nakamura et al., 1993a; Okamura et al., 1995a; Ushio et al., 1996c). IL-18 alone only induces low levels of IFN-γ and requires a co-stimulant, including LPS and IL-12, to significantly enhance IFN-γ production (Munder et al., 1998). IL-18 acts synergistically with IL-12 by IL-12 inducing the expression of IL-18R1 (Yoshimoto et al., 1998). IL-18 has also been reported to induce synthesis of IL-2, granulocyte macrophage colony-stimulating factor, and TNF-α from T cells, and to suppress the production of IL-10 (Ushio et al., 1996a).

1.7.5.1. IL-18 Binding Protein

IL-18 binding protein (IL-18BP) is a secreted, naturally occurring IL-18 inhibitor, which binds directly to IL-18, thereby preventing it from binding to its receptor. IL-18BP

specifically inhibits IL-18 activity *in vitro* by blocking the production of IFN-γ in a dose-dependent manner. Mice injected with IL-18BP and LPS have undetectable levels of IFN-γ in their splenocytes compared to those injected with LPS alone (Novick et al., 1999; Aizawa et al., 1999). IL-18BP is constitutively expressed in mononuclear cells and has been shown to be significantly elevated in serum during sepsis (Novick et al., 2001). Expression of IL-18BP is indirectly regulated by IL-18, as IFNγ, which is induced by IL-18, induces IL-18BP (Muhl et al., 2000a; Paulukat et al., 2001a; Veenstra et al., 2002). As well as four other regulatory elements, two IFN-γ-induced transcription factor response elements, including a silencer element involved in transcriptional regulation, have been identified within the promoter region of IL-18BP (Hurgin et al., 2002).

Four different isoforms of IL-18BP (IL-18BPa-d) are generated through alternative splicing. In four human cDNA librares screened, IL-18BPa was the most abundant of the isoforms, while IL-18BPb was found only in monocyte and Jurkat libraries, IL-18BPc only in Jurkat and spleen libraries, and IL-18BPd in Jurkat libraries only (Novick et al., 1999). As shown by IFNγ production from PBMCs and a natural killer cell line, only IL-18BP isoforms A and C are able to inhibit the biological activity of IL-18, with IL-18BPa being the most efficient. In a BIAcore sensor chip assay, IL-18BPb and IL-18BPd, which lack a complete Ig domain, did not bind to IL-18. IL-18BPa and IL-18BPc bound only to mature IL-18, and not proIL-18. IL-18BPc had an approximately 10-fold lower affinity for IL-18 than IL-18BPa (Kim et al., 2000).

Because only two of the four isoforms of IL-18BP are able to bind and inhibit IL-18 with high affinity, the proportions in which they are expressed will affect the levels of active IL-18, and in turn the inflammatory response.

1.7.5.2. IL-18 Receptor Complex

IL-18 induces a signalling cascade in a similar manner to IL-1β, through a receptor complex of a binding receptor (IL-18R1) (Torigoe et al., 1997), and a signalling, non-ligand binding co-receptor (IL-18RAP) (Born et al., 1998). IL-18R1 is a member of the IL-1 receptor gene family, and IL-18RAP is related to IL-1RAcP.

1.7.6. Other IL-1 Family Members

1.7.6.1. Ligands

In addition to the classical IL-1 ligands IL-1α, IL-1β, and IL-18, a number of other genes belonging to the IL-1 family have been discovered (Mulero et al., 1999; Smith et al., 2000; Kumar et al., 2000; Busfield et al., 2000; Pan et al., 2001; Lin et al., 2001; Schmitz et al., 2005). The nomenclature of IL-1F5-IL-1F11, numerically assigned according to publication date, (Sims et al., 2001; Schmitz et al., 2005) will be used in this thesis. These new family members have similar genomic structures to each other and to the classic IL-1 ligands (Taylor et al., 2002). Like IL-1α and IL-1β none contain a signal peptide, and only IL-1F7 and IL-11 contain a pro-domain and are cleaved by caspase-1 (Kumar et al., 2002; Schmitz et al., 2005).

Through NF-κB-dependent reporter assays, it was shown that IL-1F6, IL-1F8 and IL-1F9 are able to initiate the signalling pathway leading to NF-κB activation in various cell types. By also transiently co-transfecting, in various combinations, plasmids containing different IL-1 receptor genes, it was shown that IL-1F6, IL-1F8, and IL-1F9 induced NF-κB activation is dependent on the formerly orphan receptor IL-1RL2, and the promiscuous co-receptor IL-1RAcP (Debets et al., 2001; Towne et al., 2004). However, a number of studies have been unable to demonstrate binding of these ligands to any IL-1 family receptors or co-receptors (Smith et al., 2000; Towne et al., 2004; Kumar et al., 2002). Conversely, while there is no evidence that IL-1F10 is able to act as an agonist (Towne et al., 2004), and at the time of writing has no known function (Dinarello et al., 2010), it has been demonstrated that it is able to bind to sIL-1R1, although with a lower affinity than IL-1Ra and IL-1β (Lin et al., 2001).

IL-1F5 has the greatest amino acid homology (50%) with IL-1Ra, including at a number of cystine residues which are evolutionarity conserved in IL-1Ra (Mulero et al., 1999). It was shown that IL-1F5 does not induce NF-κB activation, but does specifically antagonise the response induced by IL-1F9, but not that induced by IL-1α. This antagonism was found to be very potent; approximately 50% inhibition achieved with IL-1F5 at similar or lower concentrations than IL-1F9, whereas a 1000-fold excess of IL-1Ra is required to give the same level of inhibition of IL-1α (Debets et al., 2001; Towne et al., 2004). It has been shown that in glial cells, IL-F5 attenuates IL-1 and LPS induced inflammatory responses through interaction with SIGIRR (section 1.7.6.2) and induction of IL-4, but has no anti-inflammatory effect in macrophages and dendritic cells (Costelloe et al., 2008).

Alternative splicing generates five different isoforms of IL-1F7, of which IL-1F7b is the most abundant. IL-1F7 is constitutively expressed in testis, uterus, thymus, and monocytes, and can be induced in PBMCs and dendritic cells (Pan et al., 2001). Both pro- and mature IL-1F7 bind to IL-18R1 (section 1.7.5.2) but do not induce IFN-γ production as there is no recruitment of IL-18RAP (section 1.7.5.2). They also do not inhibit IL-18 induced IFNγ production, suggesting that IL-18R1 preferentially binds IL-18 (Pan et al., 2001; Kumar et al., 2002). IL-1F7b, does however also bind to IL-18BP (section 1.7.5.1), and increases its ability to antagonise IL-18 (Bufler et al., 2002). Following cleavage by caspase-1, mature IL-1F7b translocates to the nucleus, where it has been demonstrated in macrophages to reduce LPS-induced expression of TNFα, IL-1α, and IL-6 (Sharma et al., 2008). It has also been shown to reduce the expression of TNFα, IL-1α, IL-1β, and IL-8 in undifferentiated and differentiated THP-1 cells, and in epithelial cells stimulated with LPS or IL-1β, with no effect on expression of the anti-inflammatory cytokines IL-1Ra and IL-10 (Nold et al., 2010).

IL-11 (also known as IL-33) is expressed as a precursor and requires cleavage by caspase-1 into the mature form. It induces NF-κB phosphorylation and MAP kinase activation through IL-1RL1 (Schmitz et al., 2005) and IL-1RAcP (Ali et al., 2007). Th1, but not Th2, cells are responsive to IL-11, upregulating expression of IL-5 and IL-13 (Schmitz et al., 2005).

IL-1F5, IL-1F6, IL-1F7, IL-1F8, and IL-1F9 are all expressed by monocytes, and upregulated following stimulation with LPS. IL-1F5, IL-1F6, IL-1F7, and IL-1F8 are also expressed in B cells, and IL-1F6 in T cells too. (Kumar et al., 2000; Smith et al., 2000; Lin et al., 2001). IL-1F6 has been shown to act as pro-inflammatory cytokine in murine models of skin inflammation (Blumberg et al., 2007), and IL-1F9 in mouse models of lung inflammation (Ramadas et al., 2010; Chustz et al., 2010).

1.7.6.2. Receptors

In addition to those previously discussed (section 1.7.4), a number of proteins, defined by Ig-like domains in the extracellular segments involved in protein-ligand and protein-protein interactions, have been identified as belonging to the IL-1 receptor family. Despite minimal amino acid homology, the Ig-like domains in the IL-1 receptor family proteins all contain two β-pleated sheets conected with intradomain disulfide bonds, formed by conserved cysteine residues located at similar positions (Subramaniam et al., 2004).

IL-1RL1 (IL-1 receptor-like 1) does not bind any of the IL-1 ligands. It has however been demonstrated, through a fusion protein of the intracellular domain of IL-1RL1 and the extracellular domain of IL-1R1, that IL-1RL1 is capable of signal transduction (Mitcham et al., 1996). Both membrane-bound, and secreted isoforms of IL-1RL1 are expressed (Li et al., 2000). The membrane bound isoform is predominantly expressed on hematopoetic cells, and the secreted isoform by fibroblasts and mast cells (Bergers et al., 1994). IL-1RL1 is constitutively expressed by Th2 cells, and is upregulated in Th2 and natural killer cells but not Th1 cells, during inflammation (Lohning et al., 1998). IL-1RL1 has been shown to be effective in limiting airway hyper-reactivity in a model of allergic asthma (Coyle et al., 1999).

Like IL-1RL1, IL-1RL2 (IL-1 receptor-like 2) is also incapable of binding IL-1 ligands, despite having a similar structure to IL-1R1 (Lovenberg et al., 1996). However, IL-1F6, IL-1F8, and IL-1F9 (section 1.7.6.1), have all been shown to induce NF-κB activation through IL-1RL2 and IL-1RAcP (Debets et al., 2001; Towne et al., 2004). IL-1RL2 is expressed in human lung epithelium, brain vasculature, keratinocyte skin cells, fibroblasts, endothelial cells and monocytes (Lovenberg et al., 1996; Debets et al., 2001).

Two members of the IL-1 receptor gene family indentified through sequence homology, are TIGIRR1 (three immunolglobulin domain-containing IL-1 receptor related protein), which has homology to IL-1R1, and TIGIRR2, which has homology to IL-1RAcP (Born et al., 2000). Neither TIGIRR protein is able to bind IL-1 ligands (Smith et al., 2000; Sana et al., 2000), or to induce a response when the intracellular domain is fused to the extracellular domains of IL-1R1 or IL-1RAcP (Born et al., 2000). SIGIRR (single immunoglobulin IL-1R related protein), which was identified as containing an IL-1R family signalling domain, is also not able to bind IL-1 ligands, or induce a response when the intracellular domain is fused to the extracellular domin of IL-1R1 (Thomassen et al., 1999).

1.7.7. IL-1 associated proteins

As mentioned previously, there are a number of proteins which are not members of the IL-1 gene family, but are integral to IL-1 activity.

1.7.7.1. Caspase-1

Caspase-1 is required to cleave IL-1β (Cerretti et al., 1992; Thornberry et al., 1992; Miller et al., 1993), IL-18 (Ghayur et al., 1997; Gu et al., 1997; Fantuzzi et al., 1998), IL-1F7 (Kumar et al., 2002), and IL-33 (Kumar et al., 2002; Schmitz et al., 2005) into their mature, active forms. Caspase-1 is itself synthesised as a precursor which undergos two internal cleavages to become an enzymatically active heterodimer. These internal cleavages take place through autoprocessing, where the caspase-1 precursor undergoes oligomerisation with itself or homologs (Wilson et al., 1994a; Gu et al., 1995). In order to cleave pro-IL-1β, two molecules of the caspase-1 heterodimer form a tetramer with two molecules of pro-IL-1β (Wilson et al., 1994b; Walker et al., 1994).

Caspase-1 deficient mice are overtly normal but do not secrete mature IL-1β following LPS stimulation, and are resistant to endotoxic shock (Li et al., 1995a; Kuida et al., 1995), and ischemic acute renal failure (Melnikov et al., 2001). In splenocytes from mice either deficient for caspase-1 or treated with a caspase-1 specific inhibitor, IL-12 induced IFNγ production is reduced, the same effect as is seen in the presence of antibodies to IL-18. Despite constitutive expression of pro-IL-18, administration of IL-12 increases circulating levels of IL-18 in wild type mice, but not in mice deficient for caspase-1 (Fantuzzi et al., 1999). In a model of chronic arthritis, caspase-1 knockout mice have reduced joint inflammation and less cartilage damage than wild type mice (Joosten et al., 2009).

1.7.7.2. Cryopyrin

Cryopyrin, also known as NALP3 (NACHT-, leucine-rich repeat (LRR)- and PYD-containing protein 3), is part of the NALP3 or IL-1β inflammasome, which is required for activation of caspase-1. The NALP3 inflammasome is a multi-protein complex comprised of NALP3, adapter protein ASC, CARD containing protein cardinal, and caspase-1. It promotes dimerisation or oligomirisation of caspase-1, leading to formation of an active enzyme (Agostini et al., 2004b).

Cryopyrin belongs to family of NOD-like receptors, which sense pathogen-associated molecular patterns (PAMPs) (Kanneganti et al., 2006a; Kanneganti et al., 2006b; Mariathasan et al., 2006) and nonmicrobial danger or damage patterns, including monosodium urate crystals or calcium pyrophosphate dihydrate (Martinon et al., 2006), heat shock proteins (Mayor et al., 2007), and adenosine-5'-triphosphate (Mariathasan et al., 2006).

1.7.8. IL-1 Family Gene Organisation

The majority of the IL-1 family candidate genes included in this thesis, are located in two clusters on chromosome 2 of the human genome (Figure 1.4). Eight of the IL-1 ligand genes, and the gene encoding IL-1Ra (*IL1RN*), are all positioned within a genomic region spanning 0.4Mb (Nicklin et al., 2002; Taylor et al., 2002). Except for *IL1RAP*, the genes for all of the investigated IL-1 receptor family members, are also located together in a gene cluster (Dale and Nicklin, 1999; Smith et al., 2004). In this thesis the two clusters are referred to as the *IL1* ligand cluster and the *IL1* receptor cluster.

1.8. IL-1 in disease

There is evidence that relatives of patients with sJIA have a higher incidence of other autoimmune diseases compared to the general population. It is believed that there may be similarities in the underlying etiopathologies of diseases which share clinical features. There may be genetic variants which predispose to autoimmunity, and are therefore common to a number of autoimmune diseases, and others which, in combination with environmental factors, determine the exact phenotype of each individual (Simmonds and Gough, 2004; Woo and Colbert, 2009). Evidence regarding the etiopathology of other immune-related diseases can therefore inform on possible candidates in sJIA (Phelan et al., 2006).

As well as being effective in treating a proportion of patients with sJIA, the recombinant IL-1Ra, Anakinra, also gives rapid and dramatic improvement in other diseases which are clinically similar to sJIA. This includes adult-onset Still's disease (AOSD), the adult-onset form of sJIA (Fitzgerald et al., 2005), Muckle-Wells syndrome (MWS) (Hawkins et al., 2004), neonatal-onset multisystem inflammatory disease (Lovell and Bowyer, 2003), and familial cold auto-inflammatory syndrome (FCAS) (Hoffman et al., 2004). Gain of function mutations in *NALP3* have also been identified in patients with MWS, FCAS, and chronic infantile neurological cutaneous and articular syndrome (Aganna et al., 2002; Feldmann et al., 2002; Dode et al., 2002; Hoffman et al., 2001). PBMCs from patients with FCAS show normal cytokine expression under basal conditions, and of pro-IL-1β following LPS stimulation, but higher secretion of mature IL-β and IL-18 following LPS-stimulation than cells from healthy controls. On treatment with a caspase-1 inhibitor this secretion is blocked with equal potentcy in cells from patients and healthy controls (Stack et al., 2005).

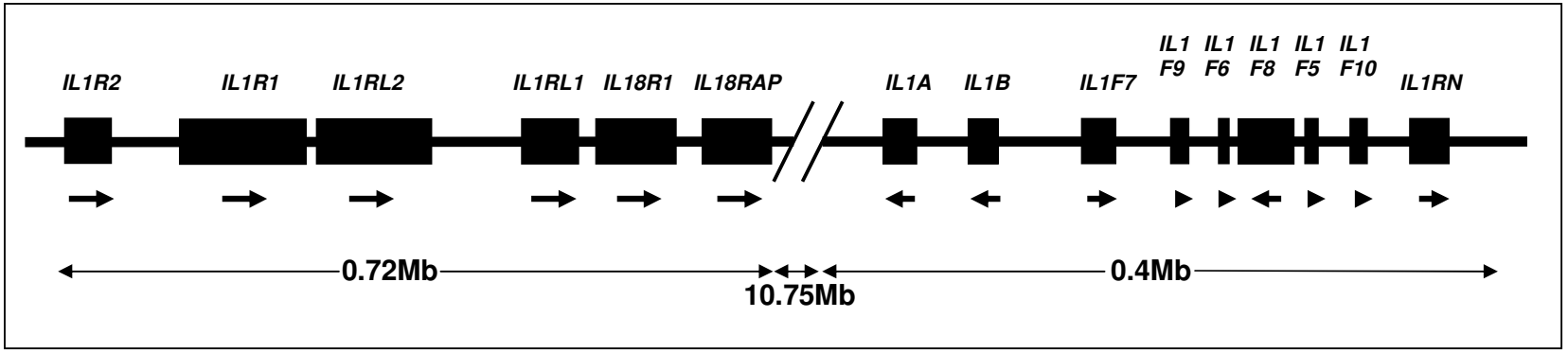


Figure 1.4 Organisation of the two IL-1 gene family clusters

Shown are the order, orientation, and relative sizes of the IL-1 gene family members located in the two clusters on chromosome 2. The two clusters are referred to in this thesis as the *IL1* receptor cluster and the *IL1* ligand cluster. The *IL1* receptor cluster spans 0.72Mb and contains six IL-1 receptor family genes. The *IL1* ligand cluster spans 0.4Mb and contains nine genes belonging to the IL-1 family. The two clusters are 10.75Mb apart on chromosome 2 of the human genome.

IL-1α has been shown to be involved in cartilage degradation (Isaev et al., 1992). In a mouse model of arthritis, levels of membrane-bound IL-1α, but not serum IL-1α or IL-1β, correlate with disease severity (Niki et al., 2004). Synoviocytes from affected joints of patients with RA have high levels of IL-1α mRNA (Buchan et al., 1988), and patients who produce anti-IL-1α antibodies develop a less destructive disease (Graudal et al., 2002). Compared to wild type mice, mice deficient for IL-1α and IL-1β, either individually or in combination, have a significantly decreased number of osteoclasts in trabelcular bone, in combination with increased femur mineral density, trabecular bone mass, and cortical thickness (Lee et al., 2010). Patients with adult polyarthritis have higher plasma levels of IL-1Ra, sIL-1R1, and sIL-1R2 than healthy controls, with the level of IL-1Ra positively correlating with measures of disease activity and joint destruction. The level of sIL-1R2 is higher in patients with non-destructive arthritis than in patients with destructive disease, and correlates negatively with indices of joint destruction (Jouvenne et al., 1998a). In collagen-induced mouse models of arthritis, overexpression of sIL-1RAcP significantly ameliorates disease activity (Smeets et al., 2003).

Psoriasis is an immune mediated skin disease characterised by inflammation and altered epidermal differentiation, leading to redness and scaling. In lesional psoriasis skin, expression of IL-1F5, IL-1F9, and IL-1RL2 is significantly increased compared to healthy skin (Debets et al., 2001). Overexpression of IL-1F6 in mouse skin results in a hyperproliferative inflammatory skin condition in newborn animals. This phenotype is dependent on the presence of IL-1RL2 and IL-1RAcP, and is exacerbated by deficiency of the antagonist IL-1F5 (Blumberg et al., 2007). In immunodeficient mice on to which human psoriatic skin has been transplanted, treatment with antibodies to IL-1RL2 resolves the inflammatory changes in the psoriatic skin (Blumberg et al., 2010).

IL-18 and IL-18BP have been implicated as having a role in a number of inflammatory and autoimmune diseases. IL-18 is present at increased levels in intestinal tissues affected with Crohn's disease (Pizarro et al., 1999; Monteleone et al., 1999) and, in colitis mouse models, treatment with an IL-18 antibody elicits a dose-dependent reduction in disease severity (Siegmund et al., 2001a). Administration of IL-18BP (Ten Hove et al., 2001) and deficiency for caspase-1 (Siegmund et al., 2001c) also reduces disease severity in mouse models of colitis. Administration of IL-18BP or antibodies to IL-18 (Plater-Zyberk et al., 2001), and IL-18 deficiency (Wei et al., 2001), significantly suppresses the effects of collagen-induced

arthritis, while administration of IL-18 increases the erosive and inflammatory components of the disease (Gracie et al., 1999a). Levels of IL-18 are increased, and IL-18BP decreased, in patients with RA compared to healthy controls, with higher levels of IL-18 and lower levels of IL-18BP in patients with active disease compared to those with inactive disease (Sivalingam et al., 2007). Levels of IL-18 are also higher in patients with RA compared to those with psoriatic arthritis (Bresnihan et al., 2002) and osteoarthritis (Gracie et al., 1999b). Elevated levels of IL-18 with no corresponding increase in IL-18BP, is found in patients with lupus nephritis (Liang et al., 2006) and active idiopathic thrombocytopenic purpura (Shan et al., 2009). This is also seen in patients with secondary hemophagocytic syndrome, characterised by uncontrolled activation of Th1 cells and macrophages, in which IL-18 levels correlate with clinical status and biological markers of disease (Mazodier et al., 2005).

1.9. IL-1 Gene Family Disease Associations

A number of polymorphisms in IL-1 family genes have been shown to be associaited with various inflammatory and autoimmune diseases.

A number of studies have shown significant association of SNPs in IL-1 genes with RA. A SNP in *IL1A* which introduces an amino acid substitution making proIL-1α more susceptible to protease cleavage, is associated with RA (Kawaguchi et al., 2007), as are a number of SNPs within *IL1B* (Buchs et al., 2001; Harrison et al., 2008). A SNP in *IL1A* is also associated with the severity of chronic polyarthritis (Jouvenne et al., 1999). An *IL18* haplotype (Gracie et al., 2005), and a number of *IL1RN* variants (Luotola et al., 2010) have also been associated with susceptibility to RA. As well as being significantly associated with RA, SNPs in *IL1F10* and *IL1RN* are also associated with levels of CRP, ESR, and bone erosion in patients from a Korean population (Jung et al., 2010). Although significant associations with RA have been demonstrated in a number of studies, other studies have been unable to replicate these results. A meta-analysis of eighteen published studies of IL-1 polymorphisms in RA, found evidence of a significant association of a SNP in *IL1B*, but only in the Asian population (Lee et al., 2009). A separate meta-analysis investigating fifteen studies of an *IL1RN* variable number tandem repeat (VNTR), found no significant disease association over all (Xue et al., 2010).

Associations of IL-1 polymorphisms with osteoarthritis (OA) have also been demonstrated, including linkage to the *IL1* receptor gene cluster (section 1.7.8) in a genome wide

association study (GWAS) (Nakki et al., 2010). SNPs in *IL1B* and *IL1RN* are associated with hip OA (Meulenbelt et al., 2004), a haplotype of SNPs in *IL1A*, *IL1B*, and *IL1RN* with knee OA (Smith et al., 2004), and SNPs in *IL1R1* with hand OA (Nakki et al., 2010). As well as being associated with disease susceptibility, polymorphisms in *IL1RN* are also significantly correlated with radiographic scores of OA disease activity (Swellam et al., 2010).

The cluster on chromosome 2 containing the *IL1* ligand genes (section 1.7.8) has also been identified in a GWAS study as a candidate region for ankylosing spondylitis (AS) (Laval et al., 2001). Further studies have demonstrated significant association of polymorphisms in *IL1B, IL1F6*, *IL1F8*, and *IL1F10* (Timms et al., 2004), and in *IL1B*, *IL1F10* and *IL1RN* (Chou et al., 2006; van der et al., 2002) with susceptibility to AS. In a Chinese Han population significant disease association, and a weak association with clinical manifestation, were identified with SNPs in *IL1F10*, *IL1B*, and *IL1RN* (Guo et al., 2010). A large study of 2,675 patients from ten countries investigated nine polymorphisms in *IL1A*, *IL1B*, *IL1F10*, and *IL1RN*, previously published as significantly associated with AS. Significant disease association of polymorphisms in *IL1A* was confirmed, with no significant heterogeneity between the centers involved (Sims et al., 2008). As well as associations with the *IL1* ligand cluster, a number of SNPs in the genomic region containing *IL1R2* have also been shown to be significantly associated with AS in a GWAS study (Reveille et al., 2010).

Polymorphisms in *IL18* have been associated with type 1 diabetes, in which elevated levels of IL-18 are found in the subclinical stage (Kretowski et al., 2002; Ide et al., 2004), coronary artery disease (Tiret et al., 2005a), systemic lupus eythematosus (SLE) (Xu et al., 2007b), Crohn's disease (CD) (Tamura et al., 2002; Aizawa et al., 2005), and procititis-type ulcerative colitis (UC) (Takagawa et al., 2005). A SNP in *IL18RAP* has also been shown to be associated with CD and UC (Zhernakova et al., 2008), a result supported by the Wellcome Trust Case Control Consortium study (Wellcome Trust Case Control Consortium, 2007). Other *IL1* family associations include SNPs in *IL1RN* and *IL1F5* with susceptibility and severity of alopecia areata (Tazi-Ahnini et al., 2002), *IL1A, IL1B*, and *IL1RN* with periodontitis (Kinane and Hart, 2003), *IL1R1* with type 1 diabetes (Bergholdt et al., 2000), and *IL1A*, *IL1B*, *IL1F7*, *IL1F8*, and *IL1F10* with psoriatic arthritis (Rahman et al., 2006).

Polymorphisms in members of the IL-1 gene family have also been associated with a wide variety of other diseases, including Graves' disease (Khalilzadeh et al., 2010; Liu et al., 2010), glaucoma (Mookherjee et al., 2010), aneurismal subarachnoid hemorrhage (Fontanella

et al., 2010), Tourette syndrome (Chou et al., 2010), schizophrenia (Xu and He, 2010), breast cancer (Han et al., 2010), sudden infant death syndrome (Ferrante et al., 2010), polycystic ovary syndrome (Yang et al., 2010), and opioid and alcohol dependence (Liu et al., 2009a).

A VNTR in *IL1RN* has been shown to be associated with JIA. This association was also significant in the subtypes of extended oligoarthritis, enthesitis-related arthritis, and other arthritis, but not sJIA, when the subtypes were examined separately (Vencovsky et al., 2001). An *IL1A* promoter polymorphism is associated with early-onset oligoarticular arthritis. This association was especially strong in patients who developed chronic iridocyclitis, and was also associated with elevated ESR (McDowell et al., 1995). This subtype-specific association was not, however, replicated in subsequent studies (Donn et al., 1999b; Donn et al., 2001a). A 13 marker haplotype in the *IL18* promoter region has been associated with JIA in a Japanese population. When the subtypes were analysed separately this association remained significant for the oligoarticular, but not the systemic, subtype (Sugiura et al., 2006). This haplotype has also been found to be significantly associated with susceptibility to AOSD (Sugiura et al., 2002; Sugiura et al., 2006), in which patients have significantly elevated IL-18 serum levels (Kawaguchi et al., 2001). Homozygous mutations of *IL1RN* were identified in all nine screened patients with neonatal onset of sterile multifocus osteomylelitis, periostitis, and pustulosis, a systemic inflammatory disease with skin and bone involvement. The mutations resulted in expression of a truncated protein which was not secreted from the cell. All of the patients treated with Anakinra showed a rapid response (Aksentijevich et al., 2009).

1.10. IL-1 in sJIA

As previously mentioned in section 1.4, there is evidence that patients with sJIA have an altered IL-1 cytokine profile, including increased expression levels of IL-1β, IL-1Ra (Madson et al., 1994; de Benedetti et al., 1995; Muller et al., 1998; Lotito et al., 2007), IL-18 (Maeno et al., 2002), and a defect in IL-18RAP phosphorylation (de Jager et al., 2009). The most compelling evidence for the involvement of IL-1 in the pathogenesis of sJIA is the dramatic improvement in both clinical symptoms and laboratory markers of disease activity seen in some patients following treatment with IL-1 blockade (section 1.2.3) (Verbsky and White, 2004; Henrickson, 2004; Irigoyen et al., 2004; Pascual et al., 2005; Quartier et al., 2006; Lequerre et al., 2008; Gattorno et al., 2008). A number of studies have investigated *IL1*

gene family polymorphisms in the systemic subtype of JIA, but none have demonstrated significant associations (McDowell et al., 1995; Vencovsky et al., 2001; Donn et al., 2001a; Sugiura et al., 2006). However, the initial aim of these studies was to investigate genetic involvement in JIA as a whole. Due to the low prevalence of the systemic subtype these studies included insufficient numbers of sJIA patients for adequate power to detect significant associations when analysed as a separate group. Additionally, most previous association studies examining *IL1* members have only investingated limited polymorphic diversity in one or two of the family members. The difficulty with this approach is that even if no significant association is detected the gene cannot be eliminated as a candidate for disease susceptibility, as it is possible that any causative polymorphism was not represented in the investigation. To fully investigate the contribution of *IL1* genetic variation in susceptibility to sJIA it is therefore important that complete screens of *IL1* gene polymorphisms are performed in studies specifically investigating the systemic subtype.

1.11. Hypothesis of the project

The hypothesis behind this thesis is that the IL-1 inflammatory response plays an important role in the pathogenesis of sJIA. That there is dysregulation of the IL-1 response towards a maintained pro-inflammatory state, and that this dysregulation occurs at the genetic level.

1.12. Aims of the project

The aims of this project are:

- To peform a case-control association study of 20 genes involved in the IL-1 response, with susceptibility to sJIA

- To perform a thorough association screen of polymorphisms in the candidate genes

- To investigate the biological implications of any polymorphisms found to be associated with sJIA

Chapter 2

Materials

2. Materials

2.1. Patient and Control Samples

In this study the term Caucasian is used to describe individuals of European descent and is used throughout.

All patient and healthy control DNA samples used in the association study were collected as part of previous or on-going projects in Professor Woo's group and collaborating groups. Ethical approval for the study was obtained (Great Ormond Street Hospital for Children NHS Trust and Institute of Child Health, Research Ethics Committee (reference 02RU06). All individuals or, when appropriate, parents gave informed consent.

2.1.1. Healthy controls

Stage-1: 151 individuals from Caucasian 16-30 year olds from a GP practice in a stable population of the west Midlands.

Stage-2: 184 Caucasian first-time blood donors from the national blood transfusion centre in London.

Additional healthy controls: In order to create a database of individuals from whom it would be possible to collect repeated blood samples DNA samples from 53 healthy individuals working in the Division of Infection and Immunity (coded as Windeyer Control 'WC' samples) were collected from blood or saliva samples as part of this project. All participants gave informed written consent (Research Ethics Committee reference: UCLH 02/0188). All whole blood samples for PBMC extraction were from individuals in this healthy control set.

2.1.2. Patients

All patients had confirmed diagnoses according to the International league against Rheumatism (ILAR) criteria for ILAR1 (systemic JIA).

Stage-1: 137 Caucasian sJIA patients collected from the British Society for Paediatric and Adolescent Rheumatology (BSPAR) National DNA repository at the Arthritis Research UK

Epidemiology Unit in Manchester and the Outpatients Departments at both Great Ormond Street Hospital and the Middlesex Hospital in London.

Stage-2: 105 Caucasian patients from Great Ormond Street Hospital, Wexham Park Hospital, Berkshire UK, and Necker Hospital in Paris.

Both patient cohorts had an approximately 1:1 ratio of females to males, 49.6% female in stage-1 (65 females, 66 males), and 55.4% female in stage-2 (56 females, 45 males, sex unavailable for one patient). The patient cohorts also had a similar distribution of age of disease onset (Figure 2.1), although the frequency of patients in whom the disease manifested at two years of age was different, 15.7% and 6.12% in stages 1 and 2 respectively. There was no age of onset data available for six of the stage-1 patients and four of the stage- 2 patients.

2.2. Laboratory reagents

2.2.1. Chemicals

Acrylamide: 30% Acrylamide/Bis solution (Avrylamid:N,N'-Methylenbisacrylamid) 19:1 (5% C) (#161-0154 , Bio-Rad)
Agarose: electrophoresis grade (#A9539, Sigma-Aldrich)
Ammonium persulfate: (#A-7460, Sigma-Aldrich)
Ampicillin: dessicated (#A-9518, Sigma-Aldrich)
Bacto-tryptone: (#T9410, Sigma-Aldrich)
Boric acid: (#15583-024, GIBCO BRL)
Chloroform: (#100776B, VWR BDH)
Diothiothreitol (DTT): (#43815, Sigma-Aldrich)
Ethanol: (#S7-16, Sigma-Aldrich-Aldrich)
Ethidium Bromide: 10mg/ml solution (#E1510, Sigma-Aldrich)
Ethylenediaminetetra-acetic acid (EDTA): (#E-5391, Sigma-Aldrich)
Glacial acetic acid: (#100015N, VWR BDH)
Glycerol: (#G5516, Sigma-Aldrich)
Isopropanol (propan-2-ol): (#20842.323, VWR BDH)
Methanol: (#101586B, VWR BDH)
Potassium hydroxide: (#10210, VWR BDH)
Sodium Chloride: (#S3014, Sigma-Aldrich)

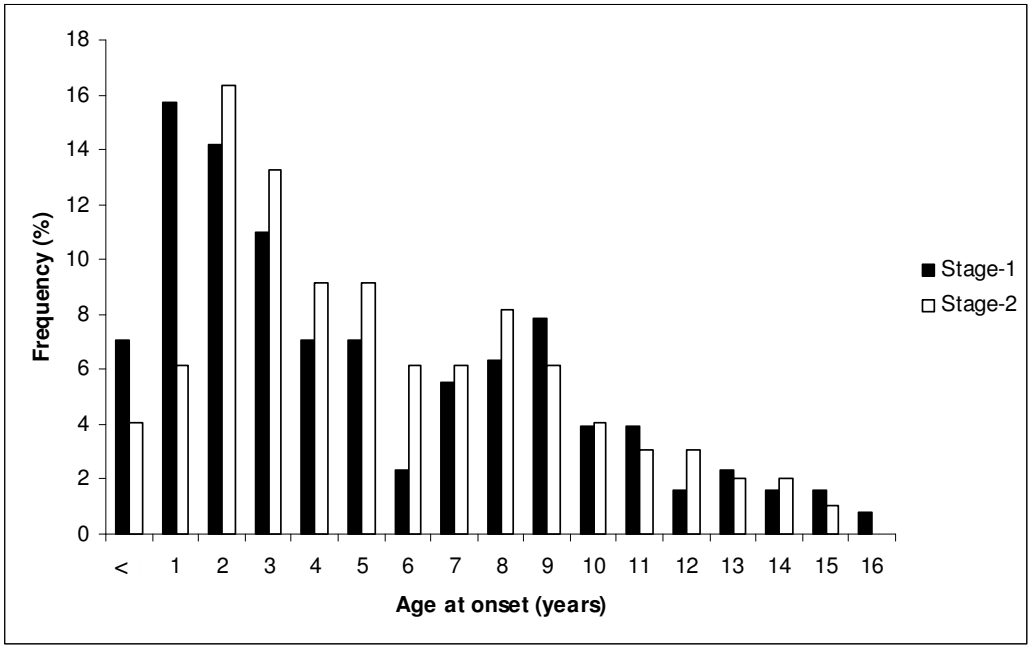


Figure 2.1 Age of disease onset of patients

The age of disease onset of the patients used in the two stages of the association study are shown as the frequency within each cohort.

Sodium hydroxide: (#102525P, VWR BDH)

Sulphuric acid: (#87003-272, VWR BDH)

NNN'N'-Tetramethylethylenediamine (TMED): (#443083G, BDH Electran)

Phosphate Buffered Saline (PBS): 10x Dulbecco's saline 0.0095M (PO_4), without Ca and Mg (#BE17-515F, Wittaker)

Polyoxyethylensorbitan monolaurate (Tween-20): (#P9416, Sigma-Aldrich)

Tris base: Trizma® base (#T4661, Sigma-Aldrich)

Tris-HCl: (#T3253, Sigma-Aldrich)

Yeast extract: (#T1625, Sigma-Aldrich)

2.2.2. Buffers

In all cases buffers were made up with reverse osmosis purified water. Where necessary, buffers were autoclaved, or sterilised by filtration through a 0.2µ filter before use.

Tris-acetate-EDTA (TAE): 50x stock (per litre)

2M Tris base	242g
Glacial acetic acid	57.1ml
0.05M EDTA	100ml 0.5M EDTA

Tris-borate-EDTA (TBE): 10x stock (per litre)

0.89M Tris base	108g
0.89M Boric acid	55g
20mM EDTA	40ml 0.5M EDTA

Tris-EDTA (TE):

10mM Tris-HCl

1mM EDTA

Adjusted to pH 7.5 or 8.0 with 10M NaOH

Cytoplasmic lysis buffer:

10mM HEPES (pH7.9)

1.5mM MgCl2

10mM KCl

0.5mM DTT

1mM Vanadate

1 complete protease inhibitor tablet/50ml

Nuclear lysis buffer:

20mM HEPES (pH7.9)

1.5mM MgCl2

25% glycerol

420mM NaCl

0.2mM EDTA (pH8)

1mM DTT

1mM Vanadate

1 complete protease inhibitor tablet/10ml

Parker Buffer:

20mM Tris-HCl (pH7.5)

10mM $MgCl_2$

100mM NaCl

1mM EDTA (pH8)

20% glycerol

1mM DTT

2.2.3. Gels

Agarose gels:

Agarose gels of 1-3% (w/v) were made by dissolving the appropriate weight of agarose in either 50 or 100ml 1x TAE by boiling in a microwave. Once cooled to hand temperature ethidium bromide was added to a final concentration of 1µg/ml. The gel was cast in a horizontal gel tray with well formers. The gel volume used was dependent on the volume of sample to be electrophoresed.

Acrylamide gels:

4% non-denaturing TBE-polyacrylamide gel (50ml):

30% ready mixed acrylamide	6.7ml
10x TBE	2.5ml
TEMED	50µl

10% APS 250µl

s.v.50ml with water

The gel was poured between two glass plates separated by 2mm spacers and allowed to set for at least one hour at room temperature. The gel was either used immediately or stored at 4°C until use.

2.2.4. DNA extraction

From Blood: Gentra Puregene blood kit (#158389, QIAGEN)

From Saliva: Oragene® DNA (DNA Genotek)

2.2.5. PCR

PCR reagents:

PCR buffer: 500mM KCl, 100mM Tris-HCl pH8.5, 25mM $MgCl_2$, 2mM each dNTP

$MgCl_2$: 25mM solution (#A3511, Promega)

dNTPs: dNTP Mix 10mM each dNTP (#U1511, Promega)

DNA polymerase: GoTaq® DNA polymerase 5u/µl (#M3175, Promega)

Primers: 100µM stock solution (Sigma-Aldrich-Genosys)

HotStart PCR reagents: Hot Start HiFidelity DNA Polymerase kit (#202602, QIAGEN)

DNA visualisation reagents:

Loading dye: Blue/Orange 6x loading dye – 0.4% orange G, 0.03% bromophenol blue, 0.03% xylene cyanol FF, 15% Ficoll® 44, 10mM tris-HCl (pH7.5, 50mM EDTA (pH8.0) (#G190A, Promega)

Ladders:

1kb DNA Step ladder (#G694A, Promega)

Bench top 100bp DNA ladder (#G829B, Promega)

25bp DNA Step ladder (#G351A, Promega)

2.2.6. DNA purification

Gel extraction: QIAquick gel extraction kit (#28704, QIAGEN)

This protocol allows the purification of DNA fragments between 70bp and 10kb.

PCR purification: QIAquick PCR purification kit (#28106, QIAGEN)
This protocol allows the purification of DNA fragments between 100bp and 10kb.

Plasmid purification: QIAprep Spin Miniprep kit (#27106, QIAGEN)
Prior to DNA purification using a silica membrane bacterial cells are lysed under alkaline conditions and the lysates separated by centrifugation.

2.2.7. Cell culture

THP-1 cell line: (ATCC® Number TIB-202, LGC Standards)
RPMI 1640: with L-glutamine and sodium bicarbonate (#R8758, Sigma-Aldrich Aldrich)
DMEM: with 4500 mg/L glucose, L-glutamine and sodium bicarbonate, without sodium pyruvate (#D5796, Sigma-Aldrich Aldrich)
Fetal calf serum (FCS): GIBCO^{TM} Fetal Bovine Serum heat-inactivated (#12319018, Invitrogen)
Trypsin-EDTA: 0.5g porcine trypsin, 0.2g EDTA·4Na/L Hank's balanced salt solution (T3924, Sigma-Aldrich Aldrich)
Trypan blue: (#T-6146, Sigma-Aldrich)
Fume hood: Class II safety cabinet (Nuaire)
Incubator: DH autoflow $C0_2$ air jacketed incubator (Nuaire)

2.2.8. PBMC isolation

Lymphoprep (Axis-Shield, Norway)
Lymphoprep solution:
9.1% sodium diatrizoate, 5.7% polysaccharide, density 1.077 ± 0.001g/ml, osmolality 290 ± 15 mOsm

2.2.9. Bacterial culture

Luria-Bertani (LB) broth per litre:

Bacto-tryptone 10g

Yeast extract 5g
Sodium chloride 10g

Adjusted to pH 7.0 with 10M KOH

On use LB broth was aliquoted at the required volume using sterile technique and 100µg/ml ampicillin added.

LB agar: LB broth plus Bacto-agar (10g/l)
Agar was melted by boiling in a microwave. Once cooled to hand temperature 100µg/ml ampicillin was added and poured into culture plates at approximately 20ml per plate. Plates were left at room temperature to solidify and then either used immediately or stored at 4°C for up to one week.

Competent cells:
Subcloning Efficiency™ DH5α™ Competent *E.coli* Cells (#18265-017, Invitrogen)
Genotype:
F$^-$ϕ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *rec*A1 *end*A1 *hsd*R17(r_K^-,m_K^+) *pho*A*sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ$^-$

pUC19 Control DNA 100 pg/µl (#1136177, Invitrogen)

2.2.10. Cloning

Vector: pGL3-Basic vector (#E1751, Promega)
A map of the vector is shown in Figure 2.2.

Calf Intestinal Alkaline Phosphatase: 1u/µl (#M182A, Promega)
10x reaction buffer:
(1x) 50mM Tris-HCl (pH 9.3), 1mM $MgCl_2$, 0.1M $ZnCl_2$, 1mM spermidine

T4 DNA Ligase: 3u/µl (#M1801, Promega)
10x Ligase buffer:
300mM Tris-HCl (pH 7.8), 100mM $MgCl_2$, 100mM DTT, 10mM ATP

Restriction enzymes: 10u/µl
*Bgl*II (#R6081, Promega)

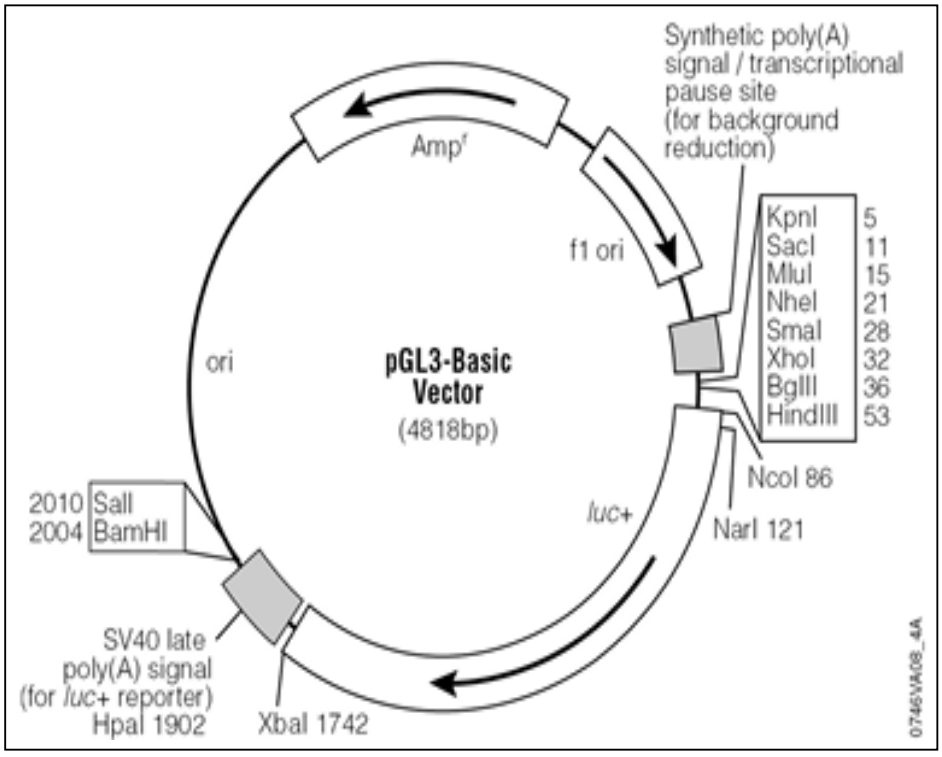


Figure 2.2 pGL3-Basic vector map

DNA sequence of interest was inserted into the pGL3-Basic vector for cloning using the restriction enzyme recognition sequences contained within the multiple cloning site of the vector.

*Mul*I (#R6381, Promega)
10x Reaction buffer D (#R9921, Promega)
(1x) 6mM Tris-HCl, 6mM $MgCl_2$, 150mM NaCl, 1mm DTT, pH7.9

2.2.11. RNA quantification

All techniques involving RNA were performed under good RNA handling practices. All plasticware was guaranteed RNase free, glassware was decontaminated using RNaseOUT™, and all reagents used were RNase-free.

RNA isolation: TRIZOL® (#15596-026, Invitrogen), Glycogen: (#10901393001, Roche)

Reverse Transcription: QuantiTect Reverse Transcription kit (#205313, QIAGEN)

qRT-PCR primer assays: All qRT-PCR primer assays were QuantiTect® Primer Assays (QIAGEN)
Reconstituted in 1.1ml TE
IL18BP (#QT00046564)
RPL32 (#QT01668198)
RPLP0 (#QT00075012)
GAPDH (#QT01192646)
ACTB (#QT01680476)

qRT-PCR: QuantiTect SYBR Green kit (# 204054, QIAGEN)

SYBR Green PCR master mix:
HotStarTaq *Plus* DNA Polymerase, QuantiFast SYBR Green PCR Buffer (containing Tris-HCl, KCl, NH_4CL, $MgCl_2$,Q-Bond (patent-pending additive which increases the affinity of *Taq* DNA polymerases for short-single stranded DNA)), dNTP mix, SYBR Green 1 fluorescent dye, ROX passive reference dye

2.2.12. ELISA

ELISA plates: EIA/RIA polystyrene high binding microplate (#3590, Corning)
Wash buffer: Phosphate buffered saline- (0.05%) Tween-20 (PBST)
Detection solution: 3,3',5,5'-tetramethylbenzidine (TMB) liquid substrate (#T8665, Sigma-Aldrich)

Stop solution: 2N H_2SO_4

Blocking solution/ Reagent diluent: 1% BSA solutions were made with sterile PBS and 0.2µ filtered prior to use.

Bovine Serum Albumin (BSA): Initial fractionation by heat shock fraction V minimum 98% (#A-7906, Sigma-Aldrich)

IL-18BPa Duoset (#DY119, R&D systems):

Capture antibody: Mouse anti-human IL18BPa (Part 841505)

Standard: Recombinant human IL-18BPa (Part 841507)

Detection antibody: Biotinylated goat anti-human IL-18BPa detection antibody (Part 841506)

Streptavidin: Streptavidin conjugated to horseradish-peroxidase (HRP) (Part 890803)

Plate washer- High Throughput System LP440 (Adil instruments)

P**late reader** – MRX TC Revelation v4.21 (Dynex Technologies)

2.2.13. EMSA

T4 polynucleotide kinase: 10u/µl with T4-kinase buffer (#M4101, Promega)

Radioactive label: ATP, [γ-^{32}P]-3000Ci/mmol 10mCi/ml EasyTide, activity 3.7MBq/10ml (NEG502A, PerkinElmer)

Probe purification columns: Illustra ProbeQuant G-50 Micro columns (#28-9034-08, GE Healthcare)

Poly(deoxyinosinic-deoxycytidylic) acid sodium salt (poly dI-dC): (#P4929-10UN, Sigma-Aldrich)

NP40: (#1747401, Roche)

Sp1 antibody: (#sc-59X, Santa Cruz Biotechnology Inc)

BCA assay: Bio-Rad Protein assay (#500-0006, Bio-RAD)

Radiographic film: Amersham Hyperfilm ^{TM} MP (#28906847, GE Healthcare)

Chapter 3

Methods

3. Methods

3.1. Association Study

A two-stage case-control association study was performed using tagging SNP selection.

3.1.1. Introduction

3.1.1.1. Tagging SNPs

The use of tagging SNPs (tSNPs) is a method to greatly reduce the number of polymorphisms required to be genotyped. This is achieved by selecting a minimal subset which act as proxies for the others, and capture all of the variation contained within the full set. Polymorphisms are able to act as proxies for each other through linkage disequilibrium (LD).

Loci on the same chromosome are in LD when they are inherited together more often than expected by chance following independent assortment. Genetic recombination during meiosis results in random-assortment of the parental variant combinations in the offspring. Theoretically, the physically further apart loci are on the chromosome, the greater the chance that recombination will have occurred. This would result in a steady decrease of LD over distance. However, recombination rates are not constant over the genome, with regions known as recombination hot spots, having much higher recombination rates than the rest of the genome. Regions of high LD separated by regions of no LD are seen across the genome, giving rise to a block-like pattern of LD (The International HapMap Consortium, 2005). Because there is not a consistent correlation between distance and LD it is therefore necessary for LD relationships between pairs of polymorphisms to be determined empirically.

The notation used in the following equations, with the common allele of a locus termed allele 1, and the rare allele as allele 2, is:

A = allele 1 of locus 1
a = allele 2 of locus 1

B = allele 1 of locus 2
b = allele 2 of locus 2

The combination of alleles at different loci on a single chromosome is termed a haplotype. The LD coefficient, D, is calculated between pairs of loci as the difference between the

observed haplotype frequency and, based on the allele frequencies, that expected following random assortment:

$$D_{AB} = f_{AB} - f_A \times f_B$$

f = frequency

A positive D value indicates that the haplotype is present in the population more often than expected by chance, and a negative D value, less often than expected. This value, however, is dependent on the allele frequency. An alternative measure is D′, which reflects the evidence that recombination has occurred between the loci investigated. Lower D′ values reflect more evidence that recombination has taken place between the loci.

$$D' = \begin{cases} \dfrac{-D}{D_{min}} & \text{When } D < 0 \\ \dfrac{D}{D_{max}} & \text{When } D > 0 \end{cases}$$

D_{min} = theoretical minimum for the observed allele frequencies
= the smaller of Ab or aB

D_{max} = theoretical maximum for the observed allele frequencies
= the larger of AB or ab

An alternative LD statistic is r^2, the square of the correlation coefficient:

$$r^2 = \frac{(f_{AB} - f_A \times f_B)^2}{f_A \times f_a \times f_B \times f_b}$$

r^2 measures the specific relationship between the alleles of two loci, i.e. how often the common alleles are present together at both loci. r^2 values range between 0, no LD (random assortment), and 1, absolute LD. If two loci have $r^2=1$ then only two of the four possible haplotypes are observed, i.e. either the common alleles are present at both loci or the rare alleles are present at both loci. When this is the case it would be redundant to genotype both loci, as genotyping one provides information on the allele present at both (G.McVean, 2007).

This relationship between polymorphisms can be utilised to select a subset of loci for genotyping. Only polymorphisms able to act as proxies for others, tSNPs, need to be investigated while those captured, tagged SNPs, are not investigated directly.

LD, and therefore tSNP selection, can be most accurately calculated when phase, the combination of alleles physically present together on the same chromosome, is known. When an individual is homozygous at one or both loci of interest, phase can be inferred with confidence as there is only one allele possible on each homologous chromosome. However, genotype data alone cannot be used to determine phase when an individual is heterozygous at both loci, as more than one haplotype combination is possible. If parental genotypes are also known this informs which haplotypes may be present. The recent generation of population genotyping databases has made this approach to genetic studies possible by providing reference panels for the calculation of LD patterns and tSNP selection.

The International human haplotype map (HapMap) Consortium (www.hapmap.ncbi.nlm.nih.gov), a collaboration of 19 research groups, was launched in 2002 to 'determine the common patterns of DNA sequence variation in the human genome and to make this information freely available in the public domain'. It is intended for use as a research tool in the identification of genetic factors involved in disease susceptibility (The International HapMap Consortium, 2003). Phase I of the project, published in 2005, encompassed approximately 1.3 million common (minor allele frequency (MAF) $\geq$0.05) SNPs distributed at approximately one SNP every 5kb across the genome. These SNPs were genotyped in 90 (30 father, mother, adult offspring trios) CEPH Utah residents of Northern and Western European descent samples (described below), designated as CEU. Three other ethnic reference groups were also included: Yoruba, Han Chinese, and Japanese. The genotyping was distributed by chromosome region across nine centres using six different technologies, employing standardised rules for data quality control. Prior to publication all data from the different collaborators underwent blinded quality assessment and was found to be >99% accurate (The International HapMap Consortium, 2005). An additional 2.1 million SNPs, in the same samples, were genotyped in the 2007 released Phase II of the project. The Phase I and II data together cover approximately 1 SNP every kb of the genome. This is estimated to be approximately 25-35% of all common SNPs in the human genome assembly (The International HapMap Consortium, 2007).

Programmes for Genomic Applications (PGA) (http://public.nhlbi.nih.gov/geneticsgenomics/home) was launched in 2000 by the National Heart, Lung and Blood Institute to '[develop] information, tools, and resources to link genes to biological function on a genomic scale'. All data generated from the constituent projects is publicly available. The SeattleSNPs project (http://.pga.gs.washington.edu) focused on genetic variants in the human inflammatory response. A variation discovery stage was performed in which a total of 327 candidate genes were sequenced in 23 unrelated CEPH samples (Utah residents). The InnateImmunity project (www.pharmgat.org/IIPGA2), investigating associations between genetic variants and airway diseases, also sequenced 83 candidate genes to identify common variable sites through sequencing 23 unrelated CEPH (Utah residents) samples.

Centre d'Etude du Polymorphisme Humain (CEPH) is a non profit research institute founded in 1983 (Dausset et al., 1990). Their aim was to compile a panel of reference families, and to make them available to the scientific community for the generation of a DNA polymorphism map. The premise was that the most efficient way to collaboratively construct a genetic map was for all researchers to use the same families, enabling data from different researchers to be complied together (Cann, 1992). The most informative samples for constructing genetic maps, due to the requirements of phase determination, are families with large sibships, and living parents and grandparents. Forty eight families in the CEPH reference panel are Utah residents of European descent. These families were selected based on them being extended pedigrees, and having no evidence of genetic diseases. The incidence of common diseases in these families are expected to be the same as in the general population (Williams et al., 1979; Botstein et al., 1980).

3.1.1.2. Two-stage study design

In a two-stage association study a proportion of the study samples are genotyped for all of the polymorphisms of interest in stage-1. In stage-2 of the study the remaining samples are then genotyped for only the polymorphisms showing the strongest evidence of frequency differences in stage-1. This study design reduces the amount of genotyping required, enabling more effective use of the resources available. It has been shown that genotyping 50% of the available samples for all markers of interest in stage-1, and 10% of the markers in stage-2, provides near optimal power, compared to a one-stage study design, while reducing the

amount of genotyping required by 45% (Satagopan et al., 2002). Performing joint analysis of the data from both stages also provides greater power than considering stage-2 as a replication study (Skol et al., 2006).

3.1.2. Stage-1

3.1.2.1. tSNP Selection

In order to capture regulatory regions in the flanking sequences, the regions selected for investigation were extended to include the surrounding sequences, both upstream and downstream to the next flanking genes. These were identified in the July 2003 human genome reference sequence (NCBI build 34), hg16 annotation track. Genotyping data for all available SNPs within each of the seven candidate regions were obtained from the International HapMap project (HapMap data release #18/phaseII Sept05, on the National Center for Biotechnology Information (NCBI) B34 assembly, dbSNPb124) for the CEU (CEPH Utah residents) population. Additional genotyping data for 16 of the candidate genes (excluding *IL1RL1, IL18RAP, CASP1* and *CIAS1*) in CEPH samples was also available from the SeattleSNPs and InnateImmunity PGA projects.

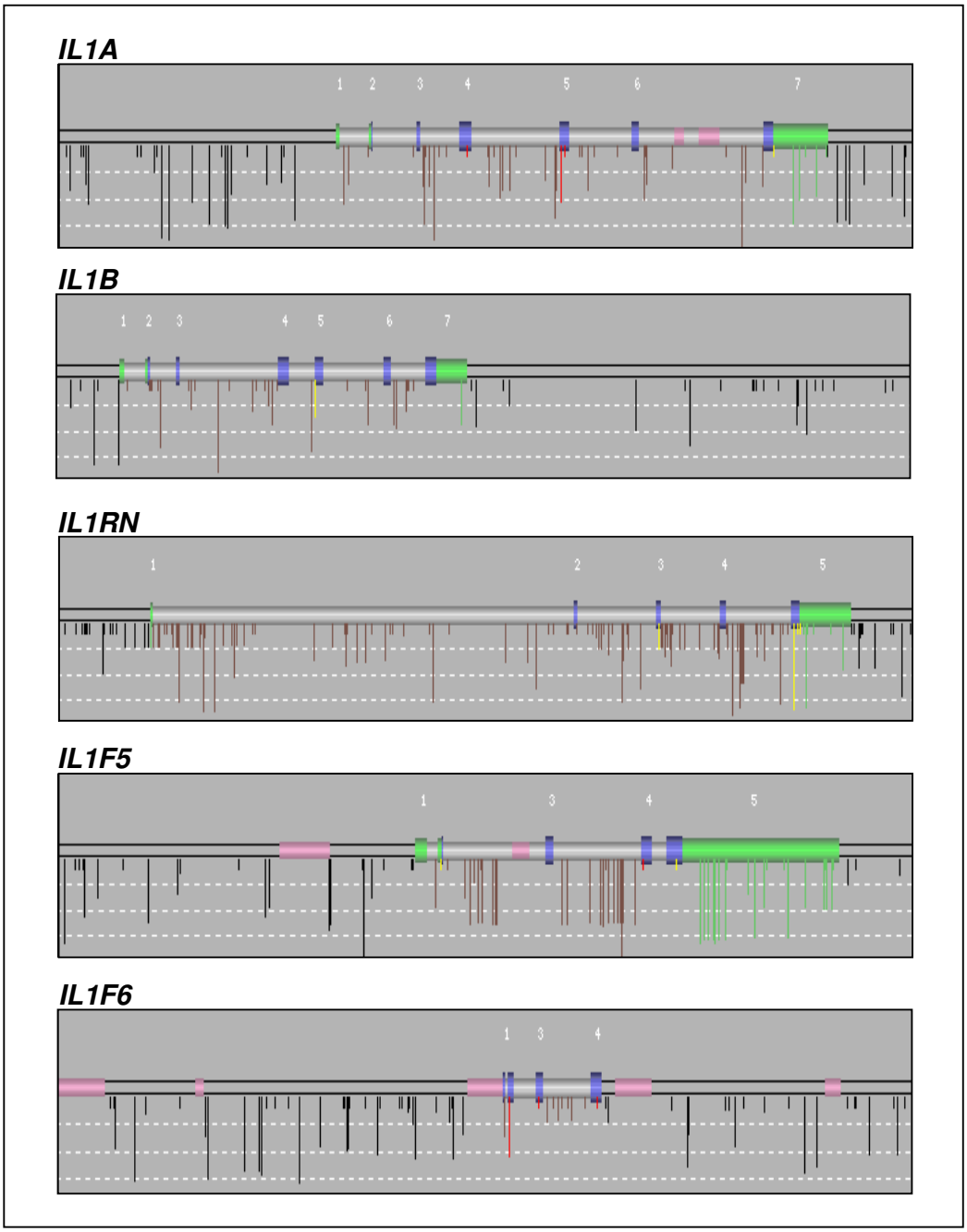
The regions sequenced for variation discovery in both PGA projects extended beyond the candidate gene. InnateImmunity sequenced 1,000bp either side of the gene, and SeattleSNPs to the next flanking gene (Table 3.1). Due to technical difficulties InnateImmunity did not scan some repetitive regions for polymorphisms (Figure 3.1). Both HapMap and PGA used CEPH reference panel samples. Nine of the 23 samples sequenced for *IL1A*, *IL1B*, *IL1R1*, *IL1R2*, *IL1RN*, *IL18R1*, *IL18*, and *IL18BP* were included in the HapMap set. All 23 samples sequenced for *IL1F5*, *IL1F6*, *IL1F7*, *IL1F8*, *IL1F9*, *IL1F10*, *IL1RL2* were genotyped by HapMap (Table 3.2). In the PGA databases polymorphisms were identified by their position within the sequenced region. The sequences surrounding the variants were BLAT aligned (BLAST Like Alignment Tool) (Kent, 2002) against the human genome to determine, where applicable, the assigned rs identifier. To ensure that relationships between polymorphisms included in only one of the databases could be investigated, the genotyping data from both sources were manually combined into one file with more complete SNP coverage.

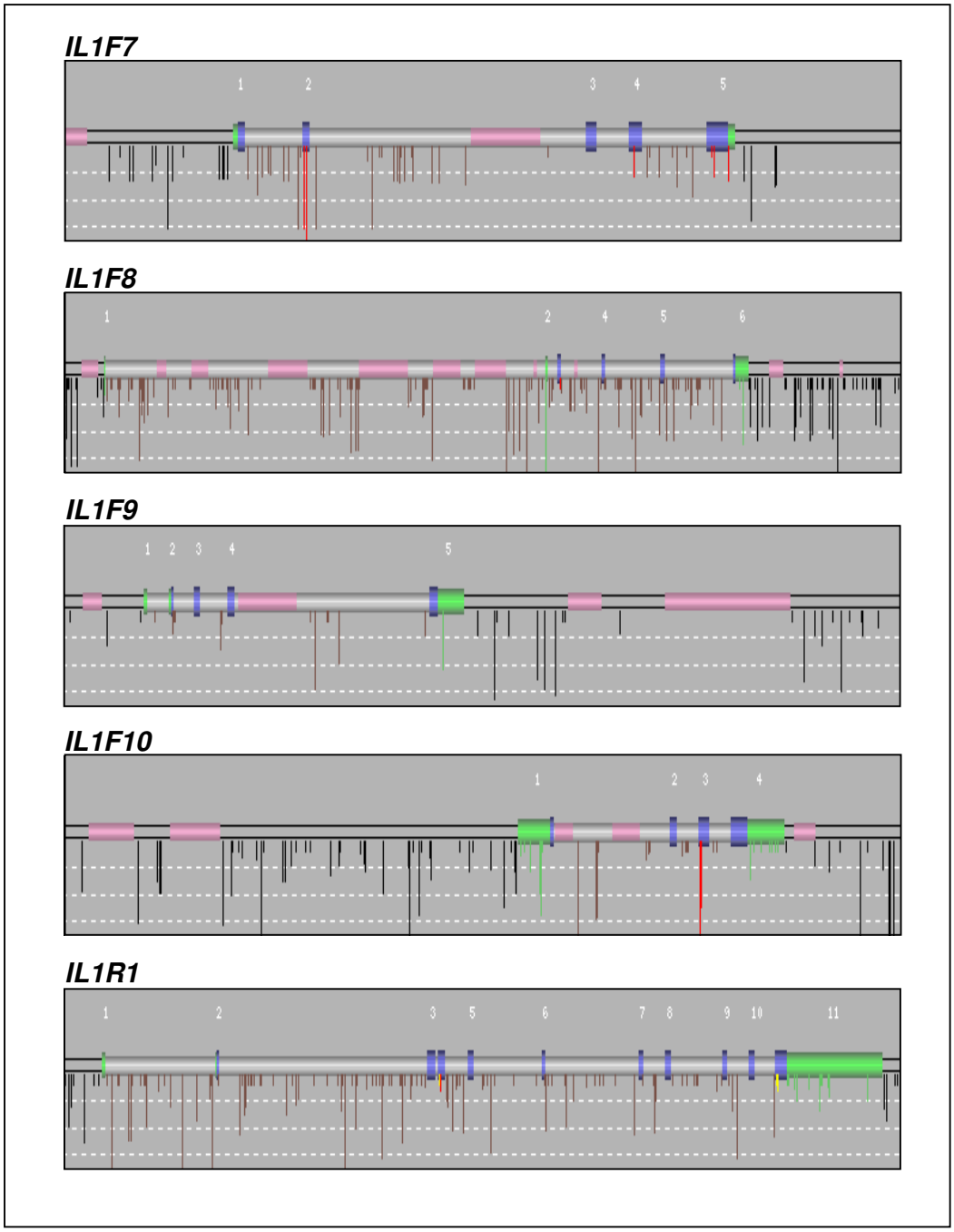
Computation of LD between each SNP pair was calculated using Haploview (Barrett et al., 2005) (version 3.2). This software employs a two marker Expectation-Maximisation

Gene	PGA project	Up-stream (bp)	Down-stream (bp)	Nº variants
IL1A	SeattleSNPs	6,028	1,836	50
IL1B	SeattleSNPs	1,305	9,117	35
IL1RN	SeattleSNPs	10,004	6,963	91
IL1F5	SeattleSNPs	4,587	1,066	54
IL1F6	SeattleSNPs	2,115	1,531	36
IL1F7	SeattleSNPs	1,739	1,968	43
IL1F8	SeattleSNPs	1,951	7,257	104
IL1F9	SeattleSNPs	1,937	10,440	17
IL1F10	SeattleSNPs	6,050	1,546	47
IL1R1	SeattleSNPs	1,295	633	85
IL1R2	SeattleSNPs	1,849	1,406	101
IL1RL2	SeattleSNPs	1,983	1,995	186
IL18	InnateImmunity	1,000	1,000	124
IL18BP	InnateImmunity	1,000	1,000	41
IL18R1	InnateImmunity	1,000	1,000	258

Table 3.1 Flanking sequences scanned for PGA variation discovery

The PGA SeattleSNPs and InnateImmunity projects performed variation discovery by sequencing candidate genes in CEPH individuals. The amount of flanking sequence (in bp) either side of the genes included in the sequenced regions are given. The number of variants identified in the CEPH samples is also shown. The distance up-stream indicated is relative from the start of transcription, and the distance down-stream relative to the end of the 3′UTR.


IL1A
1 2 3 4 5 6 7
IL1B
1 2 3 4 5 6 7
IL1RN
1 2 3 4 5
IL1F5
1 3 4 5
IL1F6
1 3 4


IL1F7
1
2
3
4
5
IL1F8
1
2
4
5
6
IL1F9
1
2
3
4
5
IL1F10
1
2
3
4
IL1R1
1
2
3
5
6
7
8
9
10
11

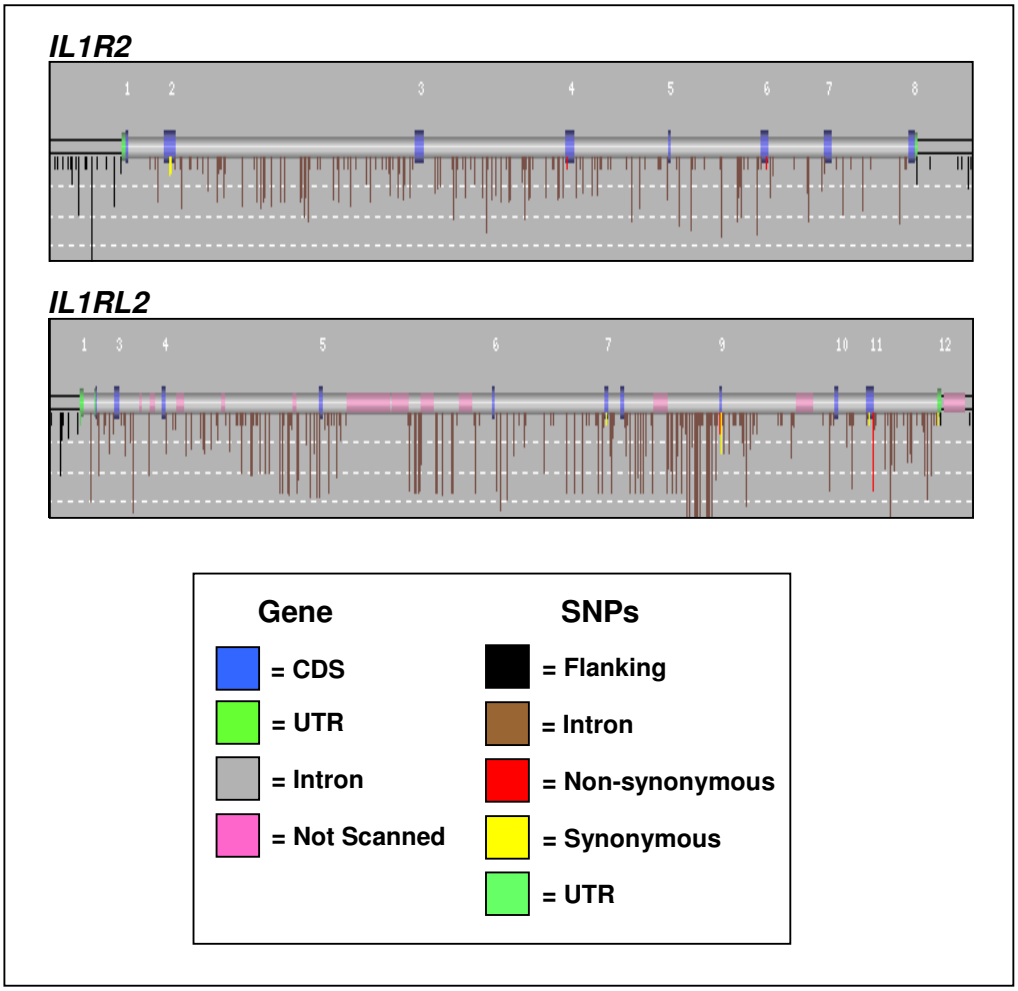


Figure 3.1 SeattleSNPs gene regions scanned for variation discovery

The PGA SeattleSNPs project performed variation discovery by sequencing candidate genes in CEPH individuals. The regions sequenced extended into the gene flanking regions. Shown are the regions sequenced with the gene features, and repetitive regions not scanned for polymorphisms, indicated. The vertical lines underneath the gene region indicate the position and frequency of each SNP.

Pictures adapted from SeattleSNPs.

Family ID	Individual ID	HapMap	Gene panel 1	Gene panel 2
1	NA12004	+	-	-
	NA12003	+	-	-
	NA10838	+	-	-
2	NA12006	+	-	-
	NA12005	+	-	-
	NA10839	+	-	-
3	NA12057	+	-	+
	NA12056	+	-	-
	NA10851	+	+	-
4	NA11830	+	-	-
	NA11829	+	-	-
	NA10856	+	-	-
5	NA11832	+	-	+
	NA11831	+	-	-
	NA10855	+	-	-
6	NA11993	+	-	+
	NA11992	+	-	+
	NA10860	+	+	-
7	NA11995	+	-	+
	NA11994	+	-	+
	NA10861	+	+	-
8	NA12236	+	-	-
	NA12154	+	-	-
	NA10830	+	+	-
9	NA12156	+	-	+
	NA12155	+	-	-
	NA10831	+	+	-
10	NA07345	+	-	-
	NA07357	+	-	-
	NA07348	+	+	-
11	NA07000	+	-	-
	NA06994	+	-	-
	NA07029	+	-	-
12	NA07056	+	-	+
	NA07022	+	-	-
	NA07019	+	+	-
13	NA11840	+	-	+
	NA11839	+	-	+
	NA10854	+	+	-
14	NA12044	+	-	-
	NA12043	+	-	-
	NA10857	+	+	-
15	NA12145	+	-	-
	NA12144	+	-	-
	NA10846	+	-	-
16	NA12239	+	-	+
	NA12146	+	-	-
	NA10847	+	-	-
17	NA12234	+	-	-
	NA12264	+	-	-
	NA10863	+	-	-
18	NA12717	+	-	-
	NA12716	+	-	-
	NA12707	+	-	-
19	NA12892	+	-	+
	NA12891	+	-	+
	NA12878	+	-	-

Family ID	Individual ID	HapMap	Gene panel 1	Gene panel 2
20	NA12813	+	-	+
	NA12812	+	-	-
	NA12801	+	-	-
21	NA12815	+	-	+
	NA12814	+	-	+
	NA12802	+	-	-
22	NA12873	+	-	-
	NA12872	+	-	-
	NA12864	+	-	-
23	NA12875	+	-	-
	NA12874	+	-	-
	NA12865	+	-	-
24	NA12761	+	-	-
	NA12760	+	-	-
	NA12752	+	-	-
25	NA12763	+	-	-
	NA12762	+	-	-
	NA12753	+	-	-
26	NA07055	+	-	+
	NA07034	+	-	-
	NA07048	+	-	-
27	NA06985	+	-	+
	NA06993	+	-	+
	NA06991	+	-	-
28	NA12751	+	-	+
	NA12750	+	-	+
	NA12740	+	-	-
29	NA11882	+	-	+
	NA11881	+	-	+
	NA10859	+	-	-
30	NA12249	+	-	-
	NA12248	+	-	-
	NA10835	+	-	-
31	NA06990	-	+	-
32	NA07349	-	+	-
33	NA10842	-	+	-
34	NA10843	-	+	-
35	NA10844	-	+	-
36	NA10845	-	+	-
37	NA10848	-	+	-
38	NA10850	-	+	-
39	NA10852	-	+	-
40	NA10853	-	+	-
41	NA10858	-	+	-
42	NA12547	-	+	-
43	NA12548	-	+	-
44	NA12560	-	+	-

Table 3.2 CEPH samples in the population genotyping databases

Family trio samples are ordered as: father, mother, offspring.

Gene panel 1 = *IL1A*, *IL1B*, *IL1R1*, *IL1R2*, *IL1RN*, *IL18R1*, *IL18*, *IL18BP*

Gene panel 2 = *IL1F5*, *IL1F6*, *IL1F7*, *IL1F8*, *IL1F9*, *IL1F10*, *IL1RL2*

algorithm, ignoring missing data, to estimate the maximum-likelihood values of phase frequencies from which the LD is calculated.

Within Haploview a number of quality control checks were carried out on the genotyping data. Thresholds of: a maximum of one Mendelian inheritance error and $p \leq 0.05$ deviation from Hardy-Weinberg equilibrium were applied. All SNPs violating these, or with MAF <0.05, were excluded. This MAF cut-off was selected because the power to detect a significant frequency difference in SNPs below this is minimal in cohort sizes of those available in this project.

Using the Haploview integrated Tagger software, sets of tSNPs were selected for each candidate region based on the LD results, examining for pairwise tagging, $r^2 \geq 0.8$. This tSNP selection strategy has been shown to reduce the number of typed SNPs required, while retaining 96% power to capture all SNPs relative to genotyping all SNPs directly (de Bakker et al., 2005). Tagger utilises the algorithm developed by Carlson et al (Carlson et al., 2004). The SNP exceeding the r^2 threshold with the maximum number of other SNPs is identified. This maximally informative SNP and the SNPs it captures are grouped together as a bin. The binning process is repeated for the remaining SNPs until all are assigned to a bin. Any SNPs $r^2<0.8$ with all others are placed in a singleton bin and require direct genotyping. Figure 3.2 illustrates the graphical representation of LD, generated in Haploview, and the selection of tSNPs.

Due to the methodology utilised in the Golden Gate genotyping platform (section 3.1.2.2.) it was not possible to directly genotype insertions, deletions, or multi-allelic SNPs. However, as the population genotyping data from PGA was generated through re-sequencing, it included all polymorphism types. In order to include these variations in the investigation all non-simple SNPs were not selected as tSNPs to be genotyped.

Prior to finalisation, the total tSNP set, for all candidate regions, underwent assay design through the manufacturer run Illumina Assay Design Tool. Based on the surrounding genomic sequence, primer sets were designed for each SNP. A SNP score, scaled between 0 and 1 to reflect how well the assay was predicted to perform on the GoldenGate platform, was calculated for each SNP assay. The SNP scores were used to identify the SNP set with the optimal predicted performance, to be genotyped, as illustrated in Figure 3.3.

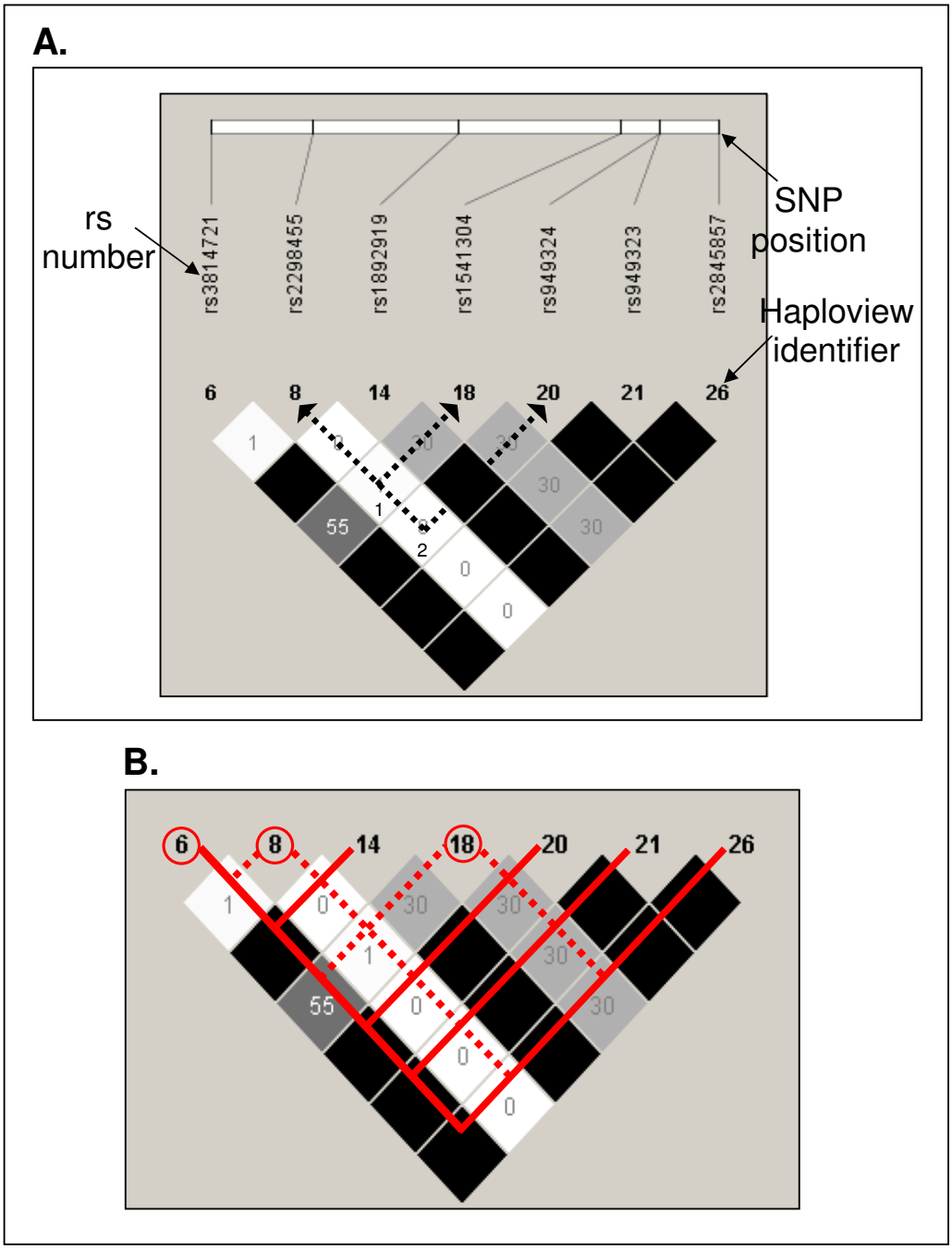


Figure 3.2 Graphical representation of LD patterns

A Graphical representation of the LD relationships between SNPs is generated in Haploview. The example shown is the *IL18BP* candidate region. (A.) Only SNPs satisfying the selection criteria are included in the LD plot. The relative position within the region of each SNP is represented by a line. Each SNP is labelled by rs number and Haploview ID number. Haploview ID numbers are assigned to each SNP according to chromosomal order in the full SNP set. Each square represents the pairwise LD between the two SNPs at the top of each diagonal column, so that square 1 shows the r^2 LD between SNPs 8 and 18, and square 2 the r^2 LD between SNPs 8 and 20. The graphical representation is shown on a grey scale, with higher r^2 LD represented by

darker blocks. The white blocks show the r^2=0 (random assortment) and the black blocks show r^2=1 (total LD). (B.) Of the seven SNPs in the candidate region satisfying the selection criteria a minimum subset of three (circled in red) are required to capture all of the variation. SNPs 8 and 18 (red dashed lines) are not in high LD ($r^2 \geq 0.8$) with any of the other SNPs (SNP 8 r^2 max=0.1, SNP 18 r^2 max=0.55) and therefore must be genotyped directly. SNPs 6, 14, 20, 21, and 26 are all in high LD ($r^2 \geq 0.8$) with each other. Therefore, it is only necessary to genotype one of them. SNP 8 was selected as the tSNP (solid red lines) but any of the five SNPs could have been selected. For clarity the r^2 values are shown in each block (not included in the results section plots).

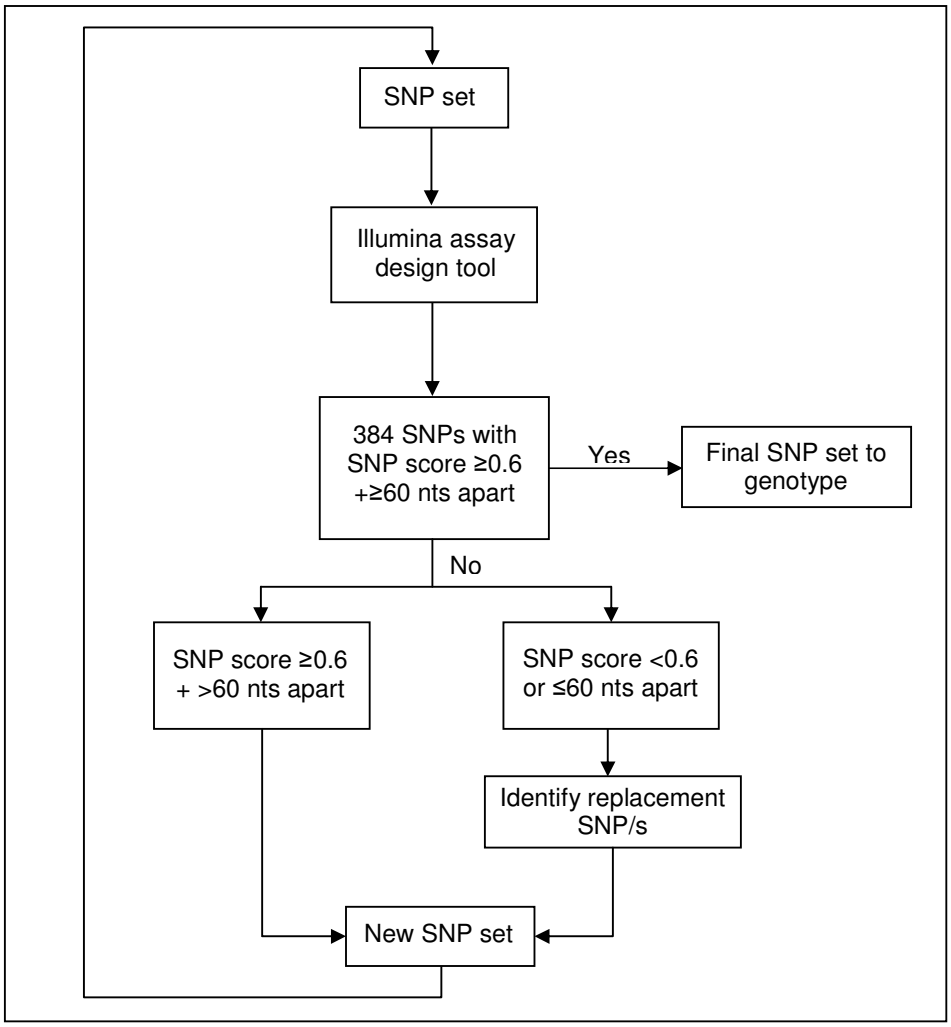


Figure 3.3 Finalisation of tSNP set

The full list of selected tSNPs was submitted to the Illumina Assay Design Tool for design and genotype performance prediction of, each SNP assay. The SNP design score was calculated based on: specificity within the genome and melting temperature of the required primers, and presence of a repetitive region. Assigned SNP scores were scaled between 0 and 1, with a manufacturer recommended cut-off of 0.6 for assays of sufficient quality to provide high quality genotyping results. Due to the design of the platform it was not possible to genotype SNPs within 60 nucleotides of each other. Any SNPs with a design score <0.6 or within 60 nucleotides of another were identified from the design report, and substituted where possible. This was achieved for each SNP in turn by repeating the tSNP selection for the set of SNPs captured, selecting an alternative tSNP able to capture all of the variation within that bin. For some SNPs there was not an alternative single SNP in sufficiently high LD with all SNPs within the bin to capture them all. In these cases two or more alternative SNPs, which between them captured all of the remaining SNPs, were selected as substitutes. This process of tSNP list submission, assay design scoring, and SNP substitution, was repeated until a final set of 384 tSNPs, all with SNP scores ≥ 0.6, was identified.

nts – nucleotides

3.1.2.2. Genotyping

Genotyping was performed using the Illumina Golden Gate assay genotyping platform (custom 384 SNP panel, 96-sample Sentrix array matrix).

The Illumina golden gate assay utilises three oligonucleotides for each SNP investigated: one locus-specific oligonucleotide (LSO) and two allele-specific oligonucleotides (ASOs). The LSOs contain a sequence complimentary to a region downstream of the SNP, an address sequence (unique to each loci), and a universal primer sequence. The ASOs contain a region complimentary to the SNP position (in the presence of one of the possible alleles), and a universal primer sequence, one for each ASO type (Figure 3.4). A pool of oligonucleotides for all SNPs included in the assay are hybridised to genomic DNA. Through an extension and ligation step they are joined together creating a full length oligonucleotide, containing information on both the locus and allele. It is these oligonucleotides, rather than the genomic DNA, which undergo PCR amplification using the universal primer sequences in each LSO and ASO. The primers complimentary to the ASO sequences are dye-labelled to allow allele discrimination. Illumina designates alleles for each SNP as either allele 'A' or allele 'B', details of this designation method are given in Illumina's ''TOP/BOT' strand and 'A/B' allele' technical note. Primer 1, corresponding to allele 'A', is Cy3 labelled, and primer 2, corresponding to allele 'B' is Cy5 labelled. The dye-labelled, amplified oligonucleotides are hybridised, via the address sequences, to beads coated in the corresponding locus-specific address sequence. The fluorescence intensity of the oligonucleotides hybridised to each bead is measured and the SNP it corresponds to determined. A map of the locations of each bead type is created prior to use of the assay, based on the address sequence present (Gunderson et al., 2004). A more detailed explanation of the platform is given in Appendix 1 A.

Genotyping was performed following manufacturer's recommended protocol. The protocol is outlined in Appendix 1 B.

3.1.2.2.1. Genotype calling

The Illumina BeadStudio genotyping module is designed for use with the GoldenGate platform. BeadStudio utilises a clustering algorithm to identify and define the parameters of three clusters of samples, corresponding to each genotype. The clusters are defined based on the fluorescence intensity at each wavelength of all DNA samples analysed. Prior to clustering each array, corresponding to one DNA sample, is self-normalised, adjusting for

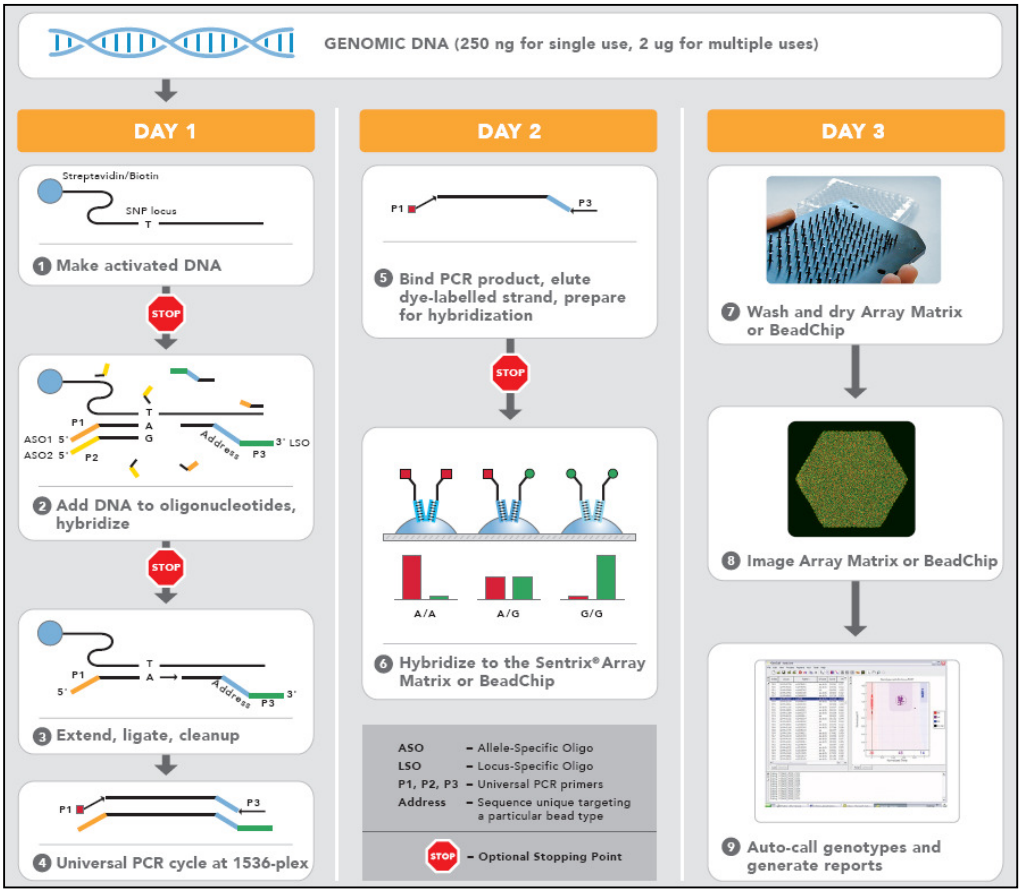


Figure 3.4 Golden Gate Assay Work-flow

The Illumina Golden Gate genotyping platform is a bead based, array of arrays. Shown is the assay work-flow demonstrating the principles utilised. Genomic DNA is bound to paramagnetic beads (1) and hybridised to allele-specific (ASO) and locus-specific oligonucleotides (LSO) (2). Each oligonucleotide contains a universal PCR primer sequence, and LSOs also contain a locus-specific address sequence. The DNA-bound oligonuclotides are extended and ligated together, creating a full length oligonucleotide containing information on both the locus and allele (3). Following removal of the genomic DNA, the full length oligonucleotides undergo PCR using universal primers (4). Each ASO complementary primer is dye-labelled according to allele. The dye-labelled PCR products are purified (5) and hybridised to the array matrix (6). The array matrix consists of beads coated in locus-specific address sequences, complementary to those in the LSOs, to which the PCR products hybridise. Following removal of mis-hybridised sequences (7) the array is imaged recording fluorescence intensity at both dye wavelengths (8). The position of beads of each locus identity (according to address sequence) is mapped previously, enabling compilation of the fluorescence data for each individual according to loci. This data is plotted for each individual as the fluorescence intensity at each wavelength against signal intensity. From this plot the clusters of data points corresponding to the three genotypes are identified and genotype calling performed (9).

Picture reproduced with permission from Illumina® Inc.

nominal intensity variation between the two colour channels, background differences between the channels, and possible cross-talk between the dyes.

A scatter plot (Genoplot) of the data is generated for each SNP (Figure 3.5) as either a Cartesian (A) or polar plot (B). In a Cartesian plot the normalised Cy3 intensity (corresponding to allele 'A') for each DNA sample is plotted against the normalised Cy5 intensity (corresponding to allele 'B'). In a polar plot the normalised theta, angle deviation of the data point from pure 'A' allele signal, is plotted against normalised R, the intensity of the fluorescence. The software determined cluster locations are circled, and the areas within which samples are assigned genotypes, based on a GenCall 0.25 cut-off (explained below), are shaded. All graph features are colour coded according to genotype: 'AA' = red, 'AB' = purple, 'BB' = blue. A GenTrain score between 0 and 1 is calculated for each SNP, based on the fluorescence intensity, angle, dispersion, and overlap between genotype clusters, with higher scores reflecting better defined clusters. The calling algorithm identifies which genotype cluster each data point corresponds to, based on fluorescence intensity values for each sample. The GenTrain score, in combination with information from the sample clustering algorithm and a model-to-fit score, is used to calculate a GenCall quality score, between 0 and the GenTrain score, for each genotype call. GenCall scores indicate the reliability and confidence of the genotype call, with samples further away from the centre of the genotype cluster being assigned lower scores. A GenCall score of 0.25 is used to define the limits of the genotype calling region, shown as the darker shaded regions in the Genoplots. All samples falling outside of this call region are not assigned genotypes and classed as failed for that SNP assay. The 50^{th} (GC50) or 10^{th} (GC10) percentile of the GenCall scores across all genotypes are used as quality scores for either SNPs (all samples for that SNP), or samples (all SNPs for that sample).

Automatic sample clustering and genotype calling were performed using Illumina BeadStudio (version 2.1.8.32932). SNP clustering was checked visually and adjusted manually where appropriate, with particular attention paid to all SNPs with low GenTrain scores. GC10 <0.5 and visual assessment of the definition and spacing of genotype clusters was used in the decision to class SNP assays as failed. Sample performance was checked by the GC10 score, with all samples with GC10 <0.5 discarded. All clustering and decisions to exclude were discussed and agreed with the Illumina European applications manager.

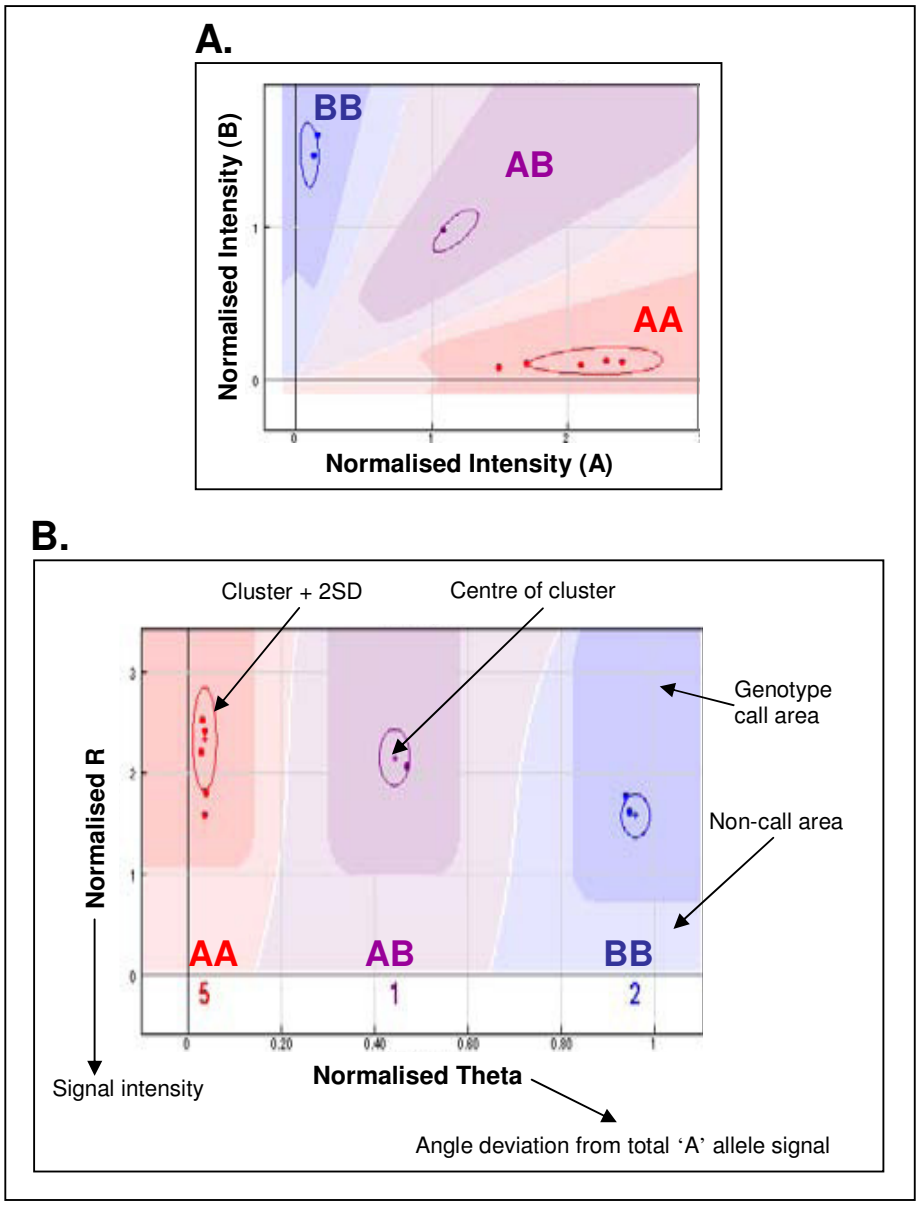


Figure 3.5 Genoplot of GoldenGate genotyping results

A scatter plot of the fluorescence data of all DNA samples is generated for each SNP investigated. These are either plotted as the Cy3 (allele 'A') signal intensity against the Cy5 (allele 'B') intensity (Cartesian plot A.) or as theta (deviation from total Cy3 signal) against R (signal intensity) (polar plot B.). DNA samples fall into three clusters, corresponding to the three possible genotypes, and are assigned genotypes depending on the cluster they fall within. These cluster locations are used to define the call areas of each genotype. All samples falling outside of these call areas (dark shaded areas) are classed as failed for that assay.

Figure adapted with permission from Illumina® Inc.

3.1.2.2.2. Golden Gate genotyping pilot study

One of the sources of patient DNA samples used in stage-1 of this study was the British Society for Paediatric and Adolescent Rheumatology (BSPAR) national DNA repository. However, these samples had been collected over a number of years and were known, from previous studies, to be of varying quality. It was therefore necessary to perform a pilot study using these DNA samples to determine which were of sufficient quality to successfully genotype on the Illumina GoldenGate platform.

Patient samples from the repository had been genotyped in three previous studies, two investigating five SNPs each, and the other study two SNPs. Not all repository DNA samples had been used in all three studies, with samples having been genotyped for two, five, seven, ten, or twelve SNPs. Eight samples were selected as representative of the cohort, covering the range of previous call rates, ranging from 100% to 20% (Table 3.3). These DNA samples were genotyped on the Illumina training array comprised of 1,082 loci.

3.1.2.3. Data Quality Controls

The control population was checked for significant deviations from Hardy-Weinberg equilibrium at each genotyped loci using either Pearson's chi-squared or Fisher's exact test, with a significant p value cut-off of 0.05.

The Hardy-Weinberg equilibrium (HWE) principle states that in a randomly mating population, allele and genotype frequencies of a polymorphism, not subjected to natural selection or mutation, will remain constant between generations. When a bi-allelic polymorphism is in HWE the frequencies of the three possible genotypes will conform to the following equation:

$$p^2 + 2pq + q^2 = 1$$

p= frequency of allele 1, q = frequency of allele 2

Where:

$$f\text{AA} = p^2$$
$$f\text{Aa} = 2pq$$
$$f\text{aa} = q^2$$

Sample	Call Freq.	SNPs tested
BRM119	100	12
GOS070	100	12
GOS018	75	12
MDX030	75	12
BRM243	50	12
GOS262	50	12
MDX004	28	7
WEX081	20	5

Table 3.3 BSPAR repository patient samples used in GoldenGate Pilot Study

Eight sJIA patient DNA samples from the BSPAR repository were included in the GoldenGate pilot study. These samples were selected as representative of the range of call rates in previous studies of all the repository samples. Shown are the total number of SNPs tested in previous studies and the call frequencies of these DNA samples.

The allele frequencies (p and q) are calculated based on the number of each genotype in the population (Templeton, 2006).

$$p = \frac{n_{AA} + (0.5 \times n_{Aa})}{n_{population}} \qquad q = \frac{n_{aa} + (0.5 \times n_{Aa})}{n_{population}}$$

n_{xx} = number of individuals of that genotype
$n_{population}$ = number of individuals in the population

Significant deviation from HWE suggests that, either the assumptions of HWE are violated or, most likely, that there are errors in the genotyping assay resulting in genotype assignment bias. Violation of HWE is commonly used in population-based studies as a test for genotyping errors (Hosking et al., 2004). Spurious genotyping data can lead to false-positive association results. It is therefore important to identify erroneous assays prior to data analysis. A disease-associated polymorphism may not be in HWE in a case population, violation of HWE in patients can even be used to detect disease associations (Lee, 2003). Therefore, it is not necessary to test for deviations from HWE in case populations as part of the data quality control. Significant violations of HWE are tested for using Pearson's chi-squared (χ^2) statistic with 1 degree of freedom. This statistic tests for deviations of the genotype numbers observed in the data from the expected genotype numbers. Expected genotype numbers are calculated from the HWE equation based on the allele frequencies. χ^2 is calculated as:

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

O = Observed frequency
E = Expected frequency

Because the χ^2 significance value calculated is an approximation it is not accurate when expected numbers are below five (Templeton, 2006). Therefore, it is not appropriate to use this statistic with polymorphisms where any genotype is below this frequency. In this situation Fisher's exact test is employed instead. With the observed and expected numbers represented as:

	AA	**Aa**	**aa**	**Row total**
Observed	*A*	*b*	*c*	*a + b + c*
Expected	*D*	*e*	*f*	*d + e + f*
Column total	*a + d*	*b + e*	*c + f*	*n*

Fisher's exact test is calculated as:

$$p = \frac{(a+b+c)!\,(d+e+f)!\,(a+d)!\,(b+e)!\,(c+f)!}{n!\,a!\,b!\,c!\,d!\,e!\,f!}$$

Due to Fisher's exact test being overly conservative and computationally intensive it is only used when chi-squared is inappropriate (Elston and Johnson, 2010).

3.1.2.4. Association Analysis

Association analysis of the genotyping data was performed using the COCAPHASE module of UNPHASED (Dudbridge, 2003) (v2.403 and v3.0). UNPHASED tests for disease association by performing unconditional logistic regression. This test performs likelihood ratio tests under a log-linear model of the probability that an allele belongs to a case rather than a control.

Both single marker and haplotype analysis, rare haplotype threshold cut-off of 0.01 in both case and control cohorts, was performed in COCAPHASE. As some haplotypes will be of uncertain phase COCAPHASE utilises an expectation-maximisation (EM) algorithm to estimate haplotype frequencies. This is an iterative process of an expectation step, calculating the log-likelihood of each haplotype phase assuming HWE and based on the frequency of know haplotypes, followed by a maximisation step, maximising the expected log-likelihood using a log-linear model of the imputed haplotypes (Clayton, 2007).

Although due to the tagging SNP selection criteria, none of the tSNPs genotyped were in $r^2>0.8$, SNPs within the same candidate region were still in some amount of LD with each other. Therefore, any SNPs in LD with a disease-associated SNP may also show evidence of significant frequency differences because of this relationship. When multiple tSNPs in the same region showed a significant frequency difference they were separated into clusters according to LD patterns. A stepwise conditional regression approach was then used within each LD cluster to identify which tSNPs showed primary (direct) effects, and which secondary (indirect) effects. In this analysis the tSNP being conditioned on was set as the null model, so that any effect of this SNP is accounted for. The tSNP being tested was then added into the model and tested for improvement of fit. If both markers represent independent primary effects, both will show a significant ($p<0.05$) effect when conditioned on the other. If

a primary effect SNP is conditioned on a secondary effect SNP, a significant effect will be shown. In the reverse analysis, a secondary effect SNP conditioned on a primary effect SNP, no additional significant effect will be observed (Cordell and Clayton, 2002).

The stepwise conditional regression was then repeated with the primary effect SNPs from each LD cluster to determine if the effects were independent, or related to each other, using the same criteria as previously. Nested conditional analysis was also performed to ensure that all of the primary tSNPs were required to fully explain the effects seen.

3.1.3. Stage-2

3.1.3.1. Genotyping

All SNPs showing evidence of significant ($p<0.05$) frequency differences between the patient and control cohorts in stage-1 of the association study were genotyped using the genotyping service provided by KBioscience UK, with no DNA water controls included in all plates. Due to assay design constraints it was not possible to directly genotype rs3773999 (*IL1RAP* SNP14). Instead rs9290935 (*IL1RAP* SNP14a), which was in complete LD ($r^2=1$) with SNP14, was genotyped in the stage-2 samples.

The genotyping platform used is based on Applied Biosystems' Taqman® chemistry, in which each SNP assay contains two primers, which bind either side of the SNP locus, and two allele-specific oligonucleotide probes. Each probe has a fluorescent dye attached to the 5′ end, VIC© on the allele 1 probe and FAM™ on the allele 2 probe, and a non-fluorescent quencher dye attached on the 3′ end of the probe. When the quencher is in close proximity to the fluorescent dye it reduces the fluorescent emission through fluorescent resonance energy transfer. During amplification the hybridised probe is cleaved by the 5′ exonuclease activity of the polymerase, releasing the reporter dye, separating it from the quencher and so increasing the signal. The resulting fluorescent signal of each dye is proportional to the amount of amplicon and used to determine the genotype.

3.1.3.2. Data Quality Controls

To check for any genotyping discrepancies between the two genotyping platforms used for the two association study stages DNA samples from 18 patients used in stage-1 were also genotyped for the stage-2 SNPs on the KBioscience platform. These samples were not

included in the analysis of this study stage.

Genotypes of the control population were checked for Hardy-Weinberg equilibrium using Pearson's chi-squared test or Fisher's exact test, with a significant p value cut-off of 0.05.

3.1.3.3. Association Analysis

Data from both study stages were analysed in the software PLINK (Purcell et al., 2007), stratifying by study stage using the Cochran-Mantel-Haenszel (CMH) test for meta-analysis and the Breslow-Day (BD) test for homogeneity of odds ratios.

The CMH test was used to analyse the relationship between disease state and SNP allele, controlling for study stage. With allele frequencies represented in a 2x2 table for each stage as:

	1	**2**	**Row total**
Case	a	b	$a + b$
Control	c	d	$c + d$
Column total	$a + c$	$b + d$	n

The CMH statistic is calculated as the squared sum of the deviations between the observed value and the expected value under the null hypothesis, over the estimate of the variance of the squared differences:

$$\chi_{CMH} = \frac{(\Sigma\ (a-(a+b)(a+c)/n))^2}{\Sigma(a+b)(a+c)(b+d)(c+d)/(n^3-n^2)}$$

The CMH statistic has a chi-squared distribution with degrees of freedom (df)= number of strata -1 (McDonald, 2009).

The BD statistic tests for significant differences between the effect size (ORs) in each strata through the sum of the squared deviations of observed and expected values, each standardised by its variance:

$$\chi_{BD} = \sum \frac{(a - E(a/OR_{CMH}))^2}{Var\ (a/\ OR_{CMH})}$$

E = expected value under the null hypothesis
Var = variance of the squared differences

The BD statistic has a chi-squared distribution with df= number of strata -1. A significant test indicates between-strata heterogeneity of association strength (Breslow and Day, 1980).

3.1.4. WTCCC control cohort analysis

Subsequent to completion of the genotyping stage of this association study, additional population genotyping data has been made available, including the Wellcome Trust Case Control Consortium (WTCCC) project (Wellcome Trust Case Control Consortium, 2007).

The WTCCC is a collaboration of over 50 UK research groups to perform genome-wide association studies (GWAS) of complex human diseases using common control cohorts. In phase 2 of the project genotyping was performed on Affymetrix v6.0 and Illumina 1.2M genotyping chips. The control cohorts used are from two different sources: the 1958 British birth cohort, and UK blood donors. The 1958 British birth cohort (National Child Development Study, http://www.cls.ioe.ac.uk/studies.asp?section=000100020003) consists of all individuals born (almost 17,500) in England, Wales, and Scotland in one week in 1958. In 2002-2004 blood samples were collected from survivors. 3,000 of these, self-reported as white ethnicity, were selected for inclusion in the WTCCC control cohort as being representative of sex and geographical region. The UK blood donor cohort was specifically collected for this project. It consists of 3,000 blood donors aged 18-69 years, selected based on sex and geographical region to mirror the distribution of the 1958 British birth cohort samples. Following the initial phase of the WTCCC project there were few significant allele frequency differences between the two control groups. The two cohorts were therefore combined for use as a single control group (Wellcome Trust Case Control Consortium, 2007). The available genotyping data from phase 2 of the WTCCC project includes genotypes for a total of 4,673 control individuals.

This resource is available for use by researchers as a control cohort in genetic association

studies as an alternative to in-house cohorts. In an additional analysis, for confirmation of the study results, the pooled stage-1 and stage-2 cases were compared with the WTCCC control cohort, where available, as a larger and independent control cohort, for all SNPs included in both stages of the association study. Association analysis was performed in UNPHASED as described in section 3.1.2.4.

The statistical power of an association study is the probability that a true disease-association will be detected under the parameters specified. CaTS association study power calculator (Skol et al., 2006) was used to determine the power of the analysis. Based on the samples sizes available the analysis would have 96% power to detect the effect of a disease associated SNP (MAF 0.3 and a relative risk of 1.7) in a multiplicative model at a significance level of 0.05.

3.1.5. Identification of additional tagged SNPs

As the tSNP selection was performed based on the HapMap2 genotyping data, and only included SNPs within the selected candidate regions, further LD analysis was performed based on the HapMap3 data (section 3.1.6.) to identify any additional SNPs in high LD with the SNPs showing evidence of significant disease association.

SNAP (Johnson et al., 2008) is a freely available internet based program for identifying nearby SNPs in LD with query SNPs (http://www.broadinstitute.org/mpg/snap). Using Haploview version 4.0 the pairwise LD, based on the HapMap datasets, between all SNPs within 500kb of each other has been pre-calculated and stored for rapid query responses. Using the proxy search application, and specifying use of the CEU population, all SNPs within 500kb of, and in $r^2<0.8$ with, the SNP of interest were requested. Genotyping data from HapMap3 data release 2 was used unless the SNP of interest was not included, in which case HapMap2 release 22 was used.

3.1.6. Validation of the tagging SNP selection results

Investigation of the LD patterns over the candidate regions, and the selection of tagging SNPs, was performed using the population genotyping data from phase 2 of the HapMap project (The International HapMap Consortium, 2005). Further population genotyping data was made available subsequent to the completion of the tSNP selection and genotyping

stages of this study. This includes phase 3 of the HapMap project, and the WTCCC project (Wellcome Trust Case Control Consortium, 2007).

Phase 3 of the HapMap project was publicly released in 2009. In this phase, 180 CEU individuals from 36 families, consisting of: one singleton, one duo, eight trios, one four-member family, seven five-member families, and eighteen six-member families, were genotyped. Non-redundant SNP assays from Phases I and II of the project were included. The WTCCC project is explained in section 3.1.4.

In order to investigate the validity of the LD analysis and the tSNP selection, generated from the HapMap phase 2 population, genotyping data from these larger population cohorts were analysed for comparison.

Genotyped SNPs from each candidate region, determined in the tSNP selection stage (based on HapMap phase 2 data) to be capturing additional SNPs, were selected as representative of the whole data set. All stage-1 genotyped SNPs which showed a significant frequency difference between the case and control populations (sections 4.2 to 4.5) were selected for comparisons. Additional SNPs, identified as tagging a number of other SNPs within the region, were also selected for investigation. For the candidate regions with no evidence of disease association in stage-1 (sections 4.6 to 4.8) all genotyped SNPs tagging other SNPs were investigated. Data from both HapMap phase 3 and WTCCC phase 2 for the selected tSNPs and all SNPs assigned as captured by them, where available, were extracted and the pairwise LD calculated in Haploview version 4.1. The r^2 values of each SNP pair calculated from the three different population genotyping resources were compared to identify any discrepancies.

3.2. Sequence feature prediction

Evidence for a genetic polymorphism being associated with disease does not necessarily mean that it is involved in disease susceptibility, it may only be a marker for the true causal variant. In order to play a role in disease susceptibility or severity, causal variants must have a biological role and effect gene expression. Therefore, only polymorphisms located within functional regions can be involved in disease susceptibility. It is therefore important that a functional role in gene regulation of polymorphisms is demonstrated for disease associations to be shown with confidence.

Many genetic variants, each with a small effect, are involved in complex diseases such as sJIA. As these variants are also found in the general population it is reasonable to assume that disease associated polymorphisms will have a small biological effect on gene regulation and/or function. Current knowledge regarding regulation of gene expression is limited, but it is accepted that a substantial proportion of regulatory regions are located in the ~53% of the human genome which is non-coding (Lander et al., 2001; Venter et al., 2001). As well as core promoters, to which the transcriptional pre-initiation complex and RNA polymerase bind, there are a number of other regulatory elements involved in the temporal, spatial, and inductive expression regulation. These include enhancer/silencer elements, methylation sites, transcription factor binding sites, DNAse hypersensitivity sites, and microRNAs. Regulatory elements can be 5′ or 3′ of a gene, within introns, within other genes (Loots, 2008), and up to megabases away from the gene (Nobrega et al., 2003; Sagai et al., 2005). It is therefore not possible to predict the probability of a disease associated polymorphism being involved in gene regulation based solely on its location relative to a gene.

A useful technique in the initial stages of investigating potential function is to use bioinformatic tools to predict the presence of functional elements through examination of the genomic sequence. As knowledge of regulatory sequence motifs is incomplete these *in silico* techniques cannot be absolute and experimental validation is required as empirical proof of biological function. They are however useful tools to inform subsequent experimental investigation.

3.2.1. Comparative Genomics

From an evolutionary perspective, sequences involved in gene expression, including coding regions and all regulatory elements, will change at a slower rate than sequences with no biological role. This is because sequence changes abrogating function are often deleterious to the organism and will therefore not spread throughout the population, resulting in preservation of the original sequence. Mutations in regions with no innate function will confer no biological impact on the organism and therefore be evolutionarily neutral. These mutations will therefore accumulate, and the sequence will constantly change over evolutionary time. Utilising the assumptions that the function of regulatory elements are conserved between species, and that functional conservation is accompanied by nucleotide sequence conservation, identification of genomic regions with potential function is possible

through identification of homologous regions across species by comparative genomics (Nardone et al., 2004). The working hypothesis for this approach is that regions with high levels of evolutionary conservation are more likely to be functionally important than regions with no conservation. Although sequence conservation in non-coding regions is suggestive of the presence of a regulatory element, it is not possible to infer the type of element present from this evidence. Further investigation of the conserved region, both sequence prediction and experimental techniques, is therefore required.

Sequence disparity between closely related species, for example mouse and rat, highlights regions where nucleotide changes are highly tolerated. As there has been little evolutionary pressure on such a sequence it is unlikely to be functionally important. Similarity in genomic sequence between more distantly related species, for example human and murine, indicates regions which have been highly constrained by natural selection, and are therefore more likely to have a functional role (Nardone et al., 2004). Aligning both mouse and rat sequences to the human genome enables simultaneous identification of regions with high homology between distantly related species and regions with low homology between closely related species. Regions highly conserved between all three species are more likely to have functional roles, while regions with no homology, or homology between human and only one murine species, have the least evidence for the presence of a functional element.

Regions in the TNFα promoter corresponding to sequences non-essential for expression in the human promoter show high variability between great apes, new world monkeys, and old world monkeys. Conversely, a 69 base pair sequence which is completely conserved between primates corresponds to part of the human TNFα promoter region which forms the transcriptionally active nucleoprotein-DNA complex, essential for gene regulation (Leung et al., 2000). Sequence homology corresponding to conserved function has also been shown between the human and mouse genome. Significant homology between these species in the proximal 5'flanking region and distal sequence at the 5' end of the β-globin gene cluster correlates to the locations of known regulatory regions in the human sequence (Hardison et al., 1997). Sequence comparison of 28 human and murine orthologous gene pairs specifically upregulated in skeletal muscle showed that all 99 experimentally defined human transcription factor (TF) binding sites in the sequences, except those for Sp-1, were within blocks of sequence conservation (Wasserman et al., 2000). A number of studies have used multispecies comparison to identify regions of sequence conservation, and determined experimentally that

they correspond to regulatory elements. This includes a conserved Ets-1 binding site in the Th2 cytokine gene locus (Strempel et al., 2010) and a 5′ regulatory element in the Pax8 gene locus (Nitsch et al., 2010). Of 16 identified multispecies conserved elements in presynaptic gene regions 75% were found to function as neuronal enhancers (Liu et al., 2009b).

Although identification of sequence homology can indicate the presence of a regulatory or functional element, it does not necessarily follow that regions with no species homology do not have a functional role. An experimentally identified transcriptionally active region in the upstream region of the human *IL6* promoter was found to not be within a highly conserved region, showing that non-conserved regions may also be functionally relevant (Samuel et al., 2008). Some regulatory elements may have a species-specific role, or compensatory sequences changes may have occurred resulting in conservation of function without sequence conservation. Following the generation of nearby new functional elements, which will happen by chance due to their small size, evolutionary constraints on the original site will be relaxed and the site may then be lost (Dermitzakis and Clark, 2002). An example of this is TF recognition sequences, which are degenerate and generally only six to eight nucleotides long. This means that TF binding can be maintained despite sequences changes, and that new binding sites can be spontaneously generated through mutation. In addition, homologous genes can be regulated by different TFs, and recognition sequences for some homologous TFs are different between species. It is therefore possible for identical TF binding sites to be present in different species but without any sequence similarity.

VISTA (Mayor et al., 2000b) is a freely available internet based program for visualising global DNA sequence conservation (http://genome.lbl.gov/vista/index.shtml). Sequences are aligned against the reference sequence using the Global Alignment System (GLASS) alignment program (Batzoglou et al., 2000). GLASS aligns sequences through an iterative hierarchical alignment approach in which long segments with exact base matching and high similarity in the flanking regions are aligned initially. Using successive smaller matching segments, alignment is repeated on the intervening regions, and finally on any remaining short unaligned regions. Following sequence alignment VISTA calculates the percentage identity over a 100bp moving window. This is plotted as a peak and trough plot with the base sequence along the x-axis, and percentage identity along the y-axis. All conserved regions with >70% identity over a minimum of 100bp are highlighted under the curve.

Genomic sequences for the genes of interest were obtained from the Ensembl genome browser (www.ensembl.org) for the following species: Human (*Homo* sapiens, Febuary 2009 assembly), Mouse (*Mus* musculus, m37 assembly 2007), and Rat (*Rattus* norvegicus, November 2004 assembly). mVISTA, for multiple species alignment, (http://genome.lbl.gov/vista) (Mayor et al., 2000a) was used to align the mouse and rat sequences against the human genome sequence.

3.2.2. Transcription Factor Binding prediction

The program AliBaba 2.1 (http://www.gene-regulation.com/pub/programs/alibaba2/.html) was used to predict transcription factor (TF) binding sites within the genomic sequence flanking the SNPs of the associated *IL18BP* haplotype. AliBaba utilises known recognition sequences from the TRANSFAC database (http://www.biobase.de/) of experimentally determined TF binding sites (Wingender, 1988; Wingender et al., 2000) to predict recognition sequences within a sequence of interest. Pairwise alignment of the known recognition sequences to the sequence of interest is performed, followed by construction of matrices for each alignment (Grabe, 2002). Predictions were performed using a minimum matrix conservation cut off of 75%. AliBaba is a widely used TF binding site prediction program, and a number of publications have shown experimental validation of TF binding which it has predicted (Chowdhury et al., 2006; Zhang et al., 2008a; Sharma et al., 2009; Reddy et al., 2009).

3.3. General experimental protocols

3.3.1. DNA extraction

3.3.1.1. From blood

To isolate the DNA from white blood cells, 5ml blood was collected into tubes containing 50µl 0.5M EDTA. Extraction was carried out using Puregene kit following manufacturer's instructions. Briefly, the blood was mixed with 15ml red blood cell lysis solution by inversion, incubated at room temperature for five-ten minutes, and centrifuged at 2,000g for five minutes to pellet the white blood cells. The supernatant was discarded and the pellet resuspended in the residual supernatant by vigorous vortexing. Then 5ml of cell lysis solution was added and the sample vortexed to lyse the cells. To remove RNA from the sample 12.5µl RNase A solution was added, mixed by inversion, incubated for 15 minutes at 37^{o}C, and

cooled by placing on ice for three minutes. To precipitate the protein, 833µl protein precipitation solution was added, mixed by vortexing until the solution was uniform, and centrifuged at 2,000g for ten minutes. To precipitate the DNA the supernatant was transferred into a fresh tube containing 2.5ml 100% isopropanol, gently inverted 50 times to mix, and centrifuged at 2,000g for three minutes to pellet the DNA. The supernatant was discarded, the pellet washed by several inversions with 2.5ml 70% ethanol, and centrifuged again for one minute. After removal of the ethanol the pellet was allowed to air dry and rehydrated in 100µl DNA hydration solution.

3.3.1.2. From saliva

DNA was extracted from buccal epithelial cells and white blood cells present in saliva using Oragene DNA kit according to manufacturer's guidelines. Immediately on collection saliva was mixed with Oragene DNA solution. At the time of extraction the mixture was incubated at 50°C for one hour, mixed 1/25 with Oragene DNA purifier, vortexed, incubated on ice for ten minutes, and centrifuged at a minimum of 3,500g for ten minutes. Following transfer of the supernatant to a clean tube an equal volume of ethanol was added and incubated at room temperature for ten minutes. The sample was centrifuged at a minimum of 3,500g for ten minutes, the supernatant discarded, and the DNA pellet washed with 1 ml 70% ethanol. The ethanol was removed and the DNA pellet rehydrated in 1ml TE buffer.

3.3.1.3. Nucleic acid Quantification

All nucleic acids were quantified by analysis of 1µl of the sample on a NanoDrop ND1000 spectrophotometer (ThermoScientific). The OD of the sample at 260nm, the absorption wavelength of nucleic acid, is read and used to calculate the concentration. The optical density (OD) at 280nm, the wavelength at which protein absorbs, is also measured and the 260:280 ratio given as an indication of sample purity.

3.3.2. Polymerase Chain Reaction (PCR)

3.3.2.1. Primer Design

Oligonucleotide primers were designed to amplify the region of interest according to the UCSC reference genomic sequence. Primers were designed using the standard parameters in the Primer 3 design tool (http://frodo.wi.mit.edu/primer3/input.htm), and manual adjustment

when appropriate. Optimal parameters for primers were: 20 nucleotides long, 50% G/C content, no runs of a single nucleotide, melting temperature (Tm) of 56°C, and no self-complementarity. It was also important that pairs of primers were not complimentary to each other to prevent the formation of primer-dimers. Complimentary sites for the primers within the genome were identified using the blastn algorithm in the NCBI Basic Local Alignment Search Tool (BLAST) (http://blast.ncbi.nlm.nih.gov/Blast) and primer pairs complimentary to multiple sites within the genome discarded.

When restriction enzyme (RE) sites were to be introduced into the amplified region to enable downstream processing they were added to the 5′ end of the primer. A minimum of three additional nucleotides were also added 5′ of the RE site to ensure the enzyme was able to bind to and cleave the DNA.

3.3.2.2. PCR reaction

The standard PCR reaction consisted of:

DNA	50ng
PCR reaction buffer	1x
Taq polymerase	1.25u
Each primer	10µM
Water	s.v. 50µl

All PCR runs included a no-DNA negative control to check for reaction contamination.

3.3.2.3. PCR cycling

The basic PCR cycle used was:

Denaturation	95°C	5 min	
Denaturation	95°C	15 sec	x cycles
Primer annealing	x°C	1 min	
Extension	72°C	2 min	
Extension	72°C	10 min	

PCR product was stored at 4°C until use.

3.3.2.4. Optimisation

PCR conditions were optimised for each primer pair by performing 12 reactions over an annealing gradient of 50°C to 65°C with reactions containing each of 1.25mM, 2.5mM, and 3.75mM. $MgCl_2$. An example to illustrate this optimisation process is shown in Figure 3.6. When necessary, further optimisation over additional $MgCl_2$ concentrations or temperature gradients was performed.

3.3.2.5. PCR product visualisation

PCR product with 1x gel loading dye was run on an ethidium bromide agarose gel alongside an appropriate DNA ladder for size estimation. The DNA was visualised under UV light.

3.3.2.6. PCR Product Purification

All DNA purification methods used in this study used QIAGEN spin columns, which are based on silica membranes within the columns. Silica specifically adsorbs DNA under high-salt conditions but not under low-salt conditions. Adsorption buffers create the correct salt and pH conditions for optimal DNA binding for each application.

Purification of PCR products from enzymatic reactions was carried out using the QIAquick PCR purification kit according to manufacturer's instructions. All centrifugation steps were at 13,00rpm for one minute in a bench-top centrifuge. Buffer PB1 was added to the sample to a final ratio of 5:1 by volume. The mixture was centrifuged in the spin column, the flow-through discarded, 0.75ml buffer PE added, and centrifuged twice discarding the flow-through between spins. 50µl water was placed onto the membrane and left to stand for one minute prior to centrifugation to elute the purified sample into a collection tube.

3.3.2.7. PCR Product Gel Extraction

DNA extraction from agarose gel was performed using the QIAquick gel extraction kit according to manufacturer's instructions. All centrifuge steps were at 13,00rpm for one minute in a bench-top centrifuge. The gel section containing the DNA fragment of interest was excised and incubated at 50°C with three volumes (100mg ≈ 100µl) buffer QG until the

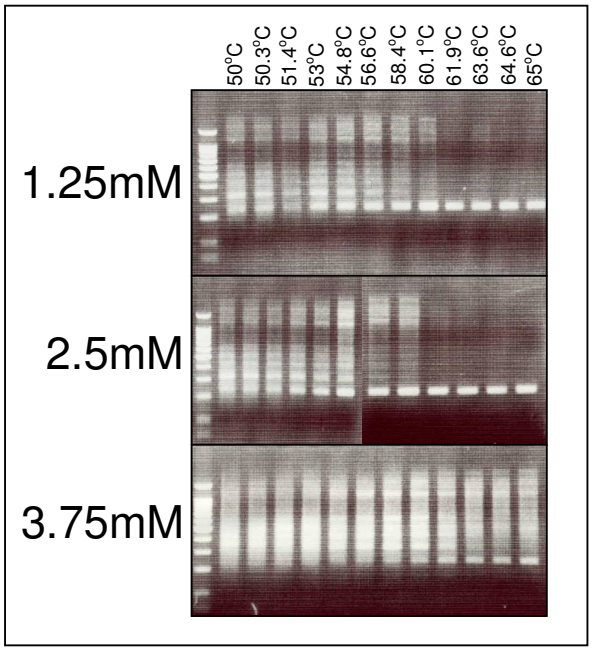


Figure 3.6 PCR condition optimisation

Optimal PCR conditions are specific to every primer pair so optimisation of each reaction is necessary. Reactions were performed over a range of primer annealing temperatures, 12 reactions ranging between 50°C and 65°C, and containing different $MgCl_2$ concentrations.

gel was completely dissolved. The sample was pipetted into the spin column, centrifuged, the flow-through discarded, 0.5ml buffer QG added to the column, and centrifuged again. 0.75ml buffer PE was added, the column centrifuged, the flow-through discarded, and the sample centrifuged again. To elute the sample, 30µl water was placed onto the membrane, left for one minute, and centrifuged.

3.3.3. Genotyping of Healthy Controls

The SNPs were genotyped using the genotyping service provided by KBioscience UK (section 3.1.3.1.), with no DNA water controls included in all the plates genotyped.

3.3.4. Cloning

3.3.4.1. Generation of insert

The region of interest was PCR amplified using specific primers including restriction enzyme recognition sequences. To generate compatible ends the genomic DNA PCR product and cloning vector were double digested with 5u each *Bgl*II and *Mlu*I. The digested vector was phosphatase treated to prevent it re-circularising during ligation using calf intestinal alkaline phosphatase (CIAP) following the manufacturer's recommended protocol for 5′ recessed ends. To all 20µl of the digestion reaction 1x CIAP reaction buffer, and 5u CIAP was added, s.v. 50µl with water, incubated at 37°C for 15 minutes and 56°C for 15 minutes. An additional 5u CIAP was added and the reaction incubated for another 15 minutes at 37°C and 15 minutes at 56°C. Both the vector and insert were purified to remove all enzymes and excised nucleotides.

3.3.4.2. PCR Product Ligation into Plasmid

The genomic PCR product was ligated into the cloning vector in a 10µl reaction containing 1x ligase buffer and 3u T4 ligase. The amount of insert included in the reaction was calculated for 100ng vector using the following formula:

$$\frac{\text{ng of vector x kb size of insert}}{\text{kb size of vector}} \times \text{molar ratio of } \frac{\text{insert}}{\text{vector}} = \text{ng of insert}$$

A ligation reaction without insert was also performed as a negative control following the

transformation step. The ligation reaction was incubated overnight at 4°C. Prior to transformation the efficiency of the ligation was checked by running 5µl of the reaction on a 1% agarose gel.

3.3.4.3. Plasmid Transformation into Competent Cells

Subcloning efficiency DH5α competent cells were used according to manufacturer's instructions. Competent cells were thawed on ice and aliquoted into cold tubes at 50µl per reaction. For both the reaction with insert and the no insert negative control 10ng of the overnight ligation reaction, or 250ng pUC19 vector as a positive transformation control, was added to the cells. The cells were incubated with the plasmids for 30 minutes on ice, heat shocked in a 42°C water bath for 20 seconds, and placed back on ice for two minutes. To each sample 250µl pre-warmed LB broth was added and incubated for one hour at 37°C with shaking at 225rpm. Of the transformed cells, 100µl were spread onto agar plates containing 100µg/ml ampicillin and incubated overnight at 37°C.

3.3.4.4. Plasmid amplification

Selected colonies from the plated transformed cells were picked using an inoculation loop and transferred to 5ml LB broth containing 100µg/ml ampicillin. Cultures were incubated at 37°C overnight with shaking at 225rpm.

3.3.4.5. Colony PCR

To confirm the cultured bacterial colonies contained the genomic insert a colony PCR (cPCR) was performed with 0.5µl of the overnight culture and the same PCR conditions used to amplify the genomic insert. For visualisation of the insert, 10µl of the PCR product was run on a 1% agarose gel.

3.3.4.6. Small scale plasmid preparation 'Miniprep'

Plasmid DNA was purified from bacterial cells using QIAprep Spin Miniprep kit following manufacturer's instructions. Unless otherwise stated all centrifugations were at 13,000 rpm using a bench-top microcentrifuge. Bacterial cells from 1.5ml overnight cultures were pelleted by centrifugation at 14,000rpm for ten minutes, resuspended in 250µl buffer P1, transferred to a microcentrifuge tube, and mixed with 250µl buffer P2 by six inversions. Folowing addition of 350µl buffer N3 the sample was inverted a further six times, the reaction centrifuged for ten

minutes, the supernatant applied to the spin column and centrifuged for one minute. The flow-through was discarded, 0.75ml buffer PE added to the spin column to wash the DNA, followed by centrifugation for one minute and the flow-through being discarded. To remove any residual wash buffer the spin column was centrifuged again for one minute. In a collection tube 50µl water was applied to the centre of the column. After one minute of incubation the column was centrifuged for one minute to elute the DNA.

3.3.4.7. Sequencing

Plasmid DNA was sent to Beckman Coulter Genomics for sequencing using the vector contained primer sites RVprimer3 and GLprimer2. The sequence chromatograms for each sequencing reaction were analysed using Gene Codes DNA sequence analysis software Sequencher v. 4.6 (www.genecodes.com). The sequences were aligned against the reference genomic sequence to identify the SNP positions.

3.3.5. Cell culture

All cell culture was performed under sterile conditions inside a class II safety cabinet in a containment level 2 tissue culture room. Cell cultures were propagated at 37°C, 5% CO_2.

3.3.5.1. Peripheral Blood Mononuclear Cells (PBMCs)

PBMCs were cultured in complete RMPI 1640 supplemented with 5% FCS.

3.3.5.2. THP-1 Cells

THP-1 cells were maintained between $5x10^4$ and $8x10^5$ viable cells/ml in complete RMPI 1640 supplemented with 10% FCS.

3.3.5.3. Cell counting

Cell suspension was mixed with an equal volume 0.4% Trypan Blue and loaded onto a hemocytometer. The intact cell membrane of live cells excludes the dye while the cell membrane of dead cells is no longer intact, enabling the dye to penetrate the cell and stain them blue. Un-stained viable cells in the central $1mm^2$ square in both chambers were counted, averaged, and multiplied by the dilution factor. The total volume of the counted area was

100nl (area 1mm^2 x depth 0.1mm). To convert this to cell concentration/ml the cell count was multiplied by 10^4.

3.3.6. PBMC isolation

PBMCs are isolated from blood through the use of density sedimentation through an isoosmotic medium with a density of 1.077g/ml. The majority of mononuclear cells have densities below 1.077g/ml and so they are retained at the sample/medium interface while erythrocytes and polymorphonuclear cells, which have higher densities, sink through the medium.

Whole blood was mixed with preservative free heparin immediately after being taken. PBMCs were isolated using the density gradient media Lymphoprep according to manufacturer's instructions. The blood was diluted with an equal volume of media and gently layered onto Lymphoprep solution at half volume of the blood dilution. The sample was centrifuged at 1,800rpm at room temperature for 25 minutes with no brake, the interphase containing the cells transferred to a new tube, mixed with medium to a total volume of 40ml, and centrifuged at 2,500 rpm for 20 minutes. The supernatant was removed, the cells resuspended in 10ml media, and counted.

3.3.7. Enzyme Linked Immunosorbant Assay (ELISA)

ELISA enables quantification of a specific antigen within a sample, immobilised on a solid support, through binding of an antibody specific for the antigen of interest. By using a biotinylated antibody to allow horse radish peroxidase (HRP) conjugation, and adding a chromogenic substrate oxidised by the peroxidise enzyme, the intensity of the resulting colour development is proportional to the amount of antigen. By including a standard at known concentrations a standard curve can be generated and the unknown antigen concentration quantified. The principle of ELISA is illustrated in Figure 3.7.

Each sample was measured in duplicate and all incubations were at room temperature. Blocking solution was used at 200µl/well and stop solution at 50µl/well, all other reagents were at 100µl/well. Each plate wash was performed in triplicate using 300µl/well 0.05% Tween-20 in PBS.

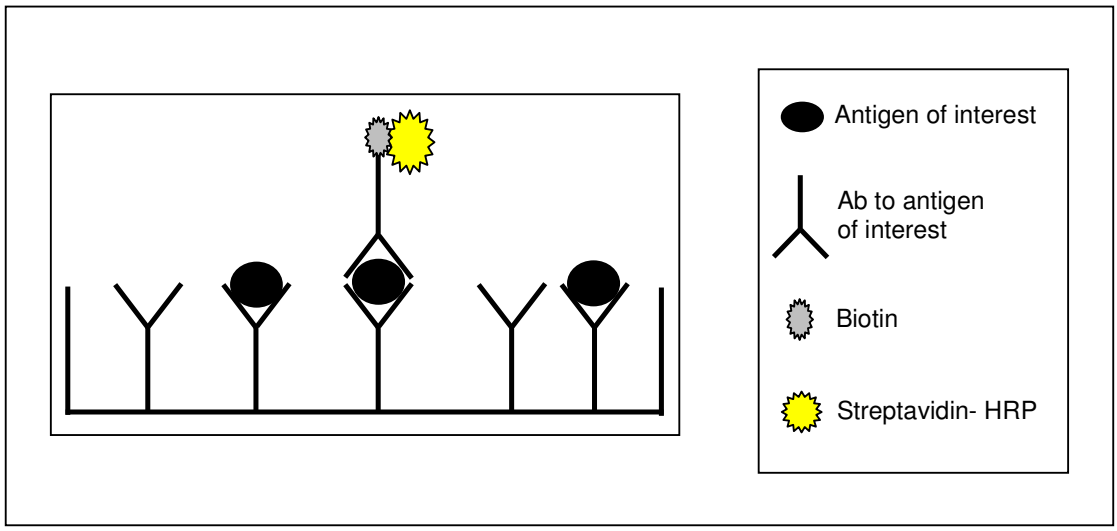


Figure 3.7 ELISA principle

In a sandwich ELISA the solid surface of the plate is coated with an antibody directed against the antigen of interest (capture antibody), to which the sample is exposed and the antigen of interest bound. A second antibody directed against the antigen of interest (detection antibody), generated in a different species to the capture antibody, is then added, also binding directly to the antigen of interest. The detection antibody is conjugated to biotin to enable subsequent conjugation of HRP, allowing colorimic quantification of the antigen of interest.

Ab- Antibody, HRP- Horseradish Peroxidase

3.3.7.1.Sandwich ELISA

Capture antibody in PBS was used to coat the wells of an ELISA plate and incubated overnight. The plate was washed and blocked for one hour with blocking solution. Following three plate washes standard or sample was added and incubated for two hours. The plate was washed again and the detection antibody added to each well and incubated for two hours. The plate was washed and incubated with streptavidin-HRP, diluted to 1 in 200 in reagent diluent, for twenty minutes out of direct light. The plate was washed and 3,3',5,5'-tetramethylbenzidine (TMB) substrate added to each well. Following incubation out of direct light for twenty minutes 50µl per well stop solution was added and the plate gently tapped to ensure through mixing. The OD of the samples was measured at 450nm and 550nm on a plate reader.

3.3.7.2. ELISA Analysis

Prior to analysis the OD readings for each replicate was corrected for imperfections in the plate surface by subtracting the OD at 550nm from the OD at 450nm. The amount of protein present in the samples was calculated using a standard curve of known concentrations.

3.3.7.2.1. Standard Curve

Standard curve analysis was carried out in GraphPad Prism v5.03 using the average OD reading of the replicates. A standard curve was generated by log transforming (X=Log(X)) the concentrations of the standards used and plotting these against the OD readings by nonlinear regression using a sigmoidal dose-response (variable slope) curve. The unknown concentrations of the samples were interpolated from their OD readings using the standard curve, and anti-Log transformed (X=10^X) into concentration.

3.3.7.2.2. Statistical analysis

The expression by haplotype results were analysed by one-way analysis of variance (ANOVA) to test for significant variation between haplotype groups. ANOVA tests if the variation between groups is greater than the residual variation within groups due to differences between individuals. The residual variation is calculated as the sum of squares of the differences between each value and the group mean, and the variation between groups as the sum of squares of the differences between the group means and the grand mean adjusted

for group size. One-way rather than two-way ANOVA was used so that post-tests could be performed to examine the variance between each group. Turkey-Kramer post-tests were used to test between all pairs of groups in a step-wise manner. It is still valid to perform post-tests on the data even if the ANOVA does not give a significant result as the post-tests are more focused and may identify differences between individual groups.

3.3.8. RNA based protocols

3.3.8.1. Total RNA extraction

Total RNA was extracted from cultured cells using ethanol precipitation following preservation in mono-phasic phenol and guanidine isothiocyanate TRIZOL solution which maintains the integrity of the RNA while disrupting cells and dissolving cell components.

Unless otherwise stated all centrifugations were at 4°C. Samples stored in TRIZOL were thawed in a 37°C water bath, centrifuged at 12,000g for ten minutes to remove any insoluble material, transferred to a new tube, and incubated at room temperature for five minutes. To each sample 0.2ml chloroform was added, shaken by hand for 15 seconds, and incubated at room temperature for three minutes. After centrifugation at 12,000g for 15 minutes the aqueous phase was removed. An additional 0.5ml chloroform was added to it and mixed, incubated, and centrifuged as before. In a new tube the aqueous phase was mixed with 1µl glycogen and 0.5ml isopropanol. This was vortexed and left at room temperature for ten minutes before centrifugation at 12,000g for 15 minutes. The supernatant was removed, 1ml 75% ethanol added to the RNA pellet, the pellet dislodged by vortexing, and centrifuged at 7,500g for five minutes. The ethanol was removed, the sample centrifuged at room temperature for one minute at 13,000 rpm, the remaining ethanol removed, and the pellet left to air dry for five minutes. The RNA pellet was then resuspended in 100µl dH_2O and quantified.

3.3.8.2. Reverse Transcription

cDNA was generated from RNA within the sample using QuantiTect Reverse Transcription kit. An optimised mix of oligo-dT and random primers is annealed to the single-stranded mRNA and the RNA-dependent DNA polymerase activity of the reverse transcriptase transcribes cDNA from the RNA template. The RNA in the resulting RNA:DNA hybrid is then degraded by the RNase activity of the enzyme.

On ice 100ng RNA was mixed with 2µl gDNA wipeout buffer (s.v. 14µl water), heated to 42°C for two minutes and placed immediately on ice. To the cDNA 1µl reverse transcriptase, 4µl RT buffer, and 1µl RT primer mix was added. The samples were incubated at 42°C for 15 minutes and 95°C for three minutes. cDNA was stored at -20°C until used.

3.3.8.3. Quantitative RT-PCR

To eliminate co-amplification of any contaminating genomic DNA, primer assays contain primers designed to cross intron/exon boundaries of the transcript of interest, based on gene sequences in the NCBI Reference Sequence database. A PCR reaction containing the fluorescent dye SYBR Green 1, which binds all double-stranded DNA and emits a fluorescent signal at 521 nm on binding, is performed using the cDNA as the template. As the number of double stranded cDNA copies increases with each amplification cycle the amount of fluorescence increases at a rate in relation to abundance in the sample.

To 2µl of each cDNA sample 10µl QuantiTect SYBR Green PCR master mix, 2µl primer assay, and 6µl water was added. Samples were run on an Eppendorf realplex Mastercycler with a PCR cycle of: 95°C for 15 minutes, 40 cycles of 94°C for 15 seconds, 55°C for 30 seconds, and 72°C for 30 seconds. Each sample was analysed in duplicate.

To check the specificity of each primer pair, a melting curve is performed on the qRT-PCR reaction end product. If the primer pair is specific all the amplicons will be identical, and therefore have the same melting temperature. The temperature of the reaction is slowly raised and the SYBR green fluorescence intensity monitored. When the correct melting temperature is reached the amplicons will separate into two single strands, resulting in a drop in fluorescence intensity. By plotting the negative first derivative of the melting curve as a percentage of the highest peak (-dI/dT (%)) as a function of temperature, a peak will be present at the melting temperature of the amplicons. In a population of identical amplicons a single peak will be produced while multiple peaks will be produced in a heterogenous amplicon population.

The end point measure of qRT-PCR is the threshold cycle, $C_{T.}$ This is the PCR cycle during exponential amplification at which the SYBR Green signal crosses an arbitrarily selected threshold. The threshold for each primer assay was kept constant across all plates. The less abundant the initial mRNA template the more PCR cycles needed to reach the threshold. This

means that C_T is inversely related to the amount of mRNA in the sample, the lower the C_T the higher the mRNA concentration. C_T for each sample was determined using the Eppendorf RealPlex software.

3.3.8.3.1. Analysis

Results of qRT-PCR assays were analysed using the comparative C_T method. To correct for any differences in total RNA amounts between samples the mean C_T for the gene of interest was expressed relative to the mean amount of a housekeeping gene, ΔC_T. The housekeeping gene was selected for being expressed at a constant level regardless of culture conditions or origin of sample, and so can be used as an internal control. Results were expressed as fold change from either the housekeeping gene or a calibrator sample. Appropriate calibrators to compare other samples to include a zero time point or an untreated sample. The equations used to calculate each of these steps are shown below.

$$\Delta C_T = \text{mean } C_T \text{ gene of interest} - \text{mean } C_T \text{ housekeeping gene}$$

$$\Delta\Delta C_T = \Delta C_T \text{ sample of interest} - \Delta C_T \text{ calibrator sample}$$

$$\text{Expression relative to housekeeping gene} = 2^{-\Delta CT}$$

$$\text{Expression relative to calibrator} = 2^{-\Delta\Delta CT}$$

3.3.9. Electrophoretic Mobility Shift Assay (EMSA)

EMSA enables visualisation of protein bound to a DNA sequence of interest, working on the principle that DNA bound to protein will migrate at a slower rate under electrophoresis than DNA alone. The principle of the EMSA technique is illustrated in Figure 3.8.

3.3.9.1. Nuclear extraction

THP-1 cells at $6x10^5$/ml were cultured in T75 flasks with 100nM (62ng/ml) PMA in DMSO for 24 hours, 2% FCS RPMI for 24 hours, and three hours either with or without 20ng/ml IFNγ. The cells were washed twice with ice cold 1x PBS and placed on ice. To each flask 3ml cytoplasmic lysis buffer was added and the cells removed using a cell scraper. The cells were split into 1ml aliquots and 90µl NP40 added. The tubes were gently shaken to mix and

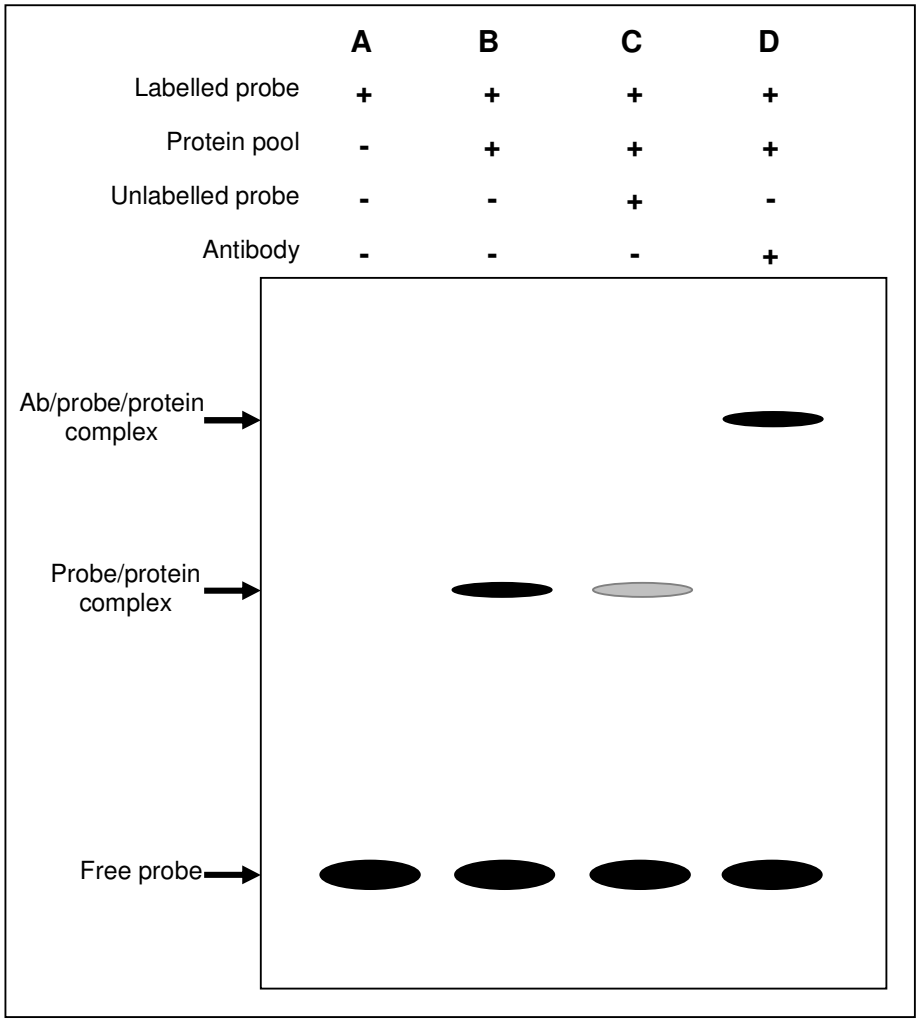


Figure 3.8 Electroporetic Mobility Shift Assay principle

The DNA sequence of interest is radioactively labelled for visualisation (free probe, lane A) and incubated with a mixture of proteins to allow binding to occur (lane B). The double stranded alternating copolymer poly d(I)D(C) is included in the binding reaction to block any proteins which non-specifically bind to DNA. To confirm specificity of any resulting bands observed an excess of unlabelled probe is also included in the binding reaction (cold competition, lane C). The DNA binding protein will preferentially bind to the more abundant unlabelled probe resulting in a decreased intensity of the band visible on the gel. As larger protein complexes will have a slower migration rate addition of an antibody directed toward the DNA binding protein causes the complex to be retarded, or shifted, in the gel compared to the DNA and binding protein alone (lane D), enabling identification of the binding protein.

left on ice for two minutes. Following centrifugation at 14,000rpm at 4°C for 30 seconds the supernatant was discarded and 40µl nuclear lysis buffer added to the nuclei pellet. The tube was gently flicked to dislodge the pellet, the buffer removed and the wash repeated. The pellet was then resuspended in 40µl nuclear lysis buffer and the reactions for each condition pooled together. Using trypan blue the mixture was examined under a microscope to gauge nuclei concentration. The nuclei were mechanically lysed using a Dounce A homogeniser for approximately 20 minutes with regular checking by microscopy until >50% of the nuclei were lysed. The mixture was centrifuged at 14,000rpm at 4°C for ten minutes and the supernatant stored in aliquots. The protein concentration of the nuclear extracts were determined using the BCA assay, in which the colour change of acidic Coomassie® Brilliant Blue G-250 dye on binding to protein is relative to the protein concentration of the sample. Six two-fold serial dilutions of the BSA standard, high standard 10µg/µl, were used to generate a standard curve. To 300µl BioRad reagent, 2µl lysate and standard were added and the OD measured at 595nm. The protein concentration was calculated using a standard curve (section 3.1.7.3.1.)

3.3.9.2. Oligonucleotide annealing

Short 30bp probes were generated by annealing two complimentary oligonucleotides of the sequence of interest. The annealing reaction contained 10µg each oligonucleotide, s.v. 50µl with water. The reaction was placed in a water bath at 95°C for five minutes, the water bath turned off and the reaction left in the water bath until cooled to room temperature. To check the reaction efficiency, 1µl of the annealing reaction was run on a 3% agarose gel at 180V along-side 1µl unannealed oligonucleotides.

3.3.9.3. Probe amplification

Probes of 203bp were PCR amplified from plasmids previously generated by Dr Wen containing the region of interest cloned from genomic DNA containing each of the two alleles (See Appendix 3).

3.3.9.4. Probe labelling

A labelling reaction of the following components was heated at 30°C for 30 minutes.

Probe	10pmol ends
T4-Kinase buffer	1x
T4 polynucleotide kinase	10u
^{32}P-γ-ATP	2µl
Water	s.v.25µl

The concentration of double stranded probe used was calculated using the following formula:

$$\text{pmol DNA} \times \frac{660\text{pg}}{\text{pmol DNA}} \times \frac{\text{pmol DNA}}{10^3\text{pg}} \times \text{DNA bp} = \text{ng DNA}$$

Unincorporated radioactive nucleotides were removed with sephadex matrix G-50 Micro column gel filtration system. Larger DNA molecules do not penetrate the matrix and are eluted while <20bp fragments, including unincorporated nucleotides, are retained within the gel. The column was vortexed to resuspend the sephadex resin and centrifuged at 735xg for one minute to remove the resin storage buffer. The labelling reaction was mixed with 25µl probe buffer and added to the top-centre of the resin. The purified probe was collected by centrifugation at 735xg for two minutes.

3.3.9.5. Protein binding

All binding reactions contained the following components in a total volume of 20µl with water:

Labelled probe	0.4pmol ends (short probes)/ 0.12pmol ends (long probe)
poly dI-dC	2.5µg
Parker Buffer	1x

Where appropriate, nuclear extract was added, and an excess unlabelled probe used for cold competition. The amount of unlabelled probe used was calculated using the formula in section 3.3.9.4. Cold competition reactions containing unlabelled probes were incubated at 21°C for 30 minutes prior to addition of labelled probe. For supershift antibody reactions the antibody, reaction buffer, and nuclear extract were incubated at 4°C overnight prior to binding to labelled probes. Final reactions were incubated for 30 minutes at 21°C. 17µl Parker buffer mixed with 3µl gel loading dye was ran in the wells either side of the samples.

3.3.9.6. Visualising DNA-protein complexes

Acrylamide gels were pre-run for one hour at 4°C prior to loading of 20µl each reaction. The gel was run at 4°C until free probe was detected in the lower buffer reservoir. The gel was dried for one hour under vacuum at 80°C and exposed to radiographic film with an intensifier screen.

3.4. *IL18BP* candidate region

3.4.1. Investigation of expression levels according to haplotype

Dr Wen within the group performed an *in vitro* transcription study of a 2kb region including both of the associated haplotype SNPs, and all of the regulatory elements in the *IL18BP* promoter experimentally validated by Hurgin et al (Hurgin et al., 2002). pGL3 expression vectors containing constructs of each of the four possible haplotypes were transiently transfected into HeLa cells, which were then either stimulated with 20ng/ml IFN-γ or left unstimulated for a further 24 hours. Luciferase reporter assays showed a significant ($p<0.0001$) effect of the haplotype in stimulated cells. The 1-1 haplotype showed the highest level of transcription, the 2-2 haplotype the lowest transcription level, and 1-2 and 2-1 haplotypes intermediate levels of transcription. The level of transcription with the construct containing the 2-2 haplotype was not significantly different from the vector without insert ($p=0.262$). Full details of this study are given in Appendix 3.

To extend these findings a study was performed as part of this thesis to investigate if the haplotype effect on expression was also evident in PBMCs of healthy individuals using the WC cohort of healthy individuals (section 2.1.1).

3.4.1.1. Haplotype determination

Based on the genotypes of the two *IL18BP* SNPs the haplotypes of each individual, homozygous for at least one SNP, were inferred manually.

3.4.1.1.1. Haplotype determination of uncertain phase

One of the WC individuals was heterozygous at both *IL18BP* SNP1 and SNP2. Therefore, haplotype phase could not be determined for this individual with any certainty; they could carry the combination of the 1-1 and 2-2 haplotypes, or the 1-2 and 2-1 haplotypes. To

ascertain which haplotypes this individual carried, a region of their genomic DNA containing both of the SNPs, was cloned into a vector and sequenced. During cloning only a single genomic PCR amplicon, a copy of one of the chromosomes, is inserted into the vector, enabling separation of the two chromosomes and identification of the alleles physically located together. The PCR primers previously designed by Dr Wen for the gene transcription study (Appendix 3) were used. The sequence amplified, primer positions and sequences, are shown in Figure 3.9. Although an expression vector was not necessary for haplotype determination the pGL3-Basic (Figure 2.2) vector was used. This is because the PCR primers had been designed for use with this vector and the cloning protocol previously optimised by Dr Wen.

The PCR reaction consisted of: 1x HotStart PCR Buffer, 1µM each IL18BP-1325F and IL18BP+577R primers, 2.5u HiFidelity polymerase, 50ng genomic DNA, 1.25u Taq polymerase, s.v. 50µl with water. 35 PCR cycles with an annealing temperature of 56°C was used. As non-specific bands were also amplified the sequence of interest was gel extracted (section 3.3.2.7.). The vector and PCR product were double digested: 1µg DNA, 1x Promeaga buffer D, 2µg BSA, 5u *BglII*, 5u *MulI*, s.v. 20µl with water, the digested vector phosphatase treated (section 3.3.4.1), and both vector and insert PCR purified (section 3.3.2.6.). The amount of insert used in the ligation reaction was calculated as:

$$\frac{\text{100ng vector x 2kb insert}}{\text{4.8 kb vector}} \text{ x 1:1 molar ratio of } \frac{\text{insert}}{\text{vector}} = \text{41.7ng of insert}$$

The ligated plasmid was transformed into competent cells and cultured (section 3.3.4.3.-4.). The cPCR reaction was altered by substitution of the HiFidelity polymerase for GoTaq® DNA polymerase, and the PCR cycles reduced to 25. Five plasmid clones containing insert were extracted by Miniprep and sequenced (section 3.3.4.6.-7.). The sequences for each of the SNP containing regions were aligned against the genomic reference sequence. The alleles present at both loci in the five clones were identified and used to infer haplotype.

3.4.1.2. PBMC stimulation

As part of the *in vitro* transcription study Dr Wen previously performed 20ng/ml IFNγ stimulation of THP-1 cells over a time course of 0, 0.5, 1, 3, 6, and 8 hours. This showed maximal RNA levels after 6 hours stimulation and maximal protein levels after 24 hours

A.

IL18BP-1325F →

gcattgg**GTTTCTAGGGAATATGGCTCAG**tagggcctgccccttcaagggcagcaagaacgaagcagg
atgttgtcactcccgtctgaataaagctccatgagctgggttggaaagctgaacatttatcttgctttcttcttcccacctctt
tccctggggtccccaggggtgtcttgcctttgttaacccccatcggagacctgagggtctgggggtggttgggcctgc
cttttgctgttcccagctgttcccctgctgacatccctaggagtggtggaatgccaggctcctgtcctgtggcctggggg
cctggggcatcattgggcctgccttcagactgaggttgctctatgttctcactgcctgaggacagtgtcccgccaccac
ctgccttcctgaggaccacactcaggctcggcttctcttccccagctgtggcttcccatggtaccctgctgcttatgcca
cctgtggggctgttcccctcctcagagaggggaatggggctggtcagaggacctagggatcacccttccttttacttttt
ccccccaccccctgtcaggctcaggtgagaagccgaggccaggctgtccaaagctcttttcatgtgtaacagttctg
agtcccttgacctcccccttggcctgactctggtctcggttccactaagacagagtcctgtcatctttccctaccatccc
ccactggactgtctctctcagaaaccccaggcagccttgc**[C/T]**tcctgacctaactcctggagatgctggtgttctt
gggaggaatttctcctgcatcttacatgctagaagcccctggggcaggaggtaggggaggaattgagaggaggaa
aaacatgaactagacacctagagaagaaggatgtgacttgtagtatcctatgtctaaattaggaatatgaatctggttttt
ctacaagaagtttgagatcacagctgactgtgttcctgatgcatccaccaaacccagttccatctgtgggcctccctggc
tctgtcaccagccgttgcaccctcccaatcacaggagtcacaaacctcagacatgcagctcctgtccacacttaatatat
gcatgcattggatcacccagccctggtctttctgcctccatggataactgcatgaccctgagagaaaacctccttagatt
tagcatcctaggttcctcacacgcctcaccctgaatcctggccctcccgcagccccagcgccatttgtcccatcagtga
caagattcatattctgatgtagactctgttgccagagccagtgttgagccagtccgcctcttccccgggaagtgcctgcc
cttccctcctgttagggttggctctcgagcttgtgtgccagttcctgggttggccgtgagagttctacagacaaggagg
aagtgctctcggtgtatttcctgtggtgggttcacacgcagctagacacagctaacttgagtcttggagctcctagagg
gaagcttctggaaaggaaggctcttcaggacctcttaggagccaggtaggagtctgggactactagtgaacctagac
ctgtggctctggccagaggggctaggatgagagacagagggtgtgatggtgggtgctgggagatgtagccgacctt
ggggctggtggctgggggagtgggtagcctgggaaaggccaggatgtggacggactggtatggcattgagcctga
agtggtccaacttggggttccccagtgcctaggaaagttgtccccttgaatgtcagtgtgaaggtgaaggaggaagca
gatgcctgttcatatggaaacaaagacctggctgtgaagaggggaggcggacaccaaagtcctgacacttgggcgg
gacagaattgatctgtgagagactcatctagttcataccctaggtgaccctgggggtggcatgggggtagattag[**A/**
C]agatcccagtctggtatcctctggagagtaggagtcc**CAGGAGCTGAAGGTTTCTGGCCACTG**aac
← **IL18BP+577R**

B.

IL18BP-1325F	CCAACGCGTttctagggaatatggctcag *MluI*
IL18BP+577R	CCAAGATCTccagaaaccttcagctcctg *BglII*

Figure 3.9 IL18BP cloning region for haplotype determination

The *IL18BP* region containing SNPs 1 and 2 (in square brackets) is shown in A. 5′ to 3′ respectively. The positions of the two primers used to amplify the region are also shown (in capitals) with direction indicated. The 5′ to 3′ sequences of the primers are given in B. The bases non-complimentary to the genome are shown in capitals with the introduced restriction enzyme recognition sequences underlined.

stimulation. These time points were used to inform the design of the PBMC experiments described below. PBMCs were used from WC cohort healthy individuals homozygous for the 1-1 haplotype (n=3), heterozyogous for 1-1 and 1-2 (n=3), heterozygous for 1-1 and 2-1 (n=2), homozygous for the 2-1 haplotype (n=2), and heterozygous for the 1-2 and 2-1 haplotypes (n=1). Isolated PBMCs (section 3.3.6.) were cultured for the stated lengths of time in 2ml RPMI supplemented with 5% FCS and 20ng/ml IFNγ. Cells were lysed in TRIZOL and stored at -80°C until use, as was the conditioned media. In this study the same PBMC samples were used for measurements of both protein and RNA levels. To ensure that maximal induction of both protein and RNA were included, stimulation time points of 0, 3, 6, 24, and 48 hours were performed.

3.4.1.3. Protein expression levels

Corbaz et al (Corbaz et al., 2002) previously published a study investigating the effect of different stimuli on the expression of IL-18BP from PBMCs after 48 hours. They investigated the effect of stimulation with IL-18 (20ng/ml), TNFα (10ng/ml), IFNγ (20ng/ml), IL-1β (10ng/ml), and LPS (100ng/ml), alone and in combination. They showed that IFNγ produced the maximal level of expression induction which was not enhanced by the addition of further cytokines.

3.4.1.3.1. Optimisation of cell concentration

Corbaz et al (Corbaz et al., 2002) used $2x10^7$ cells/ml from healthy individuals and showed IL-18BP expression of approximately 12ng/ml after 48hrs stimulation.

An initial experiment to optimise the PBMC concentration was performed using PBMCs from two individuals not used in the expression experiment, and one individual included in the expression by haplotype analysis. The PBMCs were cultured at $5x10^6$, $2.5x10^6$, and $1x10^6$ cells per ml.

The sandwich ELISA protocol (section 3.3.7.1.) was followed using the IL-18BPa Duoset ELISA kit with manufacturer's recommended solutions and dilutions. The blocking and reagent diluent solutions used were both 1% BSA. Capture antibody was used at 2μg/ml, detection antibody at 200ng/ml, and a streptavidin-HRP step included.

The protein concentration in the conditioned media was measured in triplicate by ELISA.

Two, seven two-fold serial dilutions with high standards of 4,000pg/ml and 3,000pg/ml of the recombinant human IL-18BPa standard were used to generate a standard curve. The mean of the replicates for each sample were taken as the final results. When dilutions of the samples were used, each replicate was corrected for the dilution factor and the average of all replicates taken.

3.4.1.3.2. Expression comparison between haplotypes

PBMCs were isolated and stimulated at $2.5x10^6$ cells/ml and the levels of IL-18BP produced by each of the samples measured by ELISA, (section 3.3.7.1. and 3.3.7.3.). Cell supernatants from samples stimulated for 0, 3, or 6 hours were plated undiluted, samples incubated for 24 and 48 hours were diluted 1 in 2 and 1 in 4 in reagent diluent. Two, seven two-fold serial dilutions with high standards of 4,000pg/ml and 3,000pg/ml were used to generate a standard curve.

Prism v5.03 was used to perform one-way ANOVA analysis (section 3.3.7.3.2.) comparing expression levels between the different haplotypes, excluding the 1-2/2-1 haplotype for which there was only one individual, for each of 24 hours and 48 hours stimulation time points.

3.4.1.4. RNA expression levels

Total RNA was extracted from TRIZOL stored PBMCs and qRT-PCR performed (section 3.3.8.1.-3.).

3.4.1.4.1. Validation of housekeeping genes

As the comparative C_T method standardising the samples to a housekeeping gene was used to analyse the RNA levels according to haplotype it was important to select a housekeeping gene expressed at a constant level, regardless of stimulation. Total RNA from one of the individuals used in the cell concentration optimisation experiment at $2.5x10^6$ PBMCs/ml was used to measure the RNA levels of five housekeeping genes.

The housekeeping genes tested were actin β (*ACTB*), glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*), hypoxanthine phosphoribosyltransferase 1 (*HPRT1*), and the large 32 and large P0 ribosomal proteins (*RPL32* and *RPLP0*).

3.4.1.4.2. Expression comparison between haplotypes

Total RNA from the stimulated PBMCs of the 11 WC individuals used in this study (section 3.4.1.2.), were used to measure the RNA levels of *IL18BP* and the housekeeping gene. The results were analysed for differences according to haplotype by one way ANOVA (section 3.3.8.3.1. and 3.3.7.3.2.).

3.4.2. Transcription factor binding

In order to experimentally investigate the predicted allele specific TF binding to *IL18BP* SNP1 (rs3814721) EMSA experiments were performed.

3.4.2.1. Probe design

Probe design was performed by Dr A.Wen.

EMSA probes of 30bp (Figure 3.10 A.) were designed to include *IL18BP* SNP1, the predicted SP1 binding recognition sequence, and additional bases either side of the recognition sequence to ensure protein binding could take place. Probes were generated by annealing complementary synthesised oligonucleotides (section 3.3.9.2.).

Larger 203bp probes, extending approximately 100bp either side of *IL18BP* SNP1 (Figure 3.10 A.), were also designed to allow investigation of any protein binding at the SNP1 locus, in the presence of other protein binding in the surrounding sequence. Primers were designed as stated in section 3.3.2.1. Probes were generated by PCR amplification of genomic sequence in plasmids previously generated by Dr Wen (section 3.3.9.3.).

3.4.2.2. Short probe EMSAs

As well as labelled probe alone, a total of eight EMSA reactions were performed (section 3.3.9.) with each of the short 30bp probes containing either the C or T allele at SNP1 (Figure 3.10 A.). Each probe was bound to nuclear extract, cold competed with both the C and T allele containing short unlabelled probes, and incubated with an SP1 antibody. All reactions were performed with nuclear extract from both unstimulated and IFNγ stimulated THP-1 cells (section 3.3.9.1.). Reaction components were used at: 0.4pmol ends labelled probe, 20µg nuclear extract, 80pmol ends unlabelled probe (200 fold excess), and 2µg SP1 antibody.

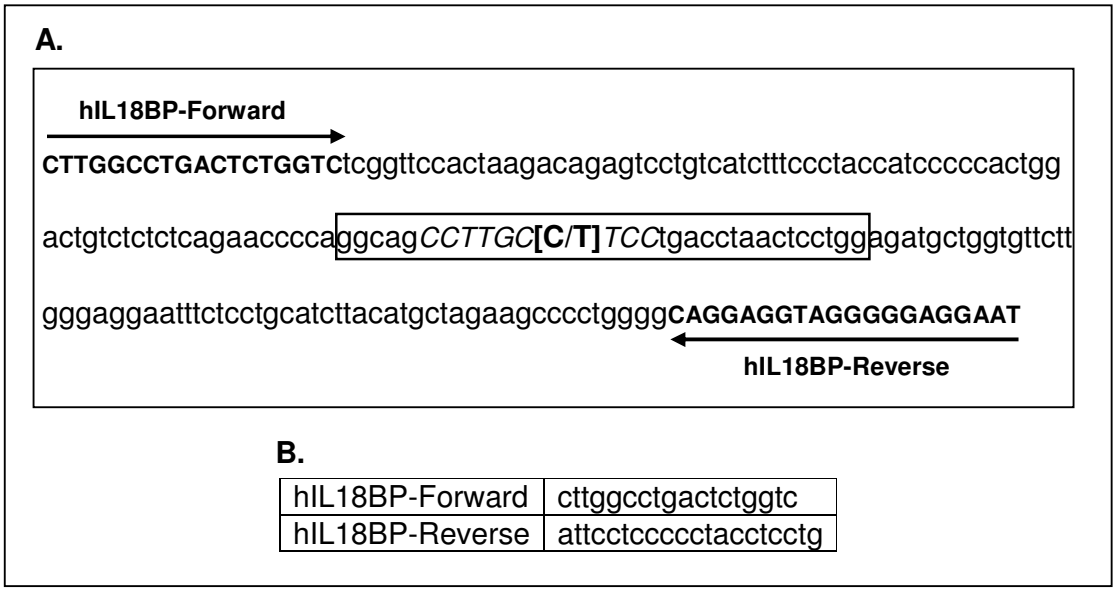


Figure 3.10 *IL18BP* SNP1 EMSA probes

To investigate the presence of allele specific protein binding, probes (A.) including SNP1 (in square brackets) and the predicted SP1 binding sequence (italicised capitals) were designed. Short 30bp probes (boxed) and long 203bp probes were generated. The position and direction of the primers used to amplify the long probe region are indicated (bold capitals). The primer sequences, written 5′ to 3′, are shown in B.

3.4.2.3. Long probe EMSAs

To ensure that all DNA-protein complexes visible with the long probes were due to sequence-specific protein binding, a titration with the non-specific competitor DNA poly dI-dC was performed. The poly dI-dC titration (1.25, 2.5, 5, 7.5, and 10μg) was performed (section 3.3.9.) with 0.12pmol ends labelled long 203bp probes (Figure 3.10 A.) and 10μg nuclear extract from unstimulated cells, s.v. 25μl with water. To ensure full separation of all DNA-protein bands for visualisation the reactions were ran on the acrylamide gel until free probe was detected in the lower buffer reservoir.

As well as labelled probe alone, a total of six EMSA reactions were performed (section 3.3.9.) with each of the long 203bp probes containing either the C or T allele at SNP1 (Figure 3.10 A.). Each probe was bound to nuclear extract from unstimulated and IFNγ stimulated THP-1 cells (section 3.3.9.1.), and cold competed (bound to nuclear extract from unstimulated cells) with both of the long and both of the short unlabelled probes. Reaction components were used at: 0.12pmol ends labelled probe, 10μg nuclear extract, 6pmol ends unlabelled long probe (50 fold excess), and 12pmol ends unlabelled short probe (100 fold excess).

3.5. Association analysis of the susceptibility alleles according to response to IL-1 blockade

Studies in which IL-1 blockade has been administered to patients with sJIA indicate that IL-1 may only be important in pathogenesis in some patients, due to the observation that only a proportion of patients show significant improvement (Quartier et al., 2006; Lequerre et al., 2008; Gattorno et al., 2008; Lovell et al., 2006; Ruperto et al., 2008). It is therefore possible to hypothesise that polymorphisms in IL-1 family genes are only involved in disease susceptibility in a proportion of the patients with sJIA. If this were the case it would be reasonable to hypothesise that significant frequency differences of IL-1 disease susceptibility SNPs would be evident between these two groups of patients. It would therefore be preferable to perform the association analysis in this study stratifying for IL-1 blockade response. However, as IL-1 blockade therapy has only been developed relatively recently, and at the time of writing is not licensed for use in sJIA, only a small proportion of sJIA patients have been treated with IL-1 blockade at Great Ormond Street hospital (GOSH). Of the 234 patients successfully genotyped in this disease susceptibility association study only

six have received treatment with Anakinra. It was therefore not possible to include response to IL-1 blockade as a parameter in the association analysis.

An open trial of Anakinra therapy in sJIA patients refractory to other treatments has recently been carried out at GOSH (EudraCT number: 2004-002439-25, Sponsor's protocol number: 04RU05). To enable genetic analysis the cohort was extended by inclusion of additional patients treated with IL-1 blockade from Utrecht University hospital, the Netherlands, Hopital Necker Enfants Malades, France, and Cincinnati Children's Hospital Medical Center, USA. These patients were not treated with IL-1 blockade as part of the same clinical trial, and are not all known to be refractory to standard therapies. Patients were classified as either responders or non-responders to treatment according to the criteria for disease remission: absence of fever and rash, no active joints, and CRP within normal range (<10).

To determine if the eight IL-1 family SNP effects identified in the disease association study segregate by response to IL-1 blockade, DNA from the 37 patients not included in the disease association study were genotyped for these alleles by KBioscience (section 3.1.2.2.), except for *IL1RAP* SNP14, which was genotyped in these patients on the Illumina GoldenGate platform as part of another project. Unphased analysis (section 3.1.1.5.) was performed to test for allele frequency differences between the two patient groups.

Chapter 4

Results

4. Results

4.1. Association Study Quality Control

4.1.1. Power Calculation

The statistical power of an association study is the probability that a true disease-association will be detected under the parameters specified. CaTS power calculator for two stage association analysis (Skol et al., 2006) was used to determine the power of the study based on the approach utilised. Disease prevalence of 0.00016 and a significance level of 0.05 were used in all calculations. Based on the samples sizes available it was specified that 50% of the DNA samples would be included in stage-1 of the study, and 5% of all the SNPs investigated, would be genotyped in stage-2. These parameters are predicted to provide 84% power to detect the effect of a disease associated SNP (MAF 0.3 and a relative risk of 1.7) in a multiplicative model. The power to detect a significant association decreases with decreasing relative risk and with decreasing risk allele frequency.

4.1.2. Golden Gate genotyping pilot study

Of the eight DNA samples included in the pilot study six showed a GC50 score >50 and were classified as successfully genotyped. Of these, five had call rates >90%, and the other 63%. The other two DNA samples had call rates <25% and GC50 scores <0.01. The two samples which did not genotype well on this platform had previous low call rates of 28% and 20%. All six of the samples with a high call frequency on the Illumina platform had previous call rates ≥50% (Table 4.1). This previous call rate of ≥50% was selected as the cut off point for a predictor of success on the Illumina platform, and used to inform selection of samples used in the association study. All samples with a previous call rate <50% were not included in the association study. This eliminated 23 of the patient samples from the repository.

4.1.3. Stage-1

Analysis of the genotyping assay results for the tSNPs in the 137 systemic JIA patients and 151 control individuals resulted in the exclusion of 12 samples due to poor genotyping call rates. This included seven patients and five controls. Twelve SNPs (3.125%) were classed as failed assays (Table 4.2) and excluded due to low call rates or confidence in clustering (GC10 Score<0.5).

	Previous studies	GoldenGate Pilot		
Sample	**Call Freq.**	**Nº Calls**	**Call Freq.**	**GC50**
BRM119	100	1049	0.97	0.83
GOS070	100	1048	0.97	0.83
GOS018	75	1047	0.97	0.83
MDX030	75	1049	0.97	0.83
BRM243	50	1011	0.93	0.79
GOS262	50	680	0.63	0.56
MDX004	28	257	0.24	<0.01
WEX081	20	159	0.15	<0.01

Table 4.1 BSPAR repository patient GoldenGate Pilot Study

Eight DNA samples from the BSPAR repository of sJIA patients were selected as representative of the range of call rates in previous studies. These samples were genotyped in a pilot study using the Illumina 1,082 SNP training array. Shown are the call frequencies of these samples in previous studies. The call frequency (out of 1,082 SNPs) and GC50 score of the same samples in the pilot study on the Illumina GoldenGate platform are also shown.

rs Number	GC10 Score
***CASP1* region**	
rs17103781	0.493
***IL1RAP* region**	
rs11916734	0
rs6798479	0
rs1195370	0.5529
rs4367119	0.7444
rs6796131	0.3483
***IL1* ligand cluster**	
rs2708947	0.4086
rs4849122	0.4975
rs2278716	0.3276
rs2234679	0.6559
rs3783512	0.3212
***IL1* receptor cluster**	
rs3218926	0.3785

Table 4.2 Genotyped SNPs classed as failed assays

Twelve SNPs genotyped on the Illumina Golden Gate platform in Stage-1 of the study were classed as failed assays due to low call rates or confidence in genotype calling. The GC10 score shown, as calculated by the Illumina BeadStudio software, is a measure of how well the samples in that SNP assay separate into three distinct genotype clusters.

Although for three of the SNPs GC10>0.5 they were still excluded because of poor cluster boundary separation (Figure 4.1). Of the SNPs that were not used in the analysis five were in the *IL1* ligand cluster, five in the *IL1RAP* region, and one in each of the *IL1* receptor cluster and *CASP1* region. Of the 276 samples with good call scores 209 had 100% call rate, and all had call rates over 96%. Of the 372 SNPs used for analysis, 279 had a 100% call rate, and all had call rates over 91%. The GC10 scores, MAF, and call rate of all 372 SNPs are shown in Appendix 2.

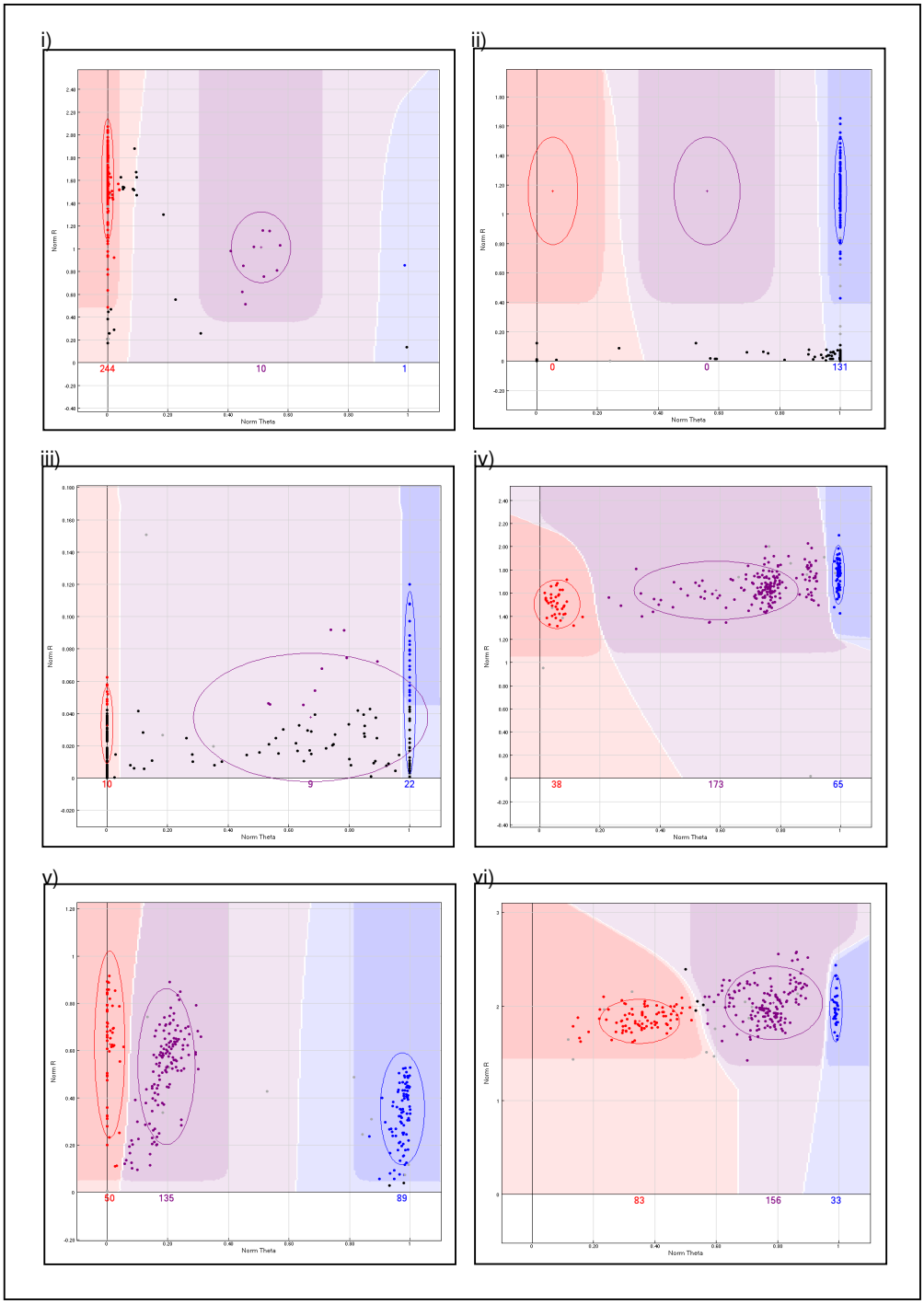
The genotype distribution of the healthy control individuals did not significantly deviate from HWE for any of the *IL18* or *IL18BP* region SNPs. SNPs within each of the other five gene regions did however violate HWE; four in the *CASP1* region, one in the *NALP3* region, three in the *IL1RAP* region, four in the *IL1* ligand cluster, and seven in the *IL1* receptor cluster. These 19 SNPs were tagging a total of 37 SNPs between them (Table 4.4).

Two of the genotyped SNPs were found to be monomorphic in the control and patient cohort used; rs6755497 in the *IL1* ligand cluster, and rs3917257 in the *IL1* receptor cluster region. Frequency data for rs6755497 has been published in dbSNP by three different sources. The International HapMap project and Perlegen, both found it to be non-polymorphic in Caucasians, while the PGA report it at a frequency of 0.053. SNP rs3917257 is also non-polymorphic in the HapMap data, but at a frequency of 0.075 in the PGA Caucasian population. All of these 33 SNPs which did not pass the quality control criteria were discarded and not included in the association analysis.

4.1.4. Stage-2

None of the 18 SNPs investigated in stage-2 of the association study significantly violated HWE. Twelve of the genotyped SNPs had call rates ≥ 98%, and all had call rates ≥ 95%. Of the 289 samples genotyped 92% had call rates ≥ 94%, with 69% of the samples having 100% call rates. The other call rates were: 89% (5 patients, 9 controls), 83% (2 patients, 4 controls), 72% (1 patient, 1 control), and 67% (2 controls).

Of the 324 genotypes (18 SNPs in 18 DNA samples) performed on the two different platforms in both stages only seven showed a discrepancy (97.8% concordance). This is within expected genotyping error rates, showing that there was no significant difference in


i)
ii)
iii)
iv)
v)
vi)
Norm R
Norm Theta
244
10
1
0
0
131
10
9
22
38
173
65
50
135
89
83
156
33

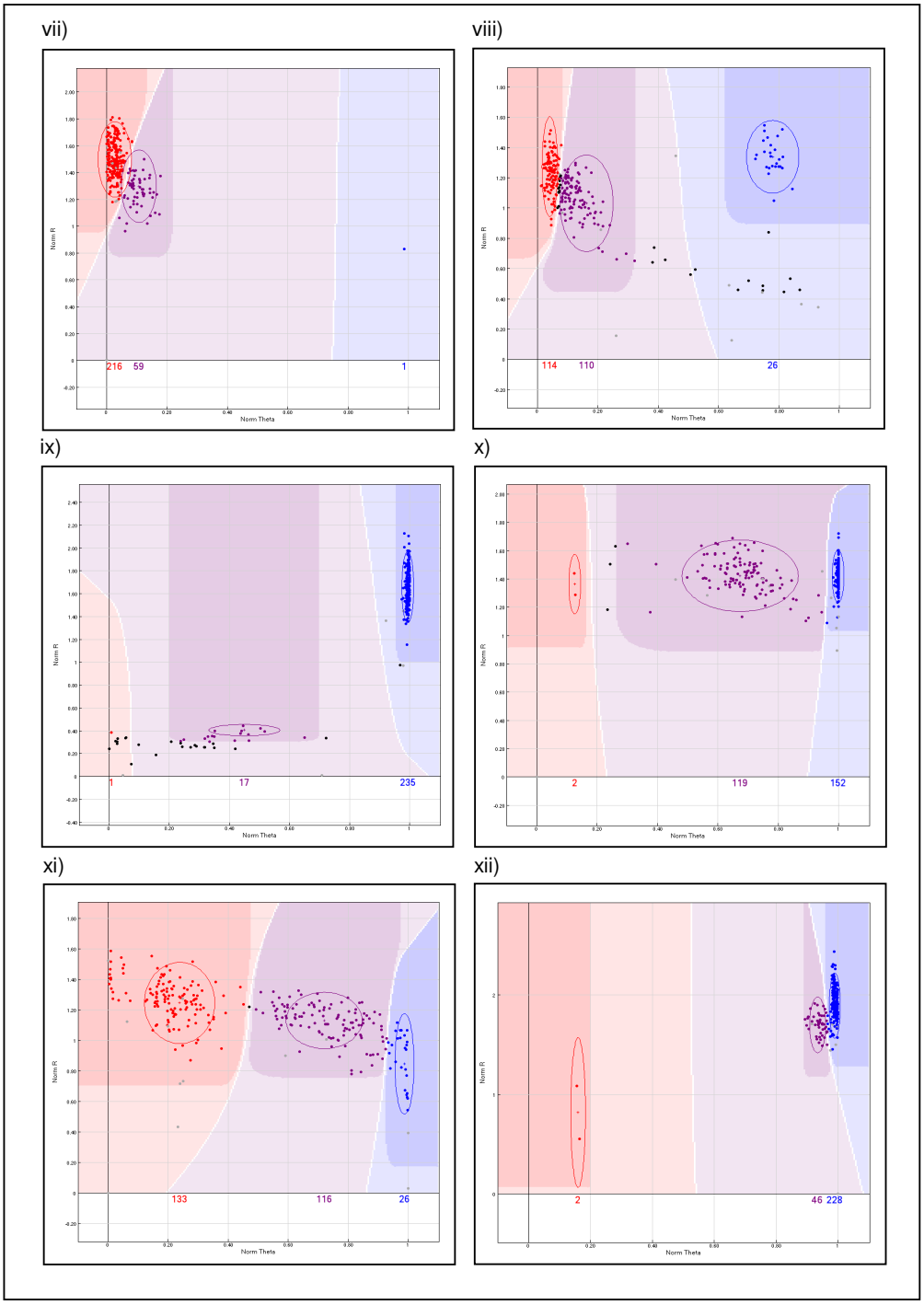


Figure 4.1 Genoplots of failed Stage-1 SNP assays

Shown are the genoplots of the Stage-1 SNP assays classed as failed. Normalised theta values (deviation from total 'A' allele signal) are plotted along the x-axes, and normalised R values (signal intensity) along the y-axes. Each data point represents one sample, the circles indicate the location of each genotype cluster, and the dark

shaded areas are the call areas. All samples falling outside these call areas are excluded from calling, shown in black.

CASP1 region locus: i) rs17103781, *ILRAP* region loci: ii) rs11916734, iii) rs6798479, iv) rs1195370, v) rs4367119, vi) rs6796131, *IL1* ligand cluster loci: vii) rs2708947, viii) rs4849122, iv) 2278716, x) rs2234679, xi) rs3783512, *IL1* receptor cluster locus: xii) rs3218926.

SNP	rs Number	HWE p value
***NALP3* region**		
5	rs7366424	0.0088
***CASP1* region**		
10	rs11604437	0.0387
11	rs7934239	0.0387
19	rs10895775	0.0457
23	rs1503392	0.028
***IL1RAP* region**		
21	rs2885370	0.0314
60	rs12634559	0.0057
77	rs6807826	0.0126
***IL1* ligand cluster**		
60	rs13005572	0.042
62	rs1867829	0.0231
65	rs162319	0.0075
69	rs419598	0.0181
***IL1* receptor cluster**		
25	rs2310175	0.0239
81	rs11123915	0.0095
104	rs2241116	0.0273
105	rs1420095	0.0176
117	rs17027415	0.0025
120	rs1861228	0.0457
123	rs2058612	0.0371

Table 4.3 Genotyped SNPs significantly deviating from Hardy-Weinberg Equilibrium

Nineteen SNPs genotyped in Stage-1 of the association study significantly deviated from HWE in the control cohort. Shown are the details of those SNPs, excluded from the association analysis, including the gene region they are located in, p-value of the chi-squared test for deviation from HWE.

Region	SNP	Individual	Stage-1	Stage-2
***IL1* ligand cluster**	1	GOS528	GT	TT
	4	GOS039	AC	CC
	4	GOS518	AC	CC
	64	GOS528	CC	AA
	80	GOS528	CC	TT
***IL1* receptor cluster**	8	GOS520	CC	CT
***IL1RAP* region**	73	GOS520	AG	AA

Table 4.4 Genotyping discrepancies between the two genotyping platforms

Eighteen patients genotyped on the Illumina Golden Gate platform in stage-1 of the association study were also genotyped by KBioscience on the TaqMan® based platform for the 18 SNPs in stage-2. Shown are the details of the seven discrepancies found between the two genotyping platforms.

genotype calling between the two platforms used. These discrepancies were in one SNP in the *IL1RAP* region, one in the *IL1* receptor cluster region, and five in the *IL1* ligand cluster region (Table 4.5). Apart from *IL1* ligand cluster SNP4, for which there were two discrepancies, there was only one discrepancy with each SNP. Three of the discrepancies were in GOS528, two in GOS520, and the remaining two with different patients. Apart from two, all the genotype discrepancies were of one allele only.

4.1.5. Validation of the tagging SNP selection results

To investigate the accuracy of the LD structure determined during the tSNP selection stage of this study from the HapMap phase 2 data (HapMap2), the LD relationships of a number of the genotyped SNPs were investigated using population genotyping data from phase 3 of the HapMap project (HapMap3) and the WTCCC control cohort (WTCCC2).

Genotyping data was available for a total of 366 pairwise LD calculations (tagging and tagged SNP pair) from at least one of HapMap3 or WTCCC2. Three hundred and seventeen SNP pairs were available from HapMap3, and 194 from WTCCC2. Of these, 147 were available from both databases. The LD relationships of each SNP pair, based on the data from all three sources, are shown in Table 4.6.

4.1.5.1. HapMap3 compared to HapMap2

The genotyping data in the HapMap3 data set replicated 89.9% (285) of the analysed tagging ($r^2 \geq 0.8$) relationships identified during tSNP selection, with 92.8% (295) showing an $r^2 \geq 0.7$ LD relationship. A total of 88.6% (281) of the analysed HapMap3 SNP pairs showed a difference in $r^2 < 0.1$ to the HapMap2 results. Of these, 60.9% (193) showed no difference in r^2 value. The average difference in calculated r^2 values between the HapMap phase 2 and 3 data was 0.0396 ± 0.0939. The largest difference seen (0.493) was between rs2072473 (*IL1* receptor cluster region SNP20) and rs2072474. The LD between these two SNPs was $r^2=1$ according to the HapMap2 data, and $r^2=0.507$ in the HapMap3 data.

4.1.5.2. WTCCC2 compared to HapMap2

The genotyping data in the WTCCC2 data set replicated 79.4% (154) of the analysed tagging ($r^2 \geq 0.8$) relationships identified during tSNP selection, with 96.9% (188) showing an $r^2 \geq 0.7$ LD relationship. Of the SNP pairs available for analysis in the WTCCC2 control cohort, 77.3% (150) showed r^2 differences to HapMap2 of <0.1. Of these, 9.28% (18) showed no

		r^2			Difference		
Tagging SNP	Tagged SNP	HapMap2	HapMap3	WTCCC2	HapMap2 vs WTCCC2	HapMap2 vs HapMap3	HapMap3 vs WTCCC2
***IL18BP* candidate region**							
SNP1 rs3814721	rs949323	1	0.626	0.435	0.374	0.565	0.191
***IL1RAP* candidate region**							
SNP3 rs1024935	rs1008800	1	1	0.965	0	0.035	0.035
SNP14 rs3773999	rs9290935	1	1	0.997	0	0.003	0.003
SNP61 rs9868955	rs4377528	0.934	0.887	n/a	0.047	n/a	n/a
	rs1393061	0.934	0.923	n/a	0.011	n/a	n/a
SNP63 rs6804748	rs1501600	1	1	0.997	0	0.003	0.003
	rs866986	1	1	0.996	0	0.004	0.004
SNP64 rs6775893	rs9868693	1	1	n/a	0	n/a	n/a
	rs9852162	1	0.918	n/a	0.082	n/a	n/a
	rs9811251	1	1	n/a	0	n/a	n/a
	rs952990	1	1	n/a	0	n/a	n/a
SNP65 rs1159213	rs1501603	0.966	0.96	1	0.006	0.034	0.04
	rs1316356	0.966	1	1	0.034	0.034	0
	rs1159211	0.966	1	n/a	0.034	n/a	n/a
	rs967512	0.934	0.927	n/a	0.007	n/a	n/a
	rs1159212	0.934	1	n/a	0.066	n/a	n/a
	rs13100545	1	1	n/a	0	n/a	n/a
	rs9877502	0.966	0.962	n/a	0.004	n/a	n/a
SNP71 rs13314422	rs11721077	1	1	0.988	0	0.012	0.012
SNP73 rs7627123	rs1393049	1	1	n/a	0	n/a	n/a
SNP75 rs9814312	rs1510902	1	1	n/a	0	n/a	n/a
***IL1* Ligand candidate region**							
SNP3 rs17561	rs1894399	1	1	0.997	0	0.003	0.003
	rs3783559	1	n/a	0.999	n/a	0.001	n/a
	rs1878321	1	1	0.998	0	0.002	0.002
	rs2856837	1	0.973	0.998	0.027	0.002	0.025
	rs1878319	1	1	0.998	0	0.002	0.002
	rs1304037	1	0.973	0.999	0.027	0.001	0.026
	rs1878320	1	1	n/a	0	n/a	n/a
	rs3783557	1	0.947	0.963	0.053	0.037	0.016
	rs2071375	1	n/a	0.999	n/a	0.001	n/a
	rs1800794	1	1	0.99	0	0.01	0.01
	rs4848300	1	0.973	0.999	0.027	0.001	0.026
	rs2856841	1	n/a	0.999	n/a	0.001	n/a
	rs1516790	1	n/a	0.999	n/a	0.001	n/a
	rs4848303	1	1	n/a	0	n/a	n/a
	rs1040193	1	n/a	0.999	n/a	0.001	n/a
	rs1800587	1	0.973	0.998	0.027	0.002	0.025
	rs3783555	1	n/a	n/a	n/a	n/a	n/a
	rs2856836	1	0.973	0.999	0.027	0.001	0.026
	rs3783520	1	1	0.998	0	0.002	0.002

Tagging SNP	Tagged SNP	HapMap2	HapMap3	WTCCC2	HapMap2 vs WTCCC2	HapMap2 vs HapMap3	HapMap3 vs WTCCC2
SNP5 rs3783516	rs2856838	1	1	1	0	0	0
	rs3783515	1	1	1	0	0	0
	rs3783547	1	n/a	0.995	n/a	0.005	n/a
	rs6746923	1	1	1	0	0	0
SNP10 rs3917368	rs1143643	1	1	n/a	0	n/a	n/a
	rs1143633	0.863	1	0.987	0.137	0.124	0.013
SNP26 rs2723192	rs2723179	1	n/a	0.996	n/a	0.004	n/a
	rs2708940	1	n/a	0.996	n/a	0.004	n/a
	rs2708957	1	1	0.994	0	0.006	0.006
	rs2723156	1	1	n/a	0	n/a	n/a
	rs2723180	1	n/a	0.996	n/a	0.004	n/a
	rs2723177	1	1	0.996	0	0.004	0.004
	rs2708924	1	0.895	0.945	0.105	0.055	0.05
	rs2723169	1	1	0.996	0	0.004	0.004
	rs2708959	1	1	0.991	0	0.009	0.009
	rs2708925	1	1	0.981	0	0.019	0.019
	rs2708943	1	1	0.993	0	0.007	0.007
	rs2708954	1	1	n/a	0	n/a	n/a
	rs2708956	1	n/a	0.991	n/a	0.009	n/a
	rs2708960	1	1	0.996	0	0.004	0.004
	rs2708965	1	1	0.99	0	0.01	0.01
	rs2723196	1	1	0.977	0	0.023	0.023
	rs2708962	1	1	0.996	0	0.004	0.004
	rs7577574	1	1	n/a	0	n/a	n/a
	rs2708958	1	1	0.996	0	0.004	0.004
	rs2723189	1	1	0.997	0	0.003	0.003
	rs2708953	1	1	0.997	0	0.003	0.003
	rs2723170	1	n/a	0.993	n/a	0.007	n/a
SNP64 rs1688075	rs315925	1	1	n/a	0	n/a	n/a
SNP69 rs419598	rs448341	1	n/a	0.996	n/a	0.004	n/a
	rs432014	1	1	0.996	0	0.004	0.004
	rs408392	1	1	0.999	0	0.001	0.001
	rs442710	1	1	1	0	0	0
	rs1794068	1	1	0.999	0	0.001	0.001
	rs447713	1	1	0.999	0	0.001	0.001
	rs128964	1	1	0.999	0	0.001	0.001
	rs495410	1	1	1	0	0	0
	rs431726	1	0.951	0.927	0.049	0.073	0.024
	rs1794065	1	n/a	0.99	n/a	0.01	n/a
	rs444413	1	1	1	0	0	0
	rs315931	0.851	0.829	0.772	0.022	0.079	0.057
	rs495282	1	1	0.998	0	0.002	0.002
	rs2853628	1	0.936	0.989	0.064	0.011	0.053
	rs423904	1	1	0.999	0	0.001	0.001
	rs315936	1	1	0.93	0	0.07	0.07
	rs392503	1	n/a	0.993	n/a	0.007	n/a
	rs1794067	1	1	0.998	0	0.002	0.002
	rs451578	0.833	0.975	0.919	0.142	0.086	0.056
	rs446433	1	1	0.999	0.001	0	0.001

Tagging SNP	Tagged SNP	HapMap2	HapMap3	WTCCC2	HapMap2 vs WTCCC2	HapMap2 vs HapMap3	HapMap3 vs WTCCC2
	rs454078	1	1	0.99	0	0.01	0.01
	rs3087263	1	n/a	0.276	n/a	0.724	n/a
SNP83 rs34862832	rs3768769	1	n/a	0.669	n/a	0.331	n/a
	rs28938790	0.876	n/a	0.779	n/a	0.097	n/a
	rs34345011	1	n/a	0.851	n/a	0.149	n/a
	rs6714534	1	n/a	0.801	n/a	0.199	n/a
	rs10496447	1	n/a	0.984	n/a	0.016	n/a
	rs28954071	0.864	n/a	0.765	n/a	0.099	n/a
	rs17660913	1	n/a	0.983	n/a	0.017	n/a
	rs13407838	1	n/a	0.771	n/a	0.229	n/a
	rs28993969	1	n/a	0.802	n/a	0.198	n/a
	rs28947170	1	n/a	0.646	n/a	0.354	n/a
IL1 Receptor candidate region							
SNP1 rs2871432	rs2310144	1	0.933	0.969	0.067	0.031	0.036
	rs2310143	1	0.96	n/a	0.04	n/a	n/a
	rs6720836	1	0.979	0.997	0.021	0.003	0.018
SNP2 rs9308849	rs2310141	1	0.955	n/a	0.045	n/a	n/a
	rs4851509	1	0.955	n/a	0.045	n/a	n/a
	rs1003376	0.962	0.908	n/a	0.054	n/a	n/a
	rs11688863	1	0.954	n/a	0.046	n/a	n/a
	rs10489971	1	0.955	n/a	0.045	n/a	n/a
SNP4 rs2190364	rs7568226	0.954	1	n/a	0.046	n/a	n/a
	rs2214890	0.954	0.973	0.986	0.019	0.032	0.013
SNP7 rs4851516	rs6726051	0.954	0.978	0.995	0.024	0.041	0.017
SNP8 rs12712122	rs1010329	0.855	0.875	n/a	0.002	n/a	n/a
	rs2190360	1	0.963	n/a	0.037	n/a	n/a
	rs740044	0.820	0.868	n/a	0.048	n/a	n/a
	rs2190361	1	0.963	n/a	0.037	n/a	n/a
SNP14 rs2110562	rs2110563	1	1	n/a	0	n/a	n/a
	rs719248	1	1	0.997	0	0.003	0.003
	rs4851527	1	1	0.982	0	0.018	0.018
	rs757918	1	1	0.978	0	0.022	0.022
SNP20 rs2072473	rs3218876	1	0.683	1	0.317	0	0.317
	rs2236926	1	n/a	0.71	n/a	0.29	n/a
	rs2236925	1	0.606	n/a	0.394	n/a	n/a
	rs2072477	1	0.702	0.711	0.298	0.289	0.009
	rs2073492	1	0.683	0.985	0.317	0.015	0.302
	rs733498	1	1	0.995	0	0.005	0.005
	rs3218879	1	0.683	1	0.317	0	0.317
	rs3218872	1	0.702	0.999	0.298	0.001	0.297
	rs3218874	1	0.683	1	0.317	0	0.317
	rs3218877	1	0.683	1	0.317	0	0.317
	rs3218953	1	n/a	0.743	n/a	0.257	n/a
	rs2072480	1	0.683	0.711	0.317	0.289	0.028
	rs3218923	1	n/a	0.706	n/a	0.294	n/a
	rs2262267	1	n/a	0.999	n/a	0.001	n/a
	rs3218873	1	0.702	0.707	0.298	0.293	0.005
	rs3218878	1	0.683	0.706	0.317	0.294	0.023
	rs2072479	1	1	0.711	0	0.289	0.289

Tagging SNP	Tagged SNP	HapMap2	HapMap3	WTCCC2	HapMap2 vs WTCCC2	HapMap2 vs HapMap3	HapMap3 vs WTCCC2
	rs2072474	1	0.507	0.795	0.493	0.205	0.288
	rs3218896	1	0.618	0.48	0.382	0.52	0.138
	rs3218885	1	0.683	0.705	0.317	0.295	0.022
	rs3218928	1	n/a	0.71	n/a	0.29	n/a
	rs3218870	1	0.683	n/a	0.317	n/a	n/a
	rs2072478	1	0.702	n/a	0.298	n/a	n/a
	rs2236928	1	n/a	0.697	n/a	0.303	n/a
	rs2073491	1	0.683	0.706	0.317	0.294	0.023
	rs2236930	1	0.683	0.705	0.317	0.295	0.022
	rs2236923	1	n/a	0.711	n/a	0.289	n/a
	rs3218903	1	n/a	0.702	n/a	0.298	n/a
	rs2236921	1	n/a	0.729	n/a	0.271	n/a
	rs2236924	1	n/a	0.711	n/a	0.289	n/a
	rs4850995	0.919	0.957	n/a	0.038	n/a	n/a
	rs3218911	1	0.795	0.706	0.205	0.294	0.089
	rs2310171	1	n/a	0.703	n/a	0.297	n/a
	rs3218892	1	0.683	0.704	0.317	0.296	0.021
	rs3218883	1	1	0.706	0	0.294	0.294
	rs2072481	1	0.724	0.711	0.276	0.289	0.013
	rs2282743	1	0.683	0.706	0.317	0.294	0.023
	rs2282744	1	0.683	0.706	0.317	0.294	0.023
	rs2298939	1	n/a	0.703	n/a	0.297	n/a
	rs2236927	1	0.691	0.714	0.309	0.286	0.023
	rs2072482	1	n/a	0.711	n/a	0.289	n/a
	rs3218909	1	n/a	0.704	n/a	0.296	n/a
	rs3218875	1	n/a	0.704	n/a	0.296	n/a
SNP66 rs2160226	rs1922289	1	1	0.999	0	0.001	0.001
	rs2302624	1	1	0.999	0	0.001	0.001
	rs6715919	1	1	1	0	0	0
	rs1558646	1	1	0.999	0	0.001	0.001
	rs995514	1	n/a	1	n/a	0	n/a
	rs7569116	1	n/a	0.999	n/a	0.001	n/a
	rs2302623	1	1	0.995	0	0.005	0.005
	rs2041752	1	1	1	0	0	0
	rs1922293	1	1	1	0	0	0
	rs3771197	1	1	0.999	0	0.001	0.001
	rs17020717	1	1	0.999	0	0.001	0.001
	rs1882514	1	1	0.998	0	0.002	0.002
	rs3771196	1	1	0.995	0	0.005	0.005
	rs3771194	1	1	0.999	0	0.001	0.001
	rs2141865	1	1	0.998	0	0.002	0.002
	rs4851002	1	1	0.999	0	0.001	0.001
	rs3755289	1	1	0.995	0	0.005	0.005
	rs7566395	1	1	0.999	0	0.001	0.001
	rs2178675	1	1	0.999	0	0.001	0.001
	rs867770	1	1	n/a	0	n/a	n/a
	rs2310236	1	1	0.999	0	0.001	0.001
	rs3771193	1	n/a	0.999	n/a	0.001	n/a
	rs3729564	1	1	1	0	0	0
	rs7605715	1	1	0.999	0	0.001	0.001

Tagging SNP	Tagged SNP	HapMap2	HapMap3	WTCCC2	HapMap2 vs WTCCC2	HapMap2 vs HapMap3	HapMap3 vs WTCCC2
	rs11688596	1	n/a	0.997	n/a	0.003	n/a
	rs870684	1	1	0.999	0	0.001	0.001
	rs2871458	1	1	0.999	0	0.001	0.001
	rs1922290	1	1	0.998	0	0.002	0.002
	rs1922295	1	1	0.999	0	0.001	0.001
	rs35917382	1	n/a	0.998	n/a	0.002	n/a
	rs3815517	1	n/a	0.997	n/a	0.003	n/a
	rs1922296	1	1	0.999	0	0.001	0.001
	rs7421641	1	1	n/a	0	n/a	n/a
	rs995515	1	1	1	0	0	0
	rs2141864	1	1	0.994	0	0.006	0.006
	rs34074031	1	n/a	0.998	n/a	0.002	n/a
	rs2310238	1	1	0.978	0	0.022	0.022
	rs6733727	1	1	1	0	0	0
	rs11686153	1	1	0.998	0	0.002	0.002
	rs3771192	1	1	0.997	0	0.003	0.003
	rs4851001	1	1	0.998	0	0.002	0.002
	rs1169062	1	1	n/a	0	n/a	n/a
	rs2041753	1	1	1	0	0	0
	rs3755290	1	1	0.999	0	0.001	0.001
	rs3074969	1	n/a	n/a	n/a	n/a	n/a
	rs2310237	1	1	0.998	0	0.002	0.002
	rs3821207	1	1	0.999	0	0.001	0.001
	rs3771195	1	1	0.999	0	0.001	0.001
SNP86 rs6706844	rs4090473	1	1	n/a	0	n/a	n/a
	rs974389	1	1	n/a	0	n/a	n/a
	rs953934	0.966	1	n/a	0.034	n/a	n/a
	rs1558622	1	1	n/a	0	n/a	n/a
	rs4399750	1	1	n/a	0	n/a	n/a
	rs1997466	1	1	n/a	0	n/a	n/a
	rs12996772	1	1	n/a	0	n/a	n/a
	rs11685424	0.932	1	n/a	0.068	n/a	n/a
SNP112 rs1880000	rs7601773	1	1	n/a	0	n/a	n/a
	rs7575867	1	1	n/a	0	n/a	n/a
	rs1523196	1	1	n/a	0	n/a	n/a
	rs7593444	1	1	n/a	0	n/a	n/a
	rs1403549	1	1	n/a	0	n/a	n/a
	rs17027341	1	1	n/a	0	n/a	n/a
	rs6749440	1	1	n/a	0	n/a	n/a
	rs1159509	1	1	n/a	0	n/a	n/a
	rs1523197	1	1	n/a	0	n/a	n/a
IL18 candidate region							
SNP1 rs10891319	rs7944155	0.871	0.845	n/a	0.026	n/a	n/a
	rs574568	0.88	0.904	0.903	0.024	0.023	0.001
SNP7 rs549908	rs243908	0.957	1	0.997	0.043	0.04	0.003
	rs360729	0.967	0.977	0.993	0.01	0.026	0.016
	rs360720	1	1	n/a	0	n/a	n/a
	rs795467	0.867	0.824	0.839	0.043	0.028	0.015
	rs360717	0.833	0.82	0.817	0.013	0.016	0.003
	rs360718	1	0.804	0.816	0.196	0.184	0.012

Tagging SNP	Tagged SNP	HapMap2	HapMap3	WTCCC2	HapMap2 vs WTCCC2	HapMap2 vs HapMap3	HapMap3 vs WTCCC2
	rs1293344	0.837	0.824	0.818	0.013	0.019	0.006
	rs11214105	0.813	0.824	0.818	0.011	0.005	0.006
SNP8 rs1834481	rs5744276	0.957	1	n/a	0.043	n/a	n/a
	rs5744258	1	1	n/a	0	n/a	n/a
	rs5744256	1	0.97	0.998	0.03	0.002	0.028
	rs5744222	1	0.97	0.974	0.026	0.03	0.004
SNP9 rs5744247	rs543810	0.831	0.78	n/a	n/a	0.051	n/a
	rs544354	0.899	0.79	n/a	0.109	n/a	n/a
SNP12 rs2043055	rs5744280	0.867	0.845	0.856	0.022	0.011	0.011
	rs7106524	0.924	1	0.994	0.076	0.07	0.006
SNP13 rs1946519	rs1946518	0.97	0.957	0.999	0.013	0.029	0.042
***CASP1* candidate region**							
SNP2 rs492859	rs572980	1	0.626	n/a	0.374	n/a	n/a
	rs547351	1	0.603	n/a	0.397	n/a	n/a
SNP4 rs1699089	rs476889	1	1	n/a	0	n/a	n/a
	rs483345	1	1	n/a	0	n/a	n/a
	rs492366	1	1	n/a	0	n/a	n/a
	rs568910	1	1	n/a	0	n/a	n/a
	rs572687	1	1	n/a	0	n/a	n/a
	rs692897	1	1	n/a	0	n/a	n/a
	rs504054	1	1	n/a	0	n/a	n/a
	rs519047	1	1	n/a	0	n/a	n/a
	rs7928549	1	1	n/a	0	n/a	n/a
	rs491191	1	1	n/a	0	n/a	n/a
	rs505901	1	1	n/a	0	n/a	n/a
	rs517736	1	1	n/a	0	n/a	n/a
	rs1503391	1	1	n/a	0	n/a	n/a
	rs507438	1	1	n/a	0	n/a	n/a
	rs507509	1	1	n/a	0	n/a	n/a
	rs546449	1	1	n/a	0	n/a	n/a
	rs692914	1	1	n/a	0	n/a	n/a
	rs572031	1	1	n/a	0	n/a	n/a
	rs516286	1	1	n/a	0	n/a	n/a
	rs542571	1	1	n/a	0	n/a	n/a
	rs490565	1	1	n/a	0	n/a	n/a
	rs500577	0.939	0.935	n/a	0.004	n/a	n/a
	rs557905	1	0.98	n/a	0.02	n/a	n/a
	rs481736	1	1	n/a	0	n/a	n/a
	rs580253	1	1	n/a	0	n/a	n/a
	rs534811	1	1	n/a	0	n/a	n/a
	rs17103597	1	1	n/a	0	n/a	n/a
	rs489715	1	1	n/a	0	n/a	n/a
SNP7 rs571593	rs526167	1	1	n/a	0	n/a	n/a
	rs17376473	1	1	n/a	0	n/a	n/a
	rs1792774	1	1	n/a	0	n/a	n/a
	rs484626	1	1	n/a	0	n/a	n/a
	rs501626	1	1	n/a	0	n/a	n/a
	rs556205	1	1	n/a	0	n/a	n/a
	rs488992	1	1	n/a	0	n/a	n/a

Tagging SNP	Tagged SNP	HapMap2	HapMap3	WTCCC2	HapMap2 vs WTCCC2	HapMap2 vs HapMap3	HapMap3 vs WTCCC2
SNP9 rs10502045	rs1977989	1	1	n/a	0	n/a	n/a
	rs576800	1	0.901	0.979	0.099	0.021	0.078
	rs10895763	1	0.974	0.994	0.026	0.006	0.02
	rs9326349	0.955	0.975	0.998	0.02	0.043	0.023
	rs11821722	1	0.974	n/a	0.026	n/a	n/a
	rs2282659	1	0.975	0.996	0.025	0.004	0.021
SNP15 rs538943	rs486123	1	1	n/a	0	n/a	n/a
SNP16 rs1623342	rs1785878	0.918	0.949	n/a	0.031	n/a	n/a
	rs529809	0.959	0.949	n/a	0.01	n/a	n/a
	rs1628434	0.958	0.949	n/a	0.009	n/a	n/a
	rs1785882	1	1	n/a	0	n/a	n/a
	rs6591110	0.959	0.95	n/a	0.009	n/a	n/a
	rs1699088	0.959	0.951	n/a	0.008	n/a	n/a
	rs4106096	0.959	0.95	n/a	0.009	n/a	n/a
	rs572717	0.957	0.948	n/a	0.009	n/a	n/a
	rs562441	1	1	n/a	0	n/a	n/a
	rs7934144	0.96	0.949	n/a	0.011	n/a	n/a
	rs1792766	0.96	0.952	n/a	0.008	n/a	n/a
	rs4638280	0.959	0.951	n/a	0.008	n/a	n/a
	rs484345	0.96	0.952	n/a	0.008	n/a	n/a
	rs1792763	0.96	0.949	n/a	0.011	n/a	n/a
SNP18 rs7109571	rs1792752	0.951	0.945	n/a	0.006	n/a	n/a
	rs1699080	0.959	0.953	n/a	0.006	n/a	n/a
	rs4587693	0.853	0.819	0.881	0.034	0.028	0.062
SNP19 rs10895775	rs7932543	0.927	0.919	n/a	0.008	n/a	n/a
	rs1483025	1	1	n/a	0	n/a	n/a
	rs1699081	0.931	0.921	n/a	0.01	n/a	n/a
	rs10895771	0.932	0.923	n/a	0.009	n/a	n/a
	rs1785872	0.933	0.925	n/a	0.008	n/a	n/a
	rs1785871	0.902	0.925	n/a	0.023	n/a	n/a
	rs1792753	0.902	0.921	n/a	0.019	n/a	n/a
	rs1785864	0.902	0.921	n/a	0.019	n/a	n/a
	rs1847293	0.933	0.921	n/a	0.012	n/a	n/a
	rs1699087	0.901	0.924	n/a	0.023	n/a	n/a
	rs3736149	0.902	0.921	n/a	0.019	n/a	n/a
	rs1699086	0.902	0.923	n/a	0.021	n/a	n/a
	rs1503399	0.933	0.925	n/a	0.008	n/a	n/a
	rs1785873	0.961	0.958	n/a	0.003	n/a	n/a
SNP22 rs1503394	rs3926237	1	1	n/a	0	n/a	n/a
SNP23 rs1503392	rs6591111	1	1	n/a	0	n/a	n/a
	rs1847298	1	1	n/a	0	n/a	n/a
	rs11226619	1	1	n/a	0	n/a	n/a
	rs7934633	1	1	n/a	0	n/a	n/a
	rs7952721	1	1	n/a	0	n/a	n/a
SNP24 rs17103781	rs3989416	1	1	n/a	0	n/a	n/a
	rs1503386	1	1	n/a	0	n/a	n/a
	rs12417470	1	1	n/a	0	n/a	n/a
	rs12417050	1	1	n/a	0	n/a	n/a

Tagging SNP	Tagged SNP	HapMap2	HapMap3	WTCCC2	HapMap2 vs WTCCC2	HapMap2 vs HapMap3	HapMap3 vs WTCCC2
	rs17103670	1	1	n/a	0	n/a	n/a
	rs12419452	1	1	n/a	0	n/a	n/a
	rs12420601	1	1	n/a	0	n/a	n/a
	rs12800151	1	1	0.894	0	0.106	0.106
	rs4755079	1	1	n/a	0	n/a	n/a
	rs12281540	1	1	n/a	0	n/a	n/a
	rs12420557	1	1	n/a	0	n/a	n/a
	rs4754124	1	1	n/a	0	n/a	n/a
	rs10895772	1	1	n/a	0	n/a	n/a
	rs12418649	1	1	n/a	0	n/a	n/a
	rs1393907	1	1	n/a	0	n/a	n/a
	rs3898011	1	1	0.895	0	0.105	0.105
	rs1503395	1	1	n/a	0	n/a	n/a
	rs475266	1	0.89	n/a	0.11	n/a	n/a
	rs12416756	1	1	n/a	0	n/a	n/a
	rs1910398	1	1	n/a	0	n/a	n/a
	rs4755078	1	1	n/a	0	n/a	n/a
	rs12801987	1	1	n/a	0	n/a	n/a
	rs17103773	1	1	n/a	0	n/a	n/a
	rs17376738	1	1	0.964	0	0.036	0.036
	rs12807436	1	1	n/a	0	n/a	n/a
	rs7930712	1	1	n/a	0	n/a	n/a
	rs7117758	1	1	0.969	0	0.031	0.031
	rs12420524	1	1	0.896	0	0.104	0.104
	rs2062807	1	1	n/a	0	n/a	n/a
	rs12418438	1	1	n/a	0	n/a	n/a
***NALP3* candidate region**							
SNP2 rs7556185	rs10737800	0.881	1	n/a	0.119	n/a	n/a
SNP8 rs12082329	rs10399672	0.805	0.663	n/a	0.142	n/a	n/a
SNP11 rs2105290	rs12130139	0.902	0.885	n/a	0.017	n/a	n/a
	rs4317844	0.809	0.789	n/a	0.02	n/a	n/a
	rs10802495	0.841	0.824	n/a	0.017	n/a	n/a
SNP12 rs10925006	rs2386547	1	1	n/a	0	n/a	n/a
SNP16 rs3738448	rs7525979	0.867	0.895	0.927	0.028	0.032	0.032

Table 4.5 Comparison of LD (r2) values from different population genotyping data sets

Population genotyping data for a subset of the SNPs included in the initial SNP selection stage of this project was extracted from the HapMap phase 3 and WTCCC phase 2 databases. Pairwise LD between SNPs, identified as being in high LD ($r^2 \geq 0.8$) in the SNP selection stage, was calculated from each data set. Only SNP pairs for which genotyping data was available for both SNPs in at least one data set is shown. The calculated LD (r^2) for each SNP pair in each of the three population genotyping data sets is shown. The difference between the three data sets of the r^2 values calculated for each SNP pair are also shown. Differences in r^2 of >0.1 between the HapMap2 data and either HapMap3 or WTCCC2, are highlighted.

difference. The average difference between the WTCCC2 and HapMap2 results was 0.0741 ± 0.128. The largest difference seen (0.724) was between rs419598 (*IL1* ligand cluster region SNP69) and rs3087263. The LD between these two SNPs was r^2=1 according to the HapMap2 data, and r^2=0.276 in the WTCCC2 data.

4.1.5.3. HapMap3 compared to WTCCC2

When compared to each other, the HapMap3 and WTCCC2 data sets showed high concordance, with differences of <0.1 in 90.48% (133) of the SNP pairs. Of these, 9.5% (14) showed exactly the same r^2 values. The average difference between the HapMap3 and WTCCC2 LD results was 0.0345 ± 0.0744.

Of the 366 SNP pairs investigated for comparison, 59 showed significant differences in r^2 (>0.1) between HapMap2 and at least one of the other data sets. However, the significant discrepancies in LD observed were not distributed evenly over the genotyped SNPs analysed. Data was available to analyse LD relationships with 49 of the genotyped SNPs. All of the r^2 differences >0.1 were found in SNP pairs involving only 11 of the genotyped SNPs. 75.8% (47) of all SNP pairs showing significant differences in LD, involved only 2 of the genotyped SNPs; SNP8 in the *IL1* ligand cluster region (rs34862832) and SNP20 in the *IL1* receptor cluster (rs2072473). Of the 43 SNPs paired with *IL1* receptor cluster SNP 20, only three did not show differences >0.1 from those based on the HapMap2 data. However, the SNP pairs including this genotyped SNP were also non-concordant between the HapMap3 and WTCCC2 data sets. Of the 24 SNP pairs available from both, only 13 had differences <0.1 between the two. For example, four of the SNP pairs were in r^2=1 in HapMap 2, r^2= 0.683 in HapMap3, and r^2=1 in WTCCC2. As the large discrepancies in LD were seen with a limited number of the SNPs investigated it is likely that they were due to SNP specific differences rather than reflecting a genome-wide difference in LD structure. However, there did not appear to be differences between the three data sets in the frequencies of these two genotyped SNPs, or the SNPs they are tagging (Table 4.6). The frequencies, in HapMap2, HapMap3, and WTCCC2 respectively, of the genotyped SNP *IL1* receptor cluster SNP20 (rs2072473) were: 0.14, 0.13, and 0.13, and 0.125, 0.125, and 0.14, for *IL1* ligand cluster region SNP83 (rs34862832). A frequency difference of 0.07 was observed for rs3218953 in the *IL1* receptor cluster region with MAF=0.083 in HapMap2 and 0.153 in WTCCC2, this SNP was not included in HapMap3. The discrepancy (0.257) in LD seen with this SNP may be attributable

SNP	HapMap2	HapMap3	WTCCC 2
rs34862832	0.125	n/a	0.144
rs3768769	0.125	n/a	0.142
rs28938790	0.140	n/a	0.157
rs34345011	0.111	n/a	0.126
rs6714534	0.140	n/a	0.161
rs10496447	0.130	n/a	0.144
rs28954071	0.083	n/a	0.158
rs17660913	0.130	n/a	0.144
rs13407838	0.139	n/a	0.158
rs28993969	0.205	n/a	0.160
rs28947170	0.125	n/a	0.140
rs2072473	0.139	0.129	0.130
rs3218876	0.107	0.108	0.10
rs2236926	0.107	n/a	0.10
rs2236925	0.107	0.082	n/a
rs2072477	0.107	0.097	0.100
rs2073492	0.107	0.108	0.101
rs733498	0.135	0.128	0.130
rs3218879	0.107	0.108	0.101
rs3218872	0.107	0.097	0.100
rs3218874	0.115	0.108	0.100
rs3218877	0.107	0.108	0.100
rs3218953	0.083	n/a	0.153
rs2072480	0.125	0.102	0.100
rs3218923	0.107	n/a	0.100
rs2262267	0.115	n/a	0.101
rs3218873	0.107	0.097	0.101
rs3218878	0.107	0.108	0.100
rs2072479	0.115	0.097	0.100
rs2072474	0.107	0.159	0.144
rs3218896	0.107	0.095	0.070
rs3218885	0.125	0.108	0.100
rs3218928	0.107	n/a	0.100
rs3218870	0.107	0.108	n/a
rs2072478	0.107	0.098	0.100
rs2236928	0.107	n/a	0.100
rs2073491	0.115	0.108	0.101
rs2236930	0.107	0.108	0.101
rs2236923	0.107	n/a	0.100
rs3218903	0.107	n/a	0.101
rs2236921	0.115	n/a	0.100
rs2236924	0.107	n/a	0.100
rs4850995	0.125	0.134	n/a
rs3218911	0.107	0.125	0.101
rs2310171	0.107	n/a	0.101
rs3218892	0.107	0.095	0.100
rs3218883	0.107	0.108	0.101
rs2072481	0.111	0.109	0.100
rs2282743	0.107	0.108	0.101
rs2282744	0.107	0.108	0.101
rs2298939	0.119	n/a	0.100
rs2236927	0.107	0.094	0.099
rs2072482	0.111	n/a	0.100
rs3218909	0.107	n/a	0.101
rs3218875	0.107	n/a	0.101

Table 4.6 Comparison of SNP frequencies in different population genotyping data sets

The frequency (MAF) of the genotyped SNPs *IL1* receptor cluster SNP20 (rs2072473), *IL1* ligand cluster region SNP83 (rs34862832), and all of the SNPs identified as being tagged by them, are shown for the HapMap2, HapMap3, and WTCCC2 data sets. All frequencies showing a difference >0.02 from the HapMap2 frequency for that SNP are highlighted.

to the frequency differences. However, over all of the SNPs there did not appear to be a general correlation between frequency and LD differences. A very similar difference of frequencies (rs28954071, 0.083 in HapMap2, 0.158 in WTCCC2) was associated with a difference in r^2 of only 0.099. There was a 0.382 and 0.52 difference in r^2 between rs2072473 (*IL1* receptor cluster SNP 20) and rs3218896, of HapMap3 and WTCCC2, respectively, compared to HapMap 2. However both of the SNPs involved had minimal frequency differences between the data sets; 0.107, 0.095, 0.070 (rs3218896), and 0.139, 0.129, 0.130 (rs2072473) in HapMap2, HapMap3, and WTCCC2 respectively.

4.1.5.4. Discussion

When the tSNP selection and genotyping stages of this study were performed HapMap2 was the only population genotyping database publicly available. However, the larger HapMap3 and WTCCC2 data sets are now available, providing more information on the genetic structure of the European population. Genotype information from both of these larger data sets showed extensive (89.9% and 79.4%) concordance with the tagging relationships ($r^2 \geq 0.8$) determined based on HapMap2 and used to inform tSNP selection in the association study. Although, due to differences in the SNPs included in the three data sets, it was not possible to compare all of the SNP pairs selected for comparison, and these were only a subset of those investigated, it is reasonable to assume that the high concordance observed here is representative of the full SNP set. It can therefore be concluded that, although performed with a more limited sample size, the tSNP selection was accurate and that the majority of the genetic variation within each candidate region was captured. The results of these comparisons have, however, also highlighted that a small number of the SNP pairs show large differences in r^2 between the data sets. It is therefore important that, where possible, the LD results from each data set is taken into account when determining which un-typed SNPs may potentially be responsible for any evidence of disease association identified.

The results of this comparison suggest that of the two population genotyping datasets the data from HapMap3, rather than WTCCC2, is more similar to HapMap2. This is as expected, given that the individuals used in both HapMap phases are from the CEU population. However, as there are differences observed between the two HapMap data releases it is evident that the additional information provided by the larger cohort size of HapMap3 contributes further data on the LD relationships. The LD patterns observed in the HapMap3 data set are, however, more similar to WTCCC2, with a high proportion (90.48%) of the

results being comparable (r^2 difference of >0.1) between the cohorts. This similarity suggests that the HapMap phase 3 population is analogous to the UK population, used by the WTCCC.

A potential cause for the differences in r^2 values seen between the data sets could be allele frequency differences between the populations studied, either due to differences in the populations, or to differences in the genotyping platforms used; Illumina and Affymetrix genotyping chips. However, this does not appear to be the case as none of the SNPs included in the comparison show frequency differences greater than would be explained by sampling differences.

The differences in LD seen between the HapMap and WTCCC data sets could be due to underlying differences in the LD structure of the populations sampled. The CEU population used in both HapMap phases are Utah residents of Northern and Western European ancestry, while the WTCCC cohort was sampled from only the UK. However, all individuals included in the WTCCC cohort are UK residents of self-reported white ethnicity, meaning that, although it is likely that the majority are, they may actually be of a more heterogeneous European, rather than purely UK ancestry. However, as the HapMap phase3 and WTCCC2 data sets show high levels of similarity it is unlikely that there is a marked difference in the underlying genetic structure of the two populations resulting in the differences seen.

Another potential source of the differences seen between the data sets is that the WTCCC population are all unrelated individuals whereas the CEU population used by HapMap consists mainly of family pedigrees. Genotyping data from members of the same family enables more precise phase prediction and therefore increases the accuracy of LD determination. However, it can also result in the overrepresentation of some haplotypes and inflated LD values. It is likely that the HapMap2 data set is less accurate than the other two as it contains information on only 90 individuals, compared to 180 in Hapmap3 and 4,673 in WTCCC2. The information in Hapmap2 is further limited by the fact that all the individuals were part of family trios, so that only 120 independent chromosomes (60 unrelated parents) were investigated. All haplotypes present in a wider population may not be represented in this limited genetic pool, leading to over-inflation of the LD.

The appropriateness of using the HapMap CEU as a reference population in tSNP selection is discussed in chapter 5.

4.2. *IL18BP* candidate region

4.2.1. Association Study

The genes flanking the candidate gene *IL18BP* were identified using the July 2003 human genome reference sequence (NCBI build 34), hg16 annotation track. The boundaries of the candidate region, chromosome 11:71434949-71443655, were defined based on the positions of the flanking genes (centre panel Figure 4.2). A total of four protein coding *IL18BP* isoforms have been described (Kim et al., 2000), two of which are shown in the figure. Flanking *IL18BP* 5′ is RING finger protein 121 (*RNF121*), and 3′, nuclear protein that associates with the mitotic apparatus (*NUMA1*), which also overlaps with the *IL18BP* sequence by 2.9kb. Although RNF121 contains a RING finger motif, present in a number of proteins known to be involved in protein-protein and protein-DNA interactions, no function for the protein has been shown. *NUMA1* is a component of the mitotic centrosome (Zeng, 2000) and has been genetically associated with breast cancer (Kammerer et al., 2005). Based on the positions of these flanking genes the candidate region extends 1.3kb (1,331bp) 5′ and 576bp 3′ of *IL18BP*. The 3′ boundary of the candidate region was selected to include, but not extend beyond, rs7938674 because the next HapMap genotyped SNP was a further 1.9kb downstream, a total of 2.5kb from the end of *IL18BP*. The whole candidate region spanned 8.7kb (8,706bp) (bottom panel Figure 4.2).

4.2.1.1. Tagging SNP selection

Genotyping data was available for a total of 31 SNPs within the selected *IL18BP* candidate region. These SNPs were all from the HapMap database (top panel Figure 4.2). Twenty one of the SNPs in the region were monomorphic in the CEU population, and a further three had a MAF<0.05. These SNPs were therefore not included in the LD analysis. A total of seven SNPs within the candidate region satisfied the inclusion criteria and were therefore included in the LD investigation and subsequent tSNP selection.

A graphical representation of the LD between these seven SNPs in the *IL18BP* candidate region is shown in Figure 4.3. SNPs 8 and 18 were not in high LD (r^2<0.8) with any of the other SNPs in the region, and therefore required direct genotyping. SNPs 6, 14, 20, 21, and 26, are all in complete LD (r^2=1) with each other.

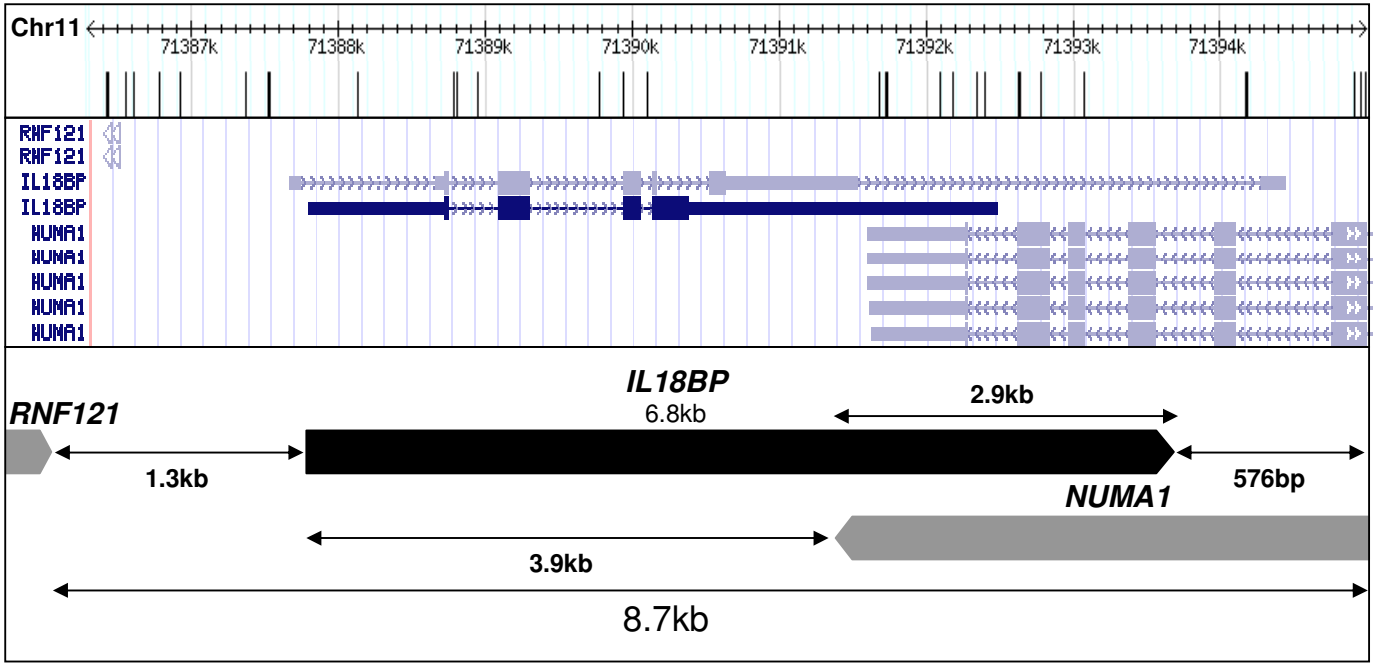


Figure 4.2 *IL18BP* candidate region

The boundaries of the *IL18BP* candidate region are the flanking genes *RNF121* and *NUMA1*, determined by the July 2003 human genome reference sequence (NCBI build 34), hg16 annotation track, shown in the centre panel. The 3′ end of *CLDN16* and the 3′ end of *OSTN* are proximal to *IL18BP*. The approximate (due to scale) positions within this region of all SNPs (irrespective of MAF) in the HapMap phase 2 database (data release #18/phaseII Sept05) are shown in the top panel. The bottom panel shows a schematic (not to scale) of the candidate region with gene orientation, size, and distances given. The gene of interest is shown in black and all other features within the region in grey. The largest isoform of each gene was used for size calculations, and, for clarity, is the only one shown. All sizes and distances are based on chromosomal positions reported in the NCBI build 34 genome reference sequence.

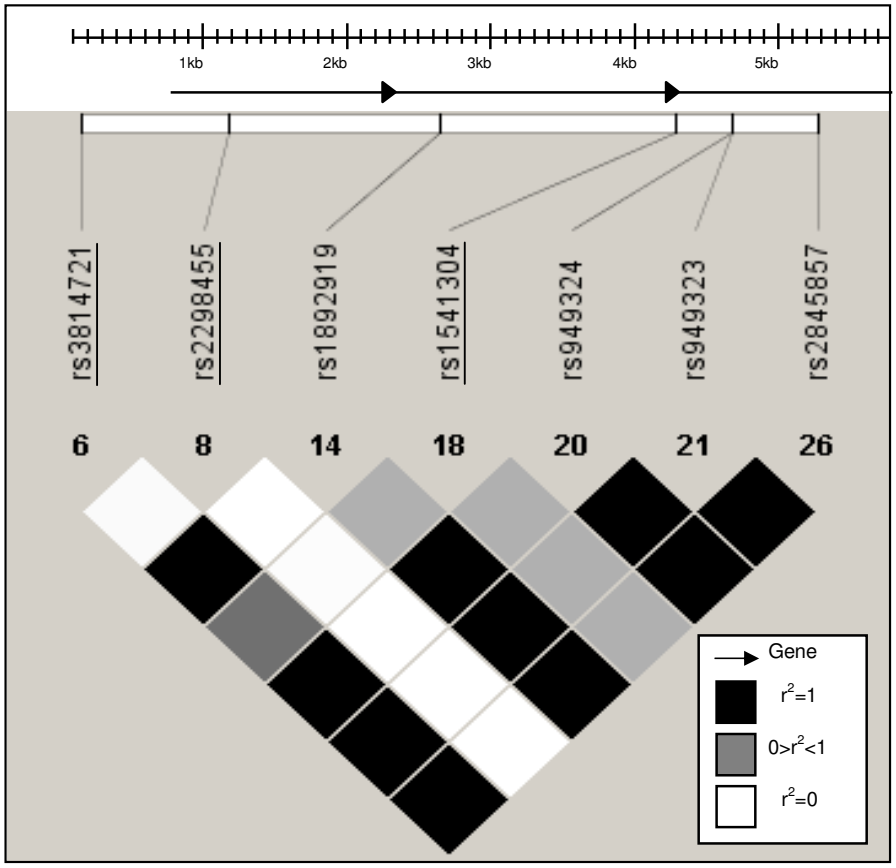


Figure 4.3 LD plot of the *IL18BP* region SNPs

The pairwise LD relationships between SNPs in the *IL18BP* candidate region were investigated to enable tSNP selection. The region shown is comprised of only the SNPs which satisfied the frequency criteria and were included in the LD analysis (n=7), not the whole of the candidate region. The gene orientation and location within the region is indicated. Each coloured square represents the pairwise LD between two SNPs. The graphical representation, generated in Haploview, is shown on a grey scale, with higher r^2 LD represented by darker blocks. The position within the region of each SNP satisfying the selection criteria is represented by a line underneath the gene position. Each SNP is labelled by rs number and Haploview ID number. The three tSNPs which were genotyped are underlined.

A total of three SNPs were selected for genotyping, two singletons (rs2298455 and rs141304) and the other (rs3814721) capturing four SNPs (rs1892919, rs949324, rs949323, and rs2845857) (Table 4.7).

4.2.1.2. Stage-1

All three of the genotyped *IL18BP* candidate region SNPs satisfied the quality control criteria (sections 3.1.2.2.1. and 3.1.2.2.2.) and were therefore analysed for evidence of disease association.

Comparison of the frequencies in the patient (n=130) and control populations (n=146) of the three tSNPs genotyped (Table 4.8, A.) showed that individually none were at a significantly different frequency, although SNP2 was tending towards significance (p=0.07). However, analysis of the SNPs as haplotypes revealed that the two-marker haplotype of SNPs 1 and 2 showed evidence of significant disease association (p=0.007) (Table 4.8, B.). Two of the individual SNP1/SNP2 haplotypes also showed a significant difference. The 1-1 haplotype, with the common allele present at both loci, was significantly less common in the patient than the control cohort (p= 0.02), and the 1-2 haplotype, with the common allele at SNP1 and the rare allele at SNP2, was significantly more frequent in the patients (p=0.01).

Conditional analysis indicated that SNPs 1 and 2 were showing two separate, independent effects. When analysed controlling for the other, both SNPs showed a significant frequency difference (SNP1 conditioning on SNP2 p=0.00868, SNP2 conditioning on SNP1 p=0.00188). Analysis of the SNP1/SNP2 haplotype with no interactions between the SNPs was not significant (p=0.19). A likelihood ratio test comparing the alternative likelihoods of the two models, with and without interactions, showed that there was a significant difference (p=0.0021) in the fit of the models, indicating that this effect was due to an interaction between the two SNPs. The genoplots of *IL18BP* SNP1 and SNP2 (Figure 4.4) show that for both SNP assays the DNA samples separated into three distinct clusters corresponding to each genotype. The GC10 scores for both SNP assays also reflect the quality of the genotype assignment, SNP1=0.871, SNP2=0.760 (Appendix 2), both being in excess of the GC10≥0.5 cut off used for quality control. These quality measurements of the SNP assays enable high confidence in the genotype assignment, and therefore that the significant difference observed is due to true frequency differences between the populations and not genotyping error.

SNP Nº	rs Number	Tagging
1	rs3814721	rs1892919,rs949324,rs949323,rs2845857
2	rs2298455	n/a
3	rs1541304	n/a

Table 4.7 *IL18BP* candidate region tSNPs

A total of three tSNPs were genotyped to capture all seven SNPs within the candidate region. The SNPs captured ($r^2 \geq 0.8$) by each genotyped tSNP are shown. For ease of identification each SNP genotyped was assigned a number according to genome order.

A.

SNP	rs Nº	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
1	rs3814721 (A/G)	1	0.931 (242)	0.938 (274)	1.132 (0.576-2.226)	0.719
		2	0.0692 (18)	0.0616 (18)		
2	rs2298455 (A/C)	1	0.838 (218)	0.890 (260)	1.565 (0.955-2.565)	0.0739
		2	0.162 (42)	0.110 (32)		
3	rs1541304 (G/A)	1	0.965 (251)	0.983 (287)	2.058 (0.681-6.222)	0.1905
		2	0.0346 (9)	0.0171 (5)		

B.

Haplotype	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	Haplotype p value	p value
1-1 (A/A)	0.774 (201.3)	0.857 (250.2)	1	0.0220	0.00777
1-2 (A/C)	0.157 (40.73)	0.0816 (23.82)	2.125 (1.231-3.669)	0.0129	
2-1 (G/A)	0.0643 (16.73)	0.0336 (9.821)	2.117 (0.918-4.884)	0.201	
2-2 (G/C)	0.00490 (1.273)	0.0280 (8.179)	0.194 (0.025-1.478)	0.0847	

Table 4.8 *IL18BP* candidate region stage-1 analysis results

Three genotyped tSNPs within the *IL18BP* candidate region were analysed for significant allele frequency differences between the patient (n=130) and control (n=146) cohorts. The frequency in each cohort of both the common (1) and rare (2) alleles of the three individual SNPs are given in A. The alternative alleles of each SNP are shown underneath the rs identifier. The allele assigned as the common (1) allele is given first and the rarer (2) allele second. The OR shown is for allele 2 (allele 1 OR=1). B. shows the frequencies of each of the four SNP1/SNP2 haplotypes. The haplotypes are presented so that the first number is the allele present at SNP1 and the second number is the allele present at SNP2. The p-value for the test for frequency differences between the cohorts is also shown.

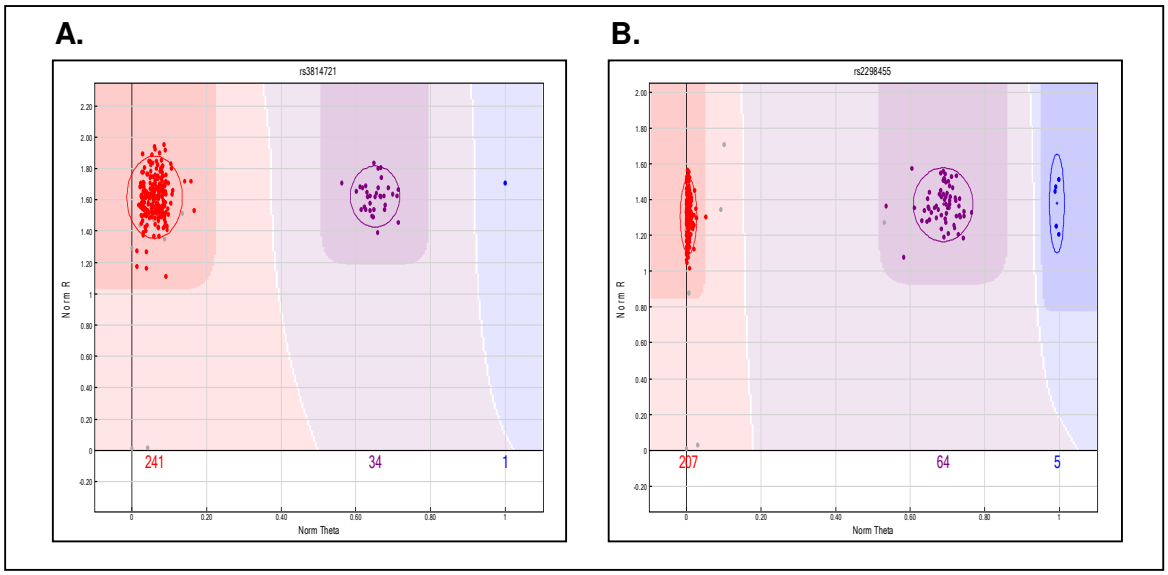


Figure 4.4 Genoplots of the significant *IL18BP* SNPs

Shown are the genoplots of the two *IL18BP* region SNPs involved in the haplotype showing evidence of disease association in stage-1 of the association study. Normalised theta values (deviation from total 'A' allele signal) are plotted along the x-axes, and normalised R values (signal intensity) along the y-axes. Each data point represents one sample, the circles indicate the location of each genotype cluster, and the dark shaded areas are the call areas.

A. SNP1 rs3814721, B. SNP2 rs2298455

4.2.1.3. Stage-2

IL18BP candidate region SNP1 and SNP2, which showed evidence of significant frequency differences in stage-1 of the association study as a haplotype, were genotyped in the stage-2 patient (n=105) and control cohorts (n=184). As in the stage-1 cohorts, neither *IL18BP* SNP showed evidence of frequency differences when analysed as individual SNPs (Table 4.9, A). Analysis of the two SNPs as a haplotype, conversely to the stage-1 results, also showed no evidence, neither overall nor as individual haplotypes, of any frequency differences between the case and control cohorts (Table 4.9, B).

4.2.1.3.1. Meta-analysis

There is no satisfactory method to perform stratified analysis of a haplotype. Ideally a Cochran-Mantel-Haenszel test, and a Breslow-Day test for homogeneity of odds, would be performed. The best option available was to pool all of the patient (n=235) and control (n=330) samples used in both stages of the study and perform the analysis on the larger cohort sizes. Analysis of the pooled genotyping data was performed as in stage-1 using UNPHASED.

There was no evidence of any significant differences in the frequencies of either SNP individually when the cohorts from both study stages were pooled Table 4.10, A. As shown in Table 4.10 B, when all the individuals genotyped in both stages of the study were pooled together a significant (p=0.012) association of the haplotype with sJIA was seen. In agreement with the results from the first stage of the study, both the 1-1 and 1-2 haplotypes showed a significant (p=0.031 and 0.032 respectively) difference in frequency between the control and patient groups.

4.2.1.4. WTCCC control cohort analysis

Both *IL18BP* candidate region SNPs involved in the associated SNP1/SNP2 haplotype had been genotyped directly in phase-2 of the WTCCC project (n=4671) and so were available for direct analysis against the pooled sJIA patient cohort (n=235).

In contrast to both the stage-1 and stage-2 analysis results using the smaller control cohorts, significant (p=0.0455) and near significant (p=0.0524) frequency differences for SNP1 and

A.

SNP	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
1	1	0.937 (193)	0.945 (346)	1.165 (0.567-2.394)	0.679
	2	0.063 (13)	0.055 (20)		
2	1	0.879 (181)	0.885 (322)	1.059 (0.625-1.795)	0.832
	2	0.121 (25)	0.115 (42)		

B.

Haplotype	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	Haplotype p value	p value
1-1	0.828 (170.6)	0.847 (308.4)	1	0.611	0.888
1-2	0.109 (22.36)	0.098 (35.56)	1.136 (0.744-2.240)	0.760	
2-1	0.050 (10.36)	0.037 (13.56)	1.380 (0.782-3.957)	0.548	
2-2	0.013 (2.643)	0.018 (6.436)	0.742 (0.026-1.485)	0.837	

Table 4.9 *IL18BP* candidate region stage-2 analysis results

The two *IL18BP* candidate region SNPs involved in the two-marker haplotype identified in stage-1 of the association study, were genotyped in the stage-2 patient (n=105) and control (n=184) cohorts. The frequency in each cohort of both the common (1) and rare (2) alleles of the three individual SNPs are given in A. The OR shown is for allele 2 (allele 1 OR=1). B. shows the frequencies of each of the four SNP1/SNP2 haplotypes. The p-value for the test for frequency differences between the cohorts is also shown.

A.

SNP	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
1	1	0.935 (435)	0.942 (620)	1.163 (0.712-1.898)	0.547
	2	0.0665 (38)	0.0578 (38)		
2	1	0.856 (399)	0.887 (582)	1.321 (0.927-1.882)	0.125
	2	0.144 (67)	0.113 (74)		

B.

Haplotype	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	Haplotype p value	p value
1-1	0.796 (370.8)	0.852 (558.7)	1	0.031	0.012
1-2	0.138 (64.16)	0.090 (59.31)	1.630 (1.109-2.395)	0.032	
2-1	0.0604 (28.16)	0.036 (23.31)	1.820 (1.006-3.292)	0.149	
2-2	0.0061 (2.842)	0.022 (14.69)	0.291 (0.056-1.510)	0.157	

Table 4.10 *IL18BP* candidate region pooled cohorts analysis results

As there is currently no method to perform meta-analysis of haplotypes, the cohorts from both stages of the association study were pooled (patient n=235, control n=330) for analysis of the two marker *IL18BP* candidate region haplotype identified in stage-1 of the association study. The frequency in each cohort of both the common (1) and rare (2) alleles of the three individual SNPs are given in A. The OR shown is for allele 2 (allele 1 OR=1). B. shows the frequencies of each of the four SNP1/SNP2 haplotypes. The p-value for the test for frequency differences between the cohorts is also shown.

SNP2 respectively were seen when analysed as individual SNPs (Table 4.11, A).

Analysis of the two SNPs as a haplotype confirmed the previous results, showing evidence of significant (p=0.00931) disease association (Table 4.11, B). As in the analysis with the pooled in-house control cohorts, the 1-1 haplotype was at a significantly higher frequency in the controls (p=0.00202), and the 1-2 haplotype at a higher frequency in the cases (p=0.0314). However, unlike the previous analyses the 2-1 haplotype also showed a significantly higher frequency (p=0.0112) in the patients. The WTCCC control cohort showed very similar frequencies of all four individual haplotypes to the control cohorts in the original analysis (Table 4.10).

4.2.1.5. Associated SNPs

In the initial tSNP selection based on the HapMap 2 genotyping data (Table 4.7) SNP1 was tagging four other SNPs with r^2=1. For ease of identification these SNPs were labelled as SNPs1a-d according to chromosomal order, SNP1a=rs1892919, SNP1b=rs949324, SNP1c=rs949323, SNP1d=rs23845857. SNP2 was not tagging any other SNPs within the candidate region. Although the initial LD analysis indicated that SNP1 was in complete LD (r^2=1) with these four SNPs, the subsequent validation using the larger HapMap3 and WTCCC2 populations did not replicate these results with regards to SNP1c, r^2=0.626 and 0.435 respectively (Table 4.5). As both larger, and therefore more accurate, datasets calculate the LD as r^2<0.8, the tag selection criteria, it was determined that SNP1 is not tagging SNP1c. As SNPs1a, b, and d were not included in HapMap3 or WTCCC2 their LD relationship with SNP1 can not be confirmed. It must therefore be assumed that they are tagged by SNP1, but viewed with caution because of the lack of validation.

4.2.1.5.1. Additional captured SNPs

Based on the HapMap3 dataset 52 SNPs were identified by SNAP (section 3.1.5.) as being in r^2>0.8 with SNP1 (Table 4.12). SNP2 is not in high LD (r^2<0.8) with any SNPs within 500kb. All SNP1 tagged SNPs are outside of the selected candidate region, in and surrounding the nearby genes. These include *NUMA1*, *RNF121*, leucine rich transmembrane and 0-methyltransferase domain containing (*LRTOMT*), chromosome 11 open reading frame 51 (*C11orf51*), and the transient receptor potential cation channel, subfamily C, member 2-like

A.

SNP	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
1	1	0.934 (435)	0.955 (8918)	1.499 (1.028-2.186)	0.0455
	2	0.0665 (31)	0.0454 (424)		
2	1	0.856 (399)	0.887 (8282)	1.312 (1.005-1.713)	0.0524
	2	0.144 (67)	0.114 (1060)		

B.

Haplotype	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	Haplotype p value	p value
1-1	0.796 (370.8)	0.852 (7953)	1	0.00202	0.00931
1-2	0.138 (64.16)	0.103 (962.7)	1.429 (1.083-1.887)	0.0314	
2-1	0.0604 (28.16)	0.0350 (326.7)	1.849 (1.224-2.803)	0.0112	
2-2	0.0061 (2.842)	0.0104 (97.34)	0.626 (0.132-2.994)	0.884	

Table 4.11 *IL18BP* candidate region analysis results with WTCCC2 controls

As an additional analysis with a larger, independent control cohort, allele frequencies in the pooled cases from both stages of the association study (n=235) were compared to the controls genotyped in phase-2 of the WTCCC project (n=4671). The frequency in each cohort of both the common (1) and rare (2) alleles of the two individual SNPs are given in A. The OR shown is for allele 2 (allele 1 OR=1). B. shows the frequencies of each of the four SNP1/SNP2 haplotypes. The p-value for the test for frequency differences between the cohorts is also shown.

SNP	r^2	Location
rs594620	0.849	downstream *LOC100133315*
rs531326	0.849	downstream *LOC100133315*
rs548127	0.849	downstream *LOC100133315*
rs674688	0.849	*LOC100133315*
rs661242	0.924	LOC100133315
rs1612492	0.849	*LOC100133315*
rs669082	1	LOC100133315
rs4945340	1	LOC100133315
rs4944232	1	*RNF121*
rs7102192	1	*RNF121*
rs10736781	0.867	*RNF121*
rs7948821	1	*RNF121*
rs4944242	1	*RNF121*
rs7931941	1	*RNF121*
rs4121395	1	*RNF121*
rs6592448	1	*RNF121*
rs10751189	1	*RNF121*
rs6592450	1	*RNF121*
rs6592451	1	*RNF121*
rs7128974	1	*RNF121*
rs6592453	1	*RNF121*
rs6592455	1	*RNF121*
rs7121260	1	*RNF121*
rs7128317	1	*RNF121*
rs1541306	1	*RNF121*
rs2298456	0.854	*NUMA1*
rs10128658	0.854	*NUMA1*
rs2298457	0.854	*NUMA1*
rs1939243	0.854	*NUMA1*
rs1573502	0.854	*NUMA1*
rs4945426	0.854	*NUMA1*
rs1939242	0.854	*NUMA1*
rs10793016	0.854	*NUMA1*
rs4338555	0.854	*NUMA1*
rs4945434	0.854	*NUMA1*
rs679926	0.854	*NUMA1*
rs560777	0.854	*NUMA1*
rs541022	0.854	*NUMA1*
rs3018302	0.854	*NUMA1*
rs3018301	0.854	*LRTOMT*
rs2511114	0.854	*LRTOMT*
rs673478	0.854	*LRTOMT*
rs2511079	0.854	*LRTOMT*
rs659513	0.854	*LRTOMT*/*C11orf51*
rs4943821	0.854	*LRTOMT*
rs17161966	0.854	*LRTOMT*
rs17161980	0.854	*LRTOMT*
rs3750908	0.854	*C11orf51*
rs3793938	0.854	*C11orf51*
rs11235440	0.854	upstream *C11orf51*
rs7113549	0.854	upstream *C11orf51*

Table 4.12 Additional SNPs tagged by *IL18BP* SNP1

Based in the HapMap3 population genotyping data all SNPs in r^2>0.8 with, and within 500kb of, the *IL18BP* SNPs showing evidence of significant disease association, were identified. The SNP identifier, r^2 value, and location of all the identified SNPs are shown. All SNPs shown are tagged by SNP 1 as SNP2 was not in r^2>0.8 with any SNPs within 500kb.

mRNA (*LOC100133315*). *LOC100133315* is upstream of *RNF121*, *LRTOMT* and *C11orf51* are upstream of *NUMA1* (Figure 4.5). Transient receptor potential channel 2 (*TRPC2*), to which *LOC100133315* is similar, is a nonselective Ca^{2+} permeable cation channel, with roles in activation of the acrosome reaction in sperm, and initiation of action potential in sensory neurons in response to pheromones (Yildirim and Birnbaumer, 2007). *LRTOMT* encodes two different proteins, a leucine-rich transmembrane protein of unknown function and an O-methyltransferase, defects in which can cause nonsyndromic deafness (Ahmed et al., 2008). *C11orf51* encodes Pdro, protein associated with detergent-resistant membranes and endosomes. Protein and mRNA levels are regulated by cellular cholesterol content, and overexpression results in reduced cholesterol cell content, suggesting involvement in the regulation of cholesterol homeostasis (Guillaumot et al., 2010).

Because of their location within non-candidate genes, which based on function are unlikely to be involved in the pathogenesis of sJIA, these additional captured SNPs were not investigated further in this study.

4.2.1.5.2. Associated SNP positions

As shown in Figure 4.6, SNP1 is upstream of all four *IL18BP* isoforms. Relative to the transcription start site (TSS) it is -665bp of isoforms A and D, -1701 of isoform B, and -837 of isoform C. SNP1 is also located within an experimentally identified silencer element involved in regulation of IFNγ-induced IL-18BP expression (Hurgin et al., 2002). SNP2 is within intron 2 of isoform A, at -495 of isoform B, in the 5′UTR of isoform C, and in intron 1 of isoform D. SNP1a is intronic in all four *IL18BP* isoforms. SNP1b is downstream of isoforms A and B, in the 3′UTR of isoform C, intron 6 of isoform D, and within the 3′UTR of the flanking gene *NUMA1*. SNP1d is downstream of *IL18BP* isoforms A, B, and C, in intron 6 of isoform D, and in intron 25 of *NUMA1*.

4.2.1.6. Summary

Following analysis of the data from both study stages, a two marker haplotype of SNP1/SNP2 (rs3814721/rs2298455) was identified as showing evidence of a significant association with sJIA. This result was confirmed when the analysis was repeated with the larger WTCCC control cohort (n=4671). SNP1, which is located within a biologically

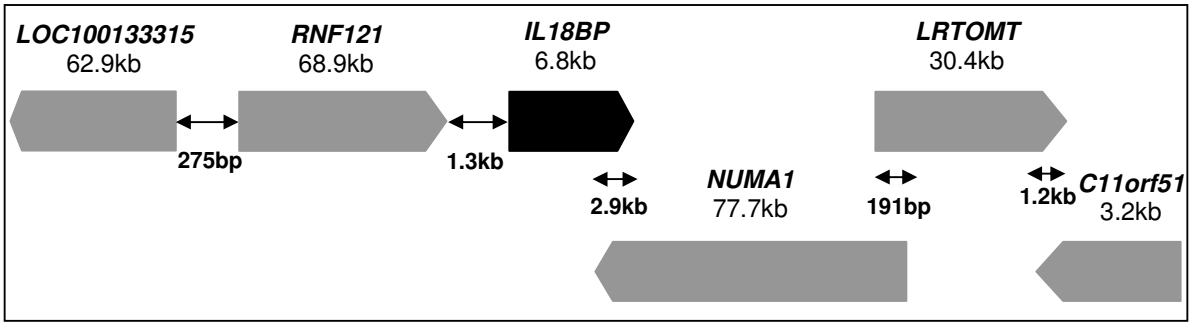


Figure 4.5 *IL18BP* flanking genes

All of the additional SNPs identified as being in high LD with *IL18BP* SNP1 are outside of the selected candidate region. Shown are the locations of the nearby genes within which these SNPs are located.

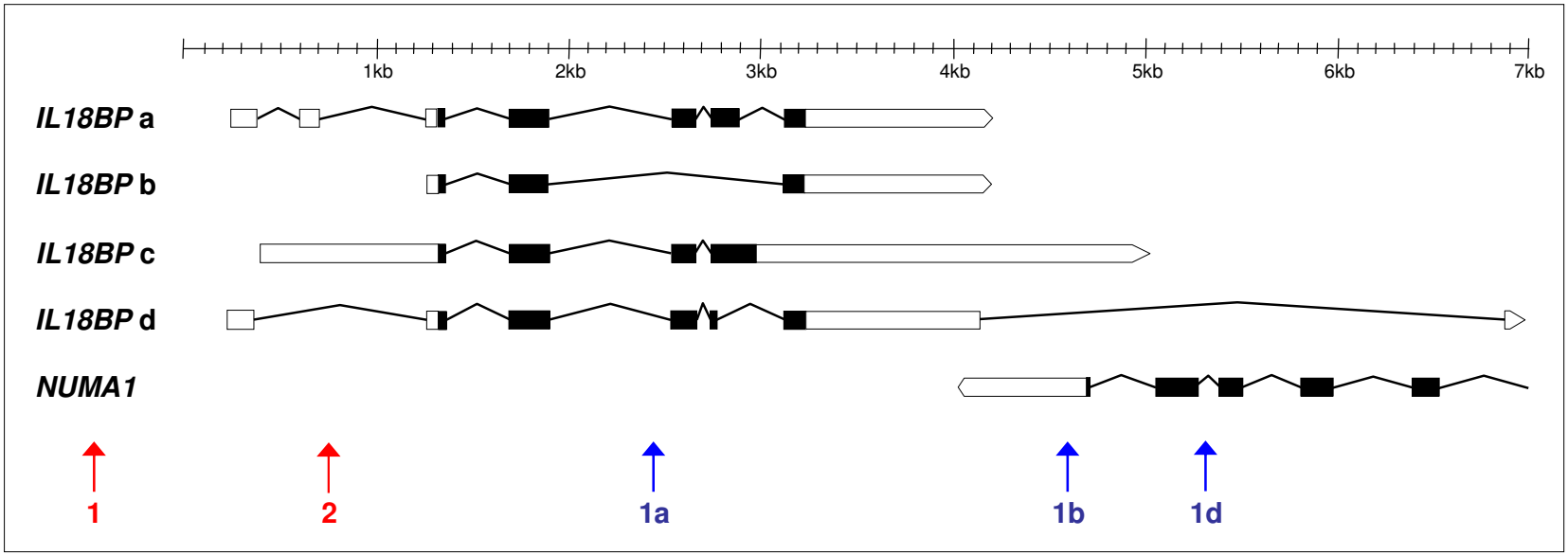


Figure 4.6 Positions of the associated SNPs in the *IL18BP* candidate region

The positions within the *IL18BP* candidate regions of the genotyped SNPs of the associated two marker haplotype, SNP1 (rs3814721), and SNP2 (rs2298455), are shown in red, and the SNPs tagged by SNP1, SNP1a (rs1892919), SNP1b (rs949324), and SNP1d (rs2845857), are shown in blue. All four *IL18BP* isoforms, and the flanking gene *NUMA1*, are shown. Untranslated regions are represented by white boxes and coding exons by black boxes.

important IFNγ response silencer element, is in high LD with three other SNPs within the candidate region, SNP2 is not tagging any other SNPs.

4.2.1.7. Discussion

IL-18BP is a specific inhibitor of the pro-inflammatory cytokine IL-18 (Novick et al., 1999; Aizawa et al., 1999). There is a regulatory feedback loop between the two proteins, IL-18 induces IFNγ production (Nakamura et al., 1993b; Okamura et al., 1995b; Ushio et al., 1996b), which in turn induces IL-18BP expression (Muhl et al., 2000b; Paulukat et al., 2001b; Veenstra et al., 2002). Elevated levels of IL-18 are therefore normally accompanied by elevated levels of IL-18BP. Although high levels of IL-18 have been reported in sJIA (Maeno et al., 2002), elevated levels of IL-18BP have not been reported. An imbalance of IL-18/IL-18BP towards a pro-inflammatory state may therefore be involved in the chronic inflammation present in sJIA. This imbalance has been demonstrated in secondary hemophagocytic syndrome, an uncontrolled activation of Th1 cells and macrophages which shares features with MAS (Mazodier et al., 2005). It was hypothesised that genetic variants in *IL18BP* altering expression regulation may be involved in sJIA pathogenesis, and in MAS secondary to sJIA.

Only a small number of studies investigating *IL18BP* variation in disease have been published. Two of these studies were unable to perform association analysis due to low MAFs of the investigated SNPs in the type 1 diabetes mellitus (Nolsoe et al., 2003) and myocardial infarction (Tiret et al., 2005b) patient populations used. *IL18BP* SNP1 (rs3814721) was investigated for involvement in schizophrenia/schizoaffective disorder, but no significant frequency differences were found (Shirts et al., 2008). *IL18BP* SNP2 (rs2298455) was however shown to be significantly associated with hepatitis C virus clearance in European Americans. An additional SNP (rs3750912) within exon 1 of *NUMA1*, which was not captured in the study presented in this manuscript, was found to be associated with virus clearance in African Americans (Mosbruger et al., 2010). The three *IL18BP* candidate region SNPs selected as part of this project (Table 4.7) were investigated in relation to coronary heart disease by a collaborating group. No frequency differences of the individual SNPs or the SNP1/SNP2 haplotype were detected between healthy individuals and patients who underwent coronary artery bypass graft surgery. Although an *IL18* haplotype was found to be related to circulating IL-18 levels, no link was observed between *IL18BP* variation and

levels of IL-18 or IL-18BP (Thompson et al., 2007), suggesting that either *IL18* has a more important role than *IL18BP* in regulation of expression, or that the *IL18BP* association is specific to only sJIA and not other inflammatory disorders studied thus far.

There are a number of possible reasons for an identified association being with a two SNP haplotype rather than an individual polymorphism. It may be that these two SNPs in combination directly alter the expression of the gene, causing disruption of the inflammatory response and resulting in disease susceptibility. It is also possible that these two SNPs are not directly responsible for the association but rather are acting as markers for another, unknown SNP which was not included in the initial SNP list, either because it has a MAF<0.05 or is outside of the gene region investigated. Another possible explanation is that both SNPs individually play a role in disease susceptibility, for which there was insufficient power to detect association, and significance was only achieved when investigated in combination. Although neither SNP showed a significant effect individually in the original analysis when the larger WTCCC control cohort was used SNP1 was significant individually and SNP2 was tending towards significance, suggesting that the latter scenario may be the case.

Although SNP2 is not in high LD with any other SNPs, SNP1 is acting as a tagging SNP for three other SNPs within the candidate region. Therefore SNP1 may not be involved in sJIA susceptibility, as it is in complete LD (r^2=1) with three other SNPs, it is possible that any of these are responsible. Based on HapMap3, SNP1 was identified as also being a proxy for a total of 52 additional SNPs, all of which are within genes flanking *IL18BP*. Based on function, these genes, and therefore their associated SNPs, are unlikely to be involved in the pathogenesis of sJIA. Despite being within other genes it is possible that these additional tagged SNPs are involved in the regulation of IL-18BP expression and therefore cannot be completely eliminated as candidates. It is not possible to discern which SNP is responsible through association analysis alone, further investigation for roles in regulation of expression would be required.

The issues of low power to detect significant associations, multiple testing, and the lack of significant findings in the stage-2 populations alone are discussed in chapter 5.

4.2.2. Comparative Genomics

In order to include all five *IL18BP* candidate region SNPs of interest, the sequence investigated was extended 850bp upstream and 1.5kb downstream of *IL18BP*. The major 'A' *IL18BP* isoform was used to annotate the gene features within the sequence.

As shown in Figure 4.7, A. all *IL18BP* isoform A exons, except for exon seven, are well conserved between human, mouse, and rat. There is however little conservation of either human UTR, and very little homology in the 1kb upstream of the start of transcription. The most notable homology is in the sequence downstream of *IL18BP*, in the 5′ UTR and coding exons of *NUMA1*.

SNP1, SNP2, and SNP1a are all located in regions with no significant conserved homology, while SNP1b and SNP1d are both within conserved regions. SNP1b is in a highly conserved region of the *NUMA1* 5′ UTR (Figure 4.7, B.), and SNP1d is in a conserved section of *NUMA1* intron two (Figure 4.7, C.).

These results can also be seen when the nucleotide alignment of the 20bp surrounding each SNP of interest is visualised (Figure 4.8). For both SNPs 1 and 1a there is no nucleotide sequence present in both murine species investigated which align to the surrounding sequences in the human genome. Despite the presence of the same nucleotide in both human and rat at the SNP2 loci there is no significant homology evident in the surrounding sequence. SNP1b is fully conserved in the three species investigated. It is the second base in a run of seven consecutive nucleotides which are fully conserved in both murine species. There is also high conservation homology over all 20 of the nucleotides flanking this SNP. The ten nucleotides upstream of SNP1d are not well conserved between the species. However, SNP1d is fully conserved between all three species, and is the first nucleotide in a run of 11 with near total conservation, only two nucleotides are different in human.

4.2.2.1. Discussion

Of the five SNPs of interest within the *IL18BP* gene region investigated, only two are within regions of evolutionary sequence conservation. Both of these SNPs are located within conserved regions of the flanking gene *NUMA1*. These regions may however also play regulatory roles in IL-18BP expression. Although no evidence for the remaining SNPs being located within regulatory regions has been demonstrated through the sequence comparison,

lack of conservation does not necessarily mean that no regulatory region is present. It is possible that *IL18BP* is differentially regulated in humans, or that different recognition sequences have evolved for some regulatory factors. Under either circumstance no evolutionary sequence conservation would be evident, despite the presence of regulatory regions. Empirical investigation is required to determine if the disease associated SNPs are located within regulatory regions.

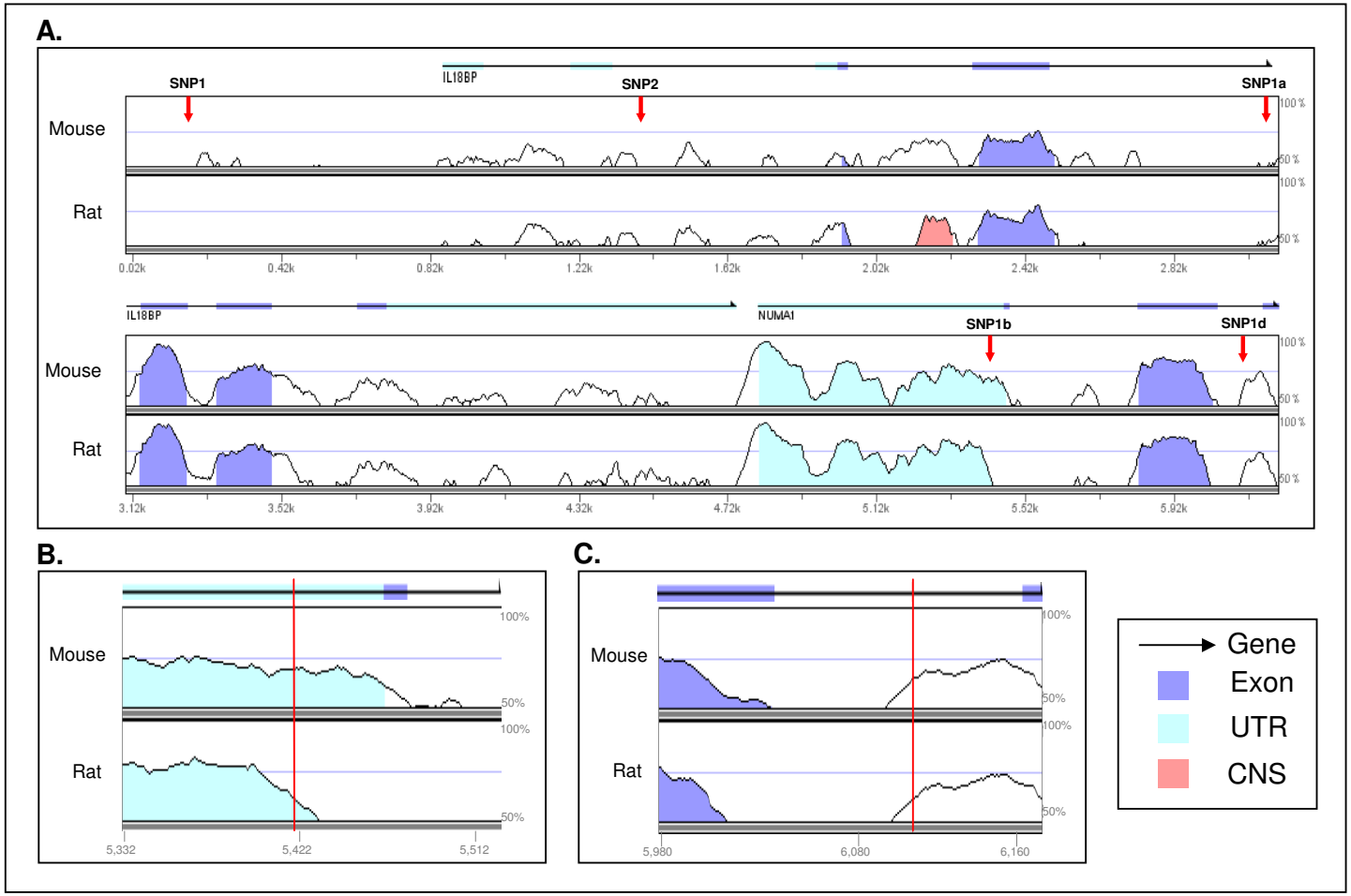


Figure 4.7 VISTA conservation homology plot of *IL18BP*

The visual alignment generated by VISTA of mouse and rat *IL18BP* and flanking regions (1kb upstream, 1.5kb downstream) aligned against the human sequence is shown as a peak-and-trough plot. The location and features, pink=CNS, blue=exon, turquoise=UTR, (*IL18BP* isoform A) of the human genes are shown above the plot. The base sequence is plotted on the x axis and percentage identity (≥50%) to the human sequence along the y axis. 75% identity indicated on the plot by the horizontal blue line. All regions of ≥70% identity to the human sequence over a minimum of 100bp are highlighted in colour under the peak. Alignment of the full gene sequence and flanking regions is depicted in A. The red arrows indicate the positions of the SNPs of interest within the alignment. Enlargements of the alignment of the sequence surrounding SNP1b (B.), and SNP1d (C.), which are indicated by a red line, are also shown.

UTR = untranslated region, CNS=conserved non-coding sequence

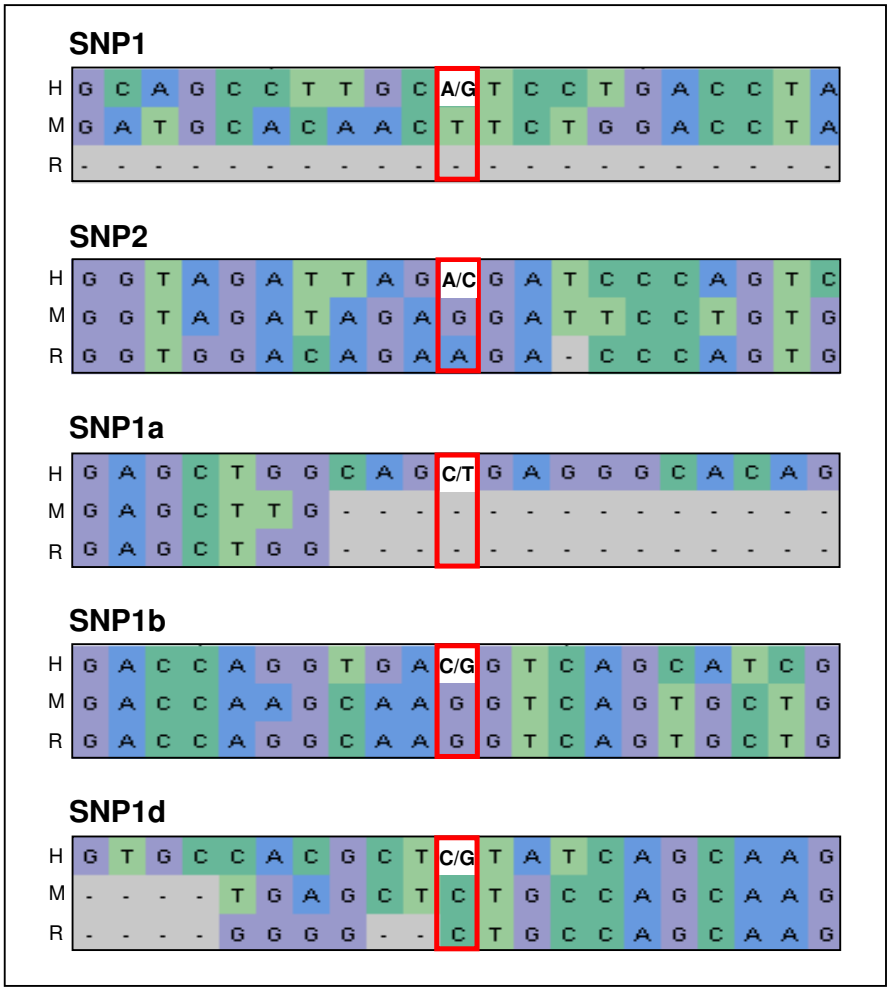


Figure 4.8 Sequence alignment of the *IL18BP* SNP nucleotides

The base pair alignment of the mouse (middle) and rat (bottom) sequences aligned against the human sequence (top) for each SNP of interest (within the red box, both alleles indicated) and 10 flanking nucleotides, are shown. Where there is no corresponding nucleotide in the mouse or rat sequences a '-' is shown.

H=Human sequence, M=Mouse sequence, R=Rat sequence

4.2.3. Investigation of expression levels according to haplotype

4.2.3.1. Haplotype determination

Of the 55 individuals collected in the WC cohort only 40 were still available for PBMC collection at the time of the expression study. The *IL18BP* SNP1 and SNP2 genotypes of the 40 available WC healthy individuals, as well as the ethnicity and sex of each individual, are shown in Table 4.13. The inferred haplotype carriage of each individual is also shown. Individual WC20 was heterozygous at both of the SNP positions and so haplotype phase could not be inferred with certainty based solely on the genotypes. All of the other individuals were homozygous for at least one SNP so haplotype carriage could be determined. Twenty six individuals carried two copies of the 1-1 haplotype, nine individuals one copy of the 1-1 haplotype and one copy of the 1-2 haplotype, two individuals one copy of the 1-1 haplotype and one copy of the 2-1 haplotype, and two individuals carried two copies of the 2-1 haplotype.

4.2.3.1.1. Haplotype determination of uncertain phase

Five clones of WC20 genomic DNA, containing the two *IL18BP* associated haplotype SNPs, were sequenced to determine haplotype phase. By aligning the sequencing results against the genomic reference sequence the alleles present at each SNP loci were identified (Figure 4.9 A. and B.). Clones one, four, and five, contained the T allele at SNP1 and the A allele at SNP2 (2-1 haplotype), and clones two and three, C alleles at both loci (1-2 haplotype) (Figure 4.9 C.). These results show that individual WC20 carries the 1-2 and 2-1 *IL18BP* haplotypes. The haplotype frequencies in this control population are therefore: 1-1 (0.79), 1-2 (0.13), and 2-1 (0.088).

4.2.3.2. Protein expression levels

4.2.3.2.1. Optimisation of cell concentration

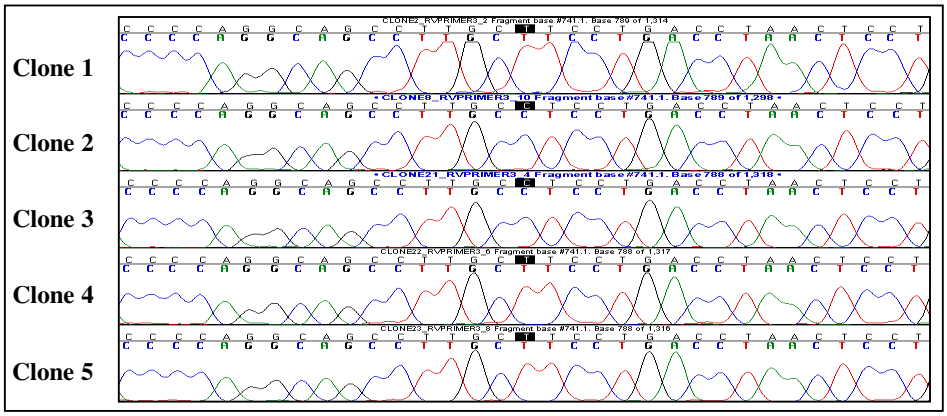
To optimise the cell concentration at which to measure IL-18BP mRNA and protein levels an initial experiment was carried out with PBMCS cultured at $5x10^6$, $2.5x10^6$, and $1x10^6$ cells per ml. The highest PBMC concentration, $5x10^6$ cells/ml, was more difficult to resuspend, and so homogenise the cell suspension, in the media. This increases the possibility of variation in cell concentration between samples. The cells were also very dense in the culture plate, forming a film of cells visible to the naked eye, therefore, the cells would not have all

	Ethnicity	Sex	SNP1	SNP2	Haplotype
WC01	Caucasian	F	1:1	1:2	1-1 & 1-2
WC02	Caucasian	F	1:1	1:1	1-1
WC03	African	F	1:1	1:2	1-1 & 1-2
WC04	Caucasian	F	1:1	1:2	1-1 & 1-2
WC05	Caucasian	M	1:1	1:1	1-1
WC07	Caucasian	M	1:2	1:1	1-1 & 2-1
WC09	Caucasian	F	1:1	1:2	1-1 & 1-2
WC11	African	M	2:2	1:1	2-1
WC13	Caucasian	F	1:1	1:1	1-1
WC14	Caucasian	M	1:1	1:1	1-1
WC15	South American	M	2:2	1:1	2-1
WC16	Caucasian	F	1:1	1:1	1-1
WC17	Caucasian	M	1:1	1:2	1-1 & 1-2
WC19	South Asian	F	1:1	1:1	1-1
WC20	Caucasian	F	1:2	1:2	1-1 & 2-2 OR 1-2 & 2-1
WC22	Caucasian	M	1:1	1:1	1-1
WC23	Caucasian	F	1:1	1:1	1-1
WC24	South Asian	F	1:1	1:1	1-1
WC25	Caucasian	M	1:1	1:1	1-1
WC27	Caucasian	M	1:1	1:2	1-1 & 1-2
WC28	Caucasian	M	1:1	1:2	1-1 & 1-2
WC29	Chinese	F	1:1	1:1	1-1
WC31	Caucasian	F	1:2	1:1	1-1 & 2-1
WC32	Caucasian	M	1:1	1:1	1-1
WC33	Chinese	F	1:1	1:1	1-1
WC34	Caucasian	F	1:1	1:1	1-1
WC35	Caucasian	F	1:1	1:2	1-1 & 1-2
WC42	Middle Eastern	F	1:1	1:2	1-1 & 1-2
WC43	Caucasian	F	1:1	1:1	1-1
WC44	South Asian	F	1:1	1:1	1-1
WC45	Caucasian	F	1:1	1:1	1-1
WC46	South Asian	F	1:1	1:1	1-1
WC47	Middle Eastern	F	1:1	1:1	1-1
WC48	Caucasian	F	1:1	1:1	1-1
WC49	Caucasian	M	1:1	1:1	1-1
WC50	Caucasian	M	1:1	1:1	1-1
WC51	Caucasian	F	1:1	1:1	1-1
WC52	Caucasian	F	1:1	1:1	1-1
WC54	Caucasian	M	1:1	1:1	1-1
WC55	Chinese	M	1:1	1:1	1-1

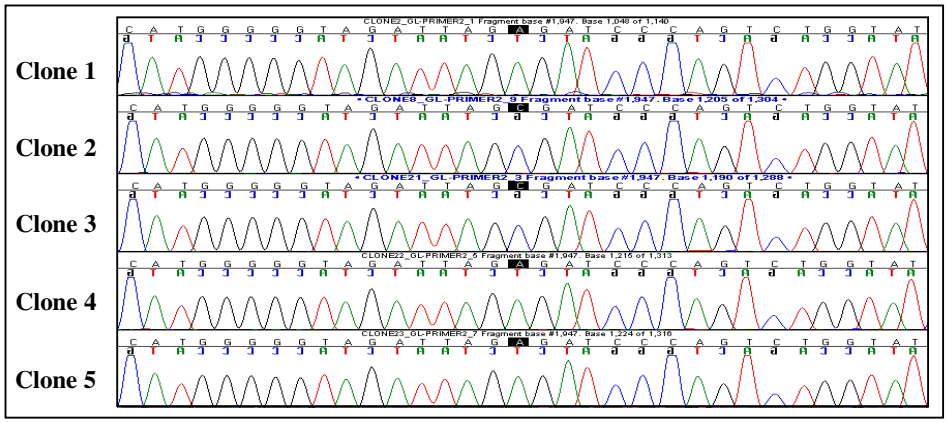
Table 4.13 *IL18BP* SNP1 and 2 genotypes of WC healthy controls

DNA samples from the WC healthy individuals were genotyped for the two sJIA associated *IL18BP* SNPs. Information only on the individuals still working within the division at the time of the expression study is shown. Genotypes are shown as the common and rare allele (SNP1 1=T, 2=C, SNP2 1=A, 2=C). The inferred haplotypes of each individual based on their genotypes are shown, when only one haplotype is shown the individual is homozygous for that haplotype. The sex and ethnicity of each individual is also shown.

A.



B.



C.

Clone	SNP1 (C=1, T=2)	SNP2 (A=1, C=2)	Haplotype
1	T (2)	A (1)	T-A (2-1)
2	C (1)	C (2)	C-C (1-2)
3	C (1)	C (2)	C-C (1-2)
4	T (2)	A (1)	T-A (2-1)
5	T (2)	A (1)	T-A (2-1)

Figure 4.9 Haplotype determination sequencing results

In order to determine haplotype phase of the individual heterozygous at both *IL18BP* SNP loci, the genomic region containing both SNPs was cloned into a vector and sequenced. The chromatograms of the sequence surrounding SNP1 (A.) and SNP2 (B.) were aligned against a reference sequence. The SNP positions in each of the five clones are highlighted. The alleles present at both SNPs in each of the five clones, and the inferred haplotypes are shown in C.

been equally exposed to the stimulant. At each of the three cell concentrations the level of IL-18BP in the media after 0, 3, and 6 hours stimulation was below the detection limit of the ELISA assay. As expected, all of the samples showed upregulation of IL-18BP after 24 hours of stimulation, further increased at 48 hours. The OD (450nm) readings for both 24 and 48 hours at $5x10^6$ cells/ml, and the 48 hour sample at $2.5x10^6$ for one individual, were in excess of the 3.5 OD reading limit of the plate reader, and so were unquantifiable. The ELISA was repeated using 1 in 2, and 1 in 4 dilutions of both the 24 and 48 hour samples. The results are shown as the average ±SD of the calculated level of IL-18BP of the three individuals (Figure 4.10). As expected, as the cell concentration increased so did the level of IL-18BP expression. With the cells at $1x10^6$ IL-18BP levels were 240 pg/ml (± 159) at 24 hours stimulation and 1140 pg/ml (± 200) at 48 hours, at $2.5x10^6$ cells/ml 1567 (±232) and 2785 pg/ml (± 863) respectively, and 2875 (± 866) and 4376 pg/ml (± 1093) respectively when the cells were at $5x10^6$ /ml. Although the highest cell concentration produced the highest levels of IL-18BP, the variation and inaccuracy introduced into the experiment due to the difficulties encountered working with this cell concentration, this was considered to be unacceptable. As a result $2.5x10^6$ cells/ml was chosen as the optimal concentration for use in this experiment, as at still produced sufficient IL-18BP to allow variations between individuals to be observed.

4.2.3.2.2. IL-18BP gene expression comparison according to haplotype

Due to the volume of blood required it was not possible to do experimental replicates of each individual at each time point. Each sample was however measured by ELISA in triplicate. The standard deviation between sample replicates were all < ±0.5 OD units. Only one 3 hour and three 6 hour stimulation samples contained detectable levels of IL-18BP, although as they are on the lower detection limits of this assay they are unlikely to be very accurate (Table 4.14). The expression levels of IL-18BP at the 24 and 48 hour time points were highly variable between individuals, ranging from 163.26 pg/ml to 3384.84 pg/ml at 24 hours, and from 441.41 pg/ml to 6244.72 pg/ml at 48 hours. This variation was evident between individuals of the same IL-18BP haplotype as well as between those of different haplotypes. For example, of the two individuals carrying two copies of the 2-1 haplotype, following 24 hours stimulation 2840pg/ml was detected from the PBMCs from WC15, while only 547pg/ml was detected for WC11. This variation between individuals within the same haplotype group is also evident in the 1-1/2-1 group in which the IL-18BP concentration at

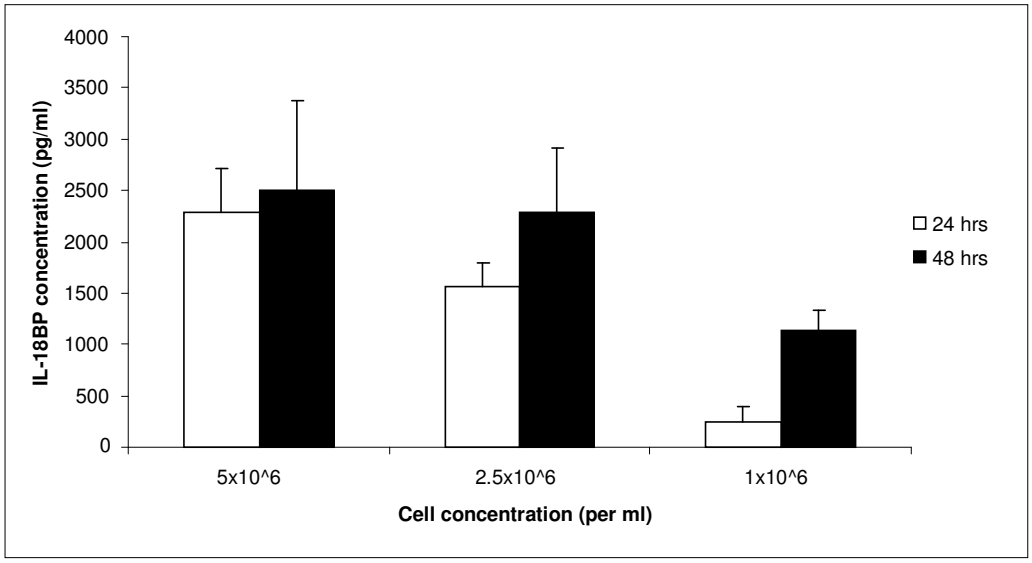


Figure 4.10 IL-18BP expression levels from stimulated PBMCs at different cell concentrations

PBMCs from healthy individuals were stimulated at three different cell concentrations with 20ng/ml IFNγ over 48 hours. The level of IL-18BP was measured by ELISA. The levels at 0, 3, and 6 hours were below the detection limit of the assay and so are not shown. Results are shown as the mean and ±standard deviation of the three individuals in pg/ml.

Haplotype	Individual	0 hrs (pg/ml)	3 hrs (pg/ml)	6 hrs (pg/ml)	24 hrs (pg/ml)	48 hrs (pg/ml)
1-1	WC14	n/a	n/a	170.09	3384.84	4196.49
	WC45	n/a	n/a	n/a	958.73	1907.34
	WC48	n/a	n/a	n/a	1059.83	2536.84
	Mean	n/a	n/a	n/a	**1801.13**	**2880.22**
	SD	n/a	n/a	n/a	**1372.46**	**1182.57**
1-1/1-2	WC01	n/a	6.76	n/a	1324.75	1922.26
	WC03	n/a	n/a	n/a	1051.90	6244.72
	WC42	n/a	n/a	n/a	1228.49	2182.79
	Mean	n/a	n/a	n/a	**1201.71**	**3449.92**
	SD	n/a	n/a	n/a	**138.38**	**2423.87**
1-1/2-1	WC07	n/a	n/a	n/a	1731.42	4423.41
	WC31	n/a	n/a	n/a	163.26	441.41
	Mean	n/a	n/a	n/a	**947.34**	**2432.41**
	SD	n/a	n/a	n/a	**1108.86**	**2815.70**
2-1	WC15	n/a	n/a	27.09	2840.77	4340.24
	WC11	n/a	n/a	n/a	547.41	1979.00
	Mean	n/a	n/a	n/a	**1694.09**	**3159.62**
	SD	n/a	n/a	n/a	**1621.65**	**1669.64**
1-2/2-1	WC20	n/a	n/a	12.65	323.15	1277.30

Table 4.14 IL-18BP protein expression levels from IFNγ stimulated PBMCs

The levels of IL-18BP in pg/ml measured by ELISA from PBMCs stimulated with 20ng/ml IFNγ for each of the individuals are shown. For the majority of the samples the levels at 0, 3, and 6 hours were below the detection limit of the assay and are shown as n/a. The mean and standard deviation at both 24 hours and 48 hours stimulation are shown for each of the haplotype groups for which a minimum of two samples were available.

both time points is ten fold higher in one individual than the other, 1731pg/ml and 163pg/ml at 24 hours, and 4423pg/ml and 441pg/ml at 48 hours. These concentrations result in the SD between the two individuals being greater than the mean concentration, 947.34±1108.86 and 2432.41±2815.70 respectively. Compared to the mean concentration of the other haplotype groups, individual WC20 carrying the two heterozygous haplotypes 1-2 and 2-1, showed lower protein expression levels at both time points (Figure 4.11).

One-way ANOVA analysis showed that haplotype did not show a significant effect on expression levels, p=0.82 at 24 hours, and p= 0.95 at 48 hours. Pairwise multiple comparisons of all haplotype groups also showed no significant difference between any two haplotype groups, $p \geq 0.05$, at either of the time points. Stimulation time was a significant factor in the levels of IL-18BP, p=0.0254.

4.2.3.3. RNA expression levels

4.2.3.3.1. Validation of house keeping genes

To ensure that correction between samples using the comparative C_T method was only for variation between cDNA levels, it was important that the expression of the housekeeping gene used does not vary over the culture conditions of the experiment. To investigate this, five different housekeeping genes were tested. Melting curves of all five gene assays were performed to test the specificity of the assays. As shown in Figure 4.12, one distinct peak was seen for each assay. This reflects the presence of one amplicon species and demonstrates that each assay only detects the RNA of interest. The two no template control (NTC) samples (light blue lines) showed no variation in flourescence intensity as there were no amplicons to detect. Expression of *GAPDH* and *ACTB* appeared to increase over time, signified by decreasing C_T values (Figure 4.13 A.), while *RPLP0*, *RPL32*, and *HPRT1* did not seem to vary. This was also apparent when the range of C_T values and the SD between time points was examined. *GAPDH* and *ACTB* showed differences between the highest and lowest values measured of 5.13 and 4.67 C_T cycles respectively.

These two genes also showed the highest SD between all time points of 1.6 and 1.67 C_T cycles respectively. *RPLP0*, *RPL32*, and *HPTR1*, all showed much lower variation over time with C_T ranges of 1.42, 0.67, and 1.85 respectively. This was also reflected in the SD between the C_T values for each gene, 0.48, 0.26, and 0.59 respectively. *RPL32* showed the

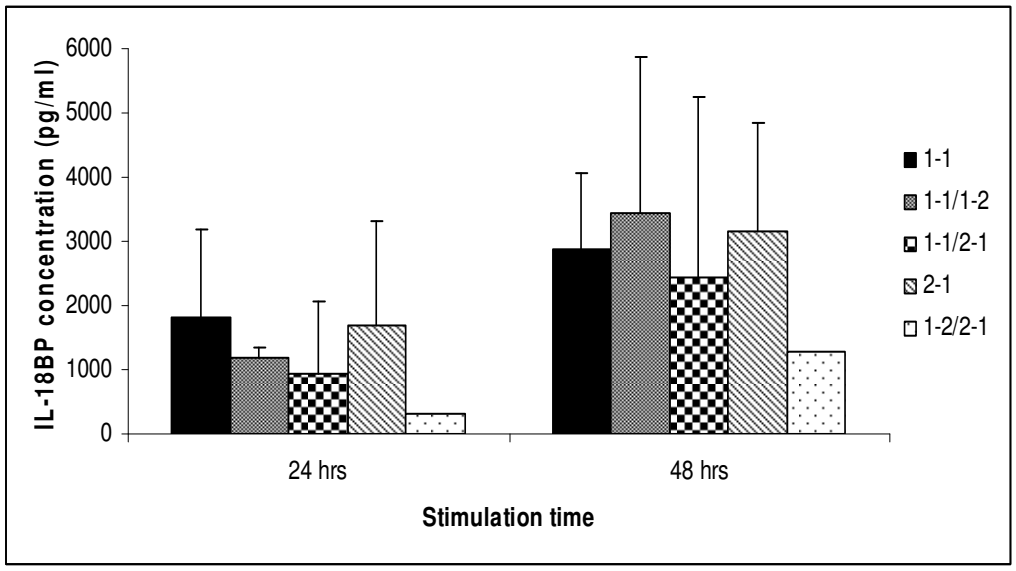


Figure 4.11 Stimulated PBMC IL-18BP protein expression levels according to haplotype

PBMCs from individuals of known haplotype of the associated SNP1/SNP2 haplotype were stimulated with 20ng/ml IFNγ for either 24 or 48 hours and the level of IL-18BP in the media measured by ELISA. The mean concentration ± SD of IL-18BP in pg/ml at both time points of all of the individuals in each haplotype group is shown. Where only one haplotype is indicated individuals were homozygous for that haplotype.

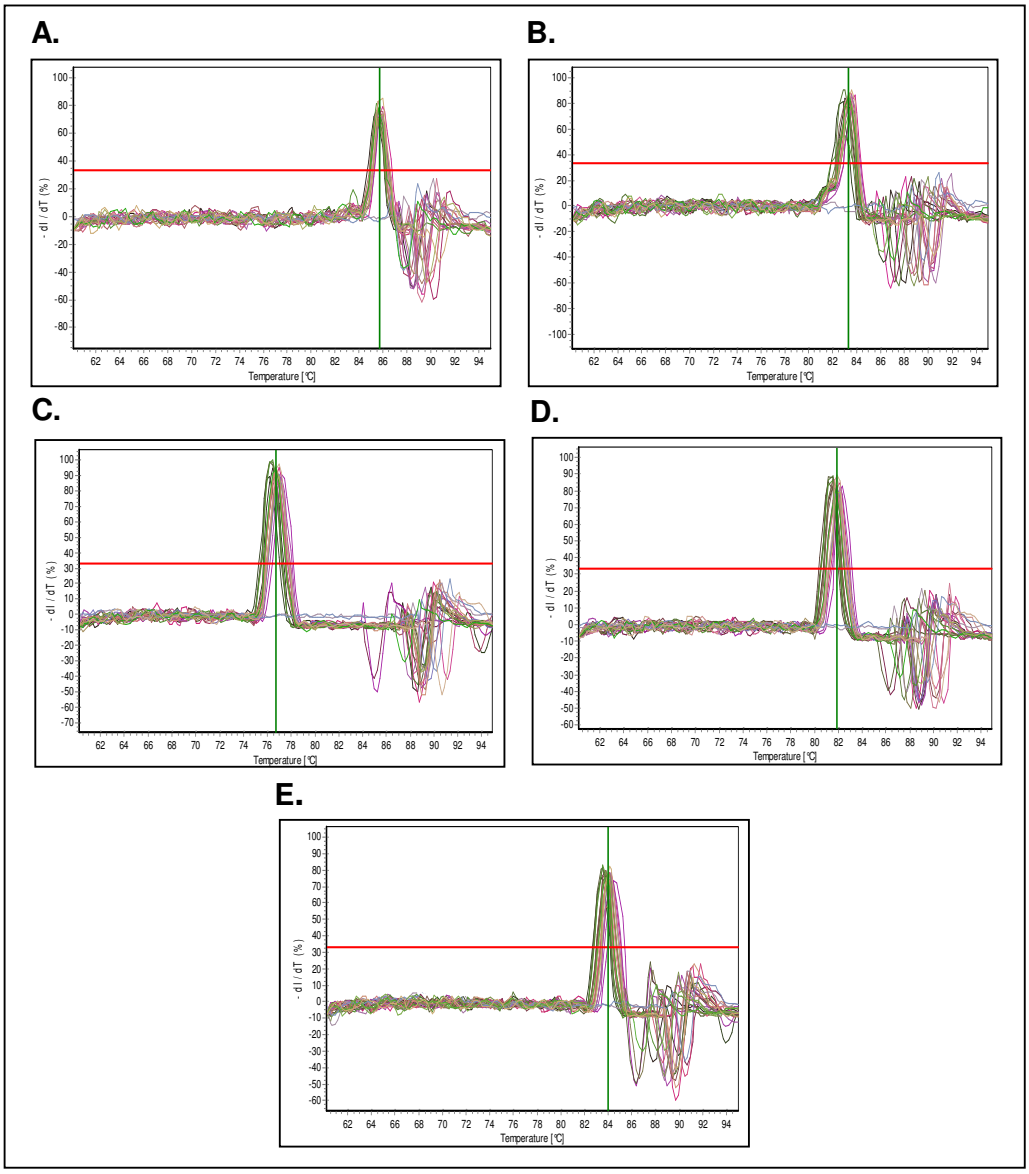


Figure 4.12 Melting curves of housekeeping qRT-PCR assay primers

Following a qRT-PCR reaction a melting curve is performed during which the SYBR green fluorescence is monitored as the temperature is increased. The results are plotted as the negative first derivative of the melting curve as a percentage of the highest peak (-dI/dT (%)) as a function of temperature. Melting curves for the *ACTB* (A.), *GAPDH* (B.), *HPRT1* (C.), *RPL32* (D.), and *RPLP0* (E.) specific primer sets are shown. The default 33% threshold reduction in flourescence is shown by the horizontal red line. The vertical green line indicates the maximum peak corresponding to the melting temperature.

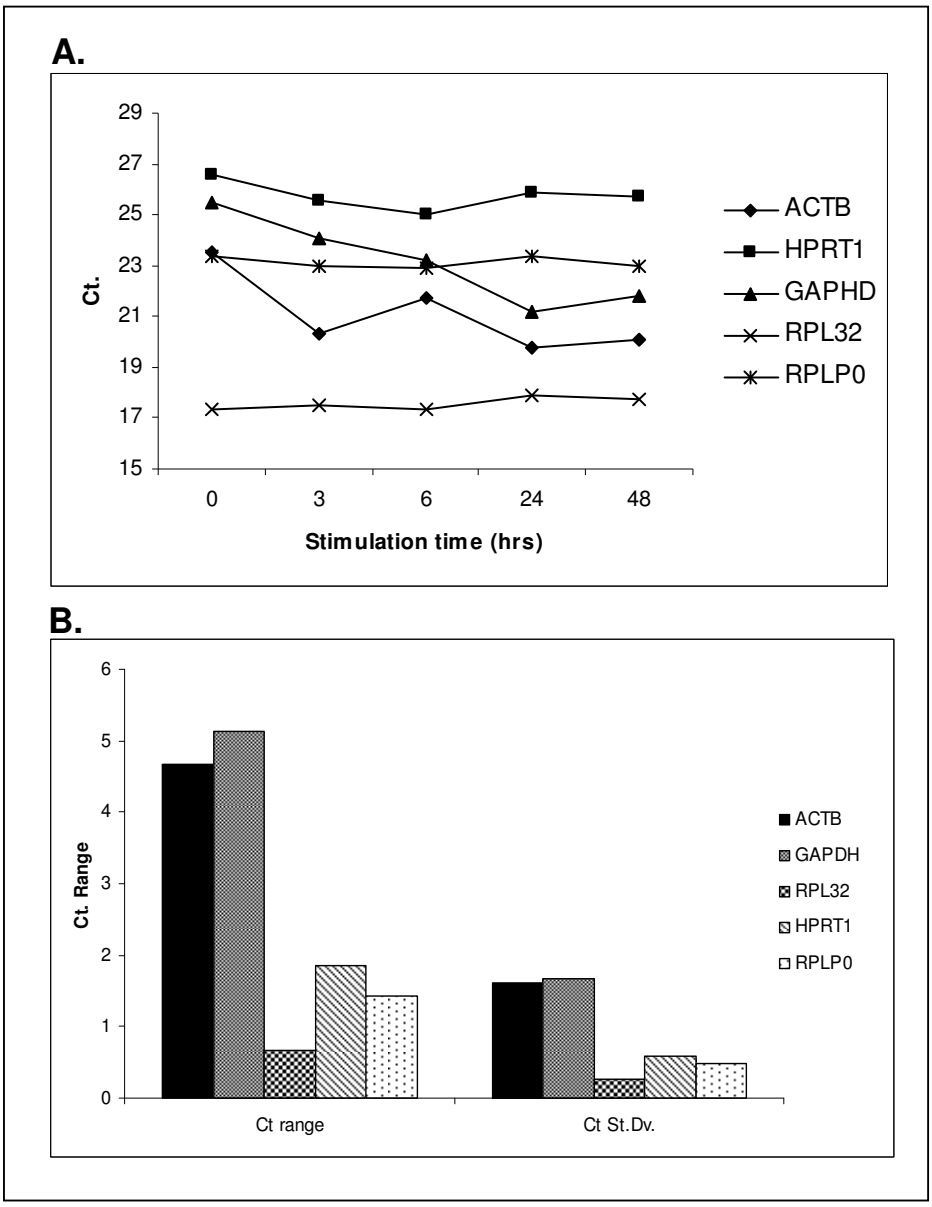


Figure 4.13 Housekeeping gene C_T variation

The RNA levels of five different housekeeping genes were measured in IFNγ stimulated PBMCs to test for variation in expression levels over the 48 hour time series used. Results are shown as the raw C_T values over time (A.). To illustrate the variation of each gene over time the range (C_T high-C_T low) and standard deviation of the raw C_T values measured are also shown.

least variation over the stimulation time course and was therefore selected as the most appropriate housekeeping gene to use under these conditions.

4.2.3.3.2. IL-18BP gene expression comparison according to haplotype

To confirm the production of only one amplicon species during the qRT-PCR reaction with the IL18BP primer assay, a melting curve of the end product was performed. As shown in Figure 4.14 one distinct peak was seen. This reflects the presence of one amplicon species and demonstrates that the assay only detects the RNA of interest. The two no template control (NTC) samples (light blue lines) showed no variation in flourescence intensity as there were no amplicons to detect. They also demonstrated that there was no primer-dimer formation in the assay reaction which would have produced a smaller peak on the graph.

qRT-PCR reactions for each sample were performed in duplicate. The standard deviation between sample replicates were all <1 C_T cycle. The 1-1/1-2 haplotype group showed the highest expression, expressed as $2^{-\Delta CT}$ relative to the housekeeping gene *RPL32*, at both the 6 and 24 hour time points. However, there was also a large SD in the mean expression at both time points, the highest SDs of all the haplotype groups. This is because individual WC42 had a higher expression level than the other two individuals in the haplotype group. At 6 hours WC42 had a relative expression level of 0.039 compared to the other individuals at 0.0081 and 0.016, and at 24 hours 0.010 compared to 0.050 and 0.047 (Table 4.15).

At the 0 and 3 hours all haplotype groups had low levels of expression (Figure 4.15). There was a significant ($p<0.05$) increase in expression after 6 hours stimulation, and again at 24 hours in all haplotype groups except 1-1/2-1 and 2-1, although this was not significant overall. Between 24 and 48 hours the *IL18BP* levels decreased ($p<0.0001$), returning to baseline levels with no significant difference compared to 0 and 3 hours. The individual carrying the 1-2/2-1 haplotype combination appeared to show lower levels of *IL18BP* expression than the other haplotype groups. This individual had the lowest expression relative to the housekeeping gene at 0, 6, and 48 hours, and levels similar to the other haplotype group not carrying the 1-1 haplotype; 2-1 homozygotes. At 3 hours the 1-2/2-1 level was 0.0051 compared to 0.0055 for the 2-1 group, and 0.0096 and 0.090 respectively at 24 hours.

ANOVA analysis showed no significant variation between haplotype groups at any time

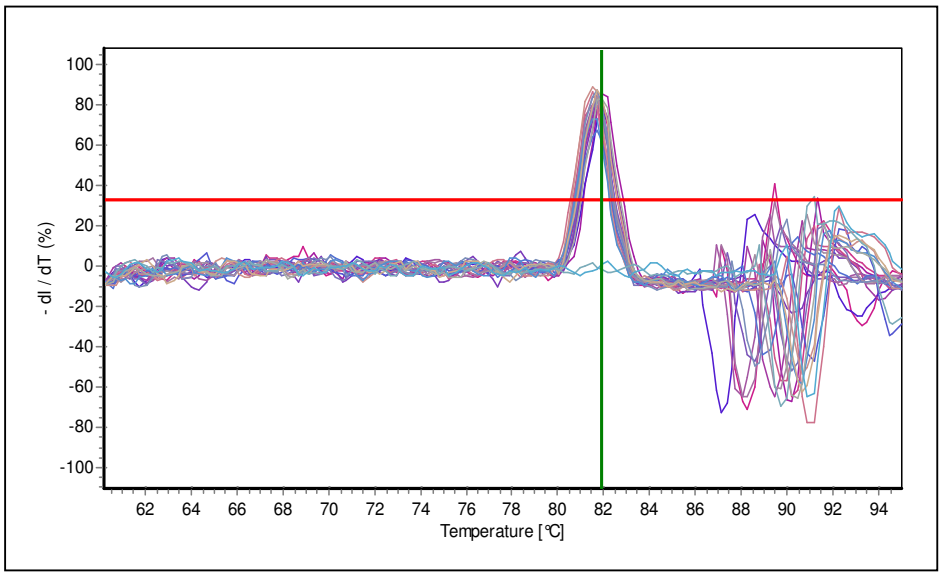


Figure 4.14 Melting curve of IL-18BP qRT-PCR assay primers

Following a qRT-PCR reaction a melting curve is performed during which the SYBR green fluorescence is monitored as the temperature is increased. The results are plotted as the negative first derivative of the melting curve as a percentage of the highest peak (-dI/dT (%)) as a function of temperature. The default threshold reduction in flourescence is 33% and shown by the horizontal red line. The vertical green line indicates the maximum peak corresponding to the melting temperature. The example shown is PBMCs stimulated with IFNγ for 6 hours.

Haplotype	Individual	0hrs	3hrs	6hrs	24hrs	48hrs
1-1	WC14	0.0054	0.0039	0.015	0.0083	0.00056
	WC45	0.0041	0.0043	0.018	0.024	0.0016
	WC48	0.0026	0.0029	0.0074	0.022	0.00042
	Mean	**0.0041**	**0.0037**	**0.014**	**0.018**	**0.00087**
	SD	**0.0014**	**0.00072**	**0.0056**	**0.0085**	**0.00066**
1-1/1-2	WC01	0.0036	0.0026	0.0081	0.050	0.007
	WC03	0.0049	0.0056	0.016	0.047	0.0047
	WC42	0.0085	0.0047	0.039	0.010	0.013
	Mean	**0.0057**	**0.0043**	**0.021**	**0.036**	**0.0081**
	SD	**0.0025**	**0.0015**	**0.016**	**0.022**	**0.0041**
1-1/2-1	WC07	0.0044	0.0025	0.0073	0.013	0.0023
	WC31	0.0044	0.0035	0.012	0.017	0.001
	Mean	**0.0044**	**0.0030**	**0.0097**	**0.015**	**0.0016**
	SD	**0.000033**	**0.00076**	**0.0033**	**0.003**	**0.00087**
2-1	WC15	0.0056	0.0052	0.020	0.01	0.0012
	WC11	0.0068	0.0057	0.010	0.0076	0.010
	Mean	**0.0062**	**0.0055**	**0.015**	**0.0090**	**0.0056**
	SD	**0.00088**	**0.00039**	**0.0072**	**0.0019**	**0.0062**
1-2/2-1	WC20	0.0028	0.0051	0.0063	0.0096	0.00048

Table 4.15 *IL18BP* RNA expression levels from IFNγ stimulated PBMCs

PBMCs from control individuals of known *IL18BP* haplotype were stimulated with 20ng/ml IFNγ. Total RNA from cells, lysed at the indicated time points, was used to perform qRT-PCR. The results are shown as $2^{-\Delta CT}$, relative to the housekeeping gene *RPL32*.

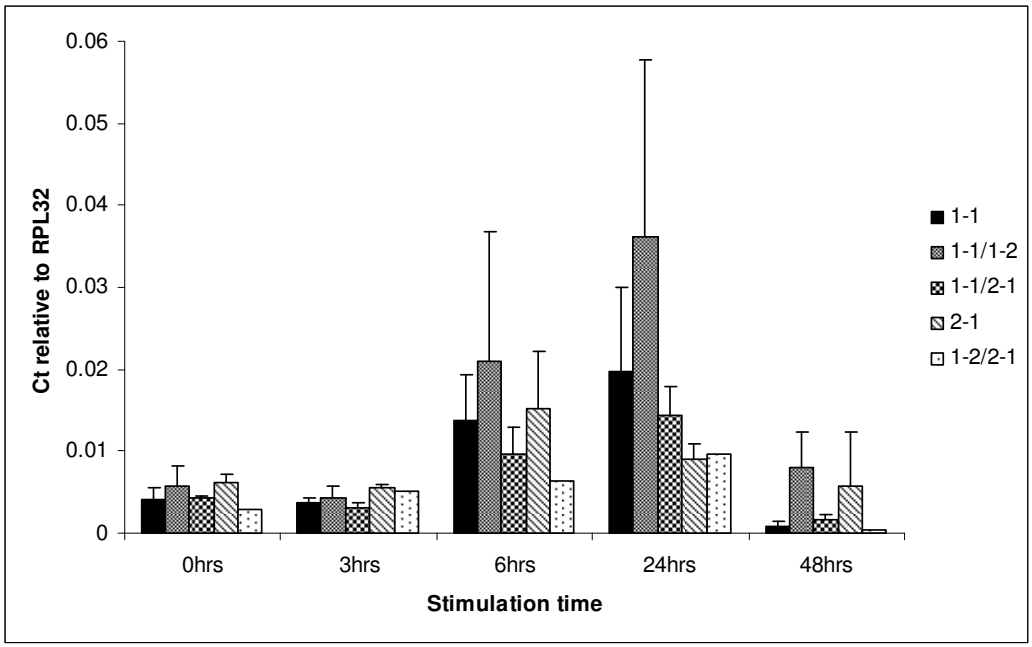


Figure 4.15 Stimulated PBMC *IL18BP* RNA expression levels according to haplotype

PBMCs from healthy individuals of known *IL18BP* haplotype were stimulated with 20ng/ml IFNγ. Total RNA from cells stored in TRIZOL at the stated time points was used to measure *IL18BP* mRNA by qRT-PCR. The results mean results ±SD for each haplotype group by stimulation time point are shown as $2^{-\Delta CT}$, relative to the housekeeping gene *RPL32*.

point: 0hrs p=0.5041, 3hrs p=0.2089, 6hrs p=0.6854, 24hrs p=0.2194, 48hrs p=0.1481. There was also no significant difference between any pair of haplotype groups at any time point.

4.2.3.4. Summary

No significant difference in IFNγ induced PBMC expression of IL-18BP according to haplotype was observed at either the RNA or protein level.

4.2.3.5. Discussion

Expression analysis with promoter constructs showed different expression levels according to the *IL18BP* SNP1/SNP2 haplotype significantly associated with sJIA (Appendix 3). The 1-1 haplotype, which is at a significantly lower frequency in the patient population showed the highest expression level, and the 1-2 haplotype which is significantly more frequent in patients, produced a significantly lower level of expression. This finding is in concordance with the hypothesis that patients with sJIA have a reduced expression of IL-18BP during inflammatory responses, resulting in decreased antagonism of circulating IL-18 increasing the resulting inflammation. This would also lead to a decrease in the negative feed-back loop required for the down-regulation of IL-18 once the infection has been cleared, leading to a maintained inappropriate inflammatory state as seen in sJIA.

The *IL18BP* SNP1/SNP2 haplotype effect on expression levels was not however seen in PBMCs from individuals of different haplotypes, suggesting that the haplotype may not be biologically relevant. Although the haplotype is involved in expression in the promoter constructs, it was investigated in isolation. However, in the environment of the genome this effect may be silenced or overpowered by other regulatory elements. It is also possible that, because transfection assays are an unnatural biological enviroment and that immortal cell lines may not behave as somatic cells would, that the results of the transfection assay are not truly representative. Due to technical difficulties in achieving sufficient transfection efficiency in the monocyte/macrophage THP-1 cell line, the epithelial cell line HeLa was used in the transcription assay instead. If the role of the *IL18BP* haplotype in regulation of expression is cell type specific, this could account for why the results of the transcription assay were not also observed in PBMCs.

It was not possible to obtain samples from healthy individuals carrying each of the four

possible haplotypes. The homozygous state for only the 1-1 (n=3) and 2-1 (n=2) haplotypes could be investigated. The rarer 1-2 (n=3) haplotype could only be investigated in individuals carrying only one copy, in combination with a copy of the 1-1 haplotype. It is possible that the presence of one copy of the 1-1 haplotype compensates for any reduced expression from an alternative haplotype. The necessity of using individuals carrying only one copy of the 1-2 haplotype may therefore have masked any haplotype expression differences. The rarest 2-2 haplotype was not represented at all in the PBMC expression analysis. However, based on the WTCCC population haplotype frequencies the 2-2 haplotype is estimated to be present at a frequency of 0.01, with approximately 1 in 10,000 individuals carrying two copies. With such a low incidence in the population this haplotype may not have a biological impact, although a deleterious effect on immune responses may be responsible for the rarity. Another issue arising from the rarity of some of the haplotypes in the population is that only small numbers of each haplotype group could be included in the study. The results therefore must be viewed with caution due to the possibility of sampling error.

It is also possible that the results have been confounded by measuring the expression levels in PBMCs rather than specific cell types. It is possible that any cell-type specific regulation of expression was not detected as it was diluted by the admixture of the cells studied.

The lack of evidence for an *IL18BP* SNP1/SNP2 haplotype effect on IL-18BP expression levels is, however, in concordance with previously published results. The three *IL18BP* candidate region SNPs selected as part of this project (Table 4.7) were investigated in relation to coronary heart disease by a collaborating group. In coronary artery bypass graft surgery patients no relationship between the IL18BP haplotype, or the SNPs individually, was found with plasma levels of IL-18 or IL-18BP (Thompson et al., 2007).

4.2.4. Transcription Factor Binding

4.2.4.1. Binding Prediction

Using a minimum matrix conservation of 75%, a total of four transcription factors (TF) were predicted to bind within the sequence flanking *IL18BP* SNP1 (Figure 4.16). Of the predicted recognition sequences, three included the SNP locus. The predicted binding of the TF Sp1 was allele-specific, with binding predicted only in the presence of the common C allele at the SNP position. Under increased stringency of the prediction algorithm to minimise false positives, 80% minimum conservation, only binding of ATF and Sp1 were predicted. Sp1 (specificity protein 1) is a promoter-specific TF (Dynan and Tjian, 1983) shown to be involved in transcription activation of a number of genes (Luzina et al., 2006; Fang et al., 2007; Hong et al., 2008; Saha et al., 2008; Okazaki et al., 2010), including of the IL-1 receptor related *SIGIRR* (Kadota et al., 2010). It has also recently been shown that IL-18 induces activation of Sp1 in adult mouse cardiocytes (Reddy et al., 2010).

No TF binding sites encompassing the SNP2 locus were predicted.

4.2.4.2. EMSA

4.2.4.2.1. Short probe EMSAs

Three protein-DNA complexes were created with the 30bp probes (Figure 4.17). Only the smallest of these protein-DNA complexes appeared to be specific to the probe nucleotide sequence as it was the only one cold competed with the unlabelled probes. This sequence-specific binding was seen with both the common C allele (lanes 1 to 9) and the rarer T allele (lanes 10 to 18) present at *IL18BP* SNP1, indicating that the binding to this sequence was not SNP1 allele specific. This finding was confirmed by the cold competition reactions. There was no difference observed when the unlabelled probe included in the reaction contained the C or T allele (lanes 3/4, 6/7, 12/13, and 16/17). The protein binding to the probes is likely to be ubiquitously expressed as the complex was present with nuclear extract from both unstimulated (lanes 2 and 11) and IFNγ stimulated THP-1 cells (lanes 6 and 15). When an antibody to the TF Sp1 was included in the reactions, no supershift of the sequence-specific protein-DNA complex was produced (lanes 5, 9, 14, and 18). This demonstrated that the protein binding to the *IL18BP* SNP1 surrounding sequence was not Sp1. Similar results were seen with two different preparations of nuclear extract.

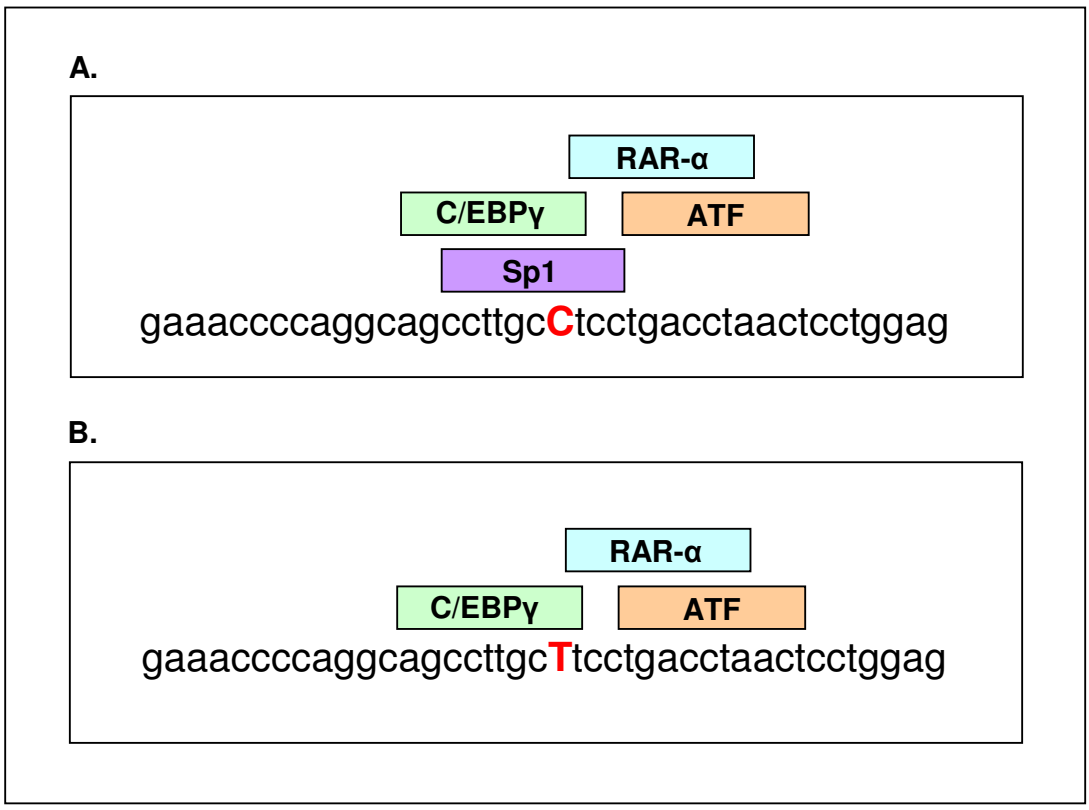


Figure 4.16 *IL18BP* SNP1 transcription factor binding prediction

A number of transcription factors were predicted by Alibaba to bind to the sequence surrounding *IL18BP* SNP1 with a minimum matrix conservation of 75%. Shown are the predicted binding positions of the transcription factors in the presence of the common C allele (A.) and the rarer T allele (B.) at SNP1.

ATF- Activating transcription factor 2, C/EBPgam- CCAAT/enhancer binding protein gamma, RAR-α- Retinoic acid receptor alpha, Sp1- Specificity protein 1.

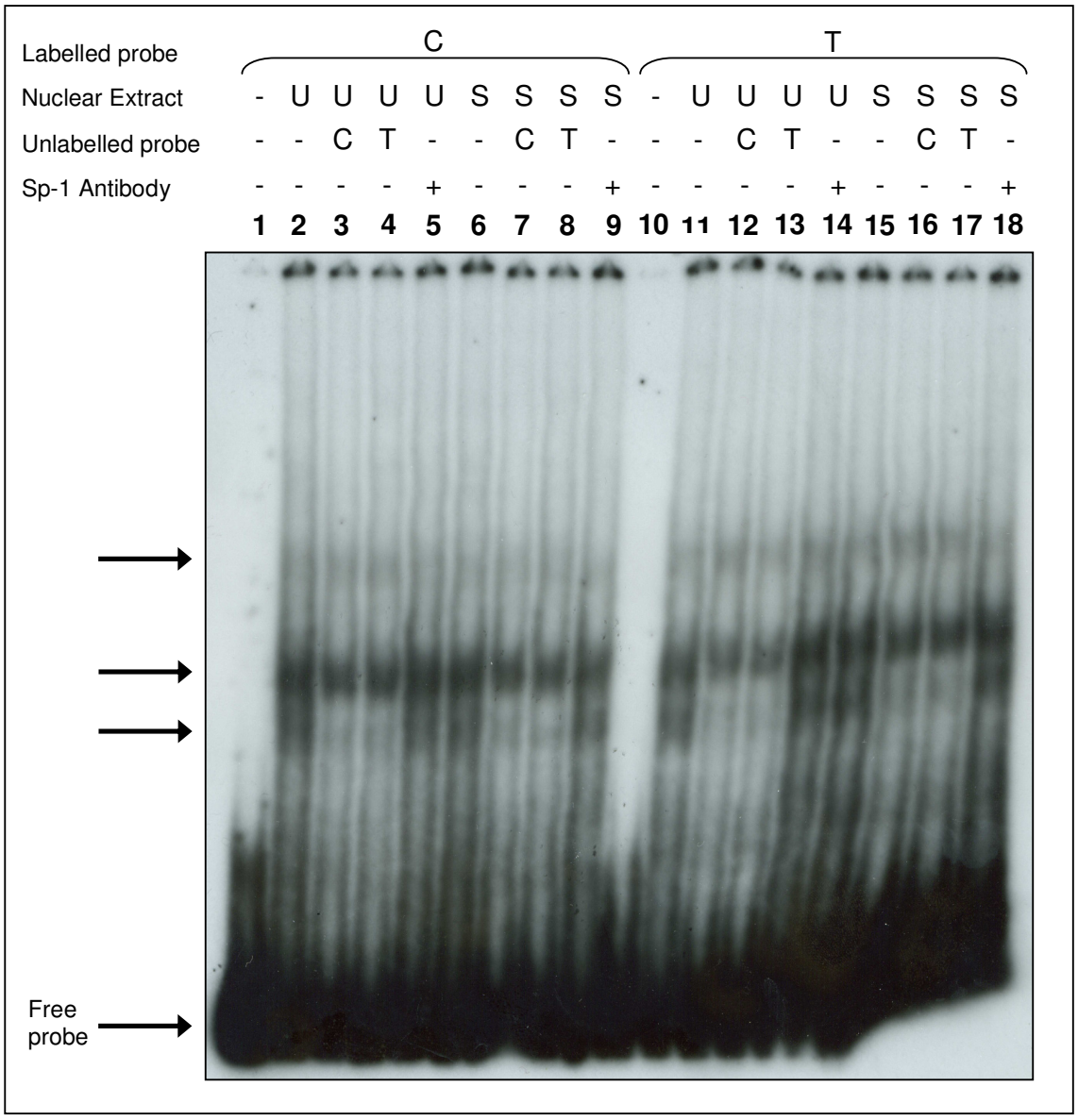


Figure 4.17 *IL18BP* short probe EMSA

EMSA probes spanning 30bp over the *IL18BP* SNP1 locus and the predicted SP1 binding site were radioactively labelled. Probes were incubated with nuclear extract from THP-1 cells, either unstimulated (U) or stimulated (S) with 20ng/ml IFNγ for three hours. Cold competition reactions were performed with 200-fold excess unlabelled probes, and supershift reactions with 2µg SP1 antibody. The allele present at *IL18BP* SNP1 in the labelled probe, and any other factors included in each binding reaction is indicated above each numbered lane of the gel. The unbound probe, and each protein-DNA complex is indicated with an arrow on the left hand side of the figure.

4.2.4.2.2. Long probe EMSAs

A total of eight DNA-protein complexes were visible with both probes, regardless of the allele present at the *IL18BP* SNP1 locus (Figure 4.18). Although there was a reduction in the background radiation present in the lanes none of the protein-DNA complexes disappeared with increasing poly dI-dC, indicating that all of the complexes created were specific to the sequence present in the probes, and that therefore the lower concentrations of poly dI-dC were sufficient to inhibit non-specific protein binding in these reactions.

Eight distinct protein-DNA complexes were created with the 203bp probes (Figure 4.19). All of the complexes appear to be specific to the probe nucleotide sequence, as unlabelled probe competed with them all, the corresponding band was either no longer visible or considerably fainter with the addition of unlabelled probe. All eight protein-DNA complexes were seen with both the common C allele (lanes 1 to 7) and the rarer T allele (lanes 8 to 14) present at *IL18BP* SNP1, indicating that none of the protein binding to this sequence was SNP1 allele specific. This finding was confirmed by the cold competition reactions. There was no difference observed when the unlabelled 203bp probe included in the reaction contained the C or T allele (lanes 4/5, and 11/12). All of the proteins binding to the probes are likely to be ubiquitously expressed as all seven complexes were present with nuclear extract from both unstimulated (lanes 2 and 9) and IFNγ stimulated THP-1 cells (lanes 3 and 10). There was also no evidence of cold competition with any of the complexes when the unlabelled probes included were the 30bp sequences including SNP1, regardless of the allele present at the SNP locus (lanes 6/7 and 13/14).

4.2.4.3. Summary

Binding of the TF Sp1 was predicted to occur in the presence of the common C allele at *IL18BP* SNP1 (rs3814721), but not in the presence of the rarer T allele. Through EMSA experiments it was demonstrated that a ubiquitously expressed nuclear protein bound to the SNP1 position in a sequence-specific manner. This binding however, was not allele-specific, and the binding protein was not determined to be Sp1.

4.2.4.4. Discussion

Transient transfection luciferase assays performed by Dr Wen showed that the disease

associated *IL18BP* two marker haplotype affected transcription levels *in vitro* (Appendix 3). These results suggest that these SNPs play a direct role in regulating *IL18BP* expression levels. Allele-specific TF binding to one of the SNPs could account for the different levels of expression seen.

It was predicted *in silico* that Sp1 would bind to the genomic sequence only in the presence of the common C allele at *IL18BP* SNP1. The C allele at this SNP is part of the 1-1 haplotype, identified in the association study as protective against disease susceptibility, which gave the highest transcription level *in vitro* (Appendix 3). If Sp1 binding at this promoter position is involved in induction of transcription it could account for the higher expression levels seen with the 1-1 haplotype. Higher expression of IL-18BP would result in increased antagonism and inhibition of the IL-18 inflammatory response. In the converse situation of the 2-2 haplotype, SP1 would be unable to bind and induce expression, resulting in the lower transcription levels seen, and therefore a maintained inflammatory state.

EMSA experiments with the 30bp probes showed that a nuclear protein, present in the cell under both stimulated and unstimulated conditions, binds to the DNA sequence surrounding the SNP locus. However, formation of DNA-protein complexes was evident regardless of the allele at the SNP position, confirmed with the longer 203bp probes. Furthermore migration of the complex was unaffected by inclusion of Sp1 antibodies. This demonstrates that *in silico* sequence based predictions, including under high stringency conditions, are fallible and should be used only to inform subsequent experimental investigations.

The protein binding to the DNA sequence immediately flanking the SNP1 locus may be any of the three other TFs also predicted under high stringency to bind to a region included in the probe sequence (Figure 4.16), regardless of the allele present at the SNP position. This is consistent with the findings of the EMSA experiments, but would require direct investigation for confirmation. As no allele-specific binding was observed, the protein binding to this locus cannot be responsible for the haplotype dependent transcription levels seen, and was therefore not investigated further as part of this project.

Despite demonstration that a nuclear protein does bind specifically to the SNP locus, the 30bp probes did not compete for binding with any of the proteins which formed complexes with the larger 203bp probes. This suggests that, although that particular protein does bind to the SNP locus when investigated in isolation, it may preferentially bind, possibly through

higher affinity interaction, to an alternative site in the nearby sequence.

The monocytic THP-1 cell line was used in the protein binding EMSA assays because monocytes are a major source of cytokines during inflammatory responses, and are therefore the most likely to be involved in sJIA pathogenesis. However, due to low transfection efficiency of THP-1 cells, the epithelial HeLa cell line was used in the transcription assay (Appendix 3). It is possible that the *IL18BP* haplotype effect on expression, and therefore transcription factor binding, is cell type specific. This could be determined by performing the EMSA experiments presented in this chapter, using nuclear extract from IFNγ stimulated HeLa cells. Due to time constraints this was not included in this thesis.

The disease association identified in the association study was with the two marker haplotype, while the protein binding experiment was investigating only SNP1. Although the two SNPs of this associated haplotype are over 1kb away from each other it is possible that *in vivo*, due to secondary structure, the two SNPs are brought into proximity enabling them to interact. Due to assay limitations it is not possible to perform EMSA experiments on the DNA sequence containing both SNP loci. It would be possible to investigate protein binding to both SNP loci simultaneously, through use of the haploChIP method in which protein-DNA complexes in living cells are crosslinked and the nuclear material immunoprecipitated (Knight, 2005). However, this technique can only be used to identify DNA sequences bound by a specific candidate transcription factor, limiting the scope of any studies. Although no TF binding was predicted at SNP2 it would be informative to investigate this experimentally to determine if allele-specific protein binding occurs at this locus.

Although individually neither of the haplotype SNPs showed evidence of significant association in the original analysis (sections 4.2.1.2. and 4.2.1.3), when analysed using the larger WTCCC control cohort, SNP1 did show a significant (p=0.0455) frequency difference individually, and SNP2 was tending towards significance (p=0.0524). Also, in the transcription assay the constructs containing the 1-2 and the 2-1 haplotypes both showed significantly lower transcription levels than the 1-1 haplotype, from which they differ at only one of the loci (Appendix 3). It would be very interesting to repeat the transcription assay investigating each SNP in isolation rather than as the haplotype. This would demonstrate if these two SNPs each play an independent role in expression regulation, or if it is a haplotype-only effect, for which the presence of both loci is required. This information would then inform any subsequent investigation into the role of this haplotype in susceptibility to sJIA.

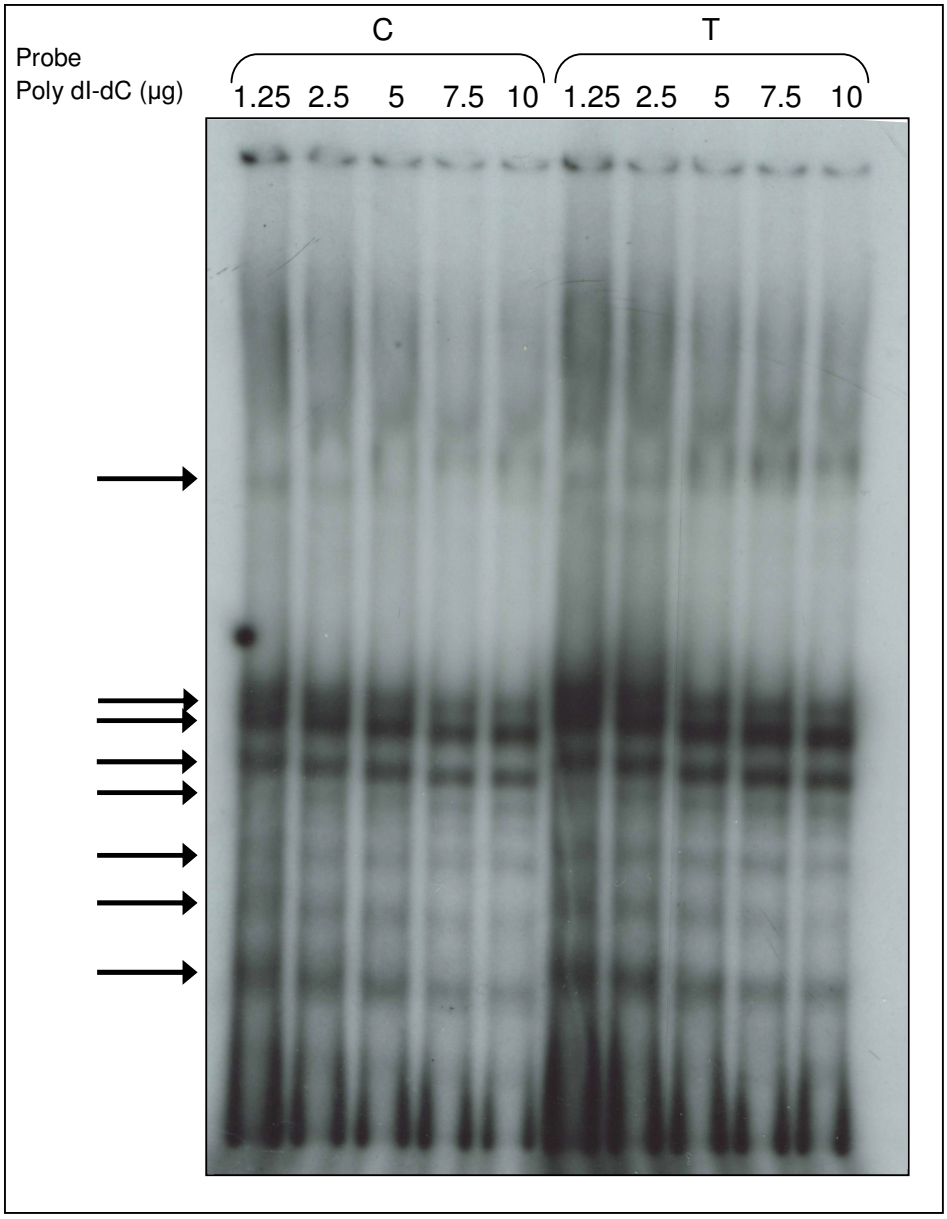


Figure 4.18 *IL18BP* long probe poly dI-dC titration EMSA

To ensure all DNA-protein bands visible following binding of nuclear extract proteins to the long 203bp probes were due to sequence-specific protein binding, a titration with the non-specific competitor DNA poly dI-dC was performed. EMSA probes were radioactively labelled and incubated with nuclear extract from unstimulated THP-1 cells. The allele at SNP1 present in the probe, and poly dI-dC concentration, used in the reaction is indicated above each lane. Each protein-DNA complex is indicated with an arrow on the left hand side of the figure.

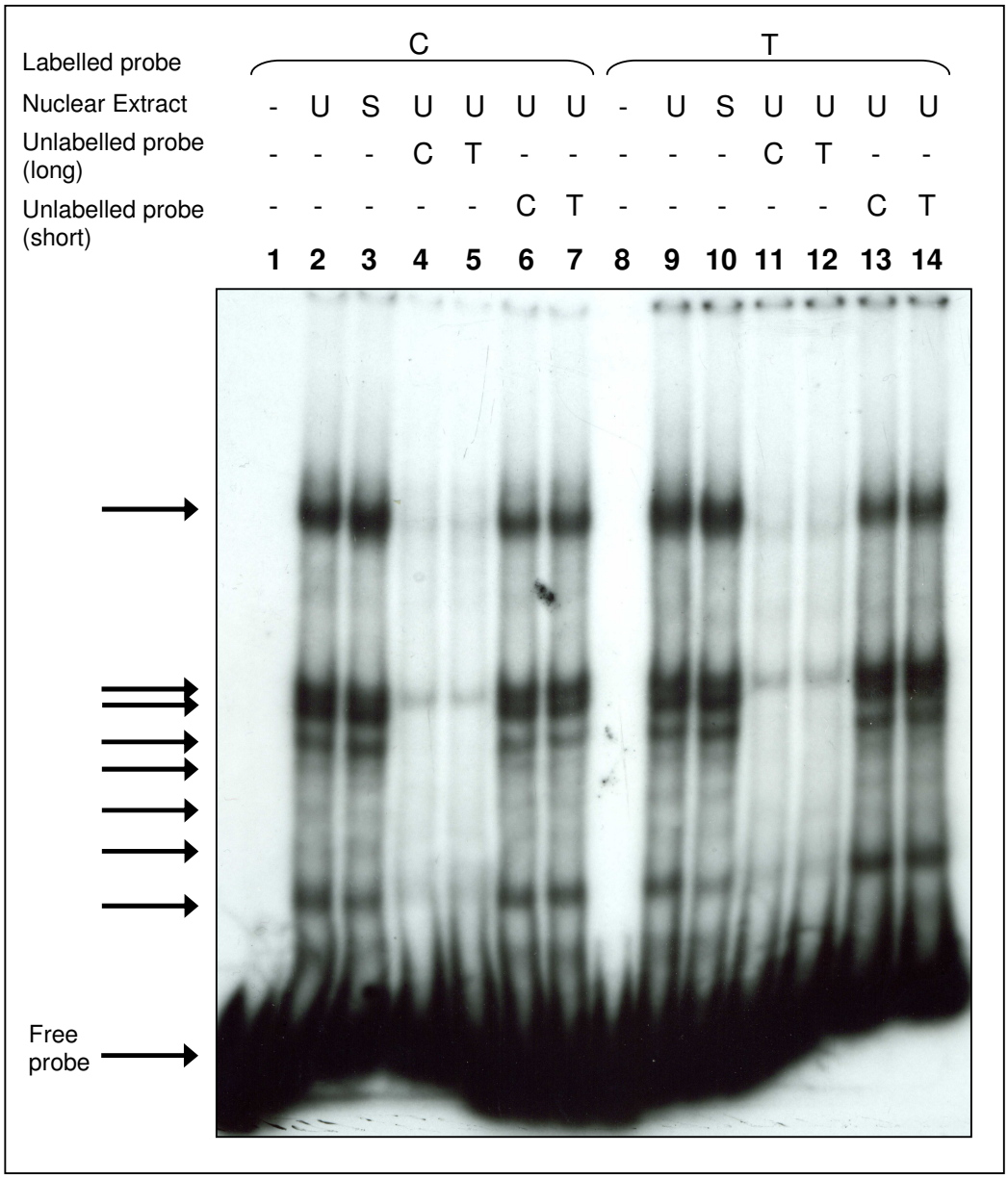


Figure 4.19 *IL18BP* long probe EMSA

EMSA probes spanning 203bp over the *IL18BP* SNP1 locus were radioactively labelled. Probes were incubated with nuclear extract from THP-1 cells, either unstimulated (U) or stimulated (S) with 20ng/ml IFNγ for three hours. Cold competition reactions were performed with 50-fold excess unlabelled long probes, and 100-fold excess short probes. The allele present at *IL18BP* SNP1 in the labelled probe, and any other factors included in each binding reaction is indicated above each numbered lane of the gel. The unbound probe and each protein-DNA complex is indicated with an arrow on the left hand side of the figure.

4.3. *IL1RAP* candidate region

4.3.1. Association Study

To determine the effect of variation within nearby regulatory elements of the *IL1RAP* candidate gene, the region investigated was extended beyond *IL1RAP* to the next flanking genes. These were identified using the July 2003 human genome reference sequence (NCBI build 34), hg16 annotation track. Based on the positions of these genes the candidate region was defined as chromosome 3:191448776-192251235 (Figure 4.20). The genes flanking *IL1RAP* were: 5′, claudin 16 (*CLDN16)*, also called paracellin-1, and 3′, osteocrin (*OSTN*). *CLDN16* is a kidney tight junction protein involved in Mg^{2+} resorption and linked to hypomagnesemia (Simon et al., 1999). *OSTN* is a soluble bone-specific regulator of osteoblast phenotype, and therefore bone metabolism (Thomas et al., 2003). Three mRNA sequences identified in cDNA libraries have also been mapped to within the selected candidate region. The provisional transmembrane protein 207 (AY358607), is upstream of *IL1RAP*. A fetal liver mRNA (AF116662), and a hippocampal mRNA (AK095107) both overlap with the *IL1RAP* gene sequence. AK095107 is weakly similar to *IL1RAP,* and potentially represents an alternative last exon and 3′ UTR of the gene.

The identified candidate region spanned 802kb (802,459bp), and extended 104kb (104,028bp) 5′, and 561kb (561,021bp) 3′, of *IL1RAP*.

4.3.1.1. tSNP selection

Genotyping data was available for a total of 286 SNPs within the selected *IL1RAP* candidate region, all of which were from HapMap (top panel Figure 4.20). Thirty five of the SNPs in the region were monomorphic in the CEU population, and a further 29 had a MAF<0.05. These 64 SNPs were not included in the LD analysis as they did not satisfy the MAF≥0.05 cut-off criteria. A total of 222 SNPs within the *IL1RAP* candidate region satisfied the inclusion criteria and were therefore included in the LD investigation and subsequent tSNP selection.

A graphical representation of the LD between these 222 SNPs in the *IL1RAP*candidate region is shown in Figure 4.21, showing eight visible blocks of high LD.

All of these LD blocks spanned a large genomic region, in genomic order, 35.7kb, 70kb, 68kb, 56kb, 77.4kb, 33.7kb, 60.2kb, and 43.8kb. LD block 1 spanned from intron 2 of both

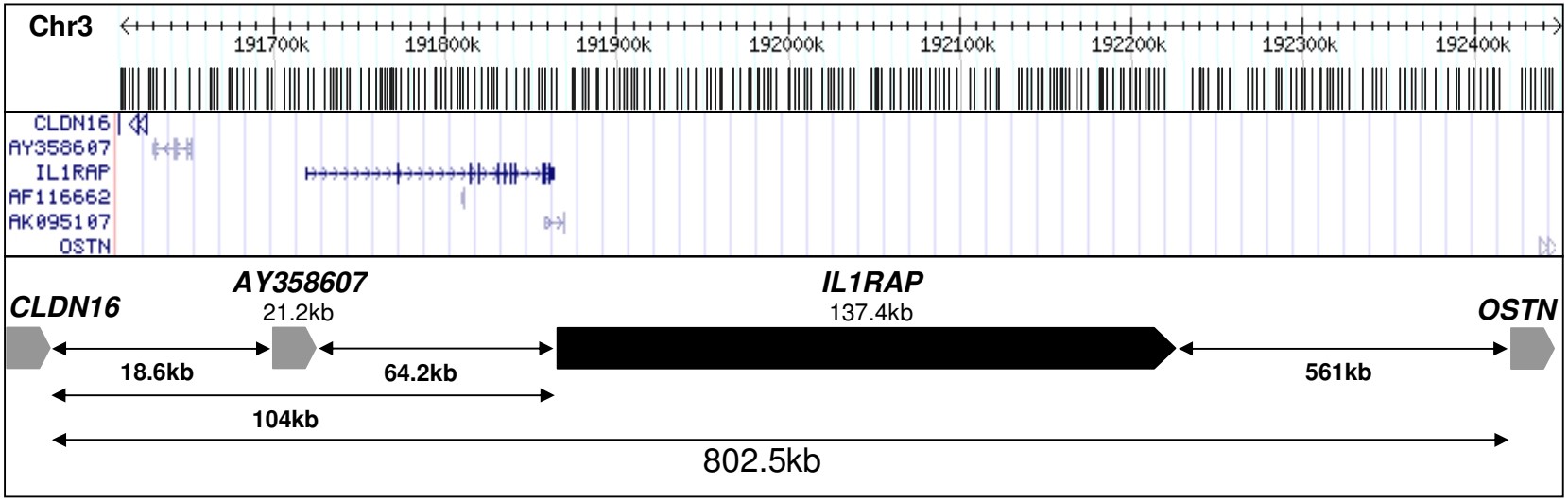


Figure 4.20 *IL1RAP* candidate region

The boundaries of the *IL1RAP* candidate region are the flanking genes *CLDN16* and *OSTN*, determined by the July 2003 human genome reference sequence (NCBI build 34), hg16 annotation track, shown in the centre panel. The 3′ end of *CLDN16* and the 5′ end of *OSTN* are proximal to *IL1RAP*. The approximate (due to scale) positions within this region of all SNPs (irrespective of MAF) in the HapMap phase 2 database (data release #18/phaseII Sept05) are shown in the top panel. The bottom panel shows a schematic (not to scale) of the candidate region with gene orientation, size, and distances given. The gene of interest is shown in black and all other features within the region in grey. The largest *IL1RAP* isoform was used for size calculations, and, for clarity, is the only one shown. All sizes and distances are based on chromosomal positions reported in the NCBI build 34 genome reference sequence.

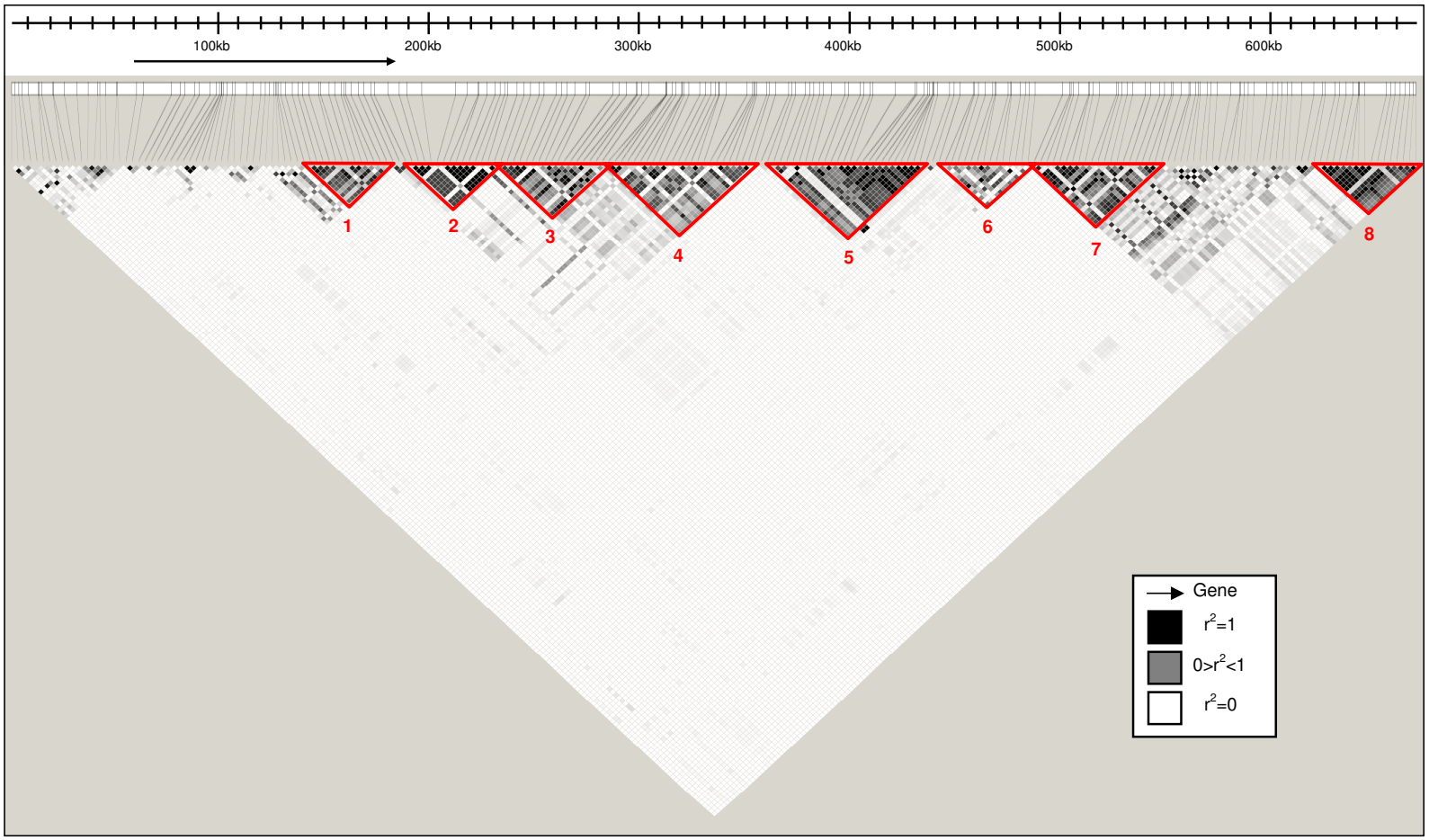


Figure 4.21 LD plot of the SNPs in the *IL1RAP* gene region

The pairwise LD relationships between SNPs in the *IL1RAP* candidate region were investigated to enable tSNP selection. The region shown is comprised of only the SNPs which satisfied the frequency criteria and were included in the LD analysis (n=222), not the whole of the candidate region. The gene orientation and location within the region is indicated. The position within the region of each SNP satisfying the selection criteria is represented by a line underneath the gene position. Each coloured square presents the pairwise LD between two SNPs. The graphical representation, generated in Haploview, is shown on a grey scale, with higher r^2 LD represented by darker blocks. Due to scale the individual SNPs are not labelled. The blocks of high LD are indicated.

IL1RAP isoforms to intron 8 of the larger membrane isoform, 6.2kb downstream of the secreted isoform, LD block 2 began 1.5kb downstream of the membrane bound isoform. All of the other regions of high LD were downstream of the gene.

Due to the LD relationships between the SNPs, a total of 103 tSNPs (46% of the total) were identified as the minimal subset required to capture the variation within all 222 SNPs. However, following assay design (section 3.1.2.1.) it was necessary to exclude three SNPs due to poor assay design scores. One of the excluded SNPs was tagging one other SNP, which was also not suitable for genotyping on the Illumina platform, and therefore also excluded. The final SNP set therefore consisted of 100 tSNPs which captured the variation within 218 SNPs. Fifty nine of the tSNPs were not in high LD ($r^2 \leq 0.8$) with any other SNP within the region and so were not tagging any additional SNPs. All of the selected tSNPs, and the SNPs captured by each, are listed in Table 4.16.

4.3.1.2. Stage-1

The genotyping assays for rs11916734, rs6798479, rs1195370, rs4367119, and rs6796131 were classed as failed (Figure 4.1 and Table 4.2). These SNPs, for which genotypes were not called and were therefore not included in the association analysis, were tagging a total of 12 SNPs. An additional three genotyped SNPs from the *IL1RAP* candidate region showed significant deviation from HWE (Table 4.3), and were therefore also not included in the association analysis. These were: rs2885370 (SNP 21), rs12634559 (SNP 60), and rs6807826 (SNP 77), none of which were tagging any other SNPs. None of the eight excluded SNPs, nor the 12 SNPs tagged by them, were in sufficiently high LD ($r^2 \geq 0.8$) with any of the analysed SNPs. They were therefore not captured by the association study, leaving a total of 202 SNPs which were captured.

Of the 92 analysed *IL1RAP* region tSNPs, 19 showed significant ($p<0.05$) frequency differences between the patient (n=130) and control (n=146) cohorts. For ease of interpretation the results of the SNPs showing significant (Table 4.17) and non-significant frequency differences (Table 4.18) are presented separately.

Of the 19 SNPs which showed evidence of frequency differences between patients and controls, for seven (SNPs 1, 13, 14, 25, 63, 64, and 65) the common allele was present at a higher frequency in the patient than the control population, giving an OR <1 for the rare

SNP Nº	rs Number	Tagged SNPs
1	rs10513852	n/a
2	rs6794669	n/a
3	rs1024935	rs1008800
4	rs3943979	n/a
5	rs10490818	n/a
6	rs9834809	n/a
7	rs7611280	n/a
8	rs11916404	rs1346016,rs9850026
9	rs9859989	n/a
10	rs9864293	rs7642497,rs7612702
11	rs9290933	n/a
12	rs766442	n/a
13	rs13070253	n/a
14	rs3773999	rs9290935
15	rs9290936	n/a
16	rs2059020	n/a
17	rs2361832	n/a
18	rs3773989	n/a
19	rs1988743	n/a
20	rs4686554	n/a
21	rs2885370	n/a
22	rs3821740	rs3773983,rs3773986
23	rs3773982	n/a
24	rs4687150	n/a
25	rs9877268	n/a
26	rs3773977	n/a
27	rs3773976	n/a
28	rs10937439	n/a
29	rs1559018	rs6800655
30	rs6444435	n/a
31	rs10513854	n/a
32	rs3773958	n/a
33	rs3773953	rs10513855
34	rs716984	rs2193874,rs2193875,rs6781037
35	rs9290939	n/a
36	rs2361836	rs1024946,rs1015704,rs4140710,rs4624606,rs4687163
37	rs2885373	n/a
38	rs7626795	n/a
39	rs4320092	rs9847868
40	rs7650510	n/a
41	rs6763652	rs11916852,rs10937444,rs11915384
42	rs6808596	n/a
43	rs6767916	rs2362586,rs10513860,rs6769569,rs6798694,rs9681809, rs6767221, rs6444450
44	rs9845129	n/a
45	rs2885546	n/a
46	rs2362601	n/a
47	rs7618955	rs9878192, rs7616748,rs9873594,rs7617445,rs6444459
48	rs2362605	n/a
49	rs2362607	rs2362595,rs4687170,rs2362615,rs2362593,rs4491961,rs9854184
50	rs10513857	n/a
51	rs6773525	rs2362627,rs2362626
52	rs9832276	rs2173911
53	rs6778424	n/a
54	rs1472597	rs6799514,rs9880171,rs4234618,rs4234619,rs4305462
55	rs10513861	n/a
56	rs9290953	rs12636955,rs7631432, rs9290954

SNP Nº	rs Number	Tagged SNPs
57	rs6444462	rs1393054,rs9828366,rs7340608,rs7640070
58	rs6790983	n/a
59	rs9682599	n/a
60	rs12634559	n/a
61	rs9868955	rs4377528,rs1393061
62	rs9290958	rs9818981,rs9821713
63	rs6804748	rs1501600, rs866986
64	rs6775893	rs9868693,rs9852162,rs12696611,rs9811251,rs952990
65	rs1159213	rs1501603,rs1316356,rs1159211,rs967512,rs1159212,rs13100545, rs9877502
66	rs13060312	n/a
67	rs9857926	n/a
68	rs1501602	n/a
69	rs6444476	n/a
70	rs9877159	n/a
71	rs13314422	rs11721077
72	rs13317447	rs9867007
73	rs7627123	rs1393049
74	rs6774178	n/a
75	rs9814312	rs1510902
76	rs7432178	n/a
77	rs6807826	n/a
78	rs1510901	rs6444485,rs2363133, rs2048417,rs1195369,rs7613308
79	rs6769335	rs13066503,rs9290973,rs9290965,rs7646824,rs7613062
80	rs2063421	n/a
81	rs1355365	rs10513863
82	rs9814786	rs1355366
83	rs7624125	rs9844117
84	rs7611804	n/a
85	rs10937457	n/a
86	rs9855788	n/a
87	rs9854473	n/a
88	rs6788413	rs2363961
89	rs1551424	n/a
90	rs1551425	n/a
91	rs2666356	rs2363970,rs6783220
92	rs11922372	n/a
93	rs7628727	rs1454690,rs7625891,rs7653182,rs7615048,rs6766436,rs12632197, rs6790793,rs7618305,rs9827324
94	rs6444517	n/a
95	rs9290978	rs9878955,rs6810355,rs9843668,rs10513864
96	rs11916734	n/a
97	rs6798479	rs2134115
98	rs1195370	n/a
99	rs4367119	rs1510907,rs6444491,rs1606051,rs6444480,rs7626377, rs6782827, rs9821532,rs10937453,rs9822035,rs2136860,rs9864904
100	rs6796131	n/a
Uncaptured	rs2241344,rs7628250,rs6767500,rs7621864	

Table 4.16 *IL1RAP* candidate region tSNPs

A total of 100 tSNPs were genotyped to capture 218 of the 222 SNPs within the candidate region. The SNPs captured ($r^2 \geq 0.8$) by each genotyped tSNP are shown. Due to poor assay design scores it was not possible to genotype or capture four of the SNPs within the region. For ease of identification each SNP genotyped was assigned a number according to genome order.

SNP	rs Nº	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
1	rs10513852 (G/A)	1 2	0.873 (227) 0.127 (33)	0.801 (234) 0.199 (58)	0.587 (0.369-0.934)	0.0225
2	rs6794669 (G/C)	1 2	0.557 (145) 0.442 (115)	0.658 (192) 0.342 (100)	1.523 (1.08-2.148)	0.0163
3	rs1024935 (A/C)	1 2	0.823 (214) 0.177 (46)	0.887 (259) 0.113 (33)	1.687 (1.041-2.733)	0.0323
13	rs13070253 (C/G)	1 2	0.646 (168) 0.354 (92)	0.548 (160) 0.452 (132)	0.664 (0.471-0.936)	0.0188
14	rs3773999 (A/G)	1 2	0.804 (209) 0.196 (51)	0.730 (213) 0.270 (79)	0.658 (0.441-0.982)	0.0390
25	rs9877268 (G/A)	1 2	0.719 (187) 0.281 (73)	0.615 (176) 0.385 (110)	0.625 (0.436-0.896)	0.0100
61	rs9868955 (A/G)	1 2	0.500 (130) 0.500 (130)	0.596 (174) 0.404 (118)	1.475 (1.052-2.066)	0.0237
63	rs6804748 (A/G)	1 2	0.619 (161) 0.381 (99)	0.527 (154) 0.473 (138)	0.686 (0.489-0.964)	0.0294
64	rs6775893 (G/A)	1 2	0.581 (151) 0.419 (109)	0.483 (141) 0.517 (109)	0.674 (0.481-0.944)	0.0213
65	rs1159213 (G/A)	1 2	0.689 (179) 0.311 (81)	0.596 (174) 0.404 (118)	0.6673 (0.469-0.948)	0.0235
68	rs1501602 (A/G)	1 2	0.815 (212) 0.185 (48)	0.914 (267) 0.086 (25)	2.418 (1.443-4.051)	0.000578
69	rs6444476 (G/C)	1 2	0.889 (231) 0.111 (29)	0.942 (275) 0.058 (17)	2.031 (1.088-3.789)	0.0234
70	rs9877159 (G/A)	1 2	0.850 (221) 0.150 (39)	0.914 (267) 0.086 (25)	1.885 (1.106-3.211)	0.0182
71	rs13314422 (C/A)	1 2	0.919 (237) 0.084 (21)	0.962 (281) 0.038 (11)	2.264 (1.07-4.79)	0.0282
72	rs13317447 (A/G)	1 2	0.927 (241) 0.073 (19)	0.973 (284) 0.027 (8)	2.799 (1.204-6.507)	0.0122
73	rs7627123 (G/A)	1 2	0.742 (193) 0.258 (67)	0.863 (252) 0.137 (40)	2.187 (1.417-3.376)	0.00033
74	rs6774178 (A/G)	1 2	0.862 (224) 0.138 (36)	0.921 (269) 0.078 (23)	1.880 (1.082-3.266)	0.0233
75	rs9814312 (G/A)	1 2	0.908 (236) 0.092 (24)	0.952 (278) 0.048 (14)	2.019 (1.021-3.992)	0.0394
85	rs10937457 (A/G)	1 2	0.650 (169) 0.350 (91)	0.733 (214) 0.267 (78)	1.477 (1.027-2.125)	0.0350

Table 4.17 Significant *IL1RAP* candidate region stage-1 analysis results

Ninety two genotyped tSNPs within the *IL1RAP* candidate region were analysed for significant allele frequency differences between the patient (n=130) and control (n=146) cohorts. The frequency in each cohort of both the common (1) and rare (2) alleles, are given. The alternative alleles of each SNP are shown underneath the rs identifier. The allele assigned as the common (1) allele is given first and the rarer (2) allele second. The OR shown is for allele 2 (allele 1 OR=1). The p-value for the test for frequency differences between the cohorts is also shown. The results for the 19 tSNPs which showed significant ($p<0.05$) frequency differences between the two cohorts are shown.

SNP Nº	rs Number	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
4	rs3943979 (G/A)	1 2	0.542 (141) 0.458 (119)	0.589 (172) 0.411 (120)	1.210 (0.863-1.695)	0.269
5	rs10490818 (A/C)	1 2	0.781 (203) 0.219 (57)	0.777 (227) 0.223 (65)	0.981 (0.655-1.467)	0.924
6	rs9834809 (T/A)	1 2	0.818 (211) 0.182 (47)	0.815 (238) 0.185 (54)	0.982 (0.637-1.513)	0.934
7	rs7611280 (A/G)	1 2	0.783 (202) 0.217 (56)	0.792 (225) 0.208 (59)	1.057 (0.700-1.596)	0.791
8	rs11916404 (G/A)	1 2	0.746 (194) 0.254 (66)	0.791 (231) 0.209 (61)	1.288 (0.866-1.916)	0.211
9	rs9859989 (A/G)	1 2	0.704 (183) 0.296 (77)	0.699 (204) 0.301 (88)	0.976 (0.677-1.405)	0.894
10	rs9864293 (A/G)	1 2	0.611 (159) 0.389 (101)	0.641 (186) 0.359 (104)	1.136 (0.804-1.606)	0.470
11	rs9290933 (A/G)	1 2	0.496 (129) 0.504 (131)	0.554 (161) 0.449 (131)	1.248 (0.893-1.745)	0.195
12	rs766442 (A/C)	1 2	0.711 (185) 0.289 (75)	0.726 (212) 0.274 (80)	1.074 (0.741-1.558)	0.705
15	rs9290936 (C/A)	1 2	0.889 (231) 0.111 (29)	0.863 (252) 0.137 (40)	0.818 (0.475-1.317)	0.366
16	rs2059020 (T/A)	1 2	0.819 (213) 0.180 (47)	0.808 (236) 0.192 (56)	0.930 (0.605-1.429)	0.740
17	rs2361832 (A/G)	1 2	0.764 (197) 0.236 (61)	0.774 (226) 0.226 (66)	1.060 (0.713-1.577)	0.773
18	rs3773989 (G/A)	1 2	0.535 (139) 0.465 (121)	0.524 (153) 0.476 (139)	0.958 (0.686-1.339)	0.803
19	rs1988743 (G/C)	1 2	0.731(190) 0.269 (70)	0.764 (223) 0.236 (69)	1.191 (0.810-1.749)	0.374
20	rs4686554 (A/C)	1 2	0.723 (188) 0.277 (72)	0.699 (204) 0.301 (88)	0.888 (0.634-1.284)	0.527
22	rs3821740 (G/A)	1 2	0.927 (241) 0.0738 (19)	0.914 (267) 0.0856 (25)	0.842 (0.452-1.567)	0.587
23	rs3773982 (G/A)	1 2	0.858 (223) 0.142 (37)	0.822 (240) 0.178 (52)	0.766 (0.484-1.212)	0.253
24	rs4687150 (G/A)	1 2	0.612 (159) 0.388 (101)	0.628 (182) 0.372 (108)	1.070 (0.758-1.511)	0.699
26	rs3773977 (A/G)	1 2	0.727 (189) 0.273 (71)	0.791 (231) 0.209 (61)	1.423 (0.961-2.106)	0.0779
27	rs3773976 (A/C)	1 2	0.842 (219) 0.158 (41)	0.843 (246) 0.157 (46)	1.001 (0.633-1.584)	0.996
28	rs10937439 (A/G)	1 2	0.638 (166) 0.362 (94)	0.569 (166) 0.431 (126)	0.746 (0.529-1.051)	0.0934
29	rs1559018 (A/G)	1 2	0.704 (183) 0.296 (77)	0.733 (214) 0.267 (78)	1.154 (0.796-1.674)	0.449
30	rs6444435 (A/G)	1 2	0.650 (169) 0.350 (91)	0.716 (209) 0.284 (83)	1.356 (0.946-1.943)	0.0971
31	rs10513854 (G/A)	1 2	0.842 (219) 0.158 (41)	0.815 (238) 0.185 (54)	0.825 (0.529-1.288)	0.397

SNP Nº	rs Number	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
32	rs3773958 (A/C)	1 2	0.838 (218) 0.162 (42)	0.822 (240) 0.172 (52)	0.889 (0.569-1.389)	0.605
33	rs3773953 (C/A)	1 2	0.892 (232) 0.108 (28)	0.894 (261) 0.106 (31)	1.016 (0.592-1.745)	0.954
34	rs716984 (A/G)	1 2	0.696 (181) 0.304 (79)	0.651 (190) 0.349 (102)	0.813 (0.569-1.162)	0.256
35	rs9290939 (G/A)	1 2	0.931 (242) 0.0692 (18)	0.935 (273) 0.0650 (19)	1.069 (0.548-2.083)	0.845
36	rs2361836 (G/A)	1 2	0.750 (195) 0.250 (65)	0.730 (213) 0.270 (79)	0.899 (0.614-1.316)	0.583
37	rs2885373 (G/A)	1 2	0.877 (228) 0.123 (32)	0.866 (253) 0.134 (39)	0.911 (0.552-1.502)	0.713
38	rs7626795 (A/G)	1 2	0.873 (227) 0.127 (33)	0.863 (252) 0.137 (40)	0.916 (0.559-1.502)	0.727
39	rs4320092 (A/C)	1 2	0.742 (193) 0.258 (67)	0.753 (220) 0.247 (72)	1.061 (0.722-1.559)	0.764
40	rs7650510 (A/G)	1 2	0.765 (199) 0.235 (61)	0.767 (224) 0.233 (68)	1.010 (0.680-1.499)	0.967
41	rs6763652 (C/A)	1 2	0.942 (245) 0.058 (15)	0.955 (279) 0.0446 (13)	1.314 (0.613-2.816)	0.482
42	rs6808596 (A/G)	1 2	0.496 (129) 0.504 (131)	0.544 (159) 0.456 (133)	1.214 (0.869-1.697)	0.256
43	rs6767916 (G/A)	1 2	0.939 (244) 0.0615 (16)	0.955 (279) 0.0446 (13)	1.407 (0.663-2.984)	0.372
44	rs9845129 (G/C)	1 2	0.689 (179) 0.311 (81)	0.699 (204) 0.301 (88)	1.049 (0.730-1.507)	0.796
45	rs2885546 (G/A)	1 2	0.731 (190) 0.269 (70)	0.767 (224) 0.233 (68)	1.214 (0.825-1.785)	0.325
46	rs2362601 (A/G)	1 2	0.846 (220) 0.154 (40)	0.831 (241) 0.169 (49)	0.894 (0.567-1.411)	0.631
47	rs7618955 (A/G)	1 2	0.519 (135) 0.481 (125)	0.507 (148) 0.493 (144)	0.952 (0.681-1.33)	0.771
48	rs2362605 (G/A)	1 2	0.671 (173) 0.329 (85)	0.676 (196) 0.324 (94)	1.024 (0.717-1.465)	0.895
49	rs2362607 (A/T)	1 2	0.704 (183) 0.296 (77)	0.729 (213) 0.271 (79)	1.134 (0.783-1.644)	0.505
50	rs10513857 (C/A)	1 2	0.842 (219) 0.158 (41)	0.832 (243) 0.168 (41)	0.928 (0.590-1.461)	0.748
51	rs6773525 (A/C)	1 2	0.819 (213) 0.181 (47)	0.839 (245) 0.161 (47)	1.150 (0.738-1.793)	0.537
52	rs9832276 (G/A)	1 2	0.631 (164) 0.369 (96)	0.630 (184) 0.370 (108)	0.997 (0.705-1.41)	0.988
53	rs6778424 (C/A)	1 2	0.873 (227) 0.127 (33)	0.863 (252) 0.137 (40)	0.916 (0.559-1.502)	0.727
54	rs1472597 (A/G)	1 2	0.813 (208) 0.187 (48)	0.836 (244) 0.164 (48)	1.173 (0.755-1.823)	0.478
55	rs10513861 (G/A)	1 2	0.954 (248) 0.0465 (12)	0.932 (272) 0.0684 (20)	0.658 (0.315-1.374)	0.259
56	rs9290953 (A/G)	1 2	0.681 (177) 0.319 (83)	0.699 (204) 0.301 (88)	1.087 (0.758-1.56)	0.651

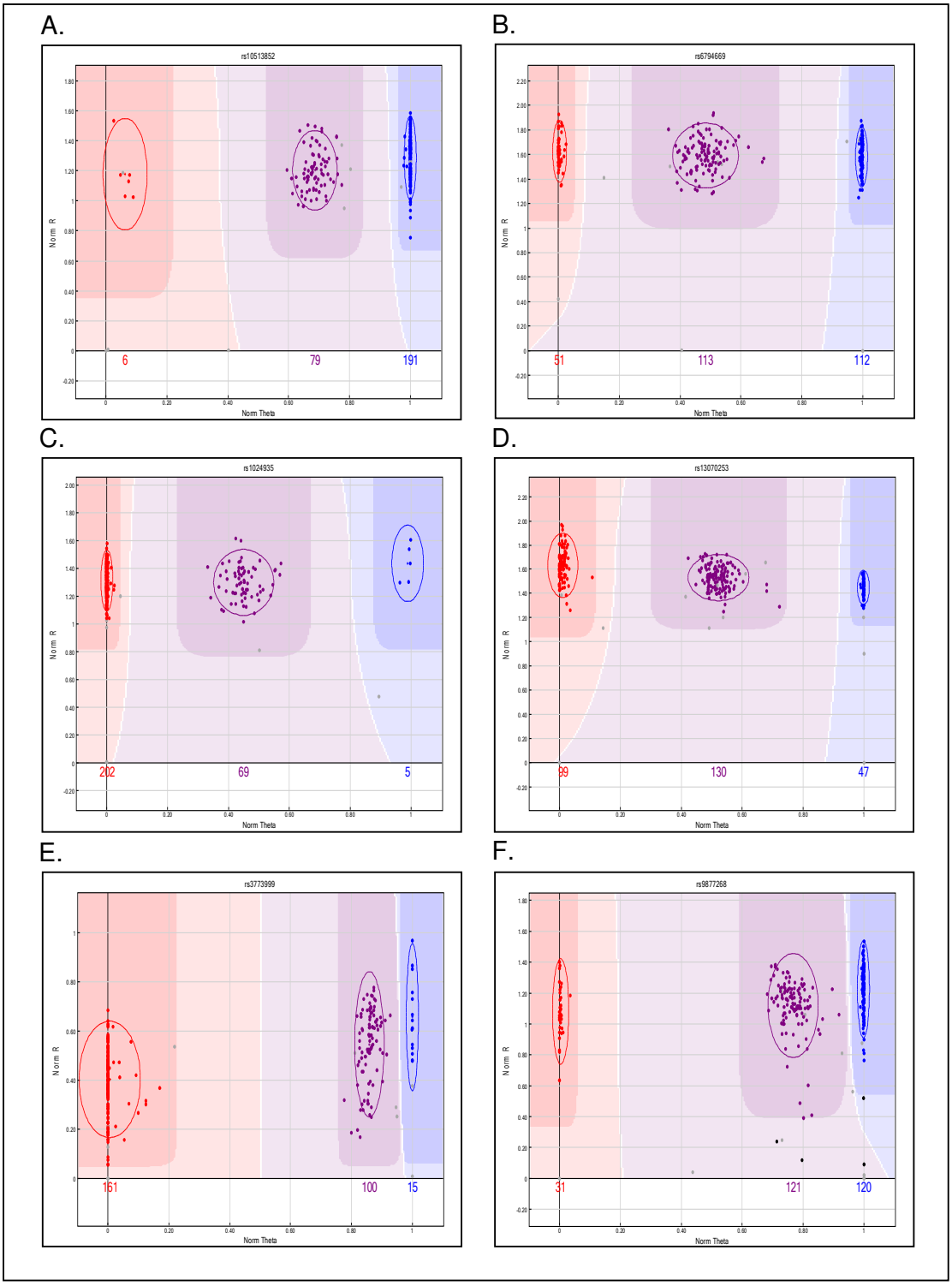
SNP Nº	rs Number	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
57	rs6444462 (G/A)	1 2	0.728 (182) 0.272 (68)	0.753 (220) 0.247 (72)	1.142 (0.777-1.678)	0.501
58	rs6790983 (G/C)	1 2	0.911 (235) 0.0895 (23)	0.883 (256) 0.117 (34)	0.737 (0.422-1.288)	0.280
59	rs9682599 (A/G)	1 2	0.873 (227) 0.127 (33)	0.873 (255) 0.127 (37)	1.002 (0.606-1.656)	0.994
62	rs9290958 (A/T)	1 2	0.892 (232) 0.108 (28)	0.890 (260) 0.110 (32)	0.981 (0.573-1.678)	0.943
66	rs13060312 (G/A)	1 2	0.888 (231) 0.112 (29)	0.880 (257) 0.120 (35)	0.922 (0.546-1.556)	0.760
67	rs9857926 (A/T)	1 2	0.912 (237) 0.0884 (23)	0.932 (272) 0.0684 (20)	1.320 (0.707-2.463)	0.383
76	rs7432178 (G/A)	1 2	0.985 (256) 0.0153 (4)	0.980 (286) 0.020 (6)	0.745 (0.208-2.669)	0.648
78	rs1510901 (G/A)	1 2	0.558 (144) 0.442 (114)	0.562 (164) 0.438 (128)	1.014 (0.724-1.421)	0.934
79	rs6769335 (G/A)	1 2	0.839 (218) 0.161 (42)	0.890 (260) 0.110(32)	1.565 (0.955-2.565)	0.0739
80	rs2063421 (G/A)	1 2	0.712 (185) 0.288 (75)	0.764 (223) 0.236 (69)	1.310 (0.895-1.917)	0.164
81	rs1355365 (A/G)	1 2	0.709 (183) 0.291 (75)	0.707 (205) 0.293 (85)	0.988 (0.683-1.43)	0.951
82	rs9814786 (G/A)	1 2	0.551 (141) 0.449 (115)	0.576 (167) 0.424 (123)	1.107 (0.789-1.554)	0.555
83	rs7624125 (G/C)	1 2	0.815 (212) 0.185 (48)	0.863 (252) 0.137 (40)	1.426 (0.903-2.254)	0.127
84	rs7611804 (A/T)	1 2	0.827 (215) 0.173 (45)	0.873 (255) 0.127 (37)	1.442 (0.900-2.311)	0.127
86	rs9855788 (A/G)	1 2	0.804 (209) 0.192 (51)	0.863 (252) 0.137 (40)	1.537 (0.978-2.412)	0.0616
87	rs9854473 (T/A)	1 2	0.873 (227) 0.127 (33)	0.890 (260) 0.110 (32)	1.181 (0.704-1.982)	0.529
88	rs6788413 (G/C)	1 2	0.573 (149) 0.427 (111)	0.524 (153) 0.476 (139)	0.820 (0.586-1.148)	0.247
89	rs1551424 (A/G)	1 2	0.639 (166) 0.361 (94)	0.716 (209) 0.284 (83)	1.426 (0.996-2.041)	0.0522
90	rs1551425 (G/A)	1 2	0.889 (231) 0.111 (29)	0.908 (265) 0.0924 (27)	1.232 (0.709-2.142)	0.459
91	rs2666356 (A/G)	1 2	0.593 (153) 0.407 (105)	0.560 (174) 0.404 (118)	1.012 (0.719-1.423)	0.946
92	rs11922372 (A/G)	1 2	0.781 (203) 0.219 (57)	0.784 (229) 0.216 (63)	1.021 (0.681-1.530)	0.921
93	rs7628727 (C/A)	1 2	0.638 (166) 0.362 (94)	0.630 (184) 0.370 (108)	0.965 (0.682-1.365)	0.839
94	rs6444517 (A/G)	1 2	0.896 (233) 0.104 (27)	0.918 (268) 0.0822 (24)	1.294 (0.727-2.305)	0.381
95	rs9290978 (A/C)	1 2	0.700 (182) 0.300 (78)	0.733 (214) 0.267 (78)	1.176 (0.812-1.704)	0.392

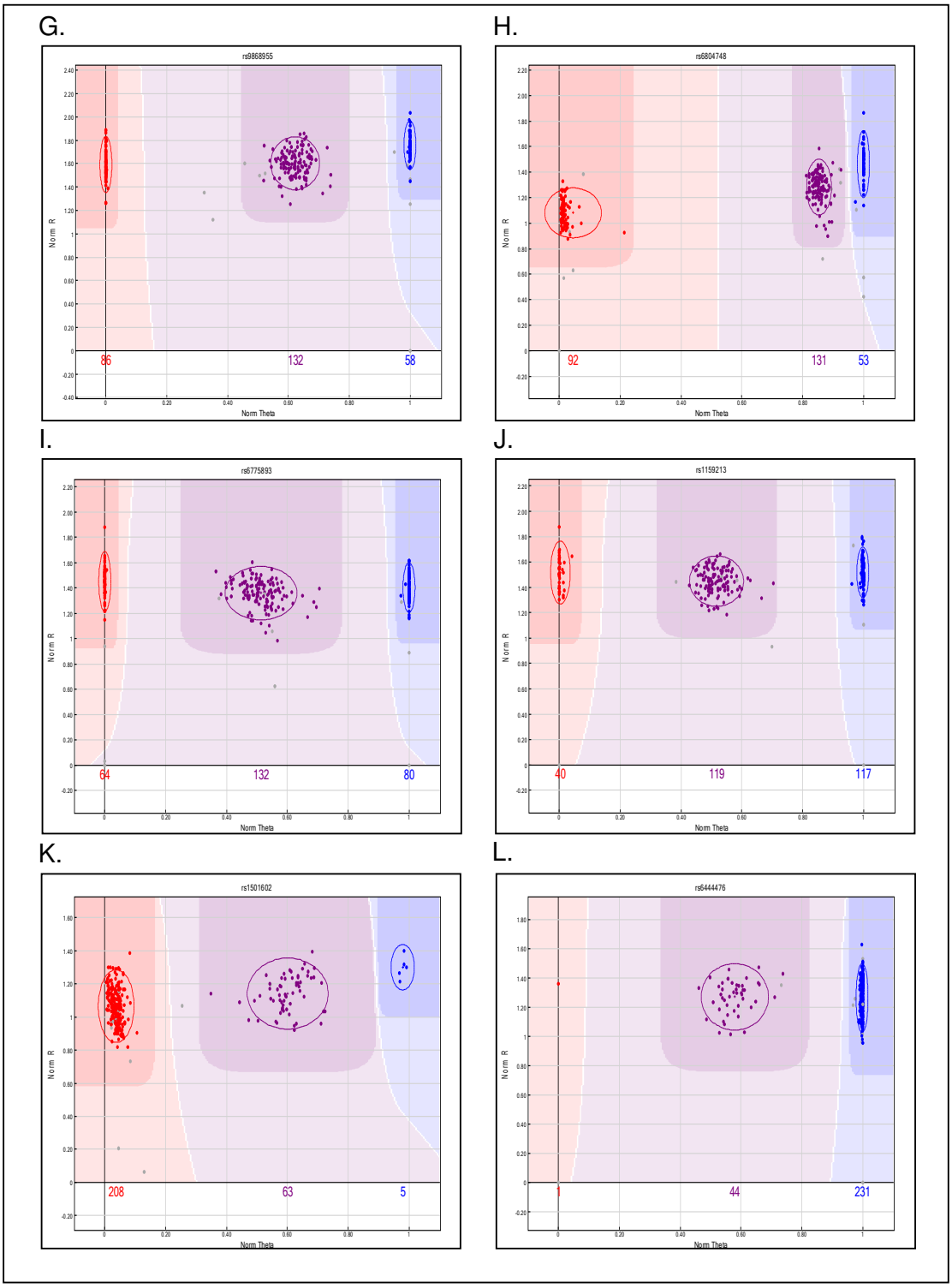
Table 4.18 Non-significant *IL1RAP* candidate region stage-1 analysis results

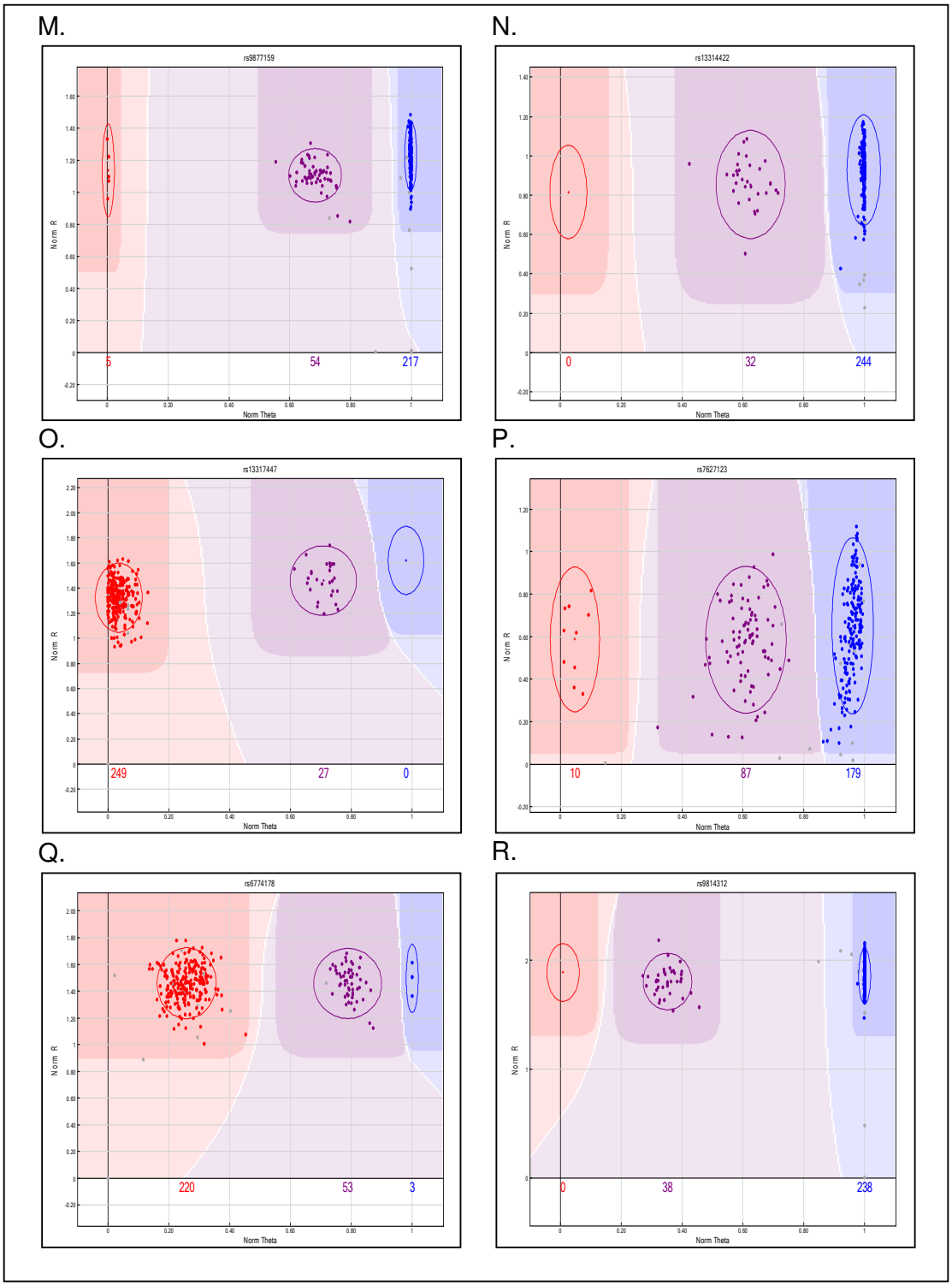
Ninety two genotyped tSNPs within the *IL1RAP* candidate region were analysed for significant allele frequency differences between the patient (n=130) and control (n=146) cohorts. The frequency in each cohort of both the common (1), and rare (2) alleles are given. The alternative alleles of each SNP are shown underneath the rs identifier. The allele assigned as the common (1) allele is given first and the rarer (2) allele second. The OR shown is for allele 2 (allele 1 OR=1). The p-value for the test for frequency differences between the cohorts is also shown. The results for the 73 tSNPs which did not show significant (p>0.05) frequency differences between the two cohorts are shown.

allele. The rare allele was present at a higher frequency in the patients in the other 12 SNPs. For SNP 64 different alleles were more frequent in the two populations. In the case population the G allele, termed as allele 1, was the most common (0.581), while in the control population this allele was only present at a frequency of 0.483, making the alternative A allele the more common. For SNP61 both alleles were present in the case population at the same frequency (0.5), while in the control population they were at frequencies of 0.596 and 0.404.

Following Bonferroni correction (0.05/92 analysed SNPs = $p<0.00054$) the frequency difference in only SNP70 ($p=0.00033$), remained significant, although SNP68 ($p=0.000578$) was tending toward significance. However, Bonferroni correction assumes that the variables tested are independent, and is therefore overly conservative in this study as the SNPs tested will be in some amount of LD with each other (Lewis, 2002). The genoplots of the 19 *IL1RAP* SNPs with significant frequency differences (Figure 4.22) show that for all the SNP assays, apart from SNPs 71 (N.), 72 (O.), and 75 (R.) for which only one homozygote cluster was observed, the DNA samples separated into three distinct clusters corresponding to each genotype. The DNA samples were dispersed along the y (normalised intensity) axis for the SNP14 (E.), 25 (F.), and 73 (P.) assays. This suggests that these assays had not performed equally well in all samples resulting in variation in the fluorescence intensity possibly the result of the strength with which the primers bound to the DNA. In the SNP 14 and 25 assays the samples within each genotype cluster showed very similar normalised theta values so each cluster could still be identified with confidence. Although the genotype clusters were more dispersed for the SNP73 assay compared to all of the others there was still clear separation between the clusters enabling genotype assignment. The GC10 scores (Appendix 2) for all 19 SNP assays also reflected the quality of the genotype assignment, all being in excess of the GC10≥0.5 cut off used for quality control. Apart from SNP74 (0.677) the GC10 score for all 19 SNPs was ≥0.724, with 63% ≥0.829. The GC10 scores for the other SNPs were: SNP1=0.838, SNP2=0.943, SNP3=0.926, SNP13=0.919, SNP14=0.742, SNP25=0.748, SNP61=0.904, SNP63=0.724, SNP64=0.941, SNP65=0.952, SNP68=0.848, SNP69=0.859, SNP70=0.793, SNP71=0.848, SNP72=0.762, SNP73=0.794, SNP75=0.829, SNP85=0.842. These quality measurements of the SNP assays the SNP assays enable high confidence in the genotype assignment, and therefore that the significant differences observed are due to true frequency differences between the populations and not genotyping error.


A.
rs10513852
Norm R
Norm Theta
6
79
191
B.
rs6794669
Norm R
Norm Theta
51
113
112
C.
rs1024935
Norm R
Norm Theta
202
69
5
D.
rs13070253
Norm R
Norm Theta
99
130
47
E.
rs3773999
Norm R
Norm Theta
161
100
15
F.
rs9877268
Norm R
Norm Theta
31
121
120


G.
rs9868955
Norm R
Norm Theta
86
132
58
H.
rs6804748
Norm R
Norm Theta
92
131
53
I.
rs6775893
Norm R
Norm Theta
64
132
80
J.
rs1159213
Norm R
Norm Theta
40
119
117
K.
rs1501602
Norm R
Norm Theta
208
63
5
L.
rs6444476
Norm R
Norm Theta
1
44
231


M.
rs9877159
Norm R
Norm Theta
5
54
217
N.
rs13314422
Norm R
Norm Theta
0
32
244
O.
rs13317447
Norm R
Norm Theta
249
27
0
P.
rs7627123
Norm R
Norm Theta
10
87
179
Q.
rs6774178
Norm R
Norm Theta
220
53
3
R.
rs9814312
Norm R
Norm Theta
0
38
238

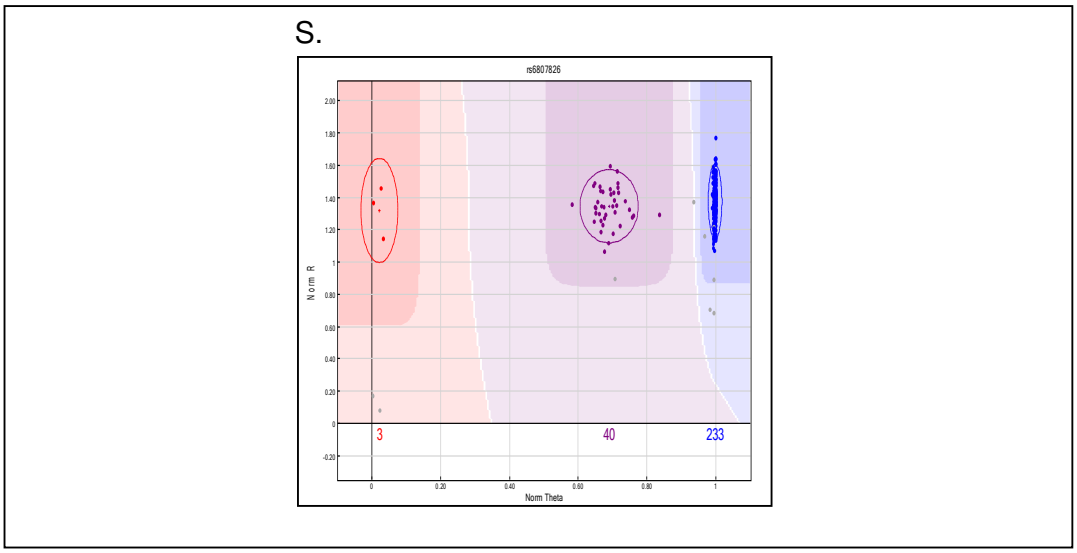


Figure 4.22 Genoplots of the significant *IL1RAP* SNPs

Shown are the genoplots of the 19 *IL1RAP* region SNPs which showed evidence of significant frequency differences in the initial stage-1 analysis. Normalised theta values (deviation from total 'A' allele signal) are plotted along the x-axes, and normalised R values (signal intensity) along the y-axes. Each data point represents one sample, the circles indicate the location of each genotype cluster, and the dark shaded areas are the call areas.

A. SNP1 rs10513852, B. SNP2 rs6794669, C. SNP3 rs1024935, D. SNP13 rs13070253, E. SNP14 rs3773999, F. SNP25 rs9877268, G. SNP61 rs9868955, H. SNP63 rs6804748, I. SNP64 rs6775893, J. SNP65 rs1159213, K. SNP68 rs1501602, L. SNP69 rs6444476, M. SNP70 rs9877159, N. SNP71 rs13314422, O. SNP72, rs13317447, P. SNP73 rs7627123, Q. SNP74 rs6774178, R. SNP75 rs9814312, S. SNP85 rs10937457.

To aid identification of secondary effects, the LD relationships between these nineteen SNPs were determined. As a tSNP approach to SNP selection was employed, none of the genotyped SNPs were in high LD with each other according to the r^2 statistic. However, r^2 is a measure of the LD relationship between alleles at different loci, but the alternative LD statistic D′ is a measure of the evidence of recombination between loci (section 3.1.1.1.). While r^2 is the most appropriate measurement in identifying SNPs which are proxies for each other, D′ is important when indentifying SNPs which are inherited together.

The D′ LD between the nineteen SNPs identified as showing significant allele frequency differences between cohorts in the initial analysis is shown in Figure 4.23. The nineteen SNPs fell into three clusters according to their LD patterns. One LD cluster was comprised of SNPs 1, 2, and 3, which were all in D′=1 with each other. SNPs 61, 63, 64, 65, 68, and 69 comprised the second LD cluster indicated in Figure 4.23. Twelve of the fifteen pairwise LD comparisons between these six SNPs were D′=1. The SNPs in this cluster not in complete LD with each other were: SNP61 and SNP68 (D′=0.86), SNP63 and SNP64 (D′=0.73), and SNP64 and SNP68 (D′=0.83). The third LD block also included six SNPs: SNPs 70, 71, 72, 73, 74, and 75. Apart from SNP74 with SNP70 (D′=0.19) and with SNP71 (D′=0.66), the SNPs within this cluster were all in D′=1 with each other. SNP13 (D′ maximum=0.65) and SNP85 (D′ maximum=0.57) were not in high LD with any of the other SNPs indentified in stage-1 of the study as being at significant frequency differences between the cohorts. Although SNP25 was in high LD with SNP69 (D′=1), it was not in high LD with any of the other SNPs within the candidate region (D′ maximum=0.53). It was therefore not included in any of the LD clusters.

In the conditional analysis, SNPs 1, 2, and 3 did not show significant effects when analysed controlling for each other (Table 4.19). This indicated that these three SNPs were all reflecting one single effect. Due to the LD between the SNPs it was not possible to elucidate with certainly, through conditional analysis, which was the primary effect. Step-wise conditional analysis of the SNPs in the second LD cluster (SNPs 61 to 69), showed that only SNP68 conferred a significant effect when each of the other SNPs were already taken into account. The other five SNPs in the cluster did not remain significant when analysed conditioning on each other, or on SNP68. This demonstrated that each of the other SNPs were showing secondary effects, due to being in LD with SNP68. A similar pattern was seen between the SNPs in the remaining LD cluster (SNPs 70 to 75). SNP73 continued to show a

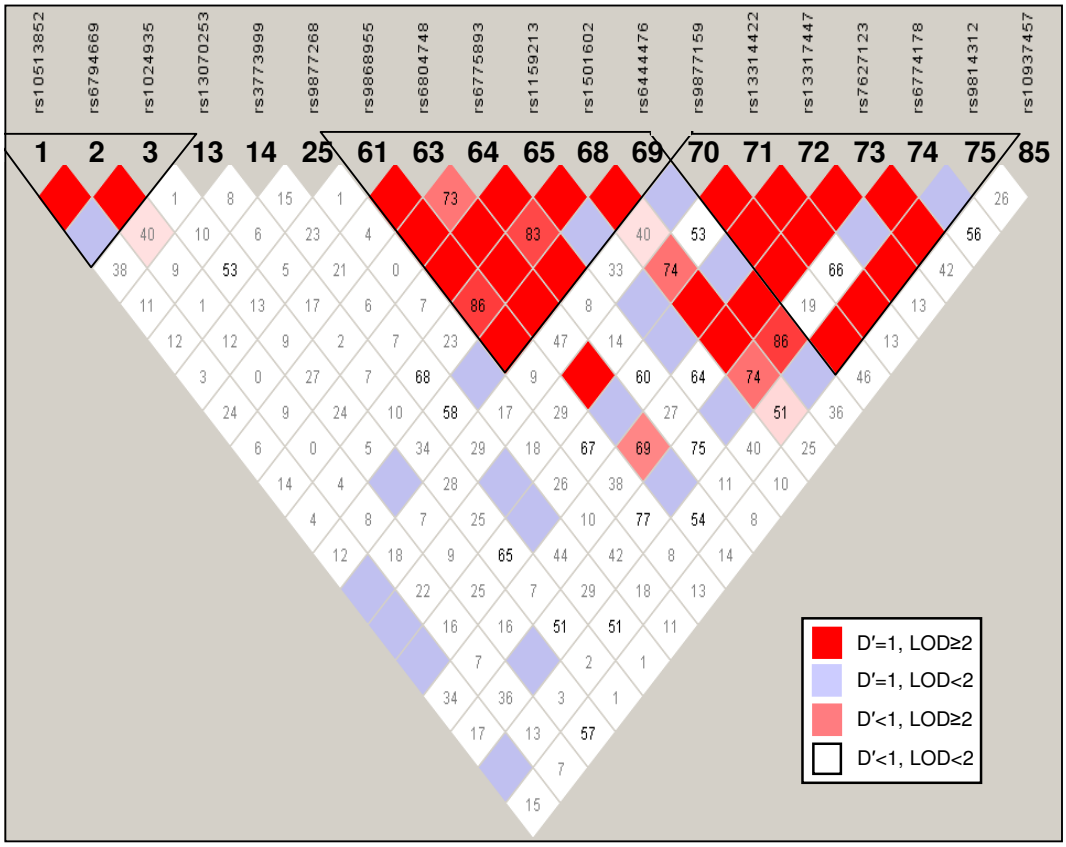


Figure 4.23 D′ relationship of associated *IL1RAP* region SNPs

In the initial analysis nineteen SNPs within the *IL1RAP* candidate region showed significant allele frequency differences between the case and control populations. The graphical representation of the D′ LD relationship between these nineteen SNPs, generated in Haploview, is shown on a sliding colour scale from white (low D′) to red (high D′), depending on LOD score. The D′ value for each pairwise comparison is shown, where no value is indicated D′=1. The LD blocks used for the conditional analysis are indicated.

LOD= logarithm of odds

Test marker	Conditional marker	p value
SNP 1, 2, and 3 cluster		
1	2	0.125
1	3	0.0539
2	1	0.0880
2	3	0.0685
3	1	0.0788
3	2	0.469
SNP 61, 63, 64, 65, 68, 69 cluster		
61	63	0.294
61	64	0.955
61	65	0.482
61	68	0.295
61	69	0.182
63	61	0.609
63	64	0.647
63	65	0.915
63	68	0.321
63	69	0.195
64	61	0.199
64	63	0.491
64	65	0.409
64	68	0.203
64	69	0.150
65	61	0.186
65	63	0.168
65	64	0.473
65	68	0.204
65	69	0.129
68	61	0.0102
68	63	0.00922
68	64	0.00772
68	65	0.00714
68	69	0.00765
69	61	0.211
69	63	0.159
69	64	0.170
69	65	0.129
69	68	0.522
SNP 70,71,72,73,74, and 75 cluster		
70	71	0.219
70	72	0.303
70	73	0.868
70	74	0.0795
70	75	0.217
71	70	0.442
71	72	1.00
71	73	0.673
71	74	0.0272
71	75	0.473
72	70	0.183
72	71	1.00
72	73	0.330
72	74	0.00856
72	75	0.153
73	70	0.0108

Test marker	Conditional marker	p value
73	71	0.00341
73	72	0.00603
73	74	0.00306
73	75	0.00329
74	70	0.0985
74	71	0.0182
74	72	0.0138
74	73	0.0584
74	75	0.0249
75	70	0.660
75	71	0.834
75	72	0.976
75	73	0.932
75	74	0.0457

Table 4.19 Conditional analysis of associated *IL1RAP* region SNPs within LD clusters

In the initial analysis nineteen SNPs within the *IL1RAP* candidate region showed significant allele frequency differences between the case and control populations. Based on the D′ LD between these SNPs (Figure 4.23) it was identified that 15 of the SNPs fell into three LD clusters (SNPs 1, 2, and 3, SNPs 61, 63, 64, 65, 68, and 69, and SNPs 70, 71, 72, 73, 74, and 75). SNPs 13, 14, 25, and 85 did not fall within an LD cluster. The SNPs within each LD cluster were tested for significant frequency differences conditioning on the other SNPs within that cluster.

significant effect when each of the other SNPs had been taken into account. However, none of the other SNPs in the LD cluster gave a significant effect, in addition to that of SNP73. SNP74 also remained significant when analysed conditioning on SNPs 71, 72, and 75, but not when conditioned on SNP70 or SNP73. These results indicate that, of the six SNPs in this LD cluster, SNP73 was the only one showing an independent, primary effect.

Analysis was then performed to determine if the associations seen in each of the three clusters, and the four SNPs not in high LD with any others, were independent from each other. Neither SNP1 nor SNP3 showed any significant difference between the cohorts when the effect of each of the other six SNPs were included in the baseline model (Table 4.20). As this indicated that these two SNPs were not independent associations they were both eliminated from the model. SNP2 and SNP13 were not significant when analysed controlling for each other (SNP2 conditioning on SNP13 p=0.0758, SNP13 conditioning on SNP2 p=0.0882). This suggests that both of these SNPs may be reflecting the same association, although they are not in high LD with each other (D′=0.4). Each of SNP2, SNP13, SNP14, SNP25, and SNP85, except for when SNP2 and SNP13 were analysed together, showed a significant difference between the cohorts after the effect of each of the other SNPs had been taken into account. This indicated that the disease association shown by these SNPs was independent to any other effects in the candidate region. Although SNP68 showed evidence of significant association, independent to each of the other SNPs included in the conditional analysis, this was abolished when analysed in addition to SNP73 (p=0.212). This indicates that SNP68 is secondary to SNP73, and that the allele frequency difference observed is a reflection of that at SNP73, which it is in high LD with (D′=1).

4.3.1.3. Stage- 2

The six *IL1RAP* candidate region SNPs which showed evidence of significant allele frequency differences following conditional analysis were included in stage-2 of the study. These were: SNPs 2, 13, 14, 25, 73, and 85 (Table 4.21). SNP14a (rs9290935) was genotyped in place of SNP14 due to assay design constraints. The LD data available at the time regarding these SNPs (based on HapMap2) showed that these two SNPs were in complete LD (r^2=1) and were therefore perfect proxies for each other. The LD relationship was subsequently calculated as r^2=1 based on HapMap3, and r^2=0.997 based on WTCCC2 (section 4.1.5.).

Test marker	Conditional marker	p value
1	13	0.156
1	14	0.581
1	25	0.0895
1	68	0.274
1	73	0.269
1	85	0.0778
2	13	0.0758
2	14	0.0126
2	25	0.0157
2	68	0.0294
2	73	0.0270
2	85	0.0131
3	13	0.902
3	14	0.0978
3	25	0.296
3	68	0.431
3	73	0.529
3	85	0.203
13	1	0.0470
13	2	0.0882
13	3	0.0334
13	14	0.0109
13	25	0.0358
13	68	0.0293
13	73	0.0274
13	85	0.0249
14	1	000462
14	2	0.0299
14	3	0.0276
14	13	0.0223
14	25	0.0146
14	68	0.0279
14	73	0.0362
14	85	0.0386
25	1	0.00476
25	2	0.00922
25	3	0.0116
25	13	0.0139
25	14	0.0126
25	68	0.0102
25	73	0.0165
25	85	0.0128
68	1	0.0019
68	2	0.00316
68	3	0.00177
68	13	0.00144
68	14	0.00195
68	25	0.00130
68	73	0.212
68	85	0.00121
73	1	0.00152
73	2	0.00532
73	3	0.00723
73	13	0.00183
73	14	0.00110
73	25	0.000527
73	68	0.0574
73	85	0.00362
85	1	0.0253

Test marker	Conditional marker	p value
85	2	0.0279
85	3	0.0285
85	13	0.0468
85	14	0.0346
85	25	0.0530
85	68	0.0145
85	73	0.00660

Table 4.20 Conditional analysis of associated *IL1RAP* region SNPs between LD clusters

The effects indentified within each LD cluster were tested to determine if they were independent of each other. The SNPs within each LD cluster, and the four SNPs which were not within a cluster, were tested for significant frequency differences conditioning on the SNPs within the other clusters.

SNP	Allele	Case freq. (N^{o})	Control freq. (N^{o})	OR (95% CI)	p value
2	1 2	0.582 (114) 0.418 (82)	0.574 (210) 0.426 (156)	0.968 (0.681-1.376)	0.857
13	1 2	0.540 (109) 0.460 (93)	0.599 (218) 0.401 (146)	1.274 (0.900-1.803)	0.172
14	1 2	0.709 (146) 0.291 (60)	0.668 (239) 0.332 (119)	0.825 (0.569-1.198)	0.311
25	1 2	0.677 (138) 0.324 (66)	0.617 (221) 0.383 (137)	0.772 (0.537-1.109)	0.159
73	1 2	0.824 (168) 0.177 (36)	0.803 (289) 0.197 (71)	0.872 (0.559-1.360)	0.544
85	1 2	0.722 (143) 0.278 (55)	0.624 (221) 0.376 (133)	0.639 (0.438-0.933)	0.0189

Table 4.21 *IL1RAP* candidate region stage-2 analysis results

The six *IL1RAP* candidate region SNPs identified as showing significant frequency differences following conditional analysis in stage-1 of the association study, were genotyped in the stage-2 patient (n=105) and control (n=184) cohorts. The frequency in each cohort of both the common (1) and rare (2) alleles of each SNP are given. The OR shown is for allele 2 (allele 1 OR=1). The p-value for the test for frequency differences between the cohorts is also shown.

Only SNP85 showed evidence of significant frequency differences between the patient (n=105) and control (n=184) cohorts used in stage-2 (p=0.0189). However, the difference was in the opposite direction to that seen in stage-1. In stage-1 allele 2 was significantly more frequent in the patients (0.35) than the controls (0.267), but in the stage-2 samples allele 2 was significantly less frequent in the patients (0.278) than the controls (0.376). The frequencies of this SNP appeared to be the opposite way around in the two stages, in that the stage-1 patients had a similar allele frequency to the stage-2 controls, and the stage-1 controls had a similar allele frequency to the stage-2 patients. This was not due to an error in sample or allele labelling during analysis, and suggests that the frequency difference observed at this loci may be due to sampling error, and not related to disease. The allele frequencies in the majority of the SNPs were similar in the populations from both stages of the study. However, in addition to the discrepancy with SNP85, some of the other SNPs showed differences between the populations used in the two stages. The MAF of SNP2 in the controls was 0.342 in the stage-1 population but 0.426 in the stage-2 population. The MAFs in the patient populations from each stage were 0.354 and 0.460 for SNP 13, and 0.196 and 0.291 for SNP14.

4.3.1.3.1. Stratified analysis

Following stratified analysis of the data from both stages of the study, stratified by study stage, three of the six *IL1RAP* SNPs showed evidence of significant frequency differences between the patient and control cohorts (Table 4.22). The frequency of the rare allele of SNP14 (CMH p=0.0323) and SNP25 (CMH p=0.00499) were significantly lower in the patient population than the control population. Neither of these SNPs showed significant differences in effect size between the stages, BD p=0.416 and 0.418 respectively. SNP73 also showed evidence of significant allele frequency differences (CMH p=0.0311). However, the BD p value for this SNP was also significant (0.00355). This indicates that the between cohort difference in allele frequency was significantly different between the populations used in the two stages. Significant heterogeneity of odds between the study stages was also shown for two of the three SNPs with no significant allele frequency differences overall, SNP13 (BD p=0.00875), and SNP85 (BD p=0.00169). Significant heterogeneity between the study stages was expected for SNP85 as significant allele frequency differences were seen in both stages individually, but in opposing directions (section 4.3.1.3.).

SNP	CMH p value	OR (95% CI)	BD p value
2	0.112	1.220 (0.955-1.558)	0.0709
13	0.472	0.915 (0.718-1.166)	0.00875
14	0.0323	0.743 (0.566-0.976)	0.416
25	0.00499	0.694 (0.537-0.896)	0.418
73	0.0311	1.396 (1.031-1.888)	0.00355
85	0.892	0.982 (0.759-1.271)	0.00169

Table 4.22 *IL1RAP* candidate region stratified analysis results

For the six SNPs investigated in both stages of the study, association meta-analysis of the data was performed using the CMH test, stratifying by study stage. A significant result for the CHM test shows that the SNP is associated with disease. The p value and OR for this test are given, the OR shown is for allele 2 (allele 1 OR=1). A BD test for homogeneity of odds ratios between the two stages was also performed. A significant BD test shows that the allele frequency difference is more prominent in one of the populations studied. Any association is therefore more likely to be due to sampling error rather than true association.

BD-Breslow-Day, CI-Confidence Interval, CMH-Cochran-Mantel-Haenszel, OR-Odds Ratio

4.3.1.4. WTCCC control cohort analysis

An additional analysis with greater statistical power to detect SNPs associated with disease was performed with the pooled stage-1 and stage-2 cases (n=235), and the WTCCC2 cohort of healthy UK individuals (n=4,671).

IL1RAP SNP2 (rs6794669) and SNP13 (rs1024935) were not available for the WTCCC2 control population. Genotyping data was also not available for SNP73 (rs7627123) in the control cohort. As a proxy for SNP73, control genotyping data for rs1393049 (SNP73a, r^2=1 in HapMap 2 and 3, section 4.1.5.) was used in the analysis.

With this larger UK wide healthy control cohort none of the tested SNPs, even the two which showed evidence of significant disease association in the stratified analysis, showed a significant allele frequency difference between the pooled stage-1 and -2 patient (n=235) and the WTCCC2 control (n=4671) cohorts (Table 4.23).

4.3.1.4. Associated SNPs

In the initial tSNP selection based on the HapMap 2 genotyping data (Table 4.16) SNP14 was tagging one other SNP with r^2=1. SNP25 was not tagging any other SNPs within the candidate region. This tagging relationship was confirmed during the subsequent validation using the larger HapMap3 and WTCCC2 populations. The LD calculated between SNP14 and rs9290935 in these larger datasets were r^2=1 from HapMap3 and r^2=0.997 from WTCCC2. SNP25 was not tagging any other SNPs within the candidate region.

4.3.1.4.1. Additional captured SNPs

Based on the HapMap3 dataset nine SNPs were identified by SNAP (section 3.1.5.) as being in r^2>0.8 with SNP14 (Table4.24). SNP25 was not in r^2>0.8 with any SNPs within 500kb. All SNP14 tagged SNPs are within *IL1RAP*.

4.3.1.4.2. Associated SNP positions

The positions of *IL1RAP* SNP14 and SNP25, which showed significant evidence of association, and the 10 SNPs tagged by SNP14, are shown in Figure 4.24. All of the associated SNPs are located within the *IL1RAP* gene sequence. SNP14 and the 10 SNPs in high LD with it are all within a 6.7kb region within *IL1RAP* intron 1. SNP25 is within

SNP	Allele	Case freq. (N^{o})	Control freq. (N^{o})	OR (95% CI)	p value
14/14a	1	0.762 (355)	0.727 (6786)	0.831	0.0908
	2	0.238 (111)	0.273 (2554)	(0.668-1.033)	
25	1	0.700 (325)	0.671 (6263)	0.871	0.179
	2	0.230 (139)	0.329 (3075)	(0.711-1.067)	
73/73a	1	0.845 (392)	0.824 (7671)	0.862	0.249
	2	0.155(72)	0.176 (1635)	(0.667-1.114)	
85	1	0.681 (312)	0.689 (6434)	1.035	0.741
	2	0.319 (146)	0.311 (2910)	(0.846-1.265)	

Table 4.23 *IL1RAP* candidate region analysis results with WTCCC2 controls

As an additional analysis with a larger, independent control cohort, allele frequencies in the pooled cases (n=235) from both stages of the association study were compared to the controls genotyped in phase-2 of the WTCCC project (n=4671). Genotyping data for rs7627123 (SNP73) in all cases, and for rs1393049 (SNP73a) (r^2=1 HapMap 2 and 3) in the controls, were used. The frequency in each cohort of both the common (1) and rare (2) alleles of the two individual SNPs are given. The OR shown is for allele 2 (allele 1 OR=1). The p-value for the test for frequency differences between the cohorts is also shown.

SNP	r^2	Location
rs4140708	1	*IL1RAP*
rs7630034	0.894	*IL1RAP*
rs3774001	1	*IL1RAP*
rs9990231	0.957	*IL1RAP*
rs11719243	0.957	*IL1RAP*
rs16865541	0.957	*IL1RAP*
rs10212103	1	*IL1RAP*
rs9848635	0.957	*IL1RAP*
rs16865546	0.833	*IL1RAP*

Table 4.24 Additional SNPs tagged by *IL1RAP* SNP14

Based in the HapMap3 population genotyping data all SNPs in r^2>0.8 with, and within 500kb of, the *IL1RAP* SNPs showing evidence of significant disease association, were identified. The SNP identifier, r^2 value, and location of all the identified SNPs are shown. All SNPs shown are tagged by SNP14 as SNP25 was not in r^2>0.8 with any SNPs within 500kb.

IL1RAP intron 3, 45.8kb downstream of rs16865546, the most 5′ of the SNP14 associated SNPs.

4.3.1.5. Summary

Following analysis of the data from both study stages, two SNPs within the *IL1RAP* candidate region, rs3773999 (SNP14) and rs9877268 (SNP25) were identified as showing evidence of significant association with sJIA. SNP14 is in high LD (r^2>0.8) with a total of ten other SNPs within the candidate region, SNP25 is not tagging any other SNPs. All of the associated SNPs are located within introns of *IL1RAP*.

4.3.1.6. Discussion

The *IL1RAP* gene encodes the IL-1 receptor accessory protein (IL-1RAcP), two isoforms of which are produced through alternative splicing, neither of which bind directly to IL-1. The larger membrane bound form (mIL-1RAcP) is the co-receptor required for all signalling through the type 1 IL-1 receptor (Hofmeister et al., 1997; Huang et al., 1997; Korherr et al., 1997; Wesche et al., 1997b; Debets et al., 2001; Towne et al., 2004). The shorter secreted isoform (sIL1RAcP) interacts with the secreted form of the type 2 IL-1 receptor creating a high affinity IL-1 scavenger, reducing circulating levels of IL-1 (Smith et al., 2003). The importance of sIL-1RAcP in inhibiting the IL-1 inflammatory response has been demonstrated in collagen-induced arthritis mouse models as overexpression of sIL-1RAcP significantly ameliorates disease activity (Smeets et al., 2003).

Two intronic *IL1RAP* SNPs have been associated with amniotic fluid IL-1β concentrations and pre-term birth in African Americans. The associations were with the maternal genotype at rs1024941, and the fetal genotype at rs3773953 (Menon et al., 2010). Both SNPs were included in this association study, rs3773953 was directly genotyped (SNP33), and rs1024941 was tagged (r^2=1) by SNP37 (based on SNAP analysis, section 3.1.5). Neither SNP however showed evidence of significant association with sJIA.

In this study two *IL1RAP* SNPs were identified as being significantly associated with susceptibility to sJIA. However, no significant allele frequency difference was seen when the analysis was repeated comparing the patients from both study stages with the large WTCCC2 control cohort.

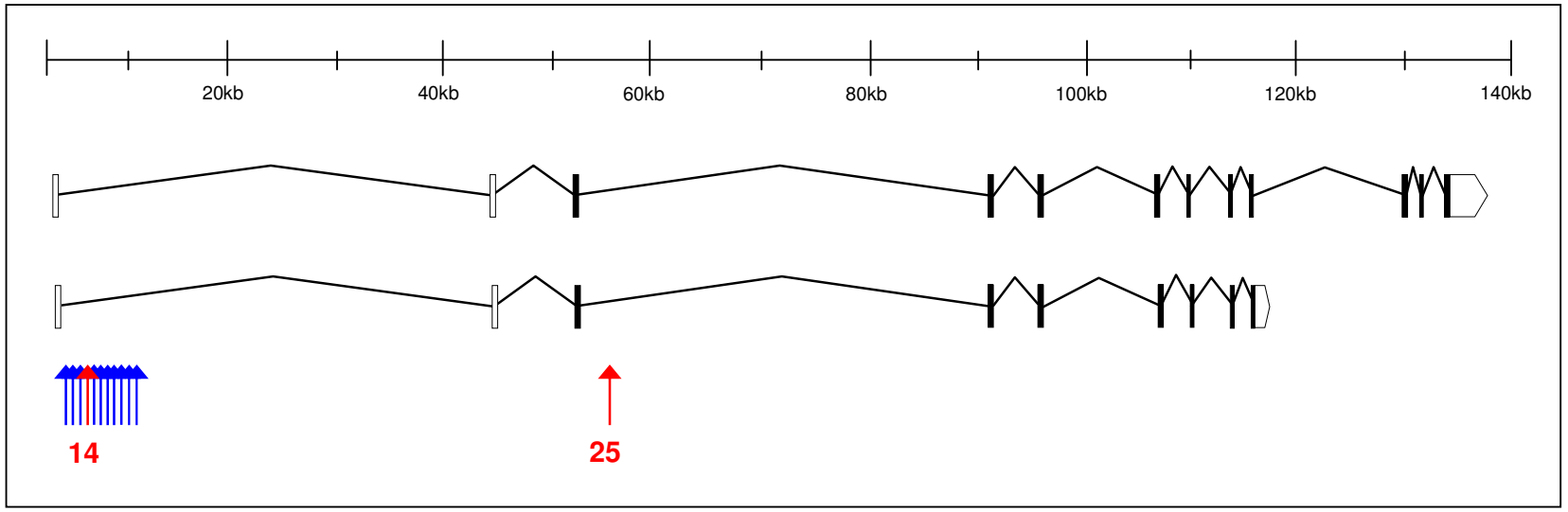


Figure 4.24 Positions of the associated SNPs in the *IL1RAP* candidate region

The positions within the *IL1RAP* candidate region of SNP14 (rs3773999) and SNP25 (rs9877268), which showed evidence of significant association, are shown in red, and the 10 SNPs tagged by SNP14 are shown in blue. Both the membrane bound (top) and secreted (bottom) isoforms of *IL1RAP* are shown. Untranslated regions are represented by white boxes and coding exons by black boxes. Due to the scale of the figure only the two genotyped SNPs are labelled. The tagged SNPs in chromosomal order are: rs4140708, rs7630034, rs3774001, rs9990231, rs11719243, rs9290935, rs16865541, rs10212103, rs9848635, rs16865546.

One of the SNPs found to be associated with disease (SNP14) is known to be tagging a total of ten other SNPs within the candidate region. All of the associated SNPs are located within introns of the gene. It is possible that these SNPs may play a role in gene splicing, and thus the proportion in which the two isoforms are expressed. It has been shown that under phorbol ester induced inflammatory response conditions *in vitro*, the splicing pattern of this gene alters so that sIL-1RAP mRNA is the more highly expressed form relative to mIL-1RAP (Jensen et al., 2000a). This suggests that splicing regulation of this gene is essential in ensuring that both isoforms are expressed at the appropriate point, and in the relevant tissue, during an inflammatory response. Any disruption of this splicing regulation could potentially result in a disturbance of the pro- and anti-inflammatory balance of the IL-1 inflammatory response.

It is important to bear in mind that SNP14 is in high LD ($r^2 \geq 0.8$) with a number of other SNPs. It is therefore possible that SNP14 is merely acting as a marker for the causal SNP related to disease susceptibility. SNP14 is only in complete LD ($r^2=1$) with, and therefore acting as a perfect proxy for, three of the 10 identified tagged SNPs. If the genotyped SNP14 is acting as an imperfect marker for the true associated SNP, the actual contribution to disease risk may be larger than that seen with SNP14. If this were the case it would be expected that the significance level would also be similarly effected, and that the modest p value ($p=0.0323$) may only be a proportion of the true value. Although not possible as part of this project due to time constraints, it would be very interesting to directly test the incompletely captured SNPs for disease association as this would assist in elucidation of the SNP/SNPs responsible for the association observed.

The alleles of one SNP in this candidate region, SNP85, were at a significantly different frequency between the stage-2 populations. The frequency difference was however in the opposite direction to that observed in the populations used in stage-1 of the study. This was not due to labelling or allele assignment error. It was also not due to genotyping error, as there were no inconsistencies in genotype assignment of the 18 patients genotyped on the two different platforms used (section 4.1.4.). Through the use of stratified meta-analysis to analyse the data from both stages of the study, this discrepancy between the stages was accounted for, resulting in a lack of

disease association over the whole study and demonstration of the significant heterogeneity between the study stages.

Although SNP73 (rs7627123) did show a significant (p=0.0311) difference in the stratified meta-analysis, it was not considered to be associated with disease susceptibility. This is because this SNP also showed considerable heterogeneity (BD p value=0.00355) between the two study stages. This heterogeneity was the result of allele frequency differences in the two patient populations, 0.258 in stage-1 compared to 0.137 in stage-2. The control populations however had similar frequencies, 0.17 and 0.19. As mentioned in section 4.3.1.2. the genoplot for this SNP assay showed more widely dispersed genotype clusters than generated with other SNPs. Although there was still clear separation between the three clusters, and the GC10 score (0.794) was in excess of the cut-off used during quality control, it is possible that the frequency difference seen in the stage-1 populations was spurious due to genotyping inaccuracy at this locus. Of the 18 patients genotyped on both of the platforms used in this study there was however only one discrepancy seen for this SNP (Table 4.4). It may therefore be possible that the genotyping is accurate and there is a genuine frequency difference of this SNP in the patient populations used. This issue, as well as low power to detect significant associations, multiple testing, and the lack of significant findings in the stage-2 populations alone, are discussed in chapter 5.

4.4. *IL1* ligand cluster candidate region

4.4.1. Association study

The genes flanking the nine candidate genes in the *IL1* ligand cluster region were identified using the July 2003 human genome reference sequence (NCBI build 34), hg16 annotation track. The boundaries of the candidate region, chromosome 2:113617332-114026720 were defined based on the positions of the flanking genes. Flanking *IL1A* 3′ is cytoskeleton associated protein 2-like (*CKAP2L),* and 5′ of *IL1RN* is pleckstrin and Sec7 domain containing 4 (*PSD4)*. PSD4, also known as exchange factor for ADP-ribosylation factor guanine nucleotide factor 6B (EFA6B) is involved in membrane recycling, and the coordination of actin cytoskeletal reorganisation through activation of the ADP ribosylation factor ARF6 (Derrien et al., 2002). The function of CKAP2L is not known. Based on the positions of the genes flanking *CASP1* the candidate region selected for investigated spanned 409kb (409,388bp). The region extended 9.3kb (9,307bp) 3′, and 40kb (39,982) 5′, of the *IL1* ligand candidate genes (bottom panel Figure 4.25).

4.4.1.1. tSNP selection

Genotyping data was available for a total of 506 SNPs within the selected *IL1* ligand cluster candidate region, 169 from HapMap (top panel Figure 4.25) and 357 from PGA. Of these, 20 were included in both population genotyping resources, and 337 had data from only PGA. Thirty of the SNPs in the region were monomorphic in the CEU population, and a further 15 had a MAF<0.05. An additional five SNPs were excluded due to low genotype call rate. These SNPs were therefore not included in the LD analysis. A total of 456 SNPs within the candidate region satisfied the inclusion criteria and were included in the LD investigation and subsequent tSNP selection.

A graphical representation of the LD between these 456 SNPs in the *IL1* ligand cluster candidate region is shown in Figure 4.26, in which seven blocks of significant LD are evident. In general these blocks of high LD correlate to the different genes within the cluster. Block 1 spanned a total of 18.5kb, from 2.2kb downstream, to 4.8kb upstream, of *IL1A*. Block 2 covered 12.3kb from 6.7kb downstream, to within intron 3, of *IL1B*. LD block 3 included *IL1F7*, extending from 24.6kb upstream to 13.6kb downstream of the gene. Block 4 however included *IL1F9*, *IL1F6*, and up to intron 3 of *IL1F8*. SNPs within *IL1F8*, from exon

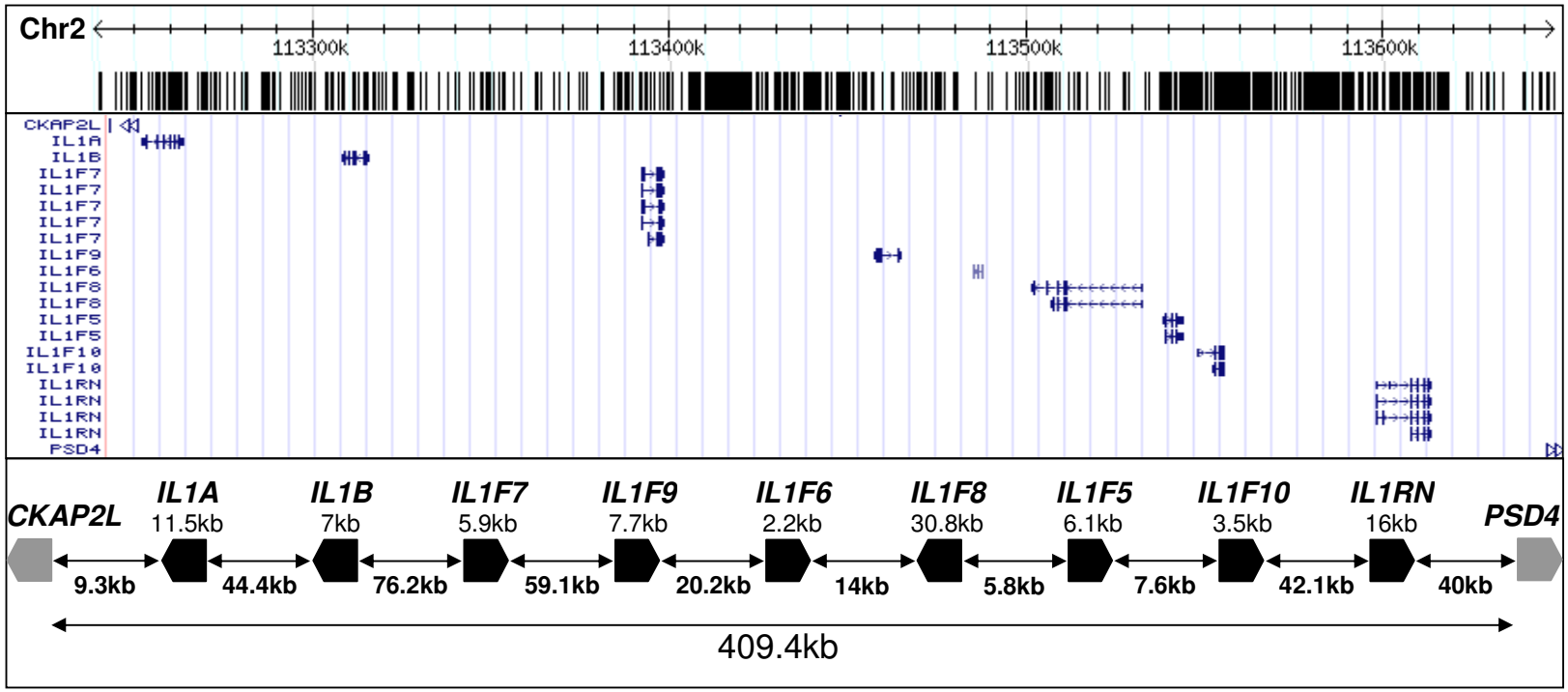


Figure 4.25 *IL1* ligand cluster candidate region

The boundaries of the *IL1* ligand cluster candidate region are the flanking genes *CKAP2L* and *PSD4*, determined by the July 2003 human genome reference sequence (NCBI build 34), hg16 annotation track, shown in the centre panel. The 5′ ends of both *CKAP2L* and *PSD4* are proximal to the gene cluster. The approximate (due to scale) positions within this region of all SNPs (irrespective of MAF) in the HapMap phase 2 database (data release #18/phaseII Sept05) are shown in the top panel. The bottom panel shows a schematic (not to scale) of the candidate region with gene orientation, size, and distances given. The genes of interest are shown in black and all other features within the region in grey. The largest isoform of each gene was used for size calculations, and, for clarity, is the only one shown. All sizes and distances are based on chromosomal positions reported in the NCBI build 34 genome reference sequence.

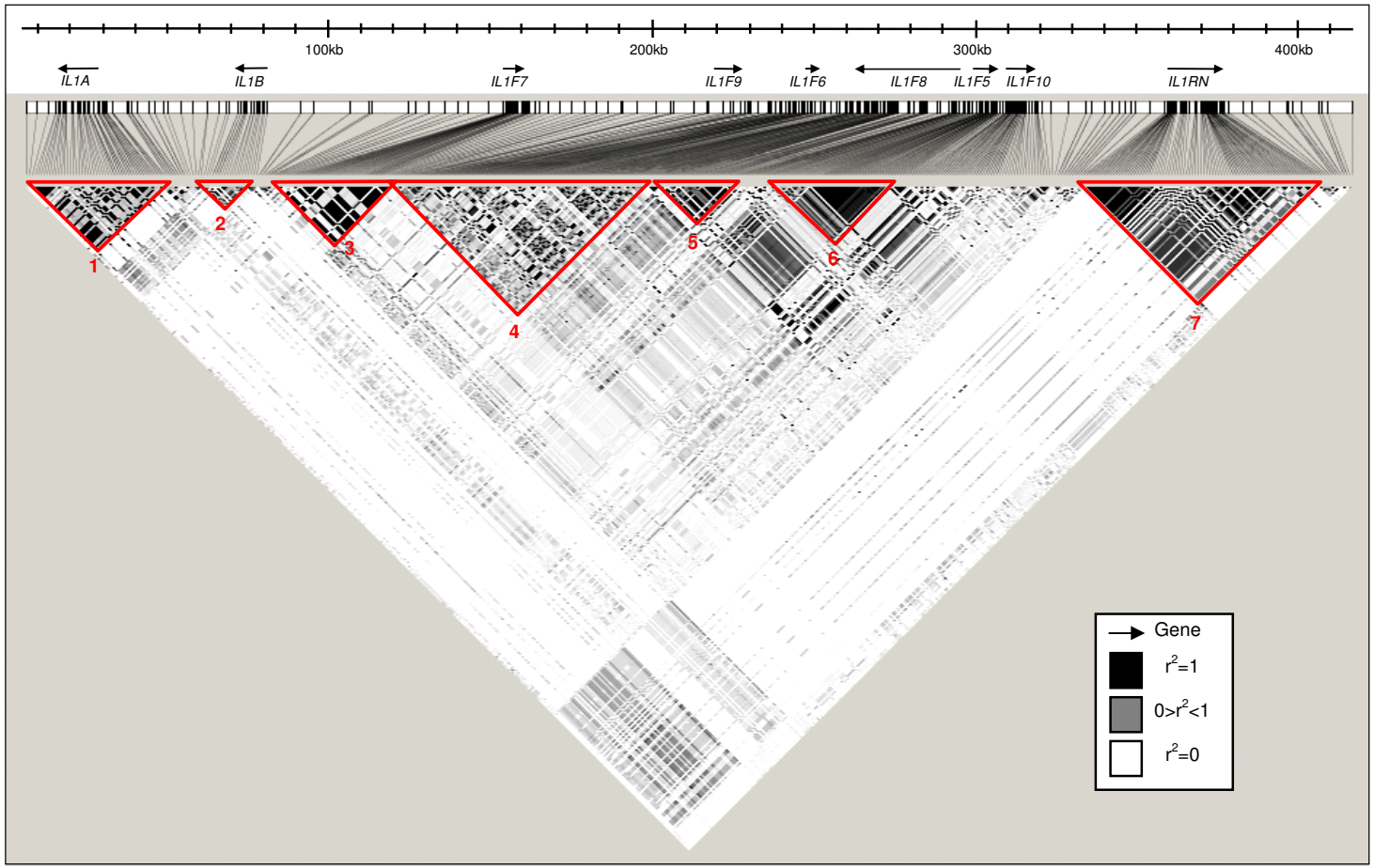


Figure 4.26 LD plot of the SNPs in the *IL1* ligand gene cluster region

The pairwise LD relationships between SNPs in the *IL1* ligand cluster candidate region were investigated to enable tSNP selection. The region shown is comprised of only the SNPs which satisfied the frequency criteria and were included in the LD analysis (n=456), not the whole of the candidate region. The gene orientation and location within the region is indicated. The position within the region of each SNP satisfying the selection criteria is represented by a line underneath the gene position. Each coloured square represents the pairwise LD between two SNPs. The graphical representation, generated in Haploview, is shown on a grey scale, with higher r^2 LD represented by darker blocks.

2 to intron 1, also separated into another LD block, the 15.1kb spanning block 5. SNPs within *IL1F5* also created a block of LD, block 6, which spanned 8.8kb from 3.8kb upstream of the gene to the 3′UTR. LD block 7 extended from 502bp upstream of *IL1RN* to 1.2kb downstream of the gene.

Due to the LD relationships between the SNPs a total of 91 tSNPs (20% of the total) were identified as the minimal subset required to capture the variation within all 456 SNPs. However, following assay design (section 3.1.2.1.) it was necessary to exclude a total of 16 SNPs, due to poor assay design scores. The final SNP set consisted of 93 tSNPs which captured the variation within 440 SNPs. Thirty six of the tSNPs were not in high LD ($r^2 \leq 0.8$) with any other SNP within the region and so were not tagging any additional SNPs. All of the selected tSNPs, and the SNPs captured by each, are listed in Table 4.25.

4.4.1.2. Stage-1

The genotyping assays for rs2708947, rs4849122, rs2278716, rs2234679, and rs3783512 were classed as failed (Figure 4.1 and Table 4.2). These SNPs, for which genotypes were not called and were therefore not included in the association analysis, were tagging a total of 45 SNPs. An additional four genotyped SNPs from the *IL1* ligand cluster candidate region showed significant deviation from HWE (Table 4.3), and were therefore also not included in the association analysis. These were: rs13005572 (SNP 60), rs1867829 (SNP 62), rs162319 (SNP 65), and rs419598 (SNP69), tagging a total of 27 SNPs. Of the 81 excluded SNPs (nine tSNPs and 72 captured SNPs) two were in sufficiently high LD ($r^2 \geq 0.8$) with two (SNPs 26 and 29) of the SNPs which were included in the analysis, that they were still captured. The remaining 79 of the SNPs however were not in high LD ($r^2 \geq 0.8$) with any of the analysed SNPs, and were therefore not captured by the association study (Table 4.26). A total of 377 SNPs were captured in the association study analysis.

Of the 93 analysed *IL1* ligand region tSNPs, nine showed significant ($p<0.05$) frequency differences between the patient (n=130) and control (n=146) cohorts. For ease of interpretation the results of the SNPs showing significant (Table 4.27) and non-significant frequency differences (Table 4.28) are presented separately. None of the SNPs satisfied the significance level following Bonferroni correction (0.05/93 SNPs analysed= $p< 0.0005$). However, Bonferroni correction assumes that the variables tested are independent, and is therefore overly conservative in this study as the SNPs tested will be in some amount of LD

SNP Nº	rs Nº	Tagged SNPs
1	rs6712572	n/a
2	rs2048874	n/a
3	rs17561	rs1894399,rs3783559,rs1878321,rs2856837,rs1878319,rs1304037,rs1878320,rs1516789,rs3783557,rs2071375, rs1800794,rs4848300,rs2856841,rs1516790,rs4848303,rs1878316,rs1040193,rs1800587,rs3783555,rs2856836, rs697,rs3783514,rs3783520
4	rs2071374	n/a
5	rs3783516	rs2856838,rs3783515,rs3783547,rs6746923
6	rs17042407	n/a
7	rs4849123	n/a
8	rs12469600	n/a
9	rs4849125	rs4849124
10	rs3917368	rs1143643,rs1143633
11	rs3917366	rs1071676,rs1143637,rs1143640,rs3917373,rs1143639
12	rs3917365	n/a
13	rs1143634	n/a
14	rs3136558	rs3136557,rs3917354
15	rs3917356	n/a
16	rs1143627	rs1143629,rs6735739,rs2708916,rs2708914
17	rs16944	n/a
18	rs2723167	rs4447608,rs2708919
19	rs12472089	n/a
20	rs2723154	n/a
21	rs4364030	rs10199311,rs2723163,rs3811042
22	rs2723168	PGAF6006535
23	rs3811046	n/a
24	rs3811048	rs11687740,rs3811045,rs6717710,rs3811047,rs4458215
25	rs2723187	n/a
26	rs2723192	rs2723179,rs2708940,rs2708957,rs2723156,rs2723180,rs2723177,rs2708924,rs2723169,rs2708959,rs2708925, rs2708943,rs2708954,rs2708956,rs2708960,rs2708965,rs28947188,rs2723196,rs2708962,rs7577574,rs2708958, rs28947201,rs2723189,rs2708953,rs2723170
27	rs3923566	rs7581834

SNP Nº	rs Nº	Tagged SNPs
28	rs879711	rs6747618,rs6758386,rs10177563,rs6761372,rs11690399
29	rs11695148	n/a
30	rs13392494	n/a
31	rs1562305	rs2305151,rs1867828
32	rs13033104	rs1084412,PGAF6014143,PGAF6014110,rs2305152,rs11695110,rs10496446,PGAF9003748,rs6723197,rs12614316
33	rs895497	rs1020404,rs1562306,rs1446514,rs1446515,rs1596894,rs1446512,rs1446511,rs7584409,rs4849136
34	rs1867827	n/a
35	rs1867831	rs11894902,rs1006122,rs6737824,rs6749299,rs4849141,rs6745709,rs6744874,rs9308681
36	rs4849140	rs1374284,rs6744288,rs12711746,rs6542108,PGAF6018828,rs13033574,rs11686530,rs6710007,rs12468503, rs4849142,rs7571656,rs13030063,rs1013477,rs6542109
37	rs7595507	n/a
38	rs12711747	rs11123155,rs4849139,rs2862772,rs11673918,rs2305150
39	rs3948121	n/a
40	rs2197578	rs6724667,rs7570483,rs4849143,rs34868654,rs4848312,rs4848311,rs7570159,rs11123156,rs3049690,rs7569284, rs4849144,rs1530554,rs11390487,rs12614012,rs2218557,rs7570058
41	rs12995447	n/a
42	rs6725997	rs11687786,rs12618939
43	rs1992761	rs11123159,rs34200521,rs35391805,rs17042721,rs34514967,rs17042750,rs28946269,rs11887249,rs28928270, rs1992762,rs11897709,rs28928294
44	rs1562302	rs34514967,rs34491153,rs12466269
45	rs2515403	rs2472188,rs1800930,rs2180235,rs2515401,rs3180234
46	rs7599662	n/a
47	rs2100071	rs6542110,rs1867832,rs2862853,rs6728769,rs1867833,rs55806606,rs2441374,rs2515404,rs1813048
48	rs921065	rs6758965,rs13407508
49	rs6755497	n/a
50	rs10206428	rs2515396,rs2515397,rs1530551,rs1374286,rs1530548,rs2515398,rs2515402,rs768627,rs3075085,rs10633611, rs2251872,rs2515394,rs2515391,rs6748448,rs1530550,rs2441376,rs3820738,rs4849146,rs6733859,rs2252007, rs2515395,rs2515399,rs957201
51	rs1446521	rs4849147
52	rs1867834	rs13019891
53	rs7570267	rs10165797
54	rs12711749	rs3811051
55	rs17042795	n/a

SNP Nº	rs Nº	Tagged SNPs
56	rs3811050	n/a
57	rs3811054	rs3811053
58	rs3827763	n/a
59	rs6743376	rs4145013
60	rs13005572	n/a
61	rs12468224	rs6759676,rs6761276,rs13386602,rs13389457,rs13398125
62	rs1867829	rs4849148,rs9308682
63	rs1542176	n/a
64	rs1688075	rs315925,rs28928309
65	rs162319	n/a
66	rs17207494	rs4496335,rs1446510,rs7574427
67	rs2592346	n/a
68	rs315934	n/a
69	rs419598	rs448341,rs432014,rs408392,rs442710,rs434792,rs1794068,rs447713,rs128964,rs495410,rs431726,rs1794065, rs444413,rs315931,rs495282,rs2853628,rs598859,rs4252008,rs423904,rs315936,rs392503,rs1794067,rs451578, rs446433,rs454078,rs3087263
70	rs2232354	n/a
71	rs380092	n/a
72	rs440286	rs579543,rs397211,rs9005,rs396201
73	rs315951	rs315953,rs315952,rs315951
74	rs4252042	rs4251991,rs973035,rs3213448,rs2071459,rs3087262,PGARN007044,rs928940,rs3087266,rs4252019,rs3181052, rs4252001
75	rs315949	n/a
76	rs315943	rs4251961,rs11123167,rs1374281
77	rs931471	rs12475781,rs6542118,rs315957
78	rs6739883	n/a
79	rs1867761	rs724496
80	rs4849159	n/a
81	rs6760120	n/a
82	rs35002769	n/a
83	rs34862832	rs13416441,rs3768769,rs6542107,rs28938790,rs28954071,rs34345011,rs17042691,rs6714534,rs10496447, rs28954071,rs28938791,rs17660913,rs13407838,rs28993969,rs28947170,rs35229022,rs13391978,rs11901453
84	rs36062386	n/a

SNP Nº	rs Nº	Tagged SNPs
85	rs28928312	n/a
86	rs28938782	rs28928288
87	rs28992498	rs2305148,rs1156701,rs2515405,rs2472189,rs17042751,rs1446522,rs6755354,rs28992497,rs1530549,rs17042709, rs3752739,rs2305149,rs996878,rs957200,rs996879,rs2515406
88	rs28928282	rs28928285
89	rs2708947	n/a
90	rs4849122	n/a
91	rs2278716	rs10165821,rs1374280
92	rs2234679	rs2234678,rs4251969,rs4251985,rs2234676,rs4251972,rs4251977,rs2234677,rs4251978,rs4251981,rs4251983, rs4251974,rs16065,rs878972,rs4251967,rs4251976,rs4251968,rs4251979,rs4251975,rs4251980,rs4251984, rs4251986,rs4251970
93	rs3783512	rs12711740,rs16109682,rs10496445,rs7567619,rs3783550,rs1969294,rs12612788,rs3783533,rs1533463,rs10496444, rs2071376,rs3783512,rs3783546,rs3783537,rs2071373,rs3783525,rs3783526,rs3783549,rs3783539,rs10167914, rs3783543
Uncaptured		rs3783548,rs1143625,rs34219867,rs36062386,rs28929172,rs28992496,rs12993057,rs1623119,rs315919,rs2637988, rs42592346,rs4251999,rs439154,rs1794066,rs1665190,rs452204

Table 4.25 *IL1* ligand cluster candidate region tSNPs

A total of 93 tSNPs were genotyped to capture 440 of the 456 SNPs within the candidate region. The SNPs captured ($r^2 \geq 0.8$) by each genotyped tSNP are shown. For ease of identification each tSNP genotyped was assigned a number according to genome order. Due to poor assay design scores it was not possible to genotype or capture 16 of the SNPs within the region.

SNP Nº	rs No	Tagging
26	rs2723192	rs2708947
64	rs1688075	rs1867829
Uncaptured		rs13005572,rs4849148,rs9308682,rs419598,rs448341,rs432014, rs408392,rs442710,rs434792,rs1794068,rs447713,rs128964,rs495410, rs431726,rs1794065,rs444413,rs315931,rs495282,rs2853628, rs598859,rs4252008,rs423904,rs315936,rs392503,rs1794067,rs451578, rs446433,rs454078,rs3087263,rs162319,rs4849122,rs2278716, rs10165821,rs1374280,rs2234679,rs2234678,rs4251969,rs4251985, rs2234676,rs4251972,rs4251977,rs2234677,rs4251978,rs4251981, rs4251983,rs4251974,rs16065,rs878972,rs4251967,rs4251976, rs4251968,rs4251979,rs4251975,rs4251980,rs4251984,rs4251986, rs4251970,rs3783512,rs12711740,rs16109682,rs10496445,rs7567619, rs3783550,rs1969294,rs12612788,rs3783533,rs1533463,rs10496444, rs2071376,rs3783546,rs3783537,rs2071373,rs3783525,rs3783526, rs3783549,rs10167914,rs3783543

Table 4.26 *IL1* ligand cluster additional tagging of analysis excluded SNPs

Nine of the genotyped tSNPs from the *IL1* ligand cluster candidate region were excluded from analysis due to poor genotyping quality or deviation from HWE. These nine SNPs were tagging 72 SNPs between them, making a total of 81 SNPs excluded from analysis. The excluded SNPs are shown, as well as the two captured ($r^2 \geq 0.8$) by two of the other genotyped SNPs.

with each other (Lewis, 2002).

The rare allele was present at a higher frequency in the patient than the control population for five SNPs (SNPs 4, 5, 7, 10, and 64) resulting in an OR>1. The common allele of the remaining SNPs was more frequent in the patient population. However, apart from SNP1 (p=0.00452), the frequency differences of these SNPs had a lower significance level (p>0.01) than the others. The highest OR (3.043) and smallest p-value (0.000899) were seen with SNP64, with the frequency of the rare allele at 0.115 in the case population and 0.0411 in the control population. The minor allele in the population was different for the cases and the controls for both SNP1 and SNP5. In the case population the SNP1 A allele, termed as allele 1, was the most common (frequency of 0.600), while in the control population this allele was only present at a frequency of 0.479, making the alternative C allele the more common. For SNP5 the G allele, termed allele 2, was the minor allele (0.485) in the cases but it was the more common allele in the control population (0.603).

The genoplots of the eight *IL1* ligand cluster SNPs showing significant frequency differences (Figure 4.27) showed that in all the SNP assays the DNA samples separated into three distinct clusters corresponding to each genotype. Although the 'AB' and 'BB' clusters in the SNP1 assay (A.) were in close proximity on the plot, all of the 'BB' cluster samples had normalised theta values of ~1, enabling discernment of the two separate clusters. In some of the other genoplots, especially for SNP4 (B.) the 'AB' samples were more widely dispersed than for either of the other genotypes. However, as all of the plots, apart from SNP1, showed well separated clusters it is possible to define them with confidence. The high quality of the genotyping assay was also reflected in the GC10 scores for the SNPs, all of which were >0.7 (SNP1=0.712, SNP4=0.909, SNP5=0.857, SNP7=0.895, SNP8=0.874, SNP10=0.798, SNP64=0.724, SNP80=0.928, SNP81=0.886) (Appendix 2). These quality measurements of the SNP assays enabled high confidence in the genotype assignment, and therefore that the significant differences observed are due to true frequency differences between the populations and not genotyping error.

To aid identification of secondary effects, the LD relationships between these nine SNPs were determined. As a tSNP approach to SNP selection was employed, none of the genotyped SNPs were in high LD with each other according to the r^2 statistic. However, r^2 is a measure of the LD relationship between alleles at different loci, but the alternative LD statistic D′ is a measure of the evidence of recombination between loci (section 3.1.1.1.). While r^2 is the most

SNP Nº	rs Number	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
1	rs6712572 (A/C)	1 2	0.600 (156) 0.400 (104)	0.479 (140) 0.521 (152)	0.614 (0.438-0.861)	0.00452
4	rs2071374 (A/C)	1 2	0.612 (158) 0.379 (100)	0.723 (211) 0.277 (81)	1.649 (1.152-2.359)	0.00606
5	rs3783516 (A/G)	1 2	0.515 (126) 0.485 (134)	0.397 (176) 0.603 (116)	1.614 (1.151-2.262)	0.00535
7	rs4849123 (A/G)	1 2	0.589 (153) 0.411 (107)	0.709 (207) 0.291 (85)	1.703 (1.196-2.425)	0.00301
8	rs12469600 (A/G)	1 2	0.788 (205) 0.212 (55)	0.709 (207) 0.291 (85)	0.653 (0.442-0.965)	0.0314
10	rs3917368 (G/A)	1 2	0.546 (142) 0.454 (118)	0.654 (191) 0.346 (101)	1.571 (1.115-2.215)	0.00963
64	rs1688075 (A/C)	1 2	0.885 (230) 0.115 (30)	0.959 (280) 0.0411 (12)	3.043 (1.524-6.079)	0.000899
80	rs4849159 (A/G)	1 2	0.869 (226) 0.131 (34)	0.805 (235) 0.195 (57)	0.620 (0.391-0.985)	0.0405
81	rs6760120 (A/G)	1 2	0.658 (171) 0.342 (89)	0.562 (164) 0.438 (128)	0.667 (0.472-0.942)	0.0209

Table 4.27 Significant *IL1* ligand cluster candidate region stage-1 analysis results

Eighty four genotyped tSNPs within the *IL1* ligand cluster candidate region were analysed for significant allele frequency differences between the patient (n=130) and control (n=146) cohorts. The frequency in each cohort of both the common (1) and rare (2) alleles are given. The alternative alleles of each SNP are shown underneath the rs identifier. The allele assigned as the common (1) allele is given first and the rarer (2) allele second. The OR shown is for allele 2 (allele 1 OR=1). The p-value for the test for frequency differences between the cohorts is also shown. The results for the nine tSNPs which showed significant ($p<0.05$) frequency differences between the two cohorts are shown.

SNP Nº	rs Number	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
2	rs2048874 (G/A)	1 2	0.869 (226) 0.131 (34)	0.880 (257) 0.120 (35)	1.105 (0.667-1.830)	0.6991
3	rs17561 (C/A)	1 2	0.746 (194) 0.254 (66)	0.702 (205) 0.298 (87)	0.802 (0.551-1.167)	0.247
6	rs17042407 (A/G)	1 2	0.665 (173) 0.335 (87)	0.736 (215) 0.264 (77)	1.404 (0.934-2.025)	0.0689
9	rs4849125 (A/G)	1 2	0.721 (186) 0.279 (72)	0.671 (196) 0.329 (96)	0.790 (0.548-1.139)	0.206
11	rs3917366 (C/A)	1 2	0.804 (209) 0.196 (51)	0.757 (221) 0.243 (71)	0.760 (0.506-1.140)	0.183
12	rs3917365 (G/A)	1 2	0.919 (239) 0.0808 (21)	0.914 (267) 0.0856 (25)	0.938 (0.512-1.720)	0.937
13	rs1143634 (G/A)	1 2	0.804 (209) 0.196 (51)	0.757 (221) 0.243 (71)	0.760 (0.506-1.140)	0.183
14	rs3136558 (A/G)	1 2	0.815 (212) 0.185 (48)	0.774 (226) 0.226 (66)	0.775 (0.511-1.176)	0.229
15	rs3917356 (A/G)	1 2	0.546 (142) 0.454 (118)	0.497 (145) 0.503 (147)	0.820 (0.586-1.146)	0.244
16	rs1143627 (A/G)	1 2	0.735 (191) 0.265 (69)	0.702 (205) 0.851 (87)	0.851 (0.587-1.235)	0.396
17	rs16944 (G/A)	1 2	0.735 (191) 0.265 (69)	0.702 (205) 0.298 (87)	0.851 (0.587-1.235)	0.396
18	rs2723167 (A/G)	1 2	0.558 (144) 0.442 (114)	0.500 (146) 0.500 (146)	0.792 (0.566-1.108)	0.173
19	rs12472089 (G/A)	1 2	0.769 (200) 0.231 (60)	0.719 (210) 0.281 (82)	0.768 (0.523-1.129)	0.179
20	rs2723154 (G/A)	1 2	0.846 (220) 0.154 (40)	0.843 (246) 0.157 (46)	0.972 (0.613-1.542)	0.905
21	rs4364030 (G/C)	1 2	0.477 (124) 0.523 (136)	0.551 (161) 0.449 (131)	1.348 (0.964-1.885)	0.0805
22	rs2723168 (G/A)	1 2	0.904 (235) 0.0962 (25)	0.890 (260) 0.110 (32)	0.864 (0.498-1.501)	0.604
23	rs3811046 (A/C)	1 2	0.709 (183) 0.291 (75)	0.678 (198) 0.322 (94)	0.863 (0.599-1.242)	0.428
24	rs3811048 (A/G)	1 2	0.712 (185) 0.288 (75)	0.675 (197) 0.325 (95)	0.841 (0.585-1.209)	0.348
25	rs2723187 (G/A)	1 2	0.904 (235) 0.0962 (25)	0.890 (260) 0.110 (32)	0.864 (0.498-1.501)	0.604
26	rs2723192 (G/A)	1 2	0.904 (235) 0.0962 (25)	0.890 (260) 0.110 (32)	0.864 (0.498-1.501)	0.604
27	rs3923566 (A/G)	1 2	0.804 (209) 0.196 (51)	0.786 (217) 0.214 (59)	0.898 (0.589-1.366)	0.614
28	rs879711 (C/A)	1 2	0.619 (161) 0.381 (99)	0.565 (165) 0.435 (127)	0.799 (0.568-1.123)	0.196
29	rs11695148 (A/G)	1 2	0.850 (221) 0.150 (39)	0.819 (239) 0.181 (53)	0.796 (0.506-1.251)	0.320
30	rs13392494 (G/A)	1 2	0.839 (218) 0.161 (42)	0.819 (239) 0.181 (53)	0.869 (0.557-1.355)	0.535

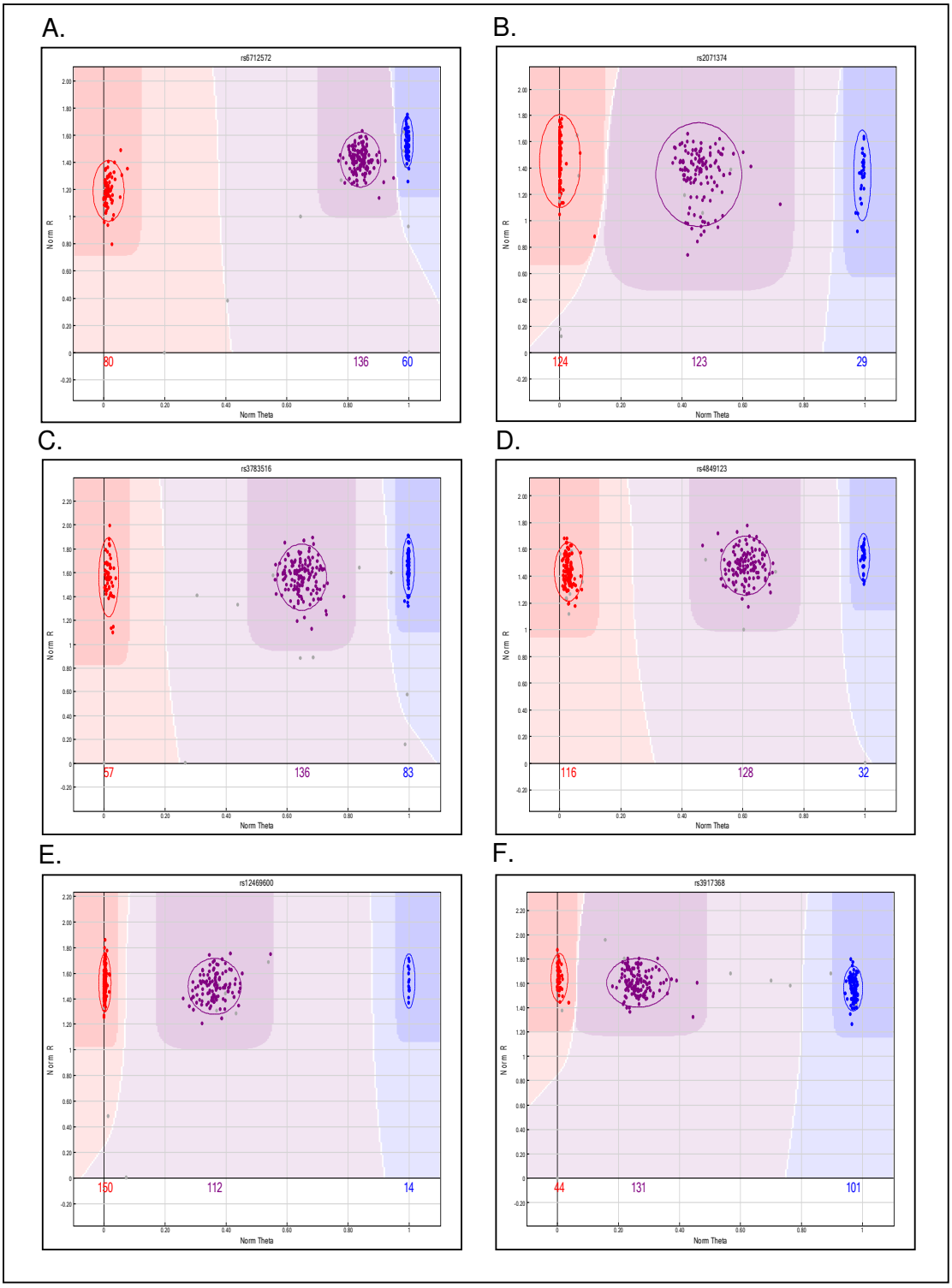
SNP Nº	rs Number	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
31	rs1562305 (G/A)	1 2	0.735 (191) 0.265 (69)	0.753 (220) 0.247 (72)	1.104 (0.753-1.619)	0.613
32	rs13033104 (A/C)	1 2	0.777 (202) 0.223 (58)	0.723 (211) 0.277 (81)	0.748 (0.507-1.103)	0.141
33	rs895497 (G/A)	1 2	0.646 (168) 0.354 (92)	0.705 (206) 0.295 (86)	1.312 (0.917-1.876)	0.137
34	rs1867827 (A/T)	1 2	0.908 (236) 0.0923 (24)	0.931 (272) 0.0685 (20)	1.383 (0.745-2.567)	0.303
35	rs1867831 (G/A)	1 2	0.648 (166) 0.352 (90)	0.726 (212) 0.274 (80)	1.437 (0.999-2.066)	0.0502
36	rs4849140 (G/C)	1 2	0.523 (136) 0.477 (124)	0.548 (160) 0.452 (132)	1.105 (0.790-1.545)	0.559
37	rs7595507 (G/A)	1 2	0.562 (146) 0.438 (114)	0.500 (146) 0.500 (146)	0.781 (0.558-1.092)	0.148
38	rs12711747 (G/A)	1 2	0.531 (138) 0.469 (122)	0.562 (164) 0.438 (128)	1.133 (0.809-1.585)	0.467
39	rs3948121 (G/C)	1 2	0.692 (180) 0.308 (80)	0.726 (212) 0.274 (80)	1.178 (0.815-1.702)	0.384
40	rs2197578 (A/G)	1 2	0.658 (171) 0.342 (89)	0.682 (199) 0.318 (93)	1.114 (0.781-1.589)	0.553
41	rs12995447 (G/A)	1 2	0.592 (154) 0.408 (106)	0.562 (164) 0.438 (128)	0.882 (0.629-1.237)	0.467
42	rs6725997 (G/A)	1 2	0.477 (124) 0.523 (136)	0.527 (154) 0.473 (138)	1.224 (0.876-1.710)	0.236
43	rs1992761 (G/A)	1 2	0.942 (245) 0.058 (15)	0.918 (268) 0.0821 (24)	0.684 (0.351-1.333)	0.259
44	rs1562302 (G/A)	1 2	0.792 (206) 0.208 (54)	0.839 (245) 0.161 (47)	1.366 (0.887-2.106)	0.157
45	rs2515403 (G/A)	1 2	0.615 (160) 0.385 (100)	0.600 (174) 0.400 (116)	0.938 (0.665-1.321)	0.712
46	rs7599662 (G/A)	1 2	0.671 (175) 0.327 (85)	0.682 (199) 0.318 (93)	1.039 (0.727-1.486)	0.833
47	rs2100071 (A/C)	1 2	0.673 (175) 0.327 (85)	0.681 (199) 0.319 (93)	1.039 (0.727-1.486)	0.833
48	rs921065 (A/G)	1 2	0.496 (128) 0.504 (130)	0.507 (148) 0.493 (144)	1.044 (0.747-1.459)	0.802
49	rs6755497 (A/G)	1 2	1.00 (260) 0.00 (0)	1.00 (292) 0.00 (0)	1.00 (1.00-1.00)	1.00
50	rs10206428 (G/A)	1 2	0.615 (160) 0.385 (100)	0.599 (175) 0.401 (117)	0.935 (0.664-1.316)	0.699
51	rs1446521 (A/G)	1 2	0.680 (174) 0.320 (82)	0.682 (199) 0.318 (93)	1.008 (0.704-1.445)	0.964
52	rs1867834 (G/A)	1 2	0.557 (145) 0.442 (115)	0.524 (153) 0.476 (139)	0.873 (0.624-1.221)	0.427
53	rs7570267 (A/G)	1 2	0.539 (140) 0.461 (120)	0.521 (152) 0.479 (140)	0.931 (0.666-1.301)	0.674
54	rs12711749 (A/G)	1 2	0.589 (153) 0.411 (107)	0.637 (186) 0.363 (106)	1.227 (0.871-1.730)	0.243
55	rs17042795 (A/G)	1 2	0.992 (258) 0.0080 (2)	0.990 (289) 0.0102 (3)	0.747 (0.124-4.505)	0.748

SNP Nº	rs Number	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
56	rs3811050 (G/A)	1 2	0.835 (217) 0.165 (43)	0.819 (239) 0.181 (53)	0.894 (0.574-1.391)	0.618
57	rs3811054 (G/A)	1 2	0.961 (246) 0.039 (10)	0.949 (260) 0.0510 (14)	0.755 (0.329-1.731)	0.504
58	rs3827763 (G/A)	1 2	0.689 (179) 0.311 (81)	0.719 (210) 0.281 (82)	1.159 (0.804-1.671)	0.430
59	rs6743376 (A/C)	1 2	0.711 (185) 0.289 (75)	0.647 (189) 0.353 (103)	0.744 (0.519-1.066)	0.106
61	rs12468224 (C/G)	1 2	0.581 (151) 0.419 (109)	0.596 (174) 0.404 (118)	1.064 (0.758-1.495)	0.719
63	rs1542176 (A/G)	1 2	0.481 (125) 0.519 (135)	0.514 (150) 0.486 (142)	1.141 (0.817-1.594)	0.440
66	rs17207494 (A/C)	1 2	0.589 (153) 0.411 (107)	0.592 (173) 0.408 (119)	1.017 (0.724-1.428)	0.924
67	rs2592346 (A/C)	1 2	0.608 (158) 0.392 (102)	0.637 (186) 0.363 (106)	1.133 (0.802-1.599)	0.478
68	rs315934 (A/G)	1 2	0.859 (220) 0.141 (36)	0.803 (233) 0.197 (57)	0.669 (0.424-1.055)	0.0814
70	rs2232354 (A/C)	1 2	0.765 (199) 0.235 (61)	0.808 (236) 0.192 (56)	1.292 (0.858-1.944)	0.219
71	rs380092 (T/A)	1 2	0.723 (188) 0.277 (72)	0.685 (200) 0.315 (92)	0.833 (0.577-1.202)	0.327
72	rs440286 (C/A)	1 2	0.739 (192) 0.262 (68)	0.771 (225) 0.229 (67)	1.189 (0.807-1.754)	0.382
73	rs315951 (C/G)	1 2	0.761 (198) 0.239 (62)	0.733 (214) 0.267 (78)	0.859 (0.584-1.263)	0.439
74	rs4252042 (G/A)	1 2	0.911 (235) 0.0892 (23)	0.877 (256) 0.123 (36)	0.696 (0.401-1.209)	0.195
75	rs315949 (G/A)	1 2	0.542 (141) 0.458 (119)	0.545 (159) 0.455 (133)	1.009 (0.721-1.411)	0.958
76	rs315943 (A/G)	1 2	0.542 (141) 0.458 (119)	0.545 (159) 0.455 (133)	1.009 (0.721-1.411)	0.958
77	rs931471 (A/G)	1 2	0.700 (182) 0.300 (78)	0.699 (204) 0.301 (88)	0.994 (0.690-1.430)	0.972
78	rs6739883 (G/A)	1 2	0.819 (213) 0.181 (47)	0.798 (233) 0.202 (59)	0.871 (0.569-1.334)	0.526
79	rs1867761 (A/G)	1 2	0.781 (203) 0.219 (57)	0.771 (225) 0.229 (67)	0.943 (0.632-1.408)	0.774
82	rs35002769 (T/A)	1 2	0.966 (251) 0.0346 (9)	0.932 (272) 0.684 (20)	0.488 (0.218-1.091)	0.0707
83	rs34862832 (G/A)	1 2	0.873 (227) 0.127 (33)	0.836 (244) 0.164 (48)	0.739 (0.458-1.193)	0.213
84	rs36062386 (G/A)	1 2	0.942 (245) 0.0577 (15)	0.962 (279) 0.0380 (11)	1.553 (0.700-3.445)	0.276
85	rs28928312 (A/G)	1 2	0.785 (204) 0.215 (56)	0.825 (241) 0.175 (51)	1.297 (0.850-1.980)	0.227
86	rs28938782 (A/G)	1 2	0.992 (258) 0.0077 (2)	0.990 (289) 0.0102 (3)	0.747 (0.124-4.505)	0.748
87	rs28992498 (C/A)	1 2	0.888 (229) 0.112 (29)	0.894 (261) 0.106 (31)	1.066 (0.624-1.823)	0.815

SNP Nº	rs Number	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
88	rs28928282 (G/A)	1 2	0.914 (234) 0.086 (22)	0.924 (268) 0.0759 (22)	1.145 (0.618-2.121)	0.666

Table 4.28 Non-significant *IL1* ligand cluster candidate region stage-1 analysis results

Eighty four genotyped tSNPs within the *IL1* ligand cluster candidate region were analysed for significant allele frequency differences between the patient (n=130) and control (n=146) cohorts. The frequency in each cohort of both the common (1) and rare (2) alleles are given. The alternative alleles of each SNP are shown underneath the rs identifier. The allele assigned as the common (1) allele is given first and the rarer (2) allele second. The OR shown is for allele 2 (allele 1 OR=1). The p-value for the test for frequency differences between the cohorts is also shown. The results for the 75 tSNPs which did not show significant ($p>0.05$) frequency differences between the two cohorts are shown.


A.
rs6712572
80
136
60
B.
rs2071374
124
123
29
C.
rs3783516
57
136
83
D.
rs4849123
116
128
32
E.
rs12469600
150
112
14
F.
rs3917368
44
131
101
Norm R
Norm Theta

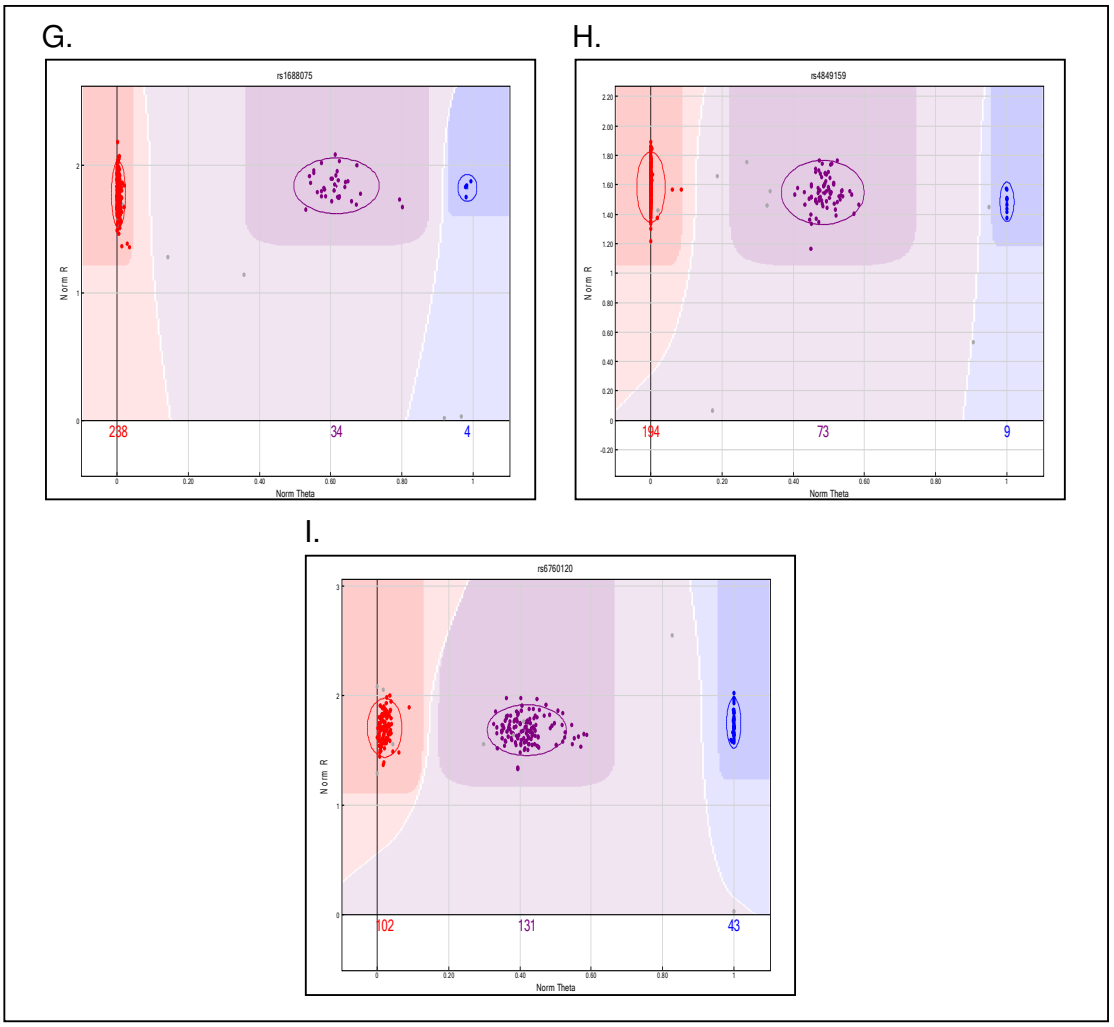


Figure 4.27 Genoplots of the significant *IL1* ligand cluster region SNPs

Shown are the genoplots of the nine *IL1* ligand cluster region SNPs which showed evidence of significant frequency differences in the initial stage-1 analysis. Normalised theta values (deviation from total 'A' allele signal) are plotted along the x-axes, and normalised R values (signal intensity) along the y-axes. Each data point represents one sample, the circles indicate the location of each genotype cluster, and the dark shaded areas are the call areas.

A. SNP1 rs6712572, B. SNP4 rs2071374, C. SNP5 rs3783516, D. SNP7 rs4849123, E. SNP8 rs12469600, F. SNP10 rs3917368, G. SNP64 rs1688075, H. SNP80 rs4849159, I. SNP81 rs6760120.

appropriate measurement in identifying SNPs which are proxies for each other, D′ is important when indentifying SNPs which are inherited together.

The D′ LD between the nine SNPs with significant allele frequency differences between cohorts in the initial analysis is shown in Figure 4.28. The nine SNPs fell into three clusters according to their LD patterns. One LD cluster was comprised of SNPs 1, 4, 5, 7, 8 and 10. Except for SNP1 with SNP8 (D′=0.36), and SNP4 with SNP10 (D′=0), these six SNPs were in notable LD (D′≥0.68) with each other. SNPs 80 and 81 were also in high LD (D′=0.95) with each other and constituted a second block of LD. SNP64 was considered as a separate cluster.

During conditional analysis of the SNPs within the LD blocks (Table 4.29) SNP8 was identified as a secondary effect and therefore removed from further analysis. The association effect demonstrated by SNP8 during the initial analysis was eliminated when each of the other SNPs were included in the baseline model. However, each of the other SNPs, apart from SNP5 which nearly reached significance (p=0.05), remained significant after SNP8 had been taken into account. When these SNPs were analysed conditioning on each other, SNPs 1, 4, 5, 7, and 10 no longer showed a significant association. This indicated that these SNPs were all reflecting one single effect, but because of the LD between them it was not possible through conditional analysis to elucidate with certainty which was responsible for the primary effect. It was therefore not possible to eliminate any more of the SNPs within this cluster. This was also seen when SNP80 and SNP81 were analysed conditioning on each other, where neither remained significant. This indicated that these two SNPs were also both reflecting the same effect, but as it could not be demonstrated that one was secondary to the other, neither could be eliminated.

Further analysis was then performed to determine if the associations in each of the three clusters were independent from each other. Each of SNPs 1, 4, 5, 7, and 10 showed a significant difference between the cohorts after the effect of SNP64 and of SNPs 80 and 81 had been accounted for (Table 4.30). This indicated that the disease association shown by these SNPs was independent to any other effects in the candidate region. The effect of SNP64 also remained significant when analysed in addition to each of the other SNPs, and was therefore also an independent effect within the candidate region. When SNPs 80 and 81 were analysed conditioning on each of SNPs 4, 5, 7, and 10, they did not reach, but were tending towards, significance (p<0.096). Both SNPs did however show a significant effect

conditioning on SNP1 (p=0.0369 and 0.0481 respectively). They were also tending towards significance when conditioned on the other SNPs in that cluster, and remained significant when analysed conditioning on SNP64. The evidence for SNP80 and SNP81 being an independent effect was therefore stronger than for them being secondary to the effects of SNPs 4, 5, 7, and 10.

Therefore neither SNP80 nor SNP81 were eliminated following the conditional analysis. The conditional analysis therefore suggested that three independent disease associations had been identified in the *IL1* ligand cluster candidate region.

4.4.1.3. Stage-2

Eight of the *IL1* ligand cluster candidate region SNPs which showed association in stage-1 of the analysis were included in stage-2 of the study. These were: SNPs 1, 4, 5, 7, 10, 64, 80, and 81. In stage-2 none of the SNPs showed any significant ($p<0.05$) frequency differences between the case (n=105) and control (n=184) cohorts used in stage-2 (Table 4.31).

4.4.1.3.1. Stratified analysis

Following stratified analysis of the data from both stages of the study, according to study stage, three of the eight *IL1* ligand cluster SNPs showed evidence of significant frequency differences between the patient and control cohorts (Table 4.32). The frequency of the rare allele of SNP1 was significantly lower in the patient population than the control population (CMH p=0.0256), and significantly higher in the patient populations for SNP4 (CMH p=0.0025) and SNP 64 (CMH p=0.00234). None of these three SNPs showed evidence of significant heterogeneity of odds between the populations used in two study stages (BD p=0.0763, 0.411, and 0.0976 respectively). SNP64 showed the most significant frequency difference, and the highest OR (2.040). Apart from SNP81 (BD p=0.0877) all of the other SNPs with non-significant CMH p values showed significant heterogeneity between the study stages (BD p= 0.0142, 0.00471, 0.0152, and 0.0218). This indicates that allele frequency differences were only present between the populations used in stage-1 of the study, and not those used in stage-2.

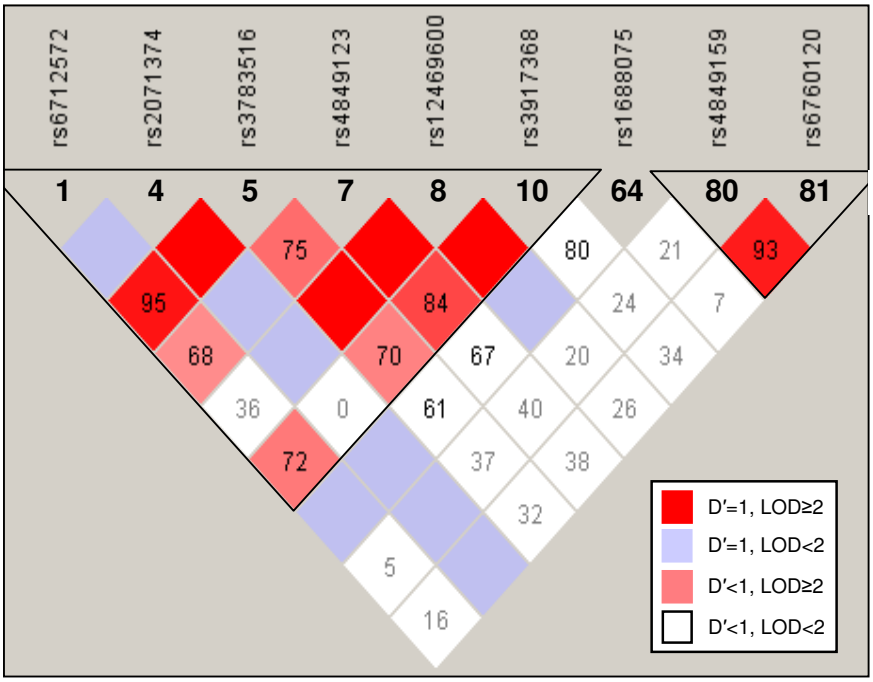


Figure 4.28 D′ relationship of associated *IL1* ligand cluster region SNPs

In the initial analysis nine SNPs within the *IL1* ligand cluster candidate region showed significant allele frequency differences between the case and control populations. The graphical representation of the D′ LD relationship between these nine SNPs, generated in Haploview, is shown on a sliding colour scale from white (low D′) to red (high D′), depending on LOD score. The D′ value for each pairwise comparison is shown, where no value is indicated D′=1. The LD blocks used for the conditional analysis are indicated.

LOD= logarithm of odds

Test marker	Conditional marker	p value
LD cluster SNPs 1, 4, 5, 7, 8, and 10		
1	4	0.143
1	5	0.325
1	7	0.321
1	8	0.0168
1	10	0.129
4	1	0.240
4	5	0.344
4	7	0.568
4	8	0.0387
4	10	0.286
5	1	0.467
5	4	0.294
5	7	0.611
5	8	0.0524
5	10	0.202
7	1	0.222
7	4	0.157
7	5	0.362
7	8	0.0234
7	10	0.101
8	1	0.0932
8	4	0.283
8	5	0.425
8	7	0.324
8	10	0.377
10	1	0.255
10	4	0.566
10	5	0.343
10	7	0.289
10	8	0.134
LD cluster SNPs 80 and 81		
80	81	0.175
81	80	0.260

Table 4.29 Conditional analysis of associated *IL1* ligand region SNPs within LD clusters

In the initial analysis nine SNPs within the *IL1* ligand cluster candidate region showed significant allele frequency differences between the case and control populations. Based on the D′ LD between these SNPs (Figure 4.28) it was identified that the all of the SNPs except SNP64 fell into two LD clusters (SNPs 1, 4, 5, 7, 8, and 10, and SNPs 80 and 81). The SNPs within each LD cluster were tested for significant frequency differences conditioning on the other SNPs within that cluster.

Test marker	Conditional marker	p value
1	64	0.00921
1	80	0.0127
1	81	0.0352
4	64	0.0122
4	80	0.0181
4	81	0.0154
5	64	0.00346
5	80	0.0223
5	81	0.0182
7	64	0.00709
7	80	0.00874
7	81	0.0322
10	64	0.0142
10	80	0.0476
10	81	0.0340
64	1	0.00391
64	4	0.00380
64	5	0.00327
64	7	0.00457
64	10	0.00340
64	80	0.000341
64	81	0.00193
80	1	0.0369
80	4	0.0540
80	5	0.0957
80	7	0.0510
80	10	0.0587
80	64	0.00115
81	1	0.0481
81	4	0.0669
81	5	0.0755
81	7	0.0732
81	10	0.0769
81	64	0.0331

Table 4.30 Conditional analysis of associated *IL1* ligand region SNPs between LD clusters

The effects indentified within each LD cluster were tested to determine if they were independent of each other. The SNPs within each LD cluster, and SNP64 which was not within a cluster, were tested for significant frequency differences conditioning on the SNPs within the other clusters.

SNP	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
1	1 2	0.524 (108) 0.476 (98)	0.512 (176) 0.488 (168)	0.951 (0.673-1.343)	0.774
4	1 2	0.654 (132) 0.347 (70)	0.715 (253) 0.285 (101)	1.328 (0.917-1.924)	0.134
5	1 2	0.617 (121) 0.383 (75)	0.585 (207) 0.415 (147)	0.873 (0.611-1.248)	0.455
7	1 2	0.714 (147) 0.286 (59)	0.669 (241) 0.331 (119)	0.813 (0.559-1.181)	0.275
10	1 2	0.668 (135) 0.332 (67)	0.631 (227) 0.369 (133)	0.847 (0.589-1.217)	0.369
64	1 2	0.916 (185) 0.084 (17)	0.937 (343) 0.063 (23)	1.370 (0.714-2.630)	0.347
80	1 2	0.757 (153) 0.243 (49)	0.800 (291) 0.200 (73)	1.277 (0.846-1.927)	0.247
81	1 2	0.575 (115) 0.425 (85)	0.581 (201) 0.419 (145)	1.025 (0.720-1.457)	0.893

Table 4.31 *IL1* ligand candidate region stage-2 analysis results

The eight *IL1* ligand candidate region SNPs identified as showing significant frequency differences following conditional analysis in stage-1 of the association study, were genotyped in the stage-2 patient (n=105) and control (n=184) cohorts. The frequency in each cohort of both the common (1) and rare (2) alleles of each SNP are given. The OR shown is for allele 2 (allele 1 OR=1). The p-value for the test for frequency differences between the cohorts is also shown.

SNP	CMH p value	OR (95% CI)	BD p value
1	0.0256	0.760 (0.597-0.967)	0.0763
4	0.00250	1.486 (1.149-1.922)	0.411
5	0.131	1.206 (0.946-1.539	0.0142
7	0.162	1.198 (0.930-1.543)	0.00471
10	0.208	1.172 (0.915-1.500)	0.0152
64	0.00234	2.040 (1.285-3.239)	0.0976
80	0.595	0.921 (0.679-1.250)	0.0218
81	0.119	0.823 (0.643-1.052)	0.0877

Table 4.32 *IL1* ligand cluster candidate region stratified analysis results

For the eight SNPs investigated in both stages of the study, association meta-analysis of the data was performed using the CMH test, stratifying by study stage. A significant result for the CHM test shows that the SNP is associated with disease. The p value and OR for this test are given, the OR shown is for allele 2 (allele 1 OR=1). A BD test for homogeneity of odds ratios between the two stages was also performed. A significant BD test shows that the allele frequency difference is more prominent in one of the populations studied. Any association is therefore more likely to be due to sampling error rather than true association.

BD-Breslow-Day, CI-Confidence Interval, CMH-Cochran-Mantel-Haenszel, OR-Odds Ratio

4.4.1.4. WTCCC control cohort analysis

An additional analysis with greater statistical power to detect SNPs associated with disease was performed with the pooled stage-1 and stage-2 cases (n=235), and the WTCCC2 cohort of healthy UK individuals (n=4,671).

Genotyping data was not available in the WTCCC2 control population for *IL1* ligand cluster SNP1 (rs6712572), and SNP64 (rs1688075). In the analysis of SNP1 in the patient cohort, rs4848298 in the WTCCC2 control cohort was used as a proxy, (r^2=1) according to HapMap3 (Table 4.35). In the analysis of SNP64 in the patient cohort, rs28928309 in the WTCCC2 control cohort was used as a proxy. These two SNPs were in high LD with each other according to both the HapMap2 (r^2=1) and HapMap3 (r^2=1) population datasets (section 4.1.5.).

The findings of the stratified analysis that SNP1, SNP4, and SNP64 were associated with disease were confirmed in the analysis of all the patients and the WTCCC2 controls (Table 4.33). The odds ratios seen in the stratified analysis, and the analysis with the WTCCC2 control cohort were very similar for both SNP1 (1.48 and 1.574 respectively) and SNP4 (0.76 and 0.74). Despite this, greater statistical significance were achieved in the analysis with the larger WTCCC2 control cohort for both SNP1 (p=0.00205) and SNP4 (p=0.00000756), compared to the stratified analysis using the smaller control cohort (CMH p=0.0256 and p=0.0025 respectively). In contrast, the OR and p value of the disease association of SNP64 were higher when the WTCCC2 control cohort was used. In the stratified analysis SNP64 showed the highest OR (2.04) and the lowest p value (p=0.00234). However, in the analysis with the WTCCC2 controls, SNP64 only just satisfied the significance level cut off (p=0.0427), with an OR of 1.398. This discrepancy may be due to sampling error in the smaller cohorts used in stage-1 and 2 of the study: such that, due to chance, the rare allele of this SNP was observed at a lower frequency than is present in the control population. It is important to note that as SNP64 had not been included in the WTCCC2 genotyping project it was necessary to use rs28928309 in this analysis as a proxy for SNP64 in the WTCCC2 population. It therefore possible that, although these two SNPs were in complete LD (r^2=1) according to the HapMap2 and HapMap3 datasets, rs28928309 is not in fact a perfect proxy for SNP64. As SNP64 was not included in the WTCCC2 project it was not possible to confirm the LD relationship between these two SNPs based on the genotyping data from this

SNP	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
1	1	0.567 (264)	0.494 (4607)	0.745 (0.618-0.899)	0.00205
	2	0.433 (202)	0.506 (4729)		
4	1	0.630 (290)	0.729 (6801)	1.574 (1.295-1.912)	0.00000756
	2	0.370 (170)	0.271 (2533)		
5	1	0.542 (247)	0.603 (5630)	1.286 (1.065-1.553)	0.00930
	2	0.458 (209)	0.397 (3704)		
7	1	0.644 (300)	0.695 (6492)	1.26 (1.037-1.53)	0.0215
	2	0.356 (166)	0.305 (2852)		
10	1	0.560 (277)	0.641 (5978)	1.19 (0.984-1.441)	0.0751
	2	0.400 (185)	0.359 (3354)		
64	1	0.898 (415)	0.925 (8627)	1.398 (1.024-1.908)	0.0427
	2	0.102 (47)	0.075 (699)		
80	1	0.820 (379)	0.773 (7222)	0.747 (0.587-0.953)	0.0155
	2	0.180 (83)	0.227 (2116)		
81	1	0.622 (286)	0.591 (5518)	0.877 (0.724-1.064)	0.182
	2	0.378 (174)	0.410 (3826)		

Table 4.33 *IL1* ligand candidate region analysis results with WTCCC2 controls

As an additional analysis with a larger, independent control cohort, allele frequencies in the pooled cases from both stages of the association study (n=235) were compared to the controls genotyped in phase-2 of the WTCCC project (n=4671). The frequency in each cohort of both the common (1) and rare (2) alleles of the two individual SNPs are given. The OR shown is for allele 2 (allele 1 OR=1). The p-value for the test for frequency differences between the cohorts is also shown.

larger population.

In contrast to the results of the stratified analysis, SNPs 5, 7, and 80, which were not significantly associated with disease in the stratified analysis, also showed evidence of disease association when analysed with the WTCCC2 control cohort. As with SNP1 and SNP4, the ORs seen in the stratified analysis and the analysis with the WTCCC2 control cohort, were very similar for both SNP5 (1.206 and 1.286 respectively) and SNP7 (1.198 and 1.26 respectively).

It had previously been demonstrated that SNPs 1, 4, 5, and 7 formed an LD cluster (Figure 4.28). The conditional analysis performed following the initial stage-1 analysis was therefore repeated to elucidate the SNPs showing primary and secondary effects (Table 4.34). As in the previous conditional analysis of these SNPs, none showed a significant difference between the case and control cohorts once the effect of any of the others had been taken into account, suggesting that these four SNPs are not demonstrating independent effects. It is most likely that only one SNP is associated with sJIA, possibly SNP4 which is the most statistically significant, but due to the D' relationship between these SNPs it is not possible to distinguish between them based on genotype frequency. Demonstration of a functional role leading to a biological effect influencing disease susceptibility would be necessary to determine which of these SNPs is the causal SNP responsible for the association observed.

4.4.1.5.Associated SNPs

In the initial tSNP selection based on the HapMap 2 genotyping data (Table 4.25) SNPs 1, 4, 7, and 80 were not tagging any other SNPs within the candidate region. SNP5 was tagging four other SNPs with r^2=1 based on the HapMap2 and HapMap3 data. SNP64 was tagging two other SNPs, and rs315925, with r^2=1. SNP64 had not been included in the WTCCC project, and rs28928309 was not included in HapMap3 so the LD relationship could not be validated. However, based on the HapMap2 dataset the r^2=1 relationship between SNPs64 and rs28928309 was confirmed.

4.4.1.5.1. Additional captured SNPs

Based on the HapMap3 dataset 12 SNPs were identified by SNAP (section 3.1.5.) as being in r^2>0.8 with SNP1, three SNPs with SNP7, and four additional SNPs with SNP64 (Table4.35).

Test marker	Conditional marker	p value
1	4	0.0591
1	5	0.228
1	7	0.190
4	1	0.111
4	5	0.198
4	7	1
5	1	0.534
5	4	0.254
5	7	0.823
7	1	0.126
7	4	1
7	5	0.0769

Table 4.34 Conditional analysis of associated *IL1* ligand region SNPs with WTCCC2 controls

Four SNPs within the *IL1* ligand cluster candidate region, previously shown to be in high the D′ LD with each other (Figure 4.28) showed significant allele frequency differences between the case and WTCCC2 control populations. These four SNPs were tested for significant frequency differences, conditioning on each of the other SNPs.

SNP4 was not included in HapMap3 and so additional tagged SNPs could not be identified. No additional tagged SNPs were identified for SNPs 5 and 10. All SNP1 tagged SNPs are within or downstream of the *IL1A* flanking gene *CKAP2L* (section 4.4.1.), and all of the SNP64 tagged SNPs are either down- or up-stream of *IL1RN*.

4.4.1.5.2. Associated SNP positions

The positions of the *IL1* ligand cluster SNPs showing significant evidence of association, and the SNPs which they are tagging, are shown in Figure 4.29 A. SNP1 is 10,814bp downstream of the 3′ end of *IL1A*, and SNP4 is within intron 4 of *IL1A*. None of the SNPs identified as being in high LD with SNP1 based on HapMap3 (Table 4.35) are indicated in the figure as they are all outside of the selected candidate region within the flanking gene *CKAP2L*. SNP5 is 2,707bp upstream of the transcription start site (TSS) of *IL1A*, and the SNPs it is tagging are at -2,863 (rs3783515), -10,205 (rs6746923) within introns 6 (rs3783547), and 3 (rs2856838) of *IL1A*. SNP7 is 18,298 downstream of the 3' end of *IL1B*, and -26,068bp upstream of *IL1A* TSS. The SNPs tagged by SNP7 are: (rs7585707) 17,992bp downstream *IL1B* and 26,347bp upstream *IL1A*, (rs17598291) 17,094bp downstream *IL1B* and 27,272bp upstream *IL1A*, and (rs62157435) 12,483bp downstream *IL1B* and 31,883bp upstream *IL1A*. Relative to *IL1RN* TSS SNP64 is at -17,274. The SNP64 captured SNPs are at: -40,731 (rs28928309), -36,184 (rs4848314), -25,484 (rs12475887), -10,103 (rs315925), -6,063 (rs315929), and -5,493 (rs315932) relative to *IL1RN* TSS. As well as being upstream of *IL1RN,* SNP64 and the SNPs it captures are also downstream of *IL1F10*. Relative to the 3′ end of *IL1F10* the SNPs are at: +1,312 (rs28928309), +5,859 (rs4848314), +16,559 (rs12475887), +24,769 (SNP64), +31,940 (rs315925), +35,980 (rs315929), and +36,550 (rs315932). SNP80 is located 33,466bp downstream of the 3′ end of *IL1RN* and 6,499bp upstream of *PSD4* TSS.

The location of SNP1, SNP4, SNP5, and three of the SNPs tagged by SNP5, relative to *IL1A* are shown in Figure 4.29 B.

4.4.1.6. Summary

Following analysis of the data from both study stages, three SNPs within the *IL1* ligand cluster candidate region, rs6712572 (SNP1), rs2071374 (SNP4), and rs1688075 (SNP64), were identified as showing evidence of significant association with sJIA. The significant

SNP	r^2	Location
SNP1 (rs6712572)		
rs908551	0.818	downstream *CKAP2L*
rs13015273	1	downstream *CKAP2L*
rs13016817	0.967	downstream *CKAP2L*
rs12622683	0.967	downstream *CKAP2L*
rs13013602	1	downstream *CKAP2L*
rs17649681	1	downstream *CKAP2L*
rs13019249	1	downstream *CKAP2L*
rs3811040	1	*CKAP2L* missense
rs3827761	1	*CKAP2L*
rs13026263	1	*CKAP2L*
rs4848298	1	*CKAP2L*
rs10200192	1	*CKAP2L*
SNP7 (rs4849123)		
rs7585707	1	upstream *IL1A,* downstream *IL1B*
rs17598291	1	upstream *IL1A,* downstream *IL1B*
rs6215435	1	upstream *IL1A,* downstream *IL1B*
SNP64 (rs1688075)		
rs4848314	1	upstream *IL1F10,* downstream *IL1RN*
rs12475887	1	upstream *IL1F10,* downstream *IL1RN*
rs315929	1	upstream *IL1RN*
rs315932	1	upstream *IL1RN*

Table 4.35 Additional SNPs tagged by *IL1* ligand cluster SNPs

Based in the HapMap3 population genotyping data all SNPs in r^2>0.8 with, and within 500kb of, the *IL1* ligand cluster SNPs showing evidence of significant disease association, were identified. The SNP identifier, r^2 value, and location of all the identified SNPs are shown. SNP4 was not included in HapMap3 so additional tagged SNPs could not be identified. SNP80 was not in r^2>0.8 with any SNPs within 500kb.

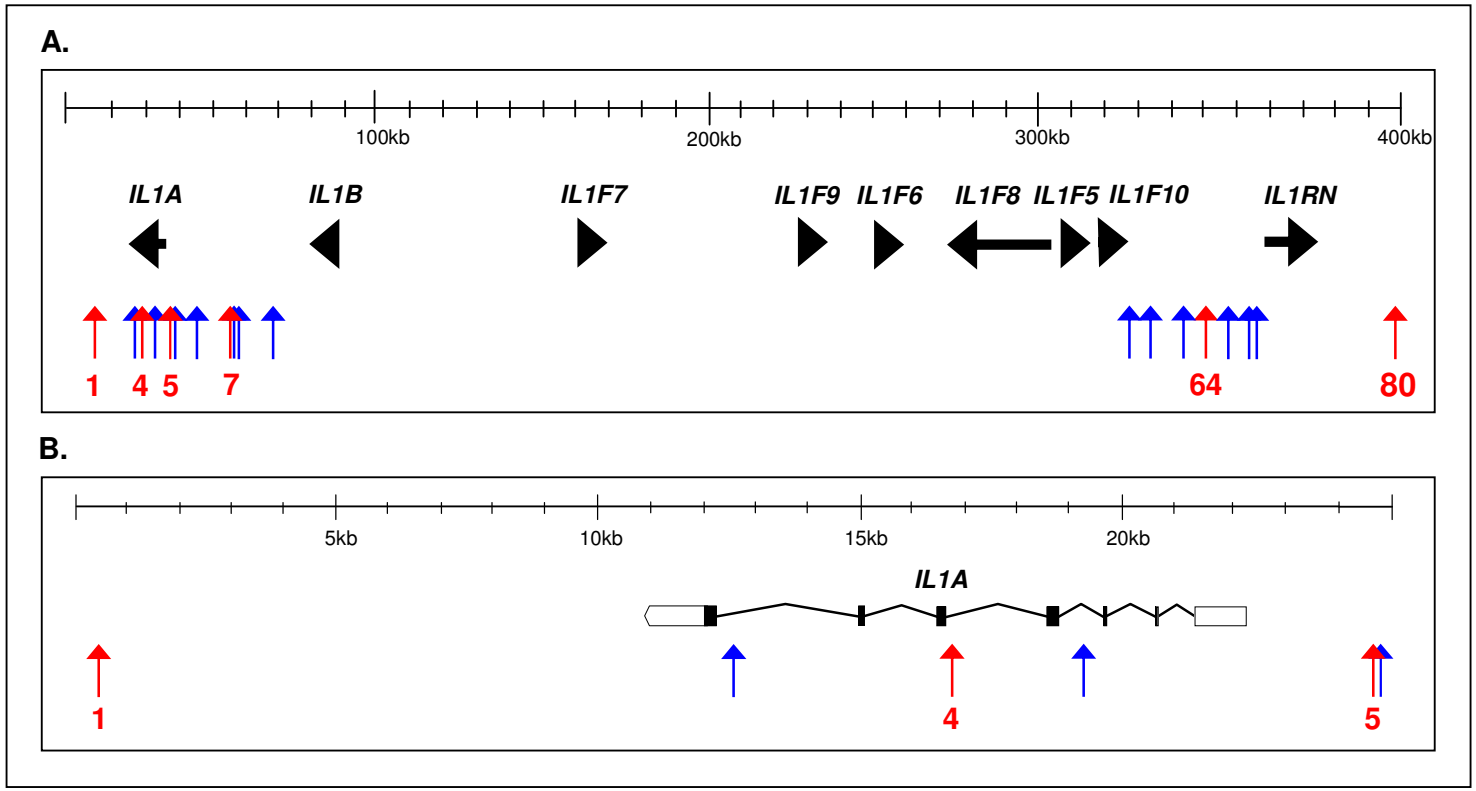


Figure 4.29 Positions of the associated SNPs in the *IL1* ligand cluster candidate region

The positions within the *IL1* ligand cluster candidate region of SNP1 (rs6712572), SNP4 (rs2071374), SNP5 (rs3783516), SNP7 (rs4849123), and SNP64 (rs1688075), which showed evidence of significant association, are shown in red, and the 13 tagged SNPs are shown in blue (A.). Due to the scale of the figure only the genotyped SNPs are labelled. The tagged SNPs in chromosomal order are: rs3783547, rs2856838, rs3783515, rs6746923, rs7585707, rs17598291, rs62157435, rs28928309, rs4848314, rs12475887, rs315925, rs315929, and rs315932. The locations of SNP1, rs2783547, SNP4, rs2856838, SNP5, and rs3783515 in relation to *IL1A* are shown in B. Untranslated regions are represented by white boxes and coding exons by black boxes.

association of each of these three SNPs was confirmed in analysis performed with the WTCCC2 control cohort. In analyses with this larger control cohort three SNPs which did not show evidence of significant association in the stratified analysis, rs3783516 (SNP5), rs4849123 (SNP7), and rs1688075 (SNP80), also showed evidence of significant association with sJIA. SNP1 is 10.8kb downstream of *IL1A* and is in high LD (r^2>0.8) with 12 SNPs, all of which are located within the flanking gene *CKAP2L*. SNP4 is within intron 4 of *IL1A* and is not known to be tagging any other SNPs within the candidate region. SNP5, 2.2kb upstream of *IL1A*, is in high LD (r^2>0.8) with two SNPs within *IL1A* introns, and two SNPs within 10.5kb of the transcription start site of *IL1A*. SNP7 is 26kb upstream of *IL1A* and in high LD (r^2>0.8) with three SNPs which are within 32kb upstream of *IL1A* and 18kb downstream of the 3' end of *IL1B*. SNP64 is 17kb upstream of *IL1RN* and in high LD (r^2>0.8) with six SNPs which are between 5.5kb and 40kb upstream of *IL1RN*. SNP80 is not known to be tagging any other SNPs and is 33.5kb downstream of the 3′ end of *IL1RN*.

The association study results for this candidate region, prior to WTCCC2 analysis, are published (Stock et al., 2008), see Appendix 4.

4.4.1.7. Discussion

Nine members of the IL-1 ligand gene family are located together in a cluster on chromosome 2. The cluster contains the genes encoding both of the classical IL-1 ligands, IL-1α and IL-1β, and the IL-1 receptor antagonist, IL-1Ra. Six additional members of the IL-1 ligand gene family are also located within this gene cluster. This includes the genes for the pro-inflammatory IL-F6, IL-1F8, and IL-1F9, the anti-inflammatory IL-1F5 and IL-1F7, and IL-1F10, a function for which has not currently been described (Dinarello et al., 2010).

A total of six SNPs in this candidate region, rs6712572 (SNP1), rs2071374 (SNP4), rs3783516 (SNP5), rs4849123 (SNP7), rs1688075 (SNP64), and rs1688075 (SNP80), showed evidence of significant disease association in this study. Four of these associated SNPs are located in, within, or nearby, *IL1A*, and the remaining two are located near to *IL1RN*. Three of these SNPs were significantly associated in the stratified meta-analysis. When the analysis was repeated, comparing the allele frequencies in the patients from both study stages with the large WTCCC2 control cohort, the results of these three SNPs were confirmed. An additional three SNPs, which were not associated in the stratified meta-analysis, were also found in this analysis to be significantly associated with sJIA. The larger

control cohort size available from the WTCCC project will have increased the power of the test, enabling identification of associations which the previous analysis had been underpowered to detect.

Polymorphisms within the genes included in this cluster have been associated with a number of diseases. The G to T nucleotide change at the *IL1A* exon 5 SNP rs17561, results in the substitution of an alanine for a serine at the 114th amino acid of pre-IL-1A. This amino acid substitution alters the response of pre-IL-1A to protease cleavage, which is required for secretion of the mature protein from the cell. There is a higher resistance to cleavage when an alanine is present at that position. Homozygosity for the G allele (alanine codon) at this SNP is associated with systemic sclerosis, and homozygosity for the T allele (serine codon) is associated with RA (Kawaguchi et al., 2007). The T allele, associated with greater processing of IL-1α to the mature form, has also been shown to be at a higher frequency in RA patients with destructive disease, and at a lower frequency in those with non-destructive disease, compared to healthy controls (Jouvenne et al., 1999). An *IL1A* promoter SNP, which is also in high LD with rs17561, was found to be associated with early-onset oligoarticular arthritis, especially in patients who developed chronic iridocyclitis, and with elevated ESR (McDowell et al., 1995). This finding was not however replicated in an independent cohort (Donn et al., 1999a), nor in a cohort including all JIA subtypes (Donn et al., 2001a). The non-synonymous SNP (rs17561) was directly genotyped in the association study presented in this thesis (SNP3), but showed no evidence of association with sJIA.

SNPs within *IL1RN* have been shown to be associated with CRP and IL-6 plasma levels in healthy individuals (rs4251961), fibrinogen levels (rs2232354), and with IL-1Ra production *ex vivo* (rs4252961) (Reiner et al., 2008). Genetic variants within *IL1RN* (rs315952, rs3213448, and rs315949) have also been found to be associated with RA (Luotola et al., 2010). All of these six SNPs were included in the association study presented in this thesis, either genotyped directly (rs2232354, SNP70, rs315949, SNP75), or tagged by genotyped SNPs (rs4251961 and rs4251961, both tagged by SNP76, rs315952, tagged by SNP73, and rs3213448, tagged by SNP74), but did not show evidence of association with sJIA. A variable number tandem repeat (VNTR) in intron two of *IL1RN* has been associated with JIA. This variation was also significantly associated with the subtypes of extended oligoarthritis, enthesitis-related arthritis and other arthritis, but not sJIA, when the subtypes were examined separately (Vencovsky et al., 2001). This *IL1RN* VNTR has also been shown to be associated

with OA of the hip (Meulenbelt et al., 2004) and knee (Smith et al., 2004). A number of studies have also found this VNTR to be associated with RA in different ethnic populations (You et al., 2007; Tolusso et al., 2006; Carreira et al., 2005; Lee et al., 2004), although in a meta-analysis of 15 studies, this association was not found to be significant overall (Xue et al., 2010). As only SNPs were included in the HapMap project, and this VNTR was not captured in the PGA polymorphism screen on *IL1A*, this VNTR was not included in the association study presented in this thesis.

A recent study by Hinks et al. included the IL-1 ligand cluster candidate region SNP1, SNP4, and SNP64, based on the published results of this study (Stock et al., 2008), in an association study of autoinflammatory genes with susceptibility to JIA (Hinks et al., 2011)(Submitted at time of writing). They found that SNP4 (rs2071374) showed evidence of significant disease association in a JIA cohort including patients of each subtype (pTrend=0.006). However, when each subtype was analysed separately SNP4 was only associated with the systemic subtype (pTrend=0.001), and when the whole JIA cohort excluding sJIA was tested there was no evidence of association. It is important to note that only seven of the 147 sJIA patients were not included in the study presented in this thesis (95.2% overlap), and that the control population used by Hinks et al for this SNP was WTCCC2. These results, although not an independent replication, do demonstrate that these IL-1 gene family associations are specific to only the systemic subtype of JIA. Because SNP1 (rs6712572) and SNP64 (rs1688075) had not been included in the WTCCC2 project, Hinks et al genotyped these two SNPs in a control cohort of 719 blood donors and individuals recruited from GP practices, which is independent from the control cohorts used in the study presented in this thesis. With this control cohort the significant association with sJIA was also seen with SNP1 (pTrend=0.009). In the control cohort the genotype distribution of SNP64 significantly deviated from HWE (p=0.03), so this SNP could not be tested for association.

Based on the D' LD relationships, and the results of the conditional analyses, it is probable that SNP1, SNP4, SNP5, and SNP7 are all reflecting one single effect. It is most likely, based on the significance levels achieved, that SNP4 is the primary effect to which the other three SNPs are secondary. However, unless a substantially larger patient cohort could be used, it is unlikely, due to the LD relationships, that further genotyping would provide additional information to determine the causal SNP at this locus. Direct demonstration of a biological effect on gene regulation or protein function would be required to determine with certainty

the SNP responsible for the disease association seen. Due to time constraints it was not possible to include this as part of this project.

Apart from SNP4 and SNP80, each of the other SNPs in this candidate region found to be associated with sJIA are in high LD with a number of other SNPs. Although all 12 of the SNPs with which SNP1 is in high LD are found within the flanking gene *CKAP2L* it is possible that they may still be involved in regulation of *IL1A* expression. Direct analysis of all of the SNPs incompletely tagged by the associated SNPs would be necessary to elucidate the validity of this potential disease association, and to determine the causal SNP.

ProIL-1α is biologically active and remains intracellularly, as well as being found on the cell membrane. Following protease cleavage into the mature form, IL-1α is released from cells, where it is a key activator of innate immunity and acute inflammatory responses. There is evidence that IL-1α plays a significant role in arthritis development and progression. It is involved in cartilage degradation (Isaev et al., 1992), levels of membrane-bound IL-1α, but not serum IL-1α or IL-1β, correlate with the severity of arthritis in a mouse model (Niki et al., 2004), high levels of IL-1α mRNA is found in synoviocytes from affected joints of patients with RA (Buchan et al., 1988), and RA patients who produce anti IL-1α antibodies develop a less destructive disease (Graudal et al., 2002). Dysregulation of IL-1A expression could lead to a perpetuation of the local inflammatory process, contributing to the development of the chronic inflammatory state in sJIA.

IL1RN encodes the IL-1 receptor antagonist IL-1Ra. IL-1 inhibitory activity is found in the serum of sJIA patients at the height of fever (Prieur et al., 1996), and patients with sJIA have been shown to have IL-1Ra plasma levels ten-fold higher than healthy control individuals (Muller et al., 1998). Alterations in the expression of this gene could lead to decreased levels of IL-1 antagonism, resulting in enhanced IL-1 cell signalling. This could account for the improvement seen in some sJIA patients when treated with the recombinant IL-1Ra, Anakinra, as the balance between activation and repression of IL-1 signalling would be restored towards a non-inflammatory state.

The issues of low power to detect significant associations, multiple testing, and the lack of significant findings in the stage-2 populations alone are discussed in chapter 5.

4.5. IL1 receptor cluster candidate region

4.5.1. Association Study

The genes flanking the six candidate genes in the *IL1* receptor cluster were identified using the July 2003 human genome reference sequence (NCBI build 34), hg16 annotation track. The boundaries of the candidate region, chromosome 2:102129849-102857620, were defined based on the positions of the flanking genes (centre panel Figure 4.30). Flanking *IL1R2* 5′ is mitogen-activated protein 4 kinase 4 (*MAP4K4*), and 3′ of *IL18RAP*, solute carrier family 9 (sodium/hydrogen exchanger), member 2 (*SLC9A2*). MAP4K4 is a key inducer of cell migration through activation of c-Jun N-terminal kinase (JNK) (Collins et al., 2006). SLC9A2, also known as Na(+)/H(+) exchanger 2 (NHE2), catalyses the counter transport of extracellular Na+ for intracellular H+, and is important in intracellular pH and cell volume regulation (Malakooti et al., 1999). Three mRNA sequences identified in cDNA libraries, have also been mapped to within the selected candidate region. *FLJ20373*, which overlaps with the *MAP4K4* sequence, is an alternative 5′ UTR for *MAP4K4*. No information is available for the mRNA clone *AY358263*, between *IL1R2* and *IL1R1*. The *AY192162* mRNA clone, the sequence for which overlaps *IL18R1*, is a putative alternative *IL18R1* transcript. Based on the positions of these flanking genes the candidate region extended 99.9kb (99,857bp) 5′ of *IL1R2* and 167kb (167,195bp) 3′ of *IL18RAP*. The whole candidate region spanned 0.7Mb (727,771bp) (bottom panel Figure 4.30).

4.5.1.1. Tagging SNP selection

Genotyping data was available for a total of 612 SNPs within the selected *IL1* receptor cluster candidate region, 293 from HapMap (top panel Figure 4.30), and 353 from PGA. Of these, 34 were included in both population genotyping resources, and 319 had data from only PGA. Forty three of the SNPs in the region were monomorphic in the CEU population, and a further 22 had a MAF<0.05. An additional six SNPs were excluded due to low genotype call rates. These SNPs were therefore not included in the LD analysis. A total of 542 SNPs within the candidate region satisfied the inclusion criteria and were included in the LD investigation and subsequent tSNP selection.

A graphical representation of the LD between these 542 SNPs in the *IL1* receptor cluster candidate region is shown in Figure 4.31. There were two visible blocks of high LD within the candidate region. One of the blocks spanned 16.8kb over *IL1R2*, from intron 1 to intron 6.

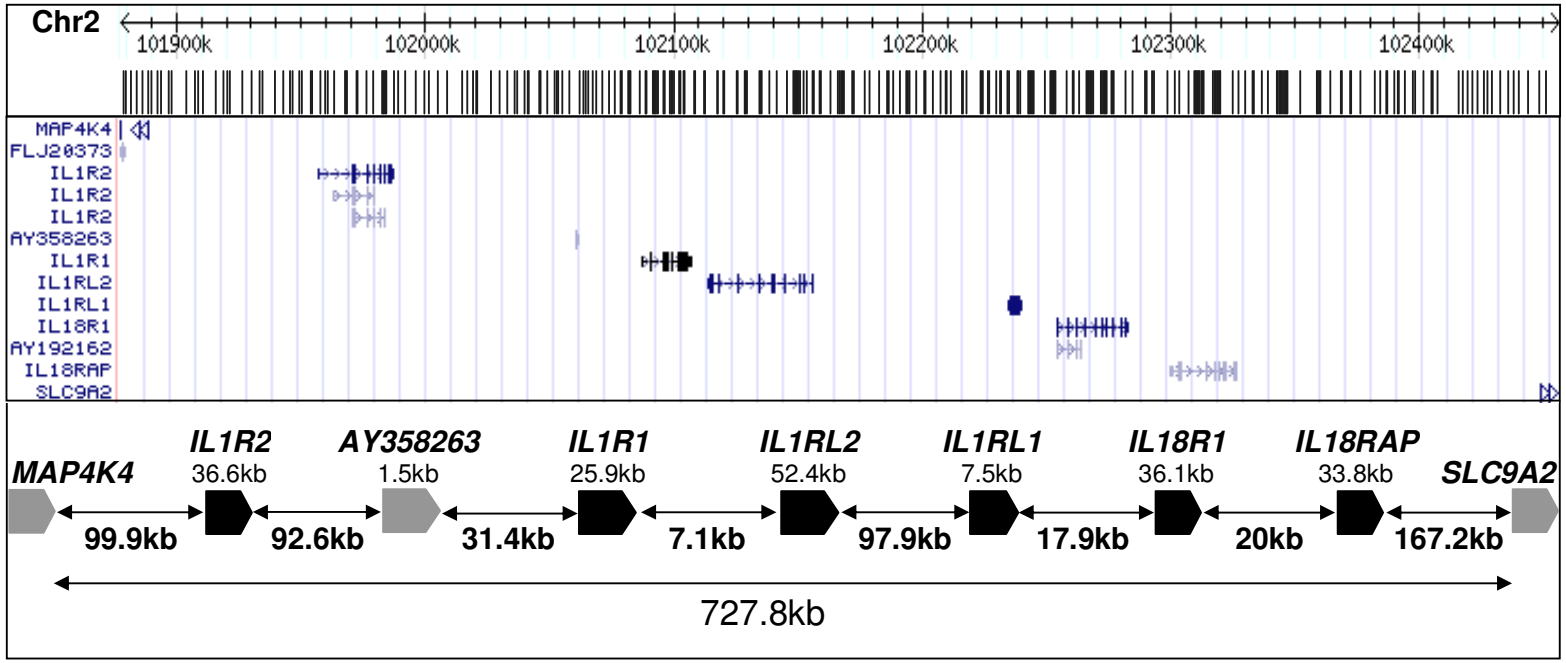


Figure 4.30 *IL1* receptor cluster candidate region

The boundaries of the *IL1* receptor cluster candidate region are the flanking genes *MAP4K4* and *SLC9A2*, determined by the July 2003 human genome reference sequence (NCBI build 34), hg16 annotation track, shown in the centre panel. The 3′ end of *MAP4K4* and the 5′ end of *SCL9A2* are proximal to the gene cluster. The approximate (due to scale) positions within this region of all SNPs (irrespective of MAF) in the HapMap phase 2 database (data release #18/phaseII Sept05) are shown in the top panel. The bottom panel shows a schematic (not to scale) of the candidate region with gene orientation, size, and distances given. The genes of interest are shown in black and all other features within the region in grey. The largest isoform of each gene was used for size calculations, and, for clarity, is the only one shown. All sizes and distances are based on chromosomal positions reported in the NCBI build 34 genome reference sequence.

The other block of high LD spanned 28.7kb of the *IL1R2* sequence, from intron 4 to intron 8. There was also a diffuse pattern of moderate to high LD on the right hand side of the region, corresponding to the region containing *IL1RL1*, *IL18R1*, and *IL18RAP*.

Due to the LD relationships between the SNPs a total of 129 tSNPs (21% of the total) were identified as the minimal subset required to capture the variation within all 612 SNPs. However, following assay design (section 3.1.2.1.) it was necessary to exclude 22 SNPs, tagged by 5 SNPs, due to poor assay design scores. The final SNP set therefore consisted of 124 tSNPs which captured the variation within 520 SNPs. Forty seven of the tSNPs were not in high LD ($r^2 \leq 0.8$) with any other SNP within the region and so were not tagging any additional SNPs. All of the selected tSNPs, and the SNPs captured by each, are listed in Table 4.36.

4.5.1.2. Stage-1

The genotyping assay for rs3218926, which was not tagging any other SNPs, was classed as failed due to poor performance (Figure 4.1 and Table 4.2). An additional seven genotyped SNPs from the *IL1* receptor cluster candidate region showed significant deviation from HWE (Table 4.3), and were therefore also excluded from the association analysis. These were: rs2310175 (SNP 25), rs11123915 (SNP 81), rs2241116 (SNP 104), rs1420095 (SNP 105), rs17027415 (SNP 117), rs1861228 (SNP 120), and rs2058612 (SNP 123), tagging a total of eight SNPs. Of the 16 excluded SNPs (eight tSNPs and eight captured SNPs) one was in sufficiently high LD ($r^2 \geq 0.8$) with SNP 20 and was therefore still captured. The remaining 15 excluded SNPs however were not in high LD ($r^2 \geq 0.8$) with any of the analysed SNPs, and were therefore not captured by the association study (Table 4.37). A total of 116 genotyped SNPs capturing 505 SNPs were included in the association study analysis.

Of the 116 analysed *IL1* receptor region tSNPs, eight showed significant ($p<0.05$) frequency differences between the patient (n=130) and control (n=146) cohorts. For ease of interpretation the results of the SNPs showing significant (Table 4.38) and non-significant frequency differences (Table 4.39) are presented separately. The rare allele was present at a higher frequency in the patient than the control population for only three SNPs (SNPs 1, 4, and 8) resulting in an OR>1, while the common allele of the remaining SNPs was more frequent in the patient population. The highest OR (1.709) and smallest p-value (0.00313) were seen with SNP8, the frequency of the rare allele 0.396 in the case population and 0.277

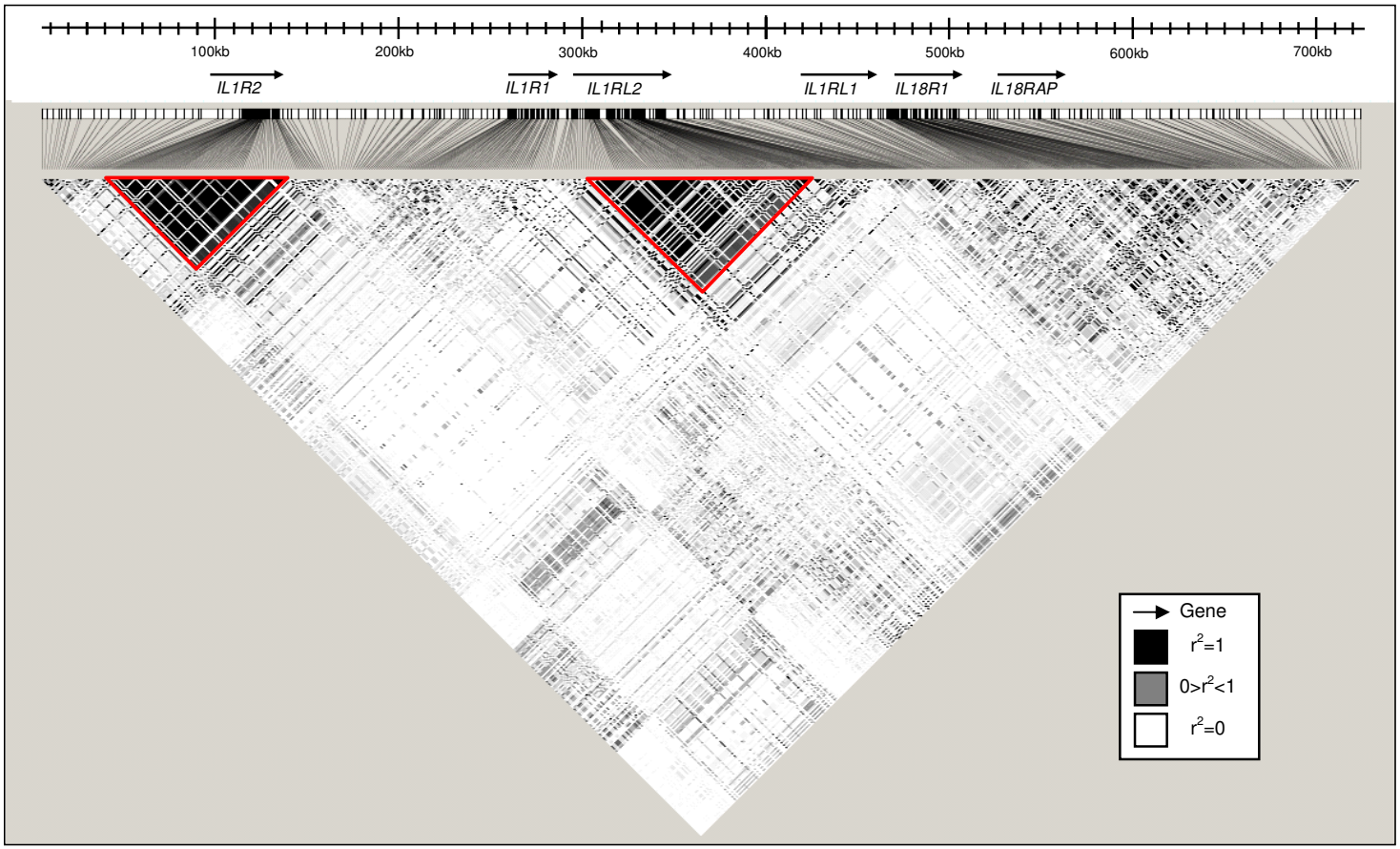


Figure 4.31 LD plot of the SNPs in the *IL1* receptor gene cluster region

The pairwise LD relationships between SNPs in the *IL1* receptor gene cluster candidate region were investigated to enable tSNP selection. The region shown is comprised of only the SNPs which satisfied the frequency criteria and were included in the LD analysis (n=542), not the whole of the candidate region. The gene orientation and location within the region is indicated. The position within the region of each SNP satisfying the selection criteria is represented by a line underneath the gene position. Each coloured square represents the pairwise LD between two SNPs. The graphical representation, generated in Haploview, is shown on a grey scale, with higher r^2 LD represented by darker blocks. Due to scale the individual SNPs are not labelled. The blocks of high LD are indicated.

SNP Nº	rs No	Tagged SNPs
1	rs2871432	rs2310144,rs2310143,rs6720836
2	rs9308849	rs2310141,rs4851509,rs1003376,rs11688863,rs10489971
3	rs2056400	n/a
4	rs2190364	rs7568226,rs2214890
5	rs1541434	n/a
6	rs2190358	n/a
7	rs4851516	rs6726051
8	rs12712122	rs1010329,rs2190360,rs740044,rs2190361
9	rs4851519	n/a
10	rs4851522	n/a
11	rs4851526	n/a
12	rs3218848	n/a
13	rs2282747	n/a
14	rs2110562	rs2110563,rs719248,rs4851527,rs757918
15	rs3218888	n/a
16	rs3218894	n/a
17	rs3218921	n/a
18	rs2072475	rs11465650,rs3218927,rs11465623,rs719250,rs3218920,rs11465660,rs719249,rs2072476,rs2302589
19	rs3218979	rs3212326,rs3218977,rs888001,rs4851535,rs2241082
20	rs2072473	rs3218876,rs2236926,rs2236925,rs2072477,rs2073492,rs733498,rs3218972,rs3218914,rs3218879, rs3218872,rs3218874,rs3218877,rs3218953,rs2072480,rs3218923,rs2262267,rs3218873,rs3218878, rs2072479,rs2072474,rs3218896,rs3218925,rs3218885,rs3218928,rs3218870,rs2072478,rs2236928, rs2073491,rs2236930,rs3218900,rs2236923,rs2236922,rs3218903,rs2236921,rs2236924,rs4850995, rs3218911,rs2310171,rs3218892,rs3218883,rs3218966,rs2072481,rs2282743,rs3218862,rs2282744, rs2298939,rs3218971,rs3218919,rs2236927,rs2072482,rs3218909,rs2236929,rs3218991,rs3218901, rs3218875
21	rs3218984	rs3218980,rs3218974,rs3218910,rs2160140,rs2072472,rs3218934
22	rs7589525	rs11688205
23	rs4851531	n/a
24	rs4851534	rs7588933,rs2310173
25	rs2310175	n/a
26	rs1558645	rs868252
27	rs17026582	n/a

SNP Nº	rs No	Tagged SNPs
28	rs7569218	n/a
29	rs2041751	n/a
30	rs4850997	rs10181737,rs1558644
31	rs871376	rs13029804
32	rs963399	rs1030023
33	rs10490571	rs1024794,rs992153,rs1861283,rs13388182
34	rs6729953	n/a
35	rs887998	n/a
36	rs1558643	rs12712127
37	rs2310186	rs6712813
38	rs4851543	n/a
39	rs1558642	rs1558642,rs9284725
40	rs6754776	n/a
41	rs3917225	rs3917239
42	rs2192752	rs101698183,rs3917246, rs3917256,rs871659,rs3771202,rs3917245
43	rs949963	rs871656,rs3755292,rs1465325,rs871657,rs13387400
44	rs2287049	rs3917238
45	rs871658	n/a
46	rs2287047	rs956730,rs6706048
47	rs3917243	n/a
48	rs3917254	rs3917328,rs13011273,rs3917320,rs3917325,rs3917314,rs3917291
49	rs3917257	n/a
50	rs3917273	rs3917317
51	rs2228139	rs34722443,rs4851553,rs11887842,rs3917299,rs17026757,rs35918372
52	rs997049	rs3917290,rs12479046,rs3771200
53	rs3917292	n/a
54	rs2160227	rs3917318,rs3917304,rs10208520,rs951192,rs3213735
55	rs3917296	n/a
56	rs951193	n/a
57	rs3171845	rs3917329,rs3917331,rs3917289,rs3917330,rs35952417,rs4851552
58	rs2110726	rs3917263

SNP Nº	rs No	Tagged SNPs
59	rs3917332	n/a
60	rs1030021	rs10184597
61	rs28385684	n/a
62	rs3771199	rs1922292
63	rs11678651	n/a
64	rs2871457	n/a
65	rs12996377	rs13014004,rs1558648,rs2241132
66	rs2160226	rs1922289,rs7369549,rs1922297,rs2302624,rs6715919,rs1558646,rs995514,rs7569116,rs2302623, rs34088260,rs2041752,rs1922293,rs34381576,rs3771197,rs17020717,rs1882514,rs3771196,rs3771194, rs2141865,rs4851002,rs3755289,rs7566395,rs2178675,rs867770,rs2310236,rs3771193,rs3729564, rs7605715,rs11688596,rs870684,rs2871458,rs1882513,rs1922290,rs1922295,rs35917382,rs3815517, rs1922296,rs7421641,rs995515,rs2141864,rs34074031,rs2310238,rs2310235,rs6733727,rs11686153, rs3771192,rs4851001,rs1169062,rs2041753,rs3755290,rs3074969,rs2310237,rs3821207,rs3771195, rs34870699
67	rs3771187	rs3771185,rs12472862,rs17637748,rs12476585,rs2310241,rs35139442,rs34778121,rs12473710, rs67445032,rs12472872
68	rs3771184	rs3771181,rs7560478,rs7421586,rs12474373,rs12468673,rs34188221,rs7572871,rs3213734,rs2287041, rs3771188,rs2302620,rs10515922,rs3755287,rs10198860,rs10166535,rs1024791,rs3755286, rs13416449,rs7583215,rs3821206,rs7572853,rs3771186,rs17026825,rs2241130
69	rs3755285	rs11123913
70	rs955754	rs2302612,rs6709635
71	rs4851561	n/a
72	rs10167431	n/a
73	rs2287040	rs3755282
74	rs13014084	rs6543112,rs12995229,rs11413518,rs12989930
75	rs11692230	n/a
76	rs6752467	n/a
77	rs1035131	rs1558626
78	rs11677452	rs9646944
79	rs1420092	n/a
80	rs1345301	rs10200410,rs2310243,rs12475055
81	rs11123915	rs1115281
82	rs1922288	rs6747153
83	rs12469892	n/a

SNP Nº	rs No	Tagged SNPs
84	rs6543113	rs4851003
85	rs1476984	rs4851567,rs917998,rs1420089
86	rs6706844	rs4090473,rs974389,rs953934,rs1558622,rs4399750,rs1997466,rs12996772,rs11685424
87	rs1420103	rs2310220
88	rs1041973	n/a
89	rs6719130	rs10515921
90	rs1921622	n/a
91	rs4988956	rs1861245,rs10176664,rs10170583
92	rs6710885	rs1882348,rs2287037,rs1420098,rs4851569
93	rs11465567	n/a
94	rs11465572	n/a
95	rs3771166	rs10182710,rs10204837,rs1420099,rs9308857,rs6543124,rs1362348,rs11123925,rs1974675,rs3755276, rs6543123
96	rs6728945	rs11465584,rs3771167,rs10197284,rs11465596
97	rs11465597	n/a
98	rs7579737	n/a
99	rs6758936	rs2041739,rs6749014,rs7594402,rs7584093,rs3755266,rs2287033,rs7556917,rs1420094,rs6731157, rs4851004,rs6706002,rs1592458,rs11465641,rs1420096,rs7559845,rs6745614,rs10204757,rs3213732, rs2310300
100	rs2270297	rs1573895,rs3771170,rs2241117,rs6733346,rs1558627,rs1568681,rs6731154,rs13006566,rs10190555, rs2058623,rs1465321,rs6753717,rs2041740,rs6727306,rs12712143
101	rs7558013	n/a
102	rs35582281	n/a
103	rs1035130	rs10490204,rs1135354,rs3771171,rs10490203,rs3771156,rs2270298,rs3771172,rs2287035,rs950880, rs2080289,rs2160202,rs1420101,rs3732125,rs4851570,rs3732126,rs3771150,rs2287034,rs11465633
104	rs2241116	n/a
105	rs1420095	n/a
106	rs3732127	rs2075186,rs3771164,rs741284,rs3771161,rs3213733
107	rs36036501	n/a
108	rs4851575	rs1807782,rs1420106,rs13015714
109	rs917997	rs2058660,rs917996,rs7559479,rs6705001,rs1558650,rs4241210,rs1468791
110	rs4140786	rs11465730,rs2310302
111	rs759381	n/a

SNP N^{o}	rs No	Tagged SNPs
112	rs1880000	rs7601773,rs7575867,rs1523196,rs7593444,rs1403549,rs17027341,rs6749440,rs1159509,rs1523197
113	rs10207579	n/a
114	rs2075190	rs4851017,rs1523198,rs1403548
115	rs1403550	rs2192758,rs2005881,rs6739426
116	rs2177317	rs10490202,rs741285,rs2015478
117	rs17027415	rs11123936,rs2080316,rs1946131
118	rs6761291	rs4851608,rs1476999
119	rs1861229	rs7559928
120	rs1861228	rs2192756
121	rs873625	rs9973883,rs2192667
122	rs7603234	rs3806581
123	rs2058612	rs741257,rs2192665,rs7602767
124	rs3218926	n/a
Uncaptured		rs3218954,rs11883987,rs2287048,rs3917265,rs3917303,rs3917324,rs2241131,rs1997504,rs34591907, rs12474258,rs1922302,rs192291,PGARL2030269,rs917994,rs2302621,rs1997502,rs3771175, rs11465570,rs11465571,rs11465613,rs11465615,rs11465620

Table 4.36 *IL1* receptor cluster candidate region tSNPs

A total of 124 tSNPs were genotyped to capture 520 of the 612 SNPs within the candidate region. The SNPs captured ($r^2 \geq 0.8$) by each genotyped tSNP are shown. For ease of identification each tSNP genotyped was assigned a number according to genome order. Due to poor assay design scores it was not possible to genotype or capture 22 of the SNPs within the region.

SNP Nº	rs Nº	Tagging
20	rs2072473	rs3218926
Uncaptured		rs11123915,rs1115281,rs2241116,rs1420095,rs17027415, rs11123936,rs2080316,rs1946131,rs1861228,rs2192756, rs2058612,rs741257,rs2192665,rs7602767,rs2310175

Table 4.37 *IL1* receptor cluster additional tagging of analysis excluded SNPs

Eight of the genotyped tSNPs from the *IL1* receptor cluster candidate region were excluded from analysis due to poor genotyping quality or deviation from HWE. These eight SNPs were tagging eight additional SNPs between them, making a total of 16 SNPs excluded from analysis. The excluded SNPs are shown, as well as the one captured ($r^2 \geq 0.8$) by another of the genotyped SNPs.

SNP	rs Nº	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
1	rs2871432 (G/A)	1 2	0.496 (129) 0.504 (131)	0.582 (170) 0.418 (122)	1.415 (1.011-1.981)	0.0428
2	rs9308849 (A/G)	1 2	0.746 (194) 0.254 (66)	0.644 (188) 0.356 (104)	0.615 (0.426-0.889)	0.00909
4	rs2190364 (G/C)	1 2	0.685 (178) 0.315 (82)	0.764 (223) 0.236 (69)	1.489 (1.022-2.168)	0.0376
6	rs2190358 (G/A)	1 2	0.758 (197) 0.242 (63)	0.664 (194) 0.336 (98)	0.633 (0.436-0.919)	0.0157
7	rs4851516 (A/C)	1 2	0.739 (192) 0.261 (68)	0.648 (188) 0.352 (102)	0.653 (0.452-0.942)	0.0219
8	rs12712122 (G/A)	1 2	0.604 (157) 0.396 (103)	0.723 (211) 0.277 (81)	1.709 (1.196-2.442)	0.00313
14	rs2110562 (A/G)	1 2	0.569 (148) 0.431 (112)	0.476 (139) 0.524 (153)	0.688 (0.491-0.962)	0.0285
23	rs4851531 (G/A)	1 2	0.664 (170) 0.336 (86)	0.554 (155) 0.446 (125)	0.627 (0.442-0.891)	0.00878

Table 4.38 Significant *IL1* receptor cluster candidate region stage-1 analysis results

One hundred and sixteen genotyped tSNPs within the *IL1* receptor cluster candidate region were analysed for significant allele frequency differences between the patient (n=130) and control (n=146) cohorts. The frequency in each cohort of both the common (1) and rare (2) alleles are given. The alternative alleles of each SNP are shown underneath the rs identifier. The allele assigned as the common (1) allele is given first and the rarer (2) allele second. The OR shown is for allele 2 (allele 1 OR=1). The p-value for the test for frequency differences between the cohorts is also shown. The results for the eight tSNPs which showed significant ($p<0.05$) frequency differences between the two cohorts are shown.

SNP N^{o}	rs Number	Allele	Case freq. (N^{o})	Control freq. (N^{o})	OR (95% CI)	p value
3	rs2056400 (C/A)	1 2	0.811 (211) 0.189 (49)	0.808 (236) 0.192 (56)	0.979 (0.639-1.498)	0.921
5	rs1541434 (G/A)	1 2	0.650 (169) 0.350 (91)	0.637 (189) 0.353 (103)	0.988 (0.696-1.402)	0.946
9	rs4851519 (G/A)	1 2	0.915 (238) 0.0846 (22)	0.884 (258) 0.116 (34)	0.701 (0.399-1.233)	0.2143
10	rs4851522 (G/A)	1 2	0.892 (232) 0.108 (28)	0.907 (265) 0.0924 (27)	1.185 (0.679-2.068)	0.551
11	rs4851526 (A/G)	1 2	0.665 (173) 0.335 (87)	0.613 (179) 0.387 (113)	0.797 (0.562-1.129)	0.201
12	rs3218848 (G/A)	1 2	0.665 (173) 0.335 (87)	0.613 (179) 0.387 (113)	0.797 (0.562-1.129)	0.201
13	rs2282747 (A/G)	1 2	0.904 (235) 0.0961 (25)	0.856 (250) 0.144 (42)	0.633 (0.374-1.072)	0.0849
15	rs3218888 (A/G)	1 2	0.904 (235) 0.0962 (25)	0.860 (251) 0.140 (41)	0.651 (0.384-1.104)	0.108
16	rs3218894 (A/G)	1 2	0.954 (248) 0.0462 (12)	0.942 (275) 0.0582 (17)	0.783 (0.367-1.671)	0.525
17	rs3218921 (G/A)	1 2	0.815 (207) 0.185 (47)	0.781 (228) 0.219 (64)	0.809 (0.513-1.232)	0.322
18	rs2072475 (G/A)	1 2	0.861 (224) 0.139 (36)	0.824 (239) 0.176 (51)	0.753 (0.474-1.198)	0.229
19	rs3218979 (G/A)	1 2	0.885 (230) 0.115 (30)	0.911 (266) 0.0890 (26)	1.045 (0.767-2.322)	0.307
20	rs2072473 (G/A)	1 2	0.885 (230) 0.115 (30)	0.836 (244) 0.164 (48)	0.663 (0.406-1.083)	0.0973
21	rs3218984 (A/G)	1 2	0.773 (198) 0.227 (58)	0.726 (212) 0.274 (80)	1.634 (0.526-1.146)	0.201
22	rs7589525 (A/C)	1 2	0.777 (202) 0.223 (58)	0.726 (212) 0.274 (80)	0.761 (0.516-1.123)	0.167
24	rs4851534 (A/T)	1 2	0.542 (141) 0.458 (119)	0.466 (136) 0.534 (156)	0.736 (0.526-1.029)	0.0724
26	rs1558645 (A/C)	1 2	0.604 (157) 0.396 (103)	0.555 (162) 0.445 (130)	0.818 (0.582-1.148)	0.244
27	rs17026582 (G/A)	1 2	0.754 (196) 0.246 (64)	0.764 (217) 0.236 (67)	1.058 (0.714-1.567)	0.780
28	rs7569218 (C/A)	1 2	0.519 (135) 0.481 (125)	0.500 (146) 0.500 (146)	0.926 (0.663-1.294)	0.652
29	rs2041751 (A/G)	1 2	0.673 (175) 0.327 (85)	0.685 (200) 0.315 (92)	1.056 (0.728-1.510)	0.766
30	rs4850997 (C/A)	1 2	0.827 (215) 0.173 (45)	0.832 (243) 0.168 (49)	0.0270 (0.666-1.619)	0.870
31	rs871376 (A/G)	1 2	0.781 (203) 0.219 (57)	0.791 (231) 0.209 (61)	1.063 (0.708-1.598)	0.768
32	rs963399 (A/C)	1 2	0.629 (161) 0.371 (95)	0.606 (177) 0.394 (115)	0.908 (0.643-1.283)	0.585
33	rs10490571 (G/A)	1 2	0.647 (167) 0.353 (91)	0.654 (191) 0.346 (101)	1.030 (0.725-1.464)	0.867

SNP Nº	rs Number	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
34	rs6729953 (A/C)	1 2	0.823 (214) 0.177 (46)	0.849 (248) 0.151 (44)	1.212 (0.771-1.904)	0.405
35	rs887998 (G/A)	1 2	0.769 (200) 0.231 (60)	0.743 (217) 0.257 (75)	0.868 (0.588-1.282)	0.476
36	rs1558643 (G/A)	1 2	0.577 (150) 0.423 (110)	0.582 (170) 0.418 (122)	1.022 (0.728-1.434)	0.900
37	rs2310186 (A/C)	1 2	0.877 (228) 0.123 (32)	0.887 (259) 0.113 (33)	1.102 (0.656-1.849)	0.714
38	rs4851543 (G/A)	1 2	0.908 (236) 0.0923 (24)	0.911 (266) 0.0890 (26)	1.040 (0.582-1.862)	0.894
39	rs1558642 (C/A)	1 2	0.761 (198) 0.239 (62)	0.750 (219) 0.250 (73)	0.939 (0.637-1.386)	0.753
40	rs6754776 (C/G)	1 2	0.877 (228) 0.123 (32)	0.884 (258) 0.116 (34)	1.065 (0.637-1.782)	0.810
41	rs3917225 (A/G)	1 2	0.565 (147) 0.435 (113)	0.565 (165) 0.435 (127)	0.999 (0.713-1.399)	0.994
42	rs2192752 (A/C)	1 2	0.761 (198) 0.239 (62)	0.750 (219) 0.250 (73)	0.939 (0.627-1.386)	0.753
43	rs949963 (G/A)	1 2	0.815 (212) 0.185 (48)	0.839 (245) 0.161 (47)	1.180 (0.758-1.837)	0.463
44	rs2287049 (A/G)	1 2	0.692 (180) 0.308 (80)	0.723 (211) 0.277 (81)	1.158 (0.802-1.672)	0.435
45	rs871658 (C/G)	1 2	0.877 (228) 0.123 (32)	0.886 (257) 0.114 (33)	1.093 (0.651-1.835)	0.737
46	rs2287047 (G/A)	1 2	0.692 (180) 0.308 (80)	0.729 (213) 0.271 (79)	1.198 (0.829-1.733)	0.336
47	rs3917243 (G/A)	1 2	0.554 (144) 0.446 (116)	0.538 (157) 0.462 (135)	0.937 (0.669-1.311)	0.703
48	rs3917254 (G/A)	1 2	0.911 (237) 0.0885 (23)	0.911 (266) 0.0890 (26)	0.993 (0.552-1.787)	0.981
49	rs3917257 (G/C)	1 2	1.00 (260) 0.00 (0)	1.00 (292) 0.00 (0)	1.00 (1.00-1.00)	1.00
50	rs3917273 (A/T)	1 2	0.604 (157) 0.396 (103)	0.610 (178) 0.390 (114)	1.024 (0.728-1.442)	0.890
51	rs2228139 (G/C)	1 2	0.938 (244) 0.0615 (16)	0.938 (274) 0.0616 (18)	0.998 (0.498-2.00)	0.996
52	rs997049 (T/A)	1 2	0.596 (155) 0.404 (105)	0.579 (169) 0.421 (123)	0.931 (0.663-1.307)	0.678
53	rs3917292 (G/A)	1 2	0.926 (239) 0.0744 (19)	0.918 (268) 0.0822 (24)	0.888 (0.474-1.661)	0.709
54	rs2160227 (C/A)	1 2	0.727 (189) 0.273 (71)	0.719 (210) 0.281 (82)	0.962 (0.662-1.398)	0.839
55	rs3917296 (A/G)	1 2	0.911 (237) 0.0885 (23)	0.894 (261) 0.106 (31)	0.817 (0.463-1.441)	0.484
56	rs951193 (G/A)	1 2	0.969 (248) 0.0313 (8)	0.942 (275) 0.0582 (17)	0.522 (0.221-1.230)	0.126
57	rs3171845 (G/A)	1 2	0.939 (244) 0.0615 (16)	0.932 (272) 0.0685 (20)	0.892 (0.452-1.760)	0.741
58	rs2110726 (G/A)	1 2	0.623 (162) 0.377 (98)	0.610 (178) 0.390 (114)	0.945 (0.669-1.332)	0.745

SNP Nº	rs Number	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
59	rs3917332 (A/T)	1 2	0.781 (203) 0.219 (57)	0.781 (228) 0.219 (64)	1.00 (0.668-1.498)	0.999
60	rs1030021 (A/C)	1 2	0.721 (186) 0.279 (72)	0.719 (210) 0.281 (82)	0.991 (0.683-1.440)	0.964
61	rs28385684 (G/A)	1 2	0.988 (257) 0.0115 (3)	1.00 (292) 0.00 (0)	1.00 (1.00-1.00)	1.00
62	rs3771199 (A/G)	1 2	0.661 (172) 0.339 (88)	0.644 (188) 0.356 (104)	0.925 (0.651-1.314)	0.663
63	rs11678651 (C/A)	1 2	0.658 (171) 0.342 (89)	0.644 (188) 0.356 (104)	0.941 (0.663-1.336)	0.733
64	rs2871457 (A/C)	1 2	0.677 (176) 0.323 (84)	0.668 (195) 0.332 (97)	0.960 (0.672-1.370)	0.820
65	rs12996377 (A/T)	1 2	0.858 (223) 0.142 (37)	0.877 (256) 0.123 (36)	1.180 (0.721-1.931)	0.511
66	rs2160226 (G/A)	1 2	0.661 (172) 0.339 (88)	0.644 (188) 0.356 (104)	0.925 (0.651-1.314)	0.663
67	rs3771187 (A/G)	1 2	0.535 (139) 0.465 (121)	0.524 (153) 0.476 (139)	0.958 (0.686-1.339)	0.803
68	rs3771184 (G/A)	1 2	0.879 (225) 0.121 (31)	0.877 (256) 0.123 (36)	0.980 (0.587-1.636)	0.938
69	rs3755285 (A/G)	1 2	0.725 (187) 0.275 (71)	0.750 (219) 0.250 (73)	1.139 (0.779-1.667)	0.503
70	rs955754 (A/G)	1 2	0.789 (205) 0.211 (55)	0.808 (236) 0.192 (56)	1.131 (0.746-1.715)	0.563
71	rs4851561 (G/A)	1 2	0.939 (244) 0.0615 (16)	0.942 (275) 0.0582 (17)	1.061 (0.525-2.145)	0.870
72	rs10167431 (G/A)	1 2	0.536 (134) 0.464 (116)	0.528 (152) 0.472 (136)	0.968 (0.689-1.359)	0.849
73	rs2287040 (G/A)	1 2	0.669 (174) 0.331 (86)	0.651 (190) 0.340 (102)	0.921 (0.647-1.310)	0.646
74	rs13014084 (A/G)	1 2	0.946 (246) 0.0545 (14)	0.959 (280) 0.041 (12)	1.328 (0.603-2.925)	0.481
75	rs11692230 (A/G)	1 2	0.592 (154) 0.408 (106)	0.579 (169) 0.421 (123)	0.946 (0.674-1.328)	0.747
76	rs6752467 (G/A)	1 2	0.873 (227) 0.127 (33)	0.849 (248) 0.151 (44)	0.819 (0.504-1.332)	0.420
77	rs1035131 (A/C)	1 2	0.627 (163) 0.373 (97)	0.640 (187) 0.360 (105)	1.060 (0.749-1.499)	0.743
78	rs11677452 (A/T)	1 2	0.785 (204) 0.215 (56)	0.815 (238) 0.185 (54)	1.210 (0.797-1.838)	0.372
79	rs1420092 (C/A)	1 2	0.862 (224) 0.138 (36)	0.877 (256) 0.123 (36)	1.143 (0.696-1.876)	0.598
80	rs1345301 (A/G)	1 2	0.565 (147) 0.435 (113)	0.541 (158) 0.459 (134)	0.906 (0.648-1.269)	0.567
82	rs1922288 (G/A)	1 2	0.742 (193) 0.258 (67)	0.750 (219) 0.250 (73)	1.041 (0.709-1.529)	0.836
83	rs12469892 (G/A)	1 2	0.842 (219) 0.158 (41)	0.870 (254) 0.130 (38)	1.251 (0.777-2.016)	0.357
84	rs6543113 (G/A)	1 2	0.735 (191) 0.265 (69)	0.723 (211) 0.277 (81)	0.941 (0.646-1.371)	0.751

SNP Nº	rs Number	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
85	rs1476984 (G/A)	1 2	0.908 (236) 0.0923 (24)	0.897 (262) 0.103 (30)	0.888 (0.505-1.562)	0.680
86	rs6706844 (G/A)	1 2	0.535 (139) 0.465 (121)	0.538 (157) 0.462 (135)	1.012 (0.724-1.415)	0.943
87	rs1420103 (C/A)	1 2	0.758 (197) 0.242 (63)	0.794 (232) 0.206 (60)	1.237 (0.828-1.847)	0.300
88	rs1041973 (C/A)	1 2	0.566 (194) 0.254 (66)	0.767 (224) 0.233 (68)	1.121 (0.759-1.654)	0.566
89	rs6719130 (G/A)	1 2	0.873 (227) 0.127 (33)	0.849 (248) 0.151 (44)	0.819 (0.504-1.332)	0.420
90	rs1921622 (A/G)	1 2	0.489 (127) 0.511 (133)	0.538 (157) 0.462 (135)	1.218 (0.871-1.702)	0.248
91	rs4988956 (G/A)	1 2	0.608 (158) 0.392 (102)	0.579 (169) 0.421 (123)	0.887 (0.631-1.247)	0.490
92	rs6710885 (A/G)	1 2	0.627 (163) 0.373 (97)	0.613 (179) 0.387 (113)	0.943 (0.668-1.330)	0.737
93	rs2058612 (A/G)	1 2	0.912 (237) 0.0885 (23)	0.914 (267) 0.0856 (25)	1.036 (0.573-1.875)	0.906
94	rs11465572 (C/A)	1 2	0.912 (237) 0.0885 (23)	0.914 (267) 0.0856 (25)	1.036 (0.57301.875)	0.906
95	rs3771166 (G/A)	1 2	0.608 (158) 0.392 (102)	0.579 (169) 0.421 (123)	0.887 (0.631-1.247)	0.490
96	rs6728945 (A/G)	1 2	0.885 (230) 0.115 (30)	0.849 (248) 0.151 (44)	0.735 (0.447-1.209)	0.223
97	rs11465597 (A/G)	1 2	0.900 (234) 0.100 (26)	0.908 (265) 0.0925 (27)	1.091 (0.619-1.922)	0.764
98	rs7579737 (A/G)	1 2	0.742 (193) 0.258 (67)	0.692 (202) 0.308 (90)	0.779 (0.537-1.131)	0.188
99	rs6758936 (G/A)	1 2	0.515 (134) 0.485 (126)	0.486 (142) 0.514 (150)	0.890 (0.637-1.244)	0.495
100	rs2270297 (G/A)	1 2	0.765 (199) 0.235 (61)	0.812 (237) 0.188 (55)	1.321 (0.876-1.991)	0.183
101	rs7558013 (C/A)	1 2	0.815 (212) 0.185 (48)	0.760 (222) 0.240 (70)	0.718 (0.475-1.085)	0.114
102	rs35582281 (A/G)	1 2	0.946 (246) 0.0539 (14)	0.986 (277) 0.0514 (15)	1.051 (0.497-2.221)	0.897
103	rs1035130 (G/A)	1 2	0.719 (187) 0.281 (73)	0.702 (205) 0.298 (87)	0.920 (0.636-1.330)	0.657
106	rs3732127 (A/C)	1 2	0.781 (203) 0.219 (57)	0.818 (239) 0.182 (53)	1.266 (0.834-1.924)	0.268
107	rs36036501 (A/C)	1 2	0.935 (243) 0.0658 (17)	0.928 (271) 0.0719 (21)	0.903 (0.466-1.751)	0.762
108	rs4851575 (A/G)	1 2	0.764 (197) 0.236 (61)	0.812 (237) 0.188 (55)	1.334 (0.885-2.012)	0.168
109	rs917997 (G/A)	1 2	0.758 (197) 0.242 (63)	0.805 (235) 0.195 (57)	1.318 (0.879-1.977)	0.181
110	rs4140786 (C/A)	1 2	0.508 (132) 0.492 (128)	0.527 (154) 0.473 (138)	1.082 (0.775-1.512)	0.644
111	rs759381 (T/A)	1 2	0.765 (199) 0.235 (61)	0.767 (224) 0.233 (68)	1.010 (0.680-1.499)	0.962

SNP Nº	rs Number	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
112	rs1880000 (A/G)	1 2	0.868 (223) 0.142 (37)	0.853 (249) 0.147 (43)	0.961 (0.597-1.545)	0.869
113	rs10207579 (A/G)	1 2	0.554 (143) 0.446 (115)	0.497 (144) 0.503 (146)	0.793 (0.567-1.111)	0.177
114	rs2075190 (T/A)	1 2	0.735 (191) 0.265 (69)	0.695 (203) 0.305 (89)	0.824 (0.568-1.194)	0.306
115	rs1403550 (G/A)	1 2	0.758 (197) 0.242 (63)	0.764 (223) 0.236 (69)	1.034 (0.699-1.529)	0.869
116	rs2177317 (G/A)	1 2	0.619 (161) 0.381 (99)	0.616 (180) 0.384 (112)	0.988 (0.701-1.394)	0.946
118	rs6761291 (A/G)	1 2	0.586 (150) 0.414 (106)	0.569 (166) 0.431 (126)	0.931 (0.663-1.308)	0.680
119	rs1861229 (G/A)	1 2	0.835 (217) 0.165 (43)	0.849 (248) 0.151 (44)	1.117 (0.706-1.766)	0.636
121	rs873625 (A/G)	1 2	0.654 (170) 0.346 (90)	0.603 (176) 0.397 (116)	0.803 (0.568-1.136)	0.215
122	rs7603234 (C/G)	1 2	0.923 (240) 0.0769 (20)	0.938 (274) 0.0616 (18)	1.269 (0.656-2.454)	0.480

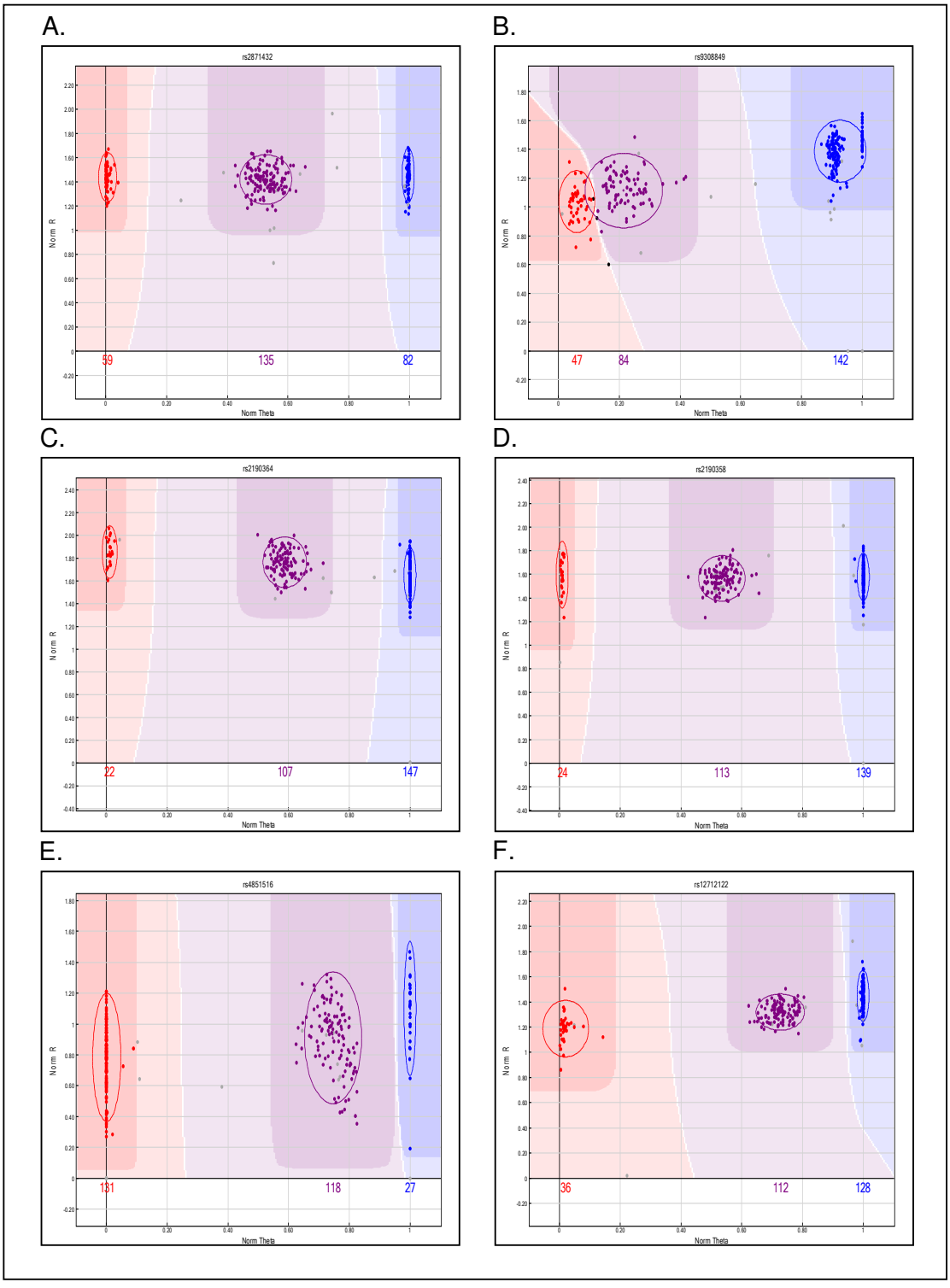
Table 4.39 Non-significant *IL1* receptor cluster candidate region stage-1 analysis results

One hundred and sixteen genotyped tSNPs within the *IL1* receptor cluster candidate region were analysed for significant allele frequency differences between the patient (n=130) and control (n=146) cohorts. The frequency in each cohort of both the common (1) and rare (2) alleles are given. The alternative alleles of each SNP are shown underneath the rs identifier. The allele assigned as the common (1) allele is given first and the rarer (2) allele second. The OR shown is for allele 2 (allele 1 OR=1). The p-value for the test for frequency differences between the cohorts is also shown. The results for the 108 tSNPs which did not show significant frequency differences (p>0.05) between the two cohorts are shown.

in the control population. The frequency differences for SNPs 1 and 4 only just reached significance, p= 0.0428 and 0.0376 respectively. None of the SNPs satisfied the significance level following Bonferroni correction (0.05/116 SNPs analysed= p< 0.0004). However, Bonferroni correction assumes that the variables tested are independent, and is therefore overly conservative in this study as the SNPs tested will be in some amount of LD with each other (Lewis, 2002).

The genoplots for each SNP showing significant frequency differences are shown in Figure 4.32. Apart from SNP2 (B) the other eight SNP assays showed three discernible, well separated clusters, enabling high confidence in the genotype assignment, and therefore that the significant differences observed are due to true frequency differences between the populations and not genotyping error. Although the GC10 score for the SNP2 assay did exceed the 0.5 threshold (GC10 = 0.565, Appendix 2) it did so only marginally, and was substantially lower than the GC10 score for the other SNPs (SNP1=0.954, SNP4=0.823, SNP6=0.965, SNP7=0.834, SNP8=0.827, SNP14=0.925, SNP23=0.731). This suggests that there may be ambiguity regarding the genotype assignment. The reason for the low GC10 score is evident from the assay genoplot, in which, rather than three distinct and well separated genotype clusters, four clusters could be discerned. Automatic clustering performed by the Beadstudio software assigned the two right hand side clusters as 'BB' and the two left hand clusters as 'AA' and 'AB'. It is possible however that the two clusters on the left hand side are 'AA', and the right hand side clusters are the 'AB' and 'BB' clusters respectively. This cluster definition results in assignment of the samples as: AA n=134, AB n=114, and BB n=28. This alternative genotype distribution does not deviate significantly from HWE, and therefore this cannot be discounted as a viable alternate cluster definition. Due to the potential inaccuracy of the genotype calling the analysis results of this SNP must be viewed with caution.

To aid identification of secondary effects, the LD relationships between these eight SNPs were determined. As a tSNP approach to SNP selection was employed, none of the genotyped SNPs were in high LD with each other according to the r^2 statistic. However, r^2 is a measure of the LD relationship between alleles at different loci, but the alternative LD statistic D′ is a measure of the evidence of recombination between loci (section 3.1.1.1.). While r^2 is the most appropriate measurement in identifying SNPs which are proxies for each other, D′ is important when indentifying SNPs which are inherited together.


A.
rs2871432
Norm R
Norm Theta
59
135
82
B.
rs9308849
Norm R
Norm Theta
47
84
142
C.
rs2190364
Norm R
Norm Theta
22
107
147
D.
rs2190358
Norm R
Norm Theta
24
113
139
E.
rs4851516
Norm R
Norm Theta
131
118
27
F.
rs12712122
Norm R
Norm Theta
36
112
128

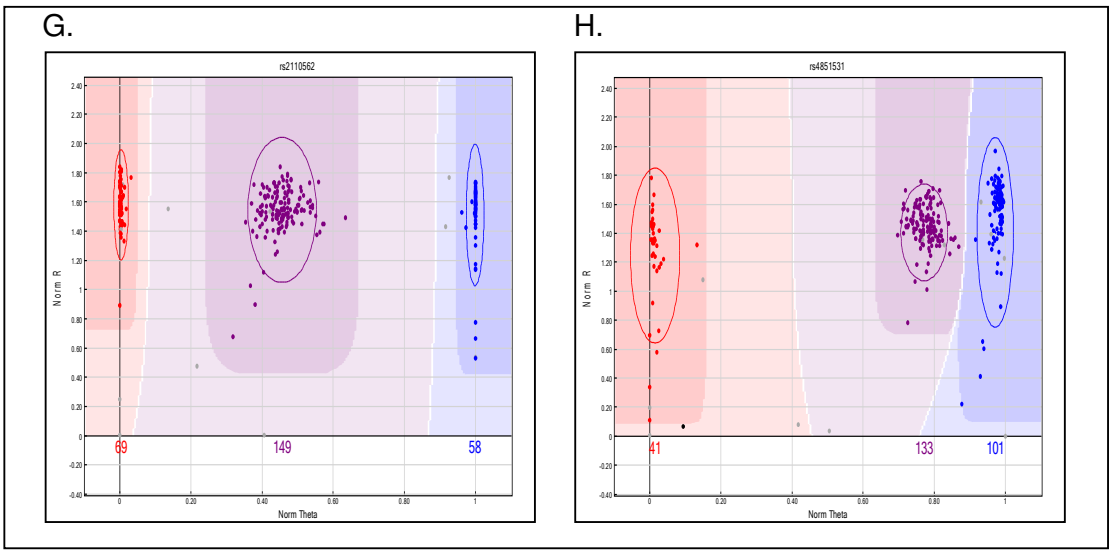


Figure 4.32 Genoplots of the significant *IL1* receptor cluster SNPs

Shown are the genoplots of the eight *IL1* receptor cluster region SNPs which showed evidence of significant frequency differences in the initial stage-1 analysis. Normalised theta values (deviation from total 'A' allele signal) are plotted along the x-axes, and normalised R values (signal intensity) along the y-axes. Each data point represents one sample, the circles indicate the location of each genotype cluster, and the dark shaded areas are the call areas.

A. SNP1 rs2871423, B. SNP2 rs9308849, C. SNP4 rs2190364, D. SNP6 rs2190358, E. SNP7 rs4851516, F. SNP8 rs12712122 G. SNP14 rs2110562, H. SNP23 rs4851531.

The D′ LD between the eight SNPs identified as showing significant allele frequency differences between cohorts in the initial analysis is shown in Figure 4.33. As the SNPs did not separate into clear blocks of LD each of the SNPs was analysed conditioning on SNP8 which showed the most significant frequency difference in the initial analysis (p=0.00313). None of the SNPs showed evidence of significant disease association once the effect of SNP8 had been accounted for (Table 4.40). Conversely, when SNP8 was tested for association conditioning on each of the other SNPs, except for SNP23, there remained a significant association effect. This indicates that the frequency differences displayed by SNPs 1, 2, 4, 6, and 7 were secondary effects, created by their LD relationship with SNP8. They were therefore not directly associated with disease state and eliminated as candidates. The SNP8 allele frequency difference did not remain significant following conditional analysis with SNP23 (p=0.128), despite SNP23 not showing a significant difference once the effect of SNP8 was taken into account (p=407). Because of the uncertain relationship between SNP23 and SNP8, SNP23 was not eliminated, and both SNPs were included in the second stage of the association study.

The tSNP selection based on the population genotyping data from HapMap2 determined that SNP8 was tagging four other SNPs within the candidate region, and that SNP23 was not tagging any other SNPs (Table 4.36). For ease of identification the SNPs captured by SNP8 were designated, according to chromosomal order, as SNP8a (rs1010329), SNP8b (rs2190361), SNP8c (rs2190360), and SNP8d (rs740044). SNP8 was in complete LD (r^2=1) with SNPs 8b and 8c, but was an imperfect proxy for SNP8a (r^2=0.855) and SNP8d (r^2=0.820). These LD relationships were confirmed in the HapMap3 data, r^2=0.875, 1, 1, and 0.868 for SNPs8a-d respectively (Table 4.5). Genotyping data was not available from WTCCC2 to further confirm the LD in that population. As the genotyped SNP8 was not a perfect proxy for SNP8a and SNP8d, meaning that the SNP frequencies will not be exactly the same and the association results may differ, these two SNPs were subsequently genotyped directly in the stage-1 populations and tested for frequency differences (Table 4.41). The minor alleles of both SNP8a and SNP8d were at a significantly higher frequency in the patient than control cohort, although SNP8d only just reached significance (p=0.04).

4.5.1.3. Stage-2

Four of the *IL1* receptor cluster candidate region SNPs genotyped in stage-1 of the association analysis were included in stage-2 of the study. These were: SNPs 8, 8a, 8d, and

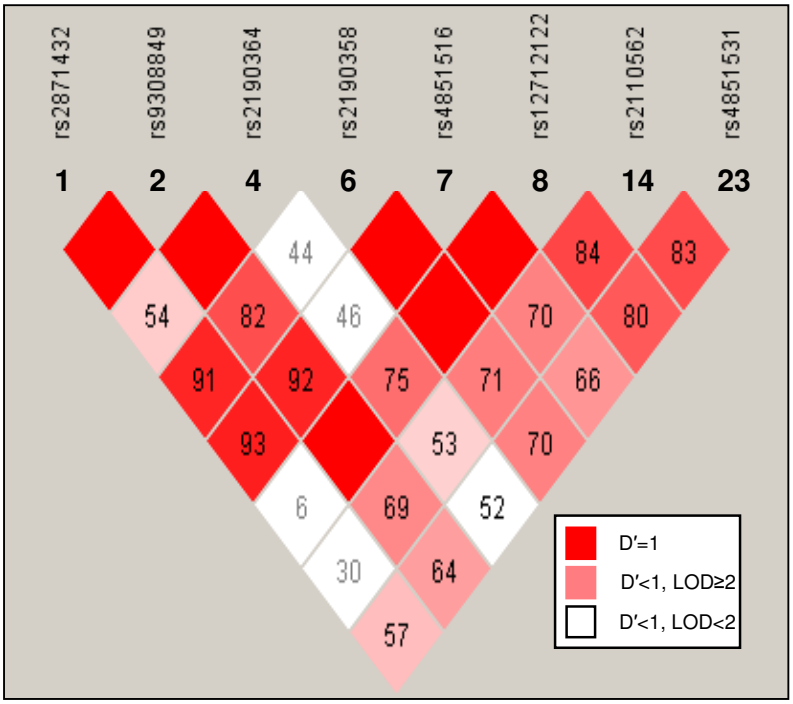


Figure 4.33 D′ relationship of associated *IL1* receptor cluster region SNPs

In the initial analysis eight SNPs within the *IL1* receptor cluster candidate region showed significant allele frequency differences between the case and control populations. The graphical representation of the D′ LD relationship between these eight SNPs, generated in Haploview, is shown on a sliding colour scale from white (low D′) to red (high D′), depending on LOD score. The D′ value for each pairwise LD comparison is shown, where no value is indicated D′=1.

LOD= logarithm of odds

Test marker	Conditioning marker	p value
1	8	0.188
2	8	0.156
4	8	0.426
6	8	0.226
7	8	0.297
14	8	0.572
23	8	0.407
8	1	0.0186
8	2	0.0451
8	4	0.0470
8	6	0.0368
8	7	0.0392
8	14	0.0445
8	23	0.128

Table 4.40 Conditional analysis of associated *IL1* receptor region SNPs

In the initial analysis eight SNPs within the *IL1* receptor cluster candidate region showed significant allele frequency differences between the case and control populations. Each of the SNPs were tested for significant frequency differences conditioning on SNP8, which had shown the most significant frequency difference (p=0.00313). SNP8 was also tested conditioning on each of the SNPs.

SNP	rs Nº	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
8a	rs1010329 (C/T)	1	0.591 (162)	0.690 (265)	1.54 (1.133-2.093)	0.00897
		2	0.409 (112)	0.310 (119)		
8d	rs740044 (G/T)	1	0.658 (179)	0.732 (287)	1.42 (1.027-1.963)	0.0409
		2	0.342 (93)	0.268 (105)		

Table 4.41 Incompletely captured *IL1* receptor cluster candidate region SNPs

SNP8 in the *IL1* receptor candidate region was acting as a tSNP for four other SNPs within the region. SNP 8 was not in complete LD (r^2<1) with SNPs 8a and 8d, and therefore not a perfect proxy. Both of these SNPs were subsequently directly genotyped in the stage-1 patient (n=130) and control (n=146) cohorts. The frequency in each cohort of both the common (1) and rare (2) alleles are given. The OR shown is for allele 2 (allele 1 OR=1). The p-value for the test for frequency differences between the cohorts is also shown.

23. None of the SNPs showed any significant ($p<0.05$) frequency differences between the case (n=105) and control (n=184) cohorts used in stage-2 (Table 4.42).

4.5.1.3.1. Stratified analysis

Following analysis of the data from both stages of the study, stratified by study stage, only one of the four *IL1* receptor cluster SNPs showed evidence of significant frequency differences between the patient and control cohorts (Table 4.43). The frequency of the SNP8 minor allele was significantly higher in the patient population than the control population, although only marginally (CMH p=0.0467). There was also however significant heterogeneity between the results for the two study stages for this SNP (BD p=0.0287). Neither of the SNPs which were in high but incomplete LD (SNP8a r^2=0.855 and SNP8d r^2=0.820) with SNP8 were significantly different between the populations over both stages (CHM p=0.114 and 0.162 respectively).

4.5.1.4. WTCCC control cohort analysis

An additional analysis with greater statistical power to detect SNPs associated with disease was performed with the pooled stage-1 and stage-2 cases (n=235), and the WTCCC2 cohort of healthy UK individuals (n=4,671).

Genotyping data were not available in the WTCCC2 control population for *IL1* receptor cluster SNP8 (rs12712122), SNP8a (rs1010329), or SNP73 (rs7627123). For analysis with SNP8 in the patient cohort, SNP8c in the WTCCC2 control cohort was used. These two SNPs were in high LD with each other according to both the HapMap2 (r^2=1) and HapMap3 (r^2=0.963) population datasets (section 4.1.5.). Control genotyping data for SNP73a (rs1393049) (SNP73a, r^2=1 in HapMap 2 and 3) was used for analysis with SNP73 in the patients.

With this larger UK wide healthy control cohort none of the tested SNPs, even the one which showed evidence of significant disease association in the stratified analysis, showed a significant allele frequency difference between the patient (n=235) and control (n=4671) cohorts (Table 4.44).

4.5.1.5. Associated SNPs

In the initial tSNP selection based on the HapMap 2 genotyping data (Table 4.36) SNP8 was

SNP	Allele	Case freq. (N^{o})	Control freq. (N^{o})	OR (95% CI)	p value
8	1 2	0.628 (123) 0.372 (73)	0.621 (226) 0.379 (138)	0.972 (0.679-1.391)	0.876
8a	1 2	0.626 (124) 0.373 (74)	0.595 (125) 0.405 (85)	0.878 (0.594-1.297)	0.521
8d	1 2	0.672 (133) 0.328 (65)	0.661 (144) 0.339 (74)	0.951 (0.625-1.446)	0.809
23	1 2	0.605 (121) 0.395 (79)	0.632 (220) 0.368 (128)	1.22 (0.785-1.605)	0.528

Table 4.42 *IL1* receptor candidate region stage-2 analysis results

The four *IL1* receptor candidate region SNPs identified as showing significant frequency differences following conditional analysis in stage-1 of the association study, were genotyped in the stage-2 patient (n=105) and control (n=184) cohorts. The frequency in each cohort of both the common (1) and rare (2) alleles of each SNP are given. The OR shown is for allele 2 (allele 1 OR=1). The p-value for the test for frequency differences between the cohorts is also shown.

SNP	CMH p value	OR (95% CI)	BD p value
8	0.0467	1.290 (1.004-1.659)	0.0287
8a	0.114	1.223 (0.952-1.57)	0.0282
8b	0.162	1.202 (0.928-1.556)	0.129
23	0.150	0.833 (0.649-1.069)	0.0226

Table 4.43 *IL1* receptor cluster candidate region stratified analysis results

For the four SNPs investigated in both stages of the study, association meta-analysis of the data was performed using the CMH test, stratifying by study stage. A significant result for the CHM test shows that the SNP is associated with disease. The p value and OR for this test are given, the OR shown is for allele 2 (allele 1 OR=1). A BD test for homogeneity of odds ratios between the two stages was also performed. A significant BD test shows that the allele frequency difference is more prominent in one of the populations studied. Any association is therefore more likely to be due to sampling error rather than true association.

BD-Breslow-Day, CI-Confidence Interval, CMH-Cochran-Mantel-Haenszel, OR-Odds Ratio

SNP	Allele	Case freq. (N^{o})	Control freq. (N^{o})	OR (95% CI)	p value
8/8c	1	0.614 (280)	0.645 (6022)	1.14	0.186
	2	0.386 (176)	0.355 (3320)	(0.939-1.383)	
8d	1	0.665 (308)	0.687 (6419)	1.108	0.305
	2	0.335 (146)	0.313 (2923)	(0.851-1.274)	
23	1	0.638 (291)	0.612 (5716)	0.894	0.261
	2	0.362 (165)	0.388 (3624)	(0.894-1.087)	

Table 4.44 *IL1* receptor candidate region analysis results with WTCCC2 controls

As an additional analysis with a larger, independent control cohort, allele frequencies in the pooled cases from both stages of the association study (n=235) were compared to the controls genotyped in phase-2 of the WTCCC project (n=4671). Genotyping data for SNP8 (rs12712122) in cases and SNP8c (rs2190360) (r^2=1 in HapMap2, 0.963 in HapMap3) in controls, and for rs7627123 (SNP73) in cases and rs1393049 (SNP73a) (r^2=1 HapMap 2 and 3) in controls, were used for analysis. The frequency in each cohort of both the common (1) and rare (2) alleles of the two individual SNPs are given. The OR shown is for allele 2 (allele 1 OR=1). The p-value for the test for frequency differences between the cohorts is also shown.

tagging four other SNPs, labelled as SNPs 8a-d. Direct genotyping and analysis of SNPs 8a and 8d eliminated them as being associated with sJIA (Table 4.43). SNPs 8b (rs2190361) and 8c (rs2190360) are in complete LD (r^2=1) with SNP8, according to both HapMap2 and HapMap3 (Table 4.5), genotyping data was not available from WTCCC2.

4.5.1.5.1. Additional captured SNPs

Based on the HapMap3 dataset nine SNPs were identified by SNAP (section 3.1.5.) as being in r^2>0.8 with SNP8 (Table 4.45). All of these additional tagged SNPs are, like SNP8, upstream of *IL1R2*.

4.5.1.5.2. Associated SNP positions

The positions of *IL1* receptor cluster SNP8, which showed significant evidence of association, and the 11 SNPs which it is tagging, are shown in Figure 4.34. SNP8 and all SNPs it is tagging are upstream of *IL1R2*. SNP8 is 16,309bp upstream of the *IL1R2* transcription start site (TSS). The SNP8 captured SNPs are at: -37,221 (SNP8b rs2190361), -36,430 (rs6756338), -34,964 (rs1859717), -34,828 (rs1859716), -33,586 (rs1019033), -33,490 (rs1019032), -33,468 (rs1019031), -32,208 (SNP8c rs2190360), -20,004 (rs12612902), -19,048 (rs12466288), and -4,007 (rs6543105) relative to *IL1R2* TSS.

4.5.1.6. Summary

Following analysis of the data from both study stages, one SNP within the *IL1* receptor cluster candidate region, rs12712122 (SNP8) was identified as showing evidence of a significant association with sJIA. This was not however replicated in analysis with the WTCCC2 control cohort. SNP8 is in high LD (r^2>0.8) with a total of eleven other SNPs within the candidate region. All of the associated SNPs are located between 37.2 and 4kb upstream of *IL1R2*.

These association study results for this candidate region, prior to the WTCCC2 analysis, are published (Stock et al., 2008), see Appendix 4.

SNP	r^2	Location
rs6756338	0.963	Upstream *IL1R2*
rs1859717	0.963	Upstream *IL1R2*
rs1859716	0.963	Upstream *IL1R2*
rs1019033	0.963	Upstream *IL1R2*
rs1019032	0.963	Upstream *IL1R2*
rs1019031	0.963	Upstream *IL1R2*
rs12612902	0.963	Upstream *IL1R2*
rs12466288	0.963	Upstream *IL1R2*
rs6543105	0.825	Upstream *IL1R2*

Table 4.45 Additional SNPs tagged by *IL1* receptor cluster SNP8

Based in the HapMap3 population genotyping data all SNPs in r^2>0.8 with, and within 500kb of, the *IL1* receptor cluster SNP showing evidence of significant disease association, were identified. The SNP identifier, r^2 value, and location of all the identified SNPs are shown.

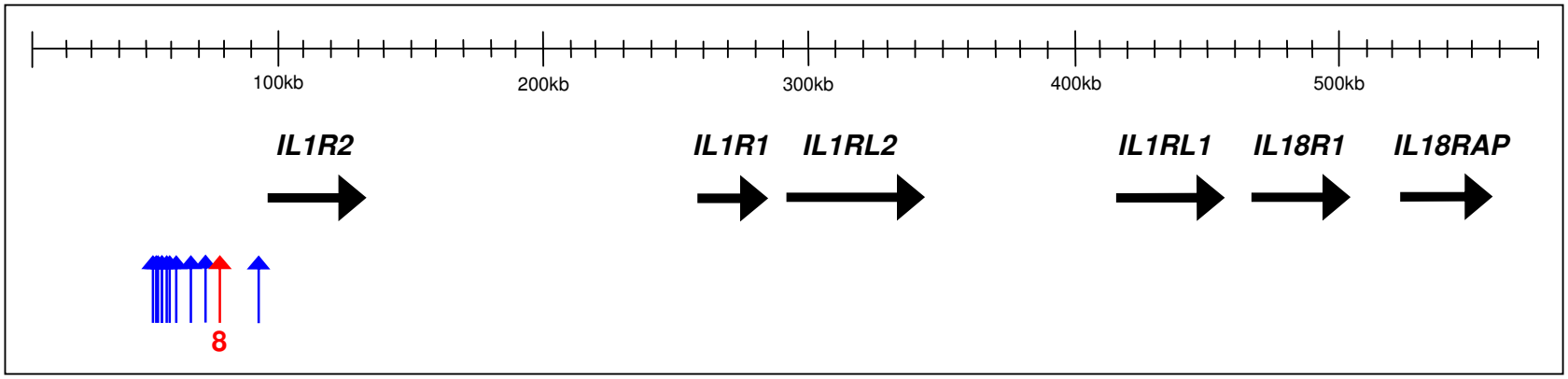


Figure 4.34 Positions of the associated SNPs in the *IL1* receptor cluster candidate region

The position within the *IL1* receptor cluster candidate region of SNP8 (rs12712122), which showed evidence of significant association, is shown in red, and the 11 SNPs tagged by SNP8 are shown in blue. Due to the scale of the figure only SNP8 is labelled. The tagged SNPs in chromosomal order are: rs2190361 (SNP8b), rs6756338, rs18599717, rs1859716, rs1019033, rs1019032, rs1019031, rs2190360 (SNP8c), rs12612902, rs12466288, and rs6543105.

4.5.1.7. Discussion

Six members of the IL-1 receptor gene family are located together in a cluster on chromosome 2. The cluster contains both the type 1 IL-1 receptor (*IL1R1*), which is required for IL-1 signal induction, and the non-signalling decoy receptor type 2 IL-1 receptor. The genes encoding the IL-18R1 and IL-18RAP receptors, both of which are required for IL-18 signal induction, are also located within the cluster. Also included in the cluster are the genes for the two IL-1 receptor-like receptors, which are involved in signal induction by other IL-1 family ligands.

Polymorphisms within the genes included in this cluster have been associated with a number of diseases. The SNP rs59339921 in *IL1R1*, was found to be significantly linked to type 1 diabetes mellitus in a family based study of a Danish population, and to have an allele dose effect on plasma levels of IL-1R1 in both patients and healthy controls (Bergholdt et al., 2000). This SNP was not included in this association study and there is no population frequency data available, suggesting that it may be at a low frequency (<0.05) in the general population. A genome wide association study in ankylosing spondylitis identified a significant association with the *IL1R2* SNP rs2310173 (Reveille et al., 2010), which was included in this study (SNP4), but did not show association with sJIA. Of nine SNPs within *IL1RL1, IL118R1*, and *IL18RAP*, associated with bronchial hyperresponsiveness (Reijmerink et al., 2008), seven were captured in this study, but were not associated with disease. The SNP rs917997, which is 1.5kb downstream of *IL18RAP*, has been associated with Crohn's disease and ulcerative colitis (Zhernakova et al., 2008), *IL18RAP* mRNA expression levels in PBMCs from untreated coeliac disease patients (Hunt et al., 2008), and, in a haplotype with rs13015714, with coeliac disease in a Hungarian population (Koskinen et al., 2009). This SNP, SNP109 in the study presented in this thesis, was not however associated with sJIA. It has been shown that patients with sJIA have a defect in the phosphorylation of IL-18RAP following binding of IL-18 (de Jager et al., 2009), however, no SNPs within this gene, or flanking regions, showed significant frequency differences between the patient and control cohorts.

Only one SNP in this candidate region, rs12712122 (SNP8) showed evidence of significant disease association in the stratified meta-analysis. However, the association of this SNP only just reached statistical significance (CMH p=0.0467) and also showed significant heterogeneity of odds between the two study stages (BD p=0.287). No significant allele

frequency difference was seen when the analysis was repeated comparing the patients from both study stages with the large WTCCC2 control cohort. However, because SNP8 was not included in the WTCCC2 study it was necessary to compare the frequency of SNP8 in the patients with that of rs2190360 in the controls. Although these two SNPs were in complete LD (r^2=1) according to the HapMap2 data, in the larger HapMap3 data the LD relationship was incomplete (r^2=0.963), so this analysis may be misleading. Due to the low significance level, the evidence for this SNP being associated with disease susceptibility is therefore not robust, and may be a spurious result due to the different populations used in the two stages of the study.

It is however also possible that SNP8 is only acting as a marker for a disease associated SNP, with which it is in incomplete LD. If this is the case it would be expected that the effect size, and also the significance level, of SNP8 would be less than the true effect size of the causal SNP. This possibility is supported by a recent study by Hinks et al. who, based on the published results of this study (Stock et al., 2008), included this polymorphism in an association study of autoinflammatory genes with susceptibility to JIA (Hinks et al., 2011) (Submitted at time of writing). Although there was no evidence of association with the JIA cohort including all seven subtypes (n=1054), when analysed separately there was significant association (pTrend=0.02) with the systemic JIA subtype (n=147). It is important to note that only seven of the 147 sJIA patients were not included in the study presented in this thesis (95.2% overlap), and is therefore not an independent replication. As the WTCCC2 control cohort, in which SNP8 had not been directly genotyped, was used in their study, Hinks et al investigated rs2190360 as a proxy for SNP8. As stated previously, these two SNPs are not in complete LD (r^2=0.963) according to the HapMap3 data set. It is therefore possible that rs2190360 is a better representation of the association at this locus, although direct testing in the whole patient cohort would be required for confirmation.

In addition to rs2190360, SNP8 tags another 12 SNPs within the candidate region, r^2=1 with two. The two SNPs (rs1010329, SNP8a, and rs740044, SNP8d) known at the time of the association study to be incompletely tagged, were themselves directly genotyped in all of the patient and control samples, but did not show significant association (Table 4.43). Direct analysis of the remaining 11 incompletely tagged SNPs, including rs2190360, would be

necessary to elucidate the validity of this potential disease association, and to determine the causal SNP. It was not possible to include this as part of this study due to time constraints.

SNP8, and all of the SNPs it is tagging, are located upstream of *IL1R2*, which encodes the non-signalling decoy IL-1 receptor. There are two isoforms, a membrane bound, and a secreted form, of this gene produced through alternative splicing, both of which are IL-1 antagonists. It has been shown that patients with adult polyarthritis have higher plasma levels of the secreted forms of both IL-1R1 (sIL-1R1) and IL-1R2 (sIL-1R2) than healthy controls. Levels of sIL-1R2 were also higher in patients with non-destructive arthritis than those with destructive disease, and were negatively correlated with indices of joint destruction (Jouvenne et al., 1998b), demonstrating the importance of IL-1R2 in ameliorating the IL-1 inflammatory response. Based on their location relative to the gene sequence it is possible that the SNP8 related SNPs may be involved in regulating expression of this gene. Reduced expression of *IL1R2* would lead to reduced IL-1 response inhibition, and potentially to a maintained inflammatory state.

The issues of low power to detect significant associations, multiple testing, and the lack of significant findings in the stage-2 populations alone are discussed in chapter 5.

4.6. *IL18* candidate region

4.6.1. Association Study

In order to determine the effect of variation within nearby regulatory elements, the region investigated was extended beyond *IL18* to the next flanking genes. These were identified using the July 2003 human genome reference sequence (NCBI build 34), hg16 annotation track. Based on the positions of these genes the candidate region was defined as chr11:111504144-111575744 (Figure 4.35). The genes flanking *IL18* are: (3′) succinate dehydrogenase complex, subunit D (*SDHD*), and (5′) testis expressed sequence 12 (*TEX12*). *SDHD* is a subunit of the succinate oxidising enzyme, involved in both the citric acid cycle and the electron transport chain (Hirawake et al., 1997). *TEX12* encodes a structural part of the synaptonemal complex, which forms between homologous chromosomes during meiosis (Hamer et al., 2006). The identified candidate region spans 71kb (71,600bp), and extends 47.5kb (47,491bp) 3′, and 3.3kb (3,280bp) 5′, of *IL18*.

4.6.1.1. tSNP selection

Genotyping data was available for a total of 114 SNPs within the selected *IL18* candidate region, 91 from HapMap and 37 from PGA. Of these, 14 were included in both population genotyping resources, and 23 had data from only PGA. Sixty four of the SNPs in the region were monomorphic in the CEU population, and a further 13 had a MAF<0.05. Due to the low population frequencies of these 77 SNPs they were not included in the LD analysis. A total of 52 SNPs within the *IL18* candidate region satisfied the inclusion criteria and were therefore included in the LD investigation and subsequent tSNP selection.

A graphical representation of the LD between these 52 SNPs in the *IL18* candidate region is shown in Figure 4.36. There was no evidence of strong LD over the candidate region. There was a block of strong LD between six SNPs at the 5′ end of *IL18*, and another between three SNPs in the most 3′, of the gene sequence, 10kb of the candidate region. These were however the only areas of maintained high LD between neighbouring SNPs. There was also a discontinuation of LD between the SNPs within and proximal to *IL18*, SNPs 36 to 115, and those 3′ of the gene. SNPs 5 to 35 showed minimal amounts of LD with the SNPs downstream. The exception to this was SNPs 34 and 35 which are in high LD with SNPs 77 and 80, and are both tagged by SNP 80. There was then a further, and more marked,

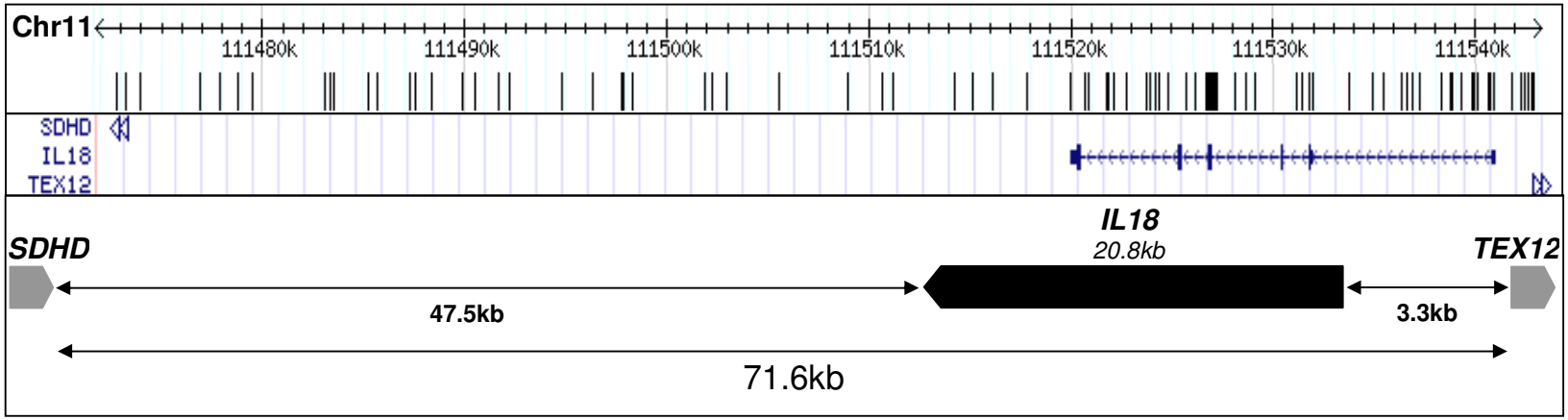


Figure 4.35 *IL18* candidate region

The boundaries of the *IL18* candidate region are the flanking genes *SDHD* and *TEX12*, as determined by the July 2003 human genome reference sequence (NCBI build 34), hg16 annotation track, shown in the centre panel. The 3′ end of *SDHD* and the 5′ end of *TEX12* are proximal to *IL18*. The approximate (due to scale) positions within this region of all SNPs (irrespective of MAF) in the HapMap phase 2 database (data release #18/phaseII Sept05) are shown in the top panel. The bottom panel shows a schematic (not to scale) of the candidate region with gene orientation, size, and distances given. The gene of interest is shown in black and all other features within the region in grey. All sizes and distances are based on chromosomal positions reported in the NCBI build 34 genome reference sequence.

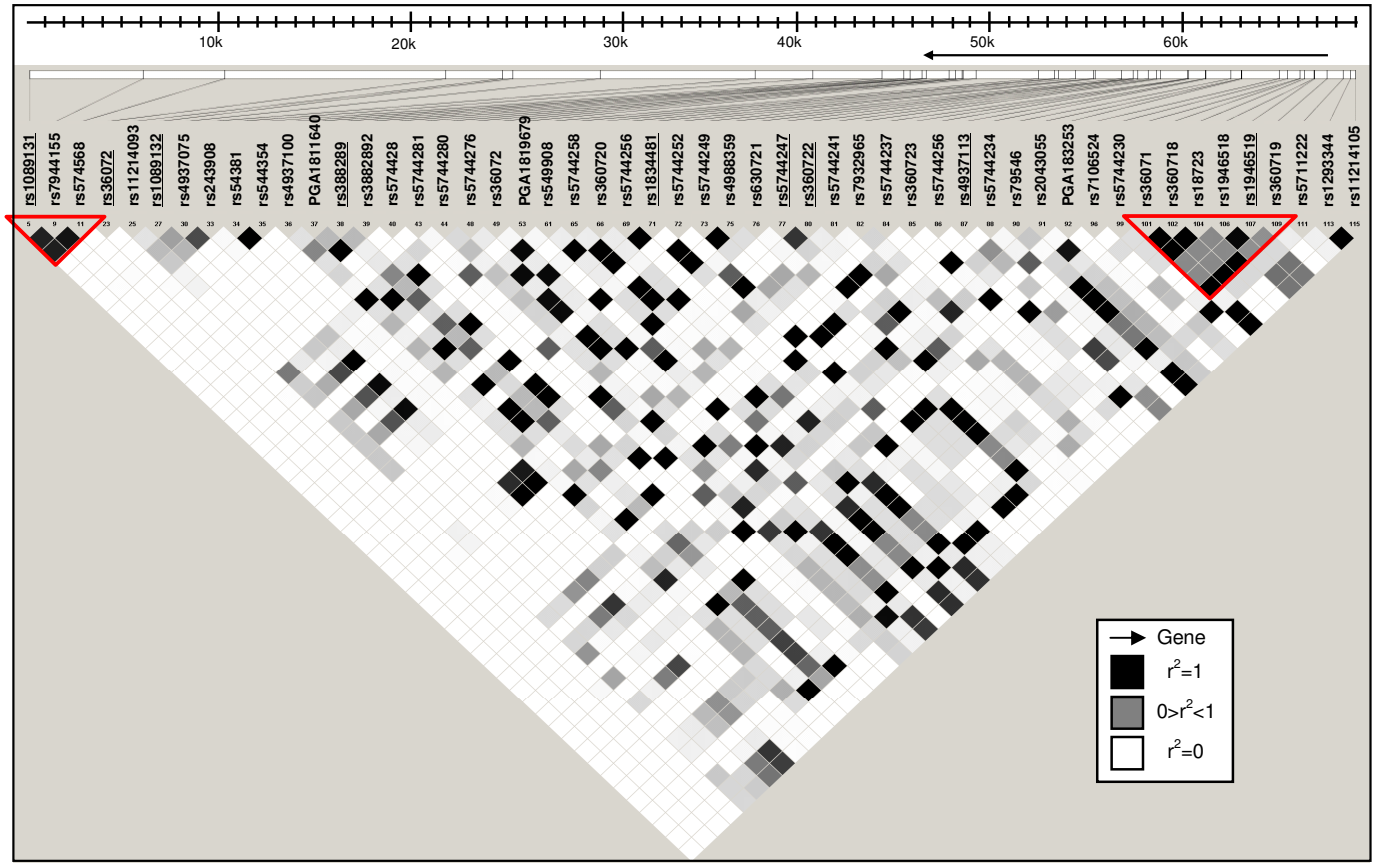


Figure 4.36 LD pattern over the *IL18* candidate region

The pairwise LD relationships between SNPs in the *IL18* candidate region were investigated to enable tSNP selection. The region shown is comprised of only the SNPs satisfying the inclusion criteria, not the whole of the candidate region. Gene orientation and location within the region is indicated. The position within the region of each SNP satisfying the selection criteria is represented by a line underneath the gene position. Each coloured square represents the pairwise LD between two SNPs. The graphical representation, generated in Haploview, is shown on a grey scale, with higher r^2 LD represented by darker blocks. Each SNP is labelled by rs number and Haploview ID number. The 13 tSNPs selected for genotyping are underlined, and the blocks of high LD indicated.

breakdown in the LD, with SNPs 5 to 23 showing no LD with any SNPs within *IL18*. SNP 23 shows no LD (r^2=0) with any of the other SNPs in the region.

Due to the LD relationships between the SNPs a total of 15 tSNPs (29% of the total) were identified as the minimal subset required to capture the variation within all 52 SNPs. However, following assay design (section 3.1.2.1.) it was necessary to exclude three of the SNPs due to poor assay design scores. The final SNP set therefore consisted of 13 tSNPs which captured the variation within 49 SNPs. Six of the tSNPs were not in high LD ($r^2 \leq 0.8$) with any other SNP within the region and so were not tagging any additional SNPs. All of the selected tSNPs, and the SNPs captured by each, are listed in Table 4.46.

4.6.1.2. Stage-1

All 13 of the genotyped *IL18* candidate region SNPs satisfied the quality control criteria (Sections 3.1.2.2.1. and 3.1.2.2.2.) and were therefore analysed for evidence of disease association.

None of the *IL18* candidate region tSNPs showed significant ($p<0.05$) frequency differences between the patient (n=130) and control (n=146) populations in either individual or haplotype analysis (Table 4.47). Therefore, further investigation of this candidate region was not pursued in stage-2 of the study.

4.6.2. Summary

No evidence was found for significant association of polymorphisms within the *IL18* candidate region and susceptibility to sJIA.

4.6.3. Discussion

IL-18 is a member of the IL-1 gene family constitutively expressed by PBMCs, mainly from monocytes/macrophages and keratinocytes. It is a pro-inflammatory cytokine which, with IL-12 as a co-stimulant, induces the production of IFNγ. It has been shown that, compared to healthy controls and patients with other JIA subtypes, patients with sJIA have highly increased levels of IL-18, which mirrors disease activity (Maeno et al., 2002). Similarly, increased levels of IL-18 have been reported in patients with adult-onset Still's disease (Kawashima et al., 2001), lupus nephritis (Liang et al., 2006),

SNP Nº	rs Number	Tagged SNPs
1	rs10891319	rs7944155,rs574568
2	rs360726	n/a
3	rs11214093	n/a
4	rs10891323	n/a
5	rs4937075	n/a
6	rs3882891	n/a
7	rs549908	rs243908,rs360729,rs360720,rs360721,rs795467,rs360717,rs360718,rs187238,rs360719,rs1293344,rs11214105
8	rs1834481	rs5744292,rs3882892,rs5744276,rs5744258,rs5744256,rs5744249,rs4988359,rs5744230,rs5744222
9	rs5744247	rs543810,rs544354,rs5744281,rs360723,rs5744232
10	rs360722	rs11267792,rs5744252,rs360723,rs5744234
11	rs4937113	n/a
12	rs2043055	rs4937100,rs5744280,rs7932965,rs7106524
13	rs1946519	rs1946518
Uncaptured		rs5744289,rs5744241, rs5744237

Table 4.46 *IL18* candidate region tSNPs

A total of 13 tSNPs were genotyped to capture 49 of the 52 SNPs within the candidate region. The SNPs captured ($r^2 \geq 0.8$) by each genotyped tSNP are shown. For ease of identification each tSNP genotyped was assigned a number according to genome order. Due to poor assay design scores it was not possible to genotype or capture three of the SNPs within the region.

SNP Nº	rs Number	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
1	rs10891319 (A/G)	1 2	0.615 (160) 0.385 (100)	0.641 (186) 0.359 (104)	1.118 (0.791-1.581)	0.529
2	rs360726 (A/G)	1 2	0.789 (205) 0.212 (55)	0.771 (225) 0.230 (67)	0.901 (0.602-1.349)	0.612
3	rs11214093 (A/G)	1 2	0.612 (158) 0.388 (100)	0.627 (183) 0.377 (109)	1.063 (0.753-1.5)	0.730
4	rs10891323 (A/G)	1 2	0.885 (230) 0.115 (30)	0.907 (263) 0.0931 (27)	1.271 (0.734-2.2)	0.392
5	rs4937075 (C/G)	1 2	0.681 (177) 0.319 (83)	0.726 (212) 0.274 (80)	1.243 (0.862-1.792)	0.245
6	rs3882891 (A/C)	1 2	0.627 (163) 0.373 (97)	0.600 (175) 0.400 (117)	0.890 (0.631-1.255)	0.506
7	rs549908 (A/C)	1 2	0.713 (184) 0.287 (74)	0.716 (209) 0.284 (83)	1.013 (0.699-1.467)	0.947
8	rs1834481 (C/G)	1 2	0.727 (189) 0.273 (71)	0.699 (204) 0.301 (88)	0.871 (0.602-1.261)	0.463
9	rs5744247 (C/G)	1 2	0.915 (238) 0.0846 (22)	0.904 (264) 0.0959 (28)	0.872 (0.485-1.565)	0.645
10	rs360722 (A/G)	1 2	0.908 (236) 0.0923 (24)	0.884 (258) 0.116 (34)	0.772 (0.445-1.34)	0.355
11	rs4937113 (A/T)	1 2	0.627 (163) 0.373 (97)	0.600 (174) 0.400 (116)	0.893 (0.633-1.259)	0.517
12	rs2043055 (A/G)	1 2	0.621 (159) 0.379 (97)	0.685 (200) 0.315 (92)	1.315 (0.932-1.888)	0.117
13	rs1946519 (A/C)	1 2	0.642 (167) 0.358 (93)	0.616 (180) 0.384 (112)	0.895 (0.633-1.265)	0.530

Table 4.47 *IL18* candidate region stage-1 results

Thirteen genotyped tSNPs within the *IL18* candidate region were analysed for significant allele frequency differences between the patient (n=130) and control (n=146) cohorts. The frequency in each cohort of both the common (1) and rare (2) alleles are given. The OR shown is for allele 2 (allele 1 OR=1). The p-value for the test for frequency differences between the cohorts is also shown.

and rheumatoid arthritis (Bresnihan et al., 2002; Gracie et al., 1999b; Sivalingam et al., 2007).

Associations of IL-18 gene polymorphisms have been found with a number of auto-inflammatory and auto-immune diseases, including: type 1 diabetes (Kretowski et al., 2002; Ide et al., 2004), procititis-type UC (Takagawa et al., 2005), inflammatory bowel disease (Aizawa et al., 2005), systemic lupus eythematosus (SLE) (Xu et al., 2007), and adult-onset Still's disease (Sugiura et al., 2002). There have however been conflicting results from association studies of *IL18* in rheumatoid arthritis (Sivalingam et al., 2003; Rueda et al., 2005; Gracie et al., 2005). In addition to disease associations, *IL18* polymorphisms have been shown to be associated with expression levels. A haplotype of two SNPs in the *IL18* promoter region, rs187238 (-607) and rs1946518 (-137), has been shown in reporter assays to affect transcription levels, and to correlate with IL-18 mRNA levels in patients with Multiple Sclerosis (Giedraitis et al., 2001). In another study the -607 SNP was shown to be correlated with IL-18 plasma concentrations in both SLE patients and healthy controls (Xu et al., 2007a). Although neither of these SNPs was directly genotyped in this association study, they were both captured. The -607 SNP is tagged by SNP 7 (r^2=1), and -137 is tagged by SNP 13 (r^2=0.97). The -607 SNP was not included in either HapMap 3 or WTCCC 2 and so the LD relationship cannot be confirmed. The tagging relationship between genotyped SNP 13 and -137 is, however, confirmed by both HapMap 3 (r^2=0.957), and WTCCC (r^2 =0.999). A different set of *IL18* SNPs have also been shown to be associated with circulating IL-18 levels in the blood of Caucasian men undergoing coronary artery bypass grafts, and in a cohort of age-matched healthy controls. The associations seen were with a single promoter region SNP, rs2043055 (-5848), and with a haplotype of five SNPs over the *IL18* region (Thompson et al., 2007). The SNPs included in the haplotype, in addition to -5848, were: rs1946519 (-9731), rs549908 (+4860), rs360729 (+8855), and rs3882891 (+11015). Apart from +8855, all of these SNPs were directly genotyped in the association study presented in this thesis: SNPs 13, 12, 7, and 6. However, +8855 was also included in this study as it is captured (r^2=0.967) by SNP 7. This was confirmed in HapMap 3 and WTCCC 2 (r^2=0.997 and 0.993 respectively). However, none of these *IL18* SNPs previously shown to be associated with expression levels showed evidence of disease association in this study ($p>0.05$). This study has therefore not shown any evidence for an association of polymorphisms within the *IL18* gene region with susceptibility to sJIA.

The lack of association with sJIA shown in this study concurs with the previous study by Sugiura et al. who found a significant association of a 13 marker *IL18* haplotype with JIA as a whole, and with the oligoarticular and polyarticular subtypes individually, but not the systemic subtype alone (Sugiura et al., 2006).

Although this study has not shown any evidence that *IL18* polymorphisms are associated with sJIA, it is possible that, due to the small cohort sizes used, the study did not have sufficient power to detect significant associations of smaller effect sizes. The issues of sample size and power in this study are discussed in chapter 5.

4.7. *CASP1* candidate region

4.7.1. Association Study

The genes flanking the candidate gene *CASP1* were identified using the July 2003 human genome reference sequence (NCBI build 34), hg16 annotation track. The boundaries of the candidate region, chromosome 11:104402687-104546387 were defined based on the positions of the flanking genes. In the NCBI build 34 reference sequence three alternatively spliced isoforms of *CASP1* are shown. However, according to build 39, the most recent at the time of writing, there are five protein coding isoforms of caspase-1. Although they were not all represented in the reference sequence version used to determine the candidate region, none of the gene sequences of these five isoforms extend beyond those in build 34, and were therefore all captured in this study. Flanking *CASP1* 3′ is caspase 5 (*CASP5),* and 5′ is caspase recruitment domain-containing protein 18 (*CARD18).* In the NCBI build 34 genome sequence the *CARD18* mRNA sequence AY358231 is shown (centre panel Figure 4.37). Caspase 5 is a member of the caspase family required to cleave the transcription regulator Max, important in cell growth, differentiation, and apoptosis (Munday et al., 1995). CARD18 (also called Iceberg) prevents the auto-activation of caspase-1, and thereby the production of active IL-1β, by preventing it binding to the adaptor RIP2 (Humke et al., 2000). Also located within the caspase-1 candidate region is the gene *COP*, encoding the caspase-1 inhibitor the gene sequence for which overlaps *CASP1*. *COP* is located within intron 1 of the long *CASP1* isoforms, and upstream of the short isoform (centre panel Figure 4.37). *COP* shares a high degree of sequence similarity with the caspase-1 prodomain, and is functionally analogous to CARD18. COP binds both caspase-1 and RIP preventing oligomerisation and activation of caspase-1 (Lee et al., 2001). Based on the positions of the genes flanking *CASP1* the candidate region selected for investigation spanned 1.4 Mb (143,700bp). The region extends 31.3kb (31,269bp) 3′, and 3.3kb 5′, of *CASP1* (bottom panel Figure 4.37).

4.7.1.1. Tagging SNP selection

Genotyping data was available for a total of 282 SNPs within the selected *CASP1* candidate region. These SNPs were all from the HapMap database (top panel Figure 4.37). Ninety nine of the SNPs in the region were monomorphic in the CEU population, and a further 43 had a MAF<0.05. These SNPs were therefore not included in the LD analysis. A total of 140 SNPs

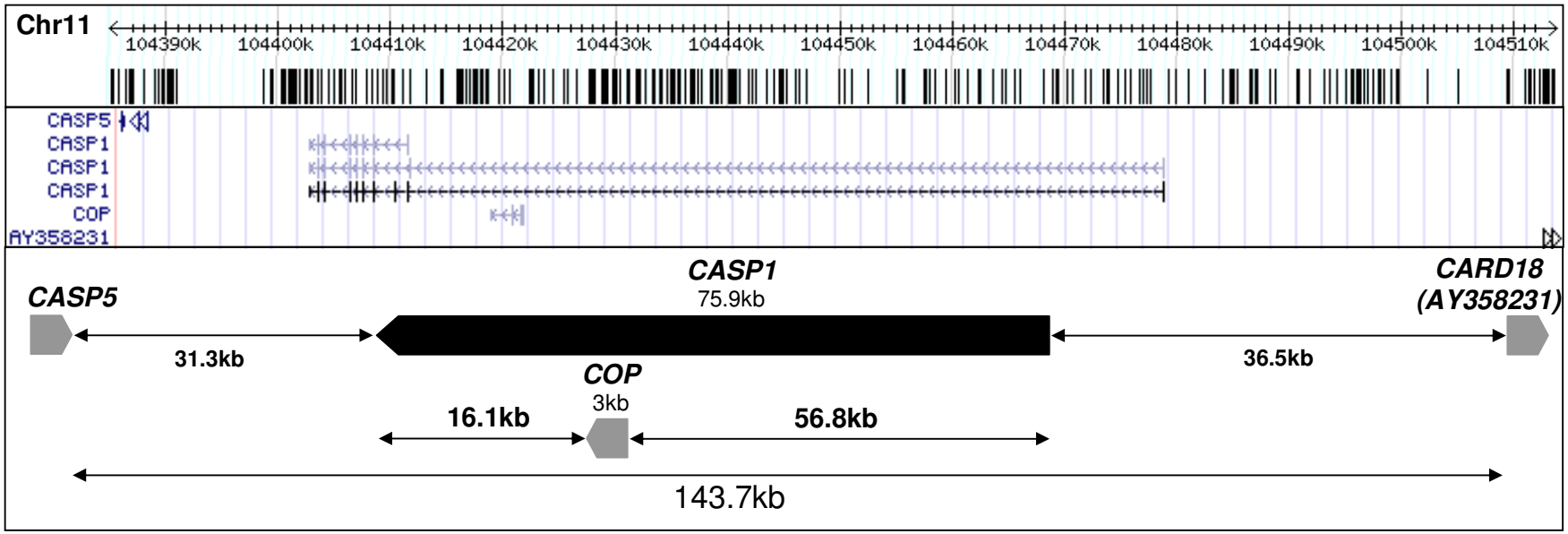


Figure 4.37 *CASP1* candidate region

The boundaries of the *CASP1* candidate region are the flanking genes *CASP5* and *CARD18* (*AY358231*), determined by the July 2003 human genome reference sequence (NCBI build 34), hg16 annotation track, shown in the centre panel. The 3′ end of *CASP5* and the 5′ end of *CARD18* are proximal to *CASP1*. The approximate (due to scale) positions within this region of all SNPs (irrespective of MAF) in the HapMap phase 2 database (data release #18/phaseII Sept05) are shown in the top panel. The bottom panel shows a schematic (not to scale) of the candidate region with gene orientation, size, and distances given. The gene of interest is shown in black and all other features within the region in grey. The largest *CASP1* isoform was used for size calculations, and, for clarity, is the only one shown. The *COP* gene sequence overlaps with *CASP1*, the distances shown from *COP* are relative to the 5′ and 3′ ends of *CASP1*. All sizes and distances are based on chromosomal positions reported in the NCBI build 34 genome reference sequence.

within the candidate region satisfied the inclusion criteria and were therefore included in the LD investigation and subsequent tSNP selection.

A graphical representation of the LD between these 140 SNPs in the *CASP1* candidate region is shown in Figure 4.38. There were two distinct LD blocks evident in the region. The LD block on the left hand side of the figure included 45 SNPs spanning 28.4kb. The block on the right hand side included 68 SNPs and spans 87kb. There was a gap of 7.9kb, between the two LD blocks of 6 SNPs which did not show high LD with the SNPs in either of the blocks. The left hand LD block spanned from 11.1kb upstream of the start of transcription of *CASP1* to 17.3kb downstream of the start of transcription, within intron 1 of *CASP1* and intron 2 of *COP*. The LD block on the right hand side of the figure spanned from 25.2kb downstream of *CASP1* start of transcription, in intron 1, to 36.3kb downstream of the 3′ end of the gene sequence. There was mainly consistent high LD in the left hand block, with most of the SNPs in the block showing high LD with the majority of the other SNPs within the block. There were however, four SNPs within this block only in high LD with each other and in very low LD with the other SNPs. Conversely, the LD within the left hand block was far more variable with each SNP being in high LD with only some of the other SNPs within the block rather than with the majority.

The minimal subset required to capture the variation within all 140 SNPs, due to the LD relationships, was 19 tSNPs (13.6% of the total). Following the assay design stage, however, it was determined that the optimal set of SNPs for genotyping consisted of 24 tSNPs, which captured all 140 SNPs within the candidate region. Of the tSNPs genotyped, 13 were not in high LD ($r^2 \leq 0.8$) with any of the other SNPs, and so it was necessary to genotype them directly. Some of the other genotyped SNPs are however tagging a large number of SNPs; SNP 4 is tagging an additional 29 SNPs, and SNP 24 an additional 31 SNPs. All of the selected tSNPs, and the SNPs captured by each, are listed in Table 4.48.

4.7.1.2. Stage-1

The genotyping assay for rs17103781 was classed as failed because there was no clear heterozygote cluster in the genoplot (Figure 4.1), resulting in a GC10 score below the 0.5 cut off (Table 4.2). This SNP, for which genotypes were not called and therefore was not included in the association analysis, was tagging 31 SNPs. An additional four genotyped SNPs from the *CASP1* candidate region showed significant deviation from HWE (Table 4.3),

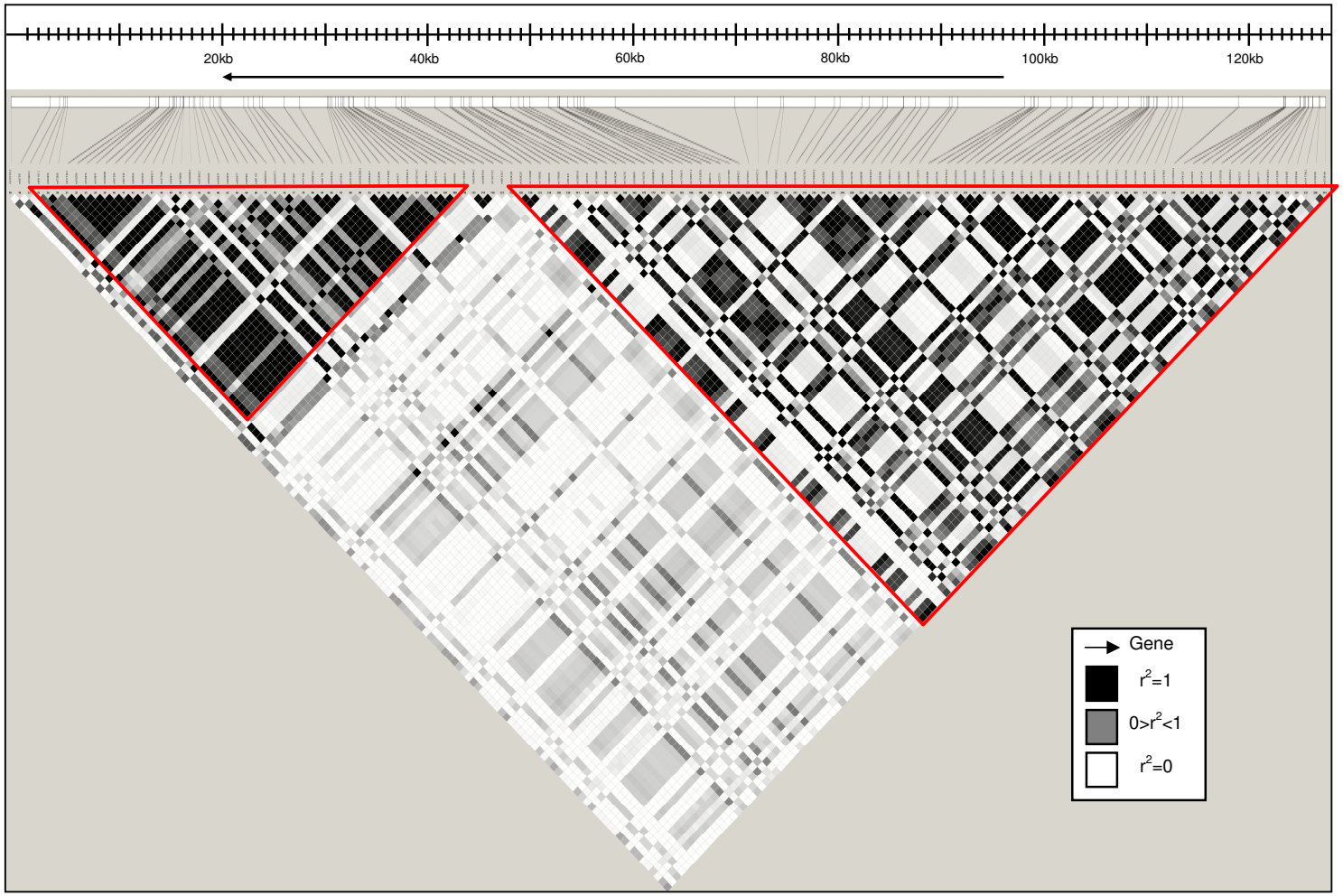


Figure 4.38 LD pattern over the *CASP1* candidate region

The pairwise LD relationships between SNPs in the *CASP1* candidate region were investigated to enable tSNP selection. The region shown is comprised of the included SNPs only, not the whole of the candidate region. The gene orientation and location within the region is indicated. The position within the region of each SNP satisfying the selection criteria is represented by a line underneath the gene position. Each SNP is labelled by rs number and Haploview ID number. Each coloured square represents the pairwise LD between two SNPs. The graphical representation, generated in Haploview, is shown on a grey scale, with higher r^2 LD represented by darker blocks. The blocks of high LD are indicated

and were therefore also not included in the association analysis. These were: rs11604437 (SNP 10), rs7934239 (SNP 11), rs10895775 (SNP19), and rs1503392 (SNP 23). SNPs 10 and 11 were not tagging any other SNPs within the candidate region, but SNPs 19 and 23 were tagging 15 and 5 additional SNPs respectively. Therefore, the five SNPs excluded from the association analysis had been selected as tSNPs for a total of 51 additional SNPs within the region. However, 48 of these 56 SNPs (5 tSNPs and 51 captured SNPs) were in sufficiently high LD ($r^2 \geq 0.8$) with two (SNPs 16 and 21) of the SNPs which were included in the analysis, that they were still captured. This included the genotyped SNPs 10, 11, and 23. Eight of the SNPs however were not in high LD ($r^2 \geq 0.8$) with any of the analysed SNPs, and were therefore not captured by the association study (Table 4.49).

Analysis of the frequencies in the patient and control cohorts was performed for 19 genotyped tSNPs, which captured a total of 132 SNPs within the *CASP1* candidate region. None of the tSNPs showed significant ($p<0.05$) frequency differences between the patient (n=130) and control (n=146) populations. The results (p-values) are given in (Table 4.50) for each tSNP. Therefore, further investigation of this candidate region was not pursued in stage-2 of the study.

4.7.2. Summary

No evidence was found for significant association of polymorphisms within the *CASP1* candidate region and susceptibility to sJIA.

4.7.3. Discussion

The caspase-1 protein encoded by *CASP1* cleaves IL-1β and IL-18 into their mature, active forms (Cerretti et al., 1992; Thornberry et al., 1992; Gu et al., 1997; Ghayur et al., 1997). Following activation by caspase-1 the mature cytokine is released from the cell. Caspase-1 is therefore an essential component of the IL-1 and IL-18 mediated inflammatory responses. The importance of caspase-1 in the production of IL-1 has been demonstrated in caspase-1 deficient mice, which do not produce mature IL-1β following injection with LPS, and are resistant to endotoxic shock (Li et al., 1995b). Addition of caspase-1 inhibitors to LPS activated murine macrophages inhibits the release of mature IL-1β, with no effect on the levels of *IL1B* mRNA or pro-IL-1β, (Miller et al., 1995), as does caspase-1 inhibitor

SNP Nº	rs Number	Tagged SNPs
1	rs3181318	n/a
2	rs492859	rs572980,rs547351
3	rs518878	n/a
4	rs1699089	rs476889,rs483345,rs492366,rs568910,rs572687,rs692897,rs504054, rs519047,rs7928549,rs491191,rs505901,rs517736,rs1503391, rs507438,rs507509,rs546449,rs692914,rs572031,rs516286,rs542571, rs7944216,rs490565,rs500577,rs557905,rs481736,rs580253, rs534811,rs17103597, rs489715
5	rs530537	n/a
6	rs501192	n/a
7	rs571593	rs526167,rs17376473,rs1792774,rs484626,rs501626,rs556205, rs488992
8	rs1613367	n/a
9	rs10502045	rs1977989,rs576800,rs10895763,rs9326349,rs11821722,rs2282659
10	rs11604437	n/a
11	rs7934239	n/a
12	rs508760	n/a
13	rs1785883	n/a
14	rs1785884	n/a
15	rs538943	rs486123
16	rs1623342	rs1785878,rs529809,rs1961599,rs1628434,rs11226601,rs1785882, rs6591110,rs1699088,rs4106096,rs572717,rs562441,rs7934144, rs1792766,rs4638280,rs484345,rs1792763
17	rs11226613	n/a
18	rs7109571	rs1792752,rs1699080,rs4587693
19	rs10895775	rs7932543,rs1483025,rs1699081,rs10895771,rs1785872,rs1785871, rs1792753,rs1785864,rs1847293,rs1699087,rs3736149,rs7936207, rs1699086,rs1503399,rs1785873
20	rs12797863	n/a
21	rs4754123	n/a
22	rs1503394	rs3926237
23	rs1503392	rs6591111,rs1847298,rs11226619,rs7934633,rs7952721
24	rs17103781	rs3989416,rs1503386,rs12417470,rs12417050,rs17103670, rs12419452,rs12420601,rs1941421,rs12800151,rs4755079, rs12281540,rs12420557,rs4754124,rs10895772,rs12418649, rs1393907,rs3898011,rs1503395,rs475266,rs12416756,rs1910398, rs4755078,rs12801987,rs17103773,rs17376738,rs12807436, rs7930712,rs7117758,rs12420524,rs2062807,rs12418438

Table 4.48 *CASP1* candidate region tSNPs

A total of 24 tSNPs were genotyped to capture all 140 SNPs within the candidate region. The SNPs captured ($r^2 \geq 0.8$) by each genotyped tSNP are shown. For ease of identification each tSNPs genotyped was assigned a number according to genome order.

SNP Nº	rs Number	Tagging
16	rs1623342	rs1792753,rs10895775,rs1785873,rs7936207,rs1699081,rs10895771, rs1503399,rs1785871,rs1785872,rs1699086,rs1785864,rs1699087, rs7932543,rs1483025,rs1847293,rs3736149
21	rs4754123	rs4754124,rs12417470,rs1503395,rs12417050,rs12281540,rs17103773,rs124 19452,rs17376738,rs3989416,rs12420557,rs475266,rs17103670, rs12416756,rs7930712,rs12418438,rs1393907,rs4755079,rs4755078, rs12800151,rs12420601,rs1941421,rs1910398,rs124186,rs1503386, rs10895772,rs12418649,rs3898011,rs475266,rs12801987,rs17103781, rs12807436,rs7117758
Uncaptured		rs11604437,rs7934239,rs1847298,rs11226619,rs7952721,rs6591111, rs7934633,rs1503392

Table 4.49 *CASP1* additional tagging of analysis excluded SNPs

Five of the genotyped tSNPs from the *CASP1* candidate region were excluded from analysis due to poor genotyping quality or deviation from HWE. The excluded SNPs, and the additional SNPs they were tagging, are shown, as well as whether or not the excluded SNPs were captured by another of the genotyped SNPs ($r^2 \geq 0.8$).

SNP Nº	rs Number	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
1	rs3181318 (G/A)	1 2	0.708 (184) 0.292 (76)	0.688 (201) 0.312 (91)	0.912 (0.634-1.313)	0.621
2	rs492859 (C/A)	1 2	0.869 (226) 0.131 (34)	0.861 (248) 0.139 (40)	0.933 (0.571-1.525)	0.781
3	rs518878 (A/G)	1 2	0.865 (225) 0.135 (35)	0.860 (251) 0.140 (41)	0.952 (0.586-1.548)	0.844
4	rs1699089 (A/G)	1 2	0.865 (225) 0.135 (35)	0.860 (251) 0.140 (41)	0.952 (0.586-1.548)	0.844
5	rs530537 (A/G)	1 2	0.608 (158) 0.392 (102)	0.613 (179) 0.387 (113)	1.023 (0.726-1.441)	0.898
6	rs501192 (G/A)	1 2	0.885 (230) 0.115 (30)	0.860 (251) 0.140 (41)	0.799 (0.483-1.321)	0.380
7	rs571593 (G/A)	1 2	0.900 (234) 0.100 (26)	0.921 (269) 0.0788 (23)	1.300 (0.722-2.339)	0.382
8	rs1613367 (A/C)	1 2	0.531 (137) 0.469 (121)	0.511 (143) 0.489 (137)	0.922 (0.657-1.293)	0.638
9	rs10502045 (T/A)	1 2	0.739 (192) 0.262 (68)	0.753 (220) 0.247 (72)	1.082 (0.737-1.589)	0.687
12	rs508760 (C/A)	1 2	0.919 (239) 0.0808 (21)	0.938 (274) 0.0616 (18)	1.338 (0.696-2.57)	0.382
13	rs1785883 (G/A)	1 2	0.914 (234) 0.0859 (22)	0.900 (261) 0.100 (29)	0.846 (0.473-1.514)	0.572
14	rs1785884 (G/A)	1 2	0.692 (180) 0.308 (80)	0.640 (187) 0.360 (105)	0.792 (0.555-1.13)	0.197
15	rs538943 (A/G)	1 2	0.915 (238) 0.0846 (22)	0.938 (274) 0.0616 (18)	1.407 (0.737-2.686)	0.299
16	rs1623342 (A/G)	1 2	0.658 (171) 0.342 (89)	0.671 (196) 0.329 (96)	1.063 (0.746-1.514)	0.737
17	rs11226613 (A/T)	1 2	0.792 (206) 0.208 (54)	0.767 (224) 0.233 (68)	0.864 (0.576-1.294)	0.476
18	rs7109571 (G/A)	1 2	0.662 (172) 0.339 (88)	0.675 (197) 0.325 (95)	1.061 (0.744-1.513)	0.744
20	rs12797863 (A/G)	1 2	0.923 (240) 0.0762 (20)	0.945 (276) 0.0548 (16)	1.437 (0.815-1.609)	0.294
21	rs4754123 (T/A)	1 2	0.923 (240) 0.0769 (20)	0.945 (276) 0.0458 (16)	1.437 (0.728-2.837)	0.294
22	rs1503394 (T/A)	1 2	0.735 (191) 0.265 (69)	0.726 (212) 0.274 (80)	0.957 (0.657-1.395)	0.821

Table 4.50 *CASP1* candidate region stage-1 results

Nineteen of the genotyped tSNPs within the *CASP1* candidate region were analysed for significant allele frequency differences between the patient (n=130) and control (n=146) cohorts. The allele frequency for each cohort is shown for both alleles. The alleles are labelled as 1, the more common allele, and 2, the rarer allele. The OR shown is for allele 2 (allele 1 OR=1). The p-value for the test for frequency differences between the cohorts is also shown.

administration to mice injected with LPS (Wannamaker et al., 2007). This inhibition of mature IL-β and IL-18 secretion, but not other cytokines, following treatment with caspase-1 inhibitors is also observed in LPS stimulated human PBMCs (Wannamaker et al., 2007). Similarly, production of IFNγ, normally induced by IL-18 and IL-12 acting as co-stimulants, is reduced by 80% in IL-12 treated splenocytes from caspase-1 deficient mice compared to cells from wild type animals. A similar reduction in IFNγ production is seen when mice are administered caspase-1 specific inhibitors prior to treatment with IL-12. In both cases pro-IL-18 is expressed normally (Fantuzzi et al., 1999). However, in addition to caspase-1, other proteases can cleave and activate IL-1β. In a model of neutrophil dominated acute arthritis caspase-1 knock out mice develop joint swelling similar to that seen in wild type mice. In a model of chronic arthritis caspase-1 knock out mice do however have less joint inflammation and cartilage damage. Joint blockade of caspase-1 and the neutrophil serine protease PR3, another activator of IL-1β, protects against cartilage and bone destruction (Joosten et al., 2009).

The role of caspase-1 in inflammatory diseases has also been demonstrated through the use of caspase-1 knock out (KO) mice, and caspase-1 inhibitors. In a model of induced acute colitis, caspase-1 KO mice show clinical scores 50% lower than wild type mice. Caspase-1 KO mice also show nearly complete protection against chronic colitis induced over 30 days. This protection against disease development is accompanied by a reduction in the release of IL-1β, IL-18, and IFNγ from colon cultures (Siegmund et al., 2001b). A 60% reduction in disease severity of type II collagen induced arthritis, accompanied by a delay in the onset of inflammation, is observed when mice are treated with caspase-1 inhibitors. This effect of reduced inflammation is also seen in mice with well established disease (Ku et al., 1996), and in models of skin inflammation (Wannamaker et al., 2007).

Prior to cleavage of IL-1β and IL-18, caspase-1 has to be activated by the IL-1 inflammasome (Martinon et al., 2002a; Agostini et al., 2004a). Gain of function mutations in cryopyrin, a component of the IL-1 inflammasome, result in over-activation of caspase-1, and thus an inappropriate inflammatory response. These mutations are responsible for the cyropyrin-associated periodic syndromes (CAPS) MWS, FCAS, and NOMID/CINCA (Hoffman et al., 2001; Feldmann et al., 2002; Aksentijevich et al., 2002; Kaipiainen-Seppanen et al., 2008a; Arostegui et al., 2010; Jeru et al., 2010), which are clinically similar to sJIA. Following stimulation with LPS PBMCs from patients with FCAS secrete significantly higher levels of

IL-1β and IL-18 than PBMCs from control individuals. Treatment with caspase-1 inhibitors however, successfully blocks secretion of the mature cytokines in both control and FCAS patient PBMCs (Stack et al., 2005). This demonstrates that the IL-1β and IL-18 over-expression is due to caspase-1 activity. It is therefore reasonable to assume that genetic variations in *CASP1*, causing over production of IL-1β and IL-18, would result in similar clinical outcomes to the CAPS, and may therefore be involved in susceptibility to sJIA.

No investigations into the involvement of *CASP1* polymorphisms in sJIA have been previously published. The minor allele of a SNP in intron 6 of *CASP1* has however been found to be significantly less frequent in myocardial infarction patients than in healthy controls with no family history of cardiovascular disease. This SNP also shows borderline significance with a reduced risk of cardiovascular mortality, reduced serum IL-18, and caspase-1 plasma levels. Further investigation of this associated SNP showed that the allele associated with reduced cardiovascular risk results in lower levels of *CASP1* mRNA expression *in vitro* (Blankenberg et al., 2006).

This study has shown no evidence that genetic variation in *CASP1* is involved in susceptibility to sJIA. The SNP identified as associated with lower caspase-1 expression (Blankenberg et al., 2006) was included in this study, SNP 6, but showed no frequency difference between the patients and control cohorts.

CASP1 is located within a cluster of caspase genes (Figure 4.37). Because the candidate region selected extended up- and down-stream of *CASP1* the flanking sequences of two other caspase genes were also included in this study. SNPs within the region down-stream of *CASP5*, activator of the transcription factor Max, and the region up-stream of the caspase-1 inhibitor *CARD18* were investigated. Also covered in this study were SNPs within, and surrounding, the gene encoding another caspase-1 inhibitor, *COP*, the sequence for which overlaps *CASP1*. As none of the SNPs potentially associated with these additional genes showed evidence of disease association, this study has also found no evidence that variation in the inhibition of caspase-1 is involved in susceptibility to sJIA.

Although this study has not shown any evidence that *CASP1* polymorphisms are associated with sJIA, it is possible that, due to the small cohort sizes used, the study did not have sufficient power to detect significant associations of smaller effect sizes. The issues of sample size and power in this study are discussed in chapter 5.

4.8. NALP3 candidate region

4.8.1. Association Study

The genes flanking the candidate gene *NALP3* were identified using the July 2003 human genome reference sequence (NCBI build 34), hg16 annotation track. The boundaries of the candidate region, chromosome 1:244441769-244561055 were defined based on the positions of the flanking genes *NALP3* is also called *NLRP3*, Nod-Like receptor family, pyrin domain containing 3. This nomenclature was used in the genome annotation track (centre panel Figure 4.39). However, for continuity *NALP3* will be used in this manuscript. As shown in the figure, there are five alternatively spliced isoforms of *NALP3*. Flanking *NALP3* 5′ is zinc finger protein 496 (*ZNF496*), and 3′, olfactory receptor, family 2, subfamily B, member 11 (*OR2B11*). ZNF496, also called Nzip1, binds to the histone methyltransferase NSD1 and acts as a transcription repressor (Nielsen et al., 2004). Members of the olfactory receptor gene family initiate the neuronal response generating the perception of smell following interaction with odorant molecules (Buck and Axel, 1991). Based on the positions of these flanking genes the candidate region extends 86kb (86,306bp) 5′ and 2Kb (2,212bp) 3′ of *NALP3*. The whole candidate region spans 0.1Mb (119,286bp) (bottom panel Figure 4.39).

4.8.1.1 Tagging SNP selection

Genotyping data was available for a total of 45 SNPs within the selected *NALP3* candidate region. These SNPs were all from the HapMap database (top panel Figure 4.39). Ten of the SNPs in the region were monomorphic in the CEU population, and a further 3 had MAF<0.05. These SNPs were therefore not included in the LD analysis. A total of 32 SNPs within the candidate region satisfied the inclusion criteria and were therefore included in the LD investigation and subsequent tSNP selection.

A graphical representation of the LD between these 32 SNPs in the *NALP3* candidate region is shown in Figure 4.40. There was very little LD between the SNPs in the candidate region, with 17 of the 32 SNPs not being in high LD ($r^2 \leq 0.8$) with any of the others. There was one block of LD present in the region, of nine sequential SNPs, SNPs 18 to 26, spanning 26kb. SNP 26, although in LD with the other SNPs in the block, was not in high LD with them (maximum $r^2 = 0.486$), reducing the span to 22.8kb. It was only necessary to genotype four of the nine SNPs in this LD block, including SNP 26. In addition to the block of LD there were four pairs of SNPs in high LD with each other. These were 3 pairs of adjacent SNPs, and one

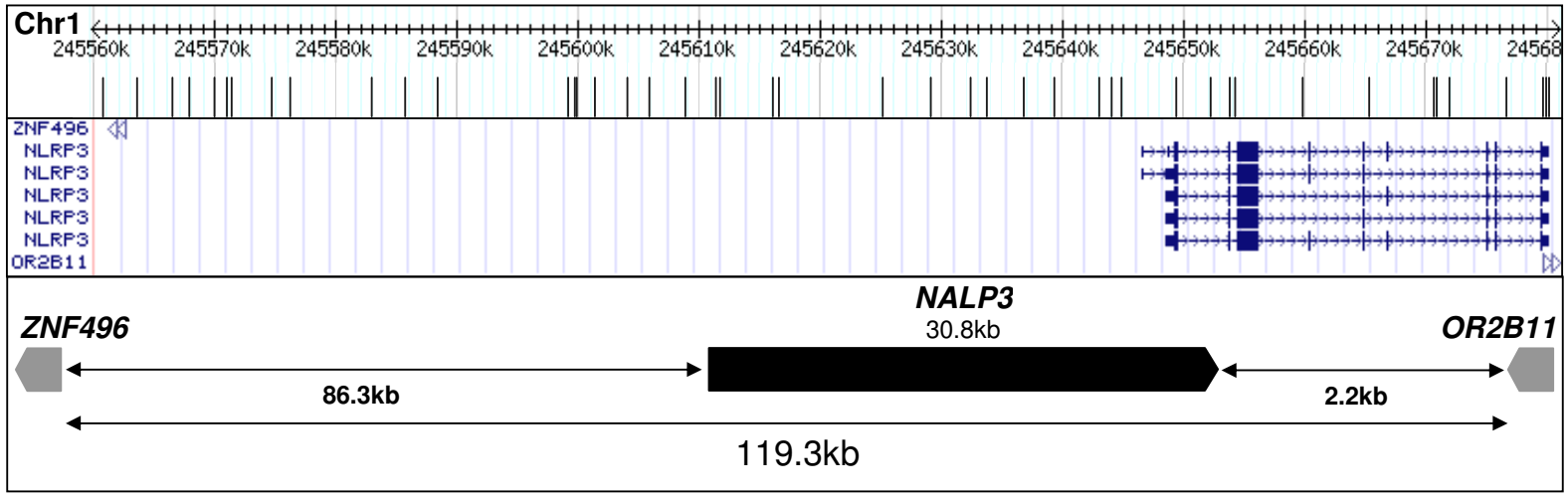


Figure 4.39 *NALP3* candidate region

The boundaries of the *NALP3* candidate region are the flanking genes *ZNF496* and *OR2B11*, determined by the July 2003 human genome reference sequence (NCBI build 34), hg16 annotation track, shown in the centre panel. *NALP3* is also called *NLRP3*, as shown here. The end 5′ of *ZNF496* and the 3′ end of *OR2B11* are proximal to *NALP3*. The approximate (due to scale) positions within this region of all SNPs (irrespective of MAF) in the HapMap phase 2 database (data release #18/phaseII Sept05) are shown in the top panel. The bottom panel shows a schematic (not to scale) of the candidate region with gene orientation, size, and distances given. The gene of interest is shown in black and all other features within the region in grey. The largest *NALP3* isoform was used for size calculations, and, for clarity, is the only one shown. All sizes and distances are based on chromosomal positions reported in the NCBI build 34 genome reference sequence.

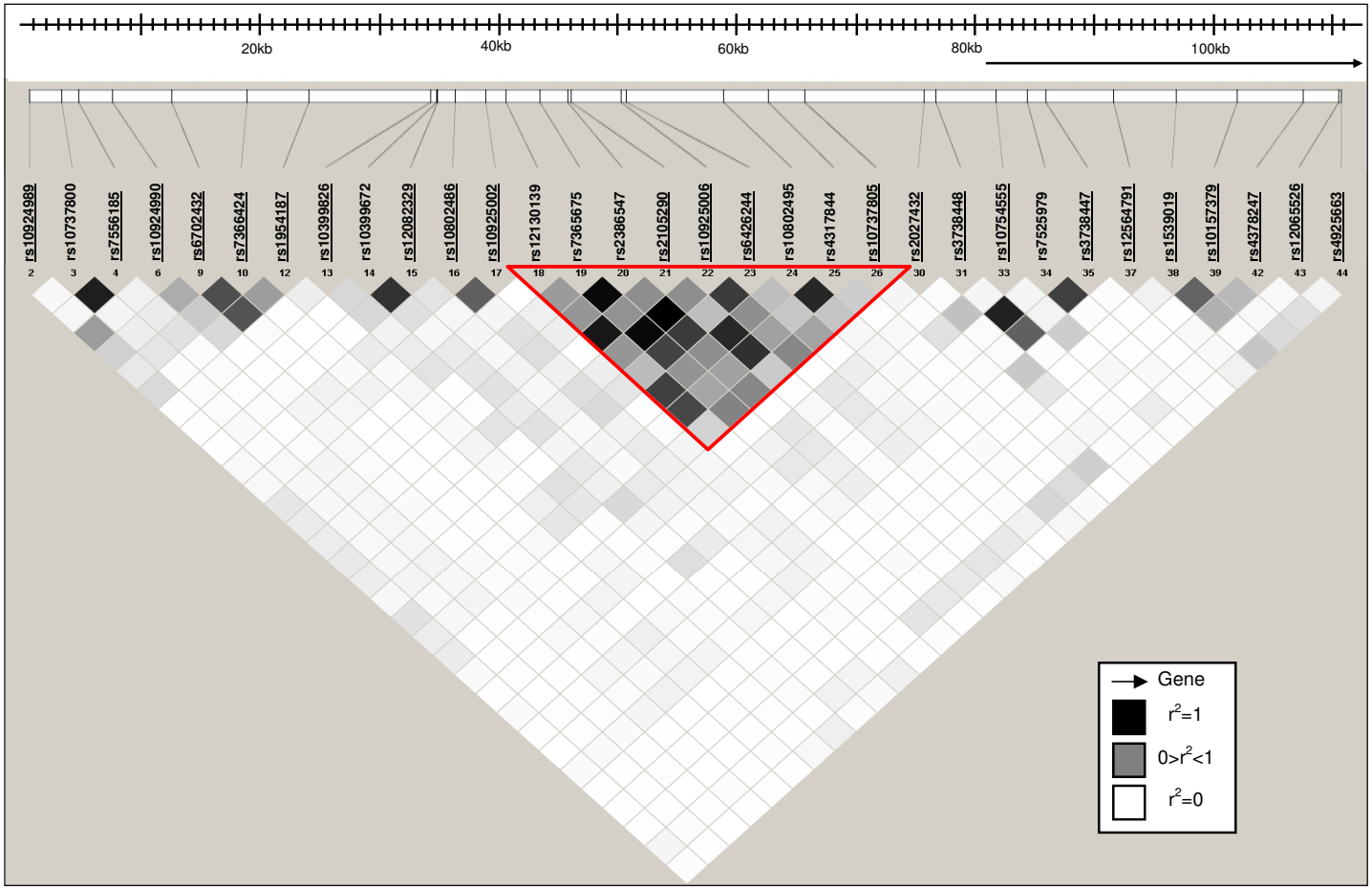


Figure 4.40 LD plot of the SNPs in the *NALP3* gene region

The pairwise LD relationships between SNPs in the *NALP3* candidate region were investigated to enable tSNP selection. The region shown is comprised of the included SNPs only, not the whole of the candidate region. The gene location within the region, and direction, is indicated. The position within the region of each SNP satisfying the selection criteria is represented by a line underneath the gene position. Each coloured square represents the pairwise LD between two SNPs. The graphical representation, generated in Haploview, is shown on a grey scale, with higher r^2 LD represented by darker blocks. Each SNP is labelled by rs number and Haploview ID number. The 24 tSNPs which were genotyped are underlined, and the block of high LD indicated.

pair which were one SNP apart (SNPs 31 and 34).

Due to the LD relationships between the SNPs a total of 24 tSNPs (75% of the total) were identified as the minimal subset required to capture the variation within all 32 SNPs. Only five of the 24 tSNPs genotyped were tagging ($r^2 \leq 0.8$) other SNPs. The remaining 19 SNPs were not in sufficient LD with any of the other SNPs and so it was necessary to genotype them all directly.

All of the selected tSNPs, and the SNPs captured by each, are listed in Table 4.51.

4.8.1.1 Stage-1

The genotyping assays for all 24 of the genotyped SNPs satisfied the quality control criteria (Sections 3.1.2.2.1.). However, rs7366424 (SNP 5) showed significant ($p<0.05$) deviation from HWE (Table 4.3) and was therefore not included in the association analysis. As SNP 5 was not in high LD ($r^2 \geq 0.8$) with any of the other SNPs within the candidate region, it was therefore not captured by any of the SNPs which were analysed. Association analysis was performed for 23 genotyped SNPs, which between them captured the variation contained in a total of 31 SNPs.

None of the *NALP3* candidate region tSNPs showed significant ($p<0.05$) frequency differences between the patient (n=130) and control (n=146) populations (Table 4.52). Therefore, further investigation of this candidate region was not pursued in stage-2 of the study.

4.8.2. Summary

No evidence was found for significant association of polymorphisms within the *NALP3* candidate region and susceptibility to sJIA.

4.8.3. Discussion

Cryopyrin (NALP3) is a central component of the NALP3 inflammasome, which activates caspase-1. Through binding to ASC and caspase-1, NALP3 activates caspase-1, thereby enabling caspase-1 to process pro-IL-1β and IL-18 into their mature, active forms (Martinon et al., 2002b; Agostini et al., 2004c). The essential role of NALP3 in IL-1 mediated

SNP Nº	rs Number	Tagged SNPs
1	rs10924989	n/a
2	rs7556185	rs10737800
3	rs10924990	n/a
4	rs6702432	n/a
5	rs7366424	n/a
6	rs1954187	n/a
7	rs10399826	n/a
8	rs12082329	rs10399672
9	rs10802486	n/a
10	rs10925002	n/a
11	rs2105290	rs12130139,rs10802495,rs4317844
12	rs10925006	rs7365675,rs2386547
13	rs6426244	n/a
14	rs10737805	n/a
15	rs2027432	n/a
16	rs3738448	rs7525979
17	rs10754555	n/a
18	rs3738447	n/a
19	rs12564791	n/a
20	rs1539019	n/a
21	rs10157379	n/a
22	rs4378247	n/a
23	rs12065526	n/a
24	rs4925663	n/a

Table 4.51 *NALP3* candidate region tSNPs

A total of 24 tSNPs were genotyped to capture all 32 52 SNPs within the candidate region. The SNPs captured ($r^2 \geq 0.8$) by each genotyped tSNP are shown. For ease of identification each genotyped SNP was assigned a number according to genome order.

SNP Nº	rs Number	Allele	Case freq. (Nº)	Control freq. (Nº)	OR (95% CI)	p value
1	rs10924989 (G/A)	1 2	0.627 (163) 0.373 (97)	0.648 (188) 0.356 (104)	1.076 (0.760-1.522)	0.680
2	rs7556185 (G/A)	1 2	0.904 (235) 0.0962 (25)	0.914 (267) 0.0856 (25)	1.136 (0.635-2.032)	0.667
3	rs10924990 (A/G)	1 2	0.565 (147) 0.435 (113)	0.542 (154) 0.458 (130)	0.911 (0.649-1.278)	0.588
4	rs6702432 (G/A)	1 2	0.589 (153) 0.412 (107)	0.551 (161) 0.449 (131)	0.860 (0.613-1.205)	0.380
6	rs1954187 (A/G)	1 2	0.646 (168) 0.354 (92)	0.593 (173) 0.408 (119)	0.796 (0.564-1.124)	0.195
7	rs10399826 (A/G)	1 2	0.846 (220) 0.154 (40)	0.870 (254) 0.130 (38)	1.215 (0.753-1.963)	0.425
8	rs12082329 (C/G)	1 2	0.512 (133) 0.489 (127)	0.517 (151) 0.483 (141)	1.023 (0.732-1.429)	0.896
9	rs10802486 (A/C)	1 2	0.708 (184) 0.292 (76)	0.706 (206) 0.295 (86)	0.989 (0.685-1.428)	0.955
10	rs10925002 (A/G)	1 2	0.600 (156) 0.400 (104)	0.610 (177) 0.390 (113)	1.044 (0.742-1.471)	0.804
11	rs2105290 (G/C)	1 2	0.558 (145) 0.442 (115)	0.534 (156) 0.466 (136)	0.910 (0.650-1.273)	0.581
12	rs10925006 (A/G)	1 2	0.735 (191) 0.265 (69)	0.750 (219) 0.250 (73)	1.084 (0.739-1.588)	0.680
13	rs6426244 (A/G)	1 2	0.669 (174) 0.331 (86)	0.688 (201) 0.312 (91)	1.092 (0.763-1.561)	0.631
14	rs10737805 (G/A)	1 2	0.862 (224) 0.139 (36)	0.865 (252) 0.137 (40)	1.013 (0.623-1.644)	0.960
15	rs2027432 (G/A)	1 2	0.823 (214) 0.177 (46)	0.832 (243) 0.168 (49)	1.066 (0.685-1.659)	0.777
16	rs3738448 (C/A)	1 2	0.939 (244) 0.062 (16)	0.904 (264) 0.0959 (28)	0.618 (0.327-1.171)	0.134
17	rs10754555 (G/C)	1 2	0.627 (163) 0.373 (97)	0.640 (187) 0.360 (105)	1.060 (0.749-1.499)	0.743
18	rs3738447 (G/A)	1 2	0.927 (241) 0.0731 (19)	0.890 (260) 0.110 (32)	0.641 (0.354-1.160)	0.137
19	rs12564791 (G/A)	1 2	0.904 (235) 0.0962 (25)	0.921 (267) 0.0793 (23)	1.235 (0.683-2.234)	0.485
20	rs1539019 (C/A)	1 2	0.624 (161) 0.376 (97)	0.658 (192) 0.343 (100)	1.157 (0.816-1.640)	0.414
21	rs10157379 (A/G)	1 2	0.636 (164) 0.364 (94)	0.634 (185) 0.366 (107)	0.991 (0.699-1.403)	0.959
22	rs4378247 (A/G)	1 2	0.842 (219) 0.158 (41)	0.877 (256) 0.123 (36)	1.331 (0.822-2.157)	0.245
23	rs12065526 (G/A)	1 2	0.884 (228) 0.116 (30)	0.907 (263) 0.093 (27)	1.282 (0.739-2.220)	0.376
24	rs4925663 (G/A)	1 2	0.585 (152) 0.415 (108)	0.555 (162) 0.445 (130)	0.885 (0.632-1.241)	0.480

Table 4.52 *NALP3* candidate region stage-1 results

Twenty three genotyped tSNPs within the *NALP3* candidate region were analysed for significant allele frequency differences between the patient and control cohorts. The allele frequency for each cohort is shown for both alleles. The alleles are labelled as 1, the more common allele, and 2, the rarer allele. The OR shown is for allele 2 (allele 1 OR=1). The p-value for the test for frequency differences between the cohorts is also shown.

inflammatory responses has been demonstrated in *NALP3* knockout mouse models. IL-1β and IL-18 are both produced in an inactive form (proIL-1β and proIL-18). Following cleavage by caspase-1 the active forms are released from the cell. It has been shown that in macrophages from *NALP3* knockout mice the production of intra-cellular pro-IL-1β and pro-IL-18 is induced following LPS stimulation. However, there is no release of the proteins from the cells, indicating that they have not been activated (Martinon et al., 2006; Mariathasan et al., 2006; Sutterwala et al., 2006).

Mutations in *NALP3* have been identified in patients with a number of auto-inflammatory diseases including MWS, FCAS, and NOMID/CINCA (Hoffman et al., 2001; Feldmann et al., 2002; Aksentijevich et al., 2002; Kaipiainen-Seppanen et al., 2008b; Arostegui et al., 2010; Jeru et al., 2010). These diseases are clinically similar to sJIA, and are collectively known as cyropyrin-associated periodic syndromes (CAPS). SNPs in *NALP3* have been shown to be associated with susceptibility to Crohn's disease, levels of *NALP3* mRNA, and IL-1β production (Villani et al., 2009). It has also been shown that levels of *NALP3* mRNA are higher in the synovium of patients with the inflammatory disease RA, compared to the non-inflammatory, erosive OA (Rosengren et al., 2005). An association with susceptibility to RA, and to a more severe disease course, has also been shown with a two marker haplotype of SNPs in *NALP3* and *CARD8*, also a component of the IL-1β inflammasome (Kastbom et al., 2008).

SNPs within *NALP3* and *MEVF*, which encodes pyrin, another mediator of caspase-1 activation, have been previously investigated for association with JIA. One SNP in each gene was shown to be associated with the psoriatic arthritis subtype, but not with JIA overall, or the sJIA subgroup (Day et al., 2008a). It has however, been shown that Turkish patients with sJIA have a significantly higher frequency of *MEVF* mutations (14.28%) , known to be found in FMF, than ethnically matched healthy controls (5%) (Ayaz et al., 2009).

The lack of significant association of SNPs within the *NALP3* candidate region and sJIA found in this study, are concordant with the results of the study by Day et al. (Day et al., 2008b). The *NALP3* SNP they found to be associated with the psoriatic arthritis JIA subtype was also genotyped in this study, SNP17. Although the other 20 *NALP3* SNPs studied by Day et al. are not stated in their manuscript so direct comparison is not possible, it is likely that the variation captured was very similar to this study, as they also performed tSNP selection using the CEU population genotyping data from HapMap. The results of this study provide no

evidence to support the hypothesis that genetic variation in *NALP3* results in inappropriate activation of caspase-1, thereby leading to excessive activation of IL-1β and IL-18 and the auto-inflammatory state seen in sJIA.

Although this study has not shown any evidence that *NALP3* polymorphisms are associated with sJIA, it is possible that, due to the small cohort sizes used, the study did not have sufficient power to detect significant associations of smaller effect sizes. The issues of sample size and power in this study are discussed in chapter 5.

4.9. Association analysis of the susceptibility alleles according to response to IL-1 blockade

A total of 43 sJIA patients were available for inclusion in this analysis. Of these 16 (37%) patients responded to IL-1 blockade according to the criteria for disease remission. The information for each patient, including response classification, is shown in Table 4.53. Three of the patients were treated with the IL-1β antibody Canakinumab, and the remaining 40 with the recombinant IL-1 receptor antagonist Anakinra. Fourteen (35%) of the patients treated with Anakinra exhibited a positive response to treatment. The proportion of patients who responded to treatment was approximately equal for each of the four centres: 41.6% of twelve patients from the UK, 33.3% of 15 patients from the USA, 50% of four patients from France, and 33.3% of the 12 patients from the Netherlands. All three of the patients treated with Canakinumab were from the Utrecht cohort. The patient cohort had a male to female ratio of 1:1.15, and an average age of onset (aao) of 6.28 years, ranging from 11 months to 14 years. The majority of the patients (35) were of Caucasian origin. The remaining eight were: Iranian (n=1), Black African (n=4), Pakistani (n=1), and Asian (n=1). There was no apparent differences in sex ratio or aao between the two treatment response groups. The male:female sex ratios for responders and non-responders respectively were: 0.77:1 and 1.077:1, average aao was 6.5 ± 4.37 years and 6.15 ± 4.11 years, and the aao range in both groups was between 1 and 14 years old.

All eight SNPs investigated were in HWE ($p>0.05$) in both the responder and non-responder groups. The results of the comparison of the allele frequencies in the two response groups, in all patients, regardless of ethnicity or treatment (n=43), of the eight IL-1 family SNPs investigated are shown in Table 4.54. Similar results were seen when only Caucasian patients (n=35), or only patients treated with Anakinra (n=40), were analysed (data not shown). None of the investigated SNPs showed a significant ($p<0.05$) difference in frequencies between the two patient groups. The two marker *IL18BP* haplotype was however tending towards significance both over all ($p=0.068$), and the 1-2 haplotype individually ($p=0.066$). Interestingly, the 1-2 haplotype, which was significantly more frequent in the patient than the control cohort in the disease susceptibility analysis, is higher in the non-responders (0.192) compared to the patients who responded to IL-1 blockade (0.0339).

For *IL1RAP* SNPs 14 and 25, *IL1* ligand cluster SNPs 4 and 64, and for the *IL1* receptor cluster SNP 8, the alleles conferring increased susceptibility to sJIA are higher in the patients

who did respond to treatment, although not statistically significant. *IL1RAP* SNP 25 shows the opposite trend, the common allele, found at a higher frequency in patients than in healthy controls, is higher in non-responders (0.8 compared to 0.7, not statistically significant), and *IL1* ligand cluster SNP 1 is at the same frequency in both patient groups. In the preliminary analysis of only the 12 patients treated at Great Ormond Street hospital, the minor allele of the IL-1 receptor cluster SNP8 was not present in patients who responded, and at a frequency of 0.5 in those who did not. However, MAFs in the larger cohort of 43 patients were 0.30 in responders, and 0.23 in non-responders. This demonstrates how small cohorts can give misleading results, and therefore the importance of using the largest possible cohort sizes.

Recruitment of further patients treated with IL-1 blockade is ongoing. As the number of patients included in the analysis is very small, no firm conclusions can be drawn from the results. The results presented here should therefore be viewed as preliminary.

4.9.1. Discussion

Treatment with the recombinant IL-1 receptor antagonist Anakinra shows remarkable improvement in approximately 40% of sJIA patients (Quartier et al., 2006; Lequerre et al., 2008; Gattorno et al., 2008). Similar results have also been reported with other IL-1 blockade therapies: 44.4% of patients administered the recombinant IL-1 receptor, IL-1 Trap, achieved ACR70 improvement scores (Lovell et al., 2006), and 57.9% of the patients treated with Canakinumab, an IL-1β antibody, achieved an improvement of ACR50, with 21% showing remission of disease (Ruperto et al., 2008). These results suggest that there are at least two subpopulations in sJIA with different underlying etiopathologies which are however convergent in the final inflammatory pathways, therefore resulting in identical clinical symptoms and laboratory findings. In one group of patients dysregulation of the IL-1 pathway may play a central role in pathogenesis, and therefore these patients respond well to IL-1 blockade. In the other group of patients, who do not respond to IL-1 blockade, a different inflammatory pathway may be central to pathogenesis, so that blockade of the IL-1 system does not correct the immune imbalance and has no effect on disease activity.

Baseline clinical information was available for 31 of the patients, all except those from the Netherlands. Prior to treatment initiation 26 patients were febrile. Following IL-1 blockade fever resolved in 23 patients, including 13 of the 16 febrile patients classed as non-responders. Similar results were seen in the 21 patients presenting with a rash at baseline. In

ID	Ethnicity	Country	Fever	Rash	Nº active joints	CRP	Responder
PID8710	Caucasian	USA	No	No	0	<0.05	Yes
PID8966	Caucasian	USA	No	No	0	<0.5	Yes
PID7549	Caucasian	USA	No	No	0	<0.05	Yes
PID8463	Caucasian	USA	No	No	0	<1.0	Yes
PID70838	Caucasian	USA	No	No	0	<0.05	Yes
PID57876	Caucasian	USA	No	No	18	7.9	No
PID916	Caucasian	USA	No	No	16	2.4	No
PID8292	Caucasian	USA	Yes	No	6	4.2	No
PID8610	Caucasian	USA	No	No	1	n/a	No
PID900	Caucasian	USA	No	No	24	2.7	No
PID1964	Caucasian	USA	No	No	23	5.6	No
PID70087	Caucasian	USA	Yes	Yes	24	8.9	No
PID8528	Caucasian	USA	No	No	23	14	No
PID7934	Caucasian	USA	No	No	32	17.8	No
PID8569	Caucasian	USA	Yes	No	25	6.8	No
FREC8B	Caucasian	France	No	No	1	18	No
FRECJL	Black African	France	No	No	0	6	Yes
FREDNR	Black African	France	No	No	0	6	Yes
FREC65	Caucasian	France	Yes	No	4	69	No
WKZ3	Caucasian	Netherlands	n/a[¶]	n/a[¶]	n/a[¶]	n/a[¶]	Yes
WKZ7	Caucasian	Netherlands	n/a[¶]	n/a[¶]	n/a[¶]	n/a[¶]	Yes
WKZ10	Caucasian	Netherlands	n/a[¶]	n/a[¶]	n/a[¶]	n/a[¶]	Yes[#]
WKZ13	Caucasian	Netherlands	n/a[¶]	n/a[¶]	n/a[¶]	n/a[¶]	Yes[#]
WKZ1	Caucasian	Netherlands	n/a[¶]	n/a[¶]	n/a[¶]	n/a[¶]	No*
WKZ9	Caucasian	Netherlands	n/a[¶]	n/a[¶]	n/a[¶]	n/a[¶]	No
WKZ2	Black African	Netherlands	n/a[¶]	n/a[¶]	n/a[¶]	n/a[¶]	No*
WKZ5	Black African	Netherlands	n/a[¶]	n/a[¶]	n/a[¶]	n/a[¶]	No*
WKZ6	Asian	Netherlands	n/a[¶]	n/a[¶]	n/a[¶]	n/a[¶]	No
WKZ11	Caucasian	Netherlands	n/a[¶]	n/a[¶]	n/a[¶]	n/a[¶]	No
WKZ12	Caucasian	Netherlands	n/a[¶]	n/a[¶]	n/a[¶]	n/a[¶]	No[#]
WKZ14	Caucasian	Netherlands	n/a[¶]	n/a[¶]	n/a[¶]	n/a[¶]	No
GOS506	Caucasian	UK	No	No	0	<1	Yes
GOS597	Caucasian	UK	No	No	0	<3	Yes
GOS666	Caucasian	UK	No	No	0	<3	Yes
GOS670	Black African	UK	No	No	2	>3	No
GOS671	Pakistani	UK	No	No	0	<5	Yes
GOS279	Caucasian	UK	No	No	1	120	No
GOS501	Caucasian	UK	No	Yes	0	7	Yes
GOS513	Iranian	UK	Yes	No	34	146	No
GOS520	Caucasian	UK	No	No	6	131	No*
GOS526	Caucasian	UK	No	No	3	37	No
GOS544	Caucasian	UK	No	No	0	44	No*
GOS641	Caucasian	UK	No	No	4	<3	No

Table 4.53 IL-1 blockade treated patient information

Forty three patients treated with IL-1 blockade were included in analysis of IL-1 polymorphism segregation according to IL-1 blockade response. Unless otherwise indicated all patients were treated with Anakinra. Patients were classified as either responders or non-responders to treatment, determined based on disease remission criteria: absence of fever and rash, no active joints, and normal CRP levels (<10) following treatment. The relevant clinical data for each patient, as well as ethnicity, country of treatment, and response classification, are shown.

[¶] = Clinical data not available, but response status defined as above, [#] = Patient was treated with Canakinumab, * = Patient experienced an adverse event during treatment, either disease flare or macrophage activation syndrome

	Allele	Responders	Non-Responders	p value	
IL18BP					
SNP1	1 2	0.875 0.125	0.942 0.0577	0.287	
SNP2	1 2	0.906 0.0938	0.808 0.192	0.211	
SNP 1- SNP 2 Haplotype	1-1 1-2 2-1 2-2	0.841 0.0339 0.0651 0.0599	0.750 0.192 0.0577 0.00	0.499 0.0664 0.338 1.00	0.0681
IL1RAP					
SNP 14	1 2	0.633 0.367	0.600 0.400	0.767	
SNP 25	1 2	0.700 0.300	0.800 0.200	0.313	
***IL1* ligand cluster**					
SNP 1	1 2	0.594 0.406	0.596 0.404	0.983	
SNP 4	1 2	0.594 0.406	0.654 0.346	0.580	
SNP 64	1 2	0.844 0.156	0.885 0.115	0.593	
***IL1* receptor cluster**					
SNP 8	1 2	0.700 0.300	0.769 0.231	0.492	

Table 4.54 Allele frequencies according to IL-1 blockade response

Forty three sJIA patients treated with IL-1 blockade were genotyped for the eight IL-1 gene family member SNPs which showed evidence of frequency differences between patients and healthy controls in the association study presented in this manuscript. Based on the criteria for disease remission patients were classified as either responders or non-responders to treatment. The allele frequencies of each SNP, and the four haplotypes of the *IL18BP* two SNP haplotype, in both patient groups (responders n=16, non-responders =27) are shown. The p-value results for differences in the allele frequencies between the two response groups are also shown.

19 of these patients the rash resolved following IL-1 blockade, including 11 of the 13 non-responders with rashes. This improvement in systemic features in the majority of sJIA patients treated with IL-1 blockade, regardless of overall response, has also been observed in a number of other studies (Gattorno M et al A&R 2008, Lequerre T et al AnRheumDis 2008). The evidence that blocking IL-1 signalling is effective in resolving rash and fever, regardless of the effect overall disease activity, suggests that in all sJIA patients, IL-1, a known endogenous pyrogen (Davidson et al., 1990), is an important mediator of these systemic features, either instigating or sustaining the response.

Although some slight differences in allele frequencies were observed between the two patient response groups they were only marginal and none reached statistical significance. The results must be viewed with extreme caution due to the small cohort size. All eight SNPs investigated for segregation according to response to IL-1 blockade were identified in the association study for disease susceptibility, regardless of response classification. It is therefore possible that these SNPs are purely related to sJIA susceptibility, with no difference in effect between the two response groups. Any SNPs in the candidate regions with a significant response-specific role may not have been identified in the initial association study due to insufficient power to detect an effect only present in a proportion of the patients. It is therefore possible that there are SNPs associated with only one response group, which were included in the initial study but the effect was not detected.

A larger genetic study investigating the differences between patients with sJIA according to response to IL-1 blockade treatment is currently underway by the UCL paediatric rheumatology group. Polymorphisms over a number of candidate genes are being investigated. The results from this may identify other important mediators in the disease eitopathology and response to IL-1 blockade. Identification of any indicators of response to IL-1 blockade would enable initiation of the appropriate treatment to those patients who would benefit significantly. This would facilitate quicker and greater resolution of disease symptoms, reducing the long term effects of disease activity, and the risk of side effects from other, ineffective, treatments. It would also be advantageous to identify patients who would not benefit from IL-1 blockade, saving time and avoiding the associated potential risks of disease flare, MAS, and infection from receiving ineffective immunosuppressive treatment.

Chapter 5
Discussion

5. Discussion

The case-control candidate gene association study presented in this thesis has identified a number of SNPs in IL-1 family genes showing evidence of association with susceptibility to sJIA. Evidence of association was seen with SNPs in the pro-inflammatory *IL1A* ligand, and the anti-inflammatory IL-1 receptor antagonist, *IL1RN*, and IL-18 antagonist, *IL18BP*. This study therefore provides evidence that genetic variation within members of the IL-1 gene family may be involved in susceptibility to sJIA, and that further investigation into the involvement of this gene family in sJIA is warranted. However, due to the small cohort sizes and testing of multiple genetic variants in this study these results must be viewed with caution. Replication in an independent and larger patient cohort is required to validate these associations. Particular caution is required with regard to the *IL18BP* haplotype association as the individual haplotypes are at low frequencies within the population, and conflicting results regarding the biological function of this haplotype have been presented within this thesis.

A number of approaches are available for investigation of genetic influences on disease susceptibility. The major variables differing between study designs are the cohorts to be included and the regions of the genome studied. Family based studies have been used previously to identify genetic variants linked to sJIA (Ogilvie et al., 2003). The major advantage of family based studies is that they are not affected by population stratification as non-affected family members are used as the control individuals. However, in diseases with low prevalence, including sJIA, the limited number of large multiplex families with multiple probands means that multicenter, and even international collaborations are essential to achieve sufficient power. In less common diseases a more feasible approach is case-control association studies, for which it is possible to collect larger cohort sizes. This study design is used widely in many disease studies. Publication in the last ten years of large databases of population genotyping data, including the International HapMap project and the Programmes for Genomic Applications projects, has enabled investigation of the linkage disequilibrium (LD) between SNPs in candidate genes. This facilitates the use of tSNP approaches, and thereby more thorough screening of candidate genes for disease associated polymorphisms.

The development of high throughput genotyping technology, such as the Illumina golden gate assay, has also made it feasible to investigate larger numbers of polymorphisms in each study. This also enables more in-depth investigation of genes to be performed. This is in contrast to many previous studies in which a limited number of variants, often only those previously demonstrated to affect expression or to be associated with a related disease, were studied. Genome wide association studies (GWAS) have become more widely used in recent years as the technology is now established and has become more affordable. GWAS is a powerful tool for the identification of positional candidate genes. However, because the coverage of each individual gene is relatively sparse, further investigation is needed for each associated region. Candidate gene studies enable a more comprehensive investigation of the gene of interest. Therefore, in studies such as that presented here, where specific genes of interest have been identified based on current knowledge of disease pathogenesis and gene function, a candidate gene approach is the most appropriate.

In this study population genotyping data from the International HapMap project was used to select tSNPs for inclusion in the association study. This strategy has become widely used in studies for a range of diseases (Dasgupta et al., 2011; Hurtado et al., 2010; Voisey et al., 2010; Currie et al., 2008; Pal et al., 2007). The employment of population genotyping data, however, means that two different populations are used; a reference population for tSNP selection, and a study population for association testing. It is therefore important that the reference population is representative of the study population, with comparable genetic ancestry and genomic LD structure. All individuals in both the patient and control cohorts included in the association study presented in this thesis were classified as being of white European descent. From the different ethnic populations included in the HapMap project, the CEU population of CEPH families was selected for use as the reference population for tSNP selection. The CEU population is comprised of individuals described as Utah residents of Northern and Western European ancestry, and is therefore of a similar geographical ancestry to the study population in this association study. A number of studies have investigated the performance of the HapMap CEU population as a reference population for a range of European populations. In a study by Lundmark et al. tSNPs were selected based on the CEU population genotypes. The selected tSNPs captured ($r^2 \geq 0.8$) between 95% and 87% of the genetic variation in Swedish, Finnish, Dutch, Australian of European ancestry, and British populations. These

five test populations and the HapMap CEU population also showed a similar LD pattern over the genomic region studied (Lundmark et al., 2008). Similar results were seen in studies of a Spanish population (Ribas et al., 2006), an Estonian population (Montpetit et al., 2006), three German, and four Italian populations (Mueller et al., 2005). SNPs with a minor allele frequency below 0.05 were however poorly captured (Montpetit et al., 2006).

Investigation of tSNPs in the association study of this project means that a number of the genotyped SNPs identified as being associated with sJIA are in high LD with other, un-genotyped SNPs. In this situation it is possible that any of the SNPs tagged by the tSNP tested for association, rather than the tSNP itself, is responsible for the disease association detected. Based on the association study results alone, it is not possible to determine which of the genotyped or tagged SNPs is directly responsible for an association. In order to determine which SNP was directly responsible for the association, functional investigations of all SNPs in the associated LD group will be required. The SNPs directly genotyped and identified as significantly associated with sJIA in this study are therefore not necessarily causative in disease susceptibility. Instead may be acting as a marker for the disease related SNP. This is, however, a possibility in all association studies.

The coverage of variation within the human genome included in the HapMap project, although extensive, is not complete. This means that there may be additional SNPs in high LD with those identified as associated with sJIA, which are the causal genetic variant. Functional assays to determine the regulatory roles of the associated SNPs would then fail to identify any biological effect resulting from the genetic association. Inclusion of genotyping data from the PGA increased the coverage of SNPs in the initial SNP set used for tSNP selection. Due to complete re-sequencing of a number of the candidate genes for SNP discovery by the PGA, the majority of polymorphisms within these genes present in the CEU population, will have been identified. However, at the time of tSNP selection for this thesis, the PGA had not sequenced *IL1RL1*, *IL18R2*, *CASP1*, or *NALP3*. The SNPs investigated in these candidate genes were therefore at a lower density than for the other candidate genes. Furthermore, it is possible that disease associated SNPs were not included in the tSNP selection stage due to occurrence in the healthy population at a frequency less than the 0.05 cut off used. Other large scale projects of common human genetic variation have recently been made publicly available. These could be utilised to

expand the list of SNPs which warrant further investigation as a result of the results presented in this thesis.

A frequent limitation in association studies of diseases with low prevalence is the necessity to use small cohorts. This leads to low power to detect significant disease associations. Using the CaTS power calculator for two stage association analysis (Skol et al., 2006) it was determined that the patient and control cohort sizes used in this study provide 84% power to detect a direct effect of a SNP with a frequency of 30% giving an odds ratio of 1.7. However, the majority of genetic associations with autoimmune diseases have ORs between 1 and 1.8, with many being towards the lower end of the range (Raj and Wakeland, 2011). Based on the same parameters as above, the power of the study to detect an association with OR=1.5 is only 61%, and 30% for an OR=1.3. The power calculation for the study presented in this thesis, based on an OR of 1.7, may therefore be overestimated. Following the initial case-control association study, the patient cohorts from both stages were re-analysed using the WTCCC control cohort of just under 5,000 healthy control individuals. This larger control cohort increases the power of the analysis to 96% power to detect a disease association under the same parameters as stated previously (OR=1.7). Inclusion of this additional analysis in the association study therefore increases for the limited power of the initial analysis. The greatest possible power for this study would have been achieved by using all of the patient samples together with the WTCCC controls in a single stage study. However, the WTCCC control data, and the patient DNA samples used in stage-2 of the study, were not available for inclusion at the time the study was started. Given the restraints of available samples at the different points throughout this study, a two-stage association study was the most appropriate at the time. Due to the limited power of this study as a result of small cohorts, it is possible that associations of smaller effect sizes were not detected. However, sJIA is a rare disease with a prevalence of approximately only 16 children in 100,000, and only approximately 20 new patients are seen each year at Great Ormond Street Hospital (personal communication P.Woo). Larger sJIA patient cohorts are consequently extremely difficult to collect. Previous case-control studies of sJIA have used similar and smaller sample sizes than that used in the study presented in this thesis. This includes patient cohorts of 92 (Fishman et al., 1998), 117 (Donn et al., 2001b), 136 (de Benedetti et al., 2003), 209 (Bukulmez et al., 2005), 172 (Fife et al., 2006), and 133 (Lamb et al., 2007). The cohort of 235 sJIA patients included in the current study is

favourably comparable to those used in previous published studies of sJIA. Following the initial case-control association study, the patient cohorts from both stages were re-analysed using the WTCCC control cohort of just under 5,000 healthy control individuals. This larger control cohort increases the power of the analysis to 96% power to detect a disease association under the same parameters as stated previously. Inclusion of this additional analysis in the association study therefore compensates for the limited power of the initial analysis.

In association studies with low prior probability, comprehensive investigation of a candidate gene requires testing of a large number of polymorphisms. This however introduces the issue of multiple testing. Using a significance cut off of $p<0.05$ for 372 SNPs, the number successfully genotyped in stage-1, it would be expected that 18.6 SNPs would show a significant difference in frequency by chance alone. However, it is widely accepted that correction for multiple testing using traditional methods such as Bonferroni correction, is not appropriate in association studies like this. This is because Bonferroni correction assumes that all variables tested are independent. As the SNPs tested in this study are in LD with each other, although $r^2<0.8$, this correction would be overly conservative (Lewis, 2002). The most reliable and convincing evidence that a genetic association is true, is replication in an independent population. Experimental demonstration that a disease associated SNP plays a direct role in regulating gene expression, and may therefore have a biological role in disease susceptibility, also adds further credence to a statistical association. This is the approach which has been employed in this study.

A two stage study design, rather than an initial and replication study, was used in this association study. The lack of any significant results when the stage-2 cohorts were analysed separately is concerning. One possible explanation for this is the significant frequency differences observed in the stage-1 analysis being false-positive results. The data from both stages of the study were analysed together by meta-analysis, stratified by study stage. The results of this analysis, coupled with tests for homogeneity of odds ratios, suggests that the stage-1 results are not false-positive. Only SNPs which showed statistically significant results in the meta-analysis of both stages, and did not show evidence of significant homogeneity of odds between the stages, were considered to be associated with sJIA. Therefore, the lack of significant association in the stage-2 cohorts

alone does not invalidate the results. The most likely explanation is that the second stage of the association study, which included a larger control cohort but a smaller patient cohort, was underpowered to detect any of the associations when analysed separately. It is also possible that the difference in the results for the two study stages is due to genetic and/or clinical heterogeneity in the patient cohorts. All of the sJIA patients included in the stage-1 cohort were UK based, of which the majority were patients at Great Ormond Street hospital in London. The stage-2 cohort of patients was comprised of three different populations; 33 patients from Great Ormond Street Hospital in London, 44 from Wexham Park Hospital in Berkshire, and 27 patients from Necker Hospital in Paris. Southern Europeans have been shown to be genetically distinct from Northern Europeans (Seldin et al., 2006). It is therefore possible that, due to inclusion of non-UK patients, that there is population stratification within the stage-2 patient cohort. Due to the small sample size of both patient cohorts there may, through sampling error, be underlying clinical differences between them. For example differences in disease severity, disease course, or clinical outcome. The patients from Wexham Park Hospital were all adults who continued to be affected by sJIA, and therefore may have a more severe disease outcome, potentially introducing patient heterogeneity into the study. Based on patient response to IL-1 blockade it appears that within sJIA there are at least two subtypes with differences in the underlying etiopathology (Quartier et al., 2006; Lequerre et al., 2008; Gattorno et al., 2008; Lovell et al., 2006; Ruperto et al., 2008). Only six of the patients included in the association study had been treated with IL-1 blockade therapy. It was therefore not possible to perform the analysis stratifying by treatment response. It is possible that the associations observed in this study are only true, or play a significant role, in a subset of sJIA patients, for example only those who respond to IL-1 blockade. By chance the stage-1 patient cohort may include a higher proportion of patients who respond to IL-1 blockade than the stage-2 cohort. This could explain the difference in the association results for the two stages of the study.

The study presented in this thesis has identified a number of SNPs as potentially associated with sJIA. Further work is necessary to confirm these results and investigate how these SNPs affect gene expression and hence disease susceptibility. Elucidation of genetic associations will provide further insight into disease pathogenesis of sJIA.

6. Future Work

In recent years there have been major advances in genotyping technology, enabling high throughput techniques, such as genome wide association studies (GWAS) to become more widely used. Although at the time of writing no GWAS studies have been published for sJIA, a study is currently underway in the USA. The results of large scale studies such as this will provide important data for researchers working on sJIA. The results will highlight disease associated regions for use in informing further studies to provide nearby positional candidate genes. Large scale studies, in concert with multi-center and international collaborations, will be invaluable for the progress of research into the genetics of sJIA.

As with all genetic association studies, replication in an independent and larger cohort will be essential for validation of the results presented in this thesis.

As discussed in chapter 5, there may be variants additional to those identified in this thesis, which are in high LD with the SNPs found to be associated with sJIA. Any additional SNPs would be candidates for involvement in disease susceptibility. Additional population genotyping databases, such as the 1000 genomes project (Durbin et al., 2010), could be utilised to further investigate the LD relationships of these SNPs. The 1000 genomes project aims to publish genome-wide sequencing results including all types of polymorphism present in the population at a frequency of at least 1%. This resource will therefore provide information not available from the HapMap project. Additional population genotyping data could also be used to further validate the LD data used in this thesis. Where additional population genotyping data is not available, further genotyping and direct association analysis of any SNPs in high, but not complete LD (<0.8 r^2 >1) with the associated SNPs, could be performed.

The *IL18BP* association results are potentially biologically interesting. Neither of the SNPs in the disease associated *IL18BP* haplotype were individually at significantly different frequencies in the patients and controls. However, in the transient transfection studies both the 1-2 and 2-1 haplotypes demonstrated significantly reduced levels of transcription *in vitro* compared to the 1-1 haplotype. It would therefore be interesting to test transcription levels according to allele at both SNPs individually. This could be carried out through generation of additional shorter constructs containing only one of the SNP loci. This would demonstrate whether the two SNPs play a role in regulating expression individually or as a haplotype.

It would be desirable to increase the number of healthy controls included in the analysis of PBMC IL-18BP expression levels according to haplotype. This would enable better interpretation of the results, either by increasing the power of the study sufficiently to detect any significant difference, or by increasing confidence in the negative findings. However, as the expected frequency of the 2-2 haplotype is only 0.01, based on the WTCCC population, it is unlikely that, without the availability of a very large cohort for screening, sufficient individuals carrying this haplotype would be identified. Another possible extension of the work presented in this thesis would be to measure the RNA and protein levels of IL-18BP in cell subsets rather than PBMCs. This would enable detection of any cell-specific haplotype effect on regulation of expression that would otherwise be masked in the admixture of cell types present in PBMC cultures.

Despite the differences in transcription shown in the transient transfection assay, no experimental evidence was found for allele-specific transcription factor binding. However, different cell types were used in the transcription and protein binding experiments. It is therefore possible that any allele-specific transcription factor binding is cell-type specific. This could be investigated by repetition of the EMSA experiment using nuclear extract from IFNγ stimulated HeLa cells.

Time constraints did not allow further investigation of the disease associated IL-1α SNPs. Further investigation into any regulatory function of these SNPs, including *in silico* sequence based predictions and transcriptional assays, could be utilised to determine the presence of any allele-specific effects on expression which may be involved in disease susceptibility.

References

Abinun,M., Flood,T.J., Cant,A.J., Veys,P., Gennery,A.R., Foster,H.E., Friswell,M., Baildam,E., Davidson,J., Southwood,T.R., Livermore,P., and Wedderburn,L.R. (2009). Autologous T cell depleted haematopoietic stem cell transplantation in children with severe juvenile idiopathic arthritis in the UK (2000-2007). Mol. Immunol. *47*, 46-51.

Adams,A. and Lehman,T.J. (2005). Update on the pathogenesis and treatment of systemic onset juvenile rheumatoid arthritis. Curr. Opin. Rheumatol. *17*, 612-616.

Aganna,E., Martinon,F., Hawkins,P.N., Ross,J.B., Swan,D.C., Booth,D.R., Lachmann,H.J., Bybee,A., Gaudet,R., Woo,P., Feighery,C., Cotter,F.E., Thome,M., Hitman,G.A., Tschopp,J., and McDermott,M.F. (2002). Association of mutations in the NALP3/CIAS1/PYPAF1 gene with a broad phenotype including recurrent fever, cold sensitivity, sensorineural deafness, and AA amyloidosis. Arthritis Rheum. *46*, 2445-2452.

Agostini,L., Martinon,F., Burns,K., McDermott,M.F., Hawkins,P.N., and Tschopp,J. (2004b). NALP3 forms an IL-1beta-processing inflammasome with increased activity in Muckle-Wells autoinflammatory disorder. Immunity *20*, 319-325.

Agostini,L., Martinon,F., Burns,K., McDermott,M.F., Hawkins,P.N., and Tschopp,J. (2004a). NALP3 forms an IL-1beta-processing inflammasome with increased activity in Muckle-Wells autoinflammatory disorder. Immunity *20*, 319-325.

Agostini,L., Martinon,F., Burns,K., McDermott,M.F., Hawkins,P.N., and Tschopp,J. (2004c). NALP3 forms an IL-1beta-processing inflammasome with increased activity in Muckle-Wells autoinflammatory disorder. Immunity *20*, 319-325.

Ahmed,Z.M., Masmoudi,S., Kalay,E., Belyantseva,I.A., Mosrati,M.A., Collin,R.W., Riazuddin,S., Hmani-Aifa,M., Venselaar,H., Kawar,M.N., Tlili,A., van der,Z.B., Khan,S.Y., Ayadi,L., Riazuddin,S.A., Morell,R.J., Griffith,A.J., Charfedine,I., Caylan,R., Oostrik,J., Karaguzel,A., Ghorbel,A., Riazuddin,S., Friedman,T.B., Ayadi,H., and Kremer,H. (2008). Mutations of LRTOMT, a fusion gene with alternative reading frames, cause nonsyndromic deafness in humans. Nat. Genet. *40*, 1335-1340.

Aizawa,Y., Akita,K., Taniai,M., Torigoe,K., Mori,T., Nishida,Y., Ushio,S., Nukada,Y., Tanimoto,T., Ikegami,H., Ikeda,M., and Kurimoto,M. (1999). Cloning and expression of interleukin-18 binding protein. FEBS Lett. *445*, 338-342.

Aizawa,Y., Sutoh,S., Matsuoka,M., Negishi,M., Torii,A., Miyakawa,Y., Sugisaka,H., Nakamura,M., and Toda,G. (2005). Association of interleukin-18 gene single-nucleotide polymorphisms with susceptibility to inflammatory bowel disease. Tissue Antigens *65*, 88-92.

Aksentijevich,I., Masters,S.L., Ferguson,P.J., Dancey,P., Frenkel,J., Royen-Kerkhoff,A., Laxer,R., Tedgard,U., Cowen,E.W., Pham,T.H., Booty,M., Estes,J.D., Sandler,N.G., Plass,N., Stone,D.L., Turner,M.L., Hill,S., Butman,J.A., Schneider,R., Babyn,P., El Shanti,H.I., Pope,E., Barron,K., Bing,X., Laurence,A., Lee,C.C., Chapelle,D., Clarke,G.I., Ohson,K., Nicholson,M., Gadina,M., Yang,B., Korman,B.D., Gregersen,P.K., van Hagen,P.M., Hak,A.E., Huizing,M., Rahman,P., Douek,D.C., Remmers,E.F., Kastner,D.L., and Goldbach-Mansky,R. (2009). An autoinflammatory disease with deficiency of the interleukin-1-receptor antagonist. N. Engl. J. Med. *360*, 2426-2437.

Aksentijevich,I., Nowak,M., Mallah,M., Chae,J.J., Watford,W.T., Hofmann,S.R., Stein,L., Russo,R., Goldsmith,D., Dent,P., Rosenberg,H.F., Austin,F., Remmers,E.F., Balow,J.E., Jr., Rosenzweig,S., Komarow,H., Shoham,N.G., Wood,G., Jones,J., Mangra,N., Carrero,H., Adams,B.S., Moore,T.L., Schikler,K., Hoffman,H., Lovell,D.J., Lipnick,R., Barron,K., O'Shea,J.J., Kastner,D.L., and Goldbach-Mansky,R. (2002). De novo CIAS1 mutations, cytokine activation, and evidence for genetic heterogeneity in patients with neonatal-onset multisystem inflammatory disease (NOMID): a new member of the expanding family of pyrin-associated autoinflammatory diseases. Arthritis Rheum. *46*, 3340-3348.

Ali,S., Huber,M., Kollewe,C., Bischoff,S.C., Falk,W., and Martin,M.U. (2007). IL-1 receptor accessory protein is essential for IL-33-induced activation of T lymphocytes and mast cells. Proc. Natl. Acad. Sci. U. S. A *104*, 18660-18665.

Allantaz,F., Chaussabel,D., Stichweh,D., Bennett,L., Allman,W., Mejias,A., Ardura,M., Chung,W., Smith,E., Wise,C., Palucka,K., Ramilo,O., Punaro,M., Banchereau,J., and Pascual,V. (2007). Blood leukocyte microarrays to diagnose systemic onset juvenile idiopathic arthritis and follow the response to IL-1 blockade. J. Exp. Med. *204*, 2131-2144.

Alsaeid,K., Haider,M.Z., and Ayoub,E.M. (2003). Angiotensin converting enzyme gene insertion-deletion polymorphism is associated with juvenile rheumatoid arthritis. J. Rheumatol. *30*, 2705-2709.

Andersson,J., Bjork,L., Dinarello,C.A., Towbin,H., and Andersson,U. (1992). Lipopolysaccharide induces human interleukin-1 receptor antagonist and interleukin-1 production in the same cell. Eur. J. Immunol. *22*, 2617-2623.

Arend,W.P., Joslin,F.G., and Massoni,R.J. (1985). Effects of immune complexes on production by human monocytes of interleukin 1 or an interleukin 1 inhibitor. J. Immunol. *134*, 3868-3875.

Arostegui,J.I., Lopez,S., Pascal,M., Clemente,D., Aymerich,M., Balaguer,F., Goel,A., Fournier,d.C., Rius,J., Plaza,S., Lopez Robledillo,J.C., Juan,M., Ibanez,M., and Yague,J. (2010). A somatic NLRP3 mutation as a cause of a sporadic case of chronic infantile neurologic, cutaneous, articular syndrome/neonatal-onset multisystem inflammatory disease: Novel evidence of the role of low-level mosaicism as the pathophysiologic mechanism underlying mendelian inherited diseases. Arthritis Rheum. *62*, 1158-1166.

Auron,P.E. (1998). The interleukin 1 receptor: ligand interactions and signal transduction. Cytokine Growth Factor Rev. *9*, 221-237.

Ayaz,N.A., Ozen,S., Bilginer,Y., Erguven,M., Taskiran,E., Yilmaz,E., Besbas,N., Topaloglu,R., and Bakkaloglu,A. (2009). MEFV mutations in systemic onset juvenile idiopathic arthritis. Rheumatology (Oxford) *48*, 23-25.

Barnes,M.G., Grom,A.A., Thompson,S.D., Griffin,T.A., Pavlidis,P., Itert,L., Fall,N., Sowders,D.P., Hinze,C.H., Aronow,B.J., Luyrink,L.K., Srivastava,S., Ilowite,N.T., Gottlieb,B.S., Olson,J.C., Sherry,D.D., Glass,D.N., and Colbert,R.A. (2009). Subtype-specific peripheral blood gene expression profiles in recent-onset juvenile idiopathic arthritis. Arthritis Rheum. *60*, 2102-2112.

Barrett,J.C., Fry,B., Maller,J., and Daly,M.J. (2005). Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. *21*, 263-265.

Batzoglou,S., Pachter,L., Mesirov,J.P., Berger,B., and Lander,E.S. (2000). Human and mouse gene structure: comparative analysis and application to exon prediction. Genome Res. *10*, 950-958.

Behrens,E.M., Beukelman,T., Paessler,M., and Cron,R.Q. (2007). Occult macrophage activation syndrome in patients with systemic juvenile idiopathic arthritis. J. Rheumatol. *34*, 1133-1138.

Ben Aribia,M.H., Leroy,E., Lantz,O., Metivier,D., Autran,B., Charpentier,B., Hercend,T., and Senik,A. (1987). rIL 2-induced proliferation of human circulating NK cells and T lymphocytes: synergistic effects of IL 1 and IL 2. J. Immunol. *139*, 443-451.

Ben Sasson,S.Z., Hu-Li,J., Quiel,J., Cauchetaux,S., Ratner,M., Shapira,I., Dinarello,C.A., and Paul,W.E. (2009). IL-1 acts directly on CD4 T cells to enhance their antigen-driven expansion and differentiation. Proc. Natl. Acad. Sci. U. S. A *106*, 7119-7124.

Benjamini,E., Coico,R., and Sunshine,G. (2000). Immunology A Short Course. (New York: Wiley-Liss).

Bergers,G., Reikerstorfer,A., Braselmann,S., Graninger,P., and Busslinger,M. (1994). Alternative promoter usage of the Fos-responsive gene Fit-1 generates mRNA isoforms coding for either secreted or membrane-bound proteins related to the IL-1 receptor. EMBO J. *13*, 1176-1188.

Bergholdt,R., Larsen,Z.M., Andersen,N.A., Johannesen,J., Kristiansen,O.P., Mandrup-Poulsen,T., Nerup,J., and Pociot,F. (2000). Characterization of new polymorphisms in the 5' UTR of the human interleukin-1 receptor type 1 (IL1R1) gene: linkage to type 1 diabetes and correlation to IL-1RI plasma level. Genes Immun. *1*, 495-500.

Beuscher,H.U. and Colten,H.R. (1988). Structure and function of membrane IL-1. Mol. Immunol. *25*, 1189-1199.

Bevan,S. and Raynes,J.G. (1991). IL-1 receptor antagonist regulation of acute phase protein synthesis in human hepatoma cells. J. Immunol. *147*, 2574-2578.

Blankenberg,S., Godefroy,T., Poirier,O., Rupprecht,H.J., Barbaux,S., Bickel,C., Nicaud,V., Schnabel,R., Kee,F., Morrison,C., Evans,A., Lackner,K.J., Cambien,F., Munzel,T., and Tiret,L. (2006). Haplotypes of the caspase-1 gene, plasma caspase-1 levels, and cardiovascular risk. Circ. Res. *99*, 102-108.

Blumberg,H., Dinh,H., Dean,C., Jr., Trueblood,E.S., Bailey,K., Shows,D., Bhagavathula,N., Aslam,M.N., Varani,J., Towne,J.E., and Sims,J.E. (2010). IL-1RL2 and its ligands contribute to the cytokine network in psoriasis. J. Immunol. *185*, 4354-4362.

Blumberg,H., Dinh,H., Trueblood,E.S., Pretorius,J., Kugler,D., Weng,N., Kanaly,S.T., Towne,J.E., Willis,C.R., Kuechle,M.K., Sims,J.E., and Peschon,J.J. (2007). Opposing activities of two novel members of the IL-1 ligand family regulate skin inflammation. J. Exp. Med. *204*, 2603-2614.

Born,T.L., Smith,D.E., Garka,K.E., Renshaw,B.R., Bertles,J.S., and Sims,J.E. (2000). Identification and characterization of two members of a novel class of the interleukin-1 receptor (IL-1R) family. Delineation of a new class of IL-1R-related proteins based on signaling. J. Biol. Chem. *275*, 29946-29954.

Born,T.L., Thomassen,E., Bird,T.A., and Sims,J.E. (1998). Cloning of a novel receptor subunit, AcPL, required for interleukin-18 signaling. J. Biol. Chem. *273*, 29445-29450.

Botstein,D., White,R.L., Skolnick,M., and Davis,R.W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. *32*, 314-331.

Bowyer,S.L., Roettcher,P.A., Higgins,G.C., Adams,B., Myers,L.K., Wallace,C., Rennebohm,R., Moore,T.L., Pepmueller,P.H., Spencer,C., Wagner-Weiner,L., Rabinovich,E., Passo,M., Lovell,D.J., McCurdy,D., Zemel,L., Schikler,K.N., Szer,I., Kurtin,P., and Lindsley,C. (2003). Health status of patients with juvenile rheumatoid arthritis at 1 and 5 years after diagnosis. J. Rheumatol. *30*, 394-400.

Breit,W., Frosch,M., Meyer,U., Heinecke,A., and Ganser,G. (2000). A subgroup-specific evaluation of the efficacy of intraarticular triamcinolone hexacetonide in juvenile chronic arthritis. J. Rheumatol. *27*, 2696-2702.

Breslow,N.E. and Day,N.E. (1980). Classical methods of analysis of grouped data. In Statistical Methods in Cancer Research, (Lyon: IARC Scientific Publications), pp. 122-156.

Bresnihan,B., Roux-Lombard,P., Murphy,E., Kane,D., FitzGerald,O., and Dayer,J.M. (2002). Serum interleukin 18 and interleukin 18 binding protein in rheumatoid arthritis. Ann. Rheum. Dis. *61*, 726-729.

Brinkman,D.M., de Kleer,I.M., ten Cate,R., van Rossum,M.A., Bekkering,W.P., Fasth,A., van Tol,M.J., Kuis,W., Wulffraat,N.M., and Vossen,J.M. (2007). Autologous stem cell transplantation in children with severe progressive systemic or polyarticular juvenile idiopathic arthritis: long-term follow-up of a prospective clinical trial. Arthritis Rheum. *56*, 2410-2421.

Buchan,G., Barrett,K., Turner,M., Chantry,D., Maini,R.N., and Feldmann,M. (1988). Interleukin-1 and tumour necrosis factor mRNA expression in rheumatoid arthritis: prolonged production of IL-1 alpha. Clin. Exp. Immunol. *73*, 449-455.

Buchs,N., di Giovine,F.S., Silvestri,T., Vannier,E., Duff,G.W., and Miossec,P. (2001). IL-1B and IL-1Ra gene polymorphisms and disease severity in rheumatoid arthritis: interaction with their plasma levels. Genes Immun. *2*, 222-228.

Buck,L. and Axel,R. (1991). A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. Cell *65*, 175-187.

Bufler,P., Azam,T., Gamboni-Robertson,F., Reznikov,L.L., Kumar,S., Dinarello,C.A., and Kim,S.H. (2002). A complex of the IL-1 homologue IL-1F7b and IL-18-binding protein reduces IL-18 activity. Proc. Natl. Acad. Sci. U. S. A *99*, 13723-13728.

Bukulmez,H., Fife,M., Tsoras,M., Thompson,S.D., Twine,N.A., Woo,P., Olson,J.M., Elston,R.C., Glass,D.N., and Colbert,R.A. (2005). Tapasin gene polymorphism in systemic

onset juvenile rheumatoid arthritis: a family-based case-control study. Arthritis Res. Ther. *7*, R285-R290.

Burns,K., Martinon,F., Esslinger,C., Pahl,H., Schneider,P., Bodmer,J.L., Di Marco,F., French,L., and Tschopp,J. (1998). MyD88, an adapter protein involved in interleukin-1 signaling. J. Biol. Chem. *273*, 12203-12209.

Busfield,S.J., Comrack,C.A., Yu,G., Chickering,T.W., Smutko,J.S., Zhou,H., Leiby,K.R., Holmgren,L.M., Gearing,D.P., and Pan,Y. (2000). Identification and gene organization of three novel members of the IL-1 family on human chromosome 2. Genomics *66*, 213-216.

Calabro,J.J., Holgerson,W.B., Sonpal,G.M., and Khoury,M.I. (1976). Juvenile rheumatoid arthritis: a general review and report of 100 patients observed for 15 years. Semin. Arthritis Rheum. *5*, 257-298.

Cann,H.M. (1992). CEPH maps. Curr. Opin. Genet. Dev. *2*, 393-399.

Canna,S., Frankovich,J., Higgins,G., Narkewicz,M.R., Nash,S.R., Hollister,J.R., Soep,J.B., and Dragone,L.L. (2009). Acute hepatitis in three patients with systemic juvenile idiopathic arthritis taking interleukin-1 receptor antagonist. Pediatr. Rheumatol. Online. J. *7*, 21.

Cao,Z., Xiong,J., Takeuchi,M., Kurama,T., and Goeddel,D.V. (1996). TRAF6 is a signal transducer for interleukin-1. Nature *383*, 443-446.

Carlson,C.S., Eberle,M.A., Rieder,M.J., Yi,Q., Kruglyak,L., and Nickerson,D.A. (2004). Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. Am. J. Hum. Genet. *74*, 106-120.

Carreira,P.E., Gonzalez-Crespo,M.R., Ciruelo,E., Pablos,J.L., Santiago,B., Gomez-Camara,A., and Gomez-Reino,J.J. (2005). Polymorphism of the interleukin-1 receptor antagonist gene: a factor in susceptibility to rheumatoid arthritis in a Spanish population. Arthritis Rheum. *52*, 3015-3019.

Carter,D.B., Deibel,M.R., Jr., Dunn,C.J., Tomich,C.S., Laborde,A.L., Slightom,J.L., Berger,A.E., Bienkowski,M.J., Sun,F.F., McEwan,R.N., and . (1990). Purification, cloning, expression and biological characterization of an interleukin-1 receptor antagonist protein. Nature *344*, 633-638.

Cerretti,D.P., Kozlosky,C.J., Mosley,B., Nelson,N., Van Ness,K., Greenstreet,T.A., March,C.J., Kronheim,S.R., Druck,T., Cannizzaro,L.A., and . (1992). Molecular cloning of the interleukin-1 beta converting enzyme. Science *256*, 97-100.

Chen,C.J., Kono,H., Golenbock,D., Reed,G., Akira,S., and Rock,K.L. (2007). Identification of a key pathway required for the sterile inflammatory response triggered by dying cells. Nat. Med. *13*, 851-856.

Ching,S., Zhang,H., Belevych,N., He,L., Lai,W., Pu,X.A., Jaeger,L.B., Chen,Q., and Quan,N. (2007). Endothelial-specific knockdown of interleukin-1 (IL-1) type 1 receptor differentially alters CNS responses to IL-1 depending on its route of administration. J. Neurosci. *27*, 10476-10486.

Chou,C.T., Timms,A.E., Wei,J.C., Tsai,W.C., Wordsworth,B.P., and Brown,M.A. (2006). Replication of association of IL1 gene complex members with ankylosing spondylitis in Taiwanese Chinese. Ann. Rheum. Dis. *65*, 1106-1109.

Chou,I.C., Lin,H.C., Wang,C.H., Lin,W.D., Lee,C.C., Tsai,C.H., and Tsai,F.J. (2010). Polymorphisms of interleukin 1 gene IL1RN are associated with Tourette syndrome. Pediatr. Neurol. *42*, 320-324.

Chowdhury,M.A., Kuivaniemi,H., Romero,R., Edwin,S., Chaiworapongsa,T., and Tromp,G. (2006). Identification of novel functional sequence variants in the gene for peptidase inhibitor 3. BMC. Med. Genet. *7*, 49.

Chustz,R.T., Nagarkar,D.R., Poposki,J.A., Favoreto,J.S., Avila,P.C., Schleimer,R.P., and Kato,A. (2010). Regulation and Function of the IL-1 Family Cytokine IL-1F9 in Human Bronchial Epithelial Cells. Am. J. Respir. Cell Mol. Biol.

Clayton,D.G. (2007). Population Association. In Handbook of Statistical Genetics, D.J.Balding, C.E.Bishop, and C.Cannings, eds. (Chichester: Wiley).

Clemens,L.E., Albert,E., and Ansell,B.M. (1985). Sibling pairs affected by chronic arthritis of childhood: evidence for a genetic predisposition. J. Rheumatol. *12*, 108-113.

Collins,C.S., Hong,J., Sapinoso,L., Zhou,Y., Liu,Z., Micklash,K., Schultz,P.G., and Hampton,G.M. (2006). A small interfering RNA screen for modulators of tumor cell motility identifies MAP4K4 as a promigratory kinase. Proc. Natl. Acad. Sci. U. S. A *103*, 3775-3780.

Colotta,F., Re,F., Muzio,M., Bertini,R., Polentarutti,N., Sironi,M., Giri,J.G., Dower,S.K., Sims,J.E., and Mantovani,A. (1993). Interleukin-1 type II receptor: a decoy target for IL-1 that is regulated by IL-4. Science *261*, 472-475.

Constantin,A., Loubet-Lescoulie,P., Lambert,N., Yassine-Diab,B., Abbal,M., Mazieres,B., de Preval,C., and Cantagrel,A. (1998). Antiinflammatory and immunoregulatory action of methotrexate in the treatment of rheumatoid arthritis: evidence of increased interleukin-4 and interleukin-10 gene expression demonstrated in vitro by competitive reverse transcriptase-polymerase chain reaction. Arthritis Rheum. *41*, 48-57.

Cooke,S.P., Forrest,G., Venables,P.J., and Hajeer,A. (1998). The delta32 deletion of CCR5 receptor in rheumatoid arthritis. Arthritis Rheum. *41*, 1135-1136.

Corbaz,A., ten Hove,T., Herren,S., Graber,P., Schwartsburd,B., Belzer,I., Harrison,J., Plitz,T., Kosco-Vilbois,M.H., Kim,S.H., Dinarello,C.A., Novick,D., van Deventer,S., and Chvatchko,Y. (2002). IL-18-binding protein expression by endothelial cells and macrophages is up-regulated during active Crohn's disease. J. Immunol. *168*, 3608-3616.

Cordell,H.J. and Clayton,D.G. (2002). A unified stepwise regression procedure for evaluating the relative effects of polymorphisms within a gene using case/control or family data: application to HLA in type 1 diabetes. Am. J. Hum. Genet. *70*, 124-141.

Costelloe,C., Watson,M., Murphy,A., McQuillan,K., Loscher,C., Armstrong,M.E., Garlanda,C., Mantovani,A., O'Neill,L.A., Mills,K.H., and Lynch,M.A. (2008). IL-1F5 mediates anti-inflammatory activity in the brain through induction of IL-4 following interaction with SIGIRR/TIR8. J. Neurochem. *105*, 1960-1969.

Coyle,A.J., Lloyd,C., Tian,J., Nguyen,T., Erikkson,C., Wang,L., Ottoson,P., Persson,P., Delaney,T., Lehar,S., Lin,S., Poisson,L., Meisel,C., Kamradt,T., Bjerke,T., Levinson,D., and Gutierrez-Ramos,J.C. (1999). Crucial role of the interleukin 1 receptor family member T1/ST2 in T helper cell type 2-mediated lung mucosal immune responses. J. Exp. Med. *190*, 895-902.

Cullinan,E.B., Kwee,L., Nunes,P., Shuster,D.J., Ju,G., McIntyre,K.W., Chizzonite,R.A., and Labow,M.A. (1998). IL-1 receptor accessory protein is an essential component of the IL-1 receptor. J. Immunol. *161*, 5614-5620.

Currie,D., Maxwell,A.P., Sadlier,D., and McKnight,A.J. (2008). Investigation of Adducin 2 (beta) DNA polymorphisms in genetic predisposition to diabetic nephropathy in Type 1 diabetes. Diabet. Med. *25*, 1001-1005.

Dale,M. and Nicklin,M.J. (1999). Interleukin-1 receptor cluster: gene organization of IL1R2, IL1R1, IL1RL2 (IL-1Rrp2), IL1RL1 (T1/ST2), and IL18R1 (IL-1Rrp) on human chromosome 2q. Genomics *57*, 177-179.

Dasgupta,S., Demirci,F.Y., Dressen,A.S., Kao,A.H., Rhew,E.Y., Ramsey-Goldman,R., Manzi,S., Kammerer,C.M., and Kamboh,M.I. (2011). Association analysis of PON2 genetic variants with serum paraoxonase activity and systemic lupus erythematosus. BMC. Med. Genet. *12*, 7.

Date,Y., Seki,N., Kamizono,S., Higuchi,T., Hirata,T., Miyata,K., Ohkuni,M., Tatsuzawa,O., Yokota,S., Joo,K., Ueda,K., Sasazuki,T., Kimura,A., Itoh,K., and Kato,H. (1999b). Identification of a genetic risk factor for systemic juvenile rheumatoid arthritis in the 5'-flanking region of the TNFalpha gene and HLA genes. Arthritis Rheum. *42*, 2577-2582.

Date,Y., Seki,N., Kamizono,S., Higuchi,T., Hirata,T., Miyata,K., Ohkuni,M., Tatsuzawa,O., Yokota,S., Joo,K., Ueda,K., Sasazuki,T., Kimura,A., Itoh,K., and Kato,H. (1999a). Identification of a genetic risk factor for systemic juvenile rheumatoid arthritis in the 5'-flanking region of the TNFalpha gene and HLA genes. Arthritis Rheum. *42*, 2577-2582.

Dausset,J., Cann,H., Cohen,D., Lathrop,M., Lalouel,J.M., and White,R. (1990). Centre d'etude du polymorphisme humain (CEPH): collaborative genetic mapping of the human genome. Genomics *6*, 575-577.

Davidson,J. (2000). Juvenile idiopathic arthritis: a clinical overview. Eur. J. Radiol. *33*, 128-134.

Davidson,J., Milton,A.S., and Rotondo,D. (1990). A study of the pyrogenic actions of interleukin-1 alpha and interleukin-1 beta: interactions with a steroidal and a non-steroidal anti-inflammatory agent. Br. J. Pharmacol. *100*, 542-546.

Davies,K. and Woo,P. (2004). The process of inflammation. In Oxford textbook of Rheumatology, D.A.Isenberg, P.J.Maddison, P.Woo, D.Glass, and F.C.Breedveld, eds. (Oxford: Oxford University Press), pp. 399-410.

Davies,U.M., Rooney,M., Preece,M.A., Ansell,B.M., and Woo,P. (1994). Treatment of growth retardation in juvenile chronic arthritis with recombinant human growth hormone. J. Rheumatol. *21*, 153-158.

Day,T.G., Ramanan,A.V., Hinks,A., Lamb,R., Packham,J., Wise,C., Punaro,M., and Donn,R.P. (2008b). Autoinflammatory genes and susceptibility to psoriatic juvenile idiopathic arthritis. Arthritis Rheum. *58*, 2142-2146.

Day,T.G., Ramanan,A.V., Hinks,A., Lamb,R., Packham,J., Wise,C., Punaro,M., and Donn,R.P. (2008a). Autoinflammatory genes and susceptibility to psoriatic juvenile idiopathic arthritis. Arthritis Rheum. *58*, 2142-2146.

de Bakker,P.I., Yelensky,R., Pe'er,I., Gabriel,S.B., Daly,M.J., and Altshuler,D. (2005). Efficiency and power in genetic association studies. Nat. Genet. *37*, 1217-1223.

de Benedetti, F., Brunner, H., Ruperto, N., Wright, S., Kenwright, A., Cuttica, R., Woo, P., Schneider, R., Lovell, D., and Martini, A. Efficacy and Safety of tocilizumab in patients with systemic Juvenile Idiopathic Arthritis (sJIA): 12-week data from the phase 3 TENDER trial. Annals of the Rheumatic Diseases 69[Supplemental 3], 146. 2010.
Ref Type: Abstract

De Benedetti,F., Massa,M., Pignatti,P., Albani,S., Novick,D., and Martini,A. (1994a). Serum soluble interleukin 6 (IL-6) receptor and IL-6/soluble IL-6 receptor complex in systemic juvenile rheumatoid arthritis. J. Clin. Invest *93*, 2114-2119.

De Benedetti,F., Massa,M., Pignatti,P., Albani,S., Novick,D., and Martini,A. (1994b). Serum soluble interleukin 6 (IL-6) receptor and IL-6/soluble IL-6 receptor complex in systemic juvenile rheumatoid arthritis. J. Clin. Invest *93*, 2114-2119.

de Benedetti,F., Massa,M., Robbioni,P., Ravelli,A., Burgio,G.R., and Martini,A. (1991b). Correlation of serum interleukin-6 levels with joint involvement and thrombocytosis in systemic juvenile rheumatoid arthritis. Arthritis Rheum. *34*, 1158-1163.

de Benedetti,F., Massa,M., Robbioni,P., Ravelli,A., Burgio,G.R., and Martini,A. (1991a). Correlation of serum interleukin-6 levels with joint involvement and thrombocytosis in systemic juvenile rheumatoid arthritis. Arthritis Rheum. *34*, 1158-1163.

de Benedetti,F., Meazza,C., Vivarelli,M., Rossi,F., Pistorio,A., Lamb,R., Lunt,M., Thomson,W., Ravelli,A., Donn,R., and Martini,A. (2003). Functional and prognostic relevance of the -173 polymorphism of the macrophage migration inhibitory factor gene in systemic-onset juvenile idiopathic arthritis. Arthritis Rheum. *48*, 1398-1407.

de Benedetti,F., Pignatti,P., Massa,M., Sartirana,P., Ravelli,A., and Martini,A. (1995). Circulating levels of interleukin 1 beta and of interleukin 1 receptor antagonist in systemic juvenile chronic arthritis. Clin. Exp. Rheumatol. *13*, 779-784.

de Jager,W., Hoppenreijs,E.P., Wulffraat,N.M., Wedderburn,L.R., Kuis,W., and Prakken,B.J. (2007). Blood and synovial fluid cytokine signatures in patients with juvenile idiopathic arthritis: a cross-sectional study. Ann. Rheum. Dis. *66*, 589-598.

de Jager,W., Vastert,S.J., Beekman,J.M., Wulffraat,N.M., Kuis,W., Coffer,P.J., and Prakken,B.J. (2009). Defective phosphorylation of interleukin-18 receptor beta causes impaired natural killer cell function in systemic-onset juvenile idiopathic arthritis. Arthritis Rheum. *60*, 2782-2793.

Debets,R., Timans,J.C., Homey,B., Zurawski,S., Sana,T.R., Lo,S., Wagner,J., Edwards,G., Clifford,T., Menon,S., Bazan,J.F., and Kastelein,R.A. (2001). Two novel IL-1 family members, IL-1 delta and IL-1 epsilon, function as an antagonist and agonist of NF-kappa B activation through the orphan IL-1 receptor-related protein 2. J. Immunol. *167*, 1440-1446.

Dermitzakis,E.T. and Clark,A.G. (2002). Evolution of transcription factor binding sites in Mammalian gene regulatory regions: conservation and turnover. Mol. Biol. Evol. *19*, 1114-1121.

Derrien,V., Couillault,C., Franco,M., Martineau,S., Montcourrier,P., Houlgatte,R., and Chavrier,P. (2002). A conserved C-terminal domain of EFA6-family ARF6-guanine nucleotide exchange factors induces lengthening of microvilli-like membrane protrusions. J. Cell Sci. *115*, 2867-2879.

Dinarello,C., Arend,W., Sims,J., Smith,D., Blumberg,H., O'Neill,L., Goldbach-Mansky,R., Pizarro,T., Hoffman,H., Bufler,P., Nold,M., Ghezzi,P., Mantovani,A., Garlanda,C., Boraschi,D., Rubartelli,A., Netea,M., van der,M.J., Joosten,L., Mandrup-Poulsen,T., Donath,M., Lewis,E., Pfeilschifter,J., Martin,M., Kracht,M., Muehl,H., Novick,D., Lukic,M., Conti,B., Solinger,A., Peyman,K., van,d., V, and Gabel,C. (2010). IL-1 family nomenclature. Nat. Immunol. *11*, 973.

Dinarello,C.A. (1996). Biologic basis for interleukin-1 in disease. Blood *87*, 2095-2147.

Dinarello,C.A. (1997). Interleukin-1. Cytokine Growth Factor Rev. *8*, 253-265.

Dinarello,C.A. (1999). IL-18: A TH1-inducing, proinflammatory cytokine and new member of the IL-1 family. J. Allergy Clin. Immunol. *103*, 11-24.

Dinarello,C.A., Cannon,J.G., Mier,J.W., Bernheim,H.A., LoPreste,G., Lynn,D.L., Love,R.N., Webb,A.C., Auron,P.E., Reuben,R.C., and . (1986). Multiple biological activities of human recombinant interleukin 1. J. Clin. Invest *77*, 1734-1739.

Dode,C., Le Du,N., Cuisset,L., Letourneur,F., Berthelot,J.M., Vaudour,G., Meyrier,A., Watts,R.A., Scott,D.G., Nicholls,A., Granel,B., Frances,C., Garcier,F., Edery,P., Boulinguez,S., Domergues,J.P., Delpech,M., and Grateau,G. (2002). New mutations of CIAS1 that are responsible for Muckle-Wells syndrome and familial cold urticaria: a novel mutation underlies both syndromes. Am. J. Hum. Genet. *70*, 1498-1506.

Donn,R., Alourfi,Z., de Benedetti,F., Meazza,C., Zeggini,E., Lunt,M., Stevens,A., Shelley,E., Lamb,R., Ollier,W.E., Thomson,W., and Ray,D. (2002). Mutation screening of the macrophage migration inhibitory factor gene: positive association of a functional polymorphism of macrophage migration inhibitory factor with juvenile idiopathic arthritis. Arthritis Rheum. *46*, 2402-2409.

Donn,R.P., Barrett,J.H., Farhan,A., Stopford,A., Pepper,L., Shelley,E., Davies,N., Ollier,W.E., and Thomson,W. (2001a). Cytokine gene polymorphisms and susceptibility to juvenile idiopathic arthritis. British Paediatric Rheumatology Study Group. Arthritis Rheum. *44*, 802-810.

Donn,R.P., Farhan,A.J., Barrett,J.H., Thomson,W., Worthington,J., and Ollier,W.E. (1999a). Absence of association between interleukin 1 alpha and oligoarticular juvenile chronic arthritis in UK patients. Rheumatology. (Oxford) *38*, 171-175.

Donn,R.P., Farhan,A.J., Barrett,J.H., Thomson,W., Worthington,J., and Ollier,W.E. (1999b). Absence of association between interleukin 1 alpha and oligoarticular juvenile chronic arthritis in UK patients. Rheumatology. (Oxford) *38*, 171-175.

Donn,R.P., Shelley,E., Ollier,W.E., and Thomson,W. (2001b). A novel 5'-flanking region polymorphism of macrophage migration inhibitory factor is associated with systemic-onset juvenile idiopathic arthritis. Arthritis Rheum. *44*, 1782-1785.

Dripps,D.J., Brandhuber,B.J., Thompson,R.C., and Eisenberg,S.P. (1991). Interleukin-1 (IL-1) receptor antagonist binds to the 80-kDa IL-1 receptor but does not initiate IL-1 signal transduction. J. Biol. Chem. *266*, 10331-10336.

Dudbridge,F. (2003). Pedigree disequilibrium tests for multilocus haplotypes. Genet. Epidemiol. *25*, 115-121.

Durbin,R.M., Abecasis,G.R., Altshuler,D.L., Auton,A., Brooks,L.D., Durbin,R.M., Gibbs,R.A., Hurles,M.E., and McVean,G.A. (2010). A map of human genome variation from population-scale sequencing. Nature *467*, 1061-1073.

Dynan,W.S. and Tjian,R. (1983). The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. Cell *35*, 79-87.

Elston,R.C. and Johnson,W.D. (2010). Some Specialized Techniques. In Basic Biostatistics for Geneticists, (Chichester: Wiley), pp. 293-317.

Evans,R.J., Bray,J., Childs,J.D., Vigers,G.P., Brandhuber,B.J., Skalicky,J.J., Thompson,R.C., and Eisenberg,S.P. (1995). Mapping receptor binding sites in interleukin (IL)-1 receptor antagonist and IL-1 beta by site-directed mutagenesis. Identification of a single site in IL-1ra and two sites in IL-1 beta. J. Biol. Chem. *270*, 11477-11483.

Fall,N., Barnes,M., Thornton,S., Luyrink,L., Olson,J., Ilowite,N.T., Gottlieb,B.S., Griffin,T., Sherry,D.D., Thompson,S., Glass,D.N., Colbert,R.A., and Grom,A.A. (2007). Gene expression profiling of peripheral blood from patients with untreated new-onset systemic juvenile idiopathic arthritis reveals molecular heterogeneity that may predict macrophage activation syndrome. Arthritis Rheum. *56*, 3793-3804.

Fang,Z., Fu,Y., Liang,Y., Li,Z., Zhang,W., Jin,J., Yang,Y., and Zha,X. (2007). Increased expression of integrin beta1 subunit enhances p21WAF1/Cip1 transcription through the Sp1 sites and p300-mediated histone acetylation in human hepatocellular carcinoma cells. J. Cell Biochem. *101*, 654-664.

Fantuzzi,G., Ku,G., Harding,M.W., Livingston,D.J., Sipe,J.D., Kuida,K., Flavell,R.A., and Dinarello,C.A. (1997). Response to local inflammation of IL-1 beta-converting enzyme-deficient mice. J. Immunol. *158*, 1818-1824.

Fantuzzi,G., Puren,A.J., Harding,M.W., Livingston,D.J., and Dinarello,C.A. (1998). Interleukin-18 regulation of interferon gamma production and cell proliferation as shown in interleukin-1beta-converting enzyme (caspase-1)-deficient mice. Blood *91*, 2118-2125.

Fantuzzi,G., Reed,D.A., and Dinarello,C.A. (1999). IL-12-induced IFN-gamma is dependent on caspase-1 processing of the IL-18 precursor. J. Clin. Invest *104*, 761-767.

Feldman,B.M., Birdi,N., Boone,J.E., Dent,P.B., Duffy,C.M., Ellsworth,J.E., Lang,B.A., Laxer,R.M., Lewkonia,R.M., Malleson,P.N., Oen,K.G., Paquin,J.D., Rosenberg,A.M., Schneider,R., and Silverman,E.D. (1996). Seasonal onset of systemic-onset juvenile rheumatoid arthritis. J. Pediatr. *129*, 513-518.

Feldmann,J., Prieur,A.M., Quartier,P., Berquin,P., Certain,S., Cortis,E., Teillac-Hamel,D., Fischer,A., and de Saint,B.G. (2002). Chronic infantile neurological cutaneous and articular syndrome is caused by mutations in CIAS1, a gene highly expressed in polymorphonuclear cells and chondrocytes. Am. J. Hum. Genet. *71*, 198-203.

Ferrante,L., Opdal,S.H., Vege,A., and Rognum,T.O. (2010). IL-1 gene cluster polymorphisms and sudden infant death syndrome. Hum. Immunol. *71*, 402-406.

Fife,M.S., Gathercole,L., Ogilvie,E.M., Stock,C.J., Mack,L.F., Donn,R.P., Thomson,W., and Woo,P. (2007). No evidence for genetic association of interferon regulatory factor 1 in juvenile idiopathic arthritis. Arthritis Rheum. *56*, 972-976.

Fife,M.S., Gutierrez,A., Ogilvie,E.M., Stock,C.J., Samuel,J.M., Thomson,W., Mack,L.F., Lewis,C.M., and Woo,P. (2006). Novel IL10 gene family associations with systemic juvenile idiopathic arthritis. Arthritis Res. Ther. *8*, R148.

Fife,M.S., Ogilvie,E.M., Kelberman,D., Samuel,J., Gutierrez,A., Humphries,S.E., and Woo,P. (2005). Novel IL-6 haplotypes and disease association. Genes Immun. *6*, 367-370.

Fishman,D., Faulds,G., Jeffery,R., Mohamed-Ali,V., Yudkin,J.S., Humphries,S., and Woo,P. (1998). The effect of novel polymorphisms in the interleukin-6 (IL-6) gene on IL-6 transcription and plasma IL-6 levels, and an association with systemic-onset juvenile chronic arthritis. J. Clin. Invest *102*, 1369-1376.

Fitzgerald,A.A., Leclercq,S.A., Yan,A., Homik,J.E., and Dinarello,C.A. (2005). Rapid responses to anakinra in patients with refractory adult-onset Still's disease. Arthritis Rheum. *52*, 1794-1803.

Fontanella,M., Rainero,I., Gallone,S., Rubino,E., Fornaro,R., Fenoglio,P., Valfre,W., Vaula,G., Benevello,C., Ducati,A., and Pinessi,L. (2010). Interleukin-1 cluster gene polymorphisms and aneurysmal subarachnoid hemorrhage. Neurosurgery *66*, 1058-1062.

Forsyth,K.D. and Levinsky,R.J. (1990). Fibronectin degradation; an in-vitro model of neutrophil mediated endothelial cell damage. J. Pathol. *161*, 313-319.

G.McVean (2007). Linkage Disequilibrium, Recombination and Selection. D.J.Balding, C.E.Bishop, and C.Cannings, eds. (Chichester: Wiley), pp. 909-940.

Galy,A.H., Dinarello,C.A., Kupper,T.S., Kameda,A., and Hadden,J.W. (1990). Effects of cytokines on human thymic epithelial cells in culture. II. Recombinant IL 1 stimulates thymic epithelial cells to produce IL6 and GM-CSF. Cell Immunol. *129*, 161-175.

Garred,P., Madsen,H.O., Petersen,J., Marquart,H., Hansen,T.M., Freiesleben,S.S., Volck,B., Svejgaard,A., and Andersen,V. (1998). CC chemokine receptor 5 polymorphism in rheumatoid arthritis. J. Rheumatol. *25*, 1462-1465.

Gattorno,M., Piccini,A., Lasiglie,D., Tassi,S., Brisca,G., Carta,S., Delfino,L., Ferlito,F., Pelagatti,M.A., Caroli,F., Buoncompagni,A., Viola,S., Loy,A., Sironi,M., Vecchi,A., Ravelli,A., Martini,A., and Rubartelli,A. (2008). The pattern of response to anti-interleukin-1 treatment distinguishes two subsets of patients with systemic-onset juvenile idiopathic arthritis. Arthritis and Rheumatism *58*, 1505-1515.

Ghayur,T., Banerjee,S., Hugunin,M., Butler,D., Herzog,L., Carter,A., Quintal,L., Sekut,L., Talanian,R., Paskind,M., Wong,W., Kamen,R., Tracey,D., and Allen,H. (1997). Caspase-1 processes IFN-gamma-inducing factor and regulates LPS-induced IFN-gamma production. Nature *386*, 619-623.

Giedraitis,V., He,B., Huang,W.X., and Hillert,J. (2001). Cloning and mutation analysis of the human IL-18 promoter: a possible role of polymorphisms in expression regulation. J. Neuroimmunol. *112*, 146-152.

Glass,D.N. and Giannini,E.H. (1999). Juvenile rheumatoid arthritis as a complex genetic trait. Arthritis Rheum. *42*, 2261-2268.

Goldenberg,J., Ferraz,M.B., Pessoa,A.P., Fonseca,A.S., Carvalho,A.C., Hilario,M.O., and Atra,E. (1992b). Symptomatic cardiac involvement in juvenile rheumatoid arthritis. Int. J. Cardiol. *34*, 57-62.

Goldenberg,J., Ferraz,M.B., Pessoa,A.P., Fonseca,A.S., Carvalho,A.C., Hilario,M.O., and Atra,E. (1992a). Symptomatic cardiac involvement in juvenile rheumatoid arthritis. Int. J. Cardiol. *34*, 57-62.

Gomez-Reino,J.J., Pablos,J.L., Carreira,P.E., Santiago,B., Serrano,L., Vicario,J.L., Balsa,A., Figueroa,M., and de Juan,M.D. (1999). Association of rheumatoid arthritis with a functional chemokine receptor, CCR5. Arthritis Rheum. *42*, 989-992.

Grabe,N. (2002). AliBaba2: context specific identification of transcription factor binding sites. In Silico. Biol. *2*, S1-15.

Gracie,J.A., Forsey,R.J., Chan,W.L., Gilmour,A., Leung,B.P., Greer,M.R., Kennedy,K., Carter,R., Wei,X.Q., Xu,D., Field,M., Foulis,A., Liew,F.Y., and McInnes,I.B. (1999a). A proinflammatory role for IL-18 in rheumatoid arthritis. J. Clin. Invest *104*, 1393-1401.

Gracie,J.A., Forsey,R.J., Chan,W.L., Gilmour,A., Leung,B.P., Greer,M.R., Kennedy,K., Carter,R., Wei,X.Q., Xu,D., Field,M., Foulis,A., Liew,F.Y., and McInnes,I.B. (1999b). A proinflammatory role for IL-18 in rheumatoid arthritis. J. Clin. Invest *104*, 1393-1401.

Gracie,J.A., Koyama,N., Murdoch,J., Field,M., McGarry,F., Crilly,A., Schobel,A., Madhok,R., Pons-Kuhnemann,J., McInnes,I.B., and Moller,B. (2005). Disease association of two distinct interleukin-18 promoter polymorphisms in Caucasian rheumatoid arthritis patients. Genes Immun. *6*, 211-216.

Granowitz,E.V., Clark,B.D., Mancilla,J., and Dinarello,C.A. (1991). Interleukin-1 receptor antagonist competitively inhibits the binding of interleukin-1 to the type II interleukin-1 receptor. J. Biol. Chem. *266*, 14147-14150.

Graudal,N.A., Svenson,M., Tarp,U., Garred,P., Jurik,A.G., and Bendtzen,K. (2002). Autoantibodies against interleukin 1alpha in rheumatoid arthritis: association with long term radiographic outcome. Ann. Rheum. Dis. *61*, 598-602.

Greenfeder,S.A., Nunes,P., Kwee,L., Labow,M., Chizzonite,R.A., and Ju,G. (1995). Molecular cloning and characterization of a second subunit of the interleukin 1 receptor complex. J. Biol. Chem. *270*, 13757-13765.

Gu,Y., Kuida,K., Tsutsui,H., Ku,G., Hsiao,K., Fleming,M.A., Hayashi,N., Higashino,K., Okamura,H., Nakanishi,K., Kurimoto,M., Tanimoto,T., Flavell,R.A., Sato,V., Harding,M.W., Livingston,D.J., and Su,M.S. (1997). Activation of interferon-gamma inducing factor mediated by interleukin-1beta converting enzyme. Science *275*, 206-209.

Gu,Y., Wu,J., Faucheu,C., Lalanne,J.L., Diu,A., Livingston,D.J., and Su,M.S. (1995). Interleukin-1 beta converting enzyme requires oligomerization for activity of processed forms in vivo. EMBO J. *14*, 1923-1931.

Guillaumot,P., Luquain,C., Malek,M., Huber,A.L., Brugiere,S., Garin,J., Grunwald,D., Regnier,D., Petrilli,V., Lefai,E., and Manie,S.N. (2010). Pdro, a protein associated with late endosomes and lysosomes and implicated in cellular cholesterol homeostasis. PLoS. One. *5*, e10977.

Gunderson,K.L., Kruglyak,S., Graige,M.S., Garcia,F., Kermani,B.G., Zhao,C., Che,D., Dickinson,T., Wickham,E., Bierle,J., Doucet,D., Milewski,M., Yang,R., Siegmund,C., Haas,J., Zhou,L., Oliphant,A., Fan,J.B., Barnard,S., and Chee,M.S. (2004). Decoding randomly ordered DNA arrays. Genome Res. *14*, 870-877.

Guo,Z., Zhang,M., An,H., Chen,W., Liu,S., Guo,J., Yu,Y., and Cao,X. (2003). Fas ligation induces IL-1beta-dependent maturation and IL-1beta-independent survival of dendritic cells: different roles of ERK and NF-kappaB signaling pathways. Blood *102*, 4441-4447.

Guo,Z.S., Li,C., Lin,Z.M., Huang,J.X., Wei,Q.J., Wang,X.W., Xie,Y.Y., Liao,Z.T., Chao,S.Y., and Gu,J.R. (2010). Association of IL-1 gene complex members with ankylosing spondylitis in Chinese Han population. Int. J. Immunogenet. *37*, 33-37.

Hafner,R. and Truckenbrodt,H. (1986). [Course and prognosis of systemic juvenile chronic arthritis--retrospective study of 187 patients]. Klin. Padiatr. *198*, 401-407.

Halle,F. and Prieur,A.M. (1991). Evaluation of methotrexate in the treatment of juvenile chronic arthritis according to the subtype. Clin. Exp. Rheumatol. *9*, 297-302.

Hamer,G., Gell,K., Kouznetsova,A., Novak,I., Benavente,R., and Hoog,C. (2006). Characterization of a novel meiosis-specific protein within the central element of the synaptonemal complex. J. Cell Sci. *119*, 4025-4032.

Han,W., Kang,S.Y., Kang,D., Park,S.K., Lee,J.Y., Kim,H., Park,A.K., and Noh,D.Y. (2010). Multiplex genotyping of 1107 SNPs from 232 candidate genes identified an association between IL1A polymorphism and breast cancer risk. Oncol. Rep. *23*, 763-769.

Hanazawa,S., Amano,S., Hanaizumi,C., Hirose,K., Ohmori,Y., Miyoshi,T., and Kitano,S. (1988). Inductive effect of human recombinant IL-1 on differentiation of a macrophage-like tumor cell line. Adv. Dent. Res. *2*, 372-375.

Hannum,C.H., Wilcox,C.J., Arend,W.P., Joslin,F.G., Dripps,D.J., Heimdal,P.L., Armes,L.G., Sommer,A., Eisenberg,S.P., and Thompson,R.C. (1990). Interleukin-1 receptor antagonist activity of a human interleukin-1 inhibitor. Nature *343*, 336-340.

Hardison,R.C., Oeltjen,J., and Miller,W. (1997). Long human-mouse sequence alignments reveal novel regulatory elements: a reason to sequence the mouse genome. Genome Res. *7*, 959-966.

Harrison,P., Pointon,J.J., Chapman,K., Roddam,A., and Wordsworth,B.P. (2008). Interleukin-1 promoter region polymorphism role in rheumatoid arthritis: a meta-analysis of IL-1B-511A/G variant reveals association with rheumatoid arthritis. Rheumatology (Oxford) *47*, 1768-1770.

Haskill,S., Martin,G., Van Le,L., Morris,J., Peace,A., Bigler,C.F., Jaffe,G.J., Hammerberg,C., Sporn,S.A., Fong,S., and . (1991). cDNA cloning of an intracellular form of the human interleukin 1 receptor antagonist associated with epithelium. Proc. Natl. Acad. Sci. U. S. A *88*, 3681-3685.

Hauser,C., Saurat,J.H., Schmitt,A., Jaunin,F., and Dayer,J.M. (1986). Interleukin 1 is present in normal human epidermis. J. Immunol. *136*, 3317-3323.

Hawkins,P.N., Lachmann,H.J., Aganna,E., and McDermott,M.F. (2004). Spectrum of clinical features in Muckle-Wells syndrome and response to anakinra. Arthritis Rheum. *50*, 607-612.

Henrickson, M. Efficacy of Anakinra in Refactory Systemic Arthritis. American College of Rheumatology . 2004.
Ref Type: Abstract

Herman,J., Dinarello,C.A., Kew,M.C., and Rabson,A.R. (1985). The role of interleukin 1 (IL 1) in tumor-NK cell interactions: correction of defective NK cell activity in cancer patients by treating target cells with IL 1. J. Immunol. *135*, 2882-2886.

Hinks,A., Barton,A., John,S., Bruce,I., Hawkins,C., Griffiths,C.E., Donn,R., Thomson,W., Silman,A., and Worthington,J. (2005). Association between the PTPN22 gene and rheumatoid arthritis and juvenile idiopathic arthritis in a UK population: further support that PTPN22 is an autoimmunity gene. Arthritis Rheum. *52*, 1694-1699.

Hinks,A., Martin,P., Thompson,S., Sudman,M., Glass,D., Stock,C., Day,T., Thomson,W., Packham,J., Childhood Arthritis Prospective Study (CAPS), BSPAR study group, Ramanan,A., and Donn,R. (2011). Autoinflammatory gene polymorphisms and susceptibility to UK juvenile idiopathic arthritis. Submitted.

Hirawake,H., Taniwaki,M., Tamura,A., Kojima,S., and Kita,K. (1997). Cytochrome b in human complex II (succinate-ubiquinone oxidoreductase): cDNA cloning of the components in liver mitochondria and chromosome assignment of the genes for the large (SDHC) and small (SDHD) subunits to 1q21 and 11q23. Cytogenet. Cell Genet. *79*, 132-138.

Hoffman,H.M., Mueller,J.L., Broide,D.H., Wanderer,A.A., and Kolodner,R.D. (2001). Mutation of a new gene encoding a putative pyrin-like protein causes familial cold autoinflammatory syndrome and Muckle-Wells syndrome. Nat. Genet. *29*, 301-305.

Hoffman,H.M., Rosengren,S., Boyle,D.L., Cho,J.Y., Nayar,J., Mueller,J.L., Anderson,J.P., Wanderer,A.A., and Firestein,G.S. (2004). Prevention of cold-associated acute inflammation in familial cold autoinflammatory syndrome by interleukin-1 receptor antagonist. Lancet *364*, 1779-1785.

Hofmeister,R., Wiegmann,K., Korherr,C., Bernardo,K., Kronke,M., and Falk,W. (1997). Activation of acid sphingomyelinase by interleukin-1 (IL-1) requires the IL-1 receptor accessory protein. J. Biol. Chem. *272*, 27730-27736.

Hong,J.S., Kim,S.W., and Koo,J.S. (2008). Sp1 up-regulates cAMP-response-element-binding protein expression during retinoic acid-induced mucous differentiation of normal human bronchial epithelial cells. Biochem. J. *410*, 49-61.

Horneff,G., Schmeling,H., Biedermann,T., Foeldvari,I., Ganser,G., Girschick,H.J., Hospach,T., Huppertz,H.I., Keitzer,R., Kuster,R.M., Michels,H., Moebius,D., Rogalski,B., and Thon,A. (2004). The German etanercept registry for treatment of juvenile idiopathic arthritis. Ann. Rheum. Dis. *63*, 1638-1644.

Hosking,L., Lumsden,S., Lewis,K., Yeo,A., McCarthy,L., Bansal,A., Riley,J., Purvis,I., and Xu,C.F. (2004). Detection of genotyping errors by Hardy-Weinberg equilibrium testing. Eur. J. Hum. Genet. *12*, 395-399.

Huang,J., Gao,X., Li,S., and Cao,Z. (1997). Recruitment of IRAK to the interleukin 1 receptor complex requires interleukin 1 receptor accessory protein. Proc. Natl. Acad. Sci. U. S. A *94*, 12829-12832.

Humke,E.W., Shriver,S.K., Starovasnik,M.A., Fairbrother,W.J., and Dixit,V.M. (2000). ICEBERG: a novel inhibitor of interleukin-1beta generation. Cell *103*, 99-111.

Hunt,K.A., Zhernakova,A., Turner,G., Heap,G.A., Franke,L., Bruinenberg,M., Romanos,J., Dinesen,L.C., Ryan,A.W., Panesar,D., Gwilliam,R., Takeuchi,F., McLaren,W.M., Holmes,G.K., Howdle,P.D., Walters,J.R., Sanders,D.S., Playford,R.J., Trynka,G., Mulder,C.J., Mearin,M.L., Verbeek,W.H., Trimble,V., Stevens,F.M., O'Morain,C., Kennedy,N.P., Kelleher,D., Pennington,D.J., Strachan,D.P., McArdle,W.L., Mein,C.A., Wapenaar,M.C., Deloukas,P., McGinnis,R., McManus,R., Wijmenga,C., and van Heel,D.A. (2008). Newly identified genetic risk variants for celiac disease related to the immune response. Nat. Genet. *40*, 395-402.

Hurgin,V., Novick,D., and Rubinstein,M. (2002). The promoter of IL-18 binding protein: activation by an IFN-gamma -induced complex of IFN regulatory factor 1 and CCAAT/enhancer binding protein beta. Proc. Natl. Acad. Sci. U. S. A *99*, 16957-16962.

Hurtado,B., Abasolo,N., Munoz,X., Garcia,N., Benavente,Y., Rubio,F., Garcia,d.F., Krupinski,J., and Sala,N. (2010). Association study between polymorphims in GAS6-TAM genes and carotid atherosclerosis. Thromb. Haemost. *104*.

Ide,A., Kawasaki,E., Abiru,N., Sun,F., Kobayashi,M., Fukushima,T., Takahashi,R., Kuwahara,H., Kita,A., Oshima,K., Uotani,S., Yamasaki,H., Yamaguchi,Y., and Eguchi,K. (2004). Association between IL-18 gene promoter polymorphisms and CTLA-4 gene 49A/G polymorphism in Japanese patients with type 1 diabetes. J. Autoimmun. *22*, 73-78.

Igarashi,K., Mitsuyama,M., Muramori,K., Tsukada,H., and Nomoto,K. (1990). Interleukin-1-induced promotion of T-cell differentiation in mice immunized with killed Listeria monocytogenes. Infect. Immun. *58*, 3973-3979.

Irigoyen, P. I., Olson, J., Hom, C., and Ilowite, N. T. Treatment of Systemic Onset Juvenile Rheumatoid Arthritis with Anakinra. American College of Rheumatology . 2004.
Ref Type: Abstract

Isaev,P.I., Hauselman,H., Tonar,E., Kuttner,K., and Zborovskii,A.B. (1992). [Metabolism of articular cartilage in the presence of interleukin-I alpha, its inhibitor and blood serum]. Biull. Eksp. Biol. Med. *114*, 269-271.

Ishikawa,S., Mima,T., Aoki,C., Yoshio-Hoshino,N., Adachi,Y., Imagawa,T., Mori,M., Tomiita,M., Iwata,N., Murata,T., Miyoshi,M., Takei,S., Aihara,Y., Yokota,S., Matsubara,K., and Nishimoto,N. (2009). Abnormal expression of the genes involved in cytokine networks and mitochondrial function in systemic juvenile idiopathic arthritis identified by DNA microarray analysis. Ann. Rheum. Dis. *68*, 264-272.

Jelusic,M., Lukic,I.K., Tambic-Bukovac,L., Dubravcic,K., Malcic,I., Rudan,I., and Batinic,D. (2007). Interleukin-18 as a mediator of systemic juvenile idiopathic arthritis. Clin. Rheumatol. *26*, 1332-1334.

Jensen,L.E., Muzio,M., Mantovani,A., and Whitehead,A.S. (2000a). IL-1 signaling cascade in liver cells and the involvement of a soluble form of the IL-1 receptor accessory protein. J. Immunol. *164*, 5277-5286.

Jensen,L.E., Muzio,M., Mantovani,A., and Whitehead,A.S. (2000b). IL-1 signaling cascade in liver cells and the involvement of a soluble form of the IL-1 receptor accessory protein. J. Immunol. *164*, 5277-5286.

Jensen,L.E. and Whitehead,A.S. (2003b). Expression of alternatively spliced interleukin-1 receptor accessory protein mRNAs is differentially regulated during inflammation and apoptosis. Cell Signal. *15*, 793-802.

Jensen,L.E. and Whitehead,A.S. (2003a). Expression of alternatively spliced interleukin-1 receptor accessory protein mRNAs is differentially regulated during inflammation and apoptosis. Cell Signal. *15*, 793-802.

Jeru,I., Marlin,S., Le Borgne,G., Cochet,E., Normand,S., Duquesnoy,P., Dastot-Le Moal,F., Cuisset,L., Hentgen,V., Fernandes,A.T., Lecron,J.C., Dhote,R., Grateau,G., Alnemri,E.S., and Amselem,S. (2010). Functional consequences of a germline mutation in the leucine-rich repeat domain of NLRP3 identified in an atypical autoinflammatory disorder. Arthritis Rheum. *62*, 1176-1185.

Johnson,A.D., Handsaker,R.E., Pulit,S.L., Nizzari,M.M., O'Donnell,C.J., and de Bakker,P.I. (2008). SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics. *24*, 2938-2939.

Jonuleit,H., Knop,J., and Enk,A.H. (1996). Cytokines and their effects on maturation, differentiation and migration of dendritic cells. Arch. Dermatol. Res. *289*, 1-8.

Joosten,L.A., Netea,M.G., Fantuzzi,G., Koenders,M.I., Helsen,M.M., Sparrer,H., Pham,C.T., van der Meer,J.W., Dinarello,C.A., and van den Berg,W.B. (2009). Inflammatory arthritis in caspase 1 gene-deficient mice: contribution of proteinase 3 to caspase 1-independent production of bioactive interleukin-1beta. Arthritis Rheum. *60*, 3651-3662.

Josephs,M.D., Solorzano,C.C., Taylor,M., Rosenberg,J.J., Topping,D., Abouhamze,A., Mackay,S.L., Hirsch,E., Hirsh,D., Labow,M., and Moldawer,L.L. (2000). Modulation of the acute phase response by altered expression of the IL-1 type 1 receptor or IL-1ra. Am. J. Physiol Regul. Integr. Comp Physiol *278*, R824-R830.

Jouvenne,P., Chaudhary,A., Buchs,N., Giovine,F.S., Duff,G.W., and Miossec,P. (1999). Possible genetic association between interleukin-1alpha gene polymorphism and the severity of chronic polyarthritis. Eur. Cytokine Netw. *10*, 33-36.

Jouvenne,P., Vannier,E., Dinarello,C.A., and Miossec,P. (1998b). Elevated levels of soluble interleukin-1 receptor type II and interleukin-1 receptor antagonist in patients with chronic arthritis: correlations with markers of inflammation and joint destruction. Arthritis Rheum. *41*, 1083-1089.

Jouvenne,P., Vannier,E., Dinarello,C.A., and Miossec,P. (1998a). Elevated levels of soluble interleukin-1 receptor type II and interleukin-1 receptor antagonist in patients with chronic arthritis: correlations with markers of inflammation and joint destruction. Arthritis Rheum. *41*, 1083-1089.

Jung,M.Y., Kang,S.W., Kim,S.K., Kim,H.J., Yun,D.H., Yim,S.V., Hong,S.J., and Chung,J.H. (2010). The interleukin-1 family gene polymorphisms in Korean patients with rheumatoid arthritis. Scand. J. Rheumatol. *39*, 190-196.

Kadota,C., Ishihara,S., Aziz,M.M., Rumi,M.A., Oshima,N., Mishima,Y., Moriyama,I., Yuki,T., Amano,Y., and Kinoshita,Y. (2010). Down-regulation of single immunoglobulin interleukin-1R-related molecule (SIGIRR)/TIR8 expression in intestinal epithelial cells during inflammation. Clin. Exp. Immunol. *162*, 348-361.

Kaipiainen-Seppanen,O., Punnonen,K., van Gijn,M.E., and Mononen,T. (2008b). Two pathogenic CIAS1 mutations and plasma cytokine profile in a Finnish patient with familial cold autoinflammatory syndrome responsive to anakinra. Scand. J. Rheumatol. *37*, 75-76.

Kaipiainen-Seppanen,O., Punnonen,K., van Gijn,M.E., and Mononen,T. (2008a). Two pathogenic CIAS1 mutations and plasma cytokine profile in a Finnish patient with familial cold autoinflammatory syndrome responsive to anakinra. Scand. J. Rheumatol. *37*, 75-76.

Kaiserlian,D., Rigal,D., Abello,J., and Revillard,J.P. (1991). Expression, function and regulation of the intercellular adhesion molecule-1 (ICAM-1) on human intestinal epithelial cell lines. Eur. J. Immunol. *21*, 2415-2421.

Kammerer,S., Roth,R.B., Hoyal,C.R., Reneland,R., Marnellos,G., Kiechle,M., Schwarz-Boeger,U., Griffiths,L.R., Ebner,F., Rehbock,J., Cantor,C.R., Nelson,M.R., and Braun,A. (2005). Association of the NuMA region on chromosome 11q13 with breast cancer susceptibility. Proc. Natl. Acad. Sci. U. S. A *102*, 2004-2009.

Kanneganti,T.D., Body-Malapel,M., Amer,A., Park,J.H., Whitfield,J., Franchi,L., Taraporewala,Z.F., Miller,D., Patton,J.T., Inohara,N., and Nunez,G. (2006a). Critical role for

Cryopyrin/Nalp3 in activation of caspase-1 in response to viral infection and double-stranded RNA. J. Biol. Chem. *281*, 36560-36568.

Kanneganti,T.D., Ozoren,N., Body-Malapel,M., Amer,A., Park,J.H., Franchi,L., Whitfield,J., Barchet,W., Colonna,M., Vandenabeele,P., Bertin,J., Coyle,A., Grant,E.P., Akira,S., and Nunez,G. (2006b). Bacterial RNA and small antiviral compounds activate caspase-1 through cryopyrin/Nalp3. Nature *440*, 233-236.

Kastbom,A., Verma,D., Eriksson,P., Skogh,T., Wingren,G., and Soderkvist,P. (2008). Genetic variation in proteins of the cryopyrin inflammasome influences susceptibility and severity of rheumatoid arthritis (the Swedish TIRA project). Rheumatology (Oxford) *47*, 415-417.

Kawaguchi,Y., Terajima,H., Harigai,M., Hara,M., and Kamatani,N. (2001). Interleukin-18 as a novel diagnostic marker and indicator of disease severity in adult-onset Still's disease. Arthritis Rheum. *44*, 1716-1717.

Kawaguchi,Y., Tochimoto,A., Hara,M., Kawamoto,M., Sugiura,T., Saito,S., and Kamatani,N. (2007). Contribution of single nucleotide polymorphisms of the IL1A gene to the cleavage of precursor IL-1alpha and its transcription activity. Immunogenetics *59*, 441-448.

Kawashima,M., Yamamura,M., Taniai,M., Yamauchi,H., Tanimoto,T., Kurimoto,M., Miyawaki,S., Amano,T., Takeuchi,T., and Makino,H. (2001). Levels of interleukin-18 and its binding inhibitors in the blood circulation of patients with adult-onset Still's disease. Arthritis Rheum. *44*, 550-560.

Kent,W.J. (2002). BLAT--the BLAST-like alignment tool. Genome Res. *12*, 656-664.

Keul,R., Heinrich,P.C., Muller-newen,G., Muller,K., and Woo,P. (1998). A possible role for soluble IL-6 receptor in the pathogenesis of systemic onset juvenile chronic arthritis. Cytokine *10*, 729-734.

Khalilzadeh,O., Anvari,M., Esteghamati,A., Momen-Heravi,F., Mahmoudi,M., Rashidi,A., Amiri,H.M., Ranjbar,M., Tabataba-Vakili,S., and Amirzargar,A. (2010). The interleukin-1 family gene polymorphisms and Graves' disease. Ann. Endocrinol. (Paris) *71*, 281-285.

Kim,S.H., Eisenstein,M., Reznikov,L., Fantuzzi,G., Novick,D., Rubinstein,M., and Dinarello,C.A. (2000). Structural requirements of six naturally occurring isoforms of the IL-18 binding protein to inhibit IL-18. Proc. Natl. Acad. Sci. U. S. A *97*, 1190-1195.

Kimura,Y., Fieldston,E., Devries-Vandervlugt,B., Li,S., and Imundo,L. (2000). High dose, alternate day corticosteroids for systemic onset juvenile rheumatoid arthritis. J. Rheumatol. *27*, 2018-2024.

Kimura,Y., Pinho,P., Walco,G., Higgins,G., Hummell,D., Szer,I., Henrickson,M., Watcher,S., and Reiff,A. (2005). Etanercept treatment in patients with refractory systemic onset juvenile rheumatoid arthritis. J. Rheumatol. *32*, 935-942.

Kinane,D.F. and Hart,T.C. (2003). Genes and gene polymorphisms associated with periodontal disease. Crit Rev. Oral Biol. Med. *14*, 430-449.

Knight,C.J. (2005). HaploChIP: an in vivo assay. Methods Mol. Biol. *311*, 49-60.

Kobayashi,Y., Yamamoto,K., Saido,T., Kawasaki,H., Oppenheim,J.J., and Matsushima,K. (1990). Identification of calcium-activated neutral protease as a processing enzyme of human interleukin 1 alpha. Proc. Natl. Acad. Sci. U. S. A *87*, 5548-5552.

Kohase,M., May,L.T., Tamm,I., Vilcek,J., and Sehgal,P.B. (1987). A cytokine network in human diploid fibroblasts: interactions of beta-interferons, tumor necrosis factor, platelet-derived growth factor, and interleukin-1. Mol. Cell Biol. *7*, 273-280.

Kopp,E.B. and Ghosh,S. (1995). NF-kappa B and rel proteins in innate immunity. Adv. Immunol. *58*, 1-27.

Korherr,C., Hofmeister,R., Wesche,H., and Falk,W. (1997). A critical role for interleukin-1 receptor accessory protein in interleukin-1 signaling. Eur. J. Immunol. *27*, 262-267.

Koskinen,L.L., Einarsdottir,E., Dukes,E., Heap,G.A., Dubois,P., Korponay-Szabo,I.R., Kaukinen,K., Kurppa,K., Ziberna,F., Vatta,S., Not,T., Ventura,A., Sistonen,P., Adany,R., Pocsai,Z., Szeles,G., Maki,M., Kere,J., Wijmenga,C., van Heel,D.A., and Saavalainen,P. (2009). Association study of the IL18RAP locus in three European populations with coeliac disease. Hum. Mol. Genet. *18*, 1148-1155.

Kretowski,A., Mironczuk,K., Karpinska,A., Bojaryn,U., Kinalski,M., Puchalski,Z., and Kinalska,I. (2002). Interleukin-18 promoter polymorphisms in type 1 diabetes. Diabetes *51*, 3347-3349.

Ku,G., Faust,T., Lauffer,L.L., Livingston,D.J., and Harding,M.W. (1996). Interleukin-1 beta converting enzyme inhibition blocks progression of type II collagen-induced arthritis in mice. Cytokine *8*, 377-386.

Kuida,K., Lippke,J.A., Ku,G., Harding,M.W., Livingston,D.J., Su,M.S., and Flavell,R.A. (1995). Altered cytokine export and apoptosis in mice deficient in interleukin-1 beta converting enzyme. Science *267*, 2000-2003.

Kumar,S., Hanning,C.R., Brigham-Burke,M.R., Rieman,D.J., Lehr,R., Khandekar,S., Kirkpatrick,R.B., Scott,G.F., Lee,J.C., Lynch,F.J., Gao,W., Gambotto,A., and Lotze,M.T. (2002). Interleukin-1F7B (IL-1H4/IL-1F7) is processed by caspase-1 and mature IL-1F7B binds to the IL-18 receptor but does not induce IFN-gamma production. Cytokine *18*, 61-71.

Kumar,S., McDonnell,P.C., Lehr,R., Tierney,L., Tzimas,M.N., Griswold,D.E., Capper,E.A., Tal-Singer,R., Wells,G.I., Doyle,M.L., and Young,P.R. (2000). Identification and initial characterization of four novel members of the interleukin-1 family. J. Biol. Chem. *275*, 10308-10314.

Kurt-Jones,E.A., Beller,D.I., Mizel,S.B., and Unanue,E.R. (1985). Identification of a membrane-associated interleukin 1 in macrophages. Proc. Natl. Acad. Sci. U. S. A *82*, 1204-1208.

Lamb,R., Thomson,W., Ogilvie,E.M., and Donn,R. (2007). Positive association of SLC26A2 gene polymorphisms with susceptibility to systemic-onset juvenile idiopathic arthritis. Arthritis Rheum. *56*, 1286-1291.

Lander,E.S., Linton,L.M., Birren,B., Nusbaum,C., Zody,M.C., Baldwin,J., Devon,K., Dewar,K., Doyle,M., FitzHugh,W., Funke,R., Gage,D., Harris,K., Heaford,A., Howland,J., Kann,L., Lehoczky,J., LeVine,R., McEwan,P., McKernan,K., Meldrim,J., Mesirov,J.P., Miranda,C., Morris,W., Naylor,J., Raymond,C., Rosetti,M., Santos,R., Sheridan,A., Sougnez,C., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Wyman,D., Rogers,J., Sulston,J., Ainscough,R., Beck,S., Bentley,D., Burton,J., Clee,C., Carter,N., Coulson,A., Deadman,R., Deloukas,P., Dunham,A., Dunham,I., Durbin,R., French,L., Grafham,D., Gregory,S., Hubbard,T., Humphray,S., Hunt,A., Jones,M., Lloyd,C., McMurray,A., Matthews,L., Mercer,S., Milne,S., Mullikin,J.C., Mungall,A., Plumb,R., Ross,M., Shownkeen,R., Sims,S., Waterston,R.H., Wilson,R.K., Hillier,L.W., McPherson,J.D., Marra,M.A., Mardis,E.R., Fulton,L.A., Chinwalla,A.T., Pepin,K.H., Gish,W.R., Chissoe,S.L., Wendl,M.C., Delehaunty,K.D., Miner,T.L., Delehaunty,A., Kramer,J.B., Cook,L.L., Fulton,R.S., Johnson,D.L., Minx,P.J., Clifton,S.W., Hawkins,T., Branscomb,E., Predki,P., Richardson,P., Wenning,S., Slezak,T., Doggett,N., Cheng,J.F., Olsen,A., Lucas,S., Elkin,C., Uberbacher,E., Frazier,M., Gibbs,R.A., Muzny,D.M., Scherer,S.E., Bouck,J.B., Sodergren,E.J., Worley,K.C., Rives,C.M., Gorrell,J.H., Metzker,M.L., Naylor,S.L., Kucherlapati,R.S., Nelson,D.L., Weinstock,G.M., Sakaki,Y., Fujiyama,A., Hattori,M., Yada,T., Toyoda,A., Itoh,T., Kawagoe,C., Watanabe,H., Totoki,Y., Taylor,T., Weissenbach,J., Heilig,R., Saurin,W., Artiguenave,F., Brottier,P., Bruls,T., Pelletier,E., Robert,C., Wincker,P., Smith,D.R., Doucette-Stamm,L., Rubenfield,M., Weinstock,K., Lee,H.M., Dubois,J., Rosenthal,A., Platzer,M., Nyakatura,G., Taudien,S., Rump,A., Yang,H., Yu,J., Wang,J., Huang,G., Gu,J., Hood,L., Rowen,L., Madan,A., Qin,S., Davis,R.W., Federspiel,N.A., Abola,A.P., Proctor,M.J., Myers,R.M., Schmutz,J., Dickson,M., Grimwood,J., Cox,D.R., Olson,M.V., Kaul,R., Raymond,C., Shimizu,N., Kawasaki,K., Minoshima,S., Evans,G.A., Athanasiou,M., Schultz,R., Roe,B.A., Chen,F., Pan,H., Ramser,J., Lehrach,H., Reinhardt,R., McCombie,W.R., de la,B.M., Dedhia,N., Blocker,H., Hornischer,K., Nordsiek,G., Agarwala,R., Aravind,L., Bailey,J.A., Bateman,A., Batzoglou,S., Birney,E., Bork,P., Brown,D.G., Burge,C.B., Cerutti,L., Chen,H.C., Church,D., Clamp,M., Copley,R.R., Doerks,T., Eddy,S.R., Eichler,E.E., Furey,T.S., Galagan,J., Gilbert,J.G., Harmon,C., Hayashizaki,Y., Haussler,D., Hermjakob,H., Hokamp,K., Jang,W., Johnson,L.S., Jones,T.A., Kasif,S., Kaspryzk,A., Kennedy,S., Kent,W.J., Kitts,P., Koonin,E.V., Korf,I., Kulp,D., Lancet,D., Lowe,T.M., McLysaght,A., Mikkelsen,T., Moran,J.V., Mulder,N., Pollara,V.J., Ponting,C.P., Schuler,G., Schultz,J., Slater,G., Smit,A.F., Stupka,E., Szustakowski,J., Thierry-Mieg,D., Thierry-Mieg,J., Wagner,L., Wallis,J., Wheeler,R., Williams,A., Wolf,Y.I., Wolfe,K.H., Yang,S.P., Yeh,R.F., Collins,F., Guyer,M.S., Peterson,J., Felsenfeld,A., Wetterstrand,K.A., Patrinos,A., Morgan,M.J., de Jong,P., Catanese,J.J., Osoegawa,K., Shizuya,H., Choi,S., and Chen,Y.J. (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Laval,S.H., Timms,A., Edwards,S., Bradbury,L., Brophy,S., Milicic,A., Rubin,L., Siminovitch,K.A., Weeks,D.E., Calin,A., Wordsworth,B.P., and Brown,M.A. (2001). Whole-genome screening in ankylosing spondylitis: evidence of non-MHC genetic-susceptibility loci. Am. J. Hum. Genet. *68*, 918-926.

Laxer,R.M. and Schneider,R. (2004). Systemic-onset Juvenile Idiopathic Arthritis. In Oxford Textbook of Rheumatology, D.A.Isenberg, P.J.Maddison, P.Woo, D.Glass, and F.C.Breeveld, eds. (Oxford: Oxford University Press), pp. 798-809.

Lee,S.H., Stehlik,C., and Reed,J.C. (2001). Cop, a caspase recruitment domain-containing protein and inhibitor of caspase-1 activation processing. J. Biol. Chem. *276*, 34495-34500.

Lee,W.C. (2003). Searching for disease-susceptibility loci by testing for Hardy-Weinberg disequilibrium in a gene bank of affected individuals. Am. J. Epidemiol. *158*, 397-400.

Lee,Y.H., Ji,J.D., and Song,G.G. (2009). Association between interleukin 1 polymorphisms and rheumatoid arthritis susceptibility: a metaanalysis. J. Rheumatol. *36*, 12-15.

Lee,Y.H., Kim,H.J., Rho,Y.H., Choi,S.J., Ji,J.D., and Song,G.G. (2004). Interleukin-1 receptor antagonist gene polymorphism and rheumatoid arthritis. Rheumatol. Int. *24*, 133-136.

Lee,Y.M., Fujikado,N., Manaka,H., Yasuda,H., and Iwakura,Y. (2010). IL-1 plays an important role in the bone metabolism under physiological conditions. Int. Immunol. *22*, 805-816.

Lequerre,T., Quartier,P., Rosellini,D., Alaoui,F., De Bandt,M., Mejjad,O., Kone-Paut,I., Michel,M., Dernis,E., Khellaf,M., Limal,N., Job-Deslandre,C., Fautrel,B., Le Loet,X., and Sibilia,J. (2008). Interleukin-1 receptor antagonist (anakinra) treatment in patients with systemic-onset juvenile idiopathic arthritis or adult onset Still disease: preliminary experience in France. Annals of the Rheumatic Diseases *67*, 302-308.

Leung,J.Y., McKenzie,F.E., Uglialoro,A.M., Flores-Villanueva,P.O., Sorkin,B.C., Yunis,E.J., Hartl,D.L., and Goldfeld,A.E. (2000). Identification of phylogenetic footprints in primate tumor necrosis factor-alpha promoters. Proc. Natl. Acad. Sci. U. S. A *97*, 6614-6618.

Lewis,C.M. (2002). Genetic association studies: design, analysis and interpretation. Brief. Bioinform. *3*, 146-153.

Li,H., Tago,K., Io,K., Kuroiwa,K., Arai,T., Iwahana,H., Tominaga,S., and Yanagisawa,K. (2000). The cloning and nucleotide sequence of human ST2L cDNA. Genomics *67*, 284-290.

Li,P., Allen,H., Banerjee,S., Franklin,S., Herzog,L., Johnston,C., McDowell,J., Paskind,M., Rodman,L., Salfeld,J., and . (1995b). Mice deficient in IL-1 beta-converting enzyme are defective in production of mature IL-1 beta and resistant to endotoxic shock. Cell *80*, 401-411.

Li,P., Allen,H., Banerjee,S., Franklin,S., Herzog,L., Johnston,C., McDowell,J., Paskind,M., Rodman,L., Salfeld,J., and . (1995a). Mice deficient in IL-1 beta-converting enzyme are defective in production of mature IL-1 beta and resistant to endotoxic shock. Cell *80*, 401-411.

Liang,D., Ma,W., Yao,C., Liu,H., and Chen,X. (2006). Imbalance of interleukin 18 and interleukin 18 binding protein in patients with lupus nephritis. Cell Mol. Immunol. *3*, 303-306.

Lin,F.H., Chang,J.B., McGuire,M.H., Yee,J.A., and Brigman,B.E. (2010). Biphasic effects of interleukin-1beta on osteoblast differentiation in vitro. J. Orthop. Res. *28*, 958-964.

Lin,H., Ho,A.S., Haley-Vicente,D., Zhang,J., Bernal-Fussell,J., Pace,A.M., Hansen,D., Schweighofer,K., Mize,N.K., and Ford,J.E. (2001). Cloning and characterization of IL-1HY2, a novel interleukin-1 family member. J. Biol. Chem. *276*, 20597-20602.

Lindsley,C.B. (1987). Seasonal variation in systemic onset juvenile rheumatoid arthritis. Arthritis Rheum. *30*, 838-839.

Liu,L., Hutchinson,M.R., White,J.M., Somogyi,A.A., and Coller,J.K. (2009a). Association of IL-1B genetic polymorphisms with an increased risk of opioid and alcohol dependence. Pharmacogenet. Genomics *19*, 869-876.

Liu,N., Li,X., Liu,C., Zhao,Y., Cui,B., and Ning,G. (2010). The association of interleukin-1alpha and interleukin-1beta polymorphisms with the risk of Graves' disease in a case-control study and meta-analysis. Hum. Immunol. *71*, 397-401.

Liu,R., Hannenhalli,S., and Bucan,M. (2009b). Motifs and cis-regulatory modules mediating the expression of genes co-expressed in presynaptic neurons. Genome Biol. *10*, R72.

Lohning,M., Stroehmann,A., Coyle,A.J., Grogan,J.L., Lin,S., Gutierrez-Ramos,J.C., Levinson,D., Radbruch,A., and Kamradt,T. (1998). T1/ST2 is preferentially expressed on murine Th2 cells, independent of interleukin 4, interleukin 5, and interleukin 10, and important for Th2 effector function. Proc. Natl. Acad. Sci. U. S. A *95*, 6930-6935.

Loiarro,M., Ruggiero,V., and Sette,C. (2010). Targeting TLR/IL-1R signalling in human diseases. Mediators. Inflamm. *2010*, 674363.

Lomater,C., Gerloni,V., Gattinara,M., Mazzotti,J., Cimaz,R., and Fantini,F. (2000). Systemic onset juvenile idiopathic arthritis: a retrospective study of 80 consecutive patients followed for 10 years. J. Rheumatol. *27*, 491-496.

Loots,G.G. (2008). Genomic identification of regulatory elements by evolutionary sequence comparison and functional analysis. Adv. Genet. *61*, 269-293.

Lotito,A.P., Campa,A., Silva,C.A., Kiss,M.H., and Mello,S.B. (2007). Interleukin 18 as a marker of disease activity and severity in patients with juvenile idiopathic arthritis. J. Rheumatol. *34*, 823-830.

Lovell, D., Giannini, E., Kinura, Y., Li, S., Hashkes, P., Reiff, A., Wallace, C., Onel, K., Nadler, D., Rosenberg, J, Osgood, G., Papadopoulos, J., and Radin, A. Preliminary Evidence for Bioactivity of IL-1 TRAP (Rilonacept), a Long Acting IL-1 Inhibitor, in Systemic Juvenile Idiopathic Arthritis. Ann.Rheum.Dis. 65(Suppl II), 693. 2006.
Ref Type: Abstract

Lovell,D.J. and Bowyer,S.L. (2003). Treatment of neonatal onset multisystem inflammatory disease (NOMID) with IL-1Ra (anakinra). Arthritis Rheum. *48*, S517.

Lovell,D.J., Giannini,E.H., Reiff,A., Jones,O.Y., Schneider,R., Olson,J.C., Stein,L.D., Gedalia,A., Ilowite,N.T., Wallace,C.A., Lange,M., Finck,B.K., and Burge,D.J. (2003). Long-term efficacy and safety of etanercept in children with polyarticular-course juvenile rheumatoid arthritis: interim results from an ongoing multicenter, open-label, extended-treatment trial. Arthritis Rheum. *48*, 218-226.

Lovenberg,T.W., Crowe,P.D., Liu,C., Chalmers,D.T., Liu,X.J., Liaw,C., Clevenger,W., Oltersdorf,T., De Souza,E.B., and Maki,R.A. (1996). Cloning of a cDNA encoding a novel interleukin-1 receptor related protein (IL 1R-rp2). J. Neuroimmunol. *70*, 113-122.

Lundmark,P.E., Liljedahl,U., Boomsma,D.I., Mannila,H., Martin,N.G., Palotie,A., Peltonen,L., Perola,M., Spector,T.D., and Syvanen,A.C. (2008). Evaluation of HapMap data in six populations of European descent. Eur. J. Hum. Genet. *16*, 1142-1150.

Luotola,K., Pietila,A., Alanne,M., Lanki,T., Loo,B.M., Jula,A., Perola,M., Peters,A., Zeller,T., Blankenberg,S., and Salomaa,V. (2010). Genetic variation of the interleukin-1 family and nongenetic factors determining the interleukin-1 receptor antagonist phenotypes. Metabolism *59*, 1520-1527.

Luzina,I.G., Tsymbalyuk,N., Choi,J., Hasday,J.D., and Atamas,S.P. (2006). CCL18-stimulated upregulation of collagen production in lung fibroblasts requires Sp1 signaling and basal Smad3 activity. J. Cell Physiol *206*, 221-228.

Lydyard,P.M., Whelan,A., and Fanger,M.W. (2000). Instant Notes in Immunology. (Oxford: BIOS Scientific Publishers Limited).

Madson,K.L., Moore,T.L., Lawrence,J.M., III, and Osborn,T.G. (1994). Cytokine levels in serum and synovial fluid of patients with juvenile rheumatoid arthritis. J. Rheumatol. *21*, 2359-2363.

Maeno,N., Takei,S., Nomura,Y., Imanaka,H., Hokonohara,M., and Miyata,K. (2002). Highly elevated serum levels of interleukin-18 in systemic juvenile idiopathic arthritis but not in other juvenile idiopathic arthritis subtypes or in Kawasaki disease: comment on the article by Kawashima et al. Arthritis Rheum. *46*, 2539-2541.

Maier,J.A., Voulalas,P., Roeder,D., and Maciag,T. (1990). Extension of the life-span of human endothelial cells by an interleukin-1 alpha antisense oligomer. Science *249*, 1570-1574.

Malakooti,J., Dahdal,R.Y., Schmidt,L., Layden,T.J., Dudeja,P.K., and Ramaswamy,K. (1999). Molecular cloning, tissue distribution, and functional expression of the human Na(+)/H(+) exchanger NHE2. Am. J. Physiol *277*, G383-G390.

Mangge,H., Gallistl,S., and Schauenstein,K. (1999). Long-term follow-up of cytokines and soluble cytokine receptors in peripheral blood of patients with juvenile rheumatoid arthritis. J. Interferon Cytokine Res. *19*, 1005-1010.

Mariathasan,S., Weiss,D.S., Newton,K., McBride,J., O'Rourke,K., Roose-Girma,M., Lee,W.P., Weinrauch,Y., Monack,D.M., and Dixit,V.M. (2006). Cryopyrin activates the inflammasome in response to toxins and ATP. Nature *440*, 228-232.

Martinon,F., Burns,K., and Tschopp,J. (2002a). The inflammasome: a molecular platform triggering activation of inflammatory caspases and processing of proIL-beta. Mol. Cell *10*, 417-426.

Martinon,F., Burns,K., and Tschopp,J. (2002b). The inflammasome: a molecular platform triggering activation of inflammatory caspases and processing of proIL-beta. Mol. Cell *10*, 417-426.

Martinon,F., Petrilli,V., Mayor,A., Tardivel,A., and Tschopp,J. (2006). Gout-associated uric acid crystals activate the NALP3 inflammasome. Nature *440*, 237-241.

Mayor,A., Martinon,F., De Smedt,T., Petrilli,V., and Tschopp,J. (2007). A crucial function of SGT1 and HSP90 in inflammasome activity links mammalian and plant innate immune responses. Nat. Immunol. *8*, 497-503.

Mayor,C., Brudno,M., Schwartz,J.R., Poliakov,A., Rubin,E.M., Frazer,K.A., Pachter,L.S., and Dubchak,I. (2000b). VISTA : visualizing global DNA sequence alignments of arbitrary length. Bioinformatics. *16*, 1046-1047.

Mayor,C., Brudno,M., Schwartz,J.R., Poliakov,A., Rubin,E.M., Frazer,K.A., Pachter,L.S., and Dubchak,I. (2000a). VISTA : visualizing global DNA sequence alignments of arbitrary length. Bioinformatics. *16*, 1046-1047.

Mazodier,K., Marin,V., Novick,D., Farnarier,C., Robitail,S., Schleinitz,N., Veit,V., Paul,P., Rubinstein,M., Dinarello,C.A., Harle,J.R., and Kaplanski,G. (2005). Severe imbalance of IL-18/IL-18BP in patients with secondary hemophagocytic syndrome. Blood *106*, 3483-3489.

McDonald,J.H. (2009). Tests for nominal variables. In Handbook of Biological Statistics, (Baltimore: Sparky House Publishing), pp. 24-95.

McDowell,T.L., Symons,J.A., Ploski,R., Forre,O., and Duff,G.W. (1995). A genetic association between juvenile rheumatoid arthritis and a novel interleukin-1 alpha polymorphism. Arthritis Rheum. *38*, 221-228.

McMahan,C.J., Slack,J.L., Mosley,B., Cosman,D., Lupton,S.D., Brunton,L.L., Grubin,C.E., Wignall,J.M., Jenkins,N.A., Brannan,C.I., and . (1991). A novel IL-1 receptor, cloned from B cells by mammalian expression, is expressed in many cell types. EMBO J. *10*, 2821-2832.

Melnikov,V.Y., Ecder,T., Fantuzzi,G., Siegmund,B., Lucia,M.S., Dinarello,C.A., Schrier,R.W., and Edelstein,C.L. (2001). Impaired IL-18 processing protects caspase-1-deficient mice from ischemic acute renal failure. J. Clin. Invest *107*, 1145-1152.

Menon,R., Fortunato,S.J., Edwards,D.R., and Williams,S.M. (2010). Association of genetic variants, ethnicity and preterm birth with amniotic fluid cytokine concentrations. Ann. Hum. Genet. *74*, 165-183.

Meulenbelt,I., Seymour,A.B., Nieuwland,M., Huizinga,T.W., van Duijn,C.M., and Slagboom,P.E. (2004). Association of the interleukin-1 gene cluster with radiographic signs of osteoarthritis of the hip. Arthritis Rheum. *50*, 1179-1186.

Miller,B.E., Krasney,P.A., Gauvin,D.M., Holbrook,K.B., Koonz,D.J., Abruzzese,R.V., Miller,R.E., Pagani,K.A., Dolle,R.E., Ator,M.A., and . (1995). Inhibition of mature IL-1 beta production in murine macrophages and a murine model of inflammation by WIN 67694, an inhibitor of IL-1 beta converting enzyme. J. Immunol. *154*, 1331-1338.

Miller,D.K., Ayala,J.M., Egger,L.A., Raju,S.M., Yamin,T.T., Ding,G.J., Gaffney,E.P., Howard,A.D., Palyha,O.C., Rolando,A.M., and . (1993). Purification and characterization of active human interleukin-1 beta-converting enzyme from THP.1 monocytic cells. J. Biol. Chem. *268*, 18062-18069.

Miossec,P., Dinarello,C.A., and Ziff,M. (1986). Interleukin-1 lymphocyte chemotactic activity in rheumatoid arthritis synovial fluid. Arthritis Rheum. *29*, 461-470.

Mitcham,J.L., Parnet,P., Bonnert,T.P., Garka,K.E., Gerhart,M.J., Slack,J.L., Gayle,M.A., Dower,S.K., and Sims,J.E. (1996). T1/ST2 signaling establishes it as a member of an expanding interleukin-1 receptor family. J. Biol. Chem. *271*, 5777-5783.

Miterski,B., Drynda,S., Boschow,G., Klein,W., Oppermann,J., Kekow,J., and Epplen,J.T. (2004). Complex genetic predisposition in adult and juvenile rheumatoid arthritis. BMC. Genet. *5*, 2.

Monteleone,G., Trapasso,F., Parrello,T., Biancone,L., Stella,A., Iuliano,R., Luzza,F., Fusco,A., and Pallone,F. (1999). Bioactive IL-18 expression is up-regulated in Crohn's disease. J. Immunol. *163*, 143-147.

Montpetit,A., Nelis,M., Laflamme,P., Magi,R., Ke,X., Remm,M., Cardon,L., Hudson,T.J., and Metspalu,A. (2006). An evaluation of the performance of tag SNPs derived from HapMap in a Caucasian population. PLoS. Genet. *2*, e27.

Mookherjee,S., Banerjee,D., Chakraborty,S., Banerjee,A., Mukhopadhyay,I., Sen,A., and Ray,K. (2010). Association of IL1A and IL1B loci with primary open angle glaucoma. BMC. Med. Genet. *11*, 99.

Moroldo,M.B., Tague,B.L., Shear,E.S., Glass,D.N., and Giannini,E.H. (1997). Juvenile rheumatoid arthritis in affected sibpairs. Arthritis Rheum. *40*, 1962-1966.

Mosbruger,T.L., Duggal,P., Goedert,J.J., Kirk,G.D., Hoots,W.K., Tobler,L.H., Busch,M., Peters,M.G., Rosen,H.R., Thomas,D.L., and Thio,C.L. (2010). Large-scale candidate gene analysis of spontaneous clearance of hepatitis C virus. J. Infect. Dis. *201*, 1371-1380.

Mosley,B., Dower,S.K., Gillis,S., and Cosman,D. (1987a). Determination of the minimum polypeptide lengths of the functionally active sites of human interleukins 1 alpha and 1 beta. Proc. Natl. Acad. Sci. U. S. A *84*, 4572-4576.

Mosley,B., Urdal,D.L., Prickett,K.S., Larsen,A., Cosman,D., Conlon,P.J., Gillis,S., and Dower,S.K. (1987b). The interleukin-1 receptor binds the human interleukin-1 alpha precursor but not the interleukin-1 beta precursor. J. Biol. Chem. *262*, 2941-2944.

Mueller,J.C., Lohmussaar,E., Magi,R., Remm,M., Bettecken,T., Lichtner,P., Biskup,S., Illig,T., Pfeufer,A., Luedemann,J., Schreiber,S., Pramstaller,P., Pichler,I., Romeo,G., Gaddi,A., Testa,A., Wichmann,H.E., Metspalu,A., and Meitinger,T. (2005). Linkage disequilibrium patterns and tagSNP transferability among European populations. Am. J. Hum. Genet. *76*, 387-398.

Muhl,H., Kampfer,H., Bosmann,M., Frank,S., Radeke,H., and Pfeilschifter,J. (2000b). Interferon-gamma mediates gene expression of IL-18 binding protein in nonleukocytic cells. Biochem. Biophys. Res. Commun. *267*, 960-963.

Muhl,H., Kampfer,H., Bosmann,M., Frank,S., Radeke,H., and Pfeilschifter,J. (2000a). Interferon-gamma mediates gene expression of IL-18 binding protein in nonleukocytic cells. Biochem. Biophys. Res. Commun. *267*, 960-963.

Mulero,J.J., Pace,A.M., Nelken,S.T., Loeb,D.B., Correa,T.R., Drmanac,R., and Ford,J.E. (1999). IL1HY1: A novel interleukin-1 receptor antagonist gene. Biochem. Biophys. Res. Commun. *263*, 702-706.

Muller,K., Herner,E.B., Stagg,A., Bendtzen,K., and Woo,P. (1998). Inflammatory cytokines and cytokine antagonists in whole blood cultures of patients with systemic juvenile chronic arthritis. Br. J. Rheumatol. *37*, 562-569.

Munday,N.A., Vaillancourt,J.P., Ali,A., Casano,F.J., Miller,D.K., Molineaux,S.M., Yamin,T.T., Yu,V.L., and Nicholson,D.W. (1995). Molecular cloning and pro-apoptotic activity of ICErelII and ICErelIII, members of the ICE/CED-3 family of cysteine proteases. J. Biol. Chem. *270*, 15870-15876.

Munder,M., Mallo,M., Eichmann,K., and Modolell,M. (1998). Murine macrophages secrete interferon gamma upon combined stimulation with interleukin (IL)-12 and IL-18: A novel pathway of autocrine macrophage activation. J. Exp. Med. *187*, 2103-2108.

Murakami,M., Hibi,M., Nakagawa,N., Nakagawa,T., Yasukawa,K., Yamanishi,K., Taga,T., and Kishimoto,T. (1993). IL-6-induced homodimerization of gp130 and associated activation of a tyrosine kinase. Science *260*, 1808-1810.

Muzaffer,M.A., Dayer,J.M., Feldman,B.M., Pruzanski,W., Roux-Lombard,P., Schneider,R., Laxer,R.M., and Silverman,E.D. (2002). Differences in the profiles of circulating levels of soluble tumor necrosis factor receptors and interleukin 1 receptor antagonist reflect the heterogeneity of the subgroups of juvenile rheumatoid arthritis. J. Rheumatol. *29*, 1071-1078.

Muzio,M., Ni,J., Feng,P., and Dixit,V.M. (1997). IRAK (Pelle) family member IRAK-2 and MyD88 as proximal mediators of IL-1 signaling. Science *278*, 1612-1615.

Nakamura,K., Okamura,H., Nagata,K., Komatsu,T., and Tamura,T. (1993a). Purification of a factor which provides a costimulatory signal for gamma interferon production. Infect. Immun. *61*, 64-70.

Nakamura,K., Okamura,H., Nagata,K., Komatsu,T., and Tamura,T. (1993b). Purification of a factor which provides a costimulatory signal for gamma interferon production. Infect. Immun. *61*, 64-70.

Nakki,A., Kouhia,S.T., Saarela,J., Harilainen,A., Tallroth,K., Videman,T., Battie,M.C., Kaprio,J., Peltonen,L., and Kujala,U.M. (2010). Allelic variants of IL1R1 gene associate with severe hand osteoarthritis. BMC. Med. Genet. *11*, 50.

Nardone,J., Lee,D.U., Ansel,K.M., and Rao,A. (2004). Bioinformatics for the 'bench biologist': how to find regulatory regions in genomic DNA. Nat. Immunol. *5*, 768-774.

Nicklin,M.J., Barton,J.L., Nguyen,M., FitzGerald,M.G., Duff,G.W., and Kornman,K. (2002). A sequence-based map of the nine genes of the human interleukin-1 cluster. Genomics *79*, 718-725.

Nielsen,A.L., Jorgensen,P., Lerouge,T., Cervino,M., Chambon,P., and Losson,R. (2004). Nizp1, a novel multitype zinc finger protein that interacts with the NSD1 histone lysine methyltransferase through a unique C2HR motif. Mol. Cell Biol. *24*, 5184-5196.

Niki,Y., Yamada,H., Kikuchi,T., Toyama,Y., Matsumoto,H., Fujikawa,K., and Tada,N. (2004). Membrane-associated IL-1 contributes to chronic synovitis and cartilage destruction in human IL-1 alpha transgenic mice. J. Immunol. *172*, 577-584.

Nishimoto,N. and Kishimoto,T. (2004). Inhibition of IL-6 for the treatment of inflammatory diseases. Curr. Opin. Pharmacol. *4*, 386-391.

Nistala,K., Adams,S., Cambrook,H., Ursu,S., Olivito,B., de Jager,W., Evans,J.G., Cimaz,R., Bajaj-Elliott,M., and Wedderburn,L.R. (2010). Th17 plasticity in human autoimmune arthritis is driven by the inflammatory environment. Proc. Natl. Acad. Sci. U. S. A *107*, 14751-14756.

Nistala,K., Woo,P., and Wedderburn,L. (2009). Juvenile Idiopathic Arthritis. In Kelly's Textbook of Rheumatology, G.S.Firestein, R.C.Budd, E.D.Harris, I.B.McInnes, S.Ruddy, and J.S.Sergent, eds. (Philadelphia: Saunders Elsevier), pp. 1657-1676.

Nitsch,R., Di,D., V, di Gennaro,A., de Cristofaro,T., Abbondante,S., De Felice,M., Zannini,M., and Di Lauro,R. (2010). Comparative genomics reveals a functional thyroid-specific element in the far upstream region of the PAX8 gene. BMC. Genomics *11*, 306.

Nobrega,M.A., Ovcharenko,I., Afzal,V., and Rubin,E.M. (2003). Scanning human gene deserts for long-range enhancers. Science *302*, 413.

Nold,M.F., Nold-Petry,C.A., Zepp,J.A., Palmer,B.E., Bufler,P., and Dinarello,C.A. (2010). IL-37 is a fundamental inhibitor of innate immunity. Nat. Immunol. *11*, 1014-1022.

Nolsoe,R.L., Pociot,F., Novick,D., Rubinstein,M., Kim,S.H., Dinarello,C.A., and Mandrup-Poulsen,T. (2003). Mutation scan of a type 1 diabetes candidate gene: the human interleukin-18 binding protein gene. Ann. N. Y. Acad. Sci. *1005*, 332-339.

Novick,D., Kim,S.H., Fantuzzi,G., Reznikov,L.L., Dinarello,C.A., and Rubinstein,M. (1999). Interleukin-18 binding protein: a novel modulator of the Th1 cytokine response. Immunity. *10*, 127-136.

Novick,D., Schwartsburd,B., Pinkus,R., Suissa,D., Belzer,I., Sthoeger,Z., Keane,W.F., Chvatchko,Y., Kim,S.H., Fantuzzi,G., Dinarello,C.A., and Rubinstein,M. (2001). A novel IL-18BP ELISA shows elevated serum IL-18BP in sepsis and extensive decrease of free IL-18. Cytokine *14*, 334-342.

Ogilvie,E.M., Fife,M.S., Thompson,S.D., Twine,N., Tsoras,M., Moroldo,M., Fisher,S.A., Lewis,C.M., Prieur,A.M., Glass,D.N., and Woo,P. (2003). The -174G allele of the interleukin-6 gene confers susceptibility to systemic arthritis in children: a multicenter study using simplex and multiplex juvenile idiopathic arthritis families. Arthritis Rheum. *48*, 3202-3206.

Ogilvie,E.M., Khan,A., Hubank,M., Kellam,P., and Woo,P. (2007). Specific gene expression profiles in systemic juvenile idiopathic arthritis. Arthritis Rheum. *56*, 1954-1965.

Okamura,H., Tsutsi,H., Komatsu,T., Yutsudo,M., Hakura,A., Tanimoto,T., Torigoe,K., Okura,T., Nukada,Y., Hattori,K., and . (1995b). Cloning of a new cytokine that induces IFN-gamma production by T cells. Nature *378*, 88-91.

Okamura,H., Tsutsi,H., Komatsu,T., Yutsudo,M., Hakura,A., Tanimoto,T., Torigoe,K., Okura,T., Nukada,Y., Hattori,K., and . (1995a). Cloning of a new cytokine that induces IFN-gamma production by T cells. Nature *378*, 88-91.

Okazaki,M., Iwasaki,Y., Nishiyama,M., Taguchi,T., Tsugita,M., Nakayama,S., Kambayashi,M., Hashimoto,K., and Terada,Y. (2010). PPARbeta/delta regulates the human SIRT1 gene transcription via Sp1. Endocr. J. *57*, 403-413.

Pal,P., Xi,H., Sun,G., Kaushal,R., Meeks,J.J., Thaxton,C.S., Guha,S., Jin,C.H., Suarez,B.K., Catalona,W.J., and Deka,R. (2007). Tagging SNPs in the kallikrein genes 3 and 2 on 19q13 and their associations with prostate cancer in men of European origin. Hum. Genet. *122*, 251-259.

Pan,G., Risser,P., Mao,W., Baldwin,D.T., Zhong,A.W., Filvaroff,E., Yansura,D., Lewis,L., Eigenbrot,C., Henzel,W.J., and Vandlen,R. (2001). IL-1H, an interleukin 1-related protein that binds IL-18 receptor/IL-1Rrp. Cytokine *13*, 1-7.

Pascual,V., Allantaz,F., Arce,E., Punaro,M., and Banchereau,J. (2005). Role of interleukin-1 (IL-1) in the pathogenesis of systemic onset juvenile idiopathic arthritis and clinical response to IL-1 blockade. J. Exp. Med. *201*, 1479-1486.

Paulukat,J., Bosmann,M., Nold,M., Garkisch,S., Kampfer,H., Frank,S., Raedle,J., Zeuzem,S., Pfeilschifter,J., and Muhl,H. (2001b). Expression and release of IL-18 binding protein in response to IFN-gamma. J. Immunol. *167*, 7038-7043.

Paulukat,J., Bosmann,M., Nold,M., Garkisch,S., Kampfer,H., Frank,S., Raedle,J., Zeuzem,S., Pfeilschifter,J., and Muhl,H. (2001a). Expression and release of IL-18 binding protein in response to IFN-gamma. J. Immunol. *167*, 7038-7043.

Perez,M.D., Abrams,S.A., Loddeke,L., Shypailo,R., and Ellis,K.J. (2000). Effects of rheumatic disease and corticosteroid treatment on calcium metabolism and bone density in children assessed one year after diagnosis, using stable isotopes and dual energy x-ray absorptiometry. J. Rheumatol. Suppl *58*, 38-43.

Petty,R.E. (1998). Classification of childhood arthritis: a work in progress. Baillieres Clin. Rheumatol. *12*, 181-190.

Petty,R.E. (2001). Growing pains: the ILAR classification of juvenile idiopathic arthritis. J. Rheumatol. *28*, 927-928.

Petty,R.E., Southwood,T.R., Manners,P., Baum,J., Glass,D.N., Goldenberg,J., He,X., Maldonado-Cocco,J., Orozco-Alcala,J., Prieur,A.M., Suarez-Almazor,M.E., and Woo,P. (2004). International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J. Rheumatol. *31*, 390-392.

Phelan,J.D., Thompson,S.D., and Glass,D.N. (2006). Susceptibility to JRA/JIA: complementing general autoimmune and arthritis traits. Genes Immun. *7*, 1-10.

Pizarro,T.T., Michie,M.H., Bentz,M., Woraratanadharm,J., Smith,M.F., Jr., Foley,E., Moskaluk,C.A., Bickston,S.J., and Cominelli,F. (1999). IL-18, a novel immunoregulatory cytokine, is up-regulated in Crohn's disease: expression and localization in intestinal mucosal cells. J. Immunol. *162*, 6829-6835.

Plater-Zyberk,C., Joosten,L.A., Helsen,M.M., Sattonnet-Roche,P., Siegfried,C., Alouani,S., van de Loo,F.A., Graber,P., Aloni,S., Cirillo,R., Lubberts,E., Dinarello,C.A., van den

Berg,W.B., and Chvatchko,Y. (2001). Therapeutic effect of neutralizing endogenous IL-18 activity in the collagen-induced model of arthritis. J. Clin. Invest *108*, 1825-1832.

Pokorny,V., McQueen,F., Yeoman,S., Merriman,M., Merriman,A., Harrison,A., Highton,J., and McLean,L. (2005). Evidence for negative association of the chemokine receptor CCR5 d32 polymorphism with rheumatoid arthritis. Ann. Rheum. Dis. *64*, 487-490.

Prahalad,S. (2006). Negative association between the chemokine receptor CCR5-Delta32 polymorphism and rheumatoid arthritis: a meta-analysis. Genes Immun. *7*, 264-268.

Prahalad,S., Bohnsack,J.F., Jorde,L.B., Whiting,A., Clifford,B., Dunn,D., Weiss,R., Moroldo,M., Thompson,S.D., Glass,D.N., and Bamshad,M.J. (2006). Association of two functional polymorphisms in the CCR5 gene with juvenile rheumatoid arthritis. Genes Immun. *7*, 468-475.

Prahalad,S., O'brien,E., Fraser,A.M., Kerber,R.A., Mineau,G.P., Pratt,D., Donaldson,D., Bamshad,M.J., and Bohnsack,J. (2004). Familial aggregation of juvenile idiopathic arthritis. Arthritis Rheum. *50*, 4022-4027.

Prahalad,S., Ryan,M.H., Shear,E.S., Thompson,S.D., Giannini,E.H., and Glass,D.N. (2000). Juvenile rheumatoid arthritis: linkage to HLA demonstrated by allele sharing in affected sibpairs. Arthritis Rheum. *43*, 2335-2338.

Prahalad,S., Shear,E.S., Thompson,S.D., Giannini,E.H., and Glass,D.N. (2002). Increased prevalence of familial autoimmunity in simplex and multiplex families with juvenile rheumatoid arthritis. Arthritis Rheum. *46*, 1851-1856.

Prieur,A.M., Roux-Lombard,P., and Dayer,J.M. (1996). Dynamics of fever and the cytokine network in systemic juvenile arthritis. Rev. Rhum. Engl. Ed *63*, 163-170.

Purcell,S., Neale,B., Todd-Brown,K., Thomas,L., Ferreira,M.A., Bender,D., Maller,J., Sklar,P., de Bakker,P.I., Daly,M.J., and Sham,P.C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. Am. J. Hum. Genet. *81*, 559-575.

Quartier, P., Lequerre, T., Rosellini, D., Sibilia, J., Debandt, M., Kone-Paut, I., and Job-Deslandre, C. Anakinra in Systemic-Onset Juvenile Idiopathic Arthritis and Adult Onset Still's Disease. Ann.Rheum.Dis. 65(Suppl II), 107. 2006.
Ref Type: Abstract

Quartier,P., Taupin,P., Bourdeaut,F., Lemelle,I., Pillet,P., Bost,M., Sibilia,J., Kone-Paut,I., Gandon-Laloum,S., LeBideau,M., Bader-Meunier,B., Mouy,R., Debre,M., Landais,P., and Prieur,A.M. (2003b). Efficacy of etanercept for the treatment of juvenile idiopathic arthritis according to the onset type. Arthritis Rheum. *48*, 1093-1101.

Quartier,P., Taupin,P., Bourdeaut,F., Lemelle,I., Pillet,P., Bost,M., Sibilia,J., Kone-Paut,I., Gandon-Laloum,S., LeBideau,M., Bader-Meunier,B., Mouy,R., Debre,M., Landais,P., and Prieur,A.M. (2003a). Efficacy of etanercept for the treatment of juvenile idiopathic arthritis according to the onset type. Arthritis Rheum. *48*, 1093-1101.

Rahman,P., Sun,S., Peddle,L., Snelgrove,T., Melay,W., Greenwood,C., and Gladman,D. (2006). Association between the interleukin-1 family gene cluster and psoriatic arthritis. Arthritis Rheum. *54*, 2321-2325.

Ramadas,R.A., Ewart,S.L., Medoff,B.D., and Levine,A.M. (2010). IL1F9 Stimulates Chemokine Production and Neutrophil Influx in Mouse Lungs. Am. J. Respir. Cell Mol. Biol.

Ramadori,G., Sipe,J.D., Dinarello,C.A., Mizel,S.B., and Colten,H.R. (1985). Pretranslational modulation of acute phase hepatic protein synthesis by murine recombinant interleukin 1 (IL-1) and purified human IL-1. J. Exp. Med. *162*, 930-942.

Ramanan,A.V. and Grom,A.A. (2005). Does systemic-onset juvenile idiopathic arthritis belong under juvenile idiopathic arthritis? Rheumatology (Oxford) *44*, 1350-1353.

Reddy,S.D., Pakala,S.B., Ohshiro,K., Rayala,S.K., and Kumar,R. (2009). MicroRNA-661, a c/EBPalpha target, inhibits metastatic tumor antigen 1 and regulates its functions. Cancer Res. *69*, 5639-5642.

Reddy,V.S., Prabhu,S.D., Mummidi,S., Valente,A.J., Venkatesan,B., Shanmugam,P., Delafontaine,P., and Chandrasekar,B. (2010). Interleukin-18 induces EMMPRIN expression in primary cardiomyocytes via JNK/Sp1 signaling and MMP-9 in part via EMMPRIN and through AP-1 and NF-kappaB activation. Am. J. Physiol Heart Circ. Physiol *299*, H1242-H1254.

Reijmerink,N.E., Postma,D.S., Bruinenberg,M., Nolte,I.M., Meyers,D.A., Bleecker,E.R., and Koppelman,G.H. (2008). Association of IL1RL1, IL18R1, and IL18RAP gene cluster polymorphisms with asthma and atopy. J. Allergy Clin. Immunol. *122*, 651-654.

Reiner,A.P., Wurfel,M.M., Lange,L.A., Carlson,C.S., Nord,A.S., Carty,C.L., Rieder,M.J., Desmarais,C., Jenny,N.S., Iribarren,C., Walston,J.D., Williams,O.D., Nickerson,D.A., and Jarvik,G.P. (2008). Polymorphisms of the IL1-receptor antagonist gene (IL1RN) are associated with multiple markers of systemic inflammation. Arterioscler. Thromb. Vasc. Biol. *28*, 1407-1412.

Reveille,J.D., Sims,A.M., Danoy,P., Evans,D.M., Leo,P., Pointon,J.J., Jin,R., Zhou,X., Bradbury,L.A., Appleton,L.H., Davis,J.C., Diekman,L., Doan,T., Dowling,A., Duan,R., Duncan,E.L., Farrar,C., Hadler,J., Harvey,D., Karaderi,T., Mogg,R., Pomeroy,E., Pryce,K., Taylor,J., Savage,L., Deloukas,P., Kumanduri,V., Peltonen,L., Ring,S.M., Whittaker,P., Glazov,E., Thomas,G.P., Maksymowych,W.P., Inman,R.D., Ward,M.M., Stone,M.A., Weisman,M.H., Wordsworth,B.P., and Brown,M.A. (2010). Genome-wide association study of ankylosing spondylitis identifies non-MHC susceptibility loci. Nat. Genet. *42*, 123-127.

Ribas,G., Gonzalez-Neira,A., Salas,A., Milne,R.L., Vega,A., Carracedo,B., Gonzalez,E., Barroso,E., Fernandez,L.P., Yankilevich,P., Robledo,M., Carracedo,A., and Benitez,J. (2006). Evaluating HapMap SNP data transferability in a large-scale genotyping project involving 175 cancer-associated genes. Hum. Genet. *118*, 669-679.

Ringwood,L. and Li,L. (2008). The involvement of the interleukin-1 receptor-associated kinases (IRAKs) in cellular signaling networks controlling inflammation. Cytokine *42*, 1-7.

Rooney,M., David,J., Symons,J., Di Giovine,F., Varsani,H., and Woo,P. (1995). Inflammatory cytokine responses in juvenile chronic arthritis. Br. J. Rheumatol. *34*, 454-460.

Rosengren,S., Hoffman,H.M., Bugbee,W., and Boyle,D.L. (2005). Expression and regulation of cryopyrin and related proteins in rheumatoid arthritis synovium. Ann. Rheum. Dis. *64*, 708-714.

Rueda,B., Gonzalez-Gay,M.A., Mataran,L., Lopez-Nevot,M.A., and Martin,J. (2005). Interleukin-18-promoter polymorphisms are not relevant in rheumatoid arthritis. Tissue Antigens *65*, 544-548.

Ruperto, N., Quartier, P., Wulffraat, N., Woo, P., Loy, A., Mouy, R., Bader-Meunier, B., Prakken, B. J., Noseda, E., Rordorf, C., and Maritini, A. for the Paediatric Rheumatology International Trials Organisation PRINTO. A phase II trial with canakinumab, a new IL-1beta blocking monoclonal antibody (ACZ885), to evaluate preliminary dosing, safety and efficacy profile in children with systemic Juvenile Idiopathic arthritis (sJIA). Pediatric Rheumatology 6. 2008.
Ref Type: Abstract

Russo,R.A. and Katsicas,M.M. (2009). Clinical remission in patients with systemic juvenile idiopathic arthritis treated with anti-tumor necrosis factor agents. J. Rheumatol. *36*, 1078-1082.

Sagai,T., Hosoya,M., Mizushina,Y., Tamura,M., and Shiroishi,T. (2005). Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb. Development *132*, 797-803.

Saha,A., Hammond,C.E., Gooz,M., and Smolka,A.J. (2008). The role of Sp1 in IL-1beta and H. pylori-mediated regulation of H,K-ATPase gene transcription. Am. J. Physiol Gastrointest. Liver Physiol *295*, G977-G986.

Saila,H., Pitkaniemi,J., Tuomilehto,J., Savolainen,A., Alakulppi,N., Tuomilehto-Wolf,E., Leirisalo-Repo,M., and Aho,K. (2004). HLA and susceptibility to juvenile idiopathic arthritis: a study of affected sibpairs in an isolated Finnish population. J. Rheumatol. *31*, 2281-2285.

Samuel,J.M., Kelberman,D., Smith,A.J., Humphries,S.E., and Woo,P. (2008). Identification of a novel regulatory region in the interleukin-6 gene promoter. Cytokine *42*, 256-264.

Sana,T.R., Debets,R., Timans,J.C., Bazan,J.F., and Kastelein,R.A. (2000). Computational identification, cloning, and characterization of IL-1R9, a novel interleukin-1 receptor-like gene encoded over an unusually large interval of human chromosome Xq22.2-q22.3. Genomics *69*, 252-262.

Satagopan,J.M., Verbel,D.A., Venkatraman,E.S., Offit,K.E., and Begg,C.B. (2002). Two-stage designs for gene-disease association studies. Biometrics *58*, 163-170.

Savage,N., Puren,A.J., Orencole,S.F., Ikejima,T., Clark,B.D., and Dinarello,C.A. (1989). Studies on IL-1 receptors on D10S T-helper cells: demonstration of two molecularly and antigenically distinct IL-1 binding proteins. Cytokine *1*, 23-35.

Savolainen,A., Saila,H., Kotaniemi,K., Kaipianen-Seppanen,O., Leirisalo-Repo,M., and Aho,K. (2000). Magnitude of the genetic component in juvenile idiopathic arthritis. Ann. Rheum. Dis. *59*, 1001.

Sawhney,S., Woo,P., and Murray,K.J. (2001a). Macrophage activation syndrome: a potentially fatal complication of rheumatic disorders. Arch. Dis. Child *85*, 421-426.

Sawhney,S., Woo,P., and Murray,K.J. (2001b). Macrophage activation syndrome: a potentially fatal complication of rheumatic disorders. Arch. Dis. Child *85*, 421-426.

Scheibel,I., Veit,T., Neves,A.G., Souza,L., Prezzi,S., Machado,S., Kohem,C., Icarelli,M., Xavier,R., Brenol,J.C., and Chies,J.A. (2008). Differential CCR5Delta32 allelic frequencies in juvenile idiopathic arthritis subtypes: evidence for different regulatory roles of CCR5 in rheumatological diseases. Scand. J. Rheumatol. *37*, 13-17.

Scheidereit,C. (2006). IkappaB kinase complexes: gateways to NF-kappaB activation and transcription. Oncogene *25*, 6685-6705.

Schmitz,J., Owyang,A., Oldham,E., Song,Y., Murphy,E., McClanahan,T.K., Zurawski,G., Moshrefi,M., Qin,J., Li,X., Gorman,D.M., Bazan,J.F., and Kastelein,R.A. (2005). IL-33, an interleukin-1-like cytokine that signals via the IL-1 receptor-related protein ST2 and induces T helper type 2-associated cytokines. Immunity *23*, 479-490.

Schneider,R. and Laxer,R.M. (1998). Systemic Onset Juvenile Rheumatoid Arthritis. In Paediatric Rheumatology, A.M.Prieur and M.Dougados, eds. Bailliere Tindall), pp. 245-272.

Seckinger,P., Lowenthal,J.W., Williamson,K., Dayer,J.M., and MacDonald,H.R. (1987). A urine inhibitor of interleukin 1 activity that blocks ligand binding. J. Immunol. *139*, 1546-1549.

Seitz,M., Loetscher,P., Dewald,B., Towbin,H., Rordorf,C., Gallati,H., and Gerber,N.J. (1996). Interleukin 1 (IL-1) receptor antagonist, soluble tumor necrosis factor receptors, IL-1 beta, and IL-8--markers of remission in rheumatoid arthritis during treatment with methotrexate. J. Rheumatol. *23*, 1512-1516.

Seldin,M.F., Shigeta,R., Villoslada,P., Selmi,C., Tuomilehto,J., Silva,G., Belmont,J.W., Klareskog,L., and Gregersen,P.K. (2006). European population substructure: clustering of northern and southern populations. PLoS. Genet. *2*, e143.

Shan,N.N., Zhu,X.J., Peng,J., Qin,P., Zhuang,X.W., Wang,H.C., and Hou,M. (2009). Interleukin 18 and interleukin 18 binding protein in patients with idiopathic thrombocytopenic purpura. Br. J. Haematol. *144*, 755-761.

Sharma,A., Kumar,M., Aich,J., Hariharan,M., Brahmachari,S.K., Agrawal,A., and Ghosh,B. (2009). Posttranscriptional regulation of interleukin-10 expression by hsa-miR-106a. Proc. Natl. Acad. Sci. U. S. A *106*, 5761-5766.

Sharma,S., Kulk,N., Nold,M.F., Graf,R., Kim,S.H., Reinhardt,D., Dinarello,C.A., and Bufler,P. (2008). The IL-1 family member 7b translocates to the nucleus and down-regulates proinflammatory cytokines. J. Immunol. *180*, 5477-5482.

Shimizu,M., Yokoyama,T., Yamada,K., Kaneda,H., Wada,H., Wada,T., Toma,T., Ohta,K., Kasahara,Y., and Yachie,A. (2010). Distinct cytokine profiles of systemic-onset juvenile idiopathic arthritis-associated macrophage activation syndrome with particular emphasis on the role of interleukin-18 in its pathogenesis. Rheumatology (Oxford) *49*, 1645-1653.

Shirakawa,F. and Mizel,S.B. (1989). In vitro activation and nuclear translocation of NF-kappa B catalyzed by cyclic AMP-dependent protein kinase and protein kinase C. Mol. Cell Biol. *9*, 2424-2430.

Shirts,B.H., Wood,J., Yolken,R.H., and Nimgaonkar,V.L. (2008). Comprehensive evaluation of positional candidates in the IL-18 pathway reveals suggestive associations with schizophrenia and herpes virus seropositivity. Am. J. Med. Genet. B Neuropsychiatr. Genet. *147*, 343-350.

Siegmund,B., Fantuzzi,G., Rieder,F., Gamboni-Robertson,F., Lehr,H.A., Hartmann,G., Dinarello,C.A., Endres,S., and Eigler,A. (2001a). Neutralization of interleukin-18 reduces severity in murine colitis and intestinal IFN-gamma and TNF-alpha production. Am. J. Physiol Regul. Integr. Comp Physiol *281*, R1264-R1273.

Siegmund,B., Lehr,H.A., Fantuzzi,G., and Dinarello,C.A. (2001b). IL-1 beta -converting enzyme (caspase-1) in intestinal inflammation. Proc. Natl. Acad. Sci. U. S. A *98*, 13249-13254.

Siegmund,B., Lehr,H.A., Fantuzzi,G., and Dinarello,C.A. (2001c). IL-1 beta -converting enzyme (caspase-1) in intestinal inflammation. Proc. Natl. Acad. Sci. U. S. A *98*, 13249-13254.

Simmonds,M.J. and Gough,S.C. (2004). Genetic insights into disease mechanisms of autoimmunity. Br. Med. Bull. *71*, 93-113.

Simon,D., Lucidarme,N., Prieur,A.M., Ruiz,J.C., and Czernichow,P. (2003). Effects on growth and body composition of growth hormone treatment in children with juvenile idiopathic arthritis requiring steroid therapy. J. Rheumatol. *30*, 2492-2499.

Simon,D.B., Lu,Y., Choate,K.A., Velazquez,H., Al Sabban,E., Praga,M., Casari,G., Bettinelli,A., Colussi,G., Rodriguez-Soriano,J., McCredie,D., Milford,D., Sanjad,S., and Lifton,R.P. (1999). Paracellin-1, a renal tight junction protein required for paracellular Mg2+ resorption. Science *285*, 103-106.

Sims,A.M., Timms,A.E., Bruges-Armas,J., Burgos-Vargas,R., Chou,C.T., Doan,T., Dowling,A., Fialho,R.N., Gergely,P., Gladman,D.D., Inman,R., Kauppi,M., Kaarela,K., Laiho,K., Maksymowych,W., Pointon,J.J., Rahman,P., Reveille,J.D., Sorrentino,R., Tuomilehto,J., Vargas-Alarcon,G., Wordsworth,B.P., Xu,H., and Brown,M.A. (2008). Prospective meta-analysis of interleukin 1 gene complex polymorphisms confirms associations with ankylosing spondylitis. Ann. Rheum. Dis. *67*, 1305-1309.

Sims,J.E., Gayle,M.A., Slack,J.L., Alderson,M.R., Bird,T.A., Giri,J.G., Colotta,F., Re,F., Mantovani,A., Shanebeck,K., and . (1993). Interleukin 1 signaling occurs exclusively via the type I receptor. Proc. Natl. Acad. Sci. U. S. A *90*, 6155-6159.

Sims,J.E., Giri,J.G., and Dower,S.K. (1994). The two interleukin-1 receptors play different roles in IL-1 actions. Clin. Immunol. Immunopathol. *72*, 9-14.

Sims,J.E., March,C.J., Cosman,D., Widmer,M.B., MacDonald,H.R., McMahan,C.J., Grubin,C.E., Wignall,J.M., Jackson,J.L., Call,S.M., and . (1988). cDNA expression cloning of the IL-1 receptor, a member of the immunoglobulin superfamily. Science *241*, 585-589.

Sims,J.E., Nicklin,M.J., Bazan,J.F., Barton,J.L., Busfield,S.J., Ford,J.E., Kastelein,R.A., Kumar,S., Lin,H., Mulero,J.J., Pan,J., Pan,Y., Smith,D.E., and Young,P.R. (2001). A new nomenclature for IL-1-family genes. Trends Immunol. *22*, 536-537.

Sivalingam,S.P., Thumboo,J., Vasoo,S., and Fong,K.Y. (2007). HLA-DRB1*04 gene polymorphisms and expressions profiles of interleukin-18 and interleukin-18 binding protein following in vitro stimulation in human peripheral blood mononuclear cells of healthy individuals and patients with rheumatoid arthritis. Life Sci. *80*, 1887-1896.

Sivalingam,S.P., Yoon,K.H., Koh,D.R., and Fong,K.Y. (2003). Single-nucleotide polymorphisms of the interleukin-18 gene promoter region in rheumatoid arthritis patients: protective effect of AA genotype. Tissue Antigens *62*, 498-504.

Skol,A.D., Scott,L.J., Abecasis,G.R., and Boehnke,M. (2006). Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. Nat. Genet. *38*, 209-213.

Smeets,R.L., van de Loo,F.A., Joosten,L.A., Arntz,O.J., Bennink,M.B., Loesberg,W.A., Dmitriev,I.P., Curiel,D.T., Martin,M.U., and van den Berg,W.B. (2003). Effectiveness of the soluble form of the interleukin-1 receptor accessory protein as an inhibitor of interleukin-1 in collagen-induced arthritis. Arthritis Rheum. *48*, 2949-2958.

Smith,A.J., Keen,L.J., Billingham,M.J., Perry,M.J., Elson,C.J., Kirwan,J.R., Sims,J.E., Doherty,M., Spector,T.D., and Bidwell,J.L. (2004). Extended haplotypes and linkage disequilibrium in the IL1R1-IL1A-IL1B-IL1RN gene cluster: association with knee osteoarthritis. Genes Immun. *5*, 451-460.

Smith,D.E., Hanna,R., Della,F., Moore,H., Chen,H., Farese,A.M., MacVittie,T.J., Virca,G.D., and Sims,J.E. (2003). The soluble form of IL-1 receptor accessory protein enhances the ability of soluble type II IL-1 receptor to inhibit IL-1 action. Immunity. *18*, 87-96.

Smith,D.E., Renshaw,B.R., Ketchem,R.R., Kubin,M., Garka,K.E., and Sims,J.E. (2000). Four new members expand the interleukin-1 superfamily. J. Biol. Chem. *275*, 1169-1175.

Sothern,R.B., Roitman-Johnson,B., Kanabrocki,E.L., Yager,J.G., Fuerstenberg,R.K., Weatherbee,J.A., Young,M.R., Nemchausky,B.M., and Scheving,L.E. (1995a). Circadian characteristics of interleukin-6 in blood and urine of clinically healthy men. In Vivo *9*, 331-339.

Sothern,R.B., Roitman-Johnson,B., Kanabrocki,E.L., Yager,J.G., Roodell,M.M., Weatherbee,J.A., Young,M.R., Nenchausky,B.M., and Scheving,L.E. (1995b). Circadian characteristics of circulating interleukin-6 in men. J. Allergy Clin. Immunol. *95*, 1029-1035.

Speckmaier,M., Findeisen,J., Woo,P., Hall,A., Sills,J.A., Price,T., Hollingworth,P., Craft,A., and Ansell,B.M. (1989). Low-dose methotrexate in systemic onset juvenile chronic arthritis. Clin. Exp. Rheumatol. *7*, 647-650.

Stack,J.H., Beaumont,K., Larsen,P.D., Straley,K.S., Henkel,G.W., Randle,J.C., and Hoffman,H.M. (2005). IL-converting enzyme/caspase-1 inhibitor VX-765 blocks the hypersensitive response to an inflammatory stimulus in monocytes from familial cold autoinflammatory syndrome patients. J. Immunol. *175*, 2630-2634.

Stevenson,F.T., Bursten,S.L., Fanton,C., Locksley,R.M., and Lovett,D.H. (1993). The 31-kDa precursor of interleukin 1 alpha is myristoylated on specific lysines within the 16-kDa N-terminal propiece. Proc. Natl. Acad. Sci. U. S. A *90*, 7245-7249.

Stock,C.J.W., Ogilvie,E.M., Samuel,J.M., Fife,M., Lewis,C.M., and Woo,P. (2008). Comprehensive association study of genetic variants in the IL-1 gene family in systemic juvenile idiopathic arthritis. Genes and Immunity *9*, 349-357.

Stoeber,E. (1981). Prognosis in juvenile chronic arthritis. Follow-up of 433 chronic rheumatic children. Eur. J. Pediatr. *135*, 225-228.

Strempel,J.M., Grenningloh,R., Ho,I.C., and Vercelli,D. (2010). Phylogenetic and functional analysis identifies Ets-1 as a novel regulator of the Th2 cytokine gene locus. J. Immunol. *184*, 1309-1316.

Stylianou,E., O'Neill,L.A., Rawlinson,L., Edbrooke,M.R., Woo,P., and Saklatvala,J. (1992). Interleukin 1 induces NF-kappa B through its type I but not its type II receptor in lymphocytes. J. Biol. Chem. *267*, 15836-15841.

Subramaniam,S., Stansberg,C., and Cunningham,C. (2004). The interleukin 1 receptor family. Dev. Comp Immunol. *28*, 415-428.

Sugiura,T., Kawaguchi,Y., Harigai,M., Terajima-Ichida,H., Kitamura,Y., Furuya,T., Ichikawa,N., Kotake,S., Tanaka,M., Hara,M., and Kamatani,N. (2002). Association between adult-onset Still's disease and interleukin-18 gene polymorphisms. Genes Immun. *3*, 394-399.

Sugiura,T., Maeno,N., Kawaguchi,Y., Takei,S., Imanaka,H., Kawano,Y., Terajima-Ichida,H., Hara,M., and Kamatani,N. (2006). A promoter haplotype of the interleukin-18 gene is associated with juvenile idiopathic arthritis in the Japanese population. Arthritis Res. Ther. *8*, R60.

Sutterwala,F.S., Ogura,Y., Szczepanik,M., Lara-Tejero,M., Lichtenberger,G.S., Grant,E.P., Bertin,J., Coyle,A.J., Galan,J.E., Askenase,P.W., and Flavell,R.A. (2006). Critical role for NALP3/CIAS1/Cryopyrin in innate and adaptive immunity through its regulation of caspase-1. Immunity *24*, 317-327.

Svantesson,H., Akesson,A., Eberhardt,K., and Elborgh,R. (1983). Prognosis in juvenile rheumatoid arthritis with systemic onset. A follow-up study. Scand. J. Rheumatol. *12*, 139-144.

Swellam,M., Mahmoud,M.S., Samy,N., and Gamal,A.A. (2010). Potential influence of interleukin-1 receptor antagonist gene polymorphism on knee osteoarthritis risk. Dis. Markers *28*, 299-305.

Symmons,D.P., Jones,M., Osborne,J., Sills,J., Southwood,T.R., and Woo,P. (1996). Pediatric rheumatology in the United Kingdom: data from the British Pediatric Rheumatology Group National Diagnostic Register. J. Rheumatol. *23*, 1975-1980.

Symons,J.A., Eastgate,J.A., and Duff,G.W. (1991). Purification and characterization of a novel soluble receptor for interleukin 1. J. Exp. Med. *174*, 1251-1254.

Symons,J.A., Young,P.R., and Duff,G.W. (1995). Soluble type II interleukin 1 (IL-1) receptor binds and blocks processing of IL-1 beta precursor and loses affinity for IL-1 receptor antagonist. Proc. Natl. Acad. Sci. U. S. A *92*, 1714-1718.

Taga,T. and Kishimoto,T. (1997). Gp130 and the interleukin-6 family of cytokines. Annu. Rev. Immunol. *15*, 797-819.

Takagawa,T., Tamura,K., Takeda,N., Tomita,T., Ohda,Y., Fukunaga,K., Hida,N., Ohnishi,K., Hori,K., Kosaka,T., Fukuda,Y., Ikeuchi,H., Yamamura,T., Miwa,H., and Matsumoto,T. (2005). Association between IL-18 gene promoter polymorphisms and inflammatory bowel disease in a Japanese population. Inflamm. Bowel. Dis. *11*, 1038-1043.

Tamura,K., Fukuda,Y., Sashio,H., Takeda,N., Bamba,H., Kosaka,T., Fukui,S., Sawada,K., Tamura,K., Satomi,M., Yamada,T., Yamamura,T., Yamamoto,Y., Furuyama,J., Okamura,H., and Shimoyama,T. (2002). IL18 polymorphism is associated with an increased risk of Crohn's disease. J. Gastroenterol. *37 Suppl 14*, 111-116.

Taylor,S.L., Renshaw,B.R., Garka,K.E., Smith,D.E., and Sims,J.E. (2002). Genomic organization of the interleukin-1 locus. Genomics *79*, 726-733.

Tazi-Ahnini,R., Cox,A., McDonagh,A.J., Nicklin,M.J., di Giovine,F.S., Timms,J.M., Messenger,A.G., Dimitropoulou,P., Duff,G.W., and Cork,M.J. (2002). Genetic analysis of the interleukin-1 receptor antagonist and its homologue IL-1L1 in alopecia areata: strong severity association and possible gene interaction. Eur. J. Immunogenet. *29*, 25-30.

Templeton,A.R. (2006). Modeling Evolution and the Hardy-Weinberg Law. In Population Genetics and Microevolutionary Theory, (Hoboken: Wiley), pp. 21-48.

Ten Hove,T., Corbaz,A., Amitai,H., Aloni,S., Belzer,I., Graber,P., Drillenburg,P., van Deventer,S.J., Chvatchko,Y., and Te Velde,A.A. (2001). Blockade of endogenous IL-18 ameliorates TNBS-induced colitis by decreasing local TNF-alpha production in mice. Gastroenterology *121*, 1372-1379.

The International HapMap Consortium (2003). The International HapMap Project. Nature *426*, 789-796.

The International HapMap Consortium (2005). A haplotype map of the human genome. Nature *437*, 1299-1320.

The International HapMap Consortium (2007). A second generation human haplotype map of over 3.1 million SNPs. Nature *449*, 851-861.

Thomas,G., Moffatt,P., Salois,P., Gaumond,M.H., Gingras,R., Godin,E., Miao,D., Goltzman,D., and Lanctot,C. (2003). Osteocrin, a novel bone-specific secreted protein that modulates the osteoblast phenotype. J. Biol. Chem. *278*, 50563-50571.

Thomassen,E., Renshaw,B.R., and Sims,J.E. (1999). Identification and characterization of SIGIRR, a molecule representing a novel subtype of the IL-1R superfamily. Cytokine *11*, 389-399.

Thompson,S.R., Novick,D., Stock,C.J., Sanders,J., Brull,D., Cooper,J., Woo,P., Miller,G., Rubinstein,M., and Humphries,S.E. (2007). Free Interleukin (IL)-18 levels, and the impact of

IL18 and IL18BP genetic variation, in CHD patients and healthy men. Arterioscler. Thromb. Vasc. Biol. *27*, 2743-2749.

Thomson,W., Barrett,J.H., Donn,R., Pepper,L., Kennedy,L.J., Ollier,W.E., Silman,A.J., Woo,P., and Southwood,T. (2002). Juvenile idiopathic arthritis classified by the ILAR criteria: HLA associations in UK patients. Rheumatology. (Oxford) *41*, 1183-1189.

Thornberry,N.A., Bull,H.G., Calaycay,J.R., Chapman,K.T., Howard,A.D., Kostura,M.J., Miller,D.K., Molineaux,S.M., Weidner,J.R., Aunins,J., and . (1992). A novel heterodimeric cysteine protease is required for interleukin-1 beta processing in monocytes. Nature *356*, 768-774.

Timms,A.E., Crane,A.M., Sims,A.M., Cordell,H.J., Bradbury,L.A., Abbott,A., Coyne,M.R., Beynon,O., Herzberg,I., Duff,G.W., Calin,A., Cardon,L.R., Wordsworth,B.P., and Brown,M.A. (2004). The interleukin 1 gene cluster contains a major susceptibility locus for ankylosing spondylitis. Am. J. Hum. Genet. *75*, 587-595.

Tiret,L., Godefroy,T., Lubos,E., Nicaud,V., Tregouet,D.A., Barbaux,S., Schnabel,R., Bickel,C., Espinola-Klein,C., Poirier,O., Perret,C., Munzel,T., Rupprecht,H.J., Lackner,K., Cambien,F., and Blankenberg,S. (2005a). Genetic analysis of the interleukin-18 system highlights the role of the interleukin-18 gene in cardiovascular disease. Circulation *112*, 643-650.

Tiret,L., Godefroy,T., Lubos,E., Nicaud,V., Tregouet,D.A., Barbaux,S., Schnabel,R., Bickel,C., Espinola-Klein,C., Poirier,O., Perret,C., Munzel,T., Rupprecht,H.J., Lackner,K., Cambien,F., and Blankenberg,S. (2005b). Genetic analysis of the interleukin-18 system highlights the role of the interleukin-18 gene in cardiovascular disease. Circulation *112*, 643-650.

Tolusso,B., Pietrapertosa,D., Morelli,A., De Santis,M., Gremese,E., Farina,G., Carniello,S.G., Del Frate,M., and Ferraccioli,G. (2006). IL-1B and IL-1RN gene polymorphisms in rheumatoid arthritis: relationship with protein plasma levels and response to therapy. Pharmacogenomics. *7*, 683-695.

Torigoe,K., Ushio,S., Okura,T., Kobayashi,S., Taniai,M., Kunikata,T., Murakami,T., Sanou,O., Kojima,H., Fujii,M., Ohta,T., Ikeda,M., Ikegami,H., and Kurimoto,M. (1997). Purification and characterization of the human interleukin-18 receptor. J. Biol. Chem. *272*, 25737-25742.

Towne,J.E., Garka,K.E., Renshaw,B.R., Virca,G.D., and Sims,J.E. (2004). Interleukin (IL)-1F6, IL-1F8, and IL-1F9 signal through IL-1Rrp2 and IL-1RAcP to activate the pathway leading to NF-kappaB and MAPKs. J. Biol. Chem. *279*, 13677-13688.

Turner,D.M., Williams,D.M., Sankaran,D., Lazarus,M., Sinnott,P.J., and Hutchinson,I.V. (1997). An investigation of polymorphism in the interleukin-10 gene promoter. Eur. J. Immunogenet. *24*, 1-8.

Ushio,S., Namba,M., Okura,T., Hattori,K., Nukada,Y., Akita,K., Tanabe,F., Konishi,K., Micallef,M., Fujii,M., Torigoe,K., Tanimoto,T., Fukuda,S., Ikeda,M., Okamura,H., and Kurimoto,M. (1996a). Cloning of the cDNA for human IFN-gamma-inducing factor,

expression in Escherichia coli, and studies on the biologic activities of the protein. J. Immunol. *156*, 4274-4279.

Ushio,S., Namba,M., Okura,T., Hattori,K., Nukada,Y., Akita,K., Tanabe,F., Konishi,K., Micallef,M., Fujii,M., Torigoe,K., Tanimoto,T., Fukuda,S., Ikeda,M., Okamura,H., and Kurimoto,M. (1996c). Cloning of the cDNA for human IFN-gamma-inducing factor, expression in Escherichia coli, and studies on the biologic activities of the protein. J. Immunol. *156*, 4274-4279.

Ushio,S., Namba,M., Okura,T., Hattori,K., Nukada,Y., Akita,K., Tanabe,F., Konishi,K., Micallef,M., Fujii,M., Torigoe,K., Tanimoto,T., Fukuda,S., Ikeda,M., Okamura,H., and Kurimoto,M. (1996b). Cloning of the cDNA for human IFN-gamma-inducing factor, expression in Escherichia coli, and studies on the biologic activities of the protein. J. Immunol. *156*, 4274-4279.

Uziel,Y., Pomeranz,A., Brik,R., Navon,P., Mukamel,M., Press,J., Barash,J., Tauber,T., Harel,L., Virgilis,D., Bibi,H., Heldenberg,D., and Wolach,B. (1999). Seasonal variation in systemic onset juvenile rheumatoid arthritis in Israel. J. Rheumatol. *26*, 1187-1189.

van der,P.M., Crusius,J.B., Garcia-Gonzalez,M.A., Baudoin,P., Kostense,P.J., Alizadeh,B.Z., Dijkmans,B.A., Pena,A.S., and van der Horst-Bruinsma IE (2002). Interleukin-1beta and interleukin-1 receptor antagonist gene polymorphisms in ankylosing spondylitis. Rheumatology. (Oxford) *41*, 1419-1423.

Veenstra,K.G., Jonak,Z.L., Trulli,S., and Gollob,J.A. (2002). IL-12 induces monocyte IL-18 binding protein expression via IFN-gamma. J. Immunol. *168*, 2282-2287.

Vencovsky,J., Jarosova,K., Ruzickova,S., Nemcova,D., Niederlova,J., Ozen,S., Alikasifoglu,M., Bakkaloglu,A., Ollier,W.E., and Mageed,R.A. (2001). Higher frequency of allele 2 of the interleukin-1 receptor antagonist gene in patients with juvenile idiopathic arthritis. Arthritis Rheum. *44*, 2387-2391.

Venter,J.C., Adams,M.D., Myers,E.W., Li,P.W., Mural,R.J., Sutton,G.G., Smith,H.O., Yandell,M., Evans,C.A., Holt,R.A., Gocayne,J.D., Amanatides,P., Ballew,R.M., Huson,D.H., Wortman,J.R., Zhang,Q., Kodira,C.D., Zheng,X.H., Chen,L., Skupski,M., Subramanian,G., Thomas,P.D., Zhang,J., Gabor Miklos,G.L., Nelson,C., Broder,S., Clark,A.G., Nadeau,J., McKusick,V.A., Zinder,N., Levine,A.J., Roberts,R.J., Simon,M., Slayman,C., Hunkapiller,M., Bolanos,R., Delcher,A., Dew,I., Fasulo,D., Flanigan,M., Florea,L., Halpern,A., Hannenhalli,S., Kravitz,S., Levy,S., Mobarry,C., Reinert,K., Remington,K., Abu-Threideh,J., Beasley,E., Biddick,K., Bonazzi,V., Brandon,R., Cargill,M., Chandramouliswaran,I., Charlab,R., Chaturvedi,K., Deng,Z., Di,F., V, Dunn,P., Eilbeck,K., Evangelista,C., Gabrielian,A.E., Gan,W., Ge,W., Gong,F., Gu,Z., Guan,P., Heiman,T.J., Higgins,M.E., Ji,R.R., Ke,Z., Ketchum,K.A., Lai,Z., Lei,Y., Li,Z., Li,J., Liang,Y., Lin,X., Lu,F., Merkulov,G.V., Milshina,N., Moore,H.M., Naik,A.K., Narayan,V.A., Neelam,B., Nusskern,D., Rusch,D.B., Salzberg,S., Shao,W., Shue,B., Sun,J., Wang,Z., Wang,A., Wang,X., Wang,J., Wei,M., Wides,R., Xiao,C., Yan,C., Yao,A., Ye,J., Zhan,M., Zhang,W., Zhang,H., Zhao,Q., Zheng,L., Zhong,F., Zhong,W., Zhu,S., Zhao,S., Gilbert,D., Baumhueter,S., Spier,G., Carter,C., Cravchik,A., Woodage,T., Ali,F., An,H., Awe,A., Baldwin,D., Baden,H., Barnstead,M., Barrow,I., Beeson,K., Busam,D., Carver,A., Center,A., Cheng,M.L., Curry,L., Danaher,S., Davenport,L., Desilets,R., Dietz,S., Dodson,K., Doup,L., Ferriera,S., Garg,N., Gluecksmann,A., Hart,B., Haynes,J., Haynes,C., Heiner,C., Hladun,S.,

Hostin,D., Houck,J., Howland,T., Ibegwam,C., Johnson,J., Kalush,F., Kline,L., Koduru,S., Love,A., Mann,F., May,D., McCawley,S., McIntosh,T., McMullen,I., Moy,M., Moy,L., Murphy,B., Nelson,K., Pfannkoch,C., Pratts,E., Puri,V., Qureshi,H., Reardon,M., Rodriguez,R., Rogers,Y.H., Romblad,D., Ruhfel,B., Scott,R., Sitter,C., Smallwood,M., Stewart,E., Strong,R., Suh,E., Thomas,R., Tint,N.N., Tse,S., Vech,C., Wang,G., Wetter,J., Williams,S., Williams,M., Windsor,S., Winn-Deen,E., Wolfe,K., Zaveri,J., Zaveri,K., Abril,J.F., Guigo,R., Campbell,M.J., Sjolander,K.V., Karlak,B., Kejariwal,A., Mi,H., Lazareva,B., Hatton,T., Narechania,A., Diemer,K., Muruganujan,A., Guo,N., Sato,S., Bafna,V., Istrail,S., Lippert,R., Schwartz,R., Walenz,B., Yooseph,S., Allen,D., Basu,A., Baxendale,J., Blick,L., Caminha,M., Carnes-Stine,J., Caulk,P., Chiang,Y.H., Coyne,M., Dahlke,C., Mays,A., Dombroski,M., Donnelly,M., Ely,D., Esparham,S., Fosler,C., Gire,H., Glanowski,S., Glasser,K., Glodek,A., Gorokhov,M., Graham,K., Gropman,B., Harris,M., Heil,J., Henderson,S., Hoover,J., Jennings,D., Jordan,C., Jordan,J., Kasha,J., Kagan,L., Kraft,C., Levitsky,A., Lewis,M., Liu,X., Lopez,J., Ma,D., Majoros,W., McDaniel,J., Murphy,S., Newman,M., Nguyen,T., Nguyen,N., and Nodell,M. (2001). The sequence of the human genome. Science *291*, 1304-1351.

Verbsky,J.W. and White,A.J. (2004). Effective use of the recombinant interleukin 1 receptor antagonist anakinra in therapy resistant systemic onset juvenile rheumatoid arthritis. J. Rheumatol. *31*, 2071-2075.

Vigers,G.P., Anderson,L.J., Caffes,P., and Brandhuber,B.J. (1997). Crystal structure of the type-I interleukin-1 receptor complexed with interleukin-1beta. Nature *386*, 190-194.

Villani,A.C., Lemire,M., Fortin,G., Louis,E., Silverberg,M.S., Collette,C., Baba,N., Libioulle,C., Belaiche,J., Bitton,A., Gaudet,D., Cohen,A., Langelier,D., Fortin,P.R., Wither,J.E., Sarfati,M., Rutgeerts,P., Rioux,J.D., Vermeire,S., Hudson,T.J., and Franchimont,D. (2009). Common variants in the NLRP3 region contribute to Crohn's disease susceptibility. Nat. Genet. *41*, 71-76.

Voiculescu,C.L., Rosu,L., and Rogoz,S. (1988). Modulation of mouse spleen natural killer (NK) cell activity by beta-interferon, interleukin-1, and prostaglandins. Lymphology *21*, 144-151.

Voisey,J., Swagell,C.D., Hughes,I.P., Lawford,B.R., Young,R.M., and Morris,C.P. (2010). Analysis of HapMap tag-SNPs in dysbindin (DTNBP1) reveals evidence of consistent association with schizophrenia. Eur. Psychiatry.

Walker,N.P., Talanian,R.V., Brady,K.D., Dang,L.C., Bump,N.J., Ferenz,C.R., Franklin,S., Ghayur,T., Hackett,M.C., Hammill,L.D., and . (1994). Crystal structure of the cysteine protease interleukin-1 beta-converting enzyme: a (p20/p10)2 homodimer. Cell *78*, 343-352.

Wallace,C.A. and Levinson,J.E. (1991). Juvenile rheumatoid arthritis: outcome and treatment for the 1990s. Rheum. Dis. Clin. North Am. *17*, 891-905.

Wallace,C.A. and Sherry,D.D. (1997). Trial of intravenous pulse cyclophosphamide and methylprednisolone in the treatment of severe systemic-onset juvenile rheumatoid arthritis. Arthritis Rheum. *40*, 1852-1855.

Wang,C., Deng,L., Hong,M., Akkaraju,G.R., Inoue,J., and Chen,Z.J. (2001). TAK1 is a ubiquitin-dependent kinase of MKK and IKK. Nature *412*, 346-351.

Wannamaker,W., Davies,R., Namchuk,M., Pollard,J., Ford,P., Ku,G., Decker,C., Charifson,P., Weber,P., Germann,U.A., Kuida,K., and Randle,J.C. (2007). (S)-1-((S)-2-{[1-(4-amino-3-chloro-phenyl)-methanoyl]-amino}-3,3-dimethyl- butanoyl)-pyrrolidine-2-carboxylic acid ((2R,3S)-2-ethoxy-5-oxo-tetrahydro-furan-3-yl)-amide (VX-765), an orally available selective interleukin (IL)-converting enzyme/caspase-1 inhibitor, exhibits potent anti-inflammatory activities by inhibiting the release of IL-1beta and IL-18. J. Pharmacol. Exp. Ther. *321*, 509-516.

Wasserman,W.W., Palumbo,M., Thompson,W., Fickett,J.W., and Lawrence,C.E. (2000). Human-mouse genome comparisons to locate regulatory sites. Nat. Genet. *26*, 225-228.

Watanabe,N. and Kobayashi,Y. (1994). Selective release of a processed form of interleukin 1 alpha. Cytokine *6*, 597-601.

Wei,X.Q., Leung,B.P., Arthur,H.M., McInnes,I.B., and Liew,F.Y. (2001). Reduced incidence and severity of collagen-induced arthritis in mice lacking IL-18. J. Immunol. *166*, 517-521.

Wellcome Trust Case Control Consortium (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature *447*, 661-678.

Wesa,A. and Galy,A. (2002). Increased production of pro-inflammatory cytokines and enhanced T cell responses after activation of human dendritic cells with IL-1 and CD40 ligand. BMC. Immunol. *3*, 14.

Wesche,H., Henzel,W.J., Shillinglaw,W., Li,S., and Cao,Z. (1997a). MyD88: an adapter that recruits IRAK to the IL-1 receptor complex. Immunity *7*, 837-847.

Wesche,H., Korherr,C., Kracht,M., Falk,W., Resch,K., and Martin,M.U. (1997b). The interleukin-1 receptor accessory protein (IL-1RAcP) is essential for IL-1-induced activation of interleukin-1 receptor-associated kinase (IRAK) and stress-activated protein kinases (SAP kinases). J. Biol. Chem. *272*, 7727-7731.

Wesche,H., Resch,K., and Martin,M.U. (1998). Effects of IL-1 receptor accessory protein on IL-1 binding. FEBS Lett. *429*, 303-306.

Williams,R.R., Skolnick,M., Carmelli,D., Maness,A.T., Hunt,S.C., Hasstedt,S., Reiber,G.E., and Jones,R.K. (1979). Utah pedigree studies: design and preliminary data for premature male CHD deaths. Prog. Clin. Biol. Res. *32*, 711-729.

Wilson,K.P., Black,J.A., Thomson,J.A., Kim,E.E., Griffith,J.P., Navia,M.A., Murcko,M.A., Chambers,S.P., Aldape,R.A., Raybuck,S.A., and . (1994a). Structure and mechanism of interleukin-1 beta converting enzyme. Nature *370*, 270-275.

Wilson,K.P., Black,J.A., Thomson,J.A., Kim,E.E., Griffith,J.P., Navia,M.A., Murcko,M.A., Chambers,S.P., Aldape,R.A., Raybuck,S.A., and . (1994b). Structure and mechanism of interleukin-1 beta converting enzyme. Nature *370*, 270-275.

Wingender,E. (1988). Compilation of transcription regulating proteins. Nucleic Acids Res. *16*, 1879-1902.

Wingender,E., Chen,X., Hehl,R., Karas,H., Liebich,I., Matys,V., Meinhardt,T., Pruss,M., Reuter,I., and Schacherer,F. (2000). TRANSFAC: an integrated system for gene expression regulation. Nucleic Acids Res. *28*, 316-319.

Woo,P. (2006). Systemic juvenile idiopathic arthritis: diagnosis, management, and outcome. Nat. Clin. Pract. Rheumatol. *2*, 28-34.

Woo,P. and Colbert,R.A. (2009). An overview of genetics of paediatric rheumatic diseases. Best. Pract. Res. Clin. Rheumatol. *23*, 589-597.

Woo,P., Laxer,R.M., and Sherry D.D (2007). Juvenile Idiopathic Arthritis (JIA). In Pediatric Rheumatology in Clinical Practice, (London: Springer), pp. 23-46.

Woo,P., Southwood,T.R., Prieur,A.M., Dore,C.J., Grainger,J., David,J., Ryder,C., Hasson,N., Hall,A., and Lemelle,I. (2000). Randomized, placebo-controlled, crossover trial of low-dose oral methotrexate in children with extended oligoarticular or systemic arthritis. Arthritis Rheum. *43*, 1849-1857.

Woo,P., Wilkinson,N., Prieur,A.M., Southwood,T., Leone,V., Livermore,P., Wythe,H., Thomson,D., and Kishimoto,T. (2005). Open label phase II trial of single, ascending doses of MRA in Caucasian children with severe systemic juvenile idiopathic arthritis: proof of principle of the efficacy of IL-6 receptor blockade in this type of arthritis and demonstration of prolonged clinical improvement. Arthritis Res. Ther. *7*, R1281-R1288.

Woolfrey,A., Storek,J., Bowyer,S., Nelson,R., Robertson,M., and Wallace,C. (2010). Long-term response of juvenile idiopathic arthritis after conditioning with 8 Gy total body irradiation followed by autologous peripheral blood stem cells: case report. Pediatr. Transplant. *14*, E65-E69.

Wulffraat,N., van Royen,A., Bierings,M., Vossen,J., and Kuis,W. (1999). Autologous haemopoietic stem-cell transplantation in four patients with refractory juvenile chronic arthritis. Lancet *353*, 550-553.

Xu,M. and He,L. (2010). Convergent evidence shows a positive association of interleukin-1 gene complex locus with susceptibility to schizophrenia in the Caucasian population. Schizophr. Res. *120*, 131-142.

Xu,Q., Tin,S.K., Sivalingam,S.P., Thumboo,J., Koh,D.R., and Fong,K.Y. (2007b). Interleukin-18 promoter gene polymorphisms in Chinese patients with systemic lupus erythematosus: association with CC genotype at position -607. Ann. Acad. Med. Singapore *36*, 91-95.

Xu,Q., Tin,S.K., Sivalingam,S.P., Thumboo,J., Koh,D.R., and Fong,K.Y. (2007a). Interleukin-18 promoter gene polymorphisms in Chinese patients with systemic lupus erythematosus: association with CC genotype at position -607. Ann. Acad. Med. Singapore *36*, 91-95.

Xue,J., Mou,L., Lu,X., and Wu,H. (2010). No association of interleukin-1 receptor antagonist VNTR polymorphism and rheumatoid arthritis susceptibility: a meta analysis. Clin. Exp. Rheumatol.

Yang,Y., Qiao,J., Tang,R.X., and Li,M.Z. (2010). Genotype and haplotype determination of interleukin (IL) 1 beta (g. -511C>T and g. +3954C>T) and IL-1RN in polycystic ovary syndrome. Fertil. Steril. *94*, 384-386.

Yildirim,E. and Birnbaumer,L. (2007). TRPC2: molecular biology and functional importance. Handb. Exp. Pharmacol. 53-75.

Yokota, S., Imagawa, T., Mori, M., Takei, S., Kawano, Y., Iwata, N., Tomiita, M., Miyoshi, M., Aihara, Y., Murata, T., Abukawa, D., Nishimoto, N., and Kishimoto, T. Rapid improvement of signs and symptoms associated with systemic Juvenile Idiopathic Arthritis (sJIA) by Interleukin-6 (IL-6) blockade - Results from a Tocilizumab sJIA Phase III clinical trial. Arthritis and Rheumatism 54[Supplemental 9]. 2006.
Ref Type: Abstract

Yokota,S., Miyamae,T., Imagawa,T., Iwata,N., Katakura,S., Mori,M., Woo,P., Nishimoto,N., Yoshizaki,K., and Kishimoto,T. (2005). Therapeutic efficacy of humanized recombinant anti-interleukin-6 receptor antibody in children with systemic-onset juvenile idiopathic arthritis. Arthritis Rheum. *52*, 818-825.

Yoshimoto,T., Takeda,K., Tanaka,T., Ohkusu,K., Kashiwamura,S., Okamura,H., Akira,S., and Nakanishi,K. (1998). IL-12 up-regulates IL-18 receptor expression on T cells, Th1 cells, and B cells: synergism with IL-18 for IFN-gamma production. J. Immunol. *161*, 3400-3407.

You,C.G., Li,J.F., Xie,X.D., Zhu,Y., Li,P.Q., and Chen,Y.R. (2007). Association of interleukin-1 genetic polymorphisms with the risk of rheumatoid arthritis in Chinese population. Clin. Chem. Lab Med. *45*, 968-971.

Yu,Y., Ge,N., Xie,M., Sun,W., Burlingame,S., Pass,A.K., Nuchtern,J.G., Zhang,D., Fu,S., Schneider,M.D., Fan,J., and Yang,J. (2008). Phosphorylation of Thr-178 and Thr-184 in the TAK1 T-loop is required for interleukin (IL)-1-mediated optimal NFkappaB and AP-1 activation as well as IL-6 gene expression. J. Biol. Chem. *283*, 24497-24505.

Zapico,I., Coto,E., Rodriguez,A., Alvarez,C., Torre,J.C., and Alvarez,V. (2000). CCR5 (chemokine receptor-5) DNA-polymorphism influences the severity of rheumatoid arthritis. Genes Immun. *1*, 288-289.

Zeng,C. (2000). NuMA: a nuclear protein involved in mitotic centrosome function. Microsc. Res. Tech. *49*, 467-477.

Zhang,F., Sha,J., Wood,T.G., Galindo,C.L., Garner,H.R., Burkart,M.F., Suarez,G., Sierra,J.C., Agar,S.L., Peterson,J.W., and Chopra,A.K. (2008a). Alteration in the activation state of new inflammation-associated targets by phospholipase A2-activating protein (PLAA). Cell Signal. *20*, 844-861.

Zhang,K., Biroschak,J., Glass,D.N., Thompson,S.D., Finkel,T., Passo,M.H., Binstadt,B.A., Filipovich,A., and Grom,A.A. (2008b). Macrophage activation syndrome in patients with systemic juvenile idiopathic arthritis is associated with MUNC13-4 polymorphisms. Arthritis Rheum. *58*, 2892-2896.

Zheng,H., Fletcher,D., Kozak,W., Jiang,M., Hofmann,K.J., Conn,C.A., Soszynski,D., Grabiec,C., Trumbauer,M.E., Shaw,A., and . (1995). Resistance to fever induction and impaired acute-phase response in interleukin-1 beta-deficient mice. Immunity *3*, 9-19.

Zhernakova,A., Festen,E.M., Franke,L., Trynka,G., van Diemen,C.C., Monsuur,A.J., Bevova,M., Nijmeijer,R.M., van 't,S.R., Heijmans,R., Boezen,H.M., van Heel,D.A., van Bodegraven,A.A., Stokkers,P.C., Wijmenga,C., Crusius,J.B., and Weersma,R.K. (2008). Genetic analysis of innate immunity in Crohn's disease and ulcerative colitis identifies two susceptibility loci harboring CARD9 and IL18RAP. Am. J. Hum. Genet. *82*, 1202-1210.

Appendix 1. Golden Gate genotyping platform

A. Methodology

The Illumina Sentrix microarray is an array of arrays. Each matrix is made up of 96 bundles, one for each sample, of 49,777 fibre optic strands (Figure 3.4, Image 7). Each strand has a chemically etched well in its end containing an individual 3-micron silica bead. Each of these beads is coated with, on average, in the order of 10^6 copies of a single-strand ~25-mer nucleotide sequence called an address sequence. There is a unique address sequence assigned to each SNP included in the array. In a separate reaction for each address type these oligonucleotides are hybridised to beads. The reactions are quantitatively mixed creating a bead pool containing beads of each address type. The bead pool is then randomly self-assembled into the wells of the array bundles. Each array bundle contains approximately 30 of each bead type, acting as replicates for each SNP. For each SNP studied there is a locus specific oligonucleotide (LSO), which binds to a unique sequence downstream of the SNP. Each LSO also contains an address sequence, complimentary to that coated on the array beads, downstream of the genomic complimentary sequence. For each SNP in the array there are also two allele specific oligonucleotides (ASOs), binding to the sequence surrounding the SNP. Each of these three oligonucleotide types also contains a different universal primer sequence.

Genomic DNA is biotinylated and immobilised on streptavidin-coated paramagnetic particles. An array specific pool of oligonucleotides, containing LSOs and ASOs, is then hybridised to the genomic DNA. Using a magnetic plate to capture the beads, to which the DNA-oligonucleotide complexes are bound, allows for stringent washes to be performed and all un-bound and mis-hybridised oligonucleotides to be removed. An extension-ligation step is performed in which the ASOs are extended to the LSOs and the two are ligated together. This forms a full-length oligonucleotide containing both SNP and allele information. Following further washes the genomic DNA and full length oligonucleotides are separated by denaturation and the non-magnetically captured full length oligonucleotides removed. These then undergo PCR amplification using the universal primer sequences contained in each ASO and LSO. The primers corresponding to those in the ASOs are Cy3 and Cy5 labelled according to allele. The dye labelled primers are also biotinylated, enabling the dye labelled strands to be captured on paragmagnetic particles, isolated from the PCR reaction, and

hybridised to the bead arrays. During hybridisation the LSO section of each full length oligonucleotide binds to the corresponding address sequence on the beads.

During assembly of the array, a map of the locations of each bead type is created by the manufacturer, enabling the results from each bead to be assigned to the corresponding SNP. This map is created through sequential hybridisations and de-hybridisations of dye-labelled decoder nucleotide sequences, complimentary to the bead address sequences, to determine the address sequence on each individual bead (Gunderson et al., 2004). This array decoding process also validates the performance of each bead as binding efficiency is tested through the signal intensity. Any fibre optic strands not containing a bead, and beads not showing high signal intensity are marked as blank on the map, and not used during genotype determination.

The fluorescence intensity at both wavelengths from each bead is read and interpreted as a genotype. Homozygosity will be shown by fluorescence at only 570nm (visualised as red) for allele 'A' and (670nm) (visualised as green) for allele 'B'. Heterozygosity will be shown by fluorescence of equal intensity at both wavelengths (visualised as yellow).

B. Genotyping Protocol

Prior to each centrifugation or vortex step plates were sealed with either adhesive film, or where appropriate, heat-sealing foil. All pulse-centrifugation steps were at 250Xg, and unless otherwise stated, all steps were at room temperature. Magnetic bead capture was performed by placing the plate onto a raised bar plate magnet for two minutes until all beads were captured.

To biotinylate the genomic DNA 5µl biotin buffer (MS1) was added to 5µl 50ng/µl DNA. The plate was pulse-centrifuged, vortexed at 2,300rpm for 20 seconds, pulse-centrifuged, heated to 95°C for 30 minutes, and pulse-centrifuged again. 5µl precipitation solution (PS1) was added to each well, the plate pulse-centrifuged, and vortexed at 2,300 rpm for 20 seconds. 15µl 2-propanol was added to each well, the plate vortexed at 1,600 rpm for 20 seconds and centrifuged for 20 minutes at 3,000Xg. After removal of the supernatant the plate was centrifuged upside-down on an absorbent pad at 8Xg for one minute and left to dry for 15 minutes. The activated DNA was re-suspended by addition of 10µl re-suspension solution (RS1), pulse-centrifugation, and vortexing at 2,300 rpm for one minute. 10µl SNP-specific primer pool (OPA), and 30µl streptavidin coated magnetic bead solution, was added to each well of a new plate. Following pulse-centrifugation 10µl activated DNA was transferred to the new plate. This was pulse-centrifuged, vortexed for one minute at 1,600 rpm, and placed on a 70°C heat block. The heat block was then immediately set to 30°C and the plate allowed to cool.

The plate was pulse-centrifuged and the magnetic beads captured. Without disturbing the beads the supernatant was discarded, and 50µl wash buffer (AM1) added to each well. To re-suspend the beads the plate was vortexed at 1,600 rpm for 20 seconds. The bead capture, supernatant removal, and wash was repeated. Two further washes were performed as previously with wash buffer UB1. To each well 37µl primer extension reaction buffer (MEL) was added, vortexed at 1,600 rpm for one minute, and incubated at 45°C for 15 minutes.

The PCR reaction mixture was prepared by adding 64µl Titanium *Taq* DNA polymerase to the MMP buffer and 30µl added to each well of a new plate and pulse-centrifuged. All magnetic beads were captured and the supernatant removed. 50µl UB1 was added and removed after a few minutes. To each well 35µl IP1 buffer was added and vortexed at 1,800

rpm for one minute to resuspend the beads. The plate was heated at 95°C for one minute to separate the oligos from the DNA. The beads were captured and 30µl supernatant transferred to the plate containing the PCR reaction mixture. The samples underwent PCR amplification at: 37°C for ten minutes, 95°C for three minutes, 34 cycles of: 95°C for 35 seconds, 56°C for 35 seconds, and 72°C for two minutes, followed by ten minutes at 72°C and five minutes at 4°C.

The plate was pulse-centrifuged and 20µl paramagnetic particle containing buffer (MPB) added to each well. The solution was mixed by pipetting and transferred to the filter plate. The plate was protected from light and left for one hour to allow capture of the biotinylated dye-labelled strands to the paramagnetic particles. The filter plate was placed on a new plate and centrifuged at 1,000 Xg for five minutes. 50µl UB2 buffer was added to each well and the plate centrifuged at 1,000 Xg for five minutes.

To each well of a new plate 30µl hybridisation buffer (MH1) was added and the filter plate placed on top so that the wells matched up. 30µl NaOH was added to the wells of the filter plate and centrifuged at 1,000 Xg for five minutes to elute the amplicons. The filter plate containing the paragmatnetic particles was discarded. The contents of the collection plate were mixed gently and 50µl transferred, in the 96 well layout, to a 384 well plate. The remaining wells of this plate were filled with 30µl UB2 in order to maintain the humidity during the hybridisation step. The plate was then centrifuged at 3,000 Xg for four minutes. To prepare the sentrix array matrix prior to hybridisation of the labelled oligios it was placed fibre bundles down in an OmniTray containing 70ml UB2 and gently agitated to remove any bubbles from the bottom of the arrays. It was left in the UB2 for three minutes, transferred to an OmniTray containing 60ml NaOH for 30 seconds, then placed back into UB2 for 30 seconds to neutralise the NaOH. The array was then placed into the labelled oligo containing plate so that each fibre bundle was in the corresponding well. The plate and array were secured together with a hybridisation cartridge and placed in a hybridisation oven at 60°C. After 30 minutes the oven was turned down to 45°C and the array left overnight to hybridise. The array was cleaned by placing in a tray containing 70ml UB2 with agitation for one minute. It was placed into a fresh tray of UB2 and agitated for a further minute, and then placed in a tray containing 70ml of the image wash buffer (IS1) for five minutes. The array was placed face up and left to air dry for 20 minutes prior to scanning by an Illumina BeadStation 500G.

Appendix 2. Genotyping assay results of Stage-1 SNPs

rs N^{o}	GenTrain Score	GC10	MAF	Call Freq. (%)
rs10157379	0.810	0.810	0.366	99.6
rs10167431	0.878	0.859	0.460	98.6
rs10206428	0.739	0.739	0.394	100
rs10207579	0.718	0.707	0.473	99.3
rs1024935	0.926	0.926	0.143	100
rs1030021	0.856	0.856	0.280	99.6
rs1035130	0.859	0.859	0.291	100
rs1035131	0.749	0.749	0.366	100
rs10399826	0.912	0.912	0.141	100
rs1041973	0.749	0.749	0.242	100
rs10490571	0.580	0.577	0.347	99.6
rs10490818	0.746	0.746	0.220	100
rs10502045	0.899	0.899	0.249	99.3
rs10513852	0.838	0.838	0.166	100
rs10513854	0.892	0.892	0.173	100
rs10513857	0.816	0.816	0.162	100
rs10513861	0.875	0.875	0.058	100
rs10737805	0.909	0.909	0.137	100
rs10754555	0.715	0.715	0.365	99.6
rs10802486	0.734	0.734	0.296	100
rs10891319	0.736	0.736	0.368	99.6
rs10891323	0.807	0.807	0.103	99.6
rs10895775	0.921	0.921	0.408	100
rs10924989	0.824	0.824	0.365	100
rs10924990	0.664	0.663	0.439	98.2
rs10925002	0.908	0.908	0.395	99.6
rs10925006	0.916	0.916	0.256	100
rs10937439	0.954	0.954	0.399	100
rs10937457	0.842	0.842	0.307	100
rs11123915	0.716	0.716	0.397	97.8
rs11214093	0.821	0.821	0.379	99.6
rs11226613	0.911	0.911	0.222	100
rs1143627	0.707	0.707	0.283	100
rs1143634	0.763	0.763	0.222	100
rs11465567	0.807	0.807	0.087	100
rs11465572	0.824	0.824	0.087	100
rs11465597	0.761	0.761	0.096	100
rs1159213	0.952	0.952	0.361	100
rs11604437	0.815	0.815	0.049	100
rs11677452	0.587	0.587	0.199	99.6
rs11678651	0.905	0.905	0.350	100
rs11692230	0.742	0.742	0.415	100
rs11695148	0.798	0.798	0.168	100
rs11916404	0.919	0.918	0.229	100
rs11922372	0.687	0.687	0.217	99.6
rs12065526	0.839	0.839	0.103	99.3
rs12082329	0.703	0.703	0.486	100
rs12468224	0.808	0.808	0.412	100
rs12469600	0.874	0.874	0.253	100

rs Nº	GenTrain Score	GC10	MAF	Call Freq. (%)
rs12469892	0.706	0.706	0.144	100
rs12472089	0.768	0.768	0.258	100
rs12564791	0.607	0.607	0.085	99.3
rs12634559	0.808	0.808	0.321	98.6
rs12711747	0.773	0.773	0.453	100
rs12711749	0.811	0.811	0.386	100
rs12712122	0.827	0.827	0.334	100
rs12797863	0.922	0.922	0.065	100
rs12995447	0.780	0.780	0.424	100
rs12996377	0.698	0.698	0.132	100
rs13005572	0.859	0.859	0.199	100
rs13014084	0.484	0.484	0.047	100
rs13033104	0.841	0.841	0.253	100
rs13060312	0.733	0.733	0.116	100
rs13070253	0.919	0.919	0.404	99.6
rs13314422	0.892	0.848	0.058	99.6
rs13317447	0.802	0.762	0.049	100
rs13392494	0.823	0.823	0.171	100
rs1345301	0.717	0.717	0.449	100
rs1355365	0.751	0.748	0.289	99.3
rs1403550	0.811	0.811	0.240	100
rs1420092	0.897	0.894	0.13	100
rs1420095	0.759	0.759	0.072	99.6
rs1420103	0.737	0.737	0.224	100
rs1446521	0.749	0.749	0.318	99.3
rs1472597	0.849	0.849	0.177	100
rs1476984	0.704	0.703	0.097	100
rs1501602	0.848	0.848	0.134	100
rs1503392	0.722	0.722	0.338	100
rs1503394	0.907	0.907	0.271	100
rs1510901	0.795	0.795	0.439	99.6
rs1539019	0.909	0.909	0.359	99.6
rs1541304	0.897	0.852	0.025	100
rs1541434	0.733	0.733	0.350	100
rs1542176	0.757	0.757	0.496	100
rs1551424	0.886	0.886	0.323	100
rs1551425	0.863	0.863	0.103	100
rs1558642	0.895	0.895	0.244	99.6
rs1558643	0.896	0.896	0.421	100
rs1558645	0.598	0.598	0.422	100
rs1559018	0.735	0.735	0.280	99.3
rs1562302	0.803	0.803	0.182	100
rs1562305	0.791	0.791	0.255	100
rs1613367	0.806	0.781	0.475	98.9
rs1623119	0.839	0.839	0.190	98.9
rs1623342	0.912	0.912	0.338	99.6
rs1688075	0.724	0.724	0.076	100
rs16944	0.743	0.743	0.283	100
rs1699089	0.713	0.713	0.137	100
rs17026582	0.897	0.897	0.240	100
rs17027415	0.708	0.708	0.042	100
rs17042407	0.859	0.859	0.298	100
rs17042795	0.844	0.801	0.009	100

rs Nº	GenTrain Score	GC10	MAF	Call Freq. (%)
rs17207494	0.884	0.884	0.408	100
rs17561	0.826	0.826	0.276	100
rs1785883	0.643	0.643	0.092	98.9
rs1785884	0.901	0.901	0.334	100
rs1834481	0.814	0.814	0.289	100
rs1861228	0.807	0.806	0.401	99.6
rs1861229	0.926	0.926	0.157	100
rs1867761	0.858	0.858	0.226	100
rs1867827	0.947	0.947	0.079	100
rs1867829	0.746	0.746	0.090	100
rs1867831	0.769	0.762	0.309	99.3
rs1867834	0.863	0.863	0.460	100
rs1880000	0.829	0.829	0.144	100
rs1921622	0.824	0.824	0.486	100
rs1922288	0.899	0.899	0.251	99.6
rs192291	0.879	0.615	0.000	100
rs1946519	0.914	0.914	0.370	100
rs1954187	0.798	0.798	0.383	100
rs1988743	0.874	0.874	0.251	100
rs1992761	0.803	0.803	0.070	100
rs2027432	0.871	0.871	0.171	100
rs2041751	0.889	0.889	0.319	100
rs2043055	0.843	0.843	0.341	98.9
rs2048874	0.929	0.929	0.125	100
rs2056400	0.803	0.803	0.190	100
rs2058612	0.934	0.934	0.440	100
rs2059020	0.942	0.942	0.186	100
rs2063421	0.749	0.749	0.262	100
rs2071374	0.909	0.909	0.329	99.3
rs2072473	0.829	0.829	0.141	100
rs2072475	0.858	0.858	0.157	100
rs2075190	0.919	0.919	0.287	100
rs2100071	0.862	0.862	0.323	100
rs2105290	0.602	0.601	0.453	99.6
rs2110562	0.925	0.925	0.478	99.6
rs2110726	0.724	0.724	0.384	100
rs2160226	0.907	0.907	0.348	100
rs2160227	0.928	0.928	0.278	100
rs2177317	0.751	0.751	0.383	100
rs2190358	0.965	0.965	0.292	100
rs2190364	0.823	0.823	0.274	100
rs2192752	0.899	0.899	0.245	100
rs2197578	0.807	0.807	0.330	100
rs2228139	0.873	0.829	0.061	100
rs2232354	0.841	0.752	0.211	99.3
rs2241116	0.612	0.610	0.220	99.6
rs2270297	0.459	0.543	0.211	100
rs2282747	0.784	0.784	0.121	100
rs2287040	0.750	0.750	0.341	100
rs2287047	0.943	0.943	0.289	100
rs2287049	0.827	0.827	0.292	100
rs2298455	0.760	0.760	0.134	100
rs2310175	0.844	0.801	0.137	100

rs Nº	GenTrain Score	GC10	MAF	Call Freq. (%)
rs2310186	0.842	0.842	0.119	100
rs2361832	0.858	0.858	0.229	99.6
rs2361836	0.899	0.899	0.260	99.6
rs2362601	0.752	0.752	0.161	99.6
rs2362605	0.792	0.790	0.327	100
rs2362607	0.950	0.950	0.282	100
rs2515403	0.813	0.813	0.392	99.6
rs2592346	0.816	0.814	0.377	100
rs2666356	0.932	0.932	0.404	99.6
rs2723154	0.912	0.912	0.155	100
rs2723167	0.649	0.649	0.471	99.6
rs2723168	0.922	0.927	0.103	100
rs2723187	0.658	0.658	0.103	100
rs2723192	0.801	0.801	0.103	100
rs28385684	0.806	0.765	0.005	100
rs2871432	0.954	0.954	0.458	100
rs2871457	0.665	0.665	0.329	100
rs2885370	0.954	0.954	0.242	100
rs2885373	0.749	0.749	0.130	100
rs2885546	0.726	0.726	0.249	100
rs28928282	0.839	0.839	0.079	98.9
rs28928312	0.819	0.819	0.193	100
rs28938782	0.926	0.880	0.009	99.6
rs28992498	0.833	0.833	0.108	99.3
rs3136558	0.802	0.802	0.206	100
rs315934	0.832	0.832	0.170	98.9
rs315943	0.879	0.879	0.455	99.6
rs315949	0.926	0.926	0.455	99.6
rs315951	0.947	0.947	0.255	100
rs3171845	0.807	0.767	0.065	100
rs3181318	0.731	0.731	0.301	100
rs3218848	0.501	0.500	0.363	100
rs3218888	0.811	0.811	0.119	100
rs3218894	0.776	0.738	0.052	100
rs3218921	0.703	0.702	0.200	99.6
rs3218979	0.537	0.376	0.101	100
rs3218984	0.776	0.776	0.251	99.3
rs34862832	0.855	0.855	0.146	100
rs35002769	0.744	0.744	0.052	99.6
rs35582281	0.916	0.870	0.052	100
rs36036501	0.795	0.795	0.069	100
rs36062386	0.568	0.539	0.047	99.6
rs360722	0.887	0.887	0.105	100
rs360726	0.831	0.831	0.220	99.6
rs3732127	0.873	0.873	0.199	99.6
rs3738447	0.922	0.922	0.092	100
rs3738448	0.873	0.873	0.079	100
rs3755285	0.778	0.778	0.262	99.6
rs3771166	0.799	0.799	0.406	100
rs3771175	0.852	0.852	0.168	100
rs3771184	0.777	0.777	0.123	99.3
rs3771187	0.813	0.813	0.473	100
rs3771199	0.912	0.912	0.348	100

rs Nº	GenTrain Score	GC10	MAF	Call Freq. (%)
rs3773953	0.739	0.739	0.106	100
rs3773958	0.725	0.725	0.170	100
rs3773976	0.789	0.789	0.159	100
rs3773977	0.905	0.905	0.238	100
rs3773982	0.784	0.784	0.161	100
rs3773989	0.859	0.859	0.471	100
rs3773999	0.742	0.742	0.236	100
rs3783516	0.857	0.857	0.453	100
rs380092	0.779	0.779	0.298	100
rs3811046	0.664	0.662	0.307	99.6
rs3811048	0.705	0.705	0.307	100
rs3811050	0.782	0.782	0.173	100
rs3811054	0.523	0.523	0.070	99.6
rs3814721	0.871	0.871	0.065	100
rs3821740	0.729	0.729	0.079	100
rs3827763	0.771	0.771	0.296	100
rs3882891	0.771	0.771	0.386	100
rs3917225	0.931	0.931	0.433	100
rs3917243	0.879	0.879	0.453	100
rs3917254	0.814	0.814	0.088	100
rs3917257	0.879	0.615	0.00	100
rs3917273	0.770	0.770	0.394	100
rs3917292	0.702	0.702	0.078	99.6
rs3917296	0.631	0.599	0.097	100
rs3917332	0.896	0.896	0.217	99.6
rs3917356	0.794	0.794	0.480	100
rs3917365	0.779	0.779	0.083	100
rs3917366	0.923	0.923	0.222	100
rs3917368	0.798	0.798	0.395	99.6
rs3923566	0.805	0.805	0.199	97.1
rs3943979	0.799	0.799	0.431	100
rs3948121	0.945	0.945	0.291	100
rs4140786	0.855	0.855	0.482	100
rs419598	0.857	0.857	0.251	100
rs4252042	0.739	0.739	0.106	99.6
rs4320092	0.899	0.899	0.253	100
rs4364030	0.745	0.745	0.484	100
rs4378247	0.918	0.918	0.139	100
rs440286	0.909	0.909	0.245	100
rs4686554	0.796	0.796	0.289	100
rs4687150	0.905	0.905	0.377	99.6
rs4754123	0.576	0.576	0.065	100
rs4849123	0.895	0.895	0.348	100
rs4849125	0.823	0.823	0.303	99.6
rs4849140	0.856	0.856	0.464	100
rs4849159	0.928	0.928	0.166	99.3
rs4850997	0.815	0.815	0.170	100
rs4851516	0.834	0.834	0.309	99.6
rs4851519	0.643	0.643	0.101	100
rs4851522	0.747	0.747	0.099	100
rs4851526	0.803	0.803	0.363	100
rs4851531	0.731	0.731	0.381	96.8
rs4851534	0.938	0.938	0.496	99.6

rs Nº	GenTrain Score	GC10	MAF	Call Freq. (%)
rs4851543	0.917	0.917	0.090	100
rs4851561	0.957	0.957	0.060	100
rs4851575	0.807	0.807	0.211	99.6
rs4925663	0.706	0.706	0.431	100
rs492859	0.699	0.699	0.134	99.3
rs4937075	0.816	0.816	0.294	100
rs4937113	0.792	0.792	0.384	99.6
rs4988956	0.841	0.841	0.406	100
rs501192	0.579	0.579	0.128	100
rs508760	0.772	0.772	0.072	100
rs518878	0.957	0.957	0.137	100
rs530537	0.772	0.772	0.388	100
rs538943	0.362	0.344	0.072	100
rs549908	0.748	0.748	0.278	98.9
rs571593	0.666	0.666	0.088	100
rs5744247	0.806	0.806	0.090	100
rs6426244	0.954	0.954	0.319	100
rs6444435	0.798	0.798	0.314	100
rs6444462	0.732	0.732	0.253	98.2
rs6444476	0.859	0.859	0.085	100
rs6444517	0.877	0.877	0.092	100
rs6543113	0.793	0.793	0.271	100
rs6702432	0.912	0.912	0.431	100
rs6706844	0.885	0.885	0.464	100
rs6710885	0.927	0.927	0.381	100
rs6712572	0.712	0.712	0.464	100
rs6719130	0.657	0.657	0.139	100
rs6725997	0.725	0.725	0.496	100
rs6728945	0.937	0.937	0.134	100
rs6729953	0.816	0.816	0.162	100
rs6739883	0.923	0.923	0.191	100
rs6743376	0.841	0.841	0.323	100
rs6752467	0.815	0.815	0.139	100
rs6754776	0.926	0.926	0.121	100
rs6755497	0.793	0.555	0.000	100
rs6758936	0.706	0.706	0.498	100
rs6760120	0.886	0.886	0.394	100
rs6761291	0.926	0.926	0.421	99.3
rs6763652	0.879	0.879	0.051	100
rs6767500	0.879	0.425	0.000	91
rs6767916	0.828	0.828	0.052	100
rs6769335	0.854	0.854	0.135	100
rs6773525	0.829	0.829	0.170	100
rs6774178	0.677	0.677	0.108	100
rs6775893	0.941	0.941	0.471	100
rs6778424	0.845	0.845	0.132	100
rs6788413	0.911	0.911	0.449	99.6
rs6790983	0.839	0.839	0.103	99.3
rs6794669	0.943	0.943	0.388	100
rs6804748	0.724	0.724	0.430	99.6
rs6807826	0.788	0.788	0.083	100
rs6808596	0.894	0.894	0.478	100
rs7109571	0.924	0.924	0.334	100

rs N^{o}	GenTrain Score	GC10	MAF	Call Freq. (%)
rs716984	0.837	0.837	0.330	100
rs7366424	0.625	0.625	0.374	100
rs7432178	0.815	0.774	0.018	100
rs7556185	0.688	0.688	0.092	100
rs7558013	0.761	0.761	0.213	100
rs7569218	0.565	0.565	0.493	100
rs7570267	0.824	0.824	0.471	100
rs7579737	0.871	0.871	0.283	100
rs7589525	0.777	0.777	0.251	100
rs759381	0.851	0.851	0.233	99.6
rs7595507	0.702	0.702	0.471	100
rs7599662	0.696	0.696	0.323	100
rs7603234	0.639	0.639	0.069	100
rs7611280	0.720	0.720	0.208	98.2
rs7611804	0.808	0.808	0.150	100
rs7618955	0.951	0.951	0.486	100
rs7624125	0.843	0.843	0.161	100
rs7626795	0.908	0.908	0.134	100
rs7627123	0.794	0.794	0.195	100
rs7628727	0.829	0.829	0.365	99.6
rs7650510	0.636	0.636	0.235	100
rs766442	0.949	0.949	0.282	100
rs7934239	0.938	0.938	0.049	100
rs871376	0.819	0.819	0.215	100
rs871658	0.761	0.761	0.119	99.3
rs873625	0.739	0.739	0.375	100
rs879711	0.921	0.921	0.410	100
rs887998	0.899	0.899	0.245	100
rs895497	0.795	0.795	0.323	100
rs917997	0.839	0.839	0.218	100
rs921065	0.728	0.728	0.496	100
rs9290933	0.714	0.714	0.475	100
rs9290936	0.829	0.829	0.125	100
rs9290939	0.642	0.642	0.067	100
rs9290953	0.918	0.918	0.309	100
rs9290958	0.946	0.946	0.108	100
rs9290978	0.862	0.862	0.282	100
rs9308849	0.565	0.565	0.309	100
rs931471	0.860	0.860	0.298	99.6
rs949963	0.722	0.722	0.171	100
rs951193	0.819	0.779	0.045	99.3
rs955754	0.855	0.855	0.202	100
rs963399	0.822	0.822	0.381	99.3
rs9682599	0.764	0.764	0.126	100
rs9814312	0.873	0.829	0.069	100
rs9814786	0.793	0.792	0.433	99.3
rs9832276	0.796	0.796	0.368	100
rs9834809	0.856	0.856	0.182	99.6
rs9845129	0.966	0.966	0.305	100
rs9854473	0.869	0.869	0.119	100
rs9855788	0.861	0.861	0.164	100
rs9857926	0.815	0.815	0.078	100
rs9859989	0.697	0.697	0.298	100

rs N^{o}	GenTrain Score	GC10	MAF	Call Freq. (%)
rs9864293	0.811	0.811	0.370	99.6
rs9868955	0.904	0.904	0.449	100
rs9877159	0.793	0.793	0.116	100
rs9877268	0.748	0.748	0.336	100
rs997049	0.927	0.927	0.413	100

Genotyping assay results of Stage-1 SNPs

Following genotype clustering each genotype is assigned a GenCall quality score, indicating the relationship of the data point to the centre of the cluster so that smaller GenCall scores reflect genotypes which are further away from the cluster, and therefore less robust. The GC10 score, the 10th percentile of all the GenCall scores across all samples for each SNP, indicates the quality and confidence in the genotype assignment for a particular SNP. The GC10 scores for all of the SNPs genotyped in stage-1 of the association study which satisfied the quality control criteria, as well as the minor allele frequency (MAF) and percentage call rate, are shown.

Appendix 3. IL-18BP haplotype transient transfection assays

A. Methods

Plasmid construction

To include all of the experimentally validated regulatory elements in the *IL18BP* promoter, (Hurgin et al., 2002) as well as the two SNPs of interest, the region from -1426 relative to start of transcription, to +577 was used for cloning, as shown in Figure 1. The region was amplified using primers CCAACGCGTttctagggaatatggctcag, including the *MluI* recognition sequence, and CCAAGATCTccagaaaccttcagctcctg containing the *BglII* recognition sequence (restriction enzyme recognition sequences underlined). Genomic DNA from an individual homozygous for the common allele (according to CEPH allele frequencies; 1 representing the common allele, 2 representing the rare allele) at each loci (C at SNP1, A at SNP2) was used as the template. The fragment was cloned into pGL3-Basic luciferase expression vector (Promega) using the enzyme restriction sites. The resulting vector is designated pGL3.1-1. To generate the remaining three haplotypes site-directed mutagenesis was performed using QuikChange II Site-Directed Mutagenesis kit (Stratagene, CA). The primer pair ccaggcagccttgcTtcctgacctaactc and gagttaggtcaggaAgcaaggctgcctgg (forced change shown in capitals) was used to force a T at SNP1; gcatgggggtagattagCgatcccagtctggt and accagactgggatcGctaatctacccccatgc to force a C at SNP2. The strategy of site directed mutagenesis implemented is shown in Figure 2, two separate mutations of pGL3.1-1 to generate pGL3.2-1 and pGL3.1-2, and a mutation of pGL3.1-2 to generate pGL3.2-2. Sequences were confirmed by sequencing with the vector primers RV primer 3, GL primer 2 (Promega), and a primer within the cloned region (aacagttctgagtcccttga).

Transient transfections and luciferase reporter gene assays

Initial qRT-PCR experiments showed optimal endogenous *IL18BP* induction in HeLa cells after 24hrs stimulation with human IFN-γ (data not shown). HeLa cells were grown in DMEM, 10% FBS, and seeded at $6x10^4$ cells per well in a 24-well plate for 24 hours prior to transfection. Using lipofectamine LTX (Invitrogen), cells were transfected with 250ng pGL3 constructs and 25ng pRL-tk and incubated for 24 hours. The cells were then either stimulated with human IFN-γ (20 ng/ml) or left unstimulated for a further 24 hours. Luciferase activities were determined using Dual-Glo® luciferase assay (Promega) in a GLOMA 96 microplate Luminometer (Promega) according to manufacturer's instructions. Promoter activities are

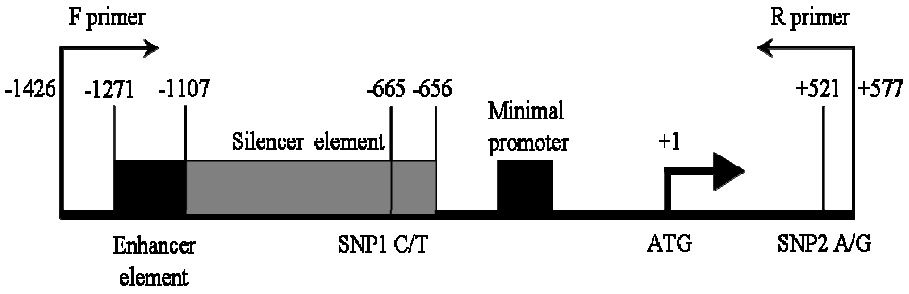


Figure 1. IL-18BP promoter region

The positions of SNP1 (rs3814721), SNP2 (rs2298455), and the silencer and enhancer elements in the Il-18BP promoter region are shown. All positions given are relative to the start of transcription. Also shown are the primers used for cloning the region.

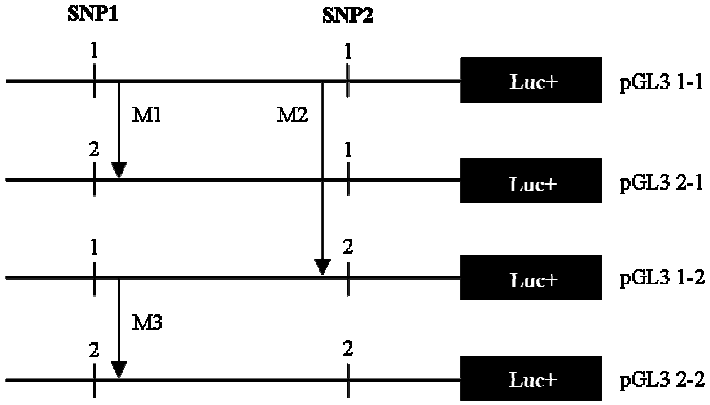


Figure 2. Site directed mutagenesis strategy for generation of pGL3 vector constructs

pGL3 1-1 cloned from genomic DNA with the common allele at both loci was used as the template sequence for generation of the other haplotype constructs. Site-directed mutagenesis was conducted sequentially on the C allele at SNP1 (Mutation 1) or the A allele at SNP2 (Mutation 2) to create the pGL3 2-1 and pGL3 1-2 vectors respectively. The pGL3 1-2 vector was then used as the template for mutagenesis at SNP1 (Mutation 3) to generate the pGL3 2-2 vector.

expressed as the ratio of firefly to renilla luciferase activity. Each sample was performed in triplicate, and the experiment repeated three times.

Statistical analysis

The ratios of firefly:renilla luciferase activity between the constructs (mean ± standard error) were analysed by one-way ANOVA using SPSS 13.0 (SPSS Inc).

B. Results

No significant difference in promoter activity between the four haplotype constructs were observed in unstimulated cells, $p=0.272$ (data not shown). Luciferase activities following IFN-γ stimulation showed that overall the promoter haplotype had a significant ($p<0.0001$) effect on transcription (data not shown).

The transcriptional activity of the individual haplotype construct is shown in Figure 3. In stimulated cells the highest transcription was seen with the pGL3.1-1 construct, the lowest with pGL3.2-2, and an intermediate level with pGL3.1-2 and pGL3.2-1. Transcriptional activity of pGL3.1-1 was significantly higher than each of the other haplotypes. The transcriptional activity of pGL3.2-2 was also significantly lower than each of the others. There was no significant difference between pGL3.1-2 and pGL3.2-1. The results of the pair-wise comparison of the constructs are shown in Table 1.

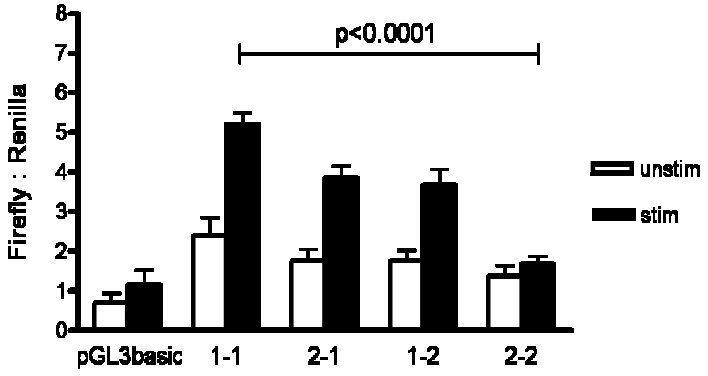


Figure 3. Luciferase activity of the haplotype constructs

HeLa cells transiently co-transfected with pGL3 *IL-18BP* promoter reporter constructs and the control pRL-tk (10:1) were either stimulated with 20ng/ml hIFN-γ for 24 hours, or left unstimulated. Shown in this figure are the mean ratios ± SE of Luciferase to Renilla activity of each of the four haplotype containing plasmids, and the pGL3basic vector without insert. Measurements are from three independent experiments, each with samples in triplicate (n=9).

	pGL3.2-1	**pGL3.1-2**	**pGL3.2-2**
pGL3.1-1	p=0.014	p=0.008	p<0.0001
pGL3.2-1	N/A	p=0.733	p=0.001
pGL3.1-2	N/A	N/A	p=0.002

Table 1. Pairwise comparisons of IFNγ-inducible luciferase activity by haplotype

The mean Luciferase:Renilla ratios for each of the four haplotype plasmid constructs were compared by one-way ANOVA. The p values are shown for each of the pairwise comparisons between the haplotype constructs.