

INFLAMMATORY BOWEL DISEASE IN
THE UK SOUTH ASIAN POPULATION:
INSIGHTS INTO CLINICAL PHENOTYPE
AND PATHOGENESIS

A thesis submitted for the degree of

Doctor of Medicine (Research)

Imperial College London

2011

DAVID GRAHAME WALKER

Department of Medicine

Imperial College London

ABSTRACT

The inflammatory bowel diseases (IBD), Crohn's disease (CD) and ulcerative colitis (UC) are significant causes of gastrointestinal morbidity. Previous research in IBD has predominantly focused on large white European/American populations, which have defined the clinical phenotype and led to important biological insights, including its genetic predisposition, and immunological and metabonomic associations. Recent studies suggest that there may be differences in the clinical phenotype of IBD in different ethnic groups. There is still limited information on the phenotype of South Asian IBD patients originating from India, Pakistan and Bangladesh, but there have been suggestions that it differs from the Northern European white population. Since the 1950s there has been substantial migration from South Asia, and currently 2.08 million South Asians live in the UK: studies of IBD in this population could provide valuable information, and may help to distinguish the degree of influence of genetic and environmental pathogenic factors. This thesis aimed to define the clinical phenotype of IBD in UK South Asians, and then to investigate the genetic, serological and metabonomic associations of IBD in South Asians and to compare the findings with a white Northern European cohort.

The phenotypic details of 367 South Asian IBD patients (273 UC: 94 CD), undergoing active follow-up in five North West London hospitals, were compared to those of 403 consecutively collected white Northern European IBD patients (188 UC: 215 CD). The phenotype of IBD differed significantly between the two populations with more extensive disease and less proctitis in the South Asian UC cohort compared to Northern Europeans ($P<0.0001$ for both). CD behaviour differed significantly between the groups with less penetrating disease compared to Northern Europeans ($P=0.01$) and a reduced need for surgery ($P=0.003$).

Genotyping, serological testing and urinary nuclear magnetic resonance (NMR) spectroscopy were used to investigate cohorts of South Asian IBD patients and healthy South Asian controls in order to investigate the different aetiological factors in South Asian IBD. Results were compared with a previously collected white Northern European cohort. The genotype frequency of *HLA-DRB1*1502* was significantly higher in South Asian patients with UC compared to both controls ($P=0.005$) and Northern European UC patients ($P<0.0001$). *HLA-DRB1*0103* was not found in the South Asian population and *CARD15/NOD2* polymorphisms were rare, significantly different to Northern Europeans with IBD. The sensitivity and specificity of serological tests (ASCA and atypical pANCA) was similar for both populations. Urinary NMR spectroscopy results from the South Asian IBD cohort showed significantly lower levels of hippurate (a co-metabolite related to gut microbial metabolism), compared to healthy South Asian controls ($P<0.0001$). Differences were also observed between the South Asian and Northern European UC cohorts, with lower levels of 4-cresol sulphate ($P<0.0001$) and hippurate ($P=0.0006$) in South Asians.

Conclusion: The phenotype and genotype of IBD in South Asians living in North West London is significantly different from that of a white Northern European IBD cohort. The emerging technique of urinary NMR spectroscopy has shown specific urinary metabolites differ between South Asian and European IBD patients which could reflect underlying differences in the gut microbiota.

DEDICATION

To Morwenna, Darcey and Zara

With lots of love

STATEMENT OF ORIGINALITY

All the work in this thesis was carried out by myself with the following exceptions:-

Horace Williams collected phenotypic data on 280 Northern European IBD patients. He also supplied the SSP-PCR genotyping results (HLA class II and *CARD15/NOD2*) for the Northern European population as well as the data for urinary metabolic profiling and serological markers in Northern Europeans.

Tariq Ahmad supplied the primer sequences required for *CARD15/NOD2* genotyping and Hiroe Sato supplied the HLA primer sequences for medium resolution HLA class II genotyping. Aaron Bancil and Pav Rai helped with pyrosequencing for *ATG16L1*, *IRGM* and *IL23R* at Hammersmith Hospital. Jaspal Kooner (Department of Cardiovascular Medicine, Imperial College London) and John Chambers (Department of Public Health, Imperial College London) supplied blood samples from 210 healthy South Asian controls.

Jane Cox acquired the urinary NMR spectra for each participant using a NMR spectrometer at the Imaging Sciences Department, Hammersmith Hospital, London.

Dr Bernard North of Imperial College London Statistical Advisory Service performed the genetic and metabonomic power calculations.

ACKNOWLEDGEMENTS

It is a pleasure to thank the many people who made this thesis possible. First and foremost, I would like to express my utmost gratitude to my supervisor Dr Tim Orchard for the continuous support he has given me throughout my research. His patience, motivation, enthusiasm, and knowledge made my research period a stimulating, enjoyable and memorable experience, I will always be thankful for his help.

I am indebted to many colleagues for providing an enjoyable environment in which I conducted my research. I am especially grateful to Horace Williams whose guidance and friendship from the outset was invaluable. Without his help in securing a research grant and guiding me through the many obstacles that I came up against, this dissertation would not have been possible.

I am very grateful to Huw Thomas, Julian Teare, Stephen Kane, Joel Mawdsley, Jayantha Arnold, Ian McNeil, Ailsa Hart, Maxton Pitcher, Andrew Milestone, Shamima Padaruth and Julian Walters who all helped me with patient recruitment. I also thank Santosh Chadda for her assistance with retrieving medical notes.

Hiroe Sato and Panos Pantelidis from the Department of Gene Therapy at the National Heart and Lung Institute were invaluable in helping me learn the techniques necessary for SSP-PCR genotyping. I am obliged to Pav Rai for teaching me how to do pyrosequencing and he was an excellent research companion. I thoroughly enjoyed supervising Aaron Bancil's BSc project and I thank him for all his help with pyrosequencing at Hammersmith Hospital.

I am very thankful to Jane Cox for her help with processing the samples for NMR spectroscopy and her helpful discussions.

Professor Sara Marshall and Professor Derek Jewell offered helpful advice with study design and I appreciated their interest in my research.

I would particularly like to thank St Mary's Paddington Charitable Trustees, the Broad Medical Research Program and Johnson & Johnson Pharmaceuticals whose generous support funded my research.

Finally, I would like to thank my wife Morwenna, and my two daughters Darcey and Zara. Without their continuing love, support and encouragement this thesis would never have been completed.

TABLE OF CONTENTS

<i>Abstract</i>	<i>1</i>
<i>Dedication</i>	<i>2</i>
<i>Statement of originality</i>	<i>3</i>
<i>Acknowledgements</i>	<i>4</i>
<i>Table of contents</i>	<i>6</i>
<i>List of tables</i>	<i>13</i>
<i>List of figures</i>	<i>16</i>
<i>List of abbreviations</i>	<i>20</i>
 CHAPTER 1: INTRODUCTION	 21
1.1 INFLAMMATORY BOWEL DISEASE	22
1.2 GLOBAL VARIATION IN THE EPIDEMIOLOGY OF INFLAMMATORY BOWEL DISEASE	22
1.2.1 A comparison of the demographics and clinical phenotype of ulcerative colitis between Asians and Caucasians	26
1.2.2 A comparison of the demographics and clinical phenotype of Crohn's disease between Asians and Caucasians	29
1.2.3 Rationale for studying migrant populations	33
1.2.4 Epidemiological studies of migrant IBD populations living in Asia	33
1.2.5 Epidemiology and clinical phenotype of UC in South Asian migrants living in the UK	35
1.2.6 Epidemiology and clinical phenotype of CD in South Asian migrants living in the UK	37

1.3	INFLAMMATORY BOWEL DISEASE GENETICS	39
1.3.1	Caspase activating recruitment domain 15 (<i>CARD15</i>)/Nucleotide-binding oligomerization domain 2 (<i>NOD2</i>)	40
1.3.2	Autophagy-Related-16-like 1 gene (<i>ATG16L1</i>)	42
1.3.3	Immunity-Related GTPase protein M (<i>IRGM</i>)	45
1.3.4	Interleukin 23 Receptor Gene (<i>IL23R</i>)	46
1.3.5	Human Leukocyte antigen (HLA)	48
1.3.6	HLA class II genes and IBD	52
1.3.7	Correlation between genotype and IBD phenotype	56
1.3.8	Variation in IBD genetics between different ethnic groups	57
1.4	METABOLIC PROFILING - EVIDENCE FOR THE ROLE OF THE MICROBIOTA IN THE PATHOGENESIS OF IBD	62
1.4.1	Principles of nuclear magnetic resonance spectroscopy (NMRS)	64
1.4.2	Clinical applications of NMRS	66
1.4.3	The role of NMRS in the characterisation of IBD	66
1.5	SEROLOGICAL MARKERS IN INFLAMMATORY BOWEL DISEASE PATIENTS	69
1.5.1	Atypical perinuclear anti-neutrophil cytoplasmic antibody (atypical pANCA)	69
1.5.2	Anti-Saccharomyces cerevisiae antibodies (ASCA)	71
1.5.3	Antibodies against Outer Membrane Porin C (anti-OmpC)	72
1.5.4	IBD phenotype associations with serological markers	72
1.5.5	The frequency of IBD serological markers in South Asian IBD patients	73

CHAPTER 2: AIMS AND HYPOTHESES 75

**CHAPTER 3: DEFINING THE CLINICAL PHENOTYPE OF INFLAMMATORY
BOWEL DISEASE IN UK SOUTH ASIANS_____ 78**

3.1	INTRODUCTION _____	79
3.1.1	Study aims _____	80
3.2	MATERIALS AND METHODS _____	81
3.2.1	Ethics Approval _____	81
3.2.2	Patient Recruitment _____	81
3.2.3	Case Definition_____	82
3.2.4	Patient demographics and clinical phenotype_____	83
3.2.5	Defining IBD disease behaviour and extent_____	84
3.2.6	Extraintestinal manifestations _____	85
3.2.7	Assessing the impact of environmental factors in the UK on disease development and phenotype _____	85
3.2.8	Statistical Analysis _____	86
3.3	RESULTS_____	87
3.3.1	UC demographics _____	87
3.3.2	UC phenotype_____	88
3.3.3	CD demographics _____	91
3.3.4	CD phenotype_____	93
3.3.5	Demographics of first-generation South Asian IBD patients _____	96
3.3.6	Phenotypic comparison of first and second-generation South Asian IBD patients _____	97
3.4	DISCUSSION_____	98
3.4.1	UC phenotype in UK South Asians _____	99
3.4.2	CD phenotype in UK South Asians _____	100
3.4.3	Potential explanations for differences in clinical phenotype _____	101
3.4.4	Conclusion _____	104

CHAPTER 4: PHENOTYPE GENOTYPE ASSOCIATIONS OF INFLAMMATORY BOWEL DISEASE IN A SOUTH ASIAN POPULATION _____ 105

4.1	INTRODUCTION _____	106
4.1.1	Study Aims _____	107
4.2	METHODS _____	109
4.2.1	Ethics approval _____	109
4.2.2	Sample Collection _____	109
4.2.3	DNA isolation _____	110
4.2.4	Determining DNA concentration _____	112
4.2.5	Single Nucleotide Polymorphism Analysis _____	113
4.2.6	Sequence Specific Primer - Polymerase Chain Reaction (SSP-PCR) _____	114
4.2.7	Pyrosequencing _____	119
4.2.8	Statistical analysis _____	123
4.3	RESULTS _____	125
4.3.1	Frequency of Crohn's disease associated <i>CARD15/NOD2</i> polymorphisms in the South Asian IBD population _____	127
4.3.2	Associations of polymorphisms in <i>ATG16L1</i> , <i>IRGM</i> and <i>IL23R</i> in the South Asian CD population _____	130
4.3.3	Comparison of the carriage frequency of <i>HLA-DRB1*0103</i> and <i>HLA-DRB1*1502</i> in South Asian IBD cases and controls _____	133
4.4	DISCUSSION _____	137
4.4.1	IBD associated genetic polymorphisms _____	137
4.4.2	HLA associations in South Asian IBD patients _____	139
4.4.3	Genetic associations in South Asian IBD patients compared to other countries _____	141
4.4.4	Conclusion _____	142

CHAPTER 5: URINARY METABOLIC PROFILING IN A SOUTH ASIAN

INFLAMMATORY BOWEL DISEASE POPULATION_____ 143

5.1	INTRODUCTION _____	144
5.1.1	Investigating changes in the gut flora using metabolic profiling _____	145
5.1.2	Study aims _____	146
5.2	METHODS_____	147
5.2.1	Ethics Approval _____	147
5.2.2	Patient recruitment _____	147
5.2.3	Experimental design_____	148
5.2.4	Urine sample preparation for NMR spectroscopy _____	150
5.2.5	In vitro NMR spectroscopy _____	151
5.2.6	Data analysis _____	151
5.2.7	Statistical methods and power calculation _____	155
5.3	RESULTS_____	157
5.3.1	NMR data analysis _____	158
5.3.2	A comparison of the urinary NMR profile of South Asian IBD patients with healthy South Asian controls_____	160
5.3.3	South Asian lifestyle & dietary comparisons _____	163
5.3.4	A comparison of the urinary NMR profile of South Asian and white Northern European IBD patients _____	164
5.3.5	Potential confounders between ethnic groups _____	167
5.4	DISCUSSION_____	168
5.4.1	Differences between South Asian IBD patients and controls_____	168
5.4.2	Comparing the metabolic profiles of IBD patients of different ethnicities_____	173
5.4.3	Potential explanations for metabolite differences between ethnic groups _____	174
5.4.4	Conclusion _____	175

CHAPTER 6: THE PREVALENCE AND DIAGNOSTIC VALUE OF

SEROLOGICAL MARKERS IN SOUTH ASIAN IBD PATIENTS _____ 177

6.1	INTRODUCTION _____	178
6.1.1	IBD phenotype associations with serological markers _____	180
6.1.2	Study aims _____	181
6.2	METHODS _____	182
6.2.1	Ethics Approval _____	182
6.2.2	Patient recruitment _____	182
6.2.3	Sample collection _____	182
6.2.4	Atypical pANCA indirect immunofluorescence assay _____	183
6.2.5	ASCA IgG and IgA Assays _____	185
6.2.6	Anti-OmpC Assay _____	188
6.2.7	Interpretation of ELISA results _____	189
6.2.8	Statistical analysis _____	189
6.3	RESULTS _____	190
6.3.1	The prevalence of ASCA, anti-OmpC and atypical pANCA in South Asian IBD patients and healthy controls _____	190
6.3.2	Sensitivity and specificity of serological markers for differentiating South Asian IBD patients from healthy South Asian controls _____	191
6.3.3	Diagnostic accuracy of ASCA and atypical pANCA in distinguishing CD from UC in South Asian patients with IBD compared to a white Northern European cohort _____	192
6.3.4	Clinical phenotype associated with serological markers in South Asian IBD patients _	194
6.4	DISCUSSION _____	196
6.4.1	The role of serological markers for differentiating UC from CD _____	197
6.4.2	Stratifying IBD phenotype using serological markers _____	198
6.4.3	Conclusion _____	200

CHAPTER 7: DISCUSSION AND SUGGESTIONS FOR FUTURE WORK	201
7.1 Summary	213
CHAPTER 8: PUBLICATIONS ARISING FROM THIS THESIS TO DATE	214
8.1 FULL PUBLICATION	215
8.2 PUBLISHED ABSTRACTS	225
CHAPTER 9: APPENDICES	227
CHAPTER 10: REFERENCE LIST	236

LIST OF TABLES

Table 1-1: Ethnic variation in Crohn's disease phenotype	32
Table 1-2: <i>HLA-DRB1</i> associations with Crohn's disease and ulcerative colitis	52
Table 1-3: Frequency of the three <i>CARD15/NOD2</i> polymorphisms associated with Crohn's disease in patients of different ethnic origin.....	59
Table 1-4: Summary of the genetic differences between UK and Japanese populations in single nucleotide polymorphisms (SNPs) associated with Crohn's disease susceptibility	60
Table 3-1: Demographic details of South Asian and white Northern European ulcerative colitis patients living in North West London, UK.....	88
Table 3-2: Phenotypic details of South Asian and white Northern European ulcerative colitis patients living in North West London, UK.....	89
Table 3-3: Demographic details of South Asian and white Northern European Crohn's disease patients living in North West London, UK.....	92
Table 3-4: Phenotypic details of South Asian and white Northern European Crohn's disease patients living in North West London, UK.....	95
Table 4-1: Single Nucleotide Polymorphism (SNP) id numbers for the 4 polymorphisms investigated in <i>CARD15/NOD2</i>	115
Table 4-2: Single Nucleotide Polymorphism (SNP) id numbers used for primer design	120
Table 4-3: South Asian participant's demographic details and disease characteristics	125

Table 4-4: White Northern European participant's demographic details and disease characteristics	126
Table 4-5: Comparison of the allele frequencies of <i>CARD15/NOD2</i> risk alleles (SNP8, 12 & 13) in South Asian and white Northern European CD patients and South Asian healthy controls.	128
Table 4-6: Results of the <i>CARD15/NOD2</i> genotyping for SNP5 in South Asian and white Northern European IBD patients and South Asian controls.....	129
Table 4-7: The frequency of CD associated SNPs in <i>ATG16L1</i> , <i>IRGM</i> and <i>IL23R</i> in South Asian CD patients and controls compared to a Northern European CD cohort	131
Table 4-8 : The frequency of <i>IL23R</i> in South Asian UC patients, controls and a white Northern European UC cohort.....	132
Table 4-9: Demographic details and disease characteristics for South Asian and white Northern European ulcerative colitis patients.....	133
Table 4-10: The associations between IBD and <i>HLA-DRB1*0103</i> and <i>DRB1*1502</i> in the South Asian population.....	134
Table 4-11: A comparison of the frequencies of <i>HLA-DRB1*0103</i> and <i>HLA-DRB1*1502</i> in South Asian and European IBD populations.....	136
Table 5-1: Peak assignments for urinary metabolites dependant on gut microbial metabolism.	154
Table 5-2: Total number of South Asian urine samples collected and processed.....	157
Table 5-3: South Asian participants demographic and phenotypic details	158

Table 5-4: Demographic details of the white Northern European and South Asian IBD cohorts including sex, age and disease location.....	164
Table 6-1: Prevalence of atypical pANCA, ASCA and anti-OmpC in Caucasian IBD patients and controls	179
Table 6-2: South Asian participant characteristics for serological typing	183
Table 6-3: Accuracy of ASCA, anti-OmpC and atypical pANCA to distinguish patients with inflammatory bowel disease from healthy controls in a South Asian population .	191
Table 6-4: Accuracy of serological markers to differentiate ulcerative colitis from Crohn's disease in South Asian and white Northern European IBD patients	193
Table 6-5: Serology results for South Asians stratified by Crohn's disease location and behaviour	194
Table 6-6: Atypical pANCA results for South Asians stratified by ulcerative colitis phenotype	195

LIST OF FIGURES

Figure 1-1: World map showing the prevalence rates (cases/ 10^5) of ulcerative colitis, based on population studies.....	24
Figure 1-2: Bar graph showing the incidence rates (cases/ 10^5 /year) of IBD in Asian countries	25
Figure 1-3: Stacked chart totalling 100%, depicting the distribution of UC in studies from Northern Europe and Asia.....	27
Figure 1-4: Structure of the <i>CARD15/NOD2</i> gene and protein showing polymorphisms <i>Arg702Trp</i> , <i>Gly908Arg</i> , and <i>Leu1007fsinsC</i> in the Leucine-rich Repeat Region (LRR).....	41
Figure 1-5: Diagram showing the role of <i>NOD2</i> proteins in NF κ B production and the importance of <i>ATG16L1</i> in autophagosome formation.	44
Figure 1-6: An illustration of the role of the Interleukin 23 receptor (IL23R) in bacterial clearance.....	47
Figure 1-7: The major histocompatibility complex genes located on chromosome 6p.	49
Figure 1-8: Structure of the HLA class II molecule showing the binding of $\alpha\beta$ heterodimers....	50
Figure 1-9: Model of IBD pathogenesis. The interaction of susceptibility genes, disease phenotype genes, the host immune response and environmental stimuli, contribute to the pathogenesis of ulcerative colitis and Crohn's Disease.....	63
Figure 1-10: A charged spinning nucleus creates a magnetic field causing it to behave like a dipole or bar magnet.....	64
Figure 1-11: A typical NMR spectra which contains signals from hundreds of metabolites.	65

Figure 1-12: Atypical pANCA staining showing a broad inhomogeneous rim-like staining of the nuclear periphery with multiple intranuclear foci	70
Figure 3-1: Bar graph comparing the ulcerative colitis phenotype of South Asian and white Northern Europeans patients.	90
Figure 3-2: Bar graph comparing Crohn's disease location in South Asians and white Northern Europeans.	94
Figure 4-1: Schematic diagram showing the process of Sequence Specific Primer-Polymerase Chain Reaction (SSP-PCR)	114
Figure 4-2: Photographs of agarose gels under ultraviolet light after electrophoresis following SSP-PCR.	118
Figure 4-3: A pyrogram showing the signal peaks that determine the nucleotide sequence	120
Figure 4-4: A pyrogram showing the IRGM189 genotype for a participant. The area of the gene representing the peptide sequence at IRGM189 is highlighted in yellow, showing a peak for T and no peak for C, therefore classed at genotype TT.	130
Figure 5-1: Genetic predisposition and environmental triggers contributing to the pathogenesis of inflammatory bowel disease.....	144
Figure 5-2: Free induction decay (FID) of the Nuclear Magnetic Resonance Spectra.	152
Figure 5-3: Fourier Transformation of the FID to give the Nuclear Magnetic Resonance (NMR) spectra in the time domain displaying metabolite peaks.	152
Figure 5-4: NMR spectra after excluding the water and urea peaks in the region 6.4 to 4.5.....	153

Figure 5-5: Schematic diagram of assignment of spectral peaks.(a) A typical spectra starts with the NMR spectrum of a biological fluid, which contains signals from hundreds of metabolites. (b) The individual spectra for each metabolite are identified according to the literature. (reproduced from Nicholson (2008)).	154
Figure 5-6: The NMR spectra of urine from a South Asian control patient (Figure 6-5i) and a South Asian Crohn's disease patient (Figure 6-5ii). The spectra in the aromatic region (δ 6.0-9.0) were magnified x4 compared to the aliphatic region (δ 0.5-4.5).	159
Figure 5-7: Urinary concentrations of (a) hippurate and (b) formate (expressed as % normalised total spectral integral) for South Asian IBD patients and South Asian controls.	161
Figure 5-8: Urinary concentrations of (a) TMAO, (b) 4-cresol sulphate and (c) DMA (expressed as % normalised total spectral integral) for South Asian IBD patients and South Asian controls.	162
Figure 5-9: Urinary concentrations of (a) hippurate and (b) 4-cresol sulphate (expressed as % normalised total spectral integral) for white Northern European and South Asian IBD patients.	165
Figure 5-10: Urinary concentrations of (a) formate and (b) TMAO (expressed as % normalised total spectral integral) for white Northern European and South Asian IBD patients.	166
Figure 5-11: The synthesis of hippurate (reproduced from Williams et al. (2010)).	169
Figure 5-12: Phylogenetic associations with specific urine metabolites (reproduced from Li et al., 2008).	171

Figure 6-1: Ultraviolet light microscopy images taken by Dr Chris Bunn, Royal Free Hospital, London. The images demonstrate (a) Atypical pANCA staining, showing a broad inhomogeneous rim-like staining of the nuclear periphery without nuclear extension; For comparison, image (b) shows typical pANCA staining with a homogeneous rim-like staining of the perinuclear cytoplasm with nuclear extension.....184

Figure 6-2: An outline of the experimental technique used to determine ASCA IgG positivity. (a) 96 well plates containing pre-attached antigens were supplied by the manufactures. (b) After adding the diluted patient sera and IgG conjugate, chromogen was added to each plate. (c) A stop solution was then added to terminate enzymatic activity and the optical density calculated. 188

Figure 6-3: Positivity of ASCA (IgA & IgG), anti-OmpC and atypical pANCA in South Asian IBD patients and healthy controls. 190

LIST OF ABBREVIATIONS

^1H	Proton
pANCA	perinuclear Antineutrophil cytoplasmic antibodies
ASCA	Anti- <i>Saccharomyces cerevisiae</i> antibodies
ATG16L1	Autophagy-related-16-like 1
CARD	Caspase activating recruitment domain
CD	Crohn's disease
EIM	Extraintestinal manifestation
ELISA	Enzyme-linked immunosorbent assay
GWA	Genome wide association
HLA	Human leucocyte antigen
IBD	Inflammatory bowel disease
IL17	Interleukin 17
IL23R	Interleukin 23 receptor gene
IRGM	Immunity-related guanosine triphosphate family, M
LD	Linkage disequilibrium
LRR	Leucine rich repeat
MHC	Major histocompatibility complex
NF κ B	Nuclear factor kappa-light-chain-enhancer of activated B cells
NMR	Nuclear magnetic resonance
NOD	Nucleotide-binding oligomerization domain
Omp-C	Outer membrane porin C
PCR	Polymerase chain reaction
SNP	Single nucleotide polymorphism
SSP	Sequence-specific primer
TB	Tuberculosis
Th-17	T helper cell type 17
TMAO	Trimethylamine-N-oxide
TSP	3-trimethylsilyl-(2,2,3,3- $^2\text{H}_4$)-1-propionate
UC	Ulcerative colitis

CHAPTER 1

INTRODUCTION

1.1 INFLAMMATORY BOWEL DISEASE

The inflammatory bowel diseases, Crohn's disease and ulcerative colitis are characterised by chronic intestinal inflammation. Ulcerative colitis (UC) is an inflammatory disorder limited to the colonic mucosa. This differs from Crohn's disease (CD), which can affect any part of the gastrointestinal tract and is characterised by chronic transmural granulomatous inflammation with a tendency to form fistulae or strictures. The incidence of UC in the UK is approximately 13.9 cases/10⁵/ year with a reported prevalence of 243 cases/10⁵ (Rubin et al., 2000). Crohn's disease affects 144/10⁵ of the UK population with approximately 8.3 cases/10⁵/year (Rubin et al., 2000). Recent studies suggest that nearly 240,000 people are affected by IBD in the UK (Carter et al., 2004). The diagnosis of UC or CD is currently based on symptoms, endoscopic, radiological and histological findings. However, in up to 18% of cases they are indistinguishable from each other (IBD unclassified) (Farmer et al., 2000; Lee et al., 1979).

1.2 GLOBAL VARIATION IN THE EPIDEMIOLOGY OF INFLAMMATORY BOWEL DISEASE

Studies from around the world suggest a widely varying incidence of IBD, with highest rates of UC and CD reported in Northern Europe and North America (Loftus, Jr., 2004). Previously, research in IBD focused predominantly on large white European/North American populations, as the reporting of detailed information on the epidemiology of IBD in areas like Asia and South America suffered from poor reporting networks, limited population data and non-uniform diagnostic criteria. The first reports of

differences in the prevalence and incidence of IBD between ethnic groups came from studies in North America which demonstrated that Ashkenazi Jews had an increased incidence and prevalence of IBD compared to white Americans (Acheson, 1960). Subsequent studies in other ethnic groups including African Americans and French Canadians (Veluswamy et al., 2010;Bhat et al., 2009), also noted variations in the epidemiology of IBD compared with Caucasian populations living in the same area.

To the present day, there is still very limited information on the epidemiology of IBD in South Asia (India, Pakistan and Bangladesh). It has been assumed that the incidence was lower than in Europe, however a population-based study of UC in North India showed that the incidence (6 cases/10⁵/year) and prevalence (44.3/10⁵) is similar to those observed in some Western countries (Sood et al., 2003;Vucelic et al., 1991;Mate-Jimenez et al., 1994) (**Figure 1-1**). The community-based epidemiological study by Sood was performed in 1999 via a house-to-house survey over a 1-year period, with suspected UC cases confirmed by endoscopy and biopsy, thus ensuring good case ascertainment.

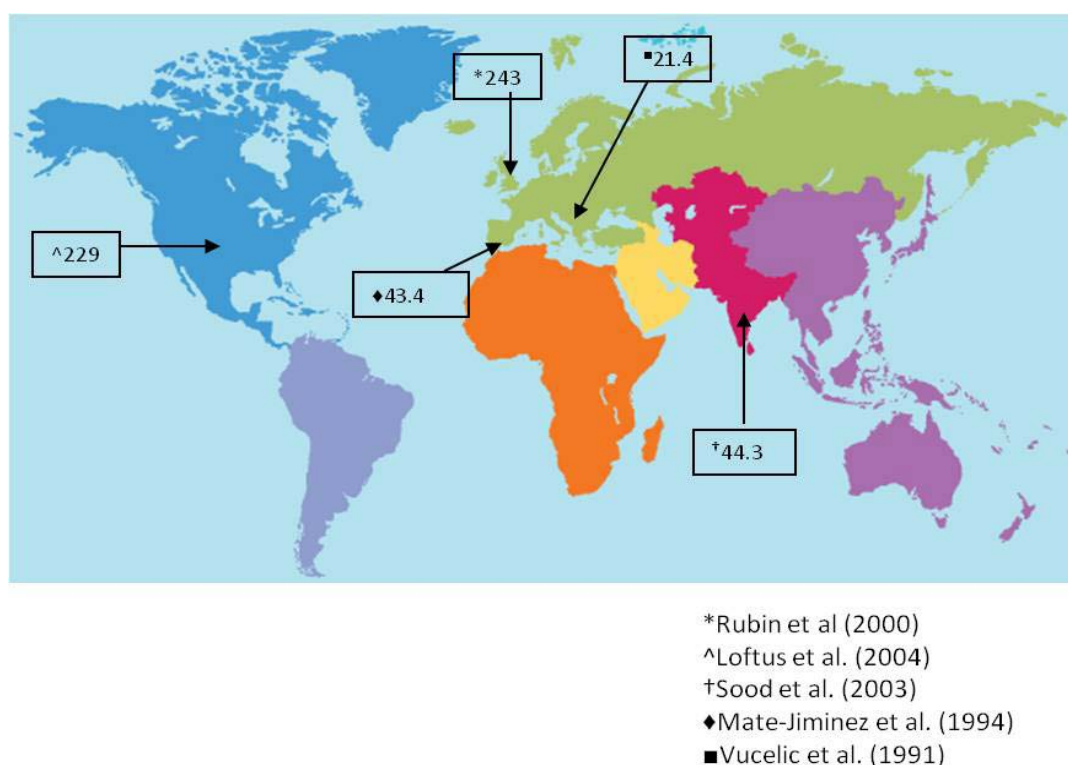


Figure 1-1: World map showing the prevalence rates (cases/10⁵) of ulcerative colitis, based on population studies.

Studies from other Asian countries have reported the incidence and prevalence of IBD using a variety of methods ranging from nationwide and hospital-based registries, with the denominator for incidence calculation based on census of the population applied to a specific region and period (Loftus, Jr., 2004; Morita et al., 1995; Leong et al., 2004; Yang et al., 2008; Niriella et al., 2010; Lok et al., 2008). The incidence rates for UC range from 0.37 – 3.08 cases/10⁵/year, much less than in India (**Figure 1-2**).

The scarcity of well-designed population-based studies in Asia, to authenticate the reported incidence and prevalence figures is a major obstacle. However, it is likely that there are racial groups among Asians that are more susceptible to IBD.

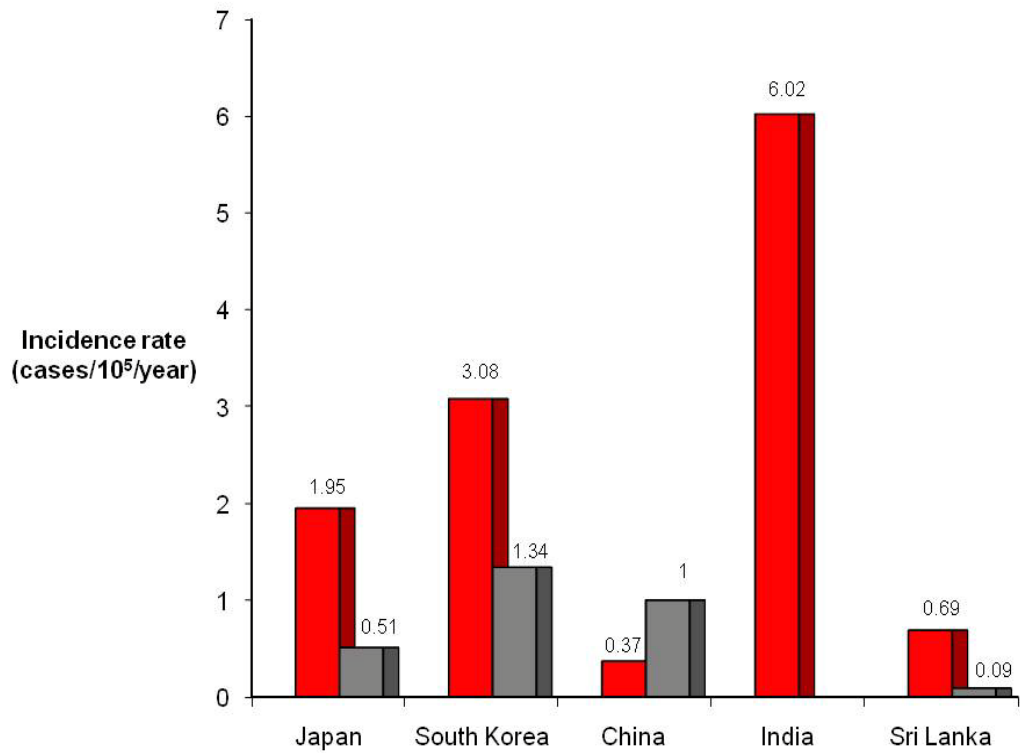


Figure 1-2: Bar graph showing the incidence rates (cases/10⁵/year) of IBD in Asian countries (■ Ulcerative colitis and ■ Crohn's disease).

(Values adapted from; Morita et al., (1995) [Japan], Yang et al., (2008) [South Korea], Lok et al., (2008) [China], Leong et al., (2004)[China], Sood et al. (2003)[India], Niriella et al., (2010) [Sri Lanka])

1.2.1 A comparison of the demographics and clinical phenotype of ulcerative colitis between Asians and Caucasians

Age & Gender

South Asian studies report a similar peak age of onset for UC as typically seen in Caucasian populations with the mean and median age at diagnosis in the third and fourth decade (Sood et al., 2003; Khosla et al., 1986; Loftus, Jr. and Sandborn, 2002). Studies from Asia have suggested a male predominance in the incidence of UC (Hou et al., 2009). This is in contrast to developed countries that generally show an equal gender distribution in UC. However, it is not clear if this is a genuine difference as reports are predominantly from hospital-based studies and could reflect a bias of the existing social structure in the region, perhaps making medical services more accessible to men than women.

UC Phenotype

Comparing the clinical features of UC across studies is fraught with difficulties, given the differences in diagnostic modalities, definitions of extent (some studies report the prevalence of limited proctitis, left sided colitis and extensive colitis but others only indicate disease extending into the sigmoid or beyond), time points of assessment, and statistical methods. This may explain why the literature appears inconsistent with several differences reported in disease location and behaviour even within the same ethnic group. For example, reports from mainland China demonstrate wide variations in the prevalence of extensive colitis (21.3% to 33.8%) (Jiang et al., 2006; APDW2004 Chinese IBD Working Group, 2006a). Studies from South Asia have described a more consistent phenotype between ethnic groups living in the same region: in a Sri Lankan

hospital-based study of 240 UC patients, 51.1% had proctosigmoiditis, 22.6% had left-sided colitis, and 22.3% had extensive colitis (Niriella et al., 2010). The prevalence of extensive colitis in the Sri Lankan study is similar to areas of India, which report the prevalence of extensive colitis to range between 16% and 20.3% (Kochhar et al., 1991; Sood et al., 2003) (**Figure 1-3**). The figures are similar to studies from Europe, with 28% of Portuguese patients having extensive colitis, but lower than Japan which reports a higher rate of extensive colitis (41%) (Fujimoto et al., 2007). However, as previously stated, the comparisons should be interpreted with caution as it is not clear whether maximal extent of disease was based on endoscopic, radiological or histological results and if full colonoscopy was performed in all cases.

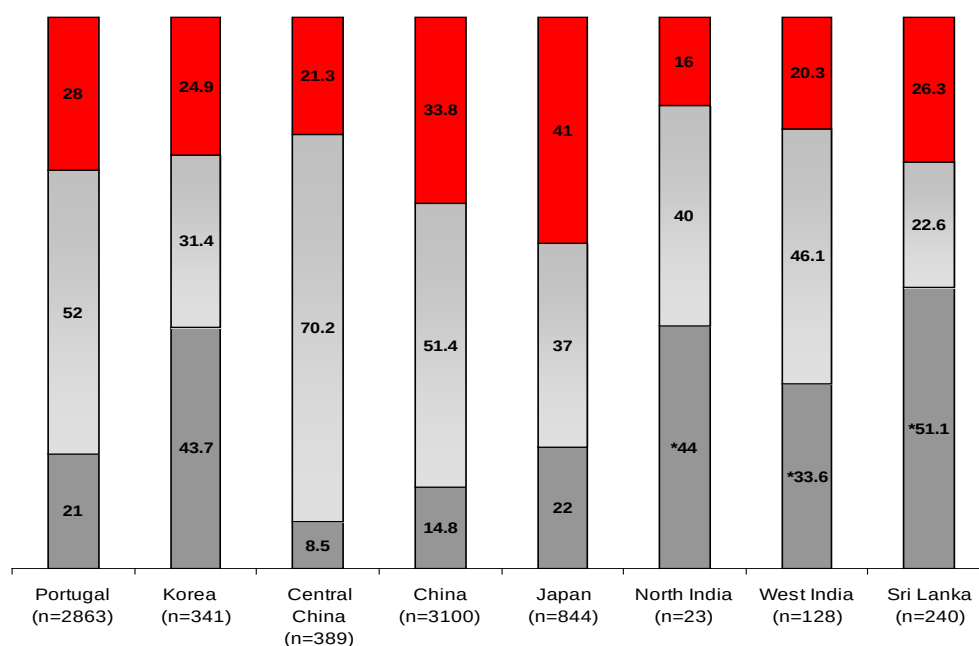


Figure 1-3: Stacked chart totalling 100%, depicting the distribution of UC in studies from Northern Europe and Asia.

(■) Extensive colitis, (■) left-sided colitis and (■) proctitis or *proctosigmoiditis

(Values adapted from Portela (2010), Sood et al., (2003); Kochhar et al., (1991), Niriella et al., (2010), Fujimoto et al., (2007) Jiang et al., (2006); APDW (2004) Chinese IBD Working Group (2006a)).

Extraintestinal manifestations in UC patients

Few papers from Asia have studied extraintestinal manifestations (EIMs) in detail. An early study by Kochhar et al.(1991) prospectively studied EIMs in 150 UC patients from the city of Chandigarh in North India. EIMs were seen in 34.7% of patients but included back pain without radiological changes, conjunctivitis and oral aphthous ulcers. Peripheral arthritis was the most frequently reported EIM (10.7%), most commonly involving the large joints. Eight percent of patients had backache with radiological evidence of sacroiliitis and 3.3% had anterior uveitis. Primary sclerosing cholangitis and skin manifestations were uncommon (1.3% and 2.7% respectively). A subsequent smaller study of 46 consecutive UC patients attending IBD clinics in the North West Indian state of Rajasthan showed only one person (2%) had peripheral arthritis and two had anterior uveitis (4%). No other EIMs were observed (Pokharna et al., 2004). The authors concluded that EIMs were rare in Indian patients in contrast to Kochhar's earlier study. A similar low frequency of EIMs has also been reported in China (6.1%) (Jiang and Cui, 2002). A large study of white British UC patients, showed 10.5% have arthritis and 3.1% anterior uveitis (Orchard et al., 1998).

Colectomy rates

A previous study from Japan reports that 33.9% of UC patients required surgery with cumulative colectomy rates of 4.1% in the first year, 6.8% in the second year and a 1–2% annual increase thereafter until the 14th year (Hiwatashi et al., 1991). This is comparable to early European studies which reported 32.4% - 37.6% of patients required colectomy, with cumulative rates of 20% after 5 years, 28% after 10 years, and 45% after 15 years (Farmer et al., 1993; Leijonmarck et al., 1990).

There are no reports of colectomy rates from the Indian subcontinent but a recent Sri Lankan study of UC patients reported a low colectomy rate of only 4.6% of cases. In this cohort, only 26.3% of patients had pancolitis and disease severity was mild in nearly 60% of cases, which may explain the lower colectomy rate. It is also possible that even those patients with extensive disease may have a disease phenotype amenable to medical therapy, although this was not specifically studied.

1.2.2 A comparison of the demographics and clinical phenotype of Crohn's disease between Asians and Caucasians

Early reports of CD in some areas of Asia were met with scepticism as major doubts persisted as to whether it was true CD as opposed to intestinal tuberculosis (TB) which is particularly common in some parts of India. However, in the past 10 years there has been a growing realisation that despite the high prevalence of TB, CD is becoming more common. This is likely to be related to increasing awareness and availability of diagnostic facilities.

Age & Gender

The peak age of onset of CD in Asian patients is reported as being in the second and third decade, similar to Caucasian studies (Yang et al., 2008; Leong et al., 2004; Yang et al., 2001). Interestingly, there appears to be an overwhelming male predominance in Asian CD patients, more than is seen in UC. In Western studies, there is generally an equal gender distribution or a moderate female predominance for CD (Vind et al., 2006; Bernstein et al., 2006). The reason for this difference is not clear but perhaps

higher rates of smoking in Asian males compared to females may influence the onset of CD in this ethnic group.

CD location and behaviour

The phenotype of CD is widely variable between white and Asian ethnic groups (**Table 1-1**). A large European study from Portugal reported that CD patients in this cohort had ileocolonic disease in 45% of cases (Magro et al., 2009) which is similar to a Canadian study that also reported a predominant phenotype of ileocolonic disease (42.5%) (Freeman, 2007). As the incidence and prevalence of CD in India is low, there are few reports on CD phenotype from this region. A single centre study from South West India (n=25) observed ileocaecal disease in 24% of patients and 76% had large bowel involvement (Pai and Khandige, 2000). A more recent multicentre study from North East India reported on 182 CD patients recruited over 7 years (Das et al., 2009). The commonest age of onset was 17-40 years (A2, as per the Montreal classification, Silverberg et al., 2005) with colonic disease (L2) being the most common disease location (41%). The predominant disease behaviour was nonstricturing, nonpenetrating (B1). Colonic disease location was also the most common finding from a Sri Lankan study of 55 CD patients of which 80% had only large intestinal involvement, the CD behaviour was not reported in this study (Niriella et al., 2010). In China, similar proportions of ileocolonic and colonic disease are observed (APDW2004 Chinese IBD Working Group, 2006b). In other areas of East Asia, the most common disease location is ileocolonic, as shown by a recent population based study from Korea (n=138). Sixty-seven percent of patients had ileocolonic CD, 21% had isolated small bowel and only 12% had colonic involvement at diagnosis (Yang et al., 2000). The findings are

consistent with a Japanese hospital-based cohort, where ileocolonic involvement predominated (65.8%) (Oriuchi et al., 2003).

Surgery for CD

Requirements for surgery are not well reported in the literature. The epidemiology study by Das et al. (2009) reports that 37% of Indian CD patients required surgical intervention, much higher than reported in a Sri Lankan study (18.2%) (Niriella et al., 2010). This may be a reflection of differences in disease location, as patients with ileal involvement are likely to develop more complicated disease behaviour requiring surgery. Only 20% of Sri Lankan patients had ileal or ileocolonic disease compared to 55% in the Indian cohort.

Extraintestinal manifestations in CD patients

There are few reports of EIMs in Asian CD patients. Das et al. (2009) showed that 13.7% of Indian CD patients had EIMs (after excluding back pain and oral ulceration). Peripheral arthropathy was present in 6% of Indian CD patients and 3% had sacroiliitis, less than a clinic-based CD population in the UK (Orchard et al., 1998).

Table 1-1: Ethnic variation in Crohn's disease phenotype (Values expressed in %)							
	Portugal (n=1413)	Canada (n=1015)	Korea (n=138)	Japan (n=183)	China (n=515)	India (n=182)	Sri Lanka (n=55)
<i>Location (%)</i>							
L1: Ileal	34	26.4	66.7	20	28.4	32	12.7
L2: Colonic	15	30.8	12.3	33	36.3	41	80
L3: Ileocolonic	45	42.5	66.7	46	35.3	23	7.3
L4: Upper GI only	1	0.3	-	1	-	4	5.5
<i>Behaviour (%)</i>							
B1: Nonstricturing, nonpenetrating	38	35.3	-	45	-	51	-
B2: Stricturing	26	33.6	-	31	-	24	-
B3: Penetrating	36	31.1	-	24	-	25	-
Perianal (p)	-	7.3	-	34	5.6	17	7.3
EIMs (%)	-	-	-	-	22.3	13.7	-
Surgery	-	-	-	-	-	37	18.2
Disease phenotype as per the Montreal Classification (Silverberg et al., 2005). EIMs, Extraintestinal manifestations, (-); information not available. <i>Information adapted from; Magro et al., (2009) [Portugal], Freeman (2007) [Canada], APDW2004 Chinese IBD Working Group, (2006)b [China], Das et al., (2009) [India], Niriella et al., (2010) [Sri Lanka], Yang et al., (2000) [Korea], Oriuchi et al., (2003) [Japan].</i>							

1.2.3 Rationale for studying migrant populations

IBD studies among migrant populations living in developed countries, with improved access to healthcare facilities, may provide more reliable data than those reported from underdeveloped parts of the world. These studies could also be helpful in determining if there are clinical differences related to ethnicity by comparing the migrant ethnic group with the host population.

The study of disease development in migrant groups who move from their place of birth to a new country can generate data related to the evaluation of environmental influences on the disease. These are of greatest value in identifying possible aetiological factors when the migrant population under investigation adopts the social, cultural and dietary habits of the host country. Provided inter-marriage with the host population is uncommon, the relative importance of genetic and environmental factors may become clearer (Probert et al., 1996b)). This epidemiological approach has been used for conditions such as diabetes and myocardial disease (Tillin et al., 2007; Chaturvedi et al., 2007; Chambers and Kooner, 2002; Chambers et al., 2001).

1.2.4 Epidemiological studies of migrant IBD populations living in Asia

Comparative prevalence data between different ethnic groups with IBD living in Singapore and Malaysia have previously been reported. In Singapore, the prevalence of UC in Asian Indians is high ($16.2/10^5$) compared to $6/10^5$ in Chinese and $4.8/10^5$ in Malaysians (Lee et al., 2000). A retrospective review from Malaysia, performed over a 13 year period (from 1985 -1998), also showed a high prevalence of UC in Indians ($17.9/10^5$) compared to $11.2/10^5$ in Chinese and $3.7/10^5$ in Malaysians (Tan and Goh,

2005). Phenotype studies have also been reported from these multi-racial countries and shown interesting differences in clinical phenotype between ethnic groups with UC. A study of 118 patients in Malaysia showed that Indian patients (n=48) had considerably higher rates of extensive colitis (60.4%: 29/48) compared to 33.3% (12/36) in Chinese and 43.3% (13/30) in Malaysia ($P=0.058$) (Hilmi et al., 2009). In keeping with other Asian countries, colectomy rates were also low in this study; Indian population 2.1% (1/48), Chinese 8.3% (3/36) and Malaysian 6.7% (2/30). The observed differences in UC phenotype seen in the Malaysian study are further strengthened by a Singaporean study showing 54% (26/48) of Indian patients had extensive colitis compared to 32% (54/169) of Chinese patients ($P=0.04$) (Ling et al., 2002).

Between 1950 and 1980 there was substantial migration of South Asians (originating from India, Pakistan and Bangladesh) to Europe, and currently 2.08 million South Asians live in the UK (Office for National Statistics 2001). This large-scale migration has opened up the possibility of epidemiological studies into environmental and genetic risk factors affecting the South Asian population. However, there are still only a few studies in the literature with reference to inflammatory bowel disease in UK South Asians.

1.2.5 Epidemiology and clinical phenotype of UC in South Asian migrants living in the UK

Epidemiology

A retrospective epidemiological study compared the incidence and prevalence of South Asian UC patients living in Leicestershire, UK between 1972-1989 with Europeans from the same area (Probert et al., 1992). One hundred and fifteen new UC cases and 29 cases of ulcerative proctitis were identified in South Asians during the study period. The standardised incidence of UC in the South Asian Hindu and Sikh population was over twice as high as that in Europeans (10.8, 16.5, 5.3 cases/ 10^5 /year respectively). This study also showed that South Asians had fewer operations for UC compared to Europeans, suggesting that UC may be less severe in UK South Asians.

Family studies show a positive history of IBD amongst Europeans more often than South Asians. In a study of 1254 patients in Leicester the risk to first degree relatives of European patients with UC was 15 fold compared to 3.5 in South Asians (Probert et al., 1993). This may indicate that environmental factors are more important than genetic factors in the South Asian IBD population.

The incidence rate for UC among indigenous Punjabi Indians (6.02/cases/ 10^5) (Sood et al., 2003), although higher than other Asian UC incidence rates, is still lower than that seen in the South Asian migrant study from Leicestershire, where UC incidence rates ranged from 13.1 to 14.3 cases/ 10^5 /year (Probert et al., 1992). Despite differences in methods and time period of assessment in both studies, this disparity in incidence rates

suggests the importance of environmental exposure, possibly acting as a triggering factor among South Asians who are genetically susceptible to IBD.

UC phenotype in UK South Asians

In Probert's original study (1992), there was no statistical difference in disease extent between Europeans and South Asians, however, South Asians tended to have more extensive disease (47%) and less proctitis (23%) compared to Europeans (39% and 36% respectively).

Differences in IBD phenotype have been reported in a small study of 82 South Asian IBD patients from Manchester, with South Asians demonstrating a higher rate of pancolonic disease (34/82: 41%) compared to a white British IBD cohort (41/158: 26%) (Rashid et al., 2008), but it was not clear whether this related to a combination of UC and CD patients or UC alone.

A later study, also from Leicestershire (Carr and Mayberry, 1999), studied the effects of migration on first and second-generation UC patients. In contrast to Probert's study (1992) that showed no differences between the disease phenotype in first- and second-generation South Asians, Carr and Mayberry suggested that second-generation patients are at a greater risk of extensive UC compared with first-generation South Asians. However, because of the small sample size (South Asian patients n=33), there is insufficient evidence to draw any firm conclusions about the influence of genetic and environmental factors.

Previous studies of UC in South Asians living in London have concentrated on the Bangladeshi community in East London (Jayanthi et al., 1992b; Tsironi et al., 2004). Tsironi's study in 2004 reported 16 new cases of UC in Bangladeshi migrants living in East London between 1997 and 2001. This correlated to an age-standardised incidence of 8.2 cases/ 10^5 /year, compared with an incidence of 2.3 cases/ 10^5 /year in the 1980s. In this study, the majority of UC patients had extensive disease (9/16; 56%) and only two patients had proctitis (12.5%), but the population size was too small to make a meaningful phenotypic comparison

1.2.6 Epidemiology and clinical phenotype of CD in South Asian migrants living in the UK

Epidemiology

There have been few studies of Crohn's disease in migrant South Asian populations. One study from the UK showed that the incidence of CD in South Asians living in Leicestershire has increased, particularly in South Asian Muslims (5.4 cases/ 10^5 /year) (Jayanthi et al., 1992a). Tsironi et al. conducted an epidemiological study to compare the incidence of IBD and abdominal tuberculosis in immigrants originating from Bangladesh (Tsironi et al., 2004). Nineteen new cases of CD in Bangladeshi migrants were reported during the study period 1997-2001. The mean age-standardised incidence of CD was 7.3 cases/ 10^5 /year in the period 1997-2001 compared with 2.3 cases/ 10^5 /year between 1981 and 1989. This indicates an increase in the incidence of CD and a fall in the incidence of abdominal tuberculosis (7.4 cases/ 10^5 /yr [1995-1989] - 2.5 cases/ 10^5 /yr [1997-2001]).

CD phenotype in UK South Asians

In Jayanthi's study (1992), the predominant disease location for both South Asians and Europeans was terminal ileal disease, and both groups had similar surgical requirements. Trironi et al. (2004) reported 47.4% (9/19) of newly diagnosed CD patients had disease limited to the terminal ileum (L1), 21% (4/19) had colonic disease only and the remaining patients (31.6%: 6/19) had ileocolonic disease. Nonpenetrating nonstricturing disease behaviour was the most common disease behaviour (9/16: 56.3%). Stricturing disease occurred in 6/19 patients (31%) and penetrating in 4 of 19 (21%).

1.3 INFLAMMATORY BOWEL DISEASE GENETICS

It is well known within the Caucasian population that IBD demonstrates a familial pattern of disease with a much higher disease frequency in first-degree relatives of affected individuals (CD: 14%, UC: 8.6%) compared to the general population (1.1%) (Peeters et al., 1996; Yang et al., 1993). Twin studies support the genetic basis for IBD, with a much higher rate of disease concordance demonstrated in monozygotic twins (11/63: 17%) compared with dizygotic twins (4/80: 5%), particularly in Crohn's disease (Thompson et al., 1996; Halfvarson et al., 2003).

Observations from twin studies led to significant interest in unravelling the genetic basis of IBD. The first major breakthrough came using linkage analysis - this technique involves identifying whether chromosomal segments are shared between family members with disease, compared to those without it. Using sibling-pair linkage, a locus was found on chromosome 16, labelled as IBD1, that was linked to CD (Hugot et al., 1996). Further susceptibility loci were found on chromosomes 3, 5, 6, 12, 14 and 19 (Hampe et al., 1999; Satsangi et al., 1996a; Rioux et al., 2001). Fine mapping studies could then be used to study specific genes within the susceptibility loci.

An alternative method to identify potential susceptibility genes in IBD is to use a candidate gene approach. In this approach biologically plausible genes are identified either by their position within regions of linkage, or by their potential function and are then studied in an association study comparing patients with healthy controls.

More recently genetic research has focused on Genome Wide Association (GWA) studies, which provide an analysis that is hypothesis-free, and can screen the whole genome for associations with IBD. GWA studies look for associations that are significantly higher than statistically expected and compare the allele frequency of patients with healthy controls. Using this technique, a number of immune-related genes have been identified, such as those involved in bacterial clearance via autophagy or the interleukin response.

To date, over 70 IBD susceptibility loci have been identified. Twenty-one have been confirmed in UC alone and 23 specifically for CD. A further 26 loci have been associated with both CD and UC. (Barrett et al., 2008; Mathew, 2008; The Wellcome Trust Case Control Consortium, 2007; Hugot et al., 2001; McGovern et al., 2010; Thompson and Lees, 2011). Genes of particular interest are discussed below.

1.3.1 Caspase activating recruitment domain 15 (*CARD15*)/Nucleotide-binding oligomerization domain 2 (*NOD2*)

The first gene contributing to Crohn's disease susceptibility was identified in 2001 by Hugot using both the candidate gene approach and fine mapping. It was located within the IBD1 susceptibility locus on chromosome 16. The gene was named caspase activating recruitment domain 15 (*CARD15*) or nucleotide-binding oligomerization domain 2 (*NOD2*) (Hugot et al., 2001). Of particular interest are three coding single nucleotide polymorphisms (SNPs) within the gene (*SNP8* [R702W], *SNP12* [G908R], *SNP13* [1007fsinsC]) (**Figure 1-4**), that are strongly and independently associated with Crohn's disease in many Caucasian populations and between them represent 82% of *CARD15/NOD2* polymorphisms in this population (Hugot et al., 2007). The *CARD15*

protein is expressed in the cytoplasm of peripheral blood monocytes, paneth cells and intestinal epithelial cells, and acts as an intracellular recognition protein for bacterial components such as peptidoglycans. When intracellular CARD15 recognises bacterial muramyl dipeptide, it activates the nuclear factor kappa B pathway (NF κ B) which regulates inflammation and therefore acts as a protective function (Ogura et al., 2001). The *CARD15/NOD2* polymorphisms Arg702Trp, Gly908Arg, and Leu1007fsinsC are associated with defective NF κ B activation by muramyl peptide (Hugot et al., 2001; Maeda et al., 2005). However, it is unclear whether the polymorphisms cause a loss of function (as shown by Ogura et al. (2001)) and thus are less effective at clearance of bacterial flora, or whether it causes an inappropriate activation of NF κ B in monocytes.

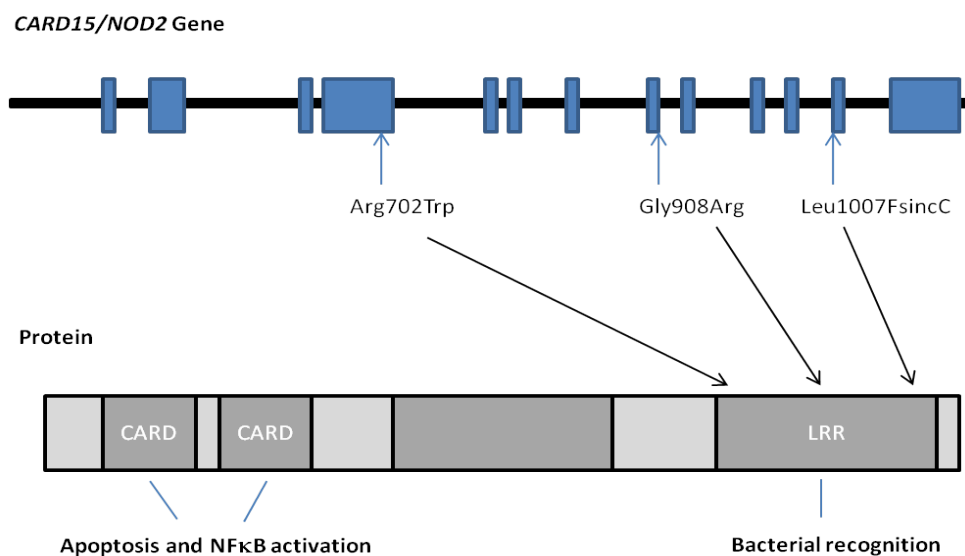


Figure 1-4: Structure of the *CARD15/NOD2* gene and protein showing polymorphisms Arg702Trp, Gly908Arg, and Leu1007fsinsC in the Leucine-rich Repeat Region (LRR).

(Adapted from Hugot et al, 2007)

In Northern European Caucasian patients with CD, 10-30% were heterozygous and 3-15% homozygous or compound heterozygous for the *CARD15/NOD2* polymorphisms. The corresponding proportions from the control population were 8-15% and 0-1% respectively (Hugot et al., 2007). A meta-analysis in 2004 calculated that the overall relative risk of developing CD in the Caucasian population was 2.4 for carriers of one *CARD15/NOD2* SNP and 17.1 for two SNPs (Economou et al., 2004).

1.3.2 Autophagy-Related-16-like 1 gene (ATG16L1)

Autophagy is a biological process of membrane trafficking in which autophagosomes engulf organelles and cytosolic macromolecules, followed by fusion of the autophagosome with a lysosome to form an autolysosome and sequestered material is degraded (Shih and Targan, 2009). Autophagy is important for cellular homeostatic functions including structural remodelling, nutrient and energy generation and degradation of damaged or long-lived cytoplasmic components. The role of autophagy as a defence against invading intracellular microorganisms has also been recognised (Nakagawa et al., 2004). It is proposed that any bacteria entering the cytoplasm by either invasion or escape from an endocytic vesicle are engulfed and compartmentalised within autophagosomes. The cell can then deploy a number of mechanisms to eliminate the bacteria eg. fusing with lysosomes and activating respiratory burst cascades (Radtke and O'Riordan, 2006). Within some cells, autophagy may also play a role in loading antigen into MHC binding groves for presentation at the cell surface (Munz, 2006). This protective role of autophagy is an important part of the innate immune response, but also links to adaptive immunity by delivery of foreign antigens necessary for immune recognition. Defective elimination of intracellular bacteria secondary to defects

in autophagy provides a plausible link with the pathogenesis of CD.

Recently, genome-wide association studies have identified a highly significant association ($P=0.039$) between Crohn's disease and the autophagy-related 16-like 1 gene (*ATG16L1*) on chromosome 2 at the *rs2241880* loci (Hampe et al., 2007). This was specific for CD and had a lack of correlation with UC ($P>0.04$). The loci *rs2241880* is a coding SNP that carries virtually all the disease risk exerted by the *ATG16L1* gene and is probably confined to individuals carrying susceptibility allele G at *rs2241880*. The frequency of the G allele was 0.6 in affected individuals and 0.53 in controls. Two additional genome wide association studies and an extension study from the Wellcome Trust consortium have been subsequently published confirming *ATG16L1* as a susceptibility gene in CD (The Wellcome Trust Case Control Consortium, 2007; Rioux et al., 2007).

The *ATG16L1* gene encodes a protein in the autophagosome pathway responsible for intracellular bacterial killing (**Figure 1-5**). The exact functional consequences of the *rs2241880* remain unclear. The human *ATG16L1* protein has an N-terminal APG16 domain consisting of coiled coils and eight C-terminal WD repeats. The *ATG16L1* polymorphism leads to an amino acid exchange (threonine to alanine) (Hampe et al., 2007) and is thought to lead to a defective immune responses in CD allowing commensal enteric bacteria to cause an overly aggressive T cell response and subsequent chronic inflammation (Hampe et al., 2007; Rioux et al., 2007; Sartor, 2008).

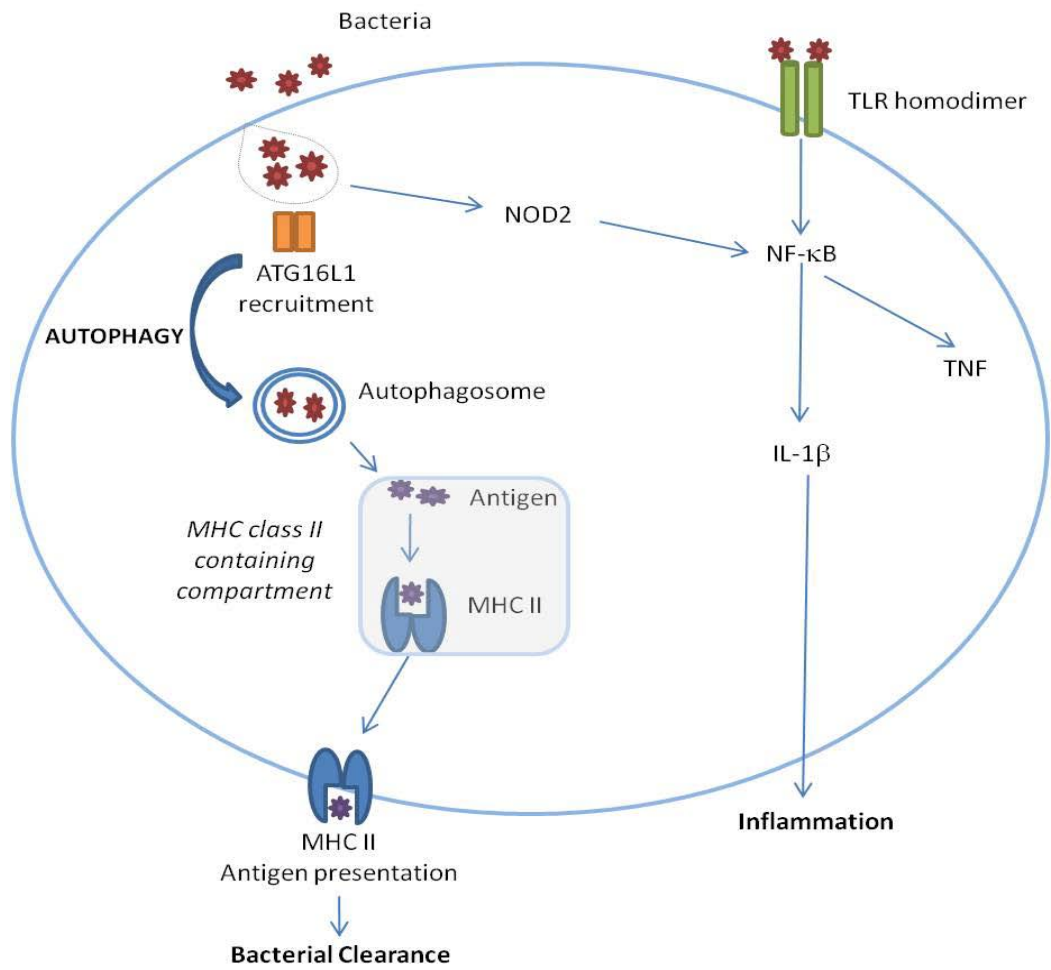


Figure 1-5: Diagram showing the role of *NOD2* proteins in NFκB production and the importance of *ATG16L1* in autophagosome formation.

(Note: It is unclear whether polymorphisms in *CARD15/NOD2* lead to loss of function or inappropriately high levels of NFκB production).

1.3.3 Immunity-Related GTPase protein M (*IRGM*)

In addition to *ATG16L1*, the Wellcome Trust Case Control Consortium (WTCCC) discovered SNPs near the Immunity-Related GTPase protein type M (*IRGM*) gene on chromosome 5q33.1 were associated with CD. The finding was replicated in an independent panel and the strongest replications were for SNPs *rs13361189* and *rs4958847* ($P = 6.6 \times 10^{-4}$ and 3.1×10^{-4} , respectively) (The Wellcome Trust Case Control Consortium, 2007). The *IRGM* gene does not seem to be involved in the predisposition to UC (Parkes et al., 2007).

The association of the *IRGM* locus with CD appears robust, but significant questions remain regarding the nature of its contribution to pathogenesis. The exact function of *IRGM* in CD aetiology and pathogenesis is unknown, but animal and human studies implicate *IRGM* as having a pivotal function in the autophagy pathway. *IRGM* deficient mice have reduced defence against intracellular pathogens such as *Toxoplasma gondii* and *Listeria monocytogenes* (Franke et al., 2008); similarly, human *IRGM* acts as its functional equivalent in the autophagy pathway and participates in conferring resistance against mycobacterial infections (Collazo et al., 2001). Initial sequencing of the *IRGM* gene by Parkes et al. did not identify any causal amino acid changes. However, a recent study by McCarroll et al. (2008) showed that a deletion polymorphism, immediately upstream of *IRGM* and in perfect linkage disequilibrium ($r^2 = 1.0$) with the most strongly CD-associated *IRGM* risk allele at the SNP *rs13361189*, could be the potentially causal variant.

For a long time it has been postulated that commensal intestinal bacteria predispose to CD but without any understanding as to how they might do so. Taken together, the genetics for *ATG16L1* and *IRGM* implicate defects in autophagy as a key part in the pathogenic pathway predisposing to CD.

1.3.4 Interleukin 23 Receptor Gene (*IL23R*)

In 2006, Duerr et al. identified the *IL23R* gene located on chromosome 1p31 as an IBD susceptibility gene in a GWA study testing 308,332 SNPs in 547 non-Jewish ileal CD cases and 548 controls. The *IL23R* SNP at *rs11209026* led to an amino acid change (Arg391Gln) and was associated with a lower risk for developing CD (Duerr et al., 2006).

This proved to have significant biological plausibility as the IL23 Receptor is stimulated by interleukin (IL) 23 and is pivotally involved in cytokine regulation via the T helper cell type (Th) 17 response. Th17 cells produce IL17 (Harrington et al., 2005; Annunziato et al., 2007), which acts on both immune and non-immune cells resulting in the recruitment of neutrophils at sites of inflammation, and triggers T cell proliferation and upregulation of a number of pro-inflammatory molecules (Awane et al., 1999; Laan et al., 2001) (**Figure 1-6**). IL17 plays a key role in host defence against bacteria, particularly at mucosal surfaces, by favouring the local recruitment of immune cells (Curtis and Way, 2009). Additionally, IL17 induces pro-inflammatory cytokine production by macrophages, thus creating a link between innate and adaptive immune responses (Kolls and Linden, 2004).

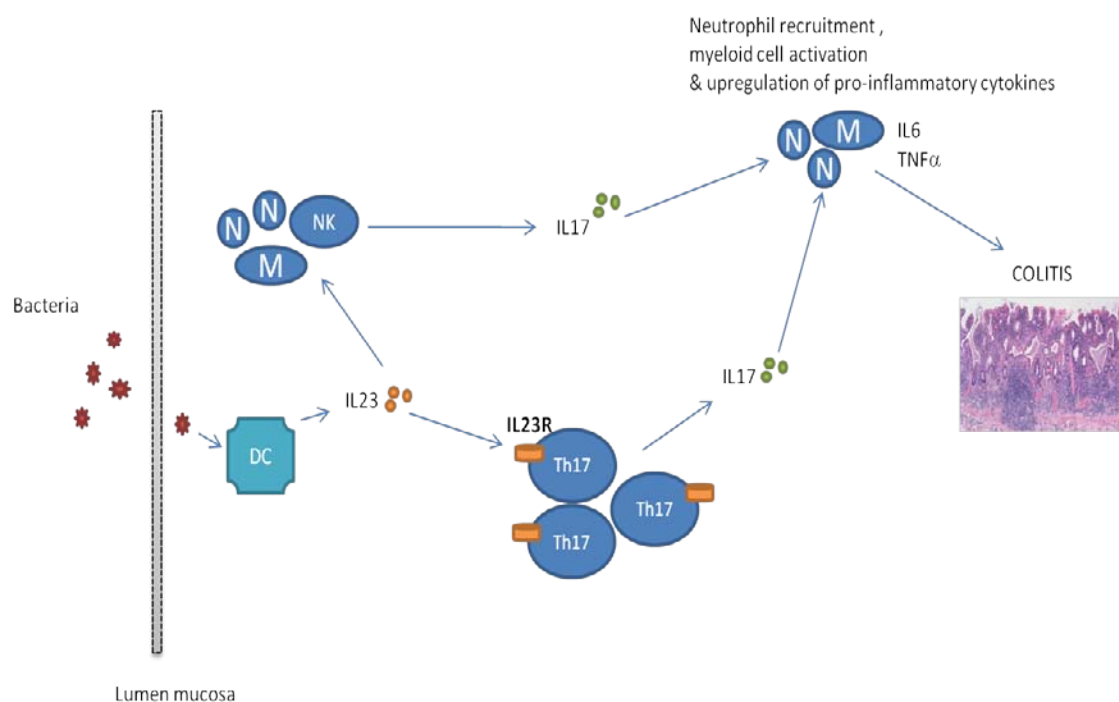


Figure 1-6: An illustration of the role of the Interleukin 23 receptor (IL23R) in bacterial clearance.

Adapted from McGovern et al. (2010).

Fuss et al. (2006) demonstrated that active inflammation in the colon of CD patients is associated with high levels of IL23 secretion and many chronically inflamed human tissues are infiltrated by highly differentiated Th17 lymphocytes (Pene et al., 2008). IL17 over-expression has been found in a number of inflammatory disorders, including IBD (Fujino et al., 2003). In addition, animal studies have suggested that IL23 may be responsible for developing intestinal inflammation within the terminal ileum (Becker et al., 2003). Therefore, a polymorphism in the IL23 receptor may reduce the subsequent Th17 and IL17 response leading to less inflammation and protection against CD.

The same SNP associated with CD (at *rs11209026*) has also been shown to confer protection in UC. (*rs11209026*, $P = 1.3 \times 10^{-8}$; OR=0.56) (Silverberg et al., 2009; McGovern et al., 2010). However, there is contrasting data on the molecular role of *IL23R* in UC. Early studies by Fujino (2003) using immunohistochemistry showed that there was higher IL17 expression in CD mucosa than in UC. In contrast, Kobayashi et al (2008) detected high levels of IL17 transcripts in both CD and UC mucosa compared to normal gut concluding that IL23 regulates the Th17 balance in UC and CD. The demonstration by Fuss et al. (2006) that IL23 is upregulated in active CD but not UC further confuses the picture, and makes it difficult to understand the precise role of the *IL23R* in UC.

1.3.5 Human Leukocyte antigen (HLA)

The IBD3 locus on the short arm of chromosome 6p, contains the genetic region known as the major histocompatibility complex (MHC). The human leukocyte antigens (HLA) are the products of genes mapped in the MHC region (**Figure 1-7**). The HLA complex is divided into three regions - class I, class II and class III. On the basis of their structure and function, class I antigens characterise the main products of HLA-A, -B and -C genes (as well as the less important (HLA-E, -F, -G) and class II antigens characterise the main products of HLA-DR, -DQ and -DP genes. In contrast, the class III region, located between class I and II genes, contains a diverse collection of over 20 genes that encode for complement, hormones, intracellular peptide processing, and other aspects of the immune response (Cassinotti et al., 2009).

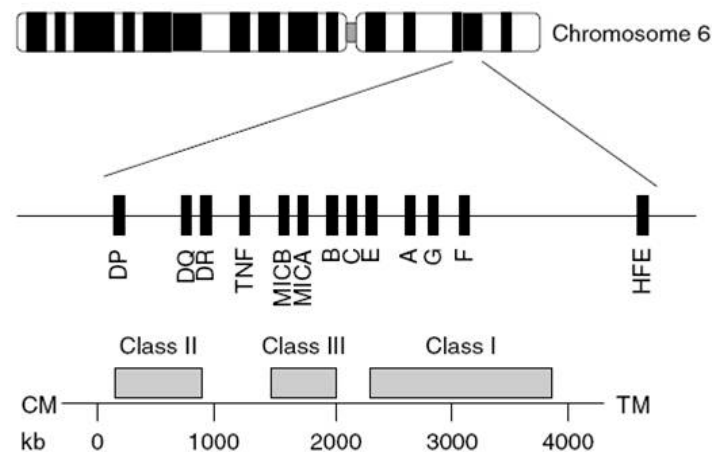


Figure 1-7: The major histocompatibility complex genes located on chromosome 6p.

(Adapted from Cassinotti et al. 2009).

For each of the HLA loci, multiple alleles exist in the population. At a single locus, the allele has a ‘co-dominant expression’, meaning that each individual will have two alleles for any given locus and both will be expressed. Class I antigens are expressed on the surface of most nucleated cells and platelets compared to Class II antigens that are expressed on B lymphocytes, monocytes, macrophages and dendritic cells. The key role of class I and II molecules is to bind to peptides and transport them to the membrane of cells for recognition by the T-cell antigen receptor. Class I molecules present mainly endogenously synthesised peptides (such as viral or tumour peptides) to CD8⁺ cytotoxic cells, whereas class II molecules present predominantly exogenous antigens (such as bacterial peptides) to CD4⁺ T helper cells.

Class I molecules are composed of a transmembrane α -chain, encoded by the HLA-A, -B or -C gene locus non-covalently bound to β -2 microglobulin. There are three types of HLA Class II molecules expressed by a single cell, namely HLA-DR, HLA-DQ and HLA-DP. These are each made up of $\alpha\beta$ heterodimers (**Figure 1-8**). The β chains are encoded by the DRB1, DQB1 and DPB1 genes respectively, which are all highly variable and the α -chains are coded for by the DPA and DQA genes which are also highly variable and the DRA gene which is invariant (Ahmad et al., 2006).

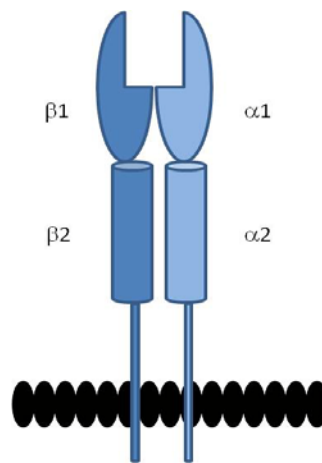


Figure 1-8: Structure of the HLA class II molecule showing the binding of $\alpha\beta$ heterodimers

The HLA system is made more complex by the fact that certain alleles occur together in the same haplotype more often than would be expected by chance, a term referred to as ‘linkage disequilibrium’ (LD). This is particularly apparent with HLA class II to the extent that HLA-DR alleles can be used to predict the HLA-DQ allele with a high degree of accuracy (Piancatelli et al., 2004).

HLA genes are highly polymorphic at each locus (DP, DQ, DR), this means that there are several alleles that may only differ by a few nucleotides in the DNA sequence, allowing many different HLA codifying sequences to exist in the human population. The combination of HLA genes on each chromosome 6 represents one HLA haplotype that is inherited from each parent.

Linkage within the HLA region has been confirmed in genome-wide scanning (*rs9469220*; $P=8.7 \times 10^{-7}$; The Wellcome Trust Case Control Consortium, 2007) and a meta-analysis combining all published linkage scans confirmed the HLA region to be the most consistent for linkage across all IBD studies (van Heel et al., 2004). The extensive linkage disequilibrium across the region has provided a great hurdle in defining the precise genes involved in disease susceptibility in the region, however, several linkage and association studies have implicated the HLA class II region in IBD susceptibility (Danze et al., 1996; Yang et al., 1999).

As HLA class II molecules bind peptides largely resulting from endosomal degradation of exogenous and endogenous proteins, they play a key role in eliciting immune responses to bacteria. This area is therefore of particular interest in IBD on the basis of its biological plausibility. The exact mechanism by which Class II genes exert their influence in IBD is unclear although one hypothesis is that in IBD, cross reactivity may exist between the peptides derived from bacterial luminal flora and from self-antigens in the gut. This may lead to the generation of auto reactive T cells, which contribute to disease pathogenesis through either stimulation, or inhibition of the immune system (Ahmad et al., 2006).

1.3.6 HLA class II genes and IBD

The *HLA-DRB1* gene is the most extensively studied HLA gene and has been implicated in determining susceptibility to both UC and CD. Although some associations have not been confirmed, others have been found more consistently (**Table 1-2**).

Table 1-2: <i>HLA-DRB1</i> associations with Crohn's disease and ulcerative colitis		
	Positive associations	Negative associations
Crohn's Disease	<i>DRB1*07</i> <i>DRB1*0103</i> <i>DRB1*04</i>	<i>DRB1*1501</i>
Ulcerative colitis	<i>DRB1*0103</i> <i>DRB1*1502</i>	<i>DRB1*04</i>
Information adapted from: Ahmad et al., 2002, Hancock et al., 2008, Stokkers et al., 1999, Yoshitake et al., 1999, Ahmad et al., 2003.		

HLA-DRB1 associations with CD

The *HLA-DRB1*0701* is the most replicated association with CD. The allele has a frequency of between 5% and 29% in European and North American populations but less than 1% in Japanese. Studies, from the United Kingdom (Ahmad et al., 2002) , Canada (Newman et al., 2004) and Spain (Fernandez et al., 2004) have shown that the *HLA-DRB1*07* allele is specifically associated with ileal disease (with or without colonic disease).

The association between *HLA-DRB1*0103* and CD has been widely replicated in European patients and is strongly associated with colonic location, particularly isolated colonic disease (Newman et al., 2004;Silverberg et al., 2003;Hancock et al., 2008). However, it is a relatively rare allele and occurs in less than 2% of the European and North American Caucasian population and is absent in the Japanese population. Studies have reported an association with *HLA-DRB1*0103* and perianal location, but this probably reflects their clinical association with colonic disease rather than an independent genetic association.

Associations with CD and the *HLA-DRB1*04* allele have been demonstrated in Japanese cohorts, in particular DRB1*0405 and *0410 (Yoshitake et al., 1999;Matake et al., 1992;Nakajima et al., 1995). These associations are not as strong in Northern Europeans/North American cohorts (Ahmad et al., 2003). *HLA-DRB1*1501* is a common allele, with a frequency of 6-25% in European and North American Caucasian populations and 6-10% in Japanese. It is negatively associated with overall CD susceptibility (Stokkers et al., 1999) and probably confers protection against CD in all ethnic groups.

HLA-DRB1 associations with UC

The *HLA-DRB1*0103* allele is the most consistently replicated association in Caucasian UC patients. The association is observed in both the *DRB1*0103-DQB1*0301* and *DRB1*0103-DQB1*0501* haplotypes, strongly implicating *DRB1*0103* in the pathogenesis of the disease (Stokkers et al., 1999). The association is strong in patients with extensive or severe disease (defined by the need for colectomy for failed medical management). However its low allele frequency (7.8% in UC patients vs 3.7% in controls) (Ahmad et al., 2003) limits its clinical value in predicting disease course.

*HLA-DRB1*1502* is associated with UC in Europeans (Ahmad et al., 2003), North Americans (Trachtenberg et al., 2000), Japanese (Futami et al., 1995; Yoshitake et al., 1999) and Korean populations (Myung et al., 2002). There is marked variation in the allele carriage rates between the populations with carriage of 20-25% in Japanese but <1% in Caucasian UC populations. Despite this, there are similar odds ratios for the gene in the disease population across the ethnic groups suggesting it is a true disease causing-variant. A small study in Japan (Masuda et al., 1994) suggested that *HLA-DRB1*1502* is associated with extensive and intractable UC but this has not been replicated in other East Asian populations (Myung et al., 2002). These results have to be interpreted with caution as in Japanese and British patients, UC has also been reportedly associated with *HLA-B*52* (Seki et al., 2001; Ahmad et al., 2003a), the class I allele in linkage disequilibrium with *DRB1*1502*.

The *HLA-DRB1*04* allele may be a protective allele as it is negatively associated with UC in North European populations (Stokkers et al., 1999). This is confined to the most prevalent subtype in the UK *HLA-DRB1*0401* but only when it is present on the

haplotypes *DRB1*0401 – DQB1*0301*.

It is not known whether a gene-dose effect operates, such that the possession of two high-risk HLA alleles confers additional risk. However, it is apparent from the literature that HLA associations vary with ethnicity, reflecting genetic heterogeneity and population-specific patterns of linkage disequilibrium across the HLA complex.

1.3.7 Correlation between genotype and IBD phenotype

Crohn's disease

Rigorous stratification of IBD patients by clinical phenotype has successfully shown an association between genotype and phenotype. In nearly all CD phenotype studies, *CARD15/NOD2* polymorphisms are significantly more prevalent in patients with ileal disease than those with CD limited to the colon (Lesage et al., 2002). A meta-analysis showed that the odds ratio for ileal compared with colon-only disease was 2.5 (95% CI 2.0 to 3.2) (Economou et al., 2004).

Parkes original association study of SNPs within the *IRGM* gene was performed in an unselected CD population which included all sub-types of CD as classified by disease extent and behaviour. On sub group analysis, there was no clear association with disease phenotype although 'pure ileal disease' tends to associate more strongly with the SNPs than 'pure colorectal' (Parkes et al., 2007). In contrast, an Italian study of 823 CD patients, Latiano et al. (2009) showed a genotype-phenotype association of the *IRGM rs4958847* polymorphism and fistulising CD behaviour ($P=0.037$). An explanation for the observed differences may relate to the much larger number of fistulising CD patients in the Italian CD cohort (25%), higher than that reported in British studies.

The first study noting the association of *IL23R* with CD was performed in patients with ileal CD, but as the authors chose to study only this subgroup (to minimise pathogenic heterogeneity) no phenotype comparisons are available. A later German study by Glas et al. included 833 CD patients of all phenotypes and 456 UC patients. In CD patients

with different *rs1209026* genotypes, there were no differences in the frequency of disease characteristics such as age of onset, disease behaviour, disease location and EIMs (Glas et al., 2007).

To date, there are no reports of association studies with *ATG* and a particular CD phenotype.

Ulcerative colitis

Phenotype studies in UC have shown the *HLA-DRB1*0103* allele is associated with extensive or severe disease, as defined by the need for colectomy for failed medical management (Ahmad et al., 2003). There is also evidence to suggest that the *HLA-DRB1*0103* allele is important in determining the extraintestinal manifestations associated with CD and UC (Orchard et al., 2002). In particular, there is a strong association between *HLA-DRB1*0103* and type 1 peripheral arthropathy (Orchard et al., 2000) with 35% of patients possessing the allele compared to 3% of controls. The allele has also been associated with pyoderma gangrenosum (Williams et al., 2007).

1.3.8 Variation in IBD genetics between different ethnic groups

Ethnic variation in CD susceptibility genes

Genetic studies in different ethnic groups have reported a widely differing carriage rates of IBD susceptibility genes. The CD-associated *CARD15/NOD2* polymorphisms are either rare or absent in Japanese, Korean and Chinese CD patients (Inoue et al., 2002; Lee et al., 2005; Leong et al., 2003). *CARD15/NOD2* polymorphism frequencies among African American and Hispanic CD children are also lower compared with white

children (Kugathasan et al., 2005). Two studies from India have reported the frequency of the three common *CARD15/NOD2* mutations in this population. The first study was a small case-control study from the Punjab of 25 South Asian CD patients. Neither the patients or controls showed any polymorphisms for *SNP 8*, *SNP12* or *SNP13* on *CARD15/NOD2* (Juyal et al., 2007). Ramakrishna's group from Vellore, Southern India, performed a larger study of 82 CD patients and 149 controls and also reported that none of the patients showed any of the three *CARD15/NOD2* polymorphisms (Pugazhendhi et al., 2008) (**Table 1-3**). Phenotypically, the patients in the study included 28 patients with ileocolonic disease and 26 patients with ileal disease, the patterns associated with the *CARD15/NOD2* mutations in Western populations. The absence of the common *CARD15/NOD2* mutations in any of these patients indicates a need to undertake more detailed and possibly non-hypothesis driven studies in this population.

Table 1-3: Frequency of the three <i>CARD15/NOD2</i> polymorphisms associated with Crohn's disease in patients of different ethnic origin						
		European (n=468)	African American (n=183)	Chinese (n=65)	North India (n=25)	South India (n=82)
<i>SNP</i>	<i>Allele</i>	%	%	%	%	%
rs2066844	C	89	96.7	100	100	100
Arg720Trp (<i>SNP8</i>)	T	11	3.3	0	0	0
rs2066845	G	94	98.9	100	100	100
Gly908Arg (<i>SNP 12</i>)	C	6	1.1	0	0	0
rs2066847	WT	88	96.2	100	100	100
Leu107fsinsC (<i>SNP13</i>)	Ins C	12	3.8	0	0	0
<i>Allele frequencies adapted from Hugot et al. (2001), Leong et al., (2003), Kugathasan et al., (2005), Juyal et al.,(2007), Pugazhendhi et al., (2008).</i>						

Yamazaki et al. (2009) recently published data on the frequency of CD associated polymorphisms in the Japanese population (Cases=484, Controls=470). Compared to results reported by the Wellcome Trust Consortium (2007) they demonstrated very variable background gene frequencies between the two populations (**Table 1-4**) and showed no association with *IL23R*, *ATG16L1* or *IRGM* in the Japanese CD patients. There are no data on the frequency of these polymorphisms in South Asian populations.

Table 1-4: Summary of the genetic differences between UK and Japanese populations in single nucleotide polymorphisms (SNPs) associated with Crohn's disease susceptibility

			European population		Japanese population	
			MAF		MAF	
SNP	Chr	Gene	Case	Control	Case	Control
rs11805303	1p31	<i>IL23R</i>	0.39	0.32	0.57	0.55
rs10210302	2q37	<i>ATG16L1</i>	0.40	0.48	0.24	0.24
rs4958847	5q33	<i>IRGM847</i>	0.15	0.11	0.51	0.50
rs13361189	5q33	<i>IRGM189</i>	0.10	0.07	0.36	0.33
Chr., chromosome; MAF, minor allele frequency as defined in the control cohorts						

As previously mentioned, the *HLA-DRB1*04* allele association is stronger in Japanese CD patients compared with Northern Europeans/North American cohorts (Ahmad et al., 2003a; Nakajima et al., 1995). The only reported data on associations of HLA-DR and DQ antigens in Indian CD patients came from a family study of 9 CD patients, 52 first degree relatives and 110 ethnically matched controls. This showed no significant differences between the groups but in view of the small study numbers, it is not possible to draw any conclusions from this study (Thakur et al., 2003).

Ethnic variation in UC susceptibility genes

Juyal's *CARD15/NOD2* study, also reported the frequency of *CARD15/NOD2* polymorphisms in Indian UC patients (Juyal et al., 2007). As with CD the frequency of SNPs 8,12 & 13 were low. However, the group subsequently resequenced all the coding regions in the *CARD15/NOD2* gene and revealed another two SNPs in this population. One of the polymorphisms at rs2066842 (*SNP5*) was more frequent among

UC cases compared to controls (17% vs 12%; Bonferroni adjusted P value=0.02) but did not reach significance for CD. SNP5 conferred modest risk for UC (OR=1.72, 95% CI: 1.17-2.52). This is interesting, as polymorphisms in *CARD15/NOD2* have been traditionally associated with ileal CD. Genotype-phenotype analysis revealed no associations between SNP5 and UC disease extent or EIMs. Although this SNP has been associated with CD in a French study ($P=0.0003$) (Lesage et al., 2002), it is of note that a previous study demonstrated this variant occurs in phase with the R702W, G908R, and 1007fs mutations and that the association was thought to be related to linkage disequilibrium (Hugot et al., 2001).

The *HLA-DRB1*0103* subtype is extremely rare in the Japanese and Korean populations however, the Japanese and Korean UC population show a significantly increased frequency of the *HLA-DRB1*1502* allele (Myung et al., 2002; Masuda et al., 1994; Futami et al., 1995). This allele is associated with UC in European, North American, Japanese and Korean populations with similar odds ratios of 2 to 4.5 despite variable background prevalence, suggesting it is a disease causing variant (Futami et al., 1995; Trachtenberg et al., 2000). The Japanese study by Masuda et al. (n=45) showed the gene frequency of *DRB1*1502* was significantly higher in patients with total colitis compared to left-sided and proctitis ($P<0.05$) but this association with UC phenotype was not observed in the Korean study (Myung et al., 2002).

There is no published data with respect to UC and HLA Class II associations in the South Asian population and there is a need for further research in this area.

1.4 METABOLIC PROFILING - EVIDENCE FOR THE ROLE OF THE MICROBIOTA IN THE PATHOGENESIS OF IBD

The clinical features of IBD are likely to be the result of interactions between genetically susceptible individuals and environmental stimuli, the key environmental trigger being the host intestinal microbiota. The gut microbiota was shown to be important in IBD after early animal models of IBD demonstrated that colitis does not occur when experimental animals are reared in germ-free conditions (Bhan et al., 1999), but the introduction of a gut bacterial load resulted in rapid mucosal inflammation (Rath et al., 2001). Subsequent human studies using fluorescent in situ hybridization (FISH) have shown that patients with IBD have an altered gut microbiota compared to controls, with bacterial changes localised to the luminal communities (Kleessen et al., 2002). Sokol et al. used flow cytometry to analyze the bacterial composition of faecal samples and showed that species of Clostridia, especially those involved in short chain fatty acid production, are reduced in both UC and CD compared with healthy controls (Sokol et al., 2006).

Therefore, the chronic intestinal inflammation seen in IBD is likely the result of an abnormal immunological response to commensal bacteria within the gut lumen in a genetically susceptible host (**Figure 1-9**).

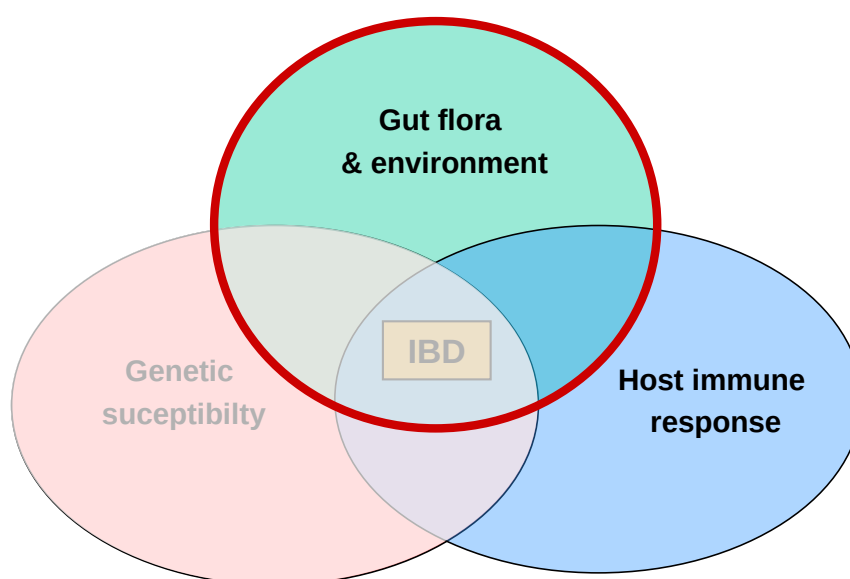


Figure 1-9: Model of IBD pathogenesis. The interaction of susceptibility genes, disease phenotype genes, the host immune response and environmental stimuli, contribute to the pathogenesis of ulcerative colitis and Crohn's disease.

As no particular species of bacteria has been shown to cause IBD (Lodes et al., 2004), this has led to the hypothesis that a dysbiosis of the microbiota may be involved either in initiating the disease or maintaining it, in a genetically susceptible host (Marchesi et al., 2007).

There is a clear need to better understand the role of the gut flora in disease pathogenesis. Nuclear Magnetic Resonance Spectroscopy (NMRS) provides a potential method for studying the relationship between patients and the intestinal flora. It allows the generation of metabolic information based on the analysis of biofluids and provides an opportunity to identify specific biomarkers of disease dependent on underlying bacterial metabolic processes.

1.4.1 Principles of nuclear magnetic resonance spectroscopy (NMRS)

NMRS is a non-invasive technique which allows the study of the molecular composition of biofluids. It is the interaction between atoms, a static magnetic field and a radiofrequency pulse that defines NMRS. The nucleus consists of protons and neutrons that determine its charge and mass. The nuclei of certain atoms, hydrogen-1 (^1H), carbon-13, (^{13}C) phosphorous -31 (^{31}P) and fluorine-19 (^{19}F) rotate around an axis and possess a quantum mechanical property known as 'spin'. The 'spin' provides angular momentum to the charged particle producing a magnetic field analogous to that produced by movement of electrons in a loop of wire. The nucleus is said to behave like a miniature bar magnet (Cobbold J, 2005) (**Figure 1-10**).

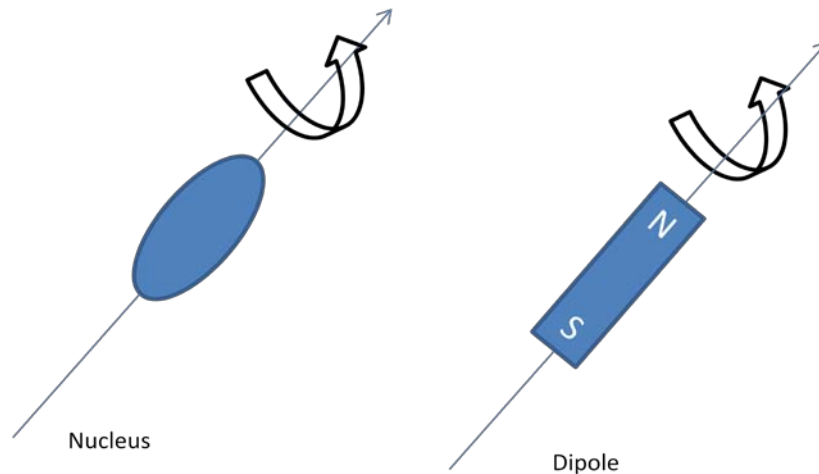


Figure 1-10: A charged spinning nucleus creates a magnetic field causing it to behave like a dipole or bar magnet.

In NMRS the sample is subjected to a strong magnetic field to align the magnetic dipoles of the protons in the sample. A radiofrequency pulse is applied at 90^0 to the external magnetic field to stimulate the nuclei into a higher energy level. When the pulse is removed, the nuclei relax to their normal energy level producing radiofrequency signals which contain information about each nucleus and therefore about the molecules in which they exist. Therefore, a nucleus within any particular molecule has a characteristic chemical shift. The signal that is detected is known as a free induction decay but is difficult to interpret in the time domain. Fourier transformation (a mathematical operation that decomposes a signal into its constituent frequencies) converts the time domain into the frequency domain producing the typical NMR spectra (Figure 1-11).

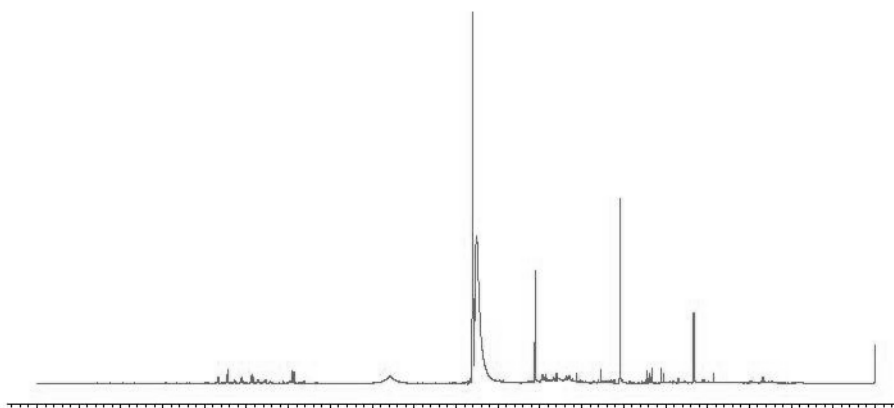


Figure 1-11: A typical NMR spectra which contains signals from hundreds of metabolites.

The technique detects a wide variety of metabolites from host and gut microbial metabolism, thus maximising opportunity for novel biomarker discovery. The use of metabolic profiling has been extensively explored because of its potential to identify biomarkers, enhance clinical diagnosis and expand hypothesis generation (Bjerrum et al., 2008).

1.4.2 Clinical applications of NMRS

NMR based spectroscopy has proven particularly successful in rapid analysis of complex biological samples. Each spectrum that is generated, yields a unique metabolic fingerprint for each biological sample and is the product of genetic and environmental contributions (eg. diet, lifestyle, gut microbial activity) under a particular set of conditions. Analyzing metabolites (molecules <1 kDa) in body fluids using spectroscopic methods provides knowledge of the metabolic phenotype. The metabolic profile provides information that cannot be obtained directly from the genotype, gene expression profiles or even the proteome of an individual.

Using a hypothesis-free approach, spectra can be compared between disease groups and controls and this has identified metabolic profiles associated with a number of diseases including diabetes, hypertension and osteoarthritis (Chung et al., 2005;Holmes et al., 2008;Lamers et al., 2005). In circumstances where previous metabolic characteristics have been identified, a hypothesis-driven approach can be used to study specific metabolites within the spectrum.

1.4.3 The role of NMRS in the characterisation of IBD

Previous experiments have demonstrated that the host microbiota significantly contributes to the overall metabolic profile generated from biofluids (Wikoff et al., 2009;Li et al., 2008;Marchesi et al., 2007), therefore any changes in the gut microbiota will be reflected in the spectroscopic profile of the biofluid being examined.

Studies using NMR spectroscopy of colonic mucosal biopsies have been able to distinguish healthy control tissue from that of IBD patients and to further distinguish ulcerative colitis from Crohn's disease (Bezabeh et al., 2001). The diagnostic accuracy was 97.9% for the differentiation of IBD tissue from healthy controls however, this technique is invasive and open to sampling error.

Non-invasive metabonomic approaches to help in the diagnosis of IBD have acquired much interest. In 2007 a group from Ireland created metabolic profiles of faecal extracts from IBD patients and healthy controls (Marchesi et al., 2007). The faecal extracts of both CD and UC patients were characterized by reduced levels of microbiota-related metabolites (butyrate, acetate, methylamine, and trimethylamine) in comparison with a control population. This is consistent with the theory that a disruption or dysbiosis of the normal gut bacterial ecology is seen in patients with IBD.

A new and exciting approach to non-invasive NMRS is the use of urine samples to analyse specific metabolites dependent on either gut bacterial metabolism alone or co-metabolism of the host and the intestinal microflora. Urinary NMR profiling has the significant advantage of having easy of sample collection. There is good evidence from animal studies to support differences in the urinary NMR profile of healthy animals and those with intestinal inflammation: Experiments comparing mice with dextran sulphate sodium (DSS) induced colitis and healthy mice, showed that the urinary metabolic profile was clearly distinguishable between the groups (Schicho et al., 2010). Large amounts of methylamines (dimethylamine [DMA], trimethylamine [TMA] and its oxidation product trimethyl-amine n-oxide [TMAO]) that are produced by gut bacteria were found to be elevated in the urine of DSS-treated mice. Metabonomic studies in an IL10 deficient colitis mouse model also displayed higher urinary levels of metabolites

produced and or modulated by the gut microflora (TMA and TMAO) compared to a cohort of healthy mice (Murdoch et al., 2008).

Studies using urinary NMR spectroscopy have recently been performed in Caucasians with IBD. NMRS of the urine has shown distinct urinary metabolic profiles for CD and UC (Williams et al., 2009). In particular, levels of hippurate (a metabolite dependent on the co-metabolism of host and gut flora) were significantly lower in IBD patients when compared to a healthy control cohort. This novel technique has never been used to study other ethnic populations with IBD. Comparing the metabolic profiles of IBD patients of different ethnicities could help determine if there is variation in the microbiota between ethnic groups. It will also help determine if the biomarkers already identified in Caucasian populations, will have potential worldwide applicability in IBD diagnosis.

1.5 SEROLOGICAL MARKERS IN INFLAMMATORY BOWEL DISEASE PATIENTS

The chronic intestinal inflammation seen in IBD is likely to be the result of an aberrant immunological response to commensal bacteria within the gut lumen. Many patients with UC and CD exhibit serological responses to microbial antigens and this could reflect the relevant microorganisms involved in pathogenesis or may simply reflect the increased permeability of the gut. In 1959, investigators first reported antibodies to colon extract in patients with ulcerative colitis (Broberger and Perlmann, 1959). Since that time, several antibodies have been described. The two most widely studied serologic markers in IBD are atypical perinuclear anti-neutrophil cytoplasmic antibody (pANCA) and anti-Saccharomyces cerevisiae antibody (ASCA). More recently, additional serum autoantibodies including anti outer membrane porin C (anti-OmpC) have been reported as showing an association with IBD (Landers et al., 2002).

1.5.1 Atypical perinuclear anti-neutrophil cytoplasmic antibody (atypical pANCA)

Perinuclear anti-neutrophil cytoplasmic antibody (pANCA) is an autoantibody directed against neutrophil granules. Testing is performed on serum by indirect immunofluorescence of peripheral blood neutrophils. Two pANCA patterns are detectable: the typical pANCA pattern of staining which results in homogeneous rim-like staining of the perinuclear cytoplasm and the atypical pANCA reactivity that is recognised by a broad non-homogeneous rim-like staining of the nuclear periphery with

multiple intranuclear foci (**Figure 1-12**).

The atypical pANCA staining is the most frequent ANCA pattern seen in patients with IBD, especially UC. In the literature, the atypical pANCA pattern may be referred to as 'pANCA with diffuse cytoplasmic staining' or 'x-ANCA'. The target antigen of atypical pANCA has not been clearly identified but it is likely that it is associated with the inner side of the neutrophil nuclear membrane (Terjung et al., 2000).

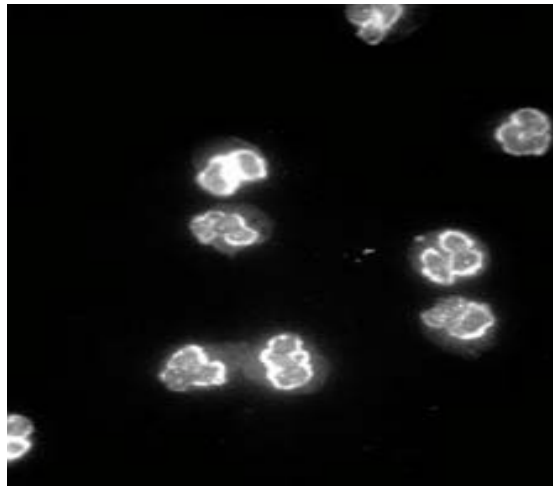


Figure 1-12: Atypical pANCA staining showing a broad inhomogeneous rim-like staining of the nuclear periphery with multiple intranuclear foci.

(Image from Ruffatti et al (2002))

Detection of atypical pANCA in the sera of UC patients remains considerably higher (40- 80%) than CD patients (5-25%) (Saxon et al., 1990;Rump et al., 1990;Bossuyt, 2006). Although the test is reasonably specific (84-95%) the sensitivity is not optimal (48-63%) (Bossuyt, 2006).

1.5.2 Anti-Saccharomyces cerevisiae antibodies (ASCA)

Anti-Saccharomyces cerevisiae antibodies (ASCA) are directed primarily against mannose residues (oligomannosides) expressed in the cell wall of the yeast *Saccharomyces cerevisiae* (Main et al., 1988). Different hypotheses concerning the origin of ASCA in IBD have been proposed but must take into account the widespread distribution of oligomannosides shared by many molecules present in a wide variety of organisms. ASCA-IgG reactive antigens have been reported in the granulomas of bowel resections and on infiltrating lymphocytes and neutrophils in inflamed tissue from CD patients (Oshitani et al., 2003). However, it is not clear whether ASCA detected by using *S. Cerevisiae* are generated by this microbe or if they represent cross-reactivity with others microorganisms carrying similar mannose structures (Sendid et al., 1996).

As with atypical pANCA, ASCA positivity is variable but studies have shown a higher positivity in CD patients (50-80%) compared with only 2-14% positivity in UC patients and 1-7% of healthy controls (Peeters et al., 2001;Quinton et al., 1998). Specificity for ASCA can be as high as 90% and sensitivity is in the range of 41-76% (Vermeire et al., 2001).

1.5.3 Antibodies against Outer Membrane Porin C (anti-OmpC)

The antibody anti-OmpC is directed against the outer membrane porin C of *Escherichia coli*. Detection of the IgG antibody is performed using ELISA. In CD, the reported seroprevalence is 55% (Landers et al., 2002). The prevalence in UC and healthy controls was low (2% and 1.3% respectively). Anti-OmpC may be of particular value to aid in the diagnosis of ASCA negative CD patients.

1.5.4 IBD phenotype associations with serological markers

Atypical pANCA

Associations between atypical pANCA and UC phenotypes have been inconsistent. A pilot study from the Mayo Clinic in 1996 showed an association with atypical pANCA and a severe left sided ulcerative colitis which is resistant to treatment. The disease had an aggressive course requiring earlier surgery (Sandborn et al., 1996). These findings have not been replicated. Interestingly, atypical pANCA has been associated with a subgroup of CD patients characterised by UC-like Crohn's colitis; the inflammation usually involves the left side of the colon and the response to therapy is generally good (Vasiliauskas et al., 1996). Complications such as stricturing or perforation were lower in this group (Klebl et al., 2003). Positivity of pANCA in this group makes it often difficult to distinguish colonic CD from UC serologically.

ASCA and anti-OmpC

The expression of antibodies to *Saccharomyces cerevisiae* (ASCA) has been associated with a CD phenotype that is remarkably similar to that of patients with the common *CARD15/NOD2* polymorphisms associated with IBD. ASCA seropositive patients present at a younger age and are more likely to have ileal disease (Dassopoulos et al., 2007;Quinton et al., 1998;Peeters et al., 2001). They are also at greater risk for complicated disease behaviour (Vasiliauskas et al., 2000) and require surgery more frequently (Forcione et al., 2004). In patients with an ASCA+ (IgG and/or IgA) /atypical pANCA- phenotype, small bowel involvement (with or without colonic disease) is more typical than pure colonic disease (Peeters et al., 2001;Mow et al., 2004). Stratification of CD patients based on the presence of anti-OmpC showed an association with penetrating disease behaviour and therefore potentially leading to a higher rate of surgical intervention (Mow et al., 2004). A subgroup of CD patients do not appear to have ASCA antibodies. Whether these individuals form a subgroup of patients with specific clinical characteristics is unknown at present.

1.5.5 The frequency of IBD serological markers in South Asian IBD patients

The only available information on the frequency of serological markers in South Asian IBD patients comes from two small Indian studies. The first, a study of 59 CD patients, 25 with UC and 21 healthy controls that recorded the frequency of ASCA (IgA or IgG) as being 61% in CD, 40% in UC and 9.5% in controls. For atypical pANCA, the prevalence was 10.1%, 32% and 0% respectively (Makharia et al., 2007). A second Indian study investigated 16 CD, 36 UC and 12 controls and recorded similar

frequencies of ASCA positivity (62% CD, 26% UC and 8% controls) (Ghoshal et al., 2007). Both studies tried to use ASCA to differentiate between CD and patients with abdominal TB but neither showed any statistical difference. To date, there have been no studies on the frequency of anti-OmpC in the South Asian population.

There have also been no studies in the literature directly comparing the frequencies of antibodies associated with IBD between different ethnic groups. It is difficult to draw any reliable comparisons between data obtained in different countries due to varying methodologies and inter-observer variability.

As serological markers for IBD become more sensitive, they may become useful in clinical practice for assisting in the diagnosis of IBD and discriminating between UC and CD. In view of this, it is of interest to determine if the sensitivity and specificity of serological markers is the same in different ethnic groups.

CHAPTER 2

AIMS AND HYPOTHESES

1. The study outlined in Chapter 3 aimed to investigate the phenotype of IBD in a cohort of UK South Asians. A second aim was to determine if there are phenotypic differences in IBD between South Asian and white Northern Europeans living in the UK.

Hypothesis: Inflammatory bowel disease phenotype is different between South Asian and white Northern European IBD patients living in the UK.

2. The aim of the study described in Chapter 4 was to examine the role of previously described genetic associations in a South Asian IBD population; specifically *CARD15/NOD2* (rs2066844, rs2066845, rs2066847 & rs2066842), *ATG16L1* (rs2241880), *IL23R* (rs11209026), *IRGM* variants (rs1000113 & rs4958847) and HLA class 2 alleles (*HLA-DRB1*0103* & *HLA-DRB1*1502*). The study also compared the frequencies of IBD associated polymorphisms in South Asians with a cohort of white Northern European IBD patients.

Hypothesis: The frequency of genetic polymorphisms associated with IBD differs between South Asians and white Northern Europeans.

3. The aim of Chapter 5 was to use urinary metabolic profiling to determine if there are differences in microbiota-related metabolites between South Asian IBD patients and South Asian controls. The NMR results from South Asian IBD patients were compared with a matched white Northern European IBD population to investigate if there are differences in urinary metabolites between IBD patients of different ethnicity.

Hypothesis: Metabolic differences seen in the urinary nuclear magnetic resonance spectra of IBD patients are independent of ethnicity and characteristic of IBD.

4. The study reported in Chapter 6 aimed to determine the diagnostic value of serological markers (atypical pANCA, ASCA IgA & IgG and anti-OmpC) in a South Asian cohort of IBD patients and to investigate any association of these serological markers with clinical phenotype. The results from the South Asian cohort were compared with a white Northern European IBD cohort to determine if there is ethnic variation in the sensitivity and specificity of these markers.

Hypothesis: The sensitivity and specificity of serological markers does not differ between ethnic groups.

CHAPTER 3

DEFINING THE CLINICAL PHENOTYPE OF INFLAMMATORY BOWEL DISEASE IN UK SOUTH ASIANS

3.1 INTRODUCTION

Epidemiological studies from around the world suggest a widely varying incidence of IBD, with highest rates of UC and CD reported in Northern Europe and North America (Loftus, Jr., 2004). Previous research in IBD has predominantly focused on large white European/American populations, but recent studies suggest there may be phenotypic differences in other ethnic groups (Hou et al., 2009; Bhat et al., 2009).

Between 1950 and 1980 there was substantial migration of South Asians (originating from India, Pakistan and Bangladesh) to Europe, and currently 2.08 million South Asians live in the UK (Office for National Statistics 2001). Studies of IBD in this population could provide valuable information that may help to distinguish the degree of influence of genetic and environmental pathogenic factors. However, to date studies have been scarce and have involved relatively small patient numbers.

A retrospective epidemiological study compared the incidence and prevalence of UC among South Asians living in the city of Leicester between 1972-1989 with Europeans from the same area (Probert et al., 1992). South Asian patients tended to have more extensive disease and less proctitis compared to Europeans but this did not reach statistical significance. However, differences in IBD phenotype were reported in a small study of 82 South Asian IBD patients from Manchester, with South Asians demonstrating a higher rate of pancolonic disease (34/82: 41%) compared to a white British IBD cohort (41/158: 26%) (Rashid et al., 2008), but it was not clear whether this related to a combination of UC and CD patients or UC alone.

There have been relatively few studies of CD in the UK South Asian population (Tsironi et al., 2004;Jayanthi et al., 1992a;Probert et al., 1993). Jayanthi et al. identified only 28 new cases of CD in the city of Leicester between 1972 and 1989 compared to 179 new cases in Europeans (Jayanthi et al., 1992a). In this study, disease behaviour, prevalence of extraintestinal manifestations and smoking history were not investigated, but the predominant disease location for both South Asians and Europeans was the terminal ileum, and both groups had similar surgical requirements. However, recent genetic studies from South Asia have demonstrated significantly lower carriage of CD-associated polymorphisms in *CARD15/NOD2* (the gene predominantly associated with ileal Crohn's disease) than in Northern Europeans (Economou et al., 2004;Juyal et al., 2007;Pugazhendhi et al., 2008), and the lack of phenotypic differences in the Leicester study may reflect its small size.

3.1.1 Study aims

Many South Asians live in London, and approximately 220,000 have settled in the North West London boroughs of Brent, Ealing, Harrow and Hounslow, making up over 22% of the population in this area (Office for National Statistics 2001, 2004). The primary aim of the study was to determine the phenotype of IBD in a South Asian population and to compare it with a white Northern European cohort from the same area. A secondary aim was to investigate the impact of environmental factors on disease phenotype by comparing that of first-generation and second-generation South Asians. A move towards a Northern European phenotype in second-generation South Asians could suggest that Western environmental stimuli are important pathogenic factors in determining disease phenotype.

3.2 MATERIALS AND METHODS

3.2.1 Ethics Approval

All patients gave informed consent and the study was approved by the St. Mary's Hospital Research and Ethics committee with site-specific approval from each participating hospital trust.

3.2.2 Patient Recruitment

To capture the maximum number of South Asian IBD patients in North West London, the IBD and gastroenterology outpatient clinics of five hospitals (St. Mary's and Hammersmith Hospitals of Imperial College Healthcare NHS Trust, St. Mark's Hospital, Ealing Hospital and West Middlesex University Hospital) were monitored for an eighteen-month period (01.08.2008 – 01.02.2010). In North West London, current practice is that IBD patients are managed in secondary care by the gastroenterology team. Primary care physicians may be involved in joint care for example, managing disease flares or blood monitoring while on immunosuppressive therapy but patients remain under regular review in secondary care. Each of the five hospitals has a dedicated IBD outpatient clinic. Within these hospitals, patients with IBD are referred to this clinic. In addition, patients seen in the IBD clinic include:- new referrals from primary care with signs and symptoms of IBD, new IBD patients referred by the consultant physician or surgeon at the time of hospitalisation, patients referred from endoscopy with IBD, IBD patients with an acute relapse and patients with quiescent

disease under regular follow-up. To ensure complete coverage of the population, all the general gastroenterology clinics were also reviewed for South Asian patients with IBD. A white Northern European population of comparable size was recruited from the IBD clinics of the same five hospitals; patients of mixed race were excluded. White Northern Europeans represent the predominant ethnic group attending the IBD clinics in North West London. Northern European patients were recruited from all five sites consecutively for a six-month period between 01.08.2009 and 01.02.2010.

3.2.3 Case Definition

The clinical notes of all patients attending the IBD/gastroenterology clinics were screened by the gastroenterologist at the time of the patient's outpatient appointment. Only patients with a confirmed diagnosis of IBD based on clinical evaluation, endoscopic/radiological appearances and histology (Van et al., 2010; Stange et al., 2008), were recruited for the study .

UC was defined on clinical history/examination with endoscopic and histological appearances of continuous mucosal inflammation of the colon, affecting the rectum and a variable extent of the colon in continuity, and characterised by a relapsing and remitting course (Stange et al., 2008). CD was defined on the basis of clinical evaluation and a combination of endoscopic findings (mucosal ulceration, skip lesions, cobblestone appearance and terminal ileal ulceration), histological appearance (discontinuous intestinal inflammation and/or non-caseating granulomas) with or without radiological evidence of skip lesions, stricturing disease, or fistulising disease. Stool culture results and mucosal histology were reviewed for all patients and those in whom there was any uncertainty regarding diagnosis (ie. those with IBD, type

unclassified (Silverberg et al., 2005) or colitis as a result of infection, radiation or ischemia) were excluded from the study.

Previous studies from the UK have defined the South Asian population as patients whose families originated from the Indian subcontinent (India, Pakistan and Bangladesh) including South Asians who had migrated via East Africa (Probert et al., 1993; Carr and Mayberry, 1999). For comparison, this study used the same definition of South Asia and only included patients both of whose parents were of South Asian origin. All patients were aged ≥ 16 years.

3.2.4 Patient demographics and clinical phenotype

A questionnaire proforma was completed prospectively by the attending doctor at the time of the outpatient appointment. Data collection included; gender, ethnic origin, place of birth, ethnicity and place of birth of the patient's parents, length of time in the UK, extraintestinal manifestations, a current drug history and a smoking history. Patients were classed as smokers if they had smoked more than seven cigarettes per week for at least 6 months immediately prior to diagnosis of IBD and/or during follow-up. Non-smokers were those who never smoked, smoked a little and occasionally, or had stopped smoking more than 6 months before IBD was diagnosed (Cosnes et al., 1996; Cottone et al., 1994; Lakatos et al., 2009). The patient's clinical notes were then carefully reviewed to gather additional demographic and phenotypic details including age at IBD diagnosis, length of follow-up, disease extent, disease behaviour, surgical history and to confirm any extraintestinal manifestations.

A second gastroenterologist (Dr Horace Williams, Imperial College London) randomly audited 200 sets of clinical notes from patients recruited to the study in order to confirm

the clinical details were accurately documented. There was disagreement regarding one case and this was resolved after discussion.

3.2.5 Defining IBD disease behaviour and extent

In the participating hospitals, it is standard protocol for patients with suspected IBD to have a full colonoscopy with ileal intubation, and biopsies taken from the terminal ileum, right colon, transverse colon, left colon and rectum. Only UC patients that had either evaluation of the entire colon by colonoscopy with biopsy or who had a colectomy with histology reports were included in the study. UC disease extent was recorded by both endoscopic appearance and maximum histological extent. When comparing the ethnic groups, maximal disease extent was determined by histology and defined in terms of the Montreal classification (ulcerative proctitis (E1), left-sided UC (E2) and extensive colitis (E3)) (Silverberg et al., 2005). Only CD patients who had undergone a full colonoscopy with ileal intubation and biopsy (when not precluded by stenosis) were included in the study. Disease extent and behaviour was based on a combination of histological extent and small bowel imaging. Patients had an oesophagogastroduodenoscopy (OGD) with biopsy when clinically indicated. Disease behaviour was considered cumulatively to the time of the most recent outpatient visit. The Montreal classification was also used to classify CD location (terminal ileal (L1), colonic (L2), ileocolonic (L3) and upper GI (L4)) and disease behaviour (nonstricturing, nonpenetrating (B1), stricturing (B2) and penetrating (B3)) (Silverberg et al., 2005). Subjects with both stricturing and penetrating disease were classified as penetrating disease (B3). Perianal fistulating disease alone did not constitute penetrating disease but was recorded as a modifier of the disease behaviour (p). Surgical history was

documented and included surgery performed in North West London and elsewhere (colectomy for UC patients, and bowel resection or treatment of related complications in CD patients).

3.2.6 Extraintestinal manifestations

Extraintestinal manifestations (EIMs) were determined from patient recall with confirmation in the clinical notes. EIMs were defined as peripheral arthropathy (Type 1 or 2) (Orchard et al., 1998), ankylosing spondylitis/sacroiliitis, skin complications (erythema nodosum or pyoderma gangrenosum), ocular complications (anterior uveitis or episcleritis), primary sclerosing cholangitis and venous thrombosis.

3.2.7 Assessing the impact of environmental factors in the UK on disease development and phenotype

To determine the potential influence of environmental factors in the UK on the development of IBD in South Asians, the rates of IBD diagnosed in the 1st, 2nd, 3rd and subsequent decades since migration were examined. In addition, the potential impact of environmental factors on disease phenotype was assessed by comparing South Asians born in the UK with first-generation South Asians diagnosed within 10 years of arriving in the UK. This cohort was specifically chosen as being representative of a group of patients that had less exposure to environmental stimuli in the UK compared to South Asians who had been born and grown up in the UK. The phenotype of second-generation South Asians was directly compared to the Northern European cohort to

determine if there was a move towards a Northern European phenotype in second-generation South Asians.

3.2.8 Statistical Analysis

Data from the South Asian cohort were compared with the Northern European cohort using the Mann-Whitney U test for continuous variables, such as duration of disease and age at diagnosis and the chi-square test for categorical variables, such as disease extent and need for surgical intervention. The disease extent in first and second-generation South Asians was compared using the chi-squared test. *P* values <0.05 were considered statistically significant.

To account for potential false positives resulting from multiple comparisons, a Bonferroni correction was applied to *P* values in the phenotypic data that showed significance ($P < 0.05$) (by a factor of 3 in the UC cohort, to allow for comparisons relating to; disease location, requirements for surgery and extraintestinal manifestations and by a factor of 4 in the CD cohort, to allow for comparisons of disease extent, disease behaviour, surgical intervention and extraintestinal manifestations). Corrected *P* values are expressed in the text as *P_c*. The data were analysed using Graph Pad Prism Software® (Version 5, U.S.A.).

3.3 RESULTS

A total of 770 well-characterised IBD patients were recruited for the study (South Asian 367, Northern European 403). In the South Asian population UC was the predominant disease (UC 74.4%: CD 25.6% [273:94]) whereas in the Northern European population there were similar proportions of UC and CD (46.7%:53.3% [188:215]).

3.3.1 UC demographics

Of the 273 South Asian UC patients identified, there was a male predominance (1.4:1), in contrast to the Northern European population which had more female patients; however, this did not reach statistical significance ($P=0.06$) (**Table 3-1**). The median age at diagnosis was 30.7 years (inter-quartile range 23.3 – 42.1 years) for the South Asian population. Two hundred and thirty-one patients (84.6%) were Indian, 39 (14.3%) Pakistani and 3 (1.1%) Bangladeshi. All patients had a diagnosis of IBD confirmed in the UK. The majority of South Asian patients (67.0%) were aged between 17 and 40 years old at diagnosis (Montreal classification A2). No significant differences were observed in smoking habits between the South Asian and European cohorts (Positive smoking history: 6.9% vs 7.4% [$P=0.85$]).

Table 3-1: Demographic details of South Asian and white Northern European ulcerative colitis patients living in North West London, UK			
	South Asian (n=273)	Northern European (n=188)	P value
Gender Male: Female, n (%)	159 (58.2) : 114 (41.8)	92 (48.9) : 96 (51.1)	$P = 0.06$
Age at diagnosis (Median [IQR]; yrs)	30.7 (23.3 - 42.1)	33.2 (25.1 - 43.2)	$^{\dagger}P = 0.11$
Age at diagnosis (Montreal): n (%)			
≤16 (A1)	10 (3.7)	9 (4.8)	$P = 0.54$
17-40 (A2)	183 (67.0)	125 (66.5)	$P = 0.96$
>40 (A3)	80 (29.3)	54 (28.7)	$P = 0.99$
Duration of follow-up (Median [IQR]; yrs)	10.5 (5.6 – 18.1)	9.3 (4.8 – 18.9)	$^{\dagger}P = 0.27$
*Smoker: n (%)	19 (6.9)	14 (7.4)	$P = 0.85$
IQR, inter-quartile range. Age at diagnosis expressed as per the Montreal classification (Silverberg et al., 2005) $^{\dagger}P$: calculated using the Mann-Whitney U test. All other P values calculated using the χ^2 test. *Smoked more than 7 cigarettes per week for at least 6 months immediately prior to diagnosis of IBD and/or during follow-up.			

3.3.2 UC phenotype

The clinical features of the South Asian patients were compared with the Northern European cohort (**Table 3-2**). The disease phenotype differed significantly between the two populations, with 63.0% of South Asian patients having extensive colitis compared to 42.5% of the Northern European cohort ($P < 0.0001$). The prevalence of left-sided disease was similar in each group (27.1% and 31.4%), but limited rectal disease was less common in South Asian UC patients (9.9% vs 26.1% in Northern European patients ($P < 0.0001$)) (**Figure 3-1**). Twenty-four (8.8%) of the South Asian UC population had

at least one extraintestinal manifestation (EIM) of IBD. This was similar to the Northern European population (6.9%). The most common EIM in South Asians was peripheral arthropathy (10 patients). In addition, forty-four South Asian UC patients (16.1%) complained of arthralgia, but with no joint swelling or radiographic changes

Table 3-2: Phenotypic details of South Asian and white Northern European ulcerative colitis patients living in North West London, UK				
	South Asian (n=273)	Northern European (n=188)	P value	Corrected P value (Pc)
Disease location (Montreal): n (%)				
Ulcerative proctitis (E1)	27 (9.9)	49 (26.1)	<i>P</i> < 0.0001	[‡]<i>P</i>_c < 0.0001
Left-sided UC (E2)	74 (27.1)	59 (31.4)	<i>P</i> = 0.36	
Extensive UC (E3)	172 (63.0)	80 (42.5)	<i>P</i> < 0.0001	[‡]<i>P</i>_c < 0.0001
Patients requiring colectomy: n (%)	18 (6.6)	20 (10.6)	<i>P</i> = 0.16	
Extra-intestinal manifestations: n (%)				
Peripheral arthropathy	10 (3.7)	5 (2.7)	<i>P</i> = 0.47	
Ankylosing spondylitis / Sacroiliitis	4 (1.5)	1 (0.5)	<i>P</i> = 0.43	
[#] Skin complications	4 (1.5)	4 (2.1)	<i>P</i> = 0.82	
Uveitis / Episcleritis	6 (2.2)	4 (2.1)	<i>P</i> = 0.82	
PSC	3 (1.1)	2 (1.1)	<i>P</i> = 0.97	
Venous Thrombosis	2 (0.7)	1 (0.5)	<i>P</i> = 0.79	
[#] Erythema nodosum and pyoderma gangrenosum: PSC, primary sclerosing cholangitis; Disease location expressed as per the Montreal classification (Silverberg et al., 2005) [‡] <i>P</i> _c : After Bonferroni correction the values remain significant. All <i>P</i> values calculated using the χ^2 test. Bold values are <i>P</i> values that showed significance (<i>P</i> < 0.05).				

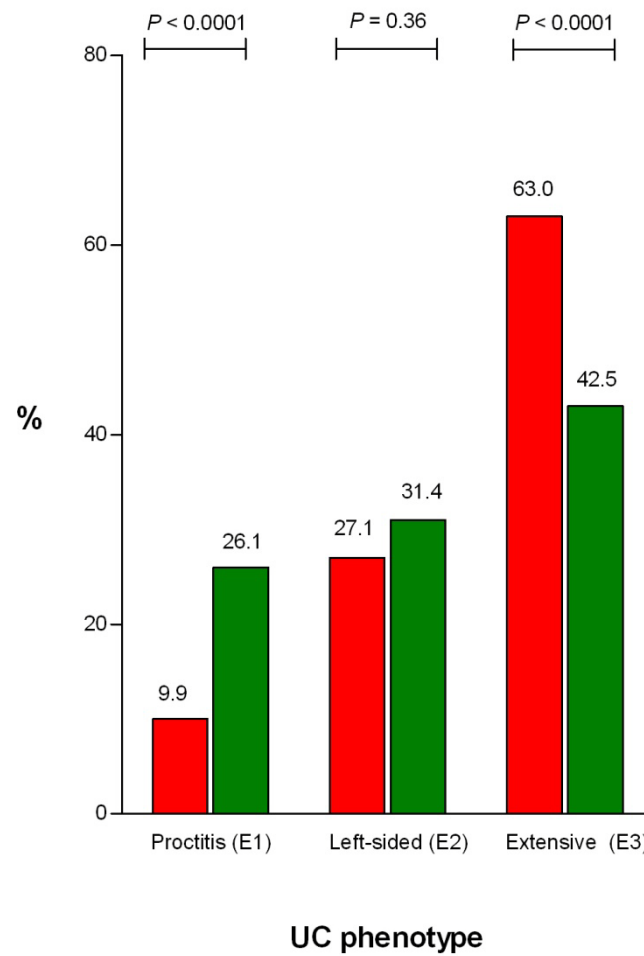


Figure 3-1: Bar graph comparing the ulcerative colitis phenotype of South Asian and white Northern Europeans patients.

(■) South Asian and (■) white Northern European patients. P values were calculated using the χ^2 test. A Bonferroni correction was applied to significant P values.

Colectomy rates in South Asians were numerically lower than the Northern European population but this did not reach statistical significance. Eighteen South Asian patients (6.6%) had undergone resections for UC (16 for failed medical management and 2 for dysplasia related to UC). Three patients had an end ileostomy and 15 had restorative ileo-anal pouches. This compared with 20 Northern European patients (10.6%) who had undergone colectomies (19 for failed medical management and 1 for dysplasia). Nearly all patients in both cohorts requiring colectomy had extensive colitis (17/18 in the South Asian cohort and 17/20 in the Northern European cohort). When patients with extensive colitis that required surgery were compared, only 17/172 South Asian patients with extensive colitis required a colectomy compared to 17/80 in the Northern European cohort (9.9% vs 21.3% $P=0.01$).

3.3.3 CD demographics

Of the 94 South Asian CD patients, the majority were Indian (87.2%). Median age at diagnosis was 30.7 years (interquartile range 21.4–42.1 years) which was similar to the Northern European cohort (**Table 3-3**). In both cohorts, the most common age range at diagnosis was 17-40 years (A2) but there were significantly fewer South Asian patients in the ≤ 16 years group (A1) ($P=0.03$). Only 10 South Asian patients had a positive smoking history (10.6%) compared to 68 Northern Europeans (31.6%) ($P<0.0001$).

Table 3-3: Demographic details of South Asian and white Northern European Crohn's disease patients living in North West London, UK

	South Asian (n=94)	Northern European (n=215)	P value
Gender (Male: Female: n (%))	48 (51.1) : 46 (48.9)	94 (43.7) : 121 (56.3)	P= 0.23
Age at diagnosis (Median [IQR]; yrs)	30.7 (21.4 – 42.1)	25.9 (20.1 - 38.2)	[†] P= 0.09
<i>Age at diagnosis (Montreal): n (%)</i>			
≤16 (A1)	5 (5.3)	30 (13.9)	P= 0.03
17-40 (A2)	61 (64.9)	139 (64.7)	P= 0.97
>40 (A3)	28 (29.8)	46 (21.4)	P= 0.11
Duration of follow-up (Median [IQR]; yrs)	13.2 (7.1 – 20.9)	11.6 (6.1 - 19.6)	[†] P= 0.39
*Smoker: n (%)	10 (10.6)	68 (31.6)	P<0.0001

IQR, inter-quartile range. Age at diagnosis expressed as per the Montreal classification (Silverberg et al., 2005)

[†]P: calculated using the Mann-Whitney U test. All other P values were calculated using the χ^2 test.

*Smoked more than 7 cigarettes per week for at least 6 months immediately prior to diagnosis of IBD and/or during follow-up.

Bold values are P values that showed significance ($P<0.05$).

3.3.4 CD phenotype

The South Asian CD phenotype was mainly colonic (L2=46.8%) and ileocolonic (L3=37.2%). Isolated ileal disease was less common (L1=16.0%). Disease in the Northern European population was predominantly ileocolonic (L3=40.5%) and although there was less isolated colonic disease than in the South Asians (34.8% vs 46.8%) this was not significantly different after correcting for multiple comparisons ($P_c=0.19$) (Figure 3-2).

The clinical behaviour pattern at last follow-up was nonstricturing, nonpenetrating (B1), in 72.3% South Asian patients, stricturing (B2) in 23.4% and penetrating (B3) in 4.3% of cases. Length of follow-up for the CD cohorts did not differ significantly. Nineteen patients (20.2%) had perianal CD, including perianal fistulae and fissures. There was significantly less B3 behaviour compared to the Northern European cohort ($P_c=0.04$). In all, 26.6% of South Asian patients had had surgical treatment for CD (including bowel resection, stricturoplasty and surgical treatment of perianal disease). This was significantly lower than in the Northern European cohort (44.7%, $P_c=0.01$).

In order to determine if surgical intervention was related to disease phenotype, a subgroup analysis was performed in all patients that required surgery. In the South Asian cohort, 40% (10/25) of surgical intervention was for B1 behaviour, 44% (11/25) for B2+B3 behaviour and 16% (4/25) for perianal disease. This was different to the Northern European population where the vast majority of operations were for B2+B3 behaviour (62.5%: 60/96) and only 16.6% (16/96) for B1 behaviour and 20.9% (20/96) for perianal disease.

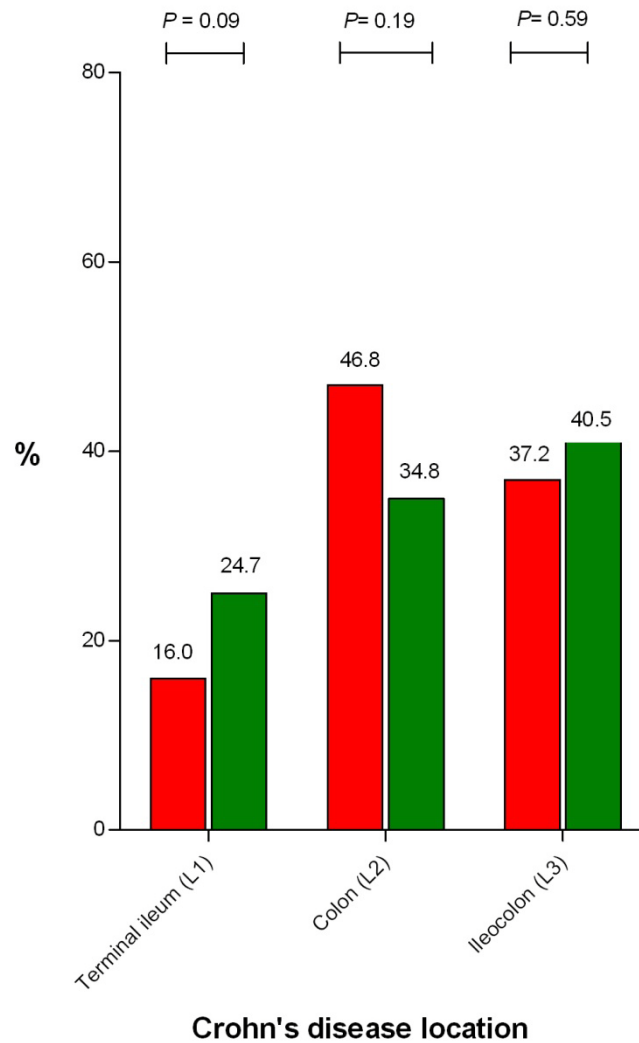


Figure 3-2: Bar graph comparing Crohn's disease location in South Asians and white Northern Europeans.

(■) South Asian and (■) white Northern European patients. *P* values were calculated using the χ^2 test. A Bonferroni correction was applied to significant *P* values.

There were no significant differences in EIMs between South Asian and Northern European cohorts ($P=0.18$). Skin complications were the most common EIM in South Asian CD patients (**Table 3-4**).

Table 3-4: Phenotypic details of South Asian and white Northern European Crohn's disease patients living in North West London, UK				
	South Asian (n=94)	Northern European (n=215)	P value	Corrected P value (<i>P_c</i>)
<i>Disease Location (Montreal): n (%)</i>				
Terminal ileum (L1)	15 (16.0)	53 (24.7)	<i>P</i> = 0.09	
Colon (L2)	44 (46.8)	75 (34.8)	<i>P</i> = 0.048	<i>P_c</i> = 0.19
Ileocolon (L3)	35 (37.2)	87 (40.5)	<i>P</i> = 0.59	
Upper GI (L4)	0	0	-	
Upper GI modifier	10 (10.6)	16 (7.4)	<i>P</i> = 0.35	
<i>Disease Behaviour (Montreal): n (%)</i>				
Nonstricturing, nonpenetrating (B1)	68 (72.3)	125 (58.1)	<i>P</i> = 0.02	<i>P_c</i> = 0.08
Stricturing (B2)	22 (23.4)	60 (27.9)	<i>P</i> = 0.41	
Penetrating (B3)	4 (4.3)	30 (14.0)	<i>P</i> = 0.01	[‡]<i>P_c</i> = 0.04
Perianal disease modifier (p)	19 (20.2)	46 (21.4)	<i>P</i> = 0.81	
Surgery for Crohn's disease n (%)	25 (26.6)	96 (44.7)	<i>P</i> = 0.003	[‡]<i>P_c</i> = 0.01
<i>Extraintestinal manifestations n (%)</i>				
Peripheral arthropathy	6 (6.4)	27 (12.6)	<i>P</i> = 0.11	
Ankylosing spondylitis / Sacroiliitis	1 (1.1)	5 (2.8)	<i>P</i> = 0.46	
*Skin complications	7 (7.4)	18 (8.4)	<i>P</i> = 0.78	
Uveitis / episcleritis	3 (3.2)	7 (3.3)	<i>P</i> = 0.98	
PSC	1 (1.1)	1 (0.5)	<i>P</i> = 0.55	
Venous Thrombosis	0	5 (2.3)	<i>P</i> = 0.14	
GI, gastrointestinal; PSC, primary sclerosing cholangitis, *Erythema nodosum and pyoderma gangrenosum Disease location expressed as per the Montreal classification (Silverberg et al., 2005) [‡]<i>P_c</i>: After Bonferroni correction the values remain significant. All <i>P</i> values calculated using the χ^2 test. Bold values are <i>P</i> values that showed significance (<i>P</i> < 0.05).				

3.3.5 Demographics of first-generation South Asian IBD patients

The exact place of birth was not available for 17 South Asian patients (10 UC patients and 7 CD patients), the analysis was performed on the remaining 350 patients (263 UC and 87 CD). The majority of South Asian patients studied were first-generation (born outside the UK) (64.6%: 226 /350) and the remainder were second-generation (born in the UK). Eighty-three percent of first-generation patients came directly from the Indian subcontinent (187/226) and the remaining 17% immigrated from East Africa (39/226). The median year that first-generation South Asian patients arrived in the UK was 1975 (IQR; 1968-1988).

To assess the effect of changing environment on the development of IBD in first-generation patients, the rates of diagnosis in the 1st, 2nd, 3rd and subsequent decades since migration were examined. Of the 226 first-generation South Asian patients (171 UC and 55 CD), three patients had a provisional diagnosis of IBD before arriving in the UK (1 UC and 2 CD). For UC, of the remaining 170 patients, more were diagnosed in the first decade (34.7%: 59/170), compared to 19.5% (33/170) in the second decade, 22.9% (39/170) in the third decade and 22.9% (39/170) in subsequent decades. Thus for UC there was no substantial increase in disease diagnosis with time spent in the UK. In contrast in the CD cohort (although small), very few of the first-generation patients were diagnosed within 10 years of arrival in the UK (7.6%: 4/53). The majority of patients were diagnosed in the 2nd (37.7%: 20/53) and 3rd decades (41.5% 22/53). The remaining 13.2% (7/53) were diagnosed after being in the UK >30 years.

3.3.6 Phenotypic comparison of first and second-generation South Asian IBD patients

To assess the potential impact of environmental factors in the UK on disease phenotype, first generation South Asians diagnosed within the first 10 years of arriving in the UK were compared to South Asians born in the UK. There was no significant difference in phenotype between the UC groups, and the predominant UC phenotype was still extensive colitis (first-generation UC diagnosed within 10 years: 34/59: [57.6%] vs second-generation 63/92 [68.5%] $P=0.17$). The numbers for CD were too small to be able to make any meaningful comparisons.

When South Asian UC patients born in the UK were compared to the Northern European UC population, there were again significantly higher rates of extensive UC in the South Asian cohort (63/92: [68.5%] vs 80/188 [42.5%]: $P<0.0001$) despite a shorter follow-up period (median duration of follow-up: 8.5 years vs 9.3 years respectively).

3.4 DISCUSSION

This is the largest phenotypic study of IBD in the UK South Asian population, examining the clinical characteristics of 367 patients. The study shows that UC is more common than CD in the North West London South Asian IBD population (UC 74.4%; CD 25.6%), and that the disease extent in South Asian UC patients is significantly different from that of a white Northern European cohort. Disease behaviour and need for surgery differed significantly between the South Asian and white Northern European CD cohorts.

Studies such as this may be prone to ascertainment bias, and hence great care was taken to collect a complete population of South Asian IBD patients attending IBD and gastroenterology outpatient clinics in a defined geographical area and to compare them with a comparable population of white Northern European IBD patients also from the same area. Potential limitations of this study are: (i) that patients with IBD-related symptoms may not have been referred to the clinic and therefore not included in the study, however this is likely to be a very small number of patients and will be the same for both ethnic groups and (ii) only patients attending hospital IBD/gastroenterology outpatient clinics were included; IBD patients under the sole care of primary care physicians or surgical outpatients would not have been included. However, in North West London, it is current practice that IBD patients are managed in secondary care by the gastroenterology team. After surgery, patients remain under joint care with the gastroenterologists. Primary care physicians may be involved in joint care of patients, but are not responsible for the sole care of IBD patients. This practice was confirmed in a survey that I conducted of 10 local primary care practices.

3.4.1 UC phenotype in UK South Asians

The frequency of extensive colitis in South Asians is significantly higher (63.0%) than in the Northern European population (42.5%) ($P < 0.0001$). These findings support the suggestion from a previous smaller study that extensive colitis is more common in South Asian patients than in Northern European patients (Rashid et al., 2008). A study by Carr and Mayberry of 33 newly diagnosed South Asian patients in Leicester City reported that UC disease extent is more extensive in second-generation UC patients (6/9; 66.6%) than first-generation (6/24; 25%). They proposed that an environmental risk factor may explain the difference. In this current study, the possibility of environmental factors in the UK influencing disease phenotype was investigated by comparing South Asians born in the UK with South Asians born in the Indian subcontinent and diagnosed within 10 years of arriving in the UK. Extensive colitis remained the predominant phenotype for each group with no significant differences between the groups. Although it is not clear how long a patient needs to be exposed to relevant environmental factors for them to be significant, the high rates of extensive colitis among South Asians, together with the absence of differences between first and second-generation UC patients, would support the idea that genetic factors may play a role in the susceptibility to extensive colitis in South Asians.

It is well recognised that patients with more extensive disease are at higher risk of colectomy (Ritchie et al., 1978; Lennard-Jones, 1983) and nearly all patients in the study population that required colectomy had extensive colitis (17/18 in the South Asian cohort and 17/20 in the Northern European cohort). It is therefore surprising that only 17 out of 172 South Asian patients with extensive colitis required a colectomy

compared to 17/80 in the Northern European cohort. This may suggest a less severe phenotype of IBD in the South Asian population and the finding is consistent with the study by Probert et al., which also suggested that South Asians with UC are at lower risk of colectomy than Europeans (Probert et al., 1992).

South Asian UC patients reported similar EIMs to the Northern European population (8.8% vs 6.9% ($P=0.49$)), with peripheral arthropathy being the most common EIM in both cohorts. The frequency of EIMs was lower than previously described in South Asians (Kochhar et al., 1991). This may reflect a more rigorous stratification in this study - patients with arthropathy were only included if they had objective evidence of an inflammatory arthropathy and patients with arthralgia or any other cause of arthritis were excluded. Patients with back pain but without radiological evidence of sacroiliitis were also excluded. Uveitis and episcleritis were the only ocular complications recorded and not conjunctivitis, which had been included as a EIM in a previous study of Indian UC patients (Kochhar et al., 1991).

3.4.2 CD phenotype in UK South Asians

Colonic disease was the most common disease location in the CD South Asian population. In both Northern European and South Asian cohorts, the predominant disease behaviour was nonstricturing nonpenetrating. Despite this, a significantly lower prevalence of penetrating behaviour was observed in the South Asian population compared to Northern Europeans ($P_c=0.04$); this may in part explain why the requirements for surgery are significantly lower in South Asians ($P_c=0.01$). The

population of South Asians with CD is smaller than the UC cohort and so further large studies would be required to confirm the differences reported.

3.4.3 Potential explanations for differences in clinical phenotype

Differences in clinical phenotype between ethnic groups are likely to be due to a combination of genetic susceptibility, environmental and immune factors. Over 70 susceptibility genes associated with IBD have been identified in the white European population over the past decade (Barrett et al., 2008;Hampe et al., 2007;The Wellcome Trust Case Control Consortium, 2007;Thompson and Lees, 2011). The *CARD15/NOD2* gene on chromosome 16q (IBD1) has been associated with isolated ileal disease in white European CD patients (Cuthbert et al., 2002;Ahmad et al., 2002;Vermeire et al., 2002). It is of note that the three disease susceptibility variants in the *CARD15/NOD2* gene associated with CD are rare or absent in South Asian IBD populations (Juyal et al., 2007;Pugazhendhi et al., 2008). Therefore, other genes or polymorphisms not yet identified may influence susceptibility to certain clinical phenotypes in this ethnic group.

Differences between ethnic groups do not necessarily imply a genetic cause but may reflect variations in environmental factors related to lifestyle in different groups. An environmental factor which may be of importance in determining CD phenotype is smoking. This study reported a significantly lower number of smokers in the South Asian CD cohort compared to the Northern European cohort (10.6% vs 31.6%; $P<0.0001$). This is consistent with previous reports that have shown significantly fewer smokers in South Asian migrants compared to Europeans (Banerjee et al., 2010;Carr

and Mayberry, 1999). Smoking has been reported to be associated with disease location, with a higher prevalence of ileal disease and a lower prevalence of colonic involvement in smokers (Lakatos et al., 2009; Lindberg et al., 1992; Mahid et al., 2007). Previous studies have also demonstrated that smoking is more frequently associated with the development of strictures or fistulae with an increased risk of surgery (Picco and Bayless, 2003; Reese et al., 2008).

Although this study did not show any differences in IBD phenotype between first- and second-generation South Asians, it has been proposed that environmental stimuli may act as a trigger for IBD in genetically susceptible individuals when moving from one part of the world to another. The 'IBD hygiene hypothesis' suggests that hygienic environments negatively affect immune development, predisposing to immune-mediated diseases like IBD (Elliott et al., 2000; de Silva et al., 2008). Reduced exposure to intestinal helminths may be a factor contributing to IBD as they are involved in host immune regulation, preventing excessive immune activation (Weinstock and Elliott, 2009), and perhaps better hygiene in the UK may be responsible for reduced exposure to enteric infections. This may result in immune dysregulation and subsequent excessive inflammation in predisposed individuals. Ethnic variation in the gut microbiota, and differences between developing and developed countries may also be involved in the dysregulation of mucosal immunity.

Previous reports have suggested that dietary constituents in the Western diet, such as refined sugars and saturated fats are linked with the development of IBD (Bernstein and Shanahan, 2008). More recently, high levels of arachidonic acid, an essential n-6 polyunsaturated fatty acid (n-6 PUFA), which is present in red meat, certain cooking

oils, and polyunsaturated margarines has also been associated with a significantly higher risk of developing UC (de Silva et al., 2010). Investigators from Leicester showed that South Asian Hindu IBD patients have significantly altered their traditional diet (Probert et al., 1996a). Dietary changes adopted by South Asians may therefore act as a potential environmental stimulus in disease development.

The lack of large population based studies of IBD in South Asia makes attempts to compare epidemiological data difficult. However, the nonstricturing, nonpenetrating colonic pattern of disease was the dominant phenotype in a recent multi-centre Indian study of 182 CD patients which may suggest that the phenotype of CD remains constant despite migration (Das et al., 2009). Two smaller epidemiological studies from India report between 16 and 20% of UC patients have extensive colitis (Kochhar et al., 1991; Sood et al., 2003). In contrast, the majority of South Asian patients living in North West London with UC had extensive disease (63.0%). The difference could relate to alterations in underlying environmental factors or to a variation in diagnostic assessment. This study based UC disease extent on a confirmed histological diagnosis rather than simply endoscopic or radiological appearances (used in the study by Kochhar et al.), but for comparison, the current data were reassessed using endoscopic extent only and extensive colitis was still the predominant phenotype (59%).

3.4.4 Conclusion

This study has shown that UC is more prevalent than CD in a UK South Asian population, and that the predominant phenotype of UC in this ethnic group is extensive colitis. This appears to be associated with a low risk of colectomy for active disease, and long-term follow-up of this group will be important, to establish whether the higher proportion of extensive colitis translates into an increased risk of dysplasia/cancer. The most prevalent CD phenotype in UK South Asians is nonstricturing, nonpenetrating colonic CD. Disease behaviour and need for surgery was significantly lower in UK South Asians compared with Northern Europeans. The phenotypic variations observed in South Asian IBD patients are likely to reflect different hereditary, environmental and behavioural factors, which may be elucidated through genetic, microbiological and immunological studies.

CHAPTER 4

**PHENOTYPE GENOTYPE ASSOCIATIONS OF
INFLAMMATORY BOWEL DISEASE IN A SOUTH ASIAN
POPULATION**

4.1 INTRODUCTION

In view of the high incidence and prevalence of IBD in parts of South Asia (Sood et al., 2003), it is of great importance to study social, environmental and host genetic factors that may underlie these findings. Both CD and UC have been shown to have a strong genetic component in Caucasian populations (Hugot et al., 2001; Duerr et al., 2006). The studies described in this chapter investigate the relevance of genetic associations previously described in other populations, in the South Asian population.

In CD, the most consistently replicated polymorphisms have been found in the *CARD15/NOD2* gene on chromosome 16 (Hugot et al., 2001). In Caucasian populations, the overall risk of developing CD was 2.4 times higher for carriers of any one of three polymorphisms in this gene (rs2066844, rs2066845 and rs2066847) (Hugot et al., 2001). These particular SNPs are most commonly associated with ileal CD (Lesage et al., 2002). The frequency of having one of the *CARD15/NOD2* SNPs in Caucasian controls is 8-15% compared to 10-30% in CD patients (Hugot et al., 2007). Recent studies of Indian IBD patients has suggested that the 3 SNPs in *CARD15/NOD2* associated with Caucasian CD patients are rare in Indian CD populations (Juyal et al., 2007; Pugazhendhi et al., 2008). However, an additional SNP in *CARD15/NOD2* (rs2066842) was associated with susceptibility to UC in this population (Juyal et al., 2007). The first part of this study was undertaken to confirm the frequency of *CARD15/NOD2* CD associated SNPs in a South Asian IBD population.

In 2006 a genome wide association (GWA) study showed an association between CD and variants in the *IL23R* gene (Duerr et al., 2006). A coding variant at rs11209026

(Arg381Gln) was reported to exert a significant protective effect in both Jewish and non-Jewish populations. The protective allele was less common in CD populations (1.9%) compared to matched controls (7%). Subsequent replication studies in a UK population confirmed the association and also reported an association with UC where the allele was again less common compared to controls (Tremelling et al., 2007). Hampe J et al.(2007) discovered an association with the autophagy 16-like 1 (*ATG16L1*) gene in a GWA study of 735 CD patients and 368 controls. This gene codes for a protein that has a crucial role in the degradation of cytoplasmic proteins and organelles. *ATG16L1* is expressed in many tissues including the colon and small bowel (Hampe et al., 2007). A single allele, the 300Ala allele of SNP rs2241880 was associated with disease susceptibility in Caucasian CD patients. There has been no confirmed association with this SNP in UC populations. Another gene associated with CD and linked to the autophagy pathway, is the *IRGM* gene on chromosome 5. This gene encodes a GTP-binding protein and is particularly involved in the elimination of intracellular bacteria, including *Mycobacterium tuberculosis* (Feng et al., 2004). SNPs in the *IRGM* gene (rs13361189 and rs4958847) are thought to result in reduced function and/or activity of this gene leading to a persistence of intracellular bacteria, consistent with existing models of CD pathogenesis (Palomino-Morales et al., 2009).

4.1.1 Study Aims

Studies from Japan suggest that SNPs in *IL23R*, *ATG16L1* and *IRGM* may not be associated with IBD in East Asian populations (Yamazaki et al., 2007; Yamazaki et al., 2009) however, there have been no previous studies of these SNPs in South Asian IBD populations. This study aimed to determine the prevalence of SNPs in *CARD15/NOD2*,

IL23R, *ATG16L1* and *IRGM* in a cohort of South Asian patients and controls and to compare the results with a white Northern European IBD cohort.

The Human Leukocyte Antigen (HLA) region on chromosome 6 contains the most consistently replicated gene loci associated with IBD. The *HLA-DRB1*0103* allele has been associated with both Caucasian CD and UC cohorts (Satsangi et al., 1996b; Trachtenberg et al., 2000). The allele has also been positively associated with disease phenotype, conferring susceptibility to extensive and severe UC (Ahmad et al., 2003a) and colonic CD (Hancock et al., 2008). In contrast to Caucasians, the *HLA-DRB1*0103* allele is not associated with IBD in East Asian populations (Yoshitake et al., 1999). However, the *HLA-DRB1*1502* allele has been associated with UC in East Asians (59.3% in UC cases vs 24.4% in healthy controls) (Yoshitake et al., 1999). This genetic association is also seen in Caucasians but the allele is rare (1.9% in cases vs 0.3% in controls) as the major allele of the serotype DR2 is *DRB1*1501* (Ahmad et al., 2003a). In a small Japanese study, the allele was associated with disease extent and severity (Masuda et al., 1994). There have been no previous reports on HLA class II associations in South Asian IBD populations. This part of the study was undertaken to determine the carriage frequency of the known IBD HLA associated alleles (*DRB1*0103* and *DRB1*1502*) in South Asian IBD patients and controls and to evaluate the influence of these alleles on disease phenotype.

4.2 METHODS

4.2.1 Ethics approval

The St. Mary's Hospital Research Ethics Committee approved this study and site-specific approval was obtained from each of the participating trusts. The South Asian control cohort was collected by Dr John Chambers and Professor Jaspal Kooner with approval from Ealing Research Ethics Committee. All participants gave written informed consent.

4.2.2 Sample Collection

South Asian patients were recruited from the inflammatory bowel disease clinics of five North West London hospitals (St. Mary's and Hammersmith Hospital of Imperial College Healthcare NHS Trust, St. Marks, West Middlesex and Ealing Hospital). Only patients with a confirmed diagnosis of Crohn's disease or ulcerative colitis were included in the study. South Asian patients were only recruited if all four grandparents were born in the Indian subcontinent.

For each participant, 10mls of peripheral blood was collected in an anticoagulant vacutainer containing di-sodium EDTA. This was immediately frozen at -80°C and used for DNA extraction and subsequent genotyping at a later date.

Dr John Chambers and Professor Jaspal Kooner supplied blood (2mls EDTA sample per patient) from a control cohort of healthy South Asian participants who each had four

grandparents born in the Indian subcontinent (n=204). These patients were recruited through the London Life Sciences Prospective Population (LOLIPOP) Study. LOLIPOP is an ongoing population-based cohort study of Indian Asian men and women aged 35–75 years in West London, U.K.

A well phenotyped and genotyped white Northern European population of 149 IBD patients (CD:99 and UC:50) previously collected by Dr. Horace Williams (Imperial College London), was used for comparison.

4.2.3 DNA isolation

DNA was extracted from peripheral blood using a two-stage procedure based upon Miller's salting out technique (Miller et al., 1988). **Appendix 1** outlines the reagents that were prepared prior to DNA extraction.

Miller's salting out technique – stage 1

For DNA extraction, the blood sample was thawed at room temperature and transferred to a 15ml polypropylene tube. 7mls of red cell lysis buffer was added and the suspension inverted several times before leaving it for 20 minutes at room temperature. Each sample was centrifuged at 2500 rpm for 20 minutes to pellet the white cell. The supernatant was discarded into Hycolin. Additional red cell lysis buffer was added to the polypropylene tube to make a total of 13mls. The tube was left to stand again for 20 minutes before centrifuging at 2500 rpm for a further 20 minutes. The supernatant was discarded and the resulting white cell pellet was either stored at -20°C for DNA

extraction at a later date or the DNA was extracted immediately by proceeding to stage 2 as outlined below.

Miller's salting out technique – stage 2

The second stage of DNA extraction involved re-suspension of the white cell pellet in 3mls of nuclear lysis buffer. Proteins were precipitated by the addition of 1ml of 6M NaCl and separated from the DNA by the addition of 2mls of 24:1 Chloroform Isoamyl alcohol. The samples were shaken to form a milky solution and then centrifuged at 3000 rpm for 30 minutes. Three distinct layers were then visible. The upper layer was aspirated and transferred into a new 15ml polypropylene tube. 10mls of 98% ethanol were added and the tube inverted several times resulting in the precipitation of DNA. The precipitated DNA was transferred into a sterile 1.5 ml eppendorf and left to dry for 30 minutes. 400µl of sterile water (Baxters®) was added to re-suspend the DNA. The sample was stored at 35°C for 12 hours to aid suspension before storing at -20°C.

Millers salting out technique requires greater than 4mls of blood to yield a significant concentration of DNA (>100ng/µl). For blood sample volumes <4mls, DNA was extracted using the QIAamp® DNA blood mini kit (Qiagen®).

QIAamp® blood mini kit technique

To extract DNA using the blood mini kit, 40µl of proteinase K and 400µl of the participant's blood was added to a 1.5 ml microcentrifuge tube. The sample was stabilised by adding 400µl of Buffer AL (all buffers supplied by Qiagen®) and incubating at 56°C for 10 min. 400µl of ethanol (96–100%) was added to the sample making a total of 1.2mls of solution, and then mixed by pulse-vortexing for 15 s. 600µl

of the final solution was transferred to a QIAamp Mini spin column in a 2ml collection tube. The tube was centrifuged at 8000rpm for 1 min. The spin column was then placed in a clean 2 ml collection tube, and the tube containing the filtrate discarded. The remaining 600µl of solution was added to the spin column and the above process repeated. 500µl Buffer AW1 was added to the spin column and the tube centrifuged at 8000rpm for 1 min. The spin column was again placed in a clean 2 ml collection tube. Buffer AW2 (500 µl) was added to the spin column and the tube centrifuged at 13,000 rpm for 3 min. The spin column was placed in a clean 1.5 ml microcentrifuge tube, and the collection tube containing the filtrate discarded.

100µl of distilled water was added to the spin column and incubated at room temperature (15–25°C) for 1 min, before finally centrifuging at 8000 rpm for 1 min. The filtrate yields approximately 6.25µg of DNA at a concentration of 60ng/µl. This is frozen at -20°C for future use.

4.2.4 Determining DNA concentration

DNA concentration was measured using a Nanodrop reader (Thermo Scientific®). A 1µl sample of DNA was pipetted onto the end of a fibre optic cable (the receiving fibre). A second fibre optic cable (the source fibre) was brought into contact with the liquid sample causing the liquid to bridge the gap between the fibre optic ends. A pulsed xenon flash lamp provides the light source and a spectrometer is used to analyze the light after passing through the sample. The instrument is controlled by PC based software, and the data was logged on a PC attached to the Nanodrop.

4.2.5 Single Nucleotide Polymorphism Analysis

The term single nucleotide polymorphism (SNP) denotes the natural nucleotide variation that exists throughout the human genome. SNPs can be single nucleotide substitutions, or single base deletions or insertions. Such changes are classified as SNPs rather than mutations when the minor variant is present in at least 1% of the population otherwise these base changes are classified as mutations. SNPs most likely to affect the gene expression or protein function are the commonest targets for study, as they may give rise to functional changes in the gene product, which may be of relevance to a disease.

SNPs are highly informative tools when used in association studies in order to elucidate the link between variants and disease phenotypes. The location of SNPs throughout the human genome is freely available from web-based databases. In this study, the NCBI SNP database (<http://snpper.chip.org/bio/snpper-enter>) was used in conjunction with SNPs reported in published materials.

Genotyping for SNPs in the *CARD15/NOD2* gene and medium resolution HLA class II typing was performed in the department of human genomics at the National Heart and Lung Institute, Imperial College London using sequence-specific primer polymerase chain reaction (SSP-PCR). Genotyping for SNPs in the *ATG16L1*, *IL23R* and *IRGM* gene were carried out using pyrosequencing at the Hammersmith Hospital, London.

4.2.6 Sequence Specific Primer - Polymerase Chain Reaction (SSP-PCR)

A sequence specific primer (SSP) is designed to have its 3' end complementary to a particular SNP variant and identifies the presence of this variant by polymerase chain reaction (PCR) amplification (Bunce et al., 1995). At a single locus, two different reactions are used. These reactions involve a shared consensus primer and an allele-specific primer, either sense or antisense, that differs at the 3' end. Each reaction also contains primers to amplify a non-polymorphic segment of DNA, which acts as a positive control to confirm amplification was successful. These primers are multiplexed (mixed in the same reaction well and amplified simultaneously). The presence of a PCR product of the expected size indicates the presence of the allele in the specimen.

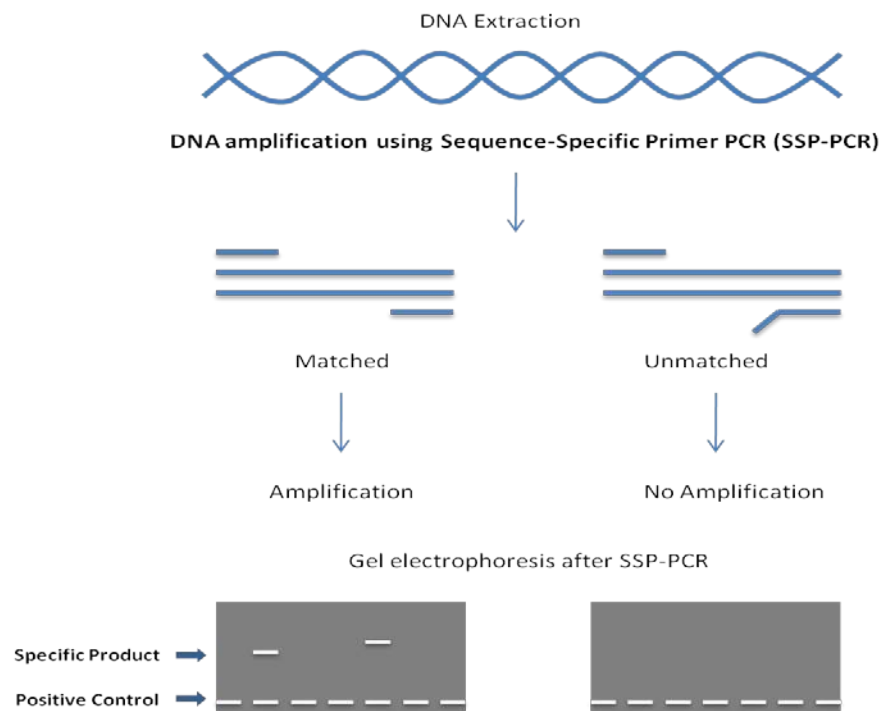


Figure 4-1: Schematic diagram showing the process of Sequence Specific Primer Polymerase Chain Reaction (SSP-PCR)

SSP-PCR is a robust and reliable method for SNP identification. As only a small number of SNPs are being examined, it has the advantage of being cost effective compared to a chip based approach. Two or more SNPs within the same gene, proven to be in tight linkage disequilibrium with each other, can be detected using allele-specific primers in the forward and reverse directions making this technique ideal for HLA genotyping.

SSP-PCR primers

Established primers were used for the 3 polymorphisms associated with CD in Caucasians and the additional polymorphism (SNP5) previously associated with IBD in South Asians (with permission from Dr. Tariq Ahmad [outlined in Ahmad *et al.*, 2002]). SNP id numbers and associated peptide changes are shown in **Table 4-1**. Primer sequences are displayed in **Appendix 2**.

Table 4-1: Single Nucleotide Polymorphism (SNP) id numbers for the 4 polymorphisms investigated in <i>CARD15/NOD2</i>			
rs number	SNPid	Location	Peptide change
rs2066844	<i>SNP8</i>	Exon4: 2104 (C>T)	R702W
rs2066845	<i>SNP12</i>	Exon8: 2722 (G>C)	G908R
rs2066847	<i>SNP13</i>	Exon11: 3020 (insC)	1007fs
rs2066842	<i>SNP5</i>	Exon4: 802 (C>T)	P268S

A 48-reaction primer set was used for medium-resolution HLA class II typing courtesy of Dr. Hiroe Sato (derived from Bunce *et al.*, 1995) (see **Appendix 3** and **Appendix 4**)

for primer sequences). For the HLA class II genotyping, haplotypes were deduced from the gel photograph by reference to known haplotypes in Indian Asian and Caucasian populations respectively using www.allelefrequencies.net (with permission from Professor Derek Middleton, Queen's University, Northern Ireland).

Control primers

The choice of the control primer is dependent on the size of the specific product. Products of 600 to 900bp were multiplexed with control primers 210/211 to produce a product size of 249bp from the Adenomatosis Polyposis Coli gene (APC). If the specific product was 250- 450bp then primers 63/64 were used. These primers amplify a constant region of *HLA-DRB1* and produce a control product of 798bp (primer sequences shown in **Appendix: 5**).

SSP-PCR technique

Primers were prepared by the manufacturers (Sigma Genosys, UK) and suspended in sterile water (Baxters, UK) to a concentration of 2µg/µl. Specific and consensus primers were titrated to obtain optimal working concentrations depending on the specificity and sensitivity of each reaction. Multiplexed primer mixes were prepared with sterile distilled water, 40µl/ml cresol red indicator (600 mg cresol red dissolved in 100mL dH₂O), which undergoes a pH related colour change from orange to purple on addition of DNA, and the appropriate amount of primer. For each DNA sample, 5µL aliquots of each multiplex primer mix are dispensed into wells of a 96 well plate and covered with 10µL mineral oil. The four *CARD15/NOD2* polymorphisms required 8 reactions and medium resolution HLA class II typing required 48 reactions.

The DNA/PCR reaction mix (8µL) is added to the primer mix and consisted of PCR buffer [(NH₄)₂SO₄], MgCl₂ (2mM), dNTP mix (300µM of each dNTP), Taq polymerase (0.32 units) and genomic DNA (dependent on DNA concentration, approx. 0.08 µg/reaction).

PCR amplification was carried out using thermal cyclers: initial denaturing temperature of 1 minute at 96°C; followed by 5 cycles of 20 seconds at 96 °C, 45 seconds at 70 °C, 25 seconds at 72 °C; 21 cycles of 25 seconds at 96 °C, 50 seconds at 65 °C, 30 seconds at 72 °C; and 4 cycles of 30 seconds at 96 °C, 60 seconds at 55 °C, 120 seconds at 72 °C.

Gel electrophoresis after SSP-PCR

Amplicons were separated on 1.5% agarose gels with 0.5 X Tris Borate EDTA buffer, stained with Sybr safe (Invitrogen, USA). 10µls of orange G loading buffer was added to each well prior to this. Electrophoresis was carried out at 200V for 20 minutes and gels were photographed over ultraviolet light (UV) using an AlphaDigiDoc digital camera (Alpha Innotech®, USA). The presence of a specific allele was determined by observing an amplicon of the expected size in association with a control PCR product. The absence of a specific allele was determined by observing only a control PCR product (**Figure 4-2**).

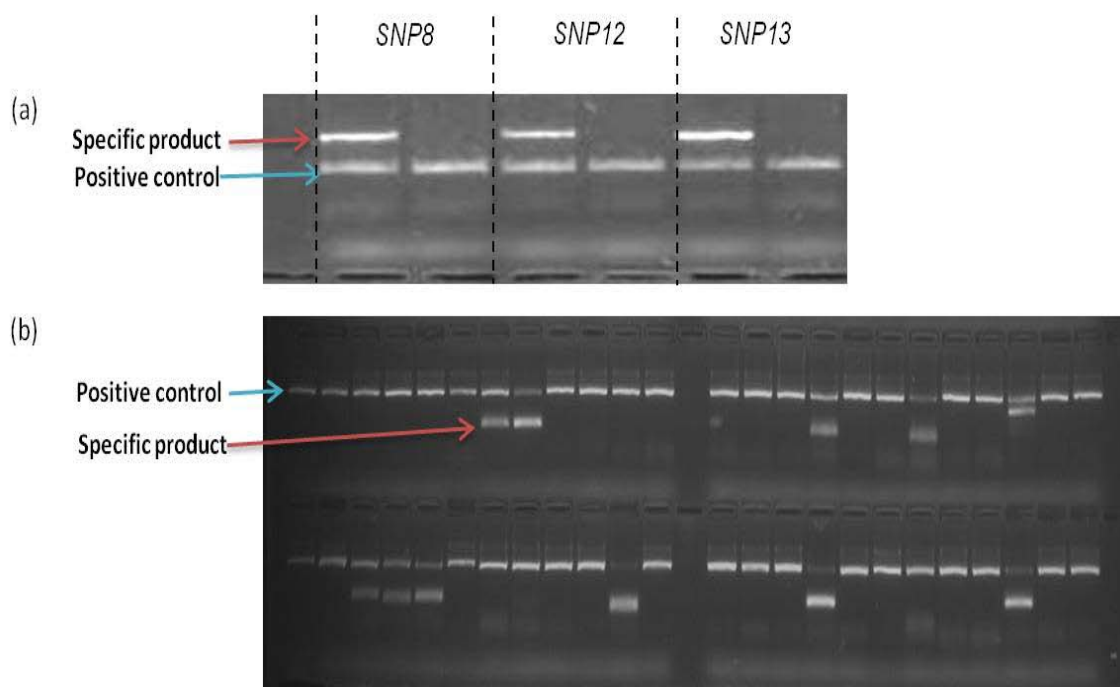


Figure 4-2: Photographs of agarose gels under ultraviolet light after electrophoresis following SSP-PCR.

(a) Products and control bands for three *CARD15/NOD2* polymorphisms (using control primers 210/211 producing a product of 249 bp). The presence of a PCR product (labelled as specific product) indicates the presence of the allele. In all three cases, the participant was homozygous for the wild type allele. (b) Products and controls bands for the 48 reactions making up the HLA class II plate (using controls 63/64 producing a product of 798 bp).

4.2.7 Pyrosequencing

Pyrosequencing has risen to prominence since being first described in 1996 (Ronaghi et al.). Instead of using the chain termination method, pyrosequencing uses a procedure known as ‘sequencing by synthesis’. The process is based on the detection of released pyrophosphate (PPi) during DNA synthesis and in a cascade of enzymatic reactions, visible light is generated that is proportional to the number of incorporated nucleotides.

Principles of pyrosequencing

After PCR amplification, double stranded DNA is first hybridised to a single stranded DNA amplicon that serves as a template. This is incubated with DNA polymerase and other enzymes required for the reaction. A deoxyribonucleotide triphosphate (dNTP) is added to the reaction. If it is complimentary to the base in the template strand, DNA polymerase incorporates the dNTP into the complimentary DNA strand. Each incorporation is accompanied by the release of pyrophosphate (PPi). The released PPi is subsequently converted to ATP by ATP-sulfurylase, which provides the energy for luciferase to oxidize luciferin and generate light. Addition of dNTPs is performed sequentially. As the process continues, the complimentary DNA strand is built up and the nucleotide sequence is determined from the signal peaks in the Pyrogram trace (**Figure 4-3**). The overall reaction from polymerization to light detection takes place within 3–4 seconds at room temperature.

Pyrosequencing has the potential advantages of speed and accuracy. Furthermore, the technique dispenses with the need for labelled primers, labelled nucleotides, and gel-electrophoresis (Ronaghi, 2001).

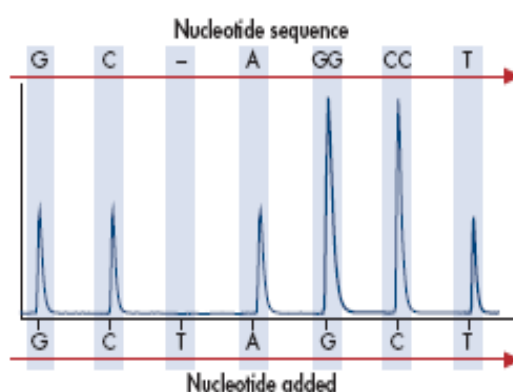


Figure 4-3: A pyrogram showing the signal peaks that determine the nucleotide sequence
(adapted from <http://www.pyrosequencing.com/DynPage.aspx?id=7454&mn1=1366&mn2=1367>).

Primer design for Pyrosequencing SNPs in ATG16L1, IL23R & IRGM

The sequence flanking the SNPs were ascertained by performing an Ensembl search (<http://www.ensembl.org>) using the SNP id (rs number) (**Table 4-2**).

Table 4-2: Single Nucleotide Polymorphism (SNP) id numbers used for primer design			
Gene	SNP id (rs number)	Chromosome	Peptide change
<i>ATG16L1</i>	rs2241880	2q37	T300A
<i>IL23R</i>	rs11209026	1p31	R381Q
<i>IRGM (SNP189)</i>	rs13361189	5q33	-
<i>IRGM (SNP847)</i>	rs4958847	5q33	-

PCR primers were designed for each SNP. A primer set consisted of a Forward (F) (200-250bp upstream of the SNP), a Sequencing (S) (final base at the SNP) and a Reverse Primer (R) (200-250bp downstream of the SNP and complimentary to the

selected sequence, read 3'-5') (**Appendix 6**). Primers were optimally designed with a length of 19 to 21 bases, a salt adjusted melting temperature of 58 to 61°C, a GC content of 45 to 55% and with no secondary structures (e.g. Hairpin loops). The online program oligonucleotide properties calculator <http://www.basic.nwu.edu/biotools/oligocalc.html> was used to determine these conditions. Once designed, primers were checked for specificity using the basic local alignment search tool (BLAST) <http://www.ncbi.nlm.nih.gov/BLAST/> from the national centre for biotechnology information (NCBI). BLAST looks for matches between the sequence you submit and a comprehensive database of all known DNA sequences. This avoids non-specific priming ensuring as far as possible that primers will only anneal to an exact site match.

PCR amplification prior to pyrosequencing

A PCR mix containing 20µl of solution was added to each well of a 96 well plate (2µl of genomic DNA (at a concentration of >100ng/µl), 2µl of 1.5 mM MgCl₂, 2µl 10x reaction Buffer, 1µl of each primer (F & R at 10 mM concentration), 1µl 2mM of each dATP, dCTP, dGTP and dTTP, 0.1µl of Biotaq polymerase (Bioline Ltd) and 10.9µl ddH₂O). Amplifications were performed in the G-Storm thermal cycler (Alpha Metrix Biotech, GmbH). The cycle consisted of 5min at 95°C, followed by 30 cycles of 30sec at 95°C, 30sec at 50-60°C and 60sec at 72°C with an additional 10min extension at 72°C.

A random selection of PCR products were analysed by standard gel electrophoresis on a 1.5% agarose gel with Sybersafe™ prior to pyrosequencing, to confirm PCR amplification.

Pyrosequencing technique

The first stage of pyrosequencing involved a binding reaction to hybridise the double stranded DNA to single stranded DNA. This required a plate containing 20µl PCR product, 38 µl of Binding Buffer (Biotage), 2 µl of Streptavidin Sepharose beads (GE Healthcare) and 20µl of dH₂O. The plate was shaken for 15 min in a plate shaker (Web Scientific) at 1300g. The Streptavidin-coated magnetic beads enabled the PCR product to be captured onto magnetic beads. After sedimentation, the remaining components of the PCR reaction were removed by washing in 70% ethanol to obtain pure double-stranded DNA followed by alkali denaturation with 0.2M NaOH to yield ssDNA. Following this, the ssDNA was washed in 10X Washing Buffer (Biotage) according to the manufactures' guidelines.

The final step before analysis is the Annealing reaction; 11.5µl Annealing Buffer (Biotage) and 0.5µl of sequence primer (S1) are aliquoted in to PyroMark™Q24 Plates (Biotage) and the bound Streptavidin Sepharose-PCR product is placed in the wells and allowed several minutes to bind. The plate is heated at 80°C for 3 min and then placed in the PyroMark Q24 (Biotage) for SNP analysis as per the manufactures' guidelines.

4.2.8 Statistical analysis

Within the South Asian CD population, the analysis was designed to examine 4 SNP's in the *CARD15/NOD2* gene, 2 SNPs in the *IRGM* gene and the single IBD associated SNP in *ATG16L1* and *IL23R*. The carriage frequency of *HLA-DRB1*0103* and *HLA-DRB1*1502* will also be determined. For the UC population the single SNP at *IL23R*, the single *CARD15/NOD2* SNP (SNP5) and the carriage frequency of *HLA-DRB1*0103* and *HLA-DRB1*1502* will be examined.

Statistical advice regarding a power calculation was discussed with the Imperial College London statistical advice service. As no data exist for the high risk allele frequencies of the SNPs in *ATG16L1*, *IL23R* and *IRGM* in the South Asian population and previous studies have shown that the 3 common *NOD2* polymorphisms are monomorphic, an accurate power calculation could not be generated for this population. If the sample size was estimated using the Western data on SNP frequencies (Balding, 2006; Sasieni, 1997) (Shaun Purcell genetic power calculator <http://pngu.mgh.harvard.edu/~purcell/gpc/>), a comparison between the CD population and normal controls requires 85 in each group to detect association (i.e. difference in allele frequencies) at 99% power with an $\alpha=0.05$. However, to allow for multiple comparisons by Bonferroni adjustment, 126 patients would be required in each group. For the UC population, using western data, 163 cases and controls are required to give a power of 87% with $\alpha=0.05$. To allow for multiple comparisons 242 participants would be required in each group.

The HLA exhibits strong linkage disequilibrium and it is generally accepted in the literature that a power calculation is not necessary for this analysis. A Bonferonni correction will be applied to the results to allow for multiple comparisons.

Genotype and allele frequencies for each SNP were calculated from the observed genotype counts. Case-control comparisons were made using the Chi-squared test. The allelic frequency of SNPs in the South Asian IBD population were compared to a white Northern European IBD cohort also using the chi-squared test. P values <0.05 were considered statistically significant. The data was analysed using Graph Pad Prism Software® (Version 5, U.S.A.).

4.3 RESULTS

217 IBD patients (61 CD, 156 UC) and 204 healthy controls were recruited during the study period. **Table 4-3** outlines the demographic details and disease characteristics of the recruited participants.

Table 4-3: South Asian participant's demographic details and disease characteristics			
	Crohn's disease	Ulcerative colitis	Healthy controls
Number (Male/Female)	61 (29/32)	156 (92/64)	204 (114/90)
Median [range] age (years)	44 [20-83]	46 [21-82]	61 [55-75]
Ethnicity			
Indian	54	137	204
Pakistani	5	18	-
Bangladeshi	2	1	-
Disease location	L1: 6	E1: 17	-
	L2: 26	E2: 40	-
	L3: 29	E3: 99	-
Disease behaviour	B1:41	-	
	B2:17	-	
	B3:3	-	
Previous surgery	17	8	-
L1=ileal disease, L2=colonic disease, L3=ileocolonic disease. E1=proctitis, E2=disease limit distal to splenic flexure, E3=disease proximal to the splenic flexure. B1=nonstricturing, nonpenetrating; B2=stricturing; B3=perforating. (Silverberg et al., 2005)			

Dr Horace Williams (Imperial College London) supplied DNA from a cohort of 149 consecutively collected white Northern European IBD patients (99 CD, 50 UC). The clinical details for this cohort are described in **Table 4-4**. These patients had already been HLA typed and genotyped for the four *CARD15/NOD2* polymorphisms as part of a previous study. DW performed additional genotyping for *IL23R*, *ATG16L1* and *IRGM* for this current study.

Table 4-4: White Northern European participant's demographic details and disease characteristics		
	Crohn's disease	Ulcerative colitis
Number (Male/Female)	99 (48/51)	50 (26/24)
Median [range] age (years)	41 [19-81]	41 [19-63]
Disease location	L1: 20	E1: 11
	L2: 35	E2: 18
	L3: 44	E3: 21
Disease behaviour	B1: 62	-
	B2: 14	-
	B3: 23	-
Previous surgery	30	1
E1=proctitis, E2=disease limit distal to splenic flexure, E3=extensive colitis (disease proximal to the splenic flexure). L1=ileal disease, L2=colonic disease, L3=ileocolonic disease. B1=nonstricturing, nonpenetrating; B2=stricturing; B3=perforating (Silverberg et al., 2005).		

4.3.1 Frequency of Crohn's disease associated *CARD15/NOD2* polymorphisms in the South Asian IBD population

All South Asian CD patients and healthy control participants were genotyped for four *CARD15/NOD2* variants. The results were compared with a cohort of previously genotyped white Northern European CD patients (n=99). Genotyping was successful in 98.4% of the South Asian cases (60/61), 99% of South Asian controls (202/204) and 99% of white Northern Europeans (98/99). In the South Asian population, analysis of the three *CARD15/NOD2* polymorphisms associated with CD in Europeans (SNP8, SNP12 & SNP13) showed that only 3.3% (2/60) of CD patients carried one of the three SNPs and 3.5% (7/202) of healthy controls. SNP13 was monomorphic in the South Asian population and no South Asian CD patients carried the SNP8 allele and only 1 South Asian control. SNP12 was also rare with 2 heterozygotes in the CD population and 6 heterozygotes among the healthy controls. There were no South Asian patients or controls with a double mutation. The difference in the minor allele frequencies between South Asian CD patients and controls was not statistically significant.

The frequencies of the SNPs 8, 12 & 13 in the South Asian population were compared with a white Northern European CD cohort. The polymorphisms were more common in the European population, 14.3% were heterozygote for SNP8 and 7.1% for SNP12. For SNP13, 6.1% were heterozygotes and 3.1% homozygotes. The minor allele frequencies for the white Northern European population were significantly higher for SNPs 8 & 13 (SNP8 $P=0.003$; SNP13 $P=0.006$). Genotypes and minor allele frequencies are shown in **Table 4-5**.

Table 4-5: Comparison of the allele frequencies of *CARD15/NOD2* risk alleles (SNP8, 12 & 13) in South Asian and white Northern European CD patients and South Asian healthy controls

SNP id	Allele frequency			¹ P value
	South Asian		Northern European	
	CD (n=60) n (%)	Controls (n=202) n (%)	CD (n=98) n (%)	
rs2066844 (SNP8)				
CC	60 (100)	201 (99.5)	84 (85.7)	
CT	0 (0)	1 (0.5)	14 (14.3)	
TT	0 (0)	0 (0)	0 (0)	
Minor allele frequency (T)	0 (0)	1 (0.2)	14 (7.1)	P=0.003
rs2066845 (SNP12)				
GG	58 (96.7)	196 (97)	91 (92.9)	
GC	2 (1.3)	6 (3)	7 (7.1)	
CC	0 (0)	0 (0)	0 (0)	
Minor allele frequency (C)	2 (1.7)	6 (1.5)	7 (3.6)	P=0.32
rs2066847 (SNP13)				
-/-	60 (100)	202 (100)	89 (90.8)	
-/insC	0 (0)	0 (0)	6 (6.1)	
insC/insC	0 (0)	0 (0)	3 (3.1)	
Minor allele frequency (insC)	0 (0)	0 (0)	12 (6.1)	P=0.006

CD: Crohn's Disease, UC: Ulcerative colitis, SNP: Single Nucleotide Polymorphisms, ins: insertion.

The minor allele is defined in the controls and its frequency in that group as well as the case sample is reported. For all 3 SNPs, the minor allele is the risk allele.

¹: Using the χ^2 test to compare minor allele frequencies of South Asian CD patients and white Northern European CD patients.

South Asian IBD patients (60 CD, 156 UC) and controls (n=202) were genotyped for *CARD15/NOD2* SNP5 and compared with a white Northern European IBD cohort (98 CD and 50 UC) (**Table 4-6**). There was no significant difference in the minor allele frequency between South Asian patients and controls (CD, $P=0.3$; UC, $P=0.3$). The T allele was significantly more frequent among the European CD cases compared to the South Asian CD cohort but not in the UC cohorts (CD; $P<0.0001$, UC; $P=0.08$).

Table 4-6: Results of the <i>CARD15/NOD2</i> genotyping for SNP5 in South Asian and white Northern European IBD patients and South Asian controls						
SNP id	Allele frequency					P value
	South Asian			Northern European		
	CD (n=60) n (%)	UC (n=156) n (%)	Controls (n=202) n (%)	CD (n=98) n (%)	UC (n=50) n (%)	
rs2066842 (SNP5)						
CC	48 (80)	121 (77.6)	145 (71.8)	47 (48)	32 (64)	
CT	10 (16.7)	32 (20.5)	53 (26.2)	39 (39.8)	17 (34)	
TT	2 (3.3)	3 (1.9)	4 (2.0)	12 (12.2)	1 (2)	
MAF (T)	14 (11.7)	38 (12.1)	61 (15.1)	63 (32.1)	19 (19)	¹ P<0.0001 ² P=0.08

CD: Crohn’s Disease, UC: Ulcerative colitis, SNP: Single Nucleotide Polymorphisms

MAF: Minor allele frequency; The minor allele is defined in the controls and its frequency in that group as well as the case sample is reported. The minor allele (T) is also the risk allele.

¹P: Using the χ^2 test to compare minor allele frequencies of South Asian CD patients and white Northern European CD patients.

²P: Using the χ^2 test to compare minor allele frequencies of South Asian UC patients and white Northern European UC patients.

4.3.2 Associations of polymorphisms in *ATG16L1*, *IRGM* and *IL23R* in the South Asian CD population

South Asian CD patients, South Asian controls and white Northern European CD patients were genotyped for SNPs in *ATG16L1*, *IRGM* and *IL23R* using pyrosequencing. An example of a pyrogram that was generated for SNP IRGM189 is displayed in **Figure 4-4**.

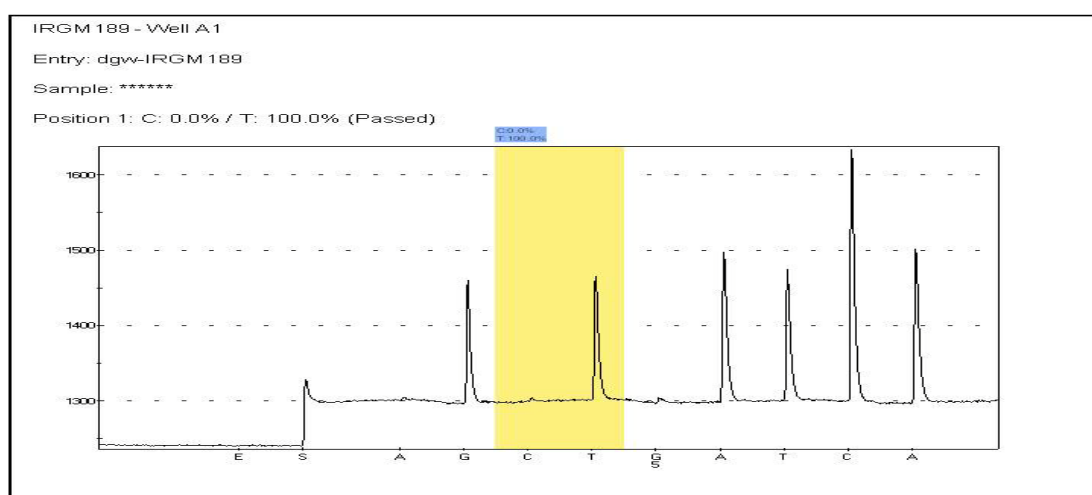


Figure 4-4: A pyrogram showing the IRGM189 genotype for a participant. The area of the gene representing the peptide sequence at IRGM189 is highlighted in yellow, showing a peak for T and no peak for C, therefore classed at genotype TT.

Genotyping was successful in 98.4% (60/61) of South Asian CD patients, 99.5% (203/204) of controls and 99% of disease controls (98/99). There was no significant difference in the frequency of any of the 4 SNPs between South Asian patients and controls. The minor allele frequency was significantly different between the ethnic groups for SNPs in *IRGM* and *IL23R* (*IRGMSNP189* $P=0.004$; *IRGMSNP847* $P=0.005$; *IL23R* $P=0.02$) (**Table 4-7**).

Table 4-7: The frequency of CD associated SNPs in <i>ATG16L1</i> , <i>IRGM</i> and <i>IL23R</i> in South Asian CD patients and controls compared to a Northern European CD cohort				
Gene (SNPId)	Allele frequency: n (%)			¹ P value
	South Asian		Northern European	
	CD (n=60)	Controls (n=203)	CD (n=98)	
ATG16L1 (rs2241880)				
GG	20 (33.3)	65 (32.0)	45 (45.9)	
AG	29 (48.3)	100 (49.3)	40 (40.8)	
AA	11 (18.4)	38 (18.7)	13 (13.3)	
*Minor allele frequency (A)	51 (42.5)	176 (43.3)	66 (33.7)	P=0.11
IRGM 189 (rs13361189)				
TT	38 (63.3)	131 (64.5)	76 (77.6)	
CT	16 (26.7)	60 (29.6)	22 (22.4)	
CC	6 (10)	12 (5.9)	0 (0)	
Minor allele frequency (C)	28 (23.3)	84 (20.7)	22 (11.2)	P=0.004
IRGM 847 (rs4958847)				
GG	33 (55)	118 (58.1)	72 (73.4)	
AG	21 (35)	68 (33.5)	23 (23.5)	
AA	6 (10)	17 (8.4)	3 (3.1)	
Minor allele frequency (A)	33 (27.5)	102 (25.1)	29 (14.8)	P=0.005
IL23R (rs11209026)				
GG	59 (98.3)	195 (96.1)	86 (87.8)	
AG	1 (1.7)	8 (3.9)	12 (12.2)	
AA	0 (0)	0 (0)	0 (0)	
Minor allele frequency (A)	1 (0.8)	8 (2.0)	12 (6.1)	P=0.02
The minor allele is defined in the controls and its frequency in that group as well as the case sample is reported. *For <i>ATG16L1</i> the risk allele is G. For <i>IRGM189</i>, <i>IRGM847</i> and <i>IL23R</i> the minor allele is the risk allele. ¹: Using the χ^2 test to compare minor allele frequencies of South Asian CD patients and white Northern European CD patients. SNP: Single Nucleotide Polymorphisms, ins: insertion.				

The single SNP in the *IL23R* gene was also analysed in a UC cohort of 156 South Asian patients, 203 controls and 48 disease controls and showed no statistical difference between any of the groups (**Table 4-8**).

Table 4-8 : The frequency of IL23R in South Asian UC patients, controls and a white Northern European UC cohort				
Gene (SNPId)	Allele frequency: n (%)			¹ P value
	South Asian		Northern European	
	UC (n=156)	Controls (n=203)	UC (n=48)	
IL23R (rs11209026)				
GG	152 (97.4)	195 (96.1)	45 (93.8)	
AG	4 (2.6)	8 (3.9)	3 (6.2)	
AA	0 (0)	0 (0)	0	
Minor allele frequency (A)	4 (1.3)	8 (2.0)	3 (3.1)	<i>P</i> =0.2
SNP: Single Nucleotide Polymorphisms ¹ : Using the χ^2 test to compare minor allele frequencies of South Asian UC patients and white Northern European UC patients.				

4.3.3 Comparison of the carriage frequency of *HLA-DRB1*0103* and *HLA-DRB1*1502* in South Asian IBD cases and controls

South Asian IBD patients and controls were genotyped for medium resolution HLA class II. Genotyping was successful in all South Asian IBD patients and controls. The results were compared with a white Northern European cohort of 50 UC patients and 76 CD patients (demographic details are shown in **Table 4-9**).

Table 4-9: Demographic details and disease characteristics for South Asian and White Northern European ulcerative colitis patients				
	Ulcerative colitis		Crohn's Disease	
	South Asian	European	South Asian	European
Number (Male/Female)	156 (92/64)	50 (26/24)	61 (29/32)	76 (39/33)
Median age (yrs) [range]	46 [21-82]	41 [19-63]	44 [20-83]	37 [19-78]
Disease location	E1: 17	E1: 11	L1: 6	L1: 17
	E2: 40	E2: 18	L2: 26	L2: 25
	E3: 99	E3: 21	L3: 29	L3: 34
Disease behaviour	-	-	B1:41	B1: 48
	-	-	B2:17	B2: 10
	-	-	B3:3	B3: 18
Previous bowel resection	8	1	17	26
E1=proctitis, E2=disease limit distal to splenic flexure, E3=extensive colitis (disease proximal to the splenic flexure). L1=ileal disease, L2=colonic disease, L3=ileocolonic disease. B1=nonstricturing, nonpenetrating; B2=stricturing; B3=perforating. (Silverberg et al., 2005).				

The carriage frequency of the *HLA-DRB1*1502* allele was significantly higher in the South Asian UC population compared to matched ethnic controls ($P=0.005$). This remained significant after Bonferroni adjustment for multiple comparisons ($P=0.01$). There was no association with *HLA-DRB1*1502* in South Asian CD patients compared to controls ($P=0.35$). None of the South Asian IBD patients or controls carried the *HLA-DRB1*0103* allele. **Table 4-10** summarise the associations between the two HLA alleles in South Asians.

Table 4-10: The associations between IBD and <i>HLA-DRB1*0103</i> and <i>DRB1*1502</i> in the South Asian population				
	Allele frequency (%)			<i>P</i> value
Allele	UC (n=156)	CD (n=61)	Controls (n=204)	
<i>DRB1*0103</i>	0	0	0	-
<i>DRB1*1502</i>	47 (30.1)	14 (22.9)	36 (17.6)	¹ $P=0.005$ ² $P=0.35$
The <i>P</i> values in this table are uncorrected. ¹P: Using the χ^2 test to compare South Asian UC patients with South Asian controls ²P: Using the χ^2 test to compare South Asian CD patients with South Asian controls				

HLA-DRB1*1502 association with clinical phenotype in South Asian IBD patients

The gene frequency of *HLA-DRB1*1502* in 31/99 patients with extensive colitis (31%) was not significantly higher than the group with left sided colitis and proctitis (16/57: 28%; $P=0.67$). Of the 8 patients requiring colectomy 3 had the *DRB1*1502* allele (37%) compared to 30% in the rest of the UC cohort ($P=0.6$). There was no association with the *HLA-DRB1*1502* allele and disease location in CD patients ($P=0.96$).

HLA class II haplotype associations with IBD in South Asians

In order to determine if a specific HLA haplotype was associated with IBD in South Asians, the carriage frequencies of all medium resolution DRB1 and DQB1 alleles were compared. The *DRB1*1502/DQB1*0601* haplotype was positively associated with UC when compared to controls ($P=0.009$).

During the analysis, it was noted that the *DRB1*1001* allele was negatively associated with UC ($P=0.001$) and this remained significant after correcting for multiple comparisons (28 variables). The only association that was found for CD was a negative association with *HLA-DRB1*1001* but this association was less robust than UC ($P=0.01$).

Differences in HLA frequencies between South Asian and Europeans IBD patients

The carriage frequency of *HLA-DRB1*0103* and *DRB1*1502* was significantly different between the ethnic groups. 18% of European UC patients and 10.5% of European CD patients carried the *DRB1*0103* allele which was not present in any of the South Asian IBD patients (UC, $P<0.0001$ and CD, $P=0.004$). The *DRB1*1502* allele was not found in any of the European UC patients compared to 30% of South Asian UC patients

($P<0.0001$). Only 3% of European CD patients carried the allele which was significantly lower than the South Asian population (23%) ($P<0.0001$) (**Table 4-11**).

Table 4-11: A comparison of the frequencies of <i>HLA-DRB1*0103</i> and <i>HLA-DRB1*1502</i> in South Asian and European IBD populations.						
HLA allele	UC		<i>P</i> value	CD		<i>P</i> value
	S. Asian (n=156)	European (n=50)		S. Asian (n=61)	European (n=76)	
DRB1*0103	0	9 (18.0)	$P<0.0001$	0	8 (10.5)	$P=0.004$
DRB1*1502	47 (30.1)	0	$P<0.0001$	14 (22.9)	2 (2.6)	$P<0.0001$

4.4 DISCUSSION

A vast amount of evidence is accumulating to suggest that host genetics determines susceptibility to IBD. Genetic studies are supplementing the knowledge gained from epidemiological studies, which suggest a difference in the genetic determinants of IBD within different ethnicities. This study aimed to investigate the contribution of *HLA class II* alleles and SNPs in *CARD15/NOD2*, *ATG16L1*, *IRGM* and *IL23R* to IBD susceptibility in a South Asian IBD population. Although these SNPs have been widely replicated in North American and European IBD patients there are no previous studies with regard to *HLA class II*, *ATG16L1*, *IRGM* or *IL23R* in the South Asian population.

4.4.1 IBD associated genetic polymorphisms

The absence of the three common *CARD15/NOD2* polymorphisms (SNP8, SNP12 and SNP13) in the South Asian population suggests that these particular *CARD15/NOD2* polymorphisms are not important in the genetic susceptibility of CD in South Asians. Although the population of CD patients was small, and this may reflect a power issue, the findings are consistent with a previous study from India that also showed these SNPs were rare in this population (Pugazhendhi et al., 2008).

SNPs in *ATG16L1* and *IRGM* (SNP189 & SNP847) have been positively associated with CD in Caucasian populations. These SNPs were analysed in a South Asian CD population and although the risk allele frequency was higher for the three SNPs compared to controls, there was no statistical differences in the SNP frequencies. A single SNP in *IL23R* (rs11209026) has been shown to be negatively associated with

both CD and UC in Western populations but there was no significant difference seen when South Asian IBD patients were compared to ethnically matched controls.

Juyal et al (2007) showed an association with SNP5 in *CARD15/NOD2* and ulcerative colitis in North Indians. This polymorphism at position 268 leads to an amino acid change Pro/Ser (C/T). In Juyal's study the T allele was significantly more frequent among UC cases compared to controls (17% vs 12%; $P=0.008$), whereas in this study there was no difference between South Asian UC patients and controls (12% vs 15%; $P=0.3$). To exclude any differences in ethnicity across the Indian subcontinent that could have affected the result, the analysis was also performed comparing only Indian UC patients with controls and this also showed no significance (11% vs 15%; $P=0.13$). The numbers in this current study (UC=156) were less than the Juyal study (UC=298) and the absence of an association could be as a result of a smaller sample size, however, the fact that the SNP frequency in this study was higher in controls than UC patients would make any potential association less likely.

An acknowledged limitation of the study is the small study number and lack of ileal CD patients within the South Asian CD population. Previous studies have shown that SNPs associated with *CARD15/NOD2*, *IRGM* and *IL23R* may be more prevalent in patients with ileal CD. As the number of patients with pure ileal disease was very small in the South Asian population ($n=6$) subgroup analysis by ileal phenotype would not produce any meaningful result. A larger study with more ileal CD patients would be required to confirm the study's findings. However, the reported results would be consistent with an East Asian study from Japan which genotyped 484 CD patients and 439 controls and showed these SNPs were not associated with CD in this ethnic group (Yamazaki et al.,

2007). For *CARD15/NOD2*, *IRGM* and *IL23R*, the SNP frequencies in the South Asian IBD cohorts were significantly different to the European disease controls showing that there are interethnic variations in IBD risk allele frequencies.

4.4.2 HLA associations in South Asian IBD patients

The study has shown an association with *HLA-DRB1*1502* and UC in a cohort of South Asian UC patients. This finding is similar to HLA studies in Japan and South Korea that have also reported an association between UC and the same DRB1 allele (Masuda et al., 1994; Myung et al., 2002; Futami et al., 1995). In contrast to the Japanese study, there was no association of *HLA-DRB1*1502* with disease phenotype in South Asians. Previous studies of this allele in a UK population have also shown that *DRB1*1502* has been associated with susceptibility to UC in Caucasian populations although with relatively low frequency (1.9% in cases vs 0.3% in controls) (Ahmad et al., 2003a). The allele was not present in any of the European UC cohort, the discrepancy in this finding is likely to relate to the small patient numbers.

Genetic associations within the MHC region are complicated by the extensive linkage disequilibrium in this region (Ahmad et al., 2003b). This is especially strong between DRB1 and DQB1 alleles, with the result that alleles are often inherited together as haplotypes. Association of an allele with a disease can occur if that allele is functionally implicated or if it is on the same haplotype as the functional gene (i.e. the allele is just a marker). This is illustrated by the positive association of the *DRB1*1502/DQ*0601* haplotype with UC in the South Asian cohort. However, it is still likely that the primary

association is with *DRB1*1502*, as *DQ*0601* can also occur with *DRB1*1501* and it is the *DRB1*1502/DQ*0601* haplotype that appears to be a risk factor in this population.

In this study the *HLA-DRB1*1001* allele was shown to be protective for development of both UC and CD. This has not been previously described in any other ethnic group and appears unique to the South Asian population and therefore this allele may have a role in IBD susceptibility in South Asian populations. Further studies are required to replicate this finding.

HLA associations are now recognised in a large number of chronic diseases. In certain pathologies eg. coeliac disease, the HLA association directly impacts the disease; more than 90% of patients express the *HLA-DQ2* molecule, and the majority of the remaining patients carry *DQ8*. The *DQ2* and *DQ8* mediated presentation of gluten peptides to CD4+ T cells is a key event in the pathogenesis of the disease (Tighe et al., 1992; Sollid et al., 1989). In IBD, HLA allele associations have become clear but whether it is the actual presentation of antigenic peptides by the associated alleles that is critical in disease pathogenesis remains unclear. Molecular studies of the HLA are beginning to focus on the peptide binding site of *HLA-DRB1* which consists of a series of peptide pockets. The specific sequence of these pockets will correspond to the antigenic peptide that predispose to disease risk and protection and may unlock the key to the role of the HLA in IBD.

4.4.3 Genetic associations in South Asian IBD patients compared to other countries

The association of *DRB1*1502* in the South Asian UC cohort and the lack of association with SNPs in *CARD15/NOD2*, *ATG16L1*, *IL23R* and *IRGM* in South Asian CD patients may suggest that this ethnic group shares a common ancestry with East Asian countries like Japan and South Korea rather than European populations. The Indian subcontinent has experienced a massive gene flow from at least two large migrations. Firstly, 15,000 years ago as part of an eastward wave of human migration that entered India and secondly, 4000 years ago, the arrival of Pastoral nomads from the central Asian steppes and Iranian plateau (Agrawal et al., 2007) . Thus a considerable amount of intermingling between races has occurred, creating variation in HLA haplotypes (Mehra et al., 1994). It was generally accepted that South Asians were more Caucasoid because of the nature of the majority of South Asian languages (Indo-European) and it was thought that they may have once shared a common language superfamily, known as Dene-Caucasian (Ayub and Tyler-Smith, 2009) which led to the belief that there may be a genetic relatedness between Caucasians and South Asians. However, the results from this study suggest more of a similarity with East Asian genetics.

4.4.4 Conclusion

The data from these results suggests four important points about the genetics of IBD in South Asians. Firstly, that the *HLA-DRB1*1502* is significantly associated with UC in the South Asian IBD population. Secondly, that the allele *HLA-DRB1*1001* may have an important protective role in the development of IBD. Thirdly the *DRB1*0103 allele* is rare amongst the South Asian population and is not responsible for the higher frequency of extensive UC in the South Asian population. Finally, the polymorphisms of *CARD15/NOD2*, *ATG16L1*, *IRGM* and *IL23R* that are associated with IBD in Western populations show no association with IBD in South Asians; however, one cannot exclude the possibility that an analysis of larger sample sizes or using sequencing to detect novel polymorphisms might unveil such an association.

CHAPTER 5

URINARY METABOLIC PROFILING IN A SOUTH ASIAN INFLAMMATORY BOWEL DISEASE POPULATION

5.1 INTRODUCTION

Although the exact pathogenesis of IBD is not clearly understood, it is widely thought to involve an abnormal genetically determined host immune response to an environmental stimulus that is likely to be bacterial (**Figure: 5-1**). Animal models have demonstrated that the induction of IBD requires the presence of bacteria (Sydora et al., 2006). No specific pathogen has been defined as a causative agent, but the gut microbiota have been shown to differ between CD and UC, and healthy control subjects (Sartor, 2008). IBD patients show an imbalance or ‘dysbiosis’ of their intestinal microbiota (Willing et al., 2009; Sartor, 2006), and this is thought in turn to lead to a breakdown in the relationship between the intestinal microbiota and the host immune system.

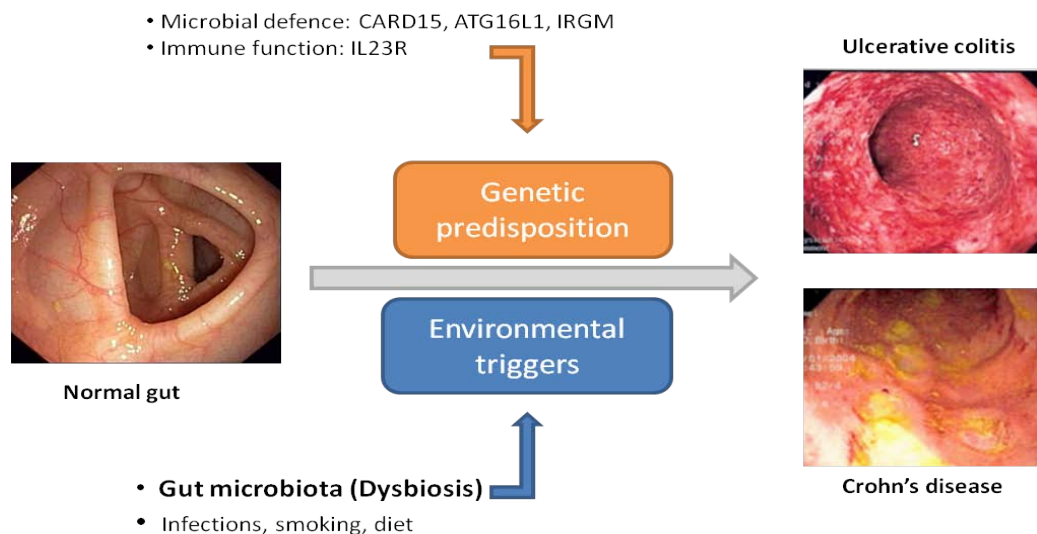


Figure 5-1: Genetic predisposition and environmental triggers contributing to the pathogenesis of inflammatory bowel disease

5.1.1 Investigating changes in the gut flora using metabolic profiling

Metabolic profiling (the study of large numbers of small molecule metabolites) using nuclear magnetic resonance spectroscopy (NMRS) has the potential to provide data regarding all metabolic processes of an organism. NMRS detects hydrogen atoms in metabolites. In a typical biological fluid, all hydrogen-containing molecules (which includes nearly all metabolites) will contribute to the ^1H NMR spectrum. The resulting spectrum is therefore an overview of all metabolites in a fluid.

Previous experiments have demonstrated that the host microbiota significantly contributes to the overall metabolic profile generated from biofluids (Wikoff et al., 2009; Li et al., 2008; Marchesi et al., 2007), therefore any changes in the gut microbiota will be reflected in the spectroscopic profile of the biofluid being examined. The measurement of metabolites in urine, serum and stool has enabled researchers to determine metabolic differences between patients in various disease states compared to healthy controls. This includes studies in pneumococcal pneumonia, multiple sclerosis hypertension and coronary artery disease (t Hart et al., 2003; Brindle et al., 2003; Brindle et al., 2002; Slupsky et al., 2009).

In IBD, NMRS has revealed gut microbial contributions to metabolite compositions in different body fluids, including serum (Schicho et al., 2010), urine (Williams et al., 2009), faecal extracts (Bezabeh et al., 2009; Marchesi et al., 2007) and colonic biopsies (Bjerrum et al., 2010). Significant differences have been observed between healthy individuals and Caucasian patients with CD and UC.

Studies using urinary NMRS in humans with IBD showed significant differences between white Northern European UC patients, CD patients and healthy controls with respect to the metabolites, hippurate (reduced in both UC and CD compared to controls), 4-cresol sulphate (reduced in CD) and formate (reduced in CD) (Williams et al., 2009). The groups were also distinguished using multivariate modelling, with a particularly important contribution from hippurate. The urinary concentration of hippurate has been shown to be modulated by gut microbes (Nicholls et al., 2003; Williams et al., 2002) and the reported differences between Northern European IBD patients and controls are likely to reflect differences in the intestinal microflora. However, it is not clear whether these differences are seen in other ethnic populations.

5.1.2 Study aims

The aims of this study were firstly to use NMR-based urinary metabolic profiling to determine if there are differences in microbiota-related metabolites between South Asian IBD patients and South Asian controls and secondly to ascertain if the urinary NMR profile observed in white Northern European IBD cohorts is similar to a cohort of South Asian IBD patients. It was hoped that the results would provide further molecular insight into disease pathogenesis and help determine which metabolic markers are determinants of disease across different ethnic populations.

5.2 METHODS

5.2.1 Ethics Approval

The St. Mary's Hospital Research Ethics Committee approved this study and site-specific approval was obtained from each of the participating trusts. All participants gave written informed consent.

5.2.2 Patient recruitment

South Asian IBD patients

South Asian patients were recruited from the inflammatory bowel disease clinics of five North West London hospitals (St. Mary's, Hammersmith, St. Marks, West Middlesex and Ealing Hospital). Only patients with a confirmed diagnosis of Crohn's disease or ulcerative colitis based on clinical evaluation, endoscopic/radiological appearances and histology (Stange et al., 2008; Van et al., 2010), were recruited for the study. Stool culture results and mucosal histology were reviewed for all patients and those in whom there was any uncertainty regarding diagnosis (ie. those with IBD, type unclassified or colitis as a result of infection, radiation or ischaemia) were excluded from the study.

Healthy South Asian Controls

A cohort of healthy South Asian controls were recruited from staff and medical students at the five participating hospitals and patient spouses who attended at the time of the

outpatient appointment. Only volunteers with no history of chronic disease were included in the study.

White Northern European IBD patients and controls

Dr Horace Williams (Imperial College London) provided the Northern European data from previously analysed urinary spectra on a cohort of 146 IBD patients (86 CD: 60 UC) and healthy controls (n=60). The spectra of 50 Northern European samples were audited (by DW) to determine if there was inter-observer variability and no significant differences were found.

5.2.3 Experimental design

Urine collection

Previous studies have demonstrated that when the diet is not standardised, the greatest variability in the metabolic profile is seen on the first-void urine sample (Lenz et al., 2003; Walsh et al., 2006). For this reason, a 10ml random mid-stream urine sample was collected from each participant at the time of their clinic appointment. The sample was immediately stored in the fridge and within 90 minutes was centrifuged at 2500g for 20 minutes and any sediment removed before storing at -80°C pending analysis.

Addressing potential confounders

Significant differences in the metabolic profile have been noted between health and disease states, but the NMR profile of biofluids also contains information on recently

consumed drugs (Sun et al., 2009), the participant's diet (Walsh et al., 2006) and physiological variation such as hormonal state and gender differences (Bollard et al., 2001). Some authors believe that when the metabolic response to disease is the primary focus of interest, it is often not necessary to consider the subtle dietary and physiological variation, as the metabolic response to disease will eclipse minor metabolic variations (Wang et al., 2005). However, dietary influences have been well documented as causing variation in the NMR profile of healthy controls (Lenz et al., 2004; Stella et al., 2006) and in order to exclude any potential confounding variables, a detailed dietary and lifestyle history was completed prospectively (by DW) (**Appendix 7**) and results compared between the groups. Data collected included: gender, ethnic origin, place of birth, disease extent, current drug history, smoking history and a 24-hour recall nutrition history.

Certain foods have been shown to particularly influence the NMR spectra (Walsh et al., 2006) and for this reason patients were asked specifically about their intake of:

1. Tea, berries and carbonated drinks, as these foods contain benzoate and may influence the urinary excretion of hippurate (Wang et al., 2005).
2. Fish, which markedly influences the metabolism of TMAO (Zuppi et al., 1998).
3. Meat and vegetables, as lactovegetarians have been shown to have higher levels of TMAO, taurine, glycine and hippurate compared to individuals on an omnivorous diet (Xu et al., 2010).

In order to replicate normal circumstances, no dietary exclusions were imposed for this study but the dietary and lifestyle data from the CD, UC and control cohorts were analysed using the following methods to determine if there were any significant differences between the groups:

- i. Mann-Whitney U tests for continuous variables such as the number of cups of tea drunk in the preceding 24 hours and the number of hours since the last meal.
- ii. Fisher's exact test or Chi-square test (as appropriate) for categorical variables, such as whether or not the subject, had eaten meat (and whether red, white or both) or fish in the preceding 24 hours and usually.

Patients with diabetes and individuals with an intercurrent illness, who were pregnant, who were taking antibiotics or who were taking pre- or probiotics were excluded from the study. Williams et al. (2009) has previously shown that gut dependent metabolites remained significantly different between IBD cases and controls in both patients taking IBD medications (5-ASAs/azathioprine/methotrexate) and patients off medications. Therefore, in this study, both patients on and off IBD medications were included.

5.2.4 Urine sample preparation for NMR spectroscopy

When required, the urine sample was thawed and 400µl of each urine sample was mixed with 200µl of 0.2M phosphate buffer (pH 7.35) (for preparation, see **Appendix 8**), to stabilise the pH. The sample was vortexed for 15 seconds and then centrifuged for 10 minutes at 13000g. 500µl of supernatant was then transferred into a 5mm NMR tube (Wilmad-Labglass, USA) and mixed with 50µl of 3-trimethylsilylsilyl (2,2,3,3 – 2H₄)-

1-propionate (TSP)/D₂O solution (for preparation see **Appendix 8**). The TSP served as an internal chemical shift reference (δ 0.00).

5.2.5 In vitro NMR spectroscopy

NMR spectroscopy was performed at the Hammersmith Hospital campus of Imperial College London using a JEOL 500MHz Eclipse+ NMR spectrometer and an 11.7 Tesla superconducting magnet. Samples from South Asian CD, UC and healthy control cohorts were analysed blinded, in a random order. A standardised protocol with water presaturation was applied. The spectral width was 15 ppm and the number of data points acquired was 64K. A pulse-collect sequence was used with a 90° pulse angle. The acquisition time was 8.74 s. A relaxation delay of 10 s (total repetition time 18.74 s) ensured full relaxation between data collects. The 16 data collects were summated.

5.2.6 Data analysis

Spectra were imported to KnowItAll Informatics System v7.9 Metabonomics Edition (Bio-Rad, USA). Free induction decays were zero-filled by a factor of two and multiplied by an exponential window function with a 0.3 Hz line-broadening factor. They were then Fourier transformed from the free induction decay (FID) (**Figure 5-2**) into a frequency spectrum (**Figure 5-3**).

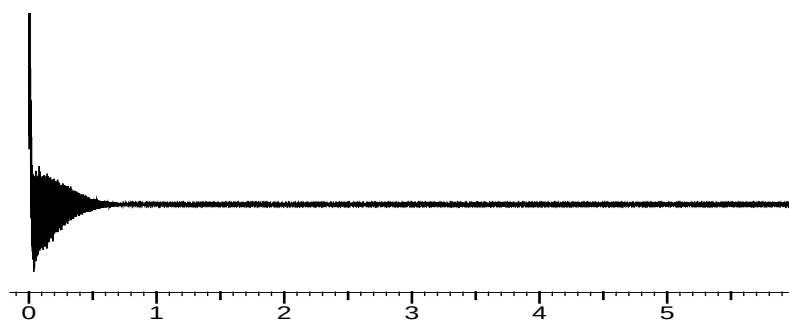


Figure 5-2: Free induction decay (FID) of the NMR Spectra.

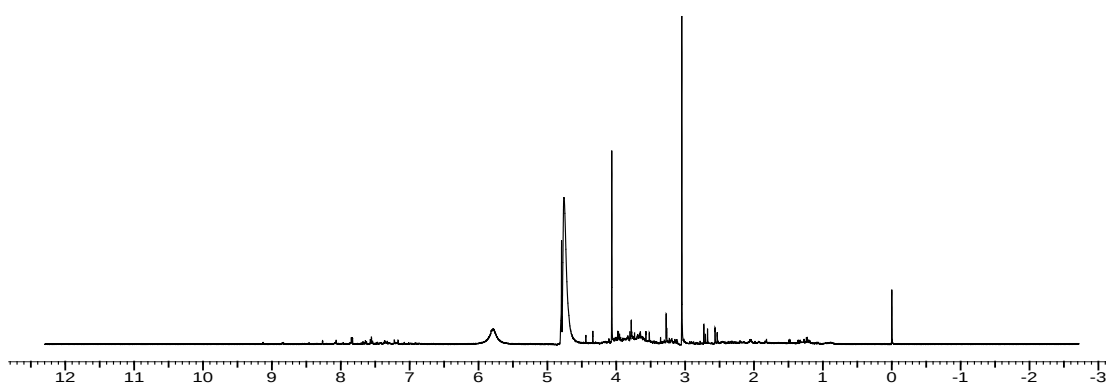


Figure 5-3: Fourier Transformation of the FID to give the NMR spectra in the time domain displaying metabolite peaks.

The reference peak for TSP was positioned at 0ppm and the spectrum was manually phased and baseline corrected. Spectral analysis included the range δ 0.2 -10.00 ppm, excluding the region δ 4.50-6.40 ppm to remove the effects of variable water suppression and variation in urea signal due to exchange of protons with the solvent (**Figure: 5-4**).

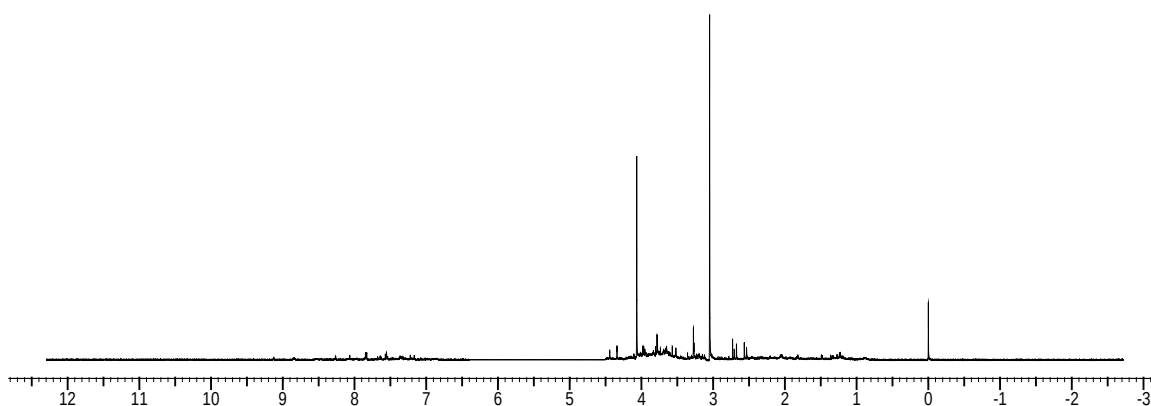


Figure 5-4: NMR spectra after excluding the water and urea peaks in the region 6.4 to 4.5.

5.2.6.1 Spectral analysis

The NMR resonances were analysed and spectral peaks assigned according to the literature (<http://www.hmdb.ca/>) (**Table 5-1**). Metabolites dependent on gut microbial metabolism were identified and quantified (formate, hippurate, 4-cresol sulphate, trimethylamine-N-oxide (TMAO) and dimethylamine (DMA)) (**Figure 5-5**).

Table 5-1: Peak assignments for urinary metabolites dependant on gut microbial metabolism.

Peak	Chemical Shift Ppm (δ)
Formate	8.46
Hippurate	7.83 , 3.97, 7.55, 7.64
4-cresol sulphate	2.35 , 7.21, 7.28
Trimethylamine-N-oxide (TMAO)	3.27
Dimethylamine (DMA)	2.73

*When a metabolite is represented at more than one point of the spectra, the spectral peak that was easiest to quantify was used (ppm marked in bold).

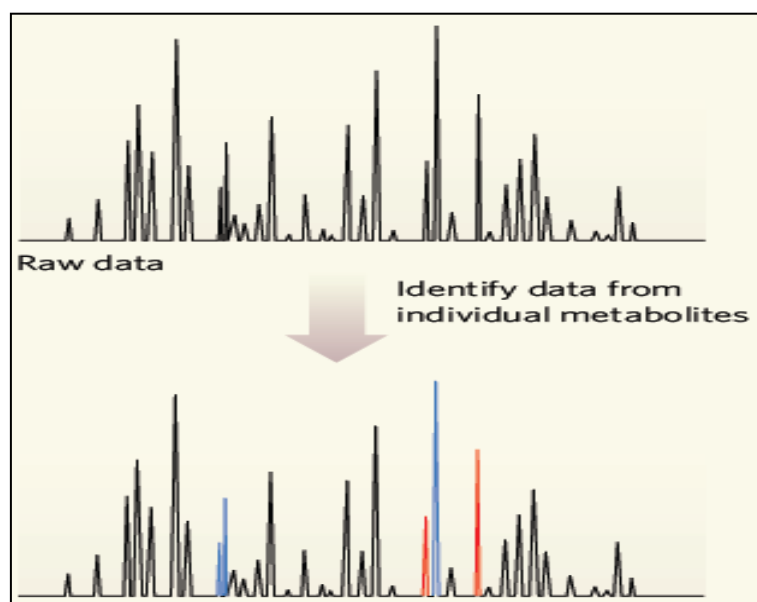


Figure 5-5: Schematic diagram of assignment of spectral peaks. (a) A typical spectra starts with the NMR spectrum of a biological fluid, which contains signals from hundreds of metabolites. (b) The individual spectra for each metabolite are identified according to the literature. (reproduced from Nicholson (2008)).

The resonances were integrated and normalised to the sum of the total spectral integral.

Values were expressed as a percentage relative to the total spectral integral.

5.2.7 Statistical methods and power calculation

Previous urinary metabonomic studies have predominantly been pilot studies and not used a traditional power calculation (Bala et al., 2004; Chung et al., 2003). After discussion with Dr Bernard North at Imperial College Statistical Advisory Service, we generated the following power calculation, which allows for multiple comparisons on univariate analysis. The calculation is based on the results of a urinary metabonomic study of IBD patients that was published in the *American Journal of Gastroenterology* (Williams et al., 2009).

- o An unpaired 2 sample t-test (equal variance) with 80% power and significance 2.5% (to adjust for multiple tests by Bonferroni) and then corrected to obtain a worst-case sample size requirement for the Mann-Whitney test by dividing by the worst case asymptotic efficiency of 0.864.
- o Controls vs CD: Mean difference (based on hippurate levels) = 1.37, SD=1.08, sample size = 13 per group for t-test: Sample size required for Mann-Whitney = $13/0.855 = 15$ per group.
- o Controls vs UC: Mean difference (based on hippurate levels) = 0.8, SD=1.17 sample size = 42 per group for t-test: Sample size required for Mann-Whitney = $42/0.855 = 49$ per group.

Therefore, 15 CD patients, 49 controls, 49 UC patients and are required to adequately power the study.

Differences in metabolite values between South Asian CD, UC and healthy control cohorts were assessed using the Mann-Whitney U test (GraphPad Prism v5.0, GraphPad

Software Inc., USA). To allow for multiple comparisons (5 metabolites) a Bonferroni correction was applied and the level of significance set at $P < 0.01$.

The South Asian CD and UC cohorts were also compared with the Northern European IBD cohorts using the Mann-Whitney U test. The dietary and lifestyle data from the CD, UC and control cohorts were analysed using the Kruskal-Wallis and Mann-Whitney U tests for continuous variables and Fisher's exact test or Chi-square test for categorical variables.

5.3 RESULTS

Urine samples were collected from a total of 34 South Asian CD patients, 121 South Asian UC patients and 52 healthy South Asian controls. The spectra were acquired for the first 22 CD samples, 53 UC samples and 36 controls. Further samples were due to be processed at the NMR spectroscopy laboratory at the Hammersmith Hospital campus of Imperial College London (**Table 5-2**). Unfortunately, the NMR facility at Hammersmith Hospital was closed down in March 2009. However, as the number of samples processed was close to the number required by the power calculation, data analysis was performed on the NMR spectra already acquired. The patient demographic details of the analysed urine samples are described in **Table 5-3**.

Table 5-2: Total number of South Asian urine samples collected and processed

	<i>Proposed number</i>	Total number of samples collected	Samples processed
Crohn's disease	15	34	22
Ulcerative colitis	49	121	53
Controls	49	52	36

Table 5-3: South Asian participants demographic and phenotypic details

	Crohn's disease	Ulcerative colitis	Healthy controls
Number (Male/Female)	22 (14/8)	53 (39/14)	36 (20/16)
Median [range] age (years)	38 [18-82]	44 [19-75]	37 [21-70]
Disease location	L1: 3	E1: 8	-
	L2: 8	E2: 10	-
	L3: 11	E3: 35	-

L1, ileal disease; L2, colonic disease; L3, ileocolonic disease; E1, proctitis; E2, disease limited to the splenic flexure; E3, disease proximal to the splenic flexure (Silverberg et al., 2005)

5.3.1 NMR data analysis

The spectral peaks of metabolites dependent on gut microbial metabolism (hippurate, formate, 4-cresol sulphate, TMAO and DMA) were identified and quantified, relative to the TSP peak at 0ppm. A representative proton NMR spectrum of urine from a South Asian control patient and a South Asian CD patient is shown in **Figure 5-6** with the spectral peaks that were measured marked accordingly.

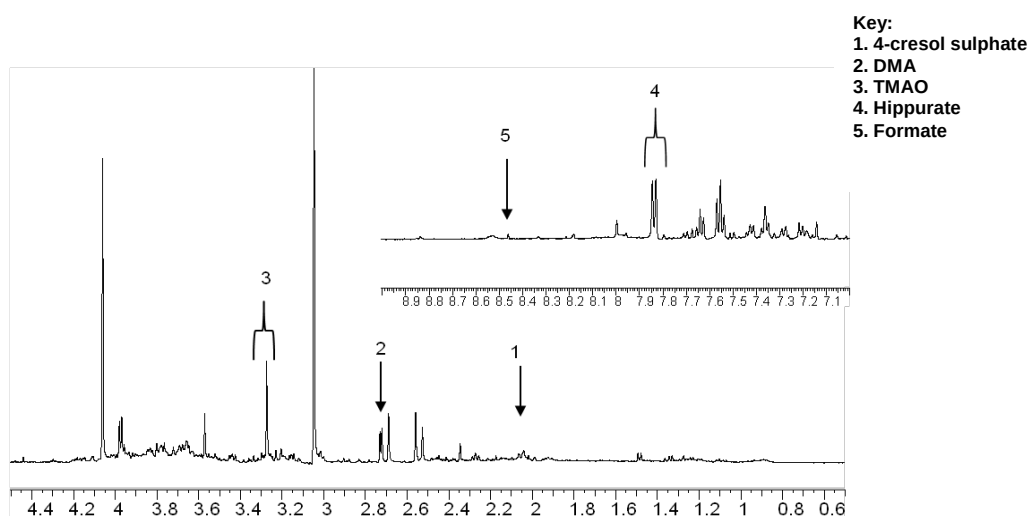


Figure: 5-6i

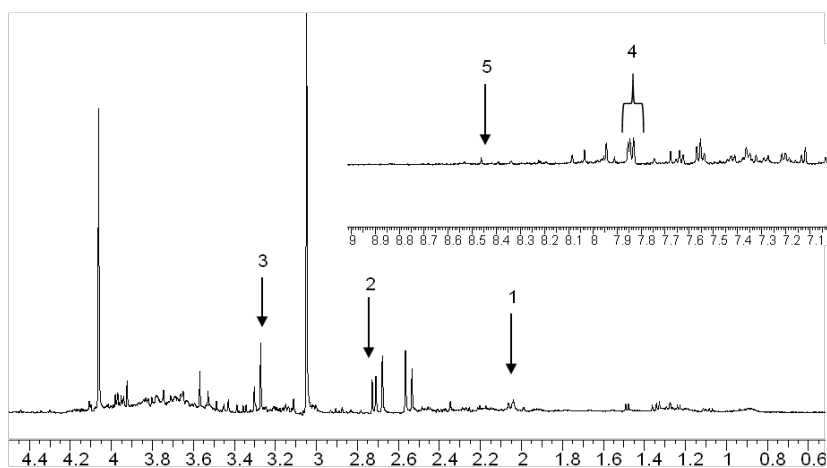


Figure: 5-6ii

Figure 5-6: The NMR spectra of urine from a South Asian control patient (Figure 5-6i) and a South Asian Crohn's disease patient (Figure 5-ii). The spectra in the aromatic region (δ 6.0-9.0) were magnified x4 compared to the aliphatic region (δ 0.5-4.5).

5.3.2 A comparison of the urinary NMR profile of South Asian IBD patients with healthy South Asian controls

The urinary spectra of the South Asian IBD patients was compared to a healthy South Asian control population to determine if there were differences in any of the five microbial dependent metabolites.

Hippurate levels were significantly lower in the South Asian UC and CD cohorts compared with South Asian controls (UC: $P < 0.0001$, CD: $P < 0.0001$). Formate levels were also significantly lower in the UC and CD cohorts compared to healthy controls levels (UC: $P = 0.01$, CD: $P = 0.004$). There was no significant difference in hippurate or formate levels between the South Asian CD and UC populations **Figure 5-7**.

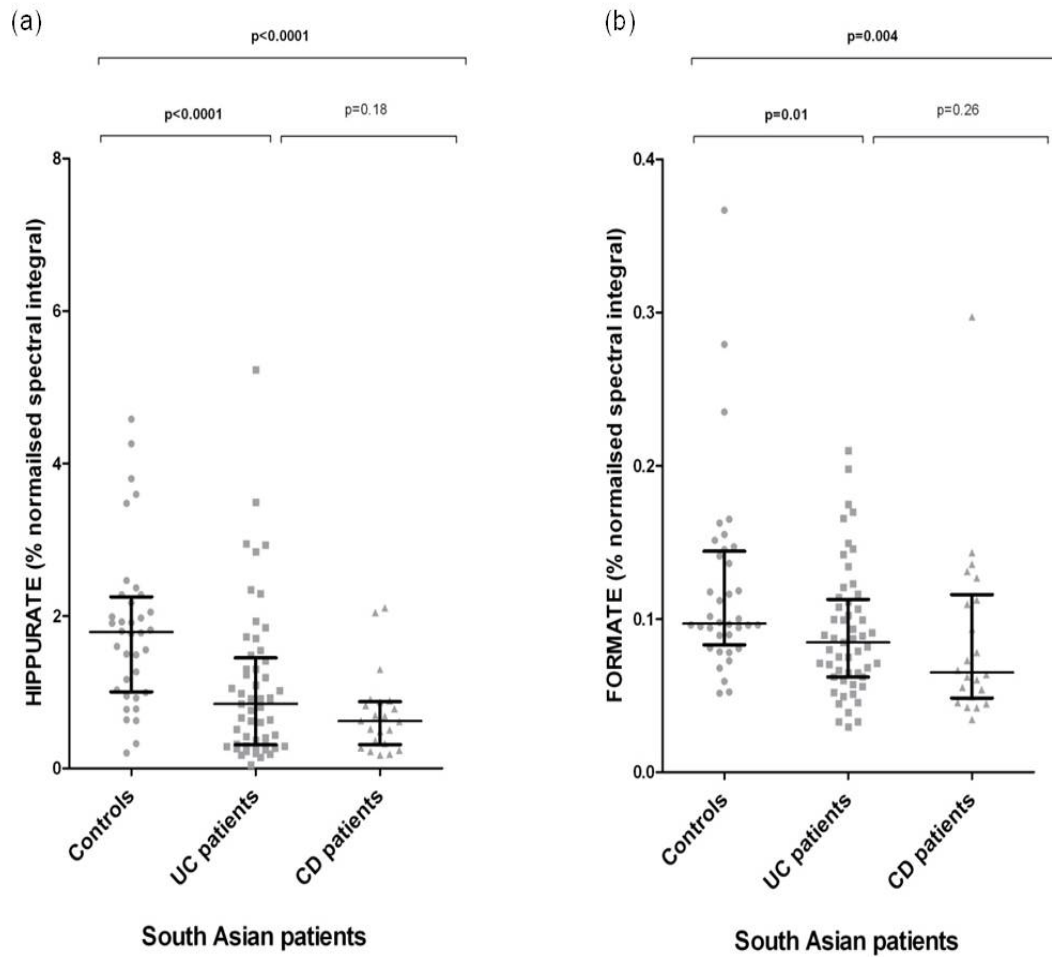


Figure 5-7: Urinary concentrations of (a) hippurate and (b) formate (expressed as % normalised total spectral integral) for South Asian IBD patients and South Asian controls. (Median and interquartile ranges shown)

TMAO was significantly lower in the UC cohort compared to controls (UC: $P=0.005$). Allowing for multiple comparisons, there was no significant difference between TMAO levels in CD patients and controls ($P=0.02$) or between the UC and CD cohorts. No significant differences were observed between the cohorts with regard to 4-cresol sulphate or DMA (Figure 5-8).

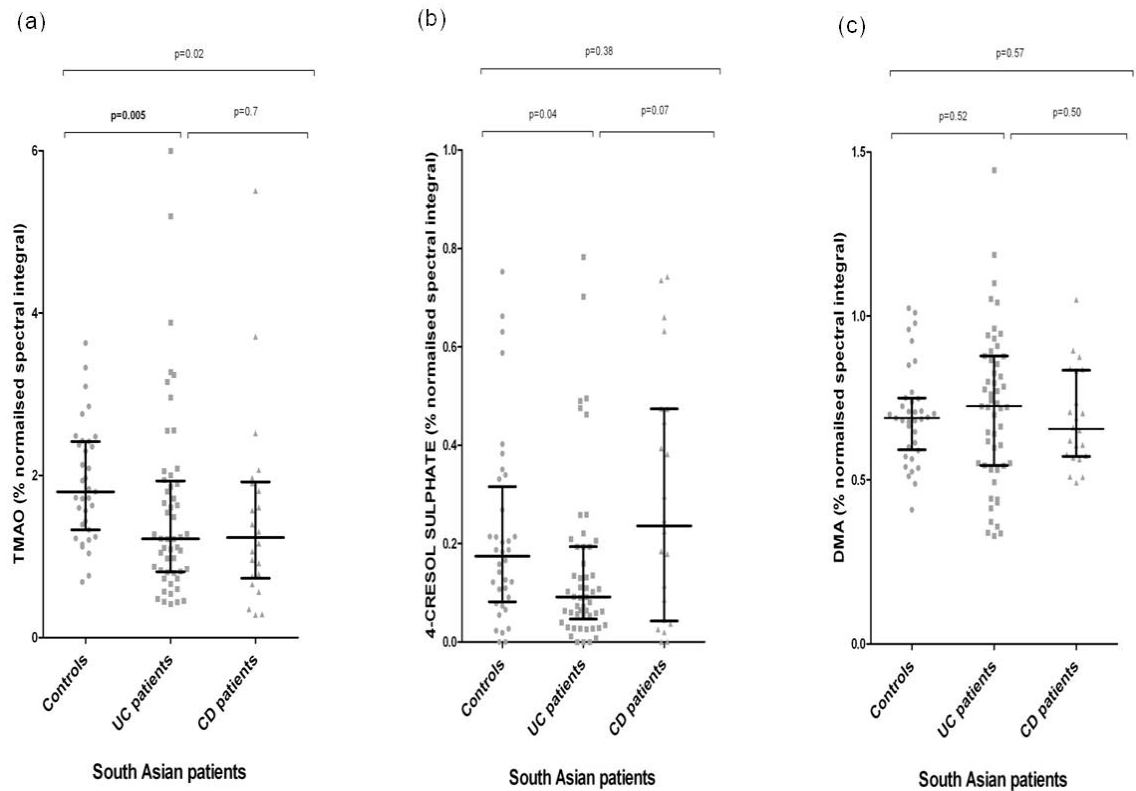


Figure 5-8: Urinary concentrations of (a) TMAO, (b) 4-cresol sulphate and (c) DMA (expressed as % normalised total spectral integral) for South Asian IBD patients and South Asian controls.

(Median and interquartile ranges shown)

5.3.3 South Asian lifestyle & dietary comparisons

Analysis of the lifestyle histories of the South Asian participants showed no significant differences between the control and disease cohorts. When the detailed dietary histories were compared, there were a significantly higher number of controls ($P=0.01$) who had consumed fish in the preceding 24 hours before urine collection, compared to the disease cohorts. There were no dietary differences in benzoate containing foods between the South Asian groups.

5.3.4 A comparison of the urinary NMR profile of South Asian and white Northern European IBD patients

The metabolites that were quantified in the South Asian cohort (hippurate, formate, TMAO, 4-Cresol Sulphate & DMA) were compared to a white Northern European IBD cohort. The demographic details of the cohorts are described in **Table 5-4**.

Table 5-4: Demographic details of the white Northern European and South Asian IBD cohorts including sex, age and disease location

	N. European UC	South Asian UC	N. European CD	South Asian CD
Number (Male/Female)	60 (30/30)	53 (39/14)	86 (47/39)	22 (14/8)
Median [range] age (years)	40 (17-66)	44 [19-75]	33 (16-66)	38 [18-82]
Disease location: n (%)	E1: 14 (23)	E1: 8 (15)	L1: 18 (21)	L1: 3 (14)
	E2: 18 (30)	E2: 10 (19)	L2: 25 (29)	L2: 8 (36)
	E3: 28 (47)	E3: 35 (66)	L3: 43 (50)	L3: 11 (50)

L1, ileal disease; L2, colonic disease; L3, ileocolonic disease; E1, proctitis; E2, disease limited to the splenic flexure; E3, disease proximal to the splenic flexure (Silverberg et al., 2005).

When the NMR spectra of Northern European and South Asian IBD patients were compared, hippurate and 4-cresol sulphate levels were significantly lower in South Asian UC patients compared to Northern European UC patients (**Figure 5-9**). There was no significant difference between the CD cohorts.

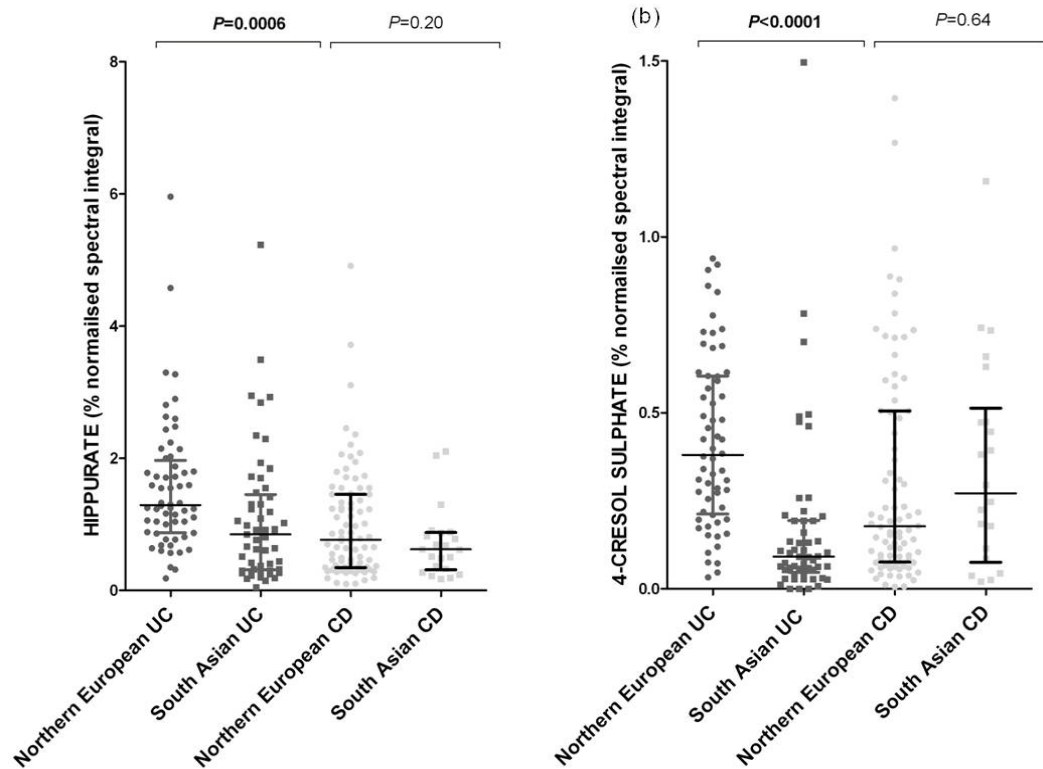


Figure 5-9: Urinary concentrations of (a) hippurate and (b) 4-cresol sulphate (expressed as % normalised total spectral integral) for white Northern European and South Asian IBD patients.

(Median and interquartile ranges shown)

No differences were observed for formate, TMAO or DMA levels between the UC or CD cohorts (**Figure 5-10**).

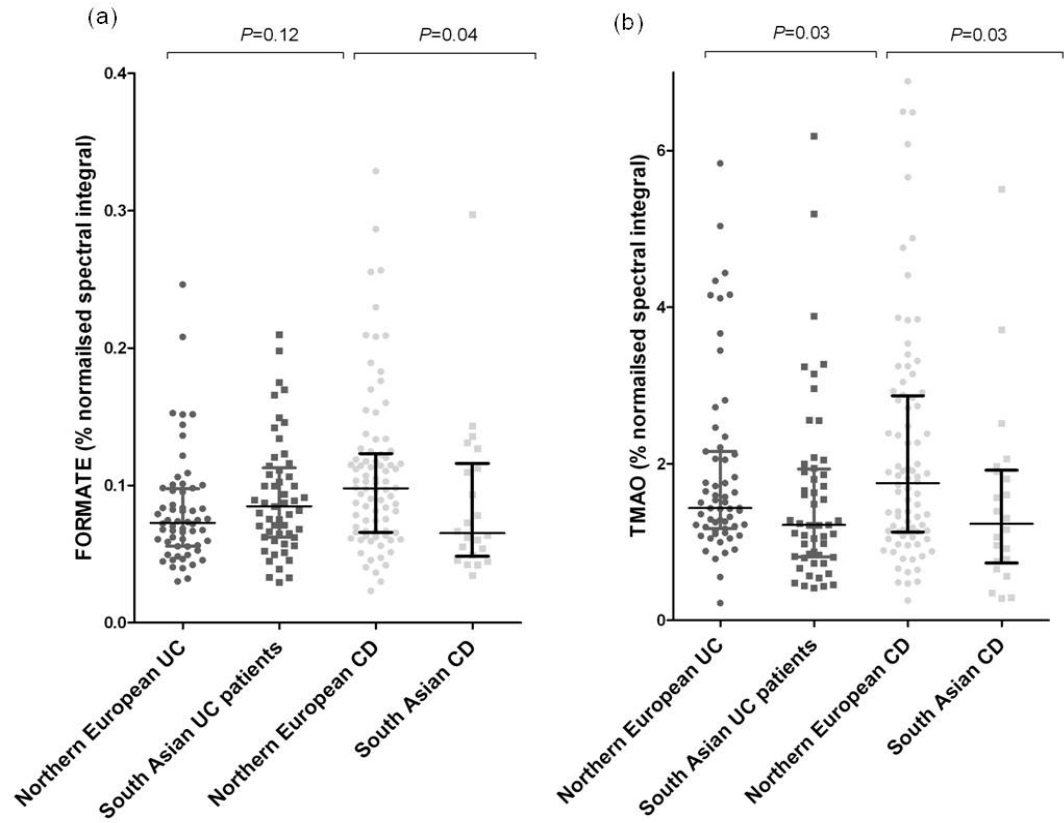


Figure 5-10: Urinary concentrations of (a) formate and (b) TMAO (expressed as % normalised total spectral integral) for white Northern European and South Asian IBD patients. (Median and interquartile ranges shown)

5.3.5 Potential confounders between ethnic groups

The lifestyle and dietary intake of the two cohorts was carefully examined to determine if there were any significant confounders between the groups. The South Asian UC cohort consisted of significantly more males than the Northern European UC cohort ($P=0.01$) and there were more South Asian UC patients that were vegetarians compared to Northern European ($P=0.02$). However, when comparing dietary intake of meat in the 24 hours prior to the urine collection, the values were not significantly different. There were no other differences between the cohorts with regard to lifestyle or dietary consumption.

When disease extent was compared between the two cohorts, there was significantly more extensive colitis in the South Asian UC cohort compared to the Northern European UC cohort ($P=0.04$). In order to identify if differences in disease phenotype was the reason for the observed variation in hippurate and 4-cresol-sulphate, the results were stratified by phenotype (E1, E2, E3) and levels of hippurate and 4-cresol sulphate compared for each of the UC phenotypes. The metabolites remained significantly different between the ethnic groups (E1: Hippurate $P=0.005$, 4-cresol sulphate $P<0.0001$; E2: Hippurate $P=0.006$, 4-cresol sulphate $P<0.0001$; E3: Hippurate $P=0.004$, 4-cresol sulphate $P<0.0001$).

5.4 DISCUSSION

It is now well established that gut bacteria play an integral part in the development of IBD and specific differences in the gut microflora between CD, UC, and healthy control subjects have been reported (Sartor, 2008). Urinary metabolic profiling is a non-invasive technique that can provide valuable insights into disease pathogenesis. It allows an integrated view of the biochemistry in complex organisms and is strongly influenced by the gut microbiota. Differences in the urinary NMR profile have been previously observed in Northern European IBD cohorts and controls (Williams et al., 2009), however this is the first study to apply urinary metabolic profiling to a different ethnic group with IBD.

5.4.1 Differences between South Asian IBD patients and controls

The first part of the study aimed to investigate if there were differences in the urinary levels of gut dependent metabolites (hippurate, TMAO, formate, DMA and 4-cresol sulphate) between South Asian UC, CD and healthy participants. Significantly lower levels of urinary hippurate and formate were observed in the South Asian UC and CD cohorts, compared to healthy South Asian controls. Lower levels of TMAO were also seen in South Asian UC patients compared to controls. No significant differences were seen between the UC and CD cohorts. The finding of a reduced level of hippurate in the UC and CD cohorts compared to controls is consistent with what was observed in Caucasian studies. However, the reduced levels of TMAO and formate seen in the South Asian IBD populations have not been described previously.

Urinary Hippurate

Urinary hippurate is the product of the conjugation of benzoate and glycine. Conjugation occurs via the formation of an intermediate, benzoyl CoA, and takes place in the mitochondria of the liver and the kidney (Schachter D and Taggart JV, 1953) (**Figure 5-11**). Benzoate is formed by the microbial degradation of dietary aromatic compounds in the intestine, such as polyphenols, purines, aromatic organic acids and amino acids (Gibson, 1967). However, it is also present in a variety of foodstuffs including most berries, fruits and fermented dairy products (Sieber R et al., 1995).

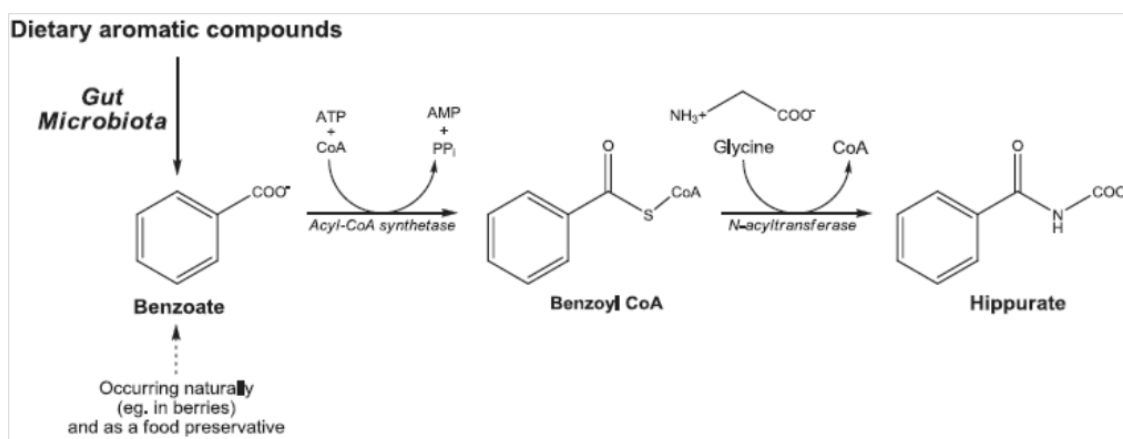


Figure 5-11: The synthesis of hippurate (reproduced from Williams et al. (2010)).

A previous human study has shown that urinary levels of hippurate are significantly lower in IBD patients when compared to healthy control individuals (Williams et al. 2009). Williams et al. (2010) conducted a subsequent study confirming that the reduced urinary hippurate levels were not simply due to differences in dietary benzoate ingestion, and that patients with CD do not have an intrinsic defect in the conjugation of benzoate with glycine. This implicates alterations in the gut microbiota in the observed

differences in hippurate excretion between IBD cases and controls and would be consistent with animal studies which have shown that urinary hippurate excretion is modulated according to the composition of the gut microbiota (Nicholls et al., 2003; Williams et al., 2002).

Recent experiments have adopted a strategy using a ‘transgenomic approach’ to link the gut microbiota with metabolic variation (Li et al., 2008). This uses a multivariate strategy to model the co-variation between the gut microbiota (using DNA fingerprinting) and the host metabolism (defined by urinary NMRS) in an attempt to functionally characterise the gut microbiota. The authors were able to correlate specific metabolites with a specific species or group of gut bacteria (**Figure 5-12**), interestingly hippurate correlated with certain species of Clostridia. Experiments using fluorescent in situ hybridization (FISH) of faecal samples have identified a reduction in Clostridia species for both UC (*Clostridium Coccoides*) and CD (*Clostridium Leptum*) (Sokol et al., 2006) which could potentially explain the reduced levels of hippurate in the IBD population.

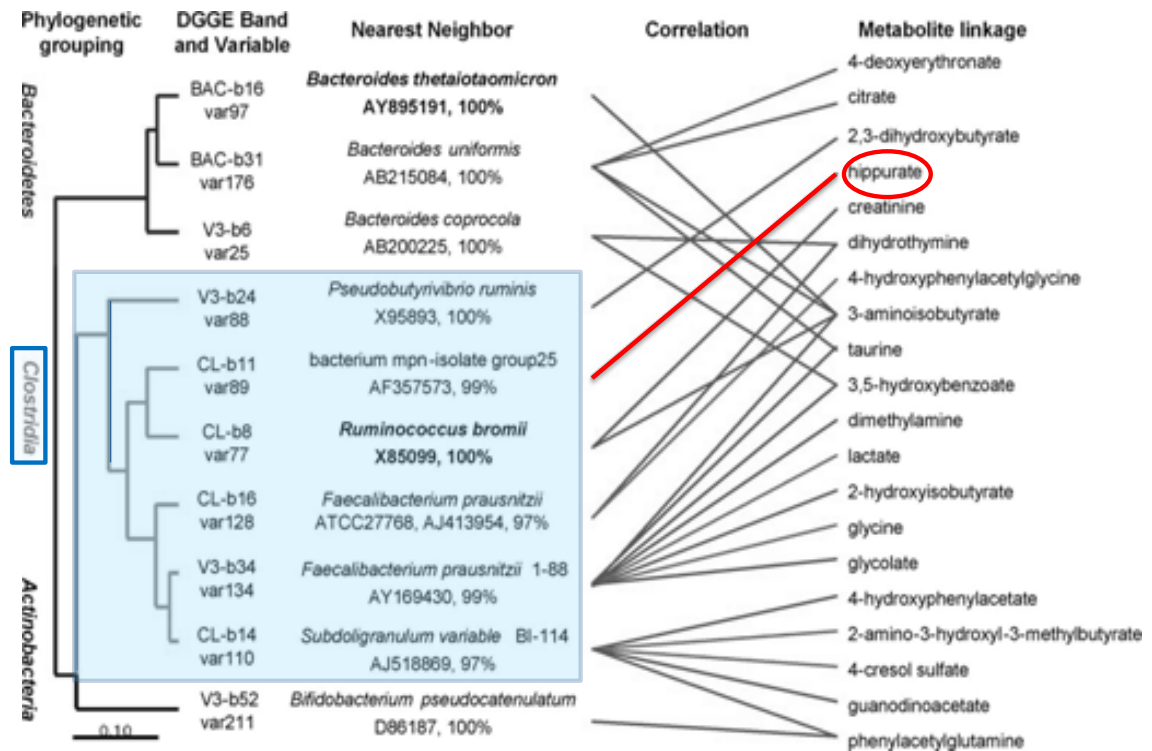


Figure 5-12: Phylogenetic associations with specific urine metabolites (reproduced from Li et al., 2008).

Urinary Formate

The finding of significantly lower levels of urinary formate in South Asian IBD patients compared to controls is in contrast to a previous study of Northern European IBD patients that reported no differences in formate levels between UC patients and controls, and significantly higher levels in Northern European CD patients compared to controls (Williams et al., 2009). Intestinal microbes have a major influence on formate metabolism as it is a characteristic product in the mixed acid fermentation of the *Enterobacteriaceae* and has been linked to *Escherichia coli* (Sartor, 2008). Further studies would be required to confirm that there are true differences in formate between the groups, but the reduced levels of this metabolite may suggest underlying differences in the gut microbiota between South Asian IBD patients and controls.

Urinary Trimethylamine-N-oxide (TMAO)

TMAO is an aliphatic amine and the oxidation product of trimethylamine (TMA), derived mainly from dietary sources, through the action of gut microbes (Al-Waiz et al., 1992). The consumption of fish products have been directly correlated with high levels of TMAO in the urinary NMR profile of humans (Bollard et al., 2005; Lenz et al., 2004). The South Asian control cohort had consumed significantly more fish in the 24 hours prior to the urine collection and this is likely to account for the observed differences rather than this being a true microbial difference between the groups.

Although the numbers in the control group were less than suggested by the power calculation, significant differences were demonstrated. However, one cannot exclude the fact that with higher participant numbers other metabolite differences between the cohorts may have been revealed.

5.4.2 Comparing the metabolic profiles of IBD patients of different ethnicities

Both Northern European and South Asian IBD populations have demonstrated lower levels of urinary hippurate when compared to controls. This would suggest that there are changes in the gut microbiota that are disease dependent and independent of ethnicity. However, when the IBD cohorts were directly compared, hippurate and 4-cresol sulphate levels were significantly lower in the South Asian UC group compared to Northern European UC patients. No metabolite differences were seen between the CD cohorts.

Similar to urinary hippurate, 4-cresol sulphate metabolism has also been linked with certain *Clostridia* species (Li et al., 2008). It is the product of the bacterial metabolism of the amino acid tyrosine and to a lesser extent phenylalanine. It is only synthesised by a few bacteria including *Clostridia* spp. and *Bacteroidetes*. (Bone et al., 1976) and a reduction in this metabolite would be consistent with already published studies showing a reduction in *Clostridia* species with UC (Sokol et al., 2006). It is of note that there was no significant difference in the level of this metabolite between South Asian UC patients and controls. However, there was a trend towards lower 4-cresol sulphate levels in South Asian UC patients ($P=0.04$) and a significant difference may be revealed in a larger study.

5.4.3 Potential explanations for metabolite differences between ethnic groups

The metabolite differences between the UC cohorts could suggest subtle variations in the gut microbiota between patients of different ethnic origin with UC. As both hippurate and 4-cresol sulphate have been associated with *Clostridia spp*, a reduction in this bacterial species may be responsible for the lower urinary metabolite levels in South Asian UC patients. Further molecular and microbiological experiments would be required to confirm this hypothesis.

NMRS studies such as this one could potentially be prone to bias from confounders like demographic, lifestyle and dietary differences between the South Asian and Caucasian populations. When the UC cohorts were compared, there was a higher male to female ratio in the South Asian cohort but no significant differences in lifestyle were noted. Gender-related differences in metabolites are debatable. One study (Li et al., 2008) described higher levels of 3-aminoisobutyrate in males, and creatine was higher in females. However, another study by Lenz (2004) showed no differences in metabolites between genders.

Dietary differences were observed between the groups with a significantly higher number of vegetarians in the South Asian cohort. However, there was no difference in the number of participants consuming meat products in the 24 hours prior to the urine collection. Long-term vegetarian status is less likely to be as important, as recently consumed foods because the metabolism and elimination of metabolites occurs quickly eg. low doses of benzoate are metabolised and eliminated rapidly, with the peak response of urinary hippurate excretion within 1-2 hrs of benzoate ingestion (Cathcart-

Rake et al., 1975). Further support for the differences in the group being real is that lacto-vegetarians have been reported as having a higher amount of urinary hippurate excretion which is the reverse of what was observed in this study (Stella et al., 2006).

Another potential confounding variable is the difference in disease extent observed between the UC cohorts, as there were a significantly higher number of patients with extensive colitis in the South Asian UC cohort. This distinction reflects the disease distribution demonstrated in Chapter 3. The influence of disease extent on the metabonomic profile has been previously addressed by Williams et al. (2009) and revealed no significant variation by disease location; this was confirmed by the subgroup analysis by disease phenotype in this study, suggesting that metabolic differences are a generic phenomenon for each disease. Studies of the bacterial flora in IBD have also found little evidence of microbial heterogeneity in different disease locations (Sokol et al., 2008).

5.4.4 Conclusion

This study has demonstrated significant differences in the levels of gut dependent metabolites in a cohort of South Asian IBD patients compared to controls. It is difficult to make definitive claims regarding the exact significance of urine metabolite changes. Epithelial injury has been well documented in IBD (McKenzie et al., 1996) and may contribute to the metabolic profile as will changes in intestinal permeability (Kennedy et al., 2000). However, if intestinal permeability was thought to have a significant contribution in IBD patients, one would have expected patients with IBD to have a higher level of metabolites excreted in the urine rather than the lower levels of hippurate

and formate that were observed. While the study shows the potential for urinary metabolic profiling to demonstrate differences between IBD patients and healthy controls, translation to clinical medicine is in part limited by the fact that there were no significant metabolite differences between the South Asian UC and CD cohorts. However, there was a trend towards lower levels of hippurate and formate in South Asian CD patients compared to UC patients and it is possible that a larger study may confirm this. What is clear is that significantly lower levels of hippurate have now been reported in both Northern European and South Asian UC and CD populations and the metabolite appears to be specific for IBD. The association of hippurate with *Clostridia* spp makes the findings biologically plausible. The search is on to better link urinary metabolites with specific microbial strains and this could provide an essential insight into the pathogenesis of IBD.

Diagnostic tools for IBD are currently limited to endoscopy with histology and radiological imaging. Numerous serological biomarkers have been proposed for the diagnosis of IBD, however, for clinical applications, none of the current markers stand-alone. More accurate tools are needed for early diagnosis of IBD in particular non-invasive approaches that can be used in place of endoscopy. In the future, urinary metabolic profiling could provide a faster, less-invasive, and cost-effective technique for diagnosis of IBD.

CHAPTER 6

THE PREVALENCE AND DIAGNOSTIC VALUE OF SEROLOGICAL MARKERS IN SOUTH ASIAN IBD PATIENTS

6.1 INTRODUCTION

The chronic intestinal inflammation seen in IBD is likely to be the result of an abnormal immunological response to commensal bacteria within the gut lumen in a genetically susceptible host. The detection of antibodies directed at microbial antigens may directly reflect the relevant microorganisms involved or may simply be related to increased gut permeability in IBD patients. The most thoroughly studied serological markers are auto-antibodies against an unidentified nuclear lamina protein present in neutrophils (atypical perinuclear anti-neutrophil cytoplasmic antibody) and antibodies against mannose epitopes from the yeast *Saccharomyces cerevisiae* (anti-*Saccharomyces cerevisiae* antibodies) (Main et al., 1988). More recently, antibodies to the outer-membrane porin C of *Escherichia coli* (anti-OmpC) has been reported in CD patients (Landers et al., 2002).

As the exact target antigen of atypical perinuclear anti-neutrophil cytoplasmic antibody (atypical pANCA) has not been identified, there are currently no sensitive or specific assays available to screen for the antibodies. Indirect immunofluorescence (IIF) of sera incubated with human peripheral blood neutrophil substrate is the only widespread method to detect the antibodies. A sample is deemed positive for atypical pANCA staining when it exhibits a broad inhomogeneous rim-like staining of the nuclear periphery with multiple intranuclear foci. However, detection rates can be highly variable due to differences in available substrate and the technique requires experienced observers for reliable interpretation (Joossens et al., 2004).

Anti-Saccharomyces cerevisiae antibodies (ASCA) (IgA and IgG) can be detected using ELISA. Approximately two-thirds of CD patients with ASCA-IgG are also positive for ASCA-IgA, but 0-19% of patients have only ASCA-IgA antibodies, suggesting that it is important to measure both type of antibody. Approximately 50-80% of CD patients are reported to be ASCA positive but only 2-14% of UC patients and 1-7% of healthy controls (Peeters et al., 2001;Quinton et al., 1998). Specificity can be as high as 90% when both ASCA IgA and IgG are positive. Sensitivity is in the range of 41-76% (Vermeire et al., 2001).

Detection of the IgG anti-OmpC antibody is also performed using ELISA. In CD, the reported seroprevalence is 55% (Landers et al., 2002). The prevalence in UC and healthy controls was low (2% and 1.3% respectively) (**Table 6-1**).

Table 6-1: Prevalence of atypical pANCA, ASCA and anti-OmpC in Caucasian IBD patients and controls			
	Crohn's disease	Ulcerative colitis	Healthy controls
Atypical pANCA (%) [*]	5-25	40-80	2.5
ASCA (IgA or IgG) (%) [*]	50-80	2-14	1.7
Anti-OmpC (%) [*]	55	2	1.3
Data adapted from: [*] Quinton et al., 1998; Peeters et al., 2001, [*] Landers et al., 2002, [*] Saxon et al., 1990; Rump et al., 1990; Bossuyt, 2006.			

6.1.1 IBD phenotype associations with serological markers

There have been few studies linking atypical pANCA with a particular UC phenotype. A small American study has reported an association with atypical pANCA and severe left sided ulcerative colitis, which is resistant to treatment (Sandborn et al., 1996) but these findings have not been repeated. Atypical pANCA has been associated with a subgroup of CD patients characterised by UC-like colitis (Vasiliauskas et al., 1996). High prevalence of pANCA positivity in this group often makes it difficult to distinguish colonic CD from UC serologically.

The expression of antibodies to *Saccharomyces cerevisiae* (ASCA) has been associated with a CD phenotype that is remarkably similar to that of patients with the common *CARD15/NOD2* polymorphisms associated with IBD (Dassopoulos et al., 2007; Quinton et al., 1998; Peeters et al., 2001). They are also at greater risk for complicated disease behaviour (Vasiliauskas et al., 2000) and require surgery more frequently (Forcione et al., 2004). Stratification of CD patients based on the presence of anti-OmpC showed an association with penetrating disease behaviour and therefore potentially leading to a higher rate of surgical intervention (Mow et al., 2004).

There is limited data on the prevalence, sensitivity and specificity of serological markers in the South Asian population. Studies have been scarce and involved small numbers (Makharia et al., 2007; Ghoshal et al., 2007).

6.1.2 Study aims

This study aimed to determine the prevalence of IBD associated serological markers in a cohort of South Asian IBD patients. A secondary aim was to calculate the accuracy of the serological markers in differentiating UC from CD in South Asians and to determine if there was any association with disease phenotype. The final study aim was to investigate if the sensitivities and specificities of the serological markers varied between ethnic groups. As it is difficult to draw reliable comparisons between serological data obtained in different laboratories due to varying methodologies and inter-observer variability, the South Asian data was compared to a white Northern European IBD cohort using the same experimental techniques. To date, there have been no studies directly comparing the frequencies of antibodies associated with IBD in different ethnic groups.

6.2 METHODS

6.2.1 Ethics Approval

The St. Mary's Hospital Research Ethics Committee approved this study and site-specific approval was obtained from each of the participating trusts. All participants gave written informed consent.

6.2.2 Patient recruitment

The same process as described in paragraph 5.2.2 was used to recruit South Asian patients and healthy South Asian controls. Dr Horace Williams (Imperial College London) provided the sera for 196 white Northern European IBD patients (114 CD: 82 UC).

6.2.3 Sample collection

A 3ml venous blood sample was collected into a plain tube for each subject (191 South Asian IBD patients [52 CD: 139 UC] and 36 healthy South Asian controls. Samples were centrifuged at 2500 rpm for 15 minutes and 500µL of serum (the supernatant) were aliquoted into microvials and stored at -80°C prior to use. The demographic and clinical characteristics of the South Asian patients and controls are displayed in **Table 6-2**.

Table 6-2: South Asian participant characteristics for serological typing			
	Crohn's disease	Ulcerative colitis	Controls
Number (Male/Female)	52 (26/26)	139 (83/56)	36 (20/16)
Disease location	L1: 5 L2: 22 L3: 25	E1: 16 E2: 34 E3: 89	- - -
Disease Behaviour	B1: 32 B2: 16 B3: 4	- - -	- - -
Patients that required surgery	19	7	-
L1=ileal disease, L2=colonic disease, L3=ileocolonic disease; B1- nonstricturing nonpenetrating; B2= stricturing; B3= penetrating; E1=proctitis, E2=disease limit distal to splenic flexure, E3=extensive colitis (Silverberg et al., 2005).			

6.2.4 Atypical pANCA indirect immunofluorescence assay

Determination of atypical pANCA was performed by Dr. Chris Bunn at the Department of Immunology, the Royal Free Hospital, London. This laboratory has expertise in using serological markers associated with IBD. All samples were anonymised and personally analysed by Dr Bunn, this avoided any inter-observer variability. The indirect immunofluorescence technique was employed, using ethanol-fixed human peripheral blood neutrophil substrate (Nova Lite, INOVA Diagnostics, USA).

Sera were diluted 1:10 and incubated for 30 minutes at room temperature, washed using phosphate buffered saline (10 mM phosphate, 150 mM NaCl, pH7.30), and incubated with fluorescein isothiocyanate conjugated rabbit anti-human IgG (DAKO, Ely, UK). Following further incubation, the wash step was repeated and the slides were examined by ultraviolet light microscopy at x400 magnification. Samples were considered to possess atypical pANCA if they produced a broad, inhomogeneous rim-like staining of the nuclear periphery without nuclear extension (**Figure 6-1**).

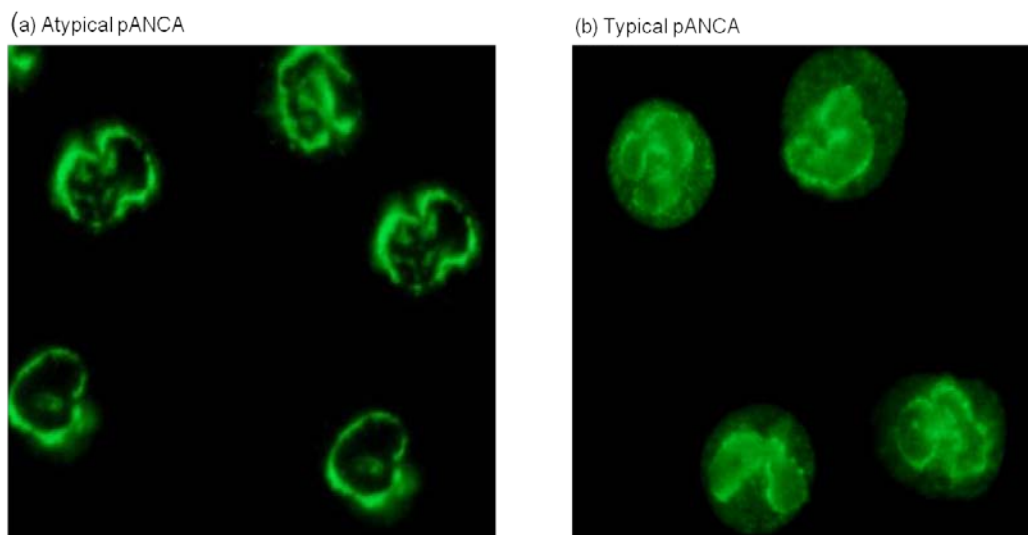


Figure 6-1: Ultraviolet light microscopy images taken by Dr Chris Bunn, Royal Free Hospital, London. The images demonstrate (a) Atypical pANCA staining, showing a broad inhomogeneous rim-like staining of the nuclear periphery without nuclear extension; For comparison, image (b) shows typical pANCA staining with a homogeneous rim-like staining of the perinuclear cytoplasm with nuclear extension.

6.2.5 ASCA IgG and IgA Assays

Principles of the technique

Quanta Lite® (INOVA Diagnostics, USA) enzyme-linked immunosorbent assays (ELISA) for ASCA-IgG, ASCA-IgA, and anti-OmpC IgA were used according to the manufacturer's instructions. Partially purified and disrupted *S. cerevisiae* antigen is bound to the wells of a polystyrene microwell plate. Pre-diluted controls and diluted patient sera are added to separate wells, allowing any ASCA antibodies to bind to the immobilised antigen. Unbound sample is washed away and an enzyme labelled with either anti human IgA or IgG conjugate is added to each well. A second incubation allows the enzyme labelled conjugate to bind to any patient's antibodies that have become attached to the microwells. A further washing step is performed and the remaining enzyme activity is measured by adding a chromogenic substrate and measuring the intensity of the colour that develops. The assay is then evaluated spectrophotometrically by measuring and comparing the colour intensity that develops in the patients' wells with the colour in the control wells.

Experimental technique: ASCA IgA assay

Experiments were performed by the author at the National Heart and Lung Institute, Imperial College London under the supervision of Dr. H. Williams and Dr. H. Sato. 96-well plates, with antigen pre-attached under conditions that preserve it in its native state, were supplied by the manufacturer. Pre-diluted controls were also supplied and included;

- ELISA negative controls, containing preservative and human serum without *S. cerevisiae* antibodies.
- ELISA ASCA IgA 'low positive', containing preservative and human serum antibodies to *S. cerevisiae*.
- ELISA ASCA IgA 'high positive', containing preservative and human serum antibodies to *S. cerevisiae*.

Preparation

Samples and reagents were brought to room temperature and thoroughly mixed. A horseradish peroxidase (HRP) wash concentrate (supplied by INOVA diagnostics) was diluted 1:40 with deionised water. A 1:101 dilution of each patient sample was prepared by adding 5µL serum to 500µL HRP sample diluent. Each sample was assayed in duplicate.

Procedure

- 100µL of the prediluted controls and patient samples were added to each well.
- The plate was covered, and incubated for 30 minutes at room temperature.
- All wells were then washed thoroughly, five times, using the diluted wash buffer.
- 100µL enzyme-labelled HRP IgA conjugate (goat anti-human IgA) was added to each well, with subsequent incubation at room temperature for 30 minutes.
- The wash step was then repeated.

- 100µL 3-3',5,5'-tetramethylbenzidine chromogen was added to each well and incubated in the dark for 30 minutes, prior to the addition of 'stop' solution (0.344 M sulphuric acid) to terminate enzymatic activity.
- The absorbance (optical density, OD) of each well was read at 450 nm within thirty minutes of stopping the reaction, using a 96-well plate reader (Dynatech MR 5000, Guernsey, Channel Islands).

As per the manufacturer's instructions, quality control checks were implemented. It was ensured that; (i) The absorbance of the 'high positive' control > 'low positive' control > negative control. (ii) The absorbance of the 'high positive' control was > 1.0, while that of the negative control < 0.2. (iii) The 'low positive' control absorbance was > twice that of the negative control, or > 0.25.

The test result was calculated by multiplying the ratio of sample to low positive control optical density with the number of units assigned to the control by the manufacturer (x 20 units). Reactivity is related to the quantity of antibody present in a nonlinear manner. Results are classified as negative (0.0– 20.0 Units), equivocal (20.1– 24.9 Units), and positive (≥ 25 U). Equivocal samples were retested prior to the allocation of a result.

Experimental technique: ASCA IgG assay

The technique for the assay of ASCA IgG was the same as that for the IgA assay, but the controls were human serum IgG antibodies, and the enzyme-labelled goat anti-human antibody was anti-IgG. **Figure 6-2** shows an overview of the methods used.

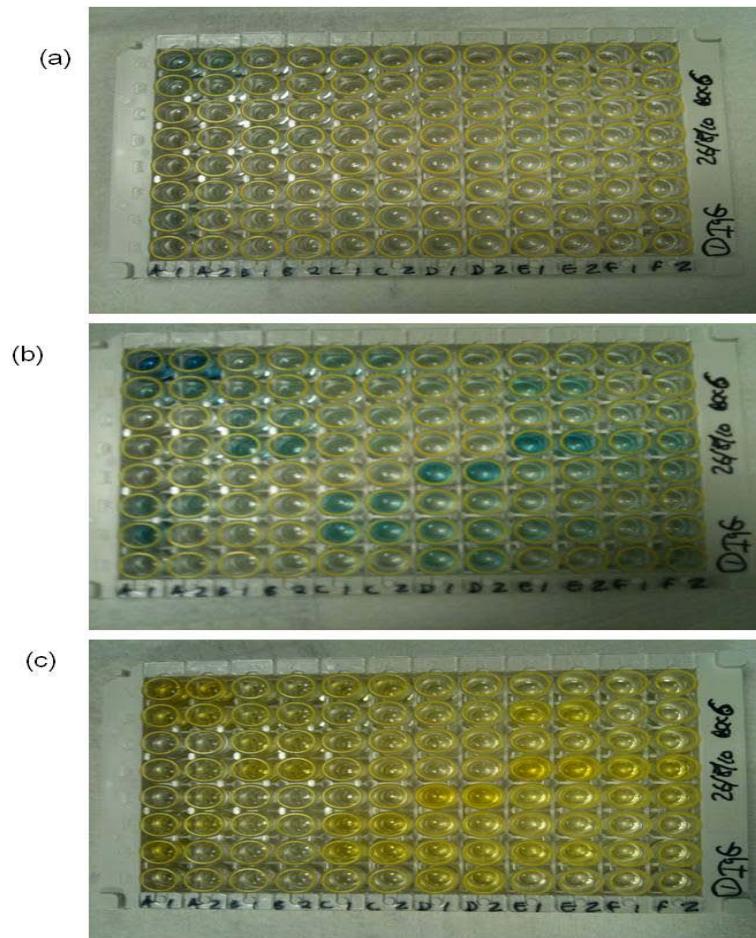


Figure 6-2: An outline of the experimental technique used to determine ASCA IgG positivity. (a) 96 well plates containing pre-attached antigens were supplied by the manufactures. (b) After adding the diluted patient sera and IgG conjugate, chromogen was added to each plate. (c) A stop solution was then added to terminate enzymatic activity and the optical density calculated.

6.2.6 Anti-OmpC Assay

To determine the presence of anti-OmpC antibodies, the same principles as above was applied but the microwell plate is bound with a proprietary lysate consisting of partially purified outer membrane and associated proteins (Omp) derived from *E.coli* and *Pseudomonas fluorescens*. An anti-human IgA conjugate is used for the enzyme label.

Experimental technique: Anti-OmpC assay

Human serum anti-OmpC antibodies are IgA, and so the controls were IgA and the enzyme-labelled goat anti-human antibody was anti-IgA. The technique for the assay was the same as for the ASCA IgA assay.

6.2.7 Interpretation of ELISA results

A cut-off for positivity was set at 25Uml, in accordance with the manufacturer's instructions and other published articles (Papp et al., 2008; Ferrante et al., 2007). The results were documented for frequency of positivity.

6.2.8 Statistical analysis

Statistical analysis was performed using GraphPad Prism Software (Version 5). The differences in prevalence of positive antibodies between groups and the relationship among antibodies and clinical parameters were studied using the Chi-square test. The sensitivity and specificity of ASCA, anti-OmpC and atypical pANCA was estimated to assess the validity of each antibody in distinguishing UC from CD. The results were compared with a white Northern European cohort also using the Chi-square test. As three variables were being compared (ASCA [IgG+IgA], anti-OmpC and atypical pANCA), a Bonferroni correction was applied and the level of significance set at $P < 0.02$.

6.3 RESULTS

6.3.1 The prevalence of ASCA, anti-OmpC and atypical pANCA in South Asian IBD patients and healthy controls

Fifty-two patients with CD were investigated; 50% (26/52) were positive for ASCA IgA or IgG (40% [21/52] for ASCA IgA and 38% [20/52] for ASCA IgG), 54% (28/52) showed anti-OmpC positivity and 15% (8/52) were positive for atypical pANCA.

In the UC cohort, 159 patients had serology performed. In all, 48% (67/139) were positive for atypical pANCA and 27% (38/139) positive for ASCA antibodies (22% [31/139] for ASCA IgA and 14% [19/139] for ASCA IgG). 35% (49/139) were positive for anti-OmpC. **Figure 6-3** outlines the number of positive results for ASCA (IgA & IgG), anti-OmpC and pANCA antibodies in CD patients, UC patients and healthy South Asian control subjects. The control cohort of 36 patients had 25% (9/36) positive for ASCA, 22% (8/36) for anti-OmpC and 3% (1/36) were positive for atypical pANCA.

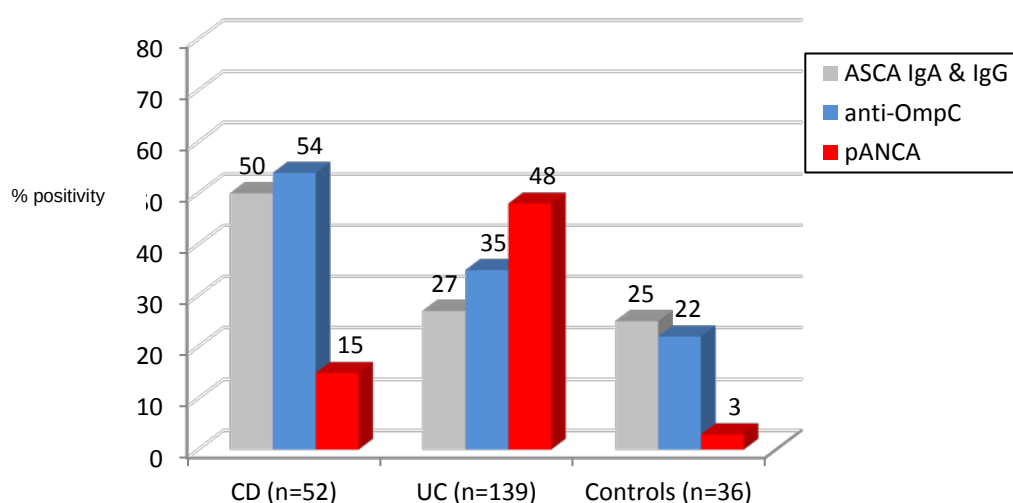


Figure 6-3: Positivity of ASCA (IgA & IgG), anti-OmpC and atypical pANCA in South Asian IBD patients and healthy controls.

The prevalence of ASCA and anti-OmpC was significantly higher in the CD cohorts compared to controls ($P=0.019$ and $P=0.003$) and significantly higher numbers of the UC cohort were positive for atypical pANCA compared to controls ($P<0.0001$).

6.3.2 Sensitivity and specificity of serological markers for differentiating South Asian IBD patients from healthy South Asian controls

The sensitivity and specificity of the serological markers were calculated to determine their potential utility in differentiating South Asian IBD patients from healthy controls. The accuracy of ASCA (IgA or IgG isotypes) and anti-OmpC for differentiation of CD from healthy controls showed a sensitivity of 50% and 54% but a higher specificity (75% & 78%). Using pANCA positivity, UC could be differentiated from controls with a sensitivity of 48% and a specificity of 97% respectively. The combination of ASCA-/pANCA+ did not show any higher sensitivity and specificity than pANCA alone for identifying UC patients from controls **Table 6-3**.

Table 6-3: Accuracy of ASCA, anti-OmpC and atypical pANCA to distinguish patients with inflammatory bowel disease from healthy controls in a South Asian population		
	Sensitivity	Specificity
ASCA+ (Total)	50% (26/52)*	75% (27/36)
OmpC	54% (28/52)*	78% (28/36)
Atypical pANCA	48% (67/139)**	97% (35/36)
ASCA+/pANCA-	42% (22/52)*	75% (27/36)
ASCA-/pANCA+	32% (45/139)**	97% (35/36)
*Specificity for CD, **Specificity for UC		

6.3.3 Diagnostic accuracy of ASCA and atypical pANCA in distinguishing CD from UC in South Asian patients with IBD compared to a white Northern European cohort

When ASCA (IgG & IgA), anti-OmpC and atypical pANCA were used to distinguish between South Asian patients with UC and CD, pANCA demonstrated the highest specificity for patients with UC (85%) and ASCA was highly specific for CD (73%). The combination of ASCA-/pANCA+ was the most specific for a diagnosis of UC (92%).

The sensitivity and specificity of the serological markers were analysed in a white Northern European IBD cohort using the same methodology (**Table 6-4**). There was no significant difference in the sensitivities of the serological markers between the South Asian and Northern European populations. However, anti-OmpC showed a significantly higher specificity in the Northern European CD population compared to the South Asian CD population ($P=0.007$).

Table 6-4: Accuracy of serological markers to differentiate ulcerative colitis from Crohn's disease in South Asian and white Northern European IBD patients

	South Asian		Northern European			
	Sensitivity % (n)	Specificity % (n)	Sensitivity % (n)	*P value	Specificity % (n)	*P value
ASCA+ (Total) for CD	50% (26/52)	73% (101/139)	54% (61/114)	<i>P=0.67</i>	84% (69/82)	<i>P=0.051</i>
Anti-OmpC for CD	54% (28/52)	65% (90/139)	39% (45/114)	<i>P=0.13</i>	82% (67/82)	<i>P=0.007</i>
Atypical pANCA for UC	48% (67/139)	85% (44/52)	37% (30/82)	<i>P=0.09</i>	83% (95/114)	<i>P=0.83</i>
ASCA+/pANCA- for CD	42% (22/52)	89% (123/139)	46% (55/114)	<i>P=0.48</i>	95% (78/82)	<i>P=0.09</i>
ASCA-/pANCA+ for UC	32% (45/139)	92% (48/52)	31% (25/82)	<i>P=0.77</i>	86% (101/114)	<i>P=0.46</i>
*When values compared to the South Asian IBD population using the Chi square test. Significance value set at <i>P</i> <0.02.						

6.3.4 Clinical phenotype associated with serological markers in South Asian IBD patients

South Asian CD patients were stratified for clinical phenotype associated with a positive ASCA or anti-OmpC result (**Table 6-5**). The presence of ASCA antibodies (either IgA or IgG) or anti-OmpC significantly correlated with small bowel CD with or without colonic involvement ($P=0.01$ & $P=0.04$ respectively). When CD patients with nonstricturing, nonpenetrating (B1) behaviour were compared to patients with more aggressive behaviour (B2 or B3), the detection of positive ASCA antibodies was associated with a more aggressive disease behaviour ($P=0.002$). ASCA and anti-OmpC positivity was significantly higher in South Asian CD patients with a history of surgery compared to CD patients without a history of surgery ($P=0.009$ for both).

Table 6-5: Serology results for South Asian stratified by Crohn's disease location and behaviour			
	ASCA+ (IgA or IgG)	Anti-OmpC	Atypical pANCA
<i>Disease location</i>			
Ileal (L1)	60% (3/5)	60% (3/5)	0% (0/5)
Colonic (L2)	23% (5/22)	45% (10/22)	23% (5/22)
Ileocolonic (L3)	64% (16/25)	60% (15/25)	12% (3/25)
<i>Disease Behaviour</i>			
Nonstricturing, nonpenetrating (B1)	34% (11/32)	53% (17/32)	19% (6/32)
Stricturing (B2)	63% (10/16)	56% (9/16)	6% (1/16)
Penetrating (B3)	75% (3/4)	50% (2/4)	25% (1/4)
Required surgery	74% (14/19)	74% (14/19)	16% (3/19)
Disease location and behaviour as per the Montreal classification (Silverberg et al., 2005). ASCA+: Ileal involvement L1 + L3 [19/30] vs no ileal involvement L2 [5/22]; $P=0.01$. Anti-OmpC+: Ileal involvement L1 + L3 [18/30] vs no ileal involvement, L2 [10/22]; $P=0.04$. ASCA+: B1 [11/32] vs B2/B3 [13/20]; $P=0.002$. ASCA+: Required surgery [14/19] vs ASCA+ but didn't require surgery [12/33]; $P=0.009$.			

The presence of atypical pANCA was not significantly associated with disease extent in UC or with requirement for surgical (**Table 6-6**).

Table 6-6: Atypical pANCA results stratified by ulcerative colitis phenotype	
	pANCA+
UC Phenotype* % (n)	
Proctitis (E1)	56% (9/16)
Left sided (E2)	53% (18/34)
Extensive (E3)	45% (40/89)
Surgery (Colectomy)	43% (3/7)
*No significant differences in pANCA positivity between disease phenotypes ($P=0.58$).	

6.4 DISCUSSION

IBD can be variable in its clinical presentation. Therefore, biomarkers that could accurately distinguish IBD patients from patients suffering from diarrhoea or other non-specific gastrointestinal symptoms would be very helpful in clinical practice. In addition, it can be a challenge to classify some IBD patients into CD or UC using accepted clinical, endoscopic, histological, and radiological criteria. This has particular relevance when initiating drug therapy or planning surgical intervention. Non-invasive markers that could accurately help separate UC and CD patients would be of great value. Serological markers associated with IBD have the potential to be a desirable clinical tool as they can be measured with little intervention and can be specific for CD and UC.

The role of serological markers associated with IBD is rapidly developing and although not currently in routine clinical practice, they are likely to be a valuable tool in the future. To date, the vast majority of studies in ASCA, anti-OmpC and atypical pANCA have been performed in white Northern European and North American populations but it is not clear if they have worldwide utility. Information on the role of ASCA and pANCA is limited for the South Asian population and this is the first study to report the association of anti-OmpC with South Asian IBD patients and the association of ASCA and atypical pANCA with disease phenotype in this population.

ASCA and anti-OmpC antibodies are known to be predominantly associated with CD, and have been shown to have a higher positivity in Caucasian CD patients compared to the background population. This current study showed a significantly higher prevalence

of these antibodies in South Asian CD patients compared to healthy South Asian controls (total ASCA, $P=0.018$; anti-OmpC, $P=0.003$). Although the markers had a high specificity for diagnosing CD (ASCA 75% and anti-OmpC 78%) their sensitivity was rather low (ASCA 54% and anti-OmpC 54%). These findings are similar to the ASCA and anti-OmpC sensitivity identified in many Caucasian studies (Devlin et al., 2007; Landers et al., 2002; Sandborn et al., 2001). Atypical pANCA positivity was significantly higher in the South Asian UC cohort compared to controls ($P<0.0001$) giving a very high specificity of 97% but again sensitivity was limited to only 48%.

Overall, the observed results would suggest that although the serological markers have good specificity, their lack of sensitivity limits their use as a screening tool in clinical practice.

6.4.1 The role of serological markers for differentiating UC from CD

Previous reports from Caucasian populations have suggested that ASCA and atypical pANCA may be of value in differentiating between UC and CD (Peeters et al., 2001; Quinton et al., 1998). Increased specificity has been obtained by assessing a combination of ASCA and pANCA, which has been shown to be helpful in differentiating UC from CD (specificity ranges from 95% to 99%). In South Asian IBD patients, pANCA has a high specificity for separating UC and CD (85% for UC) and this was increased further to 92% specificity using a combination of pANCA+/ASCA-, although this was associated with a reduction in sensitivity (48% → 32%). There was no difference in the sensitivity or specificity of pANCA between the ethnic groups. ASCA antibodies also had similar sensitivities and specificities in South

Asian and white Northern European IBD cohorts. However, anti-OmpC was significantly more specific for differentiating CD from UC in the white Northern European population compared with the South Asian cohort. As the specificity of this marker is significantly lower in the South Asian CD population, it may be a less valuable serological marker for differentiating CD from UC in this ethnic group. Larger studies are required to confirm this finding.

6.4.2 Stratifying IBD phenotype using serological markers

The presence of ASCA and anti-OmpC could potentially be used for disease stratification in Caucasian populations as they were associated with a more aggressive disease phenotype (Arnott et al., 2004) with a higher risk for surgery (Forcione et al., 2004). In this present study, a univariate analysis demonstrated that the presence of ASCA antibodies was associated with a more aggressive B2/B3 disease phenotype in CD ($P=0.002$), small bowel involvement ($P=0.01$) and an increased risk for surgery ($P=0.009$). Anti-OmpC was also associated with ileal disease ($P=0.04$) and an increased risk of surgery ($P=0.009$). The knowledge of an IBD patient's disease course could be extremely helpful in planning medical therapy. There are no current markers available in clinical practice and although research has suggested that genetic markers may predict a disease course (eg. *CARD15/NOD2* associated with more aggressive disease behaviour), these are non-specific and as allele frequencies vary between ethnic groups, they are therefore unhelpful in predicting disease course. However, in view of the strong associations of ASCA and anti-OmpC with disease phenotype in both white Northern European and South Asian populations, these serological markers could be potentially helpful in the future. It is clear from the literature that antibodies remain

relatively constant during the course of the disease and are independent of disease activity (Vasiliauskas et al., 1996; Vasiliauskas et al., 2000). Therefore, antibodies are not suitable for monitoring disease activity.

In this study, atypical pANCA was not associated with a particular UC phenotype. This is in contrast to a pilot study by Sandborn et al. (1996) which demonstrated that pANCA+ UC patients have a higher probability of developing left sided UC, which has a more aggressive course requiring early surgery. It is difficult to compare the studies, as Sandborn's study was small and focused on patients with left-sided disease rather than all UC phenotypes. The observed differences are unlikely to be related to ethnic differences as no such association with left-sided disease was noted for the white Northern European cohort examined in this study.

6.4.3 Conclusion

Based on these results, it would suggest that ASCA, anti-OmpC and atypical pANCA antibodies have a limited clinical value for the primary diagnostic workup of IBD in the South Asian population. The determination of antibodies may be a more useful diagnostic tool in patients with IBD unclassified as it may help guide medical or surgical therapy. At present, serology can only be recommended as an endoscopy adjunct to assist in the differential diagnosis and prognosis of IBD. Both ASCA and atypical pANCA have similar sensitivities and specificities in South Asian and white Northern European cohorts. However, anti-OmpC had increased specificity for distinction between individuals with CD and UC in the white Northern European cohort compared to South Asian IBD patients.

In addition to playing a role in IBD diagnosis, there is a great opportunity for serological markers to help identify patients at the time of diagnosis who will or will not develop a disabling disease course. This study has shown the potential role of ASCA and anti-OmpC in determining CD phenotype and this would greatly enhance the ability to tailor therapy at the time of diagnosis based on the risk of complicating disease.

As new serological markers are identified and additional studies examine the diagnostic performance of various combinations of markers, the utility of these markers in daily practice may improve. However, given the results from this study, it will be important to study any new markers in different ethnic groups to ascertain their worldwide application.

CHAPTER 7

DISCUSSION AND SUGGESTIONS FOR FUTURE WORK

This thesis describes a number of studies to give insights into the question of whether inflammatory bowel disease in South Asians is clinically, genetically and immunologically distinct from IBD in white Northern Europeans. This in turn may give insights into the pathogenesis of IBD overall.

IBD phenotype in UK South Asians

In Chapter 3, I have shown that the phenotype of IBD is different between South Asians and Northern Europeans living in the UK. A significantly higher proportion of South Asian patients had extensive UC and a lower number had proctitis when compared with the Northern European UC cohort. Disease behaviour was significantly different between the CD cohorts with less penetrating behaviour in the South Asian population and a lower requirement for surgery. This is the largest phenotyping study performed in UK South Asians and lends weight to the argument that there is ethnic variation in the phenotype of IBD.

Phenotyping studies such as this one may be prone to ascertainment bias and great care was taken to recruit the cohorts from the same hospitals in a defined geographic area therefore recruiting similar proportions of patients from secondary and tertiary care into each group. A potential bias may be that the South Asian and European cohorts were not collected over the same time frame: 18 months for the South Asian group and 6 months for the Northern European group. We aimed to collect populations of similar sizes to study, as Northern Europeans are the predominant ethnic group attending IBD clinics in North West London, the collection of this population was faster than that of

the South Asian population. A potential bias that could result from this would be that Northern European patients under annual review with less severe disease and on no regular medications may not have been included in the study. This would have been a relatively small number of patients and would have been likely to include more patients with limited left-sided disease/proctitis than patients with extensive colitis, therefore unlikely to impact on the observed differences in UC phenotype.

Future directions: IBD phenotyping in UK South Asians

Although the number of UC patients collected was large in comparison to previous South Asian phenotyping studies in the UK, the study had only a limited number of CD patients. Studies from Leicestershire have reported a significantly higher prevalence of UC in the UK South Asian population compared with Europeans, while the prevalence of CD in South Asian Hindus was significantly lower than Europeans (Probert et al., 1993). This study reflected these reported differences with 273 UC patients recruited over the study period compared with only 94 CD patients. This limited the analysis of first and second-generation CD patients. Future phenotyping studies in UK South Asians would benefit from a multi-centre approach for patient recruitment and should aim to have higher numbers of CD patients.

The highly significant differences in clinical phenotype between the ethnic groups prompted further investigations into potential underlying differences in genetic susceptibility and environmental triggers that could account for the observations.

Genetic differences between South Asian and white Northern European IBD patients

The potential influence of genetic susceptibility to IBD in South Asians was assessed by determining the frequency of single nucleotide polymorphisms and HLA alleles previously associated with IBD in Caucasians, in a South Asian IBD cohort (*IL23R*, *HLA-DRB1*1502* & *DRB1*0101* in UC patients; *CARD15/NOD2*, *ATG16L1*, *IRGM*, *IL23R*, *HLA-DRB1*1502* & *DRB1*0101* in CD patients). The results were compared with cohorts of South Asian healthy controls, and white Northern European IBD patients. For UC, the carriage frequency of *HLA-DRB1*1502* was significantly higher in South Asian UC cases compared with controls. However, this allele did not correlate with clinical phenotype. No differences were observed between South Asian UC cases and controls for the *IL23R* polymorphism. This result is consistent with a very recent report from a GWAS in India that showed no association of *IL23R* with UC in Indian Asians (Juyal et al., 2011). There was also no difference between South Asian CD cases and controls for IBD associated polymorphisms in *CARD15/NOD2*, *ATG16L1*, *IRGM* or *IL23R*.

When South Asian IBD patients were compared to white Northern European patients, there were significant differences between the cohorts, with higher frequencies of the *HLA-DRB1*1502* allele in South Asians compared to Northern Europeans and a significantly lower frequency of the *HLA-DRB1*0103* allele in the South Asian UC population. *CARD15/NOD2* polymorphisms associated with CD in Northern Europeans were rare in the South Asian CD population. These polymorphisms are known to be associated with terminal ileal disease (Lesage et al., 2002) but in view of the low

frequency of *CARD15/NOD2* polymorphisms seen in South Asian CD patients, they are less likely to influence disease phenotype in this population. As this study did not recruit a large number of South Asian patients with ileal CD, larger studies would be required to confirm this finding.

This study was under-powered with regard to CD, so it is possible with higher numbers differences may be reported. However, the findings from this thesis are consistent with a small study from India showing that IBD associated polymorphisms in *CARD15/NOD2* are not associated with CD susceptibility in this ethnic group (Pugazhendhi et al., 2008) . In addition, studies from other parts of Asia (Japan and Korea) have shown no association with *ATG16L1* and *IRGM* (Yamazaki et al., 2007; Yang et al., 2009).

Future directions: Insights into genetic susceptibility in South Asian IBD patients

As no significant differences were observed between CD cases and controls in the South Asian population, it would be informative to sequence the *CARD15/NOD2*, *ATG16L1*, *IL23R* and *IRGM* genes to determine if there are other loci that could confer susceptibility to CD for this population. Larger patient numbers are required to validate the genetic results observed in this study. The Wellcome Trust Consortium is conducting a UK wide GWA study of susceptibility genes in South Asian IBD patients living in the UK. DNA from patients involved in this study will form an integral part of the study.

It is widely accepted that the pathogenesis of IBD is likely to be due to a combination of variation in genetic susceptibility, environmental and immune factors. If the failure to

show any differences in the frequency of CD associated polymorphisms between South Asian cases and controls is not simply a matter of power then it raises the possibility that either CD in this ethnic group is associated with new polymorphisms of known susceptibility genes, or other susceptibility loci that have yet to be discovered or there may be less of a genetic effect and a stronger environmental influence on CD pathogenesis in South Asians.

Urinary metabolic profiling in UK South Asian IBD patients

An important potential environmental stimulus associated with IBD is the variation in gut microbiota, which is essential in triggering mucosal inflammation (Bhan et al., 1999). It was hypothesised that there would be differences in the microbiota between South Asian IBD cases and controls. The microbiota is difficult to investigate, as culture techniques are time-consuming and unreliable. Studies using fluorescent in situ hybridization on surgical resections have suggested alterations in the gut microbiota and this has been further supported by studies showing variation in the bacterial composition of faecal samples in UC and CD patients compared to controls (Sokol et al., 2006).

In this thesis, urinary metabolic profiling was used to investigate potential variations in the microbiota. Spectral resonances generated by Nuclear Magnetic Resonance (NMR) spectroscopy allow metabolites that are influenced by the gut microbiota to be quantified and therefore act as a surrogate marker of bacterial composition. Significant differences were found between South Asian IBD patients and healthy controls in the

urinary excretion of the microbial-influenced metabolites; hippurate, formate and TMAO. Dietary intake may influence the excretion of certain metabolites and therefore a careful dietary history was taken from each participant. Variation in the intake of fish products may explain the differences seen in TMAO between the groups but there were no dietary differences to account for the lower levels of formate and hippurate seen in the IBD populations. The highly significant difference in urinary hippurate observed between South Asian IBD patients and controls is consistent with what is reported in Northern European cohorts, suggesting that hippurate is a specific determinant of IBD that is consistent across ethnic groups. When the metabolic profiles of South Asian IBD patients were directly compared with a Northern European cohort, no differences were reported for CD but significantly lower levels of hippurate and 4-cresol sulphate were seen in the South Asian cohort compared with Northern Europeans. The metabolite differences observed between South Asian and Northern Europeans UC patients may reflect subtle differences in the gut flora between the ethnic groups, perhaps increasing susceptibility to UC in South Asians. Although it is possible that the observed differences could simply reflect differences in diet or different/additional environmental factors between these populations, it is of great interest that there are very consistent changes in hippurate excretion in South Asian and Northern European IBD patients.

Unfortunately, despite collecting a large number of samples for this study, the results were limited by the number of samples that could be processed due to the closure of the NMR facility at the Hammersmith Hospital. Efforts were made to try and get the remaining samples analysed in another facility with a similar NMR spectrometer (JOEL 500MHz Eclipse+). No other 500Hz spectrometers were available in the UK and for

this reason a collaboration was set up with Dr. Kishore Bakoo's research team at the Imperial NMR facility in Singapore. The remaining samples were couriered to Singapore for spectral acquisition, and the data sent back to the UK for analysis. After analysing 100 samples from Singapore it became apparent that the water suppression differed significantly in these samples, compared to those samples acquired in the UK (**Appendix 9**). This made it impossible to analyse the data comparatively. A new collaboration has been established, and all of the samples will be re-analysed at the new NMR facility at Imperial College in South Kensington London. Using these data, it will be possible to further validate the results.

Future directions: Learning more about the gut microbiota in South Asians

Data from epidemiological studies would suggest that as individuals migrate from a developing to a developed country the incidence of IBD rises. This underlines the relative importance of environmental factors and it would be informative to investigate if there are differences in the gut microbiota between first- and second-generation South Asians as this may give further insights into IBD pathogenesis.

In this thesis, I used a univariate approach to study metabolites that are known to be associated with the gut microbiota. The next stage would be to use multivariate statistical techniques to determine which metabolites differ between sample classes with a non-targeted approach. This will be useful in terms of hypothesis-generation and investigating potential pathophysiological explanations for the aetiology of IBD.

Molecular techniques have revolutionised the study of the microbiota and the recently described transgenomic methodology for exploring microbial host metabolic

relationships have revealed exciting results, allowing researchers to make a direct link between urinary metabolites and particular species of gut bacteria (Li et al., 2008). The lower levels of hippurate are likely to correspond with a reduced number of *Clostridia* species in the microflora of IBD patients. Future studies combining NMRS with ribosomal RNA typing are likely to identify further microbial candidates that have strong host metabolic connectivity.

As NMR techniques become more sophisticated, there is the potential that urinary metabolic profiling could provide a faster, less invasive, and cost-effective technique for diagnosing IBD.

Other potential environmental factors affecting IBD pathogenesis in UK South Asians

Smoking

Smoking represents one of the most consistently observed environmental influences on CD. It is hypothesised that smoking modulates outcomes in CD through effects either on vascular endothelium, on mucin production, or other toxic effects of smoking by-products on gut mucosal homeostasis. Chapter 3 reported significantly higher numbers of smokers in the Northern European CD cohort compared to South Asian CD patients. Previous studies have demonstrated that smoking is more frequently associated with the development of strictures or fistulae with an increased risk of surgery (Picco and Bayless, 2003; Reese et al., 2008), therefore smoking could represent an important contributory influence on established disease. The low number of smokers in the South

Asian CD cohort may in part explain the differences in disease behaviour between the CD cohorts.

Diet and Vitamin D

This thesis did not specifically investigate detailed dietary differences between the ethnic cohorts but previous studies have reported that dietary influences may be important in IBD pathogenesis (de Silva et al., 2010); this would be an important area for future research.

There is mounting evidence that Vitamin D possesses a wide range of immunomodulatory properties (Brennan et al., 1987) and epidemiological/experimental evidence implicates Vitamin D insufficiency as a risk factor for developing IBD (Raman et al., 2011; Wang et al., 2010). As Vitamin D levels are related to sun exposure and skin colour, they are therefore likely to vary between ethnic groups. To evaluate this further, we formed a collaboration with Dr Andrew Milestone and Dr Ailsa Hart at St. Mark's Hospital, London. A study was conducted of 215 CD patients (70 South Asian and 145 Northern European), mean 25-hydroxyvitamin D levels (25-OHD) varied significantly with ethnic origin (Northern European 31.1ng/mL, South Asian 21.4ng/mL; $P < 0.01$). A greater proportion of South Asians (45.8%) than Northern Europeans (24%) had 25-OHD levels < 20 ng/mL. Further observational and therapeutic studies are needed to determine the exact role of Vitamin D in IBD pathogenesis as this may be important factor contributing to IBD pathogenesis in this ethnic group.

The role of serological markers in South Asian IBD patients

The idea that in IBD the intestinal flora triggers and drives an aberrant immune response is further supported by the occurrence of antibodies directed at microbial antigens and enteric bacteria (Cohavy et al., 2000). Chapter 6 examined the prevalence of serological markers in a South Asian population and the association of autoantibodies with clinical phenotype. This study concentrated on the most studied serological markers (ASCA, anti-OmpC and atypical ANCA). Significant differences in the positivity of antibodies for ASCA and anti-OmpC were observed when the South Asian CD patients were compared to controls. Significantly higher numbers of atypical pANCA+ patients were reported in the South Asian UC cohort compared to controls. The exact target of the atypical pANCA antibody is not known but ASCA reactive antigens have been reported in the granulomas of bowel resections and on infiltrating lymphocytes and neutrophils in inflamed tissue from CD patients (Oshitani et al., 2003). The results would therefore suggest that there are differences in enteric bacteria between South Asian IBD patients and controls.

Serological markers have been previously investigated in the clinical setting, to aid in the diagnosis of IBD or to differentiate UC from CD. The sensitivity of the serological markers did not vary between ethnic groups but the specificity of anti-OmpC was significantly lower in the South Asian population for differentiating CD from UC.

Future directions: The role of serological markers in different ethnic groups

This study only concentrated on well known serological markers. Further studies are warranted to investigate if newer serological markers including anti-glycans antibodies

(ALCA, AMCA & ACCA) or anti-I2 are more sensitive and specific in determining the presence of disease and disease behaviour in the South Asian population.

In both Caucasian and South Asian populations, currently available serological markers (ASCA, anti-OmpC and atypical pANCA) have only a limited use in clinical practice and are only modestly helpful in distinguishing between a diagnosis of CD or UC. However, the study demonstrates that if serological markers advance into routine clinical practice, it will be important to determine their utility in different ethnic populations.

7.1 SUMMARY

This thesis has demonstrated differences between two distinct ethnic groups with regard to disease phenotype and provided insights into potential differences in genetic susceptibility and the gut microbiota between the cohorts. The exact reason for the phenotypic differences between the cohorts remains unclear and it is likely that no single susceptibility gene or bacterial species is directly associated with a particular phenotype but rather it is a combination of genetic and environmental factors. Studies in migrant populations such as this one can provide vital clues into underlying disease pathogenesis. A better understanding of the mechanisms responsible for disease pathogenesis, as well as development of less invasive diagnostic techniques, will help improve the quality of life of the many individuals that suffer from IBD.

CHAPTER 8

PUBLICATIONS ARISING FROM THIS THESIS TO DATE

8.1 FULL PUBLICATION

Walker DG, Williams HRT, Kane SP, Mawdsley JE, Arnold J, McNeil I, Thomas HJW, Teare JP, Hart AL, Pitcher MCL, Walters JR, Marshall SE, Orchard TR. Differences in inflammatory bowel disease phenotype between South Asians and Northern Europeans living in North West London, United Kingdom. *American Journal of Gastroenterology* (2011), July; 106 (7):1281-9.

Differences in Inflammatory Bowel Disease Phenotype between South Asians and Northern Europeans Living in North West London, UK

David G. Walker, MRCP^{1,2}, Horace R.T. Williams, PhD, MRCP^{1,2}, Stephen P. Kane, MD, FRCP³, Joel E. Mawdsley, MD, MRCP³, Jayantha Arnold, FRCP⁴, Ian McNeil, MD, FRCP⁴, Huw J.W. Thomas, PhD, FRCP², Julian P. Teare, MD, FRCP², Ailsa L. Hart, PhD, MRCP^{5,6}, Maxton C.L. Pitcher, MD, FRCP⁵, Julian R.F. Walters, MA, FRCP^{1,7}, Sara E. Marshall, PhD, FRCPath⁸ and Timothy R. Orchard, MD, FRCP^{1,2}

- OBJECTIVES:** The incidence and prevalence of inflammatory bowel disease (IBD) is increasing throughout Asia. Since the 1950s, there has been substantial migration from South Asia (India, Pakistan, and Bangladesh) to the United Kingdom. The aim of this study was to define the clinical phenotype of IBD in UK South Asians living in North West London, and to compare the results with a white Northern European IBD cohort.
- METHODS:** The phenotypic details of 367 South Asian IBD patients (273 ulcerative colitis (UC) and 94 Crohn's disease (CD)), undergoing active follow-up in five North West London hospitals, were compared with those of 403 consecutively collected white Northern European IBD patients (188 UC and 215 CD).
- RESULTS:** The phenotype of IBD differed significantly between the two populations. 63.0% of South Asian UC patients had extensive colitis compared with 42.5% of the Northern European cohort ($P < 0.0001$). Proctitis was uncommon in South Asian UC patients (9.9 vs. 26.1% in Northern European patients, $P < 0.0001$). In the South Asian CD cohort, disease location was predominantly colonic (46.8%). CD behavior differed significantly between the groups, with less penetrating disease compared with Northern Europeans ($P = 0.01$) and a reduced need for surgery ($P = 0.003$).
- CONCLUSIONS:** The phenotype of IBD in South Asians living in North West London is significantly different from that of a white Northern European IBD cohort. Knowledge of ethnic variations in disease phenotype may help to identify key genetic, environmental, and behavioral factors contributing to the development of IBD.

Am J Gastroenterol 2011; 00:000–000; doi:10.1038/ajg.2011.85

INTRODUCTION

The inflammatory bowel diseases (IBDs), ulcerative colitis (UC) and Crohn's disease (CD), are common causes of gastrointestinal disease in the United Kingdom that are associated with substantial morbidity. Epidemiological studies from around the world suggest a widely varying incidence of IBD, with highest rates of UC and CD reported in Northern Europe and North America (1). Previous research in IBD has predominantly focused on large white European/American populations, but recent studies suggest there may be phenotypic differences in other ethnic groups (2,3). There is still limited information on ethnic groups from South Asia, but

a small population-based study of UC in North India showed that the incidence (6 per 100,000) and prevalence (44.3 per 100,000) are similar to those observed in some Western countries (4).

Between 1950 and 1980, there was substantial migration of South Asians (originating from India, Pakistan, and Bangladesh) to Europe; currently, 2.08 million South Asians live in the United Kingdom (5). Studies of IBD in this population could provide valuable information that may help to distinguish the degree of influence of genetic and environmental pathogenic factors. However, to date studies have been scarce and involved relatively small patient numbers.

¹Department of Medicine, Imperial College London, London, UK; ²Department of Gastroenterology, Imperial College Healthcare NHS Trust, St. Mary's Hospital, London, London, UK; ³Department of Gastroenterology, West Middlesex University Hospital, London, UK; ⁴Department of Gastroenterology, Ealing Hospital, London, UK; ⁵Department of Gastroenterology, St. Mark's Hospital, London, UK; ⁶Department of Surgery and Cancer, Imperial College London, London, UK; ⁷Department of Gastroenterology, Imperial College Healthcare NHS Trust, Hammersmith Hospital, London, UK; ⁸Department of Clinical Immunology, University of Dundee, Dundee, UK. **Correspondence:** David G. Walker, MRCP, Department of Gastroenterology, Imperial College Healthcare NHS Trust, St Mary's Hospital, Praed Street, London W2 1NY, UK. E-mail: david.walker1@imperial.ac.uk
Received 27 September 2010; accepted 3 February 2011

A retrospective epidemiological study compared the incidence and prevalence of UC among South Asians living in the city of Leicester between 1972 and 1989 with Europeans from the same area (6). South Asian patients tended to have more extensive disease and less proctitis compared with Europeans but this did not reach statistical significance. However, differences in IBD phenotype were reported in a small study of 82 South Asian IBD patients from Manchester, with South Asians demonstrating a higher rate of pancolonic disease (34/82: 41%) compared with a white British IBD cohort (41/158: 26%) (7), but it was not clear whether this related to a combination of UC and CD patients or UC alone.

There have been relatively few studies of CD in the UK South Asian population (8–10). Jayanthi *et al.* identified only 28 new cases of CD in the city of Leicester between 1972 and 1989 compared with 179 new cases in Europeans (9). In this study, disease behavior, prevalence of extraintestinal manifestations (EIMs), and smoking history were not investigated, but the predominant disease location for both South Asians and Europeans was the terminal ileum, and both groups had similar surgical requirements. However, recent genetic studies from South Asia have demonstrated significantly lower carriage of CD-associated polymorphisms in *CARD15/NOD2* (the gene predominantly associated with ileal CD) than in Northern Europeans (11–13), and the lack of phenotypic differences in the Leicester study may reflect its small size.

Many South Asians live in London, and ~220,000 have settled in the North West London boroughs of Brent, Ealing, Harrow and Hounslow, making up over 22% of the population in this area (5). On the basis of previous clinical and genetic studies, it was hypothesized that disease phenotype would vary between South Asian IBD patients and white Northern Europeans. The primary aim of the study was to determine the phenotype of IBD in a South Asian population and to compare it with a Northern European cohort from the same area. A secondary aim was to investigate the impact of environmental factors on disease phenotype by comparing that of first-generation and second-generation South Asians. A move toward a Northern European phenotype in second-generation South Asians could suggest that Western environmental stimuli are important pathogenic factors in determining disease phenotype.

METHODS

Patient recruitment

To capture the maximum number of South Asian IBD patients in North West London, the IBD and gastroenterology outpatient clinics of five hospitals (St. Mary's and Hammersmith Hospital of Imperial Healthcare NHS Trust, St. Mark's Hospital, Ealing Hospital, and West Middlesex University Hospital) were monitored for an 18-month period (1 August 2008 to 1 February 2010). In North West London, current practice is that IBD patients are managed in secondary care by the gastroenterology team. Primary care physicians may be involved in joint care, e.g., managing disease flares or blood monitoring while on immunosuppressive therapy but patients remain under regular review in secondary care. Each of the five hospitals has a dedicated IBD outpatient clinic. Within these hospitals, patients with IBD are referred to this clinic. In

addition, patients seen in the IBD clinic include new referrals from primary care with signs and symptoms of IBD, new IBD patients referred by the attending internal physician or surgeon at the time of hospitalization, patients referred from endoscopy with IBD, IBD patients with an acute relapse, and patients with quiescent disease under regular follow-up. To ensure complete coverage of the population, all the general gastroenterology clinics were also reviewed for South Asian patients with IBD.

The white Northern European population was recruited from the IBD clinics of the same five hospitals; patients of mixed race were excluded. White Northern Europeans represent the predominant ethnic group attending the IBD clinics in North West London. Northern European patients were recruited from all five sites consecutively for a 6-month period between 1 August 2009 and 1 February 2010.

Case definition

The clinical charts of all patients attending the IBD/gastroenterology clinics were screened by the attending gastroenterologist at the time of the patient's outpatient appointment. Only patients with a confirmed diagnosis of IBD based on clinical evaluation, endoscopic/radiological appearances, and histology (14,15) were recruited for the study.

Previous studies from the United Kingdom have defined the South Asian population as patients whose families originated from the Indian subcontinent (India, Pakistan, and Bangladesh) including South Asians who had migrated via East Africa (10,16). For comparison, this study used the same definition of South Asia and only included patients both of whose parents were of South Asian origin. All patients were aged ≥ 16 years.

UC was defined on clinical history and examination with endoscopic appearances of continuous mucosal inflammation of the colon affecting the rectum and a variable extent of the colon in continuity, and characterized by a relapsing and remitting course (14). CD was defined on the basis of clinical evaluation and a combination of endoscopic findings (mucosal ulceration, skip lesions, cobblestone appearance, and terminal ileal ulceration), histological appearance (discontinuous intestinal inflammation and/or non-caseating granulomas) with or without radiological evidence of skip lesions, stricturing disease, or fistulizing disease. Stool culture results and mucosal histology were reviewed for all patients and those in whom there was any uncertainty regarding diagnosis (i.e., those with IBD, type unclassified (17), or colitis as a result of infection, radiation, or ischemia) were excluded from the study.

Patient demographics and disease phenotype

A questionnaire proforma was completed prospectively by the attending doctor at the time of the outpatient appointment. Data collection included gender, ethnic origin, place of birth, ethnicity and place of birth of the patients' parents, length of time in the United Kingdom, EIMs, a current drug history, and a smoking history. Patients were classed as smokers if they had smoked more than seven cigarettes per week for at least 6 months immediately before diagnosis of IBD and/or during follow-up. Non-smokers were those who never smoked, smoked a little and occasionally,

or had stopped smoking more than 6 months before IBD was diagnosed (18–20). The patient's clinical charts were then carefully reviewed by author DW to gather additional demographic and phenotypic details, including age at IBD diagnosis, length of follow-up, disease extent, disease behavior, surgical history, and to confirm any EIMs.

A second gastroenterologist (HW) randomly audited 200 sets of clinical charts from patients recruited to the study to confirm the clinical details were accurately documented. There was disagreement regarding one case and this was resolved after discussion between authors DW, HW, and TO.

Defining IBD disease behavior and extent

In the participating hospitals, it is standard protocol for patients with suspected IBD to have a full colonoscopy with ileal intubation, and biopsies taken from the terminal ileum, right colon, transverse colon, left colon, and rectum. Only UC patients that had either evaluation of the entire colon by colonoscopy with biopsy or who had a colectomy with histology reports were included in the study. UC disease extent was recorded by both endoscopic appearance and maximum histological extent. When comparing the ethnic groups, maximal disease extent was determined by histology and defined in terms of the Montreal classification (17) (ulcerative proctitis (E1), left-sided UC (E2), and extensive colitis (E3)). Only CD patients who had undergone a full colonoscopy with ileal intubation and biopsy (when not precluded by stenosis) were included in the study. Disease extent and behavior was based on a combination of histological extent and small bowel imaging. Patients had an oesophagoduodenoscopy with biopsy when clinically indicated. Disease behavior was considered cumulatively to the time of the most recent outpatient visit. The Montreal classification was also used to classify CD location (terminal ileal (L1), colonic (L2), ileocolonic (L3), and upper GI (L4)) and disease behavior (nonstricturing, nonpenetrating (B1), stricturing (B2), and penetrating (B3)). Subjects with both stricturing and penetrating disease were classified as penetrating disease (B3). Perianal fistulating disease alone did not constitute penetrating disease but was recorded as a modifier of the disease behavior (p). Surgical history was documented and included surgery performed in North West London and elsewhere (colectomy for UC patients, and bowel resection or treatment of related complications in CD patients).

EIMs

EIMs were determined from patient recall with confirmation in the clinical notes. EIMs were defined as peripheral arthropathy (type 1 or 2) (21), ankylosing spondylitis/sacroiliitis, skin complications (erythema nodosum or pyoderma gangrenosum), ocular complications (anterior uveitis or episcleritis), primary sclerosing cholangitis, and venous thrombosis.

Assessing the impact of environmental factors in the United Kingdom on disease development and phenotype

To determine the potential influence of environmental factors in the United Kingdom on the development of IBD in South Asians,

the rates of IBD diagnosed in the first, second, third, and subsequent decades since migration were examined. In addition, the potential impact of environmental factors on disease phenotype was assessed by comparing South Asians born in the United Kingdom with first-generation South Asians diagnosed within 10 years of arriving in the United Kingdom. This cohort was specifically chosen, as it was felt that it would best represent a group of patients that had less exposure to environmental stimuli in the United Kingdom compared with South Asians who had been born and grown up in the United Kingdom. The phenotype of second-generation South Asians was directly compared with the Northern European cohort to determine whether there was a move toward a Northern European phenotype in second-generation South Asians.

Statistical analysis

Data from the South Asian cohort were compared with the Northern European cohort using the Mann–Whitney *U*-test for continuous variables, such as duration of disease and age at diagnosis, and the χ^2 -test for categorical variables, such as disease extent and need for surgical intervention. The disease extent in first- and second-generation South Asians was compared using the χ^2 -test. *P* values < 0.05 were considered statistically significant.

To account for potential false positives resulting from multiple comparisons, a Bonferroni correction was applied to *P* values in the phenotypic data that showed significance (*P* < 0.05; by a factor of 3 in the UC cohort to allow for comparisons relating to disease location, requirements for surgery, and EIMs, and by a factor of 4 in the CD cohort to allow for comparisons of disease extent, disease behavior, surgical intervention, and EIMs). Corrected *P* values are expressed in the text as *P_c*. The data were analyzed using Graph Pad Prism Software (Version 5, La Jolla, CA).

Ethical Statement

All patients gave informed consent, and the study was approved by the St. Mary's Hospital Research and Ethics committee with site-specific approval from each participating hospital trust.

RESULTS

A total of 770 well-characterized IBD patients were recruited for the study (South Asian 367 and Northern European 403). In the South Asian population, UC was the predominant disease (UC 74.4%; CD 25.6% (273:94)), whereas in the Northern European population, there were similar proportions of UC and CD (46.7:53.3% (188:215)).

UC demographics

Of the 273 South Asian UC patients identified, there was a male predominance (1.4:1) in contrast to the Northern European population that had more female patients; however, this did not reach statistical significance (*P* = 0.06; Table 1). The median age at diagnosis was 30.7 years (inter-quartile range 23.3–42.1 years) for the South Asian population. A total of 231 patients (84.6%) were Indian, 39 (14.3%) Pakistani, and 3 (1.1%) Bangladeshi. All patients had a diagnosis of IBD confirmed in the United Kingdom.

Table 1. Demographic details of South Asian and white Northern European UC patients living in North West London, United Kingdom

	South Asian (n=273)	Northern European (n=188)	P value
Gender (male/female), n (%)	159 (58.2)/114 (41.8)	92 (48.9)/96 (51.1)	P=0.06
Age at diagnosis (median (IQR); years)	30.7 (23.3–42.1)	33.2 (25.1–43.2)	P=0.11 ^a
Age at diagnosis (Montreal) (17), n (%)			
≤16 (A1)	10 (3.7)	9 (4.8)	P=0.54
17–40 (A2)	183 (67.0)	125 (66.5)	P=0.96
>40 (A3)	80 (29.3)	54 (28.7)	P=0.99
Duration of follow-up (median (IQR); years)	10.5 (5.6–18.1)	9.3 (4.8–18.9)	P=0.27 ^a
Smoker, n (%) ^b	19 (6.9)	14 (7.4)	P=0.85

IBD, inflammatory bowel disease; IQR, inter-quartile range; UC, ulcerative colitis.

^aP value calculated using the Mann-Whitney U test. All other P values calculated using the χ^2 test.

^bSmoked more than seven cigarettes per week for at least 6 months immediately before diagnosis of IBD and/or during follow-up.

The majority of South Asian patients (67.0%) were aged between 17- and 40-years old at diagnosis (Montreal classification A2). No significant differences were observed in smoking habits between the South Asian and European cohorts (positive smoking history: 6.9 vs. 7.4%; $P=0.85$).

UC phenotype

The clinical features of the South Asian patients were compared with the Northern European cohort (Table 2). The disease phenotype differed significantly between the two populations, with 63.0% of South Asian patients having extensive colitis compared with 42.5% of the Northern European cohort ($P<0.0001$). The prevalence of left-sided disease was similar in each group (27.1 and 31.4%), but limited rectal disease was less common in South Asian UC patients (9.9 vs. 26.1% in Northern European patients; $P<0.0001$; Figure 1).

Colectomy rates in South Asians were numerically lower than the Northern European population but this did not reach statistical significance. A total of 18 South Asian patients (6.6%) had undergone resections for UC (16 for failed medical management and 2 for dysplasia related to UC). In all, 3 patients had an end ileostomy and 15 had restorative ileoanal pouches. This compared with 20 Northern European patients (10.6%) who had undergone colectomies (19 for failed medical management and 1 for dysplasia). Nearly all patients in both cohorts requiring colectomy had extensive colitis (15/16 in the South Asian cohort and 17/20 in the Northern European cohort).

A total of 24 (8.8%) of the South Asian UC population had at least one EIM of IBD. This was similar to the Northern European population (6.9%). The most common EIM in South Asians was

peripheral arthropathy (10 patients). In addition, 44 South Asian UC patients (16.1%) complained of arthralgia, but with no joint swelling or radiographic changes.

CD demographics

Of the 94 South Asian CD patients, the majority were Indian (87.2%). Median age at diagnosis was 30.7 years (inter-quartile range 21.4–42.1 years), which was similar to the Northern European cohort (Table 3). In both cohorts, the most common age range at diagnosis was 17–40 years (A2), but there were significantly fewer South Asian patients in the ≤16-year group (A1; $P=0.03$). Only 10 South Asian patients had a positive smoking history (10.6%) compared with 68 Northern Europeans (31.6%; $P<0.0001$).

CD phenotype

The South Asian CD phenotype was mainly colonic (L2 = 46.8%) and ileocolonic (L3 = 37.2%; Table 4). Isolated ileal disease was less common (L1 = 16%). Disease in the Northern European population was predominantly ileocolonic (L3 = 40.5%) and although there was less isolated colonic disease than in the South Asians (34.8 vs. 46.8%), this was not significantly different after correcting for multiple comparisons ($P=0.19$; Figure 2).

The clinical behavior pattern at last follow-up was nonstricturing, nonpenetrating (B1) in 72.3% South Asian patients; stricturing (B2) in 23.4% and penetrating (B3) in 4.3% of cases. Length of follow-up for the CD cohorts did not differ significantly. A total of 19 patients (20.2%) had perianal CD, including perianal fistulae and fissures. There was significantly less B3 behavior compared with the Northern European cohort ($P=0.04$). In all, 26.6% of South Asian patients had had surgical treatment for CD (including bowel resection, stricturoplasty, and surgical treatment of perianal disease). This was significantly lower than in the Northern European cohort (44.7%; $P=0.01$).

There were no significant differences in EIMs between South Asian and Northern European cohorts ($P=0.18$). Skin complications were the most common EIM in South Asian CD patients.

Demographics of first-generation South Asian IBD patients

The exact place of birth was not available for 17 South Asian patients (10 UC patients and 7 CD patients); hence, the analysis was performed on the remaining 350 patients (263 UC and 87 CD). The majority of South Asian patients studied were first-generation (born outside the United Kingdom; 64.6%: 226/350) and the remainder was second generation (born in the United Kingdom). In all, 83% of first-generation patients came directly from the Indian subcontinent (187/226) and the remaining 17% immigrated from East Africa (39/226). The median year that first-generation South Asian patients arrived in the United Kingdom was 1975 (inter-quartile range; 1968–1988).

To assess the effect of changing environment on the development of IBD in first-generation patients, the rates of diagnosis in the first, second, third, and subsequent decades since migration were examined. Of the 226 first-generation South Asian patients (171 UC and 55 CD), 3 patients had a provisional diagnosis of IBD before arriving in the UK (1 UC and 2 CD). For UC, of the remaining

Table 2. Phenotypic details of South Asian and white Northern European UC patients living in North West London, United Kingdom

	South Asian (n=273)	Northern European (n=188)	P value	Corrected P value (P _c)
<i>Disease location (Montreal) (17), n (%)</i>				
Ulcerative proctitis (E1)	27 (9.9)	49 (26.1)	P = <0.0001	P_c = <0.0001*
Left-sided UC (E2)	74 (27.1)	59 (31.4)	P = 0.36	
Extensive UC (E3)	172 (63.0)	80 (42.5)	P = <0.0001	P_c = <0.0001*
Patients requiring colectomy, n (%)	18 (6.6)	20 (10.6)	P = 0.16	
<i>EIMs, n (%)</i>				
Peripheral arthropathy	10 (3.7)	5 (2.7)	P = 0.47	
Ankylosing spondylitis/sacroiliitis	4 (1.5)	1 (0.5)	P = 0.43	
Skin complications ^b	4 (1.5)	4 (2.1)	P = 0.82	
Uveitis/episcleritis	6 (2.2)	4 (2.1)	P = 0.82	
PSC	3 (1.1)	2 (1.1)	P = 0.97	
Venous Thrombosis	2 (0.7)	1 (0.5)	P = 0.79	

EIM, extraintestinal manifestation; PSC, primary sclerosing cholangitis; UC, ulcerative colitis.

*P_c: after Bonferroni correction, the values remain significant.^bErythema nodosum and pyoderma gangrenosum.All P values were calculated using the χ^2 test.

Bold values are P values that showed significance (P < 0.05).

170 patients, more were diagnosed in the first decade (34.7%: 59/170) compared with 19.5% (33/170) in the second decade, 22.9% (39/170) in the third decade, and 22.9% (39/170) in subsequent decades. Thus, for UC, there was no substantial increase in disease diagnosis with time spent in the United Kingdom. In contrast, in the CD cohort (although small), very few of the first-generation patients were diagnosed within 10 years of arrival in the United Kingdom (7.6%: 4/53). The majority of patients were diagnosed in the second (37.7%: 20/53) and third decades (41.5%: 22/53). The remaining 13.2% (7/53) were diagnosed after being in the United Kingdom > 30 years.

Comparison of first- and second-generation South Asian IBD patients

To assess the potential impact of environmental factors in the United Kingdom on disease phenotype, first-generation South Asians diagnosed within the first 10 years of arriving in the United Kingdom were compared with South Asians born in the United Kingdom. There was no significant difference in phenotype between the UC groups, and the predominant UC phenotype was still extensive colitis (first-generation UC diagnosed within 10 years: 34/59 (57.6%) vs. second generation 63/92 (68.5%); P = 0.17). The numbers for CD were too small to be able to make any meaningful comparisons.

When South Asian UC patients born in the United Kingdom were compared with the Northern European UC population, there were again significantly higher rates of extensive UC in the South Asian cohort (63/92 (68.5%) vs. 80/188 (42.5%); P < 0.0001) despite a shorter follow-up period (median duration of follow-up: 8.5 vs. 9.3 years, respectively).

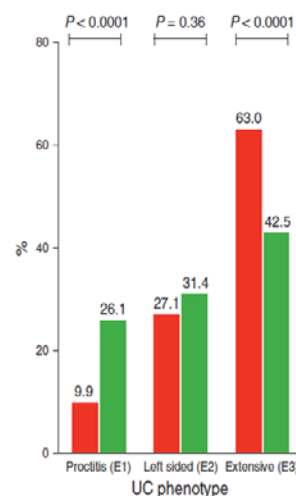


Figure 1. Ulcerative colitis (UC) phenotype for (■) South Asian and (■) white Northern European patients. P values were calculated using the χ^2 -test. A Bonferroni correction was applied to significant P values.

DISCUSSION

This is the largest study of IBD in the UK South Asian population, examining the phenotype of 367 patients. UC is more common than CD in the North West London South Asian IBD population (UC 74.4%:CD 25.6%), and disease extent in South Asian UC patients is significantly different from that of a white Northern European cohort. Studies such as this may be prone to

Table 3. Demographic details of South Asian and white Northern European CD patients living in North West London, United Kingdom

	South Asian (n=94)	Northern European (n=215)	P value
Gender (male/female), n (%)	48 (51.1)/46 (48.9)	94 (43.7)/121 (56.3)	P=0.23
Age at diagnosis (median (IQR); years)	30.7 (21.4–42.1)	25.9 (20.1–38.2)	P=0.09 ^a
Age at diagnosis (Montreal) (17), n (%)			
≤16 (A1)	5 (5.3)	30 (13.9)	P=0.03
17–40 (A2)	61 (64.9)	139 (64.7)	P=0.97
>40 (A3)	28 (29.8)	46 (21.4)	P=0.11
Duration of follow-up (median (IQR); years)	13.2 (7.1–20.9)	11.6 (6.1–19.6)	P=0.39 ^a
Smoker, n (%) ^b	10 (10.6)	68 (31.6)	P<0.0001

IBD, inflammatory bowel disease; CD, Crohn's disease; IQR, inter-quartile range.

^aP value calculated using the Mann-Whitney U test. All other P values were calculated using the χ^2 test.^bSmoked more than seven cigarettes per week for at least 6 months immediately before diagnosis of IBD and/or during follow-up.

Bold values are P values that showed significance (P<0.05).

ascertainment bias; hence, great care was taken to collect a complete population of South Asian IBD patients attending IBD and gastroenterology outpatient clinics in a defined geographical area and to compare them with a comparable population of white Northern European IBD patients also from the same area. Potential limitations of this study are: (i) that patients with IBD-related symptoms may not have been referred to the clinic and therefore not included in the study; however, this is likely to be a small number of patients and will be the same for both ethnic groups and (ii) only patients attending hospital IBD/gastroenterology outpatient clinics were included. IBD patients under the sole care of primary care physicians or surgical outpatients would not have been included. However, in North West London, it is current practice that IBD patients are managed in secondary care by the gastroenterology team. After surgery, patients remain under joint care with the gastroenterologists. Primary care physicians may be involved in joint care of patients, but are not responsible for the sole care of IBD patients. This practice was confirmed in a survey that we conducted of 10 local primary care practices.

UC phenotype in UK South Asians

The frequency of extensive colitis in South Asians is significantly higher (63.0%) than the Northern European population (42.5%) (P<0.0001). These findings support the suggestion from a previous smaller study that extensive colitis is more common in South Asian patients than in Northern European patients (7). A previous study by Carr and Mayberry (16) of 33 newly diagnosed South Asian patients in Leicester City reported that the disease

extent in UC is more extensive in second-generation UC patients (6/9, 66.6%) compared with first-generation (6/24, 25%). They proposed that an environmental risk factor may explain the difference. In this current study, the possibility of environmental factors in the United Kingdom influencing disease phenotype was investigated by comparing South Asians born in the United Kingdom with South Asians born in the Indian subcontinent and diagnosed within 10 years of arriving in the United Kingdom. Extensive colitis remained the predominant phenotype for each group, with no significant differences between the groups. Although it is not clear how long a patient needs to be exposed to relevant environmental factors for them to be significant, the high rates of extensive colitis among South Asians, together with the absence of differences between first- and second-generation UC patients, would support the idea that genetic factors may have a role in the susceptibility to extensive colitis in South Asians.

It is well recognized that patients with more extensive disease are at higher risk of colectomy (22,23), and nearly all patients in the study population that required colectomy had extensive colitis (17/18 in the South Asian cohort and 17/20 in the Northern European cohort). It is therefore surprising that only 17 out of 172 South Asian patients with extensive colitis required a colectomy compared with 17/80 in the Northern European cohort (9.9 vs. 21.3%; P=0.01). This may suggest a less severe phenotype of IBD in the South Asian population, and the finding is consistent with the findings of Probert *et al.* (6), which also suggested that South Asians with UC are at a lower risk of colectomy than Europeans.

South Asian UC patients reported EIMs similar to the Northern European population (8.8 vs. 6.9%; P=0.49), with peripheral arthropathy being the most common EIM in both cohorts. The frequency of EIMs was lower than previously described in South Asians (24). This may reflect a more rigorous stratification in this study—patients with arthropathy were only included if they had objective evidence of an inflammatory arthropathy and patients with arthralgia or any other cause of arthritis were excluded. Patients with back pain but without radiological evidence of sacroiliitis were also excluded. Uveitis and episcleritis were the only ocular complications recorded and not conjunctivitis that had been included as an EIM in a previous study of Indian UC patients (24).

CD phenotype in UK South Asians

Colonic disease was the most common disease locating in the CD South Asian population. In both Northern European and South Asian cohorts, the predominant disease behavior was non-stricturing and nonpenetrating. Despite this, a significantly lower prevalence of penetrating behavior was observed in the South Asian population compared with Northern Europeans (P<0.04); this may, in part, explain why the requirements for surgery are significantly lower in South Asians (P<0.01). The population of South Asians with CD is smaller than the UC cohort and so further large studies would be required to confirm the differences reported.

Potential explanations for differences in disease phenotype

Differences in disease phenotype between ethnic groups are likely to be due to a combination of genetic susceptibility, environmental,

Table 4. Phenotypic details of South Asian and white Northern European CD patients living in North West London, United Kingdom

	South Asian (n=94)	Northern European (n=215)	P value	Corrected P value (P _c)
<i>Disease location (Montreal) (17), n (%)</i>				
Terminal ileum (L1)	15 (16.0)	53 (24.7)	P=0.09	
Colon (L2)	44 (46.8)	75 (34.8)	P=0.048	P _c =0.19
Ileocolon (L3)	35 (37.2)	87 (40.5)	P=0.59	
Upper GI (L4)	0	0	—	
Upper GI modifier	10 (10.6)	16 (7.4)	P=0.35	
<i>Disease behavior (Montreal) (17), n (%)</i>				
Nonstricturing, nonpenetrating (B1)	68 (72.3)	125 (58.1)	P=0.02	P _c =0.08
Stricturing (B2)	22 (23.4)	60 (27.9)	P=0.41	
Penetrating (B3)	4 (4.3)	30 (14.0)	P=0.01	P_c=0.04*
Perianal disease modifier (p)	19 (20.2)	46 (21.4)	P=0.81	
Surgery for CD, n (%)	25 (26.6)	96 (44.7)	P=0.003	P_c=0.01*
<i>EIMs, n (%)</i>				
Peripheral arthropathy	6 (6.4)	27 (12.6)	P=0.11	
Ankylosing spondylitis/sacroiliitis	1 (1.1)	5 (2.8)	P=0.46	
Skin complications ^b	7 (7.4)	18 (8.4)	P=0.78	
Uveitis/episcleritis	3 (3.2)	7 (3.3)	P=0.98	
PSC	1 (1.1)	1 (0.5)	P=0.55	
Venous thrombosis	0	5 (2.3)	P=0.14	

CD, Crohn's disease; EIM, extraintestinal manifestation; GI, gastrointestinal; PSC, primary sclerosing cholangitis.

*P_c: after Bonferroni correction, the values remain significant.

^bErythema nodosum and pyoderma gangrenosum.

All P values were calculated using the χ^2 test.

Bold values are P values that showed significance (P<0.05).

and immune factors. Over 30 susceptibility genes associated with IBD have been identified in the white European population over the past decade (25–27). The *CARD15/NOD2* gene on chromosome 16q (IBD1) has been associated with isolated ileal disease in white European CD patients (28–30). It is of note that the three disease susceptibility variants in the *CARD15/NOD2* gene associated with CD are rare or absent in South Asian IBD populations (12,13). Therefore, other genes or polymorphisms not yet identified may influence susceptibility to certain disease phenotypes in this ethnic group.

Differences between ethnic groups do not necessarily imply a genetic cause but may reflect variations in environmental factors related to lifestyle in different groups. An environmental factor that may be of importance in determining CD phenotype is smoking. This study reported a significantly lower number of smokers in the South Asian CD cohort compared with the Northern European cohort (10.6 vs. 31.6%, $P<0.0001$). This is consistent with previous studies that have shown that there are significantly fewer smokers in South Asian migrants compared with Europeans (16,31). Smoking has been reported to be associated with disease location, with a higher prevalence of ileal disease and a lower prevalence of

colonic involvement in smokers (20,32,33). Previous studies have also demonstrated that smoking is more frequently associated with the development of strictures or fistulae, with an increased risk of surgery (34,35).

Although this study did not show any differences in IBD phenotype between first- and second-generation South Asians, it has been proposed that environmental stimuli may act as a trigger for IBD in genetically susceptible patients when moving from one part of the world to another. The 'IBD hygiene hypothesis' suggests that hygienic environments negatively affect immune development, predisposing to immune-mediated diseases such as IBD (36,37). Reduced exposure to intestinal helminths may be a factor contributing to IBD as they are involved in host immune regulation, preventing excessive immune activation (38), and perhaps better hygiene in the United Kingdom may be responsible for reduced exposure to enteric infections. This may result in immune dysregulation and subsequent excessive inflammation in predisposed individuals. Ethnic variation in the gut microbiota may also be involved in the dysregulation of mucosal immunity—a potential area for research.

Dietary constituents associated with the western diet, such as refined sugars and saturated fats, have also been linked with IBD

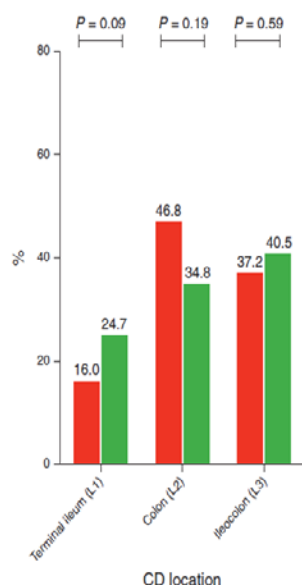


Figure 2. Crohn's disease (CD) location for South Asian and white Northern European patients. *P* values were calculated using the χ^2 -test. A Bonferroni correction was applied to significant *P* values.

(39). Investigators from Leicester showed that South Asian Hindu IBD patients have significantly altered their traditional diet (40). Dietary changes adopted by South Asians may therefore act as a potential environmental stimulus in disease development.

There is a lack of large population-based studies of IBD in South Asia that makes attempts to compare epidemiological data difficult. However, the nonstricturing, nonpenetrating colonic pattern of disease was the dominant phenotype in a recent multicentered Indian study of 182 CD patients, which may suggest that the phenotype of CD remains constant despite migration (41). Two smaller epidemiological studies from India report that between 16 and 20% of UC patients have extensive colitis (4,24). In contrast, the majority of South Asian patients living in North West London with UC had extensive disease (63.0%). The difference could relate to alterations in underlying environmental factors or to a variation in diagnostic assessment. This study based UC disease extent on a confirmed histological diagnosis rather than simply endoscopic or radiological appearances (which was used in the study by Kochhar *et al.* (24)), but for comparison, the current data were reassessed using endoscopic extent only, and extensive colitis was still the predominant phenotype (59%).

This study has shown that UC is more prevalent than CD in a UK South Asian population and that the predominant phenotype of UC in this ethnic group is extensive colitis. This appears to be associated with a low risk of colectomy for active disease, and long-term follow-up of this group will be important to establish whether the higher proportion of extensive colitis translates into

an increased risk of dysplasia/cancer. The phenotypic variations observed in South Asian IBD patients are likely to reflect different hereditary, environmental, and behavioral factors, which may be elucidated through further genetic, microbiological, and immunological studies.

ACKNOWLEDGMENTS

The authors thank the patients for participating in this study. They also thank Santosh Chada RN (St. Mary's Hospital) and Shamima Padaruth RN (Ealing Hospital) for their assistance with retrieving medical notes. Dr Iain Beveridge (West Middlesex University Hospital) and Dr Andrew Milestone (St. Mark's Hospital) assisted with patient recruitment. Professor Derek P. Jewell (University of Oxford) assisted with the critical revision of the document. The IBD database was developed with assistance from Dr Sebastian Zeki (Imperial College London).

CONFLICT OF INTEREST

Guarantor of the article: David G. Walker, MRCP.

Specific author contributions: Study design: David G. Walker, Horace R.T. Williams, Sara E. Marshall, and Timothy R. Orchard; data collection: David G. Walker and Horace R.T. Williams; data analyses: David G. Walker and Timothy Orchard; report drafting: David G. Walker, Horace R.T. Williams, and Timothy R. Orchard. All authors participated in the critical revision of this report and approved the submitted final draft. Patients were under the care of Timothy R. Orchard, Huw J.W. Thomas, Julian P. Teare, Stephen P. Kane, Joel E. Mawdsley, Jayantha Arnold, Ian McNeil, Ailsa L. Hart, Maxton C.L. Pitcher, and Julian R. Walters, who all participated in active recruitment.

Financial support: This study was funded by the generous support of the Broad Medical Research Program (Grant no. IBD-0187) and St. Mary's Paddington Charitable Trustees. DGW, HRTW, and TRO are grateful to the NIHR Biomedical Facility at Imperial College London for infrastructure support.

Potential competing interests: None.

Study Highlights

WHAT IS CURRENT KNOWLEDGE

- ✓ Currently, 2.08 million South Asians live in the United Kingdom.
- ✓ Incidence and prevalence of inflammatory bowel disease (IBD) is increasing in the UK South Asian population.
- ✓ There have been suggestions that the phenotype of IBD may differ between ethnic groups in the United Kingdom, but no large studies have been performed until now.

WHAT IS NEW HERE

- ✓ In UK South Asians, ulcerative colitis (UC) disease extent differs significantly from white Northern European patients.
- ✓ The majority of UK South Asian patients with UC have extensive colitis; nonstricturing, nonpenetrating colonic Crohn's disease (CD) is the most prevalent CD phenotype.
- ✓ UC phenotype does not differ between first- and second-generation UK South Asians.

REFERENCES

- Loftus EV Jr. Clinical epidemiology of inflammatory bowel disease: incidence, prevalence, and environmental influences. *Gastroenterology* 2004;126:1504–17.
- Bhat M, Nguyen GC, Pare P *et al.* Phenotypic and genotypic characteristics of inflammatory bowel disease in French Canadians: comparison with a large North American repository. *Am J Gastroenterol* 2009;104:2233–40.
- Hou JK, El-Serag H, Thirumurthi S. Distribution and manifestations of inflammatory bowel disease in Asians, Hispanics, and African Americans: a systematic review. *Am J Gastroenterol* 2009;104:2100–9.
- Sood A, Midha V, Sood N *et al.* Incidence and prevalence of ulcerative colitis in Punjab, North India. *Gut* 2003;52:1587–90.
- Office for National Statistics 2001. Census. <http://www.statistics.gov.uk/ccinugget.asp?id=455> 2009.
- Probert CS, Jayanthi V, Pinder D *et al.* Epidemiological study of ulcerative proctocolitis in Indian migrants and the indigenous population of Leicestershire. *Gut* 1992;33:687–93.
- Rashid ST, Bharucha S, Jamallulail SI *et al.* Inflammatory bowel disease in the South Asian population of Northwest England. *Am J Gastroenterol* 2008;103:242–3.
- Tsironi F, Feakins RM, Probert CS *et al.* Incidence of inflammatory bowel disease is rising and abdominal tuberculosis is falling in Bangladeshis in East London, United Kingdom. *Am J Gastroenterol* 2004;99:1749–55.
- Jayanthi V, Probert CS, Pinder D *et al.* Epidemiology of Crohn's disease in Indian migrants and the indigenous population in Leicestershire. *Q J Med* 1992;82:125–38.
- Probert CS, Jayanthi V, Hughes AO *et al.* Prevalence and family risk of ulcerative colitis and Crohn's disease: an epidemiological study among Europeans and south Asians in Leicestershire. *Gut* 1993;34:1547–51.
- Economou M, Trikalinos TA, Loizou KT *et al.* Differential effects of NOD2 variants on Crohn's disease risk and phenotype in diverse populations: a metaanalysis. *Am J Gastroenterol* 2004;99:2393–404.
- Juyal G, Amre D, Midha V *et al.* Evidence of allelic heterogeneity for associations between the NOD2/CARD15 gene and ulcerative colitis among North Indians. *Aliment Pharmacol Ther* 2007;26:1325–32.
- Pugazhendhi S, Amte A, Balamurugan R *et al.* Common NOD2 mutations are absent in patients with Crohn's disease in India. *Indian J Gastroenterol* 2008;27:201–3.
- Stange EF, Travis SP, Vermeire S *et al.* European evidence-based consensus on the diagnosis and management of ulcerative colitis: definitions and diagnosis. *J Crohns Colitis* 2008;2:1–23.
- Van Assche G, Dignass A, Panes J *et al.* The second European evidence-based consensus on the diagnosis and management of Crohn's disease: definitions and diagnosis. *J Crohns Colitis* 2010;4:7–27.
- Carr I, Mayberry JE. The effects of migration on ulcerative colitis: a three-year prospective study among Europeans and first- and second-generation South Asians in Leicester (1991–1994). *Am J Gastroenterol* 1999;94:2918–22.
- Silverberg MS, Satsangi J, Ahmad T *et al.* Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: Report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. *Can J Gastroenterol* 2005;19 (Suppl A): 5–36.
- Cosnes J, Carbonnel F, Beaugerie L *et al.* Effects of cigarette smoking on the long-term course of Crohn's disease. *Gastroenterology* 1996;110:424–31.
- Cottone M, Rosselli M, Orlando A *et al.* Smoking habits and recurrence in Crohn's disease. *Gastroenterology* 1994;106:643–8.
- Lakatos PL, Czegledi Z, Szamosi T *et al.* Perianal disease, small bowel disease, smoking, prior steroid or early azathioprine/biological therapy are predictors of disease behavior change in patients with Crohn's disease. *World J Gastroenterol* 2009;15:3504–10.
- Orchard TR, Wordsworth BP, Jewell DP. Peripheral arthropathies in inflammatory bowel disease: their articular distribution and natural history. *Gut* 1998;42:387–91.
- Ritchie JK, Powell-Tuck J, Lennard-Jones JE. Clinical outcome of the first ten years of ulcerative colitis and proctitis. *Lancet* 1978;1:1140–3.
- Lennard-Jones JE. The clinical outcome of ulcerative colitis depends on how much of the colonic mucosa is involved. *Scand J Gastroenterol Suppl* 1983;88:48–53.
- Kochhar R, Mehta SK, Nagi B *et al.* Extraintestinal manifestations of idiopathic ulcerative colitis. *Indian J Gastroenterol* 1991;10:88–9.
- Barrett JC, Hansoul S, Nicolae DL *et al.* Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet* 2008;40:955–62.
- Hampe J, Franke A, Rosenstiel P *et al.* A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. *Nat Genet* 2007;39:207–11.
- The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447:661–78.
- Cuthbert AP, Fisher SA, Mirza MM *et al.* The contribution of NOD2 gene mutations to the risk and site of disease in inflammatory bowel disease. *Gastroenterology* 2002;122:867–74.
- Ahmad T, Armuzzi A, Bunce M *et al.* The molecular classification of the clinical manifestations of Crohn's disease. *Gastroenterology* 2002;122: 854–66.
- Vermeire S, Wild G, Kocher K *et al.* CARD15 genetic variation in a Quebec population: prevalence, genotype-phenotype relationship, and haplotype structure. *Am J Hum Genet* 2002;71:74–83.
- Banerjee S, Biram R, Chataway J *et al.* South Asian strokes: lessons from the St Mary's stroke database. *QJM* 2010;103:17–21.
- Lindberg E, Jarnerot G, Huitfeldt B. Smoking in Crohn's disease: effect on localisation and clinical course. *Gut* 1992;33:779–82.
- Mahid SS, Minor KS, Stevens PL *et al.* The role of smoking in Crohn's disease as defined by clinical variables. *Dig Dis Sci* 2007;52:2897–903.
- Picco ME, Bayless TM. Tobacco consumption and disease duration are associated with fistulizing and stricturing behaviors in the first 8 years of Crohn's disease. *Am J Gastroenterol* 2003;98:363–8.
- Reese GE, Nanidis T, Borysiewicz C *et al.* The effect of smoking after surgery for Crohn's disease: a meta-analysis of observational studies. *Int J Colorectal Dis* 2008;23:1213–21.
- Elliott DE, Urban JF Jr, Argo CK *et al.* Does the failure to acquire helminthic parasites predispose to Crohn's disease? *FASEB J* 2000;14: 1848–55.
- de Silva HJ, de Silva NR, de Silva AP *et al.* Emergence of inflammatory bowel disease 'beyond the West': do prosperity and improved hygiene have a role? *Trans R Soc Trop Med Hyg* 2008;102:857–60.
- Weinstock JV, Elliott DE. Helminths and the IBD hygiene hypothesis. *Inflamm Bowel Dis* 2009;15:128–33.
- Bernstein CN, Shanahan F. Disorders of a modern lifestyle: reconciling the epidemiology of inflammatory bowel diseases. *Gut* 2008;57:1185–91.
- Probert CS, Bhakta P, Bhamra B *et al.* Diet of South Asians with inflammatory bowel disease. *Arq Gastroenterol* 1996;33:132–5.
- Das K, Ghoshal UC, Dhali GK *et al.* Crohn's disease in India: a multicenter study from a country where tuberculosis is endemic. *Dig Dis Sci* 2009;54:1099–107.

8.2 PUBLISHED ABSTRACTS

Walker DG, Bancil AS, Williams HRT, Bunn C, Orchard TR. How helpful are serological markers in differentiating Crohn's disease from ulcerative colitis in Indian Asian inflammatory bowel disease patients. *Gut* 2011; 60 (Supp I) A222.

Walker DG, Bancil AS, Rai PS, Williams HR, Chambers JC, Kooner JS, Orchard TR. Ethnic variation in the frequency of IBD related polymorphisms in IRGM, ATG16L1 and IL23R. *Gut* 2011; 60 (Supp I) A223.

Walker DG, Bancil AS, Sato H, Williams HR, Pantelidis P, Lee K, Chambers JC, Kooner JS, Orchard TR. HLA-DRB1*1502 but not HLA-DRB1*0103 confers susceptibility to ulcerative colitis in Indian Asians. *Gut* 2010; 59 (Supp III) A405

Walker DG, Williams HRT, Pantelidis P, Sato H, Lee K, Milestone AN, Chambers JC, Kooner JS, Orchard TR. The three common *CARD15/NOD2* polymorphisms associated with Crohn's disease are rare or absent in a UK South Asian Crohn's disease population. *Gastroenterology*, 2010: Vol 138 Issue 5, Supp I, S684.

Walker DG, Williams HRT, Bunn C, Orchard TR. The Prevalence and Clinical Significance of Atypical Perinuclear Anti-Neutrophil Cytoplasmic Antibody in UK South Asian Inflammatory Bowel Disease Patients. *Gastroenterology* 2010: Vol. 138, Issue 5, Supplement 1, Page S-527.

Milestone AN, **Walker DG**, Walters JR, Gabe SM, Pitcher MCL, Jacyna MR, Hart A. Vitamin D Status in Crohn's Disease. *Gastroenterology* 2010 Vol. 138, Issue 5, Supplement 1, Page S-202.

Walker DG, Williams HRT, Cox J, Orchard TR. Urinary metabolic profiling suggests differences in the intestinal microbiota between South Asians and Caucasians with ulcerative colitis. *Gut* 2009; Vol 58 Suppl. II A312

Walker DG, Williams HRT, Cox J, Orchard TR: Urinary metabolic profiling in Crohn's disease: is there a difference between ethnic groups? *Gut* 2009; 58 Suppl I A60

Walker DG, Williams HRT, Arnold J, Kane S, Thomas HJW, Teare JP, Jewell DP, Orchard TR. The clinical phenotype of inflammatory bowel disease in the UK South Asian population. *Gut*, Oct 2008: Vol 57, Suppl II A134

CHAPTER 9

APPENDICES

Appendix 1: Reagents required for DNA extraction using Millers Salting out technique	
<i>Red Cell Lysis Buffer</i>	Composition: 0.144 M NH ₄ Cl, 1.00 mM NaHCO ₃ . Preparation: 15.405 g NH ₄ Cl and 1.68 g NaHCO ₃ were added to 2 L deionised water and mixed.
<i>Nuclear Lysis Buffer</i>	Composition: 10 mM Tris-HCl pH 8.2, 0.4 M NaCl, 2 mM Na ₂ EDTA pH 8.0. Preparation: 23.37 g NaCl were dissolved in 900 mL deionised water. 10 mL 1M Tris-HCl pH 8.2 and 10 mL 200 mM Na ₂ EDTA pH 8.0 were added and made up to 1 L with deionised water.
<i>6 M NaCl</i>	Preparation: 350.64 g NaCl were dissolved in 800 mL deionised water, and then topped up to 1L with deionised water.

Appendix 2: Primers used for identification of the three common CD associated CARD15/NOD2 polymorphisms using sequence specific primers (courtesy of Dr Tariq Ahmad).				
NCBI reference	SNP Position	SNP id	Sense/ Anti Sense	Primers (5'-3')
rs2066844	2104C 2104T 2104	SNP 8	[as] [as] [s]	CTG AGA AGG CCC TGC TCC CTG AGA AGG CCC TGC TCT GTC ACC TCT GAT TTG GTG TTG
rs2066845	2722G 2722C 2722	SNP 12	[s] [s] [as]	GTC ACC CAC TCT GTT GCC GTC ACC CAC TCT GTT GCG GTA TTG GGA TTA CAC AAG TGA GC
rs2066847	3020- 3020Ci 3020	SNP 13	[as] [as] [s]	GGT GTC ATT CCT TTC AAG GGC GGT GTC ATT CCT TTC AAG GGG GGA GAA GAG GGC TTG TTC AG

Appendix 3: HLA class II medium resolution primer sequences (courtesy of Dr. Hiroe Sato, derived from Bunce et al. (1995)).																								
No.	Primer Sequence																							
36	5	,	T	T	G	T	G	G	C	A	G	C	T	T	A	A	G	T	T	T	G	A	A	T
37	5	,	C	T	G	C	A	C	T	G	T	G	A	A	G	C	T	C	T	C	A	C		
38	5	,	C	T	G	C	A	C	T	G	T	G	A	A	G	C	T	C	T	C	C	A		
39	5	,	C	C	G	C	C	T	C	T	G	C	T	C	C	A	G	G	A	G				
40	5	,	C	C	C	G	C	T	C	G	T	C	T	T	C	C	A	G	G	A	T			
41	5	,	T	C	C	T	G	T	G	G	C	A	G	C	C	T	A	A	G	A	G			
44	5	,	T	A	C	T	T	C	C	A	T	A	A	C	C	A	G	G	A	G	G	A	G	A
46	5	,	G	A	C	G	G	A	G	C	G	G	G	T	G	C	G	G	T	A				
47	5	,	G	T	T	T	C	T	T	G	G	A	G	C	A	G	G	T	T	A	A	A	C	A
48	5	,	C	C	T	G	T	G	G	C	A	G	G	G	T	A	A	G	T	A	T	A		
49	5	,	C	C	C	G	T	A	G	T	T	G	T	G	T	C	T	G	C	A	C	A	C	
50	5	,	A	G	T	A	C	T	C	T	A	C	G	G	G	T	G	A	G	T	G	T	T	
51	5	,	C	T	G	C	A	G	T	A	G	G	T	G	T	C	C	A	C	C	A	G		
53	5		C	G	G	T	T	G	C	T	G	G	A	A	A	G	A	C	G	C	G			
54	5	,	C	T	G	G	C	T	G	T	T	C	C	A	G	T	A	C	T	C	C	T		
58	5	,	T	C	T	G	C	A	A	T	A	G	G	T	G	T	C	C	A	C	C	T		
61	5	,	G	T	T	T	C	T	T	G	C	A	G	C	A	G	G	A	T	A	A	G	T	A
68	5	,	G	T	T	T	C	T	T	G	G	A	G	T	A	C	T	C	T	A	C	G	T	C
69	5	,	T	T	T	C	T	T	G	G	A	G	C	T	G	C	G	T	A	A	G	T	C	
70	5	,	G	T	T	T	C	T	T	G	G	A	G	C	T	G	C	T	T	A	A	G	T	C
76	5	,	G	G	A	G	T	A	C	C	G	G	G	C	G	G	T	G	A	G				
77	5	,	G	C	T	A	C	T	T	C	A	C	C	A	A	C	G	G	G	A	C	C		
78	5	,	T	G	G	T	A	G	T	T	G	T	G	T	C	T	G	C	A	T	A	C	G	
79	5	,	A	C	G	G	A	G	C	G	C	G	T	G	C	G	G	G	G					
82	5	,	G	T	G	C	G	T	C	T	T	G	T	G	A	G	C	A	G	A	A	G		
102	5	,	T	G	T	T	C	C	A	G	T	A	C	T	C	G	G	C	G	C	T			
104	5	,	C	C	C	G	C	C	T	G	T	C	T	T	C	C	A	G	G	A	A			
107	5	,	C	T	G	T	T	C	C	A	G	T	G	C	T	C	C	G	C	A	G			
111	5	,	C	C	G	C	G	G	A	A	C	G	C	C	A	C	C	T	C					
112	5	,	C	G	T	G	C	G	G	A	G	C	T	C	C	A	A	C	T	G				

151	5	C	C	G	T	A	G	T	T	G	T	G	T	C	T	G	C	A	G	T	A	A
152	5	C	C	C	G	C	G	G	T	A	C	G	C	C	A	C	C	T	C			
181	5	G	A	C	G	G	A	G	C	G	C	G	T	G	C	G	T	T	A			
252	5	C	C	A	C	C	G	C	G	G	C	C	C	G	C	G	C					
255	5	G	T	C	C	A	C	C	C	G	G	C	C	C	C	G	C	T				
256	5	C	T	G	T	T	C	C	A	G	G	A	C	T	C	G	G	C	G	A		
258	5	A	C	C	G	C	G	G	C	C	C	G	C	C	T	G	T	C				
259	5	T	C	C	A	C	C	G	C	G	G	C	C	C	G	C	T	C				
263	5	G	T	T	C	C	T	G	G	A	C	A	G	A	T	A	C	T	T	C	C	
268	5	C	T	G	C	A	G	T	A	A	T	T	G	T	C	C	A	C	C	C	G	
270	5	G	A	T	C	G	T	T	C	G	T	G	T	C	C	C	C	A	C	A	A	
273	5	T	T	C	T	T	G	G	A	G	T	A	C	T	C	T	A	C	G	G	G	
283	5	G	A	T	C	G	T	T	C	G	T	G	T	C	C	C	C	A	C	A	G	
314	5	C	T	G	G	T	A	C	T	C	C	C	C	C	A	G	G	T	C	A		
347	5	T	T	T	C	G	T	G	C	T	C	C	A	G	T	T	T	A	A	G	G	C
348	5	G	A	C	G	T	G	G	G	G	G	T	G	T	A	C	C	G	C			
349	5	G	G	A	G	C	G	C	G	T	G	C	G	T	C	T	T	G	T	A		
350	5	T	G	C	A	C	A	C	C	G	T	G	T	C	C	A	A	C	T	C		
351	5	T	G	C	A	C	A	C	C	C	T	G	T	C	C	A	C	C	G			
353	5	G	T	G	C	G	T	C	T	T	G	T	G	A	C	C	A	G	A	T	A	
485	5	T	C	C	A	C	C	G	C	G	G	C	C	C	G	C	T	T				
491	5	C	T	C	C	G	T	C	A	C	C	G	C	C	C	G	G	T				
492	5	G	C	T	G	T	T	C	C	A	G	T	A	C	T	C	A	G	C	G		
495	5	C	T	C	C	G	T	C	A	C	C	G	C	C	C	G	G	A				
831	5	T	G	G	C	T	G	T	T	C	C	A	G	T	A	C	T	C	G	G	C	G
866	5	C	G	C	C	T	G	G	T	A	C	T	C	C	C	C	A	G				
1465	5	G	G	A	C	G	G	A	G	C	G	C	G	T	G	C	G	T	C	T		
1466	5	G	C	T	G	T	T	C	C	A	G	T	A	C	T	C	G	G	C	G	T	
2489	5	A	T	A	C	T	T	C	C	A	T	A	A	C	C	A	G	G	A	G	G	A
2510	5	G	G	C	T	G	T	T	C	C	A	G	T	A	C	T	C	G	G	C	A	T

Appendix 4: HLA CLASS II medium resolution primer sequences and reaction combination						
Reaction No.	HLA Allelic Specificity	Primer 1	µg/mL	Primer 2	µg/mL	Control
1	DRB1*0101/2/4	36	10	39	10	63/64
2	DRB1*0102,0104	36	10	38	10	63/64
3	DRB1*0103	36	10	40	10	63/64
4	DRB1*1501-1505	41	10	252	2	63/64
5	DRB1*1501,1503-5 (not 1502)	41	10	38	10	63/64
6	DRB1*1601-6	41	15	491	15	63/64
7	DRB1*0301-5,*1107	68	10	255	5	63/64
8	DRB1*0301,0304	46	10	38	10	63/64
9	DRB1*0302,0305,*1302,1305,*1109, 1120,*1402,1403,1409,1413,1419	44	10	37	10	63/64
10	DRB1*0401-0422,*1410,*1122	47	10	37, 38	10	63/64
11	DRB1*1122,*0415	47	10	54	10	63/64
12	DRB1*0701	48	10	49	10	63/64
13	DRB1*0801-0811,*1415	50	10	51, 102	10	63/64
14	DRB1*0901	46	10	49	10	63/64
15	DRB1*1001	53	10	37	10	63/64
16	DRB1*1101-1121(not 1122)*1411	68, 273	10	54	10	63/64
17	DRB1*1109,1116,1120	44	5	54	5	63/64
18	DRB1*1201-1203	50	16	256	16	63/64
19	DRB1*1101-16,1118-22,*1201-3,*1301/2/4/5/6/9/10/11/14/15/16/17/18/20/21/22	68, 47, 273	10	495	10	63/64
20	DRB1*0302,0303,*1117,*1401/2/3/5/6/7/8/9/12/13/14/16/18/19/20	68	10	491	10	63/64
21	DRB1*1301/2/4/8/15/16/17/19/20/22,*1102/3/11/14/16/20/21,*1416	68, 273	10	259	10	63/64
22	DRB1*1301,1302,*1416	263	12	259, 252	12 of 259; 2.5 of 252	63/64
23	DRB1*1301,1302, (1306),1310,1315,1316,*1419,1421,*1116,1121	44	10	40, 485	10	63/64
24	DRB1*1305,1306,1310,*1109,*1419,1421	44	8	258, 485	8	63/64
25	DRB1*1303,1310,*1419,1421	68	10	485	10	63/64
26	DRB1*1303,1304,1312,1321,*1413	68	5	102	5	63/64
27	DRB1*1305,1307,1311,1314,1318,1321,*1101/4/6/9/10/(11)/12/15	68	10	104	10	63/64
28	DRB1*03	68	10	2510	10	63/64
29	DRB1*03,13,14	2489	20	38	20	63/64
30	DRB1*1401/7/16	68	10	107	10	63/64
31	DRB1*1117, 1401/5/7/8/14/18/23/26	68	20	58	20	63/64
32	DRB1*03,1402 etc	2489	10	51, 39	10	63/64
33	DRB5*0101,0102,0201,0202,0203	61	10	492	10	63/64
34	DRB3*0101	69	20	268	20	63/64
35	DRB3*0201,0202,0203	151	10	76	10	63/64
36	DRB3*0201,0301	70	13.2	38	13.2	63/64

37	DRB4*0101101,0102,0103	283	4	314	5	63/64
38	DRB4*0103102N	270	10	866	10	63/64
39	DQB1*0305	77	10	350	10	63/64
40	DQB1*02	82	10	112	10	63/64
41	DQB1*04	77	10	78	10	63/64
42	DQB1*05	79	3	152	3	63/64
43	DQB1*0601/2/3	347, 348	10	111	10	63/64
44	DQB1*0602/3/7/8	348	10	350	10	63/64
45	DQB1*0603-9	349	10	350, 351	10	63/64
46	DQB1*0301,0304	181	10	112	10	63/64
47	DQB1*0302	353	10	831	10	63/64
48	DQB1*03032	1465	6	1466	6	63/64

Appendix 5: Control Primers used for SSP-PCR	
Control primers	Primer (5' – 3')
210	ATG ATG TTG ACC TTT CCA GGG
211	TTC ATC AGT TGC TGC CCC TC
63	TGC CAA CTG GAG CAC CCA A
64	GCA TCT TGC TCT GTG CAG AT

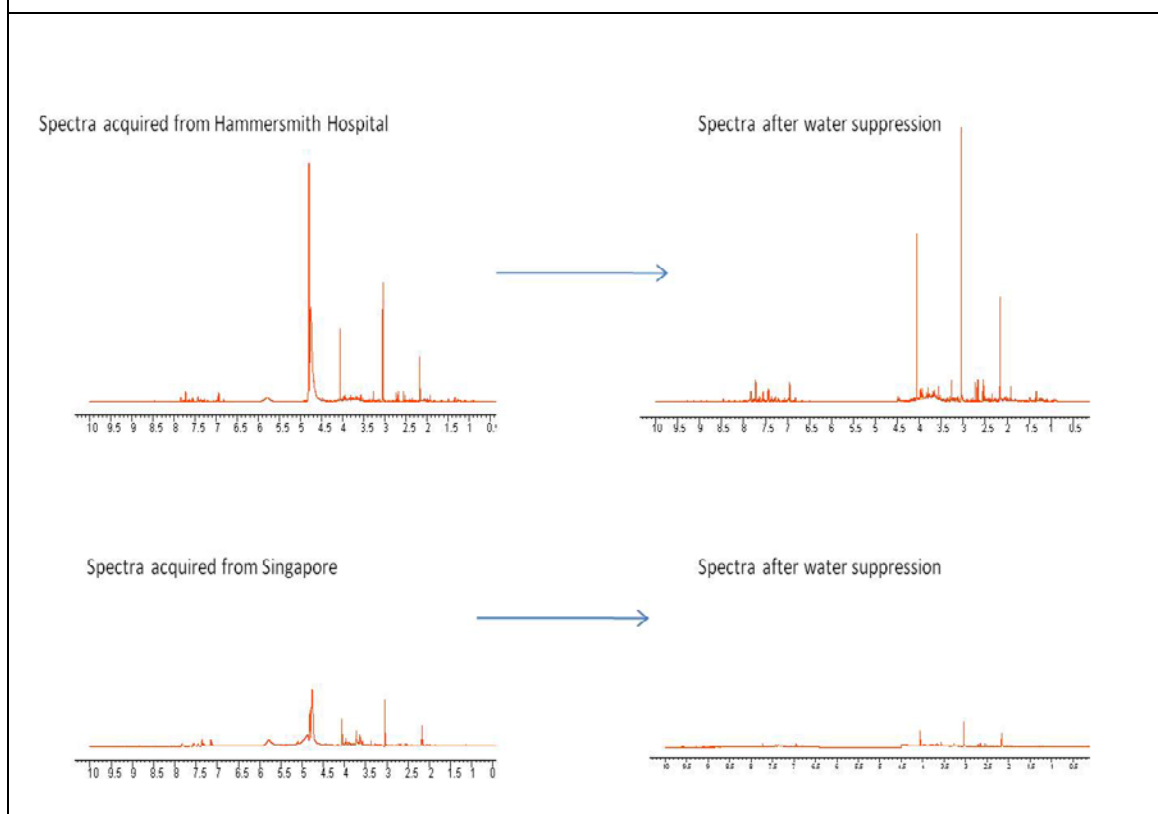
Appendix 6: Primer sequences for SNPs analysed using pyrosequencing				
Gene	NCBI reference	SNP id		Base Sequence
ATG16L1	rs2241880	ATG	F1	CCCCCAGGACAATGTGGAT
			R1	CGAAGACACACAAGGCAGTAGCT
			S1	ACCAGAACCAGGATGA
IL23R	rs11209026	IL23R	F1	TGCTGTTATGTTGTCAATTCTTTC
			S1	TGGGATATTTAACAGATCAT
			R1	GTAAAATAGGCCAGGAAGCTTTTC
IRGM	rs13361189	IRGM189	F1	CGTGTCGTACCCAAGCAGAGT
			R1	CCCTTTCTAAACTGTACCCGCTGA
			S1	CAGAGTGTGCTTGAAAAT
IRGM	rs4958847	IRGM847	F1	ATGGGTGAGGTCATCTTAACAGTC
			R1	AAATCTATTTTAGGGAGTAGTGCA
			S1	ATTCATTGCCCAATAT
F1, Forward primer; R1, Reverse primer; S1, sequence primer				

Appendix 7: Urinary metabonomic questionnaire for inflammatory bowel disease patients

Ealing Hospital  NHS Trust																	
<u>Inflammatory Bowel Disease Questionnaire</u>																	
Name..... Age..... Date.....																	
Ethnicity: Indian ☐ Pakistan ☐ Sri Lanka ☐ Caucasian ☐ Duration in UK.....																	
Weight.....																	
1. CURRENT MEDICATIONS AND DOSES (including Herbal remedies)?																	
2. PAST MEDICAL HISTORY:																	
No: ☐ Yes: ☐																	
3. DIET:																	
	<table border="0" style="width: 100%;"> <tr> <th style="text-align: left;">Usually</th> <th style="text-align: left;">In past 24 hours</th> </tr> <tr> <td>Vegetarian? ☐</td> <td>☐</td> </tr> <tr> <td>Meat (red/white)? ☐</td> <td>☐ Type:.....</td> </tr> <tr> <td>Fish (type)? ☐</td> <td>☐ Type:.....</td> </tr> </table>	Usually	In past 24 hours	Vegetarian? ☐	☐	Meat (red/white)? ☐	☐ Type:.....	Fish (type)? ☐	☐ Type:.....								
Usually	In past 24 hours																
Vegetarian? ☐	☐																
Meat (red/white)? ☐	☐ Type:.....																
Fish (type)? ☐	☐ Type:.....																
<table border="0" style="width: 100%;"> <tr> <th style="text-align: left;">In past 24 hours</th> <th style="text-align: left;">In past 24 hours</th> </tr> <tr> <td>Cheese? ☐</td> <td>Yoghurt? ☐</td> </tr> <tr> <td>Grapefruit? ☐</td> <td>Berries? ☐</td> </tr> <tr> <td>Cherries? ☐</td> <td>Fizzy drinks? ☐</td> </tr> <tr> <td>Liquorice? ☐</td> <td>Coffee? ☐ How many cups:.....</td> </tr> <tr> <td>Walnuts? ☐</td> <td>Tea? ☐ How many cups:.....</td> </tr> <tr> <td>Vanilla? ☐</td> <td>Herbal Tea? ☐ How many cups:.....</td> </tr> <tr> <td>Milk? ☐</td> <td></td> </tr> </table>		In past 24 hours	In past 24 hours	Cheese? ☐	Yoghurt? ☐	Grapefruit? ☐	Berries? ☐	Cherries? ☐	Fizzy drinks? ☐	Liquorice? ☐	Coffee? ☐ How many cups:.....	Walnuts? ☐	Tea? ☐ How many cups:.....	Vanilla? ☐	Herbal Tea? ☐ How many cups:.....	Milk? ☐	
In past 24 hours	In past 24 hours																
Cheese? ☐	Yoghurt? ☐																
Grapefruit? ☐	Berries? ☐																
Cherries? ☐	Fizzy drinks? ☐																
Liquorice? ☐	Coffee? ☐ How many cups:.....																
Walnuts? ☐	Tea? ☐ How many cups:.....																
Vanilla? ☐	Herbal Tea? ☐ How many cups:.....																
Milk? ☐																	
4. LAST FOOD YOU HAD TO EAT TODAY and how many hours ago?																	
5. ALCOHOL:																	
No: ☐ Yes: ☐ Past 24hrs: Past week:.....																	
6. SMOKING:																	
No: ☐ Yes: ☐ Smoking at time of diagnosis:.....																	
7. EXERCISE in past 24 hours:																	
None: ☐ Light: ☐ Medium: ☐ Hard: ☐																	
8. GENERAL HEALTH in last 24 hours:																	
i. General well being: ☐ well ☐ slightly under par ☐ poor ☐ very poor ☐ terrible																	
ii. Bowel frequency during day: ☐ 1-3 ☐ 4-6 ☐ 7-9 ☐ >9																	
iii. Bowel frequency at night: ☐ 0 ☐ 1-3 ☐ 4-6																	
iv. Urgency: ☐ none ☐ hurry ☐ immediately ☐ incontinence																	
v. Blood in stool: ☐ none ☐ trace ☐ occasionally frank ☐ usually frank																	
vi. Complications: ☐ arthritis ☐ iritis ☐ skin rashes ☐ mouth ulcers ☐ anal fissure																	
vii. Abdominal pain: ☐ none ☐ mild ☐ moderate ☐ severe																	
Thank you for your help																	

Appendix 8: The preparation of 0.2M phosphate buffer and TSP/D2O solution	
0.2M phosphate buffer	1. Add together; • 3.8mls of 0.2M monohydrate solution ($[\text{NaH}_2\text{PO}_4]$ (1.39g NaH_2PO_4 mixed with 50mls distilled water) • 16.2mls 0.2M Dibasic Anhydrous solution (Na_2HPO_4) (1.42g Na_2HPO_4 mixed with 50mls distilled water) • 20mls dH_2O 2. Vortex for 10 seconds 3. Check with pH meter solution @pH 7.35
TSP / D2O solution	1. Add 20mg of Sodium Trimethylsilyl-Propionate to 20mls of Deuterated water (deuterium oxide D_2O). 2. Vortex for 10 seconds (the mixture is stored at room temperature in a falcon tube covered with foil).

Appendix 9: A comparison of NMR spectra acquired using the NMR facility at Hammersmith Hospital, Imperial College with spectra acquired from Singapore



CHAPTER 10

REFERENCE LIST

't Hart BA, Vogels JT, Spijksma G, Brok HP, Polman C, van der GJ (2003) 1H-NMR spectroscopy combined with pattern recognition analysis reveals characteristic chemical patterns in urines of MS patients and non-human primates with MS-like disease. *J Neurol Sci* 212:21-30.

Acheson ED (1960) The distribution of ulcerative colitis and regional enteritis in United States veterans with particular reference to the Jewish religion. *Gut* 1:291-293.

Agrawal S, Khan F, Bharadwaj U (2007) Human genetic variation studies and HLA class II loci. *Int J Immunogenet* 34:247-252.

Ahmad T, Armuzzi A, Bunce M, Mulcahy-Hawes K, Marshall SE, Orchard TR, Crawshaw J, Large O, de SA, Cook JT, Barnardo M, Cullen S, Welsh KI, Jewell DP (2002) The molecular classification of the clinical manifestations of Crohn's disease. *Gastroenterology* 122:854-866.

Ahmad T, Armuzzi A, Neville M, Bunce M, Ling KL, Welsh KI, Marshall SE, Jewell DP (2003a) The contribution of human leucocyte antigen complex genes to disease phenotype in ulcerative colitis. *Tissue Antigens* 62:527-535.

Ahmad T, Neville M, Marshall SE, Armuzzi A, Mulcahy-Hawes K, Crawshaw J, Sato H, Ling KL, Barnardo M, Goldthorpe S, Walton R, Bunce M, Jewell DP, Welsh KI (2003b) Haplotype-specific linkage disequilibrium patterns define the genetic topography of the human MHC. *Hum Mol Genet* 12:647-656.

Al-Waiz M, Mikov M, Mitchell SC, Smith RL (1992) The exogenous origin of trimethylamine in the mouse. *Metabolism* 41:135-136.

Annunziato F, Cosmi L, Santarlaschi V, Maggi L, Liotta F, Mazzinghi B, Parente E, Fili L, Ferri S, Frosali F, Giudici F, Romagnani P, Parronchi P, Tonelli F, Maggi E, Romagnani S (2007) Phenotypic and functional features of human Th17 cells. *J Exp Med* 204:1849-1861.

APDW2004 Chinese IBD Working Group (2006a) A retrospective study of 3100 inpatients with ulcerative colitis in China. *Chin J Dig* 26:368-372.

APDW2004 Chinese IBD Working Group (2006b) Retrospective analysis of 515 cases of Crohn's disease hospitalization in China: nationwide study from 1990 to 2003. *J Gastroenterol Hepatol* 21:1009-1015.

Arnott ID, Landers CJ, Nimmo EJ, Drummond HE, Smith BK, Targan SR, Satsangi J (2004) Sero-reactivity to microbial components in Crohn's disease is associated with disease severity and progression, but not NOD2/CARD15 genotype. *Am J Gastroenterol* 99:2376-2384.

Awane M, Andres PG, Li DJ, Reinecker HC (1999) NF-kappa B-inducing kinase is a common mediator of IL-17-, TNF-alpha-, and IL-1 beta-induced chemokine promoter activation in intestinal epithelial cells. *J Immunol* 162:5337-5344.

Ayub Q, Tyler-Smith C (2009) Genetic variation in South Asia: assessing the influences of geography, language and ethnicity for understanding history and disease risk. *Brief Funct Genomic Proteomic* 8:395-404.

Bala L, Nagana Gowda GA, Ghoshal UC, Misra A, Bhandari M, Khetrapal CL (2004) ¹H NMR spectroscopic method for diagnosis of malabsorption syndrome: a pilot study. *NMR Biomed* 17:69-75.

Balding DJ (2006) A tutorial on statistical methods for population association studies. *Nat Rev Genet* 7:781-791.

Banerjee S, Biram R, Chataway J, Ames D (2010) South Asian strokes: lessons from the St Mary's stroke database. *QJM* 103:17-21.

Barrett JC, et al. (2008) Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet* 40:955-962.

Becker C, Wirtz S, Blessing M, Pirhonen J, Strand D, Bechthold O, Frick J, Galle PR, Autenrieth I, Neurath MF (2003) Constitutive p40 promoter activation and IL-23 production in the terminal ileum mediated by dendritic cells. *J Clin Invest* 112:693-706.

Bernstein CN, Shanahan F (2008) Disorders of a modern lifestyle: reconciling the epidemiology of inflammatory bowel diseases. *Gut* 57:1185-1191.

Bernstein CN, Wajda A, Svenson LW, MacKenzie A, Koehoorn M, Jackson M, Fedorak R, Israel D, Blanchard JF (2006) The epidemiology of inflammatory bowel disease in Canada: a population-based study. *Am J Gastroenterol* 101:1559-1568.

Bezabeh T, Somorjai RL, Smith IC (2009) MR metabolomics of fecal extracts: applications in the study of bowel diseases. *Magn Reson Chem* 47 Suppl 1:S54-S61.

Bezabeh T, Somorjai RL, Smith IC, Nikulin AE, Dolenko B, Bernstein CN (2001) The use of ¹H magnetic resonance spectroscopy in inflammatory bowel diseases: distinguishing ulcerative colitis from Crohn's disease. *Am J Gastroenterol* 96:442-448.

Bhan AK, Mizoguchi E, Smith RN, Mizoguchi A (1999) Colitis in transgenic and knockout animals as models of human inflammatory bowel disease. *Immunol Rev* 169:195-207.

Bhat M, et al. (2009) Phenotypic and genotypic characteristics of inflammatory bowel disease in French Canadians: comparison with a large North American repository. *Am J Gastroenterol* 104:2233-2240.

Bjerrum JT, Nielsen OH, Hao F, Tang H, Nicholson JK, Wang Y, Olsen J (2010) Metabonomics in ulcerative colitis: diagnostics, biomarker identification, and insight into the pathophysiology. *J Proteome Res* 9:954-962.

Bjerrum JT, Nielsen OH, Wang YL, Olsen J (2008) Technology insight: metabonomics in gastroenterology-basic principles and potential clinical applications. *Nat Clin Pract Gastroenterol Hepatol* 5:332-343.

Bollard ME, Holmes E, Lindon JC, Mitchell SC, Branstetter D, Zhang W, Nicholson JK (2001) Investigations into biochemical changes due to diurnal variation and estrus cycle in female rats using high-resolution (1)H NMR spectroscopy of urine and pattern recognition. *Anal Biochem* 295:194-202.

Bollard ME, Stanley EG, Lindon JC, Nicholson JK, Holmes E (2005) NMR-based metabonomic approaches for evaluating physiological influences on biofluid composition. *NMR Biomed* 18:143-162.

Bone E, Tamm A, Hill M (1976) The production of urinary phenols by gut bacteria and their possible role in the causation of large bowel cancer. *Am J Clin Nutr* 29:1448-1454.

Bossuyt X (2006) Serologic markers in inflammatory bowel disease. *Clin Chem* 52:171-181.

Brennan A, Katz DR, Nunn JD, Barker S, Hewison M, Fraher LJ, O'Riordan JL (1987) Dendritic cells from human tissues express receptors for the immunoregulatory vitamin D3 metabolite, dihydroxycholecalciferol. *Immunology* 61:457-461.

Brindle JT, Antti H, Holmes E, Tranter G, Nicholson JK, Bethell HW, Clarke S, Schofield PM, McKilligin E, Mosedale DE, Grainger DJ (2002) Rapid and noninvasive diagnosis of the presence and severity of coronary heart disease using 1H-NMR-based metabonomics. *Nat Med* 8:1439-1444.

Brindle JT, Nicholson JK, Schofield PM, Grainger DJ, Holmes E (2003) Application of chemometrics to 1H NMR spectroscopic data to investigate a relationship between human serum metabolic profiles and hypertension. *Analyst* 128:32-36.

Broberger O, Perlmann P (1959) Autoantibodies in human ulcerative colitis. *J Exp Med* 110:657-674.

Bunce M, O'Neill CM, Barnardo MC, Krausa P, Browning MJ, Morris PJ, Welsh KI (1995) Phototyping: comprehensive DNA typing for HLA-A, B, C, DRB1, DRB3,

DRB4, DRB5 & DQB1 by PCR with 144 primer mixes utilizing sequence-specific primers (PCR-SSP). Tissue Antigens 46:355-367.

Carr I, Mayberry JF (1999) The effects of migration on ulcerative colitis: a three-year prospective study among Europeans and first- and second- generation South Asians in Leicester (1991-1994). Am J Gastroenterol 94:2918-2922.

Carter MJ, Lobo AJ, Travis SP (2004) Guidelines for the management of inflammatory bowel disease in adults. Gut 53 Suppl 5:V1-16.

Cassinotti A, Birindelli S, Clerici M, Trabattoni D, Lazzaroni M, Ardizzane S, Colombo R, Rossi E, Bianchi Porro G (2009) HLA and autoimmune digestive disease: A clinically oriented review for gastroenterologists. Am J Gastroenterol 104: 195-217.

Cathcart-Rake W, Porter R, Whittier F, Stein P, Carey M, Grantham J (1975) Effect of diet on serum accumulation and renal excretion of aryl acids and secretory activity in normal and uremic man. Am J Clin Nutr 28:1110-1115.

Chambers JC, Eda S, Bassett P, Karim Y, Thompson SG, Gallimore JR, Pepys MB, Kooner JS (2001) C-reactive protein, insulin resistance, central obesity, and coronary heart disease risk in Indian Asians from the United Kingdom compared with European whites. Circulation 104:145-150.

Chambers JC, Kooner JS (2002) Diabetes, insulin resistance and vascular disease among Indian Asians and Europeans. Semin Vasc Med 2:199-214.

Chaturvedi N, Coady E, Mayet J, Wright AR, Shore AC, Byrd S, McG Thom SA, Kooner JS, Schalkwijk CG, Hughes AD (2007) Indian Asian men have less peripheral arterial disease than European men for equivalent levels of coronary disease. Atherosclerosis 193:204-212.

Chung YL, Rider LG, Bell JD, Summers RM, Zemel LS, Rennebohm RM, Passo MH, Hicks J, Miller FW, Scott DL (2005) Muscle metabolites, detected in urine by proton spectroscopy, correlate with disease damage in juvenile idiopathic inflammatory myopathies. Arthritis Rheum 53:565-570.

Chung YL, Wassif WS, Bell JD, Hurley M, Scott DL (2003) Urinary levels of creatine and other metabolites in the assessment of polymyositis and dermatomyositis. Rheumatology (Oxford) 42:298-303.

Cobbold J (2005) What is NMR? *Submitted for publication*

Cohavy O, Bruckner D, Gordon LK, Misra R, Wei B, Eggena ME, Targan SR, Braun J (2000) Colonic bacteria express an ulcerative colitis pANCA-related protein epitope. Infect Immun 68:1542-1548.

Collazo CM, Yap GS, Sempowski GD, Lusby KC, Tessarollo L, Woude GF, Sher A, Taylor GA (2001) Inactivation of LRG-47 and IRG-47 reveals a family of interferon gamma-inducible genes with essential, pathogen-specific roles in resistance to infection. *J Exp Med* 194:181-188.

Cosnes J, Carbonnel F, Beaugerie L, Le QY, Gendre JP (1996) Effects of cigarette smoking on the long-term course of Crohn's disease. *Gastroenterology* 110:424-431.

Cottone M, Rosselli M, Orlando A, Oliva L, Puleo A, Cappello M, Traina M, Tonelli F, Pagliaro L (1994) Smoking habits and recurrence in Crohn's disease. *Gastroenterology* 106:643-648.

Curtis MM, Way SS (2009) Interleukin-17 in host defence against bacterial, mycobacterial and fungal pathogens. *Immunology* 126:177-185.

Cuthbert AP, Fisher SA, Mirza MM, King K, Hampe J, Croucher PJ, Mascheretti S, Sanderson J, Forbes A, Mansfield J, Schreiber S, Lewis CM, Mathew CG (2002) The contribution of NOD2 gene mutations to the risk and site of disease in inflammatory bowel disease. *Gastroenterology* 122:867-874.

Danze PM, Colombel JF, Jacquot S, Loste MN, Heresbach D, Ategbo S, Khamassi S, Perichon B, Semana G, Charron D, Cezard JP (1996) Association of HLA class II genes with susceptibility to Crohn's disease. *Gut* 39:69-72.

Das K, Ghoshal UC, Dhali GK, Benjamin J, Ahuja V, Makharia GK (2009) Crohn's disease in India: a multicenter study from a country where tuberculosis is endemic. *Dig Dis Sci* 54:1099-1107.

Dassopoulos T, Frangakis C, Cruz-Correa M, Talor MV, Burek CL, Datta L, Nouvet F, Bayless TM, Brant SR (2007) Antibodies to *saccharomyces cerevisiae* in Crohn's disease: higher titers are associated with a greater frequency of mutant NOD2/CARD15 alleles and with a higher probability of complicated disease. *Inflamm Bowel Dis* 13:143-151.

de Silva HJ, de Silva NR, de Silva AP, Jewell DP (2008) Emergence of inflammatory bowel disease 'beyond the West': do prosperity and improved hygiene have a role? *Trans R Soc Trop Med Hyg* 102:857-860.

de Silva PS, Olsen A, Christensen J, Schmidt EB, Overvaad K, Tjonneland A, Hart AR (2010) An association between dietary arachidonic acid, measured in adipose tissue, and ulcerative colitis. *Gastroenterology* 139:1912-1917.

Devlin SM, Yang H, Ippoliti A, Taylor KD, Landers CJ, Su X, Abreu MT, Papadakis KA, Vasilias EA, Melmed GY, Fleshner PR, Mei L, Rotter JI, Targan SR (2007) NOD2 variants and antibody response to microbial antigens in Crohn's disease patients and their unaffected relatives. *Gastroenterology* 132:576-586.

Duerr RH, et al. (2006) A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* 314:1461-1463.

Economou M, Trikalinos TA, Loizou KT, Tsianos EV, Ioannidis JP (2004) Differential effects of NOD2 variants on Crohn's disease risk and phenotype in diverse populations: a metaanalysis. *Am J Gastroenterol* 99:2393-2404.

Elliott DE, Urban JF, Jr., Argo CK, Weinstock JV (2000) Does the failure to acquire helminthic parasites predispose to Crohn's disease? *FASEB J* 14:1848-1855.

Farmer M, Petras RE, Hunt LE, Janosky JE, Galandiuk S (2000) The importance of diagnostic accuracy in colonic inflammatory bowel disease. *Am J Gastroenterol* 95:3184-3188.

Farmer RG, Easley KA, Rankin GB (1993) Clinical patterns, natural history, and progression of ulcerative colitis. A long-term follow-up of 1116 patients. *Dig Dis Sci* 38:1137-1146.

Feng CG, Collazo-Custodio CM, Eckhaus M, Hieny S, Belkaid Y, Elkins K, Jankovic D, Taylor GA, Sher A (2004) Mice deficient in LRG-47 display increased susceptibility to mycobacterial infection associated with the induction of lymphopenia. *J Immunol* 172:1163-1168.

Fernandez L, Mendoza JL, Martinez A, Urcelay E, Fernandez-Arquero M, Garcia-Paredes J, Pena AS, az-Rubio M, de la Concha EG (2004) IBD1 and IBD3 determine location of Crohn's disease in the Spanish population. *Inflamm Bowel Dis* 10:715-722.

Ferrante M, Henckaerts L, Joossens M, Pierik M, Joossens S, Dotan N, Norman GL, Altstock RT, Van SK, Rutgeerts P, Van AG, Vermeire S (2007) New serological markers in inflammatory bowel disease are associated with complicated disease behaviour. *Gut* 56:1394-1403.

Forcione DG, Rosen MJ, Kisiel JB, Sands BE (2004) Anti-Saccharomyces cerevisiae antibody (ASCA) positivity is associated with increased risk for early surgery in Crohn's disease. *Gut* 53:1117-1122.

Franke A, Balschun T, Karlsen TH, Hedderich J, May S, Lu T, Schuldt D, Nikolaus S, Rosenstiel P, Krawczak M, Schreiber S (2008) Replication of signals from recent studies of Crohn's disease identifies previously unknown disease loci for ulcerative colitis. *Nat Genet* 40:713-715.

Freeman HJ (2007) Application of the Montreal classification for Crohn's disease to a single clinician database of 1015 patients. *Can J Gastroenterol* 21:363-366.

Fujimoto T, Kato J, Nasu J, Kuriyama M, Okada H, Yamamoto H, Mizuno M, Shiratori Y (2007) Change of clinical characteristics of ulcerative colitis in Japan: analysis of 844 hospital-based patients from 1981 to 2000. *Eur J Gastroenterol Hepatol* 19:229-235.

Fujino S, Andoh A, Bamba S, Ogawa A, Hata K, Araki Y, Bamba T, Fujiyama Y (2003) Increased expression of interleukin 17 in inflammatory bowel disease. *Gut* 52:65-70.

Fuss IJ, Becker C, Yang Z, Groden C, Hornung RL, Heller F, Neurath MF, Strober W, Mannon PJ (2006) Both IL-12p70 and IL-23 are synthesized during active Crohn's disease and are down-regulated by treatment with anti-IL-12 p40 monoclonal antibody. *Inflamm Bowel Dis* 12:9-15.

Futami S, Aoyama N, Honsako Y, Tamura T, Morimoto S, Nakashima T, Ohmoto A, Okano H, Miyamoto M, Inaba H, . (1995) HLA-DRB1*1502 allele, subtype of DR15, is associated with susceptibility to ulcerative colitis and its progression. *Dig Dis Sci* 40:814-818.

Ghoshal UC, Ghoshal U, Singh H, Tiwari S (2007) Anti-Saccharomyces cerevisiae antibody is not useful to differentiate between Crohn's disease and intestinal tuberculosis in India. *J Postgrad Med* 53:166-170.

Gibson DT (1967) Microbial degradation of aromatic compounds. *Science* 161:1093-1097.

Glas J, et al. (2007) rs1004819 is the main disease-associated IL23R variant in German Crohn's disease patients: combined analysis of IL23R, CARD15, and OCTN1/2 variants. *PLoS One* 2:e819.

Halfvarson J, Bodin L, Tysk C, Lindberg E, Jarnerot G (2003) Inflammatory bowel disease in a Swedish twin cohort: a long-term follow-up of concordance and clinical characteristics. *Gastroenterology* 124:1767-1773.

Hampe J, et al. (2007) A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. *Nat Genet* 39:207-211.

Hampe J, Shaw SH, Saiz R, Leysens N, Lantermann A, Mascheretti S, Lynch NJ, MacPherson AJ, Bridger S, van DS, Stokkers P, Morin P, Mirza MM, Forbes A, Lennard-Jones JE, Mathew CG, Curran ME, Schreiber S (1999) Linkage of inflammatory bowel disease to human chromosome 6p. *Am J Hum Genet* 65:1647-1655.

Hancock L, Beckly J, Geremia A, Cooney R, Cummings F, Pathan S, Guo C, Warren BF, Mortensen N, Ahmad T, Jewell D (2008) Clinical and molecular characteristics of isolated colonic Crohn's disease. *Inflamm Bowel Dis* 14:1667-1677.

Harrington LE, Hatton RD, Mangan PR, Turner H, Murphy TL, Murphy KM, Weaver CT (2005) Interleukin 17-producing CD4⁺ effector T cells develop via a lineage distinct from the T helper type 1 and 2 lineages. *Nat Immunol* 6:1123-1132.

Hilmi I, Singh R, Ganesanathan S, Yatim I, Radzi M, Chua AB, Tan HJ, Huang S, Chin KS, Menon J, Goh KL (2009) Demography and clinical course of ulcerative colitis in a multiracial Asian population: a nationwide study from Malaysia. *J Dig Dis* 10:15-20.

Hiwatashi N, Yamazaki H, Kimura M, Morimoto T, Watanabe H, Toyota T (1991) Clinical course and long-term prognosis of Japanese patients with ulcerative colitis. *Gastroenterol Jpn* 26:312-318.

Holmes E, Loo RL, Stamler J, Bictash M, Yap IK, Chan Q, Ebbels T, De IM, Brown IJ, Veselkov KA, Daviglus ML, Kesteloot H, Ueshima H, Zhao L, Nicholson JK, Elliott P (2008) Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature* 453:396-400.

Hou JK, El-Serag H, Thirumurthi S (2009) Distribution and manifestations of inflammatory bowel disease in Asians, Hispanics, and African Americans: a systematic review. *Am J Gastroenterol* 104:2100-2109.

Hugot JP, Chamaillard M, Zouali H, Lesage S, Cezard JP, Belaiche J, Almer S, Tysk C, O'Morain CA, Gassull M, Binder V, Finkel Y, Cortot A, Modigliani R, Laurent-Puig P, Gower-Rousseau C, Macry J, Colombel JF, Sahbatou M, Thomas G (2001) Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 411:599-603.

Hugot JP, Laurent-Puig P, Gower-Rousseau C, Olson JM, Lee JC, Beaugerie L, Naom I, Dupas JL, Van GA, Orholm M, Bonaiti-Pellie C, Weissenbach J, Mathew CG, Lennard-Jones JE, Cortot A, Colombel JF, Thomas G (1996) Mapping of a susceptibility locus for Crohn's disease on chromosome 16. *Nature* 379:821-823.

Hugot JP, Zaccaria I, Cavanaugh J, Yang H, Vermeire S, Lappalainen M, Schreiber S, Annese V, Jewell DP, Fowler EV, Brant SR, Silverberg MS, Cho J, Rioux JD, Satsangi J, Parkes M (2007) Prevalence of CARD15/NOD2 mutations in Caucasian healthy people. *Am J Gastroenterol* 102:1259-1267.

Inoue N, Tamura K, Kinouchi Y, Fukuda Y, Takahashi S, Ogura Y, Inohara N, Nunez G, Kishi Y, Koike Y, Shimosegawa T, Shimoyama T, Hibi T (2002) Lack of common NOD2 variants in Japanese patients with Crohn's disease. *Gastroenterology* 123:86-91.

Jayanthi V, Probert CS, Pinder D, Wicks AC, Mayberry JF (1992a) Epidemiology of Crohn's disease in Indian migrants and the indigenous population in Leicestershire. *Q J Med* 82:125-138.

Jayanthi V, Probert CS, Pollock DJ, Baithun SI, Rampton DS, Mayberry JF (1992b) Low incidence of ulcerative colitis and proctitis in Bangladeshi migrants in Britain. *Digestion* 52:34-42.

Jiang L, Xia B, Li J, Ye M, Yan W, Deng C, Ding Y, Luo H, Hou W, Zhao Q, Liu N, Ren H, Hou X, Xu H (2006) Retrospective survey of 452 patients with inflammatory bowel disease in Wuhan city, central China. *Inflamm Bowel Dis* 12:212-217.

Jiang XL, Cui HF (2002) An analysis of 10218 ulcerative colitis cases in China. *World J Gastroenterol* 8:158-161.

Joossens S, Daperno M, Shums Z, Van SK, Goeken JA, Trapani C, Norman GL, Godefridis G, Claessens G, Pera A, Pierik M, Vermeire S, Rutgeerts P, Bossuyt X (2004) Interassay and interobserver variability in the detection of anti-neutrophil cytoplasmic antibodies in patients with ulcerative colitis. *Clin Chem* 50:1422-1425.

Juyal G, Amre D, Midha V, Sood A, Seidman E, Thelma BK (2007) Evidence of allelic heterogeneity for associations between the NOD2/CARD15 gene and ulcerative colitis among North Indians. *Aliment Pharmacol Ther* 26:1325-1332.

Juyal G, Prasad P, Senapati S, Midha V, Sood A, Amre D, Juyal RC, Thelma BK (2011) An investigation of genome-wide studies reported susceptibility loci for ulcerative colitis shows limited replication in north Indians. *PLoS One* 6:e16565.

Kennedy RJ, Hoper M, Deodhar K, Erwin PJ, Kirk SJ, Gardiner KR (2000) Interleukin 10-deficient colitis: new similarities to human inflammatory bowel disease. *Br J Surg* 87:1346-1351.

Khosla SN, Girdhar NK, Lal S, Mishra DS (1986) Epidemiology of ulcerative colitis in hospital and select general population of northern India. *J Assoc Physicians India* 34:405-407.

Klebl FH, Bataille F, Berteau CR, Herfarth H, Hofstadter F, Scholmerich J, Rogler G (2003) Association of perinuclear antineutrophil cytoplasmic antibodies and anti-*Saccharomyces cerevisiae* antibodies with Vienna classification subtypes of Crohn's disease. *Inflamm Bowel Dis* 9:302-307.

Kleessen B, Kroesen AJ, Buhr HJ, Blaut M (2002) Mucosal and invading bacteria in patients with inflammatory bowel disease compared with controls. *Scand J Gastroenterol* 37:1034-1041.

Kobayashi T, Okamoto S, Hisamatsu T, Kamada N, Chinen H, Saito R, Kitazume MT, Nakazawa A, Sugita A, Koganei K, Isobe K, Hibi T (2008) IL23 differentially regulates the Th1/Th17 balance in ulcerative colitis and Crohn's disease. *Gut* 57:1682-1689.

Kochhar R, Mehta SK, Nagi B, Bhatia V, Goenka MK, Malik AK (1991) Extraintestinal manifestations of idiopathic ulcerative colitis. *Indian J Gastroenterol* 10:88-89.

Kolls JK, Linden A (2004) Interleukin-17 family members and inflammation. *Immunity* 21:467-476.

Kugathasan S, Loizides A, Babusukumar U, McGuire E, Wang T, Hooper P, Nebel J, Kofman G, Noel R, Broeckel U, Tolia V (2005) Comparative phenotypic and CARD15 mutational analysis among African American, Hispanic, and White children with Crohn's disease. *Inflamm Bowel Dis* 11:631-638.

Laan M, Lotvall J, Chung KF, Linden A (2001) IL-17-induced cytokine release in human bronchial epithelial cells in vitro: role of mitogen-activated protein (MAP) kinases. *Br J Pharmacol* 133:200-206.

Lakatos PL, Czegledi Z, Szamosi T, Banai J, David G, Zsigmond F, Pandur T, Erdelyi Z, Gemela O, Papp J, Lakatos L (2009) Perianal disease, small bowel disease, smoking, prior steroid or early azathioprine/biological therapy are predictors of disease behavior change in patients with Crohn's disease. *World J Gastroenterol* 15:3504-3510.

Lamers RJ, van Nesselrooij JH, Kraus VB, Jordan JM, Renner JB, Dragomir AD, Luta G, van der GJ, DeGroot J (2005) Identification of an urinary metabolite profile associated with osteoarthritis. *Osteoarthritis Cartilage* 13:762-768.

Landers CJ, Cohavy O, Misra R, Yang H, Lin YC, Braun J, Targan SR (2002) Selected loss of tolerance evidenced by Crohn's disease-associated immune responses to auto- and microbial antigens. *Gastroenterology* 123:689-699.

Latiano A, Palmieri O, Cucchiara S, Castro M, D'Inca R, Guariso G, Dallapiccola B, Valvano MR, Latiano T, Andriulli A, Annese V (2009) Polymorphism of the IRGM gene might predispose to fistulizing behavior in Crohn's disease. *Am J Gastroenterol* 104:110-116.

Lee GH, Kim CG, Kim JS, Jung HC, Song IS (2005) [Frequency analysis of NOD2 gene mutations in Korean patients with Crohn's disease]. *Korean J Gastroenterol* 45:162-168.

Lee KS, Medline A, Shockey S (1979) Indeterminate colitis in the spectrum of inflammatory bowel disease. *Arch Pathol Lab Med* 103:173-176.

Lee YM, Fock K, See SJ, Ng TM, Khor C, Teo EK (2000) Racial differences in the prevalence of ulcerative colitis and Crohn's disease in Singapore. *J Gastroenterol Hepatol* 15:622-625.

Leijonmarck CE, Persson PG, Hellers G (1990) Factors affecting colectomy rate in ulcerative colitis: an epidemiologic study. *Gut* 31:329-333.

Lennard-Jones JE (1983) The clinical outcome of ulcerative colitis depends on how much of the colonic mucosa is involved. *Scand J Gastroenterol Supp* 48-53.

Lenz EM, Bright J, Wilson ID, Hughes A, Morrisson J, Lindberg H, Lockton A (2004) Metabonomics, dietary influences and cultural differences: a ¹H NMR-based study of urine samples obtained from healthy British and Swedish subjects. *J Pharm Biomed Anal* 36:841-849.

Lenz EM, Bright J, Wilson ID, Morgan SR, Nash AF (2003) A ¹H NMR-based metabonomic study of urine and plasma samples obtained from healthy human subjects. *J Pharm Biomed Anal* 33:1103-1115.

Leong RW, Armuzzi A, Ahmad T, Wong ML, Tse P, Jewell DP, Sung JJ (2003) NOD2/CARD15 gene polymorphisms and Crohn's disease in the Chinese population. *Aliment Pharmacol Ther* 17:1465-1470.

Leong RW, Lau JY, Sung JJ (2004) The epidemiology and phenotype of Crohn's disease in the Chinese population. *Inflamm Bowel Dis* 10:646-651.

Lesage S, Zouali H, Cezard JP, Colombel JF, Belaiche J, Almer S, Tysk C, O'Morain C, Gassull M, Binder V, Finkel Y, Modigliani R, Gower-Rousseau C, Macry J, Merlin F, Chamaillard M, Jannot AS, Thomas G, Hugot JP (2002) CARD15/NOD2 mutational analysis and genotype-phenotype correlation in 612 patients with inflammatory bowel disease. *Am J Hum Genet* 70:845-857.

Li M, et al. (2008) Symbiotic gut microbes modulate human metabolic phenotypes. *Proc Natl Acad Sci U S A* 105:2117-2122.

Lindberg E, Jarnerot G, Huitfeldt B (1992) Smoking in Crohn's disease: effect on localisation and clinical course. *Gut* 33:779-782.

Ling KL, Ooi CJ, Luman W, Cheong WK, Choen FS, Ng HS (2002) Clinical characteristics of ulcerative colitis in Singapore, a multiracial city-state. *J Clin Gastroenterol* 35:144-148.

Lodes MJ, Cong Y, Elson CO, Mohamath R, Landers CJ, Targan SR, Fort M, Hershberg RM (2004) Bacterial flagellin is a dominant antigen in Crohn disease. *J Clin Invest* 113:1296-1306.

Loftus EV, Jr. (2004) Clinical epidemiology of inflammatory bowel disease: Incidence, prevalence, and environmental influences. *Gastroenterology* 126:1504-1517.

Loftus EV, Jr., Sandborn WJ (2002) Epidemiology of inflammatory bowel disease. *Gastroenterol Clin North Am* 31:1-20.

Lok KH, Hung HG, Ng CH, Kwong KC, Yip WM, Lau SF, Li KK, Li KF, Szeto ML (2008) Epidemiology and clinical characteristics of ulcerative colitis in Chinese population: experience from a single center in Hong Kong. *J Gastroenterol Hepatol* 23:406-410.

Maeda S, Hsu LC, Liu H, Bankston LA, Iimura M, Kagnoff MF, Eckmann L, Karin M (2005) Nod2 mutation in Crohn's disease potentiates NF-kappaB activity and IL-1beta processing. *Science* 307:734-738.

Magro F, Portela F, Lago P, Ramos de DJ, Vieira A, Peixe P, Cremers I, Cotter J, Cravo M, Tavares L, Reis J, Goncalves R, Lopes H, Caldeira P, Ministro P, Carvalho L, Azevedo L, da Costa-Pereira A (2009) Crohn's disease in a southern European country: Montreal classification and clinical activity. *Inflamm Bowel Dis* 15:1343-1350.

Mahid SS, Minor KS, Stevens PL, Galandiuk S (2007) The role of smoking in Crohn's disease as defined by clinical variables. *Dig Dis Sci* 52:2897-2903.

Main J, McKenzie H, Yeaman GR, Kerr MA, Robson D, Pennington CR, Parratt D (1988) Antibody to *Saccharomyces cerevisiae* (bakers' yeast) in Crohn's disease. *BMJ* 297:1105-1106.

Makharia GK, Sachdev V, Gupta R, Lal S, Pandey RM (2007) Anti-*Saccharomyces cerevisiae* antibody does not differentiate between Crohn's disease and intestinal tuberculosis. *Dig Dis Sci* 52:33-39.

Marchesi JR, Holmes E, Khan F, Kochhar S, Scanlan P, Shanahan F, Wilson ID, Wang Y (2007) Rapid and noninvasive metabonomic characterization of inflammatory bowel disease. *J Proteome Res* 6:546-551.

Masuda H, Nakamura Y, Tanaka T, Hayakawa S (1994) Distinct relationship between HLA-DR genes and intractability of ulcerative colitis. *Am J Gastroenterol* 89:1957-1962.

Matake H, Okabe N, Naito S, Yao T (1992) An HLA study on 149 Japanese patients with Crohn's disease. *Gastroenterol Jpn* 27:496-501.

Mate-Jimenez J, Munoz S, Vicent D, Pajares JM (1994) Incidence and prevalence of ulcerative colitis and Crohn's disease in urban and rural areas of Spain from 1981 to 1988. *J Clin Gastroenterol* 18:27-31.

Mathew CG (2008) New links to the pathogenesis of Crohn disease provided by genome-wide association scans. *Nat Rev Genet* 9:9-14.

McCarroll SA, Huett A, Kuballa P, Chilewski SD, Landry A, Goyette P, Zody MC, Hall JL, Brant SR, Cho JH, Duerr RH, Silverberg MS, Taylor KD, Rioux JD, Altshuler D, Daly MJ, Xavier RJ (2008) Deletion polymorphism upstream of IRGM associated with altered IRGM expression and Crohn's disease. *Nat Genet* 40:1107-1112.

McGovern DP, et al. (2010) Genome-wide association identifies multiple ulcerative colitis susceptibility loci. *Nat Genet* 42:332-337.

McKenzie SJ, Baker MS, Buffinton GD, Doe WF (1996) Evidence of oxidant-induced injury to epithelial cells during inflammatory bowel disease. *J Clin Invest* 98:136-141.

Mehra NK, Bouwens AG, Naipal A, Rajalingam R, Grubic Z, Taneja V, Tilanus MG, Giphart MJ (1994) Asian Indian HLA-DR2-, DR4-, and DR52-related DR-DQ genotypes analyzed by polymerase chain reaction based nonradioactive oligonucleotide typing. Unique haplotypes and a novel DR4 subtype. *Hum Immunol* 39:202-210.

Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 16:1215.

Morita N, Toki S, Hirohashi T, Minoda T, Ogawa K, Kono S, Tamakoshi A, Ohno Y, Sawada T, Muto T (1995) Incidence and prevalence of inflammatory bowel disease in Japan: nationwide epidemiological survey during the year 1991. *J Gastroenterol* 30 Suppl 8:1-4.

Mow WS, Vasilias EA, Lin YC, Fleshner PR, Papadakis KA, Taylor KD, Landers CJ, breu-Martin MT, Rotter JI, Yang H, Targan SR (2004) Association of antibody responses to microbial antigens and complications of small bowel Crohn's disease. *Gastroenterology* 126:414-424.

Munz C (2006) Autophagy and antigen presentation. *Cell Microbiol* 8:891-898.

Murdoch TB, Fu H, MacFarlane S, Sydora BC, Fedorak RN, Slupsky CM (2008) Urinary metabolic profiles of inflammatory bowel disease in interleukin-10 gene-deficient mice. *Anal Chem* 80:5524-5531.

Myung SJ, Yang SK, Jung HY, Chang HS, Park B, Hong WS, Kim JH, Min I (2002) HLA-DRB1*1502 confers susceptibility to ulcerative colitis, but is negatively associated with its intractability: a Korean study. *Int J Colorectal Dis* 17:233-237.

Nakagawa I, Amano A, Mizushima N, Yamamoto A, Yamaguchi H, Kamimoto T, Nara A, Funao J, Nakata M, Tsuda K, Hamada S, Yoshimori T (2004) Autophagy defends cells against invading group A Streptococcus. *Science* 306:1037-1040.

Nakajima A, Matsuhashi N, Kodama T, Yazaki Y, Takazoe M, Kimura A (1995) HLA-linked susceptibility and resistance genes in Crohn's disease. *Gastroenterology* 109:1462-1467.

Newman B, Silverberg MS, Gu X, Zhang Q, Lazaro A, Steinhart AH, Greenberg GR, Griffiths AM, McLeod RS, Cohen Z, Fernandez-Vina M, Amos CI, Siminovitch K (2004) CARD15 and HLA DRB1 alleles influence susceptibility and disease localization in Crohn's disease. *Am J Gastroenterol* 99:306-315.

Nicholls AW, Mortishire-Smith RJ, Nicholson JK (2003) NMR spectroscopic-based metabonomic studies of urinary metabolite variation in acclimatizing germ-free rats. *Chem Res Toxicol* 16:1395-1404.

Nicholson JK, Lindon JC (2008) Systems biology: Metabonomics. *Nature* 455:1054-1056.

Niriella MA, de Silva AP, Dayaratne AH, Ariyasinghe MH, Navarathne MM, Peiris RS, Samarasekara DN, Satharasinghe RL, Rajindrajith S, Dassanayake AS, Wickramasinghe AR, de Silva HJ (2010) Prevalence of inflammatory bowel disease in two districts of Sri Lanka: a hospital based survey. *BMC Gastroenterol* 10:32.

Office for National Statistics 2001 (2004) Census.

Ogura Y, Bonen DK, Inohara N, Nicolae DL, Chen FF, Ramos R, Britton H, Moran T, Karaliuskas R, Duerr RH, Achkar JP, Brant SR, Bayless TM, Kirschner BS, Hanauer SB, Nunez G, Cho JH (2001) A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* 411:603-606.

Orchard TR, Chua CN, Ahmad T, Cheng H, Welsh KI, Jewell DP (2002) Uveitis and erythema nodosum in inflammatory bowel disease: clinical features and the role of HLA genes. *Gastroenterology* 123:714-718.

Orchard TR, Thiyagaraja S, Welsh KI, Wordsworth BP, Hill Gaston JS, Jewell DP (2000) Clinical phenotype is related to HLA genotype in the peripheral arthropathies of inflammatory bowel disease. *Gastroenterology* 118:274-278.

Orchard TR, Wordsworth BP, Jewell DP (1998) Peripheral arthropathies in inflammatory bowel disease: their articular distribution and natural history. *Gut* 42:387-391.

Oriuchi T, Hiwatashi N, Kinouchi Y, Takahashi S, Takagi S, Negoro K, Shimosegawa T (2003) Clinical course and longterm prognosis of Japanese patients with Crohn's disease: predictive factors, rates of operation, and mortality. *J Gastroenterol* 38:942-953.

Oshitani N, Hato F, Suzuki K, Sawa Y, Matsumoto T, Maeda K, Higuchi K, Kitagawa S, Arakawa T (2003) Cross-reactivity of yeast antigens in human colon and peripheral leukocytes. *J Pathol* 199:361-367.

Pai CG, Khandige GK (2000) Is Crohn's disease rare in India? *Indian J Gastroenterol* 19:17-20.

Palomino-Morales RJ, Oliver J, Gomez-Garcia M, Lopez-Nevot MA, Rodrigo L, Nieto A, Alizadeh BZ, Martin J (2009) Association of ATG16L1 and IRGM genes polymorphisms with inflammatory bowel disease: a meta-analysis approach. *Genes Immun* 10:356-364.

Papp M, Altorjay I, Dotan N, Palatka K, Foldi I, Tumpek J, Sipka S, Udvardy M, Dinya T, Lakatos L, Kovacs A, Molnar T, Tulassay Z, Miheller P, Norman GL, Szamosi T, Papp J, Lakatos PL (2008) New serological markers for inflammatory bowel disease are associated with earlier age at onset, complicated disease behavior, risk for surgery, and NOD2/CARD15 genotype in a Hungarian IBD cohort. *Am J Gastroenterol* 103:665-681.

Parkes M, et al. (2007) Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nat Genet* 39:830-832.

Peeters M, Joossens S, Vermeire S, Vlietinck R, Bossuyt X, Rutgeerts P (2001) Diagnostic value of anti-Saccharomyces cerevisiae and antineutrophil cytoplasmic autoantibodies in inflammatory bowel disease. *Am J Gastroenterol* 96:730-734.

Peeters M, Nevens H, Baert F, Hiele M, de Meyer AM, Vlietinck R, Rutgeerts P (1996) Familial aggregation in Crohn's disease: increased age-adjusted risk and concordance in clinical characteristics. *Gastroenterology* 111:597-603.

Pene J, Chevalier S, Preisser L, Venereau E, Guilleux MH, Ghannam S, Moles JP, Danger Y, Ravon E, Lesaux S, Yssel H, Gascan H (2008) Chronically inflamed human tissues are infiltrated by highly differentiated Th17 lymphocytes. *J Immunol* 180:7423-7430.

Piancatelli D, Canossi A, Aureli A, Oumhani K, Del BT, Di RM, Liberatore G, Tessitore A, Witter K, El AR, Adorno D (2004) Human leukocyte antigen-A, -B, and -Cw polymorphism in a Berber population from North Morocco using sequence-based typing. *Tissue Antigens* 63:158-172.

Picco MF, Bayless TM (2003) Tobacco consumption and disease duration are associated with fistulizing and stricturing behaviors in the first 8 years of Crohn's disease. *Am J Gastroenterol* 98:363-368.

Pokharna RK, Kabra PK, Sharma R, Kochar DK (2004) Extraintestinal manifestations of idiopathic ulcerative colitis in northwestern India. *Indian J Gastroenterol* 23:89-90.

Portela F, Magro F, Lago P, Cotter J, Cremers I, de DJ, Vieira A, Lopes H, Caldeira P, Barros L, Reis J, Carvalho L, Goncalves R, Campos MJ, Ministro P, Duarte MA, Amil J, Rodrigues S, Azevedo L, Costa-Pereira A (2010) Ulcerative colitis in a Southern European country: a national perspective. *Inflamm Bowel Dis* 16:822-829.

Probert CS, Bhakta P, Bhamra B, Jayanthi V, Mayberry JF (1996a) Diet of South Asians with inflammatory bowel disease. *Arq Gastroenterol* 33:132-135.

Probert CS, Jayanthi V, Hughes AO, Thompson JR, Wicks AC, Mayberry JF (1993) Prevalence and family risk of ulcerative colitis and Crohn's disease: an epidemiological study among Europeans and south Asians in Leicestershire. *Gut* 34:1547-1551.

Probert CS, Jayanthi V, Pinder D, Wicks AC, Mayberry JF (1992) Epidemiological study of ulcerative proctocolitis in Indian migrants and the indigenous population of Leicestershire. *Gut* 33:687-693.

Probert CS, Jayanthi V, Rampton DS, Mayberry JF (1996b) Epidemiology of inflammatory bowel disease in different ethnic and religious groups: limitations and aetiological clues. *Int J Colorectal Dis* 11:25-28.

Pugazhendhi S, Amte A, Balamurugan R, Subramanian V, Ramakrishna BS (2008) Common NOD2 mutations are absent in patients with Crohn's disease in India. *Indian J Gastroenterol* 27:201-203.

Quinton JF, Sendid B, Reumaux D, Duthilleul P, Cortot A, Grandbastien B, Charrier G, Targan SR, Colombel JF, Poulain D (1998) Anti-Saccharomyces cerevisiae mannan antibodies combined with antineutrophil cytoplasmic autoantibodies in inflammatory bowel disease: prevalence and diagnostic role. *Gut* 42:788-791.

Radtke AL, O'Riordan MX (2006) Intracellular innate resistance to bacterial pathogens. *Cell Microbiol* 8:1720-1729.

Raman M, Milestone AN, Walters JR, Hart AL, Ghosh S (2011) Vitamin D and gastrointestinal diseases: inflammatory bowel disease and colorectal cancer. *Therap Adv Gastroenterol* 4:49-62.

Rashid ST, Bharucha S, Jamallulail SI, Banait GS, Kemp K, Makin A, Newman WG (2008) Inflammatory bowel disease in the South Asian population of Northwest England. *Am J Gastroenterol* 103:242-243.

Rath HC, Schultz M, Freitag R, Dieleman LA, Li F, Linde HJ, Scholmerich J, Sartor RB (2001) Different subsets of enteric bacteria induce and perpetuate experimental colitis in rats and mice. *Infect Immun* 69:2277-2285.

Reese GE, Nanidis T, Borysiewicz C, Yamamoto T, Orchard T, Tekkis PP (2008) The effect of smoking after surgery for Crohn's disease: a meta-analysis of observational studies. *Int J Colorectal Dis* 23:1213-1221.

Rioux JD, et al. (2001) Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. *Nat Genet* 29:223-228.

Rioux JD, et al. (2007) Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nat Genet* 39:596-604.

Ritchie JK, Powell-Tuck J, Lennard-Jones JE (1978) Clinical outcome of the first ten years of ulcerative colitis and proctitis. *Lancet* 1:1140-1143.

Ronaghi M (2001) Pyrosequencing sheds light on DNA sequencing. *Genome Res* 11:3-11.

Ronaghi M, Karamohamed S, Pettersson B, Uhlen M, Nyren P (1996) Real-time DNA sequencing using detection of pyrophosphate release. *Anal Biochem* 242:84-89.

Rubin GP, Hungin AP, Kelly PJ, Ling J (2000) Inflammatory bowel disease: epidemiology and management in an English general practice population. *Aliment Pharmacol Ther* 14:1553-1559.

Ruffatti A, Sinico RA, Radice A, Ossi E, Cozzi F, Tonello M, Grypiotis P, Todesco S (2002) Autoantibodies to proteinase 3 and myeloperoxidase in systemic sclerosis. *J Rheumatol* 29:918-923.

Rump JA, Scholmerich J, Gross V, Roth M, Helfesrieder R, Rautmann A, Ludemann J, Gross WL, Peter HH (1990) A new type of perinuclear anti-neutrophil cytoplasmic antibody (p-ANCA) in active ulcerative colitis but not in Crohn's disease. *Immunobiology* 181:406-413.

Sandborn WJ, Landers CJ, Tremaine WJ, Targan SR (1996) Association of antineutrophil cytoplasmic antibodies with resistance to treatment of left-sided ulcerative colitis: results of a pilot study. *Mayo Clin Proc* 71:431-436.

Sandborn WJ, Loftus EV, Jr., Colombel JF, Fleming KA, Seibold F, Homburger HA, Sendid B, Chapman RW, Tremaine WJ, Kaul DK, Wallace J, Harmsen WS, Zinsmeister AR, Targan SR (2001) Evaluation of serologic disease markers in a population-based cohort of patients with ulcerative colitis and Crohn's disease. *Inflamm Bowel Dis* 7:192-201.

Sartor RB (2008) Microbial influences in inflammatory bowel diseases. *Gastroenterology* 134:577-594.

Sartor RB (2006) Mechanisms of disease: pathogenesis of Crohn's disease and ulcerative colitis. *Nat Clin Pract Gastroenterol Hepatol* 3:390-407.

Sasieni PD (1997) From genotypes to genes: doubling the sample size. *Biometrics* 53:1253-1261.

Satsangi J, Parkes M, Louis E, Hashimoto L, Kato N, Welsh K, Terwilliger JD, Lathrop GM, Bell JI, Jewell DP (1996a) Two stage genome-wide search in inflammatory bowel disease provides evidence for susceptibility loci on chromosomes 3, 7 and 12. *Nat Genet* 14:199-202.

Satsangi J, Welsh KI, Bunce M, Julier C, Farrant JM, Bell JI, Jewell DP (1996b) Contribution of genes of the major histocompatibility complex to susceptibility and disease phenotype in inflammatory bowel disease. *Lancet* 347:1212-1217.

Saxon A, Shanahan F, Landers C, Ganz T, Targan S (1990) A distinct subset of antineutrophil cytoplasmic antibodies is associated with inflammatory bowel disease. *J Allergy Clin Immunol* 86:202-210.

Schachter D, Taggart JV (1953) Benzoyl coenzyme A and hippurate synthesis. *JBiol Chem* 925-934.

Schicho R, Nazyrova A, Shaykhutdinov R, Duggan G, Vogel HJ, Storr M (2010) Quantitative metabolomic profiling of serum and urine in DSS-induced ulcerative colitis of mice by (1)H NMR spectroscopy. *J Proteome Res* 9:6265-6273.

Seki SS, Sugimura K, Ota M, Matsuzawa J, Katsuyama Y, Ishizuka K, Mochizuki T, Suzuki K, Yoneyama O, Mizuki N, Honma T, Inoko H, Asakura H (2001) Stratification analysis of MICA triplet repeat polymorphisms and HLA antigens associated with ulcerative colitis in Japanese. *Tissue Antigens* 58:71-76.

Sendid B, Colombel JF, Jacquinet PM, Faille C, Fruit J, Cortot A, Lucidarme D, Camus D, Poulain D (1996) Specific antibody response to oligomannosidic epitopes in Crohn's disease. *Clin Diagn Lab Immunol* 3:219-226.

Shih DQ, Targan SR (2009) Insights into IBD Pathogenesis. *Curr Gastroenterol Rep* 11:473-480.

Sieber R, Bütikofer U, Bosset JO (1995) Benzoic acid as a natural product in cultured dairy products and cheese. *Int Dairy Journal* 227-246.

Silverberg MS, et al. (2009) Ulcerative colitis-risk loci on chromosomes 1p36 and 12q15 found by genome-wide association study. *Nat Genet* 41:216-220.

Silverberg MS, Mirea L, Bull SB, Murphy JE, Steinhart AH, Greenberg GR, McLeod RS, Cohen Z, Wade JA, Siminovitch KA (2003) A population- and family-based study of Canadian families reveals association of HLA DRB1*0103 with colonic involvement in inflammatory bowel disease. *Inflamm Bowel Dis* 9:1-9.

Silverberg MS, et al. (2005) Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: Report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. *Can J Gastroenterol* 19 Suppl A:5-36.

Slupsky CM, Rankin KN, Fu H, Chang D, Rowe BH, Charles PG, McGeer A, Low D, Long R, Kunimoto D, Sawyer MB, Fedorak RN, Adamko DJ, Saude EJ, Shah SL, Marrie TJ (2009) Pneumococcal pneumonia: potential for diagnosis through a urinary metabolic profile. *J Proteome Res* 8:5550-5558.

Sokol H, Lay C, Seksik P, Tannock GW (2008) Analysis of bacterial bowel communities of IBD patients: what has it revealed? *Inflamm Bowel Dis* 14:858-867.

Sokol H, Seksik P, Rigottier-Gois L, Lay C, Lepage P, Podglajen I, Marteau P, Dore J (2006) Specificities of the fecal microbiota in inflammatory bowel disease. *Inflamm Bowel Dis* 12:106-111.

Sollid LM, Markussen G, Ek J, Gjerde H, Vartdal F, Thorsby E (1989) Evidence for a primary association of celiac disease to a particular HLA-DQ alpha/beta heterodimer. *J Exp Med* 169:345-350.

Sood A, Midha V, Sood N, Bhatia AS, Avasthi G (2003) Incidence and prevalence of ulcerative colitis in Punjab, North India. *Gut* 52:1587-1590.

Stange EF, Travis SP, Vermeire S, Reinisch W, Geboes K, Barakauskiene A, Feakins R, Flejou JF, Herfarth H, Hommes DW, Kupcinskis L, Lakatos PL, Mantzaris GJ, Schreiber S, Villanacci V, Warren BF (2008) European evidence-based Consensus on the diagnosis and management of ulcerative colitis: Definitions and diagnosis. *J Crohns Colitis* 2:1-23.

Stella C, Beckwith-Hall B, Cloarec O, Holmes E, Lindon JC, Powell J, van der OF, Bingham S, Cross AJ, Nicholson JK (2006) Susceptibility of human metabolic phenotypes to dietary modulation. *J Proteome Res* 5:2780-2788.

Stokkers PC, Reitsma PH, Tytgat GN, Van Deventer SJ (1999) HLA-DR and -DQ phenotypes in inflammatory bowel disease: a meta-analysis. *Gut* 45:395-401.

Sun J, Schnackenberg LK, Beger RD (2009) Studies of acetaminophen and metabolites in urine and their correlations with toxicity using metabolomics. *Drug Metab Lett* 3:130-136.

Sydora BC, Martin SM, Lupicki M, Dieleman LA, Doyle J, Walker JW, Fedorak RN (2006) Bacterial antigens alone can influence intestinal barrier integrity, but live bacteria are required for initiation of intestinal inflammation and injury. *Inflamm Bowel Dis* 12:429-436.

Tan YM, Goh KL (2005) Ulcerative colitis in a multiracial Asian country: racial differences and clinical presentation among Malaysian patients. *World J Gastroenterol* 11:5859-5862.

Terjung B, Spengler U, Sauerbruch T, Worman HJ (2000) "Atypical p-ANCA" in IBD and hepatobiliary disorders react with a 50-kilodalton nuclear envelope protein of neutrophils and myeloid cell lines. *Gastroenterology* 119:310-322.

Thakur S, Ranjan P, Ghoshal UC, Muller-Myhsok B, Khan F, Talwar S, Agarwal S (2003) Association of human leucocyte DR and DQ antigens in Crohn's disease in Asian Indians: a family study. *Trop Gastroenterol* 24:185-188.

The 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.

Thompson AI, Lees CW (2011) Genetics of ulcerative colitis. *Inflamm Bowel Dis* 17:831-848.

Thompson NP, Driscoll R, Pounder RE, Wakefield AJ (1996) Genetics versus environment in inflammatory bowel disease: results of a British twin study. *BMJ* 312:95-96.

Tighe MR, Hall MA, Barbado M, Cardi E, Welsh KI, Ciclitira PJ (1992) HLA class II alleles associated with celiac disease susceptibility in a southern European population. *Tissue Antigens* 40:90-97.

Tillin T, Dhutia H, Chambers J, Malik I, Coady E, Mayet J, Wright AR, Kooner J, Shore A, Thom S, Chaturvedi N, Hughes A (2007) South Asian men have different patterns of coronary artery disease when compared with European men. *Int J Cardiol*.

Trachtenberg EA, Yang H, Hayes E, Vinson M, Lin C, Targan SR, Tyan D, Erlich H, Rotter JI (2000) HLA class II haplotype associations with inflammatory bowel disease in Jewish (Ashkenazi) and non-Jewish caucasian populations. *Hum Immunol* 61:326-333.

Tremelling M, et al. (2007) IL23R variation determines susceptibility but not disease phenotype in inflammatory bowel disease. *Gastroenterology* 132:1657-1664.

Tsironi E, Feakins RM, Probert CS, Rampton DS, Phil D (2004) Incidence of inflammatory bowel disease is rising and abdominal tuberculosis is falling in Bangladeshis in East London, United Kingdom. *Am J Gastroenterol* 99:1749-1755.

van Heel DA, Fisher SA, Kirby A, Daly MJ, Rioux JD, Lewis CM (2004) Inflammatory bowel disease susceptibility loci defined by genome scan meta-analysis of 1952 affected relative pairs. *Hum Mol Genet* 13:763-770.

Van AG, Dignass A, Panes J, Beaugerie L, Karagiannis J, Allez M, Ochsenkuhn T, Orchard T, Rogler G, Louis E, Kupcinskis L, Mantzaris G, Travis S, Stange E (2010) The second European evidence-based Consensus on the diagnosis and management of Crohn's disease: Definitions and diagnosis. *J Crohns Colitis* 4:7-27.

Vasiliauskas EA, Kam LY, Karp LC, Gaiennie J, Yang H, Targan SR (2000) Marker antibody expression stratifies Crohn's disease into immunologically homogeneous subgroups with distinct clinical characteristics. *Gut* 47:487-496.

Vasiliauskas EA, Plevy SE, Landers CJ, Binder SW, Ferguson DM, Yang H, Rotter JJ, Vidrich A, Targan SR (1996) Perinuclear antineutrophil cytoplasmic antibodies in patients with Crohn's disease define a clinical subgroup. *Gastroenterology* 110:1810-1819.

Veluswamy H, Suryawala K, Sheth A, Wells S, Salvatierra E, Cromer W, Chaitanya GV, Painter A, Patel M, Manas K, Zwank E, Boktor M, Baig K, Datti B, Mathis MJ, Minagar A, Jordan PA, Alexander JS (2010) African-American inflammatory bowel disease in a Southern U.S. health center. *BMC Gastroenterol* 10:104.

Vermeire S, Joossens S, Peeters M, Monsuur F, Marien G, Bossuyt X, Groenen P, Vlietinck R, Rutgeerts P (2001) Comparative study of ASCA (Anti-Saccharomyces cerevisiae antibody) assays in inflammatory bowel disease. *Gastroenterology* 120:827-833.

Vermeire S, Wild G, Kocher K, Cousineau J, Dufresne L, Bitton A, Langelier D, Pare P, Lapointe G, Cohen A, Daly MJ, Rioux JD (2002) CARD15 genetic variation in a Quebec population: prevalence, genotype-phenotype relationship, and haplotype structure. *Am J Hum Genet* 71:74-83.

Vind I, Riis L, Jess T, Knudsen E, Pedersen N, Elkjaer M, Bak A, I, Wewer V, Norregaard P, Moesgaard F, Bendtsen F, Munkholm P (2006) Increasing incidences of inflammatory bowel disease and decreasing surgery rates in Copenhagen City and County, 2003-2005: a population-based study from the Danish Crohn colitis database. *Am J Gastroenterol* 101:1274-1282.

Vucelic B, Korac B, Sentic M, Milicic D, Hadzic N, Juresa V, Bozikov J, Rotkvic I, Buljevac M, Kovacevic I, . (1991) Epidemiology of Crohn's disease in Zagreb, Yugoslavia: a ten-year prospective study. *Int J Epidemiol* 20:216-220.

Walsh MC, Brennan L, Malthouse JP, Roche HM, Gibney MJ (2006) Effect of acute dietary standardization on the urinary, plasma, and salivary metabolomic profiles of healthy humans. *Am J Clin Nutr* 84:531-539.

Wang TT, Dabbas B, Laperriere D, Bitton AJ, Soualhine H, Tavera-Mendoza LE, Dionne S, Servant MJ, Bitton A, Seidman EG, Mader S, Behr MA, White JH (2010) Direct and indirect induction by 1,25-dihydroxyvitamin D3 of the NOD2/CARD15-defensin beta2 innate immune pathway defective in Crohn disease. *J Biol Chem* 285:2227-2231.

Wang Y, Tang H, Nicholson JK, Hylands PJ, Sampson J, Holmes E (2005) A metabonomic strategy for the detection of the metabolic effects of chamomile (*Matricaria recutita* L.) ingestion. *J Agric Food Chem* 53:191-196.

Weinstock JV, Elliott DE (2009) Helminths and the IBD hygiene hypothesis. *Inflamm Bowel Dis* 15:128-133.

Wikoff WR, Anfora AT, Liu J, Schultz PG, Lesley SA, Peters EC, Siuzdak G (2009) Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proc Natl Acad Sci U S A* 106:3698-3703.

Williams HR, Cox IJ, Walker DG, Cobbold JF, Taylor-Robinson SD, Marshall SE, Orchard TR (2010) Differences in gut microbial metabolism are responsible for reduced hippurate synthesis in Crohn's disease. *BMC Gastroenterol* 10:108.

Williams HR, Cox IJ, Walker DG, North BV, Patel VM, Marshall SE, Jewell DP, Ghosh S, Thomas HJ, Teare JP, Jakobovits S, Zeki S, Welsh KI, Taylor-Robinson SD, Orchard TR (2009) Characterization of inflammatory bowel disease with urinary metabolic profiling. *Am J Gastroenterol* 104:1435-1444.

Williams HRT, Sato H, Welsh KI, Teare JP, Thomas HJW, Marshall SE, Jewell DP, Orchard TR (2007) Clinical and genetic associations of pyoderma gangrenosum in patients with inflammatory bowel disease: evidence for a shared predisposition to extraintestinal manifestations. *Gut* 56 (Supplement III):A21.

Williams RE, Eyton-Jones HW, Farnworth MJ, Gallagher R, Provan WM (2002) Effect of intestinal microflora on the urinary metabolic profile of rats: a (1)H-nuclear magnetic resonance spectroscopy study. *Xenobiotica* 32:783-794.

Willing B, Halfvarson J, Dicksved J, Rosenquist M, Jarnerot G, Engstrand L, Tysk C, Jansson JK (2009) Twin studies reveal specific imbalances in the mucosa-associated microbiota of patients with ileal Crohn's disease. *Inflamm Bowel Dis* 15:653-660.

Xu J, Yang S, Cai S, Dong J, Li X, Chen Z (2010) Identification of biochemical changes in lactovegetarian urine using 1H NMR spectroscopy and pattern recognition. *Anal Bioanal Chem* 396:1451-1463.

Yamazaki K, Onouchi Y, Takazoe M, Kubo M, Nakamura Y, Hata A (2007) Association analysis of genetic variants in IL23R, ATG16L1 and 5p13.1 loci with Crohn's disease in Japanese patients. *J Hum Genet* 52:575-583.

Yamazaki K, Takahashi A, Takazoe M, Kubo M, Onouchi Y, Fujino A, Kamatani N, Nakamura Y, Hata A (2009) Positive association of genetic variants in the upstream region of NKX2-3 with Crohn's disease in Japanese patients. *Gut* 58:228-232.

Yang H, McElree C, Roth MP, Shanahan F, Targan SR, Rotter JI (1993) Familial empirical risks for inflammatory bowel disease: differences between Jews and non-Jews. *Gut* 34:517-524.

Yang H, Plevy SE, Taylor K, Tyan D, Fischel-Ghodsian N, McElree C, Targan SR, Rotter JI (1999) Linkage of Crohn's disease to the major histocompatibility complex region is detected by multiple non-parametric analyses. *Gut* 44:519-526.

Yang SK, Hong WS, Min YI, Kim HY, Yoo JY, Rhee PL, Rhee JC, Chang DK, Song IS, Jung SA, Park EB, Yoo HM, Lee DK, Kim YK (2000) Incidence and prevalence of ulcerative colitis in the Songpa-Kangdong District, Seoul, Korea, 1986-1997. *J Gastroenterol Hepatol* 15:1037-1042.

Yang SK, Loftus EV, Jr., Sandborn WJ (2001) Epidemiology of inflammatory bowel disease in Asia. *Inflamm Bowel Dis* 7:260-270.

Yang SK, Park M, Lim J, Park SH, Ye BD, Lee I, Song K (2009) Contribution of IL23R but not ATG16L1 to Crohn's disease susceptibility in Koreans. *Inflamm Bowel Dis* 15:1385-1390.

Yang SK, Yun S, Kim JH, Park JY, Kim HY, Kim YH, Chang DK, Kim JS, Song IS, Park JB, Park ER, Kim KJ, Moon G, Yang SH (2008) Epidemiology of inflammatory bowel disease in the Songpa-Kangdong district, Seoul, Korea, 1986-2005: a KASID study. *Inflamm Bowel Dis* 14:542-549.

Yoshitake S, Kimura A, Okada M, Yao T, Sasazuki T (1999) HLA class II alleles in Japanese patients with inflammatory bowel disease. *Tissue Antigens* 53:350-358.

Zuppi C, Messana I, Forni F, Ferrari F, Rossi C, Giardina B (1998) Influence of feeding on metabolite excretion evidenced by urine ¹H NMR spectral profiles: a comparison between subjects living in Rome and subjects living at arctic latitudes (Svaldbard). *Clin Chim Acta* 278:75-79.