# Identification of novel genetic determinants in the high prevalence early-onset inflammatory bowel disease population in Scotland

**Johan Emiel Van Limbergen**

PhD thesis presented to The University of Edinburgh – 2010

# Contents

# Dedication

For their enduring love, support and understanding this thesis is dedicated to

*Leticia Balcarse, the love of my life*

*Lucas Van Limbergen, our dear son*

*Suzanne Van Nuffel, my aunt and guide*

*Lucie Van Nuffel(†), my mother and inspiration*

*'Everyone knows the difficulty of things that are exquisite and well done – so to have facility in such things gives rise to the greatest wonder. In order to be worthy, exceptional virtuosity should be conveyed with an unaffected, effortless dignity. Sprezzatura is the art of making something difficult look easy.'*

*Baldesar Castiglione in 'Il Libro del Cortegiano' (The Book of the Courtier, 1528)*

*'Our imagination is stretched to the utmost, not, as in fiction, to imagine things which are not really there, but just to comprehend those things which are there.'*

*Richard Feynman in 'The Character of Physical Law' (1965)*

# List of Publications arising from this thesis

Permission from publishers and co-authors has been obtained for publications included in this thesis. (All papers included on CD attached to thesis)

Articles

1. Genetics of the innate immune response in inflammatory bowel disease. Van Limbergen J, Russell RK, Nimmo ER, Ho GT, Arnott IDR, Wilson DC, Satsangi J. Inflammatory Bowel Diseases 2007; 13(3):338-55.

2. Contribution of the NOD1/CARD4 insertion/deletion polymorphism +32656 to inflammatory bowel disease in Northern Europe. Van Limbergen J, Russell RK, Nimmo ER, Törkvist L, Lees CW, Drummond HE, Smith L, Anderson NH, Gillett PM, McGrogan P, Hassan K, Weaver LT, Bisset WM, Mahdi G, Arnott ID, Sjöqvist U, Lördal M, Farrington SM, Dunlop MG, Wilson DC, Satsangi J. Inflammatory Bowel Diseases 2007;13(7):882-9

3. IL23R Arg381Gln is associated with childhood onset inflammatory bowel disease in Scotland. Van Limbergen J, Russell RK, Nimmo ER, Drummond HE, Smith L, Davies G, Anderson NH, Gillett PM, McGrogan P, Hassan K, Weaver L, Bisset WM, Mahdi G, Wilson DC, Satsangi J. Gut 2007;56(8): 1173-4.

4. Investigation of NOD1/CARD4 variation in Inflammatory Bowel Disease using a haplotype-tagging strategy. Van Limbergen J, Nimmo ER, Russell RK, Drummond HE, Smith L, Anderson NH, Davies G, Arnott ID, Wilson DC, Satsangi J. Human Molecular Genetics 2007;16(18):2175-86.

5. The Genetics of Inflammatory Bowel Disease. Van Limbergen J, Russell RK, Nimmo ER, Satsangi J. Am J Gastroenterol 2007;102(12):2820-32.

6. The autophagy gene ATG16L1 influences susceptibility and disease location but not childhood onset in Crohn’s disease in Northern Europe. Van

Limbergen J, Russell RK, Nimmo ER, Drummond HE, Smith L, Anderson NH, Davies G, Gillett PM, McGrogan P, Weaver LT, Bisset WM, Mahdi G, Arnott ID, Wilson DC, Satsangi J. Inflammatory Bowel Diseases 2008; 14(3):338-46.

7. Definition of phenotypic characteristics of childhood-onset inflammatory bowel disease. Van Limbergen J, Russell RK, Drummond HE, Aldhous MC, Round NK, Nimmo ER, Smith L, Gillett PM, McGrogan P, Weaver LT, Bisset WM, Mahdi G, Arnott ID, Satsangi J, Wilson DC. Gastroenterology 2008;135(4):1114-22.

8. Detailed assessment of NOD2/CARD15 exonic variation in inflammatory bowel disease in Scotland: implications for disease pathogenesis. Russell RK, Drummond HE, Wilson DC, Anderson NH, Arnott IDR, Van Limbergen JE, Satsangi J, Nimmo ER. Genes & Immunity 2008;9(6):556-60.

9. Faecal Calprotectin complements routine laboratory investigations in Diagnosing Childhood Inflammatory Bowel Disease. Quail MA, Russell RK, Van Limbergen JE, Rogers P, Drummond HE, Wilson DC, Gillett PM. Inflammatory Bowel Diseases 2009;15(5):756-9.

10. Germline variants of IRGM in childhood-onset Crohn's disease. Van Limbergen J, Russell RK, Nimmo ER, Drummond HE, Wilson DC, Satsangi J. Gut 2009;58(4):610-1.

11. Diverse genome-wide association studies associate the IL12/IL23 pathway with Crohn Disease. Wang K, Zhang H, Kugathasan S, Annese V, Bradfield JP, Russell RK, Sleiman PM, Imielinski M, Glessner J, Hou C, Wilson DC, Walters T, Kim C, Frackelton EC, Lionetti P, Barabino A, Van Limbergen J, Guthery S, Denson L, Piccoli D, Li M, Dubinsky M, Silverberg M, Griffiths A, Grant SF, Satsangi J, Baldassano R, Hakonarson H. American Journal of Human Genetics 2009;84:399-405.

12. Van Limbergen JE, Stevens C, Nimmo ER, Wilson DC, Satsangi J. Autophagy: from basic science to clinical application. Mucosal Immunology 2009;2(4):315-330.

13. Filaggrin loss-of-function variants are associated with atopic co-morbidity in paediatric inflammatory bowel disease. Van Limbergen J, Russell RK, Nimmo E, Zhao YW, Liao HH, Drummond HE, Smith L, Anderson NH, Davies G, Gillett PM, McGrogan P, Weaver L, Bisset MW, Mahdi G, Wilson DC, McLean I, Satsangi J. Inflammatory Bowel Diseases 2009;15(10):1492-1498.

14. Reply to letter of the Editor. Van Limbergen J, Wilson DC, Russell RK, Drummond HE, Satsangi J. Gastroenterology 2009;136(7):2409-10.

15. The genetics of Crohn's disease. Van Limbergen J, Wilson DC, Satsangi J. Annual Review of Genomics and Human Genetics 2009;10:89-116.

16. Detailed haplotype-tagging study of MUC19 in Inflammatory Bowel Disease. Phillips A, Nimmo ER, Van Limbergen J, Drummond HE, Smith L, Satsangi. Inflammatory Bowel Diseases 2009 DOI 10.1002/ibd.21074

17. Common variants at five new loci associated with early-onset inflammatory bowel disease. Imielinski M, Baldassano RN, Griffiths A, Russell RK, Annese V, Dubinsky M, Kugathasan S, Bradfield JP, Walters TD, Sleiman P, Kim CE, Muise A, Wang K, Glessner JT, Saeed S, Zhang H, Frackelton EC, Hou C, Flory JH, Otieno G, Chiavacci RM, Grundmeier R, Castro M, Latiano A, Dallapiccola B, Stempak J, Abrams DJ, Taylor K, McGovern D; Western Regional Research Alliance for Pediatric IBD; International IBD Genetics Consortium, Heyman MB, Ferry GD, Kirschner B, Lee J, Essers J, Grand R, Stephens M, Levine A, Piccoli D, Van Limbergen J, Cucchiara S, Monos DS, Guthery SL, Denson L, Wilson DC, Grant SF, Daly M, Silverberg MS, Satsangi J, Hakonarson H, Silber G, Wrobel I, Quiros A, Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, Rioux JD, Brant SR, Barmada MM, Bitton A, Dassopoulos T, Datta LW, Green T, Kistner EO, Murtha MT,

Regueiro MD, Rotter JI, Schumm LP, Steinhart AH, Targan SR, Xavier RJ; NIDDK IBD Genetics Consortium, Libioulle C, Sandor C, Lathrop M, Belaiche J, Dewit O, Gut I, Heath S, Laukens D, Mni M, Rutgeerts P, Van Gossum A, Zelenika D, Franchimont D, Hugot JP, de Vos M, Vermeire S, Louis E; Belgian-French IBD consortium; Wellcome Trust Case Control Consortium, Cardon LR, Anderson CA, Drummond H, Nimmo E, Ahmad T, Prescott NJ, Onnie CM, Fisher SA, Marchini J, Ghori J, Bumpstead S, Gwillam R, Tremelling M, Deloukas P, Mansfield J, Jewell D, Mathew CG, Parkes M, Georges M. Nat Genet 2009;41(12):1335-40.

# List of published abstracts

1. Tumour Necrosis Factor α Promoter Polymorphisms: Influence on phenotype and severity in childhood IBD and regression analysis of the relative contribution to Crohn's disease. Van Limbergen J, Russell RK, Nimmo ER, Drummond HE, Anderson N, Smith L, Wilson DC, Gillett PM, McGrogan P, Hassan K, Weaver LT, Bisset WM, Mahdi G, Satsangi J.
   Gut 2006;55 (Suppl II):A10(034).
   Gastroenterology 2006;130(4):A3.

2. NOD1/CARD4 insertion/deletion polymorphism +32656 and IBD: influence on susceptibility and disease phenotype in the Scottish childhood onset inflammatory bowel disease (IBD) population and regression analysis of the relative contribution to Crohn's disease. Van Limbergen J, Russell RK, Nimmo ER, Drummond HE, Anderson NH, Smith L, Wilson DC, Gillett PM, McGrogan P, Hassan K, Weaver LT, Bisset WM, Mahdi G, Satsangi J. Gastroenterology 2006;130(4):A3.
   Arch Dis Child 2006;91(Suppl I):A12(G15).
   JPGN 2006;42(5):E107-E108.
   JPGN 2006;43 Suppl 2:S11.

3. Association of a complex insertion/deletion polymorphism of NOD1/CARD4 with susceptibility to inflammatory bowel disease in the Scottish population. Van Limbergen J, Lees CW, Nimmo ER, Russell RK, Drummond HE, Wilson DC, Arnott ID, Satsangi J.
   Gastroenterology 2006;130(4):A64.

4. Analysis of CCL20 variants in IBD provides further evidence for genetic heterogeneity in disease susceptibility. Lees CW, Nimmo ER, Russell RK, Van Limbergen J, Smith A, Drummond HE, Satsangi J.
   Gut 2006;55(Suppl. II):A1.

5. Germline variation of NOD1/CARD4 does not determine susceptibility to inflammatory bowel disease: results of a detailed haplotype-tagging investigation. Van Limbergen JE, Nimmo ER, Russell RK, Drummond HE, Anderson NH, Tenesa A, Lees CW, Arnott ID, Wilson DC, Satsangi J.
   Gut 2007;56:A122.
   Gastroenterology 2007;132(4):A151.

6. Disease phenotype but not genotype differs in children and adults with inflammatory bowel disease at diagnosis: analysis of more than 1500 patients from a homogenous patient cohort. Russell RK, Van Limbergen JE, Drummond HE, Nimmo ER, Arnott IDR, Gillett PM, McGrogan P, Hassan K, Weaver LT, Bisset WM, Mahdi G, Wilson DC, Satsangi J.
   JPGN 2007; 44(6):e17.
   Paediatrics and Child Health 2007;17(10):419.

7. Faecal Calprotectin - a significant advance in the diagnosis of IBD in childhood. Quail M, Russell RK, Van Limbergen J, Rogers P, Drummond HE, Wilson DC, Gillett PM.
   Paediatrics and Child Health 2007;17(10):412.

8. IL23R Arg381Gln is associated with childhood onset inflammatory bowel disease in Scotland. J Van Limbergen, Russell RK, Nimmo ER, Drummond HE, Smith L, Anderson NH, Gillett PM, McGrogan P, Hassan K, Weaver LT, Bisset WM, Mahdi G, Wilson DC, Satsangi J.
   JPGN 2007; 44 (6):e92.

9. A detailed haplotype-tagging investigation into the role of RICK/RIPK2 germline variation in childhood onset inflammatory bowel disease
   Van Limbergen J, Russell RK, Nimmo ER, Drummond HE, Smith L, Anderson NH, Davies G, Gillett PM, McGrogan P, Hassan K, Weaver LT, Bisset WM, Mahdi G, Wilson DC, Satsangi J.
   Gastroenterology 2007;132(4):A59.
   JPGN 2007; 44(6):e109.

10. The contribution of germline variation in the autophagy genes ATG16L1 and IRGM to childhood inflammatory bowel disease. Van Limbergen J, Russell RK, Nimmo ER, Drummond HE, Smith L, Anderson NH, DaviesG, Gillett PM, McGrogan P, Weaver LT, Bisset WM, Mahdi G, Arnott ID, Wilson DC, Satsangi J.
    Gut 2008;57(Suppl 1):A151-2.
    Gastroenterology 2008;134(4):A53.

11. A population-based ecological study identifies population density, urban/rural location, and spatial clustering as determinants of susceptibility to early-onset inflammatory bowel disease in Scotland. Wilson DC, Pollock KGJ, Van Limbergen J, Armitage EL, Aldhous MC, Drummond HE, Ternent HE, Innocent GT, Satsangi J.
    Gastroenterology 2008;134(4):A190.

12. Loss-of-function variants of the epithelial barrier protein filaggrin predispose to marked atopic co-morbidity in paediatric inflammatory bowel disease. Van Limbergen J, Russell RK, Nimmo ER, Zhao Y, Liao H, Drummond HE, Smith L, Anderson NH, Davies G, Gillett PM, McGrogan P, Weaver LT, Bisset WM, Mahdi G, Wilson DC, McLean WHI, Satsangi J.
    Gastroenterology 2008;134(4):A706.

13. A detailed investigation into epidemiological risk factors for childhood onset inflammatory bowel disease in Scotland. Van Limbergen J, Hobbs EA, Russell RK, Nimmo ER, Drummond HE, Smith L, Anderson NH, Davies G, Gillett PM, McGrogan P, Weaver LT, Bisset WM, Mahdi G, Satsangi J, Wilson DC.
    Arch Dis Child 2008;93(Suppl 1):A18.
    Gut 2008;57(Suppl 1):A150-1.
    Gastroenterology 2008;134(4):A189.

14. A detailed haplotype tagging investigation of the IL23R gene confirms gene-wide association with childhood onset IBD and CD. Van Limbergen J, Russell RK, Nimmo ER, Drummond HE, Smith L, Anderson NH, Davies G, Gillett PM, McGrogan P, Weaver LT, Bisset WM, Mahdi G, Wilson DC, Satsangi J.
    Gut 2008;57(Suppl 1):A151.
    Gastroenterology 2008;134(4):A459.

15. Childhood inflammatory bowel disease has a more extensive disease phenotype in comparison to adult onset IBD using the Montreal classification. Van Limbergen JE, Russell RK, Drummond HE, Round NK, Nimmo ER, Arnott ID, Gillett PM, McGrogan P, Weaver LT, Mahdi G, Bisset WM, Wilson DC, Satsangi J.
    Gastroenterology 2008;134(4):A189.

16. Germline variation of a novel NOD2/CARD15 interacting protein, GALNT2, is associated with genetic susceptibility to Crohn's disease (CD). AM Phillips, J Van Limbergen, G Davies, HE Drummond, L Smith, A Smith, J Satsangi, ER Nimmo.
    Gastroenterology 2009

17. Detailed haplotype-tagging study of germline variation of MUC19 in Inflammatory Bowel Disease. AM Phillips, J Van Limbergen, ER Nimmo, HE Drummond, L Smith, J Satsangi.

    Gut 2009;58(suppl 1):A56-57.

    Gastroententerology 2009

18. A Yeast Two-Hybrid Screen using CARD15/NOD2 implicates novel pathways in IBD susceptibility. Nimmo ER, Smith AJ, Aldhous MC, Quail M, Soo K, Drummond H, Van Limbergen J, Russell RK, Noble CL, Wilson DC, Satsangi J.
    Gut 2009;58(Suppl 1):A61.

19. Epidemiological risk factors for childhood onset inflammatory bowel disease in Scotland: a case-control study. J Van Limbergen, H Spiers, R Farhadi, ML Wilson, RK Russell, G Mahdi, J Satsangi, DC Wilson.

    Gut 2009;58 (Suppl 1):A64.

    Gastroenterology 2009

# Declaration of Originality

I declare that all the work in this thesis is entirely my own, unless otherwise indicated, performed in the Gastrointestinal Unit laboratory between July 2005-2008. This work has not been submitted for any other professional degree or professional qualification.

The experimental work and subsequent analysis was all carried out by Johan Van Limbergen with the exception of:

1. Dr Niall Anderson and Dr Gail Davies gave assistance with the Transmission disequilibrium testing analysis for TNF-alpha, NOD1/CARD4, IL23R, and ATG16L1.
2. TaqMan genotyping was performed by the staff of the Wellcome trust clinical research facility, Western General Hospital, Edinburgh.
3. DNA sequencing was performed in the MRC Human Genetics Unit, University of Edinburgh.
4. The genotype data for NOD2/CARD15, IBD5, DLG5 was obtained by Dr Richard K Russell. Genotypic data in the Scottish adult IBD population and healthy controls was obtained by Dr Ian Arnott (NOD2/CARD15) and Dr Colin Noble (IBD5 and DLG5).
5. The genotype data of the loss-of-function alleles of Filaggrin, described in Chapter 8, were obtained by Irwin McLean and colleagues at the Epithelial Genetics Group, Division of Molecular Medicine, College of Life Sciences and Medicine, Dentistry & Nursing, University of Dundee, UK.

This work was funded by a Research Training Fellowship from Action Medical Research, The Gay-Ramsay-Steel-Maitland or Stafford Trust and the Hazel M Wood Charitable Trust. Additinional funding for the project was received from a Wellcome Trust Programme Grant (072789/Z/03/Z). Financial assistance was also provided by Schering-Plough and the GI/Nutrition Research Fund, Child Life and Health, University of Edinburgh.

# Acknowledgements

This thesis would not have been possible without the help of a large group of people. They have all been supportive and encouraging during the three years of this project and I would like to thank them sincerely.

I am indebted to my supervisors Professor Jack Satsangi, Dr Elaine Nimmo and Dr David Wilson.

Professor Jack Satsangi has supported and inspired me to strive for the excellence he pursues in his research unit. His impressive know-how in the field of Inflammatory Bowel Disease genetics will continue to influence me, long after leaving his group.

Dr Elaine Nimmo has been a tremendously understanding and friendly companion through the ups and downs of this research project. When faced with the day-to-day business of helping me through my PhD, she always managed to match her breadth of scientific knowledge with kind advice, and both have been nothing short of amazing.

I am extremely fortunate to have had Dr David Wilson as a mentor and friend during the past seven years. Not only has he given me the opportunity to train in both paediatric gastroenterology and fundamental research for which I am very grateful. Equally important for me have been our conversations, during which he shared with me his admirable and rare mix of cutting-edge research and clinical savoir-faire, paralleled by pure savoir-vivre.

I would like to thank Dr Richard Russell for both his hard work in setting up this research repository and his friendly advice during the past four years.

Dr Peter Gillett has been there for me during the past seven years with loads of encouragement and support and I thank him wholeheartedly for it.

The other members of the GI Team at the Royal Hospital for Sick Children in Edinburgh (Pam Rogers, Catherine Paxton, Mary Smith, Julie Hardie and Jenny Livingstone) have all offered tremendous personal support and encouragement.

I would like to thank all of the members of the Scottish Paediatric Gastroenterology Hepatology and Nutrition Group: Paraic McGrogan, Kamal Hassan, Lawrence Weaver, Michael Bisset and Gamal Mahdi for 100% personal support and enrolment of patients into the study.

From the team at the Molecular Medicine Centre, I would like to give a special 'thank you' to Hazel Drummond for her amazing patience and advice in all matters relating to databases; Marian Aldhous for great help and advice particularly with the analysis of the phenotypic data, Colette McColl and Amanda Smith for their friendly support throughout my PhD; Gail Davies and Niall Anderson for never tiring of giving statistical advice; Linda Smith for all those years of recruiting families into the study.

I would to thank Professor Jürgen Schwarze from Child Life and Health at the University of Ediburgh and the Centre for Inflammation Research for his guidance and support which were invaluable in the completion of this thesis.

From the Bone Research-group at the Molecular Medicine Centre: the late Dr Stuart Bear provided the primers and PCR protocol for the Vitamin D Receptor PCR used in the validation of the saliva DNA quality.

From the Cancer Research Centre: Craig Stevens for great help in writing the introduction to the autophagy chapter.

I am indebted to the MRC Human Genetics Unit and the Wellcome Trust Clinical Research Facility (Angie Fawkes and Lee Murphy) for their support and help with the validation of the saliva DNA extraction method by generously donating the 6 Illumina 550k chips and performing the analysis.

I would also like to acknowledge the help of all patients and parents who participated in the study together with the specialist nurses, dieticians and secretaries in each of the teaching hospitals as well as the paediatricians, practice nurses and GPs throughout Scotland whose support for the study was invaluable.

# Abbreviations

| | |
|---|---|
| ANCA : | Anti-Neutrophil Cytoplasmic Antibodies |
| ASCA: | Anti-Saccharomyces Cerevisiae Antibodies |
| BMI: | Body mass index |
| CARD: | Caspase Activating Recruitment Domain |
| CD: | Crohn's disease |
| CDAI: | Crohn's disease activity index |
| CRP: | C- reactive protein |
| DAF: | Decay-Accelerating Factor |
| DepCat: | Deprivation category |
| DLG5: | Discs, large homolog 5 (Drosophila) |
| DNA: | Deoxyribonucleic Acid |
| dNTP | Deoxyribonucleotide triphosphate |
| DSS: | Dextran Sodium Sulphate |
| EDTA: | Ethylenediaminetetraacetic acid |
| ESR: | Erythrocyte sedimentation rate |
| GCSF: | Granulocyte Colony-Stimulating Factor |
| GM-CSF: | Granulocyte-Macrophage Colony-Stimulating Factor |
| GRIM19: | Gene associated with retinoid-IFN-induced mortality 19 |
| HC: | Healthy Controls |
| IBD: | Inflammatory Bowel Disease |
| IC: | Indeterminate Colitis |
| IEC: | Intestinal epithelial cell |
| IFN-γ: | Interferon-Gamma |
| IL: | Interleukin |

| | |
|---|---|
| IQR: | Inter Quartile Range |
| IRAKs: | IL-1 receptor associated kinases |
| IRF1: | Interferon regulatory factor-1 |
| LD: | Linkage disequilibrium |
| LRR: | Leucine-rich-repeat |
| MDP: | Muramyl Dipeptide |
| MDR1: | Multi-Drug Resistance 1 |
| mM: | Millimolar |
| MUC: | Mucin |
| NF-κβ: | Nuclear factor kappa-beta |
| NIDDK: | National Institute of Diabetes and Digestive and Kidney Diseases |
| NOD: | Nucleotide Oligomerisation Domain |
| OCTN: | Organic Cation Transporter |
| OR: | Odds Ratio |
| P4HA2: | Prolyl 4-hydroxylase alpha-2 subunit precursor |
| PAMP: | Pathogen Associated Molecular Pattern |
| PCR: | Polymerase chain reaction |
| PDLIM4: | PDZ and LIM domain protein 4 |
| RFLP: | Restriction Fragment Length Polymorphism |
| RR: | Relative risk |
| sCD14: | Soluble Cluster of Differentiation 14 |
| TAK1: | TGFβ-activating kinase 1 |
| TDT: | Transmission disequilibrium testing |
| TE: | 10mM Tris pH 7.6 |
| TFF: | Trefoil Factor |

| | |
|---|---|
| TGF-β: | Transforming Growth Factor - Beta |
| TLR: | Toll-like Receptor |
| TNF-α: | Tumour Necrosis Factor-Alpha |
| SNP: | Single Nucleotide Polymorphism |
| UACL: | Ulcer Associated Cell Lineage |
| UC: | Ulcerative colitis |
| UK: | United Kingdom |
| WTCCC: | Wellcome Trust Case Control Consortium |
| WTCRF: | Welcome Trust Clinical Research Facility |

# Abstract

**Background & aims:** The inflammatory bowel diseases (IBD), Crohn's disease (CD) and ulcerative colitis (UC), are common causes of chronic gastrointestinal morbidity, affecting up to 1 in 250 of the general population in Northern Europe. Up to 25% of IBD is diagnosed during childhood or adolescence. The aims for this thesis were to study the epidemiology, natural history and novel genetic determinants of childhood onset IBD in Scotland.

**Methods:** The existing repository of childhood onset and adult onset IBD patients, established at the Western General Hospital in Edinburgh, was used and expanded. Thus, anatomical location and behaviour of disease were assessed in 416 childhood onset (276 CD, 99 UC, 41 IBDU diagnosed before 17th birthday) and 1297 adult patients (596 CD, 701 UC) using the Montreal classification. Additional phenotypic (at diagnosis and at regular follow-up intervals) and epidemiological data were gathered. In this cohort, genotyping of germline variants in putative susceptibility genes (NOD1/CARD4, IL23R, ATG16L1, IRGM, FLG) was performed to enable single variant and haplotype-tagging association studies. Genotypic data of population-matched healthy controls were obtained locally (n=342) and from the Wellcome Trust Case Control Consortium (n=2937).

**Results:** Compared with adults, childhood-onset CD was characterized by a more extensive, "panenteric" phenotype (ileocolonic plus upper GI; $p<0.0001$ OR23.3; 95% CI (13.4–40.6) with less isolated ileal ($p<0.0001$ OR 0.06 (0.03–0.1) or colonic disease ($p<0.0001$, OR 0.3 (0.2–0.5)). In 39%, the anatomic extent increased within 2 years.

UC was also more extensive in children at diagnosis vs adults ($p<0.0001$ OR 5.1 (2.7–9.4)). In population-matched and age, sex and postcode-matched case-control analysis, childhood onset IBD and CD was associated with asthma ($p<0.0001$ OR 1.7 (1.3-2.1) and ($p=0.005$ OR 2.5 (1.3-4.8), respectively).

Inherited variation of NOD1/CARD4 was not a strong determinant of disease susceptibility in the Scottish population (both in single marker and haplotype-tagging studies, all p>0.05 after Bonferroni correction).

We found that the allelic frequency of rs11209026*A located within the IL23R gene, differed significantly between IBD / CD cases and controls (p=0.01 OR 0.51(0.3-0.9) and p=0.04 OR 0.5 (0.3-0.98)). Using a gene-wide haplotype-tagging strategy, we demonstrated that the multiple association signals of the IL23R locus are independent of rs11209026 in childhood onset IBD and CD.

In Scottish children, the effect of germline variation of ATG16L1 and IRGM on CD susceptibility was relatively small (OR< 1.4), and appeared less than in adult disease. Genotype–phenotype analysis demonstrated an association of pure ileal disease with the ATG16L1 rs2241880G-allele (p=0.02 OR 1.3 (1.03–1.7)). Using binary logistic regression analysis, we confirmed the effect of rs2241880 genotype (GG) on ileal disease versus colonic disease (p=0.03 OR 2.4 (1.05–5.6)).

Null alleles of the epithelial barrier protein FLG have no important effect on IBD susceptibility (p>0.4), but contribute to the high prevalence of atopy, notably co-existent eczema and food allergy (p=0.0003 OR 3.3 (1.7–6.6) and p=0.0001 OR 4.5 (2.0–10.0), respectively).

**Conclusion:** Childhood onset IBD is characterised by extensive intestinal involvement and progression of disease after diagnosis. Genetic association studies in childhood and adult IBD have provided evidence for a large number of new genomic loci. These loci encode genes involved in a number of homeostatic mechanisms: innate pattern recognition receptors, the differentiation of Th17-lymphocytes, autophagy, maintenance of epithelial barrier integrity and the orchestration of the secondary immune response.

# 1 Introduction to the genetics of the innate immune response in inflammatory bowel disease

## *Introduction*

### *1.1.1 Paediatric Inflammatory Bowel Disease*

The inflammatory bowel diseases (IBD), Crohn's disease (CD [MIM 266600]) , Ulcerative Colitis (UC[MIM 191390]), are common causes of chronic gastrointestinal morbidity, affecting up to 1 in 250 of the general population in Northern Europe.[1] Up to 25% of IBD is diagnosed during childhood or adolescence.[2] Not only does the well-documented rising incidence of childhood onset IBD in Scotland and elsewhere in the UK, pose a significant public health problem affecting delivery of care from primary to quaternary level.[3-6] The high disease burden often affects children, young adults (peak age of onset in the third decade of life) and their families in a truly pervasive manner.

Early onset IBD most commonly presents in early adolescence, impacting heavily on all aspects of physical, psychological and sexual development.[7] IBD activity influences linear growth, accrual of an adequate peak bone mass, education and future employment prospects.

Crohn's disease is characterised by transmural inflammation and granuloma formation, affecting the gastrointestinal tract from mouth to anus. The inflammation in CD is often patchy, causing so-called 'skip-lesions'. The behaviour of disease is dynamic and influenced strongly by the initial location of disease: over time the majority of patients will require surgery for stricturing or penetrating (perianal and/or intestinal) complications.[8]

The inflammation in ulcerative colitis is typically limited to the mucosa, affecting the large bowel continuously from the rectum to more proximally to a varying degree: from isolated proctitis to pancolitis. Disease activity in UC is also highly variable: Turner et al. recently reported more than 1 in 4 children with UC resident in the Greater Toronto Area required admission for intravenous corticosteroid therapy during a 10 year period, 61% of whom progressed to require colectomy after 6-years follow-up.[9]

In 10-15% of patients, IBD presents with inflammation limited to the colon without clear features to distinguish between CD and UC. IBD- type unclassified (previously

indeterminate colitis) is then the preferred diagnosis. In all types of IBD, extraintestinal manifestations can occur involving the skin, joints or eyes.

Treatment of the disease is often cumbersome and hampered by a small evidence base from appropriately designed trials in childhood.(Wilson DC et al, DDW 2008) Indeed, most of the "evidence" in childhood is distilled from adult studies. Improving our understanding of disease pathogenesis in childhood will hopefully lead to more appropriately designed treatment regimens for children and adults with IBD alike.

### *1.1.2 Innate immunity: a paradigm shift*

Until a few years ago, the perceived dichotomy between innate and adaptive/acquired immunity on the one hand and the strong evidence for a T-cell driven disease process, susceptible to treatment with immunomodulators, on the other hand, led researchers to believe a dysregulation of the adaptive immune system was the underlying mechanism in the pathogenesis of the inflammatory bowel diseases (IBD). The discovery of NOD2/CARD15 as the first susceptibility gene in Crohn's disease (followed by the implication in IBD of variant alleles of OCTN, DLG5 and TLRs) has shifted the focus of research firmly to the innate immune response and the integrity of the epithelial barrier.

The importance of the commensal bacterial flora and its communication strategies with the host is illustrated by the finding of gastrointestinal manifestations similar to IBD in several innate immunodeficiency syndromes. Efficient communication depends on adequate recognition by the host of the luminal micro-organisms. Increasing knowledge of ligands for the main pattern-recognition receptors of the innate immune system, Toll-like Receptors (TLRs) and CATERPILLER-proteins (including NOD1 and NOD2), has shown innate immunity to be much more complex than a mere mechanism to distinguish 'self' from 'non-self'.

The picture that is emerging from an exponentially enlarging body of research is that of an integrated system capable of maintaining the integrity of the epithelial barrier,

recognising commensals as well as potentially dangerous pathogens, inducing tolerance or mounting a first line of defence. If necessary, the innate immune system can also orchestrate a secondary adaptive immune response while initiating resolution of the inflammatory response.

In this first chapter, I summarize the recent advances in our understanding of the influence germline variants of genes regulating the innate immunse system have on the development of IBD. Most interest has been concentrated on the family of genes coding for pattern-recognition receptors (e.g. TLRs and NOD-LRRs), genes involved in epithelial integrity (e.g. MDR1, DLG5 and OCTNs) and mucin genes.

### *From innate immunodeficiencies to defective innate immunity*

Crohn's disease (CD) is now widely thought to arise from a dysregulated response to commensal micro-organisms in the normal mucosal flora.[10] However, recent data have suggested that the precipitating event for the chain of immunological responses leading to IBD may not be a loss of tolerance towards commensal luminal bacteria but rather an immunodeficiency.[11-13] The discovery of NOD2/CARD15 as the first susceptibility gene in CD, the elucidation of its functional properties and the association of NOD2/CARD15 homozygous variants with a younger age at diagnosis, have strengthened this hypothesis further.[14-22]

Furthermore, enterocolitides resembling IBD have been found in a myriad of rare (congenital) immunodeficiency syndromes (see Table 1). Many of these rare immunodeficiency syndromes share a quantitative or qualitative neutrophil function defect.[23-26] Although a primary neutrophil defect in CD remains a contentious area of research, various neutrophil abnormalities have been described in CD. Neutrophil chemotaxis (triggered by IL-8) was found to be impaired in CD, irrespective of NOD2/CARD15 genotype.[27-30] However, other groups have presented conflicting results with regards to IL-8 secretion in CD.[31-33] Superoxide generation defects have been demonstrated, not related to IBD activity.[34-36] Phagocytosis and microbial killing too were found to be affected in CD.[37;38] A defect of adequate neutrophil-mediated clearance of mucosal micro-organisms may lead to chronic inflammation

and the perpetuation of the T-cell driven process in IBD.[27] In therapeutic trials, Granulocyte-Macrophage Colony-Stimulating Factor (GM-CSF), a hematopoietic growth factor which stimulates cells of the intestinal innate immune system, has been used successfully in moderate-to-severe Crohn's disease.[39;40]

**Table 1-1: Immunodeficiency syndromes with IBD like enterocolitis**

Table 1-1 – Immunodeficiency syndromes with IBD like enterocolitis

| |
|---|
| Bruton's X-linked agammaglobulinemia[41-43] |
| Common variable immunodeficiency[44-46] |
| Wiskott-Aldrich syndrome[47-50] |
| Chronic granulomatous disease[11;51-55] |
| Glycogen storage disease I non a[11;23;56-58] |
| Leukocyte adhesion deficiency[24] |
| Chediak-Higashi syndrome[59] |
| Hermansky-Pudlak syndrome[60] |
| Turner syndrome[61] |
| Congenital neutropenia[62;63] |
| Cyclic neutropenia[64] |
| Autoimmune neutropenia[65] |

## *Epithelial barrier: interface between intestinal flora and mucosal immunity*

### *1.1.3 Intestinal flora*

Considerable evidence now suggests that the immune response in IBD is not directed to specific pathogens but rather to ubiquitous luminal bacteria.[10;66;67] In health, the intestinal flora consists of up to 100 trillion bacteria and undergoes very few changes after it is established during infancy.[68;69] The initial colonisation of the gut is facilitated by the delayed production of gastric acid and pancreatic proteases by the newborn infant, glycopeptides (promoting the growth of *Bifidobacterium* and *Lactobacillus* and limiting *Bacteroides* species) and immunomodulating agents (eg TGF-β, IL-10, sCD14, GCSF) in human milk.[70-76] Recent, animal data have demonstrated that the intestinal antimicrobial peptides undergo a complete developmental switch in innate immune effector expression and anatomical distribution shortly after birth.[77] The expression of cathelicidin-related antimicrobial peptide was limited to the first two weeks after birth and gradually disappeared with the onset of increased stem cell proliferation and epithelial cell migration along the crypt-villus axis.[77] After the neonatal period, Paneth-cell derived enteric antimicrobial peptides provide protection from intestinal infection and maintenance of enteric homeostasis, as discussed later in this chapter.

Breastfed infants are less often colonised with bacteria other than *Bifidobacteria*, compared with formula-fed infants.[78] However, the protective effect of breastfeeding on the development of IBD has recently been under discussion.[79-81] Other factors influencing the colonisation of the gut in the newborn like maternal smoking and neonatal infections have also been implicated but results have been conflicting.[81-86]

Clear differences have been observed between CD patients compared to healthy controls with regards to faecal flora composition, protection provided by the mucus layer and mucosal colonisation.[66;87-90] Aggressive species (eg *adherent-invasive E.Coli*, *Enterobacteriaceae* and *Bacteroides* species) are abundant on the mucosal

surface in comparison with protective genera, both in adults and children with IBD.[66;91-93] In comparison, a recent study on commensal flora in healthy subjects, reported that bacteria were only observed at the luminal side of the mucus layer and the bacterial microflora present in the faeces was similar to that in the mucus layer of the terminal ileum and colon.[94] Frank et al. have recently provided further evidence for the unique changes which occur in the microflora of IBD: a depletion of commensal bacteria, notably members of the phyla *Firmicutes* and *Bacteroidetes*.[95]

The above findings suggest that changes in both the bacterial microflora and the mucus layer are important in IBD pathogenesis.[96] Episodes of infectious gastroenteritis lead to an increased risk of developing IBD, especially in the first year post-gastroenteritis.[97] Furthermore, using pre-, pro- and synbiotics to manipulate the enteric microflora (and restore the beneficial predominance of *Lactobacillus* and *Bifidobacterium* at the expense of eg *Enterobacteriaceae* and *Bacteroides* species) has proven to be particularly successful in the treatment of pouchitis and mild-to-moderate ulcerative colitis.[98-100] There is also presumptive evidence that antibiotics (metronidazole +/- ciprofloxacin) are effective in pouchitis, perianal fistulae, Crohn's colitis and ileocolitis but not in isolated ileal disease.[98]

### *1.1.4 Intestinal Mucins*

The mucus layer acts as a lubricant and protective physical barrier between the mucosal surface and the luminal contents. It consists of mucus glycoproteins (mucins) and trefoil factors, secreted by goblet cells.[101;102] Mucins are very large filamentous molecules containing tandemly repeated peptide domains which are highly *O*-glycosylated.[101] Large numbers of sulphate and sialic acid residues added to these carbohydrate chains are responsible for the negative surface charge of mucins.[103] To date, more than twenty mucin-type glycoproteins have been identified.[104;105]

Mucins can be divided into three main categories:

1) secretory, gel-forming mucins: MUC2, MUC5AC, MUC5B and MUC6: all located in the gene cluster on 11p15
2) membrane-bound mucins: MUC1 (1q22), MUC3A, MUC3B, MUC11, MUC12, MUC17 (all on 7q22 (an IBD susceptibility locus), MUC4 (3q29 (also an IBD susceptibility locus))
3) a group of as yet of non-classifiable mucins (e.g. MUC7, MUC8, MUC15).[103;105-109]

Secretory mucins are produced in specialised mucous cells of glandular tissues and goblet cells of the gastrointestinal tract.[110] MUC2 is the predominant mucin in the small and large intestine, located in bulky apical granules of goblet cells.[111-113] MUC5B is expressed in small quantities in a subset of goblet cells at the bottom of colonic crypts.[114] MUC5AC and MUC6 are expressed in the gastric epithelium while MUC6 is also expressed in Brunner's glands of the duodenum.[115;116] The membrane-bound mucins are widely expressed in non-specialised serous or epithelial cells, like enterocytes.[117]

In IBD, the colonic mucus layer varies in thickness compared with controls.[118] The reduction of the thickness in ulcerative colitis is believed to reflect the depletion of goblet cells.[101] The increase in bacteria in the mucus layer of IBD patients contributes to further mucus degradation by the production of mucinase, sulphatase and glycosidase.[119-121] Recently, altered colonic glycoprotein expression in unaffected identical twins of IBD patients was found to be associated with pre-clinical NFκB activation.[122] A previous report had found these changes in glycoprotein expression were only present in UC.[123]

The influence of cytokines, bacteria (including probiotics), bacterial components and bacterial metabolites (e.g. butyrate) on mucin expression and structure was first shown *in vitro*.[124-129] IL-10 deficient ($^{-/-}$) mice, a widely accepted mouse model of IBD, crossed to human MUC1-transgenic mice, were recently shown to develop $MUC1^{+}$ IBD characterized by an earlier age of onset, higher inflammation scores, and a much higher incidence and number of colon cancers compared with IL-$10^{-/-}$

mice.[130] In animal models of inflammation using Dextran Sodium Sulphate (DSS) – treated mice, MUC2 gene expression was not modified by colitis, whereas MUC1, 3 and 4 were increased after acute colitis.[131] MUC2 knockout ($MUC2^{-/-}$) mice experiments demonstrated that MUC2 deficiency leads to inflammation of the colon and contributes to the onset and perpetuation of experimental (DSS) colitis.[102] In the same study, Van Der Sluis et al. demonstrated *de novo* expression of MUC6 in the goblet cells of $MUC2^{-/-}$ mice reflecting a possible compensatory mechanism through the induction of an Ulcer Associated Cell Lineage (UACL) from intestinal stem cells.[132] Another murine model was crucial in demonstrating the influence of IL-10 on MUC2 synthesis.[133] Dietary manipulations including prebiotics, probiotics or amino acids supplements were shown to influence mucin expression in animal models.[134-136] Recently, a recombinant cysteine-rich, Epidermal Growth Factor-like domain of MUC3 was suggested as a novel therapeutic agent to stimulate intestinal wound healing.[137]

MUC2 protein synthesis, secretion and sulphation are decreased in active ulcerative colitis although Myerscough et al. found MUC2 and MUC4 mRNA to be similar to controls.[118;138-141] MUC1 and MUC3 mRNA were reduced in ulcerative colitis in this study.[141] In Crohn's disease, decreased mRNA levels of MUC3, MUC4 and MUC5B were found in healthy and inflamed ileal mucosa compared with controls, whereas MUC1 mRNA was reduced in inflamed mucosa only.[142] Using different assays, Buisine et al. reported normal expression of MUC2 and MUC3 in non-inflamed ileal mucosa from CD patients compared with controls. In inflamed ileal biopsies, heterogeneous staining was observed although overall levels of hybridisation were still similar to controls.[143] In the UACL in the inflamed ileum from CD patients, expression of MUC1, 3, 4, 5AC, 5B and 6 but not MUC2 were demonstrated.[143;144] This MUC6 expression pattern matches findings in the $MUC2^{-/-}$ mice.[102] MUC5AC expression was also demonstrated in inflamed biopsies from paediatric IBD patients.[145]

Different genetic variations in the MUC3A have been associated with UC and CD.[146;147] Polymorphisms of variable number of tandem repeats (VNTRs) within this gene in Japanese and Caucasian ulcerative colitis samples were described.[146] In this study, Kyo et al found the frequency of rare VNTR alleles in UC patients to be increased relative to controls. In further work, the same group described how non-synonymous single nucleotide polymorphisms of MUC3A, involving a tyrosine residue with a proposed role in cell signaling, may confer genetic predisposition to familial, but not sporadic CD.[147] Recently, DNA sequence changes in MUC2 have been associated with CD whereas allelic variation of MUC4 and MUC13 genes was found to be associated with UC.[148]

### *1.1.5 Trefoil factors*

The integrity of the epithelial barrier not only depends on adequate protection by a stable mucus layer but also on swift restitution of epithelial continuity after injury. The process of restitution entails epithelial cell migration across the site of injury and involves the rapid disassembly of cell-cell and cell-substratum adhesions, de-differentiation as well as spreading of surface cells.[149] Secretory TFF (trefoil factor family) peptides TFF1 (also known as pS2), TFF2 (also called spasmolytic peptide (SP)) and TFF3 (or intestinal trefoil factor (ITF)) are key players in this early repair sequence.[150] They are co-localised with mucin in mucus-secreting cells and abundantly secreted by these cells onto the mucosal surface of the gastrointestinal tract.[144;151] The compact structure of TFFs results from the formation of characteristic intra-chain disulphide bonds and renders these small molecules resistant to protease degradation.[152] All three TFFS are encoded together in a cluster on 21q22.3.[153;154] During inflammatory states, the normally tight control of regional specific expression is lost.[149]

TFF1 is normally found in foveolar cells from the body to the pyloric sphincter of the stomach mucosa.[155] TFF1 expression is comparable in normal gastric tissue,

superficial gastritis, atrophic gastritis and intestinal metaplasia.[156] TFF1 knockout mice displayed gastric and small intestinal mucosa abnormalities and tumorigenesis.[157] TFF1 acts as a gastric tumor suppressor and is also involved in differentiation pathways of distinct gastric progenitor cells and the reduction of induced cell apoptosis.[155] TFF1 deficiency leads to small intestinal inflammatory disorders, possibly due to loss of mucin polymerisation.[155;158] Both in healthy mucosa and in IBD, TFF1 is co-packaged and co-secreted with MUC5AC in mucous and neuroendocrine cells.[144;159;160] To date, association studies between genetic (promoter) variants of TFF1 and IBD have not been performed.

TFF2 is expressed (together with MUC6) in the antral and pyloric glands of the stomach, within Brunner's glands of the duodenum and in the UACL.[144;161] TFF2 knockout mice displayed only mild phenotypic changes unless chemically challenged.[162] However, in these TFF2 knockout mice alterations in both antigen presentation through MHC Class I molecules and mouse defensin (cryptidin) expression were observed.[163;164] Another Paneth cell product, cysteine-rich intestinal protein (CRIP) mRNA, was also upregulated in TFF2 deficient mice.[163] CRIP plays a role in regulation of the cytokine balance and consequently the immune response.[165] These observations have led to a reappraisal of the role of TFF2, and TFFs in general, as modulators of the inflammatory response rather than merely secretory proteins, interacting with mucins. TFF2's ability to regulate nitric oxide production and modulate expression of adhesion molecules is illustrative of this more complex role.[166;167] These findings explain previous observations in a rat colitis model where exogenous TFF2 reduced the inflammatory infiltrate.[168] So far, no association studies between allelic (promoter) variants of TFF2 and IBD have been performed.

TFF3 is present in large amounts within the apical storage compartment of goblet cells and is secreted, together with MUC2, throughout the small and large intestine.[144;149] Their co-localisation in the mucus layer has an additive effect on the protection against a variety of chemical and microbial insults when compared with the their individual protective properties.[169;170] In the TFF3 knockout mouse model,

DSS administration resulted in widespread, superficial colonic ulceration and restitution was absent.[171] Like TFF2, TFF3 also modulates NO production through inducible nitric oxide synthase in a mucin-dependent manner.[172] In intestinal epithelial cells, TFF3 can also modulate the expression of Decay-Accelerating Factor (DAF / CD55), a plasma membrane protein regulating the communication between the complement system and T cell immunity.[173;174] The TFF3 peptide is thought to have a protective role against complement activation through the induction of DAF in intestinal epithelial cells.[164;174] Through its effect on the complement system, DAF is a negative modulator of T cell immunity.[175] In the human intestinal tract, DAF expression is upregulated during inflammation.[176-178] Recent reports on reduced IL-8 production and increased IL-10 secretion in intestinal epithelial cells and monocytes, respectively, provide further evidence for the immunoregulatory properties of TFF3.[179;180] No association studies have been performed between genetic (promoter) variants of TFF3 and IBD. Therapeutic use of TFF3 via enema and via a novel delivery mechanism using *Lactococcus lactis* has been reported.[181-184]

Inflammatory cytokines IL-1β, IL-6 and TNF-α are important regulators of TFF gene transcription.[185;186] Whereas TNF-α has been shown to negatively regulate TFF3 gene transcription *in vitro* via NK-κB, the *in vivo* effect of IL-6 seems to be an induction of TFF3 via STAT3 (signal transducer and activator of transcription 3).[187;188] However, at least for IL-6 the effect appears to depend critically on which signal transduction pathway is activated.[185;186;188] On the other hand, Th2 cytokines IL-4 and IL-13 increase TFF3 (and MUC2) expression, dependent on STAT6.[189] In gastric cells, ligands of the peroxisome proliferators-activated receptor γ (PPARγ) induce TFF1 and TFF2 transcription.[190] The influence on TFFs of PPARγ variants, which were recently associated with susceptibility to Crohn's disease, has not yet been investigated.[191]

### *1.1.6 Genetics and the intestinal epithelial barrier*

Until recently, intestinal epithelial cells (IECs) were considered passive mechanical barrier components of the innate immune system.[192] However, the discovery of nucleotide binding oligomerisation domain 2 (NOD2/CARD15) / caspase-recruitment domain family member 15 (CARD15) has changed this.[14-16] The characterisation of a multitude of pathogen associated molecular pattern (PAMP) receptors, able to recognise bacterial, viral and fungal as well as the ability of at least some of these receptors to sense cellular stress signals have put the IECs forward as active members of the first line of defence.[193-196]

Pro-inflammatory cytokines (eg TNF-α , IFN-γ and IL-13) cause epithelial barrier abnormalities through apoptosis and apoptosis-independent effects on tight junctions and transcellular absorption.[197-204] IL-10, MDR-1 and N-Cadherin knockout mice develop intestinal inflammation, presumably due to changes in epithelial permeability.[205-209] Recently, IEC –specific cytolytic CD8+ effector T cells were identified as the earliest initiators of the inflammatory chain reaction leading to IEC apoptosis and disruption of the mucosal barrier.[210;211]

Increased permeability has been described in IBD patients, as well as in healthy relatives and spouses.[212-219] Similarly, subclinical inflammation has been found in healthy relatives of IBD patients.[220] *In vitro* studies and intra-familial work by Buhner et al suggested that NOD2/CARD15 variants are involved in determining the intestinal permeability in CD patients and healthy relatives, thus integrating genetic discoveries and environmental risk factors into the model of an altered intestinal barrier as central to IBD pathogenesis.[221;222]

#### 1.1.6.1 MDR1(ABCB1) and PXR/NR1I2

A number of genes involved in epithelial cell function have been studied as candidate susceptibility genes in human IBD. The strongest evidence to date favours MDR1. The MDR1 gene (Multidrug Resistance 1/ATP Binding Cassette, subfamily B, member 1), which encodes P-glycoprotein 170, is located in an IBD susceptibility locus on Chromosome 7q21.[107;108] The MDR1 gene encodes an efflux pump of amphipathic toxins and is highly expressed at the apical surface of epithelia of the colon and distal small bowel.[223] Even under specific pathogen-free conditions, MDR-1 knockout mice are susceptible to developing a severe, spontaneous intestinal inflammation which is preventable by and treatable with antibiotics.[207] As is the case with many of the other candidate genes studied so far, studies into the MDR1 C3435T have revealed significant heterogeneity in patient and control allelic frequency in different populations. Associations between 3435T and UC have been reported in the German and Scottish population.[224-226] In contrast, replication studies from Germany, Great Britain, North America, Slovenia, Italy, Spain and the Netherlands have been negative.[227-232] However, a meta-analysis of 9 association studies of C3435T showed a significant association of the 3435T allele with UC, but not CD.[233] The 2677G variant was found by Brant et al. to be associated with IBD.[228] Ho and colleagues did not find an association between this polymorphism and IBD but a haplotype 3435T/2677G was significantly associated with UC.[226] Contrastingly, 2677T was associated with UC by Potočnik et al. and Onnie et al.[229;233] Using a gene-wide haplotype tagging strategy, Ho et al. observed a highly significant association between the common MDR1 haplotypes and UC, but not CD.[234] This association was critically dependent on one haplotype tagging SNP, intronic variant rs3789243. The effect of this tagging SNP was independent of C3435T. In this study, the association was strongest with the phenotype of extensive disease.[234]

The hypothesis that a defective mucosal detoxification system might be predisposing to increased risk of intestinal inflammation was further supported by DNA microarray analysis involving non-affected colonic tissue of CD and UC.[235] In this study, Langmann et al showed marked downregulation of a number of detoxification genes (of the glutathione and sulfo-transferase family) and ABC transporters

(including MDR1) in UC, together with the near complete loss of the transcriptional regulator pregnane X receptor (PXR).[236] PXR is a nuclear hormone receptor for a large number of structurally and pharmacological diverse endogenous and exogenous compounds (including pregnanes, corticosteroids, rifampicin and bile acids).[237-240] PXR regulates the induction of many genes including CYP3A4 and the ABC transporter family genes (including MDR1).[241] Dring and colleagues showed that germ-line variants of the PXR/NR1I2 gene located on 3q13.33 (an IBD susceptibility locus), confered susceptibility to IBD.[107;242] This effect was most significant for two polymorphisms in the promoter region of this gene (rs3814055 and rs1523127) in the Irish population.[242] However, a gene wide association study using a haplotype tagging strategy to assess the overall contribution of this gene to disease susceptibility in the Scottish population by Ho et al. did not replicate this finding.[243]

#### 1.1.6.2 DLG5

Further fine mapping of an IBD susceptibility locus on the pericentromeric region of Chromosome 10 led to the identification of the DLG5 (Drosophila Discs Large Homolog 5) gene, a member of the MAGUK (Membrane Associated Guanylate Kinase) family, as a CD suscepitiblity gene.[244;245] MAGUKs are known to form scaffolds for proteins involved in intracellular signal transduction. DLG5 is important in maintaining the epithelial structure, and the genetic variants in DLG5 could therefore interfere with the epithelial barrier.[246] The DLG5 is expressed most strongly in placental tissue and less so in heart, skeletal muscle, liver, small bowel, and colon.[247] Stoll and colleagues first identified two extended DLG5 haplotypes that influenced disease susceptibility in the German population.[244] The first haplotype was characterised by the presence of a polymorphism (G113A) that resulted in an amino change at position 30 from arginine to glutamine (R30Q). Stoll et al found carriage of the 113A variant to be associated with CD in a case control study and on transmission disequilibrium testing (TDT) but did not replicate this finding in a second TDT population reported in the same publication.[244] In silico analysis suggests that the 113A (R30Q) variant may impair DLG5 scaffolding function, but as yet no expression or functional studies in IBDs have been conducted. Evidence of

epistasis between the 113A variant of DLG5 and NOD2/CARD15 variants was also observed in the CD cohort. A second haplotype, haplotype A was tagged by eight marker SNPs and was observed to be significantly under transmitted in the IBD group, suggesting the haplotype may be protective.[244] Since the publication of the paper by Stoll et al., these variant alleles of the DLG5 gene have been the subject of a large number of association studies (summarised in Table 1-2), generally with conflicting results with at best a very modest effect on CD susceptibility and phenotype.[248-262]

**Table 1-2: DLG5 association studies**

| Table 1-2a - DLG5 susceptibility studies examining the 113G/A variant: case-control studies | | | | | |
|---|---|---|---|---|---|
| Population studied | Number of IBD patients | Allele frequency (cases) | Number of controls | Allele frequency (controls) | Resultant p value |
| European[244]* | 525 | 13.2% | 515 | 9.0% | 0.001 |
| Canadian/Italian[261] | 332 | 11.0% | 202 | 5.9% | 0.003 |
| UK[261] | 689 | 9.3% | 493 | 9.7% | ns |
| Scottish[248] | 652 | 11.4% | 255 | 13.1% | 0.30 |
| German[251] | 970 | 10.1% | 972 | 10.8% | ns |
| USA[263] | 281 | 8.5% | 479 | 10.3% | ns |
| German[255] | 400 | 8.0% | 422 | 10.9% | 0.19 |
| Hungarian[255] | 268 | 8.9% | 205 | 11.2% | 0.26 |
| Italian[253] | 227 | 8.6% | 160 | 11.1% | 0.22 |
| Greek[259] | 205 | 0% | 100 | 0% | ns |
| Norwegian[256] | 386 | 9.2% | 226 | 10.9% | 0.34 |
| English[254] | 1104 | 10.1% | 750 | 11.6% | 0.32 |
| Belgium[260] | 577 | 11.8% | 301 | 10.8% | ns |
| Hungarian[252] | 773 | 11.3% | 150 | 15.0% | 0.06 |
| Japanese[262] | 484 | 0% | 345 | 0% | ns |
| UK[264] | 1148 | 9.9% | 749 | 10.1% | ns |
| Canadian[257] | 581 | 8.2% | 537 | 8.7% | 0.62 |
| New Zealand[265] | 820 | 11.5% | 416 | 10.0% | 0.27 |

* Cases were European, controls were German only

| Table 1-2b - DLG5 susceptibility studies examining the 113G/A variant: family-based association studies | | | |
|---|---|---|---|
| Population studied | Number of IBD trios | IBD susceptibility | CD susceptibility |
| German[244] | 457 | 0.004 | 0.04 |
| German/UK[244] | 485 | 0.09 | 0.07 |
| Canada [261] | 182 | ns* | not given |
| UK [261] | 124 | 0.017 | not given |
| Pooled UK/Canada [261] | 306 | 0.018 | not given |
| Belgium[260] | 373 | 0.01** | 0.006** |
| Scottish[266] | 270 | 0.045 | 0.18 |

* Exact value not stated. ** Significant *undertransmission*.[1]Footnote to table 2b.

#### 1.1.6.3 IBD5 locus / OCTN1&2

The first detailed evidence for a susceptibility locus on 5q31 was obtained through linkage analysis of Canadian affected sibling pair families, especially in early-onset CD.[267] Multimarker analysis across multiple haplotype blocks defined a 250-kb region that confered susceptibility to CD.[268] This region with extensive linkage disequilibrium contains a cytokine gene cluster with several genes important in maintenance of epithelial integrity and/or immunoregulation: interferon regulatory factor-1 (IRF1), interleukin-4,5 and 13, OCTN1&2 (organic cation transporters 1&2), PDLIM4 (PDZ and LIM domain protein 4), P4HA2 (Prolyl 4-hydroxylase alpha-2 subunit precursor).[269] As long range regulatory elements may exist with the

[1] Footnote to table 2b: The TDT results to date for variants of the DLG5 113G/A SNP in different publications demonstrate a modest effect on IBD susceptibility in 3 independent populations out of 7 studied. Only one study[244] has demonstrated significant disease susceptibility for Crohn's disease.

IBD5 risk haplotype, the nearby location (<0.5Mb) of GM-CSF precursor CSF2 could also be important.[270] Several studies, including 2 recent childhood-onset cohort analyses, have replicated the association with CD (see Table 3).[251;271-279] Although evidence regarding CD location is conflicting, IBD5 may be associated with more severe disease behaviour.[260;269;271] This is reflected in childhood onset CD by an association of the IBD5 risk haplotype with indices of faltering growth.[278] The IBD5 risk haplotype has also been associated with increased susceptibility to UC in adults and children.[272;278] There is also evidence for epistasis with NOD2/CARD15 variant carriage in both CD and UC.[272;273;280]

The organic cation transporter genes (OCTN1&2) are within a single haplotype block (block 7) of the IBD5 locus. Peltekova et al. identified mutations within OCTN1 (SLC22A4 1672C/T) and OCTN2 (SLC22A5–207G/C)) and found an association between a combined TC- haplotype of these two variant alleles with increased susceptibility to CD, independent of the IBD5 risk haplotype as assessed by IGR2078 (located in block 4, ~100 kb and 3 recombination hotspots away from 1672C/T and -207G/C).[281] However, more detailed analysis by inclusion of haplotype tagging SNPs on block 5-7 have shown the contribution of the OCTN variants is not independent of the extended IBD5 haplotype, leaving the question which is the causative IBD5 variant as yet unanswered.[271;278;279;282] Association studies of the IBD5 locus with IBD, to date, have been summarised in Table 3.

**Table 1-3: IBD5 locus association studies**

| Population studied | Number of IBD/CD/UC patients | CD/UC susceptibility | Number of controls | Study type |
|---|---|---|---|---|
| Canadian[268] | 276/256/20 | CD | Not applicable | Family association |
| UK[275] | 787/330/457 | CD | 870 | Case-control |
| UK[283] | 1104/496/512 | CD/UC | 750 | Case-control |
| Belgian[260] | 373/297/66 | None | Not applicable | Family association |
| Belgian[260] | 608/472/120 | None | 305 | Case-control |
| Scottish[278] | 271/197/74 | CD | Not applicable | Family association |
| Scottish [278] | 299/200/74 | CD/UC | 256 | Case-control |
| Scottish[271] | 679/374/305 | CD | 294 | Case-control |
| UK/German[273] | 831/511/320 | CD | Not applicable | Family association |
| UK/German[273] | 1072/684/388 | CD | 701 | Case-control |
| UK[274] | 534/282/252 | CD | Not applicable | Family association |
| Japanese[274] | 178 CD | None | 156 | Case-control |
| Canadian[281] | 503 CD | CD | 390 | Case-control |
| Swedish[284] | 178 CD | CD | 143 | Case-control |
| German[251] | 988/625/363 | CD | 1012 | Case-control |

| | | | | |
|---|---|---|---|---|
| Canadian/USA[285] | 1879/1283/565 | CD/UC | Not applicable | Family association |
| German[272] | 555/368/187 | CD/UC | Not applicable | Family association |
| Japanese[286] | 488/241/247 | None | 270 | Case-control |
| Japanese[262] | 484 CD | None | 345 | Case-control |
| USA[279] | 264 CD | CD | 527 | Case-control and Family association |
| Italian[253] | 227/134/93 | None | 164 | Case-control |
| Hungarian[287] | 74 CD | None | 49 | Case-control |
| Italian[288] | 386/200/186 | CD | 347 | Case-control |
| Italian[288] | 217/108/109 | None | Not applicable | Family association |
| European[282] | 1200 | CD | 1200 | Case-control |
| Italian[289] | 1199/570/629 | CD | 357 | Case-control |
| Italian[290] | 899/444/455 | CD/UC | 611 | Case-control |
| Spanish[291] | 309 CD | CD | 408 | Case-control |

#### 1.1.6.4 Myosin IX B

Recently, variant alleles of the myosin IX B gene (MYO9B, located on 19p13, an IBD susceptibility locus), have been associated with IBD (especially UC) in a large multicentre study and more recently, in two studies from Southern Europe.[292-294] Myosin IX B is expressed in leukocytes and epithelial cells and thought to influence intestinal permeability. Although, MYO9B variants have previously been associated with celiac disease in the Dutch population, replication studies in Scandinavian and British celiac disease cohorts were negative.[295-297]

### *Molecular mechanisms of recognition of pathogen-associated motifs*

The innate immune response is not non-specific, as was originally thought, but rather is able to discriminate between 'healthy self' and a variety of pathogenic/ 'danger' signals.[298-300] As ever increasing numbers of immunostimulatory ligands are being identified, the ability of the innate immune system to recognise tissue-derived signals in the absence of pathogens as well as its role in normal physiological processes, such as transport of blood lipids and development, has dramatically increased the relevance of a thorough understanding of pattern recognition by the innate immune system.[299;301;302] The innate immune system exerts its effects through three different kinds of receptors: secreted receptors for opsonization of pathogens (mannan-binding lectin), receptors for phagocytosis (macrophage scavenger receptor) and pattern-recognition receptors (PRRs).[303] Mannan binding lectin (MBL) gene polymorphisms have been associated with lower plasma levels of MBL and are protective against UC (but not CD).[304] Reports have been conflicting regarding their association with antibodies to *Saccharomyces cerevisiae* (ASCA).[305;306]

The identification of the trinity of cellular PRRs comprising of Toll-like receptors (TLRs), NOD-like receptors (NLRs) and RIG-I-like receptors (RLRs) has led to significant advances in our understanding of innate immunity (see Figure 1).[307] To date, TLRs have been implicated in recognition of bacteria, viruses, fungi, protozoa as well as 'danger' signals such as heat shock proteins, β-defensins, hyaluronan, cardiolipin, surfactant protein and fibronectin.[299] NLRs (also known as CATERPILLERs: CARD (caspase-recruitment domain) transcription enhancer, R (purine)-binding, pyrin, lots of leucine repeats) recognise bacterial degradation products as well as cellular stress signals such as uric acid.[308-311] RLRs are intracellular viral sensors.[307] All PRRs recognise microbial components, also called pathogen-associated molecular patterns (PAMPs), that are essential for the survival of the microorganism and are therefore difficult for the microorganism to alter.[298]

### *1.1.7 Toll-like receptors*

Since the first description of a mutant 'Toll' fruitfly in 1985 and the characterisation of Toll paralogues in mammals in the mid-1990s, a total of 13 members of the mammalian TLR family have been identified.[298;312-314] Medzhitov and Janeway made the breakthrough discovery that one of these TLRs activates NK-κB, suggesting that these receptors may link innate and adaptive immunity.[313] Shortly thereafter, bacterial lipopolysaccharide (LPS) was identified as the ligand for TLR4.[315;316] To date, ligands have been identified for nine of ten human TLRs.[317]

#### 1.1.7.1 Ligand specificity

TLRs are type I integral membrane glycoproteins characterised by an extracellular domain containing varying numbers of leucine-rich-repeat (LRR) motifs and a cytoplasmic signalling domain homologous to that of the interleukin-1 receptor (IL-1R), called the Toll/IL-1R homology (TIR) domain.[298;318] Although TLRs can be further divided into several subgroups, each of which recognising related PAMPs, ligand promiscuity (eg TLR2 and TLR4) has been puzzling for Toll-researchers and mucosal immunologists alike.[299;319] For example, TLR4 recognises a very divergent collection of ligands such as LPS, the fusion protein of respiratory syncytial virus, a subunit of bacterial fimbriae, fibronectin, β-defensin and heat-shock proteins.[298;299] One attractive explanatory hypothesis proposed by Matzinger et al. is that these receptors have evolved to recognise the hydrophobic portions of molecules when these become exposed after injury.[299] LPS, peptidoglycan, lipoteichoic acid and many other immunostimulatory microbial products (e.g. flagellae) have large hydrophobic portions that are exposed when microorganisms are damaged.[299] Hydrophobic portions are also integral parts of endogenous immunostimulators like defensins and uric acid.[320;321]

The subfamily of TLR1, TLR2 and TLR6 (individually or in heterodimers) recognises lipids.[298] TLR2 plays a major role in detecting Gram-positive bacteria and is involved in the recognition of a variety of microbial components, including lipoteichoic acid and lipoproteins which are found in cell walls of Gram positive bacteria.[298] TLR2 is able to dimerise with TLR1 and TLR6, thereby increasing the

discriminatory capacity to recognise subtle changes in the lipid portion of lipoproteins.[322-324] Reports by Fukase et al and Travassos et al have recently questioned the previously held assumption that peptidoglycan recognition was TLR2 – dependent.[317;325;326] Similarly, impurities of LPS preparations have caused confusion regarding its TLR2 – stimulating activity.[327;328] To date, only chemically pure lipopeptides have been unambiguously demonstrated to be TLR2 stimulators.[298]

Crystallographic studies followed by mutational analysis were successful in identifying the specific ligand binding site of dsRNA to TLR3.[329;330] LPS (from Gram-negative bacteria) triggers signalling through TLR4 after associating with LPS binding protein, CD14 (a glycosylphosphatidylinositol linked protein expressed on the cell surface of phagocytes) and MD-2.[331;332]

CD14 is located on the IBD susceptibility locus on Chromosome 5q13.[107] The CD14 promoter polymorphism -159C/T has been associated with CD in a Greek (with evidence for epistasis with NOD2/CARD15/CARD15) and German study.[333;334] In a Japanese IBD cohort, this polymorphism has been associated with UC.[335] No significant effect of CD14-159C/T was observed in an Australian IBD population.[336] In Scottish and Irish CD patients, Arnott et al. found that the CD14-159C/T variant frequency did not differ from controls.[337]

TLR5 binds the constant domain of flagellin, the major protein constituent of bacterial flagellae.[338;339] Although this domain is relatively conserved among different species, important gastrointestinal pathogens like *Helicobacter pylori* and *Campylobacter jejuni* avoid flagellin-specific immune responses through the production of flagellins lacking pro-inflammatory properties.[340] TLR3, TLR7, TLR8 and TLR9 recognise nucleic acids. Unmethylated CpG dinucleotides from bacterial DNA are recognised by TLR9.[341;342]

#### 1.1.7.2 Expression

Expression of TLRs is modulated by pathogens, cytokines and environmental stresses. While TLRs 1,2,4,5 and 6 are expressed on the cell surface, TLRs 3,7,8 and 9 are found in intracellular compartments such as endosomes.[298] TLRs are expressed

by a variety of cells throughout the gastrointestinal tract, including intestinal epithelial cells.[331;343-348] To avoid continuing recognition of commensal bacterial ligands by healthy epithelium, TLR2 and TLR4 expression is downregulated at the luminal surface and maybe limited to the basolateral membrane for TLR5.[345-348] Thus, flagellin would only be recognised by the host when bacteria have invaded the epithelium.[298] However, Bambou et al. have shown in polarised cell lines and in *ex vivo* models that luminal stimulation of intestinal cells with flagellin does result in signalling through TLR5.[349] During inflammation, TLR2 and TLR4 expression is upregulated in intestinal macrophages.[344] In active IBD, TLR4 expression is increased in intestinal epithelial cells.[346]

#### 1.1.7.3 Signal transduction

Upon ligand binding, TLR signalling is initiated by dimerization of TLRs, which can form homodimers (such as TLR4) or heterodimers (such as TLR1, 2 and 6).[332] Dimerisation leads to conformational changes necessary for the recruitment of one of the TIR (Toll/interleukin-1 receptor)-domain containing adaptor molecules: MyD88, TIRAP/MAL, TRIF/TICAM1, TRAM and the recently identified negative regulator of TRIF-dependent signalling SARM.[314;350-354] Except for TLR3, all TLR signalling is MyD88-dependent. The diversity and specificity of TLR function is determined by the selective use of these intracellular adaptor molecules.[332] A collection of downstream kinases (IRAKs: IL-1 receptor associated kinases) is then recruited, ultimately leading to the degradation of the IKK complex and activation of NF-κB. In addition to the effect on NF-κB, TLR activation also results in activation of a wide variety of transcription factors including IRF-3, IRF-5, IRF-7, ELK1, IκBζ and AP-1.[314;332;355-357]

This is further illustrated by the findings of Rakoff-Nahoum et al in MyD88 knockout mice.[358] MyD88 knockout mice, and to a lesser extent TLR2 and TLR4 knockout mice, showed massive ulceration and denudation in the epithelial cell layer very soon after DSS administration due to a greatly decreased capacity to produce cytoprotective factors such as interleukin-6 and KC (Keratinocyte Chemoattractant)-1 chemokine in response to TLR signalling by commensal organisms.[359] On the other

hand, treatment with a synthetic TLR4 antagonist, blocking the interaction of LPS with the immune system, was able to inhibit the development of moderate-to-severe disease in two mouse models of colonic inflammation: the dextran sodium sulphate (DSS) model and multidrug resistance gene 1a-deficient mice.[360]

#### 1.1.7.4 Regulation of TLR signalling

Close regulation of TLR signalling takes place 1) in the extracellular space, 2) by transmembrane proteins, 3) intracellularly, 4) through reduction of TLR expression.[332]

1) Extracellularly, soluble decoy TLRs (eg sTLR2, sTLR4) may act to prevent overactivation of the host response against microbial products through competitive inhibition. Although multiple TLR4 mRNAs have been detected in humans (from a single copy of the TLR4 gene on Chromosome 9q33), indicating that sTLR4 could exist in humans, only murine sTLR4 has been characterised conclusively.[361] A soluble form of human TLR2 was identified by LeBouder et al. sTLR2 coimmunoprecipitates with sCD14 (a TLR co-receptor) in plasma and human milk.[362]

2) Transmembrane protein regulators of TLR signalling are ST2 (IL-1Receptor Like1/FIT-1), SIGIRR (single immunoglobulin and toll-interleukin 1 receptor (TIR) domain) and TRAILR (TNF-related apoptosis-inducing ligand receptor 1).[332] Of these only SIGIRR is located in an IBD susceptibility locus on Chromosome 11p15.5 but no germline variant studies have been undertaken to date.[107] Like ST2, SIGIRR is an orphan receptor that does not induce NF-κB activation.[363-365] *In vitro* SIGIRR has been shown to interact with TLR4, IRAK and TRAF6.[363] *In vivo*, SIGIRR-deficient mice displayed increased susceptibility to endotoxin shock.[363;366] Recent *ex vivo* patient data showed increased SIGIRR expression during sepsis.[367] SIGIRR is highly expressed in gut epithelial cells and immature dendritic cells.[368] Its important role in regulation of intestinal inflammation is illustrated by the development of more

severe DSS-induced commensal-driven colitis in SIGIRR-knockout mice compared with wildtype mice.[364;366]

3) Intracellular negative regulators of TLR signalling include MyD88s (the short form of MyD88), IRAKM, SOCS1, NOD2/CARD15, PI3K, TOLLIP and A20.

SOCS1 (Suppressor Of Cytokine signalling 1) is one of eight members of the SOCS family.[369] Macrophages from SOCS1 deficient mice produce increased levels of pro-inflammatory cytokines in response to stimulation with TLR4 and TLR9 ligands.[370;371] The gastrointestinal phenotype of SOCS1/T-cell receptor α (TCRα) double knockout mice is characterised by the earlier development of a more severe colitis compared with TCRα-knockout mice dependent on IFNγ and IL-4, resembling human UC.[372] On the other hand, transgenic mice overexpressing SOCS1 also developed spontaneous colitis with age and were more susceptible to TNBS (2,4,6-trinitrobenzene sulphonic acid) induced colitis associated with increased expression of IFNγ and TNF α and reduced levels of transforming growth factor β.[373] The SOCS1 gene is located on chromosome 16p13.13 in the IBD1 locus.[107;245;374;375] To date, association studies of sequence variants of SOCS1 with IBD have not been performed.

TOLLIP (Toll-interacting protein) is critical in maintaining the intestinal epithelium hyporesponsive to TLR2 ligands.[376] Overexpression of TOLLIP resulted in inhibition of TLR2- and TLR4- mediated NF-κB activation dependent on the N-terminal C2 domain of TOLLIP.[377-380] A mutation of a lysine residue to glutamic acid (K150E) within this C2 domain causes an inability of TOLLIP to inhibit LPS-induced NF-κB activation.[380] The TOLLIP gene (Chr 11p15.5) lies in the IBD locus on chromosome 11.[107;381;382] No genetic variants of TOLLIP have been studied in IBD so far.

Another intracellular regulator of TLR signalling is NOD2/CARD15. The NOD2/CARD15 gene is situated in the IBD1 locus (Chromosome 16q12.1).[107;374]

Watanabe et al described how Muramyl Dipeptide (MDP) (the ligand of NOD2/CARD15) suppressed TLR2-ligand-induced Th1-cell responses in wild-type mice but not in NOD2/CARD15-deficient mice.[383] In contrast, two recent reports did not support this inhibitory hypothesis.[384;385]A further study by Netea et al provided evidence for a complex regulation of TLR signalling through induction of both pro- and anti-inflammatory cytokines in response to NOD2/CARD15 activation.[386]

4) Reduction of TLR expression is achieved by degradation of TLRs through ubiquitinylation, inhibition of TLR expression by anti-inflammatory cytokines (eg TGF-β and IL-10) and TLR regulated apoptosis.[332;387] In health, intraepithelial cells demonstrate decreased expression of TLR4 and MD-2 resulting in tolerance to commensal bacteria.[196;345;388]

#### 1.1.7.5 Germline variants of TLR genes in IBD

DNA sequence variants of TLR genes in IBD have received substantial research interest.[389] TLR2 (4q31.3), TLR3 (4q35.1), TLR4 (9q33.1) and TLR9 (3p21.3) are all located in regions associated with IBD by genome wide searches (see Table 4).[107] A detailed study analysing single nucleotide polymorphisms in the coding sequence or promoter region of TLRs 1,2,5,6,7 in a Flemish case-control study did not show any association with IBD susceptibility, but genotype-phenotype associations between TLR1R80T and TLR2R753G variants and pancolitis in UC were reported.[303] The same authors also described negative associations between UC with proctitis only and TLR6S249P and the TLR1S602I variant and ileal disease involvement in CD.[303]

The TLR4 Asp299Gly polymorphism leads to altered recognition of LPS by the extracellular domain of TLR4.[390] This TLR4 polymorphism was associated with Crohn's disease and ulcerative colitis in a Belgian study.[391] In Greek, German and Dutch populations the association of the Asp299Gly polymorphism with Crohn's disease was replicated and an association with colonic disease was described.[333;392-394] Further support for an association of this variant with CD and IBD (OR 1.45

(1.11-1.90) and OR 1.36 (1.01-1.84), respectively) came from a recent meta-analysis by Browning et al.[395] In one German cohort, an association was demonstrated between ulcerative colitis and the TLR4 Thr399Ile polymorphism.[396] However, there is substantial heterogeneity between populations, and no association was noted between the Asp299Gly polymorphism and inflammatory bowel disease in either a Hungarian study or the Scottish dataset.[337;397] Recent data suggest that neither of these two variants is causal, but they are merely in linkage disequilibrium with as yet unidentified causal variants.[398]

Polymorphisms of the TLR5 gene were associated with severity of Crohn's disease in Ashkenazi Jews.[399] In another Jewish cohort, a dominant-negative TLR5 C1174T polymorphism was protective against development of CD, but not UC.[400] However, this inactivating mutation is also present in 5-10% of healthy individuals.[401]

A variant allele of TLR9, -1237C was found to be associated with CD in a German cohort.[396] In a recent study, a combined carriership of the alleles TLR9 -1237C and CD14 -260T was increased in the chronic relapsing pouchitis group when compared with UC patients with infrequent pouchitis.[402]

### *1.1.8 Caterpillers – Nod-LRRs – Nacht-LRRs – NLRs*

Following on from the success in characterization of Toll-like receptors, a new family of intracellular pattern recognition receptors, the CATERPILLER (CARD (caspase-recruitment domain) transcription enhancer, R (purine)-binding, pyrin, lots of leucine repeats) gene family was described in early 2000.[403;404] The MHC class II transactivator (MHC2TA - a protein with a nucleotide-binding domain followed by leucine-rich repeats (LRRs)) offered the blueprint for the identification of an increasingly large family of proteins – found in both plants and mammals.[405] This family of proteins has also been called NACHT (domain present in NAIP, C2TA, HET-E and TP1)-leucine-rich repeat (LRR).[406] In general, members of this family contain an amino (N) -terminal domain that consists of protein-protein interaction cassettes (eg CARDs and pyrin domains), a central nucleotide-binding oligomerization domain (NOD), which facilitates self-oligomerization and has ATPase activitiy, and a carboxy (C) -terminal LRR domain, which is involved in

ligand recognition.[407] The 4 main subgroups within the CATERPILLER family are distinguished by their N-terminal domains: acidic (e.g. MHC2TA), pyrin (e.g. NALP3, NOD6, NOD8), CARD (e.g. NOD1/CARD4, NOD2/CARD15), baculoviral inhibitory repeat (eg NAIP) and unclassified (e.g. NOD3, NOD9).[311] NOD1/CARD4 contains a single CARD at the N-terminus, whereas NOD2/CARD15 contains two CARDs.

A high number of genes within this family are associated with immunological disorders. Allelic variants of the MHC2TA gene/promoter region are associated with bare lymphocyte syndrome, rheumatoid arthritis, multiple sclerosis and myocardial infarction.[408;409] NALP3 (also known as CIAS1/cryopyrin) binds uric acid crystals and is associated with gout and pseudogout.[310] NALP3 variants are associated with familial-cold autoinflammatory syndrome, Muckle-Wells syndrome and neonatal-onset multisystem inflammatory disease.[410;411] Although not part of the CATERPILLER family, the pyrin protein (encoded by the Mediterranean Fever gene (MEFV)), which shares the pyrin domain found in the largest subgroup of CATERPILLER proteins, is noteworthy.[412;413] Recently, variants in this gene, located in the IBD locus on Chromosome 16p13.3 were found to be associated with IBD.[107;414]

Two previous studies in Jewish CD patients had failed to show a significant association.[415;416]

#### 1.1.8.1 Ligand specificity

Bacterial peptidoglycan (PGN) derivatives have now confidently been identified as the ligands for NOD1/CARD4 and NOD2/CARD15. NOD1/CARD4 recognises intracellular γ-D-glutamyl-meso-diaminopimelic acid (iE-DAP), a breakdown product of PGN from Gram-negative bacteria and a limited number of Gram-positive bacteria.[417-420]

NOD2/CARD15 is an intracellular PAMP receptor for Muramyl Dipeptide (MDP), a breakdown product of bacterial PGN of both Gram-negative and Gram-positive

bacteria.[418;421;422] In contrast to early reports, it has become evident that neither NOD1/CARD4 nor NOD2/CARD15 sensed LPS directly but rather detected contaminations with the above PGN derivatives of non-purified LPS preparations.[331;423;424] Recently, monomeric flagellin was identified as the ligand for CARD12 (IPAF).[425-428] Similar to TLRs on the cell surface, NOD proteins may have a role in the early detection of hydrophobic danger signals in the cytoplasm.[299;429;430]

These PGN motifs can enter the cell through phagocytosis (as is suggested for APCs), active transport or through pathogen induced mechanisms.[407] It has recently been shown that recognition of *Helicobacter Pylori* by NOD1/CARD4 is dependent on 'injection' of PGN into epithelial cells through a type IV secretion system, encoded in *H. Pylori*'s cag pathogenicity island.[431] MDP binds directly to the LRR region of the NOD2/CARD15 molecule, after transport across the cell membrane of IECs by hPEPT1.[417;432] Membrane recruitment of NOD2/CARD15 in intestinal epithelial cells is necessary for MDP recognition and does not occur in subjects carrying the Leu1007fsinsC mutation.[433]

#### 1.1.8.2 Expression

Both NOD1/CARD4 and NOD2/CARD15 are expressed mainly in the cytosol.[429] NOD1/CARD4 is constitutively and inducibly expressed in a wide variety of tissues and immunoregulatory cells.[194;434] NOD2/CARD15 has been shown to be expressed by antigen-presenting cells (APCs, e.g. macrophages, dendritic cells) and epithelial cells, inducible by TNF-α and IFN-γ.[195;435] Whereas most intestinal epithelial cell lines (IECs) transcribe the NOD2/CARD15 gene to mRNA, translation to protein is absent or limited.[193;195] Significant expression of NOD2/CARD15 seems to be limited to Paneth cells, which are located at the base of the intestinal crypts.[436] Consistent with upregulation by inflammatory cytokines, NOD2/CARD15 expression is increased in involved intestine in both epithelial cells and macrophages of patients with CD.[193;437;438]

#### 1.1.8.3 Signal transduction

Ligand binding to the LRR domain is followed by a complex conformational change in the CATERPILLER molecule, making the NOD domain available for self-oligomerization.[439;440] This facilitates recruitment-activation of RICK (receptor-interacting serine/threonine kinase; also known as RIPK2 (Receptor-interacting serine/threonine-protein kinase 2) or CARDIAK) through a CARD-CARD interaction at the plasma membrane.[441-443] This activation then further initiates various signalling pathways, including activation of NF-κB and mitogen-activated protein kinases (MAPKs), leading to a variety of immune responses.[385;424;437;444]

A rapidly increasing number of intracellular proteins are being identified that influence downstream signalling after NOD1/CARD4 or NOD2/CARD15 activation. By means of a Yeast Two-Hybrid screen, Podolsky et al. identified GRIM19 (gene associated with retinoid-IFN-induced mortality 19, located on Chromosome 19p13.2, an IBD susceptibility locus) as a novel interacting protein with NOD2/CARD15 (but not NOD1/CARD4).[107;445] GRIM19 was required for regulation of NF-κB activation and expression is decreased in inflamed mucosa of IBD patients.[445] These findings are puzzling as GRIM19 is a mitochondrial protein and NOD2/CARD15 does not seem to co-localise with this organelle.[446] In a recent report, Ferreira et al. did not find evidence for association of germline variation of GRIM19 with CD.[447]

Through biochemical and Yeast Two-Hybrid screens, Erbin (Erbb2-interacting protein) was identified as a novel NOD2/CARD15 inhibitory protein.[448;449] NOD2/CARD15 and Erbin co-localize at the entry foci of *Shigella Flexneri* and maximal affinity between the two proteins was observed after 30-40 minutes of infection. The Erbin gene is located on Chromosome 5q12.3 in an IBD susceptibility locus.[107]

TAK1 (TGFβ-activating kinase 1, located within the IBD locus on Chromosome 3p25 – also known as NR2C2 (nuclear receptor subfamily 2, group C, member 2)) interacts with NOD2/CARD15 through the LRR domain and reciprocally negatively regulates NOD2/CARD15 mediated NF-κB activation.[107;446;450] TAK1 knockout mice develop age-related colitis through the loss of regulatory T-cells.[451] None of the

genes associated with these proteins have been the subject of association studies in IBD so far.

CARD-CARD binding of NOD2/CARD15 to procaspase-1 leads to increased secretion of IL-1β.[452] Although CARD12 activation leads to the induction of expression of IL-1β, NOD-NOD binding of NOD1/CARD4 or NOD2/CARD15 to activated CARD12 may lead to inhibition of NOD1/CARD4- and NOD2/CARD15-mediated activation of NF-κB and production of IL-1β.[453;454] CARD6 competitively inhibits NOD1/CARD4 (but not NOD2/CARD15) mediated NF-κB activation through RICK.[455] Recently, another adaper protein CARD9 was demonstrated to have a critical function of NOD2/CARD15 mediated activation of p38 and Jnk in innate immune responses to intracellular pathogens by means of challenging CARD9 deficient mice with *Listeria Monocytogenes* and in wild-type cells.[456] Another level of complexity was unveiled by Rosenstiel et al. when they described the peculiar interaction of NOD2/CARD15 with a transcript isoform of NOD2/CARD15, NOD2-S, generated by the skipping of the third exon, which encodes for a protein that is truncated within the second CARD.[457] NOD2-S is preferentially expressed in the human colon and is upregulated by the anti-inflammatory cytokine IL-10. NOD2-S interacts with NOD2/CARD15 and RICK/RIPK2 and inhibits the oligomerisation of NOD2/CARD15, thus downregulating NF-κB activation and IL-8 release.

Taken together, these findings are likely to change our view on NOD-protein signalling from a highly discrete response to a mosaic of responses by a network of interacting family members.[407] The discovery that MDP is able to trigger IL-1β secretion through NALP3-mediated procaspase-1 activation, adds further support to this hypothesis.[458]

#### 1.1.8.4 Germline variants of CATERPILLER genes in IBD

The NOD1/CARD4 gene is located within the putative IBD susceptibility locus on chromosome 7p14.3 (see Table 4).[108] McGovern et al. described the association of the deletion variant of a complex intronic insertion/deletion polymorphism (32656)

of NOD1/ CARD4 with susceptibility to IBD, demonstrating a gene-dosage effect of the deletion variant on age of onset in CD and IBD.[459;459] In contrast, other groups studying variants of the coding-sequence of NOD1/CARD4 or the 32656 ins/del polymorphism have not found an association with IBD.[460-464] Heterogeneity of the control populations for the NOD1/CARD4+32656 deletion allele frequency between England, Scotland and Germany underlie the negative replication data.[462-464] These findings will be discussed in greater detail in the chapter on the role of inherited variation of NOD1/CARD4 in IBD.

The discovery of NOD2/CARD15 is widely accepted as a landmark in complex disease genetics. The NOD2/CARD15 gene is located within the IBD1 locus on the pericentromeric region of Chromosome 16.[14-16] The 3 major CD-associated polymorphisms (Arg702Trp, Gly908Arg and Leu1007fsinsC) are located in or near the LRR domain, interfering with the ability of NOD2/CARD15 to recognize MDP.[18;389;437] These variants have also been associated with increased transplant related mortality in bone marrow transplant patients.[465] NOD2/CARD15 variant alleles associated with Blau syndrome and early-onset sarcoidosis are located within the NOD region and lead to increased basal and MDP-induced NF-κB activity.[466;467] A recent meta-analysis (of more than 40 association studies) and a pooled analysis of more than 7000 patients, confirmed the association of NOD2/CARD15 allelic LRR variants with susceptibility to Crohn's disease as well as genotype-phenotype associations with small bowel involvement and a more aggressive disease behaviour.[468;469] In the genome-wide association studies to date and a recent meta-analysis have confirmed the importance of NOD2/CARD15 in CD in the Western World, population-specific effects are noteworthy.[470-472] The small and absent contribution of NOD2/CARD15 status to CD in Northern European and Japanese CD, respectively, is reflective of the genetic heterogeneity inherent to complex polygenic diseases.[256;337;389;473-475]

NOD2/CARD15 deficient mice do not demonstrate spontaneous intestinal inflammation.[383;385] In addition, colonic inflammation following DSS in

NOD2/CARD15 deficient mice is not different from wildtype.[385] However, enhanced colonic inflammation after DSS exposure does occur in mice expressing the Leu1007fsinsC NOD2/CARD15 variant, mediated by IL-1β.[384] Mouse macrophages carrying this NOD2/CARD15 variant show increased NF-κB activity compared to NOD2/CARD15 wild type macrophages, in contrast to human studies of NOD2/CARD15 variant function.[17;18;384;386;421;476]

#### 1.1.8.5 TLR – NOD synergy

The complexity of the response to innate immune ligands is demonstrated by the synergistic effects NOD1/CARD4 and NOD2/CARD15 stimulation have on TLR responses, both pro- and anti-inflammatory.[17;383;385;386;477-483] Conflicting reports have been published on the influence of the CD-associated Leu1007fsinsC NOD2/CARD15 variant on NOD1/CARD4 activation in response to iE-DAP.[484;485]

Recent work by Watanabe et al. in a mouse model of experimental colitis demonstrated that MDP activation of NOD2/CARD15 regulates innate responses to intestinal microflora by downregulating multiple TLR responses, not just TLR2. Thus, they obtained *in vivo* confirmation of *in vitro* experiments with Dendritic Cells which showed that prestimulation of cells with MDP reduces cytokine responses to multiple TLR ligands and that this reduction was dependent on enhanced IFN regulatory factor 4 activity.[486]

Netea and colleagues have reported that peripheral blood mononuclear cells from patients with Crohn's disease homozygous for the frameshift NOD2/CARD15 mutation displayed a loss of cross-tolerance between NOD2/CARD15 and TLR4, leading to uninhibited release of TNF-α by TLR4 ligands and intestinal bacteria. Crohn's disease carrying a wild-type NOD2/CARD15 showed no loss of this cross-tolerance to TLR4 agonists after preincubation with NOD2/CARD15 ligands, compared to controls.[487]

Discerning the precise contribution each individual PAMP receptor makes to the innate immune response has been complicated by the presence of contaminants in the

preparations used and by the milieu (e.g. patients vs. animal models, cell lines vs. mononuclear cells)-specific effects of these receptors. This is illustrated by the conflicting results reported on the interaction between TLR-2 and NOD2/CARD15 signalling by Watanabe et al. (who studied splenocytes from NOD2/CARD15 deficient mice, macrophages of mice expressing the Leu1007fsinsC NOD2/CARD15 variant and performed a luciferase reporter gene experiment in a differentially transfected HT-29 intestinal cell line) and Netea et al. (who studied human mononuclear cells, murine peritoneal macrophages and hamster ovary transfected cell lines).[383;386] As there is emerging evidence that the mucosal immune system keeps the systemic immune system largely ignorant of the (commensal) bacteria it interacts with, observations in transfected cell lines or even peripheral blood mononuclear cells may only be of limited relevance to further our understanding of the primary pathways in IBD pathogenesis.[488]

### *Human defensins*

At present there is great interest in the putative importance of defective production of anti-microbial peptides by epithelial cells in the pathogenesis of IBD. Antimicrobial proteins of the innate immune system include defensins, cathelicidin and the newly discovered RegIIIγ (regenerating islet 3 gamma), a C-type lectin.[13;489;490]

Defensins are a family of evolutionarily related antimicrobial peptides with a characteristic β-sheet-rich fold and a framework of six disulphide-linked cysteines, first identified in humans more than twenty years ago.[491-493] In addition to their antimicrobial properties, defensins have also been implicated in cell differentiation processes as well as in immunomodulation (bridging innate and adaptive immunity) through regulation of chemotaxis and the production of pro-inflammatory cytokines.[494]

There has been rapid evolution of the more than forty defensin genes identified so far in humans.[495] This is believed to reflect an evolutionary response of the immune system to the ever-changing environmental microbial flora.[496;497] Distinct defensin

gene clusters have been identified on Chromosome 6p21 (an IBD locus), 8p21-23, 20p13 (an IBD locus) and 20q11 (an IBD locus).[107;494]

In humans, α-defensins are expressed constitutively and inducibly in leukocytes, the reproductive tract and small intestinal Paneth cells.[489;495] Human neutrophil α-defensins (Human Neutrophil Peptides HNPs, 1-4) are stored in granules, which fuse with phagosomes during phagocytosis of microbes.[495] HD-5 and HD-6 are human enteric α-defensins expressed in Paneth cells. Pro-HD-5 requires Paneth cell trypsin to be converted to active HD-5 after its release from apical storage granules upon stimulation by bacterial products such as MDP (via NOD2/CARD15) and bacterial DNA (via TLR9).[435;436;498;499] Disruption of matrix metalloproteinase-7 (matrilysin, the enzyme responsible for activation of α-defensins in mice) led to increased virulence of orally administered *Salmonella typhimurium*.[500] Transgenic mice overexpressing HD-5 were markedly resistant to oral challenge with *S. typhimurium*.[501]

Following the observation that the production of α-defensins may be reduced in ileal CD, Wehkamp et al. also described this reduction was more pronounced in carriers of CD associated NOD2/CARD15 mutations.[502;503] As CD associated NOD2/CARD15 variants also lead to loss of synergy with TLR9 in mononuclear cells, the negative effect of NOD2/CARD15 variants on Paneth cell function may also be TLR9 mediated.[478] So far, germline variation of HD-5 and HD-6 in IBD has not been studied. Copy number variation does not influence HD-5 and HD-6 expression as the haploid genome contains a single copy of these innate immune genes.[504]

β-defensins are expressed in a wide variety of tissues including epithelial cells of the gastrointestinal tract.[489] Whereas HBD-1 (human β-defensin 1) is constitutively expressed in the lower gastrointestinal tract without significant up-regulation in response to inflammation, HBD-2,3 and 4 are inducibly expressed in inflammatory

conditions (e.g. infection and in response to pro-inflammatory cytokines).[505] HBD-2,3 and 4 are up-regulated in UC with no induction or even down-regulation in CD.[506-508] TLR2 and TLR4 have been implicated in inducible expression of HBD-2 in intestinal epithelial cells after stimulation with PGN and LPS through an NF-κB dependent pathway.[509;510] Pathogens like *Salmonella* and *Cryptosporidium* have developed a virulence strategy to suppress intestinal epithelial α- and β-defensin expression, respectively.[511;512] Voss et al. identified defective HBD-2 production by human embryonic kidney 293 cells transfected with NOD2/CARD15 variants in response to MDP.[513]

In contrast to HD-5 & 6, the DNA copy number of the β-defensin gene cluster on 8p23.1 is highly polymorphic within the healthy population. Recently, DEFB4 (encoding the HBD-2 protein) gene copy number of <4, has been associated with diminished mucosal HBD-2 mRNA expression and increased susceptibility to colonic CD.[514] In this study by Fellermann et al. healthy individuals, as well as UC patients and CD patients with ileal resections/disease had a median of 4 (range 2-10) HBD-2 copies per genome. Intriguingly, a higher copy number of β-defensin genes was shown to be associated with an increased risk of psoriasis.[515]

### *Concluding remarks*

The contribution of germline variants of genes regulating the innate immune response to development of IBD has received a great deal of research interest since the discovery of NOD2/CARD15/CARD15 as the first susceptibility gene in CD, 7 years ago. The intestinal epithelial cell is increasingly recognised as a key player in this innate immune response through the production of mucins and defensins as well as the expression of pattern-recognition receptors. Germline variants of MUC-, NOD-LRR- and TLR- genes have been associated with IBD in different populations

but replication of these findings in independent cohorts has proved troublesome. The relative risk effect of many of these variants is likely to be small making large case-control cohorts necessary. Pooling of datasets in collaborative efforts such as the National Institute of Diabetes and Digestive and Kidney Diseases (National Institutes of Health) NIDDK(NIH)) and European IBD genetics consortia has overtaken single centre studies with relatively low statistical power. However, controversies such as those surrounding the contribution of DLG5, MDR1 and TLR4 to development of IBD have taught us that pooling of genotyping data from studies in different populations has its risks and might mask population-specific effects. This has become increasingly important when interpreting genome wide association data.

**Table 1-4: IBD candidate genes involved in the innate immune response**

| Gene | Genomic location | IBD susceptibility locus |
|---|---|---|
| MUC3A | 7q22 | ✓ |
| MDR1/ABCB1 | 7q21 | ✓ |
| PXR/NR1I2 | 3q13 | ✓ |
| DLG5 | 10q22 | ✓ |
| OCTN1&2 | 5q31 | ✓ |
| MyosinIX B | 19p13 | ✓ |
| TLR2 | 4q31 | ✓ |
| TLR3 | 4q35 | ✓ |
| TLR4 | 9q33 | ✓ |
| TLR5 | 1q42 | |
| TLR6 | 4p14 | |
| TLR9 | 3p21 | ✓ |
| SIGIRR | 11p15 | ✓ |
| SOCS1 | 16p13 | ✓ |
| TOLLIP | 11p15 | ✓ |
| MEFV | 16p13 | ✓ |
| NOD1/CARD4 | 7p14 | ✓ |
| NOD2/CARD15 | 16q12 | ✓ |
| GRIM19 | 19p13 | ✓ |
| Erbin | 5q12 | ✓ |
| TAK1/NR2C2 | 3p25 | ✓ |
| HD-5&6 and HBD-2 | 8p23 | |

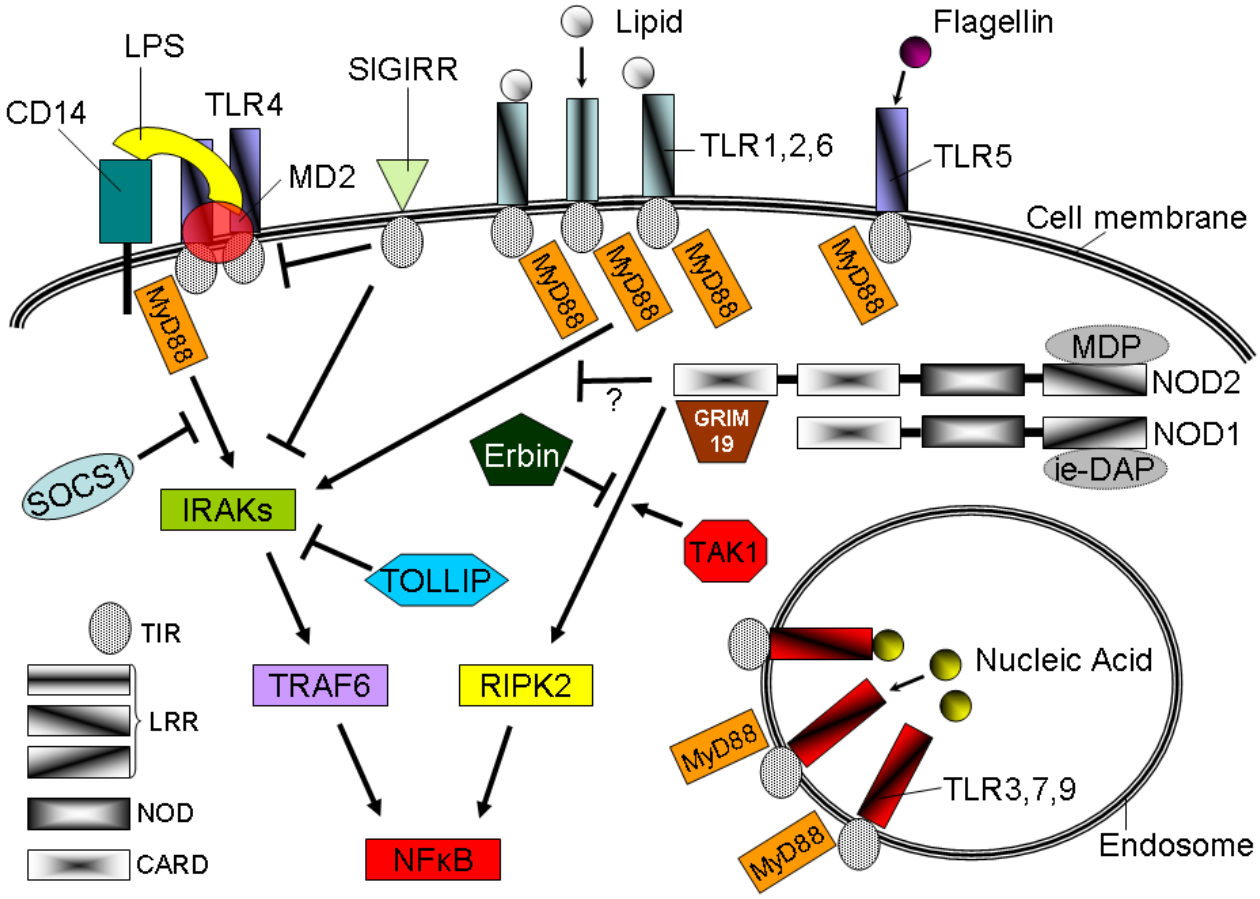



**Figure 1-1: Innate immunity pattern – recognition receptors and signal transduction pathways of relevance to IBD pathogenesis.**

TLR: Toll-like receptor. LPS (Lipopolysaccharide) triggers signalling through TLR4 after associating with LPS binding protein, CD14 (a glycosylphosphatidylinositol linked protein expressed on the cell surface of phagocytes) and MD-2. Except for TLR3, all TLR signalling is MyD88-dependent. SIGIRR (single immunoglobulin and toll-interleukin 1 receptor (TIR)) and SOCS1 (Suppressor Of Cytokine signalling 1) inhibit TLR4 signalling. TOLLIP (Toll-interacting protein) is critical in maintaining the intestinal epithelium hyporesponsive to TLR2 ligands. Activation of a collection of downstream kinases (IRAKs: IL-1 receptor associated kinases) leads to the degradation of the IKK complex and activation of NF-κB via TRAF-6 (TNF receptor-associated factor 6). NOD1/CARD4 (Nucleotide binding Oligomerisation

Domain 1/ Caspase Recruitment Domain-containing protein 4) recognises intracellular γ-D-glutamyl-meso-diaminopimelic acid (iE-DAP). NOD2/CARD15 (Nucleotide binding Oligomerisation Domain 2/ Caspase Recruitment Domain-containing protein 15) is an intracellular receptor for Muramyl Dipeptide (MDP). RIPK2 (Receptor-interacting serine/threonine-protein kinase 2) interacts with NOD2/CARD15 and NOD1/CARD4 through a CARD-CARD interaction. This activation then further initiates various signalling pathways, including activation of NF-κB. GRIM19 (gene associated with retinoid-IFN-induced mortality 19) interacts with NOD2/CARD15 (but not NOD1/CARD4). Erbin (Erbb2-interacting protein) was identified as a novel NOD2/CARD15 inhibitory protein. TAK1 (TGFβ-activating kinase 1) interacts with NOD2/CARD15 and is required for NOD2/CARD15 mediated NF-κB activation.

# 2 Patients, materials and methods

### *Definitions*

#### *2.1.1 "Early-onset" patients*

Definitions of childhood, adolescence and adulthood are highly dependent on sociocultural and socioeconomical trends. The definition of early-onset IBD is therefore equally arbitrary. The definitions used for early onset patients in IBD have varied between studies and have included <16 years,[267] <17 years[516], <19 years,[517] <20 years,[518] <21years,[519] < 22 years[520], <25 years[459] and even <40 years.[521] Rioux et al. defined an age cut off at less than 16 years in the Canadian IBD population. This study demonstrated a bimodal age distribution of patients with IBD and used Gaussian distribution curves to set this age limit.[267]

Our project has targeted children with an age at diagnosis of IBD less than 17 years. Clinical follow-up of IBD patients in the paediatric gastroenterology centres across Scotland is typically till the end of secondary education. It is important to recognise that a large number of older teenagers will never be referred to the paediatric gastroenterology services due to admission guidelines in Accident & Emergency departments (eg <13 years at RHSC Edinburgh) or primary referral to adult surgical or gastroenterology services. A UK-wide study into the investigation and management of newly diagnosed IBD in childhood (June 1998-1999) showed more than 50% of teenagers with IBD are looked after by adult services.[522] Although it is important to point out the radical overhaul that has taken place of the care of these young patients since the introduction of dedicated paediatric GI teams across Scotland (second half of the 1990s), our study can not be considered truly population-based. This has important implications for both interpretation of genetic data as well as anthropometric data. Older teenagers who have not been affected by pubertal/growth delay and who are managed successfully without high-end immunosuppression are therefore less likely to be referred primarily or secondarily to paediatric GI services across Scotland.

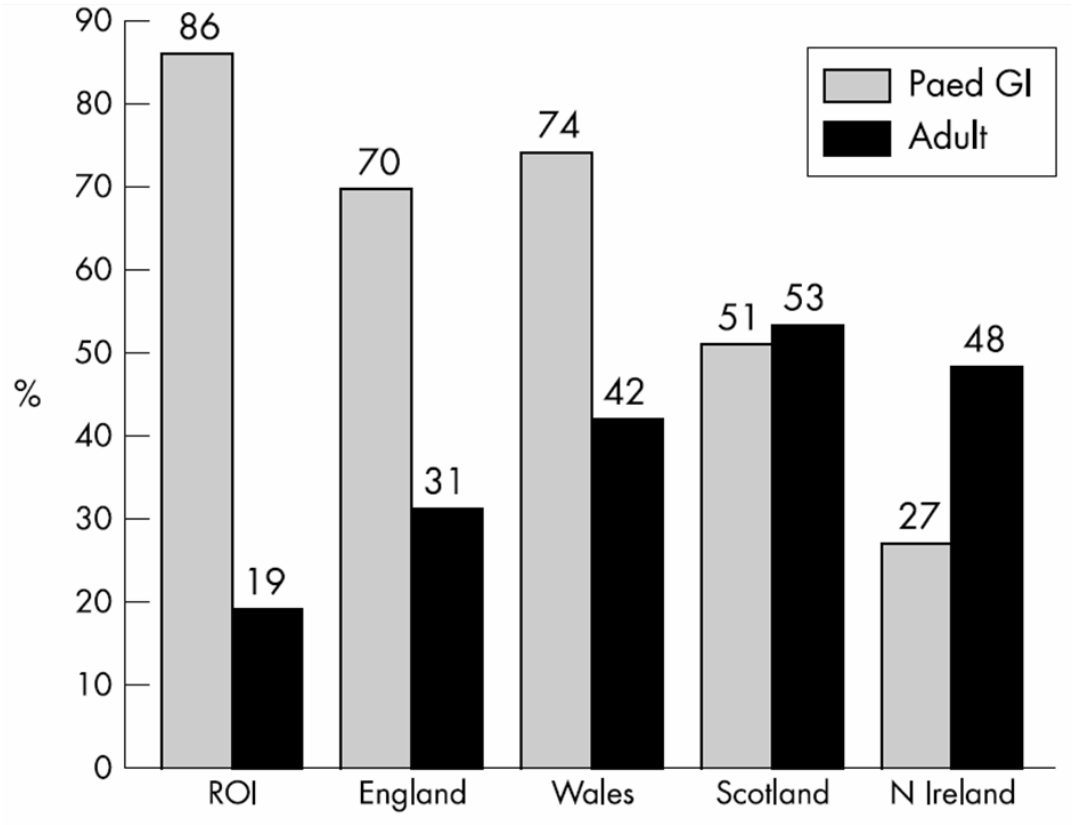



**Figure 2-1: Specialist involvement in IBD care of childhood onset patients by region in the UK.**

Figure 2.1: Adapted from Sawczenko et al.[522] Selected specialist involvement by country. Paed GI, paediatric gastroenterologist; adult, adult gastroenterologist and/or adult surgeon; ROI, Republic of Ireland. Variations between countries for Paed GI and Adult, both $p < 0.001$, $\chi 2$ test. Scotland has the highest level in the UK of involvement of adult services in the care of newly diagnosed paediatric IBD (n=71), assessed from June 1998-June 1999.

### *2.1.2 Disease definition*

The diagnosis of IBD was based on standard criteria as set out by Lennard-Jones.[523] After exclusion of enteric infection (standard coproculture +/- *Clostridium Difficile* toxin assay), these criteria were used to classify a patient as having either ulcerative colitis or Crohn's disease. A checklist based on these criteria was constructed by Dr Richard Russell and Hazel Drummond to ensure each patient entering the study definitely fulfilled the criteria for having Crohn's disease (appendix 1): only if they met the definitions in three or more section criteria, were children labelled as CD. After discussion with Hazel Drummond, exceptions to this rule were patients who

had isolated colonic Crohn's disease who had granulomas and a pancolitis with inflammatory type disease behaviour: using these criteria a score of only 1 or 2 may be generated by such patients but the presence of epitheliod granulomas in these circumstances was deemed enough to classify the patient as having Crohn's disease, as noted in the classification itself.

Patients were diagnosed with ulcerative colitis when they had uniform inflammation that commenced in the rectum (although macroscopic rectal sparing was acceptable) and continued for a varying extent more proximally. In addition to chronic inflammation on histology the patients had evidence of ulceration, cryptitis, crypt abscess formation together with goblet cell/mucin depletion without any epitheliod granulomata.

A patient was categorised as having "indeterminate colitis"/IBD type unclassified (in the Montreal classification) if definite evidence of chronic inflammatory bowel disease affecting the colon only was present, but the patient remained unclassifiable as either CD or UC, after considering all clinical, radiological, endoscopic and pathological findings and discussion with Hazel Drummond to limit interobserver variability as previously described in other IBD cohorts. [524;525]

### *Study participants*

### *2.1.3 Patients and parents*

Patients and parents were recruited from the three tertiary paediatric gastroenterology centres in Scotland based in the specialist childrens' hospitals in Edinburgh, Glasgow and Aberdeen between 2002 and 2008 (by Dr Richard Russell and Ms Linda Smith from 2002-August 2005) and from January 2006 also from the paediatric gastroenterology clinic in Ninewells Hospital, Dundee. Patients attending the Western General Hospital (Edinburgh) diagnosed with IBD under the age of 17 years were also recruited to the study. Age-appropriate information leaflets explaining the background to and the purpose of the study were supplied to all patients and parents prior to approaching them for study enrolment (appendices 2 and 3 respectively). In those patients and families who agreed to take part written consent was then obtained

from patients and parents prior to participation in the study (appendices 4 and 5 respectively).

Adult IBD patients were recruited from the Gastrointestinal Medicine Department at the Western General Hospital, Edinburgh. Informed consent from all study participants was obtained before recruitment into the study.

### *2.1.4 Blood collection*

Each patient had up to 10 millilitres (mls) of blood collected into EDTA and Lithium Heparin tubes. The blood was usually collected when the patients were having blood taken as part of their on going clinical care as requested by their supervising clinician and not merely for participation in the research study, as stated in the information leaflets. A topical local anaesthetic cream [Eutectic Mixture of Lidocaine and Prilocaine (EMLA® Astra-Zeneca) or Amethocaine base 4.0%w/w (Ametop® Smith & Nephew)] was applied if requested by the patient or family, prior to venapuncture. Parents each gave 10mls of whole blood collected into an EDTA tube collected in a Vacutainer® (Nu-care) system. In the case of "parents" who were not the genetic parents of the child, as disclosed in confidence by the guardian to the researcher, "sham" blood taking was undertaken by placing a plaster on the carer's arm, to avoid disclosure of any potentially sensitive information to the child in the course of the family agreeing to take part in the study. On rare occasions, the parent's GP kindly took the blood sample and information leaflets and consent sheets were mailed.

### *2.1.5 Healthy control recruitment*

Healthy adult controls were recruited from healthy volunteers and from volunteers recruited from the Scottish blood transfusion service. The number of controls used in each of the genetic studies varied and the exact number used in the study of each of the candidate genes is listed individually in the subsequent chapters.

### 2.1.6 Ethical approval

Ethical approval was given for all participating centres (Edinburgh, Lothian Research Ethics Committee (LREC) Reference Number: LREC/2002/6/18; Aberdeen, Grampian Research Ethics Committee (GREC) reference number: GREC 03/0273; Glasgow, Yorkhill Research Ethics Committee (YREC) reference number: YREC number P12/03). In order to extend our recruitment to Ninewells Hospital (Dundee), a new COREC application was submitted in January 2006 with a request for a Site Specific Assessment of Ninewells Hospital (local Principal Investigator: Dr Gamal Mahdi). The application received a favourable ethical opinion and the Lothian Research Ethics Committee was henceforth named as the Main REC.

Additional favourable ethical opinions were obtained for substantial amendments pertaining to the accrual of a paediatric control cohort (age-, sex-, postcode-matched to our paediatric IBD patients), both for epidemiological and genetic studies (using non-invasive saliva-kits to extract DNA from). In order to comply with Caldicott-regulations, GPs were asked to forward recruitment packs to matched healthy children from their practice database. Age-appropriate information leaflets and consent sheets were designed and approved by the LREC. (see Appendix 6)

A favourable ethical opinion by the LREC had previously been obtained for the recruitment of adult patients with IBD at the Western General Hospital, Edinburgh. (put down LREC number – get it from Colette)

### 2.1.7 Data collection

Retrospective data were collected on patients using case note review and a questionnaire (previously devised by Dr Richard Russell). All paediatric IBD patients and their parents completed a questionnaire during face-to-face interview with a member of the research team (Dr Richard Russell, Ms Linda Smith or JVL), collecting information on patient and parental smoking details, ethnicity and a detailed family history of IBD and other diseases (see appendix 7). Additional clinical data was collected on patient demographics, age at IBD symptom onset and

diagnosis, medications, extraintestinal manifestations, history of atopic disease (eczema, asthma, allergic rhinitis and food allergy) and need for surgery.

### *Disease phenotype*

All childhood patients recruited to the study were phenotyped by either Dr Richard Russell, Dr Johan Van Limbergen, Mrs Hazel Drummond, Miss Nikki Round (a fourth year medical student, as part of her SSC4 project under supervision of Dr David Wilson and JVL). Data entry onto a Microsoft Access database (designed by Dr Richard Russell and Mrs Hazel Drummond; Microsoft Corporation, Redmond, WA, USA) and quality control were performed by Mrs Hazel Drummond (database manger). Previously designed proforma sheets were used to collect the data in a consistent manner and to further reduce the chances of inter-observer error for Crohn's disease and Ulcerative/Indeterminate Colitis (Appendices 8 and 9, respectively). The phenotypic criteria used were the Montreal classification for patients with Crohn's disease [516] as well as a more detailed scoring system of each anatomical location affected by IBD including extra-intestinal manifestations and bone health.

Phenotypic information was collected at diagnosis and reassessed in 2 yearly intervals for CD patients for the first 10 years post diagnosis and then collected at 5 year intervals subsequently. In patients with UC/IC phenotypic details were collected at diagnosis and the time of last follow-up.

Detailed phenotyping of adult IBD patients occurred under the supervision of Mrs Hazel Drummond by a number of researchers from our unit (including among others Dr Ian Arnott, Dr Gwo-Tzer Ho and Dr Colin Noble).

### *2.1.8 The Montreal classification of Crohn's disease*

A summary of the classification appears in Table 2-1.

| | | | |
|---|---|---|---|
| **Age at diagnosis** | | | |
| A1 | 16 y or younger | | |
| A2 | 17 – 40 years | | |
| A3 | Over 40 years | | |
| **Disease location** | | | |
| L1 | Terminal Ileum (TI) | L1 + L4 | TI + Upper GI |
| L2 | Colon | L2 + L4 | Colon + Upper GI |
| L3 | Ileocolon | L3+L4 | Ileocolon + Upper GI |
| L4 | Upper GI | | |
| **Disease behaviour** | | | |
| B1 | Non-stricturing, non-penetrating | B1p | Non-stricturing, non-penetrating + perianal |
| B2 | Stricturing | B2p | Stricturing + Perianal |
| B3 | Penetrating | B3p | Penetrating + Perianal |

**Table 2-1: The Montreal classification of Crohn's Disease**

#### 2.1.8.1 Age at diagnosis

The age (A) category is assigned by the patient's age at diagnosis. The Montreal classification has addressed some of the difficulties of previous systems in classifying paediatric IBD: a childhood onset category was introduced (<17 years at diagnosis, A1); A2 for patients diagnosed between 17 and 40 years of age; A3 for patients > 40 years old at diagnosis.

#### 2.1.8.2 Disease location

The Montreal classification divided disease location into four different disease locations prefixed by the letter L for location: L1 for ileal disease and was defined as disease in the distal third of small bowel (on our database more strictly interpreted as disease limited to the terminal ileum), with or without caecal involvement; L2 for CD limited to the colon only; L3 for ileocolonic disease; L4 represented disease in the upper gastrointestinal tract anywhere proximal to terminal ileum. In addition to the four disease locations described in the Vienna classification (ileal, colonic, ileo-colonic and upper GI, L1-4) the Montreal classification added three further subcategories; these were the first three categories (L1-L3) with the addition of L4 where the disease sites coexisted, meaning the L4 disease location no longer negated the presence of often substantial lower gastrointestinal disease allowing for more in depth phenotypic analysis.[516;521]

It is noteworthy that patients with CD limited to the mouth and/or perianal, were not classifiable using the Montreal classification. This shortcoming of the Montreal classification will be further discussed in the chapter on the distinct phenotype of childhood onset IBD.

Disease extent was defined as the maximum disease involvement prior to the first surgical resection. Macroscopically, the minimum criteria for involvement of a disease location were ulceration or the presence of an aphthous lesion. Erythema and/or oedema did not suffice to score a site as affected by CD.

#### 2.1.8.3 Disease behaviour

The Montreal classification divided disease behaviour into three different categories prefixed by the letter B: B1 for inflammatory (thus non-stricturing/non-penetrating) CD; B2 for stricturing disease (defined as a constant luminal narrowing on radiology, at the time of endoscopy or surgery with a pre-stenotic dilatation or if accompanied by clinical signs and symptoms of intestinal obstruction); B3 for penetrating disease defined as intra-abdominal fistulas, inflammatory masses plus or minus abscess formation at any time excluding complications arising in the immediate post

operative period but not penetrating perianal disease. In the Montreal classification, recognition was given to the fact that enteric and perianal fistulae may represent different disease phenotypes. Thus, perianal disease was removed from the B3 category and instead added as a suffix "p" to any of the disease categories, B1-3. Like disease location, disease behaviour can not be down graded at follow-up.

### *2.1.9 The Montreal classification of ulcerative colitis*

In patients with ulcerative colitis disease extent was divided into three categories: E1 for patients with a proctitis (inflammation limited to the rectum), E2 for left sided disease distal to the splenic flexure and E3 for extensive disease proximal to the splenic flexure.

**Table 2-2: The Montreal classification of ulcerative colitis**

| **Disease extent in ulcerative colitis** | |
|---|---|
| E1 | Proctitis |
| E2 | Left sided (distal) |
| E3 | Extensive (pancolitis) |

### *2.1.10 Detailed anatomical location of paediatric IBD*

In the absence of a recognised paediatric classification, a more detailed scoring system of anatomical involvement of the GI tract was also used. In patients with Crohn's disease, location was categorised and analysed according to individual disease locations, regardless of disease involvement elsewhere in the gastrointestinal tract in contrast to the hierarchical model of the Montreal classification. The extent of ulcerative colitis was also scored in greater detail than suggested by the Montreal classification. Although extremely useful to explore detailed genotype-phenotype relationships, this more detailed model was used sparingly during the analysis of genotypic data to avoid stastical errors because of multiple testing.

#### 2.1.10.1Crohn's disease

2.1.10.1.1 Disease Location

For each anatomical location (based on findings from endoscopy, biopsy, surgery or barium follow through) within the gastrointestinal tract (oral, oesophageal, gastric antrum and body, duodenal, jejunal, ileal, caecal, ascending, transverse, descending and sigmoid colon, rectum and perianal area), presence or absence of CD was noted.

Oral Crohn's disease was defined by macroscopic changes (mucosal tags, deep linear ulceration, cobblestoning, lip swelling and fissuring) after examination by a paediatric dentist/oral medicine specialist only. The presence of epitheloid granulomata on buccal biopsy was also classified as oral Crohn's disease. The term oral CD was preferred to the term oro-facial granulomatosis.

Patients were diagnosed with Crohn's disease of the upper GI tract (oesophagus, stomach, duodenum or jejunum) when biopsies from any of these sites confirmed the presence of epitheliod granulomas (not merely in the presence of chronic inflammation) and/or when features of macroscopic disease were present as defined above. The presence of a chronic gastritis in the absence of other features of Crohn's disease was *not* sufficient to indicate disease involvement of the gastric body/antrum.

Perianal disease was defined by the presence of fissures, fistulae, abscess formation or perianal ulceration but not by the presence of skin tags alone.

#### 2.1.10.2Disease behaviour

The definitions for disease behaviour were the same as those used in the Montreal classification.[516] Perianal disease, using the suffix p, was only scored when case note review indicated the presence of penetrating complications such as deep ulceration, abscess formation or fistulisation. In the analysis, Montreal penetrating perianal disease was analysed separately (B1p-or-B2p-or-B3p versus B1-or-B2-or-B3).

#### 2.1.10.3Ulcerative colitis

##### 2.1.10.3.1 Disease location

The Montreal classification has offered the first widely accepted system to classify disease extent in ulcerative colitis. In order to prepare for more detailed genotype-phenotype relationship analyses, the disease extent of UC was further divided into 6 categories depending on the maximum disease extent defined by either macroscopic appearance or histology (rectum, sigmoid, descending, transverse, ascending and pancolitis when inflammation involved the whole colon). It was also noted whether the patient had a comprehensive assessment in line with investigative guidelines summarised in the Porto-criteria of the European Society of Paediatric Gastroenterology Hepatology and Nutrition (colonoscopy, upper GI endoscopy +/- Barium Follow Through) or a limited assessment (i.e. sigmoidoscopy) as well as a subjective assessment of whether the examination was adequate (eg good bowel preparation – success of ileal intubation / extent of colonic visualised during the procedure).[526]

#### 2.1.10.4Surgery

Surgery was defined as any operative intervention for IBD, thus excluding examination under anaesthesia, and divided into resective surgery, drainage procedures or other types of operations (eg fistula-, pouch-surgery).

#### 2.1.10.5Growth data

The basic anthropometric data collected in all patients were weight (in kilograms) and height (in metres) at diagnosis and at follow-up. All patients had their height measured on a wall mounted stadiometer and weight measured on regularly calibrated scales. Quetelet's body mass index (BMI) was then calculated by dividing the patients weight by their height squared. Standard UK centile charts (based on the UK 1990 population data (© Child Growth Foundation 1996)) were used to allocate a centile band for each of these three parameters (<0.4$^{th}$, 0.4$^{th}$-2$^{nd}$, 2$^{nd}$-9$^{th}$, 9$^{th}$-25$^{th}$ , 25$^{th}$-50$^{th}$, 50$^{th}$-75$^{th}$, 75$^{th}$-91$^{st}$, 91$^{st}$ -98$^{th}$,98$^{th}$-99.6$^{th}$ and >99.6$^{th}$). The collected anthropometric data were further expressed as z-scores (British 1990 Growth Reference for height, weight and BMI ©Child Growth Foundation, 2 Mayfield Avenue, London W4 1PW) to assess deviation from the mean, expressed in units of its distribution's standard deviation prefixed by + or -.

Assessment of pubertal status was also recorded but the inconsistency of case note documentation of this parameter made it unsuitable for further detailed analysis.

#### 2.1.10.6Haematological and biochemical disease markers

Using proforma sheets, data were collected for all patients who had blood tests performed at the time of diagnosis and at designated times of follow-up i.e. 2 yearly for patients with CD and at diagnosis and last follow up for patients with UC/IC (appendix 8). Abnormal blood tests were defined as follows: Albumin <35 g/l, Erythrocyte sedimentation rate (ESR) > 10mm in the first hour. C- reactive protein (CRP) was recorded as normal or high (the specific units used varied between various laboratories used by the participating centres in the study). Abnormal values

for haemoglobin concentration were age dependent: age 1-3 years <110 g/l, 3-6years <117g/l, >6 years <120g/l and platelets >400 $x10^9$/l. The normal ranges for vitamin $B_{12}$ were 193–982 nanograms/$l^{-1}$ and for serum folate 3–7μg/$l^{-1}$. Faecal Calprotectin was considered abnormal at > 50 μg/g.

## *Patients*

### *2.1.11 Demographics*

The demographics of childhood onset IBD patients are shown in Table 2.3.

**Table 2-3: Demographics of Scottish childhood onset IBD patients**

| Childhood IBD N=416 | CD | UC | IBDU |
|---|---|---|---|
| N | 276 | 99 | 41 |
| M / F | 164 / 112 | 48 / 51 | 19 / 22 |
| Median Age at diagnosis (Q1-Q3) | 11.5 years (8.9-13.2) | 10.9 years (8.8-10.8) | 10 years (7.6-12.5) |
| Caucasian | 97.8 % (270/276) | 94.9% (94/99) | 97.5% (40/41) |
| Median Duration of follow-up (Q1-Q3) | 3.7 years (1.7 – 6.0) | 3.5 years (1.1-4.8) | 2.5 years (0.4-4.1) |
| Smoking at diagnosis (Yes/No/Ex-smoker) | 1.6% / 95.3% / 3.1% | | |

Several genetic determinants were analysed in children as well as adults with IBD (n=1297). Detailed demographics of the adult IBD cohort from the Western General Hospital are presented in Table

**Table 2-4: Demographics of Scottish adult onset IBD patients**

| Adult IBD n=1297 | CD | UC |
|---|---|---|
| N | 596 | 701 |
| M / F | 216 / 380 | 342 / 359 |
| Median Age at diagnosis (Q1-Q3) | 29.7 years (23.7-43.5) | 34.5 years (26.0-50.0) |
| Caucasian | 99.3% (578/582) | 97.3% (673/688) |
| Median Duration of follow-up (Q1-Q3) | 10.3 years (3.8-20.6) | 8.9 years (4.2-16.5) |
| Smoking at diagnosis (Yes/No/Ex-smoker) | 43.9% / 44.4% / 11.7% | 19.2% / 49.9% / 30.9% |

### *Database*

All study data were entered into a Microsoft Access 2000© database. Anonymised data were stored in this password-secure Access database, which was designed and maintained for the purpose of this study and the parallel genetic studies in adult IBD patients by Dr Richard Russell and Mrs Hazel Drummond. All data were backed up on compact disc weekly, or after a period of large data entry before being stored in secure fire-resistant container. A dedicated computer without connection to the internet was used for storage of the study data.

## *Statistics*

### *2.1.12 General statistics*

Minitab statistical software version 13 (Minitab ltd, Coventry, UK) and GraphPad Instat (version 3.06 for Windows 98, GraphPad Software, San Diego California USA, www.graphpad.com) were used to analyse genotypic data and genotype-phenotype associations using Chi-squared or Fisher's exact test (for n<5) for unifactorial analysis. Non-parametric data were analysed using the Mann-Whitney U test or Kruskall-Wallis test (for analysis of age at diagnosis as a function of genotype) in Minitab.

Multifactorial analysis applying binary logistic regression analysis was performed using Minitab software. Interaction between the genetic factors entered in the multifactorial models was also assessed.

Log-likelihood analysis was carried out using PM software (1000 permutations – random seeds).[527]

Power calculations were performed using QUANTO version 1.2 (http://hydra.usc.edu/gxe).[528;529]

### *2.1.13 PedCheck*

The PedCheck programme (Department of Human Genetics, University of Pittsburgh, Pittsburgh, PA) was used to check for genotyping error and potential cases of non-paternity.[530] Any family trios with inconsistent results incompatible with normal patterns of inheritance were excluded from further genetic analyses.

### *2.1.14 Transmission disequilibrium testing*

Transmission disequilibrium testing analysis was performed using the TRANSMIT package (Version 2.5, David Clayton, Cambridge, UK).[531] This method makes full use of the data available using inferred genotypes if genotypic data from both parents are not available. Assistance with data analysis with the TRANSMIT programme was provided by Dr Niall Anderson (Senior lecturer in statistical genetics, Edinburgh

University). Latterly, single and multiple marker TDT analyses were performed using FBAT software (version 1.7.3, available from http://www.biostat.harvard.edu/~fbat/fbat.htm). [532] Odds Ratios (ORs) (99.5% Confidence Interval) for each of the SNPs were calculated using TDTHAP, run in R 2.5.0.[533]

For the analysis of Filaggrin, DAB1, TTC29 and IRGM SNPs, PLINK software was used for TDT analysis (http://www.journals.uchicago.edu/doi/abs/10.1086/519795).[534]

### *2.1.15 Integration of case-control and TDT-analyses*

The method for integrating results from case-control studies and TDT to provide a combined estimate of disease-marker association was described by Kazeem and Farall.[535] As this approach is only useful when the samples of subjects used in the two analyses are independent of each other, yet share the same genealogy, we based this calculation on our TDT cohort of early onset IBD and our samples of adult onset IBD patients in the case-control analysis.

## *Cellular methods*

### *2.1.16 DNA extraction*

Until January 2006, genomic DNA was extracted from blood using a modified salting-out technique (Appendix 10).[536] Thereafter, the Nucleon kit (Tepnel Life Sciences PLC, Manchester, United Kingdom) was used by staff at the MRC Human Genetics Unit to extract DNA from peripheral blood. DNA samples were then subjected to the PicoGreen® dsDNA Quantitation Reagent (an ultra-sensitive fluorescent nucleic acid stain for quantitating double-stranded DNA (dsDNA)) at the Wellcome Trust Clinical Research Facility. The solution was then re-suspended in an appropriate volume of 1xTE (10mM Tris (pH8.0), 1m EDTA (pH8.0)) to give a final DNA concentration of 100ng/µl.

### *2.1.17 Validation of DNA quality extracted from saliva*

The problems associated with obtaining a peripheral blood sample of children with IBD after they have entered remission, children with difficult venous access or needle phobia and the ethical considerations rendering the establishment of a paediatric (blood) DNA control cohort unacceptable, made us consider the use of kits to extract DNA from a saliva sample.

DNA Genotek (DNA Genotek Inc., 29 Camelot Drive, Ottawa, Ontario, Canada, http://www.dnagenotek.com/) produce the Oragene DNA Self-collection kit. I performed an initial validation of the performance of this extraction method based on 1) DNA size, 2) performance in PCR/sequencing and 3) concordance of blood- and saliva-derived DNA on the Illumina 550k chip.

1) 1 microgram of DNA from 10 individuals, extracted from their saliva using the Oragene DNA protocol, was run on a standard agarose gel for 90 minutes together with a Lambda Hind III DNA ladder (http://www.abgene.com) (Figure 2-2). The dsDNA fragments all proved to be greater than 20,000 basepairs in length.

2) Next, I performed a standard PCR reaction of the DNA sequence flanking the Vitamin D receptor BSM I allele, baring in mind the necessity of good quality DNA for this PCR to be successful (background, protocol and primers obtained from Dr Stuart Bear, Bone Research Group (Professor Stuart Ralston), Molecular Medicine Centre)).[537] For this experiment, paired blood and salivary DNA from 6 individuals was used (Figure 2-3). Consistent DNA product lengths were obtained in unevenly numbered columns (containing blood derived DNA) and evenly numbered columns (containing saliva derived DNA). Both performed equally well when reading the DNA sequence.

3) Lastly, Angie Fawkes and Lee Murphy from the WTCRF kindly offered to assess the performance of saliva derived DNA on the Illumina 550k chip. From 3 individuals, paired blood and saliva derived DNA was analysed. The call rates and reproducibility rates are shown in Table 2-5.

**Table 2-5: Call rates and reproducibility rates for paired blood/saliva DNA samples on Illumina550k Chip**

| Sample Nr | Call Rate Saliva/Blood DNA | Reproducibility Rate |
|---|---|---|
| RR1096 | 95.03% / 98.85% | 99.7% |
| RR1284 | 98.15% / 98.82% | 99.9% |
| RR1285 | 98.12% / 99.48% | 99.9% |

**Figure 2-2: Lambda Hind III dsDNA ladder versus salivary DNA**

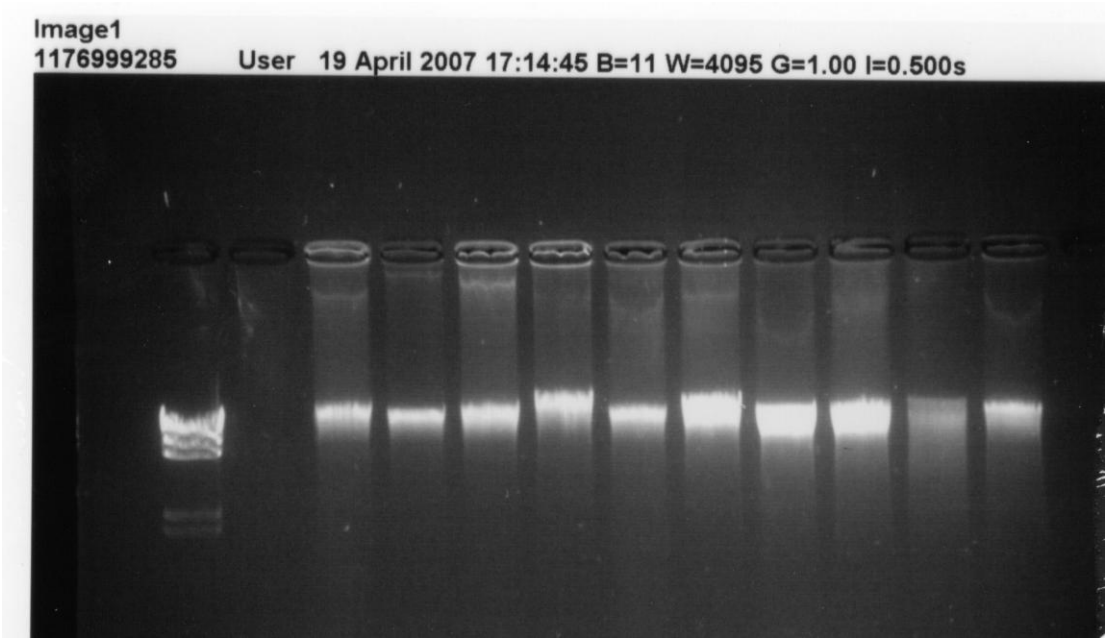
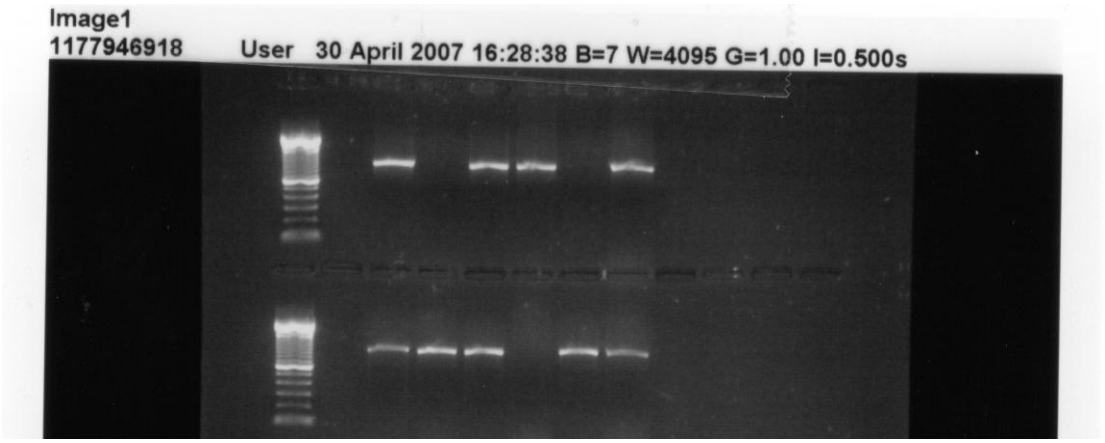



Figure 2-2: Standard Agarose Gel run for 90 minutes as discussed below. Lambda Hind III ds DNA ladder was run in the first column. The next 10 column contain 1 microgram of DNA extracted from saliva of 10 volunteers (parents and children with IBD). The DNA size was shown to be greater than 20 kilobases.

**Figure 2-3: PCR of Vitamin D receptor BSM 1 allele flanking sequence**




Colums 1 2 3 4 5 6

**Figure 2-3: PCR of Vitamin D Receptor BSM1 allele flanking sequence using paired blood/salivary DNA samples from 6 individuals, shows consistent fragment lengths between uneven (containing blood derived DNA) and even (containing saliva derived DNA) columns.**

### *2.1.18 Lymphocyte and Plasma extraction*

Peripheral blood lymphocytes were extracted from the Lithium Heparin samples collected from children with IBD. The lymphocytes were then cryopreserved. Plasma was isolated at the time of peripheral blood lymphocytes extraction and samples were stored at -80$^{o}$C until use. The lymphocyte extraction and separation of plasma was carried out by the staff in the Wellcome Trust clinical research facility, University of Edinburgh.

### *2.1.19 PCR*

All standard Polymerase chain reactions (PCR) were run on a Techne Touchgene Gradient machine or on a Peltier Thermal Cycler (DNA Engine Tetrad 2). The

reagents used and the specific reaction conditions are discussed for the individual SNPs in the subsequent chapters.

### *2.1.20 DNA sequencing*

Direct sequencing was performed on the 7900 HT sequence detection system (Applied Biosystems, Foster City, Ca, USA) by the Technical Services section of the MRC Human Genetics Unit, Edinburgh. DNA sequences were analysed using Sequencher v 4.5 (Gene Codes Corporation, Ann Arbor, MI, USA).

### *2.1.21 SNP selection for haplotype -tagging approach*

Haplotype-tagging SNPs were selected using genotypic data from the CEU-study group (Thirty U.S. trios with Northern and Western European ancestry, collected by the Centre d'Etude du Polymorphisme Humain (CEPH)), available from the HapMap project (Release 21, July 2006; http://www.hapmap.org/cgi-perl/gbrowse/hapmap_B35/). SNPs were selected using Haploview software (version 3.32, freely available from http://www.broad.mit.edu/mpg/haploview/). SNPs were identified using a haplotype-tagging strategy based on solid spine of linkage disequilibrium ($r^2$>0.8, haplotype frequency > 5%, minor allele frequency > 10%). A more detailed description of the SNP selection process for each of the genes analysed by means of a haplotype-tagging approach are given in the respective chapters.

### *2.1.22 TaqMan*

The majority of genotyping was performed using TaqMan® (© Biosystems) and was performed in the Welcome Trust Clinical Research Facility (WTCRF) located within the Western General Hospital, Edinburgh. This PCR-based assay uses laser scanning technology that excites fluorescent dyes present in specifically designed probes. The system includes a built-in thermal cycler, a laser to induce fluorescence, CCD (charge-coupled device) detector, real-time sequence detection software, and TaqMan® reagents for the fluorogenic 5' nuclease assay. The cycle-by-cycle

detection of the increase in the amount of PCR product is quantified in real time as the special probes, "reporter dye", fluoresces when the "quencher" is removed from the fragment during the PCR extension cycle.

#### 2.1.22.1SNP design

TaqMan genotyping was ordered using primers from an Assay on Demand Assay for SNPs that had previously been genotyped or, the target sequence for each SNP was submitted to the Assay by Design Service for Custom SNP Genotyping Assays. The Assays from these services consist of a mix of unlabelled primers and TaqMan Minor Groove Binder probes. Genotyping was performed using DNA plated in 384 (4x96) well-plates, using the TaqMan polymerase chain reaction-based method. In each plate, negative control wells are used to check the quality of reaction.

#### 2.1.22.2PCR reaction

The final volume PCR reaction was 5 micro litres (µl) using 20 nanograms of genomic DNA, 2.5 µl of TaqMan Master Mix and 0.125µl of 40x Assay By design Genotyping Assay Mix, or 0.25µl of 20x if using Assay On Demand Genotyping Assay. The cycling parameters were as follows: 95° for 10 minutes, followed by 40 cycles of denaturation at 92° for 15 seconds and annealing/extension at 60° for 1 minute. PCR plates were then read on ABI PRISM 7900HT (Applied Biosystems) instrument with Sequence Detection System (Applied Biosystems) version 2.1 software.

#### 2.1.22.3Allele discrimination

Allelic discrimination using this chemistry is based on the design of two TaqMan probes, specific for the wild type allele and the mutant allele separated using laser scanning technology. Each of the two probes is labelled with a different fluorescent tag 6-carboxy-fluorescein (FAM) and VIC®. Each probe is designed with the gene mutation affecting the middle part of the probe sequence. The binding efficiency of the wild type TaqMan probe to the mutant allele and vice versa is low, due to the

mismatch within the TaqMan probe and the target sequence; therefore, mismatched binding is highly reduced.

#### 2.1.22.4Results

Allocation of allele status as homozygous wild type, heterozygote or homozygous mutant was made by personnel in the WTCRF blinded to the study aims. The results from TaqMan studies were then collected onto an Excel 2000© spreadsheet before being entered onto the Access 2000© study database, with results matched using pre-assigned unique study numbers.

### *Primers*

Primers for sequencing the TNF-α promoter region, the insertion/deletion polymorphism +32656 of NOD1/CARD4, and the exons of IRGM gene were designed using data from Ensembl (www.ensembl.org). Primers for the sequencing of the CpG-island of the promoter region of NOD2/CARD15 were designed by entering data from Ensemble into the MethPrimer Program (http://www.urogene.org/methprimer/ ). The primers used in analysis of the above genes are listed in the relevant chapters.

### *Haplotypes*

Haplotypes were calculated using Haploview version 3.32.

### *Gels used for PCR reactions*

### *2.1.23 Preparation*

Multipurpose agarose (2.25g) was added to 150 millilitres (ml) of half strength TBE (53g of Tris base, 27.5g o boric acid and 20ml of 0.5M EDTA for 1 l of 5X solution, then diluted to 0.5 X) into a 250 ml beaker. This was then microwaved on full power

for between 2 and 3 minutes agitating half way through until all of the agarose had dissolved into solution. The agarose was then cooled to around 50°C by running under a cold water tap, before 10 microlitres (µl) of ethidium bromide, latterly 15 microliters of SYBR® Green (http://www.invitrogen.com ) was added and then agitated further to ensure even distribution. The agarose solution was then poured into a transparent gel-casting tray, which was sealed at each end using autoclave tape. Four combs containing either 12 or 16 wells were then inserted into the gel. The gel was then left to cool and solidify at room temperature for at least 30 minutes.

### *2.1.24 Electrophoresis*

The gel was removed from its tray and case removed before being placed into an electrophoresis tank filled with half strength TBE. Before loading the PCR mix in question 5 µl of orange G loading buffer (20g of Sucrose in 40ml water, dissolve 100mg of Orange G in above solution, q.s. to 50ml with water) had been added to 20µl of PCR mix. The mixture was then pipetted into the wells in the gel before running the electrophoresis. The running time varied dependent on optimal conditions for the PCR reaction in question but a standard running time of 30 minutes (depending on the size of the DNA fragment) and 150 Watts of power was used for most gels.

### *2.1.25 Interpretation*

After the gel had completed electrophoresis it was then removed from the tank and placed in a covered ultraviolet hood. Images were then recorded using a digital camera image loaded onto a designated computer.

# 3 Phenotypic characteristics of childhood onset IBD

### *Introduction*

The inflammatory bowel diseases (IBD), Crohn's disease (CD) and ulcerative colitis (UC), are complex polygenic diseases, and are a major cause of morbidity in Europe and North America, where up to 1/250 of the general population are affected.[1] The direct economic costs are estimated as £720 million per year for patients in the United Kingdom.[1;538]

As many as 25% of patients first present during childhood or adolescence.[2] In children, these diseases have been marked by a steady rise in incidence over the last 4 decades, with disease presenting most commonly immediately before the start of the teenage years, thereby impacting on emotional and physical development as well as affecting linear growth, education and future employment prospects.[4]

It has remained a topic of great debate whether childhood-onset disease is etiologically distinct from adult-onset disease – a debate recently catalysed by the search for susceptibility genes in CD and UC. Clinical experience, including the early requirement for second line immunomodulatory drugs [539;540] in childhood-onset disease, suggests that childhood-onset disease may have a more "severe" phenotype. This hypothesis has been supported to some extent by the limited available data suggesting that childhood-onset CD may be characterised by extensive intestinal involvement at presentation.[541-543] However, rigorous studies investigating the progression of intestinal involvement and behaviour in childhood disease are not available to help explore this hypothesis further.

In contrast to childhood-onset disease, the development of internationally accepted disease classification systems in recent years has facilitated longitudinal studies addressing changes in disease phenotype in adults with IBD.[521] A series of careful studies have consistently demonstrated the stability of disease location over time, as well as the progression of the behaviour phenotype from purely inflammatory

disease to either stricturing or penetrating disease.[544] The most recent classification system, proposed by the Montreal working party, attempted to integrate the increasing knowledge basis that had evolved since previous groups in Rome and Vienna had addressed these issues.[516] Although best considered a work in progress, the *Montreal* classification has formed the basis of phenotypic classification in many published adult clinical and genetic studies since 2005.

In the present study we have described the presenting phenotype and progression of disease phenotype in childhood-onset CD and UC. We have used the phenotypic sub-classification system for disease behaviour and location suggested in the report of the Montreal working party, and we have used the age of 16 years or younger at diagnosis to define childhood-onset disease (A1 in the Montreal classification). We also studied the phenotype of IBD in the subgroup of children diagnosed in early childhood – before the eighth birthday. We evaluated the need for immunosuppressive medication and surgery in childhood-onset disease, and compared time to first surgery in childhood-onset and adult-onset IBD. Finally, we assessed whether the phenotype of childhood-onset IBD may be different from adult-onset IBD by comparing these phenotypic characteristics in the Scottish population.

## *Methods*

### *3.1.1 Subjects*

416 children with IBD diagnosed before their 17th birthday were recruited from all the paediatric gastroenterology centres in Scotland as part of an ongoing childhood onset IBD genetics project. Children were investigated according to the ESPGHAN 'Porto'-criteria for diagnosis of IBD.[526;545] Detailed demographics are presented in Table 2-3.

1297 patients with adult onset IBD were recruited from the Western General Hospital, Edinburgh. Detailed demographics are presented in Table 2-4.

The study was approved by the local Research Ethics Committee. Informed personal or parental consent was obtained.

### 3.1.2 Classification of IBD

The diagnosis of IBD was based on standard criteria as set out by Lennard-Jones.[523] All patients were phenotyped using the Montreal classification (Tables 2-1 and 2-2) which has addressed some of the difficulties of previous systems in classifying paediatric IBD: a childhood onset category was introduced (<17 years at diagnosis) and the Vienna-classification was modified so that upper gastrointestinal (GI) disease could 'co-exist' with lower GI locations of disease (see Figure 3-1).[516;546] A dedicated database manager (HE Drummond) performed quality control of phenotypic data entered onto a Microsoft Access database (Microsoft Corporation, Redmond, WA). Oral Crohn's disease was defined by macroscopic/biopsy changes after examination by a paediatric dentist/oral medicine specialist only.

We scored the progression of anatomical involvement of the GI tract and disease behaviour in 2-yearly intervals to 4 years in childhood CD. Adult CD location was assessed at last clinical follow-up and adult CD behaviour was analysed after 5-years clinical follow-up. Childhood and adult onset UC was assessed at diagnosis and at last clinical follow-up.

When comparing childhood and adult onset IBD, the location was defined as the maximum extent by the time of last follow-up but before the first resection.[516]

### 3.1.3 Statistical analysis

In order to avoid statistical errors generated by multiple testing, we opted to analyse our data first as contingency tables consisting of columns of Montreal categories versus rows of different time points of clinical follow-up. Following this, unifactorial analyses were performed using $\chi^2$ or Fisher's exact test (when n<5 per field of 2x2 table) to calculate Odds Ratios (OR) and 95% Confidence Intervals (CI). Minitab software (Release 13.20, Minitab Inc., State College, PA, USA) and GraphPad Instat software (v 3.06, San Diego, Ca, USA) were used. Kaplan-Meier curves to assess

time from diagnosis to 1st surgical resection were calculated using GraphPad Prism software (v 4.0, San Diego, Ca, USA). Analysis of time to use of immunomodulatory drug therapy (first use of any of azathioprine, methotrexate or biological therapy) was also assessed and is presented as a Kaplan-Meier curve in the childhood-onset group only. Differences between Kaplan-Meier curves were compared using a logrank test automatically calculated in Prism.

**Figure 3-1: The Montreal classification of Crohn's disease**

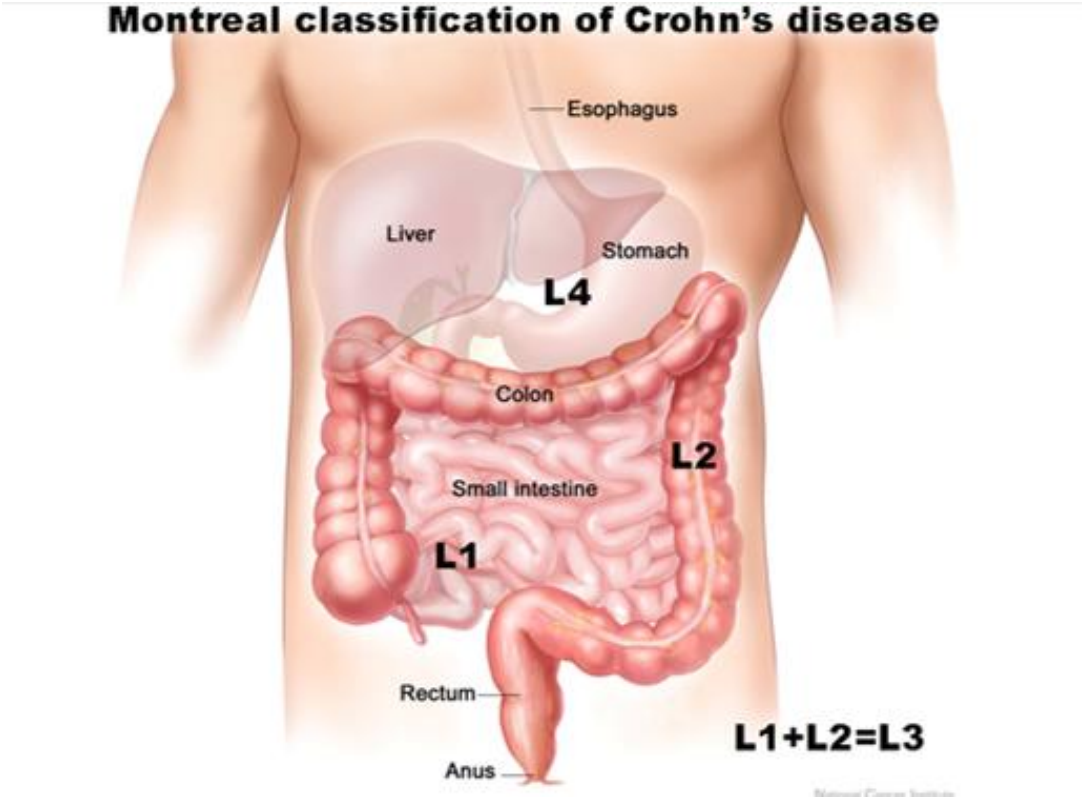



**Figure 3-1: The Montreal classification of Crohn's disease. Adapted from an original image by the National Cancer Institute (http://www.cancer.gov/).**

## *Results*

### *3.1.4 Phenotypic characteristics of childhood onset CD*

*CD at diagnosis* - CD affected the small as well as large bowel (L3+/-L4) in 50.5%, the colon (L2+/-L4) in 36.3% and the ileum (L1+/-L4) in 5.9%. More than half of children (50.9%) were affected by CD proximal to the terminal ileum at diagnosis (any L4). Follow-up of at least 2 years was completed in 196/273 (71.8%). Follow-up data of at least 4 years were available in 132/273 (48.4%).(Table 3-1)

*CD location* changed with time. Detailed analysis of the 196 childhood-onset CD patients who had at least two years follow-up demonstrated a large proportion of these children experienced extension of anatomical involvement of the gastrointestinal tract. At diagnosis, 53/196 (27.0%) already had the maximum disease extent (L3+L4). Of the 143 children who could therefore change CD location at 2-years follow-up, 56 (39.1%) children did. These changes were due to the inclusion of findings on upper GI endoscopy in 15/56 (26.8%) only; in the majority (41/56, 73.2%), changes were due to extension from localised disease to more extensive disease involving the lower GI tract. This extension of anatomical involvement was more likely to be due to increasing ileal involvement over time (27/56 (48.2%) patients (ileal: 23; ileocolonic: 4)) rather than colonic involvement with time (11/56 (19.6%) (p=0.001 OR 3.81 (1.64-8.84)).

*CD behaviour* also changed with time. Disease behaviour progressed significantly from diagnosis to four years follow-up (p=0.001). By two years follow-up, the behaviour of CD changed with an increase in stricturing disease (±perianal) from 4.4% to 9.6% (p=0.02 OR 0.43 (0.20-0.91)) and a decrease in inflammatory disease (±perianal) from 91.2% to 82.7% (p=0.005 OR 2.16 (1.24-3.78)). By 4 years after diagnosis, both stricturing disease (±perianal) and penetrating disease (±perianal) increased from 4.4% to 12.9% (p=0.001 OR 0.31 (0.14-0.67)) and from 4.4 % to 11.4%. (p=0.008 OR 0.36 (0.16-0.79)), respectively. Inflammatory disease

(±perianal) decreased from 91.2% at diagnosis to 75.8%. (p <0.0001 OR 3.32 (1.86-5.92)). Perianal disease increasingly complicated the other behaviour phenotypes over time, rising from 13.9% at diagnosis to 22.2% after 4 years follow-up (p=0.04 OR 0.57 (0.34-0.98)).

**Table 3-1: The phenotype of childhood onset CD at diagnosis and follow-up**

| Childhood onset CD | At diagnosis | At follow-up ≥ 2 years | At follow-up ≥ 4 years |
|---|---|---|---|
| CD phenotype – location* | | | |
| L1 | 10 / 273 (3.7%) | 5 / 196 (2.6%) | 4 / 132 (3.0%) |
| L2 | 57 / 273 (20.9%) | 30 /196 (15.3%) | 24 / 132 (18.2%) |
| L3 | 53 / 273 (19.4%) | 32 / 196 (16.3%) | 20 / 132 (15.2%) |
| L1+L4 | 6 / 273 (2.2%) | 4 / 196 (2.0%) | 3 / 132 (2.3%) |
| L2+L4 | 42 / 273 (15.4%) | 32 / 196 (16.3%) | 24 / 132 (18.2%) |
| L3+L4 | 85 / 273 (31.1%) | 86 / 196 (43.9%) | 51 / 132 (38.6%) |
| L4 | 6 / 273 (2.2%) | 2 / 196 (1.0%) | 2 / 132 (1.5%) |
| ***P*** | | ***0.61*** | |
| CD phenotype – behaviour | | | |
| B1(±p) | 215 / 273 (no p) (78.8%)<br>249 / 273 (± p) (91.2%) | 136 / 197 (69.0%)<br>163 / 197 (82.7%) | 81 / 132 (61.4%)<br>100 / 132 (75.8%) |
| B2(±p) | 9 / 273 (no p) (3.3%)<br>12 / 273 (±p) (4.4%) | 14 / 197 (7.1%)<br>19 / 197 (9.6%) | 10 / 132 (7.6%)<br>17 / 132 (12.9%) |
| B3(±p) | 11 / 273 (no p) (4.0%)<br>12 / 273 (±p) (4.4%) | 14 / 197 (7.1%)<br>15 / 197 (7.6%) | 12 / 132 (9.1%)<br>15 / 132 (11.4%) |
| ***P*** | | ***0.001*** | |

**Table 3-1: Comparison of Montreal disease location and behaviour at diagnosis, at follow-up of at least 2 years and follow-up of at least 4 years in paediatric Crohn's disease patients. *At diagnosis 14 children had oral (7)/ oral+perianal (4)/ perianal (3) disease only. At ≥2 year follow-up 5 children had oral (1) / oral+perianal (4) disease only. At ≥4 year follow-up 4 children had oral (1) / oral+perianal (3) disease.**

### *3.1.5 Phenotypic characteristics of childhood onset UC*

UC phenotype was assessed at diagnosis and at last follow-up (Table 4). Pancolitis/extensive colitis was present in 74.7% at diagnosis. 73/95 of children with childhood onset UC had follow-up data recorded (median duration of follow-up 3.5 years). At last follow-up, the proportions of disease location had not changed significantly.

Of those children with follow-up data available, 49/73 of childhood onset UC already had the maximum extent (E3) at diagnosis. Of the 24 who could possibly increase extent at follow-up, 46% extended their disease location from less extensive to extensive colitis/pancolitis (11/24).

**Table 3-2: The phenotype of childhood onset UC at diagnosis and follow-up**

| UC phenotype | At diagnosis | At last follow-up | *p* |
|---|---|---|---|
| E3 | 71/95 (74.7%) | 60/73 (82.2%) | |
| E2 | 20/95 (21.1%) | 12/73 (16.4%) | *0.39* |
| E1 | 4/95 (4.2%) | 1/73 (1.4%) | |

**Table 3-2: Comparison of Montreal disease location and behaviour at diagnosis and at last follow-up in paediatric ulcerative colitis patients.**

### 3.1.6 Is there a distinct phenotype of childhood IBD diagnosed before 8 years of age?

82/408 (20.1%, 8/416 children had incomplete data) children were diagnosed with IBD before their 8th birthday (age cut-off as previously described by Heyman et al. [541]): 53 CD (64.6%), 19 UC (23.2%) and 10 IBDU (12.2%). In children older than 8 at the time of diagnosis (n=326, 79.9%), CD was diagnosed in 220/326 (67.5%), UC in 78/326 (23.9%) and IBDU in 28/326 (8.6%). At diagnosis in children under 8, IBD was more likely to be localised to the colon (L2+UC+IBDU) than in older children (47/82 (57.3%) vs 145/326 (44.5%), p=0.03 OR 1.68 (1.03-2.73)). CD presenting below the age of 8 was characterised by more isolated colonic disease (L2) (34.0% vs 17.7%, p=0.009 OR 2.39 (1.23-4.64)), less ileal involvement (L1±L4 and L3±L4: 21/53 (39.6%) vs 133/220 (60.5%), p=0.006 OR 0.43 (0.23-0.79)) and less 'panenteric' CD (L3+L4) (18.9% vs 33.6%, p=0.03 OR 0.46 (0.22-0.96)). CD limited to the mouth and/or perianal region occurred more commonly in children younger than 8 years at diagnosis (7/53 (13.2%) vs 6/220 (2.7%), p=0.001 OR 5.43 (1.74-16.90)).

### 3.1.7 Immunomodulator usage in childhood-onset IBD

The time to use of immunomodulatory drug therapy (first use of any of azathioprine, methotrexate or biological therapy) was assessed and presented as Kaplan-Meier curves in the childhood-onset group only (Figure 3.2). These data involve 408 cases, with missing data for the 8 (1.9 %) remaining cases. There was a significant difference between these curves (logrank test p=0.008), with the median time to any immunomodulator usage 17 months for CD, 28 months for UC and 63 months for IBDU. At 120 months of follow-up, the proportion of cases of CD, UC and IBDU who had not been exposed to any immunomodulation were 11.5%, 40.5% and 38.4% respectively.

**Figure 3-2: The time to use of immunomodulatory drug therapy in childhood-onset IBD**

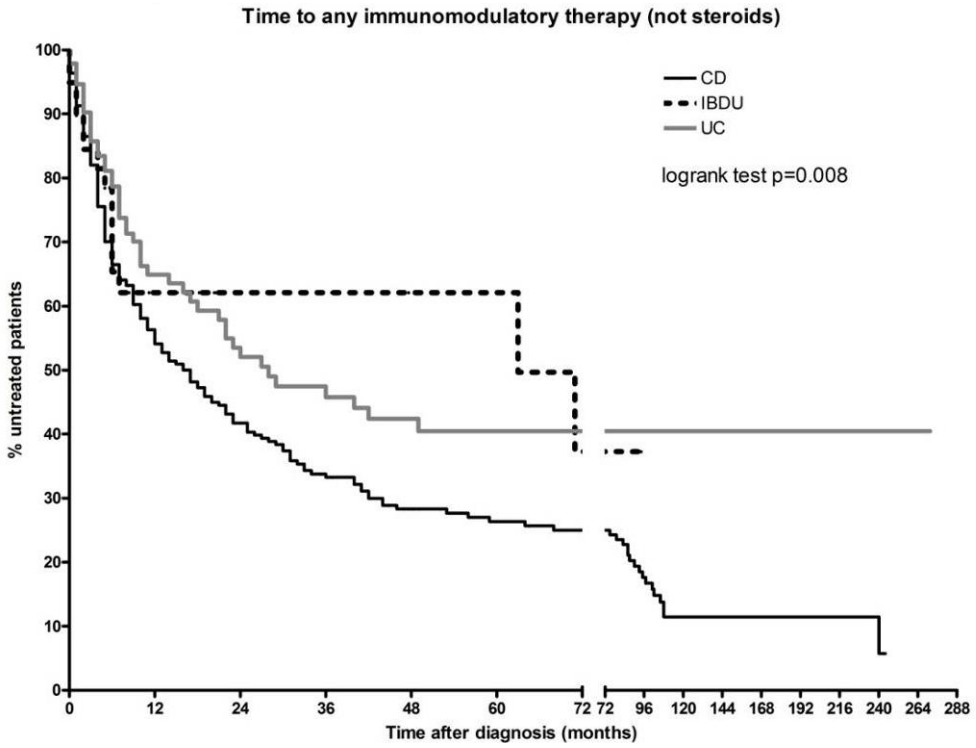



**Figure 3-2: Kaplan-Meier curves showing time from diagnosis to first use of immunomodulatory drug therapy (any of azathioprine, 6-mercaptopurine, methotrexate or biological therapy) in the cohort of childhood-onset IBD patients.**

### *3.1.8 Need for resectional surgery in childhood-onset and adult-onset IBD*

Kaplan-Meier curves assessing the time from diagnosis to first resection/colectomy are shown for childhood-onset and adult-onset CD (data available for n=585) and UC (data available for n=654). (Figures 3.3&3.4)

Significant differences were seen in the Kaplan-Meier curves between the childhood-onset and adult-onset cohorts, in CD and UC (logrank test $p<0.001$ and $p=0.03$, respectively).

17.1% of childhood-onset CD patients required surgery compared with 52.8% of adult-onset CD patients ($p<0.0001$ OR 0.18 (0.13-0.26)). 5 years after diagnosis 20.2% of childhood-onset CD patients had undergone surgery versus 42.8% of adult onset CD patients. By 10 years follow-up these percentages were 34.5% and 55.5%, respectively. The median time to surgery in childhood-onset and adult-onset CD was 13.7 and 7.8 years respectively.(Figure 3.3)

**Figure 3-3: Time from diagnosis to surgery in childhood-onset CD and adult-onset CD.**

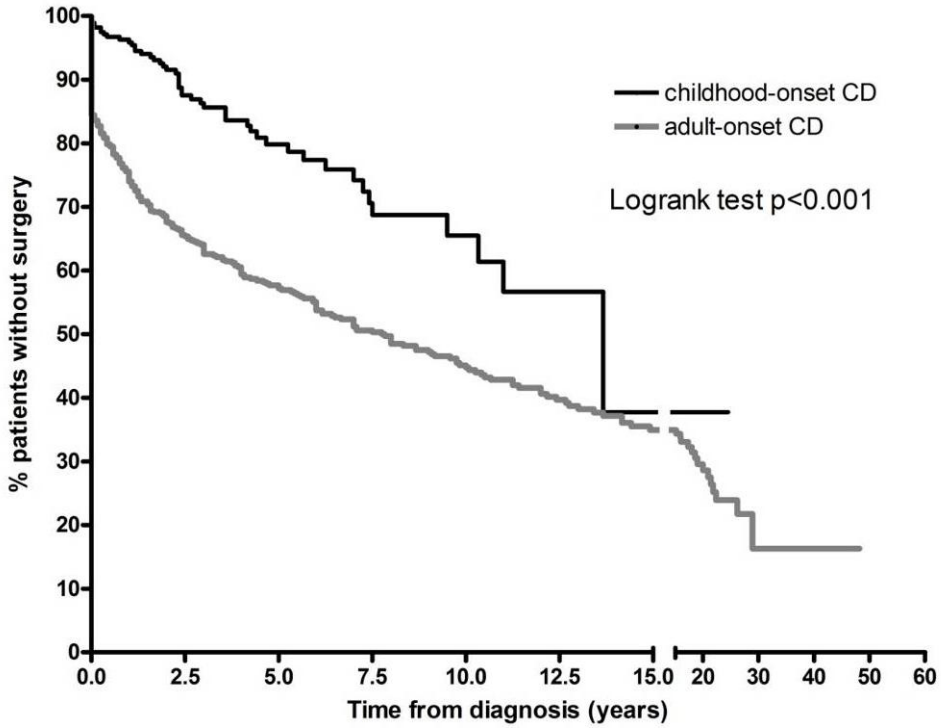



**Figure 3-3: Kaplan-Meier curves showing time from diagnosis to surgery in childhood-onset CD and adult-onset CD.**

20.0% of childhood-onset UC patients underwent colectomy compared with 21.7% of adult-onset UC patients (p=0.70, OR 0.90 (0.53–1.54)). 5 years after diagnosis, 26.1% of childhood-onset UC patients had undergone surgery versus 15.5% of adult-onset UC patients. By 10 years follow-up, 40.9% of childhood-onset UC patients and 19.9% of adult-onset UC patients had had a colectomy. Based on the Kaplan-Meier analysis, the median time to surgery was 11.1 years in childhood-onset UC compared with >50 years in adult-onset UC.(Figure 3.4)

**Figure 3-4: Time from diagnosis to surgery in childhood-onset UC and adult-onset UC.**

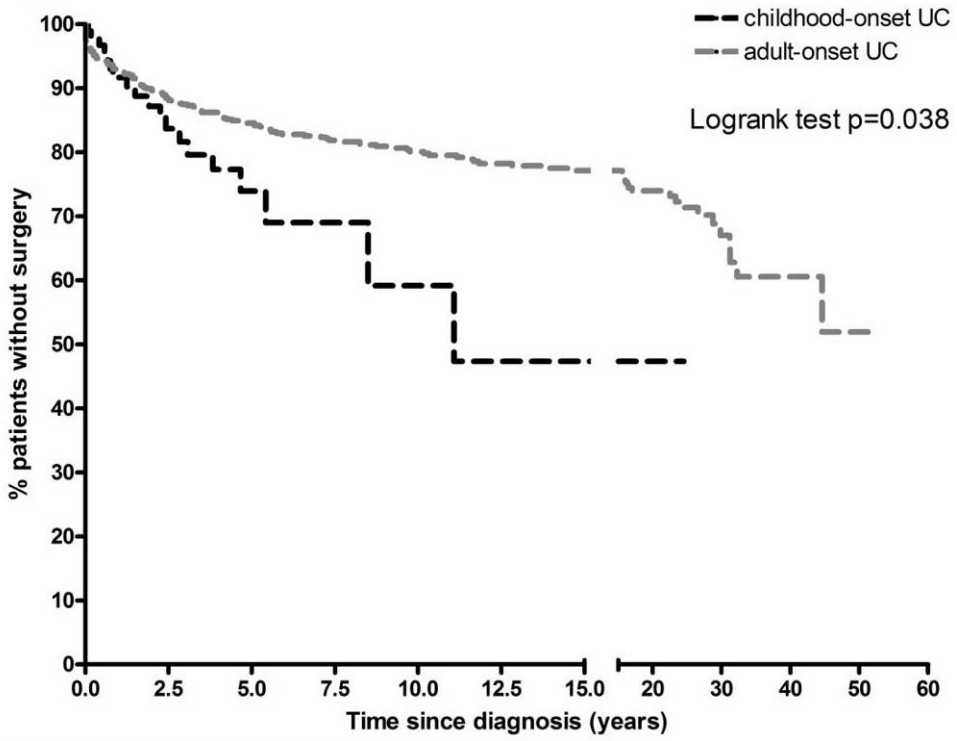



**Figure 3-4: Kaplan-Meier curves showing time from diagnosis to surgery in childhood-onset UC and adult-onset UC.**

### 3.1.9 Is the phenotype of childhood onset IBD different to that of adult onset IBD?

Phenotypic data on 416 children and 1297 adults with IBD were compared at the time of last follow-up (Table 3-3). In childhood-onset CD, there was a clear male predominance compared to adult-onset CD (59.4% vs 36.2%, p<0.0001 OR 2.58 (1.92-3.45)). UC was equally likely to occur in both sexes in childhood-onset and adult-onset disease. A family history of IBD (in any member of the family) was more common in childhood-onset than adult-onset disease (33.9% vs 19.6%, p<0.0001 OR 2.11 (1.65-2.69)).

Location of CD (defined by the Montreal system as maximum extent by the time of last follow-up but before the first resection) showed statistically significant differences between childhood-onset and adult-onset disease (p<0.001). Childhood-onset CD was characterised by less isolated ileal involvement (L1) 2.6% versus 31.5% (p<0.0001 OR 0.06 (0.03-0.12)) and less isolated colonic disease (L2) 15.0% versus 36.1% (p <0.0001 OR 0.31 (0.21 -0.46)). Childhood-onset IBD was characterised by more 'panenteric'/extensive Crohn's disease (L3+L4) 43.2% vs 3.2% (p<0.0001 OR 23.36 (13.45-40.59)).

In order to overcome the confounding factor of different investigation protocols in adult-onset and childhood-onset CD with regards to the use of upper GI endoscopy and small bowel assessment, we also analysed the location of CD controlling for involvement of the GI tract proximal to the terminal ileum (L4) by means of a contingency table containing the categories L1+/- L4, L2 +/- L4, L3 +/-L4 and L4.

The extent of involvement of the GI tract in childhood-onset CD remained statistically different from adult-onset CD both when assessed at last follow-up and when only the patients with a minimum of 5 years follow-up were included (both p<0.001, also see Table 3-4). There is a highly significant association of ileal disease location with stricturing/penetrating disease behaviour and of colonic disease

location with inflammatory disease behaviour when disease behaviour is stratified for disease location in both childhood-onset and adult-onset CD after at least 5 years follow up (Table 3-4; p<0.001).

Comparing disease behaviour at 5 year follow-up, the behaviour of CD was similar in the two groups (p=0.17).

UC in the childhood-onset cohort was more extensive (p<0.001), with extensive colitis (E3) in 82.2% of childhood-onset but only 47.6% of adult-onset UC (p<0.0001, OR 5.08 (2.73-9.45)).

**Table 3-3: Comparison of location and behaviour between childhood onset and adult onset IBD patients using the Montreal classification**

| | Scottish childhood onset | Scottish adult onset |
|---|---|---|
| N | 416 | 1297 |
| CD phenotype – location* | At last follow-up | At last follow-up |
| L1 | 7 / 273 (2.6%) | 160 / 507 (31.5%) |
| L2 | 41 / 273 (15.0%) | 183 / 507 (36.1%) |
| L3 | 52 / 273 (19.0%) | 99 / 507 (19.5%) |
| L1+L4 | 5 / 273 (1.8%) | 21 / 507 (4.1%) |
| L2+L4 | 43 / 273 (15.8%) | 8 / 507 (1.6%) |
| L3+L4 | 118 / 273 (43.2%) | 16 / 507 (3.2%) |
| L4 | 2 / 273 (0.7%) | 13 / 507 (2.6%) |
| ***p*** | ***<0.001*** | |

| CD phenotype – behaviour | At 4 years | At 5 years |
|---|---|---|
| B1(±p) | 99/136 (72.7%) | 249/377 (66.0%) |
| B2(±p) | 20/136 (14.7%) | 54/377 (14.3%) |
| B3(±p) | 17/136 (12.6%) | 74/377 (19.7%) |
| *p* | ***0.17*** | |
| UC phenotype | At last follow-up | At last follow-up |
| E3 | 60/73 (82.2%) | 271/569 (47.6%) |
| E2 | 12/73 (16.4%) | 201/569 (35.3%) |
| E1 | 1/73 (1.4%) | 97/569 (17.0%) |
| ***p*** | ***<0.001*** | |

**Table 3-3: Comparison of location and behaviour between childhood onset and adult onset IBD patients using the Montreal classification: * 5 children had oral (1)/ oral+perianal disease (4). 7 adults had oral (4)/ oral+perianal (1)/ perianal (2) disease.**

**Table 3-4: CD phenotype – behaviour stratified for disease location**

| | Childhood-onset CD | | | Adult-onset CD | | |
|---|---|---|---|---|---|---|
| | B1(+/-p) | B2(+/-p) | B3(+/-p) | B1(+/-p) | B2(+/-p) | B3(+/-p) |
| L1 +/- L4 | 1/95 | 2/20 | 4/17 | 63/246 | 35/54 | 40/74 |
| L2 +/- L4 | 36/95 | 7/20 | 6/17 | 122/246 | 4/54 | 9/74 |
| L3 +/- L4 | 57/95 | 10/20 | 7/17 | 58/246 | 11/54 | 24/74 |
| L4 | 1/95 | 1/20 | 0/17 | 3/246 | 4/54 | 1/74 |
| p | ^ | | | *<0.001* | | |
| CD phenotype – behaviour stratified for disease location | | | | | | |

| (childhood-onset and adult-onset CD) | | | |
|---|---|---|---|
| | B1(+/-p)<br>(n=341) | B2(+/-p)<br>(n=74) | B3(+/-p)<br>(n=91) |
| L1 +/- L4 | **64**<br>*(98)* | **37**<br>*(21)* | **44**<br>*(26)* |
| L2 +/- L4 | **158**<br>*(124)* | **11**<br>*(27)* | **15**<br>*(33)* |
| L3 +/- L4 | **115**<br>*(113)* | **21**<br>*(24)* | **31**<br>*(30)* |
| L4 | **4**<br>*(7)* | **5**<br>*(1)* | **1**<br>*(2)* |
| p | *<0.001* | | |

**Table 3-4: Comparison of childhood-onset CD with adult-onset CD. CD location assigned at last follow-up (at least 5 years). Analysis of behaviour after 5 years stratified for disease location is presented. ^: excluded L4 category for 3x3 $\chi 2$ analysis due to low numbers: $\chi 2$ – statistic 15.65 (due to small numbers, calculation of p-value not reliable). In the analysis of behaviour stratified for disease location, the sum of childhood-onset and adult-onset CD patients is given in bold. The number between brackets in italics is the expected number. The highly significant p-value is due to the association of ileal disease location with stricturing/penetrating disease behaviour association whereas colonic disease location is associated with inflammatory disease behaviour.**

### *Discussion*

The present study represents not only a detailed application of the Montreal classification of IBD in a large cohort of patients with childhood-onset disease, but also provides new robust data on disease evolution, as well as need for immunomodulation and surgery in childhood-onset disease. Rigorous follow-up of 416 children with IBD in the three paediatric gastroenterology centres in Scotland has allowed us to document these aspects, notably the progression of disease location and behaviour. Data were collected by structured case notes review, and measures taken to overcome inter-observer bias.

A number of our findings are likely to be important and pertinent both to basic research, as well as to clinical practice, and may be useful in patient counselling. We have demonstrated that the presenting phenotype of childhood-onset IBD was characterised by extensive anatomical involvement, with strikingly high rates of 'panenteric' CD (small bowel, large bowel and upper GI tract involvement) as well as extensive UC. Furthermore, disease extent was remarkably dynamic in childhood-onset disease: even within 2 years of diagnosis, childhood-onset CD progressed to more extensive anatomical involvement in greater than 1 in 3 patients. Disease behaviour rapidly progressed to complications of stricture formation or the development of fistulae. Disease extent in 46% of the few cases of childhood-onset UC also showed extension within the period of follow-up.

We have provided compelling evidence for the high prevalence of an extensive or panenteric disease phenotype at presentation in childhood-onset CD. Recent data from the North American Pediatric IBD Consortium Registry are consistent with our findings, albeit using a distinct disease classification system.[543] Gupta et al. presented their findings in 600 paediatric CD patients at diagnosis: 61.5% of children had small bowel and colon involvement (compared with 50.5% in our study). Baldassano et al.

recently described a small cohort of 142 children with CD: 86% had involvement of the small bowel and colon (± upper GI tract).[547] Cucchiara et al. showed in an Italian cohort of paediatric CD patients with at least 1 year follow-up (n=200), that 58% had ileocolonic CD (using the Vienna classification).[548]

Our data in UC are also consistent with the hypothesis that childhood-onset disease has a very high bias towards extensive disease. This observation is also consistent with other datasets. Previously reported rates of pancolitis in patients with UC, from prospective childhood studies (80-90%) [541;549;550] are far higher than those from prospective adult studies (24% pancolitis[551] or 33% extensive colitis [552]). Our study not only provides further replication of these data but also directly compares a childhood-onset and adult-onset cohort within the same population using the Montreal classification (82% of childhood-onset UC patients had extensive colitis at last follow-up compared with 47% of adult-onset UC).

Moreover, we provide evidence that disease phenotypic characteristics are dynamic, and changeable, rather than stable, in childhood-onset disease. Detailed analysis of our patients with childhood-onset CD demonstrated that the anatomical extent of disease progressed to more extensive involvement soon after diagnosis in 39%. In the vast majority, these changes were not due to the inclusion of findings on upper GI endoscopy but rather to progression from limited disease (oral/perianal, L1, L2 or L4) to involve both small and large bowel (L3). In almost 50% of patients who progressed in anatomical involvement, this was due to involvement of the ileum. The data contrast with the stability of disease location repeatedly reported in adult CD.[553;554] Louis et al. followed 125 adult patients with CD and found at 10 years following diagnosis only 15% had changed disease location (location was defined using the Vienna classification).[553] Henriksen et al. observed a change of (Vienna) CD location in 13.5% after 5 years follow-up (n=200).[554]

In our CD cohort, our data intriguingly suggest heterogeneity even within childhood-onset phenotype. We have shown that involvement of the ileum is age-dependent: children under 8 years at time of diagnosis had significantly less involvement of the ileum and more isolated colonic disease than children older than 8 years at diagnosis, thus confirming previous studies.[541;542;555] In a large study of nearly 1400 North American early onset patients, Heyman et al demonstrated by multifactorial analysis that a colonic predominant phenotype exists in IBD diagnosed under the age of 8 years.[541] Paul et al. studied 413 paediatric IBD patients and also demonstrated a greater tendency for very young patients to present with colonic disease.[542]

In addition to the dynamic nature of disease location, we have also demonstrated that childhood-onset CD behaviour was not stable over time: inflammatory disease behaviour progressed with the development of stricturing/intestinal penetrating complications. This progression is also seen in adult-onset disease.(see Table 5) In the landmark study of disease behaviour in 2000 CD patients, Cosnes and colleagues demonstrated that 40% had penetrating disease as defined by the Vienna criteria at 5 years and 70% by 20 years.[8] Similar data from Belgium and Scotland were reported subsequently.[556;557] Detailed analysis of disease behaviour stratified for disease location demonstrated a highly significant association of stricturing/penetrating disease complications with ileal CD and of inflammatory disease with colonic CD (both in childhood-onset and adult-onset CD; Table 3-4).

Direct comparison of the phenotypic characteristics in adult-onset and childhood-onset disease also emphasised the extensive intestinal involvement in children, although these comparisons are limited by the retrospective nature of the present study. The increased intensity of investigation in childhood-onset disease compared with adult-onset IBD is noteworthy.[526;545;558] This line of argument has been previously suggested to underlie the high prevalence of upper GI disease reported in children, when compared with adults – upper GI endoscopy is rarely performed in adult IBD assessments and historic (adult) datasets typically have less small bowel assessment. In the absence of clear upper GI-related symptoms,

oesophagogastroduodenoscopy (OGD) is rarely performed in the investigation of adult-onset CD.[558;559] However, involvement of the upper GI tract has been reported to be as high as 75% in prospective adult CD studies where OGD was routinely performed.[560;561]

Notwithstanding this issue, which is difficult to resolve without a longitudinal prospective study both in children and adults, analysis based on the Montreal categories which treats upper GI disease as a modifier (ileal, colonic or ileocolonic disease each with or without CD proximal to the terminal ileum), confirmed the statistically significant differences between CD location in childhood-onset and adult-onset onset CD.

In an attempt to further define the "severity" of childhood-onset IBD, we evaluated the need for both immunomodulation and surgery (Figures 3.1-3). By 12 months from diagnosis, 45.9 % of childhood-onset CD, 37.9% of IC and 35.1% of UC have commenced immunomodulator therapy. However, interpretation of these data as a surrogate for severity is problematic, and confounded by multiple factors, most notably the variability amongst individual physicians in thresholds for the use of these agents. Physician preferences have undoubtedly altered in the last decade, both for adults and in children. Increasing use of immunomodulatory therapy early in the course of disease has become a well-established treatment paradigm in paediatric IBD practice, following the landmark publication by Markowitz et al. [562]; in adults, data from our own centre clearly illustrate similar temporal trends towards early use - significantly higher rates of 6MP/Azathioprine usage of 47.9% at 12 months were seen in an inception cohort diagnosed between 2003-7 compared to a rate of 13.3% in a cohort diagnosed in 1998-2002.[563]

Many may regard the need for surgery after diagnosis as a potential "gold-standard" marker of disease severity. For this reason, we have analysed need for surgical intervention in our cohort. Intriguingly, these data suggest that surgical intervention may occur earlier in childhood-onset UC than in adult-onset UC: but that the opposite relationship is seen for CD. However, there are multiple possible

confounding variables in using these data as a surrogate for severity, especially when comparing practice for children and teenagers with CD against practice in adults. The more extensive panenteric disease of childhood-onset CD may preclude referral for early surgery, as do the frequently strong reservations of children and their families concerning early surgical intervention (especially if formation of a stoma is required). We suggest that these factors impact significantly on the data in childhood-onset CD, and lead to the escalation of medical therapy rather than surgery, notwithstanding disease severity. In UC, the higher surgical rates we demonstrate in childhood-onset disease are interesting, in the context of the high prevalence of extensive colonic involvement in these children, and may well reflect "severity" more accurately. Notwithstanding these controversies, the different trends in time to first surgery in childhood-onset CD and UC are of interest and require further evaluation.

The findings of our study may have clinical implications for management in childhood-onset IBD. It might be considered that early use of biological agents and immunosuppressants may ideally be targeted preferentially on children with aggressive disease. However, it is increasingly apparent that other factors need also be considered in highlighting either all or a sub-group of childhood-onset IBD as the group for whom "top-down" intervention is appropriate - most notably the recent concerns regarding the development of hepatosplenic T cell lymphoma in patients treated with combined immunomodulation, together with the lack of any evidence-based strategies for discontinuation of anti-TNF therapies, or even thiopurines.[564]

Thus, we propose that the initial application of our data may be in the translational research setting. Identifying specific laboratory markers of disease susceptibility or phenotype - genetic, serological or proteomic - critically depends on the availability of an appropriate classification system. Our study has applied the Montreal classification in a large cohort of childhood-onset IBD patients, and demonstrates the potential usefulness of this system in studies of pediatric IBD.

# 4 Environmental risk factors for childhood-onset IBD in Scotland: a case-control study

## *Introduction*

### *4.1.1 The incidence of IBD in Scotland*

The increasing incidence of IBD poses a real public health problem.[565] In Scotland, incidence of both CD and UC in patients aged ≤16 rose by 30% between 1981 and 1995. [4] According to a prospective study, Scotland now has the highest incidence of IBD in patients <16 in the British Isles, at 6.5/100 000. [5] This is roughly twice the incidence reported in 1983.[3] Paediatric CD (4.2/100 000) predominates over UC (1.8/100 000) in Scotland as in most parts of the world.[2] A study of juvenile-onset CD in Scotland found increased incidence in northern as compared to southern Scotland and an association with increased affluence as determined by the postcode-linked Carstairs Deprivation Category (DepCat) score (replaced in 2004 with the Scottish Index of Multiple Deprivation (http://www.scotland.gov.uk/Topics/Statistics/SIMD/). (Figure 4-1)[566]

Numerous studies from across Western Europe have published a rapid increase in incidence of IBD in children, this trend being strongest for childhood CD. [567-569] The north-south gradient described for CD incidence in Scotland has also been observed in France.[570]

However, getting a true estimate of changing childhood incidence is problematic since definitions of "childhood" vary between studies. The increasing incidence reported could reflect increased awareness among health professionals. Nevertheless, evidence for increasing incidence among children is compelling. Most of the available incidence and prevalence data in childhood IBD is around a decade old, so updated figures for adults and children are eagerly awaited. [571]

Although recent advances in elucidating the genetic factors predisposing to IBD have led to great optimism, the increasing incidence of IBD during the last 4 decades is clearly due to environmental factors or their epigenetic effects.[3;4;566;572] The study of environmental factors associated with inflammatory bowel disease has led to a great deal of confusion/frustration among researchers and the public alike, as illustrated most poignantly by the debate surrounding the Measles/Mumps/Rubella vaccine.[573;574]

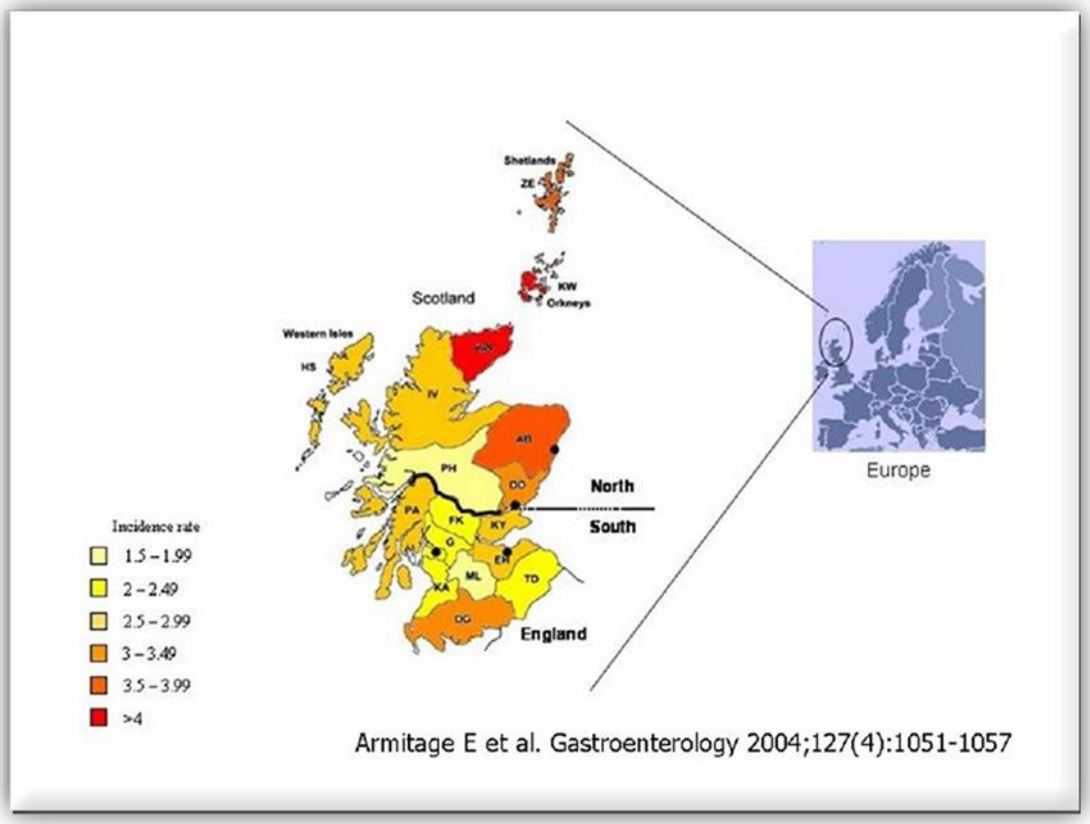



**Figure 4-1: The rising incidence of childhood onset Crohn's disease in Scotland: association with northern latitude and affluence.**

Figure 4-1: The rising incidence of childhood onset Crohn's disease in Scotland: association with northern latitude and affluence (adapted from Armitage et al. Gastroenterology 2004)

### *4.1.2 Aetiology of IBD: nature vs nurture*

Strong evidence for the genetic contribution to IBD susceptibility came from twin studies conducted in the UK, Sweden and Denmark. Combined results give a concordance rate for CD of 36% in monozygotic twins[575-577] compared with 4% in dizygotic twins. Equivalent results for UC were 16% and 4%, thus suggesting that the genetic contribution is stronger in CD than in UC. [578] Were IBD exclusively genetic in aetiology, we would expect concordance between monozygotic twins to approach 100%. The rise in CD in the past century and the marked geographical variation in incidence of IBD cannot be explained by genetics alone.[574]

In utero – and early childhood events influence health and disease in later life. Obesity, coronary heart disease, and diabetes are but a few of the conditions for which an increasing evidence base exists relating to their programming early in development.[579] As the first exposure to colonising micro-organisms and enteropathogenic bacteria occurs post-natally, the development of the intestinal immune system is likely to be influenced/(epigenetically) modified during these first contacts through e.g. breastfeeding. Additional environmental factors can also be crucial during childhood with clear relevance to development of IBD later in life, e.g. passive smoking, appendicectomy/appendicitis and the presence of atopic disease.

Oral contraceptive medication (OCP) has clearly been shown recently to influence the subsequent development of both CD and UC.[580] The pooled relative risk (RR) for CD for women currently taking the OCP was 1.51 (95% confidence interval [CI] 1.17–1.96, P = 0.002), and 1.46 (95% CI 1.26–1.70, P <0.001), adjusted for smoking. The RR for UC in women currently taking the OCP was 1.53 (95% CI 1.21–1.94, P = 0.001), and 1.28 (95% CI 1.06–1.54, P = 0.011), adjusted for smoking. The RR for CD increased with the length of exposure to OCP. In our exclusively paediatric cohort and control group, we were unable to study these effects adequately and have therefore opted to focus on passive smoking, appendicectomy, atopic disease and breastfeeding.

#### 4.1.2.1 Passive smoking

In adults, smoking significantly protects against UC and improves its course.[581] This also manifests as a slightly elevated risk in the first year following smoking cessation.[581] The mechanism of action remains unclear. A meta-analysis suggests that ex-smokers are 70% more likely than those who never smoked to develop UC. In contrast, smoking predisposes to CD being associated with the development and increased severity of CD.[581;582] Bernstein et al suggests that smoking represents an important contributory influence on established disease but is not a risk factor.[583]

Childhood exposure to tobacco smoke may influence the eventual development of IBD, especially CD, where a few studies have shown childhood passive smoking to increase risk of IBD. [83;584-586]

For UC, the relationship is more controversial; some studies show passive smoking as a risk factor [83], others indicate it has no effect [81;585;587] and one other shows passive smoking to have a protective influence.[82] Few studies examine passive smoking as a risk factor for *paediatric* IBD specifically. Those that do, present opposing results.[81;83;84] A recent meta-analysis by Jones et al. showed that there was no clear relationship between childhood passive smoke exposure and the later development of IBD, either CD or UC, with an important caveat of drawing conclusions from the small number of heterogeneous studies performed to date.[588]

If anything, these studies have illustrated the importance of investigating tobacco smoke exposure during different periods, from prenatal to time of IBD diagnosis. This could help elucidate which type(s) of exposure are most significant to the aetiology of paediatric IBD.

#### 4.1.2.2 Appendicectomy

Appendicectomy has been identified as reducing risk of UC. Gilat et al first provided evidence for this in a 1987 multicentre study[589]. The association has since been confirmed in case-control studies from many different countries.[590] [581;591;592] Several studies report that the reduction in risk is greatest if appendicectomy is performed

before the age of 20 years.[590;591;593] Additionally, Kurina et al noted the association was strongest for those undergoing appendicectomy between 10 and 14 years of age.[590] This group suggested that it is the appendicitis rather than the appendicectomy that confers protection.[590] Frisch and colleagues provided robust evidence for the protective effect of appendicectomy associated with appendicitis or mesenteric adenitis.[594] They studied a large cohort of 709 353 Swedish (1964-2004) and Danish (1977-2004) patients who had undergone appendicectomy and who were followed up for subsequent ulcerative colitis. The impact of appendicectomy on risk was also studied in 224 483 people whose parents or siblings had inflammatory bowel disease. During 11.1 million years of follow-up in their appendicectomy cohort, 1192 patients developed ulcerative colitis (10.8 per 100 000 person years). Appendicectomy without underlying inflammation was not associated with reduced risk (standardised incidence ratio 1.04, 95% confidence interval 0.95 to 1.15). Before the age of 20, however, appendicectomy for appendicitis (0.45, 0.39 to 0.53) or mesenteric lymphadenitis (0.65, 0.46 to 0.90) was associated with significant risk reduction. A similar pattern was seen in those with affected relatives, whose overall risk of ulcerative colitis was clearly higher than the background risk (1404 observed v 446 expected; standardised incidence ratio 3.15, 2.99 to 3.32). In this cohort, appendicectomy without underlying appendicitis did not modify risk (rate ratio 1.04, 0.66 to 1.55, v no appendicectomy), while risk after appendicectomy for appendicitis was halved (0.49, 0.31 to 0.74).[594]

In Crohn's disease, the evidence for an association with appendicectomy appears less straightforward. However, the effect of underlying inflammation as the precipitating event to undertake appendicectomy, is also increasingly recognised. Radford-Smith et al. found that appendicectomy delays the onset and mitigates the clinical course of both UC and CD.[592] Kaplan and colleagues reported the findings of a large study of 709 353 appendectomy patients in Sweden and Denmark: they concluded that the increased risk of CD after appendicectomy was likely due to diagnostic bias.[595] No effect was observed when appendicectomy occurred before the age of 10. Long-term increased CD risk up to 20 years after the appendectomy was seen only in appendectomy patients without appendicitis or mesenteric lymphadenitis.[595]

#### 4.1.2.3 Atopic disease

In adults, IBD has been linked with other immune-mediated conditions such as asthma.[596] Incidence of IBD and other immune-mediated conditions increased significantly over the last 50 years while infectious disease like rheumatic fever, hepatitis A and tuberculosis decreased.[597] (Figure 4-2)

In this context, it is important to consider the "hygiene hypothesis" and the gut-equivalent of this theory: "the old friends hypothesis".[574;598] Children who are less exposed to (low grade) infectious organisms (including helminths) may be more susceptible to immune-mediated disorders. In the recent KOALA-study, Penders et al. showed that differences in gut microbiota (eg *E. Coli* and *C. Difficile*) composition precede the development of atopy.[78;599] In this model, it is important to note that socioeconomic status, passive smoking and breastfeeding are confounding factors.

Putative links between IBD and tonsillectomy[600], measles infection[601], measles vaccination and early gastroenteritis[602] have all been challenged.

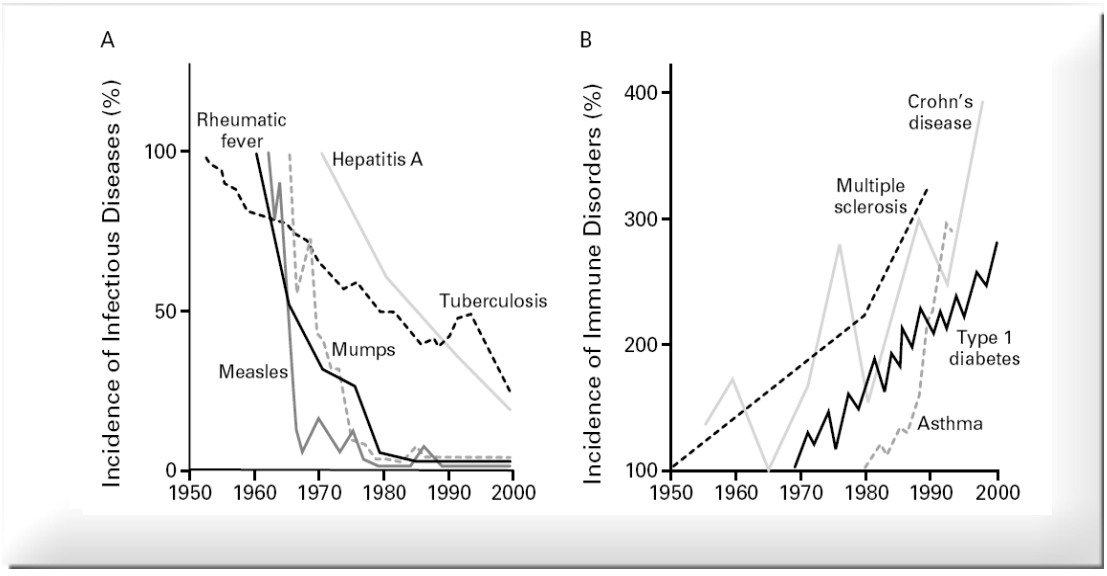



**Figure 4-2: The decreasing incidence of infectious diseases over the last 5 decades is accompanied by an increase in immune-mediated disorders.**

**Figure 4-2: The decreasing incidence of infectious diseases over the last 5 decades is accompanied by an increase in immune-mediated disorders. Adapted from Bach J-F. N Engl J Med 2002;347(12).**

#### 4.1.2.4 Breastfeeding

The effect of breastfeeding on the subsequent susceptibility to paediatric IBD is controversial. There are few exclusively paediatric studies. Infant feeding practices may play a role in the development of IBD by affecting the early exposure to dietary antigens, through alteration of the enteric microflora and thus altered development of the enteral mucosa-associated lymphoid system, as discussed in Chapter 1.[574]

Breastfeeding is shown as a protective factor in most studies, particularly in relation to CD [79;603-605]. The large multicenter study by Gilat et al. found no significant difference between IBD patients and control subjects in the frequency of breastfeeding.[589] One study has suggested breastfeeding may be associated with an increased risk of developing IBD.[81]

### *Aims*

Using data from publicly available databases and our own data obtained via face-to-face interview of young IBD patients and their families and matched control data obtained via postal questionnaire, this study aims to investigate the relationship between childhood IBD and atopic disease and to assess the effect of putative environmental risk factors (such as breastfeeding, immunisations, active and passing smoking exposure), surgical and medical history and family history of bowel cancer or IBD on susceptibility to paediatric IBD in Scotland.

## *Methods*

### *4.1.3 Cases*

The cohort of paediatric IBD patients as described in detail in Chapter 3 was used. The following data had been obtained during face-to-face interview: ethnicity, postcode, sex, age at symptom onset and disease diagnosis, personal and parental smoking, breastfeeding, vaccination history, other medical problems (including asthma, eczema, hayfever/allergic rhinitis and food allergy) and family history (IBD, bowel cancer, coeliac disease and autism).

### *4.1.4 Controls*

The initial study design was to obtain controls, defined as a child of the same age (± 1 year), postcode (as a proxy for affluence as determined by the DepCat score) and sex as the matched case who does not have IBD at a defined point in time. GPs of cases would be contacted and asked to identify two controls per case (due to expected response rate of 50-60%). Data collection would be via a questionnaire asking for detailed environmental information, similar to the one used for IBD patients (see Appendices). We devised age-appropriate cover letters and consent forms for controls, based on readability indices for children older and younger than 13 years. To ensure the study complied with the Data Protection Act, we planned to ask GPs to forward questionnaires to controls.

Due to a delay in obtaining full ethical approval, we consulted publicly available population data.

From the *Scottish Health Survey 2003,* we gained data for:

- Doctor-diagnosed asthma prevalence in children aged 2-15 from across the whole of Scotland.

The *Information Services Division* (ISD) (NHS Scotland) provided population data for:

- Breastfeeding prevalence for children born in 1996 from Lothian and Greater Glasgow
- Immunisation prevalence at 24 months for children born Jan-Dec 1993 in Lothian, Greater Glasgow and Grampian.

After obtaining ethical approval via a substantial amendment to our project (LREC/2002/6/18), the study as planned originally ran from 3/08 to 7/08. We hoped to achieve one population control for each case therefore only the first reply was used for matching. The responses were recorded and stored in a specially designed Microsoft Access database (designed by Richard K Russell).
A previous project conducted by Richard K Russell and Rana Farhadi had recruited 62 matched controls (55% matching rate), using similar methodology in SE Scotland alone in August-October 2004 (Rana Farhadi, SSC4 project). There were 308 cases remaining whose GPs had not been yet been contacted. Of these 308 cases identified, 193 GPs were willing and able to select controls (63% participation rate). Therefore 193 controls were identified with 64 unique matches responding in time to be included in the analysis (33% response rate). Thus, the second, Scotland-wide, phase of this epidemiology project achieved a 21% matching rate (64/308). Combined with the data obtained by Farhadi and Russell, this gives an overall matching rate of 30% (126/416).

### *4.1.5 Matched cases and controls – demographics*

Of the 126 pairs of matched IBD cases and controls, median age at IBD diagnosis was 9.8 years (Q1-Q3: 7.5-11.7), representing an earlier onset sample compared to our complete database of childhood onset IBD patients (Table 4-1). Genders were matched between cases and controls. Except one control of Pakistani ethnicity, all were Caucasian.

**Table 4-1: Median Age at diagnosis and ethnicity of all IBD patients in database**

| | **CD** | **UC** | **IC** |
|---|---|---|---|
| **Median Age at diagnosis** (years) (Q1-Q3) | 11.5 (8.9-13.2) | 10.9 (8.8-10.8) | 10 (7.6-12.5) |
| **Caucasian** | 97.8 % (270/276) | 94.9% (94/99) | 97.5% (40/41) |

**Table 4-2: Gender of matched IBD cases**

| | **Male** | **Female** |
|---|---|---|
| **IBD** | 67 | 57 |
| **CD** | 46 | 35 |
| **UC** | 15 | 15 |

### *4.1.6 Statistics*

Using Pearson's chi-squared test, population data were compared with patient data to assess the association between IBD and different environmental exposures. Uncorrected p-values, odds ratios (OR) and 95% Confidence Intervals (CI) are given. Due to interdependence of the assessed variables, a stringent Bonferroni correction for multiple testing would be too conservative. For borderline associations, we commented on the biological/statistical reliability of the association compared with our other findings. We analysed data for IBD and for UC and CD separately.

### *Results*

In sections 4.5.1-4.5.3, the results of our population-matched case-control analysis will be presented. The data on the subsequent sex, age and postcode-matched case-control analysis on 126 childhood onset IBD patients, will be presented from section 4.5.4 onwards.

## *4.1.7 Population-matched case-control analysis: Asthma*

In the control group (taken from the *Scottish Health Survey 2003*, in which children aged 2-15 throughout Scotland were surveyed), 14% of girls (203/1449) and 21% of boys (318/1516) had doctor-diagnosed asthma. The total weighted percentage of Scottish children aged 2-15 with doctor-diagnosed asthma is thus 17.6% (521/2965).(Table 4-3)

Asthma is significantly associated with paediatric IBD in the Scottish population ($p<0.0001$ OR 1.67 (1.32-2.10). This association is driven exclusively by the high prevalence of self-reported asthma in CD patients ($p=0.0002$ OR 1.74 (1.31-2.30)).

**Table 4-3: Population-matched case-control analysis: asthma prevalence in IBD**

**Table 4-3: Number of asthma sufferers and percentage prevalence: cases versus population-matched controls**

| **Asthma** | **Number of asthma sufferers** | **p** | **OR (95% CI)** |
|---|---|---|---|
| Population data | 521/2965 (17.6%) | <0.0001 | 1.67 (1.32-2.10) |
| All IBD patients | 115/439 (26.2%) | | |
| CD patients | 74/274 (27.0%) | 0.0002 | 1.74 (1.31-2.30) |
| UC patients | 22/101 (21.8%) | 0.3383 | 1.31 (0.81-2.12) |

### *4.1.8 Population-matched case-control analysis: Immunisations history*

After correction for multiple testing, there were no significant differences between cases and controls regarding vaccination against Diptheria, Tetanus, Pertussis, HIb3 or MMR. Results for IBD overall and stratified for CD and UC are given (Table 4-4).

| **Diphtheria** | | | |
|---|---|---|---|
| | **Number of children immunised** | **p** | **OR (95% CI)** |
| Population data | 21946/22545 (97.3%) | 0.4626 | 1.37 (0.68-2.78) |
| All IBD patients | 403/411 (98.1%) | | |
| CD patients | 260/268 (97.0%) | 0.6011 | 0.89 (0.44 -1.80) |
| UC patients | 95/95 (100%) | 0.1971 | |
| **HIb3** | | | |
| Population data | 21876/22545 (97.0%) | 0.2583 | 0.78 (0.47-1.29) |
| All IBD patients | 406/422 (96.2%) | | |
| CD patients | 250/264 (94.7%) | 0.0166 | 0.55 (0.32-0.94) |
| UC patients | 93/95 (97.9%) | 0.8483 | 1.42 (0.35-5.78) |
| **MMR** | | | |
| Population data | 21159/22545 (93.9%) | 0.4827 | 1.19 (0.77-1.84) |
| All IBD patients | 401/423 (94.8%) | | |
| CD patients | 249/264 (94.3%) | 0.8536 | 1.09 (0.64-1.84) |
| UC patients | 91/97 (93.8%) | 0.8202 | 0.99 (0.43-2.27) |

| **Tetanus** | | | |
|---|---|---|---|
| | **Number of children immunised** | **p** | **OR (95% CI)** |
| Population data | 21948/22545 (97.4%) | 0.4687 | 1.37 (0.68-2.77) |
| All IBD patients | 403/411 (98.1%) | | |
| CD patients | 260/268 (97.0%) | 0.5943 | 0.88 (0.44-1.79) |
| UC patients | 95/95 (100%) | 0.1982 | - |

| **Pertussis** | | | |
|---|---|---|---|
| | **Number of children immunised** | **p** | **OR (95% CI)** |
| Population data | 21474/22545 (95.2%) | 0.0208 | 2.23 (1.15-4.43) |
| All IBD patients | 402/411 (97.8%) | | |
| CD patients | 260/268 (97.0%) | 0.2268 | 1.62 (0.80-3.28) |
| UC patients | 95/95 (100%) | 0.0531 | - |

**Table 4-4: Population-matched case-control analysis: Immunisations history**

**Table 4-4: Population-matched case-control analysis: Immunisations history**

### *4.1.9 Population-matched case-control analysis: Breastfeeding*

Breastfeeding is defined as exclusive breastfeeding or mixed breast- and bottle-feeding.

Breastfeeding data from cases was included only where breastfeeding duration was recorded. Breastfeeding was defined as having been breastfed for a) 1 week or more and b) 2 weeks or more. This provides the most accurate comparison with controls, for which breastfeeding was recorded at the health visitor's visit at 10 days. Control data were available for children born in 1996.

Results were analysed for all IBD and stratified for CD and UC. Results were stratified by deprivation using the Carstairs Deprivation Score. This score uses the following variables to calculate deprivation category by postcode: overcrowding in the home, male unemployment, low social class (by type of occupation) and 'no car'.

For IBD, CD and UC, there were no significant differences between cases and controls regarding breastfeeding (Tables 4-5 and 4-6). This was true for breastfeeding lasting 1 week or more and for breastfeeding lasting 2 weeks or more. No significant differences were found when results were stratified by deprivation or geographical region.

Analysis of controls showed marked differences in breastfeeding rates (for more than 1 week) between Lothian and Greater Glasgow (4368/8196 (53.3%) vs. 2270/6775 (33.5%) $p<10^{-4}$ OR 2.26, 95%CI 2.12-2.42)

Breastfeeding rates in IBD cases were significantly different between DepCat 1-2 and DepCat 6-7 (53/79 (67.1%) vs. 17/67 (25.4%), p<10-4 OR 6.00 CI (2.91-12.36)). This difference was also present in controls (4226/6648 (63.6%) vs. 1768/7750 (22.8%) $p<10^{-4}$ OR 5.90 CI 5.49-6.35).

**Table 4-5: Breastfeeding for 1 week or more: IBD cases compared with controls**

## Table 4-5: Breastfeeding for 1 week or more: IBD cases compared with population-matched controls

| | **Cases** | **Controls** | | |
|---|---|---|---|---|
| | Number breastfed | Number breastfed | **p** | **OR**<br>**95% CI** |
| **Lothian** | 106/218<br>(48.6%) | 4368/8196<br>(53.3%) | 0.15 | 0.83<br>(0.63-1.09) |
| **Greater Glasgow** | 68/169 (40.2%) | 2270/6775<br>(33.5%) | 0.08 | 1.34<br>0.98-1.82 |

**Table 4-6: Breastfeeding at 2 weeks: IBD cases compared with controls.**

## Table 4-6: Breastfeeding at 2 weeks: IBD cases compared with population-matched controls

| | **Cases** | **Controls** | | |
|---|---|---|---|---|
| | Number breastfed | Number breastfed | **p** | **OR**<br>**95% CI** |
| **Lothian** | 105/218<br>(48.2%) | 4368/8196<br>(53.3%) | 0.11 | 0.81<br>(0.62-1.07) |
| **Greater Glasgow** | 61/169<br>(36.1%) | 2270/6775<br>(33.5%) | 0.53 | 1.12<br>(0.82- 1.54) |

**Table 4-7: Breastfeeding for one week or more by diagnosis: cases compared with population-matched controls.**

| | **Lothian** | | | |
|---|---|---|---|---|
| | **Cases** | **Controls** | | |
| | Number breastfed | Number breastfed | **p** | **OR**<br>**95% CI** |
| **Crohn's Disease** | 82/152 (53.9%) | 4368/8196 (53.3%) | 0.93 | 1.03<br>(0.74-1.42) |
| **Ulcerative colitis** | 14/35 (40%) | 4368/8196 (53.3%) | 0.08 | 0.58<br>(0.3-1.15) |

| | **Greater Glasgow** | | | |
|---|---|---|---|---|
| | **Cases** | **Controls** | | |
| | Number breastfed | Number breastfed | **p** | **OR**<br>**95% CI** |
| **Crohn's Disease** | 35/93 (37.6%) | 2270/6775 (33.5%) | 0.46 | 1.20<br>(0.78-1.83) |
| **Ulcerative colitis** | 18/41 (43.9%) | 2270/6775 (33.5%) | 0.21 | 1.55<br>(0.84-2.88) |

**Table 4-7: Breastfeeding for one week or more by diagnosis: cases compared with population-matched controls.**

**Table 4-8: Breastfeeding for 2 weeks or more by diagnosis: cases compared with population-matched controls.**

| | **Lothian** | | | |
|---|---|---|---|---|
| | **Cases** | **Controls** | | |
| | Number breastfed | Number breastfed | **p** | **OR**<br>**95% CI** |
| **Crohn's Disease** | 81/152 (53.3%) | 4368/8196 (53.3%) | 0.93 | 1.00<br>(0.72-1.38) |
| **Ulcerative colitis** | 14/35 (40%) | 4368/8196 (53.3%) | 0.08 | 0.58<br>(0.3-1.15) |

| | **Greater Glasgow** | | | |
|---|---|---|---|---|
| | **Cases** | **Controls** | | |
| | Number breastfed | Number breastfed | **p** | **OR**<br>**95% CI** |
| **Crohn's Disease** | 39/93 (41.9%) | 2270/6775 (33.5%) | 0.10 | 1.43<br>(0.95-2.17) |
| **Ulcerative colitis** | 16/41 (39.0%) | 2270/6775 (33.5%) | 0.56 | 1.27<br>(0.68-2.38) |

**Table 4-8: Breastfeeding for 2 weeks or more by diagnosis: cases compared with population-matched controls.**

**Table 4-9: Breastfeeding for 1 week or more by deprivation category: cases compared with population-matched controls**

| | Cases | Controls | | |
|---|---|---|---|---|
| **Depcat score*** | Number breastfed | Number breastfed | **p** | **OR 95% CI** |
| **1-2** | 53/79 (67.1%) | 4226/6648 (63.6%) | 0.59 | 1.17 (0.73-1.87) |
| **3-5** | 119/254 (46.9%) | 9046/22475 (40.2%) | 0.03 | 1.31 (1.02- 1.68) |
| **6-7** | 17/67 (25.4%) | 1768/7750 (22.8%) | 0.72 | 1.15 (0.66-2.00) |

**Table 4-9: Breastfeeding for 1 week or more by deprivation category: cases compared with population-matched controls. * 1= least deprived, 7= most deprived**

**Table 4-10: Breastfeeding for 2 weeks or more by deprivation category: cases compared with population-matched controls**

| | Cases | Controls | | |
|---|---|---|---|---|
| **Depcat score*** | Number breastfed | Number breastfed | **p** | **OR 95% CI** |
| **1-2** | 52/79 (65.8%) | 4226/6648 (63.6%) | 0.76 | 1.10 (0.69-1.76) |
| **3-5** | 111/254 (43.7%) | 9046/22475 (40.2%) | 0.29 | 1.15 (0.90- 1.48) |
| **6-7** | 17/67 (25.4%) | 1768/7750 (22.8%) | 0.61 | 1.15 (0.66-2.00) |

**Table 4-10: Breastfeeding for 2 weeks or more by deprivation category: cases compared with population-matched controls. * 1= least deprived, 7= most deprived**

### *4.1.10 Matched case-control analysis: past medical history*

Invididual past medical histories did not differ significantly between IBD cases and age, sex and postcode-matched controls, except for atopic disease which will be discussed in detail below. 1 control suffered from coeliac disease. 1 from each group suffered from autism. Minor surgical procedures were occurred in both groups: 6 cases (4.8%) versus 12 controls (9.7%) have undergone tonsillectomy (p=0.14). 5 cases (4.1%) compared to 1 control (0.8%) have undergone appendectomy (p=0.09).

### *4.1.11 Matched case-control analysis: Family History (FH)*

Of the IBD patients, 40 (32.5%) compared to 14 (12.6%) controls had a positive FH (OR3.34; 95% CI (1.70-6.56), p=0.0003). In CD 28 (35.0%) patients had a positive FH (OR 3.34 (1.48-7.51) P=0.002).

Coeliac disease (n cases=5; n control=2), and autism (n cases=4; n control=3) did not differ significantly between families of IBD sufferers and matched control families (p=0.31 and 0.80, respectively).

Family History of Bowel cancer (n cases=21; n control=11) was more common in the families affected by a child with IBD compared with matched control families but this did not reach statistical significance on univariate analysis.(p=0.06).

### *4.1.12 Matched case-control analysis: Atopic Disease*

Overall, children with IBD were significantly more affected by self-reported atopic disease than matched controls (p=0.01, OR 1.88 (1.13-3.13)). It should be noted that, in order to facilitate data recording using the questionnaires in both cases and matched controls, no distinction was made between allergic asthma, exercise-induced asthma, cough-variant asthma etc.(see Appendices) Asthma (p=0.005 OR 2.48 (1.28-4.78)), eczema (p=0.001 OR 2.83 (1.50-5.36)) and food allergy (p=0.03 OR 2.98 (1.04-8.54)) were all more common in children with IBD than in matched controls.(Table 4-11)

**Table 4-11: Self-reported atopic diseases: cases compared with matched controls**

| | IBD Cases | Matched controls | p value | OR |
|---|---|---|---|---|
| **Any atopy** | 66/124 (53.2%) | 46/122 (37.7%) | 0.01 | 1.88 (1.13-3.13) |
| **Asthma** | 34/124 (27.4%) | 16/121 (13.2%) | 0.005 | 2.48 (1.28-4.78) |
| **Eczema** | 39/124 (31.5%) | 17/122 (13.9%) | 0.001 | 2.83 (1.50-5.36) |
| **Hayfever** | 23/124 (18.5%) | 29/122 (23.8%) | 0.31 | |
| **Food allergy** | 14/124 (11.3%) | 5/122 (4.1%) | 0.03 | 2.98 (1.04-8.54) |

**Table 4-11: Self-reported atopic diseases: cases compared with matched controls**

When results were stratified for CD as in Table 4-12, the association was as strong as for IBD. Atopy, asthma, eczema and food allergy all showed significant associations with CD. Atopy was 2.28 times more likely in those with CD, a higher number than found with combined IBD data.

**Table 4-12: Self-reported atopic Disease: CD cases versus matched controls**

| | CD Cases | Matched Controls | p value | OR (95% CI) |
|---|---|---|---|---|
| **Any atopy** | 45/81 (55.6%) | 28/79 (35.4%) | 0.01 | 2.28 (1.21-4.30) |
| **Asthma** | 25/81 (30.9%) | 11/78 (14.1%) | 0.01 | 2.72 (1.23-6.01) |
| **Eczema** | 27/81 (33.3%) | 10/79 (12.7%) | 0.001 | 3.45 (1.54-7.74) |
| **Hayfever** | 14/81 (17.3%) | 16/79 (20.3%) | 0.63 | |
| **Food allergy** | 8/81 (9.9%) | 1/79 (1.3%) | 0.03 | 8.54 (1.04-70.06) |

**Table 4-13: Self-reported atopic Disease: CD cases versus matched controls**

### *4.1.13 Matched case-control analysis: Immunisations History*

**Table 4-13: Self-reported immunisations history: IBD and CD cases compared with matched controls**

| **Immunisation** | | | **Matched Controls** | **p value** | **OR (95% CI)** |
|---|---|---|---|---|---|
| **DPT** | **IBD** | 119/121 (98.3%) | 108/124 (87.3%) | 0.0007 | 8.81 (1.98-39.23) |
| | **CD** | 77/78 (98.7%) | 68/81 (84.0%) | 0.001 | 14.72 (1.88-#) |
| **HIB** | **IBD** | 116/118 (98.3%) | 68/124 (54.8%) | <0.0001 | 47.76 (11.30 – #) |
| | **CD** | 75/76 (98.7%) | 38/81 (46.9%) | <0.0001 | 84.87 (11.25-#) |
| **Men C** | **IBD** | 115/117 (98.3%) | 88/124 (71.0%) | <0.0001 | 23.52 (5.51-#) |
| | **CD** | 72/74 (97.3%) | 57/81 (70.4%) | <0.0001 | 15.16 (3.44-66.84) |
| **MMR** | **IBD** | 119/121 (98.3%) | 116/124 (93.5%) | 0.05 | 4.10 (0.85-19.73) |
| | **CD** | 78/78 (100%) | 74/81 (91.4%) | 0.007 | |

**Table 4-13: Immunisations history: IBD and CD cases compared with matched controls**

We compared the rate of self-reported immunisation history for all common childhood immunisations between IBD cases and matched controls. For IBD cases, these data were obtained during an interview with one of the researchers. When the parents were unable to recall precisely which immunisations had been given, one of the researchers would then ask whether any immunisations such as MMR were ever refused. For matched controls, as no such interview was possible logistically and in view of Data Protection Regulations, these data were gathered from a list of immunisations which the control parents/control subjects had to tick to confirm that the respective immunisations had been administered. Significant differences between IBD/CD cases and matched controls were found regarding self-reported immunisation history against DPT, HIB and Men C. Self-reported MMR showed a borderline association with IBD but a significant difference in immunisation rate between CD cases and matched controls.

### *4.1.14 Matched case-control analysis: Breastfeeding*

Results were analysed for all IBD and stratified for CD. For IBD and CD, there were no significant differences between cases and controls regarding breastfeeding (IBD: 52.5% vs. 54.9% in controls, p=0.71, CD: 57.3% vs. 60.0% in controls, p=0.73). Duration of breastfeeding was not different between IBD cases and matched controls: in the IBD cases the median duration was 20 weeks (Q1-Q3: 12-36) and for matched controls the median was 17 weeks (Q1-Q3: 7.75-40), p=0.52, Mann Whitney U test).

### *4.1.15 Matched case-control analysis: Passive smoking*

None of the matched IBD cases and only 3/123 controls were current smokers. 1 control was an ex-smoker.

Parental smoking was assessed as smoking around the time of pregnancy, birth, currently smoking or ex-smoker. Analysis was performed on the basis of whether the mother and/or the father smoked (from now on referred to as 'parental smoking') as this represents passive smoking exposure in the household.

**Table 4-14: Parental smoking during pregnancy, at birth and current status: IBD cases compared with matched controls**

| | | IBD Cases | Matched Controls | p, OR (95% CI) |
|---|---|---|---|---|
| **Parental smoking** | **Pregnancy** | 54/113 (47.8%) | 40/121 (33.1%) | 0.02 OR 1.85 (1.09-3.15) |
| | **Birth** | 53/112 (47.3%) | 38/121 (31.4%) | 0.01 OR 1.96 (1.15-3.35) |
| | **Current** | 48/112 (42.9%) | 31/119 (26.1%) | 0.007 OR 2.13 (1.22-3.71) |
| | **Ex** | 26/64 (40.6%) | 31/87 (35.6%) | 0.53 |
| **Mother** | **Pregnancy** | 29/121 (24.0%) | 25/123 (20.3%) | 0.49 |
| | **Birth** | 31/121 (25.6%) | 24/123 (19.5%) | 0.25 |

| | | | | |
|---|---|---|---|---|
| | **Current** | 37/121 (30.6%) | 23/122 (18.9%) | 0.03 OR 1.90 (1.04-3.44) |
| | **Ex** | 20/84 (23.8%) | 21/97 (21.6%) | 0.72 |
| **Father** | **Pregnancy** | 39/114 (34.2%) | 32/122 (26.2%) | 0.18 |
| | **Birth** | 38/113 (33.6%) | 30/121 (24.8%) | 0.13 |
| | **Current** | 30/113 (26.5%) | 24/119 (20.2%) | 0.25 |
| | **Ex** | 20/83 (24.1%) | 23/96 (24.0%) | 0.98 |

**Table 4-14: Parental smoking during pregnancy, at birth and current status: IBD cases compared with matched controls**

On univariate analysis parental smoking at pregnancy, birth or currently significantly correlates with the later development of IBD in the child. When maternal and paternal smoking were examined separately, only maternal current smoking retained significance (p=0.03). In this context, it is noteworthy that the overall IBD cohort (ie not just the IBD cases for whom a matched control was available), the maternal smoking rate during pregnancy and after birth was higher at 123/440 (28.0%) and 138/439 (31.4%), respectively. On analysis of the matched CD case-control cohort, no significant differences were noted with regards to parental or maternal smoking although the passive smoking rate was consistently higher in the CD cases.(Table 4-15)

**Table 4-15: Parental smoking during pregnancy, at birth and current status: CD cases compared with matched controls**

| | | CD cases | Matched controls | P value |
|---|---|---|---|---|
| **Parental smoking** | **Pregnancy** | 32/73 (43.8%) | 27/80 (33.8%) | 0.20 |
| | **Birth** | 31/72 (43.1%) | 25/80 (31.3%) | 0.13 |
| | **Current** | 29/74 (39.2%) | 21/79 (26.6%) | 0.09 |
| **Mother** | **Pregnancy** | 18/78 (23.1%) | 15/80 (18.8%) | 0.50 |
| | **Birth** | 20/78 (25.6%) | 14/80 (17.5%) | 0.21 |
| | **Current** | 24/78 (30.8%) | 14/79 (17.7%) | 0.05 |

**Table 4-15: Parental smoking during pregnancy, at birth and currently: CD cases compared with matched controls**

### *4.1.16 Matched case-control analysis: Multifactorial Analysis*

Multifactorial analysis was performed using binary logistic regression. Significant findings from unifactorial analysis were entered in the model.(Table 4-16) The factors contributing significantly to IBD were FH of IBD, FH of bowel cancer and an individual's history of asthma and eczema.

The significant findings on unifactorial analysis of association with parental smoking during pregnancy, at birth and currently were not replicated. Entering the FH of IBD as an interacting variable with the smoking history did not alter this. For CD, the same conclusions were drawn.(Table 4-17)

**Table 4-16: Matched IBD case-control analysis: Multifactorial analysis**

| | p | OR | 95% CI |
|---|---|---|---|
| **Parental Smoking**<br>• During pregnancy | 0.353 | 0.41 | 0.06-2.68 |
| **Parental Smoking**<br>• At birth | 0.148 | 4.45 | 0.59-33.65 |
| **Parental Smoking**<br>• At time of diagnosis | 0.335 | 1.57 | 0.63-3.89 |
| **Atopy** | 0.237 | 0.59 | 0.24-1.42 |
| **Asthma** | 0.045 | 2.59 | 1.02-6.58 |
| **Eczema** | 0.015 | 3.09 | 1.25-7.64 |
| **Food Allergy** | 0.105 | 2.86 | 0.80-10.20 |
| **FH of IBD** | <0.001 | 4.07 | 1.89-8.75 |
| **FH of bowel cancer** | 0.033 | 2.72 | 1.09-6.81 |

**Table 4-16: Matched IBD case-control analysis: Multifactorial analysis**

**Table 4-17 Matched CD case-control analysis: Multifactorial analysis**

| | p | OR | 95% CI |
|---|---|---|---|
| **Asthma** | 0.045 | 3.52 | 1.03-12.06 |
| **Eczema** | 0.029 | 3.89 | 1.15-13.13 |
| **FH of IBD** | 0.003 | 4.41 | 1.68-11.52 |
| **FH of bowel cancer** | 0.025 | 4.14 | 1.20-14.32 |

**Table 4-17: Matched CD case-control analysis: Multifactorial analysis**

### *Discussion*

Using both population-matched and sex,age and postcode-matched control data, we have shown the novel association of childhood onset IBD and CD with asthma in the high incidence Scottish population. We did not observe any association between paediatric IBD/CD/UC and either immunisation history or breastfeeding after stratification for geographical region and affluence using population-matched data.

Before considering our findings in view of the relevant literature for each of our main findings, it is pertinent to address some of the methodological issues studies like this are faced with.

In the first phase of this study, we used population-matched data as control data. For immunisation history and breastfeeding, we matched the area of residence of the cases and controls. We selected control data from Lothian, Grampian and Greater Glasgow (to match the place of residence of our cases). We also endeavoured to match cases and controls by age. 92% of our cases were born between 1985 and 2000: 40% between 1985 and 1990; 39% between 1991 and 1995 and 13% between 1996 and 2000. Control immunization data came from children born in 1993.

For the asthma data, controls are from the whole of Scotland as breakdown by health-board was unavailable. Our asthma controls were children aged 2-15 in 2003, i.e. children born between and 1988 and 2001. However, given that asthma prevalence is increasing, matching children individually for date of birth is preferable and was be performed in the second phase of this study which ran from March till July 2008. The definition of asthma merits further discussion: in cases, results relied on parental reporting of the child's "asthma" whereas for controls, population-data

(from the surveyed GP surgeries) on "doctor-diagnosed asthma" was used. Furthermore, no distinction was made between true atopic/allergic asthma, exercise-induced asthma, cough-variant asthma and nocturnal asthma.

Although the issue of self-reporting compared with doctor-diagnosed asthma could have influenced our finding of increased asthma prevalence among cases, our subsequent matched case-control analysis confirmed this association was a true finding.

Breastfeeding population control data came from children born in 1996 (this was the only complete set of population-data available) and were stratified according to geographical location (Glasgow vs Lothian) and according to Deprivation Category. In the cohort of IBD cases, recall bias may influence our breastfeeding data. Parents were asked during face-to-face interview about the duration of breastfeeding (in weeks). For population-matched controls, breastfeeding in the different health boards was recorded by a health visitor (at 10 days). These considerations should be viewed in the light of the increased power of our study due to the use of large population-based datasets. As breastfeeding rates also tend to change over time, the matched case-control analysis in a subset of IBD patients was conducted to address this.

This type of study is subject to recall bias. Parents of IBD cases were asked about their child's exposure to environmental risk factors. Our population-matched control group for immunisations used GP records to establish immunisation status, thus limiting recall bias. However, recall bias may exist in the case group and hence affect comparisons. Although it would be methodologically preferable to use GP immunisation records for both cases and controls, we opted, for logistical and Data Protection reasons, to use the same questionnaire for cases and controls in the second phase of this study.

We then encountered additional methodological difficulties due to the use of a postal questionnaire to ascertain immunisations history in our sex, age and postcode-matched controls. This led to the marked underreporting of administered

immunisations as control subjects/control parents had to tick a list of immunisations to confirm they had been administered, which was crucially different to the way the questionnaire was filled out during the recruitment of IBD cases (when a member of the research team would ask if the parents had ever refused for their children to have an immunisation). We therefore feel that these data should be interpreted with great caution, especially when compared with the better-powered case-control analysis employing population-matched data based on GP records.

### *4.1.17 Asthma*

The current study found asthma to be more prevalent in paediatric IBD patients than in the Scottish paediatric population as a whole ($p < 0.0001$ OR 1.67 (1.32-2.10) and in age, sex and postcode matched controls ($p=0.005$ OR 2.48 (1.28-4.78)).

This is consistent with findings of two recent case-control studies. Smaller studies find respiratory symptoms to be more common in IBD patients than in the general population.[606-608] In the paediatric IBD population, the finding of an association between asthma and IBD is novel.

**Table 4-14: Prevalence of asthma in IBD patients: overview of available literature**

**Table 4-18: Prevalence of asthma in IBD patients: overview of available literature**

| **Asthma prevalence** | **Study participants** | **All IBD** OR (95% CI) | **CD** OR (95% CI) | **UC** OR (95% CI) |
|---|---|---|---|---|
| Weng et al (2007)[609] | 12,601 cases | 1.5 (1.5-2.3) | | |
| Bernstein et al (2005)[596] | 8072 cases | - | 1.43 (1.26-1.62) | 1.66 (1.46-1.88) |

Becker et al. have shown immune-mediated disease susceptibility loci exhibit non-random clustering, so there may be a common genetic basis for these disorders.[610] A number of studies note the clustering of immune-mediated diseases other than asthma with IBD. Psoriasis [596;609], rheumatoid arthritis[596;609], multiple sclerosis[611;612], chronic sinonasal disease[613] and pericarditis[596] all appear to be more prevalent among IBD patients.

The association of IBD with asthma should be viewed in light of the recent paradigm shift in the regulation of the T-helper lymphocytic response. Recent advances in the biology of the T-helper cell regulation, have demonstrated that categorising immune-mediated diseases into those driven by Th1 cells (CD, type 1 diabetes and multiple sclerosis) and those driven by Th2 cells (asthma, allergic rhinitis and atopic dermatitis), is not doing the complex regulation of these responses justice. Indeed, recent data in suggest that even within Crohn's disease (both in human tissue and in the mouse model SAMP/Yit), which was long thought to be a Th1-driven disease process, time of assessment relative to the onset of the disease process is crucial in understanding the relative contribution of Th1- and Th2 cytokines to Crohn's disease.[614;615]

Epidemiological evidence showing the coincidence of diseases driven by $T_H$ 1 and by $T_H$ 2 suggests there may be a common pathway for both types of immune-mediated disease, perhaps even with sex-specific effects. [616;617] The association between CD (Th1) and asthma (Th2) in our exclusively paediatric study as well as in recent reports by Weng et al. and Bernstein et al. provide further epidemiological evidence for common regulatory pathways of the inflammatory response.[596;609] In addition to advances in our knowledge of the T-regulatory cell, it has been the characterisation of Th17 cells, the elucidation of the role which IL23 plays in the differentiation of these cells, together with functional data on interleukin 17 (IL-17) in immune-mediated tissue injury that has produced a true schism with the old school doctrine of a Th1/Th2 dichotomy.[618] IL17 is involved in organ-specific immunity in the

intestine, allergic disorders of lungs and skin, and microbial infections of the intestines and nervous system.[618] The advances in our understanding of Th17 biology only preceded the new genetic susceptibility data implicating the IL23 receptor by a few years. [470]

In addition to the influence of germline variation on immune sytem regulation, the association of IBD with immune-mediated diseases such as asthma, also necessitates the investigation of environmental triggers (with putative epigenetic effects) in view of the "hygiene/old friends hypothesis". The "hygiene hypothesis" is based on observations that exposure to infections in early life reduces the incidence of bronchial asthma. [619] Recent data by Green and colleagues have suggested that exposure to enteric infections in early life affords protection against development of IBD.[620] .

In many epidemiological studies surrogate markers for childhood hygiene are assessed: number of siblings and birth order, crowding in the home, urban versus rural environment, contact with farm animals[621], pets[584], drinking of unpasteurised milk[584] and availability of a hot water tap [591;622]. Socioeconomic status is often used as an indicator of less crowded living space and better hygiene.[566] Many of these studies show positive associations between increased childhood hygiene and risk of IBD. It is postulated that people protected from infections in childhood are ultimately more susceptible to chronic intestinal inflammation due to altered regulation of the immune response at the time of infections later in life.[622]

It is prudent to point out that no study has found all markers of childhood hygiene to be associated with disease and some studies have contradicted the theory. [591;623] One case-control study even suggested that poor hygiene contributes to the pathogenesis of IBD. [624]

### *4.1.18 Breastfeeding*

The current study finds breastfeeding is not associated with paediatric IBD. It also shows that breastfeeding prevalence decreases with increasing deprivation, both for cases and controls. We were careful to stratify our data also for this confounding factor but found no difference between breastfeeding rates in cases and controls.

Our results present important negative findings in the current climate of confusion regarding this issue. Studies examining association between breastfeeding and IBD exhibit large heterogeneity, even when only the highest-quality studies are compared.[625]

Breast milk provides the neonate with direct immunological protection and contains components that enhance the development of the immune system. It also affects the pattern of exposure to micro-organisms in the neonatal period. A protective effect of breastfeeding may be the result of the development of oral tolerance to specific microflora and food antigens. Additionally, oligosaccharides present in human milk have been to shown to have an anti-inflammatory effect, inhibiting leucocyte rolling and adhesion. This could help to explain the lower incidence of inflammatory conditions in breastfed infants, although a recent systematic review suggests the protective effect of breastfeeding is mainly due to its protective effect against infections early in life, rather than against the development of allergic conditions later.[626-628] The role of breastfeeding in the establishment of the intestinal microflora should also be considered with regards to its potential of influencing subsequent development of allergic diseases and IBD (also see Chapter 1). As yet unknown environmental factors influence the establishment of this gut flora very early in life (as illustrated by the diversity of flora between Estonian and Swedish children at one week of age, in the study by Bjorksten et al.) and less microbial diversity has been demonstrated in atopic infants, with increased levels of *Clostridium*.[627;629-631] Penders et al reported the association of *C.difficile* in the stool of infants at 1 month of age with the later development of atopy in the KOALA-birth cohort study.[78;599]

It is noteworthy that supplementation with probiotics (from before birth to 12 months of age) is able to influence the development the allergic (IgE associated) eczema, with the only negative study commencing probiotic treatment postnatally.[627;632;633] Even more striking is the observation that the specific constituents of the maternal vaginal flora at 22-24 weeks gestation (e.g. *Ureaplasma* and *Staphylococci*) are associated with subsequent asthma in the children.[634] This effect of maternal flora is further illustrated by the association of subsequent development of asthma with caesarean section.[78;635-637] The risk of developing childhood diabetes is possibly also associated with alterations of the maternal flora, as suggested by the increased risk of childhood diabetes associated with the use of certain antimicrobials (e.g. macrolides and quinolones).[638] A previous study from our group by Russell et al. associating perinatal exposure to smoking with the later development of IBD, could therefore also be viewed from the perspective maternal smoking has on the composition of the vaginal microflora.[84;639] In contrast, maternal smoking early in pregnancy has been associated with a protective effect on later development of IBD in children.[86]

Conversely, breastfeeding may predispose to IBD. Delayed infection occurring at weaning, leading to inappropriate immune response and persistence of intestinal inflammation could explain this.[81] Breast milk pollution in industrialised areas may predispose to IBD. Industrial chemicals and environmental contaminants have indeed been found in breast milk. [640] Fine and ultrafine particles may be present in breastmilk and are known to be potent adjuvants in the antigen-mediated immune response but their role in IBD remains unclear.

Few studies have investigated breastfeeding as a risk factor for *juvenile-onset* IBD specifically. Table 4-10 summarises paediatric case-control studies to date.

Our study shows a lack of association between breastfeeding and both CD and UC. Two studies support our findings; Koletzko et al[641] and Baron et al[81] showed no association between breastfeeding and UC. However, our finding of a lack of association between breastfeeding and *paediatric* CD is novel. Other studies show

breastfeeding to have a protective effect (stronger in CD than UC).[603;605] In contrast, Baron et al showed breastfeeding as a risk factor for paediatric CD.[81]

One study found the relative risk of IBD to decrease as duration of breastfeeding increased.[605] The current study was unable to address this question.

**Table 4-19: Paediatric case-control studies testing association of breastfeeding with CD and UC**

**Table 4-19: Paediatric case-control studies testing association of breastfeeding with CD and UC**

| Authors, year of publication | Number of cases/ controls | Age at diagno-sis | Definition of breastfee-ding | OR | 95% CI | Conclusion |
|---|---|---|---|---|---|---|
| **Crohn's Disease** | | | | | | |
| Koletzko et al, 1989 [603] | 114/180 | <18 years | No exact definition | 0.26 | 0.11-0.67 | Breastfeeding protective |
| Rigas et al, 1993[605] | 68/202 | Unspec ified. "Paediatr ic and adolesce nt patients" | No exact definition | 0.48 | 0.27-0.85 | Breastfeeding protective |

| | | | | | | |
|---|---|---|---|---|---|---|
| Baron et al, 2005[81] | 222/222 | <17 years | Nonexclu-sive breastfee-ding | 2.1 | 1.3-3.4 | Breastfeeding is a risk factor |

**Ulcerative Colitis**

| | | | | | | |
|---|---|---|---|---|---|---|
| Koletzko et al, 1991[641] | 93/138 | Unspecified | No exact definition | 0.59 | 0.27-1.3 | No association |
| Rigas et al, 1993[605] | 39/202 | Unspecified. "Paediatric and adolescent patients" | No exact definition | 0.56 | 0.3-1.03 | Breastfeeding protective |

Other studies examining association between breastfeeding and IBD use patients with disease-onset *at any age* as their study population. Like the paediatric studies, the results of these investigations are heterogeneous. Several studies suggest an inverse association between breastfeeding and IBD. For example, Acheson and Truelove[642] and Whorwell et al[602] found breastfeeding to decrease the risk of UC but not CD. Bergstrand and Hellers found breastfeeding to decrease the risk of CD[643]. However, many studies demonstrate no such association[584;587;589;644;645]

A meta-analysis of 17 studies, including data from both children and adults, showed breastfeeding to reduce the risk of CD and UC. When the 8 highest quality studies were statistically analysed, the pooled odds ratio was 0.45 (95% CI: 0.26, 0.79) for CD and 0.56 (95% CI ; 0.38,0.81) for UC.[79] However, Klement et al. acknowledge that reviewing studies on this subject is problematic. Exclusivity and duration of breastfeeding are often ill-defined. Additionally, recall bias affects all the studies, thus ORs quoted in this meta-analysis should not be considered definitive.

### *4.1.19 Immunisations*

The current study presents important negative findings regarding the association of vaccinations with paediatric IBD, CD or UC. We found no significant differences between cases and controls for vaccination against diphtheria, tetanus, pertussis, MMR or HIb3 after correction for multiple testing in our large sample size study. We have identified clear methodological issues precluding the interpretation of immunisation history data from our age, sex and postcode-matched cohort.

Of all vaccinations, MMR has caused most controversy in relationship to IBD; specifically the live attenuated measles element of this vaccination. In 1995, Thompson et al presented evidence that measles vaccination was linked to increased odds of developing UC and CD, giving ORs of 2.53 and 3.01 respectively. [645] Interest in the subject increased and new studies emerged. It was proposed that measles virus can persist in intestinal tissue and that IBD could in fact result from granulomatous vasculitis in response to chronic infection.[601;646;647] Additionally, evidence suggested that measles virus can cause prolonged disruption of immune function, particularly in the helper T cell response. [648] It was therefore suggested that measles infection and vaccination might predispose to IBD.[646]

However, methodological concerns were raised about the Thompson study and, since 1995, studies have contradicted this evidence. A UK case-control study (1997) found no correlation between measles vaccination and either CD, UC or all IBD combined.[649] Indeed, no convincing link between either wild-type or attenuated measles virus and IBD has been established.[650] Our more reliable, population-matched results are in agreement with this.

Ecological studies also support our findings. For example, Hermon-Taylor et al showed that the increase in CD in the UK started some 20 years prior to the

introduction of the measles vaccine, suggesting that the vaccination is not the cause of the increase.[651]

In 1998, a report reignited public concern because it detailed 12 children with non-specific colitis, ileal-lymphoid-nodular hyperplasia and developmental disorders attributed to MMR vaccination.[652] However, concerns about this report include potential selection biases, recall bias, lack of clear case definition and lack of a comparison group.[653;654]

Our population-matched results supplement the strong evidence base for lack of association between MMR and IBD. For health professionals and the general public alike, the safety-affirmation of childhood immunisations with regards to IBD, particularly MMR, could contribute to the prevention of a further decline in vaccination uptake and subsequent rise in vaccine-preventable diseases.[655;656] Whether specific sub-groups of children (e.g. with metabolic disorders) are more at risk than others of immunisation-related adverse events, remains the topic of intense debate in high impact journals, as illustrated recently in Pediatrics and the New England Journal of Medicine.[657-659]

# 5 Germline variation of NOD1/CARD4 in IBD in Northern Europe

### *Introduction*

The breakthrough in IBD genetics when NOD2/CARD15 was discovered as the first susceptibility gene in CD heralded extensive analyses of signalling pathways of the innate immune system implicated in the pathogenesis of CD.[406;407] NOD2/CARD15 is an intracellular pathogen-associated molecular pattern receptor for muramyl dipeptide (MDP), a breakdown product of bacterial peptidoglycan of both Gram-negative and Gram-positive bacteria.[418;421;422] The contribution of NOD2/CARD15 variants to CD susceptibility in the Scottish population is markedly smaller than in other populations studied (Central and Southern Europe, North America and Canada), but comparable to the Scandinavian population.[337;473;660-664]

Like NOD2/CARD15, NOD1/CARD4 is a member of the phylogenetically conserved Nod-like receptor family.[665;666] The NOD1/CARD4 gene contains 14 exons (54 kb) and is located within the putative IBD susceptibility locus on chromosome 7p14.3 which was described in the only UK genome sib-pair analysis.[108]

Structurally the NOD1/CARD4 protein is similar to NOD2/CARD15, but only has one CARD domain.[311] At its Leucine Rich Repeat (LRR)-domain NOD1/CARD4 binds intracellular γ-D-glutamyl-meso-diaminopimelic acid (iE-DAP), a breakdown product of peptidoglycan from Gram-negative bacteria and a limited number of Gram-positive bacteria (e.g. *Listeria monocytogenes).*[417-420;667] In contrast to NOD2/CARD15, NOD1/CARD4 is constitutively expressed in epithelial cells throughout the gastrointestinal tract.[194;434;668]. Recent work has implicated NOD1/CARD4 in the response against *Helicobacter pylori* and Gram- negative enteric pathogens like *Escherichia coli* which avoid recognition by Toll-like Receptors (TLRs).[431;668;669] Like NOD2/CARD15, NOD1/CARD4 initiates a pro-inflammatory response largely dependent on NF-κB activation.[418;424;446] Investigators from the Philpott lab have shown that stimulation alone was sufficient to drive antigen-specific immunity with a predominant Th2 polarization profile.[670] In conjunction with TLR stimulation, however, NOD1/CARD4 triggering was required to instruct the onset of Th1 and Th2 as well as Th17 immune pathways. Pretreatment of macrophages with TLR ligands revealed that innate immune responses induced by

bacterial infection relied on NOD1/CARD4 and NOD2/CARD15 and their adaptor RICK/RIPK2.[671] Cells outside of the hematopoietic lineage provided the early signals necessary to orchestrate the development of Nod1-dependent immune responses.[446;672] These findings highlight NOD1/CARD4 as a key innate immune trigger in the local tissue microenvironment that drives the development of adaptive immunity.

Taken together, these observations make NOD1/CARD4 a strong functional and positional IBD candidate gene.

A number of groups have investigated whether germline variation of NOD1/CARD4 is associated with increased susceptibility to IBD.[459-461;463;464;673] Of these, the positive study by McGovern et al has received the greatest attention. These investigators from Oxford suggested the association of the deletion variant of a complex intronic insertion*2/deletion*1 polymorphism (32656) of NOD1/CARD4 with susceptibility to IBD using a combination of TDT and case-control analysis.[459] In a cohort of 556 trios, these investigators observed distortion of transmission of this NOD1/CARD4 variant in IBD and UC, but not CD. Single SNP case-control analysis (664 IBD patients and 335 controls) showed significant associations with IBD and CD (especially onset <25 years), but not UC. They went on to perform a sliding haplotype TDT analysis, a new approach in IBD genetics, obtaining significance for a two-marker haplotype spanning a large section of the LRR domain. In addition, a gene-dosage effect of the deletion variant on age of onset in CD and IBD was demonstrated.[459] The same group also described an association of the insertion allele of this complex insertion/deletion with high IgE levels and asthma.[434]

However, well-designed replication studies (including the data presented later in this chapter) involving a combined total of over 10,000 subjects have been negative, thereby shedding doubt on the contribution of this single NOD1/CARD4 SNP.[463;464;674] Tremelling et al. assessed 1370 IBD patients and 760 controls from the Cambridge area and did not find evidence for association of this NOD1/CARD4 variant with IBD (p=0.74).[463] Franke and colleagues were also unable to replicate the suggested association in a German case-control analysis of 1015 IBD patients and 886 controls (p=0.51) and a TDT of 775 IBD trios (p=0.24).[464] These replication

studies had >95% power to replicate the effect size (odds ratio 2.0 for CD), as calculated in the initial study by the Oxford group.

In this study, our aims were twofold. First, we set out to assess the contribution of the NOD1/CARD4 insertion/deletion polymorphism to susceptibility to adult- and childhood-onset IBD in Northern Europe and to examine genotype-phenotype associations. We also sought to investigate the interaction between this NOD1/CARD4 polymorphism and NOD2/CARD15 variant alleles in determining IBD susceptibility and phenotype.

Our second aim was to assess the overall contribution of NOD1/CARD4 to IBD susceptibility, using a robust gene-wide haplotype tagging approach. We examined the influence of NOD1/CARD4 haplotypic variation on IBD phenotype and investigated the interaction between germline variants of NOD1/CARD4 and NOD2/CARD15 in determining IBD susceptibility and phenotype. Together, these two lines of investigation provide strong evidence that inherited variation of NOD1/CARD4 is not a strong determinant of disease susceptibility in the Scottish population, based on detailed power calculations for single SNP and haplotype analysis.

## *Methods*

### *5.1.1 Subjects*

#### 5.1.1.1 NOD1/CARD4+32656 ins/del polymorphism

3962 Northern European individuals were studied in total, comprising of 1791 IBD patients, 522 parents (293 mothers and 229 fathers) of IBD patients under the age of 16 at diagnosis and 1649 healthy controls (HCs). 97% of patients were Caucasian. Patient demographics are presented in Table 5-1.

*Scottish early onset IBD / TDT Cohort*

The group of IBD patients aged under 16 years of age at diagnosis was recruited from the three tertiary paediatric gastroenterology centres in Scotland (Edinburgh,

Glasgow and Aberdeen) and at the Western General Hospital in Edinburgh, a tertiary referral centre for IBD.[516] There were 217 complete family trios (69%).

*Scottish adult onset IBD Cohort*

The group of Scottish IBD patients (male 389 / female 437) was recruited from the IBD clinic at the Western General Hospital.

*Scottish Control Cohort*

The control group consisted of 1372 healthy Caucasian subjects (male 555 / female 557): Scottish blood donors (n=260), healthy volunteers (as assessed by questionnaire) recruited as part of the local study of IBD genetics (n=110) or healthy subjects recruited in the course of a Scotland-wide study of colorectal cancer genetics (n=1002). Data regarding the first two have been reported in previous studies.[248;271] The latter group consists of Caucasian subjects resident in Scotland with detailed ancestry (over 66% having at least three grandparents born in Scotland).

*Swedish adult onset IBD Cohort*

The group of Swedish IBD patients (362 male / 270 female) was recruited from hospitals in Stockholm County, Sweden.

*Swedish Control Cohort*

The Swedish control group consisted of 277 Caucasian healthy subjects (male 126/ female 151: healthy volunteers from the Karolinska University Hospital and orthopaedic day case surgery patients undergoing arthroscopic procedures; the majority for sports related complaints. All were free of any known chronic medical disease).

**Table 5-1: Demographics of IBD populations genotyped for the NOD1/CARD4+32656 ins/del**

| | Scottish early onset | Scottish adult onset | Swedish adult onset |
|---|---|---|---|
| N | 313 | 826 | 632 |
| M / F | 171 / 142 | 389 / 437 | 362 / 270 |
| Median Age at diagnosis (Q1-Q3) | 11.1 (8.6–12.9) | 31.1 (23.3 – 46.0) | 26.6 (19.7-39.6) |
| CD | 205 (65.5%) | 344 (41.6%) | 244 (38.6%) |
| UC | 81 (25.9%) | 482 (58.4%) | 388 (61.4%) |
| IBDU | 27 (8.6%) | - | - |
| Asthma | 82/309 (26.5%) | - | - |
| Any atopic disease | 153/309 (49.5%) | - | - |
| Caucasian | 97.6% | 98.3% | 99.4% |

#### 5.1.1.2 Haplotype-tagging approach

2296 subjects comprising of 1323 IBD patients (of which 356 children with IBD) (662 CD, 628 UC, 33 IBDU), 603 parents and 370 controls were genotyped. IBD patients (male 624/ female 699; median age at diagnosis [Quartile 1 – Quartile 3] 26.08 years [14.83 years- 39.91 years]) were recruited from the Western General Hospital and the three tertiary paediatric gastroenterology centres in Scotland (Edinburgh, Glasgow and Aberdeen). The cohort of children with IBD used in the TDT analysis was also used for the Scottish IBD case-control analysis (combined and childhood onset). Control subjects consisted of Scottish blood donors (n=260) and healthy volunteers (as assessed by questionnaire) recruited as part of the local

study of IBD genetics (n=110). More than 97% of the study population (patients and controls) was Caucasian. There were 246 complete family trios (69%).

### 5.1.2 Data collection

The diagnosis of IBD was based on standard criteria.[523] Data were collected retrospectively by review of patient files and questionnaire as previously described.[271;660] The Montreal classification was used to classify disease location and behaviour for CD and UC.[516]

### 5.1.3 Haplotype-tagging SNPs selection

The NOD1/CARD4 gene is located between positions 30,430,672 and 30,484,833 (Ensembl Release 43 - ENSG00000106100). Haplotype-tagging SNPs were selected using genotypic data from the CEU-study group (Thirty U.S. trios with Northern and Western European ancestry, collected by the Centre d'Etude du Polymorphisme Humain (CEPH)), available from the HapMap project (Release 21, July 2006; http://www.hapmap.org/cgi-perl/gbrowse/hapmap_B35/ ). (see Figure 5-1 and Table 5-2) SNPs were selected using Haploview software (version 3.32, http://www.broad.mit.edu/mpg/haploview/). SNPs were identified using a haplotype-tagging strategy based on solid spine of linkage disequilibrium ($r^2$>0.8, haplotype frequency > 5%, minor allele frequency > 10%). Tagging of the 5' region of NOD1/CARD4 was achieved by tagging the rs2529447 variant of (block 5 - Fig 1) at position 30,501,735. The 3' region was tagged to position 30,437,394 (rs10267377, block 3 - Fig 1).

**Figure 5-1: NOD1/CARD4 genomic region with haplotype-tagging SNPs**

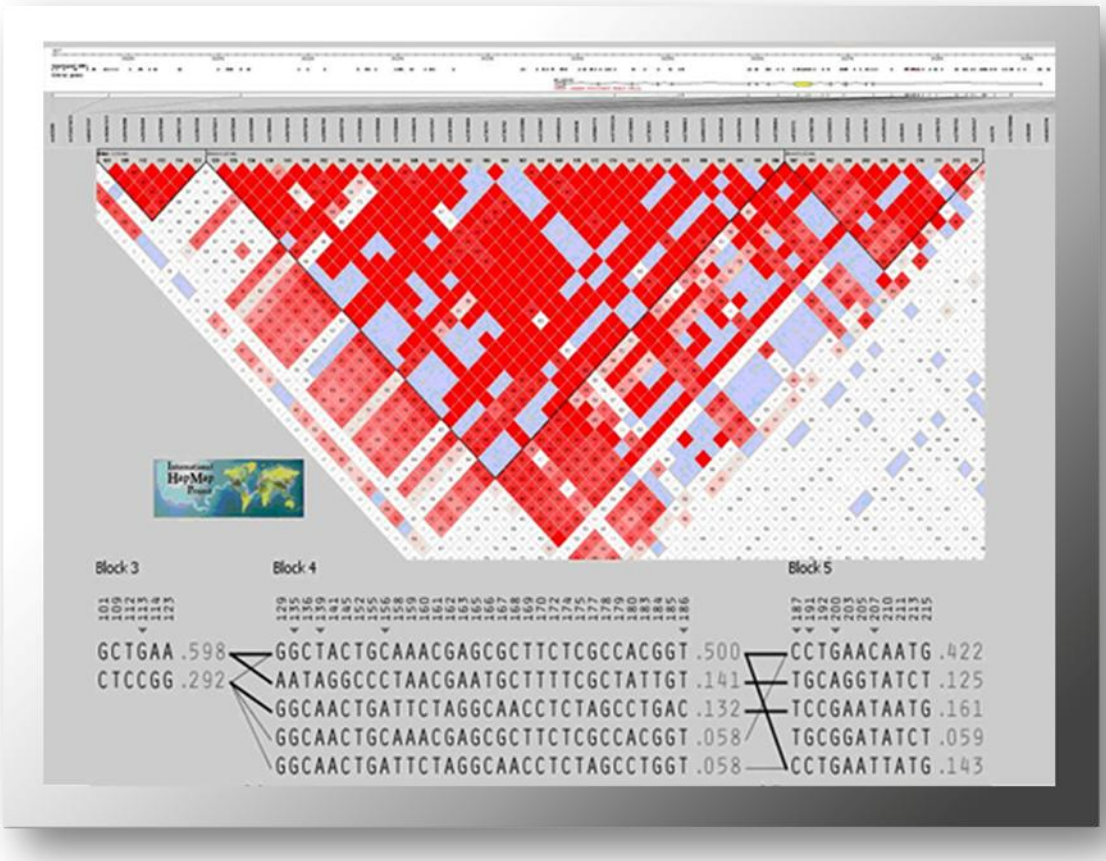



**Figure 5-1:View of the genomic region containing NOD1/CARD4 generated from HapMap (Release 21) project data using Haploview software. The haplotypic variation of the NOD1/CARD4 is tagged by 9 SNPs shown in the lower panel, covering 3 haplotype blocks (blocks 3-5).**

**Table 5-2: NOD1/CARD4 haplotype-tagging SNPs**

| NOD1/CARD4 haplotype-tagging SNPs | | |
|---|---|---|
| rs-number | position | Sequences |
| rs2970500 | 30,445,010 | GAGAGAGAAGGGAGAGAAAGTGAGA[C/G]AGCTTAACT GAGCTTAAATGTTAAC |
| rs2075820 | 30,458,762 | CGCAGCAGGAAGGCAAACACCTCCT[C/T]GGGGTCCCGC TCTGGGTAGCAGTAG |
| rs7789045 | 30,460,547 | TTGCTGACTGGTGGTCTCTTCCAGC[A/T]GACTTGAAGCT CCCTGAGGGCAGGA |
| rs1558066 | 30,469,712 | GCAACCAGGACTCTGTCCCTGAATC[A/C]TGGGGTGGCC CCAGAGAAACCACAG |
| rs4720004 | 30,478,744 | CACATGGTGAGGGAGAACCCCCGAT[C/T]TGATGGTGTT TACAGATAGGGCTGT |
| rs932272 | 30,479,684 | CACTCTAATAATGTTAGTCAATATC[A/G]GTATCAGATTT TTTTTTCCAGTTGA |
| rs2709799 | 30,481,428 | CCTAACCTGTATGGTGGCAATGACA[C/G]AACAGAAGGC CAAGGGACATCAACT |
| rs2529445 | 30,496,482 | GTGTCCATGTATTTATTTCTAGTAT[A/G]TAGGAATGCAA TTGATTTTTTATGT |
| rs38403 | 30,498,178 | GTAAAGTGTATTGTATTTACCCATA[C/T]TTTTGCTCACA GTGATCATTTTTTC |

**Table 5-2: NOD1/CARD4 tagging SNPs: rs-number, genomic location and flanking sequences**

### *5.1.4 Genotyping*

The NOD1/CARD4+32656 insertion*2/deletion*1 and the nine NOD1/CARD4 tagging SNPs were genotyped using TaqMan (7900HT sequence detection system; Applied Biosystems, Foster City, CA, USA). Failed TaqMan samples for the ins/del polymorphism were genotyped by direct sequencing and analysed by Sequencher v4.5 (Gene Codes Corporation, Ann Arbor, MI, USA). (5'-3'

AATCTCTGAGGTTGGGTGAGTAGAA; 3'-5' GGTCCTTCTGGTGTACTG ATGTATG; probes CCCCCACACACACAG / CCCCCCACACACAG). In 252 samples, NOD1/CARD4+32656 ins/del genotyping by TaqMan was compared with results obtained by direct sequencing with 100% genotype concordance.

### *5.1.5 Statistics*

#### 5.1.5.1 NOD1/CARD4+32656 insertion/deletion

The NOD1/CARD4+32656 insertion*2/deletion*1 polymorphism was analysed for association with IBD, CD, UC and IBDU by case-control study and TDT. Genotype-phenotype analyses based on the Montreal classification were performed.[516] Allelic, carrier and genotype frequencies were compared between patients and healthy controls in the association study. In the genotype-phenotype analyses, these parameters were compared between patients with a particular phenotype and patients without that phenotype. For subgroup analyses, CD patients were further stratified according to NOD2/CARD15 variant carriage.

Power calculations for the case-control analyses were based on a comparison by allele (carriage versus no carriage) and those for the TDT analyses used the calculation method of Knapp under a multiplicative genotypic risk model.[675] Both sets of calculations assumed a population risk allele frequency of 0.7, a population prevalence of 0.004 for IBD or 0.002 for CD or UC and a significance level of 0.05.

SPSS (Version 13.0; SPSS Inc., Chicago, IL, USA) and Minitab (Release 13.20, Minitab Inc., State College, PA, USA) statistical packages were used. Case-control analysis was performed using $\chi^2$ or Fisher exact test as appropriate. Multifactorial genotype-phenotype analyses were performed applying a binary logistic regression model. Gene-gene interactions were assessed as interaction categorical variables in the logistic regression model. We assessed the influence of NOD1/CARD4+32656 genotype on age at diagnosis by comparing the median age at diagnosis for each genotype, applying the Kruskal-Wallis test. Transmission disequilibrium testing analysis (TDT) was performed using the TRANSMIT package (Version 2.5, David

Clayton, Cambridge, UK).[531] This method makes full use of the data available using inferred genotypes if data from both parents are not available.

#### 5.1.5.2 NOD1/CARD4 haplotype-tagging approach

Allelic, genotype and haplotype frequencies were compared between cases and controls using $\chi^2$ (Minitab software, Release 13.20, Minitab Inc., State College, PA, USA and GraphPad Instat software version 3.06, San Diego, Ca, USA) and Haploview software (version 3.32). Single and multiple marker TDT analyses were performed using FBAT software (version 1.7.3, available from http://www.biostat.harvard.edu/~fbat/fbat.htm). [532] Odds Ratios (ORs) (99.5% Confidence Interval) for each of the SNPs were calculated using TDTHAP, run in R 2.5.0.[533] Log-likelihood analysis was carried out using PM software (1000 permutations – random seeds).[527] Power calculations were performed using QUANTO version 1.2 (http://hydra.usc.edu/gxe).[528;529] The method for integrating results from case-control studies and TDT to provide a combined estimate of disease-marker association was described by Kazeem and Farall (Table 7).[535] As this approach is only useful when the samples of subjects used in the two analyses are independent of each other, yet share the same genealogy, we based this calculation on our TDT cohort of early onset IBD and our samples of adult onset IBD patients in the case-control analysis.

## *Results*

### *5.1.6 Power calculations*

#### 5.1.6.1 NOD1/CARD4+32656 insertion/deletion

All Scottish IBD patients (childhood onset + adult onset) were analysed in a case-control analysis with 1372 control samples.The case-control analysis had 80% power to detect an effect size of OR 1.21 for IBD and OR 1.27 for CD/UC. Assuming a multiplicative risk model, the TDT analysis was adequately powered (>80%) to detect an effect size of Odds Ratio (OR) 1.5 for IBD, OR 2 for CD and OR 2.5 for UC.

#### 5.1.6.2 NOD1/CARD4 haplotype-tagging approach

We have first analysed the power associated with the analysis of individual ***SNPs***, as the power of each analysis is dependent upon the allelic frequency. We estimated the population risk to IBD as 0.004 (0.002 for CD and UC), based on population-based data by Rubin et al.[676] In this study, the age-sex adjusted point prevalence for ulcerative colitis on 1st January 1995 was 243.4/100 000 (Confidence Interval (CI): 217.4-269.4) and for Crohn's disease 144.8/100 000 (CI: 124.8-168.8). The most recent childhood onset IBD prevalence data in the Scottish population were derived from hospital discharges and date back to the period before 1997 (symptom onset 1981-1995).[4] Armitage et al. estimated the combined (CD+UC) crude prevalence of childhood onset IBD in Scotland to be 22.9/100 000 (13.7/100 000 CD, 9.2/100 000 UC) of the population aged 0-16 years. It is pertinent to point out that both the Rubin and Armitage study reflect reported disease prevalence more than a decade ago. Since then, the incidence of IBD has continued to rise and more up-to-date data from the Scottish Executive are awaited.

In comparison with dominant and recessive risk models, we assessed the multiplicative risk model to be more appropriate for the investigation of a complex polygenic disease. Using a multiplicative risk model, our case-control analysis had 90% power to detect an OR of 1.5 to IBD and 1.6 for CD based on case-control analysis of the tagging SNPs with $\alpha = 0.005$ to allow for multiple testing. (Table 5-7) By comparison, the index paper by McGovern et al. quoted an OR 2.0 for CD.[459] Effect sizes excluded (calculated as OR), are given in Tables 5-x and 5-x for the IBD population overall, as well as for the early onset IBD cohort.

In addition, we have revised the power calculation concerning the ***haplotype analysis***, and are able to demonstrate that for the nine-marker haplotypes with frequency >5% (77 % of haplotypes), our study was adequately powered to exclude OR of 1.8 for IBD with 90% power. For this calculation $\alpha$ was again set at 0.005 to stringently correct for multiple testing of the 9 most common haplotypes. (Table 5-9)

For all analyses, the 99.5% CIs are given to ensure compatibility with the statistical threshold of significance, $\alpha = 0.005$, as obtained after Bonferroni correction.

The true power of this study to demonstrate an effect of each SNP on disease susceptibility is best estimated by integrating case-control analysis and TDT analysis. Results of this analysis, which was first described by Kazeem and Farall, are given in Table 5-11.

### *5.1.7 NOD1/CARD4+32656 insertion/deletion polymorphism*

#### 5.1.7.1 NOD1/CARD4+32656: Case – control analysis Scottish IBD

Genotyping was successful in 95% of Scottish IBD patients and 90% of Scottish controls. All Scottish IBD patients (childhood onset + adult onset) were analysed in a case-control analysis with 1372 control samples. The case-control analysis had 80% power to detect an effect size of OR 1.21 for IBD and OR 1.27 for CD/UC. Results of the Scottish case-control analysis (including stratification for age at diagnosis) are shown in Tables 5-3 and 5-4.

None of the genotypes studied differed significantly between IBD (p=0.96), CD (p=0.81) or UC (p=0.81) patients and controls. Subgroup analyses of genotype frequencies per age at diagnosis (according to Montreal classification) of IBD, CD, UC and IBDU in patients versus controls, showed no significant differences (p>0.15 for all age groups). (Table 5-4)

Using the Kruskal-Wallis test, no influence of NOD1/CARD4+32656 genotype on age at diagnosis of IBD was observed in Scottish IBD patients (CD p=0.69, UC p=0.51). No significant effect of NOD2/CARD15 variant carriage on the influence of NOD1/CARD4+32656 genotype on CD susceptibility was observed (p=0.49 and p=0.14 for CD patients carrying a NOD2 variant and NOD2 wildtype CD patients versus controls stratified for NOD2 variant carriage, respectively).

**Table 5-3: Case-control analysis of NOD1/CARD4+32656 in Scottish IBD**

| | *1 Allele | *1*1 | *1*2 | *2*2 | p |
|---|---|---|---|---|---|
| Controls (n=1233) | 1816/2466 (73.6%) | 661 (53.6%) | 494 (40.1%) | 78 (6.3%) | |
| IBD (n=1079) | 1596/2158 (73.9%) | 584 (54.1%) | 428 (39.6%) | 67 (6.3%) | 0.96 |
| CD (n=515) | 758/1030 (73.6%) | 279 (54.2%) | 200 (38.8%) | 36 (7.0%) | 0.81 |
| UC (n=537) | 794/1074 (73.9%) | 287 (53.4%) | 220 (41.0%) | 30 (5.6%) | 0.81 |
| IBDU < 17 (n=27) | 45/54 (83.3%) | 19 (70.4%) | 7 (25.9%) | 1 (3.7%) | 0.22 |

**Table 5-3: Genotype frequencies in Scottish IBD patients and controls of NOD1/CARD4+32656*1: deletion allele - *2: insertion allele. Allelic and genotype frequencies are shown for IBD, CD, UC, IBDU and controls**

**Table 5-4: Genotype frequencies of NOD1/CARD4+32656 in Scottish IBD patients stratified for age at diagnosis**

| | *1 Allele | *1*1 | *1*2 | *2*2 | p |
|---|---|---|---|---|---|
| Controls (n=1233) | 1816/2466 (73.6%) | 661 (53.6%) | 494 (40.1%) | 78 (6.3%) | |
| IBD < 17 (n=364) | 539/728 (74%) | 204 (56.0%) | 131 (36.0%) | 29 (8.0%) | 0.26 |
| IBD 17-40 (n=470) | 680/940 | 237 | 206 | 27 | 0.36 |

| | | | | | |
|---|---|---|---|---|---|
| | (72.3%) | (50.5%) | (43.8%) | (5.7%) | |
| IBD > 40 (n=245) | 377/490 (76.9%) | 143 (58.4%) | 91 (37.1%) | 11 (4.5%) | 0.29 |
| CD < 17 (n=228) | 333/456 (73.0%) | 123 (54.0%) | 87 (38.1%) | 18 (7.9%) | 0.63 |
| CD 17-40 (n=201) | 294/402 (73.1%) | 106 (52.7%) | 82 (40.8%) | 13 (6.5%) | 0.97 |
| CD > 40 (n=83) | 128/166 (77.1%) | 49 (59.0%) | 30 (36.1%) | 4 (4.9%) | 0.60 |
| UC < 17 (n=90) | 135/180 (75%) | 53 (58.9%) | 29 (32.2%) | 8 (9.0%) | 0.27 |
| UC 17-40 (n=269) | 386/538 (71.7%) | 131 (48.7%) | 124 (46.1%) | 14 (5.2%) | 0.18 |
| UC > 40 (n=162) | 249/324 (76.9%) | 94 (58.0%) | 61 (37.6%) | 7 (4.4%) | 0.43 |

**Table 5-4: Genotype frequencies in Scottish IBD patients stratified for age at diagnosis and controls - NOD1/CARD4+32656*1: deletion allele - *2: insertion allele. Allelic and genotype frequencies are shown for IBD, CD, UC and controls. Patients were stratified for age at diagnosis according to the Montreal classification.[516]**

#### 5.1.7.2 NOD1/CARD4+32656: Case-control analysis in Swedish IBD

Genotyping was successful in 99% of Swedish IBD patients and 100% of Swedish controls. In the independent Swedish adult onset case-control analysis no significant association was observed between NOD1/CARD4+32656 and IBD (p=0.64), CD (p=0.75) or UC (p=0.64). (Table 5-5)

**Table 5-5: Case-control analysis of NOD1/CARD4+32656 in Swedish IBD**

| | *1 Allele | *1*1 | *1*2 | *2*2 | p |
|---|---|---|---|---|---|
| Controls (n=277) | 425/554 (76.7%) | 164 (59.2%) | 97 (35.0%) | 16 (5.8%) | |
| IBD (n=632) | 945/1264 (74.5%) | 353 (56.0%) | 239 (37.8%) | 40 (6.3%) | 0.64 |
| CD (n=244) | 367/488 (75.2%) | 137 (56.1%) | 93 (38.1%) | 14 (5.8%) | 0.75 |
| UC (n=388) | 578/776 (74.4%) | 216 (55.7%) | 146 (37.6%) | 26 (6.7%) | 0.64 |

**Table 5-5: Genotype frequencies in Swedish IBD patients and controls. NOD1/CARD4+32656*1: deletion allele - *2: insertion allele. Allelic and genotype frequencies are shown for IBD, CD, UC and controls.**

#### 5.1.7.3 NOD1/CARD4+32656 TDT analysis in early onset IBD trios

Assuming a multiplicative risk model, the TDT analysis was adequately powered (>80%) to detect an effect size of Odds Ratio (OR) 1.5 for IBD, OR 2 for CD and OR 2.5 for UC. No distortion of transmission towards affected offspring was seen for NOD1/CARD4/CARD4 +32656 in IBD (p=0.47), CD (p=0.62) or UC (p=0.40). (Table 5-6)

**Table 5-6: NOD1/CARD4+32656 TDT analysis in Scottish early onset IBD trios**

| TDT NOD1/CARD4+32656 | | IBD | CD | UC |
|---|---|---|---|---|
| # Affected Offspring | | 309 | 202 | 80 |
| Allele Transmission (Observed/ Expected) | Allele1 | 458 / 453.2 | 293/295.6 | 119/116.2 |
| | Allele2 | 160 / 164.8 | 111 / 108.4 | 41/43.8 |
| Global Chi-squared (1 df) | | 0.484 | 0.219 | 0.602 |
| Global p-value for transmission | | 0.471 | 0.628 | 0.408 |

**Table 5-6: Transmission Disequilibrium Test of NOD1/CARD4+32656 in early onset IBD trios (n=315). The ratio of observed transmission over expected transmission for the NOD1/CARD4+32656*1 allele is shown.**

#### 5.1.7.4 Genotype – Phenotype Analysis – Scottish IBD

Detailed phenotypic data allowing the use of the Montreal classification were available for 84% (697/826) and 91% (263/286) of Scottish adults and children with IBD, respectively. Analysing adult CD patients' allelic, carriage and genotype frequencies using the Montreal classification, did not demonstrate any significant influence of NOD1/CARD4+32656 on disease location or behaviour (p>0.08 for all).

In childhood CD, multifactorial analysis (also controlling for NOD2/CARD15 variant carriage) showed that carriage of the NOD1/CARD4+32656 insertion*2 allele was associated with gastric body disease (p=0.01 OR 4.77 CI 1.32-17.27) and ileal disease (p=0.03 OR 3.92 CI 1.11-13.89) at diagnosis and with perianal disease (as defined by the Montreal classification) at follow-up (p=0.03 OR 12.09 CI 1.27-114.61). However, these results are based on relatively small numbers of patients and do not remain significant after stringent correction for multiple testing. NOD1/CARD4+32656 did not significantly influence disease extent in UC (p>0.12

for all). We did not find evidence for a significant association between NOD1/CARD4+32656 and extra-intestinal manifestations of disease in CD/UC.

### *5.1.8 NOD1/CARD4 gene-wide haplotype-tagging approach*

#### 5.1.8.1 NOD1/CARD4 Haplotype-tagging SNPs: Case-control analysis

The haplotype structure of the genomic region spanning NOD1/CARD4 with the position of identified tagging SNPs is provided in Figure 5-1 (based on HapMap data Release 21). Nine SNPs were genotyped in 2296 subjects of which 1323 were IBD patients, including 356 children with IBD and their parents. For all patients and controls, all SNPs conformed to the Hardy-Weinberg equilibrium ($p>0.01$).

We assessed association with IBD, CD and UC by using single- and multiple marker case-control analysis, Transmission Disequilibrium Testing (TDT) and log-likelihood analysis.

No significant associations were observed between any of the nine NOD1/CARD4 SNPs studied and IBD, CD or UC when assessing allelic and genotype frequencies (after Bonferroni correction $p>0.05$ for all) (Table 5-7). Although, the rs932272 showed weak association with IBD and CD ($p=0.04$ in combined adults + paediatric IBD/CD, $p=0.02$, $p=0.01$ in paediatric IBD and CD, respectively), these findings did not retain significance after Bonferroni correction.

Similarly, when disease susceptibility was studied in case-control analysis of haplotype frequencies, no significant effect of NOD1/CARD4 on IBD susceptibility was observed (corrected $p>0.05$ for IBD, CD and UC) (Table 5-9).

After stratification of CD patients and controls for carriage of the three common NOD2/CARD15 variant alleles (R702W, G908R and Leu1007fsinsC), none of the haplotype variants of NOD1/CARD4 was associated with CD (corrected $p>0.05$).

The two-marker haplotype rs2970500/rs2075820 overlaps with the area of strongest association in the report by McGovern et al (rs6958571/rs2907748).[459] Case-control analysis of this two-marker haplotype was non-significant in IBD, CD and UC ($p>0.16$, $p>0.33$ and $p>0.11$, respectively).

We also assessed whether germline variation of NOD1/CARD4 influences the phenotype of IBD. We used the Montreal guidelines for classification of disease behaviour and disease location of CD and UC in our analyses. No significant influence of NOD1/CARD4 haplotype variants on phenotype was evident.[516] Stratification for NOD2/CARD15 variant carriage (R702W, G908R or Leu1007fsinsC) did not influence the observed lack of effect of NOD1/CARD4 on phenotype in CD or UC.

#### 5.1.8.2 NOD1/CARD4 Haplotype-tagging SNPs: childhood onset IBD case-control analysis

We assessed in detail whether the haplotypic variants of NOD1/CARD4 were associated with increased susceptibility in our cohort of 356 children.[459] Sub-group susceptibility analysis of patients with early onset IBD (diagnosed at < 17 years of age), did not show any of the tagging SNPs (either assessed as single – marker or in haplotypes) to be associated significantly with IBD, CD or UC (corrected $p>0.05$ for all) (Table 5-8). Also in early onset CD patients, NOD2/CARD15 variant carriage did not play a significant role in determining the contribution of NOD1/CARD4 to disease susceptibility (corrected $p>0.05$).

Genotype-phenotype analysis in childhood onset CD and UC, based on the Montreal classification, did not show any significant effect of NOD1/CARD4 haplotypic variation on disease phenotype at diagnosis.

#### 5.1.8.3 NOD1/CARD4 Haplotype-tagging SNPs: Transmission Disequilibrium Test

Results of single marker TDT analysis for IBD, CD and UC are shown in Table 5-10. No significant distortion of transmission was observed for any of the markers, after

correction for multiple analyses. Multiple marker TDT, followed by permutation analysis, in IBD, CD and UC trios yielded no significant distortion of transmission of any of the haplotypes studied (p=0.74 for IBD, p=0.77 for CD and p=0.68 for UC - haplotype frequencies > 2% were studied comprising all 9 tagging SNPs). The two-marker haplotype rs2970500/rs2075820 which overlaps with the area of strongest association in the report by McGovern et al (rs6958571/rs2907748), showed no distortion of transmission in IBD (p=0.25), CD (p=0.47) and UC (p=0.42).[459]

#### 5.1.8.4 Log-likelihood analysis

Using log-likelihood analysis (heterogeneity model – 1000 permutations), set $\chi^2$ statistics (degrees of freedom=511) were reached 103, 396 and 35 times for IBD, CD and UC, respectively (corresponding uncorrected p-values 0.10, 0.39 and 0.03).

**Table 5-7 - Genotype frequencies of nine NOD1/CARD4 haplotype-tagging SNPs in healthy controls and patients with IBD, CD and UC. p-values were obtained by 2x2 $\chi^2$-analysis of allelic frequency. Uncorrected p-values and Odds Ratio (OR) (99.5% Confidence Interval (CI)) are shown, followed by effect with OR excluded with 90% power using a multiplicative risk model [in brackets], assuming a population risk of 0.004 for IBD (0.002 CD/UC). To allow for multiple testing, α was set to 0.005 for the power calculation.**

**Table 5-8: Genotype frequencies of nine NOD1/CARD4 haplotype-tagging SNPs in healthy controls and children/adolescents with IBD, CD and UC, diagnosed before age 17. p-values were obtained by 2x2 $\chi^2$-analysis of allelic frequency. Uncorrected p-values and OR (99.5% CI) are shown, followed by effect with Odds Ratio (OR) excluded with 90% power using a multiplicative risk model [in brackets], assuming a population risk of 0.00023 for IBD (0.00014 CD, 0.00009 UC). To allow for multiple testing, α was set to 0.005 for the power calculation.**

**Table 5-9: Haplotypes consisting of 9 NOD1/CARD4 tagging SNPs (haplotype frequency >1%) in Scottish IBD patients (adult + pediatric combined) and controls. p-values, Odds Ratios (OR) and 99.5% Confidence Intervals (CI) are given. Effect sizes (90% power) excluded for each of the haplotypes are given (expressed as OR). For this power calculation, α was set to 0.005 to account for multiple testing of the 9 most common haplotypes.**

**Table 5-10: Results of TDT analysis in families with early onset IBD – analysis using FBAT software package. Odds Ratios (ORs) (99.5% Confidence Interval) for each of the SNPs were calculated using TDTHAP.[533]**

**Table 5-11: A combined estimate of disease-marker association by integration of TDT and case-control analysis is given. This calculation is based on our TDT cohort of early onset IBD and our samples of adult onset IBD patients in the case-control analysis.[535] OR, Variance (Var) and 99.5 % CI are shown. Allelic frequencies of 1 and 2 allele in IBD cases and controls are given: a = (IBD cases allele 1), b = (IBD cases allele 2), c = (controls allele 1), d = (controls allele 2).**

**Table 5-7: NOD1/CARD4 haplotype tagging SNPs genotype and allelic frequencies in Scottish IBD**

| | | rs2970500 | rs2075820 | rs7789045 | rs1558066 | rs4720004 | rs932272 | rs2709799 | rs2529445 | rs38403 |
|---|---|---|---|---|---|---|---|---|---|---|
| Control | 11 | 191/345 (55.4%) | 174/331 (52.6%) | 77/338 (22.8%) | 242/355 (68.2%) | 255/343 (74.4%) | 81/327 (24.8%) | 26/343 (7.6%) | 198/342 (57.9%) | 121/344 (35.2%) |
| | 12 | 141/345 (40.9%) | 129/331 (39.0%) | 186/338 (55.0%) | 104/355 (29.3%) | 81/343 (23.6%) | 185/327 (56.6%) | 152/343 (44.3%) | 132/342 (38.6%) | 167/344 (48.5%) |
| | 22 | 13/345 (3.7%) | 28/331 (8.4%) | 75/338 (22.2%) | 9/355 (2.5%) | 7/343 (2.0%) | 61/327 (18.6%) | 165/343 (48.1%) | 12/342 (3.5%) | 56/344 (16.3%) |
| | Allele 2 | 167/690 (24.2%) | 185/662 (27.9%) | 336/676 (49.7%) | 122/710 (17.2%) | 95/686 (13.8%) | 307/654 (46.9%) | 482/686 (70.3%) | 156/684 (22.8%) | 279/688 (40.6%) |
| IBD | 11 | 668/1214 (55.0%) | 674/1203 (56.0%) | 325/1195 (27.2%) | 847/1223 (69.2%) | 873/1213 (72.0%) | 379/1187 (31.9%) | 96/1233 (7.8%) | 782/1233 (63.4%) | 395/1184 (33.3%) |
| | 12 | 468/1214 (38.6%) | 454/1203 (37.7%) | 619/1195 (51.8%) | 352/1223 (28.8%) | 315/1213 (26.0%) | 606/1187 (51.1%) | 487/1233 (39.5%) | 399/1233 (32.4%) | 613/1184 (51.8%) |
| | 22 | 78/1214 (6.4%) | 75/1203 (6.3%) | 251/1195 (21.0%) | 24/1223 (2.0%) | 25/1213 (2.0%) | 202/1187 (17.0%) | 650/1233 (52.7%) | 52/1233 (4.2%) | 176/1184 (14.9%) |
| | Allele 2 | 624/2428 (25.7%) | 604/2406 (25.1%) | 1121/2390 (46.9%) | 400/2446 (16.4%) | 365/2426 (15.0%) | 1010/2374 (42.5%) | 1787/2466(72.4%) | 503/2466 (20.4%) | 965/2368 (40.7%) |
| | *P* | 0.42, OR 1.08 (0.81–1.43) [1.47] | 0.13, OR 0.86 (0.65–1.14) [1.47] | 0.19, OR 0.89 (0.70–1.14) [1.42] | 0.60, OR 0.94 (0.68–1.29) [1.57] | 0.43, OR 1.10 (0.77–1.56) [1.60] | 0.04, OR 0.84 (0.65–1.07) [1.42] | 0.25, OR 1.11 (0.85–1.45) [1.45] | 0.17, OR 0.87 (0.64–1.16) [1.50] | 0.92, OR 1.01 (0.78–1.29) [1.42] |
| CD | 11 | 338/607 (55.7%) | 331/596 (55.5%) | 156/595 (26.2%) | 418/613 (68.2%) | 438/609 (71.9%) | 192/589 (32.6%) | 46/614 (7.5%) | 385/614 (62.7%) | 207/609 (34.0%) |
| | 12 | 232/607 (38.2%) | 225/596 (37.8%) | 307/595 (51.6%) | 186/613 (30.3%) | 160/609 (26.3%) | 302/589 (51.3%) | 241/614 (39.3%) | 202/614 (32.9%) | 313/609 (51.4%) |
| | 22 | 37/607 (6.1%) | 40/596 (6.7%) | 132/595 (22.2%) | 9/613 (1.5%) | 11/609 (1.8%) | 95/589 (16.1%) | 327/614 (53.2%) | 27/614 (4.4%) | 89/609 (14.6%) |
| | Allele 2 | 306/1214 (25.2%) | 305/1192 (25.6%) | 571/1190 (48.0%) | 204/1226 (16.6%) | 182/1218 (14.9%) | 492/1178 (41.8%) | 895/1228 (72.9%) | 256/1228 (20.8%) | 491/1218 (40.3%) |
| | *P* | 0.62, OR 1.05 (0.77–1.44) [1.55] | 0.26, OR 0.88 (0.65–1.20) [1.55] | 0.47, OR 0.93 (0.71–1.22) [1.47] | 0.75, OR 0.96 (0.67–1.36) [1.65] | 0.51, OR 1.09 (0.74–1.60) [1.67] | 0.03, OR 0.81 (0.61–1.06) [1.47] | 0.22, OR 1.14 (0.84–1.52) [1.52] | 0.31, OR 0.89 (0.64–1.23) [1.57] | 0.91, OR 0.99 (0.75–1.30) [1.50] |
| UC | 11 | 308/576 (53.5%) | 326/576 (56.6%) | 157/569 (27.6%) | 406/578 (70.2%) | 412/573 (71.9%) | 177/569 (31.1%) | 48/587 (8.2%) | 376/589 (63.8%) | 179/544 (32.9%) |
| | 12 | 229/576 (39.8%) | 217/576 (37.7%) | 300/569 (52.7%) | 157/578 (27.2%) | 147/573 (25.7%) | 290/569 (51.0%) | 230/587 (39.2%) | 188/589 (31.9%) | 281/544 (51.7%) |
| | 22 | 39/576 (6.7%) | 33/576 (5.7%) | 112/569 (19.7%) | 15/578 (2.6%) | 14/573 (2.4%) | 102/569 (17.9%) | 309/587 (52.6%) | 25/589 (4.3%) | 84/544 (15.4%) |
| | Allele 2 | 307/1152 (26.6%) | 283/1152 (24.6%) | 524/1138 (46.0%) | 187/1156 (16.2%) | 175/1146 (15.3%) | 494/1138 (43.4%) | 848/1174 (72.2%) | 238/1178 (20.2%) | 449/1088 (41.3%) |
| | *P* | 0.24, OR 1.13 (0.83–1.55) [1.55] | 0.11, OR 0.84 (0.61–1.14) [1.55] | 0.13, OR 0.86 (0.65–1.13) [1.47] | 0.57, OR 0.93 (0.65–1.33) [1.65] | 0.40, OR 1.12 (0.76–1.65) [1.67] | 0.14, OR 0.87 (0.65–1.14) [1.47] | 0.36, OR 1.10 (0.81–1.48) [1.52] | 0.18, OR 0.86 (0.61–1.18) [1.57] | 0.76, OR 1.03 (0.78–1.36) [1.50] |

**Table 5-8: NOD1/CARD4 haplotype tagging SNPs genotype and allelic frequencies in Scottish children with IBD**

| | | rs2970500 | rs2075820 | rs7789045 | rs1558066 | rs4720004 | rs932272 | rs2709799 | rs2529445 | rs38403 |
|---|---|---|---|---|---|---|---|---|---|---|
| Control | 11 | 191/345 (55.4%) | 174/331 (52.6%) | 77/338 (22.8%) | 242/355 (68.2%) | 255/343 (74.4%) | 81/327 (24.8%) | 26/343 (7.6%) | 198/342 (57.9%) | 121/344 (35.2%) |
| | 12 | 141/345 (40.9%) | 129/331 (39.0%) | 186/338 (55.0%) | 104/355 (29.3%) | 81/343 (23.6%) | 185/327 (56.6%) | 152/343 (44.3%) | 132/342 (38.6%) | 167/344 (48.5%) |
| | 22 | 13/345 (3.7%) | 28/331 (8.4%) | 75/338 (22.2%) | 9/355 (2.5%) | 7/343 (2.0%) | 61/327 (18.6%) | 165/343 (48.1%) | 12/342 (3.5%) | 56/344 (16.3%) |
| | Allele 2 | 167/690 (24.2%) | 185/662 (27.9%) | 336/676 (49.7%) | 122/710 (17.2%) | 95/686 (13.8%) | 307/654 (46.9%) | 482/686 (70.2%) | 156/684 (22.8%) | 279/688 (40.5%) |
| IBD < 17 years | 11 | 190/323 (58.8%) | 189/321 (58.9%) | 98/325 (30.2%) | 224/332 (67.5%) | 236/331 (71.3%) | 110/321 (34.3%) | 22/330 (6.7%) | 216/328 (65.9%) | 106/330 (32.1%) |
| | 12 | 107/323 (33.2%) | 112/321 (34.9%) | 166/325 (51.0%) | 103/332 (31.0%) | 91/331 (27.5%) | 160/321 (49.8%) | 128/330 (38.8%) | 102/328 (31.1%) | 177/330 (53.6%) |
| | 22 | 26/323 (8.0%) | 20/321 (6.2%) | 61/325 (18.8%) | 5/332 (1.5%) | 4/331 (1.2%) | 51/321 (15.9%) | 180/330 (54.5%) | 10/328 (3.0%) | 47/330 (14.3%) |
| | Allele 2 | 159/646 (24.6%) | 152/642 (23.7%) | 288/650 (44.3%) | 113/664 (17.0%) | 99/662 (15.0%) | 262/642 (40.8%) | 488/660 (73.9%) | 122/656 (18.6%) | 271/660 (41.1%) |
| | *P* | 0.86, OR 1.02 (0.71–1.46) [1.65] | 0.07, OR 0.80 (0.56–1.14) [1.62] | 0.04, OR 0.80 (0.59–1.09) [1.57] | 0.93, OR 0.99 (0.66–1.47) [1.72] | 0.56, OR 1.09 (0.70–1.69) [1.80] | 0.02, OR 0.78 (0.56–1.06) [1.57] | 0.13, OR 1.20 (0.85–1.69) [1.60] | 0.05, OR 0.77 (0.52–1.13) [1.67] | 0.84, OR 1.02 (0.74–1.39) [1.57] |
| CD < 17 years | 11 | 117/211 (55.5%) | 125/209 (59.8%) | 62/213 (29.10%) | 146/223 (65.5%) | 158/222 (71.2%) | 78/214 (36.4%) | 16/216 (7.4%) | 144/217 (66.4%) | 71/220 (32.3%) |
| | 12 | 78/211 (37.0%) | 70/209 (33.5%) | 107/213 (50.2%) | 75/223 (33.6%) | 62/222 (27.9%) | 102/214 (47.7%) | 75/216 (34.7%) | 64/217 (29.5%) | 116/220 (52.7%) |
| | 22 | 16/211 (7.5%) | 14/209 (6.7%) | 44/213 (20.7%) | 2/223 (0.9%) | 2/222 (0.9%) | 34/214 (15.9%) | 125/216 (57.9%) | 9/217 (4.1%) | 33/220 (15.0%) |
| | Allele 2 | 110/422 (26.1%) | 98/418 (23.4%) | 195/426 (45.8%) | 79/446 (17.7%) | 66/444 (14.9%) | 170/428 (39.7%) | 325/432 (75.2%) | 82/434 (18.9%) | 182/440 (41.4%) |
| | *P* | 0.48, OR 1.10 (0.74–1.64) [1.72] | 0.10, OR 0.79 (0.52–1.18) [1.70] | 0.20, OR 0.85 (0.60–1.21) [1.65] | 0.81, OR 1.04 (0.66–1.62) [1.82] | 0.63, OR 1.09 (0.66–1.76) [1.90] | 0.01, OR 0.74 (0.52–1.06) [1.65] | 0.07, OR 1.29 (0.87–1.90) [1.70] | 0.11, OR 0.79 (0.51–1.21) [1.75] | 0.78, OR 1.03 (0.73–1.46) [1.65] |
| UC <17 years | 11 | 51/81 (63.0%) | 47/81 (58.0%) | 24/81 (29.6%) | 55/77 (71.4%) | 55/78 (70.5%) | 22/78 (28.2%) | 4/82 (4.9%) | 51/81 (63.0%) | 26/79 (32.9%) |
| | 12 | 22/81 (27.1%) | 30/81 (37.0%) | 47/81 (58.0%) | 19/77 (24.7%) | 21/78 (26.9%) | 44/78 (56.4%) | 37/82 (45.1%) | 29/81 (35.8%) | 42/79 (53.2%) |
| | 22 | 8/81 (9.9%) | 4/81 (5.0%) | 10/81 (12.4%) | 3/77 (3.9%) | 2/78 (2.6%) | 12/78 (15.4%) | 41/82 (50.0%) | 1/81 (1.2%) | 11/79 (13.9%) |
| | Allele 2 | 38/162 (23.5%) | 38/162 (23.5%) | 67/162 (41.4%) | 25/144 (17.4%) | 25/156 (16.0%) | 68/156 (43.6%) | 119/164 (72.6%) | 32/162 (19.8%) | 64/158 (40.5%) |
| | *P* | 0.84, OR 0.96 (0.53–1.70) [2.15] | 0.24, OR 0.79 (0.44–1.40) [2.12] | 0.05, OR 0.71 (0.43–1.17) [2.07] | 0.95, OR 1.01 (0.51–1.99) [2.27] | 0.48, OR 1.19 (0.59–2.35) [2.40] | 0.45, OR 0.87 (0.52–1.44) [2.07] | 0.56, OR 1.12 (0.64–1.92) [2.10] | 0.40, OR 0.83 (0.45–1.53) [2.20] | 0.99, OR 1.00 (0.60–1.65) [2.07] |

**Table 5-9: Haplotypes consisting of nine NOD1/CARD4 haplotype -tagging SNPs in Scottish IBD patients (adult + childhood onset combined) and controls**

| IBD ($n = 1323$) | | Control ($n = 370$) | *P*, OR (99.5% CI) | OR for IBD excluded |
|---|---|---|---|---|
| 111111212 | 33.3% | 33.2% | 0.89, OR 1.01 (0.78–1.30) | 1.45 |
| 122112121 | 15.2% | 16.9% | 0.10, OR 0.83 (0.60–1.14) | 1.60 |
| 212222211 | 13.5% | 13.3% | 0.89, OR 1.01 (0.71–1.44) | 1.64 |
| 111111211 | 8.0% | 7.9% | 0.62, OR 1.08 (0.68–1.70) | 1.80 |
| 111112111 | 7.1% | 5.9% | 0.34, OR 1.18 (0.71–1.94) | 1.87 |
| 122111211 | 4.7% | 4.6% | 0.95, OR 1.01 (0.57–1.79) | 2.05 |
| 212111212 | 3.4% | 3.1% | 0.79, OR 1.06 (0.54–2.09) | 2.25 |
| 222112121 | 2.9% | 3.9% | 0.50, OR 0.85 (0.43–1.66) | 2.40 |
| 212211211 | 2.0% | 1.8% | 0.92, OR 0.97 (0.41–2.29) | 2.80 |

**Table 5-10: NOD1/CARD4 haplotype-tagging SNPs: TDT analysis**

| rs-number | Allelic frequency | Number of informative families | Transmissions seen | Transmissions expected | p | OR (99.5% CI) |
|---|---|---|---|---|---|---|
| rs2970500 | 0.738 | 107 | 146 | 142 | 0.48 | 1.25 (0.74–2.09) |
| rs2075820 | 0.751 | 110 | 161 | 151 | 0.07 | 1.48 (0.89–2.45) |
| rs7789045 | 0.530 | 127 | 140 | 130.5 | 0.14 | 1.33 (0.84–2.09) |
| rs1558066 | 0.813 | 94 | 138 | 130.5 | 0.15 | 1.55 (0.86–2.78) |
| rs4720004 | 0.841 | 76 | 115 | 109 | 0.21 | 1.66 (0.88–3.10) |
| rs932272 | 0.575 | 116 | 144 | 131.5 | 0.04 | 1.47 (0.92–2.33) |
| rs2709799 | 0.265 | 116 | 77 | 83 | 0.32 | 0.77 (0.47–1.24) |
| rs2529445 | 0.797 | 103 | 155 | 145.5 | 0.08 | 1.60 (0.93–2.72) |
| rs38403 | 0.596 | 124 | 144 | 151.5 | 0.23 | 0.92 (0.58–1.46) |

**Table 5-11: NOD1/CARD4 haplotype-tagging SNPs: a combined estimate of disease-marker association by integration of TDT and case-control analyses**

| Marker | TDT | | Adult case–control | | | | | | | | Combined estimate | | | | | | Homogeneity *P*-value |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | OR | Var | a | b | c | d | OR | Var | 99.5% CI | | Weight 1 | Weight 2 | Combined Var | Combined OR | 99.5% CI | | |
| RS2970500 | 1.245 | 0.034 | 1317 | 465 | 523 | 167 | 0.904 | 0.011 | 0.675 | 1.211 | 29.3950 | 92.5077 | 0.0082 | 0.977 | 0.758 | 1.260 | 0.13 |
| RS2075820 | 1.481 | 0.032 | 1312 | 452 | 477 | 185 | 1.126 | 0.010 | 0.845 | 1.500 | 31.0388 | 95.4524 | 0.0079 | 1.204 | 0.938 | 1.545 | 0.18 |
| RS7789045 | 1.328 | 0.026 | 907 | 833 | 340 | 336 | 1.076 | 0.008 | 0.834 | 1.388 | 38.2244 | 121.6488 | 0.0063 | 1.132 | 0.906 | 1.413 | 0.25 |
| RS1558066 | 1.553 | 0.043 | 1495 | 287 | 588 | 122 | 1.081 | 0.014 | 0.775 | 1.507 | 23.1134 | 71.1712 | 0.0106 | 1.181 | 0.885 | 1.577 | 0.13 |
| RS4720004 | 1.656 | 0.050 | 1498 | 266 | 591 | 95 | 0.905 | 0.017 | 0.630 | 1.300 | 19.9529 | 60.0770 | 0.0125 | 1.052 | 0.769 | 1.440 | 0.01 |
| RS932272 | 1.468 | 0.027 | 984 | 748 | 347 | 307 | 1.164 | 0.008 | 0.899 | 1.507 | 36.8758 | 117.7533 | 0.0065 | 1.230 | 0.982 | 1.542 | 0.21 |
| RS2709799 | 0.769 | 0.029 | 507 | 1299 | 204 | 482 | 0.922 | 0.010 | 0.699 | 1.216 | 33.9130 | 102.8927 | 0.0073 | 0.882 | 0.694 | 1.121 | 0.35 |
| RS2529445 | 1.600 | 0.036 | 1429 | 381 | 528 | 156 | 1.108 | 0.012 | 0.819 | 1.500 | 27.6923 | 85.9944 | 0.0088 | 1.212 | 0.931 | 1.577 | 0.09 |
| RS38403 | 0.923 | 0.027 | 1014 | 694 | 409 | 279 | 0.997 | 0.008 | 0.770 | 1.290 | 37.4400 | 118.2546 | 0.0064 | 0.978 | 0.781 | 1.225 | 0.68 |

### *Discussion*

We have demonstrated in three independent cohorts (from two Northern European populations with similar small contribution of NOD2/CARD15 variants to CD), that NOD1/CARD4+32656 genotype does not influence susceptibility to IBD, CD or UC, independent of NOD2/CARD15 variant carriage.[337;473;660] We did not observe an effect of NOD1/CARD4+32656 genotype on age at diagnosis of CD or UC. For this variant, we have reliably excluded an effect size of OR greater than 1.21 for IBD and 1.27 for CD/UC in the Scottish IBD population. Our data concur with recently published reports from East-Anglia and Germany, but contrast with the conclusion of McGovern et al. from Oxford.[459;463;464]

The counterintuitive concept of association of the most common allele with disease susceptibility and the influence of the analysis used (TDT versus case-control) was reviewed recently by Mitchell et al.[677] An extensive review of the literature by these authors (identifying 29 TDT and 182 case-control analyses) demonstrated that 59% of TDT analyses showed an association between the most common allele and susceptibility to disease versus only 31% of case-control analyses ($p=3.6x10^{-6}$). In their opinion, the tendency of the TDT to identify the most common allele as a risk factor is a direct effect of the presence of undetected genotyping errors. We have been very careful to minimise the possibility of genotyping error in our study: we used a high-throughput centre for our TaqMan analysis (Wellcome Trust Clinical Research Facility, Western General Hospital campus) and validated the TaqMan assay by comparing it with direct sequencing in 252 samples with 100% concordance between the two analyses.

In accordance with the common disease-common allele hypothesis, the NOD1/CARD4+32656 deletion allele (or a marker in linkage disequilibrium with this variant) may yet be associated with IBD in our study populations, but the effect size is clearly smaller than OR 1.3 (compared with OR 2 (1.2-3.5) for CD, as calculated by McGovern et al.).[459;678] The recent reports from East-Anglia and Germany, both corroborating our results, have further cast doubt on the importance of this NOD1/CARD4 polymorphism in determining susceptibility to IBD.[463;464] In

addition, for a disease-predisposing allele to become the most common allele in a population, clear effects on reproductive ability must be present. The relatively weak association with asthma and IgE of the NOD1/CARD4+32656 insertion allele is probably not important enough in this respect.[434] It is noteworthy that the clustering of chronic inflammatory conditions (e.g. asthma) and IBD, as observed in our childhood onset cohort, has been reported previously in adult IBD.[596]

The study of the contribution of NOD1/CARD4 to IBD has seen replication studies unable to confirm a positive index study, mirroring the experience in other complex diseases.[679] The significant differences between frequencies of the deletion allele in control populations, in the UK studies, are noteworthy, as it seems that these differences, rather than differences between allelic frequencies in patients underlie the discrepancies between centres.[674] Whether true genetic heterogeneity, phenotypic differences between IBD cohorts or population admixture underlie these differences has often been difficult to elucidate. In IBD, only for NOD2/CARD15 has true genetic heterogeneity been demonstrated with certainty.[337;389;474]

The difference (68.2% vs 74.3%, p=0.003) between the control allelic frequency of the NOD1/CARD4+32656 deletion allele in the Oxford and Cambridge studies seems more likely to be explained by a statistical false positive result or population stratification rather than by genetic heterogeneity. Although true heterogeneity between different populations has been demonstrated for NOD2/CARD15, considering the short geographical distance between the two centres, this possibility seems highly unlikely compared to admixture of the control population.[337] Indeed, the single variant TDT analysis, not subject to this confounding element, was negative for CD in the Oxford study. It is noteworthy that in the most recent regression analysis by McGovern and colleagues from Oxford, carriage of the deletion allele of NOD1/CARD4+32656 did not contribute significantly to the diagnosis of CD.[680]

The present study of almost four thousand individuals does not implicate the NOD1/CARD4+32656 variant as important in determining susceptibility to IBD in these genetically homogenous, high-incidence populations.

We therefore pursued the investigation of NOD1/CARD4 as an IBD susceptibility gene further by performing the first gene-wide haplotype tagging approach to assess the contribution of germline variation of NOD1/CARD4 to IBD susceptibility and phenotype. We have made use of data of the second phase of the HapMap project to select SNPs tagging the haplotypic variation of the NOD1/CARD4 gene.[681] This has allowed us to assess the contribution of NOD1/CARD4 to IBD in a more robust manner compared to previously performed single SNP analysis. In this study comprising 2296 subjects, we have used three different methods (case-control analysis [single marker and haplotype], TDT [single marker and haplotype] and log-likelihood analysis) to provide robust support for a lack of association between germline variation of NOD1/CARD4 and IBD. Detailed power calculations, both for single SNP and haplotype analysis, have demonstrated our study was adequately powered to detect an OR of 1.5 and 1.6 for IBD and CD, respectively, after applying stringent criteria for multiple testing ($\alpha = 0.005$). This compares favourably with the calculated OR for CD of 2.0 in the index paper from Oxford, reporting association of NOD1/CARD4+32656 with IBD/CD.[459] Furthermore, access to our large cohort of patients with childhood onset IBD (n=356), characterised by a small but definite contribution of NOD2/CARD15, has permitted us to conclude that common haplotypic variation of NOD1/CARD4 is not an important genetic determinant of early onset disease. Detailed analysis of a two-marker haplotype overlapping with the haplotype of strongest association in the Oxford study, also failed to show any association with IBD, CD or UC.

Notwithstanding these conflicting genetic findings, NOD1/CARD4 has remained a strong functional IBD candidate gene in view of its important role in regulating the innate immune response. The expression pattern of NOD1/CARD4 across the gastrointestinal tract as well as the presence of different splice variants with altered binding properties of the LRR domain, illustrate the importance of NOD1/CARD4 in the ability of the mucosa to sense the microbial content of the gut lumen.[434;670] Functional work has also provided evidence for interaction with Toll-like receptor pathways, again suggestive of a crucial role for NOD1/CARD4 in the innate immune response against bacteria.[431;477] This is further supported by the recent association of

serological markers in IBD patients with the insertion/deletion polymorphism of +32656 of NOD1/CARD4.[682] Indeed, the presence of an altered mucosa-associated microflora in IBD patients is indicative of an innate immune response which is either unable to recognise these micro-organisms appropriately (due to altered expression or altered ligand binding) or unable to orchestrate an effective second line immune reaction to clear them.[91;93;666]

The problems associated with analysis of single SNPs, have stimulated researchers to turn to haplotype-tagging based approaches to assess the gene-wide contribution to the disease under study. The use of a gene-wide haplotype-tagging strategy has already proven successful in the analysis of other IBD candidate genes, in which equivocal evidence had been provided by single SNP analyses in different populations.

Applying this approach to capture the haplotypic variations of the Multi Drug Resistance 1 (MDR1) gene, Ho et al. observed a highly significant association in the Scottish adult population between the common haplotypes of MDR1 and UC ($p=4.22 \times 10(-7)$) but not CD ($p=0.22$).[234] These data provided really clear evidence of an important contribution to susceptibility and phenotype, in the face of inconsistent reports from other populations, and have been supported by meta-analysis.[223] Ho and colleagues also confidently refuted the role of the pregnane X receptor gene (PXR/NR1I2) in IBD in the same population after application of a gene-wide haplotype tagging approach.[243]

In our study, a tagging approach based on HapMap data available for this region, has excluded an important role of common germline NOD1/CARD4 variation in our high-incidence population. Increasing data support the validity of this approach.

There is now good evidence that tagging SNP identification based on sample sizes like those of the HapMap project is appropriate to capture common variants (mean allelic frequency > 5%) with little loss of power if a causal variant is represented by tagging rather than by direct genotyping.[683;684] The transferability of tagging SNPs in genetic association studies across different populations has also been studied.[685] In recent reports by de Bakker et al and Conrad et al, the portability of tagging SNP

selection based on HapMap data across multiple populations was demonstrated, in particular for Caucasian populations.[684;686]

To date, analysis of candidate genes in our early-onset Scottish IBD population has demonstrated only weak effects of the IBD5 locus (OCTN1&2 genes), DLG5 and NOD2/CARD15 on disease susceptibility.[266;278;660] In view of the high incidence of disease, it is likely other important determinants are present, underlying susceptibility in our and other populations.

It is noteworthy that none of the genome-wide association studies (GWAS) to date report association with NOD1/CARD4 in Crohn's disease. Thus, the North-American genome-wide association study did not find any evidence for association between ileal CD and any of the marker SNPs in the region of NOD1/CARD4 (1991 cases and 1214 controls were studied).[687] (Judy Cho, personal communication) Our own detailed analysis of WTCCC data does not reveal association with CD in the British population (1746 cases vs. 10603 controls).[472] (http://www.wtccc.org.uk/) Although we did not have access to the primary data, the studies by Hampe et al. and Libioulle et al. also do not report any association between CD and NOD1/CARD4 SNPs.[471;688] The meta-analysis of the NIDDK/WTCCC/Belgian-French CD cohorts also concluded germline NOD1/CARD4 variation is not associated with CD, with increased power compared our cohort and the separate GWAS cohorts.

In summary, we have examined in detail, both in adult- and childhood-onset IBD, the role of germline variation of NOD1/CARD4. Although the importance of NOD1/CARD4 in maintaining intestinal epithelial homeostasis through adequate recognition of mucosa-associated bacterial flora is undeniable, the present data argue strongly against a contribution of common haplotypic variation to inherited susceptibility to IBD.

# 6 The contribution of germline variation of IL23R to the genetic susceptibility to IBD

## *Introduction*

The impetus for the detailed study of the IL23 signalling pathway in CD susceptibility came from the discovery of association of germline variation of the IL23R gene with CD in the first published Genome Wide Association Study (GWAS) by the North American IBD Genetics Consortium in 2006.[470] The North American IBD genetics consortium performed a genome-wide association study testing 308,332 markers in 567 ileal CD patients and 571 controls of non-Jewish European ancestry.[470] The same authors demonstrated replication of the protective effect of the rare IL23R allele (Arg381Gln, frequency 1.9% in non-Jewish ileal CD vs. 7.0% in non-Jewish controls) in an independent case-control association study of Jewish ileal CD patients, as well as in a family-based association analysis (n=833).

IL23 belongs to a family of cytokines which also consists of soluble cyliary neutrotrophic factor receptor/cardiothrophin-like cytokine (sCNTFR/CLC), CLC/cytokine-like factor-1, IL12, IL27 and IL35.[689] The members of this family, which use interchangeable parts for both ligand and receptor, therefore have signalling characteristics in common as well as discrete functional capacity rendering the elucidation of their biology complex (see Table 6-1).[689] IL23 was described by Oppmann et al. as a novel heterodimeric cytokine, consisting of a p40 subunit (shared with IL12) and a unique p19 subunit (see Figure 6-1).[690] IL23 uses a heterodimeric receptor, composed of IL12Rβ1 and IL23R (similar to gp130) to exert its effect on the terminal differentiation (and stabilisation) of the Th17 T-cell population.[691] IL23R is mainly expressed by T-cells, natural killer cells, and to a lower extent by monocytes and dendritic cell populations.[691]

**Table 6-1: The members of the IL12 family, their components and their impact on immune regulation (adapted from Tato et al.)[689]**

| Designation | Components | Receptor | Primary function |
|---|---|---|---|
| IL12 | p35 and p40 | IL12Rβ1 and IL12Rβ2 | Differentiation of Th1 T-cells |
| IL23 | p19 and p40 | IL12Rβ1 and IL23R | Terminal differentiation of Th17 T-cells |
| IL27 | p28 and EBI3 | WSX-1 and gp130 | Induction of early Th1 as well as strong anti-inflammatory effects on both Th1 and Th17 |
| IL35 | p35 and EBI3 | unknown | Immune suppression |

EBI3: Epstein-Barr virus-induced gene 3

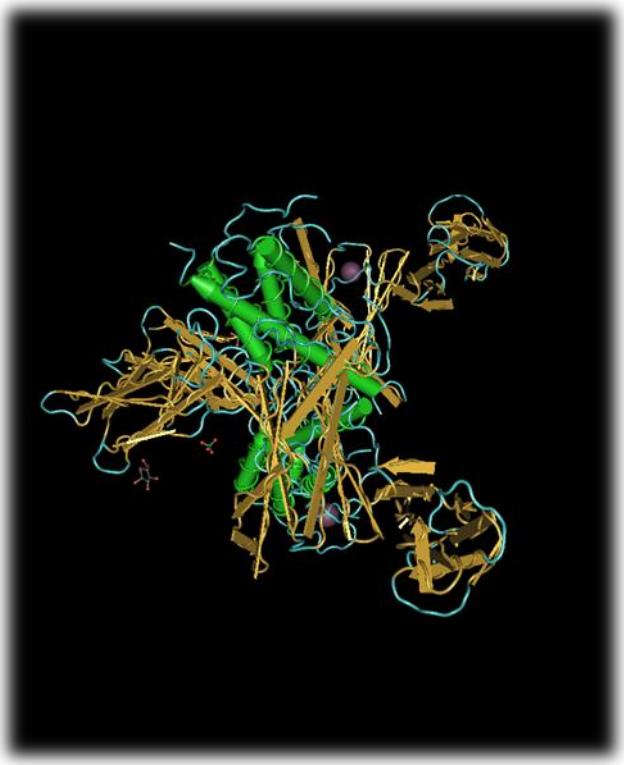


**Figure 6-1: Crystal structure of IL23 – adapted from Beyer et al.[692]**

The signal transduction pathway downstream of IL23R contains JAK2 and STAT3.[693] (see Figure 6-2) Recently, germline variation of both JAK2 and STAT3 has been implicated in the pathogenesis of CD by the meta-analysis of GWAS.[694] It is noteworthy that, before the studies discussed below implicated the IL23 signalling pathway in CD, functional data in the mouse model characterised by ileal inflammation, the Samp1/Yit mouse already demonstrated STAT3 activation via IL6 (which is homologous to the p35 subunit of IL12).[689;695]

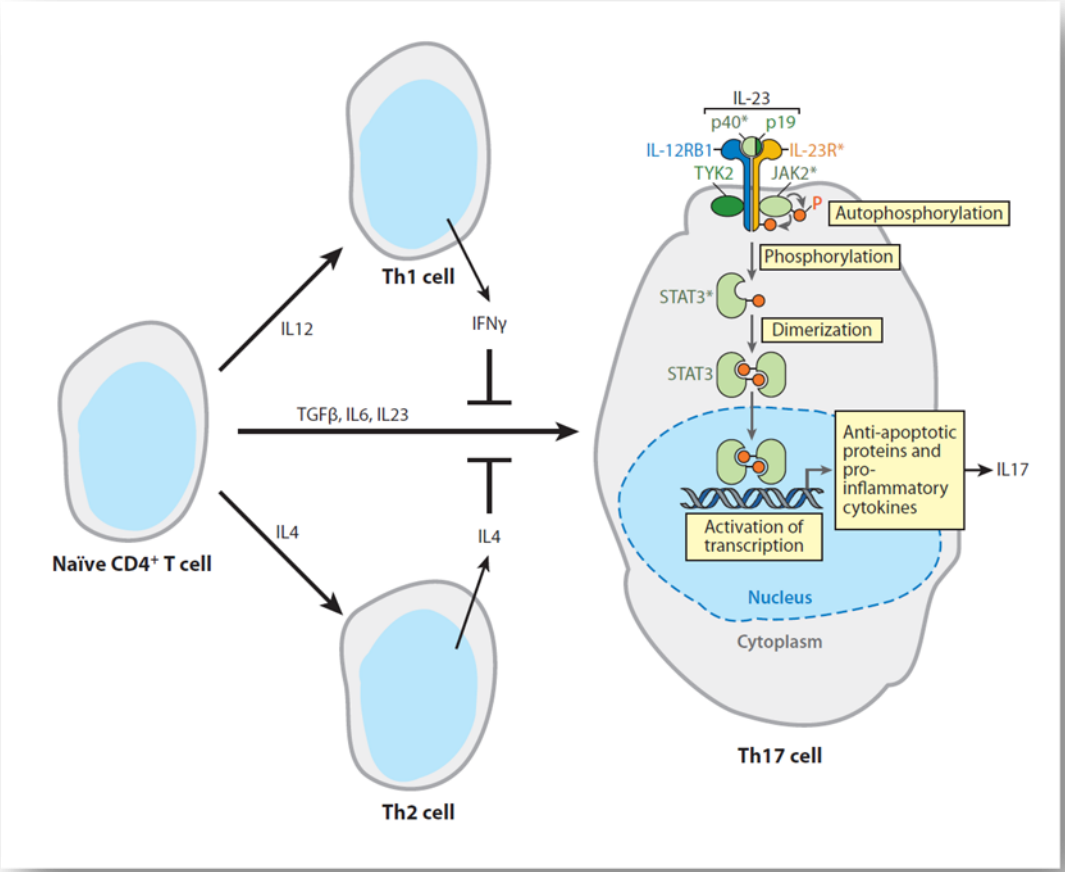



**Figure 6-2: Differentiation of CD4+ T-cell lymphocytes under the influence of IL12, IL4, TGFβ, IL6 and IL23. The signalling pathway downstream of IL23R is illustrated.**

The description of IL23R germline variation as a confirmed determinant of inherited susceptibility to IBD occurred almost simultaneously with the discovery of a new subset of effector T-cells, Th17 cells. IL23 is involved in the terminal differentiation, stabilisation and pathogenic capacitation of this subset of CD4+ T cells. Th17 cells are characterised by the ability to produce a unique set of inflammatory cytokines, such as IL17A, IL17F, IL6 and TNFα under the influence of the transcription factor RORγt (Retinoid-related orphan receptor γ t, also known as RORγ2) and JAK2/STAT3 activation, in response to stimulation with TGFβ and IL6.[689]

This latter observation is particularly relevant for our understanding of IBD pathogenesis. Increasingly a novel dichotomy, reminiscent of the antiquated Th1/Th2 paradigm, is being proposed between T-regulatory cells and Th17 cells. T-regulatory cells maintain intestinal immunotolerance/anergy by secreting TGFβ and IL10. [696-698] It is noteworthy that the tissue with the highest expression of IL23 is the gut, with IL23 perhaps acting as a switch factor inhibiting the function of TGFβ-dependent T-regulatory cells while enhancing the activities of inflammatory Th17 cells, thus promoting chronic inflammation.[689] Zhou et al. recently demonstrated that a precise regulation of the development of naive CD4+ T-cells requires a tight regulation of the cytokine milieu, notably TGFβ, IL6, IL21and IL23.[698] The first therapeutic trial of a monoclonal antibody, ustekinumab, targeting the common p40 subunit of IL12 and IL23 (encoded by the IL12B gene which will be discussed later), has shown promising results in moderate-to-severe CD.[699]

After the index report in *Science* by the North American consortium, the first replication studies of the IL23R association came from the UK. The Wellcome Trust Case Control Consortium (WTCCC) replicated the protective effect of the rare IL23R allele in CD (n=1902) and UC (n=975) versus healthy controls (n=1345).[700] Although initially identified in a cohort of ileal CD, IL23R germline variation has subsequently also been associated with UC.[700-703] Since these initial studies, the association of CD with this locus has been replicated widely and confirmed recently in the meta-analysis of GWAS where the rs11465804 variant (D'=1 and $r^2$=0.88 with the Arg381Gln variant (rs11209026)) achieved a highly significant p value of $10^{-63}$ in the combined cohort, with in the case-control analysis an effect with OR 2.50, second only to the NOD2/CARD15 locus.[694]

The index study by Duerr et al. had already shown that the CD association signal of this locus was not limited to the Arg381Gln variant of the IL23R gene. Raelson et al. and Taylor et al. have since performed studies to confirm this effect of gene-wide variation of IL23R on the genetic susceptibility to CD.[704;705]

In this chapter, we describe the results in our own cohort of childhood onset IBD, using single marker and haplotype-tagging association analyses, both case-control and family-based. In brief, we confirm the role of IL23R as a genetic determinant of childhood onset IBD and we show that, in order to appreciate fully the contribution of this locus to the genetic susceptibility of IBD, a gene-wide approach is more appropriate than single marker analysis.[706;707]

## *Methods*

### *6.1.1 Subjects and genotyping*

#### 6.1.1.1 IL23R Arg381Gln single marker association study

To assess the contribution of the Arg381Gln variant (rs11209026) of IL23R in determining susceptibility and phenotype in childhood onset IBD in Scotland, 1294 subjects comprising 358 IBD patients < 17 years at diagnosis (Table 1), 594 parents and 342 controls were genotyped for rs11209026 G/A using TaqMan (7900HT sequence detection system; Applied Biosystems, Foster City, CA, USA). We also sought to investigate the interaction between carriage of any of the 3 common NOD2/CARD15 variants, previously genotyped by Dr Russell, and carriage of this IL23R variant in determining CD susceptibility.[708;709]

Demographics and IBD phenotype of IBD patients diagnosed before 17 years of age, based on Montreal guidelines for classification of CD/UC, are shown in Table 6-2, and were described in greater detail in Chapter 3.[516;710]

**Table 6-2: Demographics and IBD phenotype of childhood onset patients studied in the single marker analysis of rs11209026 G/A (Arg381Gln).**

| | | | | |
|---|---|---|---|---|
| n | 358 | | | |
| M / F | 205 / 153 | | | |
| Median Age at diagnosis (Q1-Q3) | 11.1 years (8.6 – 12.9) | | | |
| CD / UC / IBD type unclassified | 239 / 88 / 31 | | | |
| Caucasian | 97 % | | | |
| Montreal CD location | L1±L4: 13 | L2±L4: 81 | L3±L4: 131 | L4: 3 |
| Montreal CD behaviour | B1(±p): 214 | B2(±p): 9 | B3(±p): 9 | Bx(p): 27 |
| Montreal UC extent | E1:4 | E2: 17 | E3: 63 | |

CD: L1: ileal; L2: colonic; L3: ileocolonic; L4: upper gastrointestinal tract;

B1: not-stricturing, not-penetrating; B2: stricturing; B3: penetrating;

p: penetrating perianal;

UC: E3: extensive colitis; E2: left-sided colitis; E1: proctitis only

#### 6.1.1.2 IL23R haplotype tagging study

In order to facilitate a better appreciation of the gene-wide contribution of germline variation of *IL23R* to childhood IBD susceptibility and phenotype, and to investigate gene-gene interaction with *NOD2/CARD15*, we performed a detailed haplotype tagging investigation in 709 subjects, consisting of 357 childhood IBD patients (233 CD, 86 ulcerative colitis, 38 IBD- type unspecified) and 352 population-matched controls, as described previously.[711] Eight IL23R haplotype tagging SNPs (rs3762318, rs4655679, rs12041056, rs6656929, rs10889668, rs10489630, rs1004819, rs790631) were identified using HapMap data (minor allelic freq >10%,

haplotype freq >5%, based on solid spine of Linkage Disequilibrium (LD), see Figure 6-2) and genotyped using TaqMan.

### *6.1.2 Analysis & Statistics*

Allelic and genotype frequency comparisons between cases and controls using $\chi^2$ and transmission disequilibrium testing (TDT) were applied to assess association of IL23R rs11209026 with IBD, CD and UC. IL23R rs11209026G/A (Arg381Gln) Transmission Disequilibrium Testing in trios with childhood onset IBD, assessed using FBAT software (version 1.7.3, available from http://www.biostat.harvard.edu/~fbat/fbat.htm ).

Eight IL23R haplotype tagging SNPs were used in allelic, genotype and haplotype frequency case-control analysis, log-likelihood analysis (using PM software as explained previously) and genotype-phenotype analyses (based on the Montreal classification).[516;527]

Subjects were stratified for carriage of any of the three common NOD2/CARD15 variant alleles in both single marker and haplotype tagging analyses.

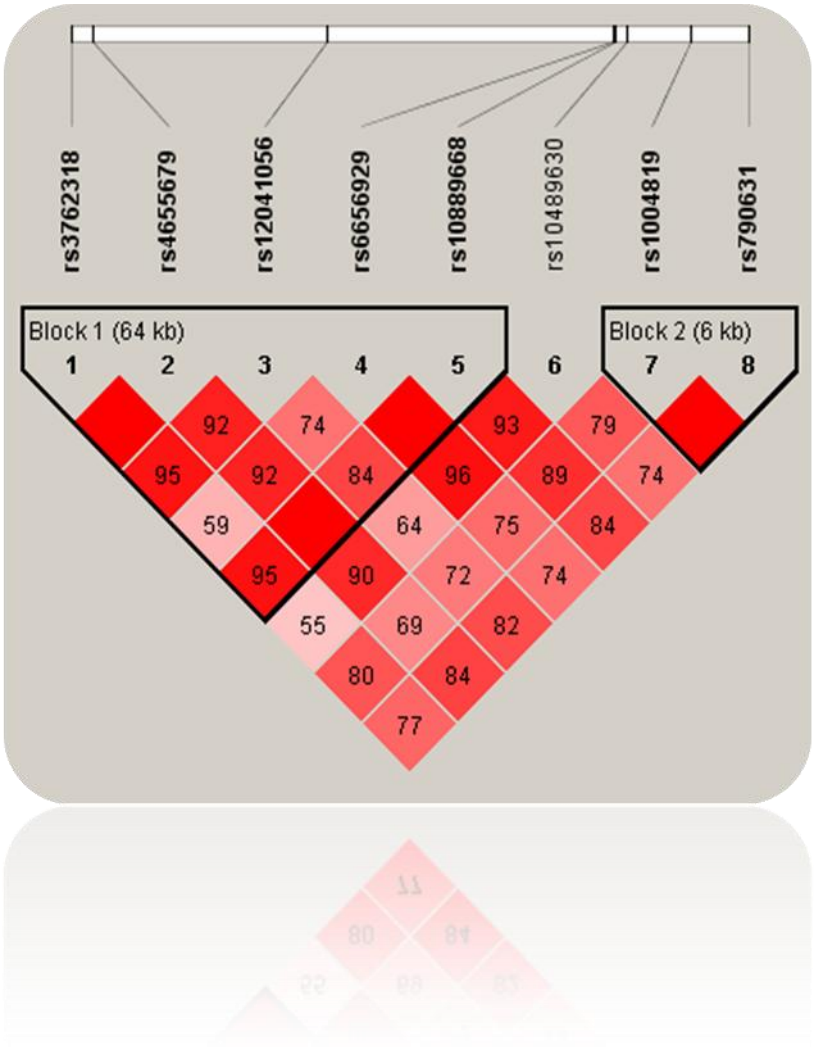



**Figure 6-3: Eight IL23R haplotype tagging single nucleotide polymorphisms (SNPs) (rs3762318, rs4655679, rs12041056, rs6656929, rs10889668, rs10489630, rs1004819, rs790631) were identified using HapMap data (minor allelic freq >10%, haplotype freq >5%, based on solid spine of LD) and visualised using Haploview software.**

**Table 6-3: Eight IL23R haplotype tagging SNPs, the non-synonymous Arg381Gln variant and their respective genomic locations**

| | | |
|---|---|---|
| rs3762318 | 67,369,707 | C/T |
| rs4655679 | 67,372,245 | C/T |
| rs12041056 | 67,399,848 | C/T |
| rs6656929 | 67,433,629 | A/T |
| rs10889668 | 67,433,832 | T/C |
| rs10489630 | 67,435,210 | T/G |
| rs1004819 | 67,442,801 | C/T |
| rs790631 | 67,449,510 | C/T |
| rs11209026 | 67,478,546 | G/A |

**Table 6-3: Eight IL23R haplotype tagging SNPs, the non-synonymous Arg381Gln variant and their respective genomic locations (based on Ensembl release 41 – October 2006) relative to the location of the IL23R gene (67,404,671 to 67,495,250)**

## *Results*

### *6.1.3 IL23R Arg381Gln analysis*

In cases and controls, rs11209026 was in Hardy-Weinberg equilibrium.

The allelic frequency of rs11209026*A differed significantly between IBD / CD cases and controls (2.9% / 3.0% vs 5.5%, p=0.01, OR 0.51 95% CI 0.30-0.88 and p=0.04, OR 0.53 CI 0.28-0.98) (Table 6-3).

**Table 6-4: IL23R rs11209026G/A (Arg381Gln) genotype and allelic frequencies in controls and patients with IBD, CD and UC (diagnosed < 17 years of age). Hardy-Weinberg equilibrium (HWE) p-values are given.**

**Table 6-4: IL23R rs11209026G/A (Arg381Gln) genotype and allelic frequencies in controls and childhood onset IBD/CD/UC.**

| Arg381Gln IL23R | Control | IBD | p | CD | p | UC | p |
|---|---|---|---|---|---|---|---|
| HWE | 0.27 | 0.56 | | 0.63 | | 0.73 | |
| GG | 304/342 (88.9%) | 337/358 (94.1%) | 0.01 | 219/233 (94.0%) | 0.03 | 80/86 (93.0%) | 0.25 |
| GA | 38/342 (11.1%) | 21/358 (5.9%) | 0.01 | 14/233 (6.0%) | 0.03 | 6/86 (7.0%) | 0.25 |
| AA | 0 | 0 | | 0 | | 0 | |
| Allele A | 38/684 (5.5%) | 21/716 (2.9%) | 0.01 | 14/466 (3.0%) | 0.04 | 6/172 (3.5%) | 0.27 |

The GG genotype was associated with increased risk of IBD / CD (p=0.01, OR 2.01 CI 1.15-3.49 and p=0.03, OR 1.96 CI 1.03-3.70). TDT analysis showed significant overtransmission of the G allele for IBD (p=0.004) and CD (p=0.04), with a trend towards significance in UC, hindered by a small number of informative UC families (Table 6-4).

**Table 6-5 - IL23R rs11209026G/A (Arg381Gln) Transmission Disequilibrium Testing in trios with childhood onset IBD, assessed using FBAT software (version 1.7.3, available from http://www.biostat.harvard.edu/~fbat/fbat.htm ).**

**Table 6-5: IL23R rs11209026G/A (Arg381Gln) TDT in trios with childhood onset IBD.**

| Arg381Gln IL23R | # Informative families | Transmissions observed | Transmissions expected | p |
|---|---|---|---|---|
| IBD | 34 | 59 | 50.5 | 0.004 |
| CD | 23 | 39 | 34 | 0.04 |
| UC | 9 | 16 | 13.5 | 0.09 |

In CD, there was no difference (p=0.94) in allelic frequency between NOD2/CARD15 wildtype and NOD2/CARD15 variant carrying patients. However, due to the small numbers of cases and controls carrying this IL23R variant, our study was not adequately powered to formally assess epistasis with NOD2/CARD15. Genotype-phenotype analysis in CD and UC based on the Montreal classification did not demonstrate any significant effect of IL23R rs11209026. Specifically we were not able to show a protective effect against ileal CD (p=0.21).[516]

### *6.1.4 IL23R haplotype tagging investigation*

We observed significant associations of four of the tagging SNPs (rs3762318, rs6656929, rs10889668, rs1004819) with IBD/CD on analysis of allelic/genotype frequency, however without correction for multiple comparisons.

**Table 6-6: Case-control analysis of eight IL23R haplotype tagging SNPs in childhood onset IBD/CD/UC (genotype and allelic frequency)**

**Table 6-6: Case-control analysis of eight IL23R haplotype tagging SNPs in childhood onset IBD/CD/UC (genotype and allelic frequency)**

| **IL23R rs3762318** | | HC HW=0.74 Failed 11 | All IBD 0.76 Failed 16 | p | All CD 0.34 Failed 10 | p | All UC 0.37 Failed 5 | p |
|---|---|---|---|---|---|---|---|---|
| | 11 | 259/352 (73.6%) | 238/357 (66.7%) | **0.04 OR 0.72 (0.52-0.99)** | 147/233 (63.1%) | **0.007 OR 0.61 (0.43-0.88)** | 68/86 | 0.29 |
| Genotype frequencies | 12 | 85/352 (24.1%) | 108/357 (30.3%) | 0.06 | 79/233 (33.9%) | **0.01 OR 1.61 (1.12-2.32)** | 16/86 | 0.27 |
| | 22 | 8/352 (2.3%) | 11/357 (3.1%) | 0.50 | 7/233 | 0.58 | 2/86 | 0.97 |
| Allele | 1 | 603/704 (85.7%) | 584/714 (81.8%) | **0.04 OR 0.75 (0.57-1.00)** | 373/466 (80.0%) | **0.01 OR 0.67 (0.49-0.92)** | 152/172 | 0.35 |

| **IL23R rs4655679** | | HC HW=0.66 Failed 17 | All IBD 0.64 Failed 17 | p | All CD 0.74 Failed 13 | p | All UC 0.55 Failed 4 | p |
|---|---|---|---|---|---|---|---|---|
| | 11 | 162/346 (46.8%) | 188/356 (52.8%) | 0.11 | 121/230 | 0.17 | 46/87 | 0.31 |
| Genotype frequencies | 12 | 152/346 (43.9%) | 139/356 (39.0%) | 0.18 | 93/230 | 0.40 | 33/87 | 0.31 |
| | 22 | 32/346 (9.2%) | 29/356 (8.1%) | 0.60 | 16/230 | 0.32 | 8/87 | 0.98 |
| Allele | 1 | 476/692 (68.8%) | 515/712 (72.3%) | 0.14 | 335/460 | 0.14 | 125/174 | 0.43 |

| **IL23R rs12041056** | | HC HW=0.95 Failed 23 | All IBD 0.11 Failed 22 | p | All CD 0.05 Failed 16 | p | All UC 0.72 Failed 5 | p |
|---|---|---|---|---|---|---|---|---|
| | 11 | 118/340 | 112/351 | 0.43 | 71/227 | 0.39 | 26/86 | 0.43 |
| Genotype frequencies | 12 | 165/340 | 185/351 | 0.27 | 124/227 | 0.15 | 44/86 | 0.66 |
| | 22 | 57/340 | 54/351 | 0.62 | 32/227 | 0.39 | 16/86 | 0.68 |
| Allele | 1 | 401/680 | 409/702 | 0.78 | 266/454 | 0.89 | 96/172 | 0.45 |

| **IL23R rs6656929** | | HC HW=0.10 Failed 18 | All IBD 0.14 Failed 16 | p | All CD 0.69 Failed 9 | p | All UC 0.39 Failed 5 | p |
|---|---|---|---|---|---|---|---|---|
| | 11 | 117/345 (33.9%) | 157/357 (44.0%) | **0.006 OR 1.53 (1.13-2.08)** | 104/234 (44.4%) | **0.01 OR 1.56 (1.11-2.19)** | 37/86 | 0.11 |
| Genotype frequencies | 12 | 180/345 (52.2%) | 150/357 (42.0%) | **0.007 OR 0.66 (0.49-0.89)** | 102/234 (43.6%) | **0.04 OR 0.71 (0.51-0.99)** | 36/86 | 0.08 |
| | 22 | 48/345 (13.9%) | 50/357 (14.0%) | 0.97 | 28/234 | 0.49 | 13/86 | 0.77 |
| Allele | 1 | 414/690 (60.0%) | 464/714 (65.0%) | 0.05 | 310/468 | **0.03 OR 1.31 (1.02-1.67)** | 110/172 | 0.65 |

| **IL23R rs10889668** | | HC HW=0.81 Failed 21 | All IBD 0.66 Failed 22 | p | All CD 0.38 Failed 11 | p | All UC 0.55 Failed 7 | p |
|---|---|---|---|---|---|---|---|---|
| | 11 | 268/342 (78.4%) | 258/351 | 0.13 | 163/232 (70.3%) | **0.02 OR 0.65 (0.45-0.96)** | 71/84 | 0.20 |
| Genotype frequencies | 12 | 69/342 (20.2%) | 87/351 | 0.14 | 65/232 (28.0%) | **0.02 OR 1.54 (1.04-2.27)** | 12/84 | 0.21 |
| | 22 | 5/342 | 6/351 | 0.79 | 4/232 | 0.80 | 1/84 | 0.84 |
| Allele | 1 | 605/684 (88.5%) | 603/702 | 0.15 | 391/464 (84.3%) | **0.04 OR 0.70 (0.50-0.99)** | 154/168 | 0.23 |

| **IL23R rs10489630** | | HC HW=0.06 Failed 18 | All IBD 0.92 Failed 17 | p | All CD 0.40 Failed 7 | p | All UC 0.70 Failed 7 | p |
|---|---|---|---|---|---|---|---|---|
| | 11 | 110/345 | 137/356 | 0.06 | 93/236 | 0.06 | 30/84 | 0.50 |
| Genotype frequencies | 12 | 184/345 | 167/356 | 0.08 | 115/236 | 0.27 | 39/84 | 0.25 |
| | 22 | 51/345 | 52/356 | 0.94 | 28/236 | 0.31 | 15/84 | 0.48 |
| Allele | 1 | 404/690 | 441/712 | 0.19 | 301/472 | 0.07 | 99/168 | 0.92 |

| **IL23R rs1004819** | | HC HW=0.67 Failed 20 | All IBD 0.80 Failed 19 | p | All CD 0.27 Failed 12 | p | All UC 0.30 Failed 6 | p |
|---|---|---|---|---|---|---|---|---|
| | 11 | 167/343 (48.7%) | 151/354 | 0.10 | 91/231 (39.4%) | **0.02 OR 0.69 (0.49-0.96)** | 45/85 | 0.48 |
| Genotype frequencies | 12 | 147/343 | 162/354 | 0.44 | 114/231 | 0.12 | 31/85 | 0.28 |
| | 22 | 29/343 | 41/354 | 0.16 | 26/231 | 0.26 | 9/85 | 0.53 |
| Allele | 1 | 481/686 (70.1%) | 464/708 | 0.06 | 296/462 (64.1%) | **0.03 OR 0.76 (0.59-0.98)** | 121/170 | 0.78 |

| **IL23R rs790631** | | HC HW=0.31 Failed 30 | All IBD 0.43 Failed 15 | p | All CD 0.87 Failed 10 | p | All UC 0.30 Failed 5 | p |
|---|---|---|---|---|---|---|---|---|
| | 11 | 185/333 | 199/358 | 0.99 | 128/233 | 0.88 | 48/86 | 0.96 |
| Genotype frequencies | 12 | 131/333 | 132/358 | 0.50 | 90/233 | 0.86 | 30/86 | 0.44 |
| | 22 | 17/333 | 27/358 | 0.18 | 15/233 | 0.49 | 8/86 | 0.14 |
| Allele | 1 | 501/666 | 530/716 | 0.60 | 346/466 | 0.70 | 126/172 | 0.59 |

In order to assess the validity of our positive findings on allelic/genotype frequency case-control analysis (bearing in mind that a stringent Bonferroni-correction for multiple comparisons is likely to be overconversative), we entered our genotypic data in a haplotype log-likelihood analysis of IBD and CD. We confirmed the gene-wide association signal of the *IL23R* gene (recessive model (1000 permutations) p=0.01 and p=0.002, respectively). No significant associations with UC were demonstrated on any of the above analyses, possibly due to the small UC cohort available and the ensuing lack of power in our investigations.

**Table 6-7: IL23R log-likelihood analysis in childhood onset IBD**

| | |
|---|---|
| 1000 permutations | |
| User-specified model chi-squared statistic (3.60) was reached 677 times<br>Recessive model chi-squared statistic (64.04) was reached 12 times<br>Dominant model chi-squared statistic (49.20) was reached 142 times<br>Model-free chi-squared statistic (64.04) was reached 14 times<br>Heterogeneity model chi-squared statistic (67.92) was reached 16 times | |
| Empirical p-values for these statistics are as follows: | |
| T1 - User specified model: | p-value = 0.6770 |
| T2 - Mendelian recessive model: | p-value = 0.0120 |
| T3 - Mendelian dominant model: | p-value = 0.1420 |
| T4 - Model-free analysis: | p-value = 0.0140 |
| T5 - Heterogeneity model: | p-value = 0.0160 |

**Table 6-8: IL23R log-likelihood analysis in childhood onset CD**

| | |
|---|---|
| 1000 permutations | |
| User-specified model chi-squared statistic (1.44) was reached 871 times<br>Recessive model chi-squared statistic (68.12) was reached 2 times<br>Dominant model chi-squared statistic (51.90) was reached 53 times<br>Model-free chi-squared statistic (68.12) was reached 3 times<br>Heterogeneity model chi-squared statistic (70.36) was reached 3 times | |
| Empirical p-values for these statistics are as follows: | |
| T1 - User specified model: | p-value = 0.8710 |
| T2 - Mendelian recessive model: | p-value = 0.0020 |
| T3 - Mendelian dominant model: | p-value = 0.0530 |
| T4 - Model-free analysis: | p-value = 0.0030 |
| T5 - Heterogeneity model: | p-value = 0.0030 |

Haplotype analysis demonstrated a significant protective effect of the 11221211 haplotype on IBD (2.1% versus 4.4% in healthy controls, p=0.02 OR 0.49 (0.26-0.92)) and CD (2.0%, p=0.02 OR 0.46 (0.22-0.97)). By then extending the haplotype analysis to include the previously genotyped Arg381Gln variant, we were able to show that the protective effect the 11221211 haplotype was independent of the Arg381Gln variant ($r^2$ with tagging SNPs $\leq 0.05$; IBD: p=0.02 OR 0.50 (0.27-0.93); CD: p=0.02 OR 0.46 (0.22-0.96)) (see Figure 6-3)

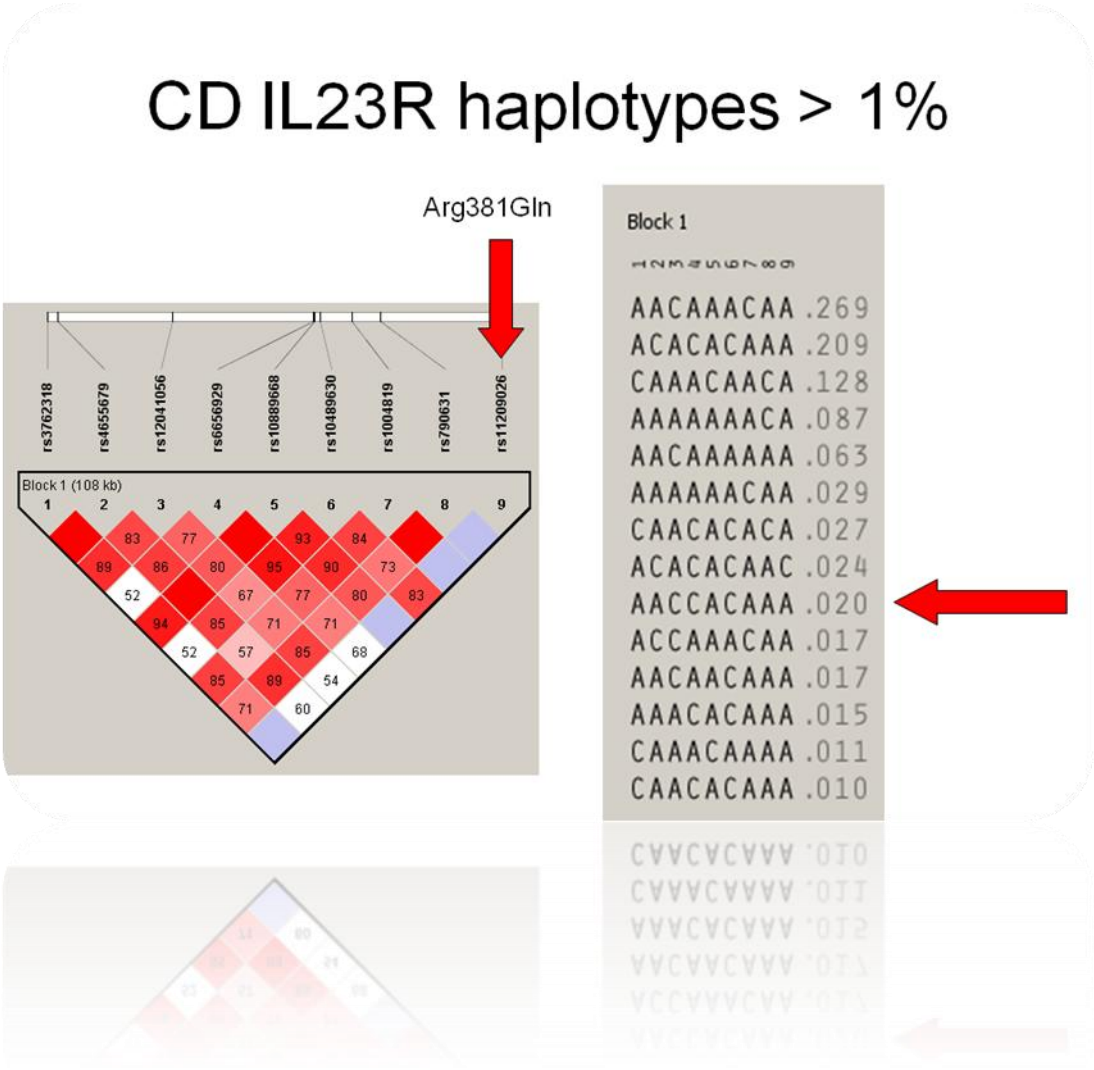



**Figure 6-4: IL23R haplotype analysis including the Arg381Gln SNP. The protective IL23R haplotype, composed of the eight tagging SNPs (11221211 or AACCACAA), is independent of the Arg381Gln variant (with very low (≤0.05) $r^2$ with the tagging SNPs). Note: 1 or A denotes the common, ancestral allele and 2 or C denotes the minor allele.**

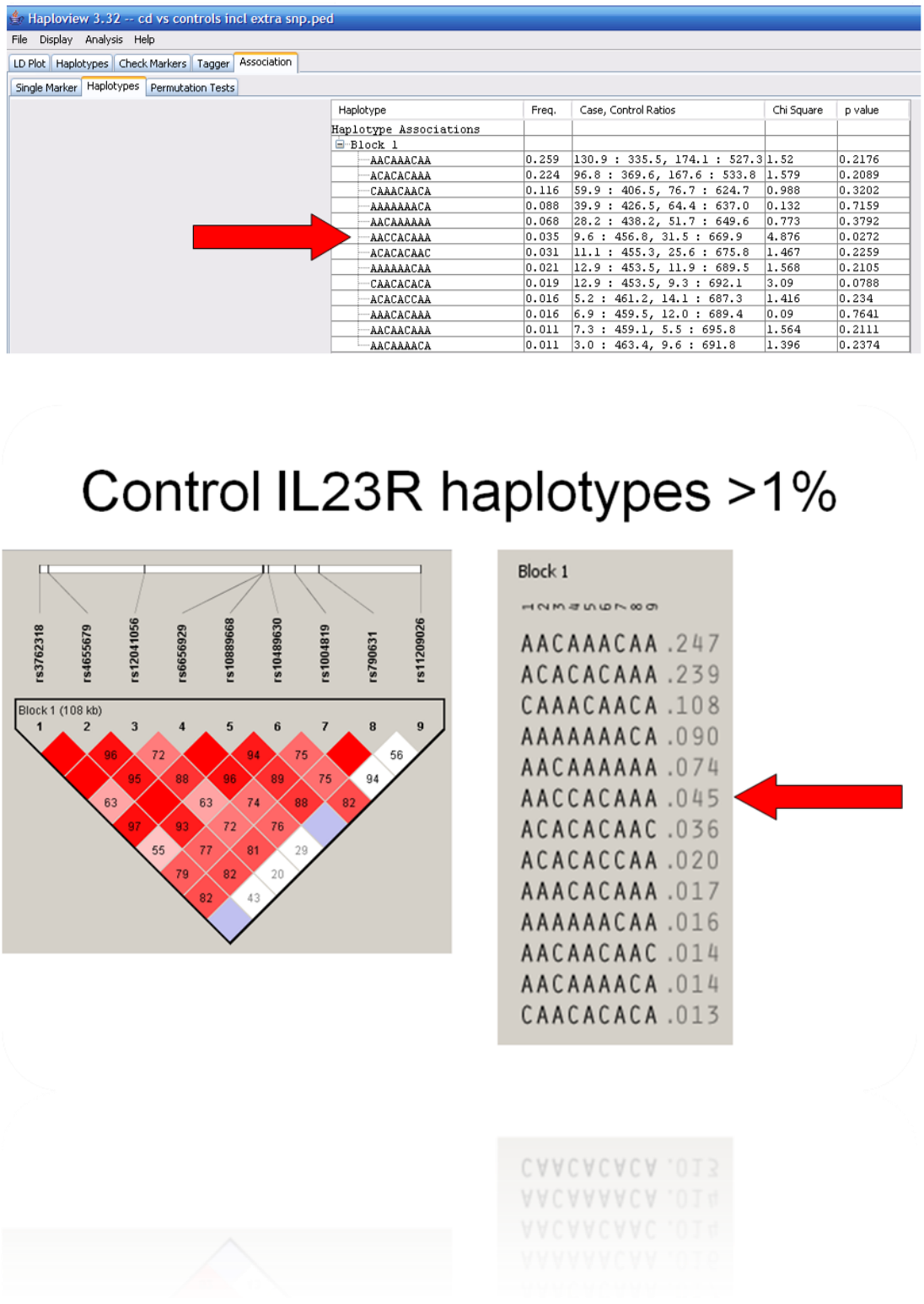






**Figure 6-4: IL23R haplotype analysis including the Arg381Gln SNP.**

When assessing the association signal for each haplotype block (based on solid spine of LD (see Figure 6-4 for CD), we observed association with both 5'- and 3'-end haplotype blocks. After correction for multiple comparisons, no significant genotype-phenotype associations were seen. In homozygous wildtype NOD2/CARD15 children affected by CD and controls, we observed association with a novel risk haplotype (21121212, 4.7% vs 0.9% p=0.002 OR 5.17 (1.49-17.90)).

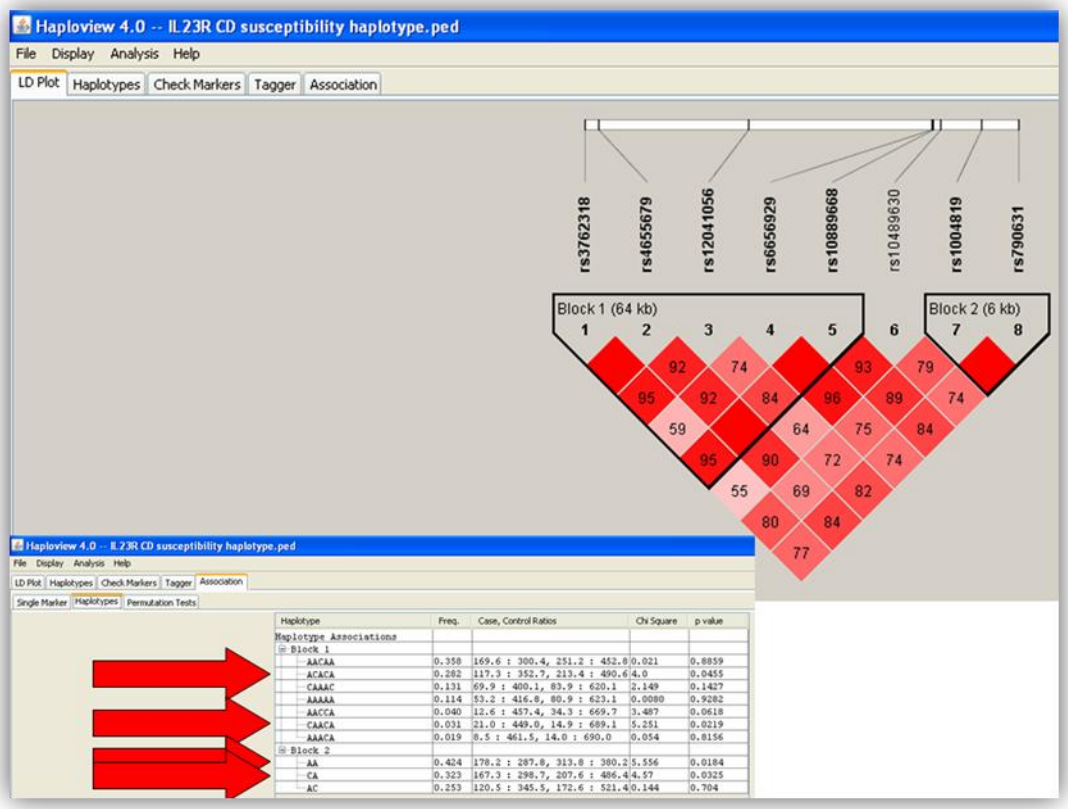



**Figure 6-5: Haplotype susceptibility analysis of eight IL23R tagging SNPs (based on solid spine of LD)**

## *Discussion*

The successful identification of IL23R as a novel IBD susceptibility gene has provided proof of principle for the applicability of *genome-wide association* studies in the genetics of complex diseases. In the present study, we show that IL23R variation influences IBD and CD susceptibility, but not phenotype, in an exclusively paediatric IBD cohort. We have also demonstrated in our high-incidence population of childhood onset IBD that the haplotypic association signals of the IL23R locus are complex and that they can not be simply attributed to LD with the Arg381Gln variant.

Medio 2009, it is pertinent to address the contribution of the IL12/IL23 pathway to the development of IBD at both the genetic (single marker and haplotypic analysis) and the functional level. In addition, we will discuss the novel avenue of pathway-based analysis briefly as a bridge between these two approaches.

In the first phase of this study, we assessed the IL23R Arg381Gln SNP, reported by Duerr et al. in their landmark *Science* publication.[470] We were able to show association with childhood onset IBD and CD using both case-control and TDT analyses. Our study was not adequately powered to comment on the association with UC.

Adequately powered replication studies in populations of North America and Europa, investigating the association of germline variation of IL23R with the genetic susceptibility to IBD (CD as well as UC) have been unequivocally positive, both in adult and childhood onset IBD.[470-472;687;694;700-706;712-732] Yamazaki et al. showed that in Japanese patients affected by CD, inherited variation of IL23R does not determine disease susceptibility to the same degree as in populations of European descent.[733]

Duerr and colleagues had already described in their index publication that the association signal at the IL23R locus is wider than just the Arg381Gln variant.[470] Various other groups have studied multiple variants across this locus and have reported strong(er) associations of other IL23R variants with CD.[701;704;705;718] This complex association signal, accompanied by a large number of IL23R splice variants, have contributed to the frustration of investigators studying the precise impact of germline variation of IL23R on the differentiation of naive CD4+ T-cells and on the development of IBD.[734;735]

Raelson et al. demonstrated that distinct association signals occur at the 5' and 3' regions of the gene.[704] These authors also reported that it is highly unlikely that the non-synonymous Arg381Gln variant fully explains the functional role of this locus in CD etiology, as the CD associated SNP does not occur consistently in all risk and protective haplotypes.

In view of this complex association signal, it may be more appropriate to select a set of tagging SNPs based on HapMap project data.[681] Taylor and colleagues adopted this approach and performed haplotypic association analysis in adult CD.[705] These researchers clearly demonstrated that several haplotype blocks across this locus have distinct risk effects, both increased and decreased.

In addition to being the first IL23R haplotype tagging study in paediatric IBD, our study adds to the literature described above, by assessing the contribution of the Arg381Gln in addition to the association signal provided by the haplotype tagging SNPs. Firstly, we confirm using haplotype association and log-likelihood analyses that *gene-wide* inherited variation of the IL23R locus is a genetic determinant of childhood onset IBD (and CD). Secondly, we show that the *gene-wide* association signal of the eight haplotype tagging SNPs is not altered by inclusion of the Arg381Gln. Indeed, the common 'risk' allele of Arg381Gln (rs11209026G) is

inferred to reside on the 'protective' haplotype (as described in Figure 6-4) without altering the strength of the association, compared with the eight-marker haplotypic protective effect. Indeed, $r^2$ of rs11209026 with the IL23R tagging SNPs is very low ($\leq 0.05$).

Two recent studies have added a further layer of complexity to the elucidation of the functional repercussions of these genetic findings. Wang et al. applied pathway analysis to data from GWAS to date in IBD, including the dataset to which we contributed our samples, of the International Pediatric IBD Genetics Consortium.[736] GWAS typically focus on single-locus analysis, which may not have the power to detect the majority of genuinely associated loci. Wang and colleagues applied pathway analysis, using Affymetrix SNP genotype data from the Wellcome Trust Case Control Consortium (WTCCC), and uncovered significant association between CD and the IL12/IL23 pathway, harboring 20 genes ($p = 8 \times 10^{-5}$). Interestingly, the pathway contains multiple genes (IL12B and JAK2) or homologs of genes (STAT3 and CCR6) that were recently identified as genuine susceptibility genes only through meta-analysis of several GWAS.[694] In addition, the pathway contains other susceptibility genes for CD, including IL18R1, JUN, IL12RB1, and TYK2, which do not reach genome-wide significance by single-marker association tests. The observed pathway-specific association signal was subsequently replicated in three additional GWAS of European and African American ancestry generated on the Illumina HumanHap550 platform, including the paediatric IBD GWAS discussed in Chapter 9. This novel pathway analysis suggests that examination beyond individual SNP hits, by focusing on genetic networks and pathways, is important to unleashing the true power of GWA studies.[736] Further evidence for this new layer of complexity was provided in a recent study by McGovern et al. describing genetic epistasis between IL17 and IL23 pathway genes.[737] These authors showed that a significant association between CD and the widely replicated IL23R variants is only seen in the presence of IL17A or IL17RA variants.

Since its description in 2000 by Oppmann et al., the role of IL23 in the inflammatory response has been widely investigated.[690] In addition to its role in dysregulated inflammation, there is also mounting evidence that IL23 and the Th17 axis mediate beneficial events in host protective immunity and barrier function in the intestine.[738] The discovery of IL23, specifically of its heterodimeric structure (IL12 p40 and an IL23 specific p19 subunit) as shown in Figure 6-1, has led to a re-evaluation of the role of IL12 in inflammatory disease.[738] Following groundbreaking work in the IL23 knockout mouse model which demonstrated that IL23 rather than IL12 was the cytokine driving experimental autoimmune encephalitis, IL23 has been identified as the causative agent in a number of inflammatory disorders in intestinal inflammation.[739-743] IL23 expression is notably increased in the intestine rather than systemically during intestinal inflammation, indicating a tissue-specific role in the inflammatory response.[738;742;744]IL23 is increasingly recognised as a pivotal player in the homeostasis of the intestinal immune response via its role in T-cell differentiation (via stabilisation of the Th17 population and the suppression of T-regulatory cell differentiation) and because of the observation that innate immune pathways can trigger its expression.

Several confirmed IBD susceptibility determinants are also regulators of IL23R expression or signalling. For example, TLR3,4 and 8 stimulation have been shown to differentially regulate IL12 and IL23 in myeloid cells (in favour of IL12).[745] By contrast, stimulation of TLR2 (alone or in combination with NOD2/CARD15 induces expression of IL23, as shown by Van Beelen and colleagues.[745;746] Dectin-1 agonists, such as curdlan, induce a striking IL23 response via CARD9.[747] Veldhoen et al. showed that zymosan, a component of the cell wall of *Saccharomyces cerevesiae* (a TLR2 and dectin-1 agonist), induces IL23 production as well as promoting Th17 responses *in vivo*.[748]

CD is associated with elevations of IL12, IL23, IFN-γ and IL17.[749-754] Together with the overwhelming evidence for a genetic association of the IL12 – IL23 – IL17 axis with CD, their is now a real need to integrate these findings. Schmechel and

colleagues analysed the effect of carriage of any of the CD associated IL23R SNPs on IL22 expression.[755] These authors demonstrated that IL23R genotypes influence IL-22 serum expression, linking genetic CD susceptibility to Th17 cell function. However, NOD2/CARD15 mutations in DCs have been linked to a decreased ability to prime Th17 responses *in vitro:* upon stimulation with TLR2 and NOD2 agonists, a failure to upregulate IL23 may be harmful rather than protective.[738;746] Our own observation that homozygous wildtype NOD2/CARD15 CD is associated with a different IL23R risk haplotype, is perhaps best assessed with these novel functional data in mind. Although an epistatic finding in a subgroup of childhood onset CD patients such as this, would not stand up against stringent correction for multiple testing, it is clear from these *in vitro* studies that functionally none of these proteins act alone. A genetic correlate of these functional interactions is therefore highly likely (as evident from small studies with pairwise assessment of epistasis, e.g. McGovern et al. in the case of IL17 and IL23), even though our latest efforts at assessing epistasis at the genome-wide level have not confirmed this.[694;737]

In summary, we have demonstrated using a single SNP and a gene-wide haplotype tagging strategy that the multiple association signals of the IL23R locus are independent of the Arg381Gln variant in childhood onset IBD and CD. In our high-incidence population characterised by low NOD2/CARD15 variant carriage, we have observed interaction of the IL23R locus with NOD2/CARD15 through the identification of a novel IL23R risk haplotype.

# 7 Autophagy in Northern European IBD: the role of ATG16L1 and IRGM

## *Introduction*

The ability to adapt to environmental change is essential for survival. This is true for the organism as a whole and for individual cells alike. The eukaryotic cell has developed myriad processes to recognise nutritional or microbial changes, both extra- and intracellularly. During cellular stress or starvation, the evolutionarily conserved autophagy response is used to recycle nutrients from non-essential or defunct cell organelles. In addition, the cell uses autophagy to clear its intracellular milieu of misfolded proteins or invading microorganisms.

Autophagy is a cellular pathway involved in protein and organelle degradation and likely evolved as a single cells adaptation to starvation. In times of nutrient starvation the cell can self-digest some non-essential components via autophagy to sustain its minimal growth requirements until a food source becomes available. The recent identification of the importance of autophagy genes in the genetic susceptibility to Crohn's disease suggests that in addition to non-selective, starvation-induced autophagy, the selective, autophagic response to microorganisms may play a crucial role in the pathogenesis of common complex immune-mediated diseases. In this chapter, we will discuss how autophagy has already been implicated in the pathogenesis of an increasing number of medical conditions (notably cancer, neurodegenerative diseases as well as inflammatory bowel disease). This recent progress has in turn led to great interest in the therapeutic potential of manipulation of both selective and non-selective autophagy in established disease.

The term *autophagy* encompasses several distinct processes involving the delivery of portions of the cytoplasm to the lysosome for degradation: chaperone-mediated autophagy, micro-autophagy and macro-autophagy. More recently, the concept of classifying autophagy into selective and non-selective has regained interest.

In *chaperone-mediated autophagy*, proteins that contain a specific pentapeptide motif are translocated directly into the lysosome for degradation and this process

requires the action of cytosolic and lysosomal chaperones. *Micro-autophagy* sequesters cytosol directly at the lysosomal surface while *macro-autophagy* sequesters cytosol into a distinct double membraned structure called the autophagosome which then fuses with and delivers its cargo to the lysosome for degradation.(Figure 8-1) Here we will focus on macro-autophagy (herein referred to as autophagy) and in particular an adaptation to autophagy for the selective removal of pathogens (bacteria, viruses, fungi and parasites) that have entered the cytosol, also termed *xenophagy*.

During autophagy, proteins and organelles that are targeted for destruction are sequestered into large double-membraned vesicles, called autophagosomes, which fuse with lysosomes to form autolysosomes where they are degraded through the action of lysosomal hydrolases.[756] In turn, the resulting macromolecules are released back into the cytosol through membrane permeases. Autophagosomes are formed by expansion of a structure called a phagophore or a pre-autophagosomal structure (PAS) but the origin of the phagophore remains unknown. Genetic studies in the yeast *Saccharomyces cerevisiae* identified the core 17 autophagy (*Atg*) genes that constitute the basic machinery for autophagosome formation and many of the corresponding proteins have homologues in mammals.

Although first described in the 1960s, the last decade really has seen autophagy research come of age through a series of molecular breakthroughs in the regulation of autophagy.[757] Most recently, investigators in North America and Europe, studying the inherited susceptibility to Crohn's disease, have recently identified germline variants of the autophagy genes *ATG16L1* (Autophagy-related 16-like 1) and *IRGM* (immunity-related GTPase family M) to be involved in the genetic predisposition to Crohn's disease.[471;472;687;688;704;758;759] These findings have led researchers in the field of IBD genetics to join the expanding group of autophagy-investigators from a myriad of research disciplines: cancer biology, apoptosis, heart disease, liver disease, myopathies, neurodegenerative diseases, innate and adaptive immunity and even lifespan extension.[756;757;760] A better understanding of macro-autophagy in particular,

offers great hope for the successful manipulation of autophagy. This therapeutic potential is notably for diseases characterized by dysregulation of the immune response against microorganisms (such as Crohn's disease), of intracellular processing of misfolded proteins (such as in Crohn's disease) or of apoptosis (implicated in Crohn's disease and cancer). I will therefore focus on macro-autophagy (herein referred to as autophagy) and specifically on selective autophagy (e.g. induced by misfolded proteins or microorganisms entering the cytosol) rather than non-selective (starvation-induced) autophagy.

**Figure 7-1: The autophagic machinery.**

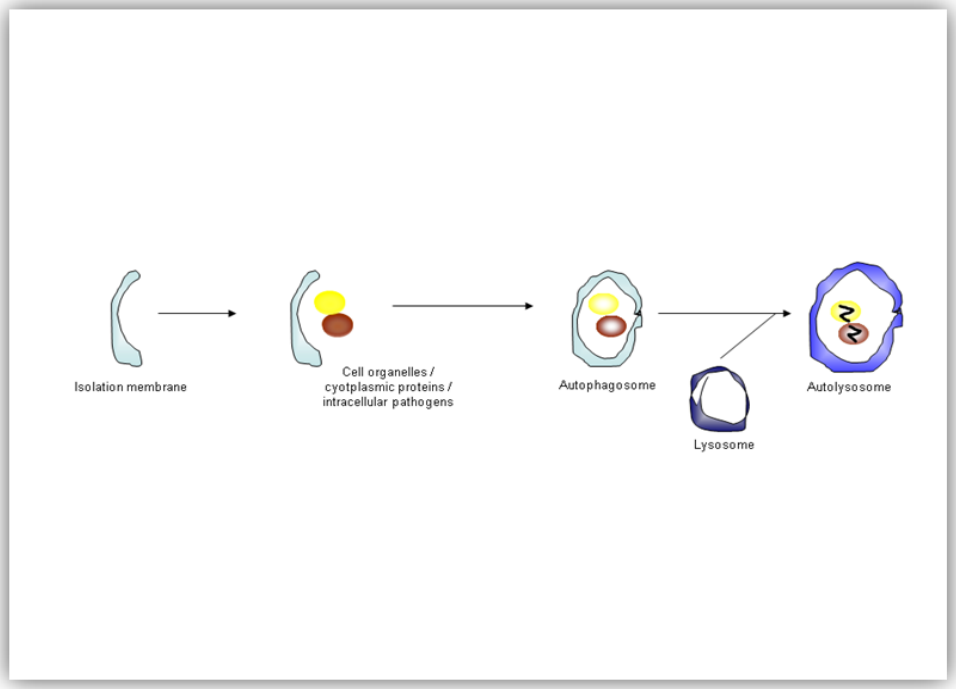



**Figure 7-1: Schematic Depiction of Autophagy. (a, b) cytosolic material is sequestered by an expanding membrane sac, the phagophore. (c) a double membrane vesicle, the autophagosome, then forms at the phagophore and encloses proteins, organelles or pathogens to be degraded. (d) the autophagosome then fuses with the lysosome to form the autolysosome where its contents are degraded.**

### 7.1.1 *The molecular mechanisms of autophagy*

Autophagy can be divided into 3 distinct stages: vesicle nucleation (formation of the so-called phagophore), vesicle elongation (growth and closure of the autophagosome) and fusion of the autophagosome with a lysosome to form an autolysosome.[761] This process is controlled by the Atg (autophagy-related) proteins.[762] Atg proteins form distinct functional complexes, including a protein serine/threonine kinase complex that responds to upstream signals such as mTOR (mammalian Target of Rapamycin), which will be discussed in detail in the section on the regulation of autophagy. Initially cytosolic material is sequestered by an expanding membrane sac, the phagophore or isolation membrane. Although the phagophore assembly site (PAS) is the proposed site for autophagosome formation, the precise mechanisms of isolation membrane formation are unknown. In *Saccharomyces cerevisiae*, the PAS is a peri-vacuolar site to which most core Atg proteins localise.[763;764] Despite the lack of comprehensive studies, the colocalisation of the core Atg proteins to the PAS has also been observed in mammalian cells.[765-767] For a comprehensive review of the core molecular machinery of autophagosome formation readers are referred to Xie and Klionsky.[762]

In brief, autophagosomal membrane-formation and expansion is mediated by two ubiquitin-like protein conjugation systems, the LC3 (Atg8) and Atg12 systems (Figure 2). Initially, the LC3 precursor is processed by the cystine protease Atg4 into the mature form of LC3 (LC3-I), which can then be modified by the ubiquitin E1-like protein Atg7 and the ubiquitin E2-like protein Atg3 to generate a smaller lipidated form of LC3 (LC3-II). Recruitment of LC3-II to the growing isolation membrane depends on the Atg12 system. Atg12 is activated by Atg7 then further modified by the ubiquitin E2-like protein Atg10 and eventually covalently linked to Atg5. The Atg5-Atg12 complex then associates with Atg16 to initiate the elongation stage of isolation membrane formation by recruiting the lipidated LC3-II. Atg16

determines the site of of LC3 lipidation through interaction with the Golgi-resident small GTPase Rab33.[768;769]

LC3 is the only Atg protein in higher eukaryotes that remains associated with the mature autophagosome. The yeast Atg8 has multiple paralogs in mammals termed LC3A (with two splice isoforms, a and b in humans), LC3B, GABARAP, GABARAPL1, and GABARAPL2 (GATE-16), encoded on different chromosomes, with the exception of LC3B and GABARAPL2, which are linked on chromosome 16 in the human genome.[770] The LC3 routinely referred to in the literature is LC3B.[770]

The conversion of LC3-I to the membrane associated, lipidated form LC3-II forms the basis for one of the assays that monitor autophagy levels.[771-773] Upon completion of the autophagosome, the Atg5-Atg12-Atg16 complex dissociates from the outer autophagosomal membrane and is recycled together with LC3 (which then needs further delipidation by Atg4).[774] In the maturation stage, the fraction of LC3 trapped on the luminal membrane of the autophagosome is degraded in the autolysosome.[761;772] Finally, a recycling pathway (comprising Atg2/Atg9/Atg18) mediates the disassembly of Atg-proteins from matured autophagosomes.

Autophagosomal membrane-formation differs between non-selective and selective autophagy.[762] Both these forms of autophagy require specific adaptations of the core autophagic machinery.[762] In non-selective autophagy, autophagosomes contain primarily bulk cytoplasmic contents.[762] In contrast, selective autophagy in yeast (e.g. containing peroxisomes or bacteria) results in autophagosomes with a configuration similar to the cargo it envelops.[775-777] The autophagic machinery is able to respond to information from the cargo to use it as a scaffold. In yeast, adaptor-proteins Atg11 and Atg19, interact with the core machinery (Atg1/LC3/Atg9) to wrap the isolation membrane around the cargo.[778] These adaptations in selective autophagy have also been observed in mammalian cells in response to invading bacteria.[779] Both PIP3, a product of the Vps34/beclin1 complex (see below), and LC3 have been suggested as

such anchors for the envelopment of the substrate by autophagosomal membrane formation.[780;781]

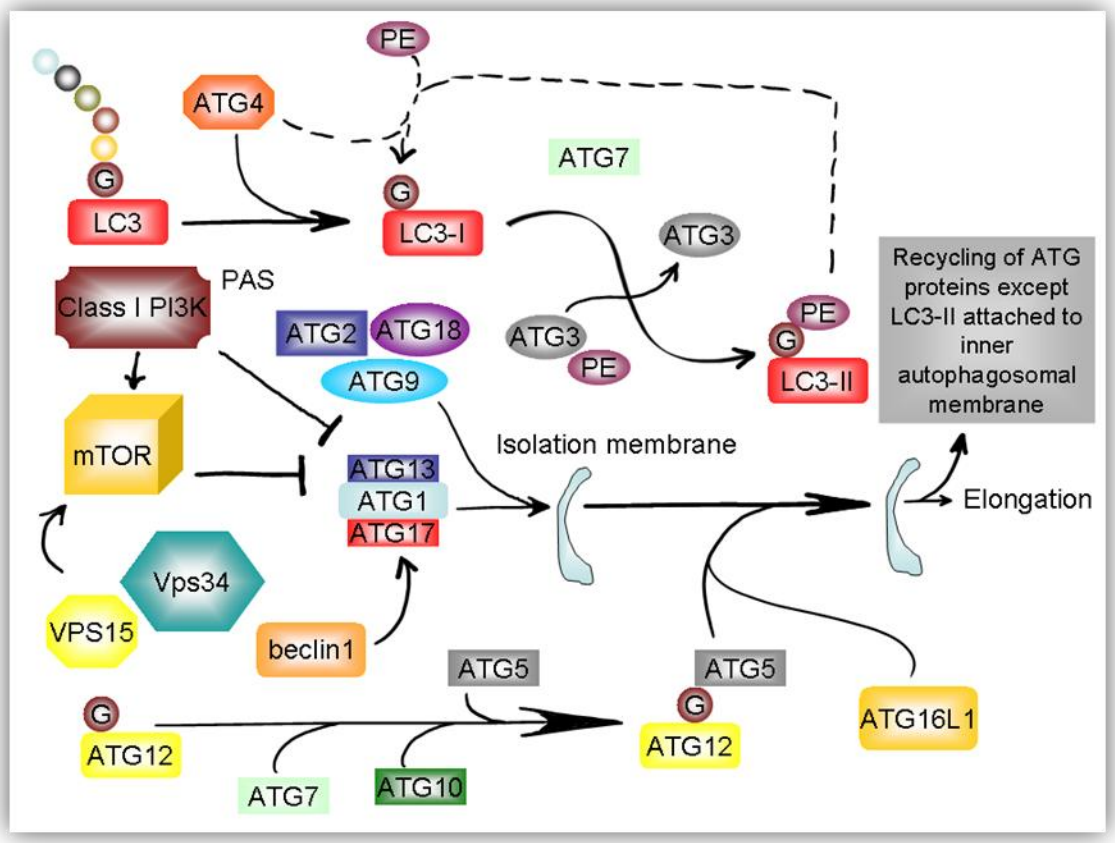



**Figure 7-2: Molecular machinery of autophagosome formation.**

**Figure 7-2: Molecular machinery of autophagosome formation. Autophagosomal membrane formation and expansion is mediated by two ubiquitin-like protein conjugation systems, the Atg8 (LC3) and Atg12 systems, while a recycling pathway mediates the disassembly of Atg proteins from matured autophagosomes (Atg2, Atg9, Atg18). PE: phosphatidylethanolamine; G: Glycine.**

### *7.1.2 The regulation of autophagy*

Early autophagy research implicated autophagy in the generation of energy during cellular starvation.[782-787] Firstly, metabolic hormones were shown to regulate autophagy, with glucagon inducing autophagy and insulin being inhibitory.[788] Later, Mortimore et al. and Seglen et al. discovered the regulatory role of amino acids, the inhibitory action of 3-methyladenine and the first evidence for the regulatory effects of protein kinases and phosphatases on autophagy.[789-791] These findings underpinned our understanding of autophagy as an energy-generating catabolic mechanism at times of cellular starvation via the catabolism of cell components after their sequestration into autophagolysosomes.

In contrast to the ubiquitin proteasome system which selectively degrades single proteins, autophagy can degrade large protein aggregates and whole organelles by both selective and non-selective mechanisms. Starvation induced autophagy is non-selective and involves large-scale degradation of the cytoplasm, therefore excessive levels of autophagy would be undesirable. On the other hand, autophagy is required to maintain minimal growth requirements and viability during starvation, therefore insufficient autophagy would also be undesirable: studies in *C. elegans* have shown that both low levels or excessively high levels of autophagy signal death in response to starvation, whilst intermediate levels signal survival.[792] Autophagy is normally maintained at a constitutive/basal level, and is activated in response to a myriad of stimuli including nutrient deprivation, pathogens, cytokines, protein aggregates and damaged organelles, thus autophagy must be a tightly regulated process.

The breakthrough in the investigation of the molecular regulation of autophagy came with the identification of the *T*arget *o*f *R*apamycin (TOR) gene and the observation that rapamycin, as an inhibitor of TOR, acts as an autophagy-inducer.[793;794] In mammalian cells, the best characterized regulatory pathways include the class I PI3K (phosphoinositide-3-kinase) and mammalian target of rapamycin (mTOR) which act

to inhibit autophagy, and the class III PI3K Vps34 which is paradoxically involved in both mTOR activation and the initiation of autophagosome formation (Figures 7-2 & 7-3).

Before mammalian cells become committed to grow and proliferate, sufficient nutrients and energy must be available. A key pathway that senses and signals this information is the mTOR pathway (Figure 7-3). The serine/threonine kinase mTOR forms two types of multi-subunit complex involving distinct partner proteins: TORC1 and TORC2.[795] The form of mTOR that directly regulates cell growth is TORC1, which contains mTOR in complex with Raptor (regulatory associated protein of mTOR), PRAS40 (proline-rich Akt substrate of 40kDa, an inhibitor of mTOR) and GβL.[795] The second complex, TORC2, differs in that it contains the binding partner Rictor (rapamycin-insensitive companion of TOR).[795] TORC1 is sensitive to rapamycin, whereas mTORC2 is insensitive to this compound. Only TORC1 (hereafter referred to as mTOR) is linked to the control of cell growth and autophagy. mTOR regulates the phosphorylation of a number of components of the translational machinery. In particular phosphorylation and activation of 4E-BP1 and S6K1 are stimulated by serum insulin and growth factors in an mTOR-dependent manner.[796]

The TSC complex is a heterodimer composed of tuberin (TSC2) and hamartin (TSC1) and is the major regulator of the mTOR signalling pathway.[797] TSC2 contains a GTPase-activating protein (GAP) domain that converts the small GTPase Rheb (Ras homolog enriched in brain) to its inactive GDP-bound form. mTOR activity is stimulated by the active GTP-bound form of Rheb, thus the TSC complex acts to inhibit mTOR function.[797] The TSC complex integrates growth factor and cellular energy signalling with mTOR activity.

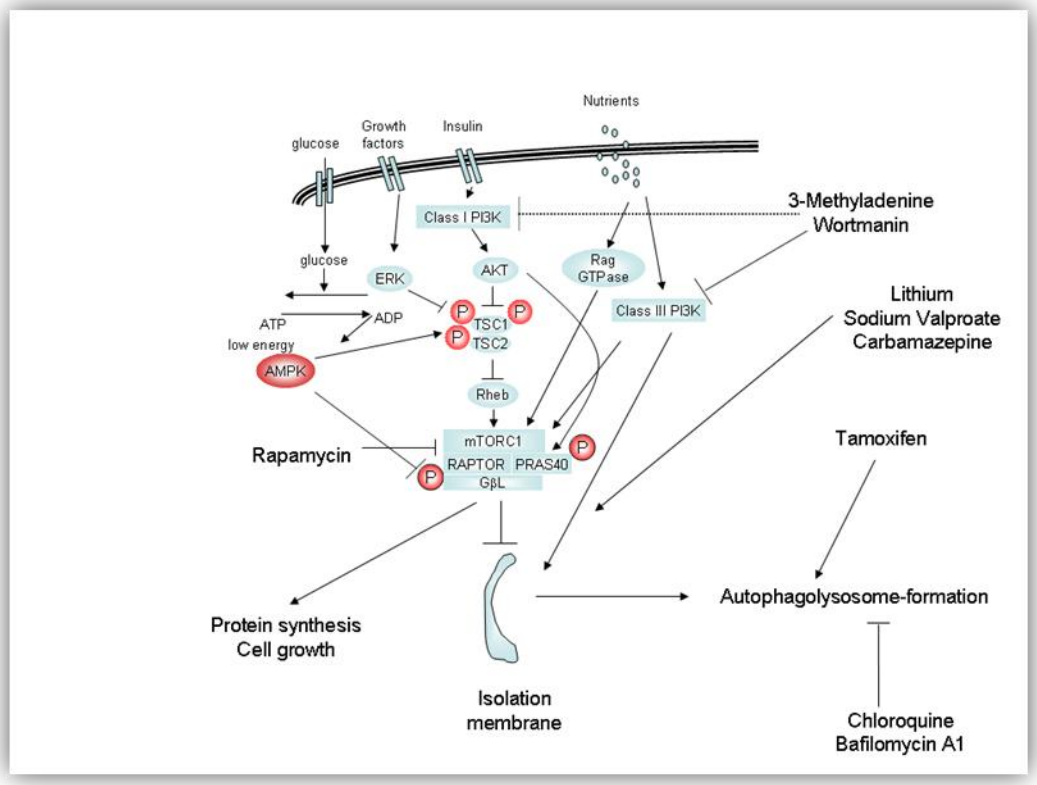



**Figure 7-3: Signalling, the regulation of non-selective autophagy and therapeutic options available to manipulate the autophagic response.**

**Figure 7-3: Signalling, the regulation of non-selective autophagy and therapeutic options available to manipulate the autophagic response. For details of the signalling pathways and inhibitors indicated in the figure the reader is referred to the main text ). P: signal transduction through direct phosphorylation.**

Signals from growth factor receptors are transduced to TSC1/2 via the class I PI3K/Akt and Ras/Mek/Erk pathways, and from cellular energy sensing via the LKB-AMPK pathway.[797] Factors which mediate nutrient (amino acid) signalling to mTOR include the class III PI3K Vps34 and the Rag family of GTPases.[798;799] Insulin stimulates Akt via the class I PI3K pathway and, once active, Akt stimulates mTOR via phosphorylation of both TSC2 (inhibiting its Rheb-GAP activity) and of

PRAS40 (relieving its inhibition of mTOR).[800] Growth factor signalling also stimulates the Ras/Mek/Erk pathway; once active, Erk phosphorylates TSC2 at different sites from Akt, also resulting in the inhibition of its Rheb-GAP activity.[801] Conversely, metabolic stresses that reduce energy levels in the cell activate the LKB-AMPK pathway. AMPK has two inhibitory effects on mTOR: it phosphorylates TSC2 and promotes its Rheb-GAP activity, and it phosphorylates Raptor.[802]

mTOR inhibits autophagy indirectly by inhibiting the activities of the serine/threonine kinase Atg1.[803;804] Starvation signals sensed by mTOR are transmitted to the Atg proteins, many of which accumulate at the PAS (Figure 2). In yeast, Atg1 interacts with Atg13 and Atg17, and its kinase activity is required for autophagy induction.[805;806] The Atg1 complex (Atg1/Atg13/Atg17) is essential for the recruitment of other Atg proteins to the PAS during non-specific autophagy, however this does not require Atg1 kinase activity.[807;808] In mammals, two Atg1 homologues have been discovered; ULK1 and ULK2. The role of ULKs in autophagy induction has not yet been properly characterised, however ULK kinase activity increases under starvation conditions, and kinase-dead mutants of ULK exert a dominant-negative effect on autophagosome formation.[809] In addition focal adhesion kinase (FAK) family interacting protein of 200kDa (FIP200), an ULK-interacting protein, is required for autophagosome formation as autophagosome formation is almost completely blocked in FIP200-deficient cells.[809] It was also observed that ULK1, ULK2 and FIP200 localize to isolation membranes, suggesting that the ULK-FIP200 complex plays an essential role in the early stages of autophagosome formation.[809]

The class III PI3K, Vps34, has been extensively studied in the context of vesicular trafficking processes, and recent work has shown that Vps34 also plays an important role in the ability of cells to respond to changes in nutrient conditions.[810] The activity of Vps34 requires the association of the protein kinase Vps15 (Figure 2).[811] When amino acids are plentiful, Vps34-Vps15 contributes to mTOR activation, thereby repressing autophagy. By contrast, the Vps34-Vps15-beclin1 (Atg6) complex

initiates autophagosome formation.[811] The Vps34-beclin1 complex can be activated by the beclin1 interacting proteins UVRAG (UV radiation resistance gene) and AMBRA1 (activating molecule in beclin1-regulated autophagy), and inhibited by another beclin1-interacting partner Bcl-2 (B-cell lymphoma 2).[811] Thus, depending on its binding partners, Vps34 is subject to different modes of regulation leading to activation or inhibition of autophagy.

Beclin1, the first identified mammalian autophagy gene interacts with several cofactors to activate the lipid kinase Vps34, thereby inducing autophagy.[812;813] Beclin1 is a BH3-domain-only protein that binds to the BH3 domain of the anti-apoptotic proteins Bcl-2/Bcl-XL.[814] Under normal conditions, beclin-1 is bound to and inhibited by Bcl-2 or the Bcl-2 homolog Bcl-XL and the dissociation of beclin1 from Bcl-2 is essential for its autophagic activity.[813] Nutrient deprivation stimulates the dissociation either by activating BH3-only proteins (such as Bad) that can competitively disrupt the interaction, or by posttranslational modification of beclin1 or Bcl-2.[813]

Although mTOR is still considered to be the central regulator of autophagy, other types of regulation independent of mTOR like Bcl-2/beclin1-complex and Atg4-regulation via c-Jun-N-terminal-kinase (JNK) and reactive oxygen species (ROS), have been described.[757;815-817] In addition, lowering the levels of the secondary messenger myo-inositol-1-4-5 triphosphate (IP3) also induces autophagy in an mTOR-independent manner. The primary function of IP3 is to release calcium from the endoplasmic reticulum (ER) to the cytoplasm, suggesting that autophagy might be modulated by calcium.

Recently, the NAD-dependent deacetylase Sirt1 has been implicated in the regulation of autophagy.[818] Transient increased expression of Sirt1 is sufficient to stimulate basal rates of autophagy. Sirt1-/- mouse embryonic fibroblasts do not fully activate autophagy under starved conditions but reconstitution with wild-type Sirt1 restores

autophagy. Sirt1 interacts with Atg5, Atg7 and Atg8 and the absence of Sirt1 leads to markedly elevated acetylation of these proteins.

### *7.1.3 Autophagy in Inflammatory bowel disease*

Both childhood and adult onset CD are characterised by a progressive disease course in terms of location of disease as well as disease behaviour (e.g. structuring or fistulising complications).[710] The involvement of the ileum is noteworthy as this is rarely a presenting feature in childhood but becomes involved as children with CD approach adulthood.[710;819] The anti-microbial Paneth cells are abundantly present in the crypts of the ileum in the healthy state, but in CD they characteristically are also present in the colon, then termed 'Paneth cell metaplasia'.

The autophagy genes Atg16L1 and IRGM are among the more than 30 susceptibility loci for CD that have been identified to date using genome-wide-association-studies (GWAS).[694] Autophagy has been implicated in limiting inflammation by eliminating pathogens, blocking cell necrosis and controlling NF-Kβ signalling.

Tumour cells that have not been successfully removed by apoptosis or autophagy are removed by necrotic cell death. Necrosis leads to activation of inflammatory responses,[820] thus preventing cell necrosis is an important mechanism through which autophagy regulates inflammation. It has been demonstrated that the removal of apoptotic cell corpses and necrotic cell debris is necessary to avoid excessive recruitment of neutrophils and to prevent chronic inflammation.[820;821] Consistent with this, mice lacking Atg5 display a defect in apoptotic corpse engulfment during embryonic development.[822] Interestingly, loss of tolerance against self-antigens due to defective clearance has been implicated in systemic lupus erythematosus pathophysiology.[823]

Another mechanism through which autophagy may regulate inflammation is via NF-Kβ. NF-Kβ is a master regulator of the inflammatory response and, as mentioned previously, autophagy can selectively degrade NIK and IKK resulting in activation of NF-Kβ signalling.[824;825] Whether any or all of inadequate removal of apoptotic/necrotic cells, misregulation of NF-Kβ signalling, or defective pathogen clearance through a defect in the autophagic machinery contribute to inflammatory bowel disease is currently the subject of intense investigation.

The different GWAS in CD and the meta-analysis of several of these GWAS have identified a number of genes involved in the regulation of the Th17-cell population, notably IL23R, JAK2, STAT3 and ICOSLG.[470-472;687;694;732] The recent report by Guo et al. demonstrating that Th17-mediated autoimmune disease (in this case a murine model of Multiple Sclerosis) is constrained by TRIF-dependent type I IFN production and its downstream signalling pathway, is also pertinent for our understanding of the pathogenesis of CD, as it provides the first evidence for a potential link between the autophagy pathway and the development of Th17 cells.[826] In the remainder of this section, we will discuss the autophagy genes ATG16L1 and IRGM as well as putative CD susceptibility genes LRRK2 and XBP1 in view of their relevance to the autophagy pathway.

#### 7.1.3.1.1 ATG16L1

Hampe et al. were the first group to implicate the autophagy pathway in CD.[688] These authors performed a genome-wide association study of 19,779 non-synonymous SNPs in 735 CD patients and 368 controls and demonstrated association with a coding variant of the ATG16L1 gene (Autophagy-related 16-like 1 gene, Ala197Thr polymorphism, rs2241880A/G) gene, on Chromosome 2q37.1. In their paper, the association was replicated in 2 other independent cohorts. Haplotype and regression analysis confirmed this variant carries virtually all the disease risk exerted by the ATG16L1 locus.[687;688]

ATG16L1 is expressed in colon, small bowel, intestinal epithelial cells, leukocytes and spleen and multiple splice variants have been reported.[687;688] The Ala197Thr variant leads to an amino acid change from the polar threonine to the non-polar alanine at the evolutionary conserved position 300 in the N-terminal WD-repeat domain of the molecule.[688]

Rioux and colleagues confirmed the association of this variant allele with ileal CD in the North-American genome-wide association study using case-control analysis and family-based association analysis.[687] Of note, all CD patients had ileal involvement as a criterion for entry into this study. These authors also provided novel functional data: utilizing oligo-based silencing RNA directed against ATG16L1 isoforms, autophagy induced by *S. Typhimurium* was significantly different in ATG16L1 knockdown HEK293 cells when compared with control cells.[687] However, Hampe et al. observed no clear difference in tissue expression of ATG16L1 between controls and CD patients and expression was independent of the rs2214880 genotype, both at protein and cDNA level.[688] Recently, Glas et al. confirmed these findings in CD patients and a mouse model of ileitis.[827]

The association of the Ala197Thr single nucleotide polymorphism (SNP) with susceptibility to CD has now been replicated in several independent cohorts, of which the key studies are summarised in Table 8-1.[471;472;547;687;730;731;827-829]

Recent studies have explored genotype-phenotype associations in adult onset CD, and to a more limited extent in paediatric disease. Prescott et al found that this ATG16L1 variant is associated specifically with ileal CD, but not with isolated colonic disease.[828] Based on a sub-group analysis of 140 CD patients diagnosed before their 17$^{th}$ birthday, these authors suggested an association between the Ala197Thr variant allele and early onset CD, as well as an effect of ATG16L1 genotype on age at diagnosis. Subsequently, Baldassano et al. replicated this association in a cohort of 142 children from Philadelphia with Crohn's disease.

As discussed in chapter 3, it is pertinent for the introduction of the genotype-phenotype analysis conducted of the ATG16L1variant, that the phenotype in early onset IBD in Scotland and Scandinavia differs from adult onset disease and is characterised by the extensive nature of disease.[830;831] Childhood onset CD is characterised by extensive colonic disease and the relative rarity of purely ileal involvement.

Recent reports by Cadwell, Saitoh and Kuballa constitute a real breakthrough in our understanding of ATG16L1 function and firmly implicate the homeostasis and anti-microbial properties of the Paneth cell.[832-834] Cadwell and colleagues showed that, in the epithelium of the ileum, ATG16L1 and ATG5 are crucial for Paneth cell biology.[832] ATG16L1-deficient Paneth cells exhibit notable abnormalities in the granule exocytosis pathway and increased expression of genes involved in peroxisome proliferator-activated receptor (PPAR) signalling as well as several acute phase reactants and adipocytokines (notably leptin and adiponectin). Crohn's disease patients homozygous for the risk allele, identified by Hampe et al., displayed Paneth cell granule abnormalities like those observed in ATG16L1-deficient mice.[832]

On the other hand, Saitoh and colleagues showed in ATG16L1-deficient mice that ATG16L1 is required to survive the period of neonatal starvation, as shown before for ATG5 and ATG7, and also an increased severity of colitis when induced by dextran sulphate sodium.[833] Loss of ATG16L1 in macrophages caused aberrant LPS-induced IL-1beta and IL-18, due to the activation of caspase-1 in a TRIF dependent manner.[833]

Using a knockdown-reconstitution strategy in human epithelial cells (HeLa and Caco2), Kuballa and colleagues demonstrated that homozygosity for the ATG16L1 risk allele, identified by Hampe et al, resulted in a reduced capture of *S. Typhimurium* within autophagosomes.[834] These authors further demonstrated that the wildtype (ATG16L1*300T) and mutant (*300A) coding variants are both fully

competent in mediating basal autophagy and that there is no difference between the two in allowing dimerisation or binding to Atg5, in HeLa and HEK293T cells respectively. Taken together, these studies implicate several distinct functions of Paneth cell biology, notably the capture of microorganisms within autophagosomses, the exocytosis of antimicrobial peptides and the orchestration of the immune response. However, the true effect of some of these genetic variants may only become clear when these genetic variants are studied in the setting of a high microbial load and when other genetic variants associated with CD susceptibility are taken into account (e.g. TLR4 and NOD2/CARD15).

**Table 7-1: Overview of the key ATG16L1 association studies published to date. Odds Ratio (OR) calculations are based on allelic frequencies in cases vs. controls of each population. *: p= 0.007 vs Scottish early onset CD; **: p=0.009 vs Scottish early onset CD.**

**Table 7-1: Overview of the key ATG16L1 association studies published to date. Odds Ratio (OR) calculations are based on allelic frequencies in cases vs. controls of each population.**

| **ATG16L1 rs2241880 A/G** | | Scottish | Baldassano[547] (USA) | Libioulle[471] (Belgium) | Hampe[688] (Germany) | Prescott[828] (UK) | Rioux[687] (North-America) *Ileal CD only* | Cummings[829] (UK) | WTCCC[472] (UK - imputed data) |
|---|---|---|---|---|---|---|---|---|---|
| Control | Allele G | 372/690 (53.9%) | 292/562 (52.0%) | 1273/2324 (54.8%) | 1484/2800 (53.0%) | 1268/2470 (51.3%) | 1286/2368 (54.3%) | 619/1190 (52.0%) | 3052/5872 (52.0%) |
| Adult CD | Allele G | 437/720 (60.7%) (OR 1.32) | | 879/1456 (60.4%) (OR 1.26) | 1480/2466 (60.0%) (OR 1.33) | 1435/2472 (58.1%) (OR 1.31) | 1656/2592 (63.9%) (OR 1.49) | 393/645 (61.0%) (OR 1.45) | 2091/3494 (59.8%) (OR 1.38) |
| Childhood CD | Allele G | 291/538 (54.1%) (OR 1.01) | 181/284 (63.7%) (OR 1.62)* | | | 178/280 (63.6%) (OR 1.65)** | | | |

7.1.3.1.2 IRGM

A few months after the genetic association of CD with ATG16L1, another autophagy gene, IRGM, was implicated by the large Wellcome Trust Case Control Consortium study. Reported by Parkes et al, these investigators performed a genome-wide association study in individuals with Crohn's (n=1748) and 2938 healthy controls (1500 individuals from the 1958 Birth cohort and 1500 individuals selected from blood donors recruited as part of the WTCCC) and detected strong association at four novel loci in addition to the previously identified NOD2/CARD15, IL23R and ATG16L1 loci.[835]

They subsequently tested 37 SNPs from these and other loci for association in an independent case-control sample (n=1182 CD vs n=5746 non-autoimmune WTCCC cases (bipolar disorder, coronary artery disease and hypertension) crucially *assuming* no overlap with Crohn's disease etiology).[472;835] Of the 37 SNPs, twelve SNPs showed a difference in allele frequency between these two groups who were then genotyped in 2,024 independent population controls from the 1958 British Birth Cohort to test formally for association. Of the new loci, the strongest replication adjacent to a known gene was for SNPs rs13361189 and rs4958847 ($p_{replication}$= 6.6 x 10-4 and 3.1 x 10-4, respectively) immediately flanking IRGM on chromosome 5q33.1. Association in the combined panels of 2,930 cases and 4,962 controls was highly significant ($p_{combined}$= 2.1 x $10^{-10}$ and 3.8 x $10^{-9}$, respectively).

The long coding exon of human IRGM encodes a 20-kDa protein of 181 amino acids.[836;837] As none of the associated SNPs were known to be functional, Parkes et al. resequenced the coding exon of IRGM and the four small putative downstream exons in 48 affected individuals homozygous or heterozygous for risk alleles. They detected two new nonsynonymous sequence variants, 51G/C (E17D) and 281C/A (T94K) and an exonic synonymous SNP, 313T/C, (rs10065172, L105). These SNPs were then genotyped in 769 unselected CD patients and 705 controls. Only the silent

313T/C variant was associated with Crohn's disease (p=0.008) and was found to be in near perfect Linkage Dysequilibrium with SNP rs13361189 ($r^2 = 0.91$). Sequencing of the IRGM coding region in another 100 unselected affected individuals and 100 controls did not detect additional variants.[472]

Parkes and colleagues therefore suggested that the causal variant(s) do not change the amino acid sequence of IRGM. They may lie in regulatory sequences in LD with the associated SNPs and affect IRGM expression, or the exonic (313T/C) SNP itself might affect transcript splicing or the rate of protein translation.[472]

IRGM is expressed in several tissues notably the colon, small bowel, peripheral blood leukocytes and U937 monocytic cell line, as assessed by PCR screening of uncloned cDNA transcription. As discussed briefly in section 8.2.2, IRGM belongs to the p47 immunityrelated GTPase family. Its mouse homolog, LRG-47 (encoded by Irgm), critically controls intracellular pathogens by autophagy, and Irgm–/– mice show markedly increased susceptibility to *Toxoplasma gondii* and *Listeria monocytogenes*.[838] Consistent with this, IRGM induces autophagy and thereby control of intracellular Mycobacterium tuberculosis in human macrophages.[837] The role of IRGM in protecting mature effector CD4+ T-lymphocytes against IFN-γ induced autophagic cell death, has demonstrated a feedback mechanism in the Th1 response that limits the detrimental effect of IFN-γ on effector T-lymphocyte survival while facilitating the antimicrobial functions of IFN-γ.[839] An alteration in IRGM regulation, due to a common deletion polymorphism in the promoter region of IRGM, that affects the efficacy of autophagy, is postulated in CD.[840]

#### 7.1.3.1.3 LRRK2

The meta-analysis of three GWAS in Crohn's disease led to the identification of 21 novel loci in addition to confirming 11 previously reported loci.[694] One of the novel loci meeting stringent criteria of genome-wide significance contains, in addition to

the MUC19 gene , also the gene LRRK2 (leucine-rich repeat kinase 2), which encodes a multi-domain protein expressed mainly in the cytoplasm of neurons, myeloid cells and monocytes. LRRK2 mutations are the single most common genetic cause of Parkinson disease, accounting for 1–40% of all cases depending on the population studied.[841] Recently, Plowey and colleagues reported that transfection of LRRK2 cDNA containing the common G2019S mutation resulted in significant decreases in neurite length, which were not observed in cells transfected with wild type LRRK2 or its kinase-dead K1906M mutation.[842] G2019S LRRK2 transfected cells also exhibited striking increases in autophagic vacuoles. RNA interference knockdown of LC3 or Atg7 reversed the effects of G2019S LRRK2 expression on neuronal process length, whereas rapamycin potentiated these effects. These results in Parkinson's and Crohn's disease offer further support for the role of autophagy in disease pathogenesis, although it remains to be elucidated whether LRRK2 affects the autophagy pathway directly or indirectly (e.g. through other signalling pathways such as MAPK).

Our group has investigated haplotypic variation of the MUC19 gene (A. Phillips et al. presented at DDW 2009).[843] The lack of association with CD in our adequately-powered Scottish adult CD cohort, further points to the LRRK2 gene as a novel susceptibility gene in CD.

### 7.1.4 Aims

In the Scottish population, we have previously demonstrated that the effect of germline variation of NOD2/CARD15, the IBD5 locus and IL23R on susceptibility is similar in adult onset and early onset disease.[271;278;337;660;700;706] In this study, we set out to address whether the ATG16L1 Ala197Thr variant and germline variation of the IRGM gene are independent determinants of susceptibility to childhood onset IBD, by studying this polymorphism in 361 childhood onset IBD patients. Following this, we genotyped these variants in 855 adult onset IBD, all from Scotland. We have also made use of the Wellcome Trust Case Control Consortium data to expand our control cohort yielding increased power to our study. We have explored in detail the relationship between disease phenotype and the Ala197Thr polymorphism of *ATG16L1* and the variants of IRGM.

### Methods

### 7.1.5 Subjects

2418 subjects consisting of 392 childhood IBD (262 CD, 96 UC, 34 IBDU; median age 11.2 years (Inter-Quartile range (Q1-Q3: 8.7-13.1)), 685 parents, 979 adult IBD (442 CD, 537 UC; median age 32.2 years (Q1-Q3: 24.8-46.7)) and 362 population-matched controls were genotyped for rs2241880 (ATG16L1) and 3 IRGM variants (rs10065172, rs13361189, rs4958847). Scottish healthy controls consisted of blood donors and locally recruited healthy controls, as described previously.[248] More than 95% of the Scottish subjects were Caucasian.

In addition, 1958 British Birth Cohort control data and WTCCC healthy blood donors were used for the analysis (n=2937).

This study was approved by the Local Research Ethics Committee.

### 7.1.6 IBD Phenotyping

IBD was diagnosed by standard criteria.[523] Detailed sub-phenotypic data based on the Montreal classification for the adult and childhood Scottish CD cohorts are given in chapter 3.[516] Of particular note, early onset CD patients were characterised by a low frequency of isolated ileal disease (L1±L4 (ileal disease with or without upper gastrointestinal tract disease) 20/257 (7.8%) compared with adult CD 134/360 (37.2%), $p<10^{-4}$ ).[516] In addition, the frequency of ileocolonic disease (L3±L4 (ileo-colonic disease with or without upper gastrointestinal tract disease) differed significantly between childhood onset CD and adult onset CD (147/257 (57.2%) and 83/360 (23.1%), respectively, $p<10^{-4}$).

Data collection in childhood onset CD also included height, weight and Body Mass Index (BMI) at diagnosis. Z-scores were calculated using nomograms from the British 1990 growth reference (Child Growth Foundation, 2 Mayfield Avenue, London W4 1PW).

### 7.1.7 SNP selection and genotyping

#### 7.1.7.1 ATG16L1

Haploview software (version 3.32) was used to assess HapMap data of the region on Chromosome 2 spanning the ATG16L1 gene (position 233,914,305 to 234,014,304).[681] This confirmed the results of the analysis by Hampe et al. that germline variation at this locus can be tagged satisfactorily by the rs2241880 variant (analysis based on solid spine of linkage disequilibrium, minor allelic frequencies >10%, haplotype frequencies >10%).[688] (Figure 8.1)

All individuals were genotyped for the ATG16L1 rs2241880A/G variant using the TaqMan system (7900HT sequence detection system; Applied Biosystems, Foster City, CA, USA) at the Wellcome Trust Clinical Research Facility (Western General Hospital campus) with call rates in all populations >91%. The rs2241880 genotype frequencies in controls were in Hardy-Weinberg equilibrium. The 3 most common NOD2/CARD15 polymorphisms (R702W, G908R and Leu1007fsinsC) and the

variant alleles of OCTN1&2 (SL22A4&5) within the IBD5 locus were genotyped as previously described.[271;278;337;660]

#### 7.1.7.2 IRGM

3 IRGM variants (rs10065172, rs13361189, rs4958847) were selected for genotyping. rs13361189 (position 150,203,580) and rs4958847 (position 150,219,780) are the SNPs with the strongest replication signal in the WTCCC, immediately flanking IRGM of Chromosome 5q33.1.[472] The synonymous, exonic rs10065172 (position 150,208,191) was selected as it was the only SNP to retain an association signal with CD, of the 3 SNPs studied in the last phase of the WTCCC study, described above.[472] In 48 CD patients, detailed sequencing of all 5 IRGM exons and exon-intron boundaries was performed. In the childhood onset trios and controls, an additional SNP (rs6869426, position 150,213,197) tagging the haplotype block containing the IRGM gene was studied (haplotype frequency >10%, minor allelic frequency (MAF)>1%, based on solid spine of linkage disequilibrium: HapMap SNP data, version 3.32 – 200kb region spanning the IRGM gene (150,207,888-150,260,491) is shown in Figure 8.2).[681] All individuals were genotyped for these variants using the TaqMan system (7900HT sequence detection system; Applied Biosystems, Foster City, CA, USA) at the Wellcome Trust Clinical Research Facility (Western General Hospital campus), with call rates >92% and satisfying the Hardy-Weinberg in all control cohorts (p>0.05).

**Figure 7-5: HapMap data analysis using Haploview software of 100kb region spanning the ATG16L1 region. Position of the rs2144880 variant is indicated (arrow). This variant tags the haplotype block, which contains only the ATG16L1 gene.**

**Figure 7-6: HapMap data spanning the 200kb region containing the IRGM gene. The tagging SNP rs6869426 is indicated (red arrow).**

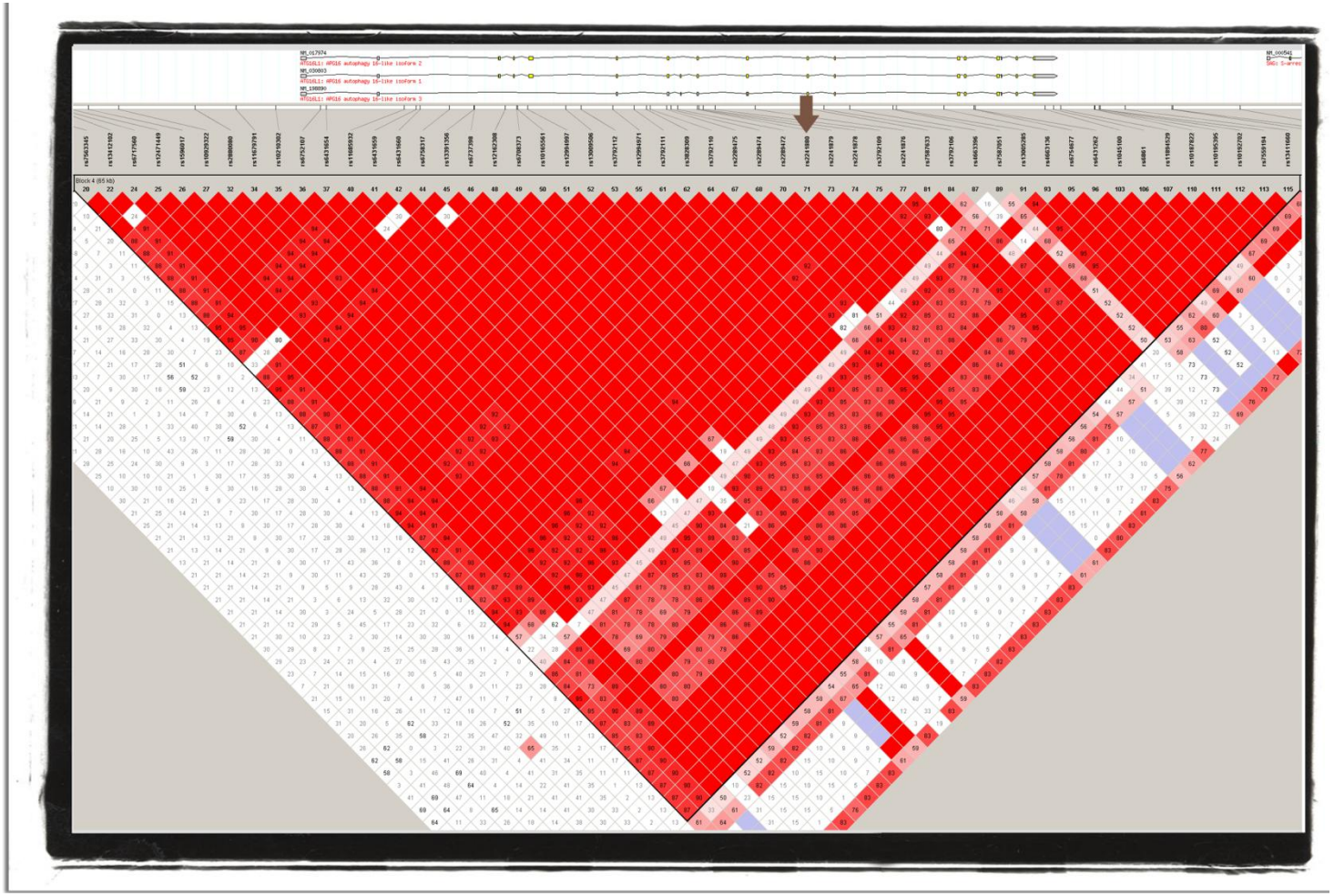


**Figure 7-4: HapMap data spanning the ATG16L1 region.**

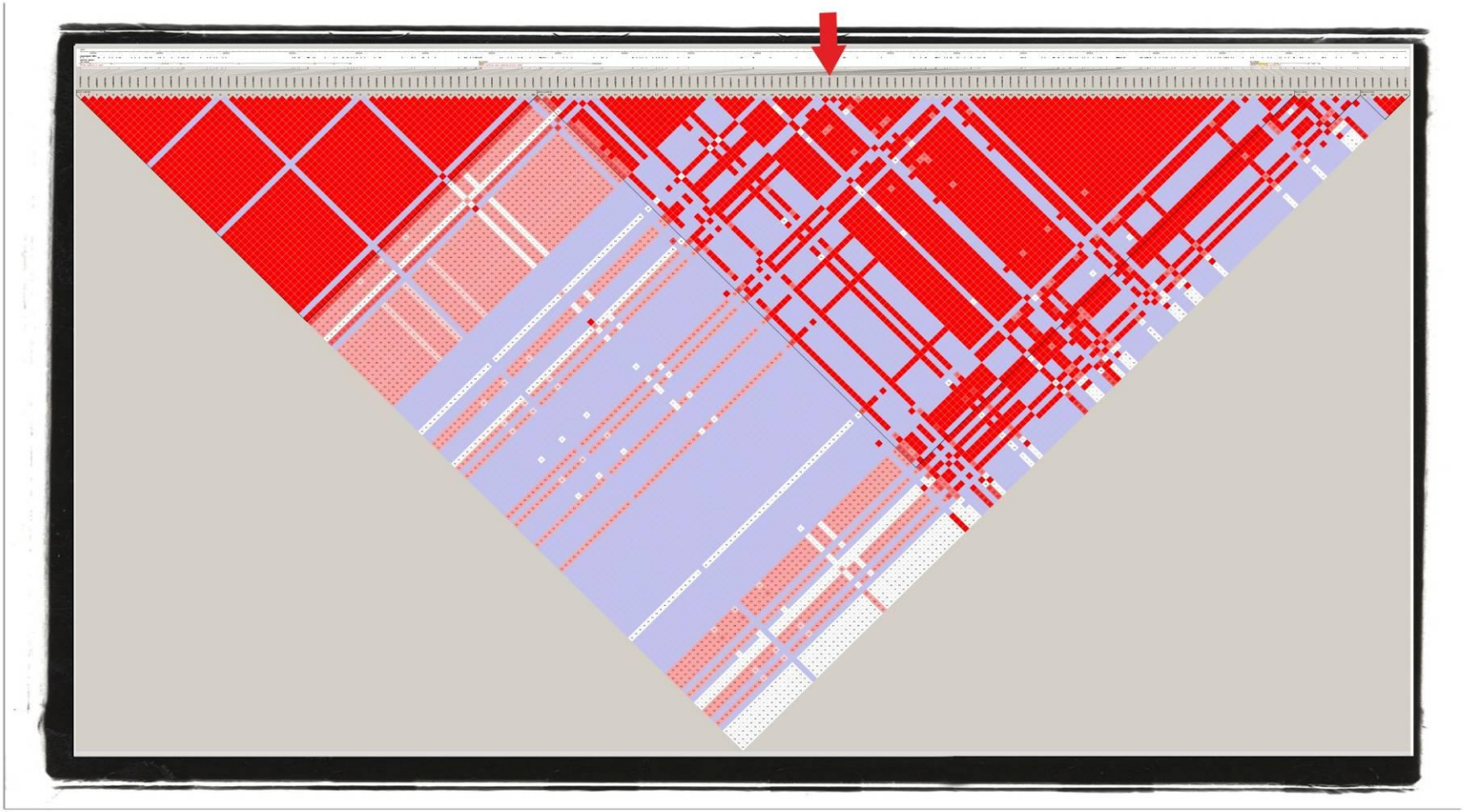


**Figure 7-5: HapMap data spanning the 200kb region containing the IRGM gene.**

### *7.1.8 Statistical analysis*

Allelic and genotype frequencies of the ATG16L1 and IRGM variants were compared between cases and controls using $\chi^2$ (Minitab software, Release 13.20, Minitab Inc., State College, PA, USA and GraphPad Instat software version 3.06, San Diego, Ca, USA). TDT analysis was performed using FBAT software (version 1.7.3, available from http://www.biostat.harvard.edu/~fbat/fbat.htm).[532] Detailed genotype-phenotype analysis was performed using Minitab. The effect of ATG16L1 rs2241880 genotype on age at diagnosis was further assessed using the Kruskal-Wallis test. ANOVA (+ Hsu's Multiple Comparisons with the Best (set to the rs2241880 AA genotype)) was used to assess the relationship between rs2241880 genotype and z-score of height, weight and BMI at diagnosis.

#### 7.1.8.1 Power calculations

Power calculations were performed using QUANTO version 1.2 (http://hydra.usc.edu/gxe).[528;529] Based on the most recent prevalence data of Scottish paediatric CD (14/100 000 of the population <17 years), our childhood onset CD study had 80% power to detect an effect with OR>1.3 for ATG16L1 (when using the sum of Scottish controls + WTCCC controls, p>0.05 when comparing allelic frequencies between two control cohorts for all genotyped variants based on a log-additive model).[4]

This compares favourably with the OR >1.60 as calculated in the British early onset cohort from Prescott et al. (n=140) and also with data derived in a cohort of North-American patients with early onset CD from Baldassano et al.(n=142) for the ATG16L1 variant.[547;828] (Table 8-1)

Our adult CD cohort had 80% power to detect an OR of 1.26 for the rs2241880 ATG16L1 variant (sum of Scottish controls + WTCCC controls, population risk

144/100 000).[1;676;835] The combined adult + childhood onset CD cohort vs the sum of WTCCC + Scottish controls had 80% power to detect an effect with OR 1.20.

Power calculations for the IRGM variants are given in Table 8.2 (80% power – prevalence data as above) and compared with the ORs obtained in the index WTCCC study:

**Table 7-1: Power calculations based on log-additive risk model for childhood onset and overall CD in Scotland for 4 IRGM variants**

| | Scottish childhood CD | Scottish CD overall | WTCCC |
|---|---|---|---|
| rs10065172 | OR 1.51 | OR 1.31 | OR 1.44 |
| rs13361189 | OR 1.46 | OR 1.33 | OR 1.51 |
| rs4958847 | OR 1.40 | OR 1.27 | OR 1.35 |
| rs6869426 | OR 1.41 | N/A | N/A |

***Results***

### *7.1.9 The influence of ATG16L1 Ala197Thr/ rs2241880 on Scottish childhood onset IBD susceptibility*

No association of ATG16L1 with CD was seen in our early onset CD case-control analysis (frequency of rs2241880 G allele in CD was 291/538 (54.1%) vs. Scottish controls 372/690 (53.9%), p=0.95 OR 1.01 (0.80-1.26)). (Table 8-3)

Comparing the allelic frequency in the two available control cohorts showed there was no statistically significant difference in allelic frequency (372/690 (53.9%) vs 3052/5872 (51.9%), p=0.33). We therefore proceeded to analyse the CD cases vs the expanded cohort consisting of the sum of WTCCC and Scottish controls.

The allelic frequency in early onset CD also did not differ significantly from the allelic frequency in the expanded control cohort. (54.1% vs 3052/5872 + 372/690= 3424/6562 (52.2%), p=0.39).

Transmission disequilibrium testing (TDT) did not show any distortion of transmission in CD trios (p=0.51). No significant differences in genotype or allelic frequency of the rs2241880 variant allele were observed between early onset UC and controls (G allele frequency 84/174 (48.3%), p=0.18 OR 0.80 (0.57-1.11) versus healthy Scottish controls). TDT analysis in UC trios was also negative (p=0.32).

**Table 7-4: Genotype and allelic frequencies of ATG16L1 rs2241880 in cases (stratified for age at diagnosis <17 years and >17 years) and controls. p-values are calculated using genotype/allelic frequency in cases vs. controls. Allelic frequency of rs2144880 G differed significantly between CD<17 and CD>17 (p=0.01).**

**Table 7-2: Genotype and allelic frequencies of ATG16L1 rs2241880 in cases (stratified for age at diagnosis <17 years and >17 years) and controls**

| **ATG16L1 rs2241880 A=1 G=2** | | Scottish Controls | CD<17 | p | CD>17 | p | Combined CD | p |
|---|---|---|---|---|---|---|---|---|
| Genotype frequencies | 11 | 71/345 (20.6%) | 58/269 (21.6%) | 0.76 | 60/360 (16.7%) | 0.18 | 118/629 (18.8%) | 0.49 |
| | 12 | 176/345 (51.0%) | 131/269 (48.7%) | 0.56 | 163/360 (45.3%) | 0.12 | 294/629 (46.7%) | 0.20 |
| | 22 | 98/345 (28.4%) | 80/269 (29.7%) | 0.71 | 137/360 (38.1%) | 0.006 OR 1.55 (1.13-2.12) | 217/629 (34.5%) | 0.05 |
| Allele | 2 | 372/690 (53.9%) | 291/538 (54.1%) | 0.95 | 437/720 (60.6%) | 0.01 OR 1.32 (1.07-1.63) | 728/1258 (57.9%) | 0.09 |

### *7.1.10 Genotype-phenotype analysis in Scottish childhood onset IBD*

In childhood onset CD, we observed the lowest frequency of the rs2144880G variant allele in isolated colonic disease (L2±L4: 87/172 (50.6%)), increasing to 55.4% (163/294) in ileocolonic disease (L3±L4) and 62.5% (25/40) in pure ileal disease (L1±L4) (p=0.28 OR 1.44 (0.74-2.80), L1±L4 vs other CD locations). Genotype-phenotype analysis of disease behaviour in children with CD showed no significant associations. ANOVA assessment of z-scores of height, weight and BMI at CD diagnosis did not show any association with ATG16L1 genotype (p=0.79, 0.61 and 0.65, respectively).

### *7.1.11 The influence of ATG16L1 Ala197Thr/rs2241880 on susceptibility to adult onset IBD*

rs2144880 was genotyped in our adult IBD cohort, comprising 360 Scottish CD and 495 Scottish UC patients. When comparing the allelic frequency of rs2241880G in the Scottish adult CD cohort with population-matched controls (n=345), a significant association of ATG16L1 with CD was present (p=0.01 OR 1.32 (1.07-1.63)).(Table 8-3). This association was confirmed when we compared Scottish adult CD with the expanded WTCCC+Scottish control cohort (60.6% vs 52.2%, p<10-4 OR 1.42 (1.21-1.66).

Adult Scottish UC patients did not differ from control subjects when assessing genotype or allelic frequency of the rs2241880 variant (G allele frequency 525/990 (53.0%), p=0.72 OR 0.97 (0.79-1.17)).

**Table 7-5: Genotype and allelic frequencies of ATG16L1 rs2241880 in adult + childhood onset CD cases and controls. p-values are calculated using genotype/allelic frequency in cases vs. controls.**

**Table 7-3: Genotype and allelic frequencies of ATG16L1 rs2241880 in adult + childhood onset CD cases and controls.**

| **ATG16L1 rs2241880 A=1 G=2** | | Controls | L1±L4 | p | L2±L4 | | L3±L4 | p | L4 | p L1±L4 vs. not L1±L4 |
|---|---|---|---|---|---|---|---|---|---|---|
| | 11 | 71/345 (20.6%) | 20/160 (12.5%) | 0.02 | 51/226 (22.6%) | 0.57 | 48/234 (20.5%) | 0.98 | 1/13 (7.7%) | 0.01 |
| Genotype frequencies | 12 | 176/345 (51.0%) | 83/160 (51.9%) | 0.85 | 97/226 (42.9%) | 0.05 | 102/234 (43.6%) | 0.07 | 6/13 (46.2%) | 0.06 |
| | 22 | 98/345 (28.4%) | 58/160 (36.3%) | 0.07 | 68/226 (30.1%) | 0.66 | 84/234 (35.9%) | 0.05 | 6/13 (46.2%) | 0.51 |
| Allele | 2 | 372/690 (53.9%) | 199/320 (62.2%) | 0.01 OR 1.41 (1.07-1.84) | 233/452 (51.5%) | 0.43 | 270/468 (57.7%) | 0.20 | 18/26 (69.2%) | 0.02 OR 1.34 (1.03-1.74) |

### 7.1.12 Genotype-phenotype analysis in adult onset IBD

In adult onset CD, we observed the lowest frequency of the rs2144880G variant allele in isolated colonic disease (55.0% (143/260) in L2±L4), increasing to 56.2% (99/176) in ileocolonic disease (L3±L4) and 62.5% (165/264) in pure ileal disease (L1±L4) (p=0.08 OR 1.30 (0.96-1.77), allelic frequency in L1±L4 vs. other CD locations). Genotype-phenotype analysis of disease behaviour in adults with CD showed no significant associations.

### 7.1.13 Genotype-phenotype analysis in combined childhood and adult onset CD cohorts

A combined genotype-phenotype analysis of early onset and adult onset CD is presented in Table 7-5. The allelic frequency of rs2144880G differed significantly between patients with L1±L4 and other CD patients (p=0.02, OR 1.34 (1.03-1.74)). Binary logistic regression analysis (correcting for NOD2/CARD15 variant carriage, IBD5 (OCTN1&2) genotype and age at diagnosis) confirmed the effect of ATG16L1 rs2241880 genotype (GG) on pure ileal disease vs. pure colonic disease (p= 0.03 OR 2.43 (1.05-5.65)), in addition to the known effect of NOD2/CARD15 variant carriage (p<0.001 OR 3.13 (1.71-5.73)) and older age at diagnosis (p=0.02 OR 1.02 (1.00-1.03)) per 1 year age increment.

There was no significant difference in frequency of the rs2241880 G allele between males and females with CD (333/584 (57.0%) vs. 399/682 (58.5%), p=0.59 OR 0.94 (0.75-1.18)).

ATG16L1 rs2241880 genotype did not influence age at CD diagnosis significantly (Kruskal-Wallis test, p=0.51).

### 7.1.14 Interaction between ATG16L1 and NOD2/CARD15

In adult CD, we found evidence for an interaction between ATG16L1 and NOD2/CARD15. Adult CD patients (>17 years at diagnosis) not carrying any of the

3 common NOD2/CARD15 polymorphisms demonstrated a higher frequency of the rs2241880G allele than patients carrying at least one variant NOD2/CARD15 allele (303/474 (63.9%) vs. 79/148 (53.3%), p=0.02 OR 1.55 (1.07-2.25)) This difference between NOD2/CARD15 variant carriers and wild-type cases was not present in early onset CD, possibly due to the small number of childhood onset CD patients carrying at least one of the NOD2/CARD15 variants (n=59). These findings mirror those by Hampe et al. that ATG16L1 germline variation is a risk factor in adult CD, even in the absence of NOD2/CARD15 mutations.[688]

### *7.1.15 The influence of IRGM germline variation on CD susceptibility*

#### 7.1.15.1Susceptibility analysis in childhood and adult onset CD

The results of the case-control association analysis are presented in Tables 7-6, 7-7, 7-8 and 7-9. No association was seen between any of the 4 studied IRGM variants and childhood onset CD. Analysis are presented as Scottish CD (overall as well as stratified for age <17 and >17 at diagnosis) compared with Scottish Healthy controls and the sum of Scottish Healthy controls + WTCCC controls. Healthy control genotyope frequencies were not significantly different between Scottish controls and WTCCC controls (p>0.05).

TDT analysis in CD trios was negative for all variants (p>0.05).

In adult onset CD, the IRGM rs13361189C allele was associated with CD (9.2% vs 6.8%, p=0.01 OR 1.39 (1.06-1.82)). However, after stringent Bonferroni-correction for multiple testing, all IRGM variants studied did not differ significantly between CD cases and controls.

**Tables 7-6, 7-7, 7-8: Case-control analysis of rs10065172, rs13361189, rs4958847 in Scottish CD vs Scottish + WTCCC healthy controls. p*: comparison Scottish CD vs Scottish healthy controls. p**: Scottish CD vs sum of WTCCC and Scottish healthy controls. p***: Scottish CD stratified for age <17 and >17 vs sum of WTCCC healthy controls and Scottish healthy controls.p****: WTCCC CD vs WTCCC controls. p*****: Scottish healthy controls vs WTCCC controls.**

**Table 7-4: Case-control analysis of IRGM rs10065172 in Scottish CD vs Scottish + WTCCC healthy controls**

| **IRGM rs10065172 C=1 / T=2** | | Scottish controls *(p******)* | All Scottish CD | p* | WTCCC Controls | WTCCC + Scottish HC | p** | Scottish CD<17 *(p***)* | Scottish CD>17 *(p***)* | WTCCC CD | p**** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 11 | 298/346 (86.1%) *(0.98)* | 524/630 (83.2%) | 0.22 | 579/672 (86.2%) | 877/1018 (86.1%) | 0.10 | 232/273 (85.0%) *(0.62)* | 290/355 (81.7%) *(0.04)* | 548/682 (80.4%) | 0.004 OR 0.66 (0.49-0.88) |
| Genotype frequencies | 12 | 46/346 (13.3%) *(0.87)* | 101/630 (16.0%) | 0.25 | 87/672 (12.9%) | 133/1018 (13.1%) | 0.09 | 39/273 (14.3%) *(0.59)* | 62/355 (17.5%) *(0.04)* | 128/682 (18.8%) | 0.003 OR 1.55 (1.16-2.09) |
| | 22 | 2/346 (0.6%) *(0.59)* | 5/630 (0.8%) | 0.70 | 6/672 (0.9%) | 6/1018 (0.6%) | 0.62 | 2/273 (0.7%) *(0.78)* | 3/355 (0.8%) *(0.60)* | 6/682 (0.9%) | 0.97 |
| Allele | 2 | 50/692 (7.2%) *(0.93)* | 111/1260 (8.8%) | 0.22 | 99/1344 (7.4%) | 149/2036 (7.3%) | 0.12 | 43/546 (7.9%) *(0.65)* | 67/710 (9.4%) *(0.07)* | 140/1364 (10.3%) | 0.007 OR 1.44 (1.10-1.88) |

**Table 7-5: Case-control analysis of IRGM rs13361189 in Scottish CD vs Scottish + WTCCC healthy controls**

| **IRGM rs13361189 T=1 / C=2** | | Scottish controls *(p******)* | All Scottish CD | p* | WTCCC Controls | WTCCC + Scottish HC | p** | Scottish CD<17 *(p***)* | Scottish CD>17 *(p***)* | WTCCC CD | p**** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 11 | 313/362 (86.5%) *(0.89)* | 542/642 (84.4%) | 0.38 | 2538/2927 (86.7%) | 2851/3289 (86.7%) | 0.12 | 237/272 (87.1%) *(0.83)* | 302/366 (82.5%) *(0.02)* | 1409/1741 (80.9%) | $<10^{-4}$ OR 0.65 (0.55-0.76) |
| Genotype frequencies | 12 | 47/362 (13.0%) *(0.94)* | 95/642 (14.8%) | 0.42 | 384/2927 (13.1%) | 431/3289 (13.1%) | 0.24 | 33/272 (12.1%) *(0.64)* | 61/366 (16.7%) *(0.05)* | 322/1741 (18.5%) | $<10^{-4}$ OR 1.50 (1.28-1.77) |
| | 22 | 2/362 (0.6%) *(0.13)* | 5/642 (0.8%) | 0.67 | 5/2927 (0.2%) | 7/3289 (0.2%) | 0.01 | 2/272 (0.7%) *(0.09)* | 3/366 (0.8%) *(0.03)* | 10/1741 (0.6%) | 0.01 |
| Allele | 2 | 51/724 (7.0%) *(0.75)* | 105/1284 (8.2%) | 0.36 | 394/5854 (6.7%) | 445/6578 (6.8%) | 0.06 | 37/544 (6.8%) *(0.97)* | 67/732 (9.2%) *(0.01)* | 342/3482 (9.8%) | $<10^{-4}$ OR 1.51 (1.30-1.76) |

**Table 7-6: Case-control analysis of IRGM rs4958847 in Scottish CD vs Scottish + WTCCC healthy controls**

| **IRGM rs4958847 G=1 / A=2** | | Scottish controls *(p******)* | All Scottish CD | p* | WTCCC Controls | WTCCC + Scottish HC | p** | Scottish CD<17 *(p***)* | Scottish CD>17 *(p***)* | WTCCC CD | p**** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 11 | 263/348 (75.6%) *(0.16)* | 464/614 (75.6%) | 0.99 | 2314/2937 (78.8%) | 2577/3285 (78.4%) | 0.11 | 196/256 (76.6%) *(0.48)* | 265/354 (74.9%) *(0.12)* | 1275/1746 (73.0%) | <10-4 OR 0.73 (0.63-0.84) |
| Genotype frequencies | 12 | 81/348 (23.3%) *(0.12)* | 139/614 (22.6%) | 0.82 | 580/2937 (19.7%) | 661/3285 (20.1%) | 0.15 | 53/256 (20.7%) *(0.82)* | 85/354 (24.0%) *(0.08)* | 428/1746 (24.5%) | 0.0001 OR 1.32 (1.15-1.52) |
| | 22 | 4/348 (1.1%) *(0.64)* | 11/614 (1.8%) | 0.23 | 43/2937 (1.5%) | 47/3285 (1.4%) | 0.49 | 7/256 (2.7%) *(0.10)* | 4/354 (1.1%) *(0.64)* | 43/1746 (2.5%) | 0.01 |
| Allele | 2 | 90/696 (12.9%) *(0.21)* | 161/1228 (13.1%) | 0.91 | 666/5874 (11.3%) | 756/6570 (11.5%) | 0.10 | 67/512 (13.1%) *(0.28)* | 93/708 (13.1%) *(0.19)* | 514/3492 (14.7%) | $<10^{-4}$ OR 1.35 (1.19-1.53) |

**Table 7-7: Case-control analysis of the IRGM rs6869426 haplotype tagging SNP in Scottish childhood onset CD vs Scottish healthy controls**

| **IRGM rs6869426 T=1 / C=2** | | Scottish controls | Childhood IBD | p | Childhood CD | p | Childhood UC | p | Childhood IBDU | p |
|---|---|---|---|---|---|---|---|---|---|---|
| | 11 | 111/344 (32.3%) | 115/355 (32.4%) | 0.97 | 74/237 (31.2%) | 0.79 | 29/87 (33.3%) | 0.84 | 12/31 (38.7%) | 0.46 |
| Genotype frequencies | 12 | 161/344 (46.8%) | 173/355 (48.7%) | 0.60 | 119/237 (50.2%) | 0.41 | 41/87 (47.1%) | 0.95 | 13/31 (41.9%) | 0.60 |
| | 22 | 72/344 (20.9%) | 67/355 (18.9%) | 0.49 | 44/237 (18.6%) | 0.48 | 17/87 (19.5%) | 0.77 | 6/31 (19.4%) | 0.83 |
| Allele | 2 | 383/688 (55.7%) | 403/710 (56.8%) | 0.68 | 267/474 (56.3%) | 0.82 | 99/174 (56.9%) | 0.77 | 37/62 (59.7%) | 0.54 |

**Table 7-9: Case-control analysis of the IRGM rs6869426 haplotype tagging SNP in Scottish childhood onset IBD (CD/UC/IBDU) vs Scottish healthy controls. Uncorrected p-values are given.**

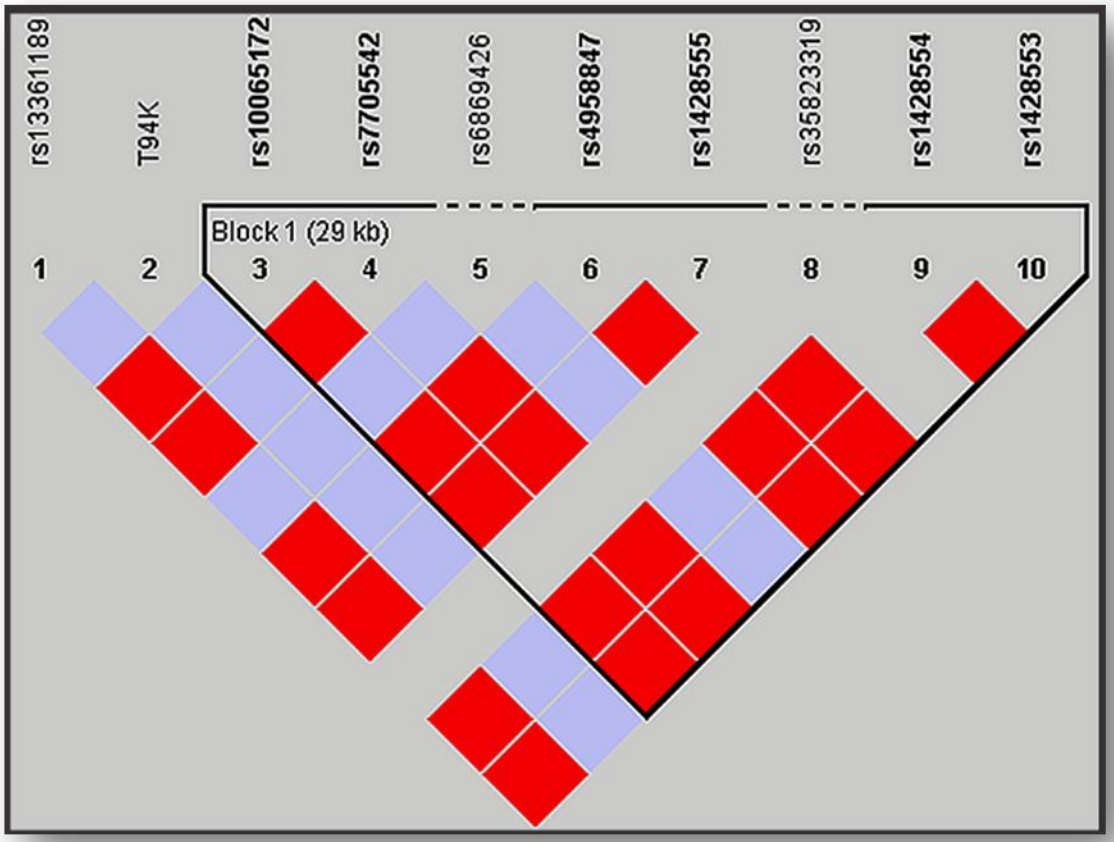
### 7.1.15.2Detailed sequencing of the 5 IRGM exons

By detailed sequencing of the 5 IRGM exons (including exon-intron boundaries) in 48 (24 adult + 24 childhood onset) CD patients, we identified 12 variants of which 7 were informative alleles in our population. (Figure 7-7 represents an overview of these variants generated in Haploview to illustrate the LD across this region) The 3 WTCCC variants genotyped in the whole Scottish CD cohort (rs10065172, rs13361189 and rs4958847) were then added to the analysis, giving 10 SNPs to perform haplotype analysis on: D' 1, paired $r^2$>0.82 except with rs4958847. Haplotypic variation (of haplotypes with frequency >1%) in this block was tagged by the two-marker haplotype consisting of rs10065172 and rs4958847. The informative alles were rs13361189 T/C (150,203,580: WTCCC SNP intronic), rs10065172 C/T (150,208,191: WTCCC SNP exon 1 (150,207,888-150,208,424) synonymous Leucine to Leucine at position 105), rs7705542 G/A (150,208,511: intronic), rs4958847 G/A (150,219,780: WTCCC SNP intronic), rs1428555 C/T (150,237,584: intronic), rs1428554 A/G (150,237,809: exon 2 (150,237,735-150,237,869) only present in 2 isoforms (IRGMc and IRGMe) however this SNP is not translated to aminoacid) and rs1428553 G/A (150,237,820: exon 2 (as for rs1428554)).

It is important to note that even on the most recent version of Ensembl (Release 49 – March 2008) the IRGM gene is not recognised as such. The reading frame of the genomic region containing the sequence corresponding to the IRGM gene, given on Ensembl, does not correspond with the work performed by Bekpen et al.[836] and by Prescott et al. (Natalie Prescott, personal communicaton unpublished data). As the numbering by Prescott et al. is the one used in the WTCCC paper, we have used this in our analysis.

**Figure 7-7: Detailed sequencing of the 5 IRGM exons (including exon-intron boundaries) in 48 CD patients. Of the 12 variants were identified of which 7 were informative in our population – genotype results of the three WTCCC SNPs were added to complement the haplotype analysis: D' 1, paired r2>0.82 except with rs4958847.**

**Figure 7-6: Detailed sequencing of the 5 IRGM exons (including exon-intron boundaries)**




### 7.1.15.3Haplotype susceptibility analysis and log-likelihood analysis

Haplotypic variation (including haplotype frequency >1%) of the genomic region containing the IRGM was tagged by the two-marker haplotype consisting of rs10065172 and rs4958847.

This two-marker haplotype was then used to perform a haplotype CD susceptibility analysis using Haploview software (10 000 permutations). No significant association of either the single markers or the two-marker haplotype was demonstrated. (in both the overall Scottish CD cohort and the adult onset CD cohort)

We then performed a log-likelihood analysis (1000 permutations) using this two-marker haplotype, as shown below for the overall Scottish CD cohort:

Log-likelihood analysis using 2-marker haplotype

| | |
|---|---|
| Chi-squared statistic for user-specified model | = 0.10, df=3, p=0.9918 |
| Chi-squared statistic for recessive model | = 1.68, df=3, p=0.6414 |
| Chi-squared statistic for dominant model | = 1.62, df=3, p=0.6548 |
| Chi-squared statistic for model-free analysis | = 1.78, df=4, p=0.7761 |
| Chi-squared statistic for heterogeneity model | = 1.70, df=3, p=0.6369 |

Random number seed = 1407

Number of replicates = 1000

User-specified model chi-squared statistic (0.10) was reached 304 times

Recessive model chi-squared statistic (1.68) was reached 428 times

Dominant model chi-squared statistic (1.62) was reached 416 times

Model-free chi-squared statistic (1.78) was reached 446 times

Heterogeneity model chi-squared statistic (1.70) was reached 423 times

Empirical p-values for these statistics are as follows:

| | |
|---|---|
| T1 - User specified model: | p-value = 0.3040 |
| T2 - Mendelian recessive model: | p-value = 0.4280 |
| T3 - Mendelian dominant model: | p-value = 0.4160 |
| T4 - Model-free analysis: | p-value = 0.4460 |
| T5 - Heterogeneity model: | p-value = 0.4230 |

Log-likelihood analysis of the adult onset CD cohort is shown below:

Chi-squared statistic for user-specified model = 0.22, df=3, p=0.9743

Chi-squared statistic for recessive model = 3.56, df=3, p=0.3131

Chi-squared statistic for dominant model = 3.34, df=3, p=0.3421

Chi-squared statistic for model-free analysis = 3.84, df=4, p=0.4281

Chi-squared statistic for heterogeneity model = 3.58, df=3, p=0.3105

Random number seed = 1306

Number of replicates = 1000

User-specified model chi-squared statistic (0.22) was reached 115 times

Recessive model chi-squared statistic (3.56) was reached 177 times

Dominant model chi-squared statistic (3.34) was reached 195 times

Model-free chi-squared statistic (3.84) was reached 195 times

Heterogeneity model chi-squared statistic (3.58) was reached 175 times

Empirical p-values for these statistics are as follows:

T1 - User specified model: p-value = 0.1150

T2 - Mendelian recessive model: p-value = 0.1770

T3 - Mendelian dominant model: p-value = 0.1950

T4 - Model-free analysis: p-value = 0.1950

T5 - Heterogeneity model: p-value = 0.1750

#### 7.1.15.4IRGM genotype-phenotype analysis in childhood and adult onset CD

Detailed genotype-phenotype analysis were performed (based on allelic frequency) in childhood and adult onset CD. In both adult onset and childhood onset, the Montreal location (at last follow-up but assigned before the first resection) was used. In childhood onset, CD behaviour was analysed in the group with at least 2 years follow-up. In adult onset disease, CD behaviour was analysed in the group with at least 5 years follow-up. Neither disease location nor disease behaviour were significantly influenced by any of the 3 IRGM variants (rs10065172, rs13361189 and rs4958847) (data not shown).

## *Discussion*

The high incidence of familial disease in children with Crohn's disease and ulcerative colitis is well-documented, and provides a strong catalyst in the search for genes that predispose to early onset disease. Our primary objectives were to establish whether the ATG16L1 variant rs2144880A/G and the IRGM variants (rs10065172, rs13361189 and rs4958847) are independent determinants of susceptibility to childhood onset CD, as suggested in datasets presented by Prescott et al. and Baldassano et al. for ATG16L1, and as a first investigation of its role in early onset CD for IRGM.[547;828]

Although our study was well-powered to replicate the findings in the Prescott and Baldassano reports, we did not observe association of the rs2144880G allele with childhood onset CD in our population. In our extension studies in the Scottish adult population, we were able to demonstrate an effect of this ATG16L1 polymorphism on CD susceptibility. Detailed genotype-phenotype analysis showed that the overall association signal was driven by a distinct effect on purely ileal involvement, which is uncommon in childhood onset disease in Northern Europe.

Inadequate power is a common reason for non-replication of positive genetic associations, in follow-up studies. Our study involving 269 children with CD had more than 97% power to replicate the association of germline variation of ATG16L1 with susceptibility to early onset CD with an effect size of OR>1.60, as calculated in the two smaller series.[547;828] (Table 7-1)

We increased the power of our study even further (to 80% power to detect an effect with OR 1.30 in childhood onset CD and OR 1.26 in adult onset CD) by adding the WTCCC control cohort to our group of controls. These two sets of controls were not significantly different and could therefore be added (comparing the allelic frequency in the two available control cohorts showed there was no statistically significant difference in allelic frequency (372/690 (53.9%) vs 3052/5872 (51.9%), p=0.33)). The allelic frequency in early onset CD also did not differ significantly from the allelic frequency in the expanded control cohort. (54.1% vs 3052/5872 + 372/690= 3424/6562 (52.2%), p=0.39).

In spite of this power, we did not replicate the association of ATG16L1 with CD in our early onset CD case-control analysis (frequency of rs214480 G allele in CD was 291/538 (54.1%) vs. controls 372/690 (53.9%), p=0.95 OR 1.01 (0.80-1.26)). (Table 7-3)

We explored two possible explanations for the apparent discordance in results compared with other reports. Firstly, by also examining the Scottish adult IBD population, we searched for genetic heterogeneity in the contribution of ATG16L1 to CD susceptibility between our population and the other populations studied to date. We hypothesised that a lack of association in adults, as well as children would provide substantial evidence for true heterogeneity between the Scottish and other populations analysed so far. In this context, we, and others have previously documented striking population-related differences in the contribution of NOD2/CARD15.[389] Recent reports have also described a similar geographical gradient for other key components of the innate immune response such as Toll-like Receptors 1,4,6 and 10.[835;844]

When comparing the allelic frequency of rs2144880G in the Scottish adult CD cohort (comprising 360 CD and 495 UC patients ) with population-matched controls (n=345), a significant association of ATG16L1 with CD was present (p=0.01 OR 1.32 (1.07-1.63)) and confirmed in the expanded control cohort (60.6% vs 52.2%, $p<10^{-4}$ OR 1.42 (1.21-1.66)).(Table 7-3)

These positive findings in adult disease are clearly not consistent with the hypothesis that the lack of observed association of ATG16L1 with Scottish early-onset CD is due to population-specific factors alone. Indeed, further analysis showed that the G variant allelic frequency in CD patients <17 years at diagnosis from Scotland, differed significantly from Scottish CD patients >17 years at diagnosis (p=0.01) and from the other early onset cohorts reported to date.(Tables 7-1 and 7-3)

We therefore investigated, and provided strong evidence in favour of our alternative hypothesis, namely that the observed lack of allelic association in our early onset cohort related directly to the characteristic disease phenotype of early onset disease, specifically the low proportion of isolated ileal disease and the high incidence of colonic involvement. The distribution of CD location in our childhood onset cohort is comparable to other populations in reports from Northern France, Sweden, Canada and the European childhood IBD registry under the auspices of ESPGHAN (European Society for Pediatric Gastroenterology and Nutrition).[541;542;569;845-847]

We found the allelic frequency of the rs2144880G variant to be lowest in isolated colonic disease (L2±L4: 87/172 (50.6%)), increasing to 55.4% (163/294) in ileocolonic disease (L3±L4) and 62.5% (25/40) in pure ileal disease (L1±L4). This trend was also present in adult onset CD.

Thus, detailed genotype-phenotype analysis demonstrated that this difference in contribution of ATG16L1 can be attributed to the relative rarity of pure ileal disease in CD diagnosed before the age of 17 years (only occurring in 20/257 (7.8%) of our early onset CD cohort).[516] Indeed, the allelic frequency of rs2144880G differed

significantly between pure ileal disease (L1±L4) CD patients and other CD patients (p=0.02, OR 1.34 (1.03-1.74)).(Table 7-4)

We confirmed the effect of ATG16L1 rs2241880 genotype (GG) on pure ileal disease versus pure colonic disease (p= 0.03 OR 2.43 (1.05-5.65)), in addition to the known effect of NOD2/CARD15 variant carriage (p<0.001 OR 3.13 (1.71-5.73)) and older age at diagnosis (p=0.02 OR 1.02 (1.00-1.03)) on multifactorial analysis using logistic regression, while also controlling for OCTN1&2 (IBD5) genotype.

We have demonstrated that replication of the effect of ATG16L1 on susceptibility to CD is critically dependent on the presence of a high ratio of ileal ('pure ileal'(L1±L4) + 'ileocolonic' (L3±L4)) disease to colonic ('pure colonic' (L2±L4)) disease in the CD population under study. In our early onset CD population this ratio is less than 2 ((20+147)/86), and there is no overall association of ATG16L1 with early onset CD in Scotland. In comparison, in the recent report from Philadelphia, Baldassano et al. replicated the association of early onset CD and germline variation of ATG16L1 in a smaller cohort characterised by a strikingly high frequency of ileocolonic CD (86.6%) and a low frequency of isolated colonic disease (5.7%).[547] The difference in phenotype between this and other paediatric cohorts bears consideration.

In our adult onset cohort, the association of CD with ATG16L1 rs2241880G is present with an OR similar to other adult CD studies reported to date. The strong effect of disease location on the susceptibility signal provided by the cohort under study is further illustrated by the strength of this signal (higher allelic frequency (G allele, 63.9%) and OR (1.49)) in the NIDDK IBD GC study which focused on 'pure ileal' CD.[687] (Table 7-1)

In contrast with the clear evidence in favour of the role of ATG16L1 in CD susceptibility in Scotland, is the picture emerging from our analysis of the 3 SNPs of greatest significance on the locus containing the IRGM gene.

Should this come as a surprise? Were we, as a single centre study focussing on childhood onset disease, not likely to be underpowered to replicate the latest findings of the genome-wide association studies? Or put differently, is the 'low-hanging fruit' been picked with the widespread replication/confirmation of NOD2/CARD15, IBD5, IL23R and ATG16L1?

The answer to the power-question is clearly negative. Our detailed power-calculations have shown that (admittedly by using the expanded WTCCC control cohort + Scottish healthy controls), our overall Scottish CD cohort had 80% power to replicate all of the strongest hits on the IRGM locus.

The composition of our Scottish CD cohort (ie the large proportion of early onset disease compared with other cohorts) is another possible explanation for the lack of association. Indeed for one of the IRGM variants, we did observe association (based on allelic frequency but without stringent Bonferroni-correction) with adult onset CD, but not childhood onset. The difference between these two cohorts' allelic frequencies was not statistically significant (rs13361189: 9.2% in adult onset CD vs 6.8% in childhood onset CD, p=0.12). The same was true for the two other IRGM variants studied.

In view of our findings for the ATG16L1 variant, we were careful to exclude a similar effect of phenotype on the susceptibility signal. For all three IRGM variants, no clear distribution of the allelic frequency based on ileal involvement was seen (neither in adult onset CD, nor in childhood CD).

In order to avoid missing population-specific low-frequency variants which would not have been sufficiently tagged by the genotyped IRGM SNPs in our CD cohort, we resequenced the 5 putative exons and exon-intron boundaries in 48 CD patients. According to our analysis, haplotype-frequencies >1% were tagged by the two-marker haplotype consisting of rs10065172 and rs4958847. Haplotype-susceptibility analysis and log-likelihood analysis (both in the overall Scottish CD cohort and the

adult onset CD cohort), confirmed the absence of association found on single-marker analysis.

Although the association with germline variation of IRGM was not reported in any of the other genome-wide association studies from North-America, Germany/UK and Belgium/France, the results of the meta-analysis (involving all except the Germany/UK non-synonymous SNP study) of these studies has confirmed IRGM as a susceptibility gene with OR 1.35 (Barrett et al, personal communication). The smaller sample size of the other genome-wide association studies in comparison with the large WTCCC is then likely to have played a role and could also explain why, in the frantic world of complex disease genetics, one year on from the first association report, replication in an independent cohort has not been published.[471;472;687;688] It is pertinent to point out that our power-calculations indicate that our study had adequate power to replicate the OR for all three IRGM variants, reported in the initial WTCCC study.

Genetic heterogeneity between the Scottish population and populations of more southern latitude, is one possible explanation for not being able to replicate the association between CD susceptibility and germline variation of IRGM. The genetic precedents illustrated above (eg NOD2/CARD15, TLRs) should be considered together with the stiking finding in the WTCCC report of geographic variation within the UK of several genomic loci with significance values equal or greater than most of the novel association signals.[835]

A more disconcerting possibility was recently raised by Terwilliger et al.[848] These authors raised doubts about some of the fundamental assumptions of the HapMap in terms of estimation of $r^2$ and importantly also about the fact that a causal association of a particular SNP with disease does not necessarily imply the absence of a three-way interaction involving disease, the causal locus, and some other marker locus in LD with it.[849] Thomas and Stram argued in response to the doubts raised by Terwilliger that '…Although there certainly are examples where gene–environment independence might be questionable in candidate gene association studies, this seems

unlikely for most situations involving genome-wide scans with SNPs for which there is no prior hypothesis about environmental modifiers...'.[849]

Whether the LD between loci and the etiological effect of the functional variant are independent of each other and whether they are statistically independent of all other etiological factors (in exposure and action), is very difficult to ascertain when we consider that the suscepitibility genes identified to date all impact on our ability to cope with infective exposures.[848] Then, these caveats may impact on the ability to replicate associations in different populations (or even sub-populations of large cohorts), who have likely been subjected in the past to different selective (infection-related) pressures. These caveats become even more important when we consider that the rise in incidence of CD over the last 4 decades is clearly environmental/epigenetic changes and when we consider how important germline variation of the autophagy pathway may have been in selecting out the European (and by virtue of migration also North-American) survivors of epidemics (as discussed in the Introduction of this chapter).

It is also pertinent to address whether the recent genome-wide association studies in CD have provided evidence for the existence of specific childhood onset genes. Several adult CD susceptibility genes have been replicated in childhood onset disease (NOD2/CARD15, IBD5 locus and IL23R), but only for NOD2/CARD15 there is some evidence for a gene-dosage effect on age at diagnosis.[278;660;706;850] It is noteworthy that relatively few early onset patients have been recruited in the consortia repositories so far (e.g. 206 children diagnosed with CD before their 17$^{th}$ birthday out of a total of 1902 in the Wellcome Trust Case Control Consortium).[700] Genome-wide association studies in CD to date have relied heavily on sporadic CD cases whereas familial disease is typically associated with younger age at diagnosis. A proxy-measure of success in identifying early onset IBD genes could thus be to assess how many genes have been identified on IBD loci. During the 1990s, these IBD loci were identified through the application of non-parametric linkage analysis in affected sibling pairs.[389] Only the NOD2/CARD15 gene (IBD1), markers in the HLA locus (IBD3) and markers on the IBD5 locus map to previously identified

susceptibility loci.[14;285] This suggests international collaborative efforts focussing on childhood onset IBD will be necessary to identify specific early onset genes on these IBD loci.

In summary, we have presented evidence for the important role of disease phenotype rather than age at diagnosis in determining the true effect of germline variation of ATG16L1 on disease susceptibility. We conclude that ATG16L1 (rs2144880) is associated with ileal CD, a relationship that is independent of age at diagnosis.

We have shown that germline variation of IRGM is not a strong determinant of CD susceptibility in Scotland. In contrast with our findings for ATG16L1, no clear phenotypic variation underlies this observation. As the genetic landscape of Crohn's disease is altered by the discovery of new determinants, these data illustrate that it remains critical to explore and define clearly genotype-phenotype relationships in at-risk populations and to assess the contribution of these novel determinants in cohorts of different latitude and age at diagnosis.

# 8 The role of Filaggrin loss-of-function variants in paediatric IBD

### *Introduction*

Great progress in our understanding of the genetic susceptibility to IBD has been made through the study of common germline variants in genome-wide association studies (GWAS).[851] Barrett et al. have recently reported on the meta-analysis of three GWAS from North America, the United Kingdom and Belgium/France: more than 30 loci have now been identified and replicated satisfying stringent genome-wide significance criteria.[694] Novel pathogenic mechanisms are being proposed and genetic evidence is accumulating for the empirical overlap with other auto-immune diseases such as psoriasis, ankylosing spondylitis and type 1 diabetes. However, the odds ratios (OR) associated with these common variants are small (1.1-1.5) and the penetrance of these variants is likely to be very small.[852] A current estimate of the variance in risk for Crohn's disease explained by the 32 identified loci is about 10% (which equates to about 20% of the genetic risk). Few of the common variants identified to date are thought to be the causal variants and while non-synonymous SNPs in linkage dysequilibrium or regulatory effects on gene expression are clearly important, they do not explain fully the genetic architecture of CD. Further fine mapping and direct sequencing are now indicated to search for rare variants on these loci.

Rare variants are more likely to have a greater contribution to the individual's genetic susceptibility to common diseases.[852] Fine mapping of regions identified through previous linkage analysis has elucidated causal variants with a low allelic frequency in two genetic success stories which are noteworthy in this context.

In Crohn's disease, variants of the intracellular pathogen-associated molecular pattern receptor NOD2/CARD15 (especially the 1007 frameshift mutation, allelic frequency in healthy controls <2%) with OR of 3.99 in the recent meta-analysis and 7.71 in our own childhood-onset CD population) dominate in explaining the variance in genetic risk to CD.[660;694;709]

In atopic dermatitis (eczema), null alleles of the epithelial barrier protein filaggrin (FLG) have been identified as the strongest genetic factor to date, for the development of atopic dermatitis.[853-855] These mutations also cause the semi-dominant skin-scaling disorder ichtyosis vulgaris.[854;856] While each of the identified FLG null-alleles are relatively rare (minor allele frequency <3% in population controls of European ancestry), collectively up to 10% of the population carry at least one null-allele.(Figure 9-1)

FLG is located within the epidermal differentiation complex on chromosome 1q21.3, a gene cluster expressed late in epidermal differentiation.[857] Previous linkage analysis studies have implicated this locus in susceptibility to atopic dermatitis as well as psoriasis and ulcerative colitis.[107;858]

Filaggrin is a crucial component of the cornified envelope in the outer layer of the epidermis.[859] The cornified envelope is a specialised membrane structure that forms during terminal differentiation of keratinocytes and consists of epidermal proteins that are cross-linked extensively by transglutaminases and enmeshed with keratinocyte-derived lipids, and is as such essential for the barrier function and hydration of the skin.[859]

Epidemiological as well as genetic evidence is mounting for an important overlap between diseases associated with epithelial barrier dysfunction. Even though atopic disease affects up to 25% of the general population, it is more prevalent still among patients with IBD.[589;596;860-862] Immune-mediated disease like atopy, psoriasis, IBD and multiple sclerosis have seen a sustained increase in incidence over the last 5 decades.[597] Most recently, the GWAS in childhood asthma has added the locus on chromosome 17q21 containing the ORMDL3 gene to the long list of shared genetic susceptibility determinants involved in both IBD and atopic disease (including among others the HLA-region, several interleukins (e.g. IL12B on the IBD5 locus),

prostaglandin-receptors, transforming growth factor beta, defensins, Toll-like receptors, NOD-like receptors, STAT3 etc).[694;863-865]

**Figure 8-1: Filaggrin molecular structure and position of loss-of-function variants**

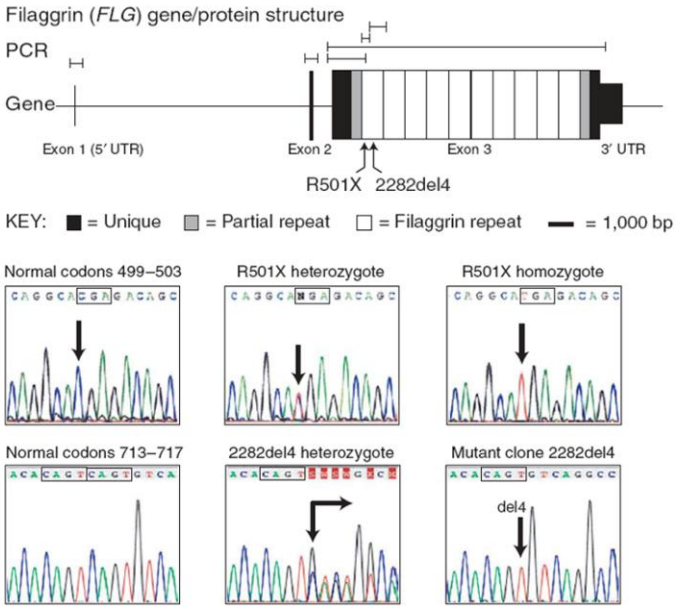



**Figure 9-1: Filaggrin molecular structure, position of loss-of-function variants ('null-alleles') and respective DNA sequence changes are illustrated (adapted from Smith et al.[856])**

### *Aims*

In the present study, we aimed to assess the contribution of the rare FLG null-alleles to the genetic susceptibility of IBD in the high incidence Scottish childhood-onset IBD population. As we have previously shown, this cohort of children with IBD is characterised by a high prevalence of atopic disease.[861] Therefore, we wanted to assess what role FLG null-alleles play in the genetic susceptibility to atopic disease in childhood-onset IBD.

## *Methods*

### *8.1.1 Subjects*

2082 individuals consisting of 403 children with IBD (263 CD, 97 UC, 44 IBD type unspecified), diagnosed before their 17th birthday (median age 11.3 years (8.8-13.1)), their parents (n=683) and 996 unselected paediatric population controls (recruited from a birth cohort in a single District General Hospital in Whitehaven, West-Cumbria, UK) were genotyped for FLG null-alleles R501X and 2282del4. More than 95% of the subjects and population controls were Caucasian.

IBD was diagnosed by standard criteria.[523] Detailed sub-phenotypic data based on the Montreal classification, have been reported previously.[516;710] During face-to-face interview with each family of which a child was diagnosed with IBD, a detailed history was taken by a member of the research team (LS, RKR or JVL) for the presence of any family-reported atopic disease: asthma, atopic dermatitis (eczema), allergic rhinitis(AR)/hayfever and food allergy (FA). Specifically before marking any food allergy, the food antigen and the allergic reaction after exposure to it were discussed with the family. As all paediatric IBD patients were recruited for an ongoing IBD genetics project, additional food challenges or specific IgE assays were not performed as they were not part of the original study protocol which was approved by the local ethics committees at the participating centres (Edinburgh, Glasgow, Aberdeen and Dundee – LREC 2002/6/18).

### 8.1.2 Genotyping

All individuals were genotyped for the FLG null-alleles R501X and 2282del4 using the TaqMan system (7900HT sequence detection system; Applied Biosystems, Foster City, CA, USA) at the Epithelial Genetics Group, Human Genetics Unit, University of Dundee. The FLG R501X and 2282del4 genotype frequencies in population controls conformed to the Hardy-Weinberg equilibrium ($p>0.15$).

The 3 most common NOD2/CARD15 polymorphisms (R702W, G908R and Leu1007fsinsC) were genotyped as previously described.[660]

### 8.1.3 Statistical analysis

Allelic, genotype and carriage of any null-alleles frequencies were compared between cases and controls using Pearson's $\chi^2$ or Fisher's Exact test (when $n<6$) (Minitab software, Release 13.20, Minitab Inc., State College, PA, USA and GraphPad Instat software version 3.06, San Diego, Ca, USA). TDT analysis was performed using PLINK software (http://www.journals.uchicago.edu/doi/abs/10.1086/519795).[534]

Uncorrected p-values are given. In IBD and CD, multifactorial analysis (correcting for NOD2/CARD15 variant carriage) was performed using Minitab. The Mann-Whitney U test was used to assess the relationship between carriage of a FLG null-alleles with z-score of height, weight and BMI at diagnosis. Z-scores were calculated using nomograms from the British 1990 growth reference (Child Growth Foundation, 2 Mayfield Avenue, London W4 1PW).

### 8.1.4 Power calculations

Power calculations were performed using QUANTO version 1.2 (http://hydra.usc.edu/gxe).[528;529] Based on the most recent prevalence data of Scottish paediatric IBD (23/100 000 of the population <17 years), our case-control analysis in early onset IBD had 80% power to detect an effect with Odds Ratio (OR) 1.41 or greater.[4]

### Results

### 8.1.5 IBD susceptibility & genotype-phenotype analysis

Association analysis based on allelic, genotype or carriage of at least one FLG null-alleles frequencies, showed no effect of FLG null-alleles on IBD, CD, UC or IBDU susceptibility. (Tables 9-1, 9-2 and 9-3). Family-based association analysis showed no distortion of transmission of either the R501X or 2282del4 null-alleles in IBD, CD or UC (R501X: all p>0.50; 2282del4: all p>0.40).

Detailed genotype-phenotype analysis based on the Montreal classification, with stratification of the CD cases for the presence of atopic disease, showed no effect of FLG null-alleles on CD location or behaviour (data not shown). In IBD and CD cases, carriage of a FLG null allele had no effect on z-scores of height, weight and Body Mass Index at diagnosis (data not shown).

**Table 8-1: Case-control analysis of FLG R501X in childhood-onset IBD, CD, UC and IBDU.**

**Table 8-2: Case-control analysis of FLG 2282del4 in childhood-onset IBD, CD, UC and IBDU.**

**Table 8-3: Case-control analysis of carriage of at least one FLG null-alleles in childhood-onset IBD, CD, UC and IBDU.**

| **Filaggrin R501X Wild-type=1 Null (X)=2** | | Controls | IBD | p | CD | p | UC | p | IBDU | p |
|---|---|---|---|---|---|---|---|---|---|---|
| | 11 | 941/996 (94.5%) | 354/381 (92.9%) | 0.27 | 224/245 (91.4%) | 0.07 | 87/92 (94.6%) | 0.97 | 32/33 (97.0%) | 0.53 |
| Genotype frequencies | 12 | 53/996 (5.3%) | 26/381 (6.8%) | 0.28 | 21/245 (8.6%) | 0.05 | 4/92 (4.3%) | 0.68 | 1/33 (3.0%) | 0.56 |
| | 22 | 2/996 (0.2%) | 1/381 (0.3%) | 0.82 | 0/245 | 0.48 | 1/92 (1.1%) | 0.12 | 0/33 | 0.79 |
| Allele | 2 | 57/1992 (2.9%) | 28/762 (3.7%) | 0.26 | 21/490 (4.3%) | 0.10 | 6/184 (3.3%) | 0.75 | 1/66 (1.5%) | 0.51 |

**Table 8-1: Case-control: FLG R501X in childhood-onset IBD, CD, UC, IBDU.**

| **Filaggrin 2282del4 Wild-type=1 delACTG=2** | | Controls | IBD | p | CD | p | UC | p | IBDU | p |
|---|---|---|---|---|---|---|---|---|---|---|
| | 11 | 899/951 (94.5%) | 366/381 (96.1%) | 0.24 | 233/244 (95.5%) | 0.54 | 88/92 (95.7%) | 0.64 | 34/34 | 0.16 |
| Genotype frequencies | 12 | 52/951 (5.5%) | 15/381 (3.9%) | 0.24 | 11/244 (4.5%) | 0.54 | 4/92 (4.3%) | 0.64 | 0/34 | 0.16 |
| | 22 | 0/951 | 0/381 | | 0/244 | | 0/92 | | 0/34 | |
| Allele | 2 | 52/1902 (2.7%) | 15/762 (2.0%) | 0.25 | 11/488 (2.3%) | 0.55 | 4/184 (2.2%) | 0.65 | 0/68 | 0.16 |

**Table 8-2: Case-control: FLG 2282del4 in childhood-onset IBD, CD, UC, IBDU.**

| Fillagrin Null allele carriage | | Controls | IBD | p | CD | p | UC | p | IBDU | p |
|---|---|---|---|---|---|---|---|---|---|---|
| At least one null allele | | 103/944 (10.9%) | 40/378 (10.6%) | 0.86 | 31/243 (12.8%) | 0.41 | 8/91 (8.8%) | 0.53 | 1/33 (3.0%) | 0.14 |

**Table 8-3: Case-control analysis of carriage of at least one FLG null-alleles in childhood-onset IBD, CD, UC and IBDU. Out of 378 IBD patients, there were only 2 compound heterozygotes (1 CD / 1 UC).**

### *8.1.6 Atopy in childhood-onset IBD*

Asthma was present in 25% of IBD patients, eczema in 29%, allergic rhinitis in 21% and food allergy in 12% of IBD cases. In total, 52% of IBD patients had at least one manifestation of atopic disease.

In IBD and CD cases, carriage of a FLG null-allele differed significantly between patients with and those without co-existent atopic disease (13.8% (27/195) vs 6.1% (11/179), p=0.01 OR 2.45 95%CI 1.18-5.11 and 16.3% (22/135) vs 7.6% (8/105), p=0.04 OR 2.36 95%CI 1.01-5.54, respectively).

The atopic manifestations driving this association in both IBD and CD cases, were eczema and food allergy. In IBD cases, stratified for the presence of eczema, carriage of a FLG null-allele occurred in 19.3% (21/109) of IBD+eczema cases compared with 6.7% (17/254) of IBD without co-existent eczema (p=0.0003 OR 3.33 95%CI 1.68-6.60). We also observed a significant difference of carriage of FLG null-allele between the subgroup of IBD+eczema patients and population controls (19.3% vs 10.9%, p=0.01 OR 1.95 95%CI 1.16-3.27).

Stratifying IBD cases for the presence of food allergy, demonstrated the association of co-existent FA in IBD with carriage of a FLG null-allele (28.2% (11/39) vs 8.1% (26/322), p=0.0001 OR 4.47 95%CI 2.00-10.00)). The frequency of carriage of a FLG null-allele differed significantly between IBD+food allergy cases and population controls (28.2% vs 10.9%, p=0.0009 OR 3.21 95%CI 1.55-6.64).

Stratification for the presence of asthma and allergic rhinits, did not show a different carriage frequency of FLG LOFV between IBD cases affected by asthma or allergic rhinitis and IBD cases without co-existent asthma or allergic rhinitis.

We demonstrated that atopic co-morbidity additional to the presence of eczema in IBD cases, increased the effect size (odds ratio) of the association. Carriage frequency of a FLG null-allele was significantly different between:

1) IBD patients with co-existent asthma and eczema (22.9% (11/48) vs IBD without asthma and eczema 7.1% (15/210), p=0.001 OR 3.86 95%CI 1.65-9.08 and vs population controls 10.9%, p=0.01 OR 2.43 95%CI 1.20-4.91)

2) IBD+asthma+eczema+allergic rhinitis (29.2% (7/24) vs IBD without asthma, eczema and allergic rhinitis (7.1% (13/184), p=0.0006 OR 5.42 95%CI 1.90-15.41 and vs population controls 10.9%, p=0.005 OR 3.16 95%CI 1.36-8.30)

3) IBD+asthma+eczema+allergic rhinitis+food allergy (45.5% (5/11) vs IBD without asthma, eczema, allergic rhinitis and food allergy 6.4% (11/172), p=0.0009 OR 12.19 95%CI 3.21-46.35 and vs population controls 10.9%, p=0.004 OR 6.80 95%CI 2.04-22.69).

In childhood-onset CD, carriage of a FLG null-allele differed between cases also affected by eczema or food allergy and CD cases without (20.3% (16/79) vs 8.7%

(14/161), p=0.01 OR 2.67 95%CI 1.23-5.79 and 29.0% (9/31) vs 9.6% (20/208), p=0.002 OR 3.85 95%CI 1.56-9.48, respectively). Like in IBD overall, co-existence of different manifestations of atopy in addition to eczema and food allergy increased the effect size associated with the association: asthma+eczema (p=0.007 OR 3.54 95%CI 1.35-9.30), asthma+eczema+allergic rhinitis (p=0.0005 OR 7.00 95%CI 2.07-23.71) and asthma+eczema+allergic rhinitis+food allergy (p=0.004 OR 12.12 95% CI 2.54-57.85). Like in IBD, these same associations were also seen when comparing paediatric CD cases with increasing atopic comorbidity in addition to eczema with population controls (data not shown).

Multifactorial analysis (also correcting for carriage of any of the three CD associated variants of NOD2/CARD15) in both IBD and CD cases, showed an association of carriage of a FLG null-allele with co-existent atopic disease (p=0.01 OR 2.4 95%CI 1.2-5.1) and co-occurrence of different forms of atopy (p=0.003 OR 3.5 95%CI 1.5-8.1). These associations did not alter when NOD2/CARD15 variant carriage and atopic disease where entered as interacting variables in the logistic regression model.

## *Discussion*

We report the first study into the contribution of FLG null-alleles to childhood-onset IBD susceptibility and to the high prevalence of co-existent atopic disease in childhood-onset IBD. We have demonstrated that the two most common FLG null-alleles and carriage of at least one FLG null allele, are not associated with increased inherited risk to the development of IBD *per se*. Our study was adequately powered to detect effect sizes greater than OR 1.40.

Our study is consistent with a previous report in adult-onset IBD by Ruether et al. in a population from Northern Germany, also showing no effect of these variants on disease susceptibility.[866] These null-alleles have also been studied in psoriasis and rheumatoid arthritis (RA).[867-869] Autoantibody formation to citrullinated (pro)filaggrin has been proven to be a highly specific serological marker for

rheumatoid arthritis and the FLG R501X and 2282del4 null-alleles may contribute to the development of humoral autoimmunity against citrullinated peptides.[870] However, as in psoriasis, the FLG LOFV do not contribute to the genetic susceptibility of RA.[867-869;871]

FLG null-alleles have consistently been shown to be a major genetic risk factor for atopic dermatitis and asthma, particularly in the context of atopic dermatitis.[854;872-887] A recent meta-analysis has added further weight to this conclusion.[855]

We have shown that FLG null-alleles influence co-existent atopic disease in IBD, particularly eczema and food allergy, with the novel finding of an increased effect size (odds ratio) with increasing atopic disease-load. Although decreasing numbers in the subgroups of IBD patients with more atopic co-morbidity lead to a widening of the confidence intervals associated with the increments in the corresponding odds ratios, this trend is noteworthy.

The contribution of the two FLG null-alleles studied in the present report to atopic co-morbidity in childhood-onset IBD, is comparable to that in the non-IBD population. In the study by Palmer et al., 22.9% (64/279) of patients affected by asthma and eczema carried at least one null-allele.[872] This is identical to the percentage of IBD patients affected by asthma+eczema carrying at least one FLG null allele in our study (11/48).

The high prevalence of atopic disease in IBD, reported by our group and others, has cast further doubt on the relevance of the Th1-Th2 dichotomy to IBD.[583;618;861] The progress made in our understanding of the differentiation of Th17-cells and the crucial role of IL23R in susceptibility to CD and UC, illustrate that merely relying on the cytokine profiles in immune mediated diseases to classify them as either Th1-driven or Th2-driven does not do their complex pathogenesis justice.[694;697;702] Epidemiological considerations, particularly the ‘hygiene hypothesis’, the geographical variation in incidence and the higher incidence of CD in more affluent

strata of the population, have further substantiated genetic and immunological objections to this old paradigm.[566;583;618;888;889] Genetic overlap between susceptibility determinants in a number of immune-mediated diseases, notably different manifestations of allergic disease and IBD, have come to light in the course of the application of GWAS to the study of these complex polygenic diseases.[890;891] For example, germline variation of the IL12B gene (on the IBD5 locus) has been associated with increased susceptibility to asthma and CD and UC.[694;702;727;864] Also in the case of the most recent addition to the list of genetic determinants of asthma, the CHI3L1 (chitinase 3-like 1) gene, there is clear evidence for its potential role in IBD pathogenesis.[892] Mizoguchi et al. have shown that CHI3L1 is up-regulated specifically in inflamed mucosa (both lamina propria and colonic epithelial cells) in several murine colitis models and ulcerative colitis and Crohn's disease patients but absent in normal controls. CHI3L1 is required for the enhancement of adhesion and internalization of bacteria in colonic epithelial cells CEC.[893]

Further functional parallels can be drawn when we consider that recent insights into the genetic determinants of asthma/eczema susceptibility have illustrated how atopy could be the inappropriate adaptive immune process secondary to an altered (innate) immune function of the epithelial barrier.[894]

Of note in our study, is the strong independent association of food allergy with carriage of a null-allele in the context of IBD. The expression of (pro)filaggrin in the gastrointestinal tract could potentially shed further light on this association. De Benedetto et al. recently demonstrated that FLG is not expressed in the oesophagus whereas our own data suggest FLG is expressed in the human colon (AT Evans, manuscript in preparation) (see Figure 9-2).[895] The parallels with the absence of filaggrin expression in the bronchial mucosa and the strongest effect on asthma in the context of atopic dermatitis are striking.[896] Allergic sensitisation through an altered epithelial barrier (either via the skin or the gastrointestinal tract (colon)) could conceivably result in local symptoms (in the case of eczema and colitis e.g. secondary to cow’s milk protein allergy (CMPA)) and only secondarily, in organs where there is no demonstrable FLG-related epithelial barrier defect via an up-

regulated adaptive immune response (e.g. atopic asthma and CMPA-driven gastrooesophageal reflux disease).[860]

Akei et al. have demonstrated that epicutaneous exposure to allergens can potently prime for the subsequent development of eosinophilic esophagitis.[897] In turn, gastrointestinal allergy can induce allergic pulmonary responses to specific and unrelated allergens.[898] The intriguing recent observation that an altered colonic flora in infancy may precede the onset of atopic disease in the KOALA Birth Cohort study, adds further support to the central role in disease pathogenesis now awarded to an altered innate immune response (including the epithelial barrier integrity).[599] The study of the innate immune response is therefore correctly receiving increasing research interest as a key contributor to the subsequent dysregulated inflammation which is of relevance to the pathogenesis of atopic disease and IBD alike.

In summary, we have demonstrated that carriage of FLG null-alleles influences the high prevalence of atopic co-morbidity in childhood-onset IBD without a significant contribution to the inherited susceptibility of IBD itself. As the contribution of FLG null-alleles to atopic co-morbidity in paediatric IBD appears to be similar to that in the non-IBD population, other determinants (genetic/environmental) likely underlie the higher prevalence of atopic disease in IBD. The identification of relatively rare variants as major determinants of inherited susceptibility in both eczema (FLG) and CD (NOD2/CARD15) illustrate the need for detailed resequencing and fine mapping of the loci identified in the recent series of GWAS. The epidemiological evidence for clustering of diseases within families and in individual patients, will undoubtedly continue to prompt interest in molecular pathways, identified in co-existent immune-mediated diseases.

**Figure 8-2: Expression of Filaggrin in the human gut (3 panels).**

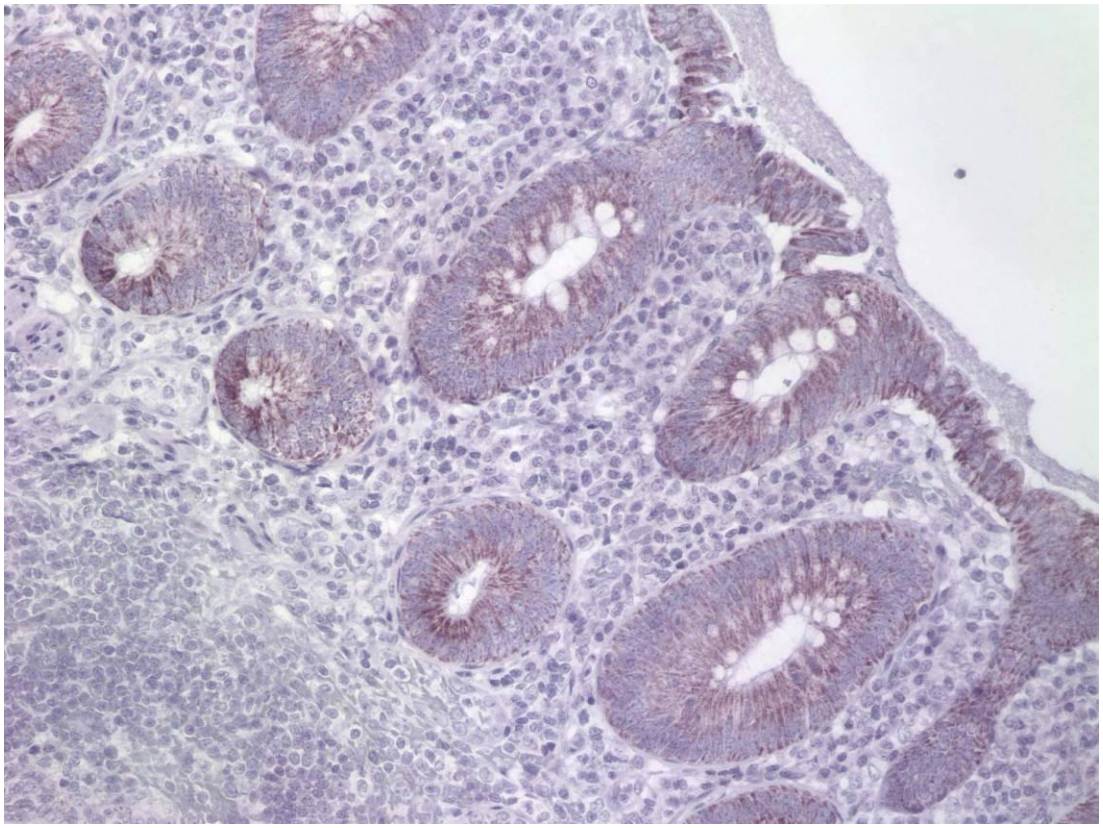
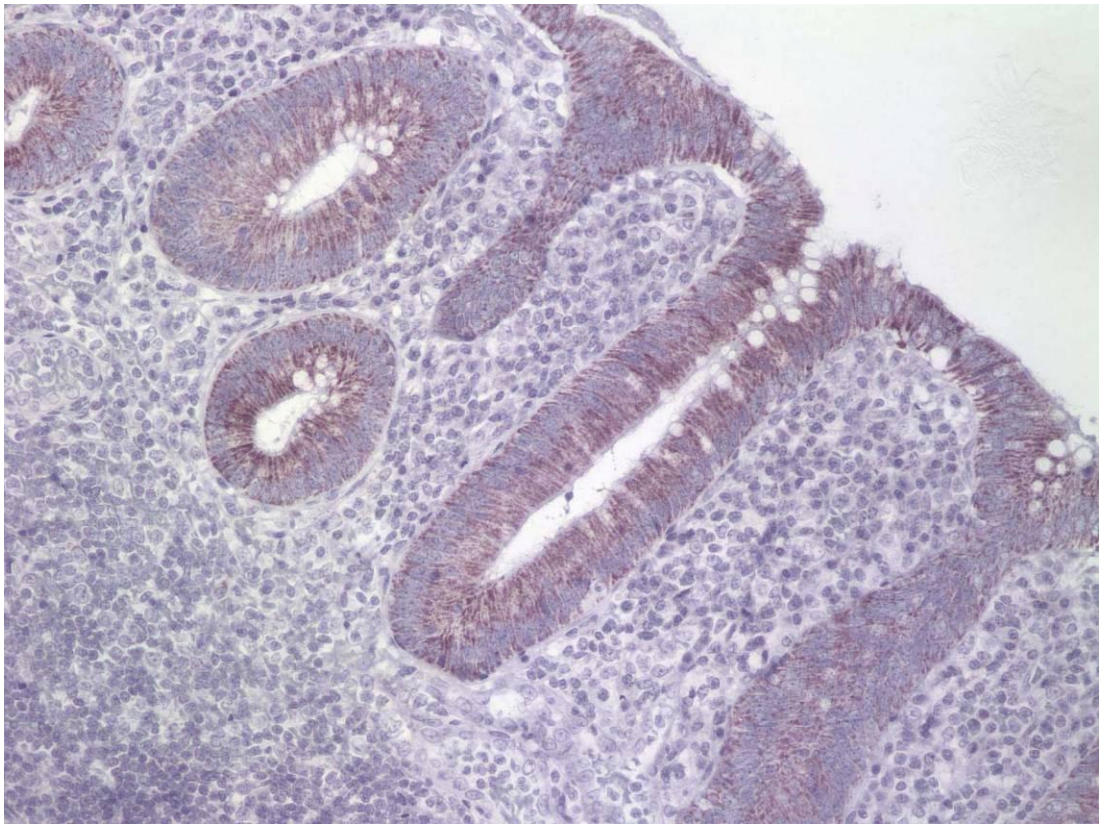
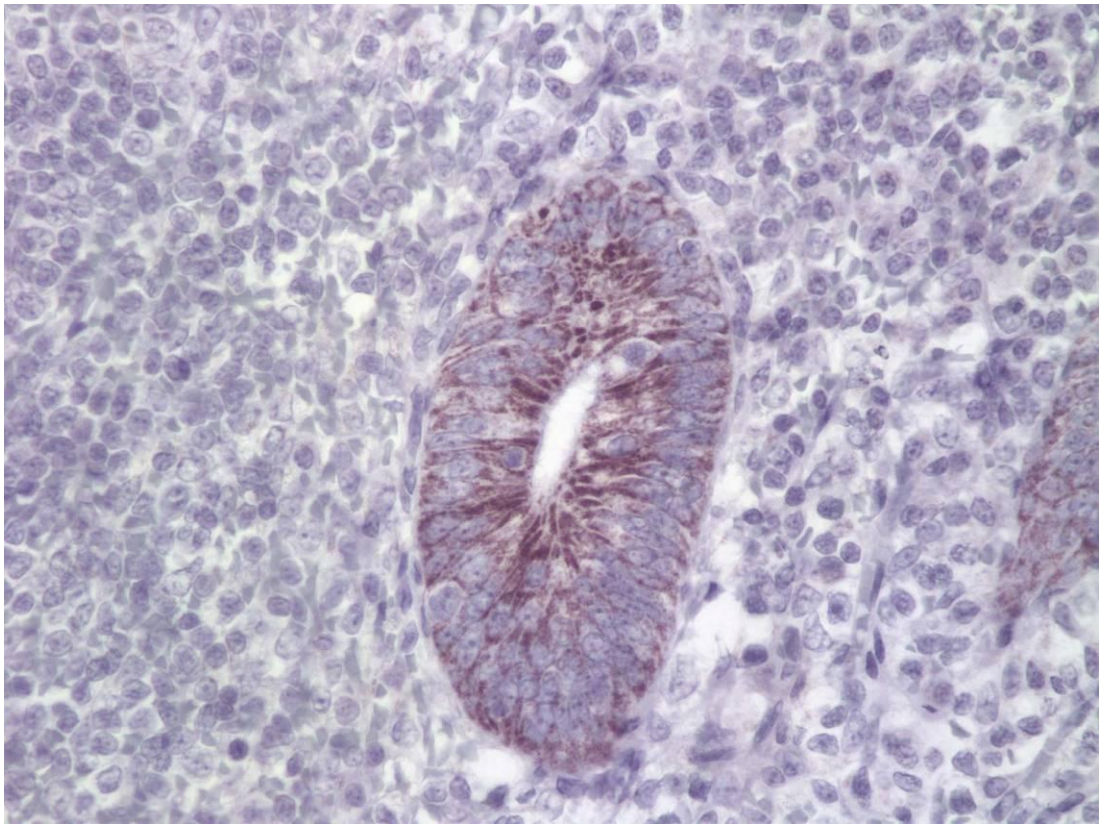


**Figure 8-2: Expression of Filaggrin in the human gut. Immunoperoxidase staining of paraffin-embedded sections was performed by using the ChemMate Peroxidase/DAB system (DakoCytomation, Hamburg, Germany). Mouse mAb 15C10 (Novocastra, Newcastle, UK) was used to detect the human filaggrin repeat unit. Antigen retrieval of paraformaldehyde-fixed, paraffin-embedded sections was performed by heating sections under pressure for 10 to 15 minutes in 10 mmol/L citrate buffer, pH 6.0. (Images by Dr Alan T Evans, Department of Pathology, Tayside University Hospitals NHS Trust, Dundee)**





# 9 An overview of recent developments in Crohn's disease genetics and concluding remarks

## *Introduction*

Crohn's disease (CD) and ulcerative colitis (UC) are chronic relapsing and remitting diseases of the gastrointestinal tract, commonly referred to as inflammatory bowel disease (IBD). Due to a sharp rise in incidence in the last third of the 20$^{th}$ century, IBD now affects up to 1 in 250 of the adult population in the Western world.[1]

CD distinguishes itself from UC in that the inflammation of CD is characteristically discontinuous, transmural (affecting all layers of the bowel wall) and often granulomatous. CD inflammation can occur anywhere in the gastrointestinal tract from mouth to anus: whereas in adults, disease is often limited to the ileum (the last portion of the small bowel) and/or colon, in children, CD is typically more extensive.[710] While treatment advances have reduced IBD mortality considerably, its residual morbidity and impact on health care resources can be significant.[565] Onset during childhood and adolescence (up to 25% of all IBD patients) can have devastating effects on all aspects of the development of the young IBD sufferer and her/his family. The characteristically extensive involvement of the gastrointestinal tract in childhood onset IBD and rapid progression of disease are important contributing factors to overall disease burden.[710]

Although the precise aetiology of CD remains elusive, epidemiological data conclusively point to a dysregulation of the immune response against the luminal flora in a genetically susceptible host. This dysregulation of the immune response is often heralded by an acute infective trigger.[97;899] Although traditional pathogens are probably not responsible for IBD, alteration of the composition and function of the microbiome (as either a primary or a secondary phenomenon) is the subject of intense investigation.[96] Other environmental factors are thought to be involved but there is only a substantive evidence base for smoking and appendicectomy, when it was performed without evidence of appendicitis or mesenteric lymphadenitis.[581;595]

Several genetic epidemiological observations provided the impetus to embark on larger-scale genetic studies. The high concordance for CD in monozygotic twins (36%, compared with 16% for UC), the higher risk of IBD to a first degree relative if the proband has CD (relative risk (RR) 5-35) rather than UC (RR 10-15) as well as the high rate of a family history of IBD (up to 30%, especially in cohorts of childhood-onset IBD, characterised by a high proportion of CD) have provided strong epidemiological evidence for a genetic contribution to CD susceptibility.[578;710]

These observations led to the application of non-parametric linkage analysis in 11 genome-wide searches for regions of the genome involved in IBD during the 1990s (Figure 9-1).[389] In these studies, cosegregation with CD of a (comparatively) limited set of markers spread across the genome was assessed in affected sibling pairs. A large number of putative and confirmed linkage regions were thus described and later subjected to a meta-analysis.(see figure 9-1)[107;389] Recently, this family-based hypothesis-free linkage approach has come under scrutiny due to the failure of the recent genome-wide association studies (GWAS) to identify disease-associated variants in many of these previously identified regions. Two issues are noteworthy in this respect: the proportion of CD patients recruited into the recent GWAS with a family history has been low (e.g. 15% in the Wellcome Trust Case Control Consortium (WTCCC)) and linkage analysis was instrumental in the initial identification of several of the strongest signals in the recent GWAS, notably IBD1 (and the subsequent discovery of NOD2/CARD15), IBD3 (including the HLA-region) and the IBD5 region.(149)

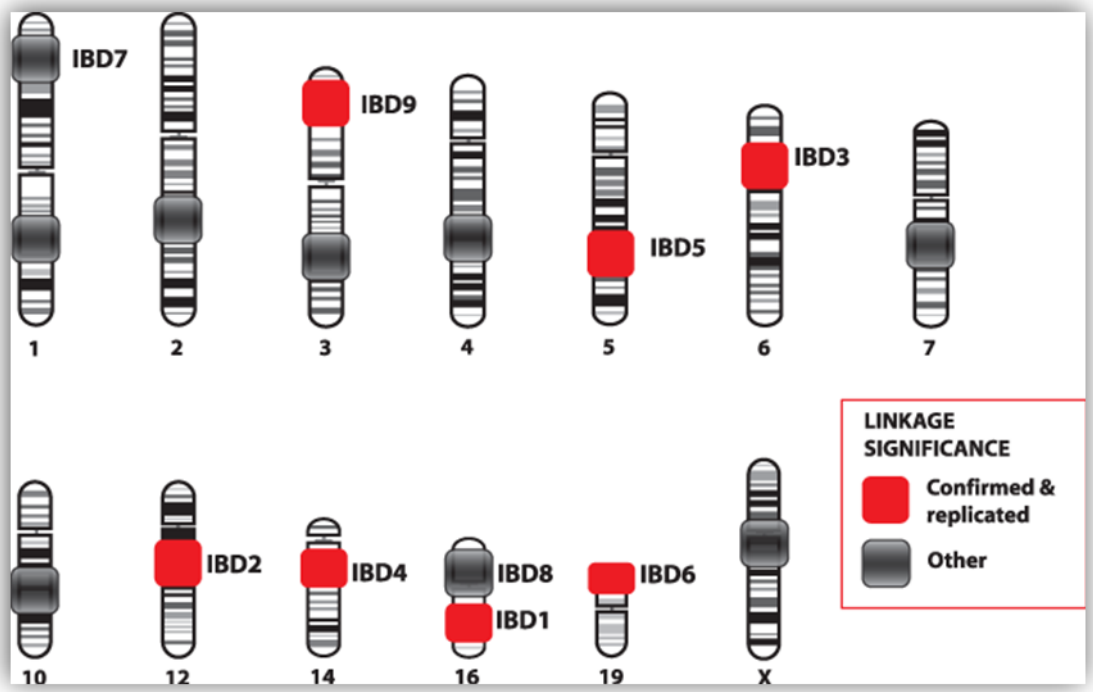



**Figure 9-1: IBD susceptibility loci and subchromosomal regions which were identified in genome-wide scans using nonparametric linkage analysis.**

**Figure 9-1: IBD susceptibility loci and subchromosomal regions which were identified in genome-wide scans using nonparametric linkage analysis**

Reductions in cost of genome-wide strategies with increasing 'coverage' of the genome have facilitated the shift from hypothesis-driven candidate gene research to hypothesis-free genome-wide studies: initially using linkage analysis but more recently GWAS. The limitations of linkage analysis (access to a large set of affected siblings and the lack of power to identify genes of weak effect), have led researchers of all common complex polygenic conditions to turn to GWAS, which involves the analysis of several hundred thousand markers, distributed across the whole genome.(see Table 10-1 for an overview of the GWAS performed to date in CD) The scale of these GWAS has necessitated the setting of stringent statistical criteria for multiple testing (so called 'Bonferroni'-correction), thus requiring p-values smaller than $5 \times 10^{-8}$.[694]

**Table 9-1: Overview of the GWAS performed to date in CD**

| Location | Novel Gene/Locus identified |
|---|---|
| Japan-UK[900] | TNFSF15 |
| Germany-UK[688] | ATG16L1 |
| North-America[470;687] | IL23R, 10q21.1 'gene desert' (EGR2), PHOX2B, NCF4, FAM92B |
| Belgium[471] | 5p13.1 'gene desert' (PGER4) |
| UK[472] | IRGM, NKX2-3, PTPN2, 3p21 1q 'gene deserts', IL12B, FLJ45139 |
| Canada-Germany[704] | 3p21, 4p16.1, 17q11, 17q23 |
| The Netherlands-UK[901] | IL18RAP, CARD9 |
| USA/Italy (Paed)[732] | TNFRSF6B, PSMG1 |

**Table 9-1: Overview of the GWAS performed to date in CD**

Following the recent 'gold rush' in identifying common variants with small effect size on the genetic susceptibility to CD, interest has been renewed interest in 'candidate-gene analysis' of loci identified through GWAS which can be very gene-dense and characterised by an extended haplotype structure (e.g. the HLA-region),

both of which can hamper the identification of causal variants using classical genetic techniques.

In this chapter, I will describe the loci, identified to date, which have made the strongest impact on our understanding of CD pathophysiology and/or provide strong therapeutic potential. These loci encode genes involved a number of homeostatic mechanisms: innate pattern recognition receptors (NOD2/CARD15, TLR4, CARD9), the differentiation of Th17-lymphocytes (IL23R, JAK2, STAT3, CCR6, ICOSLG), autophagy (ATG16L1, IRGM, LRRK2), maintenance of epithelial barrier integrity (IBD5, DLG5, PTGER4, ITLN1, DMBT1 and XBP1) and the orchestration of the secondary immune response (HLA-region, TNFSF15/TL1A, IRF5, PTPN2, PTPN22, NKX2-3, IL12B, IL18RAP, MST1).

Complementary to the GWAS performed to date focussed on adult CD, the assessment of individuals with a younger age of disease onset and a characteristically more extensive phenotype, is likely to prove valuable, as shown recently for asthma.(95) I will therefore also discuss the relevance of GWAS in determining the genetic susceptibility to childhood-onset IBD, as we have contributed nearly 500 samples to the GWAS performed by the international pediatric IBD genetics consortium, which is due to report shortly on its findings.

## *Innate pattern recognition*

### *9.1.1 NOD2/CARD15*

Further fine mapping of the region on chromosome 16 (IBD1), which was first identified by Hugot and colleagues in 1996, led to the identification of the NOD2/CARD15 gene (nucleotide-binding oligomerisation domain 2)/CARD15 (caspase activation recruitment domain 15).[14] Its discovery was one of the first success stories in complex polygenic disease genetics and has been responsible for a renewed interest in the innate immune response in IBD.[666]

NOD2/CARD15 is a member of a family of intracellular pattern-recognition receptors (PRRs) which recognise microbial components. Muramyl dipeptide (MDP), a degradation product of peptidoglycan derived from the cell wall of Gram-negative and Gram-positive bacteria, are recognised by the Leucine Rich Repeat (LRR) domain of NOD2/CARD15.(see Figure 9-2) NOD2/CARD15 expression appears to be limited to Paneth cells, which are located at the base of the intestinal crypts, and influenced by inflammatory cytokines like Tumour-Necrosis-Factor-α.[193;435] The three main CD associated NOD2/CARD15 variants (accounting for >80% of the identified germline variants) are situated in or close to the LRR-domain and are thought to interfere with bacterial recognition.[18;20] The association of NOD2/CARD15 variants with increased susceptibility to CD was confirmed in two comprehensive meta-analyses.[468;469] Initial optimism about the proportion of population attributable risk explained by these variants has been tempered by the observation that homozygosity/compound heterozygosity for the disease associated alleles is not underrepresented in Caucasian healthy controls.[902] Nevertheless, germline variation of NOD2/CARD15 has remained the strongest genetic determinant of genetic CD susceptibility (Odds Ratio 3.99 in the recent GWAS meta-

analysis).[694] The association of NOD2/CARD15 variant carriage with disease phenotype has been driven especially by ileal disease and the reduced secretion of α-defensins by ileal Paneth cells in patients carrying NOD2/CARD15 variants is important in this respect. [468;502]

Although functional analyses in human peripheral mononuclear cells point to a 'loss of function' due to the CD associated polymorphisms, related to an MDP sensing defect, NOD2/CARD15 signalling is but one instrument in an orchestra of innate immunity signalling pathways.[17;666] Illustrative of this complex regulation were the conflicting early results reported on the interaction between TLR-2 and wildtype NOD2/CARD15 signalling by Watanabe et al. and Netea et al. [383;386] More recently, additional experiments by Clark et al. have shown the crucial role of a MAP3K, MEKK4, in regulating the signal transduction beyond NOD2/CARD15 through the activation of RIP2 and NFkB.[903] Uninhibited NOD2/CARD15 signalling is associated with hyporesponsiveness to both TLR2 and TLR4 ligands.[903;904] This induction of cross-tolerance to TLR4 signalling is lost in CD patients homozygous for the 1007frameshift (also called 3020insC) mutation which truncates the LRR domain of NOD2/CARD15.[487]

Two additional ways of NOD2/CARD15 regulation have recently been described. A short isoform of NOD2/CARD15, NOD2-S, which lacks the third exon (thus encoding for a protein that is truncated within the second CARD domain), is preferentially expressed in the human colon and is up-regulated by the anti-inflammatory IL10.[457] Overexpression of NOD2-S down-regulates NOD2/CARD15-induced NFkB activation, IL8 and pro-IL1β downstream of NOD2/CARD15 and its adaptor molecule RIP2.[457]

A second type of regulation of NOD2/CARD15 function depends on the recently described 5' UTR splice forms of NOD2/CARD15.[905] These isoforms not only show a TNF-α/intestinal inflammation-sensitive expression pattern but also link

NOD2/CARD15 regulation with the autophagy response (which will be discussed later), through a rapamycin-sensitive inhibitory effect of these novel upstream Open Reading Frames on translation efficacy.[905]

In view of the recent associations of CD genetic susceptibility with TLR4 and CARD9 polymorphisms (as discussed below and see figure 9-2), these single or dual pathway experiments are clearly not appropriate to understand the full complexity of signalling variation related to sequence changes.[395;901;906]

The failure to reliably identify genetic interaction between the large number of genetic variants identified to date, does not necessarily mean an absence of interaction at the functional level.[694] As the physiological mucosal immune response keeps the systemic immune response largely ignorant of the luminal epitopes and flora, and local immune responses are orchestrated in a tissue-specific manner in a complex interaction between intestinal epithelial cells and immune-effector cells, observations in transfected cell lines or even human peripheral blood mononuclear cells may prove to be of limited relevance.[488;672;744;907;908]

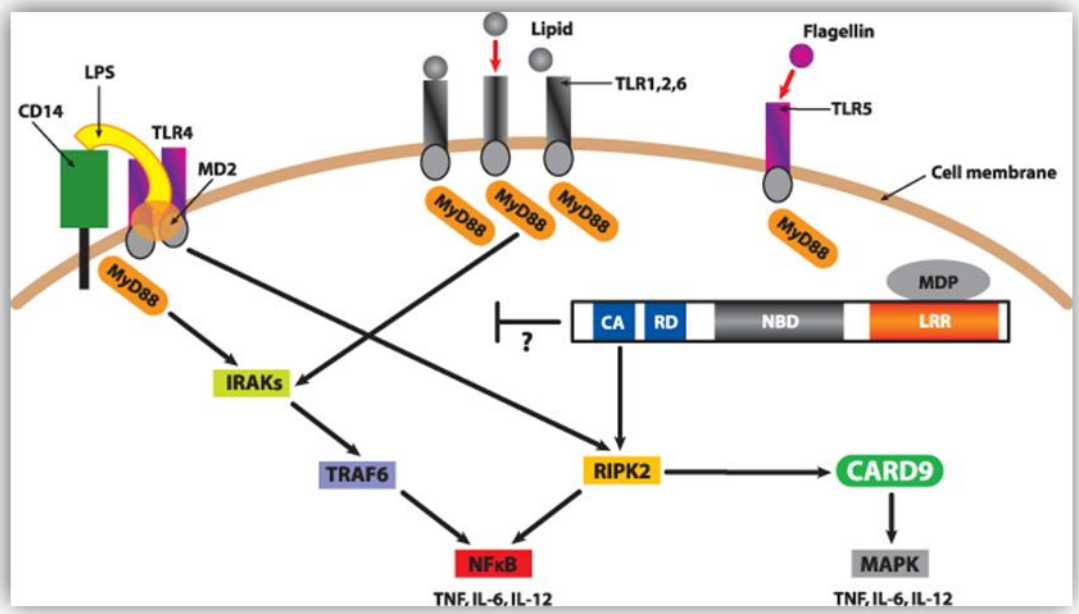



**Figure 99-2: Innate immunity pattern – recognition receptors and signal transduction pathways of relevance to IBD pathogenesis.**

**Figure 9-2: Innate immunity pattern – recognition receptors and signal transduction pathways of relevance to IBD pathogenesis. TLR: Toll-like receptor. LPS (Lipopolysaccharide) triggers signalling through TLR4 after associating with LPS binding protein, CD14 (a glycosylphosphatidylinositol linked protein expressed on the cell surface of phagocytes) and MD2. Except for TLR3, all TLR signalling is MyD88-dependent. Activation of a collection of downstream kinases (IRAKs: IL-1 receptor associated kinases) leads to the degradation of the IKK complex and activation of NF-κB via TRAF-6 (TNF receptor-associated factor 6). NOD2/CARD15 (Nucleotide binding Oligomerisation Domain 2/ Caspase Recruitment Domain-containing protein 15) is an intracellular receptor for Muramyl Dipeptide (MDP). RIPK2 (Receptor-interacting serine/threonine-protein kinase 2) interacts with NOD2/CARD15 through a CARD-CARD interaction. This activation then further initiates various signalling pathways, including activation of NF-κB. CARD9 is located at a crossroads between TLR-NOD2/CARD15 receptors and connects NOD2/CARD15 also to MAPK signalling.**

### *9.1.2 TLR4*

Germline variation of the TLR genes in CD has received much interest.[389] TLR2 (4q31.3), TLR3 (4q35.1), TLR4 (9q33.1) and TLR9 (3p21.3) are all located in regions associated with IBD by genome wide searches.[107] While TLRs 1,2 and 6 have been associated with extensive colonic IBD in a Flemish case-control study, it has been the inherited variation of TLR4 that has been studied most.[303] Two independent meta-analyses have recently confirmed the association of the TLR4 Asp299Gly polymorphism with CD.[395;909] Genetic interaction with specific NOD2/CARD15 haplotypes, strengthening the association of this polymorphism with colonic disease, has been described.[906]

TLR4, located on the cell membrane, binds a very divergent collection of ligands such as LPS, the fusion protein of respiratory syncytial virus, a subunit of bacterial fimbriae, fibronectin, β-defensin and heat-shock proteins.[666] Matzinger et al. have proposed that these receptors have evolved to bind to specific hydrophobic portions of molecules like defensins, uric acid, LPS, peptidoglycan, lipoteichoic acid and many other immunostimulatory microbial products (e.g. flagellae) which are exposed when microorganisms are damaged.[299]

In addition to the interaction between TLR4 signalling and NOD2/CARD15 activation described above, TLR4 signal transduction is highly complex (Figure 10-2).[666] Of the large number of interacting molecules, associations of CD with germline variation of CD14, TIRAP (TIR domain-containing adapter protein) and MAST3 (microtubule-associated serine/threonine-protein kinase gene-3) have been described but are awaiting further replication.[666;909;910] Functional studies into the effects of TLR4 polymorphisms (mainly focussed on single marker analysis rather than the Asp299Gly/Thr399Ile-haplotype, which is more common in populations of European ancestry) have been legion but not conclusive to date.[911] The cytokine phenotype of

the Asp299Gly/Thr399Ile-haplotype does not differ from the wildtype TLR4 whereas the Asp299Gly/wt haplotype (found almost exclusively in Africa) has a stronger pro-inflammatory cytokine profile compared with wildtype TLR4. One hypothesis states that TLR4 Asp299Gly has evolved as a protective allele against malaria, explaining its high prevalence in sub-Saharan Africa, whereas the same allele could have been disadvantageous after migration of modern humans into Eurasia (e.g. because of increased susceptibility to severe bacterial infections).[912]

### *9.1.3 CARD9*

In their candidate-gene analysis of 85 genes involved in innate immune pathways in 1851 IBD patients (1062 CD, 789 UC) and 1936 controls, Zhernakova and colleagues identified the locus on 9q34.3 containing the CARD9 gene (rs10870077) and confirmed replication in the Wellcome Trust Case Control Consortium(WTCCC) GWAS dataset.[901] The implication of CARD9 as a CD susceptibility gene has been particularly interesting in view of its functional properties, linking TLR, NOD2/CARD15 and C-type lectins like Dectin-1 (and thus innate immunity against fungi e.g. *C.Albicans*) with the adaptive immune response, notably T-cell differentiation and the development of Th17-cells (via IL23).[747;913-915] Furthermore, CARD9 is located at a crossroads between TLR-NOD2/CARD15 receptors on the one hand (via RIP2 it connects NOD2/CARD15 to MAPK signalling), and immunoreceptor tyrosine-based activation motif (ITAM)-tyrosine kinase receptors (like Dectin-1) to NFkB-activation, on the other hand (Figure 10-2).[916]

Acosta-Rodriguez et al. demonstrated that *C.Albicans* preferentially induced the differentiation of Th17 cells which expressed CCR6, a mucosa-homing receptor, and CCR4, a skin-homing receptor, as well as the transcription factor RORγt, which is essential for the differentiation of Th17 cells.[917] Germline variation of CCR6 has also been implicated in CD susceptibility in the GWAS meta-analysis (see below).[694]

### *9.1.4 NLRP3*

Another member of the CATERPILLER family of genes, encoding proteins that comprise a nucleotide-binding domain and a leucine-rich repeat domain, NLRP3 (on Chromosome 1q44) encodes cryopyrin.[311;446] Cryopyrin is involved in the inflammasome signalling platform by regulating caspase-1 activity and IL-1β processing.[918]

The importance of cryopyrin in inflammation is highlighted by gain-of-function mutations within its NOD domain that are associated with three hereditary periodic fever syndromes: Muckle-Wells syndrome, familial cold autoinflammatory syndrome and neonatal-onset multisystem inflammatory disease.[410;411;919] Hyperproduction of IL-1β is thought to be a central event leading to symptoms in these three syndromes. Consistent with these observations is the successful use of IL-1β targeted therapy for treating Muckle-Wells syndrome and familial cold autoinflammatory syndrome.[411;919] Contrary to the gain-of-function mutations of these conditions, Villani and colleagues, using a candidate gene approach, uncovered a regulatory region downstream of NLRP3 that contributes to Crohn's disease susceptibility and is associated with hypoproduction of IL-1β and decreased NLRP3 expression. The risk allele of rs6672995, located in a predicted regulatory region, was associated with a decrease in LPS-induced IL-1β production, and the risk allele of rs4353135 was associated with a decrease in baseline NLRP3 expression in two independent sample sets of healthy donors.[918;920]

### *The differentiation of Th17-lymphocytes: IL23R, JAK2, STAT3, CCR6, ICOSLG*

In response to the signalling of the innate immune response, the next level of coordination of defence against specific pathogens and the mediation of different types of tissue-specific inflammation is at the level of T helper (Th) cells which constitute an important arm of the adaptive immune system.[921] Within this family of T helper cells, the recently discovered Th17 subgroup has received intense research interest of late.[618] The antiquated Th1/Th2 dichotomy has now been replaced with a more appropriate set of CD4+ lymphocyte populations distinguished on the basis of Cluster of Differentiation-markers, effector cytokines-profile and transcription factors. The differentiation of naïve CD4+ T cells can take several distinct paths. In Figure 10-3, we have illustrated the differentiation into pro-inflammatory Th1, Th2 and Th17 cells. It is however important to remember that the same naïve CD4+ T cells can differentiate into regulatory and anti-inflammatory T cell populations such as the transcription factor FOXP3-expressing T regulatory cell population (Treg) and the FOXP3-negative T regulatory type 1 (Tr1) population.[921;922] TGF-β (Transforming growth factor β) via SMAD molecules (Small Mothers Against Decapentaplegic), IL27 and ICOS (inducible T cell co-stimulator receptor) signalling are crucial in the regulation of this type of differentiation.[696;697;922] ICOSLG (inducible T cell co-stimulator ligand), a co-stimulatory molecule expressed on intestinal (and other) epithelial cells, may therefore have a role in their antigen presentation to and regulation of mucosal T lymphocytes.[923] Germline variation of ICOSLG has recently been identified as a contributor to the genetic susceptibility for CD in the meta-analysis of genome-wide association studies in CD.[694]

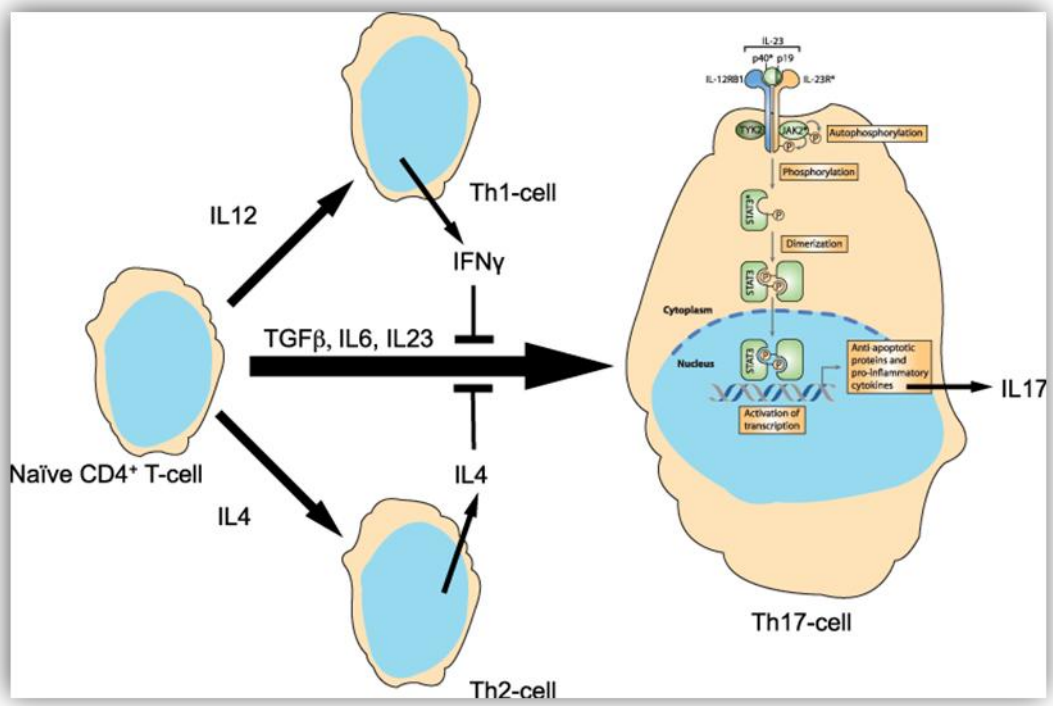



**Figure 9-3: T helper cell differentiation from naïve T cell to Th1, Th2 and Th17 cells.**

**Figure 9-3: T helper cell differentiation from naïve T cell to Th1, Th2 and Th17 cells.** **Th1 differentiate into Th1 cells, under the influence of IL12 and IFNγ. Th2 cells differentiate from naïve CD4+ T cells under the influence of IL4 and IL13. Th17 cells (characterised by the transcription factor ROR- γ-t) differentiate under the influence of TGF-β and IL6. Th17 cells produce not only IL17 (IL17A, IL17F) but also IL22 and IL21. The role attributed to IL23 is not that of active induction of differentiation of naïve T cells, but rather to stabilise the already differentiated Th17 subpopulation.**

Along the pro-inflammatory axis, naïve CD4+ T cells can differentiate into Th1 cells, under the influence of IL12 and interferon-γ (IFN-γ).[921] They typically produce interferon-γ (IFN-γ), regulated by the transcription factor T-bet, and mediate protection against intracellular pathogens (e.g. through the induction of autophagy as illustrated later).[921] Th2 cells differentiate from naïve CD4+ T cells under the

influence of IL4 and IL13. Their effector cytokines are IL4, IL5 and IL13, produced under the influence of the transcription factor, GATA3. Th2 cells mediate clearance of extracellular pathogens and IgE-mediated immune responses and allergy.[860] Th17 cells (characterised by the transcription factor ROR- γ-t) differentiate under the influence of TGF-β and IL6.(see Figure 10-3) Th17 cells produce IL17A, IL17F, IL22 and IL21. The role attributed to IL23 is not that of active induction of differentiation of naïve T cells, but rather to stabilise the already differentiated Th17 subpopulation.[921]

IL23 is a binary complex of a four-helix bundle cytokine (p19) and a soluble class I cytokine receptor p40. IL23 and IL12 share p40 as an alpha-receptor subunit (encoded by the IL12B gene which will be discussed later), yet show only 15% sequence homology between their four-helix cytokines p19 and p35, respectively, and signal through different combinations of shared receptors.[924] IL23 signals through a heterodimer-receptor composed of the IL12R β1 chain (also a portion of the IL12 receptor) and a specific IL23R subunit. Downstream of the IL23 receptor, autophosphorylation of JAK2 followed by homodimerisation of STAT3 are required for the STAT3-homodimer to translocate to the nucleus and exert its effects on gene transcription.(see Figure 10-3) The Th17 transcription factor ROR- γ-t induces the expression of IL23R in response to IL21, which itself depends on IL6-mediated activation of ROR- γ-t and STAT3.[925]

In animal models, the Th17 T cell subset has been shown to mediate chronic and autoimmune inflammatory conditions, with a central role for IL23 in the development of intestinal disease.[741] Increased expression of IL17 and IL23 has been demonstrated in IBD.[750;754] A wide range of microorganisms, including Gram-positive and Gram-negative bacteria, *Mycobacterium tuberculosis* as well as fungi (e.g. *Candida Albicans*) can trigger a Th17 response.[921] The NOD2/CARD15 ligand, MDP, can also promote the production of Th17 cells.[746] A swift influx of Th17 cells at sites of inflammation could indicate a crucial role in connecting the innate and adaptive immune responses. Khader and colleagues have shown that this is also the

case after *Mycobacterium tuberculosis* challenge: the Th17 response is required to bring the Th1 cells into the infected tissue.[926]

The impetus for the detailed study of the IL23 signalling pathway in CD susceptibility came from the discovery of association of germline variation of the IL23R gene with CD in the first published GWAS by the North American IBD Genetics Consortium in 2006.[470] Although initially identified in a cohort of ileal CD, IL23R germline variation has subsequently also been associated with UC.[700-702] Recently, germline variation of both JAK2 and STAT3 has been implicated in the pathogenesis of CD by the meta-analysis of GWAS.[694] It is noteworthy that, before the studies discussed below implicated the IL23 signalling pathway in CD, functional data in the mouse model characterised by ileal inflammation, the Samp1/Yit mouse already demonstrated STAT3 activation via IL6.[695] The first therapeutic trial of a monoclonal antibody, ustekinumab, targeting the common p40 subunit of IL12 and IL23 (encoded by the IL12B gene which will be discussed later), has shown promising results in moderate-to-severe CD.[699]

The North American IBD genetics consortium performed a genome-wide association study testing 308,332 markers in 567 ileal CD patients and 571 controls of non-Jewish European ancestry.[470] The same authors demonstrated replication of the protective effect of the rare IL23R allele (Arg381Gln, frequency 1.9% in non-Jewish ileal CD vs. 7.0% in non-Jewish controls) in an independent case-control association study of Jewish ileal CD patients, as well as in a family-based association analysis (n=833). The first replication studies of this variant came from the UK. The Wellcome Trust Case Control Consortium (WTCCC) replicated the protective effect of the rare IL23R allele in CD (n=1902) and UC (n=975) versus healthy controls (n=1345).[700] In our own cohort of childhood onset IBD, case-control analysis (IBD n=358, controls n=342) and family-based association analysis confirmed these findings.[706] Since these initial studies, the association of CD with this locus has been replicated widely and confirmed recently in the GWAS where the rs11465804 variant (D'=1 and $r^2$=0.88 with the Arg381Gln variant (rs11209026))achieved a

highly significant p value of $10^{-63}$ in the combined cohort, with in the case-control analysis an effect with OR 2.50, second only to the NOD2/CARD15 locus.[694]

The index study by Duerr et al. had already shown that the signal of this locus was not limited to the Arg381Gln variant of the IL23R gene. Our own group and others have since performed haplotype-tagging strategies to confirm this effect of gene-wide variation of IL23R on the genetic susceptibility to CD.[704;705;707] As several variants across the IL23R gene contribute to the CD susceptibility signal, it is difficult to identify their genotype-phenotype relationship and their effect on IL23 receptor function. Similar to the IL23R locus, the loci containing the JAK2 and STAT3 genes are currently the subject of detailed resequencing efforts and functional studies to identify the causal variants.

IL17 producing T-cells have also been detected in the memory population of peripheral blood mononuclear cells, characterised by the combined expression of Chemokine Receptors 4 and 6 (CCR4 and CCR6).[917] As mentioned above, germline varation of CCR6, which encodes a member of the G protein–coupled chemokine receptor family, has been found to predispose to CD in the meta-analysis of GWAS.[694] CCR6 is a homing receptor, expressed by immature dendritic cells and memory T cells and is important for B-cell differentiation and tissue-specific migration of dendritic and T cells during epithelial inflammatory and immunological responses.[927] The ligand of this receptor is CCL20, also known as macrophage inflammatory protein 3 a (MIP-3a). Colonic epithelial cells have been shown to secrete CCL20 in response to IL17, in order to recruit Th17 cells.[928]

As illustrated before for the innate immunity against fungi, linking C-type lectins with CARD9 signalling and the differentiation of Th17 cells, the innate immune system's priming of the secondary immune response is slowly being unravelled. Recently CCL20 dependent signalling (together with β-defensin-3) has been shown as crucial in the development of murine intestinal lymphoid follicles (ILFs) in

response to NOD1/CARD4-mediated recognition of the commensal intestinal flora.[929] In the absence of these ILFs, the composition of the intestinal bacterial community is profoundly altered, demonstrating that intestinal bacterial commensals and the immune system communicate through an innate detection system to generate adaptive lymphoid tissues and maintain intestinal homeostasis.[929]

### *Autophagy: ATG16L1, IRGM, LRRK2*

Several loci containing autophagy genes (ATG16L1, IRGM and LRRK2) achieved genome-wide significance in the CD GWAS meta-analysis.[694] Autophagy denotes several distinct processes involving the delivery of portions of the cytoplasm to the lysosome for degradation (chaperone-mediated autophagy, micro-autophagy and macro-autophagy) with emerging evidence for distinct regulation of selective (directed against micro-organisms or protein aggregates) and non-selective autophagy (induced in response to cellular starvation).[756] Autophagy has been studied in a myriad of homeostatic mechanisms relevant to CD pathogenesis: containment of inflammation by eliminating pathogens and controlling NF-kB signalling, tumour biology, resistance against pathogen-induced cell death and recently also in the biology of the Paneth cell.[832;833] The complex regulation of both selective and non-selective autophagy is illustrated in Figure 10-4.

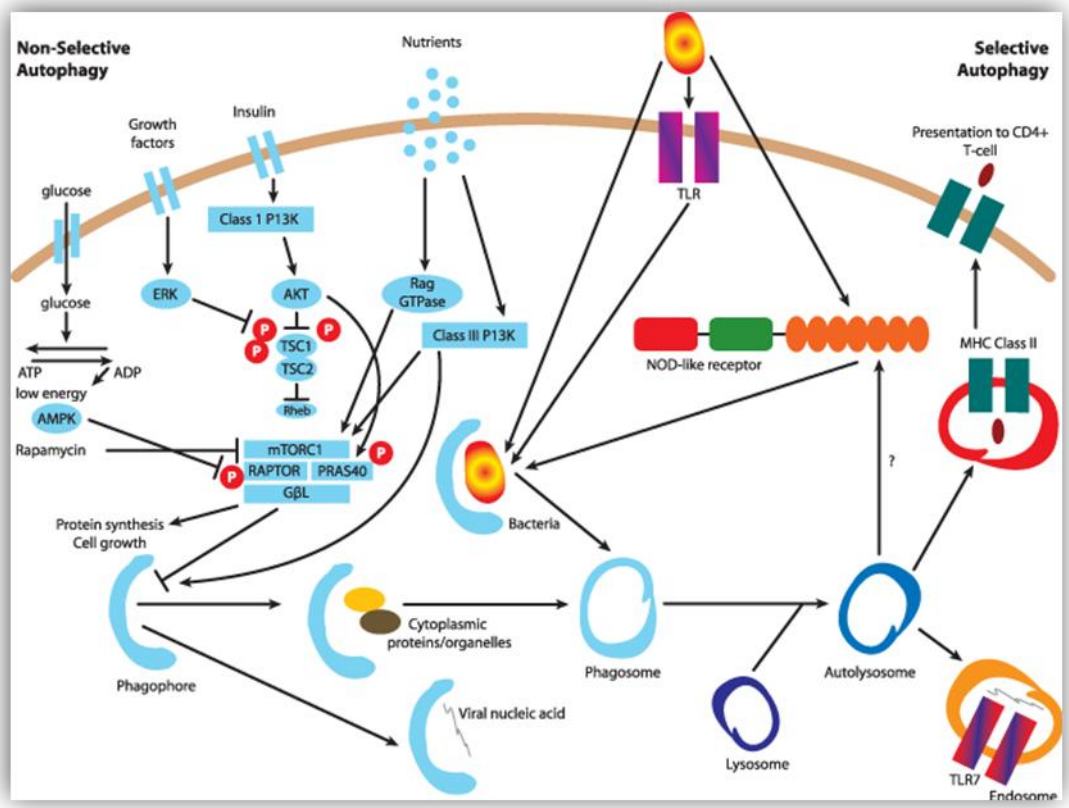



**Figure 9-4: The complex regulation of non-selective and selective autophagy.**

**Figure 9-4: The complex regulation of non-selective and selective autophagy. A wide range of non-selective stimuli (e.g. glucose, insulin, nutritients and growth factors) can trigger non-selective autophagy. Rapamycin is one of the increasing number of available options to modulate the autophagy response. The autophagic response to bacteria, the interaction with pattern-recognition receptors (Toll-like and NOD-like) and the antigen-presentation to the adaptive immune system via MHC Class II molecules are illustrated**

Autophagy can eliminate invading microbes in a highly specific manner in a process termed xenophagy (Figure 10-4), rather than causing the organism to dispose of the infected cell.[761] While autophagy has been implicated in the response against a wide range of invasive pathogens, the autophagic response against *Mycobacterium tuberculosis* is particularly relevant in CD pathogenesis.[930] Singh et al. showed that

the murine Irgm1 protein (Immunity-related p47 guanosine triphosphatase, also called LRG47) is critical in the IFN-γ-induced autophagy response against *M. tuberculosis*.[837]The intracellular survival of *M. tuberculosis* is dependent upon its ability to arrest phagolysosome biogenesis, avoid direct killing mechanisms in macrophages, and block efficient antigen processing and presentation.[931] The *Mycobacterium*-imposed block in phagolysosomal maturation can be overcome by activating cellular autophagy, either through starvation or inhibition of mTOR.[930]

Recent reports by Sanjuan and Xu help explain how pathogens that have gained access to the cytosol are recognised by the autophagic machinery. Sanjuan et al. showed the link between Toll-like Receptor (TLR) signalling in macrophages and Atg5/Atg7-dependent autophagy and phagocytosis.[779] Previously, Xu and colleagues had demonstrated that TLR4-induced autophagy was regulated through a Toll-interleukin-1-receptor-domain-containing-adaptor-inducing-interferon-beta (TRIF)-dependent signalling pathway also involving Receptor-interacting-protein-1 (RIP1) and p38 mitogen-activated-protein-kinase (MAPK), independent of the signal transduction molecule MyD88 (myeloid-differentiation-factor-88).[932]

Interaction between NOD-like receptor signalling and autophagy was demonstrated by Nuñez et al: *Shigella*-induced caspase-1 activation and cell death in macrophages are mediated through Ipaf, a cytosolic pattern-recognition receptor of the NLR-family and the adaptor protein apoptosis-associated speck-like protein containing a C-terminal caspase recruitment domain.[933] How NOD2/CARD15-signalling and autophagy interact is currently the subject of intense investigation.

Autophagy provides another type of protection from bacterial pathogenesis during *Vibrio cholerae* infection.[760] Gut epithelial cell lines resist cell death in response to the haemolytic exotoxin cytolysin of this this noninvasive enteropathogen through the induction of autophagy.[934;935]

Autophagy also plays a role in the priming of the adaptive immune response. Autophagy promotes MHC class II presentation of cytosolic antigens whether endogenous, bacterial, viral or tumour-derived.[936-940] This additional role of autophagy is important in the presentation of self-antigens in the newborn thymus (enabling thymic epithelial cells to present self-antigens to lymphocytes during positive and negative selection) and in antigen presentation by B-cells, dendritic and epithelial cells.[760;941-943]

### *9.1.5 ATG16L1*

Hampe et al. demonstrated association of CD with a coding variant of the Atg16L1 gene (Autophagy-related 16-like 1 gene).[688] Rioux et al. later showed that autophagy induced by *S. Typhimurium* was significantly different in Atg16L1 knockdown conditions.[687] Glas et al. confirmed these findings in CD patients and in murine ileitis.[827] This genetic association has now been replicated in several independent cohorts.[472;687;694;731;732;827;944] Our own group and Fowler et al. have provided evidence that the widely replicated association of CD with ATG16L1 is driven by ileal disease.[945;946] In the study by Hancock and colleagues into the genetic susceptibility to colonic CD, no association with ATG16L1 was demonstrated.[947]

Recent reports by Cadwell, Saitoh and Kuballa constitute a real breakthrough in our understanding of ATG16L1 function.[832] Cadwell and colleagues showed that, in the epithelium of the ileum, ATG16L1 and ATG5 are crucial for Paneth cell biology.[832-834] ATG16L1-deficient Paneth cells exhibit notable abnormalities in the granule exocytosis pathway and increased expression of genes involved in peroxisome proliferator-activated receptor (PPAR) signalling as well as several acute phase reactants and adipocytokines (notably leptin and adiponectin). Crohn's disease patients homozygous for the risk allele, identified by Hampe et al., displayed Paneth cell granule abnormalities like those observed in ATG16L1-deficient mice.[832] Saitoh and colleagues, on the other hand, showed in ATG16L1-deficient mice that

ATG16L1 is required to survive the period of neonatal starvation, as shown before for ATG5 and ATG7, and that these mice displayed an increased severity of dextran sulphate sodium-induced colitis.[833] Loss of ATG16L1 in macrophages caused aberrant LPS-induced IL-1beta and IL-18 production in a caspase-1 dependent manner.[833]

Kuballa and colleagues demonstrated how in human epithelial cells, the CD-associated ATG16L1 coding variant shows impairment in the capture of internalized Salmonella within autophagosomes.[834] Together these studies show a number of autophagic defects in distinct cell types of the intestinal epithelium, thus attributing the increased risk of Crohn's disease to both impaired bacterial handling intracellularly in epithelial cells as well as due to altered exocytosis of Paneth cells.

### *9.1.6 IRGM*

Another autophagy gene, IRGM, was implicated by the WTCCC study.[472] IRGM belongs to the p47 immunity related GTPase family. Its mouse homolog, LRG-47 (encoded by IRGM), critically controls intracellular pathogens by autophagy, and $Irgm^{-/-}$ mice show markedly increased susceptibility to *Toxoplasma gondii* and *Listeria monocytogenes*.[838] The role of IRGM in protecting mature effector CD4+ T-lymphocytes against IFN-γ induced autophagic cell death, has demonstrated a feedback mechanism in the Th1 response that limits the detrimental effect of IFN-γ on effector T-lymphocyte survival while facilitating the antimicrobial functions of IFN-γ.[839]

IRGM expression in macrophages is stimulated in vitro by bacterial LPS, one of the TLR4 ligands, via STAT1 and IFN-β.[948] LPS-stimulated IRGM-deficient macrophages display enhanced phosphorylation of p38, a downstream response associated with TLR4/MyD88 rather than IFN-β /STAT-1 signalling.

Bafica and colleagues therefore suggested that in LPS-stimulated (murine) macrophages, IRGM is induced by IFN-β and negatively regulates TLR4 signalling

to prevent excess proinflammatory cytokine production and shock.[948] The parallel with NOD2/CARD15-mediated downregulation of TLR signalling, as discussed above, is noteworthy. A CD-related alteration in IRGM regulation, due to a common deletion polymorphism, identified recently by McCarroll et al. in the promoter region of IRGM, that affects the efficacy of autophagy or the downregulation of TLR4-signalling, remains the subject of investigation.[840]

Our own data and a paediatric GWAS have not found an association between germline variation of IRGM and paediatric CD.[732;949] This indicates that the contribution of IRGM to the genetic susceptibility is smaller than in adult CD (Odds Ratio <1.40). Larger pediatric cohorts are therefore necessary to confirm or exclude IRGM as a paediatric CD susceptibility gene.

### *9.1.7 LRRK2*

The meta-analysis of GWAS in CD identified the gene LRRK2 (leucine-rich repeat kinase 2).[694] Little is known about the role of LRRK2 in autophagy. The role of LRRK2 mutations in the pathogenesis of CD also remains undefined. Conversely, LRRK2 mutations are the single most common genetic cause of Parkinson disease, another autophagy-mediated condition.[841] Transfection of LRRK2 cDNA containing the common G2019S mutation resulted in significant decreases in neurite length and increased autophagic vacuoles.[842] RNA interference knockdown of key autophagy molecules reversed the effects of G2019S LRRK2 expression on neuronal process length, whereas rapamycin (an inducer of autophagy via inhibition of the mTOR regulatory molecule) potentiated these effects.

## *Maintenance of epithelial barrier integrity: IBD5, DLG5, PTGER4, ORMDL3, ITLN1, DMBT1 and XBP1*

### *9.1.8 IBD5*

Mutations within the organic cation transporter genes (OCTN1&2), located in the IBD5 locus on Chromosome 5q31, were first identified by Peltekova et al.[281] These authors identified mutations within OCTN1 (1672C/T) and OCTN2 (207G/C)) and found an association of the TC-haplotype and increased susceptibility to CD in the Canadian population.[281] A large number of positive replication studies and identification in the meta-analysis of GWAS, have confirmed IBD5 as a CD susceptibility locus.[694;950] Detailed analysis in the Canadian and other populations, by inclusion of other polymorphisms on the IBD5 locus, has demonstrated that several other genes within this cytokine gene cluster are as likely to be the causative gene.[269;285] However, Barrett et al. showed, using eQTL (expression quantitative trait locus) analysis, an approach which seeks to correlate genotype and expression patterns and takes into account that such functional relationships need not respect the specific boundaries of LD around the association, that CD-associated SNPs were associated with decreased SLC22A5 mRNA expression.[694] Several studies in adult and childhood onset CD, have suggested the IBD5 risk haplotype is associated with a more severe disease phenotype.[278;951]

### *9.1.9 DLG5*

DLG5 (Drosophila Discs Large Homolog 5) gene, a member of the MAGUK (Membrane Associated Guanylate Kinase) family located on the pericentromeric IBD locus on Chromosome 10, was first identified as a CD suscepitiblity gene by Stoll et al.[244] MAGUK proteins are known to form scaffolds for other proteins involved in intracellular signal transduction and could therefore interfere with the integrity of the epithelial barrier.[950] Since the index publication, DLG5 variation in

CD has been studied in a large number of populations but replication has only been demonstrated in a few.[950]

It is noteworthy that DLG5 was not confirmed as a susceptibility gene in the meta-analysis.[694] True genetic heterogeneity, phenotypic differences between patient populations and/or stratification of control groups remain possible explanations but a meta-analysis of most published studies suggests it does not have a major role to play in IBD susceptibility.[265] A more recent meta-analysis of DLG5 variation showed a gender-stratified effect: the 30Q allele (rs 1248696) was found to be associated with a mildly decreased risk of CD in women. This gender-specific effect was confirmed in a paediatric CD cohort by Biank et al.[263;952]

### *9.1.10 PTGER4, ORMDL3 and the Gene Deserts*

A Belgian-French collaboration first identified a novel 'gene desert' on Chromosome 5p13.1 associated with CD.[471] Although there have been no known genes or CpG islands described within this region, eQTL analysis showed that expression of one of the flanking genes PTGER4 (prostaglandin receptor EP4) was regulated by CD associated polymorphisms.[471] The meta-analysis by Barrett et al. later confirmed the association signal on this locus.[694]

It is noteworthy that this locus and the PTGER4 gene have previously also been associated with asthma susceptibility.[953] Genetic overlap between susceptibility determinants in a number of immune-mediated diseases, notably different manifestations of allergic disease and IBD, have come to light in the course of the application of GWAS to the study of these complex polygenic diseases.[890;954] A GWAS in childhood asthma, followed by several replication studies, has added the locus on chromosome 17q21 containing the ORMDL3 gene to the long list of shared genetic susceptibility determinants involved in both IBD and atopic disease (including among others the HLA-region, several interleukins (e.g. IL12B on the

IBD5 locus), prostaglandin-receptors, TGF-β, defensins, Toll-like receptors, NOD-like receptors, STAT3 etc).[694;863;864;955]

In the meta-analysis by Barrett et al., the most significant CD-associated eQTL reported affected the ORMDL3 gene on chromosome 17. This suggests that the same polymorphisms might underlie susceptibility to both CD and asthma, possibly by perturbing ORMDL3 expression.[694] Germline variation of ORMDL3, which encodes a protein of unknown function belonging to a family of transmembrane proteins anchored in the endoplasmic reticulum, was recently shown by Bouzigon and colleagues to confer increased risk of asthma, with a notable restriction to early-onset asthma.[956] This risk was further increased by early-life exposure to environmental tobacco smoke.[956] These findings are particularly intruiging when one considers the detrimental effect of smoking on CD susceptibility (notably age at diagnosis) and its influence on CD phenotype.[556;582]

At least 4 other gene deserts have satisfied stringent criteria of genome-wide significance in the meta-analysis.[694] Detailed analysis of regions flanking these gene deserts is now ongoing to subject these intriguing deserts to eQTL as applied successfully in the case of the 5p13 locus. Whether these deserts contain long-range regulatory elements, remains under investigation.

### *9.1.11 ITLN1*

Intelectin (ITLN1) is a mammalian Ca2+-dependent, D-galactosyl-specific lectin expressed in Paneth and goblet cells of the small intestine and proposed to serve a protective role in the innate immune response.[957;958] Structurally identical to the intestinal lactoferrin receptor, ITLN1 has been demonstrated in secretory granules of lysozyme-positive Paneth cells in the bottom of the crypts as well as goblet cells along the crypt-villus axis, but quantitatively, the major site of intelectin deposition was the enterocyte brush border.[958] This strategic localization suggests that the

intelectin serves as an organizer and stabilizer of the brush border membrane, preventing loss of digestive enzymes to the gut lumen and protecting the glycolipid microdomains from pathogens.[958] The first genetic association with CD was recently demonstrated in the meta-analysis by Barrett et al.[694] Detailed functional analysis of this locus is eagerly awaited.

### *9.1.12 DMBT1*

The heavily sulfated membrane glycoprotein mucin-like glycoprotein (Muclin) is a DMBT1 (Deleted in Malignant Brain Tumours 1) product that is strongly expressed in organs of the gastrointestinal (GI) system.[959] A deletion allele of DMBT1 on chromosome 10q with a reduced number of scavenger receptor cysteine-rich domain coding exons was found by Renner et al. to be associated with an increased risk of CD, but not UC.[960] In this study, Dmbt1(-/-) mice displayed enhanced susceptibility to dextran sulfate sodium-induced colitis and elevated TNFα, IL6, and NOD2/CARD15 expression levels during inflammation. More recently, de Lisle et al. demonstrated that DMBT1 deficiency was associated with an impairment of exocrine pancreas whereas no significant difference in DSS-induced colitis-severity was observed between Muclin-deficient mice and wildtype.[959]

DMBT1 is strongly up-regulated in the inflamed intestinal mucosa of Crohn's disease patients with wild-type, but not with mutant NOD2/CARD15.[961] DMBT1 inhibits cytoinvasion of *Salmonella enterica* and LPS- and muramyl dipeptide-induced NF-kappaB activation and cytokine secretion in vitro. Dysregulated intestinal DMBT1 expression, due to mutations in the NOD2/CARD15 gene, may contribute to the complex pathophysiology of barrier dysfunction in CD.[961]

### *9.1.13 XBP1 and Endoplasmic Reticulum stress*

When cells are stimulated to secrete large amounts of protein, an excess of unfolded proteins accumulates in the endoplasmic reticulum (ER), triggering ER stress and activating three proteins (IRE1, PERK, and ATF6) that mediate the unfolded protein response, with a key role for the transcription factor XBP1 (X-box binding protein 1).[962] These signalling pathways expand the protein folding capacity of the cell in order to restore ER homeostasis. If the ER damage is extensive or prolonged, cells typically undergo programmed cell death. The loss of key proteins of the unfolded protein response also results in so-called "unresolved ER stress" and apoptosis. For example, in the absence of XBP1 (*Xbp1*−/−), unresolved ER stress leads to increased expression of select genes of the unfolded protein response (ATF4 and ATF6) and increased JNK signalling. These elevated signals increase cell susceptibility to programmed cell death and induce increased expression of proinflammatory genes leading to intestinal inflammation.

Proof of principle has been provided in animal models (e.g. Mucin2 mutant mice and $IL10^{-/-}$ mice) with replication in intestinal tissue from IBD patients.[963;964] Not only is XBP1 deletion in IECs associated with spontaneous enteritis and increased susceptibility to induced colitis (secondary to both Paneth cell dysfunction and an epithelium that is overly reactive to mediators of inflammation in IBD such as flagellin and TNFα, also genetic association of XBP1 tagging variants with both CD and UC has been described by Kaser et al.[965] In a total of 5322 controls, 2762 CD and 1627 UC patients, these authors identified a total of six SNPs associated with IBD that were robust after correction for multiple testing and significant after 10,000 permutations. Multiple logistic regression analysis of the entire IBD panel including gender as a covariate revealed a best model fit with SNPs rs5997391 and rs35873774 (intron 4/5 of XBP1). Deep sequencing of the XBP1 locus, characterised by weak linkage disequilibrium, in 282 CD, 282 UC and 282 healthy controls revealed a number of rare alleles of which two non-synonymous variants were only present in IBD patients and hypomorphic in detailed functional experiments.[965] Taken together,

these studies provide functional and genetic evidence for the role of the ER stress response in the pathogenesis of IBD.

## *Orchestration of the secondary immune response: HLA-region, TNFSF15/TL1A, IRF5, PTPN2, PTPN22, NKX2-3, IL12B, IL18RAP, MST1*

### *9.1.14 The HLA-region*

The region on chromosome 6p (IBD3) containing the Human Leukocyte Antigen (HLA) – genes as well as a wide range of functional IBD candidate genes like TNFα, met the strong criteria for genome-wide significance in the meta-analysis by van Heel et al. of nearly 2000 affected sibling pairs and several SNPs achieved nominal significance in the GWAS meta-analysis by Barrett et al.[107;694] The extensive linkage disequilibrium across the gene-dense region has complicated the identification of causal variants.[966] In IBD, the HLA-DRB1 is the most studied region with the strongest associations observed between HLA-DRB1*0103 and colonic CD and even more so with severe, extensive UC in Caucasians.[702] HLA-DRB1*07 is the most consistently replicated association between the MHC and CD, and more specifically with ileal disease.[967]

Fernando and colleagues recently performed an extensive pooled analysis and confirmed significant association signals arising from alleles/haplotypes related to HLA-DRB1*0103, HLA-DRB1*04, HLA-DR7, and HLA-DRB3*0301. In addition, these authors described association with HLA-B18, HLA-B21, HLA-DR6 (encompassing HLA-DRB1*1401), HLA-DR8 (including HLA-DRB1*0802 and *0803) and HLA-DR10.[967]

Current treatment options for CD include several monoclonal antibodies (e.g. infliximab, adalimumab and certolizumab) against TNFα. It is therefore noteworthy that Fernando et al substantiate reports of association of several of the TNF promoter polymorphism (TNF-1031C, TNF-863A and TNF-857T with CD), as our group had previously reported.[254;967-969] As the TNFα gene lies at only 1Mb from the CD-associated HLA class II alleles, it is unclear whether these effects are independent due to the complex linkage disequilibrium of the HLA region.

### *9.1.15 TNFSF15*

The association between CD and TNFSF15 (Tumour Necrosis Factor Superfamily, member 15 on 9q33.1 also called TNF superfamily ligand A, TL1A or vascular endothelial cell growth inhibitor, VEGI) was described in the first GWAS in CD and has since been widely replicated, including in the recent GWAS meta-analysis.[694;900;970-973] TNFSF15/TL1A is a TNF-like factor expressed in endothelial cells and dendritic cells with evidence for upregulation of mRNA and protein levels in macrophages and $CD4^+$/$CD8^+$ lymphocytes of the intestinal lamina propria of IBD patients, especially CD. TNFSF15/TL1A synergizes with both the cytokine-dependent IL12/IL18 pathway and with low-dose stimulation of the T cell receptor to significantly induce the secretion of IFN-gamma via an IL18-independent pathway and is involved in the development of chronic mucosal inflammation by enhancing Th1 and Th17 cell functions.[974;975] Pappu et al have recently shown that the expression of a TNF receptor family member, death receptor 3 (DR3; also known as TNFRSF25), is selectively elevated in Th17 cells, and that its cognate ligand, TNFSF15/TL1A, can promote the proliferation of effector Th17 cells.[976] Another member of this receptor family TNFRSF6B (tumor necrosis factor receptor superfamily member 6B) has recently been associated with childhood-onset CD and will be discussed later.[732]

### *9.1.16 IRF5*

The interferon regulatory factor 5 (IRF5) gene encodes a transcription factor that plays an important role in the innate as well as in the cell-mediated immune responses through the downstream regulation of the TLR-MyD88 signalling pathway for gene induction of proinflammatory cytokines, such as IL6, IL12 and TNFα.[355;977] The IRF5 gene has been shown to be associated with several other immune-mediated conditions including systemic lupus erythematosus, rheumatoid arthritis, multiple sclerosis and Sjögren syndrome. An association between CD and a 5 bp indel (CGGGG) polymorphism in the promoter region of the IRF5 gene was recenty described by Dideberg et al.[977] The insertion of one CGGGG unit is predicted to create an additional binding site for the transcription factor SP1.[977] To date, no further replication studies have been published.

### *9.1.17 PTPN2 and PTPN22*

Protein tyrosine phosphatases (PTPs) are important regulators of many cellular functions and a growing number of PTPs have been implicated in human disease conditions, such as developmental defects, neoplastic disorders, immunodeficiency and autoimmunity.[978] PTPN2 and PTPN22 (protein tyrosine phosphatase, nonreceptor types 2 and 22) have both been associated in the recent meta-analysis of GWAS.[694] Interestingly, the 602W allele of PTPN22 is protective against CD, but confers increased risk for Type 1 Diabetes and rheumatoid arthritis.[694]

### *9.1.18 NKX2-3*

Germline variants of the homeodomain-containing transcription factor NKX 2-3 (NK transcription factor related, locus 3) have been implicated in the genetic

susceptibility to both CD and UC.[694;702;727;979] Mice deficient in the homeobox gene Nkx2-3, expressed in developing visceral mesoderm, showed a complex intestinal malabsorption phenotype, striking abnormalities of gut-associated lymphoid tissue and spleen suggestive of deranged leukocyte homing and down-regulation of mucosal addressin cell adhesion molecule-1 (MAdCAM-1) in endothelial cells in which Nkx2-3 is normally expressed.[980;981] MAdCAM-1 is a member of the immunoglobulin superfamily, acting as an endothelial cell ligand for leukocyte homing receptors L-selectin and α4β7-integrin. The humanised monoclonal antibody directed against α4-integrin, natalizumab, has been shown to be effective in moderate to severe CD in adolescents and adults, but concerns about its safety profile, notably in combination with other immunosuppressive treatment in view of case reports on the reactivation of JC-virus and the development of progressive multifocal leukoencephalopathy, have so far limited its use.[982-984]

### *9.1.19 IL12B*

The IL12B gene encodes the p40 subunit shared by the cytokines IL12 and IL23. IL12 is composed of two subunits p40and p35, resulting in a p70 bioactive cytokine.[985] Recent data suggest that the IL12 p40 is also biologically active, notably via the induction of lymphotoxin-α in a IL12 receptor β1 (IL12 β1) dependent manner.[985] The IL12 receptor β1 subunit is also a component of the heterodimer IL23 Receptor. In contrast with the pro-inflammatory signal through IL12R β1, is the finding that signalling via IL12Rβ2 regulates both the number and functional maturity of Treg cells, thus illustrating the complex regulation of autoimmune diseases by the IL-12 pathway.[986]

The identification of IL12B as a susceptibility gene, shared by CD and UC, has added further significance to dedicating research interest to the elucidation of the complex IL12/IL23 pathways.[694;702;727] In addition to the role of IL23 signalling in the differentiation of Th17 cells (as discussed above), IL12 signalling is crucial in the

orchestration of the immune response. Recently, Kang and colleagues demonstrated that the production of IFN-γ (which is elevated in CD mucosa and correlates with IL23p19 mRNA expression) by clustered NK cells was dependent on IL12 and IL18, and critical for the activation and maturation of colocalized monocytes to TNF- and inducible nitric oxide synthase-producing dendritic cells.[751;987] Therapeutic targeting of the IL12 pathway (e.g. using the monoclonal antibody ustekinumab against the p40 subunit) is receiving renewed research interest through the identification of this pathway in the genetic susceptibility to CD.[699]

### *9.1.20 IL18RAP*

The IL18RAP (IL18 receptor accessory protein) on chromosome 2q12.1 has been identified as a CD susceptibility gene in a candidate gene study by Zhernakova et al and as a nominally replicating CD risk locus in the meta-analysis by Barrett et al.[694;901] IL18 signalling requires IL18receptor α, which binds specifically to the IL18 and contains sequence homology to IL1 receptor and TLRs, and the IL18RAP.[988] As is the case for many genes discussed above, it is pertinent that the IL18RAP locus (residing in a block of strong linkage disequilibrium which also contains IL1RL1, IL18R1 and SLC9A4) has also been implicated in a number of other immune-mediated conditions, notably coeliac disease.[989] Hunt et al. showed a large *cis* effect of rs917997 genotypes (the same SNP was identified in the studies by Zhernakova, Barrett and Hunt) on the level of IL18RAP mRNA expression in whole blood from treated individuals with celiac disease with a dosage effect on expression.[989]

IL18, which is produced by antigen presenting cells, unstimulated T-cells and NK-cells, is of interest in CD pathogenesis for several reasons.[990] Th1 cells lacking IL18RAP fail to produce IFNγ in response to IL18, IL18RAP-deficient neutrophils fail to respond to IL18-induced activation and cytokine production and IL18RAP is required for NK-mediated cytotoxicity induced by in vivo IL18 stimulation.[988] IL18 expression is increased in the mucosa of CD patients compared with controls and in

areas of macroscopic CD compared with normal mucosa in CD.[901;990;991] Lastly, blockade of IL18 has been demonstrated to attenuate inflammation in murine models of colitis.[992;993]

### *9.1.21 MST1*

The MST1 gene encodes macrophage-stimulating protein (MSP), a protein regulating the innate immune responses to several bacterial ligands.[994] Goyette and colleagues first identified a non-synonymous coding variant (rs3197999, R689C) in the macrophage-stimulating 1 (MST1) gene that accounts for the association signal on the chromosome 3p21-22 locus, and showed association with both CD and UC.[472;704;994] This finding has since been replicated in CD and UC.[694;702] MSP and its receptor MST1R have been reported to be involved in macrophage chemotaxis and activation, as well as inhibition of inflammatory responses (e.g. IL12p40 production) in response to LPS and other pro-inflammatory signals.[994] While the association results and the identification of a coding variant suggest a potential role for MST1 in susceptibility to both CD and UC, other germline variants on this gene dense locus could also have an impact on disease susceptibility.[994] Indeed, the MST1R and TRAIP (TRAF interacting protein, involved in TNFα signalling), which also reside on the 3p21 locus, have also recently been implicated in the genetic susceptibility to CD and UC.[995]

## *Germline variants involved in susceptibility to childhood-onset CD*

Paediatric DNA repositories have long suffered from difficulties in recruiting enough children and adolescents with relatively uncommon conditions to allow for adequately powered genetic association studies. Even though up to a quarter of IBD patients present during childhood, they have been relatively underrepresented in the association studies to date. Our own group and others have shown that many variants

identified to date in adult-onset CD, also contribute to childhood-onset CD.[278;660;706;732]

Kugathasan and colleagues published preliminary results of a subset of patients of the International Paediatric IBD Genetics Consortium. These authors carried out a GWAS in a cohort of 1,011 individuals with pediatric-onset IBD and 4,250 matched controls and identified previously unreported loci on chromosomes 20q13 and 21q22 located close to the TNFRSF6B and PSMG1 genes, respectively.[732] These loci contributed to both CD and UC susceptibility.

Given the fact that that all the significantly associated SNPs occurred in noncoding regions, Kugathasan and colleagues considered the TNFRSF6B gene the most compelling candidate. mRNA expression of TNFRSF6B was shown to be markedly different in colonic biopsies obtained from individuals with IBD compared to disease-free controls and TNFRSF6B mRNA expression correlated with the degree of mucosal inflammation within the colon.[732]

The protein product for TNFRSF6B acts as a decoy receptor (DCR3) in preventing FasL-induced cell death which is pertinent in view of the resistance to FasL-dependent apoptosis which has previously been shown for T lymphocytes in CD.[996] Serum DCR3 concentration differed between individuals with IBD and controls and, within the IBD group, between those with and without the identified at-risk variants captured by the TNFRSF6B tagging SNPs.[732] Funke et al. have recently demonstrated increased expression of DCR3 in ileal biopsies of CD patients both at actively inflamed and at non-active sites and elevation of DCR3 serum levels in patients with active and non-active CD as compared to healthy controls.[997] Furthermore, these authors showed that expression of DCR3 in intestinal epithelial cells was induced by TNFα and that increased DCR3 expression was associated with activation of NF-kappaB and consequent protection of intestinal epithelial cells and lamina propria T cells from CD95L-induced apoptosis.[997] They concluded therefore

that DCR3 may promote inflammation in CD by inhibiting CD95L-induced apoptosis of epithelial and immune cells as well as by inducing NF-kappaB activation.[997]

The international pediatric IBD genetics consortium, to which we have contributed 374 cases (after quality control), will report on its findings in the near future. The analysis involved a total of 3426 patients and 11963 genetically matched controls, ascertained through international collaborations in Europe and North America. Several new loci will be reported including the discovery of 16p11 (near *IL27*) and four additional regions with susceptibility to pediatric onset IBD: 22q12, 10q22, 2q37 and 19p13.11. In addition, the association of 21 of 32 previously confirmed adult-onset CD loci and 8 of 17 adult-onset UC loci with susceptibility to childhood-onset disease is described, highlighting the close pathogenetic relationship with adult-onset IBD.(Imielinski et al., Nature Genetics in press)

### *Concluding remarks, reflection on my PhD and future directions*

From the early epidemiological observations of concordance in affected twins to the multi-million dollar efforts associated with GWAS, three decades of intensive research into the genetic susceptibility to CD have led to an exponentially increasing body of genetic knowledge. Somewhat disappointly, Barrett et al estimated in the meta-analysis of GWAS in CD that 20% of the genetic risk (about 10% of the overall risk) has been explained by the germline variants identified to date. This merely illustrates that in spite of recent progress, there is a lot of work ahead. The large number of confirmed susceptibility loci now requires detailed sequencing to ascertain the contribution of rare variants, which may not have been tagged well with the SNP set available on the different genotyping platforms. Due to its complex polygenic nature, unravelling of the pathogenesis of CD will require elucidation of gene-gene interactions using novel polygenic models rather than assessments of

pairwise interactions applied to date, which did not provide convincing evidence for interaction.

A large number of CD susceptibility loci have been identified through the combined application of linkage analysis, fine mapping studies, candidate gene studies and most recently GWAS. Genetic association analysis has implicated several key mediators of the innate immune system, notably innate pattern-recognition receptors and their associated signal transductors, leading to a dysregulation of the mucosal immune respone against luminal flora. Disruption of epithelial barrier integrity has been implicated in CD pathogenesis through a wide range of mechanisms encompassing structural, metabolic and innate immune pathways. Several autophagy genes have been identified through GWAS triggering a number of recent functional studies into the role of autophagy in CD susceptibility. The discovery of the role of the IL23R gene in CD susceptibility has added further support to the large body of functional studies describing the central role of the Th17 cell in chronic immune-mediated conditions like CD.

The investigations performed in the course of my PhD have contributed in a number of ways to the different aspects of paediatric IBD research. The key focus of my PhD has been the use of the existing cohort of Scottish paediatric IBD patients for the study of genetic loci, implicated in IBD. Thus, I was able to study TNF-α, RICK/RIPK2, NOD1/CARD4, IL23R, ATG16L1, IRGM and FLG variants. From my first work on TNF- α promoter variants, it became clear that our paediatric sample size (even though at the start of my PhD, it was one of the largest repositories of paediatric IBD samples in the world) would not allow for small effect sizes to be picked up. Although later meta-analysis corroborated my findings with regards to TNF- α promoter variants, I chose not to include this as the analysis would not hold up to stringent correction for multiple testing. I was careful during later investigations to substantiate my genetic statistical findings with adequate power calculations. With the arrival of the GWAS, it became clear that even a large repository as the one we had collected over the preceding 4 years, would not allow us

to embark on a genome-wide investigation of genetic susceptibility to IBD in Scotland. Initially I focussed on studying the contribution of IL23R, ATG16L1 and IRGM in childhood-onset IBD. I was hoping to expand my work into the study of the epigenetic regulation of susceptibility loci. To this extent, I optimised the bisulphite-sequencing protocol for the study of the CpG-island in the promoter region of NOD2/CARD15. My preliminary investigations were unable to show a difference in methylation status between CD patients and controls. I realised that practical difficulties with sequencing of this bisulphite-treated DNA and the limited scope of my investigation (looking only at about 35 promoter CpG sites), would not allow me to make firm conclusions about the role of epigenetics in IBD susceptibility.

In the final months of my PhD, I helped in the preparation of samples for the international pediatric IBD genetics consortium, which will report its findings in the autumn of 2009. This collaboration made me appreciate some of the strengths of our repository even more, particularly with regards to our phenotypic data.

My predecessor, Dr Richard Russell, had set up a comprehensive database including phenotypic and epidemiological data of paediatric IBD patients recruited for the genetic studies in Scotland, in collaboration with Mrs Hazel Drummond. Building on this work, I was able to increase the number of cases available for phenotypic and epidemiological studies to well over 400. Rigourous phenotyping at regular intervals allowed us to chart the natural evolution of paediatric IBD and to compare it with adult-onset IBD. This work was published in the journal Gastroenterology and contributed to the discussions, at the recent Pediatric IBD meeting in Paris (Sept 2009), about the development of a paediatric IBD classification system.

My PhD has given me to opportunity to take my first steps as a clinician scientist. I have really enjoyed dedicating myself to research linked with the care for my paediatric patients. I successfully obtained a clinical and research fellowship in paediatric gastroenterology at the Sick Kids hospital in Toronto where I plan to dedicate myself to the study of innate immune immunology.

Suggestions for the future research agenda:

- The large number of loci identified in the genetic susceptibility to CD to date will require detailed deep sequencing to identify (rare) causal variants.

- Tissue-specific functional studies will be required to appreciate the complex regulation of many signalling cascades identified in CD pathogenesis.

- The study of the altered intestinal flora in CD, notably the mucosa-associated microbiome, will require integration of immunology and genetics to appreciate the complex interaction with the host mucosal immune response in CD patients.

- Genetic susceptibility analysis of childhood-onset CD has identified additional loci, offering the real possibility of genetic modelling. This could allow the prediction of the course of disease and the stratification of young patients according to their genetic profile (and possibly microbiome) who would benefit from more aggressive treatment options.

Whether these ongoing genetic efforts result in additional benefit for our patients, will depend on the success of integration of this progress in genetics with comparable paradigm shifts in the field of (mucosal) immunology (notably the regulation of T helper cell populations and autophagy) and the study of the microbiome, particularly the mucosa-associated flora. Eventually, the insights gained from these genetic and functional studies should help identify key environmental triggers and the establishment of a comprehensive gene-environmental model.

# 10 Appendices

*Appendix 1: Lennard-Jones Criteria-check list*

**Study Number** ☐

| | | |
|---|---|---|
| **Chronic granulomatous lip** | **1** | **Yes / No / Unknown** |
| **Pyloroduodenal disease** | | **Yes / No / Unknown** |
| **Small bowel disease** | | **Yes / No / Unknown** |
| **Chronic anal lesion** | | **Yes / No / Unknown** |
| **…….** | | |
| **Discontinuous lesions** | **2** | **Yes / No** |
| **…….** | | |
| **Fissuring ulcers** | **3** | **Yes / No** |
| **Abscess** | | **Yes / No** |
| **Fistula** | | **Yes / No** |
| **…….** | | |
| **Strictures** | **4** | **Yes / No** |
| **…….** | | |
| **Aphthoid ulcers** | **5** | **Yes / No / Unknown** |
| **Lymphoid aggregates** | | **Yes / No / Unknown** |
| **…….** | | |
| **Colonic mucin retention** | **6** | **Yes / No / Unknown** |
| **…….** | | |
| **Granulomata** | **7** | **0 / 1-2 / 2-5 /6-10 / > 10** |
| **……..** | | |

**Number of section criteria met** ☐

**Total number of criteria met** ☐

## *Appendix 2: Patient information leaflet*

**Patient Information Leaflet**

**The genetic influence in the development of Crohn's disease and colitis in children**

*We invite you to participate in research which we think may be important. The information which follows tells you about it. Try to make sure you understand what will happen to you if you decide to take part. Whether or not you do take part is entirely your choice. Please discuss this with your family and ask any questions you want to about the research and we will try our best to answer them.*

We would like to understand more about why people, specially children, develop Crohn's disease and colitis. These diseases are becoming more common in children, specially teenagers, and we want to try and work out why this is. We know that having a relative (either a parent or brother or sister) with Crohn's disease or colitis, gives you a higher chance of getting the disease. We want to examine the things that are passed onto you by your parents (your genes) and how these may result in the development of Crohn's disease or colitis.

You are going to undergo tests on your blood, having been seen by one of the doctors in the clinic.

### *What we are asking*

We would like to ask your permission to collect some extra specimens (two to four teaspoons of blood) which we can use for research in our laboratories. We are also asking both of your parents for blood samples too.

***What are we going to do with them?***

We are going to use these specimens to help research into your genes and the part they play in the development of Crohn's disease and colitis. Some of these specimens will be tested in the laboratory soon after being taken while others may be stored in the freezer before testing. Any specimens stored will only be used for studies which examine your genes and the development of Crohn's disease and colitis.

All our research is intended to increase understanding of how and why these diseases occur, how treatments work and how we may be able to improve treatments in the future.

We also plan to study how genetic make-up may lead to children getting Crohn's disease or colitis and whether this make-up can alter how effective the treatments are.

***How will this affect me?***

We will always collect specimens needed for your planned tests before any research specimens. If there are any problems during the collection of these specimens ie in collecting blood then the doctors will not collect specimens for research.

Taking extra blood will not cause you any problems. The research itself is not relevant to your care. If in the future we would like to perform research that may be directly relevant to you we will contact you to see if you agree.

***Do the specimens have my name on them?***

The people in the laboratory will not know your name as your specimens will have a number. Only Dr Van Limbergen or ward or outpatient staff supervised by him will be able to know this and link it to specimens collected.

***What if I don't want to be included or change my mind?***

You are entirely free to decide not to participate or drop out at any time and this will not affect your care in any way. If a specimen has been collected and you change your mind it will be destroyed at your request.

Version 2 16/01/06

## *Appendix 3: Parent information sheet*

**PARENT/CARER'S INVITATION TO PARTICIPATE IN A RESEARCH STUDY:**

**THE GENETIC INFLUENCE IN THE DEVELOPMENT OF CROHN'S DISEASE AND COLITIS IN CHILDREN**

**Dept of Paediatric Gastroenterology, Ninewells Hospital, Dundee**

**Dept of Paediatric Gastroenterology, Royal Hospital for Sick Children Edinburgh**

**Gastrointestinal Unit, Western General Hospital, Edinburgh**

*We invite your child to participate in research which we think may be important. The information which follows tells you about it. It is important that you understand what is in this leaflet. It says what will happen if you take part and what the risks might be. Try to make sure you know what will happen to you if you decide to take part. Whether or not you do take part is entirely your choice. Please ask any questions you want to about the research and we will try our best to answer them.*

The Department of Paediatric Gastroenterology at the Royal Hospital for Sick Children together with the Department of Gastroenterology at the Western General Hospital and other paediatric gastroenterology units in Scotland have a major role in leading research into diseases of the bowel in children. Diseases which we see often in the clinics and have on-going research interest include inflammatory bowel disease namely Crohn's disease and ulcerative colitis.

Your child is due to undergo tests involving collection of a blood sample as part of their on going care.

### What we are asking

We would like your permission to collect extra blood (up to 20 ml = 4 teaspoons) which we will use for ongoing projects within our laboratories. Some of these specimens may be stored in a freezer for future projects. If specimens are stored they will only be used for studies relating to genetics and Crohns disease and ulcerative colitis. We will also collect and store important and relevant clinical details (e.g. age, sex, medication, disease, findings from the test) for analysis with the specimens.

## How will this affect my child ?

Priority will always be given to collecting specimens needed for our planned tests. If there are any difficulties during the collection of these specimens e.g. in collecting blood then the doctors will not collect specimens for research. Taking extra blood will not mean any significant extra risk or harm to your child.

## What research is being done ?

Broadly speaking we are planning to study how the genetic make-up of children with inflammatory bowel disease (Crohn's disease or Colitis) may lead to them getting the disease.

The research included within this application will not have any impact on the care of your child (including any direct benefit or harm). All our research is aimed to increase understanding of how and why diseases occur, how treatments work and how we may be able to improve treatments in the future. If any studies may affect the care of individual children then a separate consent form will be produced.

## Can I help too ?

Yes you can by giving blood too. We would like wherever possible to collect blood from both parents as well as your child to help get maximum information from your child's blood sample. Your blood sample will expand the potential information we will gain from taking your children's blood. We appreciate your help with the study and so will arrange to take blood samples at a time and place that is convenient to you.

## Feedback

As the information found is not of any clinical relevance to your child we will not routinely be feeding back any information to you directly. We will however feed the collective results back at local support group meetings (CICRA).

## Is this anonymous ?

While collection of these specimens will not be anonymous only doctors or nurses involved in his/her clinical care will have access to his/her name. This allows us to go back to the hospital notes if we need to identify important and relevant information later on. Within the laboratory anonymity will be maintained by allocating a number (rather than his/her name) to the specimens. Only Dr Van

Limbergen or other staff from the wards/clinics under his supervision will be able to link his/her name to the specimen.

## What happens if I or my child are not keen or change our minds ?

You and your child are entirely at liberty to decide not to participate or drop out at any time and this will not affect his/her care in any way. If after specimens have been collected you change your minds contact us and the specimens will be discarded if they have not already been used.

## Is this ethical ?

We have gained approval from the Local Ethics Committee for the collection and storage of these specimens and data.

**Any other Questions ?** Any other questions can be addressed by writing to Dr Van Limbergen at the Dept of Paediatric Gastroenterology, Sciennes Road , Edinburgh or by contacting his secretary (0131 536 0615).

We also have an independent adviser Professor S J Forsyth, a consultant paediatrician who can answer any concerns you have. Professor Forsyth is based at the Ninewells Hospital in Dundee but is not a member of the research team.

Version 1 – 4/1/6

## *Appendix 4: Patient consent form*

**CHILD'S INVITATION TO PARTICIPATE IN A RESEARCH STUDY: CONSENT**

**THE GENETIC INFLUENCE IN THE DEVELOPMENT OF CROHN'S DISEASE AND COLITIS IN CHILDREN**

**Name of Patient/Volunteer:**

**Address:**

- ***I HAVE READ AND KEPT THE INFORMATION SHEET…***
- ***SPOKEN TO THE DOCTOR AND MY FAMILY AND ASKED ANY QUESTIONS I WOULD LIKE…***
- ***I REALISE I DO NOT HAVE TO DO THIS …***
- ***I REALISE ANY BLOOD SAMPLE STORED WILL ONLY BE USED TO STUDY THE EFFECTS OF GENES ON CROHNS DISEASE AND COLITIS …***
- ***AND CAN STOP OR WITHDRAW AT ANY STAGE WITHOUT AFFECTING MY TREATMENT….***
- ***I AM HAPPY TO GO AHEAD.***

***Signed*** **……**

***Name*** **……**

***Witnessed*** **……**

***Signature*** **……**

***Date*** **……**

Version 1 – 4/1/6

## *Appendix 5: Example of parent/older patient consent form*

**PATIENT WRITTEN CONSENT FORM:**

**Title of research proposal:**

**THE GENETIC INFLUENCE IN THE DEVELOPMENT OF CROHN'S DISEASE AND COLITIS IN CHILDREN**

**LREC Number: LREC/2002/6/18**

**Name of Patient/Volunteer:**

**Address:**

- The study organisers have invited me to take part in this research. **Yes/No**
- I understand what is in the leaflet about the research. I have a copy of the leaflet to keep. **Yes/No**
- I have had the chance to talk and ask questions about the study. **Yes/No**
- I know what my part will be in the study and I know how long it will take. **Yes/No**
- I know I will gain no direct benefit from taking part in the study. **Yes/No**
- I know how the study may affect me. I have been told if there are possible risks. **Yes/No**

- I understand that I should not **actively** take part in more than 1 research study at a time. **Yes/No**

- I know that the local Lothian Research Ethics Committee has seen and agreed to this study. **Yes/No**

.

- I understand that personal information is strictly confidential: I know the only people who may see information about my part in the study are the research team or an official representative of the organisation which funded the research. **Yes/No**

- I freely consent to be a subject in the study. No-one has put pressure on me. **Yes/No**

- I know that I can stop taking part in the study at any time. **Yes/No**

- I know if I do not take part I will still be able to have my normal treatment. **Yes/No**

- I understand that the sample may be stored and held for future research aimed at understanding the genetic influence on inflammatory bowel disease only, and not for other studies without my permission. **Yes/No**

- I know that if there are any problems, I can contact: **Yes/No**

  Dr Van Limbergen

  Tel. No. 0131 536 0615

Patient Signature: .....

Witness's Name .....

Witness's Signature: .....

Date: .....

The following should be signed by the Clinician/Investigator responsible for obtaining consent

As the Clinician/Investigator responsible for this research or a designated deputy, I confirm that I have explained to the patient/volunteer named above the nature and purpose of the research to be undertaken.

Clinician's Name: .....

Clinician's Signature: ..... Date: .....

## *Appendix 6: Patient/parent questionnaire*

**Genetics of Inflammatory Bowel Disease**

**Confidential**

**QUESTIONNAIRE FOR PATIENTS WITH IBD**

**Patient Identification Number** **Sex: Male / Female**

Date of Birth: Hospital ID:

Ethnic origin

| | | | |
|---|---|---|---|
| White European | □ | Hispanic | □ |
| Jewish | □ | Afro-Caribbean | □ |
| Japanese | □ | Asian | □ |
| Other/Unknown | □ | | |

Postcode:

Name of present hospital consultant: Hospital:

Diagnosis: Ulcerative Colitis/ Crohn's disease / indeterminate Colitis/OFG

Date at symptom onset (month/year)

Date of diagnosis: (month/year)

Age at diagnosis:

Year of diagnosis

Years of disease at assessment:

School year:

Other medical problems:

**SMOKING**

**1. Patient**

Does patient smoke cigarettes? YES NO

**If Yes,** How many cigarettes a day?

What age did you start smoking?

**If No,** Have you ever smoked? YES NO

How many cigarettes a day?

What age did you start smoking? (Year) What age did you stop? (Year)

**2. Mother:**

Did you smoke during pregnancy? YES NO

**If Yes,** How many cigarettes a day?

Where you smoking when patient born? YES NO

**If Yes,** How many cigarettes a day?

Do you currently smoke cigarettes? YES NO

**If Yes,** How many cigarettes a day?

What age did you start smoking?

**If No,** Have you ever smoked? YES NO

How many cigarettes a day?

**3. Father:**

Did you smoke during pregnancy? YES NO

If Yes, How many cigarettes a day?

Where you smoking when patient born? YES NO

**If Yes,** How many cigarettes a day?

Do you currently smoke cigarettes? YES NO

**If Yes,** How many cigarettes a day?

What age did you start smoking?

**If No,** Have you ever smoked? YES NO

How many cigarettes a day?

**Pregnancy and Breast-feeding**

Birth weight kg

Complications in pregnancy

Was patient breast-fed at all? YES NO

If so for how long? (Months)

**Vaccinations**

Have you given your child the following vaccinations? (Tick all that apply)

DPT HIB MMR

Hep B Men c

**Family History**

**Do you have?**

Coeliac disease? YES NO

Colon cancer? YES NO

Autistic spectrum disorder? YES NO

Asthma/eczema/hay fever/food allergy? YES NO

**Does your mother or father have?** (List affected family members/age of onset)

Crohn's disease? YES NO

Ulcerative colitis? YES NO

Coeliac disease? YES NO

Colon cancer? YES NO

Autistic spectrum disorder? YES NO

**Brothers and sisters**

1.In total, how many brothers and sisters do you have? Sibling rank

**Do your brother or sisters have?**(List affected family members/age of onset)

| | | |
|---|---|---|
| Crohn's disease? | YES | NO |
| Ulcerative colitis? | YES | NO |
| Coeliac disease? | YES | NO |
| Colon cancer? | YES | NO |
| Autistic spectrum disorder? | YES | NO |

**Do your grandparents, aunts, uncles, or cousins have?** (List affected family members and age of onset)

| | | |
|---|---|---|
| Crohn's disease? | YES | NO |
| Ulcerative colitis? | YES | NO |
| Coeliac disease? | YES | NO |
| Colon cancer? | YES | NO |
| Autistic spectrum disorder? | YES | NO |

**Surgery**

| | | |
|---|---|---|
| Has patient had any operations for IBD? | YES | NO |

**If yes,** please list below, with year if possible. Date of surgery Type of operation

| | | |
|---|---|---|
| Have you had your tonsils removed? | YES | NO |

**If yes,** when? Age/year

| | | |
|---|---|---|
| Have you had your appendix removed? | YES | NO |

**If yes**, when? Age/year

| | | |
|---|---|---|
| Any other surgery? | YES | NO |
| Have you had any broken bones? | YES | NO |

**If yes,** how many?

## *Appendix 7: Crohn's disease phenotypic data collection form*

**Study Number** ☐

**Crohn's disease site** **Presentation / Follow-up** **Year**

| **Site** | **Behaviour** |
|---|---|
| Oral | Inflammatory B1 |
| Oesophagus | Stricturing B2 |
| Gastric body | Penetrating B3 |
| Gastric antrum | **Give details below** |
| Duodenal | Strictures |
| Jejunal | Abscesses |
| Ileal TI | Fistulae |
| Caecal | Inflammatory mass |
| Ascending | Perianal ulcer |
| Transverse | |
| Descending | |
| Sigmoid | |
| Rectal | |
| Perianal | |

**Investigations**

| | | |
|---|---|---|
| DEXA | yes / no | |
| EIM's | yes / no | give details |

## *Appendix 8: UC/IC phenotypic data collection form*

**Study Number** ☐

Presentation / Follow-up Date of Last Follow-up Clinic:

Date of last investigation:

Extent of UC

Examination satisfactory / unsatisfactory

Examination type colonoscopy / sigmoidoscopy

Examination stopped at

Ulceration yes / no

Crypt abscesses yes / no

Cryptitis yes / no

Goblet cell depletion yes / no

Disease years ☐

Extent changed yes / no

Colectomy yes / no

Comments

DEXA yes / no

EIM's yes / no give details

## *Appendix 9 : Laboratory data collection form*

**Study Number** | |

Presentation / Follow-up Year

| | |
|---|---|
| ESR | |
| Hb | |
| PLATELETS | |
| CRP | |
| ALBUMIN | |
| CALPROTECTIN | |
| VITAMIN B12 | |
| FOLATE | |

For CRP please choose from - normal or raised.

Clinic Date: Age at Diagnosis:

Height : Ht centile/Z-score

Weight: Wt centile/Z-score

BMI: BMI centile/Z-score

Puberty progressing normally / delayed / not assessed / not appropriate

## *Appendix 10: The salting out method*

All steps were performed having taken universal precautions before handling blood and all stages were then performed in an appropriate "hood" within the laboratory.

1. 10 mls of whole blood collected previously into an EDTA tube were added into a 50ml conical tube.
2. 40 mls of red cell lysis buffer (RCLB) were then added to the conical tube.
3. The resulting mixture was then mixed and re-suspended for 5 minutes.
4. The solution was centrifuged at 3000 rpm for 10 minutes.
5. The supernatant was then removed leaving a pellet in the "cone" of the tube.
6. The pellet was re-suspended with a further 40 mls of RCLB and then re-centrifuged at the same speed for a further 5 minutes.
7. The resulting "white blood cell" pellet was then further re-suspended in 3 mls of nuclear lysis buffer and sodium dodecyl sulphate.
8. To this solution 1ml of 6M sodium chloride and 3mls of chloroform were then added.
9. The resulting mixture was then aggressively agitated to allow mixing.
10. The solution was then centrifuged at 3000 rpm for 20 minutes.
11. The solution then separated into 3 layers. The middle layer was then carefully pipetted out to avoid mixing with the other 2 layers.
12. The layer was added to 20 mls of 100% ethanol resulting in precipitation of DNA.
13. This DNA pellet was removed from the alcohol and dried in room air for 5 minutes.
14. The DNA was transferred into 0.5 mls of TE and stored at 4ºC until dissolves.

Reference List


1. Stone MA, Mayberry JF, Baker R. Prevalence and management of inflammatory bowel disease: a cross-sectional study from central England. *European Journal of Gastroenterology & Hepatology* 2003;**15**:1275-80.

2. Griffiths AM. Specificities of inflammatory bowel disease in childhood. *Best Practice & Research in Clinical Gastroenterology* 2004;**18**:509-23.

3. Barton JR, Gillon S, Ferguson A. Incidence of inflammatory bowel disease in Scottish children between 1968 and 1983; marginal fall in ulcerative colitis, three-fold rise in Crohn's disease. *Gut* 1989;**30**:618-22.

4. Armitage E, Drummond HE, Wilson DC, Ghosh S. Increasing incidence of both juvenile-onset Crohn's disease and ulcerative colitis in Scotland. *European Journal of Gastroenterology & Hepatology* 2001;**13**:1439-47.

5. Sawczenko A, Sandhu BK, Logan RFA, Jenkins H, Taylor CJ, Mian S *et al*. Prospective survey of childhood inflammatory bowel disease in the British Isles. *Lancet* 2001;**357**:1093-4.

6. Williams JG, Roberts SE, Ali MF, Cheung WY, Cohen DR, Demery G *et al*. Gastroenterology services in the UK. The burden of disease, and the organisation and delivery of services for gastrointestinal and liver disorders: a review of the evidence. *Gut* 2007;**56**:1-113.

7. Dubinsky M. Special issues in pediatric inflammatory bowel disease. *World Journal of Gastroenterology* 2008;**14**:413-20.

8. Cosnes J, Cattan S, Blain A, Beaugerie L, Carbonnel F, Parc R *et al*. Long-Term Evolution of Disease Behavior of Crohn's Disease. *Inflammatory Bowel Diseases* 2002;**8**:244-50.

9. Turner D, Walsh CM, Benchimol EI, Mann EH, Thomas KE, Chow C *et al*. Severe paediatric ulcerative colitis: incidence, outcomes and optimal timing for second-line therapy. *Gut* 2008;**57**:331-8.

10. Bouma G,.Strober W. The immunological and genetic basis of inflammatory bowel disease. *Nature Reviews.Immunology* 2003;**3**:521-33.

11. Korzenik JR,.Dieckgraefe BK. Is Crohn's disease an immunodeficiency? A hypothesis suggesting possible early events in the pathogenesis of Crohn's disease. *Digestive Diseases and Sciences* 2000;**45**:1121-9.

12. Folwaczny C, Glas J, Török HP. Crohn's disease: an immunodeficiency? *European Journal of Gastroenterology & Hepatology* 2003;**15**:621-6.

13. Strober W. Immunology. Unraveling gut inflammation. *Science* 2006;**313**:1052-4.

14. Hugot JP, Chamaillard M, Zouali H, Lesage S, Cezard JP, Belaiche J *et al*. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 2001;**411**:599-603.

15. Ogura Y, Bonen DK, Inohara N, Nicolae DL, Chen FF, Ramos R *et al*. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* 2001;**411**:603-6.

16. Hampe J, Cuthbert A, Croucher PJ, Mirza MM, Mascheretti S, Fisher S *et al*. Association between insertion mutation in NOD2 gene and Crohn's disease in German and British populations. *Lancet* 2001;**357**:1925-8.

17. van Heel DA, Ghosh S, Butler M, Hunt KA, Lundberg AM, Ahmad T *et al*. Muramyl dipeptide and toll-like receptor sensitivity in NOD2-associated Crohn's disease. *Lancet* 2005;**365**:1794-6.

18. van Heel DA, Hunt KA, King K, Ghosh S, Gabe SM, Mathew CG *et al*. Detection of muramyl dipeptide-sensing pathway defects in patients with Crohn's disease. *Inflammatory Bowel Diseases* 2006;**12**:598-605.

19. Bairead E, Harmon DL, Curtis AM, Kelly Y, O'Leary C, Gardner M *et al*. Association of NOD2 with Crohn's disease in a homogenous Irish population. *European Journal of Human Genetics* 2003;**11**:237-44.

20. Lesage S, Zouali H, Cezard JP, Colombel JF, Belaiche J, Almer S *et al*. CARD15/NOD2 mutational analysis and genotype-phenotype correlation in 612 patients with inflammatory bowel disease. *American Journal of Human Genetics* 2002;**70**:845-57.

21. Ahmad T, Armuzzi A, Bunce M, Mulcahy-Hawes K, Marshall SE, Orchard TR *et al*. The molecular classification of the clinical manifestations of Crohn's disease. *Gastroenterology* 2002;**122**:854-6.

22. Brant SR, Picco MF, Achkar JP, Bayless TM, Kane SV, Brzezinski A *et al*. Defining complex contributions of NOD2/CARD15 gene mutations, age at onset, and tobacco use on Crohn's disease phenotypes. *Inflammatory Bowel Diseases* 2003;**9**:281-9.

23. Couper R, Kapelushnik J, Griffiths AM. Neutrophil dysfunction in glycogen storage disease Ib: association with Crohn's-like colitis. *Gastroenterology* 1991;**100**:549-54.

24. D'Agata ID, Paradis K, Chad Z, Bonny Y, Seidman E. Leucocyte adhesion deficiency presenting as a chronic ileocolitis. *Gut* 1996;**39**:605-8.

25. Baehner RL. Chronic granulomatous disease of childhood: clinical, pathological, biochemical, molecular, and genetic aspects of the disease. *Pediatric Pathology* 1990;**10**:143-53.

26. Lopez-Osuna M, Vega-Avila E, Salamanca F, Kretschmer RR. Defective polymorphonuclear chemotaxis in patients with Turner's syndrome (45,X). *Clinical Genetics* 1988;**34**:165-71.

27. Marks DJ, Harbord MW, MacAllister R, Rahman FZ, Young J, Al Lazikani B *et al*. Defective acute inflammation in Crohn's disease: a clinical investigation. *Lancet* 2006;**367**:668-78.

28. Segal AW,.Loewi G. Neutrophil dysfunction in Crohn's disease. *Lancet* 1976;**2**:219-21.

29. O'Morain C, Segal AA, Walker D, Levi AJ. Abnormalities of neutrophil function do not cause the migration defect in Crohn's disease. *Gut* 1981;**22**:817-22.

30. Harbord MW, Marks DJ, Forbes A, Bloom SL, Day RM, Segal AW. Impaired neutrophil chemotaxis in Crohn's disease relates to reduced production of chemokines and can be augmented by granulocyte-colony stimulating factor. *Alimentary Pharmacology & Therapeutics* 2006;**24**:651-60.

31. Brandt E, Colombel JF, Ectors N, Gambiez L, Emilie D, Geboes K *et al*. Enhanced production of IL-8 in chronic but not in early ileal lesions of Crohn's disease. *Clinical and Experimental Immunology* 2000;**122**:180-5.

32. van Lierop PPE, Damen GM, Escher JC, Samsom JN, Nieuwenhuis EES. Defective acute inflammation in Crohn's disease. *The Lancet* 2006;**368**:578.

33. Netea MG, van der Meer JW, Ferwerda G, Kullberg BJ. Defective acute inflammation in Crohn's disease. *The Lancet* 2006;**368**:577-8.

34. Verspaget HW, Pena AS, Weterman IT, Lamers CB. Diminished neutrophil function in Crohn's disease and ulcerative colitis identified by decreased oxidative metabolism and low superoxide dismutase content. *Gut* 1988;**29**:223-8.

35. Curran FT, Allan RN, Keighley MR. Superoxide production by Crohn's disease neutrophils. *Gut* 1991;**32**:399-402.

36. Gionchetti P, Campieri M, Guarnieri C, Belluzzi A, Brignola C, Bertinelli E *et al*. Respiratory burst of circulating polymorphonuclear leukocytes and plasma elastase levels in patients with inflammatory bowel disease in remission. *Digestive Diseases & Sciences* 1994;**39**:550-4.

37. Curran FT, Youngs DJ, Allan RN, Keighley MR. Candidacidal activity of Crohn's disease neutrophils. *Gut* 1991;**32**:55-60.

38. Wandall JH. Function of exudative neutrophilic granulocytes in patients with Crohn's disease or ulcerative colitis. *Scandinavian Journal of Gastroenterology* 1985;1151-6.

39. Dieckgraefe BK,.Korzenik JR. Treatment of active Crohn's disease with recombinant human granulocyte-macrophage colony-stimulating factor. *Lancet* 2002;**360**:1478-80.

40. Korzenik JR, Dieckgraefe BK, Valentine JF, Hausman DF, Gilbert MJ, Sargramostim in Crohn's Disease Study Group. Sargramostim for active Crohn's disease. *New England Journal of Medicine* 2005;**352**:2193-201.

41. Teahon K, Webster AD, Price AB, Weston J, Bjarnason I. Studies on the enteropathy associated with primary hypogammaglobulinaemia. *Gut* 1994;**35**:1244-9.

42. Washington K, Stenzel TT, Buckley RH, Gottfried MR. Gastrointestinal pathology in patients with common variable immunodeficiency and X-linked agammaglobulinemia. *American Journal of Surgical Pathology.* 1996;1240-52.

43. Cellier C, Foray S, Hermine O. Regional enteritis associated with enterovirus in a patient with X-linked agammaglobulinemia. *New England Journal of Medicine* 2000;**342**:1611-2.

44. Sneller MC, Strober W, Eisenstein E, Jaffe JS, Cunningham-Rundles C. NIH conference. New insights into common variable immunodeficiency. *Ann Intern Med* 1993;**118**:720-30.

45. John HA, Sullivan KE, Smith C, Mulberg AE. Enterocolitis in infantile common variable immunodeficiency - A case report and review of the literature. *Digestive Diseases and Sciences* 1996;**41**:621-3.

46. Cunningham-Rundles C,.Bodian C. Common variable immunodeficiency: clinical and immunological features of 248 patients. *Clinical Immunology* 1999;**92**:34-48.

47. Remold - O'Donnell E, Rosen FS, Kenney DM. Defects in Wiskott-Aldrich syndrome blood cells. *Blood* 1996;**87**:2621-31.

48. Featherstone C. The many faces of WAS protein. *Science* 1997;**275**:27-8.

49. Akman IO, Ostrov BE, Neudorf S. Autoimmune manifestations of the Wiskott-Aldrich syndrome. *Seminars in Arthritis & Rheumatism* 1998;**27**:218-25.

50. Folwaczny C, Ruelfs C, Walther J, Konig A, Emmerich B. Ulcerative colitis in a patient with Wiskott-Aldrich syndrome. *Endoscopy* 2002;**34**:840-1.

51. Werlin SL, Chusid MJ, Caya J, Oechler HW. Colitis in chronic granulomatous disease. *Gastroenterology* 1982;**82**:328-31.

52. Newman SL,.Lindahl JA. Chronic granulomatous disease: a model for gastrointestinal manifestations of immunodeficiency. *Journal of Clinical Gastroenterology* 1985;**7**:470-1.

53. Isaacs D, Wright VM, Shaw DG, Raafat F, Walker-Smith JA. Chronic granulomatous disease mimicking Crohn's disease. *Journal of Pediatric Gastroenterology & Nutrition* 1985;**4**:498-501.

54. Ahlin A, De Boer M, Roos D, Leusen J, Smith CI, Sundin U *et al*. Prevalence, genetics and clinical presentation of chronic granulomatous disease in Sweden. *Acta Paediatrica* 1995;**84**:1386-94.

55. Sloan JM, Cameron CH, Maxwell RJ, McCluskey DR, Collins JS. Colitis complicating chronic granulomatous disease. A clinicopathological case report. *Gut* 1996;**38**:619-22.

56. Roe TF, Thomas DW, Gilsanz V, Isaacs H, Jr., Atkinson JB. Inflammatory bowel disease in glycogen storage disease type Ib. *Journal of Pediatrics* 1986;**109**:55-9.

57. Sanderson IR, Bisset WM, Milla PJ, Leonard JV. Chronic inflammatory bowel disease in glycogen storage disease type 1B. *Journal of Inherited Metabolic Disease* 1991;**14**:771-6.

58. Hoover EG, DuBois JJ, Samples TL, McCullough JS, Chenaille PJ, Montes RG. Treatment of chronic enteritis in glycogen storage disease type IB with granulocyte colony-stimulating factor. *Journal of Pediatric Gastroenterology & Nutrition* 1996;**22**:346-50.

59. Ishii E, Matui T, Iida M, Inamitu T, Ueda K. Chediak-Higashi syndrome with intestinal complication. Report of a case. *Journal of Clinical Gastroenterology* 1987;**9**:556-8.

60. Grucela AL, Patel P, Goldstein E, Palmon R, Sachar DB, Steinhagen RM. Granulomatous enterocolitis associated with Hermansky-Pudlak syndrome. *American Journal of Gastroenterology* 2006;**101**:2090-5.

61. Knudtzon J,.Svane S. Turner's syndrome associated with chronic inflammatory bowel disease. A case report and review of the literature. *Acta Medica Scandinavica* 1988;**223**:375-8.

62. Vannier JP, Arnaud-Battandier F, Ricour C, Schmitz J, Buriot D, Griscelli C *et al*. Chronic neutropenia and Crohn's disease in childhood. Report of 2 cases. *Archives Francaises de Pediatrie* 1982;**39**:367-70.

63. Herberhold C,.Folwaczny C. Ulcerating proctitis in a patient with congenital neutropenia. *American Journal of Gastroenterology* 2003;**98**:1204-6.

64. Lamport RD, Katz S, Eskreis D. Crohn's disease associated with cyclic neutropenia. *American Journal of Gastroenterology* 1992;**87**:1638-42.

65. Stevens C, Peppercorn MA, Grand RJ. Crohn's disease associated with autoimmune neutropenia. *Journal of Clinical Gastroenterology* 1991;**13**:328-30.

66. Guarner F. The intestinal flora in inflammatory bowel disease: normal or abnormal? *Current Opinion in Gastroenterology.21(4):414-8,* 2005.

67. Korzenik JR,.Podolsky DK. Evolving knowledge and therapy of inflammatory bowel disease. *Nature Reviews.Drug Discovery* 2006;**5**:197-209.

68. Ley RE, Peterson DA, Gordon JI. Ecological and Evolutionary Forces Shaping Microbial Diversity in the Human Intestine. *Cell* 2006;**124**:837-48.

69. Mackie RI, Sghir A, Gaskins HR. Developmental microbial ecology of the neonatal gastrointestinal tract. *Am J Clin Nutr* 1999;**69**:1035S-1045.

70. Goldman AS. Modulation of the gastrointestinal tract of infants by human milk. Interfaces and interactions. An evolutionary perspective. *J.Nutr.* 2000;**130**:426S-31S.

71. Labbok MH, Clark D, Goldman AS. Breastfeeding: maintaining an irreplaceable immunological resource. *Nature Reviews.Immunology* 2004;**4**:565-72.

72. Blais DR, Harrold J, Altosaar I. Killing the messenger in the nick of time: persistence of breast milk sCD14 in the neonatal gastrointestinal tract. *Pediatr Res* 2006;**59**:371-6.

73. Liepke C, Adermann K, Raida M, Magert HJ, Forssmann WG, Zucht HD. Human milk provides peptides highly stimulating the growth of bifidobacteria. *European Journal of Biochemistry* 2002;**269**:712-8.

74. Beerens H, Romond C, Neut C. Influence of breast-feeding on the bifid flora of the newborn intestine. *Am J Clin Nutr* 1980;**33**:2434-9.

75. Laitinen K, Hoppu U, Hamalainen M, Linderborg K, Moilanen E, Isolauri E. Breast milk fatty acids may link innate and adaptive immune regulation: analysis of soluble CD14, prostaglandin E2, and fatty acids. *Pediatr Res* 2006;**59**:723-7.

76. LeBouder E, Rey-Nores JE, Raby AC, Affolter M, Vidal K, Thornton CA *et al*. Modulation of neonatal microbial recognition: TLR-mediated innate

immune responses are specifically and differentially modulated by human milk. *Journal of Immunology* 2006;**176**:3742-52.

77. Menard S, Forster V, Lotz M, Gutle D, Duerr CU, Gallo RL *et al*. Developmental switch of intestinal antimicrobial peptide expression. *J.Exp.Med.* 2008;**205**:183-93.

78. Penders J, Thijs C, Vink C, Stelma FF, Snijders B, Kummeling I *et al*. Factors Influencing the Composition of the Intestinal Microbiota in Early Infancy. *Pediatrics* 2006;**118**:511-21.

79. Klement E, Cohen RV, Boxman J, Joseph A, Reif S. Breastfeeding and risk of inflammatory bowel disease: a systematic review with meta-analysis. *Am J Clin Nutr* 2004;**80**:1342-52.

80. Jantchou P, Turck D, Balde M, Gower-Rousseau C. Breastfeeding and risk of inflammatory bowel disease: results of a pediatric, population-based, case-control study. *Am J Clin Nutr* 2005;**82**:485-6.

81. Baron S, Turck D, Leplat C, Merle V, Gower-Rousseau C, Marti R *et al*. Environmental risk factors in paediatric inflammatory bowel diseases: a population based case control study. *Gut* 2005;**54**:357-63.

82. Sandler RS, Sandler DP, McDonnell CW, Wurzelmann JI. Childhood exposure to environmental tobacco smoke and the risk of ulcerative colitis. *Am.J.Epidemiol.* 1992;**135**:603-8.

83. Lashner BA, Shaheen NJ, Hanauer SB, Kirschner BS. Passive smoking is associated with an increased risk of developing inflammatory bowel disease in children. *American Journal of Gastroenterology* 1993;**88**:356-9.

84. Russell RK, Farhadi R, Wilson M, Drummond H, Satsangi J, Wilson DC. Perinatal passive smoke exposure may be more important than childhood exposure in the risk of developing childhood IBD. *Gut* 2005;**54**:1500-1.

85. Lashner BA,.Loftus EV, Jr. True or false? The hygiene hypothesis for Crohn's disease. *American Journal of Gastroenterology* 2006;**101**:1003-4.

86. Aspberg S, Dahlquist G, Kahan T, Kallen B. Fetal and perinatal risk factors for inflammatory bowel disease. *Acta Paediatrica* 2006;**95**:1001-4.

87. Swidsinski A, Ladhoff A, Pernthaler A, Swidsinski S, Loening-Baucke V, Ortner M *et al*. Mucosal flora in inflammatory bowel disease. *Gastroenterology* 2002;**122**:44-54.

88. Ott SJ,.Schreiber S. Reduced microbial diversity in inflammatory bowel diseases. *Gut* 2006;**55**:1207.

89. Manichanh C, Rigottier-Gois L, Bonnaud E, Gloux K, Pelletier E, Frangeul L *et al*. Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut* 2006;**55**:205-11.

90. Schultsz C, Van Den Berg FM, Ten Kate FW, Tytgat GN, Dankert J. The intestinal mucus layer from patients with inflammatory bowel disease harbors high numbers of bacteria compared with controls. *Gastroenterology.117(5):1089-97,* 1999.

91. Darfeuille-Michaud A, Boudeau J, Bulois P, Neut C, Glasser AL, Barnich N *et al*. High prevalence of adherent-invasive Escherichia coli associated with ileal mucosa in Crohn's disease. *Gastroenterology* 2004;**127**:412-21.

92. Mylonaki M, Rayment NB, Rampton DS, Hudspith BN, Brostoff J. Molecular characterization of rectal mucosa-associated bacterial flora in inflammatory bowel disease. *Inflammatory Bowel Diseases* 2005;**11**:481-7.

93. Conte MP, Schippa S, Zamboni I, Penta M, Chiarini F, Seganti L *et al*. Gut-associated bacterial microbiota in paediatric patients with inflammatory bowel disease. *Gut* 2006;**55**:1760-7.

94. van der Waaij LA, Harmsen HJ, Madjipour M, Kroese FG, Zwiers M, van Dullemen HM *et al*. Bacterial population analysis of human colon and terminal ileum biopsies with 16S rRNA-based fluorescent probes: commensal bacteria live in suspension and have no direct contact with epithelial cells. *Inflammatory Bowel Diseases* 2005;**11**:865-71.

95. Frank DN, St Amand AL, Feldman RA, Boedeker EC, Harpaz N, Pace NR. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc.Natl Acad.Sci.U.S.A* 2007;**104**:13780-5.

96. Sartor RB. Microbial influences in inflammatory bowel diseases. *Gastroenterology* 2008;**134**:577-94.

97. Garcia Rodriguez LA, Ruigomez A, Panes J. Acute gastroenteritis is followed by an increased risk of inflammatory bowel disease. *Gastroenterology* 2006;**130**:1588-94.

98. Sartor RB. Therapeutic manipulation of the enteric microflora in inflammatory bowel diseases: antibiotics, probiotics, and prebiotics. *Gastroenterology* 2004;**126**:1620-33.

99. Kuhbacher T, Ott SJ, Helwig U, Mimura T, Rizzello F, Kleessen B *et al*. Bacterial and fungal microbiota in relation to probiotic therapy (VSL#3) in pouchitis. *Gut* 2006;**55**:833-41.

100. Andoh A,.Fujiyama Y. Therapeutic approaches targeting intestinal microflora in inflammatory bowel disease. *World Journal of Gastroenterology* 2006;**12**:4452-60.

101. Shirazi T, Longman RJ, Corfield AP, Probert CSJ. Mucins and inflammatory bowel disease. *Postgraduate Medical Journal* 2000;**76**:473-8.

102. Van der SM, De Koning BA, De Bruijn AC, Velcich A, Meijerink JP, Van Goudoever JB *et al*. Muc2-deficient mice spontaneously develop colitis, indicating that MUC2 is critical for colonic protection. *Gastroenterology* 2006;**131**:117-29.

103. Einerhand AW, Renes IB, Makkink MK, Van der SM, Buller HA, Dekker J. Role of mucins in inflammatory bowel disease: important lessons from experimental models. *European Journal of Gastroenterology & Hepatology* 2002;**14**:757-65.

104. Duraisamy S, Ramasamy S, Kharbanda S, Kufe D. Distinct evolution of the human carcinoma-associated transmembrane mucins, MUC1, MUC4 AND MUC16. *Gene* 2006;**373**:28-34.

105. Porchet N,.Aubert JP. MUC genes: mucin or not mucin? That is the question. *Medecine Sciences.* 2004;569-74.

106. Pigny P, Guyonnet-Duperat V, Hill AS, Pratt WS, Galiegue-Zouitina S, d'Hooge MC *et al*. Human mucin genes assigned to 11p15.5: identification and organization of a cluster of genes. *Genomics* 1996;**38**:340-52.

107. van Heel DA, Fisher SA, Kirby A, Daly MJ, Rioux JD, Lewis CM *et al*. Inflammatory bowel disease susceptibility loci defined by genome scan meta-analysis of 1952 affected relative pairs. *Hum.Mol.Genet.* 2004;**13**:763-70.

108. Satsangi J, Parkes M, Louis E, Hashimoto L, Kato N, Welsh K *et al*. Two stage genome-wide search in inflammatory bowel disease provides evidence for susceptibility loci on chromosomes 3, 7 and 12. *Nature Genetics* 1996;**14**:199-202.

109. Koscinski I, Viville S, Porchet N, Bernigaud A, Escande F, Defossez A *et al*. MUC4 gene polymorphism and expression in women with implantation failure. *Hum.Reprod.* 2006;e-pub 28-6/2006.

110. Van Klinken BJ, Dekker J, Buller HA, de Bolos C, Einerhand AW. Biosynthesis of mucins (MUC2-6) along the longitudinal axis of the human gastrointestinal tract. *American Journal of Physiology* 1997;**273**:G296-G302.

111. Chang SK, Dohrman AF, Basbaum CB, Ho SB, Tsuda T, Toribara NW *et al*. Localization of mucin (MUC2 and MUC3) messenger RNA and peptide

expression in human normal intestine and colon cancer. *Gastroenterology* 1994;**107**:28-36.

112. Velcich A, Yang W, Heyer J, Fragale A, Nicholas C, Viani S *et al*. Colorectal cancer in mice genetically deficient in the mucin Muc2. *Science* 2002;**295**:1726-9.

113. Yang W, Velcich A, Lozonschi I, Liang J, Nicholas C, Zhuang M *et al*. Inactivation of p21WAF1/cip1 enhances intestinal tumor formation in Muc2-/- mice. *Am J Pathol* 2005;**166**:1239-46.

114. Van Klinken BJ, Dekker J, van Gool SA, van Marle J, Buller HA, Einerhand AW. MUC5B is the prominent mucin in human gallbladder and is also expressed in a subset of colonic goblet cells. *American Journal of Physiology* 1998;**274**:G871-G878.

115. Ho SB, Roberton AM, Shekels LL, Lyftogt CT, Niehans GA, Toribara NW. Expression cloning of gastric mucin complementary DNA and localization of mucin gene expression. *Gastroenterology* 1995;**109**:735-47.

116. Toribara NW, Roberton AM, Ho SB, Kuo WL, Gum E, Hicks JW *et al*. Human gastric mucin. Identification of a unique species by expression cloning. *J.Biol.Chem.* 1993;**268**:5879-85.

117. Gendler SJ,.Spicer AP. Epithelial mucin genes. *Annual Review of Physiology* 1995;**57**:607-34.

118. Pullan RD, Thomas GAO, Rhodes M, Newcombe RG, Williams GT, Allen A *et al*. Thickness of Adherent Mucus Gel on Colonic Mucosa in Humans and Its Relevance to Colitis. *Gut* 1994;**35**:353-9.

119. Dwarakanath AD, Campbell BJ, Tsai HH, Sunderland D, Hart CA, Rhodes JM. Faecal mucinase activity assessed in inflammatory bowel disease using 14C threonine labelled mucin substrate. *Gut* 1995;**37**:58-62.

120. Tsai HH, Dwarakanath AD, Hart CA, Milton JD, Rhodes JM. Increased faecal mucin sulphatase activity in ulcerative colitis: a potential target for treatment. *Gut* 1995;**36**:570-6.

121. Rhodes JM. Unifying hypothesis for inflammatory bowel disease and associated colon cancer: sticking the pieces together with sugar. *Lancet* 1996;**347**:40-4.

122. Bodger K, Halfvarson J, Dodson AR, Campbell F, Wilson S, Lee R *et al*. Altered colonic glycoprotein expression in unaffected monozygotic twins of inflammatory bowel disease patients. *Gut* 2006;**55**:973-7.

123. Tysk C, Riedesel H, Lindberg E, Panzini B, Podolsky D, Jarnerot G. Colonic glycoproteins in monozygotic twins with inflammatory bowel disease. *Gastroenterology* 1991;**100**:419-23.

124. Van S, I, Pigny P, Perrais M, Porchet N, Aubert JP. Transcriptional regulation of the 11p15 mucin genes. Towards new biological tools in human therapy, in inflammatory diseases and cancer? *Frontiers in Bioscience* 2001;**6**:D1216-D1234.

125. Campbell BJ, Rowe GE, Leiper K, Rhodes JM. Increasing the intra-Golgi pH of cultured LS174T goblet-differentiated cells mimics the decreased mucin sulfation and increased Thomsen-Friedenreich antigen (Gal beta1-3GalNac alpha-) expression seen in colon cancer. *Glycobiology* 2001;**11**:385-93.

126. Gouyer V, Wiede A, Buisine MP, Dekeyser S, Moreau O, Lesuffleur T *et al*. Specific secretion of gel-forming mucins and TFF peptides in HT-29 cells of mucin-secreting phenotype. *Biochimica et Biophysica Acta* 2001;**1539**:71-84.

127. Enss ML, Cornberg M, Wagner S, Gebert A, Henrichs M, Eisenblatter R *et al*. Proinflammatory cytokines trigger MUC gene expression and mucin release in the intestinal cancer cell line LS180. *Inflammation Research* 2000;**49**:162-9.

128. Gaudier E, Jarry A, Blottiere HM, de Coppet P, Buisine MP, Aubert JP *et al*. Butyrate specifically modulates MUC gene expression in intestinal epithelial goblet cells deprived of glucose. *American Journal of Physiology - Gastrointestinal & Liver Physiology* 2004;**287**:G1168-G1174.

129. Mack DR, Ahrne S, Hyde L, Wei S, Hollingsworth MA. Extracellular MUC3 mucin secretion follows adherence of Lactobacillus strains to intestinal epithelial cells in vitro. *Gut* 2003;**52**:827-33.

130. Beatty PL, Plevy SE, Sepulveda AR, Finn OJ. Cutting Edge: Transgenic Expression of Human MUC1 in IL-10-/- Mice Accelerates Inflammatory Bowel Disease and Progression to Colon Cancer. *J Immunol* 2007;**179**:735-9.

131. Hoebler C, Gaudier E, de Coppet P, Rival M, Cherbut C. MUC genes are differently expressed during onset and maintenance of inflammation in dextran sodium sulfate-treated mice. *Digestive Diseases & Sciences* 2006;**51**:381-9.

132. Wright NA, Pike C, Elia G. Induction of a novel epidermal growth factor-secreting cell lineage by mucosal ulceration in human gastrointestinal stem cells. *Nature* 1990;**343**:82-5.

133. Schwerbrock NM, Makkink MK, Van der SM, Buller HA, Einerhand AW, Sartor RB *et al*. Interleukin 10-deficient mice exhibit defective colonic Muc2 synthesis before and after induction of colitis by commensal bacteria. *Inflammatory Bowel Diseases* 2004;**10**:811-23.

134. Faure M, Mettraux C, Moennoz D, Godin JP, Vuichoud J, Rochat F *et al*. Specific amino acids increase mucin synthesis and microbiota in dextran sulfate sodium-treated rats. *J.Nutr.* 2006;**136**:1558-64.

135. Kleessen B,.Blaut M. Modulation of gut mucosal biofilms. *British Journal of Nutrition* 2005;**93**:S35-S40.

136. Smirnov A, Perez R, Amit-Romach E, Sklan D, Uni Z. Mucin Dynamics and Microbial Populations in Chicken Small Intestine Are Changed by Dietary Probiotic and Antibiotic Growth Promoter Supplementation. *J.Nutr.* 2005;**135**:187-92.

137. Ho SB, Dvorak LA, Moor RE, Jacobson AC, Frey MR, Corredor J *et al*. Cysteine-rich domains of Muc3 intestinal mucin promote cell migration, inhibit apoptosis, and accelerate wound healing. *Gastroenterology* 2006;**131**:1501-17.

138. Tytgat KM, Van der Wal JW, Einerhand AW, Buller HA, Dekker J. Quantitative analysis of MUC2 synthesis in ulcerative colitis. *Biochemical & Biophysical Research Communications* 1996;**224**:397-405.

139. Tytgat KM, Opdam FJ, Einerhand AW, Buller HA, Dekker J. MUC2 is the prominent colonic mucin expressed in ulcerative colitis. *Gut* 1996;**38**:554-63.

140. Van Klinken BJ, Van der Wal JW, Einerhand AW, Buller HA, Dekker J. Sulphation and secretion of the predominant secretory human colonic mucin MUC2 in ulcerative colitis. *Gut* 1999;**44**:387-93.

141. Myerscough N, Warren B, Gough M, Corfield A. Expression of mucin genes in ulcerative colitis. *Biochemical Society Transactions* 1995;**23**:536S.

142. Buisine MP, Desreumaux P, Debailleul V, Gambiez L, Geboes K, Ectors N *et al*. Abnormalities in mucin gene expression in Crohn's disease. *Inflammatory Bowel Diseases* 1999;**5**:24-32.

143. Buisine MP, Desreumaux P, Leteurtre E, Copin MC, Colombel JF, Porchet N *et al*. Mucin gene expression in intestinal epithelial cells in Crohn's disease. *Gut* 2001;**49**:544-51.

144. Longman RJ, Douthwaite J, Sylvester PA, Poulsom R, Corfield AP, Thomas MG *et al*. Coordinated localisation of mucins and trefoil peptides in the ulcer associated cell lineage and the gastrointestinal mucosa. *Gut* 2000;**47**:792-800.

145. Shaoul R, Okada Y, Cutz E, Marcon MA. Colonic expression of MUC2, MUC5AC, and TFF1 in inflammatory bowel disease in children. *Journal of Pediatric Gastroenterology and Nutrition* 2004;**38**:488-93.

146. Kyo K, Parkes M, Takei Y, Nishimori H, Vyas P, Satsangi J *et al*. Association of ulcerative colitis with rare VNTR alleles of the human intestinal mucin gene, MUC3. *Hum.Mol.Genet.* 1999;**8**:307-11.

147. Kyo K, Muto T, Nagawa H, Lathrop GM, Nakamura Y. Associations of distinct variants of the intestinal mucin gene MUC3A with ulcerative colitis and Crohn's disease. *Journal of Human Genetics* 2001;**46**:5-20.

148. Moehle C, Ackermann N, Langmann T, Aslanidis C, Kel A, Kel-Margoulis O *et al*. Aberrant intestinal expression and allelic variants of mucin genes associated with inflammatory bowel disease. *Journal of Molecular Medicine* 2006;**84**:1055-66.

149. Taupin D,.Podolsky DK. Trefoil factors: Initiators of mucosal healing. *Nature Reviews Molecular Cell Biology* 2003;**4**:721-32.

150. Hoffmann W. TFF (trefoil factor family) peptide-triggered signals promoting mucosal restitution. *Cellular & Molecular Life Sciences* 2005;**62**:2932-8.

151. Ruchaud-Sparagano MH, Westley BR, May FE. The trefoil protein TFF1 is bound to MUC5AC in human gastric mucosa. *Cell Mol.Life Sci.* 2004;**61**:1946-54.

152. Kinoshita K, Taupin DR, Itoh H, Podolsky DK. Distinct pathways of cell migration and antiapoptotic response to epithelial injury: Structure-function analysis of human intestinal trefoil factor. *Molecular & Cellular Biology* 2000;**20**:4680-90.

153. Gott P, Beck S, Machado JC, Carneiro F, Schmitt H, Blin N. Human trefoil peptides: Genomic structure in 21q22.3 and coordinated expression. *European Journal of Human Genetics* 1996;**4**:308-15.

154. Seib T, Blin N, Hilgert K, Seifert M, Theisinger B, Engel M *et al*. The three human trefoil genes TFF1, TFF2, and TFF3 are located within a region of 55 kb on chromosome 21q22.3. *Genomics* 1997;**40**:200-2.

155. Tomasetto C,.Rio M-C. Pleiotropic effects of Trefoil Factor 1 deficiency. *Cellular & Molecular Life Sciences* 2005;**62**:2916-20.

156. Luqmani Y, Bennett C, Paterson I, Corbishley CM, Rio MC, Chambon P *et al*. Expression of the Ps2 Gene in Normal, Benign and Neoplastic Human Stomach. *International Journal of Cancer* 1989;**44**:806-12.

157. Lefebvre O, Chenard M-P, Masson R, Linares J, Dierich A, LeMeur M *et al*. Gastric mucosa abnormalities and tumorigenesis in mice lacking the pS2 trefoil protein. *Science* 1996;**274**:259-62.

158. Tomasetto C, Masson R, Linares J-L, Wendling C, Lefebvre O, Chenard M *et al*. pS2/TFF1 interacts directly with the VWFC cysteine-rich domains of mucins. *Gastroenterology* 2000;**118**:70-80.

159. Poulsom R, Chinery R, Sarraf C, Van Noorden S, Stamp GW, Lalani EN *et al*. Trefoil peptide gene expression in small intestinal Crohn's disease and dietary adaptation. *Journal of Clinical Gastroenterology* 1993;**17**:S78-S91.

160. Wright NA, Poulsom R, Stamp G, Van Noorden S, Sarraf C, Elia G *et al*. Trefoil peptide gene expression in gastrointestinal epithelial cells in inflammatory bowel disease. *Gastroenterology* 1993;**104**:12-20.

161. Hanby AM, Poulsom R, Elia G, Singh S, Longcroft JM, Wright NA. The expression of the trefoil peptides pS2 and human spasmolytic polypeptide (hSP) in 'gastric metaplasia' of the proximal duodenum: Implications for the nature of 'gastric metaplasia'. *Journal of Pathology* 1993;**169**:355-60.

162. Farrell JJ, Taupin D, Koh TJ, Chen D, Zhao C-M, Podolsky DK *et al*. TFF2/SP-deficient mice show decreased gastric proliferation, increased acid secretion, and increased susceptibility to NSAID injury. *J.Clin.Invest.* 2002;**109**:193-204.

163. Baus-Loncar M, Schmid J, Lalani E-N, Rosewell I, Goodlad RA, Stamp GWH *et al*. Trefoil factor 2 (Tff2) deficiency in murine digestive tract influences the immune system. *Cellular Physiology & Biochemistry* 2005;**16**:31-42.

164. Baus-Loncar M, Kayademir T, Takaishi S, Wang T. Trefoil factor family 2 deficiency and immune response. *Cellular & Molecular Life Sciences* 2005;**62**:2947-55.

165. Lanningham-Foster L, Green CL, Langkamp-Henken B, Davis BA, Nguyen KT, Bender BS *et al*. Overexpression of CRIP in transgenic mice alters cytokine patterns and the immune response. *American Journal of Physiology - Endocrinology & Metabolism* 2002;**282**:E1197-E1203.

166. Giraud AS, Pereira PM, Thim L, Parker LM, Judd LM. TFF-2 inhibits iNOS/NO in monocytes, and nitrated protein in healing colon after colitis. *Peptides* 2004;**25**:803-9.

167. Soriano-Izquierdo A, Gironella M, Massaguer A, May FE, Salas A, Sans M *et al*. Trefoil peptide TFF2 treatment reduces VCAM-1 expression and leukocyte recruitment in experimental intestinal inflammation. *Journal of Leukocyte Biology* 2004;**75**:214-23.

168. Tran CP, Cook GA, Yeomans ND, Thim L, Giraud AS. Trefoil peptide TFF2 (spasmolytic polypeptide) potently accelerates healing and reduces inflammation in a rat model of colitis. *Gut* 1999;**44**:636-42.

169. Kindon H, Pothoulakis C, Thim L, Lynchdevaney G, Podolsky DK. Trefoil Peptide Protection of Intestinal Epithelial Barrier Function - Cooperative Interaction with Mucin Glycoprotein. *Gastroenterology* 1995;**109**:516-23.

170. Dignass A, Lynch-Devaney K, Kindon H, Thim L, Podolsky DK. Trefoil peptides promote epithelial migration through a transforming growth factor beta-independent pathway. *J.Clin.Invest.* 1994;**94**:376-83.

171. Mashimo H, Wu DC, Podolsky DK, Fishman MC. Impaired defense of intestinal mucosa in mice lacking intestinal trefoil factor. *Science* 1996;**274**:262-5.

172. Tan XD, Liu QP, Hsueh W, Chen YH, Chang H, Gonzalez-Crussi F. Intestinal trefoil factor binds to intestinal epithelial cells and induces nitric oxide production: priming and enhancing effects of mucin. *Biochem J* 1999;**338**:745-51.

173. Liu J, Miwa T, Hilliard B, Chen Y, Lambris JD, Wells AD *et al*. The complement inhibitory protein DAF (CD55) suppresses T cell immunity in vivo. *Journal of Experimental Medicine* 2005;567-77.

174. Andoh A, Kinoshita K, Rosenberg I, Podolsky DK. Intestinal trefoil factor induces decay-accelerating factor expression and enhances the protective activities against complement activation in intestinal epithelial cells. *Journal of Immunology* 2001;**167**:3887-93.

175. Heeger PS, Lalli PN, Lin F, Valujskikh A, Liu J, Muqim N *et al*. Decay-accelerating factor modulates induction of T cell immunity. *Journal of Experimental Medicine.* 2005;1523-30.

176. Makidono C, Mizuno M, Nasu J, Hiraoka S, Okada H, Yamamoto K *et al*. Increased serum concentrations and surface expression on peripheral white blood cells of decay-accelerating factor (CD55) in patients with active ulcerative colitis. *Journal of Laboratory & Clinical Medicine* 2004;**143**:152-8.

177. Uesu T, Mizuno M, Inoue H, Tomoda J, Tsuji T. Enhanced expression of decay accelerating factor and CD59/homologous restriction factor 20 on the colonic epithelium of ulcerative colitis. *Laboratory Investigation* 1995;**72**:587-91.

178. Berstad AE,.Brandtzaeg P. Expression of cell membrane complement regulatory glycoproteins along the normal and diseased human gastrointestinal tract. *Gut* 1998;**42**:522-9.

179. Zhu YQ,.Tan XD. TFF3 modulates NF-{kappa}B and a novel negative regulatory molecule of NF-{kappa}B in intestinal epithelial cells via a mechanism distinct from TNF-{alpha}. *American Journal of Physiology - Cell Physiology* 2005;**289**:C1085-C1093.

180. X.Yio , J.Yio , J.Zhang , L.Shao , G.Roda , S.Dahan *et al*. Human trefoil factor family-3 induces monocyte production of IL-10 and inhibits T cell proliferation. *DDW 2006* 2006;S1671.

181. Mahmood A, Melley L, Fitzgerald AJ, Ghosh S, Playford RJ. Trial of trefoil factor 3 enemas, in combination with oral 5-aminosalicylic acid, for the treatment of mild-to-moderate left-sided ulcerative colitis. *Alimentary Pharmacology & Therapeutics* 2005;**21**:1357-64.

182. Steidler L,.Vandenbroucke K. Genetically modified Lactococcus lactis: novel tools for drug delivery. *International Journal of Dairy Technology* 2006;**59**:140-6.

183. Vandenbroucke K, Hans W, Van Huysse J, Neirynck S, Demetter P, Remaut E *et al*. Active delivery of trefoil factors by genetically modified Lactococcus lactis prevents and heals acute colitis in mice. *Gastroenterology* 2004;**127**:502-13.

184. Playford RJ,.Ghosh S. Cytokines and growth factor modulators in intestinal inflammation and repair. *Journal of Pathology* 2005;**205**:417-25.

185. Baus-Loncar M,.Giraud AS. Multiple regulatory pathways for trefoil factor (TFF) genes. *Cellular & Molecular Life Sciences* 2005;**62**:2921-31.

186. Dossinger V, Kayademir T, Blin N, Gott P. Down-regulation of TFF expression in gastrointestinal cell lines by cytokines and nuclear factors. *Cellular Physiology & Biochemistry* 2002;**12**:197-206.

187. Loncar MB, Al azzeh ED, Sommer PS, Marinovic M, Schmehl K, Kruschewski M *et al*. Tumour necrosis factor alpha and nuclear factor kappaB inhibit transcription of human TFF3 encoding a gastrointestinal healing peptide. *Gut* 2003;**52**:1297-303.

188. Tebbutt NC, Giraud AS, Inglese M, Jenkins B, Waring P, Clay FJ *et al*. Reciprocal regulation of gastrointestinal homeostasis by SHP2 and STAT-mediated trefoil gene activation in gp130 mutant mice.[see comment]. *Nature Medicine* 2002;**8**:1089-97.

189. Blanchard C, Durual S, Estienne M, Bouzakri K, Heim MH, Blin N *et al*. IL-4 and IL-13 up-regulate intestinal trefoil factor expression: requirement for STAT6 and de novo protein synthesis. *Journal of Immunology* 2004;**172**:3775-83.

190. Shimada T, Koitabashi A, Kuniyoshi T, Hashimoto T, Yoshiura K, Yoneda M *et al*. Up-regulation of TFF expression by PPARgamma ligands in gastric epithelial cells. *Alimentary Pharmacology & Therapeutics* 2003;**18**:119-25.

191. Sugawara K, Olson TS, Moskaluk CA, Stevens BK, Hoang S, Kozaiwa K *et al*. Linkage to peroxisome proliferator-activated receptor-gamma in SAMP1/YitFc mice and in human Crohn's disease.[see comment]. *Gastroenterology* 2005;**128**:351-60.

192. Schreiber S. Slipping the barrier: how variants in CARD15 could alter permeability of the intestinal wall and population health. *Gut* 2006;**55**:308-9.

193. Rosenstiel P, Fantini M, Brautigam K, Kuhbacher T, Waetzig GH, Seegert D *et al*. TNF-alpha and IFN-gamma regulate the expression of the NOD2 (CARD15) gene in human intestinal epithelial cells. *Gastroenterology* 2003;**124**:1001-9.

194. Hisamatsu T, Suzuki M, Podolsky DK. Interferon-gamma augments CARD4/NOD1 gene and protein expression through interferon regulatory factor-1 in intestinal epithelial cells. *J.Biol.Chem.* 2003;**278**:32962-8.

195. Hisamatsu T, Suzuki M, Reinecker HC, Nadeau WJ, McCormick BA, Podolsky DK. CARD15/NOD2 functions as an antibacterial factor in human intestinal epithelial cells. *Gastroenterology* 2003;**124**:993-1000.

196. Abreu MT, Fukata M, Arditi M. TLR Signaling in the Gut in Health and Disease. *J Immunol* 2005;**174**:4453-60.

197. Gitter AH, Bendfeldt K, Schmitz H, Schulzke JD, Bentzel CJ, Fromm M. Epithelial barrier defects in HT-29/B6 colonic cell monolayers induced by tumor necrosis factor-alpha. *Annals of the New York Academy of Sciences* 2000;**915**:193-203.

198. Ye D, Ma I, Ma TY. Molecular mechanism of tumor necrosis factor-alpha modulation of intestinal epithelial tight junction barrier. *American Journal of Physiology - Gastrointestinal & Liver Physiology* 2006;**290**:G496-G504.

199. Gitter AH, Bendfeldt K, Schulzke JD, Fromm M. Leaks in the epithelial barrier caused by spontaneous and TNF-{alpha}-induced single-cell apoptosis. *FASEB J.* 2000;**14**:1749-53.

200. Bruewer M, Luegering A, Kucharzik T, Parkos CA, Madara JL, Hopkins AM *et al*. Proinflammatory Cytokines Disrupt Epithelial Barrier Function by Apoptosis-Independent Mechanisms. *J Immunol* 2003;**171**:6164-72.

201. Laukoetter MG, Bruewer M, Nusrat A. Regulation of the intestinal epithelial barrier by the apical junctional complex. *Current Opinion in Gastroenterology* 2006;**22**:85-9.

202. Soderholm JD, Streutker C, Yang PC, Paterson C, Singh PK, McKay DM *et al*. Increased epithelial uptake of protein antigens in the ileum of Crohn's disease mediated by tumour necrosis factor alpha. *Gut* 2004;**53**:1817-24.

203. Heller F, Florian P, Bojarski C, Richter J, Christ M, Hillenbrand B *et al*. Interleukin-13 is the key effector Th2 cytokine in ulcerative colitis that affects epithelial tight junctions, apoptosis, and cell restitution. *Gastroenterology* 2005;**129**:550-64.

204. Zeissig S, Bojarski C, Buergel N, Mankertz J, Zeitz M, Fromm M *et al*. Downregulation of epithelial apoptosis and barrier repair in active Crohn's disease by tumour necrosis factor alpha antibody treatment. *Gut* 2004;**53**:1295-302.

205. Madsen KL, Malfair D, Gray D, Doyle JS, Jewell LD, Fedorak RN. Interleukin-10 gene-deficient mice develop a primary intestinal permeability defect in response to enteric microflora. *Inflammatory Bowel Diseases* 1999;**5**:262-70.

206. Kuhn R, Lohler J, Rennick D, Rajewsky K, Muller W. Interleukin-10-deficient mice develop chronic enterocolitis. *Cell* 1993;**75**:263-74.

207. Panwala CM, Jones JC, Viney JL. A novel model of inflammatory bowel disease: mice deficient for the multiple drug resistance gene, mdr1a, spontaneously develop colitis. *Journal of Immunology* 1998;**161**:5733-44.

208. Hermiston ML,.Gordon JI. Inflammatory bowel disease and adenomas in mice expressing a dominant negative N-cadherin. *Science* 1995;**270**:1203-7.

209. Wirtz S,.Neurath MF. Mouse models of inflammatory bowel disease. *Advanced Drug Delivery Reviews* 2007;**59**:1073-83.

210. Nancey S, Holvoet S, Graber I, Joubert G, Philippe D, Martin S *et al*. CD8(+) cytotoxic T cells induce relapsing colitis in normal mice. *Gastroenterology* 2006;**131**:485-96.

211. Westendorf AM, Fleissner D, Deppenmeier S, Gruber AD, Bruder D, Hansen W *et al*. Autoimmune-mediated intestinal inflammation-impact and regulation of antigen-specific CD8(+) T cells. *Gastroenterology* 2006;**131**:510-24.

212. DeMeo MT, Mutlu EA, Keshavarzian A, Tobin MC. Intestinal permeation and gastrointestinal disease. *Journal of Clinical Gastroenterology* 2002;**34**:385-96.

213. May GR, Sutherland LR, Meddings JB. Is small intestinal permeability really increased in relatives of patients with Crohn's disease? *Gastroenterology* 1993;**104**:1627-32.

214. Katz KD, Hollander D, Vadheim CM, McElree C, Delahunty T, Dadufalza VD *et al*. Intestinal permeability in patients with Crohn's disease and their healthy relatives. *Gastroenterology* 1989;**97**:927-31.

215. Teahon K, Smethurst P, Levi AJ, Menzies IS, Bjarnason I. Intestinal permeability in patients with Crohn's disease and their first degree relatives. *Gut* 1992;**33**:320-3.

216. Peeters M, Geypens B, Claus D, Nevens H, Ghoos Y, Verbeke G *et al*. Clustering of increased small intestinal permeability in families with Crohn's disease. *Gastroenterology* 1997;**113**:802-7.

217. Soderholm JD, Peterson KH, Olaison G, Franzen LE, Westrom B, Magnusson KE *et al*. Epithelial permeability to proteins in the noninflamed ileum of Crohn's disease? *Gastroenterology* 1999;**117**:65-72.

218. Soderholm JD, Olaison G, Lindberg E, Hannestad U, Vindels A, Tysk C *et al*. Different intestinal permeability patterns in relatives and spouses of patients with Crohn's disease: an inherited defect in mucosal defence? *Gut* 1999;**44**:96-100.

219. Gitter AH, Wullstein F, Fromm M, Schulzke JD. Epithelial barrier defects in ulcerative colitis: characterization and quantification by electrophysiological imaging. *Gastroenterology* 2001;**121**:1320-8.

220. Thjodleifsson B, Sigthorsson G, Cariglia N, Reynisdottir I, Gudbjartsson DF, Kristjansson K *et al*. Subclinical intestinal inflammation: an inherited abnormality in Crohn's disease relatives? *Gastroenterology* 2003;**124**:1728-37.

221. Buhner S, Buning C, Genschel J, Kling K, Herrmann D, Dignass A *et al*. Genetic basis for increased intestinal permeability in families with Crohn's disease: role of CARD15 3020insC mutation? *Gut* 2006;**55**:342-7.

222. Begue B, Dumant C, Bambou JC, Beaulieu JF, Chamaillard M, Hugot JP *et al*. Microbial induction of CARD15 expression in intestinal epithelial cells via toll-like receptor 5 triggers an antibacterial response loop. *Journal of Cellular Physiology* 2006;241-52.

223. Annese V, Valvano MR, Palmieri O, Latiano A, Bossa F, Andriulli A. Multidrug resistance 1 gene in inflammatory bowel disease: a meta-analysis. *World Journal of Gastroenterology* 2006;**12**:3636-44.

224. Schwab M, Schaeffeler E, Marx C, Fromm MF, Kaskas B, Metzler J *et al*. Association between the C3435T MDR1 gene polymorphism and susceptibility for ulcerative colitis. *Gastroenterology* 2003;**124**:26-33.

225. Glas J, Torok HP, Schiemann U, Folwaczny C. MDR1 gene polymorphism in ulcerative colitis. *Gastroenterology.126(1):367*, 2004;**126**:367.

226. Ho GT, Nimmo ER, Tenesa A, Fennell J, Drummond H, Mowat C *et al*. Allelic variations of the multidrug resistance gene determine susceptibility

and disease behavior in ulcerative colitis. *Gastroenterology* 2005;**128**:288-96.

227. Croucher PJ, Mascheretti S, Foelsch UR, Hampe J, Schreiber S. Lack of association between the C3435T MDR1 gene polymorphism and inflammatory bowel disease in two independent Northern European populations. *Gastroenterology* 2003;**125**:1919-20.

228. Brant SR, Panhuysen CI, Nicolae D, Reddy DM, Bonen DK, Karaliukas R *et al*. MDR1 Ala893 polymorphism is associated with inflammatory bowel disease. *American Journal of Human Genetics* 2003;**73**:1282-92.

229. Potocnik U, Ferkolj I, Glavac D, Dean M. Polymorphisms in multidrug resistance 1 (MDR1) gene are associated with refractory Crohn disease and ulcerative colitis. *Genes & Immunity* 2004;**5**:530-9.

230. Palmieri O, Latiano A, Valvano R, D'Inca R, Vecchi M, Sturniolo GC *et al*. Multidrug resistance 1 gene polymorphisms are not associated with inflammatory bowel disease and response to therapy in Italian patients. *Alimentary Pharmacology & Therapeutics* 2005;**22**:1129-38.

231. Urcelay E, Mendoza JL, Martin MC, Mas A, Martinez A, Taxonera C *et al*. MDR1 gene: susceptibility in Spanish Crohn's disease and ulcerative colitis patients. *Inflammatory Bowel Diseases* 2006;**12**:33-7.

232. Oostenbrug LE, Dijkstra G, Nolte IM, van Dullemen HM, Oosterom E, Faber KN *et al*. Absence of association between the multidrug resistance (MDR1) gene and inflammatory bowel disease. *Scandinavian Journal of Gastroenterology* 2006;**41**:1174-82.

233. Onnie CM, Fisher SA, Pattni R, Sanderson J, Forbes A, Lewis CM *et al*. Associations of allelic variants of the multidrug resistance gene (ABCB1 or MDR1) and inflammatory bowel disease and their effects on disease behavior: a case-control and meta-analysis study. *Inflammatory Bowel Diseases* 2006;**12**:263-71.

234. Ho GT, Soranzo N, Nimmo ER, Tenesa A, Goldstein DB, Satsangi J. ABCB1/MDR1 gene determines susceptibility and phenotype in ulcerative colitis: discrimination of critical variants using a gene-wide haplotype tagging approach. *Hum.Mol.Genet.* 2006;**15**:797-805.

235. Langmann T,.Schmitz G. Loss of detoxification in inflammatory bowel disease. *Nature Clinical Practice Gastroenterology & Hepatology* 2006;**3**:358-9.

236. Langmann T, Moehle C, Mauerer R, Scharl M, Liebisch G, Zahn A *et al*. Loss of detoxification in inflammatory bowel disease: dysregulation of pregnane X receptor target genes. *Gastroenterology* 2004;**127**:26-40.

237. Staudinger JL, Goodwin B, Jones SA, Hawkins-Brown D, MacKenzie KI, LaTour A *et al*. The nuclear receptor PXR is a lithocholic acid sensor that protects against liver toxicity. *Proceedings of the National Academy of Sciences of the United States of America* 2001;**98**:3369-74.

238. Chrencik JE, Orans J, Moore LB, Xue Y, Peng L, Collins JL *et al*. Structural disorder in the complex of human pregnane X receptor and the macrolide antibiotic rifampicin. *Molecular Endocrinology* 2005;1125-34.

239. Schuetz EG, Strom S, Yasuda K, Lecureur V, Assem M, Brimer C *et al*. Disrupted bile acid homeostasis reveals an unexpected interaction among nuclear hormone receptors, transporters, and cytochrome P450. *J.Biol.Chem.* 2001;**276**:39411-8.

240. Willson TM,.Kliewer SA. PXR, CAR and drug metabolism. *Nature Reviews.Drug Discovery* 2002;**1**:259-66.

241. Kliewer SA, Goodwin B, Willson TM. The nuclear pregnane X receptor: a key regulator of xenobiotic metabolism. *Endocrine Reviews* 2002;**23**:687-702.

242. Dring MM, Goulding CA, Trimble VI, Keegan D, Ryan AW, Brophy KM *et al*. The Pregnane X Receptor Locus Is Associated With Susceptibility to Inflammatory Bowel Disease. *Gastroenterology* 2006;**130**:341-8.

243. Ho GT, Soranzo N, Tate SK, Drummond H, Nimmo ER, Tenesa A *et al*. Lack of association of the pregnane X receptor (PXR/NR1I2) gene with inflammatory bowel disease: parallel allelic association study and gene wide haplotype analysis. *Gut* 2006;**55**:1676-7.

244. Stoll M, Corneliussen B, Costello CM, Waetzig GH, Mellgard B, Koch WA *et al*. Genetic variation in DLG5 is associated with inflammatory bowel disease. *Nature Genetics* 2004;**36**:476-80.

245. Hampe J, Schreiber S, Shaw SH, Lau KF, Bridger S, MacPherson AJ *et al*. A genomewide analysis provides evidence for novel linkages in inflammatory bowel disease in a large European cohort. *American Journal of Human Genetics* 1999;**64**:808-16.

246. Vermeire S. Dlg5 and octn. *Inflammatory Bowel Diseases* 2004;**10**:888-90.

247. Shah G, Brugada R, Gonzalez O, Czernuszewicz G, Gibbs RA, Bachinski L *et al*. The cloning, genomic organization and tissue expression profile of the human DLG5 gene. *BMC Genomics.3(1):6,* 2002;**3**:6.

248. Noble CL, Nimmo ER, Drummond H, Smith L, Arnott IDR, Satsangi J. DLG5 variants do not influence susceptibility to inflammatory bowel disease in the Scottish population. *Gut* 2005;**54**:1416-20.

249. Russell RK, Drummond HE, Nimmo ER, Anderson N, Smith L, Wilson DC *et al*. Socio-economic status critically influences penetrance of the DLG5 113A variant in early-onset IBD: A novel gene-environmental interaction. *Gastroenterology* 2006;**130**:A53.

250. Tenesa A, Noble C, Satsangi J, Dunlop M. Association of DLG5 and inflammatory bowel disease across populations. *European Journal of Human Genetics* 2006;**14**:259-60.

251. Torok HP, Glas J, Tonenchi L, Lohse P, Muller-Myhsok B, Limbersky O *et al*. Polymorphisms in the DLG5 and OCTN cation transporter genes in Crohn's disease. *Gut* 2005;**54**:1421-7.

252. Lakatos PL, Fischer S, Claes K, Kovacs A, Molnar T, Altorjay I *et al*. DLG5 R30Q is not associated with IBD in Hungarian IBD patients but predicts clinical response to steroids in Crohn's disease. *Inflammatory Bowel Diseases* 2006;**12**:362-8.

253. Ferraris A, Torres B, Knafelz D, Barabino A, Lionetti P, de Angelis GL *et al*. Relationship between CARD15, SLC22A4/5, and DLG5 polymorphisms and early-onset inflammatory bowel diseases: an Italian multicentric study. *Inflammatory Bowel Diseases* 2006;**12**:355-61.

254. Tremelling M, Waller S, Bredin F, Greenfield S, Parkes M. Genetic variants in TNF-alpha but not DLG5 are associated with inflammatory bowel disease in a large United Kingdom cohort. *Inflammatory Bowel Diseases* 2006;**12**:178-84.

255. Buning C, Geerdts L, Fiedler T, Gentz E, Pitre G, Reuter W *et al*. DLG5 variants in inflammatory bowel disease. *American Journal of Gastroenterology* 2006;**101**:786-92.

256. Medici V, Mascheretti S, Croucher PJ, Stoll M, Hampe J, Grebe J *et al*. Extreme heterogeneity in CARD15 and DLG5 Crohn disease-associated polymorphisms between German and Norwegian populations. *European Journal of Human Genetics* 2006;**14**:459-68.

257. Newman WG, Gu X, Wintle RF, Liu X, Van Oene M, Amos CI *et al*. DLG5 variants contribute to Crohn disease risk in a Canadian population. *Human Mutation* 2006;**27**:353-8.

258. Friedrichs F, Brescianini S, Annese V, Latiano A, Berger K, Kugathasan S *et al*. Evidence of transmission ratio distortion of DLG5 R30Q variant in general and implication of an association with Crohn disease in men. *Human Genetics* 2006;**119**:305-11.

259. Gazouli M, Mantzaris G, Archimandritis AJ, Nasioulas G, Anagnou NP. Single nucleotide polymorphisms of OCTN1, OCTN2, and DLG5 genes in

Greek patients with Crohn's disease. *World Journal of Gastroenterology* 2005;**11**:7525-30.

260. Vermeire S, Pierik M, Hlavaty T, Claessens G, van Schuerbeeck N, Joossens S *et al*. Association of organic cation transporter risk haplotype with perianal penetrating Crohn's disease but not with susceptibility to IBD. *Gastroenterology* 2005;**129**:1845-53.

261. Daly MJ, Pearce AV, Farwell L, Fisher SA, Latiano A, Prescott NJ *et al*. Association of DLG5 R30Q variant with inflammatory bowel disease. *European Journal of Human Genetics* 2005;**13**:835-9.

262. Yamazaki K, Takazoe M, Tanaka T, Ichimori T, Saito S, Iida A *et al*. Association analysis of SLC22A4, SLC22A5 and DLG5 in Japanese patients with Crohn disease. *Journal of Human Genetics* 2004;**49**:664-8.

263. Biank V, Friedrichs F, Babusukumar U, Wang T, Stoll M, Broeckel U *et al*. DLG5 R30Q variant is a female-specific protective factor in pediatric onset Crohn's disease. *American Journal of Gastroenterology* 2007;**102**:391-8.

264. Pearce AV, Fisher SA, Prescott NJ, Onnie CM, Pattni R, Green P *et al*. Investigation of association of the DLG5 gene with phenotypes of inflammatory bowel disease in the British population. *Int.J Colorectal Dis.* 2007;**22**:419-24.

265. Browning BL, Huebner C, Petermann I, Demmers P, McCulloch A, Gearry RB *et al*. Association of DLG5 variants with inflammatory bowel disease in the New Zealand caucasian population and meta-analysis of the DLG5 R30Q variant. *Inflammatory Bowel Diseases* 2007;**13**:1069-76.

266. Russell RK, Drummond HE, Nimmo ER, Anderson N, Wilson DC, Gillett PM *et al*. The contribution of the DLG5 113A variant in early-onset inflammatory bowel disease. *Journal of Pediatrics* 2007;**150**:268-73.

267. Rioux JD, Silverberg MS, Daly MJ, Steinhart AH, McLeod RS, Griffiths AM *et al*. Genomewide search in Canadian families with inflammatory bowel disease reveals two novel susceptibility loci. *American Journal of Human Genetics* 2000;**66**:1863-70.

268. Rioux JD, Daly MJ, Silverberg MS, Lindblad K, Steinhart H, Cohen Z *et al*. Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. *Nature Genetics* 2001;**29**:223-8.

269. Reinhard C,.Rioux JD. Role of the IBD5 susceptibility locus in the inflammatory bowel diseases. *Inflammatory Bowel Diseases* 2006;**12**:227-38.

270. Loots GG, Locksley RM, Blankespoor CM, Wang ZE, Miller W, Rubin EM *et al*. Identification of a coordinate regulator of interleukins 4, 13, and 5 by cross-species sequence comparisons. *Science* 2000;**288**:136-40.

271. Noble CL, Nimmo ER, Drummond H, Ho G-T, Tenesa A, Smith L *et al*. The contribution of OCTN1/2 variants within the IBD5 locus to disease susceptibility and severity in Crohn's disease. *Gastroenterology* 2005;**129**:1854-64.

272. Giallourakis C, Stoll M, Miller K, Hampe J, Lander ES, Daly MJ *et al*. IBD5 is a general risk factor for inflammatory bowel disease: replication of association with Crohn disease and identification of a novel association with ulcerative colitis. *American Journal of Human Genetics* 2003;**73**:205-11.

273. Mirza MM, Fisher SA, King K, Cuthbert AP, Hampe J, Sanderson J *et al*. Genetic evidence for interaction of the 5q31 cytokine locus and the CARD15 gene in Crohn disease. *American Journal of Human Genetics* 2003;**72**:1018-22.

274. Negoro K, McGovern DP, Kinouchi Y, Takahashi S, Lench NJ, Shimosegawa T *et al*. Analysis of the IBD5 locus and potential gene-gene interactions in Crohn's disease. *Gut* 2003;**52**:541-6.

275. Armuzzi A, Ahmad T, Ling KL, de Silva A, Cullen S, van Heel D *et al*. Genotype-phenotype analysis of the Crohn's disease susceptibility haplotype on chromosome 5q31. *Gut* 2003;**52**:1133-9.

276. Urcelay E, Mendoza JL, Martinez A, Fernandez L, Taxonera C, Diaz-Rubio M *et al*. IBD5 polymorphisms in inflammatory bowel disease: association with response to infliximab. *World Journal of Gastroenterology* 2005;**11**:1187-92.

277. Newman B, Gu X, Wintle R, Cescon D, Yazdanpanah M, Liu X *et al*. A risk haplotype in the Solute Carrier Family 22A4/22A5 gene cluster influences phenotypic expression of Crohn's disease. *Gastroenterology* 2005;**128**:260-9.

278. Russell RK, Drummond HE, Nimmo ER, Anderson NH, Noble CL, Wilson DC *et al*. Analysis of the influence of OCTN1/2 variants within the IBD5 locus on disease susceptibility and growth indices in early onset inflammatory bowel disease. *Gut* 2006;**55**:1114-23.

279. Babusukumar U, Wang T, McGuire E, Broeckel U, Kugathasan S. Contribution of OCTN variants within the IBD5 locus to pediatric onset Crohn's disease. *American Journal of Gastroenterology* 2006;**101**:1354-61.

280. McGovern DP, van Heel DA, Negoro K, Ahmad T, Jewell DP. Further evidence of IBD5/CARD15 (NOD2) epistasis in the susceptibility to ulcerative colitis. *American Journal of Human Genetics* 2003;**73**:1465-6.

281. Peltekova VD, Wintle RF, Rubin LA, Amos CI, Huang QQ, Gu XJ *et al*. Functional variants of OCTN cation transporter genes are associated with Crohn disease. *Nature Genetics* 2004;**36**:471-5.

282. Fisher SA, Hampe J, Onnie CM, Daly MJ, Curley C, Purcell S *et al*. Direct or indirect association in a complex disease: the role of SLC22A4 and SLC22A5 functional variants in Crohn disease. *Human Mutation* 2006;**27**:778-85.

283. Waller S, Tremelling M, Bredin F, Godfrey L, Howson J, Parkes M. Evidence for association of OCTN genes and IBD5 with ulcerative colitis. *Gut.55(6):809-14,* 2006.

284. Torkvist L, Noble CL, Lordal M, Sjoqvist U, Lindforss U, Nimmo ER *et al*. Contribution of the IBD5 locus to Crohn's disease in the Swedish population. *Scandinavian Journal of Gastroenterology* 2007;**42**:200-6.

285. Silverberg MS, Duerr RH, Brant SR, Bromfield G, Datta LW, Jani N *et al*. Refined genomic localization and ethnic differences observed for the IBD5 association with Crohn's disease. *Eur J Hum Genet* 2007;**15**:328-35.

286. Tosa M, Negoro K, Kinouchi Y, Abe H, Nomura E, Takagi S *et al*. Lack of association between IBD5 and Crohn's disease in Japanese patients demonstrates population-specific differences in inflammatory bowel disease. *Scandinavian Journal of Gastroenterology* 2006;**41**:48-53.

287. Bene J, Magyari L, Talian G, Komlosi K, Gasztonyi B, Tari B *et al*. Prevalence of SLC22A4, SLC22A5 and CARD15 gene mutations in Hungarian pediatric patients with Crohn's disease. *World Journal of Gastroenterology* 2006;**12**:5550-3.

288. Cucchiara S, Latiano A, Palmieri O, Staiano AM, D'Inca R, Guariso G *et al*. Role of CARD15, DLG5 and OCTN genes polymorphisms in children with inflammatory bowel diseases. *World Journal of Gastroenterology* 2007;**13**:1221-9.

289. Latiano A, Palmieri O, Valvano RM, D'Inca R, Vecchi M, Ferraris A *et al*. Contribution of IBD5 locus to clinical features of IBD patients. *American Journal of Gastroenterology* 2006;**101**:318-25.

290. Palmieri O, Latiano A, Valvano R, D'Inca R, Vecchi M, Sturniolo GC *et al*. Variants of OCTN1-2 cation transporter genes are associated with both Crohn's disease and ulcerative colitis. *Alimentary Pharmacology & Therapeutics* 2006;**23**:497-506.

291. Martinez A, Del Carmen MM, Mendoza JL, Taxonera C, Diaz-Rubio M, de la Concha EG *et al*. Association of the organic cation transporter OCTN genes with Crohn's disease in the Spanish population. *European Journal of Human Genetics* 2006;**14**:222-6.

292. van Bodegraven AA, Curley CR, Hunt KA, Monsuur AJ, Linskens RK, Onnie CM *et al*. Genetic variation in myosin IXB is associated with ulcerative colitis. *Gastroenterology* 2006;**131**:1768-74.

293. Nunez C, Oliver J, Mendoza JL, Gomez-Garcia M, Pinero A, Taxonera C *et al*. MYO9B polymorphisms in patients with inflammatory bowel disease. *Gut* 2007;**56**:1321-2.

294. Latiano A, Palmieri O, Valvano MR, DΓÇÖINCA R, CAPRILLI R, Cucchiara S *et al*. The association of MYO9B gene in Italian patients with inflammatory bowel diseases. *Alimentary Pharmacology & Therapeutics* 2008;**27**:241-8.

295. Monsuur AJ, de Bakker PI, Alizadeh BZ, Zhernakova A, Bevova MR, Strengman E *et al*. Myosin IXB variant increases the risk of celiac disease and points toward a primary intestinal barrier defect. *Nature Genetics* 2005;**37**:1341-4.

296. Amundsen SS, Monsuur AJ, Wapenaar MC, Lie BA, Ek J, Gudjonsdottir AH *et al*. Association analysis of MYO9B gene polymorphisms with celiac disease in a Swedish/Norwegian cohort. *Human Immunology* 2006;**67**:341-5.

297. Hunt KA, Monsuur AJ, McArdle WL, Kumar PJ, Travis SP, Walters JR *et al*. Lack of association of MYO9B genetic variants with coeliac disease in a British cohort. *Gut* 2006;**55**:969-72.

298. Akira S, Uematsu S, Takeuchi O. Pathogen recognition and innate immunity. *Cell* 2006;**124**:783-801.

299. Seong SY,.Matzinger P. Hydrophobicity: an ancient damage-associated molecular pattern that initiates innate immune responses. *Nat Rev Immunol* 2004;**4**:469-78.

300. Meylan E, Tschopp J, Karin M. Intracellular pattern recognition receptors in the host response. *Nature* 2006;**442**:39-44.

301. Takeda A, Hamano S, Yamanaka A, Hanada T, Ishibashi T, Mak TW *et al*. Cutting Edge: Role of IL-27/WSX-1 Signaling for Induction of T-Bet Through Activation of STAT1 During Initial Th1 Commitment. *J Immunol* 2003;**170**:4886-90.

302. Sugiyama T,.Wright SD. Soluble CD14 mediates efflux of phospholipids from cells. *Journal of Immunology* 2001;**166**:826-31.

303. Pierik M, Joossens S, Van Steen K, Van Schuerbeek N, Vlietinck R, Rutgeerts P *et al*. Toll-like receptor-1, -2, and -6 polymorphisms influence disease extension in inflammatory bowel diseases. *Inflammatory Bowel Diseases* 2006;**12**:1-8.

304. Rector A, Lemey P, Laffut W, Keyaerts E, Struyf F, Wollants E *et al*. Mannan-binding lectin (MBL) gene polymorphisms in ulcerative colitis and Crohn's disease. *Genes & Immunity* 2001;**2**:323-8.

305. Seibold F, Konrad A, Flogerzi B, Seibold-Schmid B, Arni S, Juliger S *et al*. Genetic variants of the mannan-binding lectin are associated with immune reactivity to mannans in Crohn's disease. *Gastroenterology* 2004;**127**:1076-84.

306. Joossens S, Pierik M, Rector A, Vermeire S, Ranst MV, Rutgeerts P *et al*. Mannan binding lectin (MBL) gene polymorphisms are not associated with anti-Saccharomyces cerevisiae (ASCA) in patients with Crohn's disease. *Gut* 2006;**55**:746.

307. Creagh EM,.O'Neill LAJ. TLRs, NLRs and RLRs: a trinity of pathogen sensors that co-operate in innate immunity. *Trends in Immunology* 2006;**27**:352-7.

308. Shi Y, Evans JE, Rock KL. Molecular identification of a danger signal that alerts the immune system to dying cells. *Nature* 2003;**425**:516-21.

309. Drenth JP,.van der Meer JW. The inflammasome--a linebacker of innate defense. *New England Journal of Medicine* 2006;**355**:730-2.

310. Martinon F, Petrilli V, Mayor A, Tardivel A, Tschopp J. Gout-associated uric acid crystals activate the NALP3 inflammasome. *Nature* 2006;**440**:237-41.

311. Ting JP, Kastner DL, Hoffman HM. CATERPILLERs, pyrin and hereditary immunological disorders. *Nature Reviews.Immunology* 2006;**6**:183-95.

312. Anderson KV, Bokla L, Nusslein-Volhard C. Establishment of dorsal-ventral polarity in the Drosophila embryo: the induction of polarity by the Toll gene product. *Cell* 1985;**42**:791-8.

313. Medzhitov R, Preston-Hurlburt P, Janeway CA, Jr. A human homologue of the Drosophila Toll protein signals activation of adaptive immunity. *Nature* 1997;**388**:394-7.

314. West AP, Koblansky AA, Ghosh S. Recognition and Signaling by Toll-Like Receptors. *Annual Review of Cell and Developmental Biology* 2006;**22**:409-37.

315. Poltorak A, He X, Smirnova I, Liu MY, Van Huffel C, Du X *et al*. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science* 1998;**282**:2085-8.

316. Hoshino K, Takeuchi O, Kawai T, Sanjo H, Ogawa T, Takeda Y *et al*. Cutting edge: Toll-like receptor 4 (TLR4)-deficient mice are hyporesponsive to lipopolysaccharide: evidence for TLR4 as the Lps gene product. *Journal of Immunology* 1999;**162**:3749-52.

317. Mitchell JA, Fitzgerald KA, Coyle A, Silverman N, Cartwright N. TOLLing away in Brazil. *Nature Immunology* 2006;**7**:675-9.

318. Bowie A,.O'neill LA. The interleukin-1 receptor/Toll-like receptor superfamily: signal generators for pro-inflammatory interleukins and microbial products. *Journal of Leukocyte Biology* 2000;**67**:508-14.

319. Stover AG, Da Silva CJ, Evans JT, Cluff CW, Elliott MW, Jeffery EW *et al*. Structure-activity relationship of synthetic toll-like receptor 4 agonists. *J.Biol.Chem.* 2004;**279**:4440-9.

320. Hristova K, Selsted ME, White SH. Critical role of lipid composition in membrane permeabilization by rabbit neutrophil defensins. *J.Biol.Chem.* 1997;**272**:24224-33.

321. Koka RM, Huang E, Lieske JC. Adhesion of uric acid crystals to the surface of renal epithelial cells. *American Journal of Physiology - Renal Physiology* 2000;**278**:F989-F998.

322. Ozinsky A, Underhill DM, Fontenot JD, Hajjar AM, Smith KD, Wilson CB *et al*. The repertoire for pattern recognition of pathogens by the innate immune system is defined by cooperation between toll-like receptors. *Proceedings of the National Academy of Sciences of the United States of America* 2000;**97**:13766-71.

323. Takeuchi O, Kawai T, Muhlradt PF, Morr M, Radolf JD, Zychlinsky A *et al*. Discrimination of bacterial lipoproteins by Toll-like receptor 6. *Int.Immunol.* 2001;**13**:933-40.

324. Takeuchi O, Sato S, Horiuchi T, Hoshino K, Takeda K, Dong Z *et al*. Cutting edge: role of Toll-like receptor 1 in mediating immune response to microbial lipoproteins. *Journal of Immunology* 2002;**169**:10-4.

325. Travassos LH, Girardin SE, Philpott DJ, Blanot D, Nahori MA, Werts C *et al*. Toll-like receptor 2-dependent bacterial sensing does not occur via peptidoglycan recognition. *EMBO Reports* 2004;**5**:1000-6.

326. Dziarski R,.Gupta D. Staphylococcus aureus peptidoglycan is a toll-like receptor 2 activator: a reevaluation. *Infection & Immunity* 2005;**73**:5212-6.

327. Werts C, Tapping RI, Mathison JC, Chuang TH, Kravchenko V, Saint G, I *et al*. Leptospiral lipopolysaccharide activates cells through a TLR2-dependent mechanism. *Nature Immunology* 2001;**2**:346-52.

328. Asai Y, Hashimoto M, Fletcher HM, Miyake K, Akira S, Ogawa T. Lipopolysaccharide preparation extracted from Porphyromonas gingivalis lipoprotein-deficient mutant shows a marked decrease in toll-like receptor 2-mediated signaling. *Infection & Immunity* 2005;**73**:2157-63.

329. Choe J, Kelker MS, Wilson IA. Crystal structure of human toll-like receptor 3 (TLR3) ectodomain. *Science* 2005;**309**:581-5.

330. Bell JK, Askins J, Hall PR, Davies DR, Segal DM. The dsRNA binding site of human Toll-like receptor 3. *Proceedings of the National Academy of Sciences of the United States of America* 2006;**103**:8792-7.

331. Cario E. Bacterial interactions with cells of the intestinal mucosa: Toll-like receptors and NOD2. *Gut* 2005;**54**:1182-93.

332. Liew FY, Xu D, Brint EK, O'neill LA. Negative regulation of toll-like receptor-mediated immune responses. *Nature Reviews.Immunology* 2005;**5**:446-58.

333. Gazouli M, Mantzaris G, Kotsinas A, Zacharatos P, Papalambros E, Archimandritis A *et al*. Association between polymorphisms in the Toll-like receptor 4, CD14, and CARD15/NOD2 and inflammatory bowel disease in the Greek population. *World Journal of Gastroenterology* 2005;**11**:681-5.

334. Klein W, Tromm A, Griga T, Fricke H, Folwaczny C, Hocke M *et al*. A polymorphism in the CD14 gene is associated with Crohn disease. *Scandinavian Journal of Gastroenterology* 2002;**37**:189-91.

335. Obana N, Takahashi S, Kinouchi Y, Negoro K, Takagi S, Hiwatashi N *et al*. Ulcerative colitis is associated with a promoter polymorphism of lipopolysaccharide receptor gene, CD14. *Scandinavian Journal of Gastroenterology* 2002;**37**:699-704.

336. Peters KE, O'Callaghan NJ, Cavanaugh JA. Lack of association of the CD14 promoter polymorphism--159C/T with Caucasian inflammatory bowel disease. *Scandinavian Journal of Gastroenterology* 2005;**40**:194-7.

337. Arnott IDR, Nimmo ER, Drummond HE, Fennell J, Smith BRK, MacKinlay E *et al*. NOD2/CARD15, TLR4 and CD14 mutations in Scottish and Irish Crohn's disease patients: Evidence for genetic heterogeneity within Europe? *Genes & Immunity* 2004;**5**:417-25.

338. Hayashi F, Smith KD, Ozinsky A, Hawn TR, Yi EC, Goodlett DR *et al*. The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. *Nature* 2001;**410**:1099-103.

339. Smith KD, Andersen-Nissen E, Hayashi F, Strobe K, Bergman MA, Barrett SL *et al*. Toll-like receptor 5 recognizes a conserved site on flagellin required for protofilament formation and bacterial motility. *Nature Immunology* 2003;**4**:1247-53.

340. Andersen-Nissen E, Smith KD, Strobe KL, Barrett SL, Cookson BT, Logan SM *et al*. Evasion of Toll-like receptor 5 by flagellated bacteria. *Proceedings of the National Academy of Sciences of the United States of America* 2005;**102**:9247-52.

341. Hemmi H, Takeuchi O, Kawai T, Kaisho T, Sato S, Sanjo H *et al*. A Toll-like receptor recognizes bacterial DNA. *Nature* 2000;**408**:740-5.

342. Barton GM, Kagan JC, Medzhitov R. Intracellular localization of Toll-like receptor 9 prevents recognition of self DNA but facilitates access to viral DNA. *Nature Immunology* 2006;**7**:49-56.

343. Fusunyan R, Nanthakumar N, Baldeon M, Walker W. Evidence for an Innate Immune Response in the Immature Human Intestine: Toll-Like Receptors on Fetal Enterocytes. *Pediatr Res* 2001;**49**:589-93.

344. Hausmann M, Kiessling S, Mestermann S, Webb G, Spottl T, Andus T *et al*. Toll-like receptors 2 and 4 are up-regulated during intestinal inflammation. *Gastroenterology* 2002;**122**:1987-2000.

345. Abreu MT, Vora P, Faure E, Thomas LS, Arnold ET, Arditi M. Decreased expression of Toll-like receptor-4 and MD-2 correlates with intestinal epithelial cell protection against dysregulated proinflammatory gene expression in response to bacterial lipopolysaccharide. *Journal of Immunology* 2001;**167**:1609-16.

346. Cario E,.Podolsky DK. Differential Alteration in Intestinal Epithelial Cell Expression of Toll-Like Receptor 3 (TLR3) and TLR4 in Inflammatory Bowel Disease. *Infect.Immun.* 2000;**68**:7010-7.

347. Gewirtz AT, Navas TA, Lyons S, Godowski PJ, Madara JL. Cutting Edge: Bacterial Flagellin Activates Basolaterally Expressed TLR5 to Induce Epithelial Proinflammatory Gene Expression. *J Immunol* 2001;**167**:1882-5.

348. Otte JM, Cario E, Podolsky DK. Mechanisms of cross hyporesponsiveness to toll-like receptor bacterial ligands in intestinal epithelial cells. *Gastroenterology* 2004;**126**:1054-70.

349. Bambou JC, Giraud A, Menard S, Begue B, Rakotobe S, Heyman M *et al*. In vitro and ex vivo activation of the TLR5 signaling pathway in intestinal

epithelial cells by a commensal Escherichia coli strain. *J.Biol.Chem.* 2004;**279**:42984-92.

350. Sanders CJ, Yu Y, Moore DA, III, Williams IR, Gewirtz AT. Humoral Immune Response to Flagellin Requires T Cells and Activation of Innate Immunity. *J Immunol* 2006;**177**:2810-8.

351. Seya T, Oshiumi H, Sasai M, Akazawa T, Matsumoto M. TICAM-1 and TICAM-2: toll-like receptor adapters that participate in induction of type 1 interferons. *International Journal of Biochemistry & Cell Biology* 2005;**37**:524-9.

352. Yamamoto M, Sato S, Hemmi H, Sanjo H, Uematsu S, Kaisho T *et al*. Essential role for TIRAP in activation of the signalling cascade shared by TLR2 and TLR4. *Nature.420(6913):324-9,* 2002.

353. Yamamoto M, Sato S, Mori K, Hoshino K, Takeuchi O, Takeda K *et al*. Cutting edge: a novel Toll/IL-1 receptor domain-containing adapter that preferentially activates the IFN-beta promoter in the Toll-like receptor signaling. *Journal of Immunology* 2002;**169**:6668-72.

354. Carty M, Goodbody R, Schroder M, Stack J, Moynagh PN, Bowie AG. The human adaptor SARM negatively regulates adaptor protein TRIF-dependent Toll-like receptor signaling. *Nat Immunol* 2006;**7**:1074-81.

355. Takaoka A, Yanai H, Kondo S, Duncan G, Negishi H, Mizutani T *et al*. Integral role of IRF-5 in the gene induction programme activated by Toll-like receptors. *Nature* 2005;**434**:243-9.

356. Honda K, Ohba Y, Yanai H, Negishi H, Mizutani T, Takaoka A *et al*. Spatiotemporal regulation of MyD88-IRF-7 signalling for robust type-I interferon induction. *Nature* 2005;**434**:1035-40.

357. Yamamoto M, Yamazaki S, Uematsu S, Sato S, Hemmi H, Hoshino K *et al*. Regulation of Toll/IL-1-receptor-mediated gene expression by the inducible nuclear protein IkappaBzeta. *Nature* 2004;**430**:218-22.

358. Rakoff-Nahoum S, Paglino J, Eslami-Varzaneh F, Edberg S, Medzhitov R. Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis. *Cell* 2004;**118**:229-41.

359. Strober W. Epithelial cells pay a Toll for protection. *Nature Medicine* 2004;**10**:898-900.

360. Fort MM, Mozaffarian A, Stover AG, Correia JdS, Johnson DA, Crane RT *et al*. A Synthetic TLR4 Antagonist Has Anti-Inflammatory Effects in Two Murine Models of Inflammatory Bowel Disease. *J Immunol* 2005;**174**:6416-23.

361. Iwami Ki, Matsuguchi T, Masuda A, Kikuchi T, Musikacharoen T, Yoshikai Y. Cutting Edge: Naturally Occurring Soluble Form of Mouse Toll-Like Receptor 4 Inhibits Lipopolysaccharide Signaling. *J Immunol* 2000;**165**:6682-6.

362. LeBouder E, Rey-Nores JE, Rushmere NK, Grigorov M, Lawn SD, Affolter M *et al*. Soluble Forms of Toll-Like Receptor (TLR)2 Capable of Modulating TLR2 Signaling Are Present in Human Plasma and Breast Milk. *J Immunol* 2003;**171**:6680-9.

363. Wald D, Qin J, Zhao Z, Qian Y, Naramura M, Tian L *et al*. SIGIRR, a negative regulator of Toll-like receptor-interleukin 1 receptor signaling. *Nature Immunology* 2003;**4**:920-7.

364. Garlanda C, Riva F, Polentarutti N, Buracchi C, Sironi M, De Bortoli M *et al*. Intestinal inflammation in mice deficient in Tir8, an inhibitory member of the IL-1 receptor family. *Proceedings of the National Academy of Sciences of the United States of America* 2004;**101**:3522-6.

365. Qin J, Qian Y, Yao J, Grace C, Li X. SIGIRR inhibits interleukin-1 receptor- and toll-like receptor 4-mediated signaling through different mechanisms. *J.Biol.Chem.* 2005;**280**:25233-41.

366. Xiao H, Gulen MF, Qin J, Yao J, Bulek K, Kish D *et al*. The Toll-interleukin-1 receptor member SIGIRR regulates colonic epithelial homeostasis, inflammation, and tumorigenesis. *Immunity* 2007;**26**:461-75.

367. Adib-Conquy MP, Adrie CM, Fitting CB, Gattolliat OM, Beyaert RP, Cavaillon JMD. Up-regulation of MyD88s and SIGIRR, molecules inhibiting Toll-like receptor signaling, in monocytes from septic patients. *Critical Care Medicine* 2006;**34**:2377-85.

368. Thomassen E, Renshaw BR, Sims JE. Identification and characterization of SIGIRR, a molecule representing a novel subtype of the IL-1R superfamily. *Cytokine* 1999;**11**:389-99.

369. Alexander WS,.Hilton DJ. The role of suppressors of cytokine signaling (SOCS) proteins in regulation of the immune response. *Annual Review of Immunology* 2004;**22**:503-29.

370. Kinjyo I, Hanada T, Inagaki-Ohara K, Mori H, Aki D, Ohishi M *et al*. SOCS1/JAB is a negative regulator of LPS-induced macrophage activation. *Immunity* 2002;**17**:583-91.

371. Nakagawa R, Naka T, Tsutsui H, Fujimoto M, Kimura A, Abe T *et al*. SOCS-1 participates in negative regulation of LPS responses. *Immunity* 2002;**17**:677-87.

372. Chinen T, Kobayashi T, Ogata H, Takaesu G, Takaki H, Hashimoto M *et al*. Suppressor of cytokine signaling-1 regulates inflammatory bowel disease in which both IFNgamma and IL-4 are involved. *Gastroenterology* 2006;**130**:373-88.

373. Inagaki-Ohara K, Sasaki A, Matsuzaki G, Ikeda T, Hotokezaka M, Chijiiwa K *et al*. Suppressor of cytokine signalling 1 in lymphocytes regulates the development of intestinal inflammation in mice. *Gut* 2006;**55**:212-9.

374. Hugot JP, Laurent-Puig P, Gower-Rousseau C, Olson JM, Lee JC, Beaugerie L *et al*. Mapping of a susceptibility locus for Crohn's disease on chromosome 16. *Nature* 1996;**379**:821-3.

375. Cho JH, Nicolae DL, Gold LH, Fields CT, LaBuda MC, Rohal PM *et al*. Identification of novel susceptibility loci for inflammatory bowel disease on chromosomes 1p, 3q, and 4q: evidence for epistasis between 1p and IBD1. *Proceedings of the National Academy of Sciences of the United States of America* 1998;**95**:7502-7.

376. Melmed G, Thomas LS, Lee N, Tesfay SY, Lukasek K, Michelsen KS *et al*. Human intestinal epithelial cells are broadly unresponsive to Toll-like receptor 2-dependent bacterial ligands: implications for host-microbial interactions in the gut. *Journal of Immunology* 2003;**170**:1406-15.

377. Burns K, Clatworthy J, Martin L, Martinon F, Plumpton C, Maschera B *et al*. Tollip, a new component of the IL-1RI pathway, links IRAK to the IL-1 receptor. *Nature Cell Biology* 2000;**2**:346-51.

378. Bulut Y, Faure E, Thomas L, Equils O, Arditi M. Cooperation of Toll-like receptor 2 and 6 for cellular activation by soluble tuberculosis factor and Borrelia burgdorferi outer surface protein A lipoprotein: role of Toll-interacting protein and IL-1 receptor signaling molecules in Toll-like receptor 2 signaling. *Journal of Immunology* 2001;**167**:987-94.

379. Zhang G,.Ghosh S. Negative regulation of toll-like receptor-mediated signaling by Tollip. *J.Biol.Chem.* 2002;**277**:7059-65.

380. Li T, Hu J, Li L. Characterization of Tollip protein upon Lipopolysaccharide challenge. *Molecular Immunology* 2004;**41**:85-92.

381. Williams CN, Kocher K, Lander ES, Daly MJ, Rioux JD. Using a genome-wide scan and meta-analysis to identify a novel IBD locus and confirm previously identified IBD loci. *Inflammatory Bowel Diseases* 2002;**8**:375-81.

382. Paavola-Sakki P, Ollikainen V, Helio T, Halme L, Turunen U, Lahermo P *et al*. Genome-wide search in Finnish families with inflammatory bowel disease provides evidence for novel susceptibility loci. *European Journal of Human Genetics* 2003;**11**:112-20.

383. Watanabe T, Kitani A, Murray PJ, Strober W. NOD2 is a negative regulator of Toll-like receptor 2-mediated T helper type 1 responses. *Nature Immunology* 2004;**5**:800-8.

384. Maeda S, Hsu LC, Liu H, Bankston LA, Iimura M, Kagnoff MF *et al*. Nod2 mutation in Crohn's disease potentiates NF-kappaB activity and IL-1beta processing. *Science* 2005;**307**:734-8.

385. Kobayashi KS, Chamaillard M, Ogura Y, Henegariu O, Inohara N, Nunez G *et al*. Nod2-dependent regulation of innate and adaptive immunity in the intestinal tract. *Science* 2005;**307**:731-4.

386. Netea MG, Ferwerda G, De Jong DJ, Jansen T, Jacobs L, Kramer M *et al*. Nucleotide-binding oligomerization domain-2 modulates specific TLR pathways for the induction of cytokine release. *Journal of Immunology* 2005;**174**:6518-23.

387. McCartney-Francis N, Jin W, Wahl SM. Aberrant Toll receptor expression and endotoxin hypersensitivity in mice lacking a functional TGF-beta 1 signaling pathway. *Journal of Immunology* 2004;**172**:3814-21.

388. Naik S, Kelly EJ, Meijer L, Pettersson S, Sanderson IR. Absence of Toll-like receptor 4 explains endotoxin hyporesponsiveness in human intestinal epithelium. *Journal of Pediatric Gastroenterology & Nutrition* 2001;**32**:449-53.

389. Gaya DR, Russell RK, Nimmo ER, Satsangi J. New genes in inflammatory bowel disease: lessons for complex diseases? *Lancet* 2006;**367**:1271-84.

390. Arbour NC, Lorenz E, Schutte BC, Zabner J, Kline JN, Jones M *et al*. TLR4 mutations are associated with endotoxin hyporesponsiveness in humans. *Nature Genetics* 2000;**25**:187-91.

391. Franchimont D, Vermeire S, El Housni H, Pierik M, Van Steen K, Gustot T *et al*. Deficient host-bacteria interactions in inflammatory bowel disease? The toll-like receptor (TLR)-4 Asp299gly polymorphism is associated with Crohn's disease and ulcerative colitis. *Gut* 2004;**53**:987-92.

392. Brand S, Staudinger T, Schnitzler F, Pfennig S, Hofbauer K, Dambacher J *et al*. The role of Toll-like receptor 4 Asp299Gly and Thr399Ile polymorphisms and CARD15/NOD2 mutations in the susceptibility and phenotype of Crohn's disease. *Inflammatory Bowel Diseases* 2005;**11**:645-52.

393. Braat H, Stokkers P, Hommes T, Cohn D, Vogels E, Pronk I *et al*. Consequence of functional Nod2 and Tlr4 mutations on gene transcription in Crohn's disease patients. *Journal of Molecular Medicine* 2005;**83**:601-9.

394. Ouburg S, Mallant-Hent R, Crusius JB, van Bodegraven AA, Mulder CJ, Linskens R *et al*. The toll-like receptor 4 (TLR4) Asp299Gly polymorphism is associated with colonic localisation of Crohn's disease without a major role for the Saccharomyces cerevisiae mannan-LBP-CD14-TLR4 pathway. *Gut* 2005;**54**:439-40.

395. Browning BL, Huebner C, Petermann I, Gearry RB, Barclay ML, Shelling AN *et al*. Has Toll-Like Receptor 4 Been Prematurely Dismissed as an Inflammatory Bowel Disease Gene? Association Study Combined With Meta-Analysis Shows Strong Evidence for Association. *The American Journal of Gastroenterology* 2007;**102**:2504-12.

396. Török HP, Glas J, Tonenchi L, Bruennler G, Folwaczny M, Folwaczny C. Crohn's disease is associated with a toll-like receptor-9 polymorphism. *Gastroenterology* 2004;**127**:365-6.

397. Lakatos PL, Lakatos L, Szalay F, Willheim-Polli C, Osterreicher C, Tulassay Z *et al*. Toll-like receptor 4 and NOD2/CARD15 mutations in Hungarian patients with Crohn's disease: phenotype-genotype correlations. *World Journal of Gastroenterology* 2005;**11**:1489-95.

398. Oostenbrug LE, Drenth JPH, De Jong DJ, Nolte IM, Oosterom E, van Dullemen HM *et al*. Association between toll-like receptor 4 and inflammatory bowel disease. *Inflammatory Bowel Diseases* 2005;**11**:567-75.

399. Rotter JI, Taylor KD, Yang HY, Mei L, Picornell Y, Hawn TR *et al*. TLR5 polymorphisms are associated with OmpC and CBir1 expression and with severity of Crohn's disease in Ashkenazi Jews. *Gastroenterology* 2006;**130**:A53.

400. Gewirtz AT, Vijay-Kumar M, Brant SR, Duerr RH, Nicolae DL, Cho JH. Dominant-negative TLR5 polymorphism reduces adaptive immune response to flagellin and negatively associates with Crohn's disease. *American Journal of Physiology - Gastrointestinal & Liver Physiology* 2006;**290**:G1157-G1163.

401. Hawn TR, Verbon A, Lettinga KD, Zhao LP, Li SS, Laws RJ *et al*. A Common Dominant TLR5 Stop Codon Polymorphism Abolishes Flagellin Signaling and Is Associated with Susceptibility to Legionnaires' Disease. *J.Exp.Med.* 2003;**198**:1563-72.

402. Lammers KM, Ouburg S, Morre SA, Crusius JB, Gionchett P, Rizzello F *et al*. Combined carriership of TLR9-1237C and CD14-260T alleles enhances the risk of developing chronic relapsing pouchitis. *World Journal of Gastroenterology* 2005;**11**:7323-9.

403. Harton JA,.Ting JP. Class II transactivator: mastering the art of major histocompatibility complex expression. *Molecular & Cellular Biology*. 2000;6185-94.

404. Inohara N,.Nunez G. The NOD: a signaling module that regulates apoptosis and host defense against pathogens. *Oncogene.* 2001;6473-81.

405. Steimle V, Otten LA, Zufferey M, Mach B. Complementation cloning of an MHC class II transactivator mutated in hereditary MHC class II deficiency (or bare lymphocyte syndrome). *Cell* 1993;**75**:135-46.

406. Martinon F,.Tschopp J. NLRs join TLRs as innate sensors of pathogens. *Trends in Immunology* 2005;**26**:447-54.

407. Strober W, Murray PJ, Kitani A, Watanabe T. Signalling pathways and molecular interactions of NOD1 and NOD2. *Nat Rev Immunol* 2006;**6**:9-20.

408. Swanberg M, Lidman O, Padyukov L, Eriksson P, Akesson E, Jagodic M *et al*. MHC2TA is associated with differential MHC molecule expression and susceptibility to rheumatoid arthritis, multiple sclerosis and myocardial infarction. *Nature Genetics* 2005;**37**:486-94.

409. Patarroyo JC, Stuve O, Piskurich JF, Hauser SL, Oksenberg JR, Zamvil SS. Single nucleotide polymorphisms in MHC2TA, the gene encoding the MHC class II transactivator (CIITA). *Genes & Immunity* 2002;**3**:34-7.

410. Hoffman HM, Mueller JL, Broide DH, Wanderer AA, Kolodner RD. Mutation of a new gene encoding a putative pyrin-like protein causes familial cold autoinflammatory syndrome and Muckle-Wells syndrome. *Nature Genetics* 2001;**29**:301-5.

411. Goldbach-Mansky R, Dailey NJ, Canna SW, Gelabert A, Jones J, Rubin BI *et al*. Neonatal-onset multisystem inflammatory disease responsive to interleukin-1beta inhibition. *New England Journal of Medicine* 2006;**355**:581-92.

412. Ancient missense mutations in a new member of the RoRet gene family are likely to cause familial Mediterranean fever. The International FMF Consortium. *Cell* 1997;**90**:797-807.

413. A candidate gene for familial Mediterranean fever. The French FMF Consortium. *Nature Genetics* 1997;**17**:25-31.

414. Villani AC, Lemire M, Louis E, Silverberg M, Renaud Y, Brunet S *et al*. The familial Mediterranean fever (FMF) gene (MEFV) as a disease susceptibility gene in inflammatory bowel disease (IBD). *Gastroenterology* 2006;**130**:A53.

415. Karban A, Dagan E, Eliakim R, Herman A, Nesher S, Weiss B *et al*. Prevalence and significance of mutations in the familial Mediterranean fever gene in patients with Crohn's disease. *Genes Immun* 2004;**6**:134-9.

416. Fidder H, Chowers Y, Ackerman Z, Pollak RD, Crusius JB, Livneh A *et al*. The familial Mediterranean fever (MEVF) gene as a modifier of Crohn's disease. *American Journal of Gastroenterology* 2005;**100**:338-43.

417. Girardin SE, Jehanno M, Mengin-Lecreulx D, Sansonetti PJ, Alzari PM, Philpott DJ. Identification of the Critical Residues Involved in Peptidoglycan Detection by Nod1. *J.Biol.Chem.* 2005;**280**:38648-56.

418. Girardin SE, Boneca IG, Carneiro LAM, Antignac A, Jehanno M, Viala J *et al*. Nod1 detects a unique muropeptide from gram-negative bacterial peptidoglycan. *Science* 2003;**300**:1584-7.

419. Chamaillard M, Hashimoto M, Horie Y, Masumoto J, Qiu S, Saab L *et al*. An essential role for NOD1 in host recognition of bacterial peptidoglycan containing diaminopimelic acid. *Nature Immunology* 2003;**4**:702-7.

420. Uehara A, Fujimoto Y, Kawasaki A, Kusumoto S, Fukase K, Takada H. Meso-diaminopimelic acid and meso-lanthionine, amino acids specific to bacterial peptidoglycans, activate human epithelial cells through NOD1. *Journal of Immunology* 2006;**177**:1796-804.

421. Inohara N, Ogura Y, Fontalba A, Gutierrez O, Pons F, Crespo J *et al*. Host recognition of bacterial muramyl dipeptide mediated through NOD2: Implications for Crohn's disease. *J.Biol.Chem.* 2003;**278**:5509-12.

422. Girardin SE. Nod2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection. *J.Biol.Chem.* 2003;**278**:8869-72.

423. Inohara N, Ogura Y, Chen FF, Muto A, Nunez G. Human Nod1 confers responsiveness to bacterial lipopolysaccharides. *J.Biol.Chem.* 2001;**276**:2551-4.

424. Girardin SE, Tournebize R, Mavris M, Page AL, Li X, Stark GR *et al*. CARD4/Nod1 mediates NF-kappaB and JNK activation by invasive Shigella flexneri. *EMBO Reports* 2001;**2**:736-42.

425. Franchi L, Amer A, Body-Malapel M, Kanneganti TD, Ozoren N, Jagirdar R *et al*. Cytosolic flagellin requires Ipaf for activation of caspase-1 and interleukin 1beta in salmonella-infected macrophages. *Nature Immunology* 2006;**7**:576-82.

426. Miao EA, Alpuche-Aranda CM, Dors M, Clark AE, Bader MW, Miller SI *et al*. Cytoplasmic flagellin activates caspase-1 and secretion of interleukin 1beta via Ipaf. *Nature Immunology* 2006;**7**:569-75.

427. Ren T, Zamboni DS, Roy CR, Dietrich WF, Vance RE. Flagellin-deficient Legionella mutants evade caspase-1- and Naip5-mediated macrophage immunity. *PLoS Pathogens* 2006;**2**:e18.

428. Molofsky AB, Byrne BG, Whitfield NN, Madigan CA, Fuse ET, Tateda K *et al*. Cytosolic recognition of flagellin by mouse macrophages restricts Legionella pneumophila infection. *Journal of Experimental Medicine.* 2006;1093-104.

429. Inohara N,.Nunez G. NODs: intracellular proteins involved in inflammation and apoptosis. *Nat Rev Immunol* 2003;**3**:371-82.

430. Krantz DD, Zidovetzki R, Kagan BL, Zipursky SL. Amphipathic beta structure of a leucine-rich repeat peptide. *J Biol.Chem.* 1991;**266**:16801-7.

431. Viala J, Chaput C, Boneca IG, Cardona A, Girardin SE, Moran AP *et al*. Nod1 responds to peptidoglycan delivered by the Helicobacter pylori cag pathogenicity island. *Nature Immunology* 2004;**5**:1166-74.

432. Vavricka SR, Musch MW, Chang JE, Nakagawa Y, Phanvijhitsiri K, Waypa TS *et al*. hPepT1 transports muramyl dipeptide, activating NF-kappaB and stimulating IL-8 secretion in human colonic Caco2/bbe cells. *Gastroenterology* 2004;**127**:1401-9.

433. Barnich N, Aguirre JE, Reinecker HC, Xavier R, Podolsky DK. Membrane recruitment of NOD2 in intestinal epithelial cells is essential for nuclear factor-{kappa}B activation in muramyl dipeptide recognition. *Journal of Cell Biology* 2005;**170**:21-6.

434. Hysi P, Kabesch M, Moffatt MF, Schedel M, Carr D, Zhang Y *et al*. NOD1 variation, immunoglobulin E and asthma. *Hum.Mol.Genet.* 2005;**14**:935-41.

435. Lala S, Ogura Y, Osborne C, Hor SY, Bromfield A, Davies S *et al*. Crohn's disease and the NOD2 gene: A role for paneth cells. *Gastroenterology* 2003;**125**:47-57.

436. Ogura Y, Lala S, Xin W, Smith E, Dowds TA, Chen FF *et al*. Expression of NOD2 in Paneth cells: A possible link to Crohn's ileitis. *Gut* 2003;**52**:1591-7.

437. Abraham C,.Cho JH. Functional consequences of NOD2 (CARD15) mutations. *Inflammatory Bowel Diseases* 2006;**12**:641-50.

438. Berrebi D, Maudinas R, Hugot JP, Chamaillard M, Chareyre F, De Lagausie P *et al*. Card15 gene overexpression in mononuclear and epithelial cells of the inflamed Crohn's disease colon. *Gut* 2003;**52**:840-6.

439. Inohara N, Koseki T, Lin J, Del Peso L, Lucas PC, Chen FF *et al*. An induced proximity model for NF-kappa B activation in the Nod1/RICK and RIP signaling pathways. *J.Biol.Chem.* 2000;**275**:27823-31.

440. Tanabe T, Chamaillard M, Ogura Y, Zhu L, Qiu S, Masumoto J *et al*. Regulatory regions and critical residues of NOD2 involved in muramyl dipeptide recognition. *EMBO Journal* 2004;**23**:1587-97.

441. Inohara N, Chamaillard M, McDonald C, Nunez G. NOD-LRR proteins: role in host-microbial interactions and inflammatory disease. *Annual Review of Biochemistry* 2005;**74**:355-83.

442. Lecine P, Esmiol S, Metais JY, Nicoletti C, Nourry C, McDonald C *et al*. The NOD2-RICK Complex Signals from the Plasma Membrane. *J.Biol.Chem.* 2007;**282**:15197-207.

443. Kufer TA, Kremmer E, Adam AC, Philpott DJ, Sansonetti PJ. The pattern-recognition molecule Nod1 is localized at the plasma membrane at sites of bacterial interaction. *Cellular Microbiology* 2008;**10**:477-86.

444. Pauleau AL,.Murray PJ. Role of nod2 in the response of macrophages to toll-like receptor agonists. *Molecular & Cellular Biology* 2003;**23**:7531-9.

445. Barnich N, Hisamatsu T, Aguirre JE, Xavier R, Reinecker HC, Podolsky DK. GRIM-19 interacts with nucleotide oligomerization domain 2 and serves as downstream effector of anti-bacterial function in intestinal epithelial cells. *J.Biol.Chem.* 2005;**280**:19021-6.

446. Carneiro L, Magalhaes J, Tattoli I, Philpott D, Travassos L. Nod-like proteins in inflammation and disease. *Journal of Pathology* 2008;**214**:136-48.

447. Ferreira AC, Gomes L, Maximo V, Amil J, Carneiro F, Machado JC *et al*. GRIM-19 mutations are not associated with Crohn's disease. *Inflammatory Bowel Diseases* 2008;**14**:434-5.

448. McDonald C, Chen FF, Ollendorff V, Ogura Y, Marchetto S, Lecine P *et al*. A role for Erbin in the regulation of Nod2-dependent NF-kappaB signaling. *J.Biol.Chem.* 2005;**280**:40301-9.

449. Kufer TA, Kremmer E, Banks DJ, Philpott DJ. Role for erbin in bacterial activation of Nod2. *Infection & Immunity* 2006;**74**:3115-24.

450. Chen CM, Gong Y, Zhang M, Chen JJ. Reciprocal Cross-talk between Nod2 and TAK1 Signaling Pathways. *J.Biol.Chem.* 2004;**279**:25876-82.

451. Sato S, Sanjo H, Tsujimura T, Ninomiya-Tsuji J, Yamamoto M, Kawai T *et al*. TAK1 is indispensable for development of T cells and prevention of colitis by the generation of regulatory T cells. *Int.Immunol.* 2006;**18**:1405-11.

452. Martinon F,.Tschopp J. Inflammatory caspases: linking an intracellular innate immune system to autoinflammatory diseases. *Cell* 2004;**117**:561-74.

453. Mariathasan S, Newton K, Monack DM, Vucic D, French DM, Lee WP *et al*. Differential activation of the inflammasome by caspase-1 adaptors ASC and Ipaf. *Nature* 2004;**430**:213-8.

454. Damiano JS, Oliveira V, Welsh K, Reed JC. Heterotypic interactions among NACHT domains: implications for regulation of innate immune responses. *Biochem J* 2004;**381**:213-9.

455. Stehlik C, Hayashi H, Pio F, Godzik A, Reed JC. CARD6 is a modulator of NF-kappa B activation by Nod1- and Cardiak-mediated pathways. *J.Biol.Chem.* 2003;**278**:31941-9.

456. Hsu YM, Zhang Y, You Y, Wang D, Li H, Duramad O *et al*. The adaptor protein CARD9 is required for innate immune responses to intracellular pathogens. *Nat Immunol* 2007;**8**:198-205.

457. Rosenstiel P, Huse K, Till A, Hampe J, Hellmig S, Sina C *et al*. A short isoform of NOD2/CARD15, NOD2-S, is an endogenous inhibitor of NOD2/receptor-interacting protein kinase 2-induced signaling pathways. *Proceedings of the National Academy of Sciences of the United States of America* 2006;**103**:3280-5.

458. Martinon F, Agostini L, Meylan E, Tschopp J. Identification of bacterial muramyl dipeptide as activator of the NALP3/cryopyrin inflammasome. *Current Biology* 2004;**14**:1929-34.

459. McGovern DPB, Hysi P, Ahmad T, van Heel DA, Moffatt MF, Carey A *et al*. Association between a complex insertion/deletion polymorphism in NOD1 (CARD4) and susceptibility to inflammatory bowel disease. *Hum.Mol.Genet.* 2005;**14**:1245-50.

460. Zouali H, Lesage S, Merlin F, Cezard JP, Colombel JF, Belaiche J *et al*. CARD4/NOD1 is not involved in inflammatory bowel disease. *Gut* 2003;**52**:71-4.

461. Ozen SC, Dagli U, Kilic MY, Toruner M, Celik Y, Ozkan M *et al*. Nod2/card15, nod1/card4, and icam-1 gene polymorphisms in turkish patients with inflammatory bowel disease. *Journal of Gastroenterology* 2006;**41**:304-10.

462. Van Limbergen J, Lees CW, Nimmo ER, Russell RK, Drummond HE, Wilson DC *et al*. Association of a complex insertion/deletion polymorphism of NOD1/CARD4 with susceptibility to inflammatory bowel disease in the Scottish population. *Gastroenterology* 2006;**130**:A64.

463. Tremelling M, Hancock L, Bredin F, Sharpstone D, Bingham SA, Parkes M. Complex insertion/deletion polymorphism in NOD1 (CARD4) is not associated with inflammatory bowel disease susceptibility in east anglia panel. *Inflammatory Bowel Diseases* 2006;**12**:967-71.

464. Franke A, Ruether A, Wedemeyer N, Karlsen TH, Nebel A, Schreiber S. No association between the functional CARD4 insertion/deletion polymorphism and inflammatory bowel diseases in the German population. *Gut* 2006;**55**:1679-80.

465. Holler E, Rogler G, Brenmoehl J, Hahn J, Herfarth H, Greinix H *et al*. Prognostic significance of NOD2/CARD15 variants in HLA-identical sibling hematopoietic stem cell transplantation: effect on long-term outcome is confirmed in 2 independent cohorts and may be modulated by the type of gastrointestinal decontamination. *Blood* 2006;**107**:4189-93.

466. Miceli-Richard C, Lesage S, Rybojad M, Prieur AM, Manouvrier-Hanu S, Hafner R *et al*. CARD15 mutations in Blau syndrome. *Nature Genetics* 2001;**29**:19-20.

467. Kanazawa N, Okafuji I, Kambe N, Nishikomori R, Nakata-Hizume M, Nagai S *et al*. Early-onset sarcoidosis and CARD15 mutations with constitutive nuclear factor-kappaB activation: common genetic etiology with Blau syndrome. *Blood* 2005;**105**:1195-7.

468. Economou M, Trikalinos TA, Loizou KT, Tsianos EV, Ioannidis JP. Differential effects of NOD2 variants on Crohn's disease risk and phenotype in diverse populations: a metaanalysis. *American Journal of Gastroenterology* 2004;**99**:2393-404.

469. Oostenbrug LE, Nolte IM, Oosterom E, van der Steege G, te Meerman GJ, van Dullemen HM *et al*. CARD15 in inflammatory bowel disease and Crohn's disease phenotypes: An association study and pooled analysis. *Digestive and Liver Disease* 2006;**38**:834-45.

470. Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, Daly MJ *et al*. A Genome-Wide Association Study Identifies IL23R as an Inflammatory Bowel Disease Gene. *Science* 2006;**314**:1461-3.

471. Libioulle C, Louis E, Hansoul S, Sandor C, Farnir F, Franchimont D *et al*. A novel susceptibility locus for Crohns disease identified by whole genome association maps to a gene desert on chromosome 5p13.1 and modulates the level of expression of the prostaglandin receptor EP4. *PloS Genetics* 2007;**3**:e58.

472. Parkes M, Barrett JC, Prescott NJ, Tremelling M, Anderson CA, Fisher SA *et al*. Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nat Genet* 2007;**39**:830-2.

473. Törkvist L, Noble CL, Lördal M, Sjöqvist U, Lindforss U, Nimmo ER *et al*. Contribution of CARD15 variants in determining susceptibility to Crohn's disease in Sweden. *Scandinavian Journal of Gastroenterology.* 2006;**41**:700-5.

474. Inoue N, Tamura K, Kinouchi Y, Fukuda Y, Takahashi S, Ogura Y *et al*. Lack of common NOD2 variants in Japanese patients with Crohn's disease. *Gastroenterology* 2002;**123**:86-91.

475. Yamazaki K, Takazoe M, Tanaka T, Kazumori T, Nakamura Y. Absence of mutation in the NOD2/CARD15 gene among 483 Japanese patients with Crohn's disease. *Journal of Human Genetics* 2002;**47**:469-72.

476. Li J, Moran T, Swanson E, Julian C, Harris J, Bonen DK *et al*. Regulation of IL-8 and IL-1beta expression in Crohn's disease associated NOD2/CARD15 mutations. *Hum.Mol.Genet.* 2004;**13**:1715-25.

477. van Heel DA, Ghosh S, Butler M, Hunt K, Foxwell BM, Mengin-Lecreulx D *et al*. Synergistic enhancement of Toll-like receptor responses by NOD1 activation. *European Journal of Immunology* 2005;**35**:2471-6.

478. van Heel DA, Ghosh S, Hunt KA, Mathew CG, Forbes A, Jewell DP *et al*. Synergy between TLR9 and NOD2 innate immune responses is lost in genetic Crohn's disease. *Gut* 2005;**54**:1553-7.

479. Tada H, Aiba S, Shibata K, Ohteki T, Takada H. Synergistic effect of Nod1 and Nod2 agonists with toll-like receptor agonists on human dendritic cells to generate interleukin-12 and T helper type 1 cells. *Infection & Immunity* 2005;**73**:7967-76.

480. Fritz JH, Girardin SE, Fitting C, Werts C, Mengin-Lecreulx D, Caroff M *et al*. Synergistic stimulation of human monocytes and dendritic cells by Toll-like receptor 4 and NOD1- and NOD2-activating agonists. *European Journal of Immunology* 2005;**35**:2459-70.

481. Yang S, Tamai R, Akashi S, Takeuchi O, Akira S, Sugawara S *et al*. Synergistic effect of muramyldipeptide with lipopolysaccharide or lipoteichoic acid to induce inflammatory cytokines in human monocytic cells in culture. *Infection & Immunity* 2001;**69**:2045-53.

482. Wolfert MA, Murray TF, Boons GJ, Moore JN. The origin of the synergistic effect of muramyl dipeptide with endotoxin and peptidoglycan. *J.Biol.Chem.* 2002;**277**:39179-86.

483. Uehara A, Yang S, Fujimoto Y, Fukase K, Kusumoto S, Shibata K *et al*. Muramyldipeptide and diaminopimelic acid-containing desmuramylpeptides in combination with chemically synthesized Toll-like receptor agonists synergistically induced production of interleukin-8 in a NOD2- and NOD1-dependent manner, respectively, in human monocytic cells in culture. *Cellular Microbiology* 2005;**7**:53-61.

484. Netea MG, Ferwerda G, De Jong DJ, Werts C, Boneca IG, Jehanno M *et al*. The frameshift mutation in Nod2 results in unresponsiveness not only to

Nod2- but also Nod1-activating peptidoglycan agonists. *J.Biol.Chem.* 2005;**280**:35859-67.

485. van Heel DA, Hunt KA, Ghosh S, Herve M, Playford RJ. Normal responses to specific NOD1-activating peptidoglycan agonists in the presence of the NOD2 frameshift and other mutations in Crohn's disease. *European Journal of Immunology* 2006;**36**:1629-35.

486. Watanabe T, Asano N, Murray PJ, Ozato K, Tailor P, Fuss IJ *et al*. Muramyl dipeptide activation of nucleotide-binding oligomerization domain 2 protects mice from experimental colitis. *J Clin Invest* 2008;**118**:545-59.

487. Kullberg BJ, Ferwerda G, De Jong DJ, Drenth JP, Joosten LA, van der Meer JW *et al*. Crohn's disease patients homozygous for the 3020insC NOD2 mutation have a defective NOD2/TLR4 cross-tolerance to intestinal stimuli. *Immunology* 2008;**123**:600-5.

488. MacPherson AJ,.Harris NL. Interactions between commensal intestinal bacteria and the immune system. *Nature Reviews.Immunology* 2004;**4**:478-85.

489. Eckmann L. Defence molecules in intestinal innate immunity against bacterial infections. *Current Opinion in Gastroenterology* 2005;**21**:147-51.

490. Cash HL, Whitham CV, Behrendt CL, Hooper LV. Symbiotic bacteria direct expression of an intestinal bactericidal lectin. *Science* 2006;**313**:1126-30.

491. Ganz T, Selsted ME, Szklarek D, Harwig SS, Daher K, Bainton DF *et al*. Defensins. Natural peptide antibiotics of human neutrophils. *J.Clin.Invest.* 1985;**76**:1427-35.

492. Selsted ME, Harwig SS, Ganz T, Schilling JW, Lehrer RI. Primary structures of three human neutrophil defensins. *J.Clin.Invest.* 1985;**76**:1436-9.

493. Ganz T. Defensins: Antimicrobial peptides of innate immunity. *Nat Rev Immunol* 2003;**3**:710-20.

494. Pazgier M, Hoover DM, Yang D, Lu W, Lubkowski J. Human beta-defensins. *Cellular & Molecular Life Sciences* 2006;**63**:1294-313.

495. Selsted ME,.Ouellette AJ. Mammalian defensins in the antimicrobial immune response. *Nature Immunology* 2005;**6**:551-7.

496. Lynn DJ, Lloyd AT, Fares MA, O'Farrelly C. Evidence of positively selected sites in mammalian alpha-defensins. *Molecular Biology & Evolution* 2004;**21**:819-27.

497. Semple CA, Gautier P, Taylor K, Dorin JR. The changing of the guard: Molecular diversity and rapid evolution of beta-defensins. *Molecular Diversity* 2006;**10**:575-84.

498. Ghosh D, Porter E, Shen B, Lee SK, Wilk D, Drazba J *et al*. Paneth cell trypsin is the processing enzyme for human defensin-5. *Nat Immunol* 2002;**3**:583-90.

499. Rumio C, Besusso D, Palazzo M, Selleri S, Sfondrini L, Dubini F *et al*. Degranulation of Paneth Cells via Toll-Like Receptor 9. *Am J Pathol* 2004;**165**:373-81.

500. Wilson CL, Ouellette AJ, Satchell DP, Ayabe T, Lopez-Boado YS, Stratman JL *et al*. Regulation of intestinal alpha-defensin activation by the metalloproteinase matrilysin in innate host defense. *Science* 1999;**286**:113-7.

501. Salzman NH, Ghosh D, Huttner KM, Paterson Y, Bevins CL. Protection against enteric salmonellosis in transgenic mice expressing a human intestinal defensin. *Nature* 2003;**422**:522-6.

502. Wehkamp J, Salzman NH, Porter E, Nuding S, Weichenthal M, Petras RE *et al*. Reduced Paneth cell alpha-defensins in ileal Crohn's disease. *Proceedings of the National Academy of Sciences of the United States of America* 2005;**102**:18129-34.

503. Wehkamp J, Harder J, Weichenthal M, Schwab M, Schaffeler E, Schlee M *et al*. NOD2 (CARD15) mutations in Crohn's disease are associated with diminished mucosal alpha-defensin expression. *Gut* 2004;**53**:1658-64.

504. Linzmeier RM,.Ganz T. Copy number polymorphisms are not a common feature of innate immune genes. *Genomics* 2006;**88**:122-6.

505. O'Neil DA, Porter EM, Elewaut D, Anderson GM, Eckmann L, Ganz T *et al*. Expression and regulation of the human beta-defensins hBD-1 and hBD-2 in intestinal epithelium. *Journal of Immunology* 1999;**163**:6718-24.

506. Wehkamp J, Harder J, Weichenthal M, Mueller O, Herrlinger KR, Fellermann K *et al*. Inducible and constitutive beta-defensins are differentially expressed in Crohn's disease and ulcerative colitis. *Inflammatory Bowel Diseases* 2003;**9**:215-23.

507. Wehkamp J, Fellermann K, Herrlinger KR, Baxmann S, Schmidt K, Schwind B *et al*. Human beta-defensin 2 but not beta-defensin 1 is expressed preferentially in colonic mucosa of inflammatory bowel disease. *European Journal of Gastroenterology & Hepatology* 2002;**14**:745-52.

508. Fahlgren A, Hammarstrom S, Danielsson A, Hammarstrom ML. beta-Defensin-3 and -4 in intestinal epithelial cells display increased mRNA

expression in ulcerative colitis. *Clinical & Experimental Immunology* 2004;**137**:379-85.

509. Birchler T, Seibl R, Buchner K, Loeliger S, Seger R, Hossle JP *et al*. Human Toll-like receptor 2 mediates induction of the antimicrobial peptide human beta-defensin 2 in response to bacterial lipoprotein. *European Journal of Immunology* 2001;**31**:3131-7.

510. Vora P, Youdim A, Thomas LS, Fukata M, Tesfay SY, Lukasek K *et al*. Beta-defensin-2 expression is regulated by TLR signaling in intestinal epithelial cells. *Journal of Immunology* 2004;**173**:5398-405.

511. Salzman NH, Chou MM, de Jong H, Liu L, Porter EM, Paterson Y. Enteric salmonella infection inhibits Paneth cell antimicrobial peptide expression. *Infection & Immunity* 2003;**71**:1109-15.

512. Zaalouk TK, Bajaj-Elliott M, George JT, McDonald V. Differential regulation of beta-defensin gene expression during Cryptosporidium parvum infection. *Infection & Immunity* 2004;**72**:2772-9.

513. Voss E, Wehkamp J, Wehkamp K, Stange EF, Schroder JM, Harder J. NOD2/CARD15 mediates induction of the antimicrobial peptide human beta-defensin-2. *J.Biol.Chem.* 2006;**281**:2005-11.

514. Fellermann K, Stange DE, Schaeffeler E, Schmalzl H, Wehkamp J, Bevins CL *et al*. A chromosome 8 gene-cluster polymorphism with low human Beta-defensin 2 gene copy number predisposes to crohn disease of the colon. *American Journal of Human Genetics* 2006;**79**:439-48.

515. Hollox EJ, Huffmeier U, Zeeuwen PLJM, Palla R, Lascorz J, Rodijk-Olthuis D *et al*. Psoriasis is associated with increased beta-defensin genomic copy number. *Nat Genet* 2008;**40**:23-5.

516. Silverberg M.S., Satsangi J, Ahmad T, Arnott IDR, Bernstein C.N., Brant SR *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**:5A-36A.

517. Singer HC, Anderson JG, Frischer H, Kirsner JB. Familial aspects of inflammatory bowel disease. *Gastroenterology* 1971;**61**:423-30.

518. Polito JM, Childs B, Mellits ED, Tokayer AZ, Harris ML, Bayless TM. Crohn's disease: influence of age at diagnosis on site and clinical type of disease. *Gastroenterology* 1996;**111**:580-6.

519. Farmer RG, Michener WM, Mortimer EA. Studies of family history among patients with inflammatory bowel disease. *Clinics in Gastroenterology* 1980;**9**:271-7.

520. Brant SR, Panhuysen CI, Bailey-Wilson JE, Rohal PM, Lee S, Mann J *et al*. Linkage heterogeneity for the IBD1 locus in Crohn's disease pedigrees by disease onset and severity. *Gastroenterology* 2000;**119**:1483-90.

521. Gasche C, Scholmerich J, Brynskov J, D'Haens G, Hanauer SB, Irvine EJ *et al*. A simple classification of Crohn's disease: report of the Working Party for the World Congresses of Gastroenterology, Vienna 1998. *Inflammatory Bowel Diseases* 2000;**6**:8-15.

522. Sawczenko A, Lynn R, Sandhu BK. Variations in initial assessment and management of inflammatory bowel disease across Great Britain and Ireland. *Arch Dis Child* 2003;**88**:990-4.

523. Lennard-Jones JE. Classification of inflammatory bowel disease. *Scandinavian Journal of Gastroenterology* 1989;**24**:2-6.

524. Oefferlbauer-Ernst A, Miehsler W, Eckmullner O, Travis S, Waldhoer T, Dejaco C *et al*. Impact of interobserver disagreement on phenotype-genotype associations in Crohn's disease. *Inflamm.Bowel Dis.* 2007;**13**:156-63.

525. Riis L, Munkholm P, Binder V, Skovgaard LT, Langholz E. Intra- and interobserver variation in the use of the Vienna classification of Crohn's disease. *Inflamm.Bowel Dis.* 2005;**11**:657-61.

526. IBD Working Group of the European Society for Paediatric Gastroenterology HaN. Inflammatory bowel disease in children and adolescents: recommendations for diagnosis--the Porto criteria. *Journal of Pediatric Gastroenterology & Nutrition* 2005;**41**:1-7.

527. Zhao JH, Curtis D, Sham PC. Model-free analysis and permutation tests for allelic associations. *Human Heredity* 2000;**50**:133-9.

528. Gauderman WJ. Candidate gene association studies for a quantitative trait, using parent-offspring trios. *Genetic Epidemiology* 2003;**25**:327-38.

529. Gauderman, W. J. and Morrison, J. M. QUANTO 1.1: A computer program for power and sample size calculations for genetic epidemiology studies. 2006.

Ref Type: Computer Program

530. O'Connell JR,.Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *American Journal of Human Genetics* 1998;**63**:259-66.

531. Clayton D. A generalization of the transmission/disequillibrium test for uncertain- haplotype transmission. *American Journal of Human Genetics* 1999;**65**:1170-7.

532. Laird N, Horvath S, Xu X. Implementing a unified approach to family based tests of association. *Genetic Epidemiology* 2000;**19**:S36-S42.

533. Clayton D,.Jones H. Transmission/disequilibrium tests for extended marker haplotypes. *American Journal of Human Genetics* 1999;**65**:1161-9.

534. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira M, Bender D *et al*. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics* 2007;**81**:559-75.

535. Kazeem GR,.Farrall M. Integrating Case-control and TDT Studies. *Annals of Human Genetics* 2005;**69**:329-35.

536. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research* 1988;**16**:1215.

537. Macdonald HM, McGuigan FE, Stewart A, Black AJ, Fraser WD, Ralston S *et al*. Large-Scale Population-Based Study Shows No Evidence of Association Between Common Polymorphism of the VDR Gene and BMD in British Women. *Journal of Bone and Mineral Research* 2006;**21**:151-62.

538. GUNESH S, Thomas GAO, Williams GT, ROBERTS A, HAWTHORNE AB. The incidence of Crohns disease in Cardiff over the last 75 years: an update for 1996-2005. *Alimentary Pharmacology & Therapeutics* 2008;**27**:211-9.

539. Turner D, Grossman AB, Rosh J, Kugathasan S, Gilman AR, Baldassano R *et al*. Methotrexate Following Unsuccessful Thiopurine Therapy in Pediatric Crohn's Disease. *The American Journal of Gastroenterology* 2007;**102**:2804-12.

540. Hyams J, Crandall W, Kugathasan S, Griffiths A, Olson A, Johanns J *et al*. Induction and maintenance infliximab therapy for the treatment of moderate-to-severe Crohn's disease in children. *Gastroenterology* 2007;**132**:863-73.

541. Heyman MB, Kirschner BS, Gold BD, Ferry G, Baldassano R, Cohen SA *et al*. Children with early-onset inflammatory bowel disease (IBD): analysis of a pediatric IBD consortium registry. *Journal of Pediatrics* 2005;**146**:35-40.

542. Paul TM, Birnbaum AM, Pal DKP, Pittman NM, Ceballos CR, LeLeiko NSP* *et al*. Distinct Phenotype of Early Childhood Inflammatory Bowel Disease. *Journal of Clinical Gastroenterology* 2006;**40**:583-6.

543. Gupta N, Bostrom AG, Kirschner BS, Ferry GD, Winter HS, Baldassano RN *et al*. Gender differences in presentation and course of disease in pediatric patients with Crohn disease. *Pediatrics* 2007;**120**:e1418-e1425.

544. Arnott IDR,.Satsangi J. Crohn's disease or Crohn's diseases? *Gut* 2003;**52**:460-1.

545. North American Society for Pediatric Gastroenterology HaN, Colitis Foundation of America, Bousvaros A, Antonioli DA, Colletti RB, Dubinsky MC *et al*. Differentiating ulcerative colitis from Crohn disease in children and young adults: report of a working group of the North American Society for Pediatric Gastroenterology, Hepatology, and Nutrition and the Crohn's and Colitis Foundation of America. *Journal of Pediatric Gastroenterology & Nutrition* 2007;**44**:653-74.

546. Satsangi J, Silverberg MS, Vermeire S, Colombel JF. The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. *Gut* 2006;**55**:749-53.

547. Baldassano RN, Bradfield JP, Monos DS, Kim CE, Glessner JT, Casalunovo T *et al*. Association of the T300A non-synonymous variant of the ATG16L1 gene with susceptibility to paediatric Crohn's disease. *Gut* 2007;**56**:1171-3.

548. Cucchiara S, Latiano A, Palmieri O, Canani RB, D'Inca R, Guariso G *et al*. Polymorphisms of Tumor Necrosis Factor-[alpha] but Not MDR1 Influence Response to Medical Therapy in Pediatric-Onset Inflammatory Bowel Disease. *Journal of Pediatric Gastroenterology & Nutrition* 2007;**44**:171-9.

549. Sawczenko A,.Sandhu BK. Presenting features of inflammatory bowel disease in Great Britain and Ireland. *Arch Dis Child* 2003;**88**:995-1000.

550. Kugathasan S, Judd RH, Hoffmann RG, Heikenen J, Telega G, Khan F *et al*. Epidemiologic and clinical characteristics of children with newly diagnosed inflammatory bowel disease in Wisconsin: a statewide population-based study. *Journal of Pediatrics* 2003;**143**:525-31.

551. Witte J, Shivananda S, Lennard-Jones JE, Beltrami M, Politi P, Bonanomi A *et al*. Disease outcome in inflammatory bowel disease: mortality, morbidity and therapeutic management of a 796-person inception cohort in the European Collaborative Study on Inflammatory Bowel Disease (EC-IBD). *Scandinavian Journal of Gastroenterology* 2000;**35**:1272-7.

552. Henriksen M, Jahnsen J, Lygren I, Sauar J, Kjellevold O, Schulz T *et al*. Ulcerative colitis and clinical course: results of a 5-year population-based follow-up study (the IBSEN study). *Inflammatory Bowel Diseases* 2006;**12**:543-50.

553. Louis E, Collard A, Oger AF, Degroote E, El Yafi FAN, Belaiche J. Behaviour of Crohn's disease according to the Vienna classification: changing pattern over the course of the disease. *Gut* 2001;**49**:777-82.

554. Henriksen M, Jahnsen J, Lygren I, Aadland E, Schulz T, Vatn MH *et al*. Clinical course in Crohn's disease: results of a five-year population-based follow-up study (the IBSEN study). *Scand.J Gastroenterol.* 2007;**42**:602-10.

555. Mamula P, Telega GW, Markowitz JE, Brown KA, Russo PA, Piccoli DA *et al*. Inflammatory bowel disease in children 5 years of age and younger. *American Journal of Gastroenterology* 2002;**97**:2005-10.

556. Louis E, Michel V, Hugot JP, Reenaers C, Fontaine F, Delforge M *et al*. Early development of stricturing or penetrating pattern in Crohn's disease is influenced by disease location, number of flares, and smoking but not by NOD2/CARD15 genotype. *Gut* 2003;**52**:552-7.

557. Smith BRKB, Arnott IDRM, Drummond HEB, Nimmo ERP, Satsangi JD. Disease Location, Anti-Saccharomyces cerevisiae Antibody, and NOD2/CARD15 Genotype Influence the Progression of Disease Behavior in Crohn's Disease. *Inflammatory Bowel Diseases* 2004;**10**:521-8.

558. Nikolaus S,.Schreiber S. Diagnostics of Inflammatory Bowel Disease. *Gastroenterology* 2007;**133**:1670-89.

559. Carter MJ, Lobo AJ, Travis SPL. Guidelines for the management of inflammatory bowel disease in adults. *Gut* 2004;**53**:v1-v16.

560. Hommes DW,.van Deventer SJ. Endoscopy in inflammatory bowel diseases. *Gastroenterology* 2004;**126**:1561-73.

561. Witte AMC, Veenendaal RA, van Hogezand RA, Verspaget HW, Lamers CBHW. Crohn's Disease of the Upper Gastrointestinal Tract: the Value of Endoscopic Examination. *Scandinavian Journal of Gastroenterology* 1998;**33**:100-5.

562. Markowitz J, Grancher K, Kohn N, Lesser M, Daum F. A multicenter trial of 6-mercaptopurine and prednisone in children with newly diagnosed Crohn's disease. *Gastroenterology* 2000;**119**:895-902.

563. Ho G-T, Hudson M, Lee HM, Han WH, Ting T, Chiam PP *et al*. The efficacy of corticosteroid therapy: Analysis of 10-year inflammatory bowel disease inception cohort (1998-2007). *Gastroenterology* 2008;**134**:A1043.

564. Mackey AC, Green L, Liang LC, Dinndorf P, Avigan M. Hepatosplenic T cell lymphoma associated with infliximab use in young patients treated for inflammatory bowel disease. *Journal of Pediatric Gastroenterology & Nutrition* 2007;**44**:265-7.

565. Odes S, Vardi H, Friger M, Wolters F, Russel MG, Riis L *et al*. Cost analysis and cost determinants in a European inflammatory bowel disease

inception cohort with 10 years of follow-up evaluation. *Gastroenterology* 2006;**131**:719-28.

566. Armitage EL, Aldhous MC, Anderson N, Drummond HE, Riemersma RA, Ghosh S *et al*. Incidence of juvenile-onset Crohn's disease in Scotland: association with northern latitude and affluence. *Gastroenterology* 2004;**127**:1051-7.

567. Auvin S, Molinie F, Gower-Rousseau C, Brazier F, Merle V, Grandbastien B *et al*. Incidence, clinical presentation and location at diagnosis of pediatric inflammatory bowel disease: a prospective population-based study in northern France (1988-1999). *J Pediatr.Gastroenterol.Nutr.* 2005;**41**:49-55.

568. Turunen P, Kolho KL, Auvinen A, Iltanen S, Huhtala H, Ashorn M. Incidence of inflammatory bowel disease in Finnish children, 1987-2003. *Inflammatory Bowel Diseases* 2006;**12**:677-83.

569. Hildebrand H, Finkel Y, Grahnquist L, Lindholm J, Ekbom A, Askling J. Changing pattern of paediatric inflammatory bowel disease in northern Stockholm 1990-2001. *Gut* 2003;**52**:1432-4.

570. Nerich V, Monnet E, Etienne A, Louafi S, Ramee C, Rican S *et al*. Geographical variations of inflammatory bowel disease in France: a study based on national health insurance data. *Inflamm.Bowel.Dis.* 2006;**12**:218-26.

571. Van Limbergen J, Nimmo ER, Russell RK, Drummond HE, Smith L, Anderson NH *et al*. Investigation of NOD1/CARD4 variation in Inflammatory Bowel Disease using a haplotype-tagging strategy. *Hum.Mol.Genet.* 2007;**16**:2175-86.

572. Armitage E, Drummond H, Ghosh S, Ferguson A. Incidence of juvenile-onset Crohn's disease in Scotland. *Lancet* 1999;**353**:1496-7.

573. McVeigh, Karen. Doctor in MMR row defends stance at disciplinary hearing. The Guardian 28th March 2008, 12. 28-3-2008.

Ref Type: Newspaper

574. Koloski NA, Bret L, Radford-Smith G. Hygiene hypothesis in inflammatory bowel disease: a critical review of the literature. *World J Gastroenterol.* 2008;**14**:165-73.

575. Tysk C, Lindberg E, Jarnerot G, Floderusmyrhed B. Ulcerative colitis and Crohn's disease in an unselected population of monozygotic and dizygotic twins. A study of heritability and the influence of smoking. *Gut* 1988;**29**:990-6.

576. Orholm M, Binder V, Sorensen TIA, Rasmussen LP, Kyvik KO. Concordance of inflammatory bowel disease among Danish twins - Results of a nationwide study. *Scandinavian Journal of Gastroenterology* 2000;**35**:1075-81.

577. Thompson NP, Driscoll R, Pounder RE, Wakefield AJ. Genetics versus environment in inflammatory bowel disease: Results of a British twin study. *British Medical Journal* 1996;**312**:95-6.

578. Russell RK,.Satsangi J. IBD: a family affair. *Best Practice & Research in Clinical Gastroenterology* 2004;**18**:525-39.

579. Barker DJP. The origins of the developmental origins theory. *Journal of Internal Medicine* 2007;**261**:412-7.

580. Cornish JA, Tan E, Simillis C, Clark SK, Teare J, Tekkis PP. The risk of oral contraceptives in the etiology of inflammatory bowel disease: a meta-analysis. *American Journal of Gastroenterology* 2008;**103**:2394-400.

581. Garcia Rodriguez LA, Gonzalez-Perez A, Johansson S, Wallander MA. Risk factors for inflammatory bowel disease in the general population. *Alimentary Pharmacology & Therapeutics* 2005;**22**:309-15.

582. Aldhous MC, Drummond HE, Anderson N, Smith LA, Arnott ID, Satsangi J. Does cigarette smoking influence the phenotype of Crohn's disease? Analysis using the Montreal classification. *American Journal of Gastroenterology* 2007;**102**:577-88.

583. Bernstein CN,.Shanahan F. Disorders of a modern lifestyle-reconciling the epidemiology of inflammatory bowel diseases. *Gut* 2008;**57**:1185-91.

584. Bernstein CN, Rawsthorne P, Cheang M, Blanchard JF. A Population-Based Case Control Study of Potential Risk Factors for IBD. *The American Journal of Gastroenterology* 2006;**101**:993-1002.

585. Mahid SS, Minor KS, Stromberg AJ, Galandiuk S. Active and passive smoking in childhood is related to the development of inflammatory bowel disease. *Inflammatory Bowel Diseases* 2007;**13**:431-8.

586. Lindberg E, Jarnerot G, Huitfeldt B. Smoking in Crohn's disease: effect on localisation and clinical course. *Gut* 1992;**33**:779-82.

587. Persson PG, Leijonmarck CE, Bernell O, Hellers G, Ahlbom A. Risk indicators for inflammatory bowel disease. *International Journal of Epidemiology* 1993;**22**:286-72.

588. Jones DT, Osterman MT, Bewtra M, Lewis JD. Passive smoking and inflammatory bowel disease: a meta-analysis. *American Journal of Gastroenterology* 2008;**103**:2382-93.

589. Gilat T, Hacohen D, Lilos P, Langman MJ. Childhood factors in ulcerative colitis and Crohn's disease. An international cooperative study. *Scandinavian Journal of Gastroenterology* 1987;**22**:1009-24.

590. Kurina LM, Goldacre MJ, Yeates D, Seagroatt V. Appendicectomy, tonsillectomy, and inflammatory bowel disease: a case-control record linkage study. *Journal of Epidemiology & Community Health* 2002;**56**:551-4.

591. Duggan AE, Usmani I, Neal KR, Logan RF. Appendicectomy, childhood hygiene, Helicobacter pylori status, and risk of inflammatory bowel disease: a case control study. *Gut* 1998;**43**:494-8.

592. Radford-Smith GL, Edwards JE, Purdie DM, Pandeya N, Watson M, Martin NG *et al*. Protective role of appendicectomy on onset and severity of ulcerative colitis and Crohn's disease. *Gut* 2002;**51**:808-13.

593. Andersson RE, Olaison G, Tysk C, Ekbom A. Appendectomy and protection against ulcerative colitis. *New England Journal of Medicine* 2001;**344**:808-14.

594. Frisch M, Pedersen BV, Andersson RE. Appendicitis, mesenteric lymphadenitis, and subsequent risk of ulcerative colitis: cohort studies in Sweden and Denmark. *BMJ* 2009;**338**:b716.

595. Kaplan GG, Pedersen BV, Andersson RE, Sands BE, Korzenik J, Frisch M. The risk of developing Crohn's disease after an appendectomy: a population-based cohort study in Sweden and Denmark. *Gut* 2007;**56**:1387-92.

596. Bernstein CN, Wajda A, Blanchard JF. The clustering of other chronic inflammatory diseases in inflammatory bowel disease: a population-based study. *Gastroenterology* 2005;**129**:827-36.

597. Bach JF. The effect of infections on susceptibility to autoimmune and allergic diseases. *New England Journal of Medicine* 2002;**347**:911-20.

598. Rook GAW,.Brunet LR. Microbes, immunoregulation, and the gut. *Gut* 2005;**54**:317-20.

599. Penders J, Thijs C, van den Brandt PA, Kummeling I, Snijders B, Stelma F *et al*. Gut microbiota composition and development of atopic manifestations in infancy: the KOALA Birth Cohort Study. *Gut* 2007;**56**:661-7.

600. Wurzelmann JI, Lyles CM, Sandler RS. Childhood infections and the risk of inflammatory bowel disease. *Digestive Diseases & Sciences* 1994;**39**:555-60.

601. Ekbom A, Wakefield AJ, Zack M, Adami HO. Perinatal measles infection and subsequent Crohn's disease. *Lancet* 1994;**344**:508-10.

602. Whorwell PJ, Holdstock G, Whorwell GM, Wright R. Bottle feeding, early gastroenteritis, and inflammatory bowel disease. *British Medical Journal* 1979;**1**:382.

603. Koletzko S, Sherman P, Corey M, Griffiths A, Smith C. Role of infant feeding practices in development of Crohn's disease in childhood. *BMJ* 1989;**298**:1617-8.

604. Corrao G, Tragnone A, CAPRILLI R, Trallori G, Papi C, Andreoli A *et al*. Risk of inflammatory bowel disease attributable to smoking, oral contraception and breastfeeding in Italy: a nationwide case-control study. Cooperative Investigators of the Italian Group for the Study of the Colon and the Rectum (GISC). *Int.J Epidemiol.* 1998;**27**:397-404.

605. Rigas A, Rigas B, Glassman M, Yen YY, Lan SJ, Petridou E *et al*. Breast-feeding and maternal smoking in the etiology of Crohn's disease and ulcerative colitis in childhood. *Annals of Epidemiology* 1993;**3**:387-92.

606. Ceyhan BB, Karakurt S, Cevik H, Sungur M. Bronchial hyperreactivity and allergic status in inflammatory bowel disease. *Respiration* 2003;**70**:60-6.

607. Songur N, Songur Y, Tuzun M, Dogan I, Tuzun D, Ensari A *et al*. Pulmonary function tests and high-resolution CT in the detection of pulmonary involvement in inflammatory bowel disease. *Journal of Clinical Gastroenterology* 2003;**37**:292-8.

608. Douglas JG, McDonald CF, Leslie MJ, Gillon J, Crompton GK, McHardy GJ. Respiratory impairment in inflammatory bowel disease: does it vary with disease activity? *Respiratory Medicine* 1989;**83**:389-94.

609. Weng X, Liu L, Barcellos LF, Allison JE, Herrinton LJ. Clustering of inflammatory bowel disease with immune mediated diseases among members of a northern california-managed care organization. *American Journal of Gastroenterology* 2007;**102**:1429-35.

610. Becker KG, Simon RM, Bailey-Wilson JE, Freidlin B, Biddison WE, McFarland HF *et al*. Clustering of non-major histocompatibility complex susceptibility candidate loci in human autoimmune diseases. *Proceedings of the National Academy of Sciences of the United States of America* 1998;**95**:9979-84.

611. Minuk GY,.Lewkonia RM. Possible familial association of multiple sclerosis and inflammatory bowel disease. *New England Journal of Medicine* 1986;**314**:586.

612. Sadovnick AD, Paty DW, Yannakoulias G. Concurrence of multiple sclerosis and inflammatory bowel disease. *New England Journal of Medicine* 1989;**321**:762-3.

613. Book DT, Smith TL, McNamar JP, Saeian K, Binion DG, Toohill RJ. Chronic sinonasal disease in patients with inflammatory bowel disease. *American Journal of Rhinology* 2003;**17**:87-90.

614. Cominelli F. Early and late gut immune responses in IBD. *Journal of Pediatric Gastroenterology & Nutrition* 2008;**46**:E20.

615. Kugathasan S, Saubermann LJ, Smith L, Kou D, Itoh J, Binion DG *et al*. Mucosal T-cell immunoregulation varies in early and late inflammatory bowel disease. *Gut* 2007;**56**:1696-705.

616. Simpson CR, Anderson WJ, Helms PJ, Taylor MW, Watson L, Prescott GJ *et al*. Coincidence of immune-mediated diseases driven by Th1 and Th2 subsets suggests a common aetiology. A population-based study using computerized general practice data. *Clinical & Experimental Allergy* 2002;**32**:37-42.

617. Tirosh A, Mandel D, Mimouni FB, Zimlichman E, Shochat T, Kochba I. Autoimmune Diseases in Asthma. *Ann Intern Med* 2006;**144**:877-83.

618. Steinman L. A brief history of TH17, the first major revision in the TH1/TH2 hypothesis of T cell-mediated tissue damage. *Nat Med* 2007;**13**:139-45.

619. McCormick P,.Manning D. Chronic inflammatory bowel disease and the 'over-clean' environment: rarity in the Irish 'traveller' community. *Irish Medical Journal* 2001;**94**:203-4.

620. Green C, Elliott L, Beaudoin C, Bernstein CN. A Population-based Ecologic Study of Inflammatory Bowel Disease: Searching for Etiologic Clues. *Am.J.Epidemiol.* 2006;**164**:615-23.

621. Radon K, Windstetter D, Poluda AL, Mueller B, von Mutius E, Koletzko S *et al*. Contact with farm animals in early life and juvenile inflammatory bowel disease: a case-control study. *Pediatrics* 2007;**120**:354-61.

622. Gent AE, Hellier MD, Grace RH, Swarbrick ET, Coggon D. Inflammatory bowel disease and domestic hygiene in infancy. *Lancet* 1994;**343**:766-7.

623. Meerwaldt R, Odink RJ, Landaeta R, Aarts F, Brunekreef B, Gerritsen J *et al*. A lower prevalence of atopy symptoms in children with type 1 diabetes mellitus. *Clinical & Experimental Allergy* 2002;**32**:254-5.

624. Amre DK, Lambrette P, Law L, Krupoves A, Chotard V, Costea F *et al*. Investigating the hygiene hypothesis as a risk factor in pediatric onset Crohn's disease: a case-control study. *American Journal of Gastroenterology* 2006;**101**:1005-11.

625. Klement E,.Reif S. Breastfeeding and risk of inflammatory bowel disease. *Am J Clin Nutr* 2005;**82**:486.

626. Kramer MS,.Kakuma R. The optimal duration of exclusive breastfeeding: a systematic review. *Advances in Experimental Medicine & Biology* 2004;63-77.

627. Bjorksten B. Impact of gastrointestinal flora on systemic diseases. *Journal of Pediatric Gastroenterology & Nutrition* 2008;**46**:E12-E13.

628. Kramer MS, Matush L, Vanilovich I, Platt R, Bogdanovich N, Sevkovskaya Z *et al*. Effect of prolonged and exclusive breast feeding on risk of allergy and asthma: cluster randomised trial. *BMJ* 2007;**335**:815.

629. Bjorksten B, Sepp E, Julge K, Voor T, Mikelsaar M. Allergy development and the intestinal microflora during the first year of life. *Journal of Allergy & Clinical Immunology* 2001;**108**:516-20.

630. Sepp E, Julge K, Mikelsaar M, Bjorksten B. Intestinal microbiota and immunoglobulin E responses in 5-year-old Estonian children. *Clinical & Experimental Allergy* 2005;**35**:1141-6.

631. Wang M, Karlsson C, Olsson C, Adlerberth I, Wold AE, Strachan DP *et al*. Reduced diversity in the early fecal microbiota of infants with atopic eczema. *Journal of Allergy & Clinical Immunology* 2008;**121**:129-34.

632. Betsi GI, Papadavid E, Falagas ME. Probiotics for the treatment or prevention of atopic dermatitis: a review of the evidence from randomized controlled trials. *American Journal of Clinical Dermatology* 2008;**9**:93-103.

633. Bjorksten B. Evidence of probiotics in prevention of allergy and asthma. *Current Drug Targets - Inflammation & Allergy* 2005;**4**:599-604.

634. Benn CS, Thorsen P, Jensen JS, Kjaer BB, Bisgaard H, Andersen M *et al*. Maternal vaginal microflora during pregnancy and the risk of asthma hospitalization and use of antiasthma medication in early childhood. *Journal of Allergy & Clinical Immunology* 2002;**110**:72-7.

635. van Beijsterveldt TC,.Boomsma DI. Asthma and mode of birth delivery: a study in 5-year-old dutch twins. *Twin Research & Human Genetics: the Official Journal of the International Society for Twin Studies* 2008;**11**:156-60.

636. Hakansson S,.Kallen K. Caesarean section increases the risk of hospital care in childhood for asthma and gastroenteritis. *Clinical & Experimental Allergy* 2003;**33**:757-64.

637. Renz-Polster H, David MR, Buist AS, Vollmer WM, O'Connor EA, Frazier EA *et al*. Caesarean section delivery and the risk of allergic disorders in childhood. *Clinical & Experimental Allergy* 2005;**35**:1466-72.

638. Kilkkinen A, Virtanen SM, Klaukka T, Kenward MG, Salkinoja-Salonen M, Gissler M *et al*. Use of antimicrobials and risk of type 1 diabetes in a population-based mother-child cohort. *Diabetologia* 2006;**49**:66-70.

639. Larsson PG, Fahraeus L, Carlsson B, Jakobsson T, Forsum U. Predisposing factors for bacterial vaginosis, treatment efficacy and pregnancy outcome among term deliveries; results from a preterm delivery study. *BMC Women's Health* 2007;20.

640. Pronczuk J, Akre J, Moy G, Vallenas C. Global perspectives in breast milk contamination: infectious and toxic hazards. *Environmental Health Perspectives* 2002;**110**:A349-A351.

641. Koletzko S, Griffiths A, Corey M, Smith C, Sherman P. Infant feeding practices and ulcerative colitis in childhood. *BMJ* 1991;**302**:1580-1.

642. ACHESON ED,.Truelove SC. Early weaning in the aetiology of ulcerative colitis. A study of feeding in infancy in cases and controls. *British Medical Journal* 1961;**2**:929-33.

643. Bergstrand O,.Hellers G. Breast-feeding during infancy in patients who later develop Crohn's disease. *Scandinavian Journal of Gastroenterology* 1983;**18**:903-6.

644. Ekbom A, Adami HO, Helmick CG, Jonzon A, Zack MM. Perinatal risk factors for inflammatory bowel disease: a case-control study. *Am.J.Epidemiol.* 1990;**132**:1111-9.

645. Thompson NP, Pounder RE, Wakefield AJ. Perinatal and childhood risk factors for inflammatory bowel disease: a case-control study. *European Journal of Gastroenterology & Hepatology* 1995;**7**:385-90.

646. Thompson NP, Montgomery SM, Pounder RE, Wakefield AJ. Is measles vaccination a risk factor for inflammatory bowel disease? *Lancet* 1995;**345**:1071-4.

647. Wakefield AJ, Pittilo RM, Sim R, Cosby SL, Stephenson JR, Dhillon AP *et al*. Evidence of persistent measles virus infection in Crohn's disease. *Journal of Medical Virology* 1993;**39**:345-53.

648. Griffin DE,.Ward BJ. Differential CD4 T cell activation in measles. *Journal of Infectious Diseases* 1993;**168**:275-81.

649. Feeney M, Ciegg A, Winwood P, Snook J. A case-control study of measles vaccination and inflammatory bowel disease. The East Dorset Gastroenterology Group. *Lancet* 1997;**350**:764-6.

650. Robertson DJ,.Sandler RS. Measles virus and Crohn's disease: a critical appraisal of the current literature. *Inflammatory Bowel Diseases* 2001;**7**:51-7.

651. Hermon-Taylor J, Ford J, Sumar N, Millar D, Doran T, Tizard M. Measles virus and Crohn's disease. *Lancet* 1995;**345**:922-3.

652. Wakefield AJ, Murch SH, Anthony A, Linnell J, Casson DM, Malik M *et al*. Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *Lancet* 1998;**351**:637-41.

653. Sabra A, Bellanti JA, Colon AR. Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *Lancet* 1998;**352**:234-5.

654. Chen RT,.DeStefano F. Vaccine adverse events: causal or coincidental? *Lancet* 1998;**351**:611-2.

655. Jansen VAA, Stollenwerk N, Jensen HJ, Ramsay ME, Edmunds WJ, Rhodes CJ. Measles Outbreaks in a Population with Declining Vaccine Uptake. *Science* 2003;**301**:804.

656. Mulholland EK. Measles in the United States, 2006. *N Engl J Med* 2006;**355**:440-3.

657. Brady MT. Immunization Recommendations for Children With Metabolic Disorders: More Data Would Help. *Pediatrics* 2006;**118**:810-3.

658. Offit PA. Vaccines and Autism Revisited -- The Hannah Poling Case. *N Engl J Med* 2008;**358**:2089-91.

659. Poling JS,.Offit PA. Vaccines and Autism Revisited. *N Engl J Med* 2008;**359**:655-6.

660. Russell RK, Drummond HE, Nimmo ER, Anderson N, Smith L, Wilson DC *et al*. Genotype-phenotype analysis in childhood-onset Crohn's disease: NOD2/CARD15 variants consistently predict phenotypic characteristics of severe disease. *Inflammatory Bowel Diseases* 2005;**11**:955-64.

661. Abreu MT, Taylor KD, Lin YC, Hang T, Gaiennie J, Landers CJ *et al*. Mutations in NOD2 are associated with fibrostenosing disease in patients with Crohn's disease. *Gastroenterology* 2002;**123**:679-88.

662. Newman B, Silverberg MS, Gu X, Zhang Q, Lazaro A, Steinhart AH *et al*. CARD15 and HLA DRB1 alleles influence susceptibility and disease localization in Crohn's disease. *American Journal of Gastroenterology* 2004;**99**:306-15.

663. Vermeire S, Wild G, Kocher K, Cousineau J, Dufresne L, Bitton A *et al*. CARD15 genetic variation in a Quebec population: prevalence, genotype-phenotype relationship, and haplotype structure. *American Journal of Human Genetics* 2002;**71**:74-83.

664. Helgason A, Hickey E, Goodacre S, Bosnes V, Stefansson K, Ward R *et al*. mtDna and the islands of the North Atlantic: estimating the proportions of Norse and Gaelic ancestry. *American Journal of Human Genetics* 2001;**68**:723-37.

665. Fritz JH, Ferrero RL, Philpott DJ, Girardin SE. Nod-like proteins in immunity, inflammation and disease. *Nature Immunology* 2006;**7**:1250-7.

666. Van Limbergen J, Russell RK, Nimmo ER, Wilson DC, Ho G-T, Arnott ID *et al*. Genetics of the innate immune response in inflammatory bowel disease. *Inflammatory Bowel Diseases* 2007;**13**:338-55.

667. Opitz B, Forster S, Hocke AC, Maass M, Schmeck B, Hippenstiel S *et al*. Nod1-mediated endothelial cell activation by Chlamydophila pneumoniae. *Circulation Research* 2005;**96**:319-26.

668. Kim JG, Lee SJ, Kagnoff MF. Nod1 is an essential signal transducer in intestinal epithelial cells infected with bacteria that avoid recognition by toll-like receptors. *Infection & Immunity* 2004;**72**:1487-95.

669. Boughan PK, Argent RH, Body-Malapel M, Park J-H, Ewings KE, Bowie AG *et al*. Nucleotide-binding Oligomerization Domain-1 and Epidermal Growth Factor Receptor - critical regulators of defensins during Helicobacter Pylori infection. *J.Biol.Chem.* 2006;**281**:11637-48.

670. Fritz JH, Le Bourhis L, Sellge G, Magalhaes JG, Fsihi H, Kufer TA *et al*. Nod1-mediated innate immune recognition of peptidoglycan contributes to the onset of adaptive immunity. *Immunity* 2007;**26**:445-59.

671. Kim YG, Park JH, Shaw MH, Franchi L, Inohara N, Nunez G. The Cytosolic Sensors Nod1 and Nod2 Are Critical for Bacterial Recognition and Host Defense after Exposure to Toll-like Receptor Ligands. *Immunity* 2008;**28**:246-57.

672. Fritz JH, Le Bourhis L, Magalhaes JG, Philpott DJ. Innate immune recognition at the epithelial barrier drives adaptive immunity: APCs take the back seat. *Trends in Immunology* 2008;**29**:41-9.

673. Molnar T, Hofner P, Nagy F, Lakatos PL, Fischer S, Lakatos L *et al*. NOD1 gene E266K polymorphism is associated with disease susceptibility but not with disease phenotype or NOD2/CARD15 in Hungarian patients with Crohn's disease. *Digestive & Liver Disease* 2007;**39**:1064-70.

674. Van Limbergen J., Russell R.K., Nimmo E.R., Törkvist L, Lees CW, Drummond H.E. *et al*. Contribution of the NOD1/CARD4 insertion/deletion polymorphism +32656 to inflammatory bowel disease in Northern Europe. *Inflammatory Bowel Diseases* 2007;**13**:882-9.

675. Knapp M. A note on power approximations for the transmission/disequilibrium test. *American Journal of Human Genetics* 1999;**64**:1177-85.

676. Rubin GP, Hungin APS, Kelly PJ, Ling J. Inflammatory bowel disease: epidemiology and management in an English general practice population. *Alimentary Pharmacology & Therapeutics* 2000;**14**:1553-9.

677. Mitchell AA, Cutler DJ, Chakravarti A. Undetected genotyping errors cause apparent overtransmission of common alleles in the transmission/disequilibrium test. *American Journal of Human Genetics* 2003;**72**:598-610.

678. Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nature Genetics* 2003;**33**:177-82.

679. Ioannidis JPA, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Replication validity of genetic association studies. *Nat Genet* 2001;**29**:306-9.

680. McGovern DP, Butler H, Ahmad T, Paolucci M, van Heel DA, Negoro K *et al*. TUCAN (CARD8) genetic variants and inflammatory bowel disease. *Gastroenterology* 2006;**131**:1190-6.

681. The International HapMap Consortium. A haplotype map of the human genome. *Nature* 2005;**437**:1299-320.

682. Henckaerts L, Pierik M, Joossens M, Ferrante M, Rutgeerts P, Vermeire S. Mutations in pattern recognition receptor genes modulate seroreactivity to microbial antigens in patients with inflammatory bowel disease. *Gut* 2007;5**6**:1536-42.

683. Zeggini E, Rayner W, Morris AP, Hattersley AT, Walker M, Hitman GA *et al*. An evaluation of HapMap sample size and tagging SNP performance in large-scale empirical and simulated data sets. *Nat Genet* 2005;**37**:1320-2.

684. de Bakker PIW, Burtt NP, Graham RR, Guiducci C, Yelensky R, Drake JA *et al*. Transferability of tag SNPs in genetic association studies in multiple populations. *Nat Genet* 2006;**38**:1298-303.

685. Need AC,.Goldstein DB. Genome-wide tagging for everyone. *Nat Genet* 2006;**38**:1227-8.

686. Conrad DF, Jakobsson M, Coop G, Wen X, Wall JD, Rosenberg NA *et al*. A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. *Nat Genet* 2006;**38**:1251-60.

687. Rioux JD. Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nature Genet.* 2007;**39**:596-604.

688. Hampe J, Franke A, Rosenstiel P, Till A, Teuber M, Huse K *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.

689. Tato CM,.Cua DJ. Reconciling id, ego, and superego within interleukin-23. *Immunological Reviews* 2008;**226**:103-11.

690. Oppmann B. Novel p19 protein engages IL-12p40 to form a cytokine, IL-23, with biological activities similar as well as distinct from IL-12. *Immunity* 2000;**13**:715-25.

691. Parham C. A receptor for the heterodimeric cytokine IL-23 is composed of IL-12R[bgr]1 and a novel cytokine receptor subunit, IL-23R. *J.Immunol.* 2002;**168**:5699-708.

692. Beyer BM, Ingram R, Ramanathan L, Reichert P, Le HV, Madison V *et al*. Crystal structures of the pro-inflammatory cytokine interleukin-23 and its complex with a high-affinity neutralizing antibody. *Journal of Molecular Biology* 2008;**382**:942-55.

693. Cho JH. The genetics and immunopathogenesis of inflammatory bowel disease. *Nature Reviews.Immunology* 2008;**8**:458-66.

694. Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, Rioux JD *et al*. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet* 2008;**40**:955-62.

695. Mitsuyama K, Matsumoto S, Rose-John S, Suzuki A, Hara T, Tomiyasu N *et al*. STAT3 activation via interleukin 6 trans-signalling contributes to ileitis in SAMP1/Yit mice. *Gut* 2006;**55**:1263-9.

696. Liu Y, Zhang P, Li J, Kulkarni AB, Perruche S, Chen W. A critical function for TGF-beta signaling in the development of natural CD4+CD25+Foxp3+ regulatory T cells. *Nat Immunol* 2008;**9**:632-40.

697. O'Garra A, Stockinger B, Veldhoen M. Differentiation of human T(H)-17 cells does require TGF-beta! *Nature Immunology* 2008;**9**:588-90.

698. Zhou L. TGF-beta-induced Foxp3 inhibits TH17 cell differentiation by antagonizing ROR-gamma-t function. *Nature* 2008;**453**:236-40.

699. Sandborn WJ, Feagan BG, Fedorak RN, Scherl E, Fleisher MR, Katz S *et al*. A randomized trial of Ustekinumab, a human interleukin-12/23 monoclonal antibody, in patients with moderate-to-severe Crohn's disease. *Gastroenterology* 2008;**135**:1130-41.

700. Tremelling M, Cummings F, Fisher SA, Mansfield J, Gwilliam R, Keniry A *et al*. IL23R variation determines susceptibility but not disease phenotype in inflammatory bowel disease. *Gastroenterology* 2007;**132**:1657-64.

701. Glas J, Seiderer J, Wetzke M, Konrad A, Torok HP, Schmechel S *et al*. 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* 2007;**2**:e819.

702. Fisher SA, Tremelling M, Anderson CA, Gwilliam R, Bumpstead S, Prescott NJ *et al*. Genetic determinants of ulcerative colitis include the ECM1 locus and five loci implicated in Crohn's disease. *Nat Genet* 2008;**40**:710-2.

703. Anderson CA, Massey DC, Barrett JC, Prescott NJ, Tremelling M, Fisher SA *et al*. Investigation of Crohn's disease risk Loci in ulcerative colitis further defines their molecular relationship. *Gastroenterology* 2009;**136**:523-9.

704. Raelson JV, Little RD, Ruether A, Fournier H, Paquin B, Van Eerdewegh P *et al*. Genome-wide association study for Crohn's disease in the Quebec Founder Population identifies multiple validated disease loci. *Proceedings of the National Academy of Sciences of the United States of America* 2007;**104**:14747-52.

705. Taylor KD, Targan SR, Mei L, Ippoliti AF, McGovern D, Mengesha E *et al*. IL23R haplotypes provide a large population attributable risk for Crohn's disease. *Inflammatory Bowel Diseases* 2008;**14**:1185-91.

706. Van Limbergen JE, Russell RK, Nimmo ER, Drummond HE, Smith L, Anderson NH *et al*. IL23R Arg381Gln is associated with childhood onset inflammatory bowel disease in Scotland. *Gut* 2007;**56**:1173-4.

707. Van Limbergen J, Russell RK, Nimmo ER, Drummond HE, Smith L, Anderson NH *et al*. A detailed haplotype tagging investigation of the IL23R gene confirms gene-wide association with childhood onset IBD and CD. *Gastroenterology* 2008;**134**:A459.

708. Van Limbergen J, Russell RK, Nimmo ER, Torkvist L, Lees CW, Drummond HE *et al*. Contribution of the NOD1/CARD4 insertion/deletion polymorphism +32656 to inflammatory bowel disease in northern Europe. *Inflammatory Bowel Diseases* 2007;**13**:882-9.

709. Russell RK, Drummond HE, Wilson DC, Anderson NH, Arnott IDR, Van Limbergen JE *et al*. Detailed assessment of NOD2/CARD15 exonic variation in inflammatory bowel disease in Scotland: implications for disease pathogenesis. *Genes Immun* 2008;**9**:556-60.

710. Van Limbergen J, Russell RK, Drummond HE, Aldhous MC, Round NK, Nimmo ER *et al*. Definition of phenotypic characteristics of childhood-onset inflammatory bowel disease. *Gastroenterology* 2008;**135**:1114-22.

711. Van Limbergen J, Nimmo ER, Russell RK, Drummond HE, Smith L, Anderson NH *et al*. Investigation of NOD1/CARD4 variation in inflammatory bowel disease using a haplotype-tagging strategy. *Hum.Mol.Genet.* 2007;**16**:2175-86.

712. Dubinsky MC, Wang D, Picornell Y, Wrobel I, Katzir L, Quiros A *et al*. IL-23 receptor (IL-23R) gene protects against pediatric Crohn's disease. *Inflammatory Bowel Diseases* 2007;**13**:511-5.

713. Baldassano RN, Bradfield JP, Monos DS, Kim CE, Glessner JT, Casalunovo T *et al*. Association of variants of the interleukin-23 receptor gene with susceptibility to pediatric Crohn's disease. *Clinical Gastroenterology & Hepatology* 2007;**5**:972-6.

714. Oliver J, Rueda B, Lopez-Nevot MA, Gomez-Garcia M, Martin J. Replication of an association between IL23R gene polymorphism with inflammatory bowel disease. *Clinical Gastroenterology & Hepatology* 2007;**5**:977-81.

715. Buning C, Schmidt HH, Molnar T, De Jong DJ, Fiedler T, Buhner S *et al*. Heterozygosity for IL23R p.Arg381Gln confers a protective effect not only against Crohn's disease but also ulcerative colitis. *Alimentary Pharmacology & Therapeutics* 2007;**26**:1025-33.

716. Leshinsky-Silver E, Karban A, Dalal I, Eliakim R, Shirin H, Tzofi T *et al*. Evaluation of the interleukin-23 receptor gene coding variant R381Q in pediatric and adult Crohn disease. *Journal of Pediatric Gastroenterology & Nutrition* 2007;**45**:405-8.

717. Amre DK, Mack D, Israel D, Morgan K, Lambrette P, Law L *et al*. Association Between Genetic Variants in the IL-23R Gene and Early-Onset Crohn's Disease: Results From a Case-Control and Family-Based Study Among Canadian Children. *The American Journal of Gastroenterology* 2008;**103**:615-20.

718. Baptista ML, Amarante H, Picheth G, Sdepanian VL, Peterson N, Babasukumar U *et al*. CARD15 and IL23R influences Crohn's disease susceptibility but not disease phenotype in a Brazilian population. *Inflammatory Bowel Diseases* 2008;**14**:674-9.

719. Lappalainen M, Halme L, Turunen U, Saavalainen P, Einarsdottir E, Farkkila M *et al*. Association of il23r, tnfrsf1a, and hla-drb1*0103 allele variants with inflammatory bowel disease phenotypes in the finnish population. *Inflammatory Bowel Diseases* 2008;**14**:1118-24.

720. Einarsdottir E, Koskinen LL, Dukes E, Kainu K, Suomela S, Lappalainen M *et al*. IL23R in the Swedish, Finnish, Hungarian and Italian populations: association with IBD and psoriasis, and linkage to celiac disease. *BMC Medical Genetics.10:8,* 2009;**10**:doi:10.1186/1471-2350-10-8.

721. Marquez A, Mendoza JL, Taxonera C, Diaz-Rubio M, de la Concha EG, Urcelay E *et al*. IL23R and IL12B polymorphisms in Spanish IBD patients: no evidence of interaction. *Inflammatory Bowel Diseases* 2008;**14**:1192-6.

722. Lakatos PL, Szamosi T, Szilvasi A, Molnar E, Lakatos L, Kovacs A *et al*. ATG16L1 and IL23 receptor (IL23R) genes are associated with disease susceptibility in Hungarian CD patients. *Digestive & Liver Disease* 2008;**40**:867-73.

723. Okazaki T, Wang MH, Rawsthorne P, Sargent M, Datta LW, Shugart YY *et al*. Contributions of IBD5, IL23R, ATG16L1, and NOD2 to Crohn's disease risk in a population-based case-control study: evidence of gene-gene interactions. *Inflammatory Bowel Diseases* 2008;**14**:1528-41.

724. Latiano A, Palmieri O, Valvano MR, D'Inca R, Cucchiara S, Riegler G *et al*. Replication of interleukin 23 receptor and autophagy-related 16-like 1 association in adult- and pediatric-onset inflammatory bowel disease in Italy. *World Journal of Gastroenterology* 2008;**14**:4643-51.

725. Borgiani P, Perricone C, Ciccacci C, Romano S, Novelli G, Biancone L *et al*. Interleukin-23R Arg381Gln is associated with susceptibility to Crohn's disease but not with phenotype in an Italian population. *Gastroenterology.133(3):1049-51* 2007;**133**:1049-51.

726. Weersma RK, Stokkers PC, Cleynen I, Wolfkamp SC, Henckaerts L, Schreiber S *et al*. Confirmation of multiple Crohn's disease susceptibility loci in a large Dutch-Belgian cohort. *American Journal of Gastroenterology* 2009;**104**:630-8.

727. Franke A, Balschun T, Karlsen TH, Hedderich J, May S, Lu T *et al*. Replication of signals from recent studies of Crohn's disease identifies previously unknown disease loci for ulcerative colitis. *Nat Genet* 2008;**40**:713-5.

728. Silverberg MS, Cho JH, Rioux JD, McGovern DP, Wu J, Annese V *et al*. Ulcerative colitis-risk loci on chromosomes 1p36 and 12q15 found by genome-wide association study. *Nature Genetics* 2009;**41**:216-20.

729. Franke A, Balschun T, Karlsen TH, Sventoraityte J, Nikolaus S, Mayr G *et al*. Sequence variants in IL10, ARPC2 and multiple other loci contribute to ulcerative colitis susceptibility. *Nat Genet* 2008;**40**:1319-23.

730. Roberts RL, Gearry RB, Hollis-Moffatt JE, Miller AL, Reid J, Abkevich V *et al*. IL23R R381Q and ATG16L1 T300A Are Strongly Associated With

Crohn's Disease in a Study of New Zealand Caucasians With Inflammatory Bowel Disease. *The American Journal of Gastroenterology* 2007;**102**:2754-61.

731. Weersma RK, Zhernakova A, Nolte IM, Lefebvre C, Rioux JD, Mulder F *et al*. ATG16L1 and IL23R Are Associated With Inflammatory Bowel Diseases but Not With Celiac Disease in The Netherlands. *The American Journal of Gastroenterology* 2008;**103**:621-7.

732. Kugathasan S, Baldassano RN, Bradfield JP, Sleiman PMA, Imielinski M, Guthery SL *et al*. Loci on 20q13 and 21q22 are associated with pediatric-onset inflammatory bowel disease. *Nat Genet* 2008;**40**:1211-5.

733. Yamazaki K, Onouchi Y, Takazoe M, Kubo M, Nakamura Y, Hata A. Association analysis of genetic variants in IL23R, ATG16L1 and 5p13.1 loci with Crohn's disease in Japanese patients. *Journal of Human Genetics* 2007;**52**:575-83.

734. Kan S, Mancini G, Gallagher G. Identification and characterization of multiple splice forms of the human interleukin-23 receptor [alpha] chain in mitogen-activated leukocytes. *Genes Immun* 2008;**9**:631-9.

735. Zhang XY, Zhang HJ, Zhang Y, Fu YJ, He J, Zhu LP *et al*. Identification and expression analysis of alternatively spliced isoforms of human interleukin-23 receptor gene in normal lymphoid cells and selected tumor cells. *Immunogenetics* 2006;**57**:934-43.

736. Wang K, Zhang H, Kugathasan S, Annese V, Bradfield JP, Russell RK *et al*. Diverse genome-wide association studies associate the IL12/IL23 pathway with Crohn Disease. *American Journal of Human Genetics* 2009;**84**:399-405.

737. McGovern DP, Rotter JI, Mei L, Haritunians T, Landers C, Derkowski C *et al*. Genetic epistasis of IL23/IL17 pathway genes in Crohn's disease. *Inflammatory Bowel Diseases* 2009;**15**:883-9.

738. Ahern PP, Izcue A, Maloy KJ, Powrie F. The interleukin-23 axis in intestinal inflammation. *Immunological Reviews* 2008;**226**:147-59.

739. Cua DJ. Interleukin-23 rather than interleukin-12 is the critical cytokine for autoimmune inflammation of the brain. *Nature* 2003;**421**:744-8.

740. Kullberg MC, Jankovic D, Feng CG, Hue S, Gorelick PL, McKenzie BS *et al*. IL-23 plays a key role in Helicobacter hepaticus-induced T cell-dependent colitis. *J.Exp.Med.* 2006;**203**:2485-94.

741. Hue S, Ahern P, Buonocore S, Kullberg MC, Cua DJ, McKenzie BS *et al*. Interleukin-23 drives innate and T cell-mediated intestinal inflammation. *J.Exp.Med.* 2006;**203**:2473-83.

742. Uhlig HH. Differential activity of IL-12 and IL-23 in mucosal and systemic innate immune pathology. *Immunity* 2006;**25**:309-18.

743. Yen D, Cheung J, Scheerens H, Poulet F, McClanahan T, McKenzie B *et al*. IL-23 is essential for T cell-mediated colitis and promotes inflammation via IL-17 and IL-6. *J.Clin.Invest.* 2006;**116**:1310-6.

744. Raz E. Organ-specific regulation of innate immunity. *Nature Immunology* 2007;**8**:3-4.

745. Langrish CL, McKenzie BS, Wilson NJ, de Waal MR, Kastelein RA, Cua DJ. IL-12 and IL-23: master regulators of innate and adaptive immunity. *Immunological Reviews.* 2004;96-105.

746. van Beelen AJ. Stimulation of the intracellular bacterial sensor NOD2 programs dendritic cells to promote interleukin-17 production in human memory T cells. *Immunity* 2007;**27**:660-9.

747. LeibundGut-Landmann S, Grosz O, Robinson MJ, Osorio F, Slack EC, Tsoni SV *et al*. Syk- and CARD9-dependent coupling of innate immunity to the induction of T helper cells that produce interleukin 17. *Nat Immunol* 2007;**8**:630-8.

748. Veldhoen M, Hocking RJ, Flavell RA, Stockinger B. Signals mediated by transforming growth factor-beta initiate autoimmune encephalomyelitis, but chronic inflammation is needed to sustain disease. *Nature Immunol.* 2006;**7**:1151-6.

749. Fuss IJ, Becker C, Yang Z, Groden C, Hornung RL, Heller F *et al*. 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. *Inflammatory Bowel Diseases* 2006;**12**:9-15.

750. Fujino S. Increased expression of interleukin 17 in inflammatory bowel disease. *Gut* 2003;**52**:65-70.

751. Kobayashi T, Okamoto S, Hisamatsu T, Kamada N, Chinen H, Saito R *et al*. IL23 differentially regulates the Th1/Th17 balance in ulcerative colitis and Crohn's disease. *Gut* 2008;**57**:1682-9.

752. Kamada N, Hisamatsu T, Okamoto S, Chinen H, Kobayashi T, Sato T *et al*. Unique CD14 intestinal macrophages contribute to the pathogenesis of Crohn disease via IL-23/IFN-gamma axis. *J.Clin.Invest.* 2008;**118**:2269-80.

753. Seiderer J, Elben I, Diegelmann J, Glas J, Stallhofer J, Tillack C *et al*. Role of the novel Th17 cytokine IL-17F in inflammatory bowel disease (IBD): upregulated colonic IL-17F expression in active Crohn's disease and analysis of the IL17F p.His161Arg polymorphism in IBD. *Inflammatory Bowel Diseases* 2008;**14**:437-45.

754. Holtta V, Klemetti P, Sipponen T, Westerholm-Ormio M, Kociubinski G, Salo H *et al*. IL-23/IL-17 immunity as a hallmark of Crohn's disease. *Inflammatory Bowel Diseases* 2008;**14**:1175-84.

755. Schmechel S, Konrad A, Diegelmann J, Glas J, Wetzke M, Paschos E *et al*. Linking genetic susceptibility to Crohn's disease with Th17 cell function: IL-22 serum levels are increased in Crohn's disease and correlate with disease activity and IL23R genotype status. *Inflammatory Bowel Diseases* 2008;**14**:204-12.

756. Mizushima N, Levine B, Cuervo AM, Klionsky DJ. Autophagy fights disease through cellular self-digestion. *Nature* 2008;**451**:1069-75.

757. Klionsky DJ. Autophagy: from phenomenology to molecular understanding in less than a decade. *Nat Rev Mol Cell Biol* 2007;**8**:931-7.

758. Franke A, Hampe J, Rosenstiel P, Becker C, Wagner F, Häsler R *et al*. Systematic association mapping identifies NELL1 as a novel IBD disease gene. *PLos ONE* 2007;**2**:e691. doi:10.1371/journal.pone.0000691.

759. Duerr RH. Genome-wide association studies herald a new era of rapid discoveries in inflammatory bowel disease research. *Gastroenterology* 2007;**132**:2045-9.

760. Schmid D,.Munz C. Innate and adaptive immunity through autophagy. *Immunity* 2007;**27**:11-21.

761. Levine B,.Deretic V. Unveiling the roles of autophagy in innate and adaptive immunity. *Nat Rev Immunol* 2007;**7**:767-77.

762. Xie Z,.Klionsky DJ. Autophagosome formation: core machinery and adaptations. *Nature Cell Biology* 2007;**9**:1102-9.

763. Kim J, Huang W-P, Stromhaug PE, Klionsky DJ. Convergence of multiple autophagy and cytoplasm to vacuole targeting components to a perivacuolar membrane compartment prior to de novo vesicle formation. *J.Biol.Chem.* 2002;**277**:763-73.

764. Suzuki K. The pre-autophagosomal structure organized by concerted functions of APG genes is essential for autophagosome formation. *EMBO J.* 2001;**20**:5971-81.

765. Mizushima N. Dissection of autophagosome formation using Apg5-deficient mouse embryonic stem cells. *J.Cell Biol.* 2001;**152**:657-68.

766. Mizushima N, Kuma A, Kobayashi Y, Yamamoto A, Matsubae M, Takao T *et al*. Mouse Apg16L, a novel WD-repeat protein, targets to the autophagic isolation membrane with the Apg12-Apg5 conjugate. *J Cell Sci* 2003;**116**:1679-88.

767. Kirisako T, Baba M, Ishihara N, Miyazawa K, Ohsumi M, Yoshimori T *et al*. Formation Process of Autophagosome Is Traced with Apg8/Aut7p in Yeast. *J.Cell Biol.* 1999;**147**:435-46.

768. Itoh T, Fujita N, Kanno E, Yamamoto A, Yoshimori T, Fukuda M. Golgi-resident Small GTPase Rab33B Interacts with Atg16L and Modulates Autophagosome Formation. *Molecular Biology of the Cell* 2008;**19**:2916-25.

769. Fujita N, Itoh T, Omori H, Fukuda M, Noda T, Yoshimori T. The Atg16L Complex Specifies the Site of LC3 Lipidation for Membrane Biogenesis in Autophagy. *Molecular Biology of the Cell* 2008;**19**:2092-100.

770. Delgado M, Singh S, De Haro S, Master S, Ponpuak M, Dinkins C *et al*. Autophagy and pattern recognition receptors in innate immunity. *Immunological Reviews* 2009;**227**:189-202.

771. Kabeya Y. LC3, a mammalian homologue of yeast Apg8p, is localized in autophagosome membranes after processing. *EMBO J.* 2000;**19**:5720-8.

772. Yorimitsu T,.Klionsky DJ. Autophagy: molecular machinery for self-eating. *Cell Death Differ.* 2005;**12**:1542-52.

773. Klionsky DJ, Abeliovich H, Agostinis P, Agrawal DK, Aliev G, Askew DS *et al*. Guidelines for the use and interpretation of assays for monitoring autophagy in higher eukaryotes. *Autophagy* 2008;**4**:151-75.

774. Mizushima N, Noda T, Yoshimori T, Tanaka Y, Ishii T, George MD *et al*. A protein conjugation system essential for autophagy. *Nature* 1998;**395**:395-8.

775. Guan J, Stromhaug PE, George MD, Habibzadegah-Tari P, Bevan A, Dunn WA, Jr. *et al*. Cvt18/Gsa12 is required for cytoplasm-to-vacuole transport, pexophagy, and autophagy in Saccharomyces cerevisiae and Pichia pastoris. *Molecular Biology of the Cell* 2001;**12**:3821-38.

776. Baba M, Osumi M, Scott SV, Klionsky DJ, Ohsumi Y. Two distinct pathways for targeting proteins from the cytoplasm to the vacuole/lysosome. *J.Cell Biol.* 1997;**139**:1687-95.

777. Ogawa M, Yoshimori T, Suzuki T, Sagara H, Mizushima N, Sasakawa C. Escape of intracellular Shigella from autophagy. *Science* 2005;**307**:727-31.

778. Shintani T, Huang WP, Stromhaug PE, Klionsky DJ. Mechanism of cargo selection in the cytoplasm to vacuole targeting pathway. *Developmental Cell* 2002;**3**:825-37.

779. Sanjuan MA, Dillon CP, Tait SWG, Moshiach S, Dorsey F, Connell S *et al*. Toll-like receptor signalling in macrophages links the autophagy pathway to phagocytosis. *Nature* 2007;**450**:1253-7.

780. Pankiv S, Clausen TH, Lamark T, Brech A, Bruun JA, Outzen H *et al*. p62/SQSTM1 Binds Directly to Atg8/LC3 to Facilitate Degradation of Ubiquitinated Protein Aggregates by Autophagy. *J.Biol.Chem.* 2007;**282**:24131-45.

781. Munz C. Enhancing Immunity Through Autophagy. *Annual Review of Immunology* 2009;**27**.

782. de Duve C,.Wattiaux R. Functions of lysosomes. *Annu.Rev.Physiol.* 1966;**28**:435-92.

783. Ashford TP,.Porter KR. Cytoplasmic components in hepatic cell lysosomes. *J.Cell Biol.* 1962;**12**:198-202.

784. Deter RL, Baudhuin P, de Duve C. Participation of lysosomes in cellular autophagy induced in rat liver by glucagon. *J.Cell Biol.* 1967;**35**:C11-C16.

785. Bolender RP,.Weibel ER. A morphometric study of the removal of phenobarbital-induced membranes from hepatocytes after cessation of treatment. *J.Cell Biol.* 1973;**56**:746-61.

786. Beaulaton J,.Lockshin RA. Ultrastructural study of the normal degeneration of the intersegmental muscles of Anthereae polyphemus and Manduca sexta (Insecta, Lepidoptera) with particular reference of cellular autophagy. *J.Morphol.* 1977;**154**:39-57.

787. Veenhuis M, Douma A, Harder W, Osumi M. Degradation and turnover of peroxisomes in the yeast Hansenula polymorpha induced by selective inactivation of peroxisomal enzymes. *Arch.Microbiol.* 1983;**134**:193-203.

788. Pfeifer U. Inhibition by insulin of the physiological autophagic breakdown of cell organelles. *Acta Biol.Med.Ger.* 1977;**36**:1691-4.

789. Mortimore GE,.Schworer CM. Induction of autophagy by amino-acid deprivation in perfused rat liver. *Nature* 1977;**270**:174-6.

790. Seglen PO,.Gordon PB. 3-methyladenine: specific inhibitor of autophagic/lysosomal protein degradation in isolated rat hepatocytes. *Proc.Natl Acad.Sci.USA* 1982;**79**:1889-92.

791. Holen I, Gordon PB, Seglen PO. Protein kinase-dependent effects of okadaic acid on hepatocytic autophagy and cytoskeletal integrity. *Biochem.J.* 1992;**284**:633-6.

792. Kang C, You YJ, Avery L. Dual roles of autophagy in the survival of Caenorhabditis elegans during starvation. *Genes & Development* 2007;**21**:2161-71.

793. Kunz J. Target of rapamycin in yeast, TOR2, is an essential phosphatidylinositol kinase homolog required for G1 progression. *Cell* 1993;**73**:585-96.

794. Blommaart EF, Luiken JJ, Blommaart PJ, van Woerkom GM, Meijer AJ. Phosphorylation of ribosomal protein S6 is inhibitory for autophagy in isolated rat hepatocytes. *J.Biol.Chem.* 1995;**270**:2320-6.

795. Wullschleger S, Loewith R, Hall MN. TOR signaling in growth and metabolism. *Cell* 2006;**124**:471-84.

796. Proud CG. Signalling to translation: how signal transduction pathways control the protein synthetic machinery. *Biochem J* 2007;**403**:217-34.

797. Huang J,.Manning BD. The TSC1-TSC2 complex: a molecular switchboard controlling cell growth. *Biochem J* 2008;**412**:179-90.

798. Kim E, Goraksha-Hicks P, Li L, Neufeld TP, Guan KL. Regulation of TORC1 by Rag GTPases in nutrient response. *Nat Cell Biol* 2008;**10**:935-45.

799. Sancak Y, Peterson TR, Shaul YD, Lindquist RA, Thoreen CC, Bar-Peled L *et al*. The Rag GTPases bind raptor and mediate amino acid signaling to mTORC1. *Science* 2008;**320**:1496-501.

800. Sancak Y, Thoreen CC, Peterson TR, Lindquist RA, Kang SA, Spooner E *et al*. PRAS40 is an insulin-regulated inhibitor of the mTORC1 protein kinase. *Molecular Cell* 2007;**25**:903-15.

801. Ma L, Chen Z, Erdjument-Bromage H, Tempst P, Pandolfi PP. Phosphorylation and functional inactivation of TSC2 by Erk implications for tuberous sclerosis and cancer pathogenesis. *Cell* 2005;**121**:179-93.

802. Gwinn DM, Shackelford DB, Egan DF, Mihaylova MM, Mery A, Vasquez DS *et al*. AMPK phosphorylation of raptor mediates a metabolic checkpoint. *Molecular Cell* 2008;**30**:214-26.

803. Reggiori F,.Klionsky DJ. Autophagosomes: biogenesis from scratch? *Current Opinion in Cell Biology* 2005;**17**:415-22.

804. Young AR, Chan EY, Hu XW, Kochl R, Crawshaw SG, High S *et al*. Starvation and ULK1-dependent cycling of mammalian Atg9 between the TGN and endosomes. *J Cell Sci* 2006;**119**:3888-900.

805. Kamada Y, Funakoshi T, Shintani T, Nagano K, Ohsumi M, Ohsumi Y. Tor-mediated Induction of Autophagy Via an Apg1 Protein Kinase Complex. *J.Cell Biol.* 2000;**150**:1507-13.

806. Suzuki K,.Ohsumi Y. Molecular machinery of autophagosome formation in yeast, Saccharomyces cerevisiae. *FEBS Lett.* 2007;**581**:2156-61.

807. Cheong H, Nair U, Geng J, Klionsky DJ. The Atg1 Kinase Complex Is Involved in the Regulation of Protein Recruitment to Initiate Sequestering Vesicle Formation for Nonspecific Autophagy in Saccharomyces cerevisiae. *Molecular Biology of the Cell* 2008;**19**:668-81.

808. Kawamata T, Kamada Y, Kabeya Y, Sekito T, Ohsumi Y. Organization of the Pre-autophagosomal Structure Responsible for Autophagosome Formation. *Molecular Biology of the Cell* 2008;**19**:2039-50.

809. Hara T, Takamura A, Kishi C, Iemura Si, Natsume T, Guan JL *et al*. FIP200, a ULK-interacting protein, is required for autophagosome formation in mammalian cells. *J.Cell Biol.* 2008;**181**:497-510.

810. Byfield MP, Murray JT, Backer JM. hVps34 is a nutrient-regulated lipid kinase required for activation of p70 S6 kinase. *J.Biol.Chem.* 2005;**280**:33076-82.

811. Backer JM. The regulation and function of Class III PI3Ks: novel roles for Vps34. *Biochem J* 2008;**410**:1-17.

812. Liang XH. Protection against fatal Sindbis virus encephalitis by Beclin, a novel Bcl-2-interacting protein. *J.Virol.* 1998;**72**:8586-96.

813. Levine B, Sinha S, Kroemer G. Bcl-2 family members: dual regulators of apoptosis and autophagy. *Autophagy* 2008;**4**:600-6.

814. Pattingre S. Bcl-2 antiapoptotic proteins inhibit Beclin 1-dependent autophagy. *Cell* 2005;**122**:927-39.

815. Scherz-Shouvai R, Shvets E, Fass E, Shorer H, Gil L, Elazar Z. Reactive oxygen species are essential for autophagy and specifically regulate the activity of Atg4. *EMBO J.* 2007;**26**:1749-60.

816. Meijer AJ,.Codogno P. Signalling and autophagy regulation in health and disease. *Mol.Aspects.Med.* 2006;**27**:411-25.

817. Wei Y, Pattingre S, Sinha S, Bassik M, Levine B. JNK1-mediated phosphorylation of Bcl-2 regulates starvation-induced autophagy. *Molecular Cell* 2008;**30**:678-88.

818. Lee IH, Cao L, Mostoslavsky R, Lombard DB, Liu J, Bruns NE *et al*. A role for the NAD-dependent deacetylase Sirt1 in the regulation of autophagy. *Proceedings of the National Academy of Sciences* 2008;**105**:3374-9.

819. Meinzer UM, Idestrom MM, Alberti CM, Peuchmaur MM, Belarbi NM, Bellaiche MM *et al*. Ileal Involvement Is Age Dependent in Pediatric Crohn's Disease. *Inflammatory Bowel Diseases* 2005;**11**:639-44.

820. Kono H,.Rock KL. How dying cells alert the immune system to danger. *Nat Rev Immunol* 2008;**8**:279-89.

821. Grossmayer GE. Removal of dying cells and systemic lupus erythematosus. *Mod.Rheumatol.* 2005;**15**:383-90.

822. Qu X. Autophagy gene-dependent clearance of apoptotic cells during embryonic development. *Cell* 2007;**128**:931-46.

823. Gaipl US. Inefficient clearance of dying cells and autoreactivity. *Curr.Top.Microbiol.Immunol.* 2006;**305**:161-76.

824. Qing G, Yan P, Qu Z, Liu H, Xiao G. Hsp90 regulates processing of NF-kappaB2 p100 involving protection of NF-kappaB-inducing kinase (NIK) from autophagy-mediated degradation. *Cell Research* 2007;**17**:520-30.

825. Qing G, Yan P, Xiao G. Hsp90 inhibition results in autophagy-mediated proteasome-independent degradation of IkappaB kinase (IKK). *Cell Research* 2006;**16**:895-901.

826. Guo B, Chang EY, Cheng G. The type I IFN induction pathway constrains Th17-mediated autoimmune inflammation in mice. *J Clin Invest* 2008;**118**:1680-90.

827. Glas J, Konrad A, Schmechel S, Dambacher J, Seiderer J, Schroff F *et al*. The ATG16L1 Gene Variants rs2241879 and rs2241880 (T300A) Are Strongly Associated With Susceptibility to Crohn's Disease in the German Population. *The American Journal of Gastroenterology* 2008;**103**:682-91.

828. Prescott NJ, Fisher SA, Franke A, Hampe J, Onnie CM, Soars D *et al*. A nonsynonymous SNP in ATG16L1 predisposes to ileal Crohn's disease and is independent of CARD15 and IBD5. *Gastroenterology* 2007;**132**:1665-71.

829. Cummings JFR, Cooney R, Pathan S, Anderson CA, Barrett JC, Beckly J *et al*. Confirmation of the role of ATG16L1 as a Crohn's disease susceptibility gene. *Inflammatory Bowel Diseases* 2007;**13**:941-6.

830. Hait E, Bousvaros A, Grand R. Pediatric inflammatory bowel disease: what children can teach adults. *Inflammatory Bowel Diseases* 2005;**11**:519-27.

831. Russell RK, Van Limbergen JE, Drummond HE, Nimmo ER, Arnott IDR, Gillett PM *et al*. Disease phenotype but not genotype differs in children and adults with inflammatory bowel disease at diagnosis: analysis of more than 1500 patients from a homogenous patient cohort. *Journal of Pediatric Gastroenterology and Nutrition* 2007;**44** :e1-e360.

832. Cadwell K, Liu JY, Brown SL, Miyoshi H, Loh J, Lennerz JK *et al*. A key role for autophagy and the autophagy gene Atg16l1 in mouse and human intestinal Paneth cells. *Nature* 2008;**456**:259-63.

833. Saitoh T, Fujita N, Jang MH, Uematsu S, Yang BG, Satoh T *et al*. Loss of the autophagy protein Atg16L1 enhances endotoxin-induced IL-1beta production. *Nature* 2008;**456**:264-8.

834. Kuballa P, Huett A, Rioux JD, Daly MJ, Xavier RJ. Impaired autophagy of an intracellular pathogen induced by a Crohn's disease associated ATG16L1 variant. *PLos ONE* 2008;**3**:e3391.

835. 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.

836. Bekpen C, Hunn JP, Rohde C, Parvanova I, Guethlein L, Dunn DM *et al*. The interferon-inducible p47 (IRG) GTPases in vertebrates: loss of the cell autonomous resistance mechanism in the human lineage. *Genome Biology.6(11):R92,* 2005;**6**:R92.1-R92.18.

837. Singh SB, Davis AS, Taylor GA, Deretic V. Human IRGM Induces Autophagy to Eliminate Intracellular Mycobacteria. *Science* 2006;**313**:1438-41.

838. Collazo CM, Yap GS, Sempowski GD, Lusby KC, Tessarollo L, Woude GF *et al*. 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* 2001;**194**:181-8.

839. Feng CG, Zheng L, Jankovic D, Bafica A, Cannons JL, Watford WT *et al*. The immunity-related GTPase Irgm1 promotes the expansion of activated CD4+ T cell populations by preventing interferon-gamma-induced cell death. *Nat Immunol* 2008;**9**:1279-87.

840. McCarroll SA, Huett A, Kuballa P, Chilewski SD, Landry A, Goyette P *et al*. Deletion polymorphism upstream of IRGM associated with altered IRGM expression and Crohn's disease. *Nat Genet* 2008;**40**:1107-12.

841. Greggio E, Zambrano I, Kaganovich A, Beilina A, Taymans JM, Daniels V *et al*. The Parkinson Disease-associated Leucine-rich Repeat Kinase 2 (LRRK2) Is a Dimer That Undergoes Intramolecular Autophosphorylation. *J.Biol.Chem.* 2008;**283**:16906-14.

842. Plowey ED, Cherra SJ, III, Liu YJ, Chu CT. Role of autophagy in G2019S-LRRK2-associated neurite shortening in differentiated SH-SY5Y cells. *Journal of Neurochemistry* 2008;**105**:1048-56.

843. Phillips AM, Van Limbergen J, Nimmo ER, Drummond H, Smith L, Satsangi J. Detailed haplotype-tagging study of germline variation of MUC19 in inflammatory bowel disease. *Gut* 2009;**58**:A56-A57.

844. Riis L, Vind I, Vermeire S, Wolters F, Katsanos K, Politi P *et al*. The prevalence of genetic and serologic markers in an unselected European population-based cohort of IBD patients. *Inflammatory Bowel Diseases* 2007;**13**:24-32.

845. Balde, M., Vernier-Massouille, G., Gower-Rousseau, C., Mouterde, O., Lerebours, E., Cortot, A., Colombel, J. F., and Turck, D. Natural history of Crohn's disease in children: a population-based cohort study in Northern France. Journal of Pediatric Gastroenterology & Nutrition 42(5), E25. 2006.

Ref Type: Abstract

846. Freeman HJ. Application of the Montreal classification for Crohn's disease to a single clinician database of 1015 patients. *Can J Gastroenterol* 2007;**21**:363-6.

847. Escher JC,.IBD Working Group of the European Society for Paediatric Gastroenterology HaN. First European database of pediatric inflammatory bowel disease (IBD). *Journal of Pediatric Gastroenterology & Nutrition* 2006;**42**:E25.

848. Terwilliger JD,.Hiekkalinna T. An utter refutation of the "Fundamental Theorem of the HapMap". *Eur.J Hum.Genet.* 2006;**14**:426-37.

849. Thomas DC,.Stram DO. An utter refutation of the "Fundamental Theorem of the HapMap" by Terwilliger and Hiekkalinna. *Eur.J Hum.Genet.* 2006;**14**:1238-9.

850. Gearry RB, Roberts RL, Burt MJ, Frampton CMA, Chapman B, Collett JA *et al*. Effect of inflammatory bowel disease classification changes on NOD2 genotype-phenotype associations in a population-based cohort. *Inflammatory Bowel Diseases* 2007;**13**:1220-7.

851. Van Limbergen J, Russell RK, Nimmo ER, Satsangi J. The Genetics of Inflammatory Bowel Disease. *The American Journal of Gastroenterology* 2007;**102**:2820-31.

852. Bodmer W,.Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. *Nat Genet* 2008;**40**:695-701.

853. Palmer CNA. Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis. *Nat.Genet.* 2006;**38**:441-6.

854. Sandilands A, Terron-Kwiatkowski A, Hull PR, O'Regan GM, Clayton TH, Watson RM *et al*. Comprehensive analysis of the gene encoding filaggrin uncovers prevalent and rare mutations in ichthyosis vulgaris and atopic eczema. *Nat Genet* 2007;**39**:650-4.

855. Baurecht H. Toward a major risk factor for atopic eczema: meta-analysis of filaggrin polymorphism data. *J.Allergy Clin.Immunol.* 2007;**120**:1406-12.

856. Smith FJD. Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. *Nat.Genet.* 2006;**38**:337-42.

857. Mischke D, Korge BP, Marenholz I, Volz A, Ziegler A. Genes encoding structural proteins of epidermal cornification and S100 calcium-binding proteins form a gene complex ("epidermal differentiation complex") on human chromosome 1q21. *J.Invest.Dermatol.* 1996;**106**:989-92.

858. Cookson WO. Genetic linkage of childhood atopic dermatitis to psoriasis susceptibility loci. *Nat.Genet.* 2001;**27**:372-3.

859. McGrath JA,.Uitto J. The filaggrin story: novel insights into skin-barrier function and disease. *Trends in Molecular Medicine* 2008;**14**:20-7.

860. Galli SJ, Tsai M, Piliponsky AM. The development of allergic inflammation. *Nature* 2008;**454**:445-54.

861. Hobbs EA, Van Limbergen JE, Russell RK, Nimmo ER, Drummond HE, Smith L *et al*. A detailed investigation into epidemiological risk factors for childhood onset inflammatory bowel disease in scotland. *Gut* 2008;**57**:A150-A151.

862. Feeney MA, Murphy F, Clegg AJ, Trebble TM, Sharer NM, Snook JA. A case-control study of childhood environmental risk factors for the development of inflammatory bowel disease. *European Journal of Gastroenterology & Hepatology* 2002;**14**:529-34.

863. Moffatt MF. Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature* 2007;**448**:470-3.

864. Randolph AG, Lange C, Silverman EK, Lazarus R, Silverman ES, Raby B *et al*. The IL12B gene is associated with asthma. *American Journal of Human Genetics* 2004;**75**:709-15.

865. Vercelli D. Discovering susceptibility genes for asthma and allergy. *Nat Rev Immunol* 2008;**8**:169-82.

866. Ruether A, Stoll M, Schwarz T, Schreiber S, Folster-Holst R. Filaggrin loss-of-function variant contributes to atopic dermatitis risk in the population of Northern Germany. *Br.J.Dermatol.* 2006;**155**:1093-4.

867. Huffmeier U, Boiers U, Lascorz J, Reis A, Burkhardt H. Loss-of-function mutations in the filaggrin gene: no contribution to disease susceptibility, but to autoantibody formation against citrullinated peptides in early rheumatoid arthritis. *Annals of the Rheumatic Diseases* 2008;**67**:131-3.

868. Giardina E, Paolillo N, Sinibaldi C, Novelli G. R501X and 2282del4 filaggrin mutations do not confer susceptibility to psoriasis and atopic dermatitis in Italian patients. *Dermatology* 2008;**216**:83-4.

869. Zhao Y, Terron-Kwiatkowski A, Liao H, Lee SP, Allen MH, Hull PR *et al*. Filaggrin null alleles are not associated with psoriasis. *Journal of Investigative Dermatology* 2007;**127**:1878-82.

870. Szekanecz Z, Soos L, Szabo Z, Fekete A, Kapitany A, Vegvari A *et al*. Anti-Citrullinated Protein Antibodies in Rheumatoid Arthritis: As Good as it Gets? *Clinical Reviews in Allergy & Immunology* 2008;**34**:26-31.

871. Weichenthal M, Ruether A, Schreiber S, Nair R, Voorhees JJ, Schwarz T *et al*. Filaggrin R501X and 2282del4 mutations are not associated with chronic plaque-type psoriasis in a German cohort. *Journal of Investigative Dermatology* 2007;**127**:1535-7.

872. Palmer CN. Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis. *Nature Genet.* 2006;**38**:441-6.

873. Palmer CN. Filaggrin null mutations are associated with increased asthma severity in children and young adults. *J.Allergy Clin.Immunol.* 2007;**120**:64-8.

874. McLean WH, Palmer CN, Henderson J, Kabesch M, Weidinger S, Irvine AD. Filaggrin variants confer susceptibility to asthma. *Journal of Allergy & Clinical Immunology* 2008;**121**:1294-5.

875. Weidinger S, O'Sullivan M, Illig T, Baurecht H, Depner M, Rodriguez E *et al*. Filaggrin mutations, atopic eczema, hay fever, and asthma in children. *Journal of Allergy & Clinical Immunology* 2008;**121**:1203-9.

876. Marenholz I. Filaggrin loss-of-function mutations predispose to phenotypes involved in the atopic march. *J.Allergy Clin.Immunol.* 2006;**118**:866-71.

877. Nomura T. Unique mutations in the filaggrin gene in Japanese patients with ichthyosis vulgaris and atopic dermatitis. *J.Allergy Clin.Immunol.* 2007;**119**:434-40.

878. Brown SJ, Sandilands A, Zhao Y, Liao H, Relton CL, Meggitt SJ *et al*. Prevalent and low-frequency null mutations in the filaggrin gene are associated with early-onset and persistent atopic eczema. *Journal of Investigative Dermatology* 2008;**128**:1591-4.

879. Ekelund E, Lieden A, Link J, Lee SP, D'Amato M, Palmer CN *et al*. Loss-of-function variants of the filaggrin gene are associated with atopic eczema and associated phenotypes in Swedish families. *Acta Dermato-Venereologica.88(1):15-9*, 2008;**88**:15-9.

880. Brown SJ, Relton CL, Liao H, Zhao Y, Sandilands A, Wilson IJ *et al*. Filaggrin null mutations and childhood atopic eczema: a population-based case-control study. *Journal of Allergy & Clinical Immunology* 2008;**121**:940-6.

881. Rogers AJ, Celedon JC, Lasky-Su JA, Weiss ST, Raby BA. Filaggrin mutations confer susceptibility to atopic dermatitis but not to asthma. *Journal of Allergy & Clinical Immunology* 2007;**120**:1332-7.

882. Henderson J, Northstone K, Lee SP, Liao H, Zhao Y, Pembrey M *et al*. The burden of disease associated with filaggrin mutations: a population-based, longitudinal birth cohort study. *Journal of Allergy & Clinical Immunology* 2008;**121**:872-7.

883. Morar N, Cookson WO, Harper JI, Moffatt MF. Filaggrin mutations in children with severe atopic dermatitis. *Journal of Investigative Dermatology* 2007;**127**:1667-72.

884. Weidinger S. Filaggrin mutations strongly predispose to early-onset and extrinsic atopic dermatitis. *J.Invest.Dermatol.* 2007;**127**:724-6.

885. Stemmler S, Parwez Q, Petrasch-Parwez E, Epplen JT, Hoffjan S. Two common loss-of-function mutations within the filaggrin gene predispose for early onset of atopic dermatitis. *Journal of Investigative Dermatology* 2007;**127**:722-4.

886. Barker JN. Null mutations in the filaggrin gene (FLG) determine major susceptibility to early-onset atopic dermatitis that persists into adulthood. *J.Invest.Dermatol.* 2007;**127**:564-7.

887. Hubiche T, Ged C, Benard A, Leaute-Labreze C, McElreavey K, de Verneuil H *et al*. Analysis of SPINK 5, KLK 7 and FLG genotypes in a French atopic dermatitis cohort. *Acta Dermato-Venereologica* 2007;**87**:499-505.

888. Weinmayr G, Weiland SK, Bjorksten B, Brunekreef B, Buchele G, Cookson WOC *et al*. Atopic Sensitization and the International Variation of Asthma Symptom Prevalence in Children. *Am.J.Respir.Crit.Care Med.* 2007;**176**:565-74.

889. Flohr C, Weiland SK, Weinmayr G, Bjorksten B, Braback L, Brunekreef B *et al*. The role of atopic sensitization in flexural eczema: findings from the International Study of Asthma and Allergies in Childhood Phase Two. *Journal of Allergy & Clinical Immunology* 2008;**121**:141-7.

890. Dubois PC,.van Heel DA. New susceptibility genes for ulcerative colitis. *Nat Genet* 2008;**40**:686-8.

891. Vercelli D. Discovering susceptibility genes for asthma and allergy. *Nat Rev Immunol* 2008;**8**:169-82.

892. Ober C, Tan Z, Sun Y, Possick JD, Pan L, Nicolae R *et al*. Effect of variation in CHI3L1 on serum YKL-40 level, risk of asthma, and lung function. *New England Journal of Medicine* 2008;**358**:1682-91.

893. Mizoguchi E. Chitinase 3-like-1 exacerbates intestinal inflammation by enhancing bacterial adhesion and invasion in colonic epithelial cells. *Gastroenterology* 2006;**130**:398-411.

894. Moffatt MF,.Cookson WO. Asthma and chitinases. *New England Journal of Medicine* 2008;**358**:1725-6.

895. De Benedetto A, Qualia CM, Baroody FM, Beck LA. Filaggrin expression in oral, nasal, and esophageal mucosa. *Journal of Investigative Dermatology* 2008;**128**:1594-7.

896. Ying S, Meng Q, Corrigan CJ, Lee TH. Lack of filaggrin expression in the human bronchial mucosa. *J.Allergy Clin.Immunol.* 2006;**118**:1386-8.

897. Akei HS, Mishra A, Blanchard C, Rothenberg ME. Epicutaneous antigen exposure primes for experimental eosinophilic esophagitis in mice. *Gastroenterology* 2005;**129**:985-94.

898. Brandt EB, Scribner TA, Akei HS, Rothenberg ME. Experimental gastrointestinal allergy enhances pulmonary responses to specific and unrelated allergens. *Journal of Allergy & Clinical Immunology* 2006;**118**:420-7.

899. Porter CK, Tribble DR, Aliaga PA, Halvorson HA, Riddle MS. Infectious gastroenteritis and risk of developing inflammatory bowel disease. *Gastroenterology* 2008;**135**:781-6.

900. Yamazaki K, McGovern D, Ragoussis J, Paolucci M, Butler H, Jewell D *et al*. Single nucleotide polymorphisms in TNFSF15 confer susceptibility to Crohn's disease. *Hum.Mol.Genet.* 2005;**14**:3499-506.

901. Zhernakova A, Festen EM, Franke L, Trynka G, van Diemen CC, Monsuur AJ *et al*. Genetic analysis of innate immunity in Crohn's disease and ulcerative colitis identifies two susceptibility loci harboring CARD9 and IL18RAP. *American Journal of Human Genetics* 2008;**82**:1202-10.

902. Hugot JP, Zaccaria I, Cavanaugh J, Yang H, Vermeire S, Lappalainen M *et al*. Prevalence of CARD15/NOD2 Mutations in Caucasian Healthy People. *The American Journal of Gastroenterology* 2007;**102**:1259-67.

903. Clark NM, Marinis JM, Cobb BA, Abbott DW. MEKK4 sequesters RIP2 to dictate NOD2 signal specificity. *Current Biology* 2008;**18**:1402-8.

904. Hedl M, Li J, Cho JH, Abraham C. Chronic stimulation of Nod2 mediates tolerance to bacterial products. *Proceedings of the National Academy of Sciences of the United States of America* 2007;**104**:19440-5.

905. Rosenstiel P, Huse K, Franke A, Hampe J, Reichwald K, Platzer C *et al*. Functional characterization of two novel 5' untranslated exons reveals a complex regulation of NOD2 protein expression. *BMC Genomics* 2007;472.

906. Hume GE, Fowler EV, Doecke J, Simms LA, Huang N, Palmieri O *et al*. Novel NOD2 haplotype strengthens the association between TLR4 Asp299gly and Crohn's disease in an Australian population. *Inflammatory Bowel Diseases* 2008;**14**:585-90.

907. Rimoldi M, Chieppa M, Larghi P, Vulcano M, Allavena P, Rescigno M. Monocyte-derived dendritic cells activated by bacteria or by bacteria-stimulated epithelial cells are functionally different. *Blood* 2005;**106**:2818-26.

908. Iliev ID, Matteoli G, Rescigno M. The yin and yang of intestinal epithelial cells in controlling dendritic cell function. *J.Exp.Med.* 2007;**204**:2253-7.

909. De Jager PL, Franchimont D, Waliszewska A, Bitton A, Cohen A, Langelier D *et al*. The role of the Toll receptor pathway in susceptibility to inflammatory bowel diseases. *Genes & Immunity* 2007;**8**:387-97.

910. Labbe C, Goyette P, Lefebvre C, Stevens C, Green T, Tello-Ruiz MK *et al*. MAST3: a novel IBD risk factor that modulates TLR4 signaling. *Genes & Immunity* 2008;**9**:602-12.

911. Ferwerda B, McCall MB, Verheijen K, Kullberg BJ, van der Ven AJ, van der Meer JW *et al*. Functional consequences of toll-like receptor 4 polymorphisms. *Molecular Medicine* 2008;**14**:346-52.

912. Ferwerda B, McCall MB, Alonso S, Giamarellos-Bourboulis EJ, Mouktaroudi M, Izagirre N *et al*. TLR4 polymorphisms, infectious diseases, and evolutionary pressure during migration of modern humans. *Proceedings of the National Academy of Sciences of the United States of America* 2007;**104**:16645-50.

913. Bertin J, Guo Y, Wang L, Srinivasula SM, Jacobson MD, Poyet JL *et al*. CARD9 is a novel caspase recruitment domain-containing protein that interacts with BCL10/CLAP and activates NF-kappa B. *J.Biol.Chem.* 2000;**275**:41082-6.

914. Gross O. Card9 controls a non-TLR signalling pathway for innate anti-fungal immunity. *Nature* 2006;**442**:651-6.

915. Hsu YM. The adaptor protein CARD9 is required for innate immune responses to intracellular pathogens. *Nat.Immunol.* 2007;**8**:198-205.

916. Colonna M. All roads lead to CARD9. *Nature Immunology* 2007;**8**:554-5.

917. Acosta-Rodriguez EV. Surface phenotype and antigenic specificity of human interleukin 17-producing T helper memory cells. *Nature Immunol.* 2007;**8**:639-46.

918. Villani AC, Lemire M, Fortin G, Louis E, Silverberg MS, Collette C *et al*. Common variants in the NLRP3 region contribute to Crohn's disease susceptibility. *Nat Genet* 2009;**41**:71-6.

919. Hoffman HM, Rosengren S, Boyle DL, Cho JY, Nayar J, Mueller JL *et al*. Prevention of cold-associated acute inflammation in familial cold autoinflammatory syndrome by interleukin-1 receptor antagonist. *Lancet* 2004;**364**:1779-85.

920. Taylor J, Tyekucheva S, King DC, Hardison RC, Miller W, Chiaromonte F. ESPERR: learning strong and weak signals in genomic sequence alignments to identify functional elements. *Genome Research* 2006;**16**:1596-604.

921. Bettelli E, Korn T, Oukka M, Kuchroo VK. Induction and effector functions of TH17 cells. *Nature* 2008;**453**:1051-7.

922. Awasthi A. A dominant function for interleukin 27 in generating interleukin 10-producing anti-inflammatory T cells. *Nature Immunol.* 2007;**8**:1380-9.

923. Nakazawa A, Dotan I, Brimnes J, Allez M, Shao L, Tsushima F *et al*. The expression and function of costimulatory molecules B7H and B7-H1 on colonic epithelial cells. *Gastroenterology* 2004;**126**:1347-57.

924. Lupardus PJ,.Garcia KC. The structure of interleukin-23 reveals the molecular basis of p40 subunit sharing with interleukin-12. *Journal of Molecular Biology* 2008;**382**:931-41.

925. Zhou L. IL-6 programs TH-17 cell differentiation by promoting sequential engagement of the IL-21 and IL-23 pathways. *Nature Immunol.* 2007;**8**:967-74.

926. Khader SA. IL-23 and IL-17 in the establishment of protective pulmonary CD4+ T cell responses after vaccination and during Mycobacterium tuberculosis challenge. *Nature Immunol.* 2007;**8**:369-77.

927. Salazar-Gonzalez RM, Niess JH, Zammit DJ, Ravindran R, Srinivasan A, Maxwell JR *et al*. CCR6-mediated dendritic cell activation of pathogen-specific T cells in Peyer's patches. *Immunity* 2006;**24**:623-32.

928. Lee JW, Wang P, Kattah MG, Youssef S, Steinman L, DeFea K *et al*. Differential Regulation of Chemokines by IL-17 in Colonic Epithelial Cells. *J Immunol* 2008;**181**:6536-45.

929. Bouskra D, Brezillon C, Berard M, Werts C, Varona R, Boneca IG *et al*. Lymphoid tissue genesis induced by commensals through NOD1 regulates intestinal homeostasis. *Nature* 2008;**456**:507-10.

930. Gutierrez MG, Master SS, Singh SB, Taylor GA, Colombo MI, Deretic V. Autophagy Is a Defense Mechanism Inhibiting BCG and Mycobacterium tuberculosis Survival in Infected Macrophages. *Cell* 2004;**119**:753-66.

931. Deretic V, Singh S, Master S, Harris J, Roberts E, Kyei G *et al*. Mycobacterium tuberculosis inhibition of phagolysosome biogenesis and autophagy as a host defence mechanism. *Cellular Microbiology* 2006;**8**:719-27.

932. Xu Y, Jagannath C, Liu XD, Sharafkhaneh A, Kolodziejska KE, Eissa NT. Toll-like receptor 4 is a sensor for autophagy associated with innate immunity. *Immunity* 2007;**27**:135-44.

933. Suzuki T,.Nunez G. A role for Nod-like receptors in autophagy induced by Shigella infection. *Autophagy* 2008;**4**:73-5.

934. Gutierrez MG, Saka HA, Chinen I, Zoppino FC, Yoshimori T, Bocco JL *et al*. Protective role of autophagy against Vibrio cholerae cytolysin, a pore-forming toxin from V. cholerae. *Proc.Natl Acad.Sci.U.S.A* 2007;**104**:1829-34.

935. Saka HA, Gutierrez MG, Bocco JL, Colombo MI. The autophagic pathway: a cell survival strategy against the bacterial pore-forming toxin Vibrio cholerae cytolysin. *Autophagy* 2007;**3**:363-5.

936. Brazil MI, Weiss S, Stockinger B. Excessive degradation of intracellular protein in macrophages prevents presentation in the context of major histocompatibility complex class II molecules. *Eur.J.Immunol.* 1997;**27**:1506-14.

937. Nimmerjahn F, Milosevic S, Behrends U, Jaffee EM, Pardoll DM, Bornkamm GW *et al*. Major histocompatibility complex class II-restricted presentation of a cytosolic antigen by autophagy. *Eur.J.Immunol.* 2003;**33**:1250-9.

938. Dorfel D. Processing and presentation of HLA class I and II epitopes by dendritic cells after transfection with in vitro-transcribed MUC1 RNA. *Blood* 2005;**105**:3199-205.

939. Dengjel J. Autophagy promotes MHC class II presentation of peptides from intracellular source proteins. *Proc.Natl Acad.Sci.USA* 2005;**102**:7922-7.

940. Paludan C. Endogenous MHC class II processing of a viral nuclear antigen after autophagy. *Science* 2005;**307**:593-6.

941. Mizushima N, Yamamoto A, Matsui M, Yoshimori T, Ohsumi Y. In vivo analysis of autophagy in response to nutrient starvation using transgenic mice expressing a fluorescent autophagosome marker. *Mol.Biol.Cell* 2004;**15**:1101-11.

942. Schmid D, Pypaert M, Munz C. Antigen-loading compartments for major histocompatibility complex class II molecules continuously receive input from autophagosomes. *Immunity* 2007;**26**:79-92.

943. Nedjic J, Aichinger M, Emmerich J, Mizushima N, Klein L. Autophagy in thymic epithelium shapes the T-cell repertoire and is essential for tolerance. *Nature* 2008;**455**:396-400.

944. Weersma RK, Stokkers PCF, van Bodegraven AA, van Hogezand RA, Verspaget HW, de Jong DJ *et al*. Molecular prediction of disease risk and severity in a large Dutch Crohn's disease cohort. *Gut* 2008;**58**:388-95.

945. Van Limbergen J, Russell RK, Nimmo ER, Drummond HE, Smith L, Anderson NH *et al*. Autophagy gene ATG16L1 influences susceptibility and disease location but not childhood-onset in Crohn's disease in Northern Europe. *Inflammatory Bowel Diseases* 2008;**14**:338-46.

946. Fowler EV, Doecke J, Simms LA, Zhao ZZ, Webb PM, Hayward NK *et al*. ATG16L1 T300A shows strong associations with disease subgroups in a large Australian IBD population: further support for significant disease heterogeneity. *American Journal of Gastroenterology* 2008;**103**:2519-26.

947. Hancock L, Beckly J, Geremia A, Cooney R, Cummings F, Pattni R *et al*. Clinical and molecular characteristics of isolated colonic Crohn's disease. *Inflammatory Bowel Diseases* 2008;**14**:1667-77.

948. Bafica A, Feng CG, Santiago HC, Aliberti J, Cheever A, Thomas KE *et al*. The IFN-inducible GTPase LRG47 (Irgm1) negatively regulates TLR4-triggered proinflammatory cytokine production and prevents endotoxemia. *Journal of Immunology* 2007;**179**:5514-22.

949. Van Limbergen J, Russell RK, Nimmo ER, Drummond HE, Davies G, Wilson DC *et al*. Germline variants of IRGM in childhood-onset Crohn's disease. *Gut* 2009;**58**:610-1.

950. Van Limbergen J, Russell RK, Nimmo ER, Satsangi J. The genetics of inflammatory bowel disease. *American Journal of Gastroenterology* 2007;**102**:2820-31.

951. Brescianini S, Trinh T, Stoll M, Schreiber S, Rioux JD, Daly MJ. IBD5 is associated with an extensive complicated Crohn's disease feature: implications from genotype-phenotype analysis. *Gut* 2007;**56**:149-50.

952. Browning BL, Annese V, Barclay ML, Bingham SA, Brand S, Buning C *et al*. Gender-stratified analysis of DLG5 R30Q in 4707 patients with Crohn disease and 4973 controls from 12 Caucasian cohorts. *J Med Genet* 2008;**45**:36-42.

953. Kurz T. Fine mapping and positional candidate studies on chromosome 5p13 identify multiple asthma susceptibility loci. *J.Allergy Clin.Immunol.* 2006;**118**:396-402.

954. Vercelli D. Discovering susceptibility genes for asthma and allergy. *Nat Rev Immunol* 2008;**8**:169-82.

955. Vercelli D. Discovering susceptibility genes for asthma and allergy. *Nat Rev Immunol* 2008;**8**:169-82.

956. Bouzigon E, Corda E, Aschard H, Dizier MH, Boland A, Bousquet J *et al*. Effect of 17q21 Variants and Smoking Exposure in Early-Onset Asthma. *N Engl J Med* 2008;**359**:1985-94.

957. Tsuji S, Uehori J, Matsumoto M, Suzuki Y, Matsuhisa A, Toyoshima K *et al*. Human Intelectin Is a Novel Soluble Lectin That Recognizes Galactofuranose in Carbohydrate Chains of Bacterial Cell Wall. *J.Biol.Chem.* 2001;**276**:23456-63.

958. Wrackmeyer U, Hansen GH, Seya T, Danielsen EM. Intelectin: a novel lipid raft-associated protein in the enterocyte brush border. *Biochemistry* 2006;**45**:9188-97.

959. De Lisle RC, Xu W, Roe BA, Ziemer D. Effects of Muclin (Dmbt1) deficiency on the gastrointestinal system. *Am J Physiol Gastrointest Liver Physiol* 2008;**294**:G717-G727.

960. Renner M, Bergmann G, Krebs I, End C, Lyer S, Hilberg F *et al*. DMBT1 confers mucosal protection in vivo and a deletion variant is associated with Crohn's disease. *Gastroenterology* 2007;**133**:1499-509.

961. Rosenstiel P, Sina C, End C, Renner M, Lyer S, Till A *et al*. Regulation of DMBT1 via NOD2 and TLR4 in Intestinal Epithelial Cells Modulates Bacterial Recognition and Invasion. *J Immunol* 2007;**178**:8203-11.

962. Ma A. Unresolved ER Stress Inflames the Intestine. *Cell* 2008;**134**:724-5.

963. Shkoda A, Ruiz PA, Daniel H, Kim SC, Rogler G, Sartor RB *et al*. Interleukin-10 blocked endoplasmic reticulum stress in intestinal epithelial cells: impact on chronic inflammation. *Gastroenterology.132(1):190-207,* 2007.

964. Heazlewood CK, Cook MC, Eri R, Price GR, Tauro SB, Taupin D *et al*. Aberrant mucin assembly in mice causes endoplasmic reticulum stress and

spontaneous inflammation resembling ulcerative colitis. *PLoS Medicine / Public Library of Science* 2008;**5**:e54.

965. Kaser A, Lee AH, Franke A, Glickman JN, Zeissig S, Tilg H *et al*. XBP1 Links ER Stress to Intestinal Inflammation and Confers Genetic Risk for Human Inflammatory Bowel Disease. *Cell* 2008;**134**:743-56.

966. de Bakker PI, McVean G, Sabeti PC, Miretti MM, Green T, Marchini J *et al*. A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nature Genetics* 2006;**38**:1166-72.

967. Fernando MM, Stevens CR, Walsh EC, De Jager PL, Goyette P, Plenge RM *et al*. Defining the role of the MHC in autoimmunity: a review and pooled analysis. *PloS Genetics* 2008;**4**:e1000024.

968. van Heel DA, Udalova IA, De Silva AP, McGovern DP, Kinouchi Y, Hull J *et al*. Inflammatory bowel disease is associated with a TNF polymorphism that affects an interaction between the OCT1 and NF(-kappa)B transcription factors. *Hum.Mol.Genet.* 2002;**11**:1281-9.

969. Van Limbergen J, Russell RK, Nimmo ER, Drummond HE, Smith L, Anderson N *et al*. Tumour necrosis factor a promoter polymorphisms: Influence on phenotype and severity in childhood IBD and regression analysis of the relative contribution to Crohn's disease phenotype. *Gastroenterology* 2006;**130**:A3.

970. Kakuta Y, Kinouchi Y, Negoro K, Takahashi S, Shimosegawa T. Association study of TNFSF15 polymorphisms in Japanese patients with inflammatory bowel disease. *Gut* 2006;**55**:1527-8.

971. Picornell Y, Mei L, Taylor K, Yang H, Targan SR, Rotter JI. TNFSF15 is an ethnic-specific IBD gene. *Inflammatory Bowel Diseases* 2007;**13**:1333-8.

972. Tremelling M, Berzuini C, Massey D, Bredin F, Price C, Dawson C *et al*. Contribution of TNFSF15 gene variants to Crohn's disease susceptibility confirmed in UK population. *Inflammatory Bowel Diseases* 2008;**14**:733-7.

973. Yang SK, Lim J, Chang HS, Lee I, Li Y, Liu J *et al*. Association of TNFSF15 with Crohn's disease in Koreans. *American Journal of Gastroenterology* 2008;**103**:1437-42.

974. Bamias G, Mishina M, Nyce M, Ross WG, Kollias G, Rivera-Nieves J *et al*. Role of TL1A and its receptor DR3 in two models of chronic murine ileitis. *Proceedings of the National Academy of Sciences of the United States of America* 2006;**103**:8441-6.

975. Takedatsu H, Michelsen KS, Wei B, Landers CJ, Thomas LS, Dhall D *et al*. TL1A (TNFSF15) regulates the development of chronic colitis by

modulating both T-helper 1 and T-helper 17 activation. *Gastroenterology* 2008;**135**:552-67.

976. Pappu BP, Borodovsky A, Zheng TS, Yang X, Wu P, Dong X *et al*. TL1A-DR3 interaction regulates Th17 cell function and Th17-mediated autoimmune disease. *Journal of Experimental Medicine.* 2008;**205**:1049-62.

977. Dideberg V, Kristjansdottir G, Milani L, Libioulle C, Sigurdsson S, Louis E *et al*. An insertion-deletion polymorphism in the interferon regulatory Factor 5 (IRF5) gene confers risk of inflammatory bowel diseases. *Hum.Mol.Genet.* 2007;**16**:3008-16.

978. Vang T, Miletic AV, Arimura Y, Tautz L, Rickert RC, Mustelin T. Protein tyrosine phosphatases in autoimmunity. *Annual Review of Immunology* 2008;**26**:29-55.

979. Yamazaki K, Takahashi A, Takazoe M, Kubo M, Onouchi Y, Fujino A *et al*. Positive association of genetic variants in the upstream region of NKX2-3 with Crohn's disease in Japanese patients. *Gut* 2008;**58**:228-32.

980. Tarlinton D, Light A, Metcalf D, Harvey RP, Robb L. Architectural Defects in the Spleens of Nkx2-3-Deficient Mice Are Intrinsic and Associated with Defects in Both B Cell Maturation and T Cell-Dependent Immune Responses. *J Immunol* 2003;**170**:4002-10.

981. Wang CC, Biben C, Robb L, Nassir F, Barnett L, Davidson NO *et al*. Homeodomain factor Nkx2-3 controls regional expression of leukocyte homing coreceptor MAdCAM-1 in specialized endothelial cells of the viscera. *Developmental Biology* 2000;**224**:152-67.

982. MacDonald JK,.McDonald JW. Natalizumab for induction of remission in Crohn's disease.[update of Cochrane Database Syst Rev. 2006;3:CD006097; PMID: 16856112]. [Review] [22 refs]. *Cochrane Database of Systematic Reviews* 2007;**1**.

983. Hyams JS, Wilson DC, Thomas A, Heuschkel R, Mitton S, Mitchell B *et al*. Natalizumab therapy for moderate to severe Crohn disease in adolescents. *Journal of Pediatric Gastroenterology & Nutrition* 2007;**44**:185-91.

984. Verbeeck J, Van Assche G, Ryding J, Wollants E, Rans K, Vermeire S *et al*. JC viral loads in patients with Crohn's disease treated with immunosuppression: can we screen for elevated risk of progressive multifocal leukoencephalopathy? *Gut* 2008;**57**:1393-7.

985. Jana M,.Pahan K. Induction of lymphotoxin-alpha by interleukin-12 p40 homodimer, the so-called biologically inactive molecule, but not IL-12 p70. *Immunology* 2008;**127**:312-25.

986. Zhao Z, Yu S, Fitzgerald DC, Elbehi M, Ciric B, Rostami AM *et al*. IL-12Rbeta2 Promotes the Development of CD4+CD25+ Regulatory T Cells. *J Immunol* 2008;**181**:3870-6.

987. Kang SJ, Liang HE, Reizis B, Locksley RM. Regulation of Hierarchical Clustering and Activation of Innate Immune Cells by Dendritic Cells. *Immunity* 2008;**29**:819-33.

988. Cheung H, Chen NJ, Cao Z, Ono N, Ohashi PS, Yeh WC. Accessory Protein-Like Is Essential for IL-18-Mediated Signaling. *J Immunol* 2005;**174**:5351-7.

989. Hunt KA, Zhernakova A, Turner G, Heap GA, Franke L, Bruinenberg M *et al*. Newly identified genetic risk variants for celiac disease related to the immune response. *Nature Genetics* 2008;**40**:395-402.

990. Corbaz A, ten Hove T, Herren S, Graber P, Schwartsburd B, Belzer I *et al*. IL-18-binding protein expression by endothelial cells and macrophages is up-regulated during active Crohn's disease. *Journal of Immunology* 2002;**168**:3608-16.

991. Pizarro TT, Michie MH, Bentz M, Woraratanadharm J, Smith MF, Jr., Foley E *et al*. IL-18, a novel immunoregulatory cytokine, is up-regulated in Crohn's disease: expression and localization in intestinal mucosal cells. *Journal of Immunology* 1999;**162**:6829-35.

992. ten Hove T, Corbaz A, Amitai H, Aloni S, Belzer I, Graber P *et al*. Blockade of endogenous IL-18 ameliorates TNBS-induced colitis by decreasing local TNF-alpha production in mice. *Gastroenterology* 2001;**121**:1372-9.

993. Sivakumar PV, Westrich GM, Kanaly S, Garka K, Born TL, Derry JM *et al*. Interleukin 18 is a primary mediator of the inflammation associated with dextran sulphate sodium induced colitis: blocking interleukin 18 attenuates intestinal damage. *Gut* 2002;**50**:812-20.

994. Goyette P, Lefebvre C, Ng A, Brant SR, Cho JH, Duerr RH *et al*. Gene-centric association mapping of chromosome 3p implicates MST1 in IBD pathogenesis. *Mucosal Immunol* 2008;**1**:131-8.

995. Beckly JB, Hancock L, Geremia A, Cummings JR, Morris A, Cooney R *et al*. Two-stage candidate gene study of chromosome 3p demonstrates an association between nonsynonymous variants in the MST1R gene and Crohn's disease. *Inflammatory Bowel Diseases* 2008;**14**:500-7.

996. Mudter J,.Neurath MF. Apoptosis of T cells and the control of inflammatory bowel disease: therapeutic implications. *Gut* 2007;**56**:293-303.

997. Funke B, Autschbach F, Kim S, Lasitschka F, Strauch U, Rogler G *et al*. Functional characterization of Decoy Receptor 3 in Crohn's disease. *Gut* 2008;**58**:483-91.