Queen Mary
University of London

Altered intestinal microbiota and blood T cell phenotype are shared by patients with Crohn's disease and their unaffected siblings.

Hedin, CR; McCarthy, NE; Louis, P; Farquharson, FM; McCartney, S; Taylor, K; Prescott, NJ; Murrells, T; Stagg, AJ; Whelan, K; Lindsay, JO

For additional information about this publication click this link.
http://qmro.qmul.ac.uk/xmlui/handle/123456789/15330

ORIGINAL ARTICLE

Altered intestinal microbiota and blood T cell phenotype are shared by patients with Crohn's disease and their unaffected siblings

Charlotte R Hedin,[1,2] Neil E McCarthy,[2] Petra Louis,[3] Freda M Farquharson,[3] Sara McCartney,[4] Kirstin Taylor,[5] Natalie J Prescott,[5] Trevor Murrells,[6] Andrew J Stagg,[7] Kevin Whelan,[1] James O Lindsay[2,8]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/gutjnl-2013-306226).

[1]Diabetes and Nutritional Sciences Division, School of Medicine, King's College London, London, UK
[2]Centre for Digestive Diseases, Blizard Institute, Queen Mary University of London, London, UK
[3]Microbiology Group, Gut Health Theme, Rowett Institute of Nutrition and Health, University of Aberdeen, Bucksburn, Aberdeen, UK
[4]Department for Gastroenterology and Clinical Nutrition, University College Hospitals NHS Foundation Trust, London, UK
[5]Department of Medical and Molecular Genetics, King's College London, London, UK
[6]National Nursing Research Unit, Florence Nightingale School of Nursing and Midwifery, King's College London, London, UK
[7]Centre for Immunology and Infectious Disease, Blizard Institute, Queen Mary University of London, London, UK
[8]Gastroenterology Division, Barts Health NHS Trust, London, UK

Correspondence to
Dr James Lindsay, Barts Health NHS Trust, Endoscopy Unit, The Royal London Hospital, Whitechapel, London E1 1BB, UK;
James.Lindsay@bartshealth.nhs.uk

CrossMark

ABSTRACT

Objective Crohn's disease (CD) is associated with intestinal dysbiosis, altered blood T cell populations, elevated faecal calprotectin (FC) and increased intestinal permeability (IP). CD-associated features present in siblings (increased risk of CD) but not in healthy controls, provide insight into early CD pathogenesis. We aimed to (1) Delineate the genetic, immune and microbiological profile of patients with CD, their siblings and controls and (2) Determine which factors discriminate between groups.

Design Faecal microbiology was analysed by quantitative PCR targeting 16S ribosomal RNA, FC by ELISA, blood T cell phenotype by flow cytometry and IP by differential lactulose-rhamnose absorption in 22 patients with inactive CD, 21 of their healthy siblings and 25 controls. Subject's genotype relative risk was determined by Illumina Immuno BeadChip.

Results Strikingly, siblings shared aspects of intestinal dysbiosis with patients with CD (lower concentrations of *Faecalibacterium prausnitzii* (p=0.048), Clostridia cluster IV (p=0.003) and *Roseburia* spp. (p=0.09) compared with controls). As in CD, siblings demonstrated a predominance of memory T cells (p=0.002) and elevated naïve CD4 T cell β7 integrin expression (p=0.01) compared with controls. FC was elevated (>50 µg/g) in 8/21 (38%) siblings compared with 2/25 (8%) controls (p=0.028); whereas IP did not differ between siblings and controls. Discriminant function analysis determined that combinations of these factors significantly discriminated between groups (χ^2=80.4, df=20, p<0.001). Siblings were separated from controls by immunological and microbiological variables.

Conclusions Healthy siblings of patients with CD manifest immune and microbiological abnormalities associated with CD distinct from their genotype-related risk and provide an excellent model in which to investigate early CD pathogenesis.

INTRODUCTION

Crohn's disease (CD) causes intestinal inflammation, occurs predominantly in young adults and results in significant morbidity and reduced quality of life. The pathogenesis of CD is unknown, but is thought to depend on abnormal interactions between the intestinal immune system and microbiota in genetically susceptible hosts. This model is supported by the identification of >130 risk loci associated with CD,[1]

Significance of this study

What is already known about this subject?
► Crohn's disease is associated with intestinal dysbiosis, T cell abnormalities and elevated genetic risk.
► The siblings of patients with Crohn's disease are at increased risk of developing disease compared with the background population.
► A proportion of the siblings of patients with Crohn's disease will have elevated faecal calprotectin and intestinal permeability.

What are the new findings?
► The siblings of patients with Crohn's disease share some aspects of the intestinal dysbiosis with their affected relatives such as reduced Firmicutes including *Faecalibacterium prauznitsii* which has been shown to be relevant to clinical course of disease.
► The siblings of patients with Crohn's disease share abnormalities in T cell phenotype with their affected relatives.
► Siblings of patients with Crohn's disease can be discriminated from healthy controls by a combination of immunological and microbiological factors.

How might it impact on clinical practice in the foreseeable future?
► This is the first study to assess multiple aspects of the Crohn's disease phenotype in patients' unaffected siblings and healthy controls.
► It identifies novel microbiological and immunological factors relevant to early disease pathogenesis.
► It raises the possibility of predicting the risk of Crohn's disease onset using multidimensional phenotyping.

including genes controlling microbial sensing and immune function. Furthermore, patients with CD exhibit altered intestinal microbiota (dysbiosis) including reduced Firmicutes such as Clostridia cluster IV and *Faecalibacterium prausnitzii*,[2] and increased *Escherichia coli*.[3] CD is characterised

immunologically by predominant Th1/Th17 mucosal T cell responses,[4] with alterations in peripheral T cell populations and their expression of intestinal homing receptors such as α4β7 integrin.[5][6] Finally, intestinal inflammation in CD is associated with a raised faecal calprotectin (FC)[7] and altered intestinal permeability (IP).[8] Individual factors such as elevated FC or lower mucosal *F prausnitzii* can predict natural history of CD disease.[2][9] However, the alterations noted may be a consequence of established disease. This study aims to delineate features of individuals at risk of CD relevant to disease pathogenesis.

First-degree relatives of patients with CD are at increased risk of developing IBD compared with background populations.[10] Monozygotic twin disease discordance highlights the importance of non-genetic factors in disease pathogenesis. Genotype alone therefore accounts for a significant but limited proportion of CD risk and biomarkers of non-genetic influences must be sought.[11][12] Studying siblings of patients with CD may provide insights into genetic and non-genetic factors relevant to CD pathogenesis, that are more readily apparent in the absence of established disease.

Several authors have sought to describe the phenotype of the 'at-risk' relatives of patients with CD. FC is increased in up to 49% of first-degree relatives,[13] and IP is reportedly increased in CD relatives, although data are inconsistent.[8][14–16] Fewer data describe the intestinal microbiota of healthy relatives of patients with CD; one study reported specific relative-associated dysbiosis distinct from that associated with CD,[17] while other data have suggested an association between dysbiosis in relatives with CD and CD-associated alleles.[18] However, the microbial characteristics of relatives' dysbiosis are inconsistent between studies.

Interaction between gut microbiota and the mucosal immune system is a cornerstone of CD pathogenesis, yet few data exist that define the immune phenotype of those at risk of CD.[19] Defects in oral tolerance induction have been reported in some first-degree relatives of patients with CD[20] but there has been no detailed analysis of the associated T cell phenotype and function.

Limitations of previous studies include recruitment of unaffected parents of patients with CD, and the determination of specific aspects of the at-risk phenotype in isolation. The inclusion of siblings is preferable to parents because relevant risk factors such as raised FC[21] and altered gut microbiota,[22] may vary with age. Furthermore, parents are usually beyond the peak age of CD incidence, reducing the power to detect biomarkers of IBD susceptibility. Finally, comparison of individuals separated by a generation may be confounded by changes in prevailing environmental conditions. These issues are circumvented in the current study by comparing phenotypes of patients and their unaffected siblings.

The current study was designed to test the hypothesis that unaffected siblings of patients with CD share features of a 'CD phenotype' that may afford insights into disease pathogenesis and provide markers of CD risk. This study is unique by defining multiple dimensions of the CD phenotype simultaneously, including microbial and immunological markers alongside FC, IP and genotype. This has allowed interrogation of the relative impact and interactions of these features.

MATERIALS AND METHODS

Patients with inactive CD (Crohn's Disease Activity Index (CDAI) <150 and C reactive protein (CRP) <5 mg/L) and their healthy siblings (both aged 16–35 years) were recruited from clinics at Barts and London, and University College Hospitals NHS Trusts, UK between June 2008 and December 2011.

Healthy controls were recruited by circular email sent to staff and students at King's College London during the same period. Patients required a confirmed diagnosis of CD for >3 months and an available sibling. For exclusion criteria see online supplementary table S1. All participants provided written informed consent. Ethical approval was provided by Bromley Local Research Ethics Committee (reference 07/H0805/46).

At screening, demographics, medical and drug exposure history, physical examination, CRP, inclusion and exclusion criteria were assessed. Written instructions regarding avoidance of prebiotics/probiotics for 4 weeks, non-steroidal anti-inflammatory drugs for 1 week and alcohol for 24 h before the study were provided. Blood samples were taken for routine haematology/biochemistry, T cell analysis and genotyping. Participants completed the 5-h urine collection for IP and underwent flexible rectoscopy without bowel cleansing. Biopsies from non-inflamed rectum were sent for routine histological analysis. Stool was obtained for and stored at −20°C before processing for FC quantification and microbiological analysis.

Laboratory methods
Fecal calprotectin
Faecal samples were transferred on ice before FC extraction and ELISA analysis (Calpro AS, Lysaker, Norway) according to the manufacturer's instructions using duplicate appropriately diluted samples. FC concentration was determined relative to standard curves and expressed as μg/g of faeces.

Gut microbiota
Gut microbiota that comprise the reported dysbiosis of CD were analysed using quantitative PCR (qPCR). Faecal DNA extraction was carried out using the FastDNA spin kit for soil and the FastPrep-24 bead homogeniser (MP Biomedicals, Solon, Ohio, USA) following the manufacturer's instructions. A SYBR Green qPCR analysis of faecal DNA was performed as described previously[23] with the following modifications: amplified bacterial 16S rRNA genes served as standard templates as previously described except for *Ruminococcus bromii* L2–63[24] for the *Ruminococcus* primers; for *E coli* primers, *E coli* XL1-Blue was used. PCR reactions were performed in quadruplicate using a 7900HT Fast Real-Time PCR machine (Life Technologies). Data were analysed with SDS V.2.4 software (Life Technologies). Faecal water content was measured by lyophilisation and concentrations of bacteria expressed as $\log_{10}$ copies/gram dry faeces. Normalised proportions of bacteria were obtained by dividing specific group quantities through universal quantities. For details of primers and PCR conditions see online supplementary table S2.

Peripheral blood T cell flow cytometry
Whole blood, collected in lithium-heparin Vacutainer tubes (BD Bioscience), was stored at room temperature for ≤4 h before labelling with fluorescently conjugated monoclonal antibodies to detect CD3 T cells, naïve (CD45RA$^+$) and memory (CD45RA$^-$) subsets of CD4 and CD8 T cells. Integrin α4β7 expression was assessed by labelling with anti-β7 (see online supplementary methods for antibodies used).

Data were acquired using a LSRII 4-colour flow cytometer (BD Bioscience) and collected using fluorescence-activated cell sorting Diva software V.4.1.2 (BD Bioscience) using Flow-Count fluorospheres (Beckman Coulter) for absolute quantitation. Colour compensation was performed offline using Winlist V.6.0 (Verity Software House).

Genotyping

Human DNA was extracted from whole blood using the phenol chloroform-isoamyl alcohol method.[25] Genotyping was performed using the Illumina Infinium Immunochip,[26] which includes 70 of the 71 CD disease risk loci described in Franke et al.[19] To increase detection of NOD2 mutations and capture the enhanced risk of NOD2 compound heterozygosity, three NOD2 (rs2066845/G908R, rs2066844/R702W and rs5743293/3020insC) were individually assessed. Cumulative genotype relative risk (GRR) for each participant was therefore calculated across a total of 72 CD risk loci. A population distribution model of CD risk was generated using the REGENT R program[27] and previously published ORs.[19] Participants were categorised into reduced, average, elevated or high genotypical risk with reference to this model.[28]

Intestinal permeability

IP tests were carried out using lactulose-rhamnose tests.[29] After overnight fasting, participants consumed 450 mL water containing 5 g lactulose, 2 g mannitol and 1 g L-rhamnose (BCM Specials, Nottingham, UK), followed by a 5 h urine collection. Aliquots from the pooled urine sample were stored at −20°C until analysis.

Urinary sugar separation and detection was carried out by liquid chromatography-tandem mass spectrometry. Following microcentrifugation, supernatant was transferred to a LC-2000 platform autosampler (Jasco, Easton, Maryland, USA), and urinary sugars separated by high performance liquid chromatography using an amino column in hydrophilic interaction liquid chromatography mode. Sugars were quantified by electrospray tandem mass spectrometry, using the API 3200 (AB Sciex, Framingham, Massachusetts, USA).

Power calculation

The hypothesis was that siblings would demonstrate multiple aspects of a CD phenotype distinct from healthy controls. In order to calculate the number of participants required, we performed a power calculation based upon detecting lower concentrations of F prausnitzii in siblings compared with healthy controls. This determined that 21 subjects would be required in each group to detect a difference in F prausnitzii of 0.3 $\log_{10}$ between siblings and healthy controls assuming a SD of 0.3 $\log_{10}$ (power 0.9, significance level 0.05).

Statistical analysis

Normality was determined by visual assessment of histograms and Shapiro-Wilk test. Parametric analyses (one-way analysis of variance or Student's t test) and non-parametric analyses (Mann-Whitney or Kruskal-Wallis tests) with Tukey's or Bonferroni post hoc corrections applied as appropriate. Correlations between continuous variables were assessed with Pearson's correlation coefficients (PCCs). Associations between categorical variables were assessed using χ^2 tests. Hierarchical multiple regression analysis determined dependence between variables, after controlling for the effect of other factors. The capacity of combined variables to discriminate between the three groups was determined using discriminant analysis (SPSS, version 18.0).

RESULTS

Participants and genotype relative risk

Twenty-two patients with CD in remission (mean CDAI 71 (SD 44.8)), 21 of their healthy siblings and 25 controls matched by group for age, gender, body mass index, ethnicity, smoking status and median number of siblings were recruited (see online supplementary table S3). Patients' disease characteristics were representative of northern European populations. All participants had normal CRP (<5 mg/L) at screening. Nine (41%) patients and four (19%) siblings carried at least one NOD2 mutation compared with four (16%) controls (p=0.054). No participant had evidence of either macroscopic or microscopic mucosal inflammation. Predictably, higher proportions of patients and siblings had an elevated or high GRR across the 72 CD risk loci assessed, compared with controls (table 1).

Faecal calprotectin was increased in patients and siblings

FC was significantly higher in patients (252 µg/g, IQR 125–775 µg/g) compared with siblings (36 µg/g, IQR 17–65 µg/g, p<0.003) and controls (17 µg/g, IQR 9–41 µg/g, p<0.003). The proportion with FC concentrations above normal (50 µg/g), was greater in patients (21, 95%) versus controls (2, 8%, p<0.001), and siblings (8, 38%) versus controls, (p=0.028). FC was not affected by smoking status or disease location and did not correlate between related sibling-patient pairs (PCC=−0.005, p=0.983).

Intestinal permeability was increased in patients but not siblings

Significantly more patients (12, 57%) had abnormal IP (ratio of urinary lactulose to rhamnose >0.05) than siblings or controls (6, (29%) and 5 (23%), respectively, p=0.044), but IP was not significantly different between siblings and controls. IP did not correlate between sibling pairs, (PCC −0.024, p=0.917), but significantly correlated with FC across all groups (PCC 0.283, p=0.023).

A Crohn's-like dysbiosis, including reduced F prausnitzii, was observed in siblings

Compared with controls, patients had significantly lower concentrations of total bacteria, F prausnitzii, Clostridia cluster IV, Ruminococcus spp., Roseburia spp., Bacteroides-Prevotella and Bifidobacterium adolescentis; a dysbiosis comparable with that previously described in CD.[3] Strikingly, siblings also had significantly lower concentrations of F prausnitzii, Clostridia cluster

Table 1 The number of patients' siblings and controls observed in each of the four genotype relative risk (GRR) categories

GRR categories, n (%)	Patients (n=22)	Siblings (n=21)	Controls (n=25)	p Values
High (GRR >4.185)	3 (14)	1 (5)	0 (0)	Pat-con 0.045
Elevated (GRR 2.273–4.185)	2 (9)	3 (14)	0 (0)	Sib-con 0.046
Average (GRR 0.445–2.277)	11 (50)	12 (57)	11 (44)	Pat-sib 0.726
Reduced (GRR <0.445)	6 (27)	5 (24)	14 (56)	

Categorised GRR in patients with CD and siblings was significantly greater compared with controls. Genotyping was performed using the Illumina Infinium Immunochip as described in the methods.
CD, Crohn's disease.

Table 2 Concentrations of bacterial groups and species in faecal samples were significantly different between patients, siblings and controls

	Concentrations of bacteria, $\log_{10}$ copies/g median (IQR)				
	Patients (n=22)	Siblings (n=21)	Controls (n=25)	p Value	Between-group comparisons p Values
Faecalibacterium prausnitzii	6.88 (5.03–9.35)	9.27 (8.12–9.78)	9.59 (9.34–10.14)	<0.001	Pat-Sib: 0.006 Pat-Con: <0.003* Sib-Con: 0.048
Clostridia cluster IV	7.76 (6.54–9.56)	9.34 (8.76–9.93)	9.69 (9.48–10.26)	<0.001	Pat-Sib: 0.045 Pat-Con: <0.003 Sib-Con: 0.030
Clostridia cluster XIVa	9.86 (9.07–10.31)	10.13 (9.64–10.46)	10.19 (9.85–10.85)	0.099	–
Cluster IV *Ruminococcus* spp.	7.05 (5.94–8.29)	8.75 (8.04–9.33)	9.55 (8.37–10.02)	<0.001	Pat-Sib: <0.003 Pat-Con: <0.003 Sib-Con: 0.084
Roseburia spp.	9.19 (7.35–9.85)	9.34 (7.21–9.70)	9.92 (9.39–10.15)	0.004	Pat-Sib: 1.000 Pat-Con: 0.027 Sib-Con: 0.009
Bacteroides-Prevotella	8.83 (8.35–10.05)	10.16 (9.50–10.70)	10.48 (10.08–10.80)	<0.001	Pat-Sib: 0.009 Pat-Con: <0.003 Sib-Con: 0.639
Escherichia coli	7.95 (6.61–9.00)	7.15 (6.95–8.07)	7.32 (6.79–8.00)	0.553	–
Bifidobacterium	9.36 (8.24–10.00)	9.84 (9.41–10.05)	9.98 (9.32–10.16)	0.154	–
Bifidobacterium longum	8.91 (6.80–9.50)	9.13 (8.67–9.56)	9.24 (8.18–9.61)	0.640	–
Bifidobacterium adolescentis	5.76 (4.69–8.78)	8.87 (5.50–9.30)	9.20 (7.88–9.56)	0.018	Pat-Sib: 0.228 Pat-Con: 0.021 Sib-Con: 0.714
Universal	10.68 (10.16–10.89)	10.84 (10.55–11.03)	10.97 (10.76–11.29)	0.010	Pat-Sib: 0.759 Pat-Con: 0.015 Sib-Con: 0.117

The CD dysbiosis broadly reflected that previously described: reduced Firmicutes (including *F prausnitzii* and Clostridia cluster IV), *Bacteroides-Prevotella*, *B adolescentis* and total bacteria. Similar reductions in Firmicutes species including *F prausnitzii* and Clostridia cluster IV also characterised the sibling dysbiosis, whereas concentrations of *Bacteroides-Prevotella*, *B adolescentis* and the overall concentration of bacteria were similar between siblings and controls. Bacterial DNA was extracted from faecal samples and detected using quantitative PCR probes as detailed in the methods and online supplementary methods.
*Due to Bonferroni correction of p values, a minimum value of <0.003 is reported rather than <0.001.
CD, Crohn's disease.

IV and *Roseburia* spp. compared with controls. However, concentrations of total bacteria, *Bacteroides-Prevotella* and *B adolescentis* were not different between siblings and controls (table 2). The microbiota was also analysed as relative proportions of each group. A similar dysbiosis was described in the patients (see online supplementary table S4), with an additional finding that proportions of *E coli* were increased in patients (0.12% SD 3.95%) compared with controls (0.02% SD 0.10%, p=0.027) but were similar between siblings (0.05% SD 0.14%) and controls (p=0.825).

Proportions of *F prausnitzii* were significantly correlated between sibling pairs (PCC 0.436, p=0.043) as were Clostridia cluster IV (PCC 0.532, p=0.011) and concentrations of Clostridia cluster XIVa (PCC 0.543, p=0.009). Smoking and immunosuppressant use were not associated with dysbiosis. Likewise, IP and FC were not independently associated with dysbiosis (data not shown). In addition to the effect of group in predicting the concentration of *Bacteroides-Prevotella*, (R^2=0.276, $F_{1,66}$=25.221, p<0.001), there was a significant association between elevated/high GRR and concentrations of *Bacteroides-Prevotella* (R^2=0.327, R^2 change=0.050, $F_{1,66}$=4.838, p=0.031), with grouping accounting for 28% and GRR for 5% of the variation. Concentrations and proportions of other bacteria did not vary according to GRR.

Memory T cells were increased in patients and their siblings
The proportion of blood T cells that were CD45RA⁻ memory T cells was significantly higher in patients and siblings compared with controls, figure 1A. The higher proportion of memory T cells observed in patients and siblings compared with controls was

significant in the CD4 T cell subset (65%, SD 13%; 69%, SD 12%; 55%, SD 13%, respectively; p=0.001), and a trend towards higher proportions of memory cells was also seen in CD8 T cells (53%, SD 19%; 57%, SD 17%; 45%, SD 15%, respectively; p=0.054). Total frequencies of peripheral CD3 T cells were significantly lower in patients compared with siblings or controls, online supplementary figure S1. This observation complements previous data indicating that blood T cell numbers are reduced in active CD.[30] Absolute quantitation of T cell subsets revealed that the higher proportion of memory T cells was attributable to lower concentrations of CD45RA naïve T cells in patients and siblings compared with controls, whereas memory T cell frequencies did not differ significantly compared with controls, figure 1B–E.

Naïve CD4 T cells exhibit enhanced gut-homing potential in patients and siblings
The relative proportion of β7⁺ and β7⁻ cells in each T cell subset was analysed individually: proportions of CD4 naïve CD45RA⁺ T cells expressing β7 was significantly higher in patients and siblings compared with controls, figure 2A–D. Proportions of T cells expressing β7 did not significantly differ between groups for any of the other T cell subsets. The proportion of memory T cells expressing β7 (PCC 0.564, p=0.006), but not the proportion of β7⁺ CD4 naïve T cells , was significantly correlated between patient-sibling pairs.

Correlation between factors in healthy siblings
As demonstrated in the whole cohort, there was a negative correlation between the concentration of *Bacteroides-Prevotella* and GRR in siblings (PCC=−0.505, p=0.019). Thus siblings with the
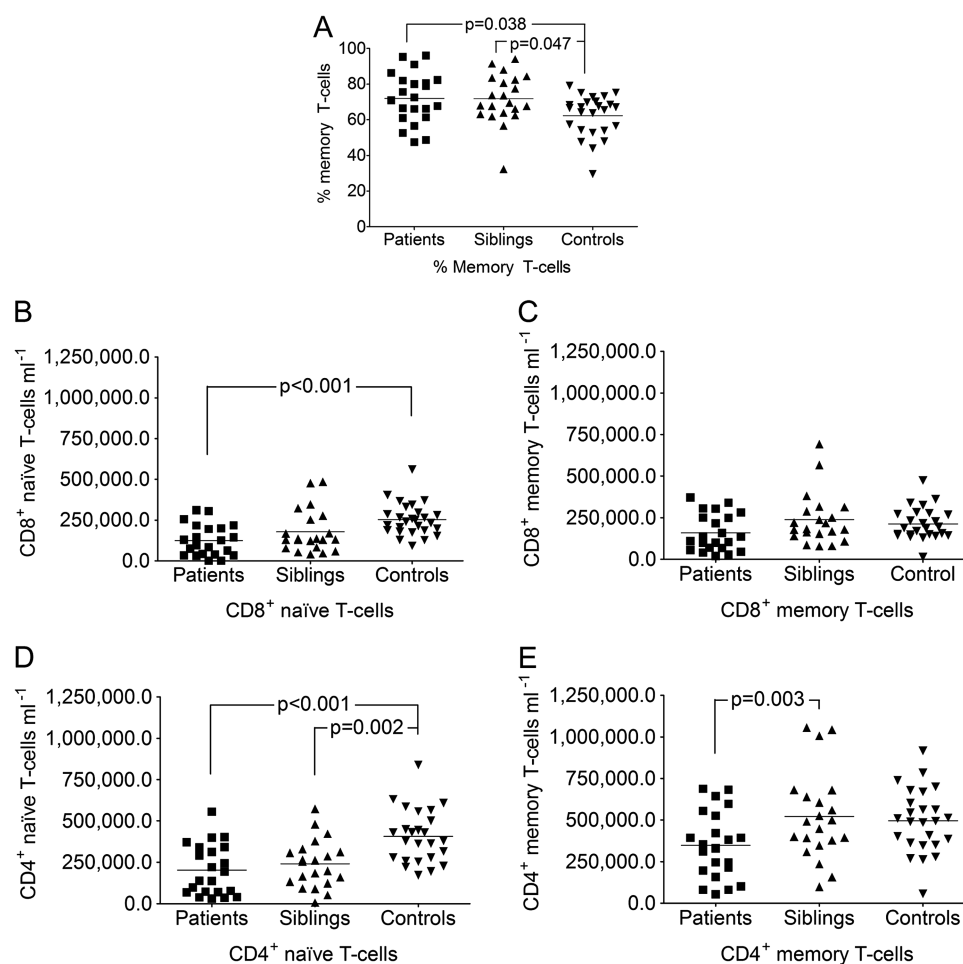
Figure 1 (A) Proportions of peripheral blood T cells with a memory (CD45RA-) phenotype were significantly greater in patients (n=22) and siblings (n=21) compared with healthy controls (n=25). (B–E) The predominance of the memory phenotype was attributable to a relative lack of naïve T cells. Although concentrations of T cells in patients were contributed to by the use of immunosuppressants by half of patients, concentrations of naïve T cells were also reduced in immunosuppressant naïve siblings, and in the CD4 naïve T cell population this difference was significant. Whole blood was labelled with fluorescently conjugated monoclonal antibodies before acquisition of data using a LSRII flow cytometer.

highest genetic risk of CD shared aspects of the CD associated dysbiosis with their affected relative. In addition, there was a negative correlation between the concentration of Clostridia cluster IV and IP (PCC=−0.514, p=0.017), the concentration of blood CD4 naïve T cells (PCC=−0.521, p=0.016) and the proportion of CD4 naïve T cells which expressed β7 integrin (PCC=−0.475, p=0.030). However, there was no correlation between Clostridia cluster IV and FC. Thus, reduced concentrations of Clostridia cluster IV in siblings were associated with several phenotypical characteristics that are associated with CD. Interestingly these correlations were not present in patients with established disease or healthy controls, which would support a role for this aspect of the dysbiosis in the pathogenesis of disease.

The multidimensional phenotype discriminates between patients, siblings and healthy controls

We next performed a discriminant analysis to determine whether combining genetic risk, FC, IP, gut microbiota and immunological parameters could distinguish between patients, 'at-risk' siblings and healthy controls, figure 3. Independent variables were entered into the model and two discriminant functions (linear combinations of the variables, whose values are as close as possible within groups and as far apart as possible between groups) were calculated, with a combined χ^2=80.409, degrees of freedom (df)=20, p<0.001. The two

discriminant functions accounted for 87% and 13% of the between-group variance, respectively. Function 1 maximally separated patients from controls. Function 2 maximally separated siblings from controls, and after removal of function 1, the discriminating power of function 2 remained significant: χ^2=15.569, (df=9), p=0.049. The loading matrix of correlations between predictor variables and discriminant functions suggested that variables contributing significantly to function 1 (providing maximal discrimination between patients and controls) were predominantly microbiological: concentrations of *F prausnitzii*, *Ruminococcus* spp., *Bacteroides-Prevotella* and Clostridia cluster IV, in addition to FC, (table 3). In contrast, the variables contributing significantly to function 2 (providing maximal discrimination between siblings and controls) were T cell attributes including the proportion of CD4 T cells expressing β7 integrin and the proportion of T cells with a memory cell phenotype, as well as concentrations of *Roseburia* spp.

DISCUSSION

This is the first study to provide detailed and simultaneous analysis of the faecal microbiology, IP, FC and blood T cell phenotype of a cohort of clinically and genetically well-characterised patients with quiescent CD. Uniquely, the equivalent detailed assessments were also performed in healthy siblings of the patients with CD and matched unrelated controls. We report

Figure 2 (A–D) Significantly higher proportions of CD4 naïve T cells expressed β7 integrin in patients (n=22) and siblings (n=21) compared with controls, (n=25). Relative abundances of β7⁺ cells did not differ between groups for the other T cell subsets. Whole blood was labelled with fluorescently conjugated monoclonal antibodies before acquisition of data using a LSRII flow cytometer.
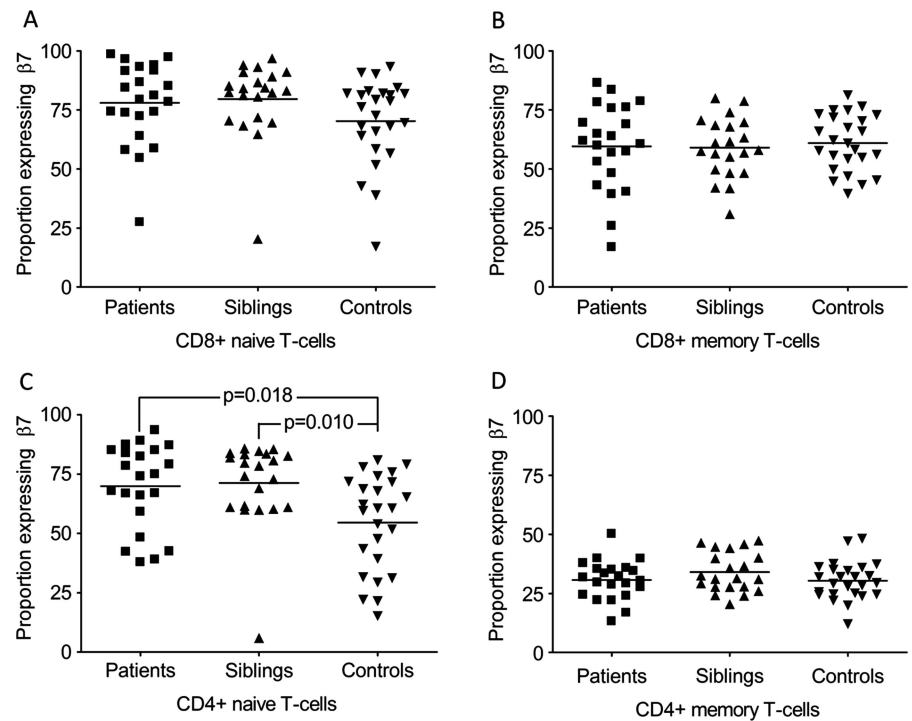
that siblings of patients with CD share multiple features of the CD phenotype in addition to genotype-related risk (table 4), including reduced Firmicutes such as *F prausnitzii*, predominance of the memory phenotype in blood T cells, and elevated FC. We hypothesise that the unaffected sibling phenotype provides a unique window into factors important in CD pathogenesis. Studying healthy, 'at-risk' siblings allows the description of pathogenic factors in the absence of the influence of disease, surgery and pharmacotherapy that may distort the phenotype in patients. Finally, this study provides proof of principle that multidimensional phenotyping can be used to discriminate a population of healthy at-risk individuals from background risk controls.

Genotype relative risk

The patient and sibling cohorts were enriched for genetic loci associated with CD. The genetic validity of the cohort is
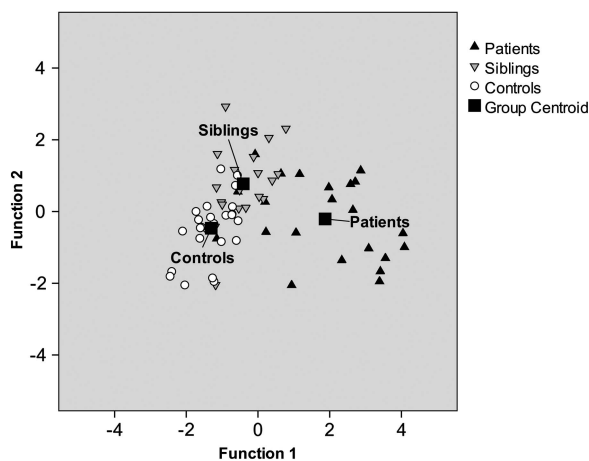

Figure 3 Direct discriminant function analysis revealed that the variables contributing significantly to function 1 (which provided maximal separation of patients with Crohn's disease from controls) were: elevated faecal calprotectin and altered faecal microbiota (reduced *Faecalibacterium prausnitzii*, cluster IV *Ruminococcus* spp., *Bacteroides-Prevotella* and Clostridia cluster IV). In contrast the variables contributing significantly to function 2 (which provided maximal discrimination of siblings from controls) were: increased β7 integrin expression by circulating naïve CD4 T cells, an increased proportion of memory CD4 T cells and reduced faecal *Roseburia* spp.

Table 3 Loading matrix of correlations between predictor variables and the two discriminant functions (figure 3) where function 1 provided most discrimination between patients and controls and function 2 provided maximal discrimination between siblings and controls

	Pooled within-groups correlations between variables and discriminant functions	
	Function 1	**Function 2**
Faecalibacterium prausnitzii, (log₁₀/g)	−0.608*	0.050
Ruminococcus spp., (log₁₀/g)	−0.544*	0.043
Faecal calprotectin, (μg/g)	0.528*	−0.283
Bacteroides-Prevotella, (log₁₀/g)	−0.473*	0.037
Clostridia cluster IV, (log₁₀/g)	−0.359*	−0.163
Genotype relative risk	0.251	0.043
Proportion of CD4⁺ naïve cells expressing β7 integrin, (%)	0.226	0.596*
Roseburia spp., (log₁₀/g)	−0.234	−0.466*
Proportion of memory T cells, (%)	0.196	0.439*

The contribution of each aspect of the multidimensional phenotype to each function is reflected by the pooled within-groups correlations between variables and discriminant functions. The values refer to the strength of the correlation of variables with each discriminant function; the positive/negative symbol refers to the direction of the correlation.
*Factors with a discriminant loading >0.300 are considered to contribute significantly to the discriminant function.
CD, Crohn's disease.

Table 4   Summary of the phenotype of patients with CD and their healthy siblings as compared with unrelated healthy controls

	Patients with CD	Unaffected siblings
Genotype relative risk	↑	↑
Faecal calprotectin	↑↑	↑
Intestinal permeability	↑	↔
Gut microbiota	↓ *Faecalibacterium prausnitzii*/Firmicutes* ↓ *Bacteroides-Prevotella* ↓ *Bifidobacterium adolescentis* ↑ *Escherichia coli*	↓ ***F prausnitzii**/Firmicutes**
Blood T cells	↑ **memory phenotype associated with naïve CD4 T cell lymphopenia** ↑ **expression of β7 integrin by naïve CD4 T cells**	↑ **memory phenotype associated with naïve CD4 T cell lymphopenia** ↑ **expression of β7 integrin by naïve CD4 T cells**

Novel observations from this study are highlighted in **bold**.
*Firmicutes which were reduced were: *F prausnitzii,* Clostridia cluster IV, *Roseburia* spp., and in patients also *Ruminococcus* spp.
CD, Crohn's disease.

supported by higher genetic risk in siblings, a NOD2 mutation rate (patients=41%, controls=16%) similar to previous data[31] and the ratio of siblings to patients carrying any NOD2 mutation. However, the amount of total disease variance that can be attributed to current genetic findings in CD is still low (<15%),[1] hence genotyping alone is inadequate to accurately identify at-risk individuals[12] and further research is required to identify features associated with predisposition towards CD. Combining such information with data on genotype may be a powerful method to detect risk.

Fecal calprotectin and intestinal permeability
Elevated FC and IP are associated with established CD, but only elevated FC appears to contribute to the sibling 'at-risk' phenotype in this study.[13] This would suggest that breach of the intestinal barrier is not an early event in CD pathogenesis. Some data suggest that increased IP may only be apparent in relatives when exposed to triggers such as aspirin ingestion,[15] and thus may not have been detected in the current study. The lack of correlation in FC between sibling pairs may imply that raised FC is not driven by genetic or early environmental factors shared between siblings.

Dysbiosis
Consistent with previous reports,[3] the CD dysbiosis in our study included reduced *F prausnitzii*, Clostridia cluster IV and *Bacteroides-Prevotella*. However, whether this dysbiosis predisposes towards CD or is an outcome of established disease cannot be determined by studying patients in isolation. Strikingly, siblings share elements of the CD dysbiosis, specifically depletion of the Firmicutes *F prausnitzii* and Clostridia cluster IV, suggesting that these elements of CD dysbiosis are not purely consequent to established disease. The focus of the sibling dysbiosis on diminished Firmicutes populations is noteworthy given the range of mechanisms by which Firmicutes contribute to gut health[32] and that phylum member *F prausnitzii* predicts CD natural history.[2] In contrast, the reduction in *Bacteroides-Prevotella* and *B adolescentis* and increased *E coli* seen in patients with CD were not found in their siblings, suggesting that these may become disturbed once disease is established. This study focused on aspects of the CD dysbiosis that have been reported previously and therefore used qPCR. Alternative methods such as 16S pyrosequencing might detect unsuspected alterations in the microbiota in siblings that had not been reported as part of the dysbiosis in CD.

Few previous studies defined the gut microbiota of relatives of patients with CD with reference to healthy controls. One suggested that the organisms comprising the dysbiosis varied between patients with CD and their relatives.[17] These findings are at odds with our study, where the sibling dysbiosis was similar to that in patients. This disparity may relate to previous studies including more heterogeneous groups of relatives. In contrast, the current study recruited only true siblings, potentially increasing the relevance of the dysbiosis we describe to early events in CD pathogenesis. Data from the current study suggest that loss of butyrate producers is common to patients with CD and their siblings, and is potentially the earliest microbiological step towards CD. Butyrate has recently been suggested to regulate the size of the colonic regulatory T cell pool and protect against murine colitis.[33]

In health, the intestinal microbiota is more similar between related than unrelated individuals.[34] However, this similarity is reduced where one relative has CD.[35] Therefore, it would be predicted that siblings in the current study, discordant for disease phenotype, would have dissimilar microbiota. This was seen with *Bacteroides-Prevotella*, *B adolescentis* and *E coli*, but intersibling similarity was preserved for three bacterial groups within the Firmicutes phylum, suggesting this aspect of the CD dysbiosis is either genetically programmed, or that an early environmental exposure has a persistent impact on the microbiota, regardless of the later development of CD. The lack of association between GRR and Firmicutes abundance in the current study may favour prenatal, childhood or adolescent environmental factors in conferring this phenotype. Embryo transfer models demonstrate a maternal effect to determine the microbiota independent of genetic influence.[36] Candidate maternally derived environmental factors include delivery by caesarean section and a protective effect of breastfeeding.[37] However, a genetic influence over a CD-specific aspect of the dysbiosis was suggested by the association between GRR and *Bacteroides-Prevotella*. It may be speculated that the development of CD may require multiple microbiological 'hits', including environmentally acquired dysbiosis (to which siblings are also exposed), as well as genetically driven dysbiosis (more prominent in patients in this study).

T cell phenotype
We demonstrate that memory T cell phenotype predominance, previously described in active CD[5] is also present in inactive CD, particularly among CD4 T cells. However, in contrast to the redistribution of T cells from naïve to memory

compartments previously reported in active CD,[5] we show that in inactive CD blood naïve T cell frequencies are depleted in isolation. In patients with CD, depletion of blood naïve CD4 T cells may therefore be a marker of disease that is independent of inflammation and may precede CD onset. Strikingly, a naïve CD4 T cell lymphopenia was demonstrated for the first time in unaffected siblings of patients with CD. It is unknown whether this naïve T cell lymphopenia reflects reduced generation of naïve T cells, their sequestration outside the circulation or redistribution towards a memory population which has limited persistence in the circulation. It is tempting to speculate that peripheral homoeostasis of naïve T cells, which is driven by low-affinity T cell receptor stimulation and interleukin 7 signalling (and may depend on the presence of the intestinal microbiota[38]), may be disrupted in patients with CD and their siblings. It is noteworthy that depletion of naïve CD4 T cells was associated with increased expression of β7 integrin, a marker of gut tropism. In lymphocyte-replete animals, expression of the integrin α4β7 on naïve T cells facilitates their entry into gut-associated lymphoid tissue.[39 40] However, under lymphopenic conditions, IL-7 is induced which enhances α4β7 expression by naïve T cells and results in their migration into the intestinal mucosa.[41] Therefore, perturbed homoeostasis of naïve CD4 T cells may be indicative of altered trafficking through organised lymphoid tissue and/or the intestinal mucosa in patients with CD and those at risk of CD.

Interactions between dimensions of the CD phenotype

A unique feature of the current study is the simultaneous measurement of a variety of dimensions of the CD phenotype that permits interrogation of relationships between factors within and between groups. Thus the dysbiosis in siblings was associated with phenotypical features associated with CD (increased permeability, the concentration of blood CD4 naïve T cells and the proportion of CD4 naïve T cells which expressed β7 integrin). Whereas, dysbiosis did not correlate with FC in either patients or siblings, indicating that it occurs independently of intestinal inflammation. The proportion of memory T cells correlated between patient-sibling pairs, but the lack of association between GRR and the proportion of memory T cells (and other features) may imply limited genetic determination. Consonant with this, a recent reanalysis of an early twin cohort estimated CD concordance in monozygotic twins to be 27% rather than the previously reported 58%.[42 43] However, the contribution of genotype to CD pathogenesis will be better described as technological advances facilitate the measurement of 'non-code' genetic effects such as epistatic, epigenetic or parent-of-origin effects.

Finally, in addition to detecting correlations between individual factors, our study has allowed a multidimensional comparison of the CD, sibling and healthy control phenotypes, enabling the identification of factors that discriminate healthy siblings of patients with CD from healthy controls. Microbiological factors such as *F prausnitzii* and *Bacteroides-Prevotella* contributed more to discrimination of patients from healthy controls but were less significant in discrimination of the sibling phenotype, which was more dependent on immunological factors. Whether this sibling immune-characterised phenotype precedes CD onset or is an alternative phenotype of individuals who do not possess sufficient cumulative microbial or other triggers required to develop CD, may only be established with longitudinal studies.

Future studies investigating CD risk should simultaneously assess multiple dimensions. The CD and at-risk phenotypes described herein must be confirmed in other populations. In view of the breadth of measures reported in this study, we cannot exclude a type II error in some circumstances. In addition, newly uncovered features of the at-risk phenotype reported in this study such as the importance of *F prausnitzii*, naïve CD4 T cell lymphopenia and altered expression of gut-homing markers provide fresh avenues of investigation into CD risk and pathogenesis. For example, the mechanisms of CD4 T cell lymphopenia in CD and 'at-risk' siblings remain to be determined, and the significance of potential pathogenic mechanisms such as Firmicutes short-chain fatty acid production, must be tested. Longitudinal studies are required to determine the capacity of the multidimensional analysis outlined in this study to predict disease in at-risk groups. Such longitudinal studies will be challenging, although it may be possible to perform repeated measurements at intervals in the offspring of patients with CD. This would have significant potential benefits to patients and their families as well as building on the insights into disease pathogenesis provided in this study.

REFERENCES

1 Jostins L, Ripke S, Weersma RK, *et al*. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012;491:119–24.

2 Sokol H, Pigneur B, Watterlot L, *et al*. *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc Natl Acad Sci USA* 2008;105:16731–6.

3 Sartor RB. Microbial influences in inflammatory bowel diseases. *Gastroenterology* 2008;134:577–94.

4 Strober W, Fuss IJ. Proinflammatory cytokines in the pathogenesis of inflammatory bowel diseases. *Gastroenterology* 2011;140:1756–67.

5 Hart AL, Kamm MA, Knight SC, *et al*. Quantitative and functional characteristics of intestinal-homing memory T cells: analysis of Crohn's disease patients and healthy controls. *Clin Exp Immunol* 2004;135:137–45.

6 Garcia de TJ, Manzano L, Leal JC, *et al*. Active Crohn's disease patients show a distinctive expansion of circulating memory CD4+CD45RO+CD28null T cells. *J Clin Immunol* 2004;24:185–96.

7 Tibble JA, Sigthorsson G, Bridger S, *et al*. Surrogate markers of intestinal inflammation are predictive of relapse in patients with inflammatory bowel disease. *Gastroenterology* 2000;119:15–22.

8 Buhner S, Buning C, Genschel J, *et al*. Genetic basis for increased intestinal permeability in families with Crohn's disease: role of CARD15 3020insC mutation? *Gut* 2006;55:342–7.

9 Meuwis MA, Vernier-Massouille G, Grimaud JC, *et al*. Serum calprotectin as a biomarker for Crohn's disease. *J Crohns Colitis* 2013;7:e678–83.

10 Calkins BM, Mendeloff AI. Epidemiology of inflammatory bowel disease. *Epidemiol Rev* 1986;8:60–91.

11 Hedin CR, Stagg AJ, Whelan K, *et al*. Family studies in Crohn's disease: new horizons in understanding disease pathogenesis, risk and prevention. *Gut* 2012;61:311–18.

12 Manolio TA, Collins FS, Cox NJ, *et al*. Finding the missing heritability of complex diseases. *Nature* 2009;461:747–53.

13 Thjodleifsson B, Sigthorsson G, Cariglia N, *et al*. Subclinical intestinal inflammation: an inherited abnormality in Crohn's disease relatives? *Gastroenterology* 2003;124:1728–37.

14 Munkholm P, Langholz E, Hollander D, *et al*. Intestinal permeability in patients with Crohn's disease and ulcerative colitis and their first degree relatives. *Gut* 1994;35:68–72.

15 Soderholm JD, Olaison G, Lindberg E, *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.

16 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.

17 Joossens M, Huys G, Cnockaert M, *et al*. Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. *Gut* 2011;60:631–7.

18 Turpin W, Moreno-Hagelsieb G, Silverberg MS, *et al*. Specific Crohn's Disease (CD) associated risk alleles are associated with the composition of the intestinal microbiota in healthy First Degree Relatives (FDR) of CD subjects. *Gastroenterology* 2013;142:S57.

19 Franke A, McGovern DP, Barrett JC, *et al*. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet* 2010;42:1118–25.

20 Kraus TA, Cheifetz A, Toy L, *et al*. Evidence for a genetic defect in oral tolerance induction in inflammatory bowel disease. *Inflamm Bowel Dis* 2006;12:82–8.

21 Joshi S, Lewis SJ, Creanor S, *et al*. Age-related faecal calprotectin, lactoferrin and tumour M2-PK concentrations in healthy volunteers. *Ann Clin Biochem* 2010;47:259–63.

22 Woodmansey EJ, McMurdo ME, Macfarlane GT, *et al*. Comparison of compositions and metabolic activities of fecal microbiotas in young adults and in antibiotic-treated and non-antibiotic-treated elderly subjects. *Appl Environ Microbiol* 2004;70:6113–22.

23 Ramirez-Farias C, Slezak K, Fuller Z, *et al*. Effect of inulin on the human gut microbiota: stimulation of Bifidobacterium adolescentis and Faecalibacterium prausnitzii. *Br J Nutr* 2009;101:541–50.

24 Dabek M, McCrae SI, Stevens VJ, *et al*. Distribution of beta-glucosidase and beta-glucuronidase activity and of beta-glucuronidase gene gus in human colonic bacteria. *FEMS Microbiol Ecol* 2008;66:487–95.

25 Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 1987;162:156–9.

26 Cortes A, Brown MA. Promise and pitfalls of the Immunochip. *Arthritis Res Ther* 2011;13:101.

27 Crouch DJ, Goddard GH, Lewis CM. REGENT: a risk assessment and classification algorithm for genetic and environmental factors. *Eur J Hum Genet* 2013;21:109–11.

28 Goddard GH, Lewis CM. Risk categorization for complex disorders according to genotype relative risk and precision in parameter estimates. *Genet Epidemiol* 2010;34:624–32.

29 Mahmood A, Fitzgerald AJ, Marchbank T, *et al*. Zinc carnosine, a health food supplement that stabilises small bowel integrity and stimulates gut repair processes. *Gut* 2007;56:168–75.

30 Selby WS, Jewell DP. T lymphocyte subsets in inflammatory bowel disease: peripheral blood. *Gut* 1983;24:99–105.

31 Mathew CG, Lewis CM. Genetics of inflammatory bowel disease: progress and prospects. *Hum Mol Genet* 2004;13 Spec No 1:R161–8.

32 Flint HJ, Scott KP, Louis P, *et al*. The role of the gut microbiota in nutrition and health. *Nat Rev Gastroenterol Hepatol* 2012;9:577–89.

33 Smith PM, Howitt MR, Panikov N, *et al*. The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science* 2013;341:569–73.

34 Turnbaugh PJ, Hamady M, Yatsunenko T, *et al*. A core gut microbiome in obese and lean twins. *Nature* 2009;457:480–4.

35 Willing B, Dicksved J, Halfvarson J, *et al*. A pyrosequencing study in twins shows that GI microbial profiles vary with inflammatory bowel disease phenotypes. *Gastroenterology* 2010;139:1844–54.

36 Spor A, Koren O, Ley R. Unravelling the effects of the environment and host genotype on the gut microbiome. *Nat Rev Microbiol* 2011;9:279–90.

37 Malmborg P, Bahmanyar S, Grahnquist L, *et al*. Cesarean section and the risk of pediatric Crohn's disease. *Inflamm Bowel Dis* 2012;18:703–8.

38 Kieper WC, Troy A, Burghardt JT, *et al*. Recent immune status determines the source of antigens that drive homeostatic T cell expansion. *J Immunol* 2005;174:3158–63.

39 Bargatze RF, Jutila MA, Butcher EC. Distinct roles of L-selectin and integrins alpha 4 beta 7 and LFA-1 in lymphocyte homing to Peyer's patch-HEV in situ: the multistep model confirmed and refined. *Immunity* 1995;3:99–108.

40 Wagner N, Lohler J, Kunkel EJ, *et al*. Critical role for beta7 integrins in formation of the gut-associated lymphoid tissue. *Nature* 1996;382:366–70.

41 Cimbro R, Vassena L, Arthos J, *et al*. Interleukin-7 induces expression and activation of integrin alpha4beta7 promoting naive T-cell homing to the intestinal mucosa. *Blood* 2012;120:2610–19.

42 Halfvarson J. Genetics in twins with Crohn's disease: less pronounced than previously believed? *Inflamm Bowel Dis* 2011;17:6–12.

43 Tysk C, Lindberg E, Jarnerot G, *et al*. 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.

GUT

Altered intestinal microbiota and blood T cell phenotype are shared by patients with Crohn's disease and their unaffected siblings

Charlotte R Hedin, Neil E McCarthy, Petra Louis, et al.

Gut 2014 63: 1578-1586 originally published online January 7, 2014
doi: 10.1136/gutjnl-2013-306226

Updated information and services can be found at:
http://gut.bmj.com/content/63/10/1578.full.html

These include:

Data Supplement	*"Supplementary Data"* **http://gut.bmj.com/content/suppl/2014/01/07/gutjnl-2013-306226.DC1.html**
References	This article cites 43 articles, 16 of which can be accessed free at: **http://gut.bmj.com/content/63/10/1578.full.html#ref-list-1**
Email alerting service	Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections	Articles on similar topics can be found in the following collections Crohn's disease (895 articles)

Notes