Single nucleotide polymorphisms in *SLC22A23* are associated with ulcerative colitis in a Canadian Caucasian cohort

Alejandra Serrano León[1, 2], Mandana Amir Shaghaghi [1, 2], Natalia Yurkova[1, 2], Charles N. Bernstein[3,4], Hani El-Gabalawy[4], Peter Eck[1, 2]

[1]Human Nutritional Sciences, University of Manitoba, [2] The Richardson Centre for Functional Foods and Nutraceuticals, University of Manitoba, Winnipeg, Manitoba, R3T 2N2, [3]University of Manitoba IBD Clinical and Research Centre, [4]Department of Internal Medicine, University of Manitoba, Winnipeg, Canada, R3E 3P4.

Author´s last name: Serrano León, Amir Shaghaghi, Yurkova, Bernstein, El-Gabalawy, Eck.

Correspondance author: Peter Eck, Department of Human Nutritional Sciences, Richardson Center for Functional Foods and Nutraceuticals, University of Manitoba, 196 Innovation Drive, Winnipeg, MB, Canada, R3T 2N2.

E-mail: peter.eck@ad.umanitoba.ca

Phone: +1-204- 272-1506; Fax: +1-204- 474-7552

Funding source: Peter Eck's Canada Research Chair and a CIHR grant to Drs. Bernstein and El-Gabalawy. Alejandra Serrano León was supported by a University of Manitoba Graduate Fellowship. Mandana Amir Shaghaghi was supported by an NSERC-FAST PhD. Fellowship.

Running title: *SLC22A23* variations associated to UC

Abbreviations: CD, Crohn's disease; HWE, Hardy-Weinberg equilibrium; IBD, Inflammatory bowel disease; LD, linkage disequilibrium; OR, Odds Ratio; PCR-RFLP, polymerase chain

reaction-restriction fragment length polymorphism; SNP, single nucleotide polymorphism; UC, ulcerative colitis;

Keywords: *SLC22A23*, polymorphisms, inflammatory bowel disease, gene-diet interaction

ABSTRACT

Background: *SLC22A23* is an orphan gene in the *SLC22* family of organic membrane transporters, and its single nucleotide polymorphism rs17309827-T was recently nominally associated with intestinal inflammation in a genome wide association study. Other polymorphisms in the *SLC22A23* gene have been associated to diseases with an inflammatory component, and polymorphisms in related genes in the *SLC22* family have been repeatedly associated to inflammatory bowel disease (IBD). **Objective:** In a candidate gene study utilizing a well phenotyped, highly monitored Manitoban Caucasian cohort we investigate whether variations in *SLC22A23* are associated with intestinal inflammation. **Design:** Selected genetic variations were genotyped by TaqMan or PCR-RFLP analysis in 160 individuals with Crohn's disease (CD), 149 individuals with Ulcerative Colitis (UC), and 142 healthy controls to determine genetic associations. **Results:** Homozygocity for SNPs rs4959235-TT and rs950318-GG was associated to IBD, where 6% (18 of 311 cases) of the patients carried these genotypes, while they were not seen in healthy controls. **Conclusion:** The here reported associations add to the emerging evidence that *SLC22A23* variants could modify IBD risk. However, the biology of the gene and the impact of the variations on the genes functions need to be tested to validate a causative role.

INTRODUCTION

Inflammatory Bowel Disease (IBD) is characterized by chronic inflammation of the gastrointestinal tract, with the principal forms Crohn´s disease (CD) and ulcerative colitis (UC). Mostly unknown environmental and genetic factors contribute to the immune dysregulation, which determine the development, maintenance and severity of the disease (1, 2).

Recent advances in genome wide association studies (GWAS) have contributed to the understanding of the genetic factors in intestinal inflammation. Although most genetic variations associated to IBD are in immunoregulatory pathways (3, 4), polymorphisms in two genes encoding organic cation transporters, *SLC22A4* and *SLC22A5,* have been repeatedly associated to IBD (5-18). Organic cation transporter proteins are involved in the absorption, distribution and excretion of therapeutic drugs, xenobiotics, and food compounds (19). Functional variations in organic cation transporter genes could impact the cellular uptake, pharmacokinetics and therefore the bioactivity of unknown organic ions, which might modulate susceptibility and severity of IBD or other related common and complex diseases. Hence, we hypothesize that organic ion imbalance as a consequence of polymorphisms in transporter genes contributes to IBD etiology.

Recently, the single nucleotide polymorphism (SNP) rs17309827-T, located in a gene in the SLC22 transporter family, *SLC22A23*, has been associated to IBD in a GWAS (3, 20), strengthening the hypothesis that imbalances of organic ions might modulate IBD risk. In addition to IBD, polymorphisms in the *SLC22A23* gene have been associated with other complex diseases which have an inflammatory component, such as endometriosis-related infertility (21), and the clearance of antipsychotic drugs (22). *SLC22A23* SNP rs17136561 was nominally associated with developing asthma in individuals with impaired allergic status, although the

46 results did not reach overall genome wide significance (OR = 1.64; 95% CI: 1.33, 2.02; p =
47 2.3×10^{-6}) (23). *SLC22A23* was also one of six genes in which expression levels could be used to
48 predict the recurrence of triple negative breast cancer in a cohort of Taiwanese woman (24).
49
50 Based on the emerging evidence associating genetic variations in *SLC22A23* with IBD and other
51 diseases having an inflammatory component, and the fact that the gene is related to *SLC22A4* and
52 *SLC22A5*, which have repeatedly been associated with IBD, further study of the gene and its
53 polymorphisms is warranted. Here we report the results of a candidate gene study we conducted
54 on a well phenotyped Manitoban Caucasian IBD cohort, which show polymorphisms of the
55 *SLC22A23* gene associated to UC, CD and IBD overall.

SUBJECTS AND METHODS

Study Population

The study population included IBD patients from the Canadian Manitoba Inflammatory Bowel Disease Cohort initiated in 2002 and described previously (25). A total of 311 age and gender matched Caucasian IBD patients (162 CD, 149 UC) and 142 healthy controls drawn from the general population who did not have personal or first degree relatives with any chronic immune disease were included in the study. CD and UC status was determined based on radiologic, endoscopic and histologic data and phenotype was assigned according to the Montreal classification (26), and are shown in **Table 1.**

Genotyping

Genomic DNA was isolated from blood as described previously (27). Polymorphism rs17309827 was genotyped by PCR-RFLP analysis, and rs9503518, rs6923667 and rs4959235 were genotyped by TaqMan® SNP genotyping assays (Applied Biosystems®, Foster City, CA).

To determine rs17309827 genotypes, 219 nucleotides (nt) were amplified between the sense primer GGAACGTACAATTCTGCA and the antisense primer GCATGTGAGCGTTTGATG using Taq Polymerase (New England BioLabs, Ipswich, MA) following the manufacturer's protocol and the following cycling conditions: initial denaturation at 95ºC for 30s, followed by 35 cycles of denaturation at 95ºC for 15s, annealing at 50ºC for 15s, extension at 68ºC for 2 min and final extension at 68ºC for 5 min. Amplicons were digested with NlaIII (New England BioLabs) using the manufacturer's recommendations, where the enzyme cuts the G variant in fragments of 171 nt and 48 nt, while the T variant does not possess a NlaIII restriction site. The restriction patterns of individual samples were determined after gel electrophoresis at a 2% UltraPure

agarose gel (Invitrogen, Burlington, ON) via ethidium bromide staining under UV light and imaged in a Gel Doc (BIO-RAD, Hercules, CA). The efficiency of the amplification and restriction digest was determined by incorporating positive and negative controls amplified from previously subcloned amplicon of known genotype in each set of RFLP analysis.

In addition, rs9503518, rs6923667, rs4959235 were genotyped by TaqMan® SNP genotyping assays catalogue number 4351379, 4351379, 4351379, respectively. (Applied Biosystems, Foster City, CA), using the TaqMan GTXpress Master Mix and the assays as recommended by the manufacturer.

Statistical analyses

Polymorphisms were tested for Hardy-Weinberg equilibrium (HWE). The case-control association of genotypes for each SNP with CD, UC, and overall IBD risk was tested, using logistic regression. Odds ratios (ORs) and 95% confidence interval (CIs) of ORs were determined for two of the three genotypes (carrier of the minor alleles) relative to the third genotype (homozygote for major allele) as a reference genotype. The analysis was carried out using SAS 9.2. The linkage disequilibrium (LD) and haplotype blocks was performed with Haploview 4.2 (28) using the default method (29).

Predicted functional impact of the SNPs and population distributions

The predicted functional impact of individual SNPs were assessed via the SNP Function Prediction tool FuncPred (30), which includes assessment of functional implications of non-synonymous SNPs, splicing regulation, stop codon changes, PolyPhen, SNPs3D, transcription factor binding sites, microRNA binding sites, regulatory potential, and conservation.

The Hapmap (31) data were assessed via the Ensembl Genome browser (32) and processed using Microsoft Excel. All databases analysis was conducted between July 10th and 22nd 2013.

RESULTS

Genotypes distributions conform to the Hardy-Weinberg equilibrium for all groups, and frequencies are listed in **Tables 2 and 3**. All frequencies are in the range reported for Caucasian populations in the HapMap (31) and SNP500 (33).

Homozygote individuals for genotypes rs4959235-TT and rs9503518-GG were exclusively found in UC and CD patients and significantly associated to the respective phenotypes (**Table 2**). Due to an overrepresentation in UC patients, the rs9503518-G allele is also associates to this phenotype (**Table 2**).

Moreover, rs4959235-TT and rs9503518-GG homozygosity is associated with IBD overall, due to the fact that both genotypes are exclusive to the patients and not healthy controls (**Table 3**). The Odds Ratios (OR) for the disease associated genotypes as calculated by logistic regression were called infinite (∞), decisively establishing significance despite the fact that p-values could not be called for the observed genotype distribution, since comparison with a group equaling zero is not permissive. Therefore we extrapolate significance through the following rationale: when the genotype distribution between IBD and Control populations is considered the overall probability of IBD is 311/ (311+142) = 0.69. Nine distinct disease individuals but no healthy controls carry the genotypes rs4959235-TT and rs9503518-GG, and the likelihood for such a distribution is 0.69^{9} =0.0354. From this very low likelihood we make the conjecture that the homozygotes for both SNPs have a significantly higher rate of affliction.

In addition, rs6923667-CT heterozygosity was significantly associated with CD, not UC, compared to controls **(Table 2)**.

131 None of the SNPs were associated with any phenotypic subtype of CD or UC.

DISCUSSION

Genotypes rs4959235-TT and rs950318-GG in the *SLC22A23* gene were associated with UC, CD, and IBD overall in our Manitoban Caucasian cohort. Both genotypes were exclusively present in IBD patients but not in healthy controls. Previously, another *SLC22A23* allele, rs17309827-T, was nominally associated to IBD in a GWAS meta-analysis (3) and independently to the risk of stricturing phenotypes in CD (34). Thus, our data add to the emerging evidence for associations of *SLC22A23* genotypes with IBD.

While, the results of our study are in agreement with results of the aforementioned GWAS meta-analysis (3), it is in contrast with reports from Swedish Caucasian cohort (35-37) and a British UC cohort (38). The discrepancy between the results may be explained by differences in the nature of the cohorts, as well as inter-study discrepancies producing different confounders, such as the selection of tag-SNPs, different ethnical backgrounds, phenotyping methodology or underpowered datasets. While our cohort is moderately sized, the validity of our results is supported by the closely monitored, well phenotyped nature of the cohort and further demonstrated by the high number of studies this cohort has produced replicating GWAS findings of IBD associations (39).

Significantly, our studies are distinguished from others investigating associations to IBD by the decisive results we report that homozygotes for the novel SNPs rs4959235-TT and rs950318-GG are exclusively present in the Manitoban Caucasian IBD patients but not at all in the healthy controls (**Tables 2, 3**). Our studies further indicate that both SNPs affect IBD risk independently since they are in distinct linkage blocks (**Figure 1**) and show distinct pattern of inheritance. Homozygotes for SNP rs4959235-TT have elevated disease risk (**Table 2**), indicative of

recessive inheritance. In contrast, both rs9503518-GG homozygosity as well as the presence of the G-allele elevates the disease risk relative to the allelic presence (**Table 2**).

Due to indication of recessive inheritance, we further investigated if SNP rs4959235-TT homozygosity could play a role in disease risk across ethnicities, using existing HapMap data. The IBD risk allele rs4959235-T is not found in Asian and sub-Saharan populations, where IBD incidence rates are lowest (**Table 4**) (40). The rs4959235-T allele reaches allelic frequencies of up to 19% in European cohorts (**Table 4**), which are known to have a higher IBD burden (41). Significantly, homozygotes for rs4959235-TT are not documented in healthy individuals of four out of five European subpopulations, despite the presence of the T-allele in 15%-19% of the individuals **(Table 4)**. This is in concordance with the data in our cohort of Manitoban Caucasians, where we did not observe rs4959235-TT homozygotes in the healthy individuals. In contrast, rs4959235-TT homozygotes are reported in cohorts of American ancestry, where C/T-heterozygocity is similar to Europeans (**Table 4**). The presence of healthy TT-homozygotes in American ancestry might indicate a less severe disease phenotype due to the distinct genetic background, which could explain the retention of both alleles throughout human evolution.

While current prediction tools do not indicate functional impact for either SNP rs4959235 (which is non-coding in intron four) or SNP rs9503518 (which is synonymous in exon 10) (30), further study is warranted to determine whether these SNPs tag causative variations. Recent next generation sequencing data confirm this possibility, as eight rare non-synonymous variations and one frame-shift mutation predicted to be detrimental have been identified within the linkage region of these SNPs **(Table 5)** (42). These should be considered in future genetic and functional investigations.

SLC22A23 is an orphan gene in the SLC22 family of organic cation and anion transporters. Its genomic organization and variations have not been comprehensively described, and its substrate is unknown (43) (http://www.ncbi.nlm.nih.gov/gene/63027). Our results, however, imply that *SLC22A23's* unknown substrate could be an intracellular antioxidant. As noted, variations in two related genes, *SLC22A4* and *SLC22A5,* encode for membrane transporters of ergothioneine (44-46) and carnitine (47-50) have consistently been associated with IBD (5, 7, 10, 14-16, 18). Since both carnitine and ergothioneine are intracellular antioxidants (51-53), we hypothesize that altered transport kinetics could lead to disturbed intracellular redox state and thus elevated IBD susceptibility. Our novel hypothesis that intracellular redox status could be compromised through altered membrane transport of antioxidants is further supported by the recent association of SNPs in the ascorbic acid transporter *SLC23A1* with CD (54).

In conclusion, we present evidence that genetic variations in the *SLC22A23* locus contribute to the susceptibility to IBD in a small fraction of Caucasian patients. Considering existing association and biological studies we theorize that genetic variations in membrane transporters contribute to disturbed intracellular redox status through altered antioxidants uptake. If future biological studies validate the proposed mechanism(s) of redox disturbance, specific dietary intervention can be developed for individuals carrying the risk genotypes. A genotype specific nutritional intervention could offer a cost-effective alternative to existing immunomodulation therapy with significantly reduced adverse effects. However, it needs to be cautioned that current knowledge suggests that applicability might be very specific and likely limited. Of the 163 independent IBD risk loci identified to date (55), the effect sizes of individual SNPs explain at best 14% of the disease phenotype (56). At present, *SLC22A23* variations, or variations in other

membrane transporters, account for significantly less effect size. Nevertheless, future investigations could reveal transporter variations having greater impact. Given that, individuals with risk genotypes can be identified cost effectively, and with IBD incidences rising worldwide, growing numbers could benefit from novel nutritional interventions in the future.

ACKNOWLEDGMENT

We would to thank Dr. Garry Crow at the Department of Animal Sciences, University of Manitoba, for providing his statistical advice. We also acknowledge the valuable editing done by Mrs. Kerry Edwards. The authors' responsibilities were as follows—PE: was the principal investigator, responsible for concept and design of the study, contributed to all drafts, edited and finalized the manuscript; ASL: conducted laboratory analysis, data collection and analysis, and drafted the manuscript; NY: contributed to the laboratory analysis; CNB: provided the samples from the Manitoba IBD Cohort, participated in editing and revising of the manuscript; CNB and HEG: provided healthy control samples. MAS contributed to the statistical analysis, the discussion and the final edits of the manuscript. ASL and MAS contributed equally to the manuscript. All authors reviewed the manuscript. PE, MAS, ASL, HEG and NY had no conflict of interest. CNB declares to have consulted in Abbott Canada, Abbvie Canada, Bristol Myers Squibb, Janssen Canada, Forest Canada, Takeda Canada and Vertex Pharmaceuticals and has received research grants from Abbott Canada and Abbvie Canada and an education grant from Aptalis. He is supported in part by the Bingham Chair in Gastroenterology.

The Manitoba IBD Cohort Study was initiated in 2001.

REFERENCES

1. Podolsky DK. Inflammatory bowel disease. *New Engl J Med.* 2002;347(6):417-29.
2. Weizman AV, and Silverberg MS. Have genomic discoveries in inflammatory bowel disease translated into clinical progress? *Current Gastroenterology Reports.* 2012;14(2):139-45.
3. Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, Rioux JD, Brant SR, Silverberg MS, Taylor KD, Barmada MM, et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet.* 2008;40(8):955-62.
4. Lees CW, Barrett JC, Parkes M, and Satsangi J. New IBD genetics: Common pathways with other diseases. *Gut.* 2011;60(12):1739-53.
5. Hradsky O, Dusatkova P, Lenicek M, Bronsky J, Duricova D, Nevoral J, Vitek L, Lukas M, and Cinek O. Two independent genetic factors responsible for the associations of the IBD5 locus with Crohn's disease in the Czech population. *Inflamm Bowel Dis.* 2011;17(7):1523-9.
6. Lakner L, Csongei V, Sarlos P, Jaromi L, Safrany E, Varga M, Orosz P, Magyari L, Bene J, Miheller P, et al. IGR2096a_1 T and IGR2198a_1 C alleles on IBD5 locus of chromosome 5q31 region confer risk for Crohn's disease in Hungarian patients. *Int J Colorectal Dis.* 2009;24(5):503-7.
7. Martinez A, Martin MC, Mendoza JL, Taxonera C, Diaz-Rubio M, de la Concha EG, and Urcelay E. Association of the organic cation transporter OCTN genes with Crohn's disease in the Spanish population. *Eur J Hum Genet.* 2006;14(2):222-6.
8. Newman B, Gu X, Wintle R, Cescon D, Yazdanpanah M, Liu X, Peltekova V, Van Oene M, Amos CI, and Siminovitch KA. A risk haplotype in the Solute Carrier Family 22A4/22A5 gene cluster influences phenotypic expression of Crohn's disease. *Gastroenterology.* 2005;128(2):260-9.
9. Noble CL, Nimmo ER, Drummond H, Ho GT, Tenesa A, Smith L, Anderson N, Arnott IDR, and Satsangi J. The Contribution of OCTN1/2 Variants Within the IBD5 Locus to Disease Susceptibility and Severity in Crohn's Disease. *Gastroenterology.* 2005;129(6):1854-64.
10. Peltekova VD, Wintle RF, Rubin LA, Amos CI, Huang Q, Gu X, Newman B, Van Oene M, Cescon D, Greenberg G, et al. Functional variants of OCTN cation transporter genes are associated with Crohn disease. *Nat Genet.* 2004;36(5):471-5.
11. Repnik K, and Potočnik U. Haplotype in the IBD5 region is associated with refractory Crohn's disease in Slovenian patients and modulates expression of the SLC22A5 gene. *Journal of Gastroenterology.* 2011;46(9):1081-91.
12. Rioux JD, Daly MJ, Silverberg MS, Lindblad K, Steinhart H, Cohen Z, Delmonte T, Kocher K, Miller K, Guschwan S, et al. Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. *Nat Genet.* 2001;29(2):223-8.
13. Rioux JD, Silverberg MS, Daly MJ, Steinhart AH, McLeod RS, Griffiths AM, Green T, Brettin TS, Stone V, Bull SB, et al. Genomewide search in Canadian families with inflammatory bowel disease reveals two novel susceptibility loci. *Am J Hum Genet.* 2000;66(6):1863-70.
14. Russell RK, Drummond HE, Nimmo ER, Anderson NH, Noble CL, Wilson DC, Gillett PM, McGrogan P, Hassan K, Weaver LT, 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(8):1114-23.
15. Silverberg MS, Duerr RH, Brant SR, Bromfield G, Datta LW, Jani N, Kane SV, Rotter JI, Philip Schumm L, Hillary Steinhart A, et al. Refined genomic localization and ethnic differences observed for the IBD5 association with Crohn's disease. *Eur J Hum Genet.* 2007;15(3):328-35.

16. Torok HP, Glas J, Tonenchi L, Lohse P, Muller-Myhsok B, Limbersky O, Neugebauer C, Schnitzler F, Seiderer J, Tillack C, et al. Polymorphisms in the DLG5 and OCTN cation transporter genes in Crohn's disease. *Gut.* 2005;54(10):1421-7.
17. Wang J, Wang X, Yang H, Wu D, Wang L, and Qian J. Contribution of the IBD5 locus to inflammatory bowel disease: a meta-analysis. *Hum Genet.* 2011;129(6):597-609.
18. Xuan C, Zhang B-B, Yang T, Deng K-F, Li M, and Tian R-J. Association between OCTN1/2 gene polymorphisms (1672C-T, 207G-C) and susceptibility of Crohn's disease: a meta-analysis. *International Journal of Colorectal Disease.* 2012;27(1):11-9.
19. Tamai I. Pharmacological and pathophysiological roles of carnitine/organic cation transporters (OCTNs: SLC22A4, SLC22A5 and Slc22a21). *Biopharmaceutics & drug disposition.* 2013;34(1):29-44.
20. Franke A, McGovern DPB, Barrett JC, Wang K, Radford-Smith GL, Ahmad T, Lees CW, Balschun T, Lee J, Roberts R, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet.* 2010;42(12):1118-25.
21. Zhao ZZ, Croft L, Nyholt DR, Chapman B, Treloar SA, Hull ML, and Montgomery GW. Evaluation of polymorphisms in predicted target sites for micro RNAs differentially expressed in endometriosis. *Mol Hum Reprod.* 2011;17(2):92-103.
22. Aberg K, Adkins DE, Liu Y, McClay JL, Bukszar J, Jia P, Zhao Z, Perkins D, Stroup TS, Lieberman JA, et al. Genome-wide association study of antipsychotic-induced QTc interval prolongation. *Pharmacogenomics J.* 2012;12(2):165-72.
23. Ramasamy A, Kuokkanen M, Vedantam S, Gajdos ZK, Couto Alves A, Lyon HN, Ferreira MA, Strachan DP, Zhao JH, Abramson MJ, et al. Genome-wide association studies of asthma in population-based cohorts confirm known and suggested loci and identify an additional association near HLA. *PLoS One.* 2012;7(9):e44008.
24. Chen LH, Kuo WH, Tsai MH, Chen PC, Hsiao CK, Chuang EY, Chang LY, Hsieh FJ, Lai LC, and Chang KJ. Identification of prognostic genes for recurrent risk prediction in triple negative breast cancer patients in Taiwan. *PLoS One.* 2011;6(11).
25. Clara I, Lix LM, Walker JR, Graff LA, Miller N, Rogala L, Rawsthorne P, and Bernstein CN. The Manitoba IBD Index: evidence for a new and simple indicator of IBD activity. *Am J Gastroenterol.* 2009;104(7):1754-63.
26. Satsangi J, Silverberg MS, Vermeire S, and Colombel JF. The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. *Gut.* 2006;55(6):749-53.
27. Cantor MJ, Nickerson P, and Bernstein CN. The role of cytokine gene polymorphisms in determining disease susceptibility and phenotype in inflammatory bowel disease. *Am J Gastroenterol.* 2005;100(5):1134-42.
28. Barrett JC, Fry B, Maller J, and Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics.* 2005;21(2):263-5.
29. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, et al. The structure of haplotype blocks in the human genome. *Science.* 2002;296(5576):2225-9.
30. Xu Z, and Taylor JA. SNPinfo: integrating GWAS and candidate gene information into functional SNP selection for genetic association studies. *Nucleic Acids Res.* 2009;37(suppl 2):W600-W5.
31. International HapMap C, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, et al. A second generation human haplotype map of over 3.1 million SNPs. *Nature.* 2007;449(7164):851-61.
32. Flicek P, Ahmed I, Amode MR, Barrell D, Beal K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fairley S, et al. Ensembl 2013. *Nucleic Acids Res.* 2013;41(Database issue):D48-55.

33. Packer BR, Yeager M, Burdett L, Welch R, Beerman M, Qi L, Sicotte H, Staats B, Acharya M, Crenshaw A, et al. SNP500Cancer: a public resource for sequence validation, assay development, and frequency analysis for genetic variation in candidate genes. *Nucleic Acids Res.* 2006;34(Database issue):D617-21.
34. Henckaerts L, Van Steen K, Verstreken I, Cleynen I, Franke A, Schreiber S, Rutgeerts P, and Vermeire S. Genetic Risk Profiling and Prediction of Disease Course in Crohn's Disease Patients. *Clin Gastroenterol H.* 2009;7(9):972-80.
35. McGovern DP, Gardet A, Torkvist L, Goyette P, Essers J, Taylor KD, Neale BM, Ong RT, Lagace C, Li C, et al. Genome-wide association identifies multiple ulcerative colitis susceptibility loci. *Nat Genet.* 2010;42(4):332-7.
36. McGovern DPB, Jones MR, Taylor KD, Marciante K, Yan X, Dubinsky M, Ippoliti A, Vasiliauskas E, Berel D, Derkowski C, et al. Fucosyltransferase 2 (FUT2) non-secretor status is associated with Crohn's disease. *Hum Mol Genet.* 2010;19(17):3468-76.
37. Törkvist L, Halfvarson J, Ong RTH, Lördal M, Sjöqvist U, Bresso F, Björk J, Befrits R, Löfberg R, Blom J, et al. Analysis of 39 Crohn's disease risk loci in Swedish inflammatory bowel disease patients. *Inflammatory Bowel Diseases.* 2010;16(6):907-9.
38. Anderson CA, Massey DCO, Barrett JC, Prescott NJ, Tremelling M, Fisher SA, Gwilliam R, Jacob J, Nimmo ER, Drummond H, et al. Investigation of Crohn's Disease Risk Loci in Ulcerative Colitis Further Defines Their Molecular Relationship. *Gastroenterology.* 2009;136(2):523-9.
39. Ryan JD, Silverberg MS, Xu W, Graff LA, Targownik LE, Walker JR, Carr R, Clara I, Miller N, Rogala L, et al. Predicting complicated Crohn's disease and surgery: phenotypes, genetics, serology and psychological characteristics of a population-based cohort. *Aliment Pharmacol Ther.* 2013;38(3):274-83.
40. Yang SK, Loftus EV, Jr., and Sandborn WJ. Epidemiology of inflammatory bowel disease in Asia. *Inflamm Bowel Dis.* 2001;7(3):260-70.
41. Molodecky NA, Soon IS, Rabi DM, Ghali WA, Ferris M, Chernoff G, Benchimol EI, Panaccione R, Ghosh S, Barkema HW, et al. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology.* 2012;142(1):46-54 e42; quiz e30.
42. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, and Sirotkin K. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* 2001;29(1):308-11.
43. Amberger J, Bocchini C, and Hamosh A. A new face and new challenges for Online Mendelian Inheritance in Man (OMIM(R)). *Hum Mutat.* 2011;32(5):564-7.
44. Gründemann D, Harlfinger S, Golz S, Geerts A, Lazar A, Berkels R, Jung N, Rubbert A, and Schömig E. Discovery of the ergothioneine transporter. *P Natl Acad Sci USA.* 2005;102(14):5256-61.
45. Nakamura T, Yoshida K, Yabuuchi H, Maeda T, and Tamai I. Functional characterization of ergothioneine transport by rat organic cation/carnitine transporter Octn1 (slc22a4). *Biological and Pharmaceutical Bulletin.* 2008;31(8):1580-4.
46. Taubert D, Grimberg G, Jung N, Rubbert A, and Schömig E. Functional role of the 503F variant of the organic cation transporter OCTN1 in Crohn's disease [7]. *Gut.* 2005;54(10):1505-6.
47. Nezu JI, Tamai I, Oku A, Ohashi R, Yabuuchi H, Hashimoto N, Nikaido H, Sai Y, Koizumi A, Shoji Y, et al. Primary systemic carnitine deficiency is caused by mutations in a gene encoding sodium ion-dependent carnitine transporter. *Nat Genet.* 1999;21(1):91-4.
48. Sekine T, Kusuhara H, Utsunomiya-Tate N, Tsuda M, Sugiyama Y, Kanai Y, and Endou H. Molecular cloning and characterization of high-affinity carnitine transporter from rat intestine. *Biochem Bioph Res Co.* 1998;251(2):586-91.

49. Tamai I, Ohashi R, Nezu JI, Yabuuchi H, Oku A, Shimane M, Sai Y, and Tsujit A. Molecular and functional identification of sodium ion-dependent, high affinity human carnitine transporter OCTN2. *Journal of Biological Chemistry.* 1998;273(32):20378-82.
50. Wang Y, Ye J, Ganapathy V, and Longo N. Mutations in the organic cation/carnitine transporter OCTN2 in primary carnitine deficiency. *P Natl Acad Sci USA.* 1999;96(5):2356-60.
51. Gülçin I. Antioxidant and antiradical activities of L-carnitine. *Life Sciences.* 2006;78(8):803-11.
52. Kalaiselvi T, and Panneerselvam C. Effect of L-carnitine on the status of lipid peroxidation and antioxidants in aging rats. *J Nutr Biochem.* 1998;9(10):575-81.
53. Cheah IK, and Halliwell B. Ergothioneine; antioxidant potential, physiological function and role in disease. *Biochimica et Biophysica Acta - Molecular Basis of Disease.* 2012;1822(5):784-93.
54. Amir Shaghaghi M, Bernstein CN, Serrano León A, El-Gabalawy H, and Eck P. Polymorphisms in the sodium-dependent ascorbate transporter gene SLC23A1 are associated with susceptibility to Crohn disease. *The American Journal of Clinical Nutrition.* 2014.
55. Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, Lee JC, Philip Schumm L, Sharma Y, Anderson CA, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature.* 2012;491(7422):119-24.
56. Beaudoin M, Goyette P, Boucher G, Lo KS, Rivas MA, Stevens C, Alikashani A, Ladouceur M, Ellinghaus D, Torkvist L, et al. Deep resequencing of GWAS loci identifies rare variants in CARD9, IL23R and RNF186 that are associated with ulcerative colitis. *PLoS Genet.* 2013;9(9):e1003723.

Table 1. Phenotypic characteristics of the Caucasian IBD cohort population

Parameters	**Crohn´s Disease (n=162)**	**Ulcerative Colitis (n= 149)**
Gender		
Female	97 (59.9%)	87 (58.4%)
Male	65 (40.1%)	62 (41.6%)
Age at diagnosis		
A1(<17 years)	17 (10.5%)	12 (8.1%)
A2 (17-40 years)	101 (62.3%)	78 (52.3%)
A3 (>40 years)	44 (27.2%)	59 (39.6%)
Location		
L1 (Ileal)	69 (42.6%)	-
L2 (Colonic)	37 (22.8%)	-
L3 (Ileocolonic)	51 (31.5%)	-
L4 (isolated upper disease)	5 (3.1%)	-
E1 (UP limited to rectum)	-	11 (7.4%)
E2 (Left sided, distal)	-	68 (45.6%)
E3 (extensive, pancolitis)	-	70 (47.0%)
Behaviour		
B1(Inflammatory)	69 (42.6%)	-
B2 (Stricturing)	54 (33.3%)	-
B3 (Penetrating/fistulizing)	39 (24.1%)	-

Note: Disease cohorts were matched for sex and age to a total of 142 non related healthy controls (80 females, 62 males). No significant difference was found between the baseline characteristics of the study populations.

Table 2. Association of *SLC22A23* SNPs with Ulcerative Colitis and Crohn's disease

SNP	Crohn´s disease *n*= 162	Ulcerative Colitis *n*= 149	Controls *n*=142
rs4959235			
CC	141 (87%)	125 (83.9%)	119 (83.8%)
CT	18 (11.1%)	18 (12.1%)	23 (16.2%)
TT	3 (1.9%)*	6 (4.0%)*	0 (0%)*
T-carrier	21 (12.9%)	24 (16.0%)	23 (16.2%)
rs6923667			
CC	51 (31.5%)	62 (41.6%)	57 (40.1%)
CT	92 (56.8%)*	66 (44.3%)	61 (43%)
TT	19 (11.7%)	21 (14.1%)	24 (16.9%)
T-carrier	111 (68.5%)	87 (58.4%)	85 (59.8%)
rs9503518			
AA	139 (85.8%)	113 (75.8%)	118 (83.1%)
AG	22 (13.6%)	28 (18.8%)	24 (16.9%)
GG	1 (0.6%)*	8 (5.4%)*	0 (0%)*
G-carrier	23 (14.2%)	36 (24.2%)*	24 (16.9%)
rs17309827			
GG	21 (13.1%)	23 (15.4%)	20 (14.1%)
GT	81 (50.6%)	60 (40.3%)	70 (49.3%)
TT	58 (36.3%)	66 (44.3%)	52 (36.6%)
T-carrier	139 (85.8%)	126 (84.6%)	122 (85.9%)

Associations determined by logistic regression and indicated through a star * ($p < 0.05$)

Table 3. Association of *SLC22A23* SNPs with IBD overall

SNP	IBD n=311	Controls *n*=142	OR (%95)
rs4959235			
CC	266 (85.5%)	119 (83.8%)	Ref
CT	36 (11.6%)	23 (16.2%)	0.70 (0.39-1.23)
TT	9 (2.9%)*	0 (0%)*	∞*
T-carrier	54 (8.7%)	23 (8.1%)	1.21 (0.86, 1.70)
rs6923667			
CC	113 (36.3%)	57 (40.1%)	Ref
CT	158 (50.8%)	61 (43%)	1.31 (0.85, 1.52)
TT	40 (12.9%)	24 (16.9%)	0.84 (0.46, 1.53)
T-carrier	238 (38.3%)	109 (38.3%)	1.08 (0.85, 1.38)
rs9503518			
AA	252 (81.0%)	118 (83.1%)	Ref.
AG	50 (16.1 %)	24 (16.9%)	0.97 (0.57, 1.66)
GG	9 (2.9%)*	0 (0%)*	∞*
G-carrier	68 (10.9%)	24 (8.4%)	1.32 (0.95, 1.85)
rs17309827			
GG	44 (14.2%)	20 (14.1%)	Ref.
GT	141 (45.6%)	70 (49.3%)	0.91 (0.50, 1.67)
TT	124 (40.2%)	52 (36.6%)	1.08 (0.58, 2.01)
T-carrier	389 (62.9%)	174 (61%)	1.14 (0.89, 1.45)

Note: Associations determined by logistic regression and significant associations were determined as followed: The Odds Ratios (OR) calculated by logistic regression were called infinite (∞) for the disease associated genotypes, decisively establishing significance despite the fact that p-values could not be called for the observed genotype distribution, since comparison with a group equaling zero is not permissive. Significance is established through the following rationale: when the genotype distribution between IBD and Control populations is considered the overall probability of IBD is 311/ (311+142) = 0.69. Nine distinct disease individuals but no healthy controls carry the genotypes rs4959235-TT and rs9503518-GG, and the likelihood for such a distribution is 0.69^{9} =0.0354. From this very low likelihood we make the conjecture that

the homozygotes for both SNPs have a significantly higher rate of affliction. The nine cases of rs4959235-TT and rs9503518-GG homozygocity are found in distinct individuals.

Significant associations are marked with a star *.

Table 4. rs4959235 allele and genotype frequencies in selected HapMap cohorts

Ancestry	Population	Allele		Genotype		
		C	T	C|C	C|T	T|T
European	CEU	0.912	0.088	0.824	0.176	
	FIN	0.898	0.102	0.806	0.183	0.011
	GBR	0.927	0.073	0.854	0.146	
	IBS	0.929	0.071	0.857	0.143	
	TSI	0.98	0.02	0.959	0.041	
American	CLM	0.867	0.133	0.75	0.233	0.017
	MXL	0.856	0.144	0.758	0.197	0.045
	PUR	0.918	0.082	0.836	0.164	
Sub Saharan	LWK	1		1		
	YRI	1		1		
African in the US	ASW	0.951	0.049	0.918	0.066	0.016
Asian	JPT	1		1		
	CHS	1		1		
	CHB	1		1		

[a]ASW: Americans of African ancestry in SW USA; CEU: Utah residents with Northern and Western European ancestry; CHB: Han Chinese in Beijing, China; CHS: Southern Han, Chinese; CLM: Colombian from Medellin; FIN: Finnish in Finland; GBR: British in England and Scotland; IBS: Iberian population in Spain; JPT: Japanese in Tokyo, Japan; LWK: Luhya in Webuye, Kenya; MXL: Mexican ancestry from Los Angeles USA; PUR: Puerto Ricans from Puerto Rico; TSI: Toscani in Italy; YRI: Yoruba in Ibadan, Nigeria.

Table 5. Non-synonymous and frameshift polymorphisms in the linkage block tagged by rs9503518

dbSNP rs No.	dbSNP allele	Protein residue	Amino acid position
rs200011775	A > G	Thr > Ala	579
rs148262614	T > C	Ser > Pro	587
rs201540017	G > A	Asp > Asn	632
rs199592727	T > C	Cys > Arg	638
rs141223516	C > A	Pro > His	644
rs201396552	A > G	Asn > Asp	666
rs150444727	A > G	Thr > Ala	670
rs112065429	A > G	Ser > Gly	671
rs34106449	Frameshift	-	634

Figure 1. Linkage of the four SNPs in the *SLC22A23* gene presented as D´ indicating the coefficient of the linkage disequilibrium for (panel A) 162 Crohn´s disease and (panel B) 149 ulcerative colitis patients.