Basic Study

Genetic association analysis of CLEC5A and CLEC7A gene single-nucleotide polymorphisms and Crohn's disease

Nagi Elleisy, Sarah Rohde, Astrid Huth, Nicole Gittel, Änne Glass, Steffen Möller, Georg Lamprecht, Holger Schäffler, Robert Jaster

Abstract

BACKGROUND
Crohn’s disease (CD) is characterized by a multifactorial etiology and a significant impact of genetic traits. While NOD2 mutations represent well established risk factors of CD, the role of other genes is incompletely understood.

AIM
To challenge the hypothesis that single nucleotide polymorphisms (SNPs) in the genes CLEC5A and CLEC7A, two members of the C-type lectin domain family of pattern recognition receptors, may be associated with CD.

METHODS
SNPs in CLEC5A, CLEC7A and the known CD risk gene NOD2 were studied using real time PCR-based SNP assays. Therefore, DNA samples from 175 patients and 157 healthy donors were employed. Genotyping data were correlated with clinical characteristics of the patients and the results of gene expression data analyses.

RESULTS
In accordance with previous studies, rs2066844 and rs2066847 in NOD2 were found to be significantly associated with CD (allelic \( P \) values = 0.0368 and 0.0474, respectively). Intriguingly, for genotype AA of rs1285933 in CLEC5A, a potential association with CD (recessive \( P \) = 0.0523; odds ratio = 1.90) was observed. There were no associations between CD and SNPs rs2078178 and rs16910631 in CLEC7A. Variants of rs1285933 had no impact on CLEC5A gene expression. In contrast, genotype-dependent differences of CXCL5 expression in peripheral blood mononuclear cells were observed. There is no statistical interaction...
between the tested SNPs of NOD2 and CLEC5A, suggesting of a novel pathway contributing to the disease.

CONCLUSION

Our data encourage enlarged follow-up studies to further address an association of SNP rs1285933 in CLEC5A with CD. The C-type lectin domain family member also deserves attention regarding a potential role in the pathophysiology of CD.

Key words: Crohn’s disease; Single nucleotide polymorphisms; NOD2; CLEC5A; Gene expression; CXCL5

©The Author(s) 2020. Published by Baishideng Publishing Group Inc. All rights reserved.

Core tip: The genetic traits of Crohn’s disease (CD) are incompletely understood. Here, we report a potential association of single nucleotide polymorphism (SNP) rs1285933 in CLEC5A, a member of the C-type lectin domain family of pattern recognition receptors, with CD. Variants of SNP rs1285933 had no impact on CLEC5A gene expression in peripheral blood mononuclear cells but correlated with the expression of CXCL5. The SNPs rs2078178 and rs16910631 in CLEC7A were not associated with the disease. The role of CLEC5A in the pathophysiology of CD deserves further attention.

INTRODUCTION

Together with ulcerative colitis, Crohn’s disease (CD) represents the most common and clinically relevant inflammatory bowel disease (IBD) [22]. While it is generally accepted that the pathogenesis of the disease is multifactorial and involves an inappropriate activation of the mucosal immune system, the precise contribution of individual environmental factors and genetic traits remains elusive [23,24]. Mutations in the NOD2 gene represent the best-characterized genetic association of CD [24-26]. Nucleotide-binding oligomerization domain 2 (NOD2) belongs to the pattern recognition receptor (PRR) family and acts as an intracellular sensor for peptidoglycan [27,28] and its fragment muramyl dipeptide [27,28]. Downstream of NOD2, the transcription factor NF-κB plays a key role in the transduction of receptor-generated signals [29].

C-type lectin domain (CLEC) receptors comprise a large family of carbohydrate-binding proteins [30,31]. Various CLEC family receptors are considered to exert functions as PRR since they recognize pathogen-associated molecules and may induce intracellular signaling pathways that regulate inflammatory processes. CLEC proteins are crucially involved in the immune response to fungal pathogens, but have also been implicated in anti-bacterial, anti-viral and anti-parasitic defense mechanisms [23-25,31,32]. Despite their functional similarities to NOD2, CLEC proteins have not been systematically studied in the context of IBD yet. Interestingly, a single nucleotide polymorphism (SNP) in the CLEC7A (DECTIN-1) gene, rs2078178, has been reported to be strongly linked to a severe form of ulcerative colitis, and this association was even stronger for the two-marker haplotype rs2078178 to rs16910631 [31]. For another CLEC gene, CLEC5A, we recently observed a CD-associated expression pattern with higher transcript levels in patient-derived peripheral blood mononuclear cells than in corresponding controls. Furthermore, CLEC5A showed a NOD2-dependent expression profile, supporting the hypothesis that both proteins may act in a regulatory network with a pathophysiological role in CD [32]. Given that defective bacterial clearance may contribute to the pathogenesis of CD [32-34], it is important to note that CLEC5A has also been suggested to be essentially involved in innate immunity through neutrophil trap formation and secretion of different proinflammatory cytokines after stimulation with Listeria monocytogenes [35]. Interestingly, the SNP rs1285933 in CLEC5A is associated with dengue severity [33,34], and
CLEC5A has been shown to be critical for dengue-virus-induced lethal disease[21].

Here, we have addressed the question if the SNPs rs2078178 and rs16910631 in CLEC7A and rs1285933 in CLEC5A are associated with CD and have analyzed effects of rs1285933 at the level of gene expression. For comparison and a positive control, the known disease-associated SNPs rs2066844 (SNP8), rs2066845 (SNP12) and rs2066847 (SNP13)[5,6] in NOD2 were included into the investigations as well.

MATERIALS AND METHODS

Patients
From October 2015 until June 2017, 175 patients (102 females and 73 males; mean age 43.1 ± 14.7 years) with CD from the Department of Gastroenterology of Rostock University Medical Center (Rostock, Germany) were included in the study. This cohort of CD patients represents an extension of a cohort that we have previously characterized regarding relationships between mutations in the NOD2 gene, the disease phenotype and anti-tumor necrosis factor-α trough levels[22].

The diagnosis of CD was based on clinical, endoscopic, histological and radiological findings of the patients. The following clinical data were collected: Age, sex, age at diagnosis, duration of the disease, disease location, disease behavior, disease activity (assessed by the Crohn’s disease activity index[23] and the Harvey–Bradshaw index[24]), disease-specific medications, and previous history of surgery (i.e., colectomy). CD was stratified via the Montreal classification[25]. Unrelated and healthy subjects from Germany (n = 157; 101 females and 56 males; mean age 25.3 ± 5.7 years) served as controls. The study was approved by the Local Ethics Board of the University of Rostock (A-2017-0137). We obtained written informed consent from all participants prior to their enrollment.

DNA extraction
EDTA whole-blood samples were subjected to DNA extraction employing the QIAamp DNA blood mini kit according to the instructions of the manufacturer (Qiagen, Hilden, Germany).

Genotyping
Genotyping was performed using TaqMan™ SNP Genotyping Allelic Discrimination Assays with VIC- and FAM-labeled probes (Thermo Fisher Scientific, Karlsruhe, Germany) for rs1285933 (CLEC5A, Assay-ID: C__9506735_10), rs2078178 (CLEC7A; Assay-ID: C__1932439_10), rs16910631 (CLEC7A; Assay-ID: C__33748498_10), rs2066844 (NOD2, SNP8, Assay-ID: C__11717468_20), rs2066845 (NOD2; SNP12, Assay-ID: C__11717466_20), and rs2066847 (NOD2, Assay-ID: SNP13 C__60383785_10). PCR was carried out in 96-well plates, employing a ViiA 7 sequence detection system (Thermo Fisher Scientific). Thermal cycling conditions were: 95 °C for 10 min, followed by 40 cycles of 15 s at 95 °C/1 min at 60 °C. After PCR, fluorescence was detected and analyzed using TaqManGenotyper software version 1.3. Alternatively, NOD2 genotypes were determined by Sanger sequencing as described before[22].

In vitro studies with peripheral blood mononuclear cells
In this study, previous data from our laboratory were re-evaluated with respect to the rs1285933 genotype[5]. Briefly, peripheral blood mononuclear cells (PBMC) had been isolated from EDTA venous blood, cultured and treated with lipopolysaccharide (1 µg/mL; Sigma-Aldrich, Deisenhofen, Germany) for 6 h. Afterwards, RNA was isolated, reversely transcribed into cDNA and subjected to real-time PCR employing standard procedures and a ViiA 7 sequence detection system. The following human-specific TaqMan™ gene expression assays with fluorescently labeled MGB probes were used to quantify target cDNA levels: Hs04398399_m1 (CLEC5A), Hs01099660_g1 (CXCL5), and Hs99999905_m1 (GAPDH). PCR conditions were as follows: 95 °C for 10 min, followed by 40 cycles of 15 s at 95 °C/1 min at 60 °C.

Statistical analysis
The data were stored and analyzed employing IBM SPSS Statistics 25.0 (International Business Machines Corporation, Armonk, New York, United States). Differences between patients and controls were assessed for distributions (genotype, allele and sex) using the χ² test or Fisher’s exact test, and for means using the t-test for independent samples (age, gene expression data), respectively. Pairwise statistical interaction between SNPs in a linear model was studied employing ANOVA. The Hardy-Weinberg equilibrium was assessed using the χ² test with 1 degree of freedom. False discovery rates were controlled by using the Benjamini-Hochberg correction.
Taiwanese children, neither rs1285933 nor other polymorphisms of CLEC5A were associated with SNP rs1285933 have not been reported yet. In a population of Other disease interactions between the dengue virus and CLEC5A receptors [20] SNP rs1285933 has also been suggested to modulate signaling pathways after the ability of ligand binding or downstream signaling. In accordance with this conclusion, e.g., with respect to its data suggest that CLEC5A might be functionally affected, itself. These CLEC5A in PBMC, but not of CXCL5, the expression of the chemokine trans we can report a reinvestigating data from our past work effect of rs1285933 on [16] CLEC5A polymorphism need to be further elucidated. To this end, effect of the beyond NOD2 into the pathogenesis of the disease. The mechanisms that underlie the potential association of SNP rs1285933 with CD. However, our findings need to be interpreted cautiously since they are based on a relatively small number of patients. To study associations of CD with SNP genotypes or allele frequencies, four genetic models (genotype, dominant, recessive, or allelic models) were employed (Table 1). As expected, significant associations with CD were found for SNPs in NOD2, specifically rs2066844 (SNP8; genotype P = 0.0498, dominant P value = 0.0219, allelic P value = 0.0368) and rs2066847 (SNP13; allelic P value = 0.0474). Intriguingly, the genotype AA of rs1285933 in CLEC5A was also potentially associated with the disease (recessive model; P = 0.0523). The corresponding odds ratios (ORs) are shown in Table 2. For NOD2, the odds of having CD might triple in the presence of the risk allele T (rs2066844; OR = 3.29), and double with allele CC (rs2066847; OR = 2.31). Increased ORs are detectable for CLEC5A, too. Genotype AA almost doubles the odds of CD (OR = 1.90). Carrying the risk allele A increases the odds of CD by 39% (OR = 1.39), whereas allele G displays a protective effect (OR = 0.72). We could not detect significant associations between CD and the two SNPs in CLEC7A (rs2078178, rs16910631) and also not for rs2066845 (SNP12) in NOD2 (Table 1). The latter finding might be explained by the rare occurrence of the risk allele C in our cohorts of small size.

We next compared patients with different genotypes of rs1285933 in CLEC5A (AA, AG and GG, respectively) regarding their clinical characteristics, employing the following parameters: Age, age at diagnosis, duration of the disease, disease location and behavior according to Montreal classification, Crohn’s disease activity index and Harvey–Bradshaw index, history of surgical treatment and treatment with drugs (including antibodies such as tumor necrosis factor-α inhibitors). There were no statistically significant differences between the three genotypes (data not shown).

To study potential functional effects of the rs1285933 polymorphism, we re-evaluated previously published gene expression data from our laboratory. In these studies, PBMC from CD patients and controls had been employed to measure the mRNA expression of a pre-selected set of genes[24]. Using a combined data set from 16 CD patients and 6 healthy controls, we observed no genotype-dependent differences of CLEC5A gene expression (Figure 1A). On the other hand, we found that the genotype GG, compared to AG, was associated with significantly lower mRNA levels of the proinflammatory chemokine CXCL5 (Figure 1B; please note that ΔCt values and expression levels show an inverse and logarithmic relationship that follows the function ΔCt = −log2(1/expressed level). Located on different chromosomes, the disease-associate SNP of CLEC5A is not correlating with disease-associated SNPs of NOD2 (data not shown). Furthermore, the pairwise contributions to the disease phenotype of the CLEC5A SNP and the other SNPs are independent from each other (Table 3).

DISCUSSION
To the best of our knowledge, the results of this study suggest for the first time a potential association of SNP rs1285933 with CD. However, our findings need to be interpreted cautiously since they are based on a relatively small number of patients from a single center.

Given that the SNP is located within the CLEC5A gene, our data implicate a PRR beyond NOD2 into the pathogenesis of the disease. The mechanisms that underlie the effect of the CLEC5A polymorphism need to be further elucidated. To this end, reinvestigating data from our past work[24] we can report a trans effect of rs1285933 on the expression of the chemokine CXCL5 in PBMC, but not of CLEC5A itself. These data suggest that CLEC5A might be functionally affected, e.g., with respect to its ability of ligand binding or downstream signaling. In accordance with this conclusion, SNP rs1285933 has also been suggested to modulate signaling pathways after interactions between the dengue virus and CLEC5A receptors[20]. Other disease associations of SNP rs1285933 have not been reported yet. In a population of Taiwanese children, neither rs1285933 nor other polymorphisms of CLEC5A were
Table 1  Genotype and allele frequencies of single nucleotide polymorphisms in the genes CLEC5A, CLEC7A and NOD2 in Crohn’s disease patients and controls

| Gene  | SNP       | Genotype | Cases (n = 175) | Controls (n = 157) | Allele | Cases (alleles) | Controls (alleles) | Genotype P value | Dominant P value | Recessive P value | Allelic P value |
|-------|-----------|----------|----------------|--------------------|--------|----------------|--------------------|-----------------|-----------------|-----------------|----------------|
|       |           |          |                |                    |        |                |                    |                 |                 |                 |                 |
|       |           |          |                |                    |        |                |                    |                 |                 |                 |                 |
| CLEC5A | rs1285933 | GG       | 35             | 36                 | G, A   | 144, 206       | 155, 159           | 0.1093          | (0.0285)        | 0.9727          | (0.5921)        |
|        |           | GA       | 74             | 83                 |        |                |                    |                 |                 | 0.0523          | (0.0091)        |
|        |           | AA       | 66             | 38                 |        |                |                    |                 |                 | 0.0900          | (0.0352)        |
|       |           |          |                |                    |        |                |                    |                 |                 |                 |                 |
| CLEC7A | rs2078178 | GG       | 104            | 100                | G, A   | 274, 76        | 251, 63            | 1.0000          | (0.5713)        | 0.9033          | (0.4320)        |
|        |           | AG       | 66             | 51                 |        |                |                    |                 |                 | 0.8344          | (0.7618)        |
|        |           | AA       | 5              | 6                  |        |                |                    |                 |                 | 0.9107          | (0.6335)        |
|       |           |          |                |                    |        |                |                    |                 |                 |                 |                 |
| CLEC7A | rs16910631| CC       | 153            | 139                | C, T   | 327, 23        | 294, 20            | 0.8078          | (0.7024)        | 0.9056          | (0.8662)        |
|        |           | CT       | 21             | 16                 |        |                |                    |                 |                 | 0.9269          | (0.6045)        |
|        |           | TT       | 1              | 2                  |        |                |                    |                 |                 | 1.0000          | (1.0000)        |
|       |           |          |                |                    |        |                |                    |                 |                 |                 |                 |
| NOD2   | rs2066844 | CC       | 146            | 148                | C, T   | 319, 31        | 305, 9             | 0.0498          | (0.0065)        | 0.0219          | (0.0019)        |
|        |           | CT       | 27             | 9                  |        |                |                    |                 |                 | 0.9583          | (0.5000)        |
|        |           | TT       | 2              | 0                  |        |                |                    |                 |                 | 0.0368          | (0.0016)        |
|       |           |          |                |                    |        |                |                    |                 |                 |                 |                 |
| NOD2   | rs2066845 | CC       | 163            | 149                | G, C   | 336, 12        | 306, 8             | 0.8481          | (0.6453)        | 0.8481          | (0.6453)        |
|        |           | GC       | 12             | 8                  |        |                |                    |                 |                 | NA              |                |
|        |           | CC       | 0              | 0                  |        |                |                    |                 |                 | 0.7874          | (0.6505)        |
|       |           |          |                |                    |        |                |                    |                 |                 |                 |                 |
| NOD2   | rs2066847 | C-C      | 147            | 143                | C, CC  | 316, 34        | 300, 14            | 0.0923          | (0.0321)        | 0.1569          | (0.0682)        |
|        |           | C-CC     | 22             | 14                 |        |                |                    |                 |                 | 0.1025          | (0.0312)        |
|        |           | CC-CC    | 6              | 0                  |        |                |                    |                 |                 | 0.0474          | (0.0103)        |

1Italic: Minor allele according to database https://www.ncbi.nlm.nih.gov/snp/.

2Numbers in brackets refer to the P value prior to Benjamini-Hochberg correction (23 tests); significant differences (P < 0.05) are indicated in bold.

3Refers to the minor allele. SNP: Single nucleotide polymorphism; NA: Not applicable (due to the absence of CC genotype).

associated with susceptibility to Kawasaki disease, coronary artery lesion formation, and intravenous immunoglobulin treatment response[26].

Of note, CLEC5A is embedded into an intronic region of another gene, MGAM, and the two transcripts are known to correlate[27], so that the effect of SNP rs1285933 is not necessarily exclusively related to the C-type lectin domain family member. Interestingly, decreased maltase activities in the small bowel mucosa are common in children with CD[28], and although this is of course no evidence for a genetic association, the role of MGAM in the context of IBD may deserve further attention as well.

We also evaluated possible associations of SNP rs1285933 with different clinical characteristics of our CD patients, including disease location, disease behavior and treatment history, but did not obtain significant results. Given that such effects have been reported for NOD2 variants[29-32], the studies are nevertheless worth to be continued in larger cohorts of patients. To this end, we conclude that the principal effect of SNP rs1285933 is modulation of CD susceptibility through a different molecular pathway than NOD2.

PRRs are key regulators of innate immune responses and inflammatory processes[13,14]. For a prominent member of this family, NOD2, a role in the pathogenesis of CD is clearly established[14]. Our results suggest an association of a polymorphism in another PRR, rs1285933 in CLEC5A, but not of rs2078178 and rs16910631 in CLEC7A, with CD. A systematic analysis of PRR functions in the context of CD might reveal novel pathomechanistic insights and help to identify new targets for diagnostic and therapy.
### Table 2 Odds ratios of genotypes and alleles of single nucleotide polymorphisms in the genes CLEC5A and NOD2

| Gene | SNP    | Genotype/allele | Odds ratio | 95%CI  | P value  |
|------|--------|-----------------|------------|--------|----------|
| CLEC5A | rs1285933 | AA              | 1.90       | 1.18-3.05 | 0.009    |
|       |         | GG              | 0.84       | 0.50-1.42 | 0.516    |
|       |         | AG              | 0.65       | 0.42-1.01 | 0.054    |
|       |         | A               | 1.39       | 1.03-1.90 | 0.034    |
|       |         | G               | 0.72       | 0.53-0.97 | 0.034    |
| NOD2  | rs2066844 (SNP8) | TT              | NA         |        |          |
|       |         | CC              | 0.31       | 0.14-0.67 | 0.003    |
|       |         | CT              | 3.00       | 1.36-6.60 | 0.006    |
|       |         | T               | 3.29       | 1.54-7.03 | 0.002    |
|       |         | C               | 0.30       | 0.14-0.65 | 0.002    |
| NOD2  | rs2066847 (SNP13) | CC-CC           | NA         |        |          |
|       |         | C-C             | 0.51       | 0.26-1.02 | 0.056    |
|       |         | C-CC            | 1.47       | 0.72-2.98 | 0.287    |
|       |         | C               | 0.43       | 0.23-0.82 | 0.011    |
|       |         | CC              | 2.31       | 1.21-4.38 | 0.011    |

1Unadjusted for multiple testing. SNP: Single nucleotide polymorphism; NA: Not applicable (missing in controls); CI: Confidence interval.

### Table 3 Pairwise statistical interaction between single nucleotide polymorphisms in a linear model

| SNP | CLEC5A | CLEC7A | NOD2 |
|-----|--------|--------|------|
|     | rs1285933 | rs2078178 | rs16910631 | rs2066844 | rs2066845 | rs2066847 |
| rs1285933 | NA      | 0.6490  | 0.7409  | 0.5266  | 0.6875  | 0.2813  |
| rs2078178 | 0.6490  | NA      | 0.1036  | 0.8573  | 0.4040  | 0.3718  |
| rs16910631 | 0.7409  | 0.1036  | NA      | 0.8980  | 0.6698  | 0.9270  |
| rs2066844 | 0.5266  | 0.8573  | 0.8980  | NA      | 2.8246e-07 | 0.9664  |
| rs2066845 | 0.6875  | 0.4040  | 0.6698  | 2.8246e-07 | NA      | 0.7399  |
| rs2066847 | 0.2813  | 0.3718  | 0.9270  | 0.9664  | 0.7399  | NA      |

1Disease—single nucleotide polymorphism (SNP) A + SNP B + SNP A: SNP B. The uncorrected P values for the last term of an ANOVA are presented in the table for all 30 interactions. The only significance was for two chromosomally neighboring SNPs within NOD2. SNP: Single nucleotide polymorphism; NA: Not applicable.
Figure 1 Effects of the rs1285933 genotype on CLEC5A and CXCL5 gene expression. Peripheral blood mononuclear cells were isolated from individuals with genotype AA (n = 8), GG (n = 5), and AG (n = 9), cultured and treated with lipopolysaccharide (1 µg/mL) for 6 h. Subsequently, the mRNA expression of the indicated genes and the house-keeping control GAPDH was analyzed by real-time PCR. Data are presented as averaged ΔCt values ± standard error of mean. *P < 0.05 vs genotype GG.

ARTICLE HIGHLIGHTS

Research background
Crohn’s disease (CD) is characterized by a multifactorial etiology and a significant impact of genetic traits. While NOD2 mutations represent well established risk factors of CD, the role of other genes is incompletely understood.

Research motivation
A better knowledge of the molecular basis of CD is considered as an essential prerequisite for a further improvement of diagnostics and therapy.

Research objectives
Previous studies from our laboratory have pointed to a possible link between CD and the expression of pattern recognition receptors of the C-type lectin domain family (specifically, CLEC5A) in peripheral blood mononuclear cells (PBMC). This observation prompted us to ask if single nucleotide polymorphisms in the genes CLEC5A and CLEC7A might be associated with the disease.

Research methods
DNA samples from patients with CD and healthy donors were subjected to the analysis of single nucleotide polymorphisms in the genes CLEC5A, CLEC7A and NOD2. For studies on gene expression, PBMC from subgroups of both cohorts were employed. Molecular findings were correlated with clinical characteristics of the patients.

Research results
For genotype AA of rs1285933 in CLEC5A, a potential association with CD and an increased odds ratio were detected. As expected, risk variants of NOD2 were associated with an increased occurrence of CD as well. Polymorphisms of rs1285933 correlated with CXCL5 gene expression but had no effect on CLEC5A expression in PBMC.

Research conclusions
SNP rs1285933 in CLEC5A may represent a novel genetic association of CD. The finding, however, needs to be reproduced in multicenter studies with larger numbers of CD patients.

Research perspectives
Pattern recognition receptors of the C-type lectin domain family deserve further attention...
regarding their potential role in the pathogenesis of CD and their relevance as diagnostic markers and therapeutic targets.

ACKNOWLEDGEMENTS

We thank Mrs. Katja Bergmann for expert technical assistance.

REFERENCES

1. Abraham C, Cho JH. Inflammatory bowel disease. *N Engl J Med* 2009; 361: 2066-2078 [PMID: 19923578 DOI: 10.1056/NEJMra0804647]  
2. Baumgart DC, Sandborn WJ. Crohn's disease. *Lancet* 2012; 380: 1590-1605 [PMID: 22914295 DOI: 10.1016/S0140-6736(12)60206-9]  
3. Mayer L. Evolving paradigms in the pathogenesis of IBD. *J Gastroenterol* 2010; 45: 9-16 [PMID: 19960355 DOI: 10.1007/s00535-009-0138-3]  
4. Naser SA, Arce M, Khaja A, Fernandez M, Naser N, Elwasila S, Thanigachalam S. Role of ATG16L, NOD2 and IL23R in Crohn's disease pathogenesis. *World J Gastroenterol* 2012; 18: 412-424 [PMID: 22342317 DOI: 10.3748/wjg.v18.i5.412]  
5. Hugot JP, Chamaillard M, Zouali H, Lesage S, Cérard JP, Belaiche J, Almer S, Tysk C, O'Morain CA, Gassull M, Binder V, Finkel Y, Cortot A, Modigliani R, Laurent-Puig P, Gower-Rousseau C, Macry J, Colombel JF, Sahaboui M, Thomas G. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 2001; 411: 599-603 [PMID: 11172600 DOI: 10.1038/35079107]  
6. Ogura Y, Bonen DK, Inohara N, Nicolae DL, Chen FF, Ramos R, Britton H, Moran T, Karalisukas R, Duerr RH, Ackar JP, Brant SR, Bayless TM, Kirschner BS, Hanauer SB, Nüeze G, Cho JH. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* 2001; 411: 603-606 [PMID: 11385577 DOI: 10.1038/35079114]  
7. Natsuka M, Uehara A, Yang S, Echigo S, Takada H. A polypeptide-water-soluble peptidoglycan exhibited both Toll-like receptor 2- and NOD2-agonistic activities, resulting in synergistic activation of human monocyte cells. *Innate Immun* 2008; 14: 298-308 [PMID: 18809654 DOI: 10.1177/175345908096518]  
8. Iyer JK, Coggeshall KM. Cutting edge: primary innate immune cells respond efficiently to polymeric peptidoglycan, but not to peptidoglycan monomers. *J Immunol* 2011; 186: 3841-3845 [PMID: 21357534 DOI: 10.4049/jimmunol.100458]  
9. Girardin SE, Boneca IG, Viala J, Chamaillard M, Labigne A, Thomas G, Philpott DJ, Sansonetti PJ. Nod2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection. *J Biol Chem* 2003; 278: 8869-8872 [PMID: 12527755 DOI: 10.1074/jbc.C200651200]  
10. Inohara N, Ogura Y, Fontalba A, Gutierrez O, Pons F, Cespo J, Fukase K, Inamura S, Kinumoto S, Hashimoto M, Foster SJ, Moran AP, AM, Fernandez-Luna JL, Nüeze G. Host recognition of bacterial muramyl dipeptide mediated through NOD2. Implications for Crohn's disease. *J Biol Chem* 2003; 278: 5509-5512 [PMID: 12514169 DOI: 10.1074/jbc.C0673200]  
11. Strober W, Murray PJ, Kitani A, Watanabe T. Signalling pathways and molecular interactions of NOD1 and NOD2. *Nat Rev Immunol* 2006; 6: 9-20 [PMID: 16493424 DOI: 10.1038/nri1747]  
12. Zelenksy AN, Greedy JE. The C-type lectin-like domain superfamily. *FEBS J* 2005; 272: 6179-6217 [PMID: 16336259 DOI: 10.1111/j.1742-4658.2005.05031.x]  
13. Hardison SE, Brown GD. C-type lectin receptors orchestrate antifungal immunity. *Nat Immunol* 2012; 13: 817-822 [PMID: 22910394 DOI: 10.1038/nri3269]  
14. Hoving JC, Boneca IG, Viala J, Chamaillard M, Labigne A, Thomas G, Philpott DJ. NOD2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection. *J Biol Chem* 2003; 278: 8869-8872 [PMID: 12527755 DOI: 10.1074/jbc.C200651200]  
15. Ilies ID, Funari VA, Taylor KD, Nguyen Q, Reyes CN, Strom SP, Brown J, Becker CA, Flesher PR, Dubinsky M, Rotter JI, Wang HL, McGovern DP, Brown GD, Underhill DM. C-type lectin receptors orchestrate antifungal immunity. *Nat Immunol* 2012; 13: 817-822 [PMID: 22910394 DOI: 10.1038/nri3269]  
16. Schäffler H, Rohde M, Rohde S, Huth A, Gittel N, Hollborn H, Koczan D, Glass Â, Lamprecht G, Jaster R. NOD2- and disease-specific gene expression profiles of peripheral blood mononuclear cells from Crohn's disease patients. *World J Gastroenterol* 2018; 24: 1196-1205 [PMID: 29568200 DOI: 10.3748/wjg.v24.i11.1196]  
17. Cooney R, Baker J, Brain O, Danis B, Pichulik T, Allan P, Ferguson DJ, Campbell BJ, Jewell D, Simmens A. NOD2 stimulation induces autophagy in dendritic cells influencing bacterial handling and antigen presentation. *Net Med* 2010; 16: 90-97 [PMID: 19996812 DOI: 10.1038/nmm.2099]  
18. Travassos LH, Carneiro LA, Ranjeeta M, Hussey S, Kim YG, Magalhães JG, Yuan L, Soares F, Chea E, Lebourhis I, Boneca IG, Alloue A, Jones NL, Nüeze G, Girardin SE, Philpott DJ, Nod1 and Nod2 direct autophagy by recruiting ATG16L1 to the plasma membrane at the site of bacterial entry. *Nature Immunol* 2010; 11: 55-62 [PMID: 19898471 DOI: 10.1038/ni.1823]  
19. Chen SY, Li FJ, Huu TY, Liang SM, Yeh YC, Liao WW, Chou TY, Chen NJ, Hsiao M, Yang WB, Hsieh SL. CLEC5A is a critical receptor in innate immunity against Listeria infection. *Nat Commun* 2017; 8: 299 [PMID: 28824166 DOI: 10.1038/s41467-017-0055-3]  
20. Xavier-Carvalho C, Cezar RDDS, Freire NM, Vasconcelos CMM, Solorzano VEF, de Toledo-Pinto TG, Fialho LG, do Carmo RF, Vasconcelos LRS, Cordeiro MT, Baptista P, de Azeredo EL, da Cunha RV, de Souza LJ, Pacheco AG, Kubelka CF, Moura PMMF, Moraes MO. Association of rs1283933 single nucleotide polymorphism in CLEC5A gene with dengue severity and its functional effects. *Hum Immunol* 2017; 78: 649-656 [PMID: 27964023 DOI: 10.1016/j.humimm.2017.07.013]  
21. Chen SY, Lin YL, Huang MT, Wu MF, Cheng SC, Lee HY, Lee CK, Chiuow TW, Wong CH, Hsieh SL. CLEC5A is a critical for dengue-virus-induced lethal disease. *Nature* 2008; 453: 672-676 [PMID: 18496526 DOI: 10.1038/nature07013]  
22. Schäffler H, Geiss D, Gittel N, Rohde S, Huth A, Glass Â, Brandhorst G, Jaster R, Lamprecht G. Mutations in the NOD2 gene are associated with a specific phenotype and lower anti-tumor necrosis factor
trough levels in Crohn's disease. *J Dig Dis* 2018; 19: 678-684 [PMID: 30284387 DOI: 10.1111/1751-2980.12677]

23 **Best WR**, Becktel JM, Singleton JW, Kern F. Development of a Crohn's disease activity index. National Cooperative Crohn's Disease Study. *Gastroenterology* 1976; 70: 439-444 [PMID: 1524701 DOI: 10.1016/S0016-5085(76)80163-1]

24 **Harvey RF**, Bradshaw JM. A simple index of Crohn's-disease activity. *Lancet* 1980; I: 514 [PMID: 6102236 DOI: 10.1016/0140-6736(80)92767-1]

25 **Silverberg MS**, Satasangi J, Ahmad T, Arnott ID, Bernstein CN, Brant SR, Caprilli R, Colombel JF, Gasche C, Geboes K, Jewell DP, Karban A, Loftus EV, Peña AS, Riddell RH, Sachar DB, Scheiber S, Steinhart AH, Targan SR, Vermeire S, Warren BF. Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. *Can J Gastroenterol* 2005; 19 Suppl A: 5A-36A [PMID: 16151544 DOI: 10.1155/2005/269076]

26 **Yang YL**, Chang WP, Hsu YW, Chen WC, Yu HR, Liang CD, Tzai YT, Huang YH, Yang KD, Kuo HC, Chang WC. Lack of association between CLEC5A gene single-nucleotide polymorphisms and Kawasaki disease in Taiwanese children. *J Biomed Biotechnol* 2012; 2012: 398628 [PMID: 22536019 DOI: 10.1155/2012/398628]

27 **Adler P**, Kolde R, Kull M, Tkachenko A, Peterson H, Reimand J, Vilo J. Mining for coexpression across hundreds of datasets using novel rank aggregation and visualization methods. *Genome Biol* 2009; 10: R139 [PMID: 19961599 DOI: 10.1186/gb-2009-10-12-r139]

28 **Wiecek S**, Wos H, Radziejewicz Winnicki I, Komraus M, Grzybowska Chlebowczyk U. Disaccharidase activity in children with inflammatory bowel disease. *Turk J Gastroenterol* 2014; 25: 185-191 [PMID: 25903060 DOI: 10.5152/tjg.2014.3994]

29 **Böning C**, Gensche J, Bührer S, Krüger S, Kling K, Dignass A, Baier P, Bochow B, Ockenga J, Schmidt HH, Lochs H. Mutations in the NOD2/CARD15 gene in Crohn's disease are associated with ileocecal resection and are a risk factor for reoperation. *Aliment Pharmacol Ther* 2004; 19: 1073-1078 [PMID: 15142196 DOI: 10.1111/j.1365-2036.2004.01967.x]

30 **Schmitzer F**, Friedrich M, Wolf C, Stallhofer J, Angelberger M, Diegelmann J, Olszak T, Tillack C, Beigel F, Göke B, Glas J, Lohse P, Brand S. The NOD2 Single Nucleotide Polymorphism rs7297653 (IVS4+10 A>C) Is a Predictor for Perianal Fistulas in Patients with Crohn's Disease in the Absence of Other NOD2 Mutations. *PLoS One* 2015; 10: e0116044 [PMID: 26147989 DOI: 10.1371/journal.pone.0116044]

31 **Alvarez-Lobos M**, Arostegui JI, Sans M, Tassies D, Plaza S, Delgado S, Lacy AM, Pique JM, Yagüe J, Panés J. Crohn's disease patients carrying Nod2/CARD15 gene variants have an increased and early need for first surgery due to strictureing disease and higher rate of surgical recurrence. *Ann Surg* 2005; 242: 693-700 [PMID: 16244543 DOI: 10.1097/01.sla.0000186173.14696.ca]

32 **Cantó E**, Ricart E, Busquets D, Monfort D, García-Planella E, González D, Balanzó J, Rodriguez-Sanchez JL, Vidal S. Influence of a nucleotide oligomerization domain 1 (NOD1) polymorphism and NOD2 mutant alleles on Crohn's disease phenotype. *World J Gastroenterol* 2007; 13: 5446-5453 [PMID: 17907287 DOI: 10.3748/wjg.v13.i41.5446]
