Published in final edited form as:

Gene. 2013 September 25; 527(2): 496–502. doi:10.1016/j.gene.2013.06.066.

Evidence of expression variation and allelic imbalance in Crohn's disease susceptibility genes *NOD2* and *ATG16L1* in human dendritic cells

Jianzhong Hu* and Inga Peter¹

Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, 1 Gustave L Levy Place, New York, NY, USA

Inga Peter: inga.peter@mssm.edu

Abstract

Human dendritic cells (DCs) play an important role in induction and progression of Crohn's disease (CD). Accumulating evidence suggests that viral infection is required to trigger CD pathogenesis in genetically predisposed individuals. *NOD2* and *ATG16L1* are among the major CD susceptibility genes implicated in impaired immune response to bacterial infection. In this study, we investigated gene expression and allelic imbalance (AI) of *NOD2* and *ATG16L1* using common variants in human monocyte-derived DCs. Significant AI was observed in ~40% and ~70% of *NOD2* and *ATG16L1* heterozygotes, respectively ($p < 0.05$). AI of *NOD2* was inversely associated with its expression level ($p = 0.015$). No correlation was detected between gene expression and AI for *ATG16L1*. When infected with Newcastle Disease Virus (NDV), *NOD2* expression in DCs was induced about four-fold ($p < 0.001$), whereas *ATG16L1* expression was not affected ($p = 0.88$). In addition, NDV infection tended to lower the variance in AI among DC populations for the *NOD2* gene ($p = 0.05$), but not the *ATG16L1* gene ($p = 0.32$). Findings of a simulation study, aimed to verify whether the observed variation in gene expression and AI is a result of sample-to-sample variability or experimental measurement error, suggested that *NOD2* AI is likely to result from a deterministic event at a single cell level. Overall, our results present initial evidence that AI of the *NOD2* and *ATG16L1* genes exists in populations of human DCs. In addition, our findings suggest that viral infection may regulate *NOD2* expression.

Keywords

Allelic imbalance; *NOD2*; *ATG16L1*; Human dendritic cells; NDV; Crohn's disease

© 2013 Elsevier B.V. All rights reserved.

*Corresponding author. Tel.: +1 2126596881; fax: +1 2126595657. jianzhong.hu@mssm.edu (J. Hu).

¹Tel.: +1 2126596881; fax: +1 2126595657.

Authors' contributions

JHU conceived and designed the study, performed the experiments and statistical analysis, and wrote the manuscript. IP participated in study design and data interpretation and reviewed the manuscript. Both authors read and approved the final version of the manuscript.

Conflict of interest

The authors have no financial interests in the subject matter or materials discussed in this manuscript.

1. Introduction

Allelic imbalance (AI) in gene expression, which refers to the differential expression of alternative alleles of the same gene, was initially characterized for epigenetic events such as gene imprinting or X chromosome inactivation. However, recent studies have found convincing evidence that AI exists in autosomal non-imprinted genes and demonstrated that AI is a common phenomenon in humans as well as in other species (Guo et al., 2004; Lo et al., 2003; Pant et al., 2006; Parker-Katiraei et al., 2008; Yan et al., 2002b). Genome-wide analysis of AI has suggested that about 4.6% of heterozygous SNP-sample pairs show evidence of AI (Heap et al., 2009). The importance of AI has been highlighted by the fact that expression variation of polymorphic alleles was implicated in disease pathogenesis and thus may have functional significance (Meyer et al., 2008; Yan et al., 2002a).

Crohn's disease (CD) is an inflammatory bowel disease resulting from defects in the regulation of mucosal immune responses to enteric bacteria in genetically susceptible individuals (Abraham and Cho, 2009; Shanahan, 2002). A recent meta-analysis has implicated 163 susceptibility loci in CD risk (Jostins et al., 2012). Notably, three of the top genes, *NOD2*, *TNFSF15* and *IRGM*, have reportedly shown AI in various human tissues and cell lines (Kakuta et al., 2009; McCarroll et al., 2008; Prescott et al., 2010; Yan et al., 2002b). Moreover, expression of another gene predisposing to CD, *PTGER4*, has been found to be associated with an upstream *cis*-acting regulatory locus suggesting AI of *PTGER4* in heterozygotes for this locus (Libioulle et al., 2007). However, AI characteristics of many CD susceptibility genes in key immune cells, as well as the mechanism of AI involvement in CD pathogenesis, have not been fully investigated.

Dendritic cells (DCs) are primary antigen-presenting cells that bridge the innate and adaptive immune responses (Banchereau et al., 2000; Reis e Sousa, 2006). DCs are thought to play an essential role in induction and progression of CD (Coombes and Powrie, 2008; Cooney et al., 2009; Levy et al., 2007; Rescigno and Di Sabatino, 2009). Several common SNPs in the *NOD2* gene have been shown to have a substantial impact on the basal characteristics of DC response to bacteria invasion (Cooney et al., 2009; Salucci et al., 2008) and the recruitment of antibacterial TH17 cells (van Beelen et al., 2007). A recent study has shown that DCs from the CD patients carrying *NOD2* or *ATG16L1* risk variants were defective in autophagy induction, bacterial trafficking and antigen presentation (Cooney et al., 2009). However, no study has reportedly assessed the expression variations of the *NOD2* and *ATG16L1* genes in DCs.

Given accumulating evidence suggesting that viral infection in the presence of certain intestinal bacteria plays an important role in CD pathogenesis in genetically susceptible individuals (Cadwell et al., 2008), the goal of this study was to investigate gene expression and AI of *NOD2* and *ATG16L1* with regard to and regardless of viral exposure in the human DCs. This study can help shed light on a complex interplay between viral exposure and expression of the key CD susceptible genes in the immune system.

2. Materials and methods

2.1. DC samples

All human research protocols for this work have been approved by the IRB of the Mount Sinai School of Medicine. The IRB of the Mount Sinai School specifically waived the need for consent due to use of discarded samples not traceable to source. DCs were obtained from 96 anonymous human blood donors following a standard protocol (Hu et al., 2011). Briefly, human monocytes from buffy coats were isolated by Ficoll density gradient centrifugation (Histopaque, Sigma Aldrich) and CD14⁺ monocytes were immunomagnetically purified by using a MACS CD14 isolation kit (Miltenyi Biotech.). CD14⁺ monocytes (0.7×10^6 cells/ml) were later differentiated into immature DCs after 5–6 day incubation in growth media [RPMI Medium 1640 (Gibco), 10% fetal calf serum (Hyclone), 2 mM of L-glutamine, 100 units/ml penicillin, 100 µg/ml streptomycin (Pen/Strep) (Invitrogen), 500 U/ml hGM-CSF (Preprotech) and 1000 U/ml hIL-4 (Preprotech)] at 37 °C.

2.2. NDV infection

The recombinant Hitchner strain of Newcastle Disease Virus (rNDV/B1) was prepared and titered as previously described. Titered NDV stock was diluted in Dulbecco's Modified Eagle Medium (DMEM) and added directly into pelleted and resuspended DCs at a multiplicity of infection (MOI) of 0.5 as previously described²⁵. After incubation for 30 min at 37 °C, fresh DC growth medium was added back to the infected DCs (1.0×10^6 cells/ml). Virus free allantoic fluid was added to additional tubes of cells to serve as a negative control.

2.3. DNA and total RNA extraction

Genomic DNA was extracted from lymphocytes of the same donors with the DNeasy kit (Qiagen, CA). Total RNA was extracted from DCs with the RNeasy kit (Qiagen, CA) and converted to cDNA by the Ovation™ RNA Amplification System (Nugen, CA). The total RNA extracted from DCs for testing cell population selection bias was converted to cDNA by MMLV reverse transcriptase 1st-strand cDNA synthesis kit (Epicentre, WI). The quantity of cDNA was determined by Nanodrop.

2.4. Read-out SNP selection, genotyping and AI measurement

We used two criteria to select the readout SNPs to track the AI: 1) SNP should be common (minor allele frequency >30%) so that there will be a sufficient number of heterozygotes in our studied sample. 2) SNP should be directly implicated in disease risk or located at the 3' UTR region to detect mainly the full length cDNA. The three major coding *NOD2* CD risk variants (rs2066844, rs2066845 and rs2066847) have low frequency in the general population (Economou et al., 2004) and therefore could not be used as readout SNPs for AI studies. Using the HapMap database for individuals of European ancestry, we assessed linkage disequilibrium (LD) between the 8 common exonic *NOD2* SNPs (rs2066842, rs2076752, rs5743266, rs61736931, rs2066843, rs2067085, rs3135499 and rs3135500) and the risk SNPs (rs2066844 and rs2066845; rs2066847 is a frame-shift mutation that is absent in the HapMap). We did not find any common SNP in LD with the risk SNPs. Therefore, we

selected the most common 3'UTR variant, rs3135499 as a read-out SNP to detect only the full length cDNA. In contrast, exonic *ATG16L1* rs2241880 is common and has been reported to be associated with CD risk (Hampe et al., 2007; Rioux et al., 2007). Genotyping and AI measurements of *NOD2* rs3135499 and *ATG16L1* rs2241880 were performed by pyrosequencing. To test the robustness of our pyrosequencing approach we examine the AI of *NOD2* using a different readout SNP, rs3135500, in high linkage disequilibrium with our initially examined *NOD2* SNP rs3135499 ($D' = 1.0$, $r^2 = 0.93$), using a subset of 9 samples. Each SNP was tested for Hardy Weinberg equilibrium (HWE). The PCR primers and pyrosequencing primer were designed by PSQ Assay Design software (Qiagen, CA). The sequences are listed in Supplementary Table 1. Targeted genomic DNA or cDNA were PCR-amplified using PyroMark PCR master mix (Qiagen, CA). The recipe contains the reaction buffer [10 μ l 2 $\times$ PCR Master Mix, 2 μ l 10 $\times$ Loading Buffer, 0.8 μ l 25 mM Mg++, 2 μ l PCR primer mix (5 μ M), and 4.2 μ l Nuclease Free Water] and 1 μ l genomic DNA or cDNA as template. PCR was performed at 95 $^{\circ}$ C for 15 min, 45 cycles of [94 $^{\circ}$ C for 30 s, 60 $^{\circ}$ C for 30 s, and 72 $^{\circ}$ C for 30 s], and a final extension at 72 $^{\circ}$ C for 10 min. PCR amplicons were checked for purity by electrophoresis in 2.0% agarose. Pyrosequencing was performed with a PyroMark24 system following the manufacturer's instruction. The detailed protocols were designed in PyroMark24 software provided by the manufacturer.

2.5. Gene expression

The expression of *NOD2* and *ATG16L1* genes in DCs was measured by quantitative real-time PCR (the sequences of PCR primers can be provided under request). Targeted cDNAs were amplified by PCR using hot start Taq polymerase (Qiagen, CA) with the optimized PCR buffer (recipe can be provided under request) and 10 ng cDNA template. The real-time PCR was performed at 95 $^{\circ}$ C for 10 min, 50 cycles of [94 $^{\circ}$ C for 10 s, 60 $^{\circ}$ C for 30 s, and 72 $^{\circ}$ C for 30 s] with a single detection at 72 $^{\circ}$ C at each cycle, and followed by melting curve analysis. Serial dilutions of genomic DNA were used to construct the standard PCR absolute quantification curves for the *NOD2* and *ATG16L1* genes.

2.6. Modeling

The simulation model proposed that each allele of a gene at a specific time point can be either on or off and the synthesized mRNA is proposed to follow an exponential decay. The amount of total mRNA from each allele in a single cell at a certain time t is given by:

$$\sum (p(t, \lambda) e^{(\frac{t_i - t}{k})} \rho_e)$$

where $p(t, \lambda)$ is a hypergeometric random function determining the transcription state at time t : 1 = active or 0 = inactive, with the pulsing frequency λ ; ρ_e is a uniform random variable from 0 to 1 for the transcription efficiency of each pulse; k is the parameter to control the mRNA exponential decay rate. For the deterministic model, the ρ_e is picked once for two alleles and kept constant in all single cells. In contrast, for the stochastic model, the ρ_e is picked randomly from the uniform distribution from 0 to 1 for two alleles in all single cells. Simulation was performed 1000 times for each sample using the *R* statistical package (version 2.10.1) and total 1000 individuals were simulated for each model.

2.7. Statistic analysis

Two sample *t*-test was performed to compare the expression levels of the *NOD2* and *ATG16L1* genes in DCs with different treatments. F test was performed to compare the equality of variances between AI of non-infected and NDV-infected samples. The association between expression levels and AI was tested by linear regression model using *R* command *glm*. Pearson correlation between the two repeated AI measurements was calculated with the *R* command *cor*. The pairwise linkage disequilibrium was assessed between *NOD2* readout SNP rs3135499 and three *NOD2* Crohn's disease risk SNPs rs2066844, rs2066845 and rs2076756 as implemented in SNAP (SNP Annotation and Proxy Search)(Johnson et al., 2008).

3. Results

3.1. Subjects

DNA samples from 96 anonymous human blood donors obtained from New York blood bank were genotyped at *NOD2* rs3135499 and *ATG16L1* rs2241880. Both SNPs conformed to HWE ($p > 0.5$). The MAF (minor allele frequency) was 0.41 for rs3135499 and 0.48 for rs2241880, consistent with the HapMap MAF reported for subjects of European ancestry.

3.2. Gene expression

Quantitative real-time PCR was performed to determine the expression levels of two key CD susceptibility genes, *NOD2* and *ATG16L1*, in human DCs with regard to and regardless of viral exposure. In this experiment, we used Newcastle Disease Virus (NDV), an avian virus that has the ability to fully trigger the human innate immune responses (Hu et al., 2007; Lopez et al., 2006). Fig. 1 shows the total expression levels of *NOD2* and *ATG16L1* with or without NDV exposure. In the non-infected cells (NI: non-infected), the basal expression levels of both *NOD2* and *ATG16L1* were on average below 100 copies. The *NOD2* gene was induced about four-fold ($p < 0.001$ by *t*-test on log transformed copy numbers, Fig. 1A) 10 hours post NDV exposure at the multiplicity of infection (MOI) = 0.5 (where approximately 50% cells were infected). In contrast, the *ATG16L1* gene was not induced by NDV ($p = 0.88$, Fig. 1B). It should be noted that, regardless of virus exposure, the sample to sample variation in expression levels for both the *NOD2* and *ATG16L1* genes was high in our studied cell population (mean number of copies $\pm$ SD: 82 ± 53 and 313 ± 292 for *NOD2* and 25 ± 25 and 25 ± 24 for *ATG16L1* in non-infected and NDV-infected samples, respectively).

3.3. Allelic imbalance

We estimated AI of the *NOD2* and *ATG16L1* genes using common SNPs located in exonic regions. AI was defined by the difference in the proportion of transcripts from each allele of the same gene (Hu et al., 2010). For each gene, expression of the two alleles in heterozygotes was determined by pyrosequencing of the cDNA amplicons containing those common variants as the read-out SNPs. In this experiment, genomic DNA from the same donors of the cDNA samples was used as control for the AI measurements with heterozygote carriers expected to have an equal amount of transcripts of each allele resulting in AI = 0%. We determined the overall AI measurement variability (mean $\pm$ SD: $-6.6\% \pm$

11.3% for *ATG16L1* and $-8.2\% \pm 7.1\%$ for *NOD2*), enabling to define a 95% confidence interval for the measurement error of the cDNA samples.

Fig. 2 shows AI levels for *NOD2* and *ATG16L1* with or without NDV infection. Non-infected samples showed higher variance in AI than NDV-infected samples for the *NOD2* gene ($p = 0.05$, by F test of the equality of variances) (Fig. 2A and B), but not for the *ATG16L1* gene ($p = 0.32$) (Fig. 2C and D). Comparing with the genomic controls, significant AI ($p < 0.05$) was observed in 17 of 40 (~40%) and 27 of 40 (~70%) non-infected heterozygous samples for *NOD2* and *ATG16L1*, respectively. Regression analysis showed a significant association of the *NOD2* expression level with its AI levels in non-infected samples ($p = 0.015$) but not in NDV-induced samples ($p = 0.44$). No association was found for *ATG16L1* in either non-infected or infected cell populations ($p = 0.85$ and 0.36 , respectively).

We performed several quality control steps to examine whether the observed AI measurements are biased due to our experimental design. Firstly, we validated the reproducibility of the pyrosequencing method by repeating the AI measurements twice using thirteen samples. The results showed a strong correlation ($r = 0.92$) between the two trials (Supplementary Fig. 1). Secondly, the AI measurements in *NOD2* were repeated using a different readout SNP in a strong LD with the initial readout SNP. The results from both experiments were highly correlated ($r = 0.85$), confirming the robustness of our experimental design. Thirdly, we tested for PCR replication error, known to occur especially in the presence of low transcript copy numbers, which can be carried over into pyrosequencing. We repeated PCR amplifications five times for three cDNA samples with low copy numbers of *ATG16L1* transcripts and tested the AI measurement consistency across those samples. The variation among the repeated measurements (mean $\pm$ SD: AI = $3.7\% \pm 6.5\%$, $6.8\% \pm 1.8\%$, $11\% \pm 6.0\%$) was consistent with the variation among genomic DNA controls, confirming an adequate reproducibility of the PCR amplification preceding the pyrosequencing approach. Fourthly, we tested whether AI is consistent across different cell samples within DCs in the same subject. For five *ATG16L1* heterozygotes, DCs from each sample were equally separated into 8–12 cell samples, and the AI of each cell sample was determined. The results showed that AI was consistent across different cell samples from the same subject (Supplementary Table 2). Considerable AI was detected in all cell samples for several subjects; however, a much higher inter-subject AI variation (42.7%) than intra-subject AI variation (4.7%) was observed. Finally, the RNA extracted from the DCs was converted to cDNA by a different cDNA conversion kit to compare whether there are any biases in RNA reverse transcription process. We concluded that the source of heterogeneity in AI is unlikely to come from the cell sub-type or processing biases and is consistent with the variation among samples.

3.4. Simulation analysis

To interpret our experimental results, we constructed two single-cell transcription pulsing models to simulate the AI variation across cell populations. Two alternative sources of AI were proposed in our model: 1. AI in gene expression as a result of the consistent but differential transcription efficiency between two alleles of the same gene, illustrated as a

deterministic model (Fig. 3A). 2. AI in gene expression as a result of random fluctuation in gene transcription from different alleles, illustrated as a stochastic model (Fig. 3B). The overall AI on the cell population level was estimated and plotted in Fig. 3B. Both models led to similar AI distributions at the cell population level consistent with our experimental observations. However, compared to the stochastic model, the deterministic model was more preferable in the presence of the large number of transcripts. This result suggests that *NOD2* AI is most likely to result from consistent differences in allelic expression. For the *ATG16L1* gene, due to its low expression level (Fig. 1), our simulations were unable to distinguish between these two models.

4. Discussion

NOD2, the first gene found to be associated with CD, is a polymorphic gene involved in innate immune responses with several common SNPs shown to have a biological impact on the basal characteristics of DC response to bacterial invasion (Cooney et al., 2009; Salucci et al., 2008). Significant *NOD2* AI has been previously observed in lymphoblastoid cell lines (Yan et al., 2002b). Our study, for the first time, showed initial evidence of a substantial AI in the *NOD2* gene in the antigen presenting cells, particularly, DCs. Furthermore, our results revealed that endogenous *NOD2* transcription is in fact significantly up-regulated by NDV infection in human DCs (Fig. 1). This virus-host gene interplay is consistent with earlier findings that *NOD2* is also a viral pattern-recognition receptor, highlighting its role in the host antiviral defense mechanisms (Sabbah et al., 2009). *NOD2* has also been known as not only the key sensor of bacteria or its components but also the important mediator of the autophagy process in DCs (Cooney et al., 2009). Our results suggest the possible role of viral infection in moderating the DC response to bacterial invasion by up-regulating the *NOD2* expression. This finding may help to guide future investigations aimed to examine whether exposure to virus affects bacterial sensing and autophagy process triggered by bacterial invasion in DCs.

Another CD-associated gene, *ATG16L1*, is known as a key component of a large protein complex essential for autophagy (Cadwell et al., 2008; Saitoh et al., 2008) playing an important role in immunity and inflammation (Levine et al., 2011). Studies of different mouse strains with *ATG16L1* deficiencies have shown that autophagy is important in regulating the secretory function of Paneth cells and the production of inflammatory cytokines in the intestine (Cadwell et al., 2008; Saitoh et al., 2008). Furthermore, several genome-wide association studies have found common *ATG16L1* SNPs to be associated with CD susceptibility (Hampe et al., 2007; Rioux et al., 2007). In our experiment, we examined the AI level of *ATG16L1* in DCs by using its CD risk variant, rs2241880, as the read-out SNP. For the first time, our results revealed that 70% of the examined DC samples showed AI (Fig. 2), indicating differential expression of the rs2241880 CD risk alleles.

We detected a high level of variation in total expression and differential expression between the two alleles of the *NOD2* and *ATG16L1* genes. To address a concern whether the observed variation resulted from experimental biases or measurement errors, we conducted a series of control experiments that suggested that the major source of the observed expression

variation is unlikely to result from a measurement or sampling error, but rather from the sample-to-sample variability.

To interpret our results in the context of experiments carried out at the cell population level, we conducted simulations on a single cell level encompassed of two distinct models assuming that gene expression of the two alleles is either consistent or random across individual cells. Our findings suggest that AI at *NOD2* is consistent with the deterministic model, implying that each individual cell would show a relatively similar response to pathogen processing. The model was inconclusive for *ATG16L1*, suggesting that both mechanisms are possible: the one described for *NOD2* and an alternative mechanism indicating that individual DCs may respond differentially to a pathogen resulting in the different levels of its maturation, migration, antigen presenting and later recruitment of T-cells. To further investigate AI mechanisms an experimental assessment at a single cell level is required.

The AI observed for *NOD2* and *ATG16L1* may also suggest either genetic or epigenetic effects on DNA transcription, leading to differential transcription efficiency between the two alleles (proposed in our simulation model above). Several studies have reported the association of DNA methylation of the *NOD2* promoter with patient's age and disease status (Fernandez et al., 2011; Kerkel et al., 2010). Studies on various diseases have shown that differential promoter specific DNA methylation plays a critical role in expression variation between two alleles (Gonsky et al., 2009; Grayson et al., 2005; Movassagh et al., 2010; Titze et al., 2009). Therefore, a possible mechanism for the variation in AI of *NOD2* and *ATG16L1* is that it results from differences in gene expression between alleles with differential DNA promoter-specific methylation. Based on that, we would expect that a higher expression level would be associated with low AI (both alleles turned on). Notably, in non-infected DCs (Fig. 2A), we observed a reverse correlation between the *NOD2* expression level and AI. In order to validate this mechanism, future work should examine the extent to which DNA methylation affects expression variation, in particular AI, in CD susceptibility genes in patients and disease-free individuals.

Limitations of our study include the fact that it has been conducted in anonymous blood donors so that we have no information on their CD status and other phenotypic characteristics. Nevertheless, CD is relatively rare in the general US population (1 in 5000), and therefore the high variability in gene expression observed in this study cannot be explained by the presence of CD cases among the blood donors. Also, given a low frequency of *NOD2* CD risk alleles in the general population (Economou et al., 2004; Hugot et al., 2007), we used a non-CD-related coding read-out SNP to determine the AI of *NOD2*. Even though a common coding variant P268S (rs2066842) is the background variant for the three major *NOD2* risk SNPs (Zhang et al., 2013), it has not been directly implicated in CD risk in individuals of European ancestry mostly represented in our sample. Therefore, we selected a 3' UTR SNP to detect the full-length cDNA. In addition, in order to fully trigger the immune responses of human DCs, we exposed the cells to NDV. However, in reality, different human viruses possess various properties to inhibit the host immune responses. Such a host-viral interplay may contribute to the regulation of the *NOD2* gene in this key immunity cell type but may not be accounted for in the present experimental setting. Moreover, we verified

that the observed variation in AI is unlikely to be due to measurement errors or cell selection, but other sources of sample-to-sample variability, such as protein translation, can also exist.

5. Conclusions

In summary, the present study examined expression variation of the selected CD susceptible genes, *NOD2* and *ATG16L1*, in human DCs and demonstrated a significant AI. Furthermore, this study demonstrated that viral infection up-regulates *NOD2* expression, but not *ATG16L1* expression in DCs. Moreover, our results suggest that high variability in AI of the *NOD2* gene is due to consistent differences in expression between two alleles across the DCs. Future studies are needed to compare AI patterns between CD patients and unaffected controls.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was funded by the New York Crohn's Disease Foundation. We thank Dr. James Wetmur for providing samples and critically reviewing the manuscript. We thank Ms. Yaomei Ding and Ms. Madhu Kumar for DC cultivation and the DNA and cDNA sample preparations.

Abbreviations

DCs	Dendritic cells
CD	Crohn's disease
NDV	Newcastle Disease Virus
AI	Allelic imbalance
<i>NOD2</i>	Nucleotide-binding oligomerization domain-containing protein 2
<i>ATG16L1</i>	Autophagy-related protein 16-1
DMEM	Dulbecco's Modified Eagle Medium
MOI	multiplicity of infection

References

- Abraham C, Cho JH. Inflammatory bowel disease. *N Engl J Med*. 2009; 361:2066–2078. [PubMed: 19923578]
- Banchereau J, et al. Immunobiology of dendritic cells. *Annu Rev Immunol*. 2000; 18:767–811. [PubMed: 10837075]
- Cadwell K, et al. A key role for autophagy and the autophagy gene *Atg16l1* in mouse and human intestinal Paneth cells. *Nature*. 2008; 456:259–263. [PubMed: 18849966]
- Coombes JL, Powrie F. Dendritic cells in intestinal immune regulation. *Nat Rev Immunol*. 2008; 8:435–446. [PubMed: 18500229]
- Cooney R, et al. *NOD2* stimulation induces autophagy in dendritic cells influencing bacterial handling and antigen presentation. *Nat Med*. 2009; 16:90–97. [PubMed: 19966812]

- Economou M, Trikalinos TA, Loizou KT, Tsianos EV, Ioannidis JP. Differential effects of NOD2 variants on Crohn's disease risk and phenotype in diverse populations: a metaanalysis. *Am J Gastroenterol*. 2004; 99:2393–2404. [PubMed: 15571588]
- Fernandez AF, et al. A DNA methylation fingerprint of 1628 human samples. *Genome Res*. 2011; 22:407–419. [PubMed: 21613409]
- Gonsky R, Deem RL, Targan SR. Distinct methylation of IFNG in the gut. *J Interferon Cytokine Res*. 2009; 29:407–414. [PubMed: 19450149]
- Grayson DR, et al. Reelin promoter hypermethylation in schizophrenia. *Proc Natl Acad Sci U S A*. 2005; 102:9341–9346. [PubMed: 15961543]
- Guo M, Rupe MA, Zinselmeier C, Habben J, Bowen BA, Smith OS. Allelic variation of gene expression in maize hybrids. *Plant Cell*. 2004; 16:1707–1716. [PubMed: 15194819]
- Hampe J, et al. A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. *Nat Genet*. 2007; 39:207–211. [PubMed: 17200669]
- Heap GA, et al. Genome-wide analysis of allelic expression imbalance in human primary cells by high-throughput transcriptome resequencing. *Hum Mol Genet*. 2009; 19:122–134. [PubMed: 19825846]
- Hu J, et al. Chromosome-specific and noisy IFNB1 transcription in individual virus-infected human primary dendritic cells. *Nucleic Acids Res*. 2007; 35:5232–5241. [PubMed: 17675303]
- Hu J, et al. A common polymorphism in the caspase recruitment domain of RIG-I modifies the innate immune response of human dendritic cells. *J Immunol*. 2010; 185:424–432. [PubMed: 20511549]
- Hu J, et al. Role of cell-to-cell variability in activating a positive feedback antiviral response in human dendritic cells. *PLoS One*. 2011; 6:e16614. [PubMed: 21347441]
- Hugot JP, et al. Prevalence of CARD15/NOD2 mutations in Caucasian healthy people. *Am J Gastroenterol*. 2007; 102:1259–1267. [PubMed: 17319929]
- Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics*. 2008; 24:2938–2939. [PubMed: 18974171]
- Jostins L, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature*. 2012; 491:119–124. [PubMed: 23128233]
- Kakuta Y, et al. TNFSF15 transcripts from risk haplotype for Crohn's disease are overexpressed in stimulated T cells. *Hum Mol Genet*. 2009; 18:1089–1098. [PubMed: 19124533]
- Kerkel K, et al. Altered DNA methylation in leukocytes with trisomy 21. *PLoS Genet*. 2010; 6:e1001212. [PubMed: 21124956]
- Levine B, Mizushima N, Virgin HW. Autophagy in immunity and inflammation. *Nature*. 2011; 469:323–335. [PubMed: 21248839]
- Levy S, et al. The diploid genome sequence of an individual human. *PLoS Biol*. 2007; 5:e254. [PubMed: 17803354]
- Libioulle C, et al. Novel Crohn disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of PTGER4. *PLoS Genet*. 2007; 3:e58. [PubMed: 17447842]
- Lo HS, et al. Allelic variation in gene expression is common in the human genome. *Genome Res*. 2003; 13:1855–1862. [PubMed: 12902379]
- Lopez CB, Yount JS, Moran TM. Toll-like receptor-independent triggering of dendritic cell maturation by viruses. *J Virol*. 2006; 80:3128–3134. [PubMed: 16537581]
- McCarroll SA, et al. Deletion polymorphism upstream of IRGM associated with altered IRGM expression and Crohn's disease. *Nat Genet*. 2008; 40:1107–1112. [PubMed: 19165925]
- Meyer KB, et al. Allele-specific up-regulation of FGFR2 increases susceptibility to breast cancer. *PLoS Biol*. 2008; 6:e108. [PubMed: 18462018]
- Movassagh M, Choy MK, Goddard M, Bennett MR, Down TA, Foo RS. Differential DNA methylation correlates with differential expression of angiogenic factors in human heart failure. *PLoS One*. 2010; 5:e8564. [PubMed: 20084101]
- Pant PV, Tao H, Beilharz EJ, Ballinger DG, Cox DR, Frazer KA. Analysis of allelic differential expression in human white blood cells. *Genome Res*. 2006; 16:331–339. [PubMed: 16467561]

- Parker-Katirae L, Bousiaki E, Monk D, Moore GE, Nakabayashi K, Scherer SW. Dynamic variation in allele-specific gene expression of Paraoxonase-1 in murine and human tissues. *Hum Mol Genet*. 2008; 17:3263–3270. [PubMed: 18678600]
- Prescott NJ, Dominy KM, Kubo M, Lewis CM, Fisher SA, Redon R, Huang N, Stranger BE, Blaszczyk K, Hudspeth B, Parkes G, Hosono N, Yamazaki K, Onnie CM, Forbes A, Dermitzakis ET, Nakamura Y, Mansfield JC, Sanderson J, Hurles ME, Roberts RG, Mathew CG. Independent and population-specific association of risk variants at the IRGM locus with Crohn's disease. *Hum Mol Genet*. 2010; 19:1828–1839. [PubMed: 20106866]
- Reis e Sousa C. Dendritic cells in a mature age. *Nat Rev Immunol*. 2006; 6:476–483. [PubMed: 16691244]
- Rescigno M, Di Sabatino A. Dendritic cells in intestinal homeostasis and disease. *J Clin Invest*. 2009; 119:2441–2450. [PubMed: 19729841]
- Rioux JD, et al. Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nat Genet*. 2007; 39:596–604. [PubMed: 17435756]
- Sabbah A, et al. Activation of innate immune antiviral responses by Nod2. *Nat Immunol*. 2009; 10:1073–1080. [PubMed: 19701189]
- Saitoh T, et al. Loss of the autophagy protein Atg16L1 enhances endotoxin-induced IL-1 beta production. *Nature*. 2008; 456(U68):264. [PubMed: 18849965]
- Salucci V, et al. Monocyte-derived dendritic cells from Crohn patients show differential NOD2/CARD15-dependent immune responses to bacteria. *Inflamm Bowel Dis*. 2008; 14:812–818. [PubMed: 18240302]
- Shanahan F. Crohn's disease. *Lancet*. 2002; 359:62–69. [PubMed: 11809204]
- Titze S, et al. Differential MSH2 promoter methylation in blood cells of Neuro-fibromatosis type 1 (NF1) patients. *Eur J Hum Genet*. 2009; 18:81–87. [PubMed: 19639020]
- Utting M, Hampe J, Platzer M, Huse K. Locking of 3' ends of single-stranded DNA templates for improved Pyrosequencing performance. *Biotechniques*. 2004; 37(70–3):66–67. [PubMed: 15283202]
- van Beelen AJ, et al. Stimulation of the intracellular bacterial sensor NOD2 programs dendritic cells to promote interleukin-17 production in human memory T cells. *Immunity*. 2007; 27:660–669. [PubMed: 17919942]
- Yan H, et al. Small changes in expression affect predisposition to tumorigenesis. *Nat Genet*. 2002a; 30:25–26. [PubMed: 11743581]
- Yan H, Yuan W, Velculescu VE, Vogelstein B, Kinzler KW. Allelic variation in human gene expression. *Science*. 2002b; 297:1143. [PubMed: 12183620]
- Zhang, W., et al. Extended haplotype association study in Crohn's disease identifies a novel, Ashkenazi Jewish-specific missense mutation in the NF-kappaB pathway gene, HEATR3. *Genes Immun*. 2013. <http://dx.doi.org/10.1038/gene.2013.19>

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.gene.2013.06.066>.

Fig. 1.

Expression of *NOD2* (A) and *ATG16L1* (B) in DCs. NI: not infected; NDV: Newcastle Disease Virus-infected. The Y axis shows the log-transformed copy number of transcripts. Copy number was calculated using the standard PCR absolute quantification curves from serial dilutions of genomic DNA. PCR result on a no template control was negative. The PCR-amplified products were pyrosequenced so that the sequence was verified for each sample. The *NOD2* gene was induced about four-fold and the difference in *NOD2* expression level between non-infected and NDV-infected samples was significant ($p < 0.001$ by *t*-test on log transformed transcript copy numbers). The *ATG16L1* gene was not induced by NDV infection ($p = 0.88$).

Fig. 2.

Allelic imbalance (AI) of *NOD2* and *ATG16L1* read-out SNP heterozygotes. Y axis shows that the AI ranges from -1 (100% minor allele) to 1 (100% major allele). A, B: *NOD2* rs3135499 heterozygote (G/T), A: non-infected; B: NDV-infected. C, D: *ATG16L1* rs2241880 heterozygote (C/T), C: non-infected; D: NDV-infected. The measured allelic imbalance of the *NOD2* and *ATG16L1* read-out SNP homozygotes was close but not equal to -1 or 1 , likely due to the pyrosequencing bias from the self-priming of ss-DNA 3'end (Utting et al., 2004).

Fig. 3. Simulation results. A illustrates the difference between the two single cell models: *Deterministic*, where each single cell has the same allelic expression pattern and *Stochastic* where the allelic expression status varies across cells. B shows the AI results in 1000 individuals combining 1000 simulations in single cells using both models with regard to AI of *ATG16L1* and *NOD2*.