

ARTICLE

Received 11 Feb 2014 | Accepted 11 Sep 2014 | Published 20 Oct 2014

DOI: 10.1038/ncomms6236

OPEN

# Characterizing the genetic basis of innate immune response in TLR4-activated human monocytes

Sarah Kim<sup>1,2,3</sup>, Jessica Becker<sup>1,2,\*</sup>, Matthias Bechheim<sup>3,\*</sup>, Vera Kaiser<sup>3</sup>, Mahdad Noursadeghi<sup>4</sup>, Nadine Fricker<sup>1,2</sup>, Esther Beier<sup>3</sup>, Sven Klaschik<sup>5</sup>, Peter Boor<sup>6</sup>, Timo Hess<sup>1,2</sup>, Andrea Hofmann<sup>1,2</sup>, Stefan Holdenrieder<sup>7</sup>, Jens R. Wendland<sup>8</sup>, Holger Fröhlich<sup>9</sup>, Gunther Hartmann<sup>7</sup>, Markus M. Nöthen<sup>1,2</sup>, Bertram Müller-Myhsok<sup>10,11,12</sup>, Benno Pütz<sup>10,\*</sup>, Veit Hornung<sup>3,\*</sup> & Johannes Schumacher<sup>1,2,\*</sup>

Toll-like receptors (TLRs) play a key role in innate immunity. Apart from their function in host defense, dysregulation in TLR signalling can confer risk to autoimmune diseases, septic shock or cancer. Here we report genetic variants and transcripts that are active only during TLR signalling and contribute to interindividual differences in immune response. Comparing unstimulated versus TLR4-stimulated monocytes reveals 1,471 expression quantitative trait loci (eQTLs) that are unique to TLR4 stimulation. Among these we find functional SNPs for the expression of *NEU4*, *CCL14*, *CBX3* and *IRF5* on TLR4 activation. Furthermore, we show that SNPs conferring risk to primary biliary cirrhosis (PBC), inflammatory bowel disease (IBD) and celiac disease are immune response eQTLs for *PDGFB* and *IL18R1*. Thus, *PDGFB* and *IL18R1* represent plausible candidates for studying the pathophysiology of these disorders in the context of TLR4 activation. In summary, this study presents novel insights into the genetic basis of the innate immune response and exemplifies the value of eQTL studies in the context of exogenous cell stimulation.

<sup>1</sup>Institute of Human Genetics, University of Bonn, Bonn 53127, Germany. <sup>2</sup>Department of Genomics, Life & Brain Center, University of Bonn, Bonn 53127, Germany. <sup>3</sup>Institute of Molecular Medicine, University of Bonn, Bonn 53127, Germany. <sup>4</sup>Division of Infection and Immunity, University College London, London WC1E 6BT, UK. <sup>5</sup>Department for Anesthesiology and Intensive Care Medicine, University of Bonn, Bonn 53127, Germany. <sup>6</sup>Institute of Pathology and Department of Nephrology, University Clinic of RWTH Aachen, Aachen 52074, Germany. <sup>7</sup>Institute for Clinical Chemistry and Clinical Pharmacology, University of Bonn, Bonn 53127, Germany. <sup>8</sup>Worldwide R&D, Pfizer Inc., Cambridge, Massachusetts 02139, USA. <sup>9</sup>Bonn-Aachen International Center for IT (B-IT), University of Bonn, Bonn 53113, Germany. <sup>10</sup>Statistical Genetics, Max Planck Institute of Psychiatry, Munich 80804, Germany. <sup>11</sup>Munich Cluster for Systems Neurology (SyNergy), Munich 80804, Germany. <sup>12</sup>Institute of Translational Medicine, University of Liverpool, Liverpool L69 3GL, UK. \* These authors contributed equally to this work. Correspondence and requests for materials should be addressed to V.H. (email: veit.hornung@uni-bonn.de) or to J.S. (email: johannes.schumacher@uni-bonn.de).

The innate immune system is the first line of defense against invading microorganisms. It is equipped with pattern recognition receptors (PRRs) that have evolved to sense conserved microbe-associated molecular patterns. Among other PRR families, Toll-like receptors (TLRs) play a pivotal role in antimicrobial defense, with TLR4 being one of the best-characterized member. TLR4 detects lipopolysaccharide (LPS) derived from Gram-negative bacteria and relays signal transduction to induce pro-inflammatory gene expression<sup>1</sup>. Apart from their antimicrobial function, TLRs also sense endogenous molecules, known as damage-associated molecular patterns<sup>2,3</sup>. Mutations and polymorphisms in TLR and TLR-signalling genes have been shown to confer susceptibility to many infectious and inflammatory diseases<sup>4,5</sup>.

Here we perform transcriptome profiling in purified monocytes and genome-wide single-nucleotide polymorphism (SNP)-genotyping in 137 healthy individuals to identify transcripts and genetic variants that contribute to interindividual differences in TLR4 signalling. This innovative approach provides new insights into the genetic control of innate immune response.

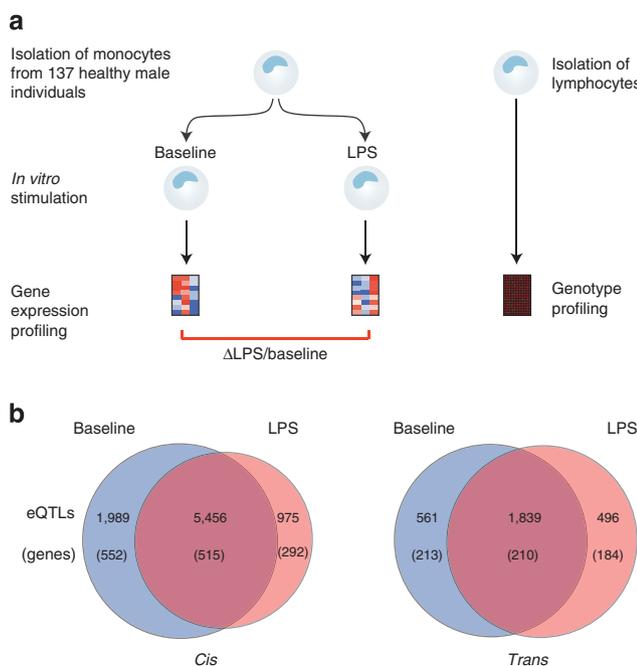
## Results

Expression profiling of unstimulated (baseline) and TLR4-stimulated monocytes revealed 9,031 differentially expressed transcripts at 90-min LPS exposure ( $P < 8.7 \times 10^{-07}$ , Bonferroni-corrected  $P$  value  $< 0.01$ ; Supplementary Fig. 1a). A duration of 90 min was chosen to study a time point that covers the primary wave of gene expression (independent of *de novo* protein synthesis) with an informative set of differentially expressed genes following PRR stimulation<sup>6</sup>. Among the most strongly differentially expressed genes (mean  $\log_2$  fold change  $> 1$ ), pathway analyses revealed a significant enrichment of genes involved in PRR signalling and their respective outputs (Supplementary Fig. 1b and Supplementary Data 1). Further

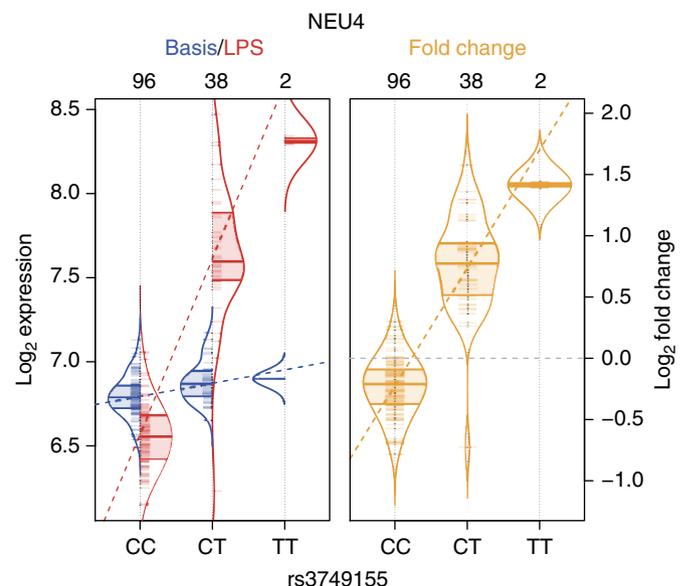
subcategorization of differentially expressed genes into subsets of coexpressed genes allowed additional refinement of the ascribed signalling pathways to specific clusters (Supplementary Fig. 1c and Supplementary Data 2).

For both conditions—baseline and TLR4 signalling—expression quantitative trait loci (eQTL) analyses were performed and a false discovery rate (FDR) of 0.01 was used to define eQTLs ( $P_{\text{Baseline}} < 8.67 \times 10^{-09}$ ,  $P_{\text{LPS}} < 7.35 \times 10^{-09}$ ). To avoid underestimation of shared eQTLs between treated and untreated monocytes, we additionally conducted a multivariate analysis of variance (MANOVA)<sup>7</sup> ( $P_{\text{Manova}} < 1.76 \times 10^{-09}$ , FDR-corrected  $P$  value  $< 0.01$ , Supplementary Fig. 2a,b). In addition, we calculated the differential expression of baseline vs TLR4-stimulated cells and performed eQTL analyses using an FDR of 0.01 ( $P_{\Delta\text{LPS}/\text{Baseline}} < 4.33 \times 10^{-10}$ ). In total, we identified 975 *cis* (corresponding to 292 genes) and 496 *trans* eQTLs (corresponding to 184 genes), which are unique to TLR4 stimulation (Fig. 1b, Supplementary Fig. 3), as well as 435 *cis* (corresponding to 74 genes) and 17 *trans* eQTLs (corresponding to 8 genes), when differential expression ( $\Delta\text{LPS}/\text{Baseline}$ ) was considered (Supplementary Fig. 4). A complete list of all significant eQTLs (baseline, LPS,  $\Delta\text{LPS}/\text{baseline}$ ) is provided by Supplementary Data 3–5.

Interestingly, analysing all eQTLs together (both *cis* and *trans*, baseline, LPS,  $\Delta\text{LPS}/\text{baseline}$ ) revealed a significant enrichment for lysosome-associated pathways, such as lysosomal enzymes and respective trafficking components (Supplementary Fig. 5, Supplementary Data 6). This is noteworthy, as lysosomes have been shown to play pivotal roles in both sensing and effector



**Figure 1 | Experimental design of the study in human monocytes of 137 healthy male donors.** (a) Overview of step-wise study design applied in human monocytes. (b) Venn diagrams of eQTLs and corresponding transcripts of baseline (Baseline) and LPS-treated (LPS) monocytes in *cis* (left) and *trans* (right). Linear regression model was used for eQTL mapping. MANOVA was used to determine shared eQTLs.



**Figure 2 | rs3749155 represents the strongest *cis* iQTL during TLR4 signalling.** Allele-specific gene expression in untreated versus LPS-treated monocytes is depicted as violin plots. The distribution of  $\log_2$  expression is shown on the left (blue: baseline, red: LPS) and  $\log_2$  fold change on the right (orange). The dashed lines indicate the regressions of the ( $\Delta$ ) gene expression values versus genotype. Individuals are indicated by small tick marks on the respective axes, with the group sizes per genotype indicated on top of each plot. The shaded ranges correspond to the 25–75% quartiles with the median indicated separately (analogous to a common boxplot). The violin plots show that *NEU4* is induced in carriers of allele T on LPS exposure, whereas *NEU4* expression is reduced in homozygotes of allele C.

function of innate immune cells, our results suggest that there is a genetic contribution to these pathways<sup>8,9</sup>. In the following we focused on eQTLs unique to TLR4 signalling, both under absolute and differential expression, and refer to them as immune response eQTLs or iQTLs.

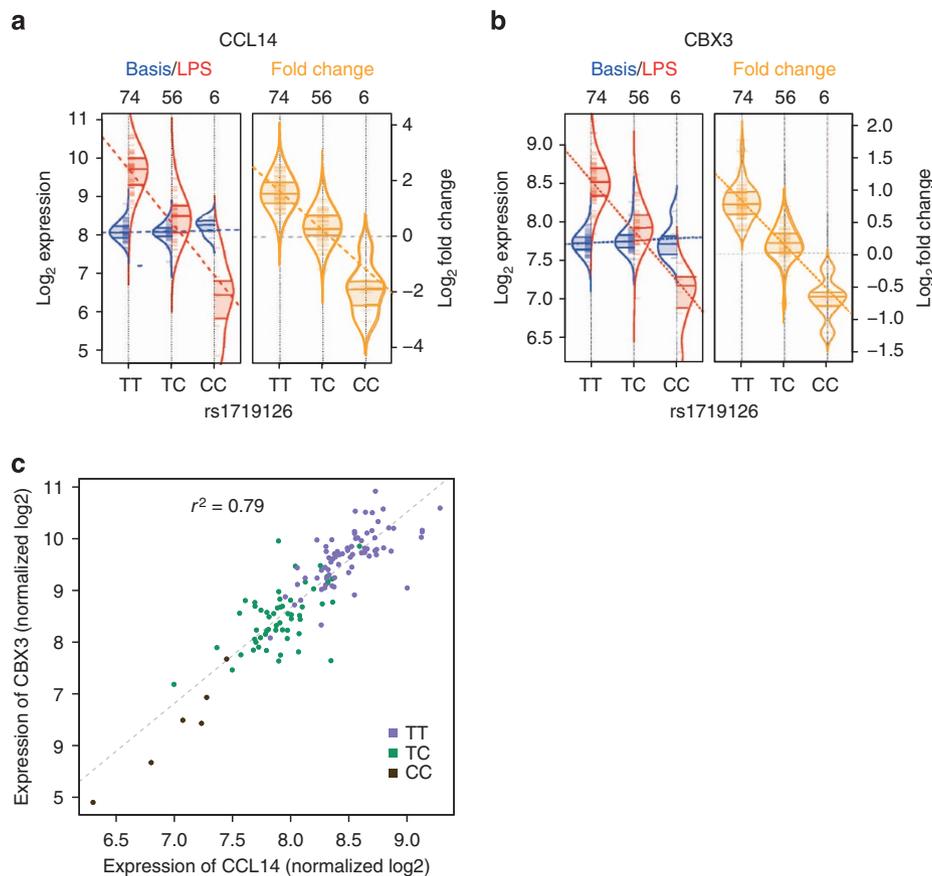
In line with the pathway analysis above, our strongest iQTL (Supplementary Table 1) is represented by the SNP rs3749155 on chromosome 2q37 and regulates the expression of *neuraminidase 4 (NEU4)*, which belongs to a group of lysosomal neuraminidases that catalyse the cleavage of sialic acids linked to glycoconjugates<sup>10</sup>. This iQTL is located in intron 3 of *NEU4* and regulates its expression in *cis* ( $P_{LPS} = 3.44 \times 10^{-44}$ , Fig. 2). In the context of innate immune signalling pathways, neuraminidases have been shown to regulate the activity of TLRs at the posttranslational level. For example, *NEU1*—the best-characterized neuraminidase—promotes the cleavage of sialic acids from the ectodomain of TLR4, which is a prerequisite for TLR activation<sup>11</sup>. Although a similar effect on TLR4 has been observed for *NEU4* (ref. 12), the relevance of *NEU4* to innate immunity has been less intensively studied so far. Our findings, however, suggest that genetic variability in *NEU4* strongly contributes to interindividual differences in the innate immune response.

SNP rs1719126 on chromosome 17q12 is the most significant *trans* iQTL and additionally represents the third strongest *cis* iQTL (Supplementary Table 1). In *cis* rs1719126 regulates the expression of the *chemokine (C-C motif) ligand 14 (CCL14)* gene, whose transcription start site (TSS) is 101 kb distant from

rs1719126 ( $P_{LPS} = 5.80 \times 10^{-30}$ , Fig. 3a). Chip-seq data from ENCODE<sup>13</sup> revealed an RNA polymerase II binding site overlapping rs1719126, which represents a potential mechanism of *CCL14* regulation (Supplementary Fig. 6). On the functional level, *CCL14* is a highly expressed chemokine, whose activity as a chemotactic signal for pro-inflammatory cells is mainly regulated by a proteolytic processing step<sup>14</sup>. Our data now reveal that *CCL14* is additionally regulated at the transcriptional level with strong genetic determination.

In *trans* rs1719126 regulates the expression of the *chromobox protein homolog 3 (CBX3)* gene on chromosome 7p15 ( $P_{LPS} = 3.48 \times 10^{-28}$ , Fig. 3b). We observed a strong correlation between *CCL14* and *CBX3* expressions among our volunteers ( $r^2 = 0.79$ , Fig. 3c), indicating that *cis*-modulation of *CCL14* results in differential expression of *CBX3*. While previous studies have found that *CBX3* acts as a heterochromatin protein<sup>15</sup>, a recent study has shown that *CBX3* directly regulates the expression and splicing of many genes<sup>16</sup>. In particular, after cytokine exposure *CBX3* is specifically recruited to inflammatory genes<sup>16</sup>.

Next, we used differential gene expression on TLR4 stimulation ( $\Delta LPS/\text{baseline}$ ) as an alternative quantitative trait to identify iQTLs that are already present in unstimulated monocytes but have additional allele-specific effects during immune response (Supplementary Fig. 4). Among these, rs10239340 on chromosome 7q32 acts in *cis* on the expression of *interferon 5 (IRF5)* and represents the most significant iQTL (Supplementary Table 2). In baseline monocytes, carriers of the T allele show elevated *IRF5*

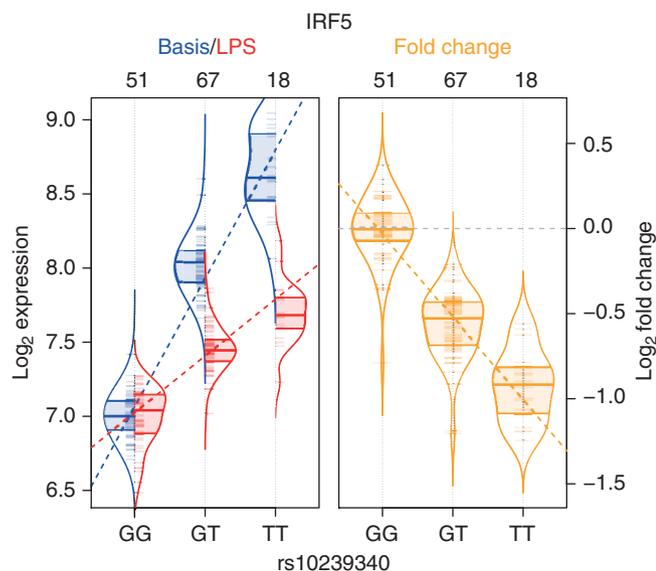


**Figure 3 | rs1719126 represents the strongest *trans* and the third strongest *cis* iQTL during TLR4 signalling.** (a) In homozygous carriers of allele T, *CCL14* expression is induced on LPS exposure and reduced in homozygotes for the opposite allele. (b) The allele-specific expression of *CBX3* is nearly identical to that observed for *CCL14*. The same homozygous T carriers show an induction of *CBX3* expression on LPS exposure, whereas a reduced expression is observed in homozygotes for the opposite allele. (c) Correlation of *CCL14* and *CBX3* expressions on LPS exposure ( $r^2 = 0.79$ ). Violin plots are depicted as in Fig. 2.

expression compared with carriers of the opposite allele ( $P_{\text{Baseline}} = 3.97 \times 10^{-59}$ ). However, on TLR4 stimulation T-allele carriers show an allele-specific downregulation of *IRF5* ( $P_{\text{ALPS/Baseline}} = 1.63 \times 10^{-35}$ , Fig. 4). *IRF5* represents a key pro-inflammatory transcription factor for the expression of many immune-relevant genes<sup>17</sup>. Thus, *IRF5* expression differences during TLR4 signalling probably affect pro-inflammatory immune response at later time points.

eQTLs for the expression of *IRF5* have been previously found in unstimulated LCLs and whole-blood samples<sup>18,19</sup>. In LCLs we observed an opposing directional eQTL effect<sup>18</sup> compared with monocytes. Here, carriers of the G allele showed elevated *IRF5* expression compared with carriers of allele T. In contrast, in whole-blood samples<sup>19</sup> we studied rs10229001, which is in perfect LD to our eQTL marker ( $r^2 = 1$ ) and observed the same expression effect as seen in monocytes. Thus—besides the context specific eQTL effect on LPS stimulation—rs10239340 represents a cell-type-specific eQTL.

Recently *lysozyme* (*LYZ*) on chromosome 12q15 has been described as a monocyte-specific master regulator eQTL<sup>20</sup>. In this study, rs10784774 was identified as *cis* eQTL for *LYZ* and as *trans* eQTL for 62 other genes under baseline condition. Analysis of hotspots in our data confirms *LYZ* as a master regulator (Supplementary Fig. 7a). *LYZ* was allele-specifically regulated in *cis* ( $P_{\text{Baseline}} = 4.38 \times 10^{-37}$ ) and a total of 108 different genes in *trans* (Supplementary Fig. 7b and Supplementary Data 7). This master regulator eQTL was also present in stimulated monocytes (Supplementary Fig. 7). However, no *LYZ*-regulated transcript showed a significant differential expression on TLR4 stimulation, indicating that this pathway does not play a prominent role in the early phase of TLR4 signalling.



**Figure 4 | rs10239340 represents the strongest iQTL during TLR4 signalling that is also functionally active in baseline monocytes and shows an additional regulatory effect on differential expression on LPS exposure.** Carriers of the T allele have elevated baseline *IRF5* expression compared with individuals carrying the G allele (left; blue violin plots). On LPS exposure a robust downregulation of *IRF5* proportional to the number of T alleles is observed, whereas no differential regulation is seen in homozygotes for the G allele (right; orange violin plots). Violin plots are depicted as in Fig. 2.

Finally, we tested whether SNPs that have been identified as risk factors for common diseases represent iQTLs and used all SNPs that have been deposited in the Catalog of Published Genome-Wide Association Studies. This revealed many GWAS-SNPs that are eQTLs under baseline and TLR4-signalling condition (Supplementary Data 8) as well as GWAS-SNPs as being active iQTLs only during TLR4 signalling.

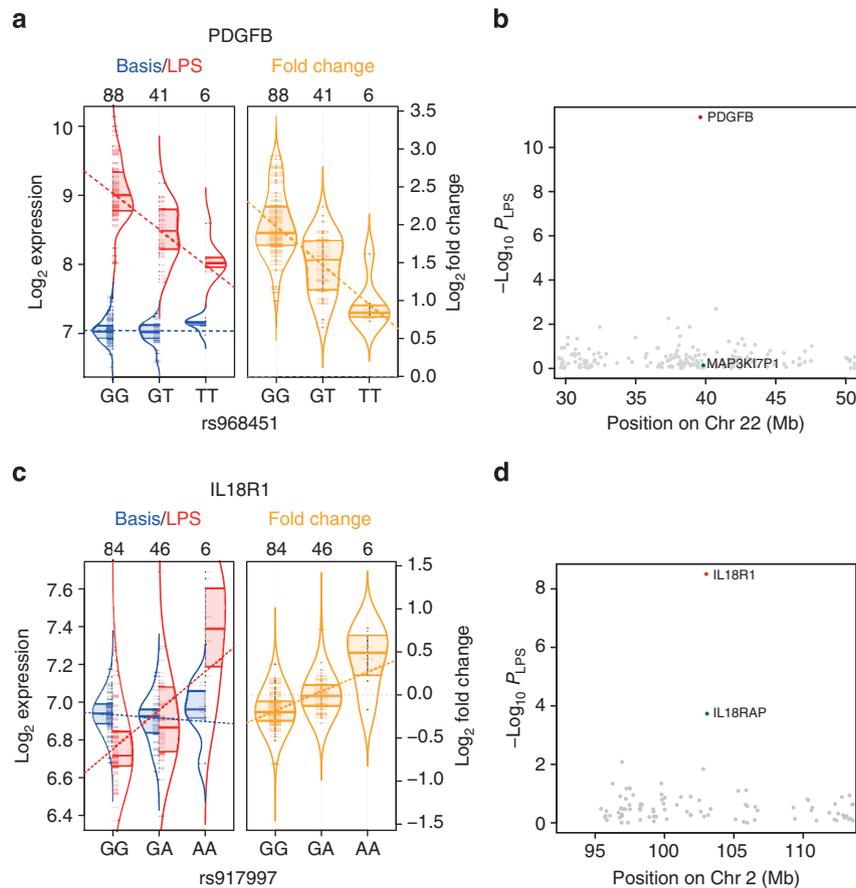
The most significant GWAS-iQTL represents rs968451 on chromosome 22q13, which regulates the expression of the *platelet-derived growth factor beta polypeptide* (*PDGFB*) gene in *cis*. The expression of *PDGFB* is induced during TLR4 signalling strongest for carriers of the common allele G ( $P_{\text{LPS}} = 4.31 \times 10^{-12}$ ,  $P_{\text{ALPS/Baseline}} = 6.09 \times 10^{-12}$ , Fig. 5a). SNP rs968451 is located 30 kb distant to the TSS of *PDGFB* and also confers risk for primary biliary cirrhosis (PBC)<sup>21</sup>, a severe autoimmune liver disease characterized by a destructive cholangitis. In this GWAS, the authors favoured *MAP3K7IP1* as being the risk-conferring PBC gene at this locus<sup>21</sup>, which is located 125 kb distant to rs968451. However, the testing of all transcripts within a 10-Mb interval of rs968451 in our data set revealed that *PDGFB* is the only transcript regulated by the PBC risk variant during TLR4 signalling (Fig. 5b).

Our data show that the PBC variant rs968451 regulates *PDGFB* expression with reduced transcript levels in risk allele carriers. This function is only active during TLR4 signalling, which fits pathophysiological concepts, according to which TLR signalling confers PBC risk<sup>22</sup>. We therefore suggest *PDGFB* as a novel PBC risk gene at this locus. Also, on the functional level *PDGFB* represents a plausible PBC risk gene, given its involvement in the induction of angiogenesis, a process that has been linked to the pathophysiology of PBC<sup>23</sup>.

The second GWAS-iQTL concerns rs917997 on chromosome 2q12, which is 91 kb distant to the TSS of the *interleukin 18 receptor 1* (*IL18R1*) gene and regulates its expression in *cis*. Only during TLR4 signalling is the expression of *IL18R1* induced in homozygous A carriers and reduced in homozygous carriers of the opposite allele ( $P_{\text{LPS}} = 3.10 \times 10^{-9}$ ,  $P_{\text{ALPS/Baseline}} = 5.99 \times 10^{-11}$ , Fig. 5c). Allele A of rs917997 has been identified as a risk factor for inflammatory bowel disease (IBD)<sup>24</sup> and celiac disease<sup>25</sup>, both inflammatory diseases of the intestine. However, based on GWAS data it was not possible to favour a risk gene, because *IL18R1* forms, together with the interleukin receptor genes *IL1R1*, *IL1R2*, *IL1RL1*, *IL1RL2* and *IL18RAP*, a cluster at this locus. We tested whether these or other nearby genes are also regulated by rs917997 during TLR4 signalling. Although not significant, *IL18RAP* is additionally *cis*-regulated within a 10-Mb interval of the risk variant ( $P_{\text{LPS}} = 1.8 \times 10^{-4}$ , Fig. 5d and Supplementary Fig. 8).

Our data show that rs917997 becomes functionally active only during TLR4 signalling and thereby influences *IL18R1* expression and to lesser extent *IL18RAP* expression. *IL18R1* and *IL18RAP* form the heterodimeric IL18 receptor complex, which relays IL18-dependent signal transduction to the NF $\kappa$ B-pathway<sup>26</sup>. As such, *IL18R1* and *IL18RAP* represent plausible risk genes for IBD and celiac disease at this locus, particularly as alterations in IL18- and TLR-mediated processes have been reported for both disorders<sup>27</sup>.

As for *IRF5*, an eQTL effect of rs917997 on the expression of *IL18RAP* has been previously described in unstimulated whole-blood samples<sup>28</sup>. In this study, opposing directional expression effects compared with monocytes have been observed. Carriers of allele G show elevated *IL18RAP* expression levels compared with carriers of allele A. This further exemplifies that gene regulation represents a complex cell-type- and context-specific process. Whether cell-type-specific eQTLs are functionally more susceptible to environmental influences, such as exogenous cell stimulation, represents an interesting hypothesis. However, future



**Figure 5 | rs968451 and rs917997 represent the strongest GWAS-iQTL during TLR4 signalling.** (a) *PDGFB* expression is allele-specifically induced on LPS exposure, strongest for carriers of allele G. Violin plots are depicted as in Fig. 2. (b) Regional association plot from stimulated monocytes shows that rs968451 regulates *PDGFB* expression selectively (red dot). In particular, *MAP3K7IP1* expression is not allele-specifically influenced by rs968451 (green dot). (c) In homozygous A carriers, *IL18R1* expression is induced on LPS exposure and reduced in G carriers. Violin plots are depicted as in Fig. 2. (d) Regional association plot from stimulated monocytes shows that rs917997 regulates *IL18R1* expression strongest (red dot). In particular the expression of genes belonging to the interleukin receptor cluster are not allele-specifically influenced by rs917997, although *IL18RAP* shows a tendency towards iQTL-regulation ( $P_{LPS} = 1.8 \times 10^{-4}$ , green dot).

studies could test this hypothesis, as eQTL data for large numbers of different cell types become increasingly available.

Recently two studies identified eQTLs under LPS stimulation in primary human monocytes<sup>29</sup> and dendritic cells<sup>30</sup>. Comparison of our study with these studies showed a significant overlap between our data and those of Fairfax *et al.*<sup>29</sup> In all, 43 and 44% of our monocyte data under baseline and LPS stimulation, respectively, were found to be genome-wide significant in Fairfax *et al.* (Supplementary Fig. 9a) and 55 (baseline) and 56% (LPS) of the data of Fairfax *et al.* showed evidence of replication in our data set (Supplementary Fig. 9b). Lee *et al.*<sup>30</sup> used a targeted approach (415-gene signature) to identify LPS-responsive eQTLs after 5-h LPS treatment in dendritic cells and showed little overlap with both monocytes works, which, however, can be explained by differences in experimental design (Supplementary Fig. 9a and b). To discover further immune-relevant eQTLs that might have been missed due to limited power, we performed a replication analysis using data from Fairfax *et al.*<sup>29</sup> This led to the identification of 672 *cis* and 237 *trans* additional iQTLs unique to TLR4 stimulation under absolute expression and 222 *cis* and 79 *trans* additional iQTLs under differential expression (Supplementary Fig. 9c and Supplementary Data 3–5). Taken together, this demonstrates the robustness of identifying immune-relevant eQTLs on LPS stimulation.

In summary, eQTL studies in the context of cellular LPS exposure provide comprehensive insights into the genetic regulation of innate immunity. Among these, functional SNPs for the expression of *NEU4*, *CCL14*, *CBX3* and *IRF5* contributed strongest to interindividual differences in immune response. Furthermore, we found that risk variants for PBC, IBD and celiac disease exclusively regulate the expression of *PDGFB* and *IL18R1* during immune response. Both genes are plausible candidates to study the disease pathophysiology in context of immune activation. Of course, follow-up studies on the biological function of the here identified eQTLs will be essential to gain further insights into the mechanisms of genetic variants regulating innate immune responses.

Our study represents an innovative approach to gain insights into gene–environmental interactions. The use of exogenous cell stimulation followed by eQTL identification might contribute to the elucidation of the physiology and pathophysiology of many cellular processes and diseases.

## Methods

**Sample collection, isolation and stimulation of CD14 + monocytes.** In total, 185 healthy male volunteers of German descent were recruited. The study was approved by the institutional review board of the University of Bonn. All volunteers signed informed consent and were between age 18 and 35 (mean 24). PBMC were obtained by Ficoll-Hypaque density gradient centrifugation of heparinized

blood. Monocytes were isolated with anti-CD14 para-magnetic beads according to the manufacturer's manual (MACS, Miltenyi Biotec). The purity of isolated monocytes was  $\geq 95\%$ . RPMI 1640 (Biochrom) supplemented with 10% heat-inactivated FCS (Invitrogen), 1.5 mM L-glutamine, 100 U ml<sup>-1</sup> penicillin, 100  $\mu$ g ml<sup>-1</sup> streptomycin (all Sigma-Aldrich) and 10 ng ml<sup>-1</sup> GM-CSF (ImmunoTools) was used to culture cells in 96-well round bottom wells at a density of 250,000 cells per well in 100  $\mu$ l overnight. Cell survival after overnight incubation was  $> 85\%$ . Cells of each volunteer were either untreated or treated with 200 ng ml<sup>-1</sup> ultrapure LPS from *Escherichia coli* (Invivogen). After 90 min cells were lysed in RLT reagent (Qiagen) and stored at  $-80^\circ\text{C}$ . C-reactive protein (CRP) levels were measured to exclude samples with elevated CRP levels. After stringent quality control (Non-smoker, no infection or vaccination 4 weeks before blood withdrawal, CRP  $< 2.5$  mg dl<sup>-1</sup>, monocyte purity  $\geq 95\%$ , monocyte survival  $> 85\%$ ), final samples from 137 individuals were further processed.

**RNA extraction.** RNA was extracted from the lysed cells using the AllPrep 96 DNA/RNA Kit from Qiagen. RNA concentrations were determined using NanoDrop (PeqLab) and a subset of samples was additionally checked for degradation in a Bioanalyser (Agilent Technologies).

**Gene expression analysis.** The Illumina TotalPrep-96 RNA Amplification Kit (Life Technologies) was used for the amplification and biotinylation of the RNA. Subsequent array-based gene expression analysis was performed on Illumina's Human HT-12 v4 Expression BeadChips (Illumina) comprising 47,231 probes. Expression profiles were normalized using function vsn2 implemented in the Bioconductor package 'vsn'<sup>31</sup>. Only probes that showed a  $P_{\text{detection}} < 0.01$  in more than 5% of at least one of the two conditions (baseline, LPS) were analysed. Probes that matched to multiple positions in the human genome and probes mapping to non-autosomal chromosomes were excluded from further analysis. Finally, probes were removed if SNPs within a probe showed an eQTL effect to the respective gene, resulting in 17,684 probes for statistical analyses.

**DNA extraction.** Genomic DNA was extracted from 10 ml blood using the Che-magic Magnetic Separation Module I (PerkinElmer Chemagen) according to the manufacturer's instructions. DNA was quantified by NanoDrop (PeqLab).

**DNA genotyping.** Genotyping was conducted on the Illumina's HumanOmniExpress BeadChips comprising 730,525 SNPs. After quality control ( $P_{\text{HWE}} > 10^{-5}$ , call rate  $> 98\%$ , MAF  $> 5\%$ ), a total of 579,090 SNPs were available for analysis. Samples showing potential admixture within the multi-dimensional scaling analysis were removed. All samples showed a call rate  $> 99\%$ .

**SNP imputation of GWAS findings.** SNPs reported in previous GWAS and not present on the HumanOmniExpress BeadChips were imputed. Imputation was performed using IMPUTE2 (ref. 32) based on 1,000 Genomes Pilot I (European) as a reference panel.

**Pathway analysis.** Data were evaluated using Ingenuity Pathway Analysis (IPA; Ingenuity Systems, Redwood City, CA, USA). IPA maps each gene within a global molecular network developed from information contained in the Ingenuity Pathways Knowledge Base. Gene networks were generated algorithmically based on their connectivity in terms of expression, activation and transcription. A 'network' in IPA is defined as a graphical representation of the molecular relationships between genes, which are represented as nodes, and the biological relationship between nodes is shown by a connecting line. All connections are supported by published data stored in the Ingenuity Pathways Knowledge Base and/or PubMed<sup>33</sup>.

Pathway enrichment analysis of differentially expressed genes, clusters of co-correlated genes and eQTL genes was performed using InnateDB<sup>34</sup>.

**eQTL-analysis.** Association tests were corrected for the top 10 multi-dimensional scaling components. As phenotypes, we mapped absolute expression values of untreated ( $P_{\text{Baseline}}$ ), LPS-treated ( $P_{\text{LPS}}$ ) and the differential expression of LPS-treated monocytes ( $P_{\text{ALPS/Baseline}}$ ) separately. Association between SNP genotypes and gene expression levels was examined by using a linear regression model of the phenotype versus minor allele dosage of the genotype (additive model). All possible SNP/phenotype combinations were tested exhaustively using PLINK v1.07 (ref. 35), downstream analyses were carried out using R with appropriate packages and in-house codes. *cis*-acting eQTLs were defined as SNPs located within 1-Mb interval on either side of a transcript to differentiate them from *trans*-acting eQTLs.

MANOVA implemented in R was used to map shared eQTLs between treated and untreated monocytes, which is equivalent to the Bayesian formulation given in a previous study<sup>7</sup>.  $P_{\text{LPS}}$  and  $P_{\text{Baseline}}$  from standard univariate analysis were compared with  $P_{\text{Manova}}$ . This revealed 3,154 genome-wide significant eQTLs (corresponding to 690 genes) at  $P_{\text{Manova}} < 1.76 \times 10^{-09}$  with an FDR of 0.01 where  $P_{\text{Manova}} < \min(P_{\text{LPS}}, P_{\text{Baseline}})$  and  $P_{\text{Baseline}} < 8.67 \times 10^{-09}$  and  $P_{\text{LPS}} < 7.35 \times 10^{-09}$ .

**Replication of eQTLs.** The study data set was compared with a recently published eQTL data set under LPS stimulation<sup>29</sup>, which used the same Illumina HumanHT-12 expression and HumanOmniExpress genotyping platform. We conducted a replication analysis of peak *cis* eQTLs (6,509 eQTLs under baseline and 6,513 eQTLs on 2-h LPS stimulation) from Fairfax *et al.*<sup>29</sup> This revealed (i) 1,731 additional *cis* eQTLs under baseline condition (corresponding to 1,092 genes;  $P_{\text{Baseline,cis}} < 4.15 \times 10^{-03}$ , FDR of 0.01), (ii) 1,659 *cis* eQTLs on LPS stimulation (corresponding to 1,083 genes;  $P_{\text{LPS,cis}} < 3.87 \times 10^{-03}$ , FDR of 0.01) and (iii) 222 *cis* iQTLs when fold change on LPS treatment is considered (corresponding to 161 genes;  $P_{\text{ALPS/Baseline,cis}} < 3.5 \times 10^{-04}$ , FDR of 0.01). Testing *trans* eQTLs (4,354 eQTLs under baseline and 2,359 eQTLs on 2-h LPS stimulation) of the same study revealed (iv) 1,671 additional *trans* eQTLs under baseline condition (corresponding to 137 genes;  $P_{\text{Baseline,trans}} < 5.82 \times 10^{-03}$ , FDR of 0.01), (v) 1,050 *trans* eQTLs on LPS stimulation (corresponding to 109 genes;  $P_{\text{LPS,trans}} < 7.32 \times 10^{-03}$ , FDR of 0.01) and (vi) 79 *trans* iQTLs when fold change on LPS treatment is considered (corresponding to 14 genes;  $P_{\text{ALPS/Baseline,trans}} < 1.71 \times 10^{-04}$ , FDR of 0.01). Replication analysis of  $P_{\text{Baseline}}$  and  $P_{\text{LPS}}$  were conducted on a per-condition basis with identical effect directions, whereas for  $P_{\text{ALPS/Baseline}}$  the union of eQTLs under baseline and LPS treatment was used. For the latter, we checked for identical effect directions in both the treated and untreated conditions.

## References

- Medzhitov, R. Recognition of microorganisms and activation of the immune response. *Nature* **449**, 819–826 (2007).
- Kono, H. & Rock, K. L. How dying cells alert the immune system to danger. *Nat. Rev. Immunol.* **8**, 279–289 (2008).
- Medzhitov, R. Origin and physiological roles of inflammation. *Nature* **454**, 428–435 (2008).
- Casanova, J. L., Abel, L. & Quintana-Murci, L. Human TLRs and IL-1Rs in host defense: natural insights from evolutionary, epidemiological, and clinical genetics. *Annu. Rev. Immunol.* **29**, 447–491 (2011).
- Netea, M. G., Wijmenga, C. & O'Neill, L. A. Genetic variation in Toll-like receptors and disease susceptibility. *Nat. Immunol.* **13**, 535–542 (2012).
- Garber, M. *et al.* A high-throughput chromatin immunoprecipitation approach reveals principles of dynamic gene regulation in mammals. *Mol. Cell.* **47**, 810–822 (2012).
- Stephens, M. A unified framework for association analysis with multiple related phenotypes. *PLoS ONE* **8**, e65245 (2013).
- Colbert, J. D., Matthews, S. P., Miller, G. & Watts, C. Diverse regulatory roles for lysosomal proteases in the immune response. *Eur. J. Immunol.* **39**, 2955–2965 (2009).
- Delgado, M. *et al.* Autophagy and pattern recognition receptors in innate immunity. *Immunol. Rev.* **227**, 189–202 (2009).
- Kelm, S. & Schauer, R. Sialic acids in molecular and cellular interactions. *Int. Rev. Cytol.* **175**, 137–240 (1997).
- Shezhetsky, A. V. & Hinek, A. Where catabolism meets signalling: neuraminidase 1 as a modulator of cell receptors. *Glycoconj. J.* **28**, 441–452 (2011).
- Finlay, T. M. *et al.* Thymoquinone-induced Neu4 sialidase activates NFkappaB in macrophage cells and pro-inflammatory cytokines in vivo. *Glycoconj. J.* **27**, 583–600 (2010).
- ENCODE Project Consortium. A user's guide to the encyclopedia of DNA elements (ENCODE) PLoS Biol. **9**, e1001046 (2011).
- Forssmann, U., Magert, H. J., Adermann, K., Escher, S. E. & Forssmann, W. G. Hemofiltrate CC chemokines with unique biochemical properties: HCC-1/CCL14a and HCC-2/CCL15. *J. Leukoc. Biol.* **70**, 357–366 (2001).
- Lomber, G., Wallrath, L. & Urrutia, R. The heterochromatin protein 1 family. *Genome. Biol.* **7**, 228 (2006).
- Smallwood, A. *et al.* CBX3 regulates efficient RNA processing genome-wide. *Genome Res.* **22**, 1426–1436 (2012).
- O'Neill, L. A. How Toll-like receptors signal: what we know and what we don't know. *Curr. Opin. Immunol.* **18**, 3–9 (2006).
- Stranger, B. E. *et al.* Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* **315**, 848–853 (2007).
- Westra, H. J. *et al.* Systematic identification of *trans* eQTLs as putative drivers of known disease associations. *Nat. Genet.* **45**, 1238–1243 (2013).
- Fairfax, B. P. *et al.* Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nat. Genet.* **44**, 502–510 (2012).
- Mells, G. F. *et al.* Genome-wide association study identifies 12 new susceptibility loci for primary biliary cirrhosis. *Nat. Genet.* **43**, 329–332 (2011).
- Harada, K. & Nakanuma, Y. Biliary innate immunity in the pathogenesis of biliary diseases. *Inflamm. Allergy Drug Targets* **9**, 83–90 (2010).
- Medina, J. *et al.* Evidence of angiogenesis in primary biliary cirrhosis: an immunohistochemical descriptive study. *J. Hepatol.* **42**, 124–131 (2005).
- Jostins, L. *et al.* Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491**, 119–124 (2012).

25. Dubois, P. C. *et al.* Multiple common variants for celiac disease influencing immune gene expression. *Nat. Genet.* **42**, 295–302 (2010).
26. Dinarello, C. A. Immunological and inflammatory functions of the interleukin-1 family. *Annu. Rev. Immunol.* **27**, 519–550 (2009).
27. Saleh, M. & Trinchieri, G. Innate immune mechanisms of colitis and colitis-associated colorectal cancer. *Nat. Rev. Immunol.* **11**, 9–20 (2011).
28. Hunt, K. A. *et al.* Newly identified genetic risk variants for celiac disease related to the immune response. *Nat. Genet.* **40**, 395–402 (2008).
29. Fairfax, B. P. *et al.* Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. *Science* **343**, 1246949 (2014).
30. Lee, M. N. *et al.* Common genetic variants modulate pathogen-sensing responses in human dendritic cells. *Science* **343**, 1246980 (2014).
31. Huber, W., von Heydebreck, A., Sultmann, H., Poustka, A. & Vingron, M. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics*. **18** (Suppl 1) S96–104 (2002).
32. Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5**, e1000529 (2009).
33. Calvano, S. E. *et al.* A network-based analysis of systemic inflammation in humans. *Nature* **437**, 1032–1037 (2005).
34. Lynn, D. J. *et al.* Curating the innate immunity interactome. *BMC Syst. Biol.* **4**, 117 (2010).
35. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).

## Acknowledgements

We thank all blood volunteers for participating to this study. We thank our laboratory technicians and colleagues responsible for database management. J.S. and V.H. received support for this work from the BONFOR research programme, individual grant O-149.0094. J.S. was supported by the NIH/DFG Research Career Transition Award. M.M.N. received support for this work from the Alfred Krupp von Bohlen und Halbach-

Stiftung. V.H. is supported by the European Research Council (ERC 243046). M.M.N. and V.H. are members of the DFG funded Excellence Cluster ImmunoSensation. P.B. was supported by the SFB/Transregio 57 (TP25 and Q1) and an individual DFG grant (BO 3755/1-1). M.N. is supported by National Institute for Health Research Biomedical Research Centre Funding to UCL/UCLH.

## Author contributions

S.Ki., J.B., G.H., V.H., J.S. initiated the study. S.Ki., J.B., B.P., M.M.N. S.Ki., A.H., H.F., J.R.W., B.M.-M., V.H., J.S. analysed and interpreted the data. S.Ki., J.B., M.B., V.K., E.B., N.F., T.H., P.B. performed the molecular genetic experiments. S.Ki., M.B., S.H., V.H., J.S. characterized the volunteers and collected blood samples. S.Ki., M.N., J.B., B.P., M.M.N., V.H., J.S. prepared the manuscript, with feedback from the other authors.

## Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Kim, S. *et al.* Characterizing the genetic basis of innate immune response in TLR4-activated human monocytes. *Nat. Commun.* 5:5236 doi: 10.1038/ncomms6236 (2014).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>