Association of CLEC16A with human common variable immunodeficiency disorder and role in murine B cells

Jin Li1*, Silje F. Jørgensen2,*, S. Melkorka Maggadottir1,3,*, Marina Bakay1, Klaus Warnatz4, Joseph Glessner1, Rahul Pandey1, Ulrich Salzer4, Reinhold E. Schmidt5, Elena Perez6, Elena Resnick7, Sigune Goldacker4, Mary Bucht4, Torsten Witte5, Leonid Padyukov8, Vibeke Videm9,10, Trine Folsøraas2,11, Faranaz Atschkezi3, James T. Elder12,13, Rajan P. Nair12, Juliane Winkelmann14,15,16,17, Christian Gieger18,19,20, Markus M. Nöther21,22, Carsten Büning23, Stephan Brand24, Kathleen E. Sullivan3,25, Jordan S. Orange26, Børre Føvings2,27, Stefan Schreiber28, Wolfgang Lieb29, Pål Aukrust2,11,31,#, Bodo Grimbacher4, Hakon Hakonarson1,32,33,#, Lennart Hammarström34,#& Eva Ellinghaus28,#, Vibeke Videm9,10, Trine Folseraas2,11, Faranaz Atschkezi3, James T. Elder12,13, Rajan P. Nair12, Juliane Winkelmann14,15,16,17, Christian Gieger18,19,20, Markus M. Nöther21,22, Carsten Büning23, Stephan Brand24, Kathleen E. Sullivan3,25, Jordan S. Orange26, Børre Føvings2,27, Stefan Schreiber28, Wolfgang Lieb29, Pål Aukrust2,11,31,#, Bodo Grimbacher4, Hakon Hakonarson1,32,33,#, Lennart Hammarström34,#& Eva Ellinghaus28,#

Common variable immunodeficiency disorder (CVID) is the most common symptomatic primary immunodeficiency in adults, characterized by B-cell abnormalities and inadequate antibody response. CVID patients have considerable autoimmune comorbidity and we therefore hypothesized that genetic susceptibility to CVID may overlap with autoimmune disorders. Here, in the largest genetic study performed in CVID to date, we compare 778 CVID cases with 10,999 controls across 123,127 single-nucleotide polymorphisms (SNPs) on the Immunochip. We identify the first non-HLA genome-wide significant risk locus at CLEC16A (rs17806056, P = 2.0 × 10⁻⁴⁵) and confirm the previously reported human leukocyte antigen (HLA) associations on chromosome 6p21 (rs1049225, P = 4.8 × 10⁻¹⁶). Clec16a knockdown (KD) mice showed reduced number of B cells and elevated IgM levels compared with controls, suggesting that CLEC16A may be involved in immune regulatory pathways of relevance to CVID. In conclusion, the CLEC16A associations in CVID represent the first robust evidence of non-HLA associations in this immunodeficiency condition.

1 Center for Applied Genomics, Children’s Hospital of Philadelphia, Abramson Research Center Suite 1216, 3615 Civic Center Boulevard, Philadelphia, Pennsylvania 19104, USA. 2 K.G. Jebsen Inflammation Research Centre, Division of Cancer Medicine, Surgery and Transplantation, Research Institute of Internal Medicine, Oslo University Hospital, Rikshospitalet, 0424 Oslo, Norway. 3 Division of Allergy and Immunology, The Children’s Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA. 4 Center for Chronic Immunodeficiency (CCI), University Medical Center Freiburg, University of Freiburg, 79106 Freiburg, Germany. 5 Clinic for Immunology and Rheumatology, Hannover Medical School, D-30625 Hannover, Germany. 6 Division of Pediatric Allergy and Immunology, University of Miami Miller School of Medicine, Miami, Florida 33136, USA. 7 Department of Medicine, Institute of Immunology, Mount Sinai School of Medicine, New York, New York 10029, USA. 8 Rheumatology Unit, Department of Medicine, Karolinska Institutet and Karolinska University Hospital Solna, SE-17176 Stockholm, Sweden. 9 Department of Laboratory Medicine, Children’s and Women’s Health, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway. 10 Department of Immunology and Transfusion Medicine, St Olav Hospital, NO-7006 Trondheim, Norway. 11 Norwegian PSC Research Center, Division of Cancer, Surgery and Transplantation, Oslo University Hospital, 0424 Oslo, Norway. 12 Department of Dermatology, University of Michigan, Ann Arbor, Michigan 48109, USA. 13 Ann Arbor Veterans Affairs Hospital, Ann Arbor, Michigan, 48105 USA. 14 Institute of Human Genetics, Helmholtz Center Munich, German Research Center for Environmental Health, 85764 Neuherberg, Germany. 15 Department of Neurology, MRI, Technische Universität München, 81675 München, Germany. 16 SyNergy Munich Cluster for Systems Neurology, Ismaninger Str. 22, 81675 München, Germany. 17 Stanford Center for Sleep Sciences and Medicine, 3165 Porter Drive, Palo Alto, CA 94301, USA. 18 Institute of Genetic Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, 85764 Neuherberg, Germany. 19 Research Unit of Molecular Epidemiology, German Research Center for Environmental Health, 85764 Neuherberg, Germany. 20 Institute of Epidemiology II, Helmholtz Zentrum München - German Research Center for Environmental Health, 85764 Neuherberg, Germany. 21 Institute of Human Genetics, University of Bonn, 53127 Bonn, Germany. 22 Department of Genomics, Life and Brain Center, University of Bonn, 53127 Bonn, Germany. 23 Department of Hepatology and Gastroenterology, Charité, Campus Mitte, 14168 Berlin, Germany. 24 University Hospital Munich-Grosshadern, Department of Medicine II, Marchioninistr. 15, D-81377 Munich, Germany. 25 Department of Pediatrics, The Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. 26 Section of Immunology, Allergy and Rheumatology, Department of Pediatric Medicine, Texas Children’s Hospital, Houston, Texas 77030, USA. 27 Section of Clinical Immunology and Infectious Diseases, Oslo University Hospital, Rikshospitalet, 0424 Oslo, Norway. 28 Institute of Clinical Molecular Biology, Christian-Albrechts-University, 24105 Kiel, Germany. 29 Institute of Epidemiology and Biobank Popgen, Christian-Albrechts-University of Kiel, 24105 Kiel, Germany. 30 Department of Clinical Immunology, Nuffield Department of Medicine, University of Oxford, Oxford OX3 9DU, UK. 31 Institute of Clinical Medicine, University of Oslo, 0424 Oslo, Norway. 32 Division of Human Genetics, The Children’s Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA. 33 The Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. 34 Division of Clinical Immunology and Transfusion Medicine, Department of Laboratory Medicine, Karolinska University Hospital, Huddinge, SE-14186 Stockholm, Sweden. * These authors contributed equally to this work. # These authors jointly supervised this work. Correspondence and requests for materials should be addressed to H.H. (email: hakonarson@email.chop.edu).
Common variable immunodeficiency (CVID) has a prevalence of ~1 in 25,000 in European populations. Recurrent bacterial respiratory tract infections constitute the predominant clinical manifestation, but a subset of CVID patients also develops gastrointestinal manifestations and lymphoid hyperplasia. In addition, various forms of autoimmune disorders affect up to 25% of the patients, of which the most common is autoimmune thrombocytopenia1. The immunological hallmark of CVID is the B-cell defect with inability to produce antibodies, monocyte/macrophage hyperactivity and signs of low-grade systemic inflammation2.

The focus of CVID genetics over the years has largely been to determine the presence of monogenic subtypes, leading to the identification of familial affection of a series of immunodeficiency genes, including CD19 (ref. 3), CD20 (ref. 4), CD81 (ref. 5), CR2 (ref. 6), ICOSL, LRBA, PLCG2 (ref. 9), PRKCD10 and TNFRSF13B11. However, most of the CVID cases are sporadic and although formal heritability estimates have not been made, a complex model of inheritance probably accounts for the majority of patients12. The human leukocyte antigen (HLA) haplotype association with CVID was found earlier by tissue typing13 and a recent genome-wide association study in CVID further supports a complex genetic heritability in CVID, with common variants within the HLA complex associating with disease development14. To what extent common genetic variations outside this region contribute to CVID susceptibility is unknown.

We hypothesize that the autoimmune comorbidity in CVID may occur on the basis of shared genetic susceptibility. We perform dense autoimmune risk loci genotyping on the Immunochip, using the Immunochip, United Kingdom and the United States (Table 1, Supplementary Fig. 1), with the Immunochip, 15. Following standard quality-control a targeted genotyping array with dense single-nucleotide poly-morphisms. In addition, various forms of autoimmune disorders affect up to 25% of the patients, of which the most common is autoimmune thrombocytopenia. The immunological hallmark of CVID is the B-cell defect with inability to produce adequate antibody responses, but patients also show other immunological abnormalities such as T-cell dysfunction, monocyte/macrophage hyperactivity and signs of low-grade systemic inflammation.

Table 1 | Overview of number of included patient and control panels before quality control, according to geography.

|          | Cases, n | Controls, n | Total, n |
|----------|----------|-------------|----------|
| Sweden   | 99       | 2096        | 2195     |
| Norway   | 112      | 1405        | 1517     |
| USA/UK   | 320      | 1405        | 1725     |
| Germany  | 351      | 6649        | 7000     |
| Total    | 886      | 11552       | 12438    |

Clec16 KD affects murine B cells. Whereas CLEC16A is highly expressed in human B cells, a role of CLEC16A in these cells has not been established. Given the paramount importance of B-cell defects in CVID development, we explored the biological implications of Clec16a KD in murine B cells isolated from splenocytes. We detected a 54.4% (± 8.4%) reduction of the total number of CD19 + B cells in the splenocyte population after induction of Clec16a KD using tamoxifen, when compared with non-tamoxifen-treated littermates (control group; P = 6.25 × 10^{-6} in two-sided T-test). In particular, the fraction of CD19 + B cells in splenocytes is reduced (17.4 ± 5.0%, P = 2.45 × 10^{-3} in two-sided T-test) (Fig. 2a and Supplementary Table 5). We did not detect significant differences in the fraction of CD27 + cells or for B-cell proliferation assaying (MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide)). However, B cells from Clec16-deficient mice exhibited an altered immunoglobulin (Ig) profile as compared with control mice with increased levels of IgM (P = 0.003 in two-sided T-test), but no changes in serum levels of IgG and IgA (Fig. 2b). In sum, the supplementary murine Clec16a data suggest an impact of CLEC16A on B-cell function.

Candidate gene association statistics. In addition to the genome-wide significant CLEC16A locus, we also found six loci exhibiting suggestive evidence for association (5 × 10^{-8} < P-value < 5 × 10^{-5}; Supplementary Fig. 5 and Supplementary Table 6). Although replication in future study panels is required...
for formally establishing these loci (FCRLA, EOMES, TNIP1, TNFAIP3, TNFSF11 and PTPN2) in CVID, prior probability for genuine association is enhanced by findings in other immune-mediated diseases (Supplementary Table 6). Furthermore, SNPs at four genes previously suggested to be involved in familial CVID subtypes, CR2, ICOS, MSH5 and TNFRSF13B, showed nominally significant association (P < 0.05) in the present analysis (Supplementary Table 7).

**HLA-DQB1 associations.** In line with previous studies[14] we detected strong associations with SNPs within the HLA complex on chromosome 6p21, peaking at rs1049225 in the 3' untranslated region of **HLA-DQB1** (P = 4.8 \times 10^{-16}; Table 2, Fig. 1a and Supplementary Fig. 6). By stepwise conditional logistic regressions, we found that the association signal consists of multiple independent effects (Supplementary Table 8 and Supplementary Fig. 7). We imputed classical HLA-A, HLA-C,

**Table 2 | Genome-wide significant (P < 5 \times 10^{-8}) associations detected by logistic regression analysis of 778 cases with CVID and 10,999 population controls.**

| SNP         | Chr. | Position (hg19) | A1/A2 | MAF cases/controls | OR (95% CI)    | P-value | Candidate gene | Additional SNPs
|-------------|------|----------------|-------|--------------------|----------------|---------|----------------|----------------|
| rs1049225   | 6p21 | 32,627,747     | A/G   | 0.16/0.26          | 0.56 (0.49, 0.64) | 4.8 \times 10^{-16} | HLA-DQB1 | 154            |
| rs17806056  | 16p13.13 | 11,192,499   | A/T   | 0.18/0.23          | 0.66 (0.57, 0.75) | 2.0 \times 10^{-9}  | CLEC16A | 21             |

A1, minor allele; A2, major allele; Chr, chromosome; CI, confidence interval; CVID, common variable immunodeficiency disorder; MAF, minor allele frequency; OR, odds ratio; SNP, single-nucleotide polymorphism.

*Most associated SNP from each locus.

†Number of additional genome-wide associated SNPs at the respective loci.

**Figure 1 | The association statistics for CVID and the relative expression level of CLEC16A in different genotype groups.** (a) A Manhattan plot of the Immunochip association statistics illustrating CVID susceptibility loci. SNP locations are plotted on the x axis according to their chromosomal position. The negative log10 of P-values per SNP derived from the association analysis are plotted on the y axis. The horizontal red line represents the genome-wide significance threshold of P = 5 \times 10^{-8}. (b) The regional association plot[48] for the CLEC16A locus. The most associated SNP (rs17806056) is indicated by the purple dot, while the colours of the remaining SNPs indicate the LD with the index SNP, as shown in the colour legend. The light blue line shows the recombination rates (HapMap project[49]) and genomic positions are from genome build hg19. The plot was generated using software LocusZoom[48]. (c) The relative expression level of CLEC16A compared between CVID cases of different genotypes at rs17806056. The mRNA level of CLEC16A was assessed by quantitative reverse transcriptase-PCR and normalized to glyceraldehyde 3-phosphate dehydrogenase (GAPDH) control. The relative fold change (y axis) was plotted against the genotype for SNP rs17806056 (x axis). Each blue dot represents the average value of three measurements from each individual sample and the black line through the dots represents the mean level among each genotype group. The number of samples in each group is n = 4 AA, n = 7 TA and n = 11 TT. P-value was determined by two-sided T-test.
Autoimmune disease associations at the CLEC16A locus.

Genome-wide association studies have detected CLEC16A associations in multiple phenotypes ranging from prototypical autoimmune disorders (for example, type 1 diabetes20 and primary biliary cirrhosis21) via immune-mediated conditions driven by exogenous antigens (for example, celiac disease22 and allergy23) to suggestive associations ($P = 1.8 \times 10^{-7}$ observed in selective IgA deficiency24. The majority of significant SNPs from these studies are trans-ethnic, localize to the same linkage disequilibrium (LD) block and are in high to moderate LD with the most strongly associated SNP within CLEC16A in the current study (rs17806056; Supplementary Table 11).

**Discussion**

The present study has demonstrated a potential role of CLEC16A in the pathogenesis of CVID by strong genetic associations and its involvement in murine B-cell function.

Because of the prominent CLEC16A SNP associations in a variety of auto-immune diseases, several studies have been carried out to examine the biological function of CLEC16A. Its *Drosophila* homologue Ema has been found to localize to the endosomal and Golgi membranes25,26. Ema mutants show defects in lysosomal degradation and protein trafficking27. It is also essential for autophagosomal growth and autophagy processes28, which are of major importance for proper immune regulation, including regulation of inflammasome activation. CLEC16A is evolutionarily conserved and may rescue the *Drosophila* Ema mutant phenotypes25,26. Human CLEC16A has been reported to be abundantly expressed in dendritic cells, natural killer cells and B cells6,17. Zouk *et al.*27 demonstrated, by using a human erythromyeloblastoid leukemia cell line, K562 cells, and human lymphoblastoid cell lines, that CLEC16A is localized with an endoplasmic reticulum marker, and that KD of CLEC16A did not affect T-cell co-stimulation. Recent murine data indicated that Clec16a is localized at endosomal membrane and forms a protein complex with Nrdp1, which is an E3 ubiquitin–protein ligase28. Clec16a may thus regulate mitophagy through the Nrdp1/Parkin pathway29. Adding to this existing knowledge, the present study has shown that CLEC16A may also be involved in B-cell function. Although we could not reproduce the complete CVID phenotype in Clec16a KD B cells, findings strongly suggest that CLEC16A, in combination with other dysregulated pathways, may contribute to the B-cell dysfunction observed in these patients.

The relationship of CLEC16A expression and its genotype status has been investigated. We have previously shown that CLEC16A is differentially expressed based on the risk allele, reported with the protective minor allele showing a higher expression than the risk allele20, and a similar observation was made in the present study. The correlation is further complicated by the existence of three CLEC16A isoforms, which could lead to isoform-specific correlations. Examining the relative expression level of the long isoforms versus short isoform, a correlation with multiple sclerosis-associated SNP rs12708716 genotype status has been observed in human thymic tissues. However, no association was found when it was examined in whole blood29. Therefore, the effect of SNP genotype status on CLEC16A expression could be both isoform specific and cell-type/tissue-type dependent30. Furthermore, the sample size in each study is small and may be underpowered to detect a correlation. At this stage, although CLEC16A is a promising candidate gene to function in the pathogenesis of CVID as demonstrated in our study, we could not fully exclude the possibility that CVID-associated SNPs in CLEC16A introns are tagging other proximal genes, although we find this is unlikely.

The identification of CLEC16A as a significant CVID risk locus reflects the notion that most genetic risk loci for immune-related diseases are pleiotropic. Significant overlap of risk loci between autoimmune disease and immune deficiency has also been found for those of rheumatoid arthritis31. The biological explanation of the overlap probably differs from locus to locus, but the overall interpretation is that of an imperfect relationship between clinical disease phenotyping and genetically determined pathophysiology. Importantly, the finding of CLEC16A associations in both autoimmunity, allergy and, as given by the present study, an immunodeficiency condition, characterized by impaired immune

**Figure 2 | The effect of inducible Clec16a KD in murine B cells. (a) Percentage of B cells (CD19+ cells) from splenocytes of inducible Clec16a KD mice after tamoxifen treatment and the littermates treated with oil as a control in two independent experiments by fluorescence-activated cell sorting analysis. Data from the two experiments were combined. Black and red dots represent data from control (Cnt) and Clec16a KD mice, respectively. The black lines through the dots represent the mean level among the ten control mice and nine Clec16a KD mice, respectively. Two-sided $t$-test was used to compare the percentage of CD19+ cells between Clec16a KD mice and control littermates. (b) Ig production from the supernatants of B cells, purified from Clec16a KD mice splenocytes, cultured with anti-mouse CD40 (100 ng ml$^{-1}$) for 6 days. Data from two independent experiments were combined. Data are mean ± s.d. of seven mice in each group. Two-sided $t$-test was used to compare the level of each Ig subtype produced from Clec16a KD and control littermates.

**Note:** All experiments were performed in compliance with the approved institutional guidelines.
response to certain capsulated bacteria, is particularly interesting, as the mechanisms leading to persistent immune activation and autoimmunity in CVID is not clear. Our findings may suggest that altered function of CLEC16A could be the 'missing link' between immunodeficiency and immune action in CVID. The current identification of novel, highly robust CLEC16A associations in CVID raises conceptually novel opportunities for exploring mechanisms underlying the autoimmune co-morbidity and the chronic inflammation observed in CVID. Hypothetically, the CLEC16A-associated aspects of adult immunodeficiency states may reveal novel concepts for the basis of immune activation in autoimmunity.

In conclusion, we have identified the first genome-wide significant non-HLA CVID risk locus at CLEC16A in the largest genetic study performed in this disease to date. CLEC16A has previously been linked to autoimmune disorders and the fact that Clec16a-deficient mice showed decreased number of B cells suggests that CLEC16A could represent a link between autoimmunity and immunodeficiency in CVID.

Methods

Subject Description. Study subjects were recruited from five countries: Sweden, Norway, the United States, the United Kingdom and Germany (Table 1 and Supplementary Table 1A). The diagnosis of CVID 37 was defined by decreased serum levels (≥ 2 s.d.) of IgG, IgA and/or IgM, and exclusion of other forms of hypogammaglobulinemia according to the World Health Organization expert group on primary immunodeficiency and the International Union of Immunological Societies. 32, 33. Controls were recruited from blood donors or population-based studies.

The Swedish CVID cases (n = 93) were recruited at the Department of Laboratory Medicine, Division of Clinical Immunology and Transfusion Medicine, Karolinska University Hospital, Huddinge, Stockholm. The Swedish controls (n = 2,096) were from a population based case–control study named Epidemiological Investigation of Rheumatoid Arthritis. 34, 35

The CVID patients in the Norwegian panel (n = 112) were recruited from the Section of Clinical Immunology and Infectious Diseases at Oslo University Hospital Rikshospitalet, Oslo, Norway. DNA samples from healthy Norwegian controls (n = 1,405) were selected from the Norwegian Bone Marrow Donor Register and the North-Trøndelag Health Study (HUNT).

The CVID patients in the United States/United Kingdom panel (n = 330) were recruited from four locations: the Immunodeficiency Clinic at Mount Sinai Medical Center, New York, NY, the Division of Allergy, Immunology and Rheumatology, All Children's Hospital, St Petersburg, FL, the Division of Allergy and Immunology at The Children's Hospital of Philadelphia, PA, and the Department of Clinical Immunology in the Nuffield Department of Medicine, Oxford Radcliffe Hospital, UK. All the US/UK cases were genotyped at the Center for Applied Genomics, Children's Hospital of Philadelphia, PA.

The US controls (n = 1,405) were recruited at the University of Michigan and genotyped at the Institute of Clinical Molecular Biology, Christian-Albrechts-University of Kiel (as part of the PAGE Immunochip data set 36).

The German CVID cases (n = 351) were recruited at the Center for Chronic Immunodeficiency, University Hospital of Freiburg, and at the Clinic for Immunology and Rheumatology, Hannover Medical School, Hannover. DNA from 2,696 German healthy controls was obtained through the Northern German biobank PopGen 36 (http://www.popgen.de) and the University Hospital Schleswig-Holstein. Genotyping of these 2,696 controls was performed at the Institute of Clinical Molecular Biology, Christian-Albrechts-University of Kiel.

One thousand nine hundred and twenty-four German controls were part of an independent population control sample from the general population living in the region of Augsburg (Cooperative Health Research in the Region of Augsburg), southern Germany 37, and were genotyped at the Helmholtz Center in Munich. One thousand and five hundred German controls were recruited from the population-based epidemiological Heinz-Nixdorf Recall study and genotyped at the Life and Brain Center at the University Clinic Bonn. Three hundred and fourteen individuals were of south German ancestry and were part of the control population recruited from the Bavarian Red Cross and 215 individuals were recruited from the Charité-Universitätsmedizin Berlin. These samples were genotyped at the University of Pittsburgh Genomics and Proteomics Core Laboratories.

Respective Regional Ethics Committees and/or controls to the study, all of which were approved: The regional ethical review board in Stockholm, The Regional Committees for Medical and Health Research Ethics South East (Norway) for Norwegian patients and The Regional Committee for Medical and Health Research Ethics Central (Norway) for the HUNT controls, Children's Hospital of Philadelphia Institutional Review Board, Mount Sinai School of Medicine Institutional Review Board, University of Oxford Institutional Review Board, University of South Florida Institutional Review Board, the University of Michigan Medical School Institutional Review Board, Ethikkommission der Universität Freiburg, Ethik-Kommission der Medizinische Hochschule Hannover, Ethik-Kommission der Medizinischen Fakultät der Christian-Albrechts-Universität zu Kiel, Ethics committee of the University Hospital S.-H., Campus Kiel (Kiel University), Ethics committee of the Faculty of Medicine, Ludwig-Maximilians-University Munich, Germany, Ethikkommission der Bayerischen Landesärztekammer for the KORA controls, Ethikkommission der Universität Duisburg-Essen for HNR controls, Ethikkommission der Charité-Universitätsmedizin Berlin for Berlin Charite controls. Written informed consent for blood sample collection, processing and genotyping were obtained by all participants.

Immunochip genotyping. DNA samples were genotyped using the Immunochip, an Illumina Select HD custom genotyping array. This BeadChip was developed for highly multiplexed SNP genotyping and the SNP content of this chip was mainly based on findings in ankylosing spondylitis, Crohn's disease, celiac disease, IgA deficiency, multiple sclerosis, primary biliary cirrhosis, psoriasis, rheumatoid arthritis, systemic lupus erythematosus, type 1 diabetes and ulcerative colitis. Genotyping was performed according to Illumina protocols, with 4 μl of genomic DNA samples at 50 ng μl−1, aliquoted to the corresponding wells of 96-well plates. The NCBI build 36 (hg19) map was used (Illumina manifest file Immuno_BeadChip_11419691_B.bpm) and normalized probe intensities were extracted for all samples passing standard laboratory QC thresholds. Genotype calling was performed with Illumina’s GenomeStudio data analysis software, the GenomeStudio GenTrain 2.0 algorithm and the cluster file generated by Trynka et al. 38 (based on the clustering of 2,000 UK samples and subsequent manual re-adjustment of cluster positions). All CVID cases from Norway, Sweden and Germany were genotyped at the Institute of Clinical Molecular Biology in Kiel, Germany. The US/UK cases were genotyped at the Center for Applied Genomics, the Children’s Hospital of Philadelphia, USA.

Quality control. Sample QC measures included sample call rate, overall heterozygosity, relatedness testing and other metrics: samples with a SNP call rate < 98% (n = 309) as well as heterozygosity outliers (n = 15), defined as beyond 5 s.d. of the mean, were excluded from analysis. Duplicate samples (n = 37) and cryptically related samples (PLHAT > 0.1875; n = 199) were identified through identity-by-state calculations with PLINK8. For each pair of duplicate or related samples, the sample with the highest SNP call rate was kept in the data set. Finally, removal of population outliers (non-Caucasians; n = 101) was performed based on principal component analysis of population heterogeneity, which is described below in detail as ‘Principal component analysis’.

In the SNP based QC SNPs with a call rate < 98%, with a minor allele frequency < 1% and markers that deviated from Hardy–Weinberg equilibrium test (P-value < 10−5) in the controls or SNPs with significant different genotyping rate between cases and controls (P < 1 × 10−5 in Fisher’s exact test) were removed. Applying the QC procedures described above resulted in 778 CVID cases and 10,999 controls available for association analysis.

Principal component analysis. Principal component analysis was conducted once to identify ethnic outliers and once again to generate covariates to control for population stratification. For the first analysis, we used HapMap samples as a reference set and 13,001 uncorrelated SNPs that passed the above QC criteria and are present in both our Immunochip data set and HapMap data set, which were LD pruned such that no pair of SNPs had r² > 0.2 and excluding problematic GC/AT SNPs and X- and Y-chromosomes. PLINK was used for LD pruning and exclusion of SNPs. All samples that did not cluster with the European samples were excluded. After removal of ethnicity outliers, second principal component analysis was performed within the remaining Immunochip samples to resolve within-Europe relationships. No population stratification in the remaining samples was observed. Principal component analyses were performed using EIGENSTRAT version 4.0.

Association analysis. For case–control association testing, logistic regression with PLINK was performed using the first three principal components from the EIGENSTRAT analysis as covariates, which are sufficient to control for population stratification as determined by the genomic inflation factor. Sub-phenotype analysis was similarly conducted via logistic regression. Step-wise conditional association analysis was performed by including the most strongly associated SNPs in the previous step as a covariate. Post-hoc power assessment showed over 99% power to detect genome-wide significant association at the HLA SNP rs1049225 (odds ratio (OR) ≥ 1.75) with 778 cases and 14-fold number of controls, and ~ 83% power for the CLEC16A SNP, rs17806056 (OR ≥ 1.5; Supplementary Fig. 8).

SNP imputation and association testing. For the CLEC16A locus on chromosome 16, haplotype pre-phasing was performed using SHAPEIT II v2. Genetic imputation was conducted using the IMPUTE2 (ref. 43) package. The 1,000 Genomes Phase I integrated variant set was used as reference panel and data were downloaded from the IMPUTE2 website (http://mathgen.stats.ox.ac.uk/ SNP imputation and association testing. For the CLEC16A locus on chromosome 16, haplotype pre-phasing was performed using SHAPEIT II v2. Genetic imputation was conducted using the IMPUTE2 (ref. 43) package. The 1,000 Genomes Phase I integrated variant set was used as reference panel and data were downloaded from the IMPUTE2 website (http://mathgen.stats.ox.ac.uk/
then crossed to the Flpo Deleter line (mouse Strain: 129S4/SvJae-Gt(ROSA)26-
the targeted insertion of the two loxP sites in the Clec16a gene locus were
cell-derived offspring were identified by PCR-based genotyping. Mice harbouring
pseudopregnant females. Male offspring were mated to C57BL/6 females and ES
were injected into C57BL/6-derived blastocysts that were then transferred to
screened for homologous recombination by Southern blot analysis. Targeted clones

Power analysis. Power analysis was conducted using software Power for
Genetic Association45, with the following settings: Genetic Model = Co-dominant
(1 df) SNP analysis; R² = 1; Disease prevalence = 0.00004; Marker allele
frequency = Disease allele frequency = 0.75; Effective degree of freedom = 123.127;
α = 0.05 and control to case ratio = 14. Relative risk was set at 1.5, 1.75 and 2,
to generate three power curves.

Animals. All animals were studied by the Institutional Animal Care and
Use Committee of the Children’s Hospital of Philadelphia. To generate Clec16aKO
mice, a 15.3-kb DNA fragment containing Clec16a exons 2–4 and flanking intronic
sequence was retrieved from C57BL/6 mouse genomic DNA and subcloned into
plasmid FLrSNiper (Ozgene) that contains an FRT-flanked PGK-driven Neomycin
cassette for negative selection. One loxP site was inserted upstream of exon 3, while a
second loxP site was placed downstream of exon 3. The linearized targeting
vector was electroporated into 86 ES cells and clones that survived selection were
screened for homologous recombination by Southern blot analysis. Targeted clones
were injected into C57BL/6-derived blastocysts that were then transferred to
pregnant females. Male offspring were mated to C57BL/6 females and ES
cell-derived offspring were identified by PCR-based genotyping. Mice harbouring
the targeted insertion of the two loxP sites in the Clec16a gene locus were
then crossed to the Flpo Deleter line (mouse strain: 129S4/SvJae-Gt(ROSA)26-
Sort2(FLP*+)) (The Jackson Laboratory) to achieve deletion of the FRT-
flanked Neomycin cassette. Clec16aKO mice were mated to UBC-Cre-ER-LBD-tg
mice (inducible Cre recombinase driven by the human ubiquitin C promoter),
to generate UBC-Cre-Clec16aKO mice.

To generate experimental groups, UBC-Cre-Clec16aKO male mice were treated
with tamoxifen (for Clec16a KD) or oil (control group). Tamoxifen (MP
Biomedical) was prepared at a concentration of 20 mg ml⁻¹ in 10% ethanol and
90% corn oil (Sigma). Four-week-old male mice received 1 mg of tamoxifen at 24-h
intervals for 5 consecutive days by gavage. Control group of mice were receiving an
equal volume of corn oil alone. After tamoxifen treatment was finished, mice were
aged an additional 2 weeks before evaluation.

Quantitative real-time PCR. The total RNA was extracted from human blood
using Trizol (Ambion, Life Technologies) and converted to complementary DNA
with High Capacity RNA-to-cDNA Kit, following the manufacturer’s protocols
(Applied Biosystems). Briefly, 1 ml of Trizol was mixed to 50 ml of whole blood,
followed by 200 ml of chloroform. The mixture was centrifuged at 12.000g for
15 min at 4°C. The aqueous phase (containing RNA) was purified with
RNAeasy kit (Qiagen). Quality of RNA was assessed by Agilent Bioanalyzer.
Fluorescence-based real-time PCR was performed in 10 µl of reaction mixture,
using qPCR TaqMan Universal Master Mix and 20 µM FAM-MGB TaqMan
assays (Applied Biosystems): Hs00323276_m1 (CLEC16A), Hs00362034_m1 (DEx)
Hs00705164_s1 (SOCS1) and Hs01275899_g1 (glyceraldehyde 3-phosphate
dehydrogenase). Relative gene expression was normalized to glyceraldehyde
3-phosphate dehydrogenase. The cycle threshold was used as an internal
control to maintain consistency across experimental and control samples.

Flow cytometry. Single-cell suspensions from murine spleens (10⁶ cells per test)
were stained with fluorochrome-conjugated monoclonal antibodies in a single
tube containing anti-mouse CD19-Allex Fluor 647 (Biolegend 115522), IgD-Allex
Fluor 488 (Biolegend, 405718), IgM-Brilliant Violet 421 (Biolegend 406517) and
CD27-PE (Biolegend 124209) antibodies (1:20 dilution) for 30 min at 4°C. Cell
-associated fluorescence was assessed with an LSR-II flow cytometer and
analyzed using software (BD FACSDiva). Data were analyzed using Flowjo
software v10.7 (Tree Star, Inc.). The effect size of the association was measured with OR and the
odds ratio interval for the OR was calculated directly from the 2 × 2 table using
the Woolf’s formula47.

HLA analyses. Conditional analysis did not demonstrate evidence for independent
association signals at the HLA-DQBI locus.

Imputation of classical alleles at HLA class I and II loci was performed using
SNP2HLA with a reference panel consisting of 5,225 unrelated individuals
collected by the Type 1 Diabetes Genetics Consortium46.

To further explore the HLA class I and HLA class II associations with CVID, HLA allele frequencies between cases and controls at a two-digit level were
analysed. The effect size of the association was measured with OR and
the odds ratio interval for the OR was calculated directly from the 2 × 2 table using
the Woolf’s formula47.

Western blot analysis. For western blot analysis, lysis of splenocytes was
performed with NP40 lysis buffer (Invitrogen). Proteins were separated on 4–12%
NuPAGE Bis-Tris gels in MOPS SDS running buffer and transferred overnight
onto nitrocellulose membranes (Invitrogen). The membranes were blocked in 3%
BSA and cut in half. The upper half of the membranes was incubated with rabbit
anti-CLEC16A polyclonal antibody (Abgent, AP6983c) at 1:1,000 dilution and
the lower half of the membranes was probed with mouse anti β-Actin monoclonal
antibody (Abcam, ab6276) at 1:1,000 dilution. The membranes were washed,
incubated with corresponding secondary antibody for 1 h and washed again;
bound antibody was detected with WesternBright ECL chemiluminescence detection
system (Advansta). Band intensities were measured using Image J software (NIH
Shared Resources). Representative western blot for Clec16a knockdown (Clec16a KD)
and control mice is shown in Supplementary Fig. 9.

References
1. Cunningham-Rundles, C. & Bodian, C. Common variable immunodeficiency:
clinical and immunological features of 248 patients. Clin. Immunol. (Orlando,
FL) 92, 34–48 (1999).
2. Aukrust, P. et al. Persistent activation of the tumor necrosis factor system in
a subgroup of patients with common variable immunodeficiency—possible
immunologic and clinical consequences. Blood 87, 671–681 (1996).
3. van Zelm, M. C. et al. An antibody-deficiency syndrome due to mutations in
the CD19 gene. N. Engl. J. Med. 354, 1901–1912 (2006).
4. Kuipers, T. W. et al. CD20 deficiency in humans results in impaired T
-cell-independent antibody responses. J. Clin. Invest. 120, 214–222 (2010).
5. van Zelm, M. C. et al. CD81 gene defect in humans disrupts CD19 complex
formation and leads to antibody deficiency. J. Clin. Invest. 120, 1265–1274
(2010).
6. Thié, J. et al. Genetic CD21 deficiency is associated with hypogamma-
globulinemia. J. Allergy Clin. Immunol. 129, 801–808 e806 (2012).
7. Grimbacher, B. et al. Homozygous loss of ICOS is associated with adult-onset
common variable immunodeficiency. Nat. Immunol. 4, 261–268 (2003).
8. Lopez-Herrera, G. et al. Deleterious mutations in LRBA are associated with a
syndrome of immune deficiency and autoimmunity. Am. J. Hum. Genet. 90,
986–1001 (2012).
9. Zhou, Q. et al. A hypermorphic missense mutation in PCLG2, encoding
phospholipase Gammaz2, causes a dominantly inherited autoinflammatory
disease with immunodeficiency. Am. J. Hum. Genet. 91, 713–720 (2012).
10. Salzer, E. et al. B-cell deficiency and severe autoimmunity caused by deficiency
of protein kinase C delta. Blood 121, 3112–3116 (2013).
11. Salzer, U. et al. Mutations in TNRF513 encoding TACI are associated with
common variable immunodeficiency in humans. Nat. Genet. 37, 820–828
(2005).
12. Vorochkovsky, I. et al. Family and linkage study of selective IgA deficiency
and common variable immunodeficiency. Clin. Immunol. Immunopathol. 77,
185–192 (1995).
13. Olruper, O., Smith, C. L., Bjorkander, J. & Hammarstrom, L. Shared HLA class
II-associated genetic susceptibility and resistance, related to the HLA-DQBI
gene, in IgA deficiency and common variable immunodeficiency. Proc. Natl
Acad. Sci. U.S.A 89, 10653–10657 (1992).
14. Orange, I. S. et al. Genome-wide association identifies diverse causes of
common variable immunodeficiency. J. Allergy Clin. Immunol. 127, 1360–1367
e1366 (2011).
15. Parkes, M. et al. Mutations in TNFRSF13B encoding TACI are associated with
HLA genotypes and antibodies to the HLA-DQBI locus.
17. Wu, C. et al. BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. Genome Biol. 10, R130 (2009).

18. Consortium, E. P. et al. An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57–74 (2012).

19. Camargo, J. F. et al. MonoMAC syndrome in a patient with a GATA2 mutation: case report and review of the literature. Clin. Infect. Dis. 57, 697–699 (2013).

20. Hakonarson, H. et al. A genome-wide association study identifies KIAA0350 as a type 1 diabetes gene. Nature 448, 591–594 (2007).

21. Mells, G. F. et al. Genome-wide association study identifies 12 new susceptibility loci for primary biliary cirrhosis. Nat. Genet. 43, 329–332 (2011).

22. Dubois, P. C. et al. Multiple common variants for celiac disease influencing immune gene expression. Nat. Genet. 42, 293–302 (2010).

23. Hinds, D. A. et al. A genome-wide association meta-analysis of self-reported allergy identifies shared and allergy-specific susceptibility loci. Nat. Genet. 45, 907–911 (2013).

24. Ferreira, R. C. et al. Association of IFIH1 and the PDRN10 gene with type 1 diabetes mellitus. Nat. Genet. 42, 777–780 (2010).

25. Kim, S., Naylor, S. A. & DiAntonio, A. The novel endosomal membrane protein Ema interacts with the class C Vps-HOPS complex to promote endosomal maturation. J. Cell Biol. 188, 717–734 (2010).

26. Kim, S., Naylor, S. A. & DiAntonio, A. Drosophila Golgi membrane protein Ema promotes autophagosomal growth and function. Proc. Natl Acad. Sci. USA 109, E1072–E1081 (2012).

27. Zou, H. et al. Functional evaluation of the role of CLEC16A at the chromosome 16p13 locus. Clin. Exp. Immunol. 12, 191–198 (2011).

28. Solemanpour, S. A. et al. The diabetes susceptibility gene Clec16a regulates mitophagy. Cell 157, 1577–1590 (2014).

29. Mero, I. L. et al. Exploring the CLEC16A gene reveals a M5-associated variant with correlation to the relative expression of CLEC16A isoforms in thymus. Genes Immun. 15, 149–155 (2014).

30. Berge, T., Leikfoss, I. & Harbo, H. From identification to characterization of the gene CLEC16A. Int. J. Mol. Sci. 14, 4476–4497 (2013).

31. Cooper, J. D. et al. Meta-analysis of genome-wide association study data identifies additional type 1 diabetes risk loci. Nat. Genet. 40, 1399–1401 (2008).

32. International Union of Immunological Societies. Primary immunodeficiency diseases. Report of an IUIS Scientific Committee. Clin. Exp. Immunol. 118 Suppl 1: 1–28 (1999).

33. Al-Herz, W. et al. Primary immunodeficiency diseases: an update on the classification from the international union of immunological societies expert committee for primary immunodeficiency. Front. Immunol. 5, 162 (2014).

34. Stolt, P. et al. Quantification of the influence of cigarette smoking on rheumatoid arthritis: results from a population based case-control study, using incident cases. Ann. Rheum. Dis. 62, 835–841 (2003).

35. Tooki, T., C.-E. et al. Identification of 15 new psoriasis susceptibility loci highlights the role of innate immunity. Nat. Genet. 44, 1341–1348 (2012).

36. Krawczak, M. et al. PopGen: population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships. Community Genet. 9, 55–61 (2006).

37. Wichmann, H. E., Gieger, C. & Illig, T. KORA-gen--resource for population-based linkage analyses. Am. J. Epidemiol. 169, 621–629 (2009).

38. Trynka, G. et al. Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. Nat. Genet. 47, 1183–1193 (2015).

39. 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).

40. Price, A. L. et al. Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. 38, 904–909 (2006).

41. Delaneau, O., Marchini, J. & Zagury, J. F. A linear complexity phasing method for thousands of genomes. Nat. Methods 9, 179–181 (2012).

42. Delaneau, O., Zagury, J. F. & Marchini, J. Improved whole-chromosome phasing for disease and population genetic studies. Nat. Methods 10, 5–6 (2013).

43. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat. Genet. 39, 906–913 (2007).

44. 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).

45. Menashe, I., Rosenberg, P. S. & Chen, B. E. PGA: power calculator for case-control genetic association analyses. BMC Genet. 9, 36 (2008).

46. Jia, X. et al. Imputing amino acid polymorphisms in human leukocyte antigens. PLoS One 8, e66833 (2013).

47. Kirkwood, B. & Sterne, J. Essential Medical Statistics 2nd edn (Blackwell Science, Oxford, UK, 2003).

48. Prum, R. J. et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics (Oxford, England) 26, 2336–2337 (2010).

49. International HapMap C. The International HapMap Project. Nature 426, 789–796 (2003).

Acknowledgements

We thank David Ellenghaus, Kristian Holm and Johannes Rokmund How for helpful discussions and analytical support, and Liv Oxsnes for analysing the subphenotype flow data for the Norwegian panel. Benedicte A. Lie and The Norwegian Bone Marrow Donor Registry at Oslo University Hospital, Rikshospitalet in Oslo, as well as Matthew A. Brown and Kristian Hveem are acknowledged for sharing the healthy Norwegian control data. Markus M. Nothen is a member of the DFG Excellence Cluster ‘Immunoscenstation’. This study was supported by the Southern and Eastern Norway Regional Health Authority, the German Federal Ministry of Education and Research (BMBF) grants 01EO1003 (CCG) and 01GM0896 (PID-NET), and the DZIF project T04.082. This work was further supported by EU grant HEALTH-F2-2008-201549 (EURO-PADnet), an Institute Development Fund from CHOP, U01HG068380, DP3 DK085708 and a donation to CAG from the Kubert Estate Foundation. The popgen 2.0 network provided the German control data and is supported by a grant from the German Ministry for Education and Research (BMBF, ID 01EY1103). Genotyping of the European case–control collection was supported by the DFG Excellence Cluster ‘Imflammation at Interfaces’ (EXC306). This work was supported by the German Federal Ministry of Education and Research (BMBF) within the framework of the e:Med research and funding concept 01ZX1036. S. Brand was supported by grants from the Deutsche Forschungsgemeinschaft (DFG, BR 19126/1) and the Else Kröner-Fresenius-Stiftung (Else Kröner-Fresenius-Stiftung 2010_EKES32). The KORA research platform was initiated and financed by the Helmholtz Zentrum München - German Research Center for Environmental Health, funded by the German Federal Ministry of Education and Research and by the State of Bavaria. KORA research was supported within the Munich Center of Health Sciences (MC Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ. The authors are responsible for the contents of this publication.

Author contributions

J.L., S.F.J., M.M., M.B., J.G., R.P. and E.E. performed data and statistical analysis. K.W., J.G., U.S., R.E.S., E.R., S.G., M.B., T.W., I.P., V.V., T.F., F.A., J.T.E., R.P.N., J.W., T.F., J.M., M.M., C.B., S.B., K.S., J.S.O., B.F., S.S., W.L., P.A., H.C., C.C.R., A.F., T.H.K., B.G., H.H. and H.H. contributed to the ascertainment of affected individuals, and/or sample and clinical data collection. T.H.K., H.H., L.H. and E.E. coordinated and supervised the project. J.L., S.F.J., M.M., P.A., T.H.K. and H.H. drafted the manuscript. All authors read and approved the final version.

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: Li, J. et al. Association of CLEC16A with human common variable immunodeficiency disorder and role in murine B cells. Nat. Commun. 6:6804 doi: 10.1038/ncomms7804 (2015).

© 2015 Macmillan Publishers Limited. All rights reserved.