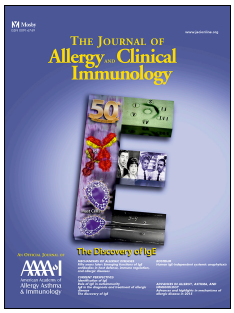
Accepted Manuscript



Loss of function NFKB1 variants are the most common monogenic cause of CVID in Europeans

Paul Tuijnenburg, MD, Hana Lango Allen, PhD, Siobhan O. Burns, MRCP PhD, Daniel Greene, MPhil, Machiel H. Jansen, BSc, Emily Staples, MRCP FRCPath PhD, Jonathan Stephens, PhD, Keren J. Carss, PhD, Daniele Biasci, PhD, Helen Baxendale, MRCP PhD, Moira Thomas, MRCP FRCPath BSc, Anita Chandra, MRCP FRCPath PhD, Sorena Kiani-Alikhan, FRCP FRCPath PhD, Hilary J. Longhurst, MD PhD, Suranjith L. Seneviratne, MD, Eric Oksenhendler, MD, Ilenia Simeoni, PhD, Godelieve J. de Bree, MD PhD, Anton T.J. Tool, PhD, Ester M.M. van Leeuwen, PhD, Eduard H.T.M. Ebberink, MSc, Alexander B. Meijer, PhD, Salih Tuna, PhD, Deborah Whitehorn, BSc, Matthew Brown, BSc, Ernest Turro, PhD, Adrian J. Thrasher, MD PhD, Kenneth G.C. Smith, PhD FMedSci, James E. Thaventhiran, MRCP FRCPath PhD, Taco W. Kuijpers, MD PhD

PII: S0091-6749(18)30286-0

DOI: 10.1016/j.jaci.2018.01.039

Reference: YMAI 13302

To appear in: *Journal of Allergy and Clinical Immunology*

Received Date: 13 April 2017

Revised Date: 15 December 2017

Accepted Date: 3 January 2018

Please cite this article as: Tuijnenburg P, Lango Allen H, Burns SO, Greene D, Jansen MH, Staples E, Stephens J, Carss KJ, Biasci D, Baxendale H, Thomas M, Chandra A, Kiani-Alikhan S, Longhurst HJ, Seneviratne SL, Oksenhendler E, Simeoni I, de Bree GJ, Tool ATJ, van Leeuwen EMM, Ebberink EHTM, Meijer AB, Tuna S, Whitehorn D, Brown M, Turro E, Thrasher AJ, Smith KGC, Thaventhiran JE, Kuijpers TW, on behalf of the NIHR-BioResource – Rare Diseases Consortium, Loss of function NFKB1 variants are the most common monogenic cause of CVID in Europeans, *Journal of Allergy and Clinical Immunology* (2018), doi: 10.1016/j.jaci.2018.01.039.

Whole Genome Sequencing:

Heterozygous pathogenic *NFKB1* variants in **4%** of European CVID cohort (*n* = 390)

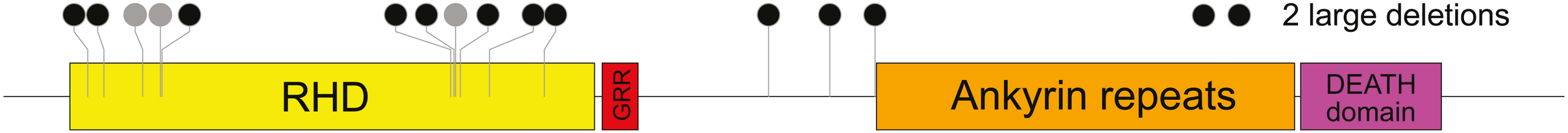
c.160-1G>A
Glu63Lysfs*64
Ile87Ser
Val98Asp
Gln99*

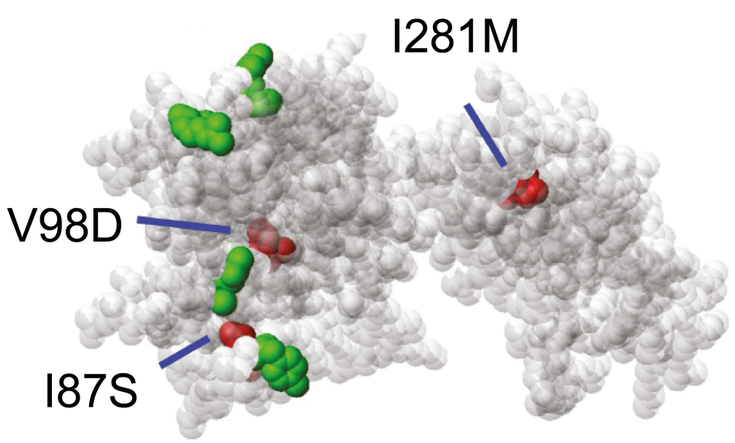
Lys278Glufs*3
c.835+2T>C
Ile281Met
Arg284*
Ser302Phefs*7
Arg336Glyfs*96

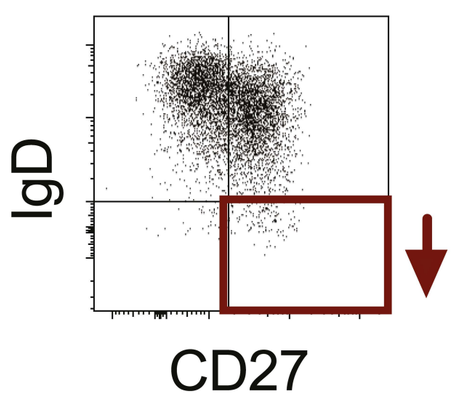
Ala475Profs*10
His513Glnfs*28
Asp541*

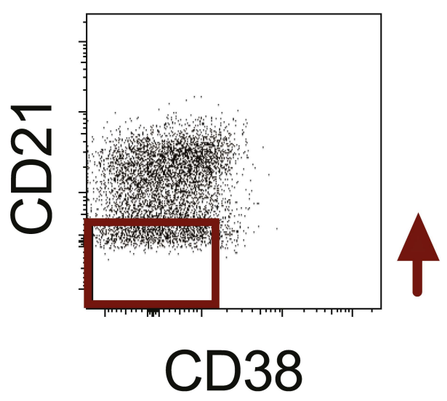
2 large deletions

CVID:	Common Variable Immunodeficiency
GRR:	Glycine-rich region
RHD:	Rel homology domain









Sinopulmonary infections

Autoimmune disease

Splenomegaly

Malignancy

Title page

Original Article

Loss of function NFKB1 variants are the most common monogenic cause of CVID in Europeans

Paul Tuijnenburg MD,[1,2]* Hana Lango Allen PhD,[3,4]* Siobhan O. Burns MRCP PhD,[5]* Daniel Greene MPhil,[3,4] Machiel H. Jansen BSc,[1,2] Emily Staples MRCP FRCPath PhD,[6] Jonathan Stephens PhD,[3,4] Keren J. Carss PhD,[3,4] Daniele Biasci PhD,[6] Helen Baxendale MRCP PhD,[6] Moira Thomas MRCP FRCPath BSc,[7] Anita Chandra MRCP FRCPath PhD,[6] Sorena Kiani-Alikhan FRCP FRCPath PhD,[8] Hilary J. Longhurst MD PhD,[9] Suranjith L. Seneviratne MD,[5] Eric Oksenhendler MD,[10] Ilenia Simeoni PhD,[3] Godelieve J. de Bree MD PhD,[11] Anton T.J. Tool PhD,[12] Ester M.M. van Leeuwen PhD,[2] Eduard H.T.M. Ebberink MSc,[13] Alexander B. Meijer PhD,[13] Salih Tuna PhD,[3,4] Deborah Whitehorn BSc,[3,4] Matthew Brown BSc,[3,4] Ernest Turro PhD,[3,4] Adrian J. Thrasher MD PhD,[14] Kenneth G. C. Smith PhD FMedSci,[6] James E. Thaventhiran MRCP FRCPath PhD,[6]** Taco W. Kuijpers MD PhD[1,2,11]**, on behalf of the NIHR-BioResource – Rare Diseases Consortium[15].

1 Department of Pediatric Hematology, Immunology and Infectious Diseases, Emma Children's Hospital, Academic Medical Center, Amsterdam, The Netherlands
2 Department of Experimental Immunology, Academic Medical Center, Amsterdam, The Netherlands
3 Department of Haematology, University of Cambridge, Cambridge, United Kingdom
4 NHS Blood and Transplant Cambridge, Cambridge Biomedical Campus, Cambridge, United Kingdom
5 Department of Immunology, Royal Free London NHS Foundation Trust, University College London Institute of Immunity and Transplantation, London, United Kingdom
6 Department of Medicine, University of Cambridge, Cambridge, United Kingdom
7 Department of Immunology, Queen Elizabeth University Hospital, Glasgow, United Kingdom
8 Department of Immunology, Royal Surrey County Hospital, Guildford, United Kingdom
9 Department of Immunology, Barts Health NHS Trust, London, United Kingdom
10 Department of Clinical Immunology, Hôpital Saint-Louis, Assistance Publique Hôpitaux de Paris (APHP), Paris, France
11 Department of Internal Medicine, Academic Medical Center, Amsterdam, The Netherlands
12 Department of Blood Cell Research, Sanquin Research, Amsterdam, The Netherlands
13 Department of Plasma Proteins, Sanquin Research, Amsterdam, The Netherlands
14 Molecular and Cellular Immunology Section, UCL Great Ormond Street Institute of Child Health and Great Ormond Street Hospital NHS Trust London, London, United Kingdom
15 NIHR-BioResource – Rare Diseases, Cambridge Biomedical Campus, Cambridge, United Kingdom
* These authors contributed equally to this work.
** These authors contributed equally to this work.

Corresponding author 1:
Taco W. Kuijpers
Meibergdreef 9
1105 AZ Amsterdam
The Netherlands
Tel: +31 20 566 2727
Fax: +31 20 566 9683
Email: t.w.kuijpers@amc.nl

Corresponding author 2:
James E. Thaventhiran
Addenbrooke's Hospital (Box 157)
Hills Rd
Cambridge
CB2 0QQ
Tel: +44 1223 762 145
Fax: +44 1223 336 849
Email: jedt2@cam.ac.uk

Funding:

This study was supported by The National Institute for Health Research England (grant number RG65966), and by the Center of Immunodeficiencies Amsterdam (CIDA). JET is supported by an MRC Clinician Scientist Fellowship (MR/L006197/1). AJT is supported by both the Wellcome Trust (104807/Z/14/Z) and by the National Institute for Health Research Biomedical Research Centre at Great Ormond Street Hospital for Children NHS Foundation Trust and University College London. EO receives personal fees from CSL Behring and MSD.

Conflicts of Interest

The authors do not have any conflicts of interest to declare.

Word count (Introduction → Discussion):

4,344

Abstract

Background: The genetic etiology of primary immunodeficiency disease (PID) carries prognostic information.

Objective: We conducted a whole-genome sequencing study assessing a large proportion of the NIHR-BioResource – Rare Disease cohort.

Methods: In the predominantly European study population of principally sporadic unrelated PID cases (*n*=846), a novel Bayesian method identified *NFKB1* as one most strongly associated with PID, and the association was explained by 16 novel heterozygous truncating, missense and gene deletion variants. This accounted for 4% of common variable immunodeficiency (CVID) cases (*n*=390) in the cohort. Amino-acid substitutions predicted to be pathogenic were assessed by analysis of structural protein data. Immunophenotyping, immunoblotting and *ex vivo* stimulation of lymphocytes determined the functional effects of these variants. Detailed clinical and pedigree information was collected for genotype-phenotype co-segregation analyses.

Results: Both sporadic and familial cases demonstrated evidence of the non-infective complications of CVID, including massive lymphadenopathy (24%), unexplained splenomegaly (48%) and autoimmune disease (48%), features prior studies correlate with worse clinical prognosis. Although partial penetrance of clinical symptoms was noted in certain pedigrees, all carriers have a deficiency in B lymphocyte differentiation. Detailed assessment of B lymphocyte numbers, phenotype and function identifies the presence of a raised $CD21^{low}$ B cell population: combined with identification of the disease-causing variant, this distinguishes between healthy individuals, asymptomatic carriers and clinically affected cases.

Conclusion: We show that heterozygous loss-of-function variants in *NFKB1* are the most common known monogenic cause of CVID that results in a temporally progressive defect in the formation of immunoglobulin-producing B cells.

Key Messages:

- Pathogenic variants in *NFKB1* are currently the most common known monogenic cause of CVID.
- There is a clear association with complications such as autoimmunity and malignancy, features associated with worse prognosis.
- These patients can be stratified by *NFKB1* protein level and the B cell phenotype.

Capsule Summary:

Whole genome sequencing revealed 16 novel pathogenic truncating, missense and gene deletion variants in *NFKB1*. Most mutations were linked to reduced protein expression, perturbed immunophenotyping and *ex vivo* stimulation assays of patients and relatives.

Keywords:

B cells, Common Variable Immunodeficiency, NF-κB1

Abbreviations used:

CADD:	Combined Annotation Dependent Depletion
CVID:	Common variable immunodeficiency
ExAC:	Exome Aggregation Consortium
IKK:	IκB kinase
IMDM:	Iscove's Modified Dulbecco's medium
InDels:	Insertions/deletions
iNKT cells:	Invariant Natural Killer T cells
LOF:	Loss-of-function
MAF:	Minor allele frequency
PBMCs:	Peripheral blood mononuclear cells
PID:	Primary immunodeficiency disease
PML:	Progressive multifocal leukoencephalopathy
RHD:	Rel homology domain
SNVs:	Single nucleotide variants
VEP:	Variant Effect Predictor

Introduction

Common variable immunodeficiency (CVID [MIM 607594]), which occurs in approximately 1:25,000 people[1-3], is a clinically and genetically heterogeneous disorder characterized by susceptibility to sinopulmonary infections, hypogammaglobulinemia and poor vaccine responses. CVID is the most common primary immune deficiency requiring life-long clinical follow up and the clinical course is highly variable with substantial excess mortality. Affected individuals frequently present with recurrent respiratory infections as well as immune dysregulatory features. The antibody deficiency is often not as marked as the agammaglobulinemia seen in the genetically defined conditions leading to B lymphocyte aplasia[4,5]. Conversely, while patients with B lymphocyte aplasia have a favorable prognosis on adequate replacement immunoglobulin treatment, the response of CVID patients is highly variable.

Past studies focused on familial cases with CVID and used techniques ranging from traditional linkage analysis to more recent exome sequencing to characterize the genetic etiology. This has revealed that monogenic gene dysfunction accounts for 10% of cases[4,5]. Several of the variants in these genes have been characterized as partially penetrant; it remains unclear whether genetic or environmental factors determine disease onset. Multiple recent studies identified variants in *NFKB1* as a monogenic cause of CVID and reported on the clinical features of these cases[6-11].

As part of this NIHR-BioResource – Rare Diseases study, we sequenced the genomes of 846 unrelated individuals with predominantly sporadic primary immunodeficiency (PID) who were recruited from across the UK and by European collaborators. Application of a recently developed statistical method BeviMed[12] to the 846 PID cases and over 5,000 control genomes, identified *NFKB1* as the gene with the highest probability of association with the disease, and with the largest number of cases explained by

variants in that gene. Further investigations revealed a series of 16 heterozygous loss-of-function (LOF) variants in *NFKB1* as the most common genetic cause of CVID.

Mutations in genes that affect NF-κB-dependent signaling are associated with a number of immunodeficiencies[13-26]. NF-κB is a ubiquitous transcription factor member of the Rel proto-oncogene family. NF-κB regulates the expression of several genes involved in inflammatory and immune responses. The classical activated form of NF-κB consists of a heterodimer of the p50/p65 protein subunits. The NF-κB family of transcription factors comprises five related proteins, c-Rel, p65 (RelA), RelB, p50 (NF-κB1) and p52 (NF-κB2) that interact to form homodimers and heterodimers with distinct gene regulatory functions[13,27-29]. Each Rel NF-κB protein has a conserved 300-amino-acid N-terminal Rel homology domain (RHD) that encompasses sequences needed by NF-κB proteins to bind DNA motifs (κB elements), to form dimers, to interact with regulatory inhibitor IκB proteins and to enter the nucleus. The ten different NF-κB dimers identified have distinct transcriptional properties[28]. In most cells, NF-κB is retained in the cytosol in a latent state through interaction with IκB proteins (such as IκBα, IκBβ and IκBε), a family of proteins with ankyrin repeats, that mediate IκB interaction with the RHD of NF-κB, masking the nuclear localization sequence and DNA-binding domains. Signal-dependent activation of an IκB kinase (IKK) complex comprising catalytic (α and β) and regulatory (NEMO) subunits, induces the phosphorylation and degradation of IκB[29], which permits NF-κB factors to enter the nucleus and regulate gene expression.

We show that variants in *NFKB1* culminate in a progressive humoral immunodeficiency indistinguishable from CVID, with a highly variable penetrance. We demonstrate the utility of an *in silico* protein prediction model for validating the predicted disease-causing substitutions, and we report on the clinical spectrum, immunological phenotype and functional consequences of heterozygous *NFKB1* variants.

Methods

Cohort

NIHR-BioResource – Rare Disease (NIHRBR-RD) study was established in the UK to further the clinical management of patients with rare diseases, by providing a national resource of whole genome sequence data. All participants provided written informed consent and the study was approved by the East of England Cambridge South national institutional review board (13/EE/0325). At the time of our analysis, the NIHRBR-RD study included whole genome sequence data from 8,066 individuals, of which 1,299 were part of the PID cohort. These were predominantly singleton cases, but additional affected and/or unaffected family members of some of the patients were also sequenced; in total there were 846 unrelated index cases.

The PID patients were recruited by specialists in clinical immunology (either trained in pediatrics or internal medicine) from 26 hospitals in the UK, and a smaller number came from the Netherlands, France and Italy. The recruitment criteria included: clinical diagnosis of CVID according to the ESID criteria (ESID Registry – Working Definitions for Clinical Diagnosis of PID, 2014, latest version: April, 25, 2017); extreme autoimmunity; or recurrent (and/or unusual) infections suggestive of severely defective innate or cell-mediated immunity. Exclusion of known causes of PID was encouraged, and some of the patients were screened for one or more PID genes prior to enrollment into the PID cohort. The ethnic make-up of the study cohort represented that of the general UK population: 82% were European, 6% Asian, 2% African, and 10% of mixed ethnicity, based on the patients' whole genome data.

Given that PID is a heterogeneous disease, with overlap in phenotypes and genetic etiologies across different diagnostic categories, we decided to perform an unbiased genetic analysis of all of 846 unrelated index cases. Whole genome sequence data was additionally available for 63 affected and 345 unaffected relatives. Within a broad range of phenotypes, CVID is the most common disease category, comprising 46% of the NIHRBR-RD PID cohort (*n*=390 index cases, range 0-93 years of age).

Sequencing and variant filtering

Whole genome sequencing of paired-end reads was performed by Illumina on their HiSeq X Ten system. Reads of 100, 125 or 150 base pairs in length were aligned to the GRCh37 genome build using Isaac aligner, variants across the samples were jointly called with AGG tool, and large deletions were identified using Canvas and Manta algorithms (all software by Illumina), as described previously[30]. Average read depth was 35, with 95% of the genome covered by at least 20 reads.

Single nucleotide variants (SNVs) and small insertions/deletions (InDels) were filtered based on the following criteria: passing standard Illumina quality filters in >80% of the genomes sequenced by the NIHRBR-RD study; having a Variant Effect Predictor (VEP)[31] impact of either MODERATE or HIGH and having a minor allele frequency (MAF) <0.001 in the Exome Aggregation Consortium (ExAC) dataset, and MAF <0.01 in the NIHRBR-RD cohort. Large deletions called by both Canvas and Manta algorithms, passing standard Illumina quality filters, overlapping at least one exon, absent from control datasets[32], and having frequency of <0.01 in the NIRBR-RD genomes were included in the analysis.

All variants reported as disease-causing in this study were confirmed by Sanger sequencing using standard protocols. Large deletions were inspected in IGV plot (Figure E1) and breakpoints were confirmed by sequencing the PCR products spanning each deletion.

Gene and variant pathogenicity estimation

In order to evaluate genes for their association with PID, we applied the BeviMed inference procedure[12] to the NIHRBR-RD whole genome dataset. BeviMed (https://CRAN.R-project.org/package=BeviMed) evaluates the evidence for association between case/control status of unrelated individuals and allele counts at rare variant sites in a given locus. The method infers the posterior probabilities of association under dominant and recessive inheritance and, conditional on such an association, the posterior

probability of pathogenicity of each considered variant in the locus. BeviMed was applied to rare variants and large rare deletions in each gene, treating the 846 unrelated PID index cases as 'cases' and 5,097 unrelated individuals from the rest of the NIHRBR-RD cohort as 'controls'. All genes were assigned the same prior probability of association with the disease of 0.01, regardless of their previously published associations with an immune deficiency phenotype. Variants which had a VEP impact labelled HIGH were assigned higher prior probabilities of pathogenicity than variants with a MODERATE impact, as described previously[12].

Immunophenotyping and B and T cell functional assays

Peripheral blood mononuclear cells (PBMCs) were isolated using standard density gradient centrifugation techniques using Lymphoprep (Nycomed, Oslo, Norway). Absolute numbers of lymphocytes, T cells, B cells and NK cells were determined with Multitest six-color reagents (BD Biosciences, San Jose, USA), according to manufacturer's instructions. For the immunophenotyping of the PBMCs we refer to the Supplemental Methods of the Online Repository.

To analyze the *ex vivo* activation of T and B cells, PBMCs were resuspended in PBS at a concentration of 5–10×10^6 cells/ml and labeled with 0.5µM CFSE (Molecular Probes) as described previously[33] and in the Supplemental Methods of the Online Repository. Proliferation of B and T cells was assessed by measuring CFSE dilution in combination with the same mAbs used for immunophenotyping. Analysis of cells was performed using a FACSCanto-II flowcytometer and FlowJo software. Patient samples were analyzed simultaneously with PBMCs from healthy controls.

ELISA

The secretion of immunoglobulins by mature B cells was assessed by testing supernatants for secreted IgM, IgG and IgA with an in-house ELISA using polyclonal rabbit anti-human IgM, IgG and IgA reagents and a serum protein calibrator all from Dako (Glostrup, Denmark), as described previously[33].

SDS PAGE and Western Blot analysis

Blood was separated into neutrophils and PBMCs. Neutrophils (5×10^6) were used for protein lysates, separated by SDS polyacrylamide gel electrophoresis and transferred onto a nitrocellulose membrane. Individual proteins were detected with antibodies against NF-κB p50 (mouse monoclonal antibody E-10, Santa Cruz Biotechnology, Texas, USA), against IκBα (rabbit polyclonal antiserum C-21, Santa Cruz Biotechnology) and against human glyceraldehyde-3-phosphate dehydrogenase, GAPDH (mouse monoclonal antibody, Merck Millipore, Darmstadt, Germany).

Secondary antibodies were either goat anti-mouse-IgG IRDye 800CW, goat anti-rabbit-IgG IRDye 680CW or goat anti-mouse-IgG IRDye 680LT (LI-COR Biosciences, Lincoln, NE, USA). Relative fluorescence quantification of bound secondary antibodies was performed on an Odyssey Infrared Imaging system (LI-COR Biosciences, Nebraska, USA), and normalized to GAPDH.

NF-κB1 protein structure

To gain structural information on the NF-κB1 RHD, a previously resolved crystal structure of the p50 homodimer (A43-K353) bound to DNA was used[34]. Ankyrin repeats of NF-κB1 (Q498-D802) were modelled employing comparative homology modelling (Modeller 9v16) using the Ankyrin repeats crystal structure of NF-κB2 as template[35,36]. There is no structural information on the region between the 6th and 7th Ankyrin repeat (F751-V771)[36] and was therefore omitted in the model. The death domain (G807-S894) structure has been resolved by Saito and co-workers using NMR (pdb: 2bdf).

Statistical analysis of lymphocyte data

Differences between groups with one variable were calculated with a non-paired Student's *t*-test or one-way ANOVA with Bonferroni post-hoc test, differences between groups with two or more variables were calculated with two-way ANOVA with Bonferroni post-hoc test using GraphPad Prism 6. A P-value less than 0.05 was considered significant.

Results

Pathogenic variants in *NFKB1* are the most common monogenic cause of CVID

In an unbiased approach to analysis, we obtained BeviMed posterior probabilities of association with PID for every individual gene in all 848 unrelated PID patients in the NIHR-BioResource – Rare Disease (NIHRBR-RD) study. Genes with posterior probabilities greater than 0.05 are shown in Figure 1, showing that *NFKB1* has the strongest prediction of association with disease status (1.000). All 13 HIGH impact variants (large deletion, nonsense, frameshift and splice site) in *NFKB1* were observed in cases only, resulting in the very high posterior probabilities of pathogenicity (mean 0.99) for this class of variants (Figure 2). On the other hand, MODERATE impact variants (missense substitutions) were observed both in cases and controls. The majority had near zero probability of pathogenicity, but three substitutions were observed in the PID cases only, and had posterior probabilities greater than 0.15 (Figure 2), suggesting their potential involvement in the disease. Genomic variants with a high Combined Annotation Dependent Depletion (CADD) score were found within both the PID and control cohorts, suggesting that this commonly used metric of variant deleteriousness cannot reliably distinguish disease-causing from benign variants in *NFKB1*. All 16 predicted likely pathogenic variants were private to the PID cohort, and further investigation revealed that all 16 individuals were within the diagnostic criteria of CVID (Table 1).

Assessment of all 390 CVID cases in our cohort for pathogenic variants showed that the next most commonly implicated genes are *NKFB2* and *BTK*, with three explained cases each (Figure E2). Importantly, based on the gnomAD dataset of 135,000 predominantly healthy individuals, none of the *NFKB1* variants reported here are observed in a single gnomAD individual, even though 90% of our CVID cohort, and all of the *NFKB1*-positive cases, had European ancestry. Therefore, our results suggest that LOF variants in *NFKB1* are the most commonly identified monogenic cause of CVID in the European

population, with 16 out of 390 CVID cases explaining up to 4.1% of our cohort. None of the variants identified here had been reported in the previously described *NFKB1* cases[6-11].

The *NFKB1* gene encodes the p105 protein that is processed to produce the active DNA-binding p50 subunit[13]. The 16 potentially pathogenic variants we identified were all located in the N-terminal p50 part of the protein (Figure 2). The effects of the three rare substitutions on NF-κB1 structure were less clear than those of the truncating and gene deletion variants, so we assessed their position in the crystal structure of the p50 protein. Their location in the inner core of the RHD (Figure 3A) suggested a potential impact on protein stability, whereas other rare substitutions in the NIHRBR-RD cohort were found in locations less likely to affect this (Figure 2 and 3A, Figure E3).

NF-κB1 LOF as the disease mechanism

Twelve patients with truncating variants (Arg284*, His513Glnfs*28, c.160-1G>A and Asp451*), one patient with gene deletion (del 103370996-103528207) and three patients with putative protein destabilizing missense variants (Ile281Met, Val98Asp and Ile87Ser) were investigated for evidence of reduced protein level. Assessment of the NF-κB1 protein level in PBMCs or neutrophils in 9 index cases and 7 *NFKB1* variant carrying relatives demonstrated a reduction in all individuals (Figure 3B, Figure E4). Relative fluorescence quantification of the bands confirmed this and demonstrated a protein level of 38 ± 4.3% (mean ± SEM) compared to healthy controls. There was no difference between clinically affected and clinically unaffected individuals (36 ± 4.4% versus 42 ± 10.1%, *n*=11 versus *n*=5, P=0.50). Our observations indicate that the pathogenic *NFKB1* variants result in LOF of the NF-κB1 p50 subunit, as reduction in protein levels was seen in all carriers regardless of their clinical phenotype, and was absent in family members that were non-carriers.

Variable disease manifestations in *NFKB1* LOF individuals

Seven individuals had evidence of familial disease (Table 1), prompting us to investigate genotype-phenotype co-segregation and disease penetrance in cases for which pedigree information and additional family members were available (Figure 4, Tables E1-3). The age at which hypogammaglobulinemia becomes clinically overt is highly variable (Figure E5), as illustrated by pedigree C in which the grandchildren carrying the c.160-1G>A splice-site variant had IgG subclass deficiency (C:III-3 and C-III-4), in one case combined with an IgA deficiency (C:III-3). Although not yet overtly immunodeficient, the clinical course of their fathers (C:II-3 and C:II-5) and grandmother (C:I-2) predicts this potential outcome, and warrants long-term clinical follow up of these children.

We also observed variants in individuals who were clinically asymptomatic. Pedigree A highlights variable disease penetrance: the healthy mother (A:II-1) carries the same Arg284* variant as two of her clinically affected children (A:III-2 and A:III-3). The identification of this nonsense variant prompted clinical assessment of the extended kindred and demonstrated that her sister (A:II-4) suffered from recurrent sinopulmonary disease and nasal polyps with serum hypogammaglobulinemia consistent with a CVID diagnosis. Overall, based on the clinical symptoms observed at the time of this study across six pedigrees, the penetrance of *NFKB1* variants with respect to the clinical manifestation of CVID is incomplete (about 60% in our cohort, 11 affected individuals among 18 variant carriers), with varied expressivity not only of age at disease onset, but of specific disease manifestations even within the same pedigree.

The clinical disease observed among the *NFKB1* variant carriers is characteristic of progressive antibody deficiency associated with recurrent sinopulmonary infections (100% of clinically affected individuals), with encapsulated microbes such as *Streptococcus pneumoniae* and *Haemophilus influenzae* species

(Table E1). The clinical spectrum of *NFKB1* LOF includes massive lymphadenopathy (24%), unexplained splenomegaly (48%) and autoimmune disease (48%) – either organ-specific and/or hematological of nature (mainly autoimmune hemolytic anemia and idiopathic thrombocytopenic purpura, Tables E1 and E2). The percentage of autoimmune complications is based on the presence of autoimmune cytopenias (autoimmune hemolytic anemia, idiopathic thrombocytopenic purpura (<50-75x10^6/mL), autoimmune neutropenia, Evans syndrome), alopecia areata/totalis, vitiligo and Hashimoto thyroiditis among the clinically affected cases. Granulomatous-lymphocytic interstitial lung disease and splenomegaly were considered lymphoproliferation. Enteropathy, liver disease, colitis and a mild decrease in platelet count (>100x10^6/mL) were neither included in those calculations nor scored separately. Histological assessment of liver disease found in three patients showed no evidence of autoimmune or granulomatous liver disease, though fibrosis and cirrhosis was observed, in these male patients. Finally, the number of oncological manifestations, predominantly hematological, was noticeable. There were two cases with solid tumors (parathyroid adenoma, breast cancer) and four cases with hematological malignancies (B-cell non-Hodgkin lymphoma, diffuse large B-cell lymphoma, follicular lymphoma, peripheral T-cell lymphoma), which add up to 6/21 cases, 28.6%.

B cell phenotype in *NFKB1* LOF individuals and immune cell activation

Index cases and family members carrying *NFKB1* variants were approached for repeat venipuncture for further functional assessment. In clinically affected individuals the B cell numbers and phenotype were indistinguishable from that described for CVID[37] (Figure 5, Figure E6). However, in clinically unaffected individuals, the absolute B cell count was often normal or raised (Figure 5A). In all individuals with *NFKB1* LOF variants the numbers of switched memory B cells were reduced (Figure 5B-5D), while a broad range of non-switched memory B cells was observed. This demonstrates that while the clinical phenotype of *NFKB1* LOF variants is partially penetrant, all carriers have a deficiency in class-switched memory B cell generation. The presence of raised numbers of the $CD21^{low}$ population described in CVID discriminates

366 between clinically affected and unaffected individuals with *NFKB1* LOF variants (Figure 5E). The B cells
367 from individuals with *NFKB1* LOF variants demonstrated impaired proliferative responses to anti-
368 IgM/anti-CD40/IL-21 and CpG/IL-2 (Figure 6A); this corresponded with the inability to generate
369 plasmablasts ($CD38^{+}/CD27^{++}$), most pronounced in the more extreme phenotypes, i.e. clinically affected
370 cases (Figure 6B, Figure E7B). Similarly, *ex vivo* IgG production was reduced in individuals with LOF
371 variants, whereas the level of IgM in the supernatants was normal (Figure 6C and 6D, Figure E7C),
372 compatible with hypogammaglobulinemia.
373
374 **T cell phenotype in *NFKB1* LOF individuals**
375 The T cell phenotype was largely normal in the subset distribution (Figure E8 and E9). Similar to the
376 knockout mouse model[38], we found an aberrant number of invariant Natural Killer T (iNKT) cells in the
377 clinically affected individuals (Figure E8). T cell proliferation was intact upon anti-CD3/anti-CD28 or IL-15
378 activation (Figure E10). Since iNKT cells have been implicated in diverse immune reactions[39], this
379 deficiency may contribute to the residual disease burden in replacement immunoglobulin treated
380 patients, some of whom had acute or chronic relapsing infection with herpes virus and, in one case, JC-
381 virus.

Discussion

In our study we show that LOF variants in *NFKB1* are present in 4% of our cohort of CVID cases, being the most commonly identified genetic cause of CVID. Furthermore, we highlight specific features of these patients that distinguish them within the diagnostic category of "CVID", which otherwise applies to an indiscrete phenotype, acquired over time, that is termed 'common' and 'variable'. The majority of the genetic variants we report here truncate or delete one copy of the gene; together with pedigree co-segregation analyses, protein expression and the B cell functional data, we conclude that *NFKB1* LOF variants are causing autosomal dominant haploinsufficiency. This has now been recognized as the genetic mode of inheritance for at least 17 known PIDs, including that associated with previously reported variants in *NFKB1*[6,40-42]. In monogenic causes of PID, incomplete penetrance has been more frequently described in haploinsufficient, relative to dominant-negative, PID disease, having been reported in more than half of the monogenic autosomal dominant haploinsufficient immunological conditions described[40]. This may be because dominant-negative gain-of-function mutations cause disease by expression of an abnormal protein at *any* level while, as seen in this study, haploinsufficiency is predicted to lead to 50% residual function of the gene product. Incomplete penetrance of a genetic illness by definition will be associated with substantial variation in the clinical spectrum of disease, and the spectrum seen in this study is consistent with prior reports; in three pedigrees with 20 individuals[6], harboring heterozygous mutant *NFKB1* alleles, the age of onset varied from 2-64 years, with a high variety of disease severity, including two mutation carriers that were completely healthy at the age of 2 and 43 years.

It is important to temper skepticism of partial penetrance of immune genetic lesions with our knowledge that individual immune genes may have evolved in response to selection pressure for host protection against specific pathogens[43]. Consequently, within the relatively pathogen-free environment of

developed countries, the relevant pathogen for triggering disease may be scarce and reports documenting partial penetrance of the clinical phenotype will increase. This makes the traditional approaches of genetics for determining causality difficult. The BeviMed algorithm used in this study prioritized both the gene *NFKB1* and individual variants within *NFKB1* for contribution to causality; the power of methods like this will increase with greater data availability. The identification of a number of rare *NFKB1* variants with high CADD scores both in PID and control datasets highlights the potential for false attribution of disease causality when the genetics of an individual case is considered outside the context of relevant control data.

Currently healthy family members carrying the same *NFKB1* LOF variant demonstrated similar reductions in p50 expression and low numbers of switched memory B cells as their relatives suffering from CVID. The longitudinal research investigation of these individuals could help identify the additional modifiers, including epigenetic or environmental factors, which influence the clinical penetrance of these genetic lesions. The similarity of results seen in patients with large heterozygous gene deletions and in those with more discrete substitutions is consistent with haploinsufficiency as the shared disease mechanism.

In patients with mild antibody deficiency it is often difficult to decide when to initiate replacement immunoglobulin therapy; this may be the case for individuals and their family members identified with LOF *NFKB1* variants. Two measures seem to correlate well with clinical disease: first, the class-switch defect and lower IgG and IgA production *ex vivo*. Immunoglobulin class-switching is known to be regulated by NF-κB. Mutations in the NF-κB essential modulator (NEMO) cause a class-switch defective hyper-IgM syndrome in humans[20] as well as in the p50 knockout mice[13,44,45]. Haploinsufficiency of NF-κB may result in defective class-switch recombination due to poor expression of activation-induced cytidine deaminase (AID), a gene regulated by NF-κB, that, when absent, is also associated with

immunodeficiency[46]. Secondly, the ability to measure the CD21^{low} B cell population is widespread in diagnostic immunology laboratories and our study identifies this marker to correlate with NF-κB-disease activity. Although the function of these cells remain to be fully elucidated[47], this laboratory test may be useful for the longitudinal assessment of clinically unaffected individuals identified with LOF *NFKB1* variants.

Apart from suffering from recurrent and severe infections (including viral disease) for which these patients had been diagnosed with PID in the first place, autoimmunity and unexplained splenomegaly are very common manifestations in our patient cohort, similar to the other heterozygous *NFKB1* cases described[6-11]. Although autoimmunity has been subject to variable percentages per cohort study[3,48,49], it seems that these complications occur more frequently in *NFKB1*-haploinsufficient patients compared to unselected CVID cohorts. In contrast to IKAROS defects, but similar to *CTLA4* haploinsufficiency, we observed that *NFKB1*-haploinsufficiency may also result in chronic and severe viral disease, as noted for CMV and JC virus infections in three of our patients. In the study of Maffucci *et al.*[11], one of the *NFKB1*-affected cases also suffered from *Pneumocystis jirovecii* and progressive multifocal leukoencephalopathy (PML), which is suggestive for JC virus infection. Whether the B cell defect in *NFKB1*-haploinsufficiency is responsible for these non-bacterial infections is unclear[50,51]. PML is most often discovered in the context of an immune reconstitution inflammatory syndrome, as seen in HIV patients on antiretroviral therapy, and in multiple sclerosis patients after natalizumab discontinuation[52]. Although the exact contribution of B cell depletion in PML pathogenesis is unknown, the increased PML risk in rituximab-treated patients[53] suggests a protective role for B cells.

Three individuals in this cohort suffered from liver failure and an additional three of transaminitis. Although autoimmunity is suspected, a non-hematopoietic origin of liver disease cannot be excluded in

454 the absence of autoantibodies and nodular regenerative disease. Mouse models have suggested a non-
455 immune role for NF-κB signaling in liver failure[13,54-56].
456
457 In the cohort of *NFKB1* patients we identified a number of malignancies. Malignancies in PID patients
458 have been cited as the second-leading cause of death after infection[57,58], and murine-models have
459 demonstrated that haploinsufficiency of NF-κB1 is a risk factor for hematological malignancy[59]. In a large
460 CVID registry study on 2,212 patients, 9% had malignancies, with one-third being lymphomas, some
461 presenting prior to their CVID diagnosis[49]. Despite the fact that our cohort is relatively small, we found
462 oncological manifestations in 29% of our cases (two-third being lymphoma), suggesting that
463 malignancies in *NFKB1*-haploinsufficiency may occur more often than in unselected CVID patients. In a
464 study in 176 CVID patients, among the 626 relatives of patients with CVID, no increase in cancer risk was
465 observed[60], suggesting that when this does occur, as in this study (three out of seven), it may be due to a
466 shared genetic lesion. Therefore, in a pedigree with a LOF variant in *NFKB1*, any relatives with cancer
467 should be suspected of sharing the same pathogenic variant.
468
469 In conclusion, previous publications[61,62] have suggested that CVID is largely a polygenic disease. Our
470 results provide further evidence that LOF variants in *NFKB1* are the most common monogenic cause of
471 disease to date, even in seemingly sporadic cases. In these patients there is a clear association with
472 complications such as malignancy, autoimmunity and severe non-immune liver disease; this is important
473 since the excess mortality seen in CVID occurs in this group[48]. The screening for defined pathogenic
474 *NFKB1* variants accompanied by B cell phenotype assessment, has prognostic management and is
475 effective in stratifying these patients.

Supplemental Data

Supplemental Data include Supplemental Methods, Figure E1-E10 and Table E1-E3.

Consortia

The members of the NIHR-BioResource – Rare Diseases PID Consortium are: Zoe Adhya, Hana Alachkar, Ariharan Anantharachagan, Richard Antrobus, Gururaj Arumugakani, Chiara Bacchelli, Helen Baxendale, Claire Bethune, Shahnaz Bibi, Barbara Boardman, Claire Booth, Michael Browning, Mary Brownlie, Siobhan Burns, Anita Chandra, Hayley Clifford, Nichola Cooper, Sophie Davies, John Dempster, Lisa Devlin, Rainer Doffinger, Elizabeth Drewe, David Edgar, William Egner, Tariq El-Shanawany, Bobby Gaspar, Rohit Ghurye, Kimberley Gilmour, Sarah Goddard, Pavel Gordins, Sofia Grigoriadou, Scott Hackett, Rosie Hague, Lorraine Harper, Grant Hayman, Archana Herwadkar, Stephen Hughes, Aarnoud Huissoon, Stephen Jolles, Julie Jones, Peter Kelleher, Nigel Klein, Taco Kuijpers (PI), Dinakantha Kumararatne, James Laffan, Hana Lango Allen, Sara Lear, Hilary Longhurst, Lorena Lorenzo, Jesmeen Maimaris, Ania Manson, Elizabeth McDermott, Hazel Millar, Anoop Mistry, Valerie Morrisson, Sai Murng, Iman Nasir, Sergey Nejentsev, Sadia Noorani, Eric Oksenhendler, Mark Ponsford, Waseem Qasim, Ellen Quinn, Isabella Quinti, Alex Richter, Crina Samarghitean, Ravishankar Sargur, Sinisa Savic, Suranjith Seneviratne, Carrock Sewall, Fiona Shackley, Ilenia Simeoni, Kenneth G.C. Smith (PI), Emily Staples, Hans Stauss, Cathal Steele, James Thaventhiran, Moira Thomas, Adrian Thrasher (PI), Steve Welch, Lisa Willcocks, Sarita Workman, Austen Worth, Nigel Yeatman, Patrick Yong.

The members of the NIHR-BioResource – Rare Diseases Management Team are: Sofie Ashford, John Bradley, Debra Fletcher, Tracey Hammerton, Roger James, Nathalie Kingston, Willem Ouwehand, Christopher Penkett, F Lucy Raymond, Kathleen Stirrups, Marijke Veltman, Tim Young.

The members of the NIHR-BioResource – Rare Diseases Enrolment and Ethics Team are: Sofie Ashford, Matthew Brown, Naomi Clements-Brod, John Davis, Eleanor Dewhurst, Marie Erwood, Amy Frary, Rachel Linger, Jennifer Martin, Sofia Papadia, Karola Rehnstrom.

The members of the NIHR-BioResource – Rare Diseases Bioinformatics Team are: William Astle, Antony Attwood, Marta Bleda, Keren Carss, Louise Daugherty, Sri Deevi, Stefan Graf, Daniel Greene, Csaba Halmagyi, Matthias Haimel, Fengyuan Hu, Roger James, Hana Lango Allen, Vera Matser, Stuart Meacham, Karyn Megy, Christopher Penkett, Olga Shamardina, Kathleen Stirrups, Catherine Titterton, Salih Tuna, Ernest Turro, Ping Yu, Julie von Ziegenweldt.

The members of the Cambridge Translational GenOmics Laboratory are: Abigail Furnell, Rutendo Mapeta, Ilenia Simeoni, Simon Staines, Jonathan Stephens, Kathleen Stirrups, Deborah Whitehorn, Paula Rayner-Matthews, Christopher Watt.

Acknowledgments

We thank the patients and their family members for participation in our study.

References

1. Primary immunodeficiency diseases. Report of a WHO scientific group. Clin Exp Immunol. 1997;109 Suppl 1:1-28.
2. Oksenhendler E, Gerard L, Fieschi C, Malphettes M, Mouillot G, Jaussaud R, et al. Infections in 252 patients with common variable immunodeficiency. Clin Infect Dis. 2008;46(10):1547-54.
3. Cunningham-Rundles C, Bodian C. Common variable immunodeficiency: clinical and immunological features of 248 patients. Clin Immunol. 1999;92(1):34-48.
4. Bousfiha A, Jeddane L, Al-Herz W, Ailal F, Casanova JL, Chatila T, et al. The 2015 IUIS Phenotypic Classification for Primary Immunodeficiencies. J Clin Immunol. 2015;35(8):727-38.
5. Picard C, Al-Herz W, Bousfiha A, Casanova JL, Chatila T, Conley ME, et al. Primary Immunodeficiency Diseases: an Update on the Classification from the International Union of Immunological Societies Expert Committee for Primary Immunodeficiency 2015. J Clin Immunol. 2015;35(8):696-726.
6. Fliegauf M, Bryant VL, Frede N, Slade C, Woon ST, Lehnert K, et al. Haploinsufficiency of the NF-kappaB1 Subunit p50 in Common Variable Immunodeficiency. Am J Hum Genet. 2015;97(3):389-403.
7. Boztug H, Hirschmugl T, Holter W, Lakatos K, Kager L, Trapin D, et al. NF-kappaB1 Haploinsufficiency Causing Immunodeficiency and EBV-Driven Lymphoproliferation. J Clin Immunol. 2016;36(6):533-40.
8. Lougaris V, Moratto D, Baronio M, Tampella G, van der Meer JW, Badolato R, et al. Early and late B-cell developmental impairment in nuclear factor kappa B, subunit 1-mutated common variable immunodeficiency disease. J Allergy Clin Immunol. 2017;139(1):349-52 e1.
9. Kaustio M, Haapaniemi E, Goos H, Hautala T, Park G, Syrjanen J, et al. Damaging heterozygous mutations in NFKB1 lead to diverse immunological phenotypes. J Allergy Clin Immunol. 2017.
10. Schipp C, Nabhani S, Bienemann K, Simanovsky N, Kfir-Erenfeld S, Assayag-Asherie N, et al. Specific antibody deficiency and autoinflammatory disease extend the clinical and immunological spectrum of heterozygous NFKB1 loss-of-function mutations in humans. Haematologica. 2016;101(10):e392-e6.
11. Maffucci P, Filion CA, Boisson B, Itan Y, Shang L, Casanova JL, et al. Genetic Diagnosis Using Whole Exome Sequencing in Common Variable Immunodeficiency. Front Immunol. 2016;7:220.
12. Greene D, BioResource N, Richardson S, Turro E. A Fast Association Test for Identifying Pathogenic Variants Involved in Rare Diseases. Am J Hum Genet. 2017;101(1):104-14.
13. Zhang Q, Lenardo MJ, Baltimore D. 30 Years of NF-kappaB: A Blossoming of Relevance to Human Pathobiology. Cell. 2017;168(1-2):37-57.
14. Yoshioka T, Nishikomori R, Hara J, Okada K, Hashii Y, Okafuji I, et al. Autosomal dominant anhidrotic ectodermal dysplasia with immunodeficiency caused by a novel NFKBIA mutation, p.Ser36Tyr, presents with mild ectodermal dysplasia and non-infectious systemic inflammation. J Clin Immunol. 2013;33(7):1165-74.
15. Courtois G, Smahi A, Reichenbach J, Doffinger R, Cancrini C, Bonnet M, et al. A hypermorphic IkappaBalpha mutation is associated with autosomal dominant anhidrotic ectodermal dysplasia and T cell immunodeficiency. J Clin Invest. 2003;112(7):1108-15.
16. Burns SO, Plagnol V, Gutierrez BM, Al Zahrani D, Curtis J, Gaspar M, et al. Immunodeficiency and disseminated mycobacterial infection associated with homozygous nonsense mutation of IKKbeta. J Allergy Clin Immunol. 2014;134(1):215-8.
17. Mousallem T, Yang J, Urban TJ, Wang H, Adeli M, Parrott RE, et al. A nonsense mutation in IKBKB causes combined immunodeficiency. Blood. 2014;124(13):2046-50.

562 18. Pannicke U, Baumann B, Fuchs S, Henneke P, Rensing-Ehl A, Rizzi M, et al. Deficiency of innate
563 and acquired immunity caused by an IKBKB mutation. N Engl J Med. 2013;369(26):2504-14.
564 19. Doffinger R, Smahi A, Bessia C, Geissmann F, Feinberg J, Durandy A, et al. X-linked anhidrotic
565 ectodermal dysplasia with immunodeficiency is caused by impaired NF-kappaB signaling. Nat Genet.
566 2001;27(3):277-85.
567 20. Jain A, Ma CA, Liu S, Brown M, Cohen J, Strober W. Specific missense mutations in NEMO result
568 in hyper-IgM syndrome with hypohydrotic ectodermal dysplasia. Nat Immunol. 2001;2(3):223-8.
569 21. Greil J, Rausch T, Giese T, Bandapalli OR, Daniel V, Bekeredjian-Ding I, et al. Whole-exome
570 sequencing links caspase recruitment domain 11 (CARD11) inactivation to severe combined
571 immunodeficiency. J Allergy Clin Immunol. 2013;131(5):1376-83 e3.
572 22. Stepensky P, Keller B, Buchta M, Kienzler AK, Elpeleg O, Somech R, et al. Deficiency of caspase
573 recruitment domain family, member 11 (CARD11), causes profound combined immunodeficiency in
574 human subjects. J Allergy Clin Immunol. 2013;131(2):477-85 e1.
575 23. Torres JM, Martinez-Barricarte R, Garcia-Gomez S, Mazariegos MS, Itan Y, Boisson B, et al.
576 Inherited BCL10 deficiency impairs hematopoietic and nonhematopoietic immunity. J Clin Invest.
577 2014;124(12):5239-48.
578 24. Jabara HH, Ohsumi T, Chou J, Massaad MJ, Benson H, Megarbane A, et al. A homozygous
579 mucosa-associated lymphoid tissue 1 (MALT1) mutation in a family with combined immunodeficiency. J
580 Allergy Clin Immunol. 2013;132(1):151-8.
581 25. McKinnon ML, Rozmus J, Fung SY, Hirschfeld AF, Del Bel KL, Thomas L, et al. Combined
582 immunodeficiency associated with homozygous MALT1 mutations. J Allergy Clin Immunol.
583 2014;133(5):1458-62, 62 e1-7.
584 26. Punwani D, Wang H, Chan AY, Cowan MJ, Mallott J, Sunderam U, et al. Combined
585 immunodeficiency due to MALT1 mutations, treated by hematopoietic cell transplantation. J Clin
586 Immunol. 2015;35(2):135-46.
587 27. Hoffmann A, Natoli G, Ghosh G. Transcriptional regulation via the NF-kappaB signaling module.
588 Oncogene. 2006;25(51):6706-16.
589 28. Oeckinghaus A, Hayden MS, Ghosh S. Crosstalk in NF-kappaB signaling pathways. Nat Immunol.
590 2011;12(8):695-708.
591 29. Vallabhapurapu S, Karin M. Regulation and function of NF-kappaB transcription factors in the
592 immune system. Annu Rev Immunol. 2009;27:693-733.
593 30. Carss KJ, Arno G, Erwood M, Stephens J, Sanchis-Juan A, Hull S, et al. Comprehensive Rare
594 Variant Analysis via Whole-Genome Sequencing to Determine the Molecular Pathology of Inherited
595 Retinal Disease. Am J Hum Genet. 2017;100(1):75-90.
596 31. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, et al. The Ensembl Variant Effect
597 Predictor. Genome Biol. 2016;17(1):122.
598 32. Zarrei M, MacDonald JR, Merico D, Scherer SW. A copy number variation map of the human
599 genome. Nat Rev Genet. 2015;16(3):172-83.
600 33. aan de Kerk DJ, Jansen MH, ten Berge IJ, van Leeuwen EM, Kuijpers TW. Identification of B cell
601 defects using age-defined reference ranges for in vivo and in vitro B cell differentiation. J Immunol.
602 2013;190(10):5012-9.
603 34. Muller CW, Rey FA, Sodeoka M, Verdine GL, Harrison SC. Structure of the NF-kappa B p50
604 homodimer bound to DNA. Nature. 1995;373(6512):311-7.
605 35. Webb B, Sali A. Comparative Protein Structure Modeling Using MODELLER. Curr Protoc
606 Bioinformatics. 2014;47:5 6 1-32.
607 36. Tao Z, Fusco A, Huang DB, Gupta K, Young Kim D, Ware CF, et al. p100/IkappaBdelta sequesters
608 and inhibits NF-kappaB through kappaBsome formation. Proc Natl Acad Sci U S A. 2014;111(45):15946-
609 51.

610 37. Wehr C, Kivioja T, Schmitt C, Ferry B, Witte T, Eren E, et al. The EUROclass trial: defining
611 subgroups in common variable immunodeficiency. Blood. 2008;111(1):77-85.
612 38. Sivakumar V, Hammond KJ, Howells N, Pfeffer K, Weih F. Differential requirement for Rel/nuclear
613 factor kappa B family members in natural killer T cell development. J Exp Med. 2003;197(12):1613-21.
614 39. Brennan PJ, Brigl M, Brenner MB. Invariant natural killer T cells: an innate activation scheme
615 linked to diverse effector functions. Nat Rev Immunol. 2013;13(2):101-17.
616 40. Rieux-Laucat F, Casanova JL. Immunology. Autoimmunity by haploinsufficiency. Science.
617 2014;345(6204):1560-1.
618 41. Kuehn HS, Boisson B, Cunningham-Rundles C, Reichenbach J, Stray-Pedersen A, Gelfand EW, et
619 al. Loss of B Cells in Patients with Heterozygous Mutations in IKAROS. N Engl J Med. 2016;374(11):1032-
620 43.
621 42. Chen K, Coonrod EM, Kumanovics A, Franks ZF, Durtschi JD, Margraf RL, et al. Germline
622 mutations in NFKB2 implicate the noncanonical NF-kappaB pathway in the pathogenesis of common
623 variable immunodeficiency. Am J Hum Genet. 2013;93(5):812-24.
624 43. Casanova JL. Severe infectious diseases of childhood as monogenic inborn errors of immunity.
625 Proceedings of the National Academy of Sciences of the United States of America. 2015;112(51):E7128-
626 37.
627 44. Sha WC, Liou HC, Tuomanen EI, Baltimore D. Targeted disruption of the p50 subunit of NF-kappa
628 B leads to multifocal defects in immune responses. Cell. 1995;80(2):321-30.
629 45. Snapper CM, Zelazowski P, Rosas FR, Kehry MR, Tian M, Baltimore D, et al. B cells from p50/NF-
630 kappa B knockout mice have selective defects in proliferation, differentiation, germ-line CH transcription,
631 and Ig class switching. J Immunol. 1996;156(1):183-91.
632 46. Revy P, Muto T, Levy Y, Geissmann F, Plebani A, Sanal O, et al. Activation-induced cytidine
633 deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2).
634 Cell. 2000;102(5):565-75.
635 47. Rakhmanov M, Keller B, Gutenberger S, Foerster C, Hoenig M, Driessen G, et al. Circulating
636 CD21low B cells in common variable immunodeficiency resemble tissue homing, innate-like B cells. Proc
637 Natl Acad Sci U S A. 2009;106(32):13451-6.
638 48. Resnick ES, Moshier EL, Godbold JH, Cunningham-Rundles C. Morbidity and mortality in common
639 variable immune deficiency over 4 decades. Blood. 2012;119(7):1650-7.
640 49. Gathmann B, Mahlaoui N, Ceredih, Gerard L, Oksenhendler E, Warnatz K, et al. Clinical picture
641 and treatment of 2212 patients with common variable immunodeficiency. J Allergy Clin Immunol.
642 2014;134(1):116-26.
643 50. Durali D, de Goer de Herve MG, Gasnault J, Taoufik Y. B cells and progressive multifocal
644 leukoencephalopathy: search for the missing link. Front Immunol. 2015;6:241.
645 51. Tan CS, Koralnik IJ. Progressive multifocal leukoencephalopathy and other disorders caused by JC
646 virus: clinical features and pathogenesis. Lancet Neurol. 2010;9(4):425-37.
647 52. Palazzo E, Yahia SA. Progressive multifocal leukoencephalopathy in autoimmune diseases. Joint
648 Bone Spine. 2012;79(4):351-5.
649 53. Carson KR, Evens AM, Richey EA, Habermann TM, Focosi D, Seymour JF, et al. Progressive
650 multifocal leukoencephalopathy after rituximab therapy in HIV-negative patients: a report of 57 cases
651 from the Research on Adverse Drug Events and Reports project. Blood. 2009;113(20):4834-40.
652 54. Tanaka M, Fuentes ME, Yamaguchi K, Durnin MH, Dalrymple SA, Hardy KL, et al. Embryonic
653 lethality, liver degeneration, and impaired NF-kappa B activation in IKK-beta-deficient mice. Immunity.
654 1999;10(4):421-9.
655 55. Pasparakis M, Luedde T, Schmidt-Supprian M. Dissection of the NF-kappaB signalling cascade in
656 transgenic and knockout mice. Cell Death Differ. 2006;13(5):861-72.

56. Geisler F, Algul H, Paxian S, Schmid RM. Genetic inactivation of RelA/p65 sensitizes adult mouse hepatocytes to TNF-induced apoptosis in vivo and in vitro. Gastroenterology. 2007;132(7):2489-503.
57. Mueller BU, Pizzo PA. Cancer in children with primary or secondary immunodeficiencies. J Pediatr. 1995;126(1):1-10.
58. Vajdic CM, Mao L, van Leeuwen MT, Kirkpatrick P, Grulich AE, Riminton S. Are antibody deficiency disorders associated with a narrower range of cancers than other forms of immunodeficiency? Blood. 2010;116(8):1228-34.
59. Voce DJ, Schmitt AM, Uppal A, McNerney ME, Bernal GM, Cahill KE, et al. Nfkb1 is a haploinsufficient DNA damage-specific tumor suppressor. Oncogene. 2015;34(21):2807-13.
60. Mellemkjaer L, Hammarstrom L, Andersen V, Yuen J, Heilmann C, Barington T, et al. Cancer risk among patients with IgA deficiency or common variable immunodeficiency and their relatives: a combined Danish and Swedish study. Clin Exp Immunol. 2002;130(3):495-500.
61. Orange JS, Glessner JT, Resnick E, Sullivan KE, Lucas M, Ferry B, et al. Genome-wide association identifies diverse causes of common variable immunodeficiency. J Allergy Clin Immunol. 2011;127(6):1360-7 e6.
62. van Schouwenburg PA, Davenport EE, Kienzler AK, Marwah I, Wright B, Lucas M, et al. Application of whole genome and RNA sequencing to investigate the genomic landscape of common variable immunodeficiency disorders. Clin Immunol. 2015;160(2):301-14.

677 **Table 1. Summary of the CVID patients' clinical presentation and their *NFKB1* variants.**

Case	*Sporadic / Familial*	*Infections*	*Auto-immunity*	*Malignancy*	*chr4 position (GRCh37)*	*Nucleotide change*	*Type of variant*	*cDNA(NM_003998.3):Amino Acid*
A	Familial	●			103504037	C>T	nonsense	c.850C>T;Arg284*
B	Familial	●	●	○	103518717	delCATGC	frameshift	c.1539_1543del;His513Glnfs*28
C	Familial	●	●	●○	103459014	G>A	splice-acceptor	c.160-1G>A;?
D	Familial	●	●		103518801	delGA	nonsense	c.1621_1622del; Asp541*
E	Sporadic	●	●	○	103504030	C>G	missense	c.843C>G;Ile281Met
F	Sporadic	●		●	103488178	T>A	missense	c.293T>A;Val98Asp
G	Sporadic	●			103488145	T>G	missense	c.260T>G;Ile87Ser
H	Familial	●	●		103501798	T>C	splice-donor	c.835+2T>C;?
I	Sporadic	●			103370996-103528207	-	large deletion	-
J	Familial	●			103517415	delG	frameshift	c.1423del;Ala475Profs*10
K	Sporadic	●	●	●	103436974-103652655	-	large deletion	-
L	Familial	●	●		103459041	delG	frameshift	c.187del;Glu63Lysfs*64
M	Sporadic	●	●		103501790	insA	frameshift	c.830dup;Lys278Glufs*3
N	Sporadic	●	●	●	103504086	insT	frameshift	c.904dup;Ser302Phefs*7
O	Sporadic	●			103488180	C>T	nonsense	c.295C>T;Gln99*
P	Sporadic	●	●		103505914	delG	frameshift	c.1005del;Arg336Glyfs*96

678 Closed dots, presence of symptoms in index patient
679 Open dots, presence of symptoms in family member of index patient
680

Figure 1. Overall BeviMed results showing that *NFKB1* has the highest posterior probability of association with disease in the NIHRBR-RD PID cohort. Genes with variants previously reported to cause PID are highlighted in red. Genes with posterior probabilities > 0.05 are shown.

Figure 2. Plot of rare missense, truncating and gene deletion *NFKB1* variants identified in the NIHRBR-RD genomes of unrelated individuals, and their location relative to *NFKB1* domains. The tracks from left to right show: number of unrelated case (red) and control (black) individuals in whom each variant was observed; the four major *NFKB1* domains; gray bars representing each exon in transcript ENST00000226574; variant annotation relative to transcript ENST00000226574 and genomic location of large deletions, with VEP HIGH impact variants and large deletions highlighted in blue; CADD scores of all nonsense, frameshift, splice and missense variants; ExAC allele frequencies; conditional probability of variant pathogenicity inferred using BeviMed. Only variants labelled as MODERATE or HIGH impact relative to the canonical transcript ENST00000226574 are shown. The initial inference that formed part of the genome-wide analysis included variant chr4:103423325G>A, which was observed in one control sample. This variant is intronic (LOW impact) relative to ENST00000226574 but is a splice variant (HIGH impact) relative to the minor transcript ENST00000505458. As variants were filtered based on the highest impact variant annotation against any Ensembl transcript, this variant was originally included in the inference. For this plot, the inference was re-run including only missense, truncating and gene deletion variants relative to the canonical transcript.

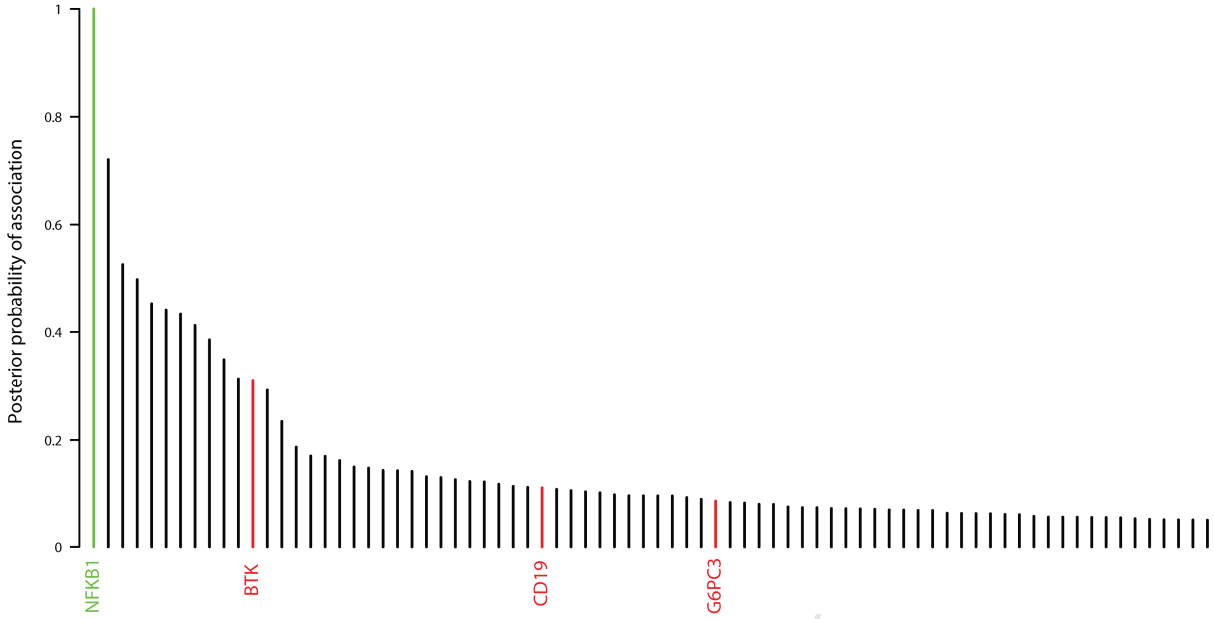
Figure 3. *NFKB1* LOF variants lead to haploinsufficiency of the p50 protein. (A) Localization of RHD substitutions with a high CADD score (>20) within the structure of the NF-κB p50 monomer. Shown is a solid (top panel) and a transparent (bottom panel) sphere representation of the NF-κB p50 monomer. Perturbed residues indicated in green were observed in a control dataset and are located on the outside of the structure, while the residues shown in red were perturbed exclusively in the PID cohort and are buried inside the structure. **(B)** Western blot analysis targeting p50, IκBα and GAPDH of *NFKB1* variant carriers. Left, representative blot of a healthy control and patient B-II:1; Right, summary of 16 *NFKB1* variant carriers showing haploinsufficiency, expressed as percentage of healthy controls on the same blot corrected for GAPDH, mean ± SEM.

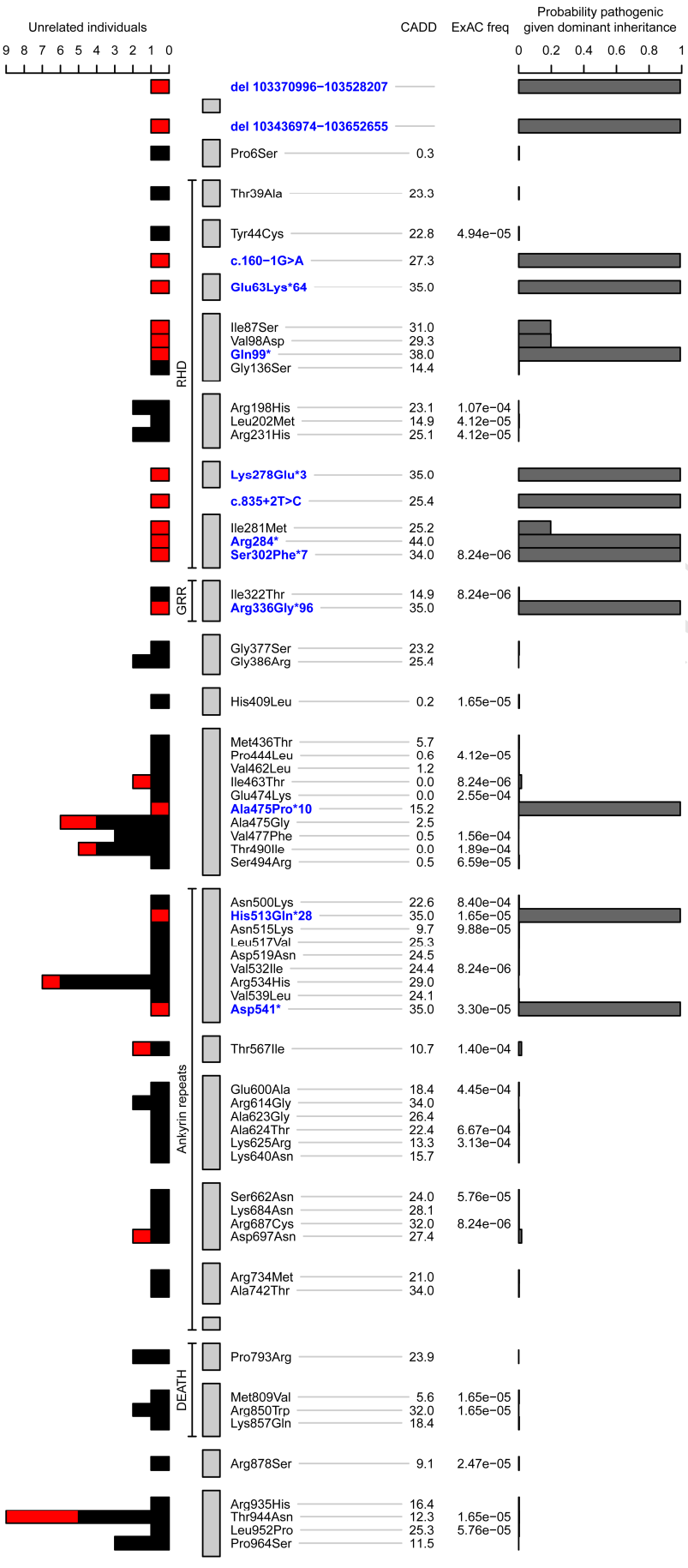
Figure 4. Pedigrees of familial *NFKB1* cases. Six affected families for which pedigree information and additional family members were available. Proband/index cases indicated with *P*.

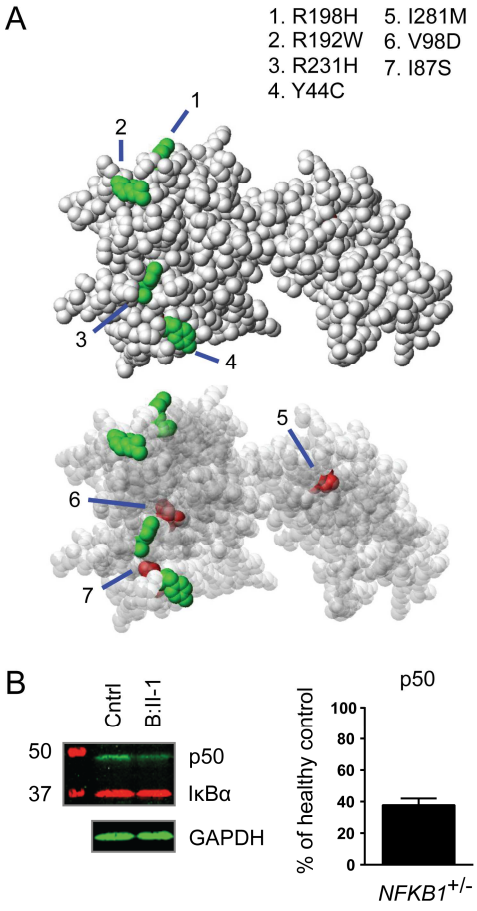
Figure 5. Decreased class-switched memory B cells and increased CD21^{low} B cells in *NFKB1* LOF variant carriers. (A) Absolute numbers of CD19^{+} B cells; each dot represents a single individual and their age. In grey are age-dependent reference values. **(B-E)** Percentages within CD19^{+}CD20^{+} B lymphocytes of **(B)** CD27^{+}IgD^{+} (non-switched memory or marginal-zone B cells) and CD27^{+}IgD^{-} (switched memory B cells), or **(C)** CD27^{+}IgG^{+}, or **(D)** CD27^{+}IgA^{+}, or **(E)** CD21^{low}CD38$^{low/dim}$. (*HD* healthy donor; *CU* clinically unaffected or *CA* clinically affected individuals with LOF variant in *NFKB1*.) Gating strategy is shown in Figure E6A. Only individuals with sufficient B cells could be analyzed. P-values were determined by one-way (Figure 5E) or two-way (Figure 5B) ANOVA with Bonferroni post-hoc test or unpaired Student's t-test (Figure 5C,D), *ns* not significant, **P≤0.01, ***P≤0.001.

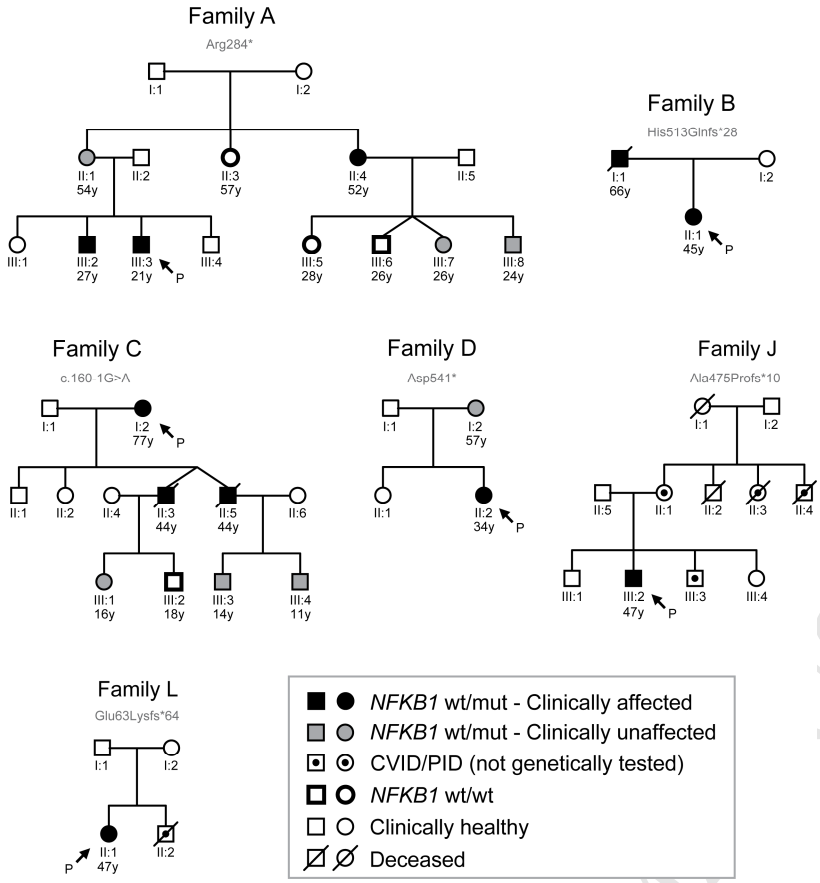
723 **Figure 6. The *ex vivo* class switch recombination defect of individuals carrying *NFKB1* LOF variants is**
724 **linked to the more extreme phenotype.** 6 day culture of CFSE-labeled lymphocytes normalized for B cell
725 number unstimulated, CpG/IL-2 (T cell independent activation) or anti-IgM/anti-CD40/IL-21 (T cell
726 dependent activation). **(A)** Percentage of divided B cells as measured by CFSE dilution **(B)** Percentage of
727 $CD27^{++}$ plasmablasts. Gating strategy is shown in Figure E7A. **(C, D)** IgM and IgG production in
728 supernatant of 6 day culture. (*HD* healthy donor; *CU* clinically unaffected or *CA* clinically affected
729 individuals with LOF variant in *NFKB1*.) Only individuals with sufficient B cells could be analyzed. P-values
730 were determined by two-way ANOVA with Bonferroni post-hoc test, *ns* not significant, **P≤0.01,
731 ***P≤0.001.

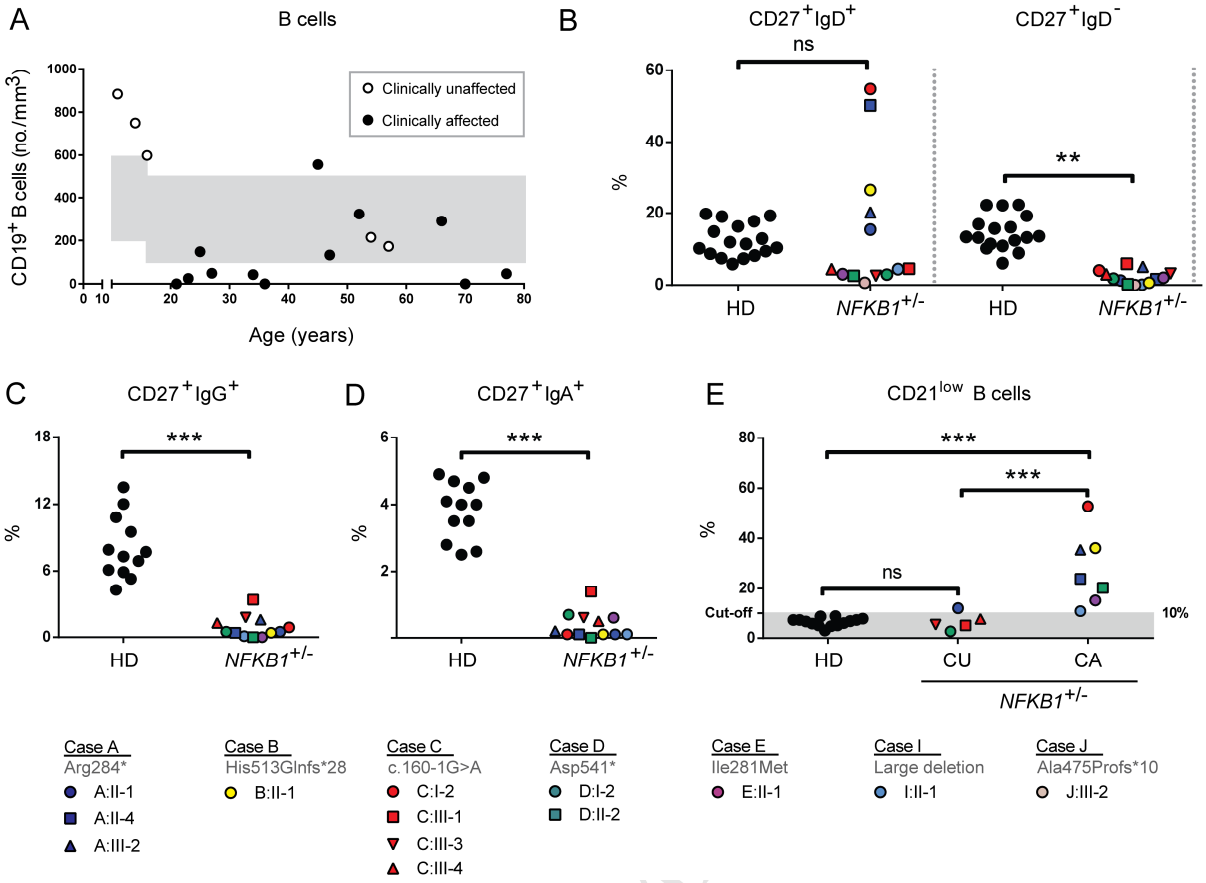
732

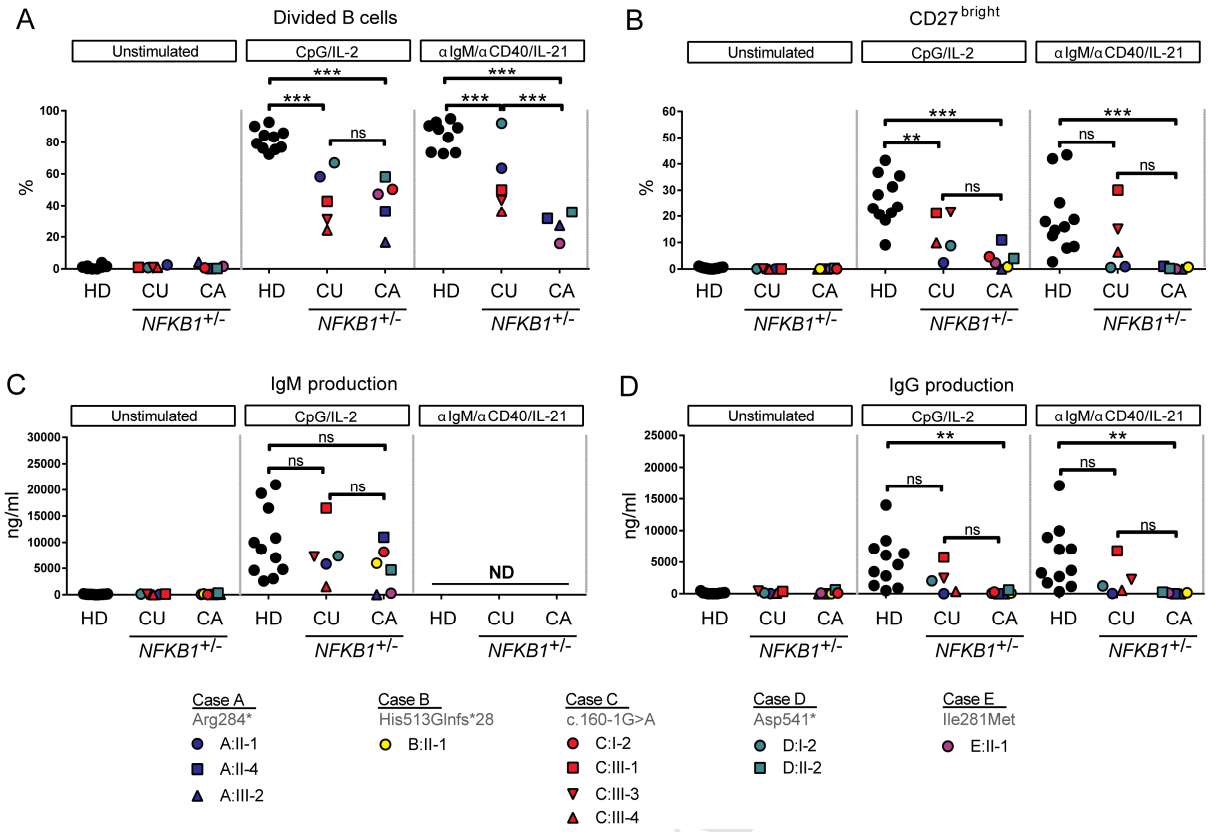












Supplemental Methods

Immunophenotyping

Peripheral blood mononuclear cells (PBMCs) were isolated using standard density gradient centrifugation techniques using Lymphoprep (Nycomed, Oslo, Norway). Absolute numbers of lymphocytes, T cells, B cells and NK cells were determined with Multitest six-color reagents (BD Biosciences, San Jose, USA), according to manufacturer's instructions. For the immunophenotyping the PBMCs were resuspended in PBS, containing 0.5% (w/v) BSA and 0.01% sodium azide and incubated with saturating concentrations of fluorescently labeled conjugated monoclonal antibodies. Analysis of cells was performed using a FACSCanto-II flowcytometer and FlowJo software. Patient samples were analyzed simultaneously with PBMCs from healthy controls. The following directly conjugated monoclonal antibodies were used: CD4 PE-Cy7 [348809], CD8 PerCP-Cy5.5 [341050], CD20 PerCP-Cy5.5 [332781], CD21 FITC [561372], CD27 APC [337169], CD38 PE-Cy7 [335825], CD45 APC-Cy7 [348815], IgD PE [555779], and TCRγδ FITC [347903] from BD (San Jose, USA), CD3 Alexa 700 [56-0038-41], CD19 Alexa 700 [56-0199-42] and CD27 APC-eFluor 780 [47-0279-42] from eBioscience (San Diego, USA), CD24 FITC [M1605] and CD27 FITC [M1764] from Sanquin (Amsterdam, the Netherlands), CD45RA (2H4-RD1) PE [6603181], TCRαβ PE [PN A39499], TCR Vα24 FITC [PM IM1589] and TCR Vβ11 PE [PN IM2290] from Beckman Coulter (Brea, USA), IgM FITC [F0317] and IgG FITC [F0158] from Dako (Glostrup, Denmark), IgA FITC [130-093-071] from Miltenyi Biotec (Bergisch Gladbach, Germany).

B and T cell functional assay

To analyze the *ex vivo* activation of T and B cells, PBMCs were resuspended in PBS at a concentration of 5–10×10^6 cells/ml and labeled with 0.5μM CFSE (Molecular Probes) in PBS for 10 minutes at 37°C under constant agitation. Cells were washed and subsequently resuspended in Iscove's Modified Dulbecco's medium (IMDM) supplemented with 10% fetal calf serum (BioWhittaker), antibiotics, and 3.57×10^-

^{4}%(v/v) β-mercaptoethanol (Merck). Labeled PBMCs containing a fixed number of 2×10^{4} (96-well plate) or 10×10^{4} B cells (48-well plate) per well in a flat-bottom plate for 6 days at 37°C and stimulated with saturating amounts of anti-IgM mAb (clone MH15; Sanquin), anti-CD40 mAb (clone 14G7; Sanquin) and 20ng/ml IL-21 (Invitrogen), or 1µg/ml CpG oligodeoxynucleotide 2006 (Invivogen) and 100U/ml IL-2 (R&D Systems), or anti-CD3 (clone 1xE) and anti-CD28 (clone 15E8), or IL-15 (R&D Systems). Proliferation of B and T cells was assessed by measuring CFSE dilution in combination with the same mAbs used for immunophenotyping and analyzed using a FACSCanto-II flowcytometer and FlowJo software.

Supplemental Figure Legends

Figure E1. IGV plot of large *NFKB1* deletion. IGV plot of patient I:II-1 showing the large *NFKB1* deletion identified by 50% reduction in the number of reads from the whole genome sequencing data mapping to that region.

Figure E2. Numbers of CVID patients per gene in which we identified a likely pathogenic variant fully explaining the patient's phenotype. Assessment of all 390 CVID cases in our cohort identified 31 patients with a monogenic defect in 11 different genes. Variants in *NFKB1* contributed to more than half of all CVID patients with a monogenic diagnosis (16/31, 52%). *NKFB2* and *BTK* were the next most commonly implicated genes, with three explained cases each.

Figure E3. Protein model of high-impact missense variants in and proximal to the ANK domain. Residues observed with missense variants containing a high CADD score (≥20) are highlighted. Most of these are situated on the protein exterior (green) and appear equally in the primary immunodeficiency cohort and non-primary immunodeficiency cohorts (**Figure 2)**. A623G (#19) appears to be located more interior, although mutation of an Alanine to Glycine is a moderate substitution. Equally notable are residues of the Ankyrin repeats which previously have been probed as interaction sites upon NF-κB dimerization. While R614 (#7), K684 (#8), L517 (#13) and R687 (#16) could be considered part of these putative interaction sites, variants at these sites are found in non-primary immunodeficiency patients (**Figure 2**).

Figure E4. Western blot analysis of all tested *NFKB1* variant carriers. Western blot analysis targeting p50, IκBα and GAPDH of *NFKB1* variant carriers. Twelve patients with truncating variants (Arg284*, His513Glnfs*28, c.160-1G>A and Asp451*), one patient with gene deletion (del 103370996-103528207) and three patients with putative protein destabilizing missense variants (Ile281Met, Val98Asp and Ile87Ser) were tested. Relative fluorescence quantification of p50 and GAPDH by Odyssey Infrared Imaging system above and below each western blot. IκBα was not targeted in the Case D western blot.

Figure E5. Serum IgM, IgG and IgA levels in serum of *NFKB1* LOF variant carrier. Each dot represents an *NFKB1* LOF variant carrier and their age. In grey age-dependent reference values.

Figure E6. Gating strategy for Figure 5B-E and additional B cell analyses. **(A)** Representative flow cytometry plots of a healthy control, patient A:II-1 (clinically unaffected) and patient A:II-4 (clinically affected). Phenotype of $CD19^+CD20^+$ B lymphocytes. Numbers represent percentages in corresponding quadrants. **(B)** Percentages of $CD27^-IgD^+$ (naïve) or $CD27^-IgD^+CD24^+CD38^+$ (transitional) B cells. (*HD* healthy donor, *NFKB1*$^{+/-}$ individual with *NFKB1* LOF variant.) Only individuals with sufficient B cells could be analyzed. P-values were determined by two-way ANOVA with Bonferroni post-hoc test (naïve B cells) or Student's t-test (transitional B cells), *ns* not significant, **P≤0.01.

Figure E7. Gating strategy Figure 6A and 6B and formation of $CD38^+$ plasmablasts and IgA production. **(A)** Representative flow cytometry plots of a healthy control and patient A:II-4 (clinically affected) after a 6 day culture of CFSE-labeled lymphocytes normalized for B cell number unstimulated, CpG/IL-2 (T cell independent activation) and anti-IgM/anti-CD40/IL-21 (T cell dependent activation). B cells gated on $CD19^+CD20^{-/+}$ and subsequently on $CD27^{++}$ (left) and CFSE and CD38 (right). Numbers represent percentages in corresponding quadrants. **(B)** Plasmablast formation measured by proliferation and CD38 upregulation ($CFSE^-CD38^+$). **(C)** IgA production in the supernatant. (*HD* healthy donor; *CU* clinically unaffected or *CA* clinically affected individuals with LOF variant in *NFKB1*.) Only individuals with sufficient B cells could be analyzed. P-values were determined by two-way ANOVA with Bonferroni post-hoc test, *ns* not significant, **P≤0.01, ***P≤0.001.

Figure E8. Additional lymphocyte numbers in individuals with *NFKB1* LOF variants. Absolute numbers of total lymphocytes ($CD45^+$), $CD4^+$ and $CD8^+$ T cells, NK cells ($CD3^-CD16^+CD56^+$) and invariant natural killer T cells ($CD3^+V\alpha24^+V\beta11^+$). Each dot represents a single individual and their age. In grey age-dependent normal values.

Figure E9. Normal T cell differentiation in individuals with NF-κB1 deficiency. (A) Representative flow cytometry dot plots of a healthy control and patient E:II-1, defining the differentiation of $CD3^+CD4^+$ and $CD3^+CD8^+$ T lymphocytes with CD27 and CD45RA. **(B, C)** Summary of subsets of **(B)** $CD4^+$ T cells and **(C)** $CD8^+$ T cells. (*HD* healthy donor, *NFKB1*$^{+/-}$ individual with *NFKB1* LOF variant.) P-values were determined by two-way ANOVA with Bonferroni post-hoc test, *ns* not significant.

Figure E10. T cells of individuals with *NFKB1* variants (*NFKB1*$^{+/-}$) show normal proliferative capacity. 6 day culture of CFSE-labeled lymphocytes unstimulated, anti-CD3/anti-CD28 (T cell receptor stimulation) or IL-15. Percentage of $CD4^+$ or $CD8^+$ T cell specific cell division as measured by CFSE dilution. (*HD* healthy donor, *NFKB1*$^{+/-}$ individual with *NFKB1* LOF variant.) P-values were determined by two-way ANOVA with Bonferroni post-hoc test, *ns* not significant.

Table E1. Clinical data of individuals carrying *NFKB1* variants

Case ID	Year of birth	Age at PID diagnosis	Clinical diagnosis	Major symptoms at onset	Infections	Autoimmunity / autoinflammation	Malignancy	Survival
Case A								
A:II-1	1961	-	Healthy	-	-	-	-	-
A:II-4	1963	2015	CVID	Sinusitis 2012 Pneumonia 2015	*S. pneumoniae*	-	-	-
A:III-2	1989	1992	XLA	-	Frequent sinopulmonary infections	-	-	-
A:III-3	1995	2005	XLA	Pneumonias 2002	EBV-related splenomegaly 2005 JC-virus 2016	- (splenomegaly)	-	-
Case B								
B:I-1	1939	1967	CVID	Appendicitis-abscess 1965 Bacterial meningitis 1967 Pneumonias 1967	*S. pneumoniae*	AIHA 2005	B-NHL (EBV-negative) 2007	Died 2007: heart attack
B:II-1	1968	2011	CVID	Pneumonias 2010	*S. pneumoniae,* *H. influenzae*	Alopecia areata 1988 Vitiligo 1989 Hypothyroiditis 2003	-	-
Case C								
C:I-2	1938	1990	CVID	Sinusitis Pneumonias ICU respiratory failure (1978) Bronchiectasis	*S. pneumoniae,* *H. influenzae*	Sister and her children have autoimmune disease (MS, IDDM1 & SLE)	Parathyroid adenoma	-
C:II-3	1972	1988	CVID	Pneumonias Bronchiectasis		IDDM1		Died in 2008: during 2nd OLT (1st OLT 2005)
C:II-5	1972	1985	CVID	OMAs Oral ulcers Sore throats	*H. influenzae* *S. pneumoniae* *C. albicans* *A. fumigatus*	-	DLBCL (EBV-neg)	Died in 2011: DLBCL
C:III-1	1999	-	Healthy	-	-	-	-	-
C:III-3	2001	-	Healthy	-	-	-	-	-
C:III-4	2004	-	Healthy	-	-	-	-	-
Case D								
D:I-2	1958	-	Healthy	-	-	Thyroid disease	-	-
D:II-2	1981	1999	CVID	ITP and anemia Hypogammaglobulinemia	Pneumonias Sinusitis	ITP Splenomegaly	-	-

				Asthma	*E. coli*, urinary tract infections	GLILD Periodontitis		
Case E								
E:II-1	1992	1999	ALPS	Bronchiectasis Hypogammaglobulinemia following multiple courses of Rituximab therapy	Chest infections Sinusitis	Autoimmune neutropenia ITP	-	-
Case F								
F:II-1	1946	2000	CVID	Severe pneumonias Bronchiectasis Lung fibrosis Cellulitis R leg	-	Hyperthyroidism	Follicular lymphoma (2005); recurrence (2008)	-
Case G								
G:II-1	1980	2001	CVID	Severe pneumonia Bronchiectasis	COPD lobectomy (2009) *M. avium*	Mild splenomegaly Chronic diarrhea	-	-
Case H								
H:II-1	1973	1997	CVID	ITP and AIHA Splenomegaly	Pneumonia Sinusitis Invasive CMV	Chronic diarrhea with villous atrophy	-	Died in 2008 CMV
Case I								
I:II-1	1991	2009	CVID	Multi-dermatomal shingles (age 12) Recurrent pneumonia Bronchiectasis Sinusitis	*S. pneumoniae* *H. influenzae*	Mild thrombocytopenia	-	-
Case J								
J:III-2	1969	2004	CVID	Recurrent pneumonia Sinusitis, otitis media Recurrent prostatitis	*M. catarrhalis* *H. influenzae* *P. aeruginosa* *S. pneumoniae* 2011 CMV DNA detected in urine	Diabetes (corticosteroid-induced)	-	-
Case K								
K:II-1	1952	1996	CVID	Recurrent pneumonias, otitis and sinusitis Pneumococcal meningitis	*P. aeruginosa* *H. influenzae* *S. pneumonia* *S. marcescens*	AIHA	Peripheral T-cell lymphoma, received CHOP	-
Case L								
L:II-1	1969	1991	CVID	Initially few symptoms – tested for IgG levels when her brother died of bacterial meningitis on		AIHA 1999 (splenectomy)	-	-

				background recurrent respiratory infections and low Ig's. Respiratory infections soon became apparent.				
Case M								
M:II-1	1985	2012	CVID	Respiratory infections, with bronchiectasis and chronic sinusitis. Chronic diarrhea	*P. aeruginosa* (bronchi/sinuses-2012,3,4,5,6), *S. aureus* 2015,6, RSV 2015, atypical myco-bacterium infection 2009, *C. difficile* - year unknown, plantar warts (HPV), ongoing Herpes Zoster, disseminated 2011 (while on azathioprine)	Evans syndrome 2009 (Rituximab) Autoimmune enteropathy 2010 Vitamin B12 deficiency (suspected pernicious anemia)	-	-
Case N								
N:II-1	1959	2015	CVID	Hypogammaglobulinemia. Generalised lymphadenopathy and splenomegaly Pancytopenia	EBV, CMV Neutropenic sepsis (without positive cultures)	Alopecia totalis	Breast cancer	
Case O								
O:II-1	1978	2001	CVID	Frequent bacterial infections Chronic diarrhea	*Giardia lamblia*	-	-	-
Case P								
P:II-1	1961	2004	CVID	AIHA (treated with Rituximab 2012). ITP and autoimmune neutropenia. Recurrent sinus and respiratory tract infections	Chronic Norovirus Rhinovirus *H. influenzae* *S. pneumoniae*	GLILD. Polyarthritis (RA-like: RF/ANA negative)	-	-

Note:

H. influenzae strains are non-typeable (= uncapsulated) unless mentioned specifically.

Abbreviations:

AIHA: Autoimmune haemolytic anaemia; B-NHL: B non-Hodgkin lymphoma; CVID: Common variable immunodeficiency; DLBCL: Diffuse Large B cell lymphoma; GLILD: Granulomatous-lymphocytic inflammatory lung disease; ITP: immune thrombocytopenia; OLT: Orthopic liver transplantation; XLA: X-linked agammaglobulinemia

Table E2. Symptoms and organ involvement in individuals carrying *NFKB1* variants

Case ID	Year of birth	Lung	Lymph nodes	Spleen	Liver function	Gastro-intestinal tract	Brain
Case A							
A:II-1	1961	-	-	-	-	-	-
A:II-4	1963	-	-	-	-	-	-
A:III-2	1989	-	-	-	-	-	-
A:III-3	1995	-	Enlarged	Splenectomy for splenomegaly, suspected malignancy	Transaminitis	-	-
Case B							
B:I-1	1939	-	-	Splenomegaly	Hepatomegaly	-	-
B:II-1	1968	-	-	-	-	-	-
Case C							
C:I-2	1938	Bronchiectasis Lung fibrosis	-	-	Transaminitis	-	-
C:II-3	1972	Bronchiectasis	Enlarged (cervical and axillary)	Enlarged	Liver fibrosis (1998) without granulomas, no signs of infection or autoimmunity	-	-
C:II-5	1972	Bronchiectasis	-	-	Liver cirrhosis (1996), suspect of hepatitis C virus infection	Duodenal partial villous blunting, no granuloma and absence of colonic plasmacells	Tremor
C:III-1	1999	-	-	-	-	-	-
C:III-3	2001	-	-	-	-	-	-
C:III-4	2004	-	-	-	-	-	-
Case D							
D:I-2	1958	-	-	-	-	-	-
D:II-2	1981	GLILD	-	Splenomegaly	Normal	-	-
Case E							
E:II-1	1992	Bronchiectasis	Enlarged as a child; biopsies unremarkable	Splenomegaly (previous ITP)	Normal	-	-
Case F							
F:II-1	1946	Bronchiectasis, fibrosis	-	-	-	-	-
Case G							
G:II-1	1980	Bronchiectasis, fibrosis	-	-	-	-	-
Case H							

H:II-1	1973	-	-	Splenomegaly	-	Chronic diarrhea with villous atrophy	CMV retinitis
Case I							
I:II-1	1991	Bronchiectasis	Enlarged (mediastinal)	Splenomegaly	Mild elevation transaminases, gamma-GT	-	-
Case J							
J:III-2	1969	Bronchiectasis (2005) Asthma	-	-	Normal	Intermittent diarrhoea and abdominal pain	-
Case K							
K:II-1	1952	Bronchiectasis (2009)	Previous peripheral T cell lymphoma: CD3+CD8+ (2015)	Splenomegaly (previous AIHA)	-	-	-
Case L							
L:II-1	1969	Chronic left lower lobe collapse, COPD (smoker), no bronchiectasis	-	Splenectomy (AIHA)	Normal	-	-
Case M							
M:II-1	1985	Bronchiectasis. No interstitial disease. Nodule top left lung	-	'Mild' splenomegaly on CT only	Normal	Autoimmune enteropathy (2011)	-
Case N							
N:II-1	1959	NAD	Generalised Lymphadenopathy	Large splenomegaly	Normal	-	-
Case O							
O:II-1	1978	NAD	-	-	Normal	-	-
Case P							
P:II-1	1961	GLILD bronchiectasis	-	Splenomegaly (2012), normalized over time	Nodular regenerative hyperplasia of the liver.	Chronic diarrhea: upon biopsies absence of plasmacells (Norovirus-positive).	-

Abbreviations:

AIHA: Autoimmune haemolytic anaemia; GLILD: Granulomatous-lymphocytic inflammatory lung disease; ITP: immune thrombocytopenia; NAD: No active disease

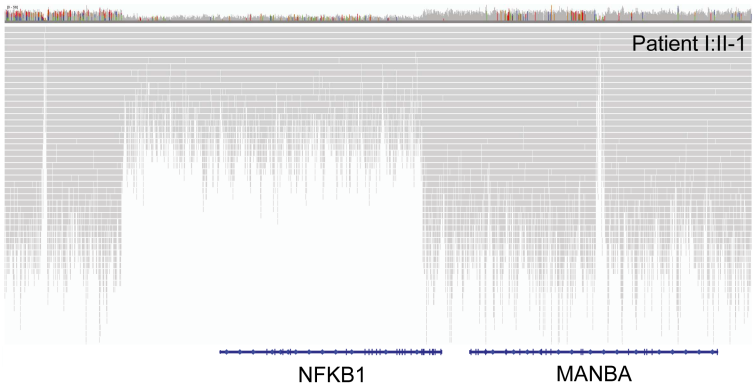
Table E3. Immunological findings in individuals carrying *NFKB1* variants

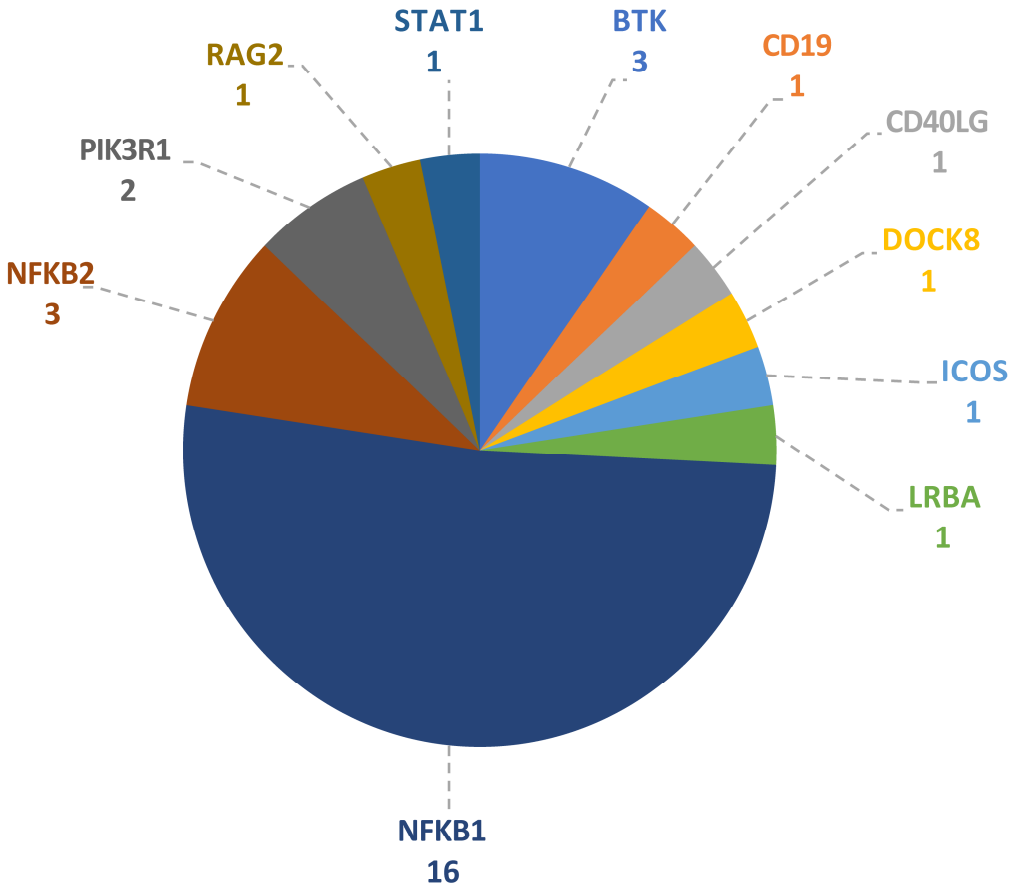
Case ID	Year of birth	Absolute lymphocyte count	Abs # CD3+ T cells	Abs # CD3/CD4+ T cells	Abs # CD3/CD8+ T cells	Abs # CD16/56+ NK cells	Abs # CD19+ B cells	IgA level (g/L)	IgG level (g/L) - prior to Ig subst.	IgM level (g/L)	ANA, other autoAbs
Case A											
A:II-1	1961	1.427	0.978 (68.5%)	0.624 (43.7%)	0.332 (23.3%)	0.228 (16.0%)	0.215 (15.0%)	0.13	1.5	0.44	ANA neg
A:II-4	1963	3.288	1.974 (60.0%)	1.305 (39.7%)	0.630 (19.2%)	0.981 (29.9%)	0.323 (9.8%)	0.03	12.1 (post IVIG)	0.44	ANA neg
A:III-2	1989	1.630	1.400 (85.9%)	0.403 (24.7%)	0.941 (57.8%)	0.153 (9.4%)	0.074 (4.5%)	<0.04	6.9	0.17	ANA neg
A:III-3	1995	2.527	2.269 (89.8%)	0.735 (29.1%)	1.401 (55.5%)	0.252 (10.0%)	0 (<0.50%)	<0.04	6.2	<0.03	ANA neg
Case B											
B:I-1	1939	1.189	0.820 (69,0%)	0.240 (20.0%)	0.540 (44.0%)	0.060 (5.0%)	0.290 (25.0%)	<0.1	<0.05	0.26	Coomb's pos
B:II-1	1968	1.936	1.551 (80.1%)	1.257 (64.9%)	0.284 (14.7%)	0.077 (4.0%)	0.291 (15.01%)	0.1	1.7	0.36	Anti-TPO pos
Case C											
C:I-2	1938	1.741	1.316 (75.6%)	0.243 (14.0%)	0.943 (54.2%)	0.375 (21.5%)	0.047 (2.7%)	<0.1	2.0	0.6	-
C:II-3	1972	0.950	0.870 (91.6%)	0.450 (47.4%)	0.410 (43.2%)	0.060 (6.3%)	0.020 (2.1%)	<0.1	1.5	0.20	IDDM1
C:II-5	1972	1.150	0.980 (85.2%)	0.650 (56.5%)	0.331 (28.8%)	0.100 (8.7%)	0.042 (3.7%)	<0.1	1.4	0.20	-
C:III-1	1999	2.959	2.027 (68.5%)	1.038 (35.1%)	0.867 (29.3%)	0.306 (10.3%)	0.597 (20.2%)	0.8	10.4	0.7	-
C:III-3	2001	3.458	1.919 (55.5%)	1.156 (33.4%)	0.591 (17.1%)	0.687 (19.9%)	0.749 (21.7%)	0.6	6.4 (low IgG2)	0.2	-
C:III-4	2004	4.054	2.615 (64.5%)	1.565 (38.6%)	0.821 (20.3%)	0.514 (12.7%)	0.883 (21.8%)	1.3	5.7 (low IgG2 and IgG3)	0.7	-
Case D											
D:I-2	1958	2.5	1.373 (54.9%)	0.761 (30.4%)	0.553 (22.1%)	NA	0.173 (6.9%)	2.3	11.0	0.84	-
D:II-2	1981	1.200	0.679 (55.6%)	0.504 (42.0%)	0.140 (11.7%)	NA	0.027 (2.2%)	<0.3	4.9	<0.1	-
Case E											

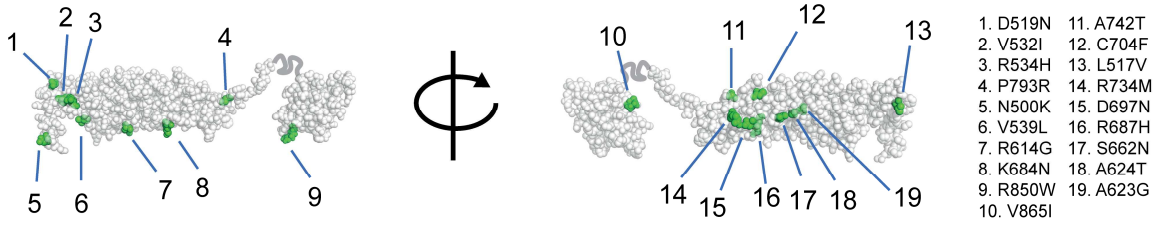
E:II-1	1992	1.046	0.963 (92.1%)	0.709 (67.8%)	0.203 (19.4%)	NA	0.022 (2.1%)	<0.1	3.9	<0.1	-
Case F											
F:II-1	1946	0.619	0.458 (74.0%)	0.174 (28.1%)	0.283 (45.7%)	NA	0	<0.1	<2	<0.1	-
Case G											
G:II-1	1980	1.022	0.684 (66.9%)	0.449 (43.9%)	0.235 (23.0%)	NA	0	<0.1	<3.9	<0.1	-
Case H											
H:II-1	1973	5.458 (splenectomy)	4.857 (89%)	2.129 (39%)	2.620 (48%)	0.262 (4.8%)	0.132 (3%)	0.02	0.1	0.03	Coomb's pos ANA pos
Case I											
I:II-1	1991	1.4	1.102 (78%)	0.717 (51%)	0.349 (25%)	0.166 (12%)	0.122 (7%)	<0.06	0.3	0.09	-
Case J											
J:III-2	1969	1.7	1.336 (78.6%)	0.706 (41.5%)	0.602 (35.4%)	0.019 (1.1%)	0.318 (18.7%)	<0.07	<1.0	0.17	-
Case K											
K:II-1	1952	0.7	0.639 (97%)	0.206 (31%)	0.388 (59%)	0.005 (1%)	0.015 (2%)	0.05	<0.1	<0.1	Coomb's pos
Case L											
L:II-1	1969	1.4	1.129 (77%)	0.687 (47%)	0.405 (28%)	0.249 (17%)	0.080 (5%)	<0.05	10.3 (post SCIG)	<0.05	Coomb's pos
Case M											
M:II-1	1985	1.0	0.902 (90%)	0.621 (62%)	0.223 (22%)	0.055 (5.5%)	0.028 (2.8%)	<0.04	Low pre Ig replacement	0.65	Coomb's pos, ANA neg
Case N											
N:II-1	1959	1.012	0.82 (81%)	0.54 (53%)	0.26 (28%)	0.15 (16%)	0.02 (2%)	<0.07	2.3	0.18	ANA neg
Case O											
O:II-1	1978	1.7	1.39 (82%)	0.97 (57%)	0.40 (24%)	0.06 (3.5%)	0.24 (14%)	0.4	4.6	0.3	ANA neg
Case P											
P:II-1	1961	0.79	0.52 (66%)	0.33 (42%)	0.18 (23%)	0.11 (14%)	0.15 (19%)	<0.1	1.9	0.1	Coomb's pos ANA neg

Abbreviations:

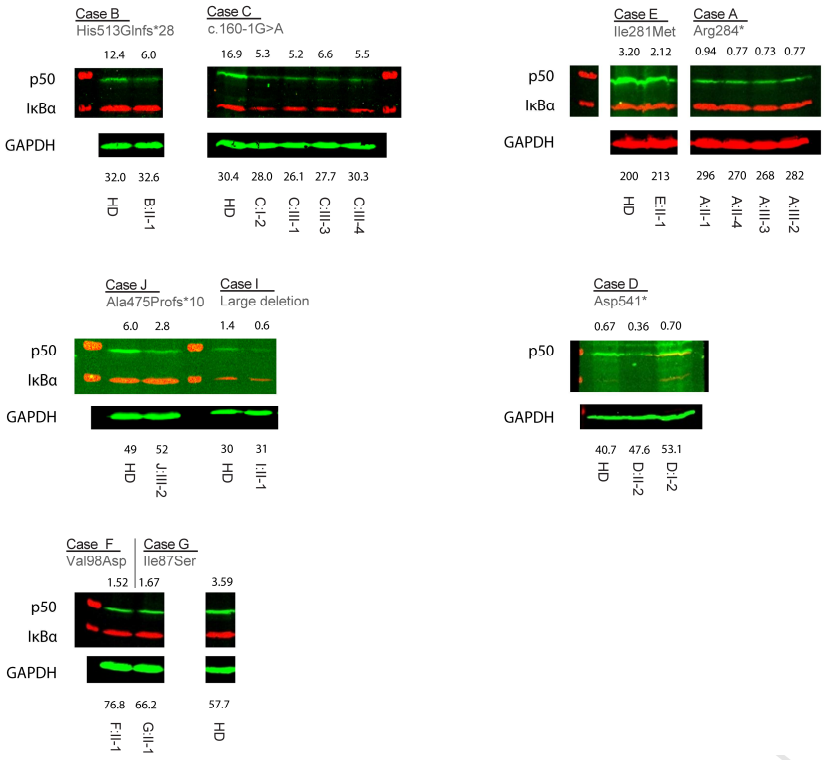
NA: Not Available

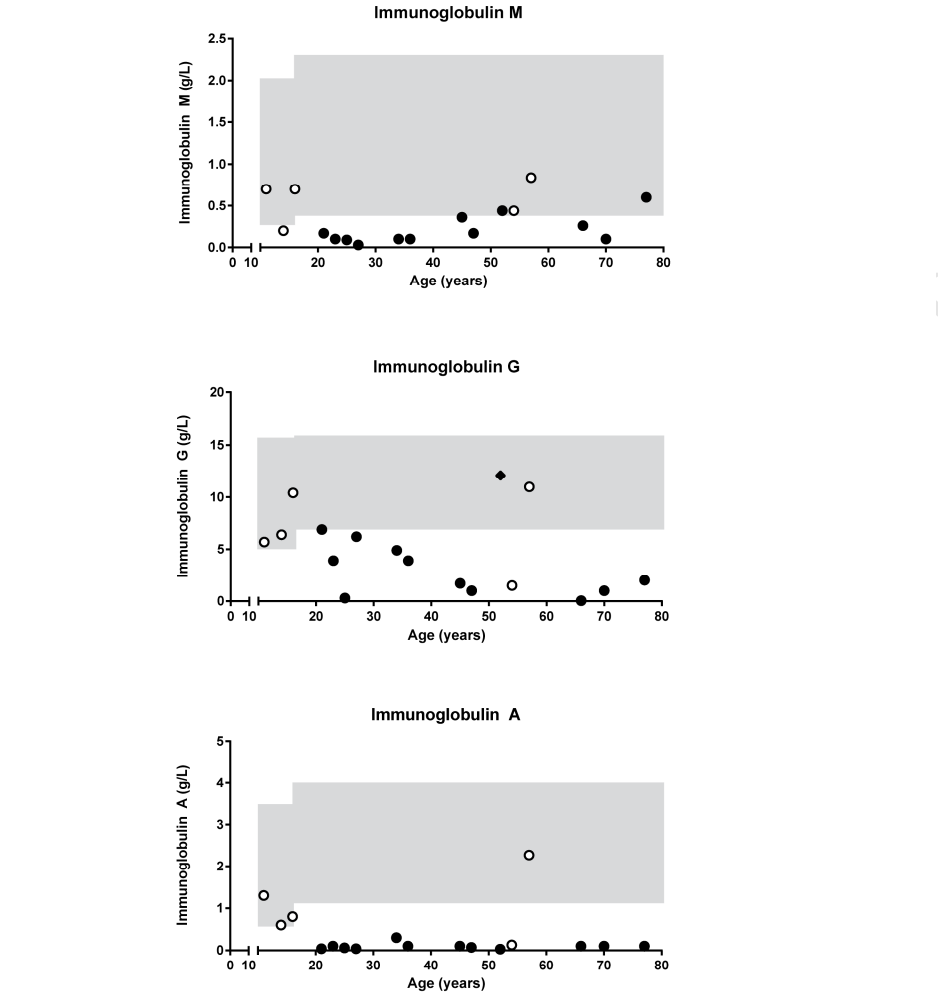


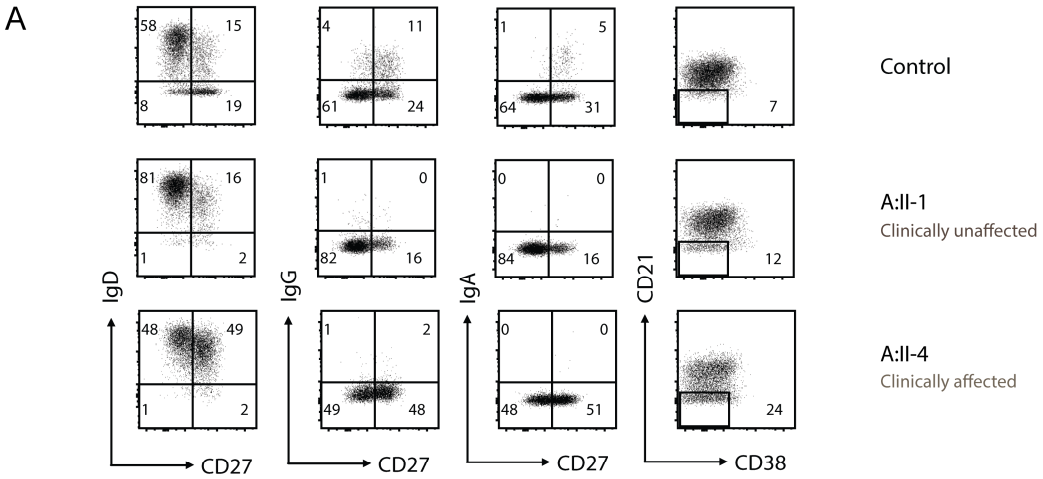


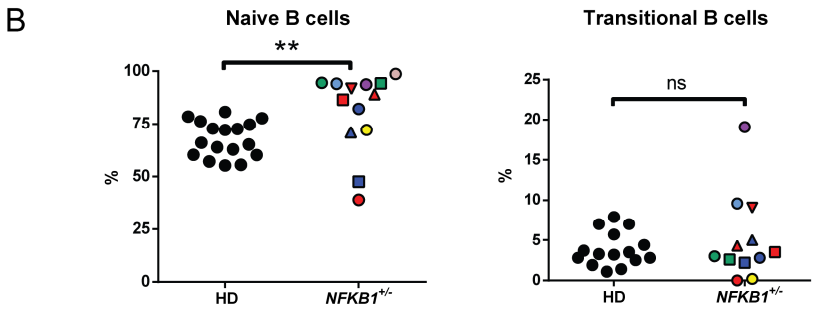


1. D519N
2. V532I
3. R534H
4. P793R
5. N500K
6. V539L
7. R614G
8. K684N
9. R850W
10. V865I
11. A742T
12. C704F
13. L517V
14. R734M
15. D697N
16. R687H
17. S662N
18. A624T
19. A623G

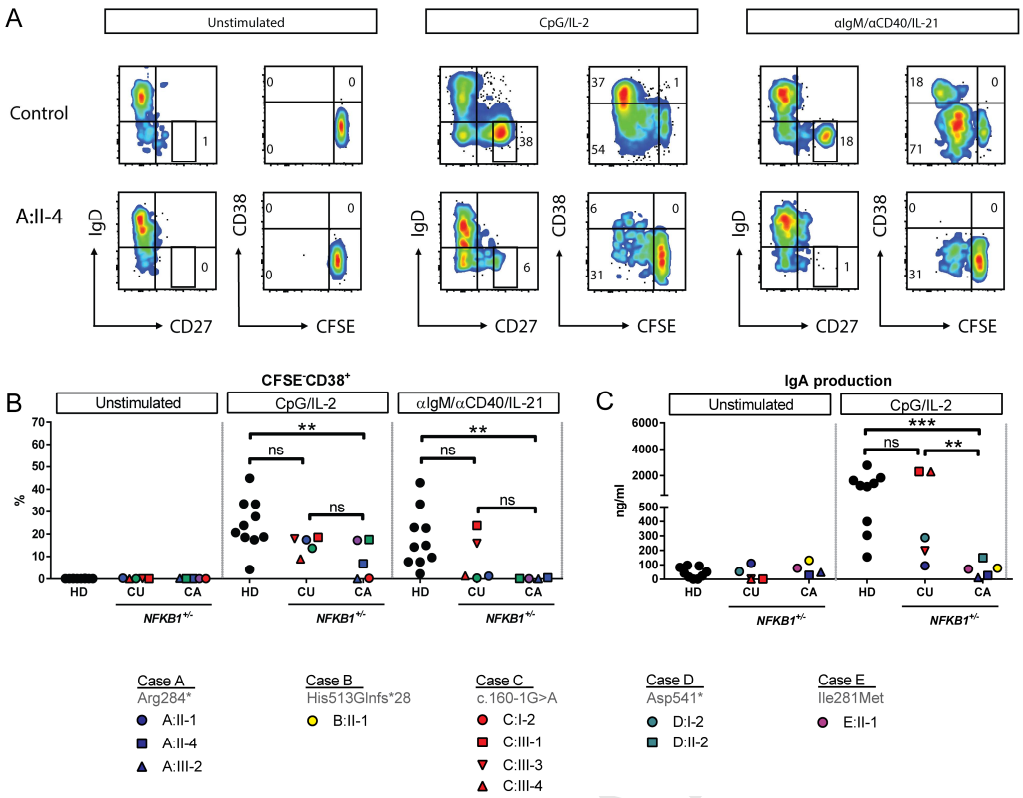



Immunoglobulin M
Immunoglobulin M (g/L)
2.5
2.0
1.5
1.0
0.5
0.0
0 10 20 30 40 50 60 70 80
Age (years)
Immunoglobulin G
Immunoglobulin G (g/L)
20
15
10
5
0
0 10 20 30 40 50 60 70 80
Age (years)
Immunoglobulin A
Immunoglobulin A (g/L)
5
4
3
2
1
0
0 10 20 30 40 50 60 70 80
Age (years)
Clinically affected
Clinically unaffected
Clinically affected (post IVIG)





Case A	Case B	Case C	Case D	Case E	Case I	Case J
Arg284*	His513Glnfs*28	c.160-1G>A	Asp541*	Ile281Met	Large deletion	Ala475Profs*10
A:II-1	B:II-1	C:I-2	D:I-2	E:II-1	I:II-1	J:III-2
A:II-4		C:III-1	D:II-2			
A:III-2		C:III-3				
		C:III-4				



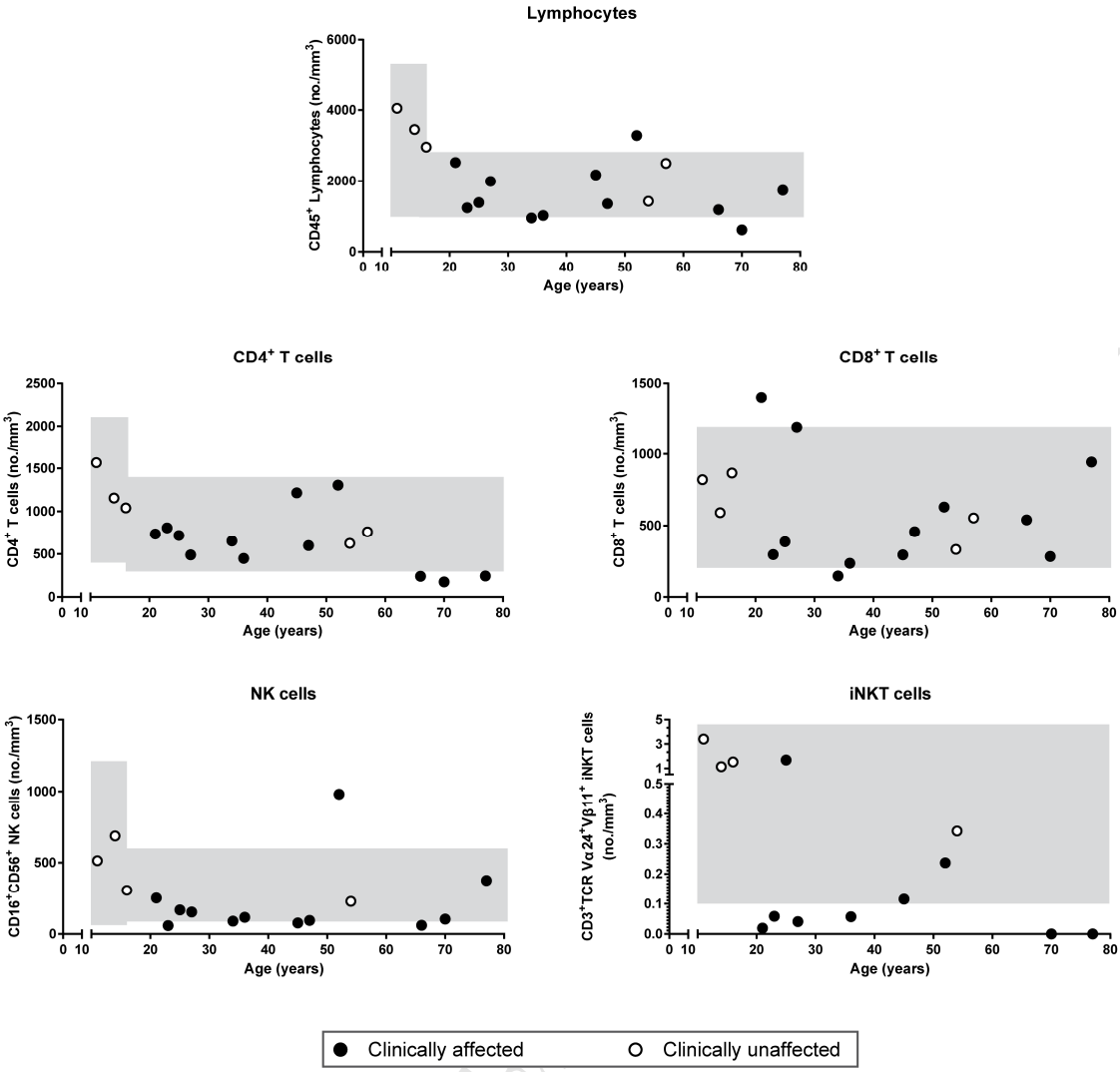
Case A
Arg284*
A:II-1
A:II-4
A:III-2

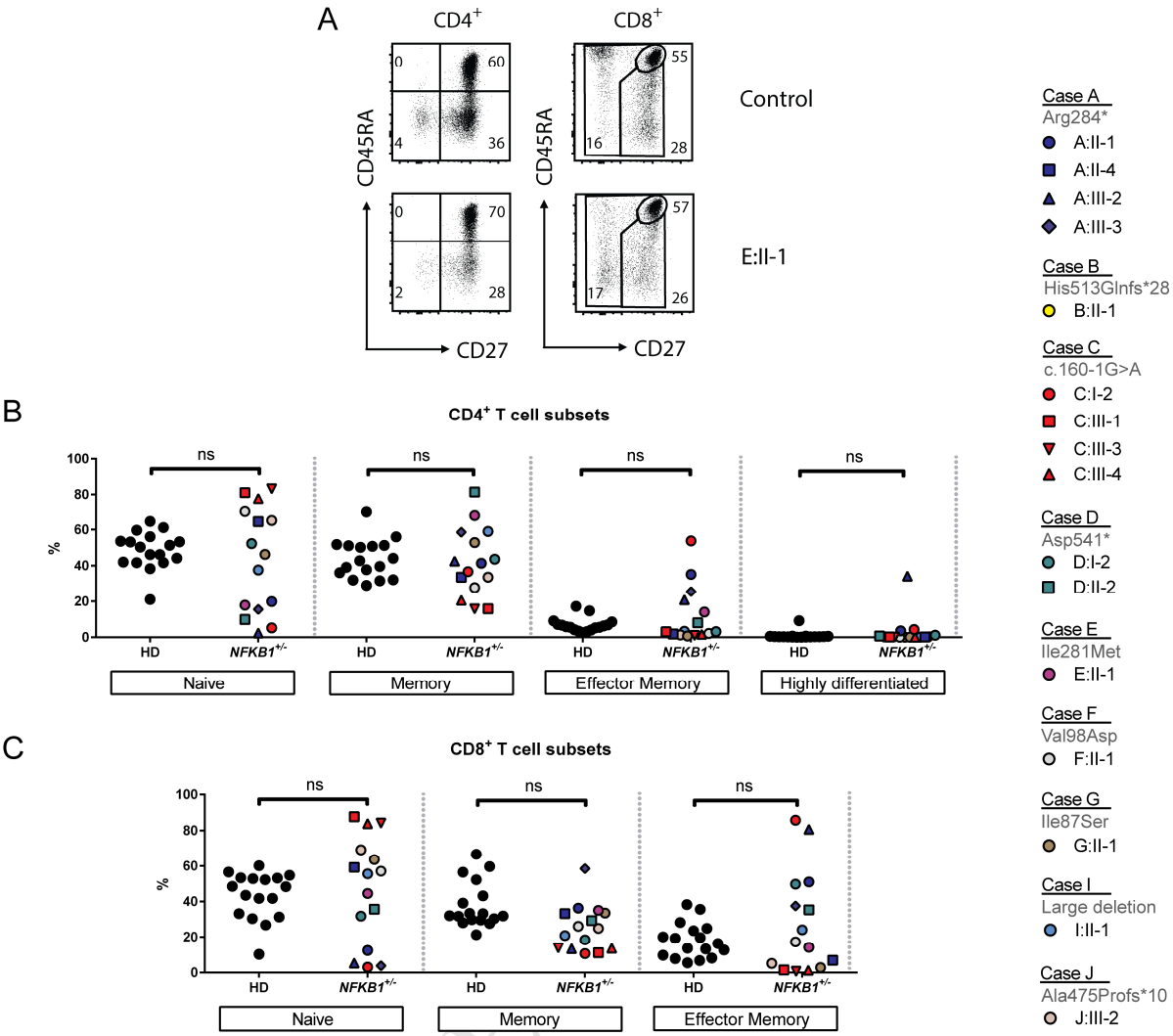
Case B
His513Glnfs*28
B:II-1

Case C
c.160-1G>A
C:I-2
C:III-1
C:III-3
C:III-4

Case D
Asp541*
D:I-2
D:II-2

Case E
Ile281Met
E:II-1





Case A
Arg284*
A:II-1
A:II-4
A:III-2
A:III-3

Case B
His513Glnfs*28
B:II-1

Case C
c.160-1G>A
C:I-2
C:III-1
C:III-3
C:III-4

Case D
Asp541*
D:I-2
D:II-2

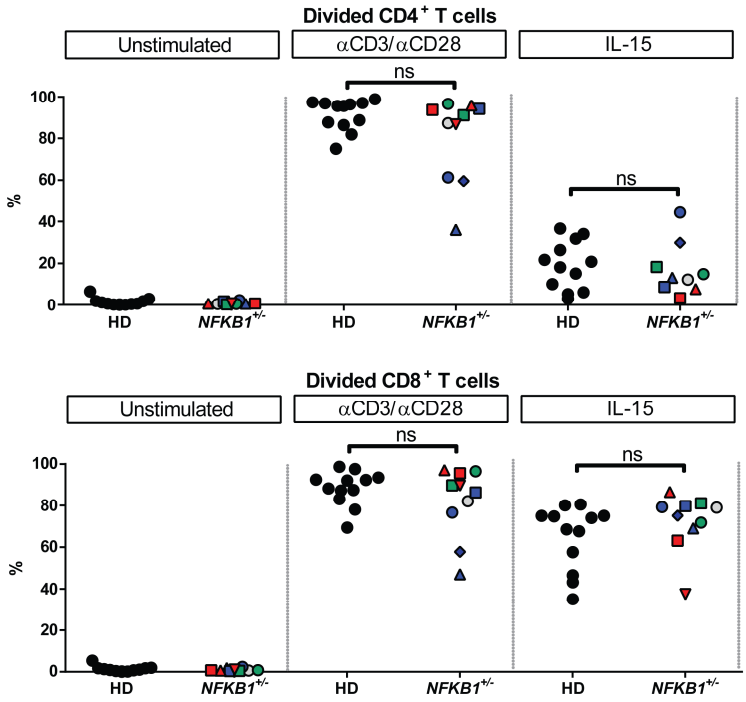
Case E
Ile281Met
E:II-1

Case F
Val98Asp
F:II-1

Case G
Ile87Ser
G:II-1

Case I
Large deletion
I:II-1

Case J
Ala475Profs*10
J:III-2



Case A
Arg284*
A:II-1
A:II-4
A:III-2
A:III-3

Case C
c.160-1G>A
C:I-2
C:III-3
C:III-4

Case D
Asp541*
D:I-2
D:II-2

Case F
Val98Asp
F:II-1