Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Glocker E-O, Kotlarz D, Boztug K, et al. Inflammatory bowel disease and mutations affecting the interleukin-10 receptor. N Engl J Med 2009;361. DOI: 10.1056/NEJMoa0907206.

Supplementary Appendix

1. Supplementary Materials and Methods

Patients and controls

All material from patients and healthy donors was obtained with informed assent/consent in accordance with the Declaration of Helsinki. This study was approved by institutional review boards at Hannover Medical School, the University of Freiburg and the Royal Free Medical School/University College London as part of human subjects' protocols to study the genetics of immunodeficiencies in humans.

Blood samples were taken from nine members of family A and six individuals from family B. In addition, we studied six patients with early-onset colitis (onset within first 12 months of life); 32 pediatric colitis patients with age of onset >12 months; 45 patients with adult-onset Crohn's disease (CD); 45 patients with adult-onset ulcerative colitis (UC); and 510 healthy controls. The 510 healthy controls comprised 180 unaffected Germans of European descent, 70 unaffected Turkish and 30 Iranian samples to test for sequence changes in exon 4 of *IL10RB* IL10R2^{Trp159X}); 100 healthy persons of German European ancestry to test for sequence changes in exon 3 of *IL10RA* (IL10R1^{Thr84Ile}); and 100 unaffected Arabic and 30 Iranian samples to test for sequence changes in exon 4 of *IL10RA* (IL10R1^{Gly141Arg}). The 280 controls used to test *IL10RB* were entirely distinct from the 230 controls used to test *IL10RA.* Ancestry was determined by physiognomological assessment.

Genotyping and analysis of genotypes in family A

We collected DNA from the six members of family A (Fig. 2A) and did a coarse genome scan using 253 microsatellite markers, followed by fine mapping near markers that showed promising segregation patterns. Reagents for some fine mapping markers were purchased from Invitrogen Research Genetics (Karlsruhe, Germany) and biomers.net GmbH (Ulm, Germany) and Qiagen GmbH (Hilden, Germany). PCRs for genotyping were performed according to the protocols accompanying the reagents; amplicons were sequenced on an ABI377 sequencer (PE Applied Biosystems, US) using COLLECTION and ANALYIS software. Integer allele lengths were assigned using GENOTYPER (PE Applied Biosystems, US) software.

Genotypes were analyzed with software specifically developed for this project to look for markers (or intervals) with perfect segregation. In the context of pedigree A, we define *perfect segregation* to mean that: 1) both affected children have the same genotype (or haplotype), 2) both affected children have a homozygous genotype (haplotype) and 3) if both parents are heterozygous, then the two unaffected children each have a genotype (haplotype) that differs from the affected siblings. We define *consistent segregation* to mean conditions 1 and 3 only because such segregation is consistent with autosomal recessive inheritance and compound heterozygous mutations. Compound heterozygosity is possible, but considered unlikely due to the consanguinity.

Due to the small family size, we expected to find (and did find) a small number (> 1) of intervals in the human genome that show perfect segregation. For this family structure, a haplotype interval of perfect segregation achieves a maximum LOD score of approximately 2.0. We expected that the genome scan would include multiple isolated markers with perfect segregation, and fine mapping would prove that some of these are within larger intervals of perfect segregation (Supplementary Table S1), while others are within intervals of consistent, but imperfect segregation.

Genome-wide homozygosity mapping and LOD score computations in index family B

DNA samples of nine family members of family B (Fig. 2C) were genotyped using the Affymetrix Genome-Wide Human SNP 6.0 mapping array following the Affymetrix standard protocol (Affymetrix, Santa Clara, CA). Briefly, 250 ng of each genomic DNA was aliquoted twice to wells of a 96-well plate following the proposed workflow in the Affymetrix Cytogenetics Copy Number Assay User Guide (Affymetrix P/N 702607 Rev. 1) and digested

with both *Nsp*I and *Sty*I on the same plate. The reaction was incubated at 37°C for 2 hours and at 65°C for 20 minutes to deactivate the enzymes. The digested DNA was then ligated to their respective *Nsp*I and *Sty*I adaptors and subsequently PCR-amplified using a single adaptor primer. PCR products with an average size of 200 – 1100 bp from both reactions were pooled and purified using Agencourt AMPure® magnetic beads (Agencourt Bioscience Corporation, Beverly, MA). Purified PCR products were fragmented, biotin-labeled and hybridized for 16 hours at 50°C. After washing and staining with the Affymetrix FS450 fluidic station, the arrays were scanned using the Affymetrix GeneChip Scanner 3000 7G. Raw image files were converted into CEL files by the Affymetrix GeneChip Command Console software Version 1.0. Genotypes of the SNPs were called by the Birdseed v2 algorithm implemented in the Genotyping Console software version 3.0.2.[1,2] The mean call rate among the nine samples was 99.5 %.

SNP genotypes were analyzed using in-house software. The software identifies intervals in which the affected individuals share the same homozygous genotypes, but in which each unaffected individual has a genotype different from the affected individuals at least some markers.

We used the software Superlink[3] to compute LOD scores for some of the 1907 SNPs in a 5Mbp region on chromosome 11q that includes a perfectly segregating 3.6Mb subinterval. *Perfect segregation* for a single SNP means that:

a. All nine available individuals have a genotyped called.
b. Both parents are heterozygous.
c. The affected child, B-II-5, is homozygous.
d. All six unaffected children have genotypes different from the genotype of B-II-5.

We used the information that the parents are second cousins, but we did not assume more consanguinity.

The disease is modeled as fully penetrant, recessive with a disease allele frequency of 0.001. The LOD scores depend on the frequency of the disease-associated marker allele. After identifying the 290 SNPs that segregate perfectly, we chose 11 SNPs at which to compute single marker LOD scores, and obtained their affected allele frequencies from HapMap.[4] These 11 SNPs were selected as described below so that no two have extreme linkage disequilibrium (see definition below). The affected allele frequency of each SNP is taken to

be the largest (worst) frequency among the four populations whose haplotypes are described in HapMap. The 11 SNPs and their LOD scores are shown in Supplementary Table S4.

We also used Superlink to compute LOD scores for multiple subsets of the SNPs shown in Supplementary Table S4. The presence of multiple perfectly segregating markers in the linkage region increases the evidence for linkage, and thus the LOD score, provided that the markers are not in extreme linkage disequilibrium. The HapMap project maintains linkage disequilibrium data for four populations: Yoruba in Ibadan, Nigeria; Japanese in Tokyo, Japan; Han Chinese in Beijing, China; and CEPH (Utah residents with ancestry from northern and western Europe). We consider two SNPs to be in *extreme linkage disequilibrium* with each other if HapMap explicitly lists an r^2-value of at least 0.8 between the SNPs in any of the four populations, or if there is a third SNP with which we consider the first and second snips are in extreme linkage disequilibrium. This means that if there is a chain of SNPs with r^2-value of at least 0.8 connecting the first and second SNP, where the r^2-value is the maximum of the four populations, then the SNPs are considered to be in extreme linkage disequilibrium.

DNA extraction, PCR and DNA sequencing

DNA of study participants was isolated by using the Gentra Puregene Blood kit (Qiagen, UK) or the Peqlab Gold Tissue DNA kit (Peqlab Germany) according to the manufacturer's instructions. Coding regions of *IL10RA* (ENSG00000110324), *IL10RB* (Ensembl No. ENSG00000159113) and other candidate genes were amplified by using the following PCR conditions: initial denaturation at 95°C for 5 min, 35 cycles of 95°C for 30 sec, 56°C for 30 sec, and 72°C for 1 min, and final elongation at 72°C for 10 min. Amplicons were purified using the QIAquick PCR purification kit (Qiagen, UK/Germany) and sequenced with the ABI PRISM BigDye Terminator cycle sequencing ready reaction kit v3.1 (Applied Biosystems, UK). The primers (all purchased from Invitrogen, UK) used for amplification and sequencing of all exons and intron-exon boundaries of the human *IL10RA* and *IL10RB* genes are listed in Supplementary Tables S5 and S6, respectively. DNA sequencing was accomplished with an ABI Prism 3130 or ABI 3130xl DNA Sequencer, respectively (Applied Biosystems, UK). The sequence reads were analyzed with Sequencing Analysis software and the DNA Sequencing Analysis software version 3.4 (Applied Biosystems) or Sequencher® software (Genecodes, US), respectively.

RNA isolation, reverse transcription and quantification of gene expression

Preparation of total RNA extracts was performed using the RNeasy Minikit (Qiagen, UK) and the Absolutely RNA™ kit (Stratagene, Germany), respectively. Subsequent cDNA synthesis was performed using Expand Reverse Transcriptase (Roche, Germany) or Omniscript Reverse Transcriptase (Qiagen, UK) according to standard protocols.

Gene expression analysis was performed on a LightCycler 2.0 using a LightCycler-FastStart DNA Master SYBR Green I kit (Roche, UK). For *SOCS3* amplification, we used *SOCS3* and *β-actin* Quantitect primer assays (Qiagen, UK). Cycling conditions consisted of an initial activation of 15 min at 95°C, 35 cycles of denaturation at 95°C for 30 s, annealing at 55°C for 30 s and elongation at 72°C for 30 s, followed by melting curve analysis to identify amplification of the correct PCR product. Quantification of *SOCS3* transcripts was accomplished by using the comparative C_T method with normalization to *β-actin*.

Flow Cytometry

Single cell suspensions were resuspended in PBS supplemented with 2% fetal calf serum (FCS, Gibco, UK/Germany), acquired on FACSCanto/LSR II flow cytometer (BD Biosciences, UK/Germany) and analyzed using FlowJo software (Tree Star, US). For cell surface staining monoclonal, PE-conjugated anitbodies recognizing IL10R1 and IL10R2, respectively, and respective isotyp controls were used according to manufacturer's instructions. Fluorescence intensity plots are shown in log scales.

Western blot analysis

Western blot analyses were performed according to standard protocols. Primary antibodies used were as follows: anti-phospho STAT3-Tyr705 and anti-phospho-STAT3-Ser727 (Cell Signaling, Germany), anti-STAT3 (BD Transduction Laboratories, Germany and Cell Signaling, UK), anti-GAPDH (Santa Cruz Biotech, Germany and R&D, UK). HRP-conjugated anti-mouse (BD Transduction Laboratories) and anti-rabbit (Cell Signaling Technology and R&D) secondary antibodies were detected using a chemiluminescent substrate (Pierce Technology, Germany or GE Healthcare, UK). Images were captured on a Chemidoc XRS Imaging System (Bio-Rad, Munich, Germany) or on Amersham Hyperfilms (Amersham, UK).

ELISAs

Enzyme-linked immunosorbent assays (ELISAs) to detect secretion of TNFα in cellular

supernatants were performed using commercially available kits (BD OptEIA Human TNF ELISA Kit, BD Biosciences, Germany or TNF development kit, Peprotech, UK). Purified peripheral blood mononuclear cells (PBMCs) or monocyte derived macrophages generated as described previously[5] were stimulated with 100 ng/ml LPS (Sigma-Aldrich, Germany) alone or in combination with recombinant human IL10 (Peprotech GmbH, Germany or R&D, UK) at a concentration of 20 ng/ml. Supernatants were analyzed in triplicates using a Tecan Sunrise ELISA microplate reader (Tecan Group, Crailsheim, Germany). The STAT3 ELISA (Cell Signaling, UK) was used according to manufacturer's instructions.

Inflammatory cytokine array

Chemiluminescence-based detection of inflammatory cytokines and chemokines in cellular supernatants was performed using a commercially available kit (RayBio© Human Inflammation Antibody Array 3, RayBiotech, Inc., USA) according to the manufacturer's instructions. The cellular supernatants used for this assay were obtained in parallel to the samples for the ELISA assays.

Retroviral expression of IL10R1 and IL10R2 in HeLa cells for studies in IL10R1 deficiency

For studying the role of *IL10RA* deficiency, human *IL10RA* wildtype or mutant cDNA was cloned into a bicistronic retroviral MMP vector[6] containing murine *cd24* as a marker gene. In addition, human *IL10RB* wildtype cDNA was cloned into the same retroviral backbone, but containing *green fluorescent protein (GFP)* as a marker gene. VSVg-pseudotyped retroviral particles were generated by tripartite transfection into the HEK 293T cell line of MMP-based vectors together with the envelope plasmid and the packaging plasmid mPD.old.gag/pol.

HeLa cells were transduced with the retroviral vector encoding either for wildtype or mutant *IL10RA*, followed by transduction with the vector encoding wildtype human *IL10RB*. Cells were cultured for 48 - 96 hours and harvested prior to cytoplasmic protein purification and Western blot analysis of Tyr705 p-STAT3 expression.

Lentiviral reconstitution of IL10R2 in patient's cells for studies in IL10R2 deficiency

Full-length human *IL10RB* cDNA was PCR amplified and cloned into a lentiviral pSIN vector expressing the reporter gene green fluorescent protein (GFP), as described previously.[7]

IL10R2-expressing lentiviral particle were generated by transfecting 293T cells and used to infect patient's A-II-3 EBV transformed B cells. Transduced cells were analyzed for GFP expression and surface expression of IL10R2 by FACS. Functionality of IL10R2 was examined by detecting STAT3 phosphorylation after stimulation of transduced cells with 100 ng/mL IL10 for 20 min.

Sequence and structural analysis of IL10R1

We performed sequence and structure analysis on wildtype, human, interleukin-10 receptor 1 (IL10R1), and on two mutated versions of the receptor, using the paradigm developed by Thusberg and Vihinen[8] for a study of ELA2, as further refined by Salzer et al.[9] in a study of TACI. Similar analyses were not performed for IL10R2 because the mutation in IL10R2 is a nonsense change.

We used RefSeq sequence NP_0015149.1 (GenBank identifier 4504633) as the amino acid sequence of the wildtype precursor to IL10R1. The substitutions we evaluated were Thr84Ile and Gly141Arg, where the coordinates 84 and 141 are with respect to the start codon of NP_00151149.1. The structure of IL10R1 bound to IL10 is provided in the protein data bank (PDB) structure 1Y6K,[10] but due to low binding affinity there is no structure including both IL10R1 and IL10R2 bound to IL10.[11,12] We mutated the PDB structure 1Y6K using Swiss-PDBViewer[11] using the protocol of Bednarski et al.[13]

2. Supplementary Results

Homozygosity mapping and LOD score computations for pedigree B

For pedigree B, we identified eight perfectly segregating intervals greater than 1Mb (Supplementary Table 3). One of these intervals spans chr11:115.3–118.9Mb (NCBI human genome build 36.3) and includes the candidate gene *IL10RA*. For LOD score computations, we used 1907 SNPs in the larger interval chr11:115.0–120.0Mb. Supplementary Table 4 shows 11 markers that segregate perfectly with affection status, but are not in extreme linkage disequilibrium. Fig. 2C shows the genotypes at these markers. Multipoint LOD score analysis of various combinations of three or four markers, such as rs21576799, rs11216831, rs511134, yields LOD scores of roughly 2.5, which is approximately the maximum score achievable for this pedigree structure.

Structural and sequence analysis of IL10R1^{WT}, IL10R1^{Thr84Ile} and IL10R1^{Gly141Arg}

Both the substitutions Thr84Ile and Gly141Arg are predicted to be deleterious. Thr84 is a highly conserved position. Moreover, structural stability methods predict that the IL10R1^{Thr84Ile} mutation causes a destabilizing change in energy. Sequence conservation methods provide insufficient information to infer that the substitution IL10R1^{Gly141Arg} is deleterious. Methods that take into account the folded structure of IL10R1 predict that IL10R1^{Gly141Arg} leads to overpacking, a charge change, a change in hydrophobicity, and a large, destabilizing change in energy; FoldX[14] predicts an energy change of more than 28 kcal/mol.

For Thr84Ile, sequence conservation methods provide strong predictions. SIFT[15], PMUT[16] and PolyPhen[17] all predict that the protein is deleterious (SIFT score 0.01; PMUT prediction score, 0.8044 with fiability 6; PolyPhen PSIC score difference 2.050). Both Conseq[18] and Consurf[19] put the position in the highest category (score 9) of conserved positions, and Conseq predicts that the position is exposed and functional.

Swiss-PdbViewer predicts a hydrogen bond between the oxygen in the side-chain of Thr84, and the nitrogen in the peptide bond between His40 and His41 (Supp. Fig. 2E). This hydrogen bond is abrogated by the substitution Thr84Ile (Supp. Fig. 2F). Probe reports severe steric clashes between the side-chain of Ile84 and both His40 and His41. The FoldX program computes a destabilizing energy change of 3.31 kcal/mol with the largest contributor to the

change being Van der Waals clashes. FoldX also predicts a destabilizing loss of sidechain hydrogen bonds. PopMusic[20] calculates a change of 0.98 kcal/mol. RankViaContact[21] predicts that Thr84Ile is stabilizing, but uses a coarse model that only considers the residues that are in contact, and not the relative orientation of the atoms within the residues.

The sequence conservation methods SIFT and PMUT did not predict IL10R1^{Gly141Arg} to be deleterious, but PolyPhen predicts that the variant is probably damaging based on sequence conservation (PSIC score difference 2.274). PolyPhen, which predicts based on multiple methods including structural analysis when PDB structures are available, also reports that the substitution results in a hydrophobicity change, a charge change, and overpacking at a buried site.

Structural stability methods predict a large change in stability resulting from the IL10R1^{Gly141Arg} substitution. RankViaContact calculates that Gly141 has the tenth strongest contact energy in the PDB structure 1Y6K, with a stabilizing contact energy of -18.05 RT units. In the mutated structure, RankViaContact concludes that the position has rank 74 and contact energy of -7.22 RT units, a destabilizing change of 10.83 RT units. Furthermore, RankViaContact identifies four residues in the mutant structure that contribute energy terms of more than 1 RT unit in magnitude, but do not make such contributions in the wildtype structure; among these Ile139 is physically closest to the Arg141 at a distance of approximately 6 Å. Conversely, Leu140 provides a stabilizing -2.98 RT unit contribution in the wildtype structure that disappears in the mutant structure. Fig. 3G and 3H show the positions of the wildtype amino acid, the mutated amino acid, and five other residues in both structures. The residues shown in Fig. 3G and 3H are in proximity to position 141 because they participate with 141 in a beta sheet.

PopMusic predicts that the mutation Gly141Arg is destabilizing to the PDB structure with a $\Delta\Delta G$ value of 2.6 kcal/mol. Molprobity[22], using the Probe[23] method, does not report severe steric clashes due to the substitution. This is consistent with the results of the mutation reported by Swiss PDB-Viewer; the substitution is destabilizing but the steric clashes are not severe. FoldX, run with the wildtype 1Y6K structure and the mutated structure generated by Swiss-PdbViewer, reports a destabilizing energy change of 28.12 kcal/mol.

3. Supplementary Tables

Supplementary Table S1. Characteristics of three intervals with perfect segregation in pedigree A.

Characteristics of the three intervals found with perfect segregation of multiple markers and disease status in the family, such that the two affected individuals are homozygous for all markers in the minimal interval.

	Chromosome 2	Chromosome 7	Chromosome 21
Genome scan marker	D2S3128	D7S3047	D21S2052
Maximal interval	113.7 – 127.8Mb	8.4 – 15.2Mb	20.5 – 33.9Mb
Minimal interval	117.3 – 126.3Mb	9.4 – 13.7Mb	23.7 – 33.5Mb
Total number of genes above minimal interval, in maximal interval	10	1	1
Total number of genes in minimal interval	26	8	69
Total number of genes below minimal interval, in maximal interval	5	0	8

Supplementary Table S2. Candidate genes in pedigree A.

Genes on chromosomes 2, 7, and 21 that were in the intervals consistent with linkage and a homozygous mutation segregating perfectly with the disease, as of November 2007 when the mutation in *IL10RB* was found. Due to improved annotation of the human genome, there have been some changes in the predicted genes since November 2007. The five genes we sequenced are marked with a * next to the gene symbol. Gene sequencing and fine mapping of the three regions were done concurrently; therefore, sequencing of *IL10RB* was deferred until we proved (Figure 2A) that *IL10RB* is in the maximal region consistent with linkage on chromosome 21.

Chromosome 2

Gene symbol	Gene Name
PAX8	paired box 8
CBWD2	COBW domain containing 2
FOXD4L1	forkhead box D4-like 1
FAM39B	family with sequence similarity 39, member B
LOC728007	similar to DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 11 (CHL1-like helicase homolog, S. cerevisiae)
RABL2A	RAB, member of RAS oncogene family-like 2A
LOC645529	similar to Hypothetical acrosin-like protease
SLC35F5	solute carrier family 35, member F5
ACTR3	ARP3 actin-related protein 3 homolog (yeast)
DPP10	dipeptidyl-peptidase 10
DDX18	DEAD (Asp-Glu-Ala-Asp) box polypeptide 18
HTR5B	5-hydroxytryptamine (serotonin) receptor 5B
CCDC93	coiled-coil domain containing 93
INSIG2	insulin induced gene 2
EN1	engrailed homeobox 1
MARCO	macrophage receptor with collagenous structure
C1QL2	complement component 1, q subcomponent-like 2
STEAP3	STEAP family member 3
DBI	diazepam binding inhibitor (GABA receptor modulator, acyl-Coenzyme A binding protein)
TMEM37	transmembrane protein 37
SCTR	secretin receptor
MGC33657	hypothetical protein MGC33657
TMEM177	transmembrane protein 177
PTPN4	protein tyrosine phosphatase, non-receptor type 4 (megakaryocyte)
EPB41L5	erythrocyte membrane protein band 4.1 like 5
TMEM185B	transmembrane protein 185B
RALB	v-ral simian leukemia viral oncogene homolog B (ras related; GTP binding protein)

INHBB	inhibin, beta B
FLJ14816	hypothetical protein FLJ14816
GLI2	GLI-Kruppel family member GLI2
TFCP2L1	transcription factor CP2-like 1
CLASP1	cytoplasmic linker associated protein 1
MKI67IP	MKI67 (FHA domain) interacting nucleolar phosphoprotein
TSN	Translin
hCG_2045614	hCG2045614
CNTNAP5	contactin associated protein-like 5
GYPC	glycophorin C (Gerbich blood group) (malaria susceptibility)
BIN1	bridging integrator 1
CYP27C1	cytochrome P450, family 27, subfamily C, polypeptide 1
ERCC3	excision repair cross-complementing rodent repair deficiency, complementation group 3 (xeroderma pigmentosum group B complementing)
MAP3K2	mitogen-activated protein kinase kinase kinase 2

Chromosome 7

Gene symbol	Gene Name
NXPH1	neurexophilin 1
NDUFA4	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 4
PHF14	PHD finger protein 14
THSD7A	thrombospondin, type I, domain containing 7A
TMEM106B	transmembrane protein 106B
FLJ14712	hypothetical protein FLJ14712
LOC389465	hypothetical gene supported by BT009862; NM_003142
*SCIN**	Scinderin
ARL4A	ADP-ribosylation factor-like 4A

Chromosome 21

Gene symbol	Gene Name
NCAM2	neural cell adhesion molecule 2
*BIC/ miRNA155**	microRNA 155
MRPL39	mitochondrial ribosomal protein L39
*JAM2**	junctional adhesion molecule 2
ATP5J	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit F6
GABPA	GA binding protein transcription factor, alpha subunit 60kDa
APP	amyloid beta (A4) precursor protein (peptidase nexin-II, Alzheimer disease)
CYYR1	cysteine/tyrosine-rich 1
ADAMTS1	ADAM metallopeptidase with thrombospondin type 1 motif, 1
ADAMTS5	ADAM metallopeptidase with thrombospondin type 1 motif, 5
C21orf94	chromosome 21 open reading frame 94
C21orf100	chromosome 21 open reading frame 100

N6AMT1	N-6 adenine-specific DNA methyltransferase 1
ZNF294	Zinc finger protein 294
RWDD2B	RWD domain containing 2B
USP16	ubiquitin specific peptidase 16
CCT8	chaperonin containing TCP1, subunit 8 (theta)
C21orf7/ TAK1L	chromosome 21 open reading frame 7
BACH1	BTB and CNC homology 1, basic leucine zipper transcription factor 1
C21orf41	chromosome 21 open reading frame 41
GRIK1	glutamate receptor, ionotropic, kainate 1
CLDN17	claudin 17
*CLDN8**	claudin 8
KRTAP24-1	keratin associated protein 24-1
KRTAP26-1	keratin associated protein 26-1
KRTAP27-1	keratin associated protein 27-1
KRTAP23-1	keratin associated protein 23-1
KRTAP13-2	keratin associated protein 13-2
KRTAP13-1	keratin associated protein 13-1
KRTAP13-3	keratin associated protein 13-3
KRTAP13-4	keratin associated protein 13-4
KRTAP15-1	keratin associated protein 15-1
KRTAP19-1	keratin associated protein 19-1
KRTAP19-2	keratin associated protein 19-2
KRTAP19-3	keratin associated protein 19-3
KRTAP19-4	keratin associated protein 19-4
KRTAP19-5	keratin associated protein 19-5
KRTAP19-6	keratin associated protein 19-6
KRTAP19-7	keratin associated protein 19-7
KRTAP6-3	keratin associated protein 6-3
KRTAP6-2	keratin associated protein 6-2
KRTAP22-1	keratin associated protein 22-1
KRTAP6-1	keratin associated protein 6-1
KRTAP20-1	keratin associated protein 20-1
KRTAP20-2	keratin associated protein 20-2
KRTAP21-2	keratin associated protein 21-2
KRTAP21-1	keratin associated protein 21-1
KRTAP8-1	keratin associated protein 8-1
KRTAP7-1	keratin associated protein 7-1
KRTAP11-1	keratin associated protein 11-1
KRTAP19-8	keratin associated protein 19-8
TIAM1	T-cell lymphoma invasion and metastasis 1
SFRS15	splicing factor, arginine/serine-rich 15
HUNK	hormonally upregulated Neu-associated kinase
C21orf45	chromosome 21 open reading frame 45
MRAP	melanocortin 2 receptor accessory protein
SNORA80	small nucleolar RNA, H/ACA box 80
C21orf119	chromosome 21 open reading frame 119
C21orf63	chromosome 21 open reading frame 63
C21orf77	chromosome 21 open reading frame 77
TCP10L	t-complex 10 (mouse)-like
C21orf59	chromosome 21 open reading frame 59
SYNJ1	synaptojanin 1
C21orf66	chromosome 21 open reading frame 66

C21orf49	chromosome 21 open reading frame 49
C21orf62	chromosome 21 open reading frame 62
OLIG2	oligodendrocyte lineage transcription factor 2
OLIG1	oligodendrocyte transcription factor 1
hCG_2045804	hCG2045804
IFNAR2	interferon (alpha, beta and omega) receptor 2
*IL10RB**	interleukin 10 receptor, beta
IFNAR1	interferon (alpha, beta and omega) receptor 1
IFNGR2	interferon gamma receptor 2 (interferon gamma transducer 1)
TMEM50B	transmembrane protein 50B
C21orf55	chromosome 21 open reading frame 55
GART	phosphoribosylglycinamide formyltransferase, phosphoribosylglycinamide synthetase, phosphoribosylaminoimidazole synthetase
SON	SON DNA binding protein
DONSON	downstream neighbor of SON

Supplementary Table 3. Genomic intervals of size > 1Mb where the markers in Pedigree B are perfectly segregating and the affected individual is homozygous.

Positions are in Mb from the upper telomere. The interval on chromosome 11 contains the mutated gene *IL10RA*.

Chromosome	Perfect interval (in Mbp)	Length of interval (in Mbp)
3	50.3 – 51.8	1.5
4	80.6 – 85.6	5
4	106.7 – 113.4	6.7
5	174.4 – 178.6	4.2
6	102.7 – 104.1	1.4
10	110.3 – 119.9	9.6
11	115.3 – 118.9	3.6
13	84.7 – 94.6	9.9

Supplementary Table S4. Segregation of markers on Chromosome 11 for the linkage analysis in pedigree B.

Markers in chr11:115.0–120.0Mbp that segregate perfectly and are not in extreme linkage disequilibrium. All but one of these markers was selected so that the less frequent allele is associated with the disease in patient B-II-5. The first column lists the NCBI reference SNP cluster identifier. The second column lists the position of the SNP on chromosome 11 of the NCBI reference assembly, build 36. The third column is the allele found to be homozygous in the affected individuals and the fourth column is the maximum frequency of this allele among the four populations mapped by HapMap. The fifth column is the single marker LOD score for this SNP.

SNP	Position (bp)	Disease-Associated Allele	Allele Frequency	LOD Score
rs4938219	115167274	G	0.744	0.87
rs2156799	115582273	C	0.058	1.88
rs12362410	115928070	T	0.289	1.27
rs4938316	116269792	A	0.156	1.51
rs1056136	116626208	G	0.364	1.17
rs2298767	116879680	T	0.158	1.51
rs11825655	117136077	C	0.200	1.42
rs11216831	117622228	T	0.100	1.68
rs525485	118060964	T	0.444	1.09
rs511134	118431145	A	0.092	1.71
rs3829261	118681709	C	0.110	1.65

Supplementary Table S5. Primers for *IL10RA* sequencing.

The first number after *IL10RA* refers to the exon; exon 7 is long and was therefore subdivided into several PCR products.

Primer identification	Oligo sequence (5' → 3')	Annealing temperature
IL10RA 1F	GACAGTGGTTCCCCGTCC	56°C
IL10RA 1R	CACTGGATGGAGAACTTTAATGG	
IL10RA 2F	GAACCTCCCTTTCTTCTTTGG	56°C
IL10RA 2R	AGGCAGGTATCTTCCCATGC	
IL10RA 3F	GGCCTCTTGCGTCTCCC	56°C
IL10RA 3R	GCAGACATGGTGAGCTATGG	
IL10RA 4F	ATTCTGGAGGCAAAGTCTCG	56°C
IL10RA 4R	AGTTCCCAATGGCACACAAG	
IL10RA 5F	CTAAAGGCCCACCAGCTCTC	56°C
IL10RA 5R	ACGCGTTTTGGATTGCAC	
IL10RA 6F	AATGGATTTCATGGGACCAG	56°C
IL10RA 6R	ACTGGCTGGGAGGAAAAGAG	
IL10RA 7_1F	CGAGCTCTCCTCCTGGG	56°C
IL10RA 7_1R	CCTCAGGTAACCCTGGAATG	
IL10RA 7_2F	TGACAGTGGCATTGACTTAGTTC	56°C
IL10RA 7_2R	GTCCAGGCAGAGGAGCAG	
IL10RA 7_3F	CCTGGGCAGCTTTAACTCAG	56°C
IL10RA 7_3R	AGGTTCCCCATGTGACCATC	
IL10RA 7_4F	GCTGAAGTCAGCTCAGACCC	56°C
IL10RA 7_4R	CAGTGCCCAGTGGCTTATC	

Supplementary Table S6. Primers for *IL10RB* sequencing.

Primer identification	Oligo sequence (5' → 3')	Annealing temperature
IL10RB 1F	AGGGTAAAGAAGACCCTCAAA	56°C
IL10RB 1R	CCTAGTTGCGTCTCAGCAG	
IL10RB 2F	AGCCATAGAGGAGAACCAAGT	56°C
IL10RB 2R	ACCTAGAGATGACAGCAGTGG	
IL10RB 3F	ACCAATAGACTTGCTCAATGC	56°C
IL10RB 3R	ACAAGGCAAGATGATGACATT	
IL10RB 4F	CTACCCTTCTTAGCCATGTCA	56°C
IL10RB 4R	TCCGATCAGATCTTTTGACTC	
IL10RB 5F	CCTTCCACTGCTTAGTCATGT	56°C
IL10RB 5R	TATGGTGTGTGAAGGACTGTG	
IL10RB 6F	GGATTGTGATGGTTAAAATGC	56°C
IL10RB 6R	CCCTTTTACAAATAGCCTTCC	
IL10RB 7F	ATAGATTTTCCAGCCAGGAGT	56°C
IL10RB 7R	GCCCTGTTTCTCACAATTAAA	

4. Supplementary References

1. Affymetrix, Inc. BRLMM: An Improved Genotype Calling Method for the Mapping 500K Array Set.
2. Korn JM, Kuruvilla FG, McCarroll SA, et. Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. Nat Genet 2008;40:1253-60.
3. Fishelson M, Geiger D. Exact genetic linkage computations for general pedigrees. Bioinformatics 2002; 18 Suppl 1:S189-98.
4. Frazer KA, Ballinger DG, Cox DR, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007;449:851-61.
5. Buettner M, Meinken C, Bastian M, et al. Inverse correlation of maturity and antibacterial activity in human dendritic cells. J Immunol 2005; 174:4203-9.
6. Klein C, Bueler H, Mulligan RC. Comparative analysis of genetically modified dendritic cells and tumor cells as therapeutic cancer vaccines. J Exp Med 2000;191:1699-708.
7. Escors D, Lopes L, Lin R, et al. Targeting dendritic cell signaling to regulate the response to immunization. Blood 2008; 111:3050-61.
8. Thusberg J, Vihinen M. Bioinformatic analysis of protein structure–function relationships: case study of leukocyte elastase (*ELA2*) missense mutations. Hum Mutat. 2006;27:1230–43.
9. Salzer U, Bacchelli C, Buckridge S, et al. Relevance of biallelic versus monoallelic *TNFRSF13B* mutations in distinguishing disease-causing from risk increasing *TNFRSF13B* variants in antibody deficiency syndromes. Blood. 2009;113:1967–76.
10. Yoon SI, Jones BC, Logsdon NJ, Walter MR. Same structure, different function crystal structure of the Epstein-Barr virus IL-10 bound to the soluble IL-10R1 chain. Structure. 2005;13:551–64.
11. Yoon SI, Logsdon NJ, Sheikh F, Donnelly RP, Walter MR. Conformational changes mediate interleukin-10 receptor 2 (IL10-R2) binding to IL-10 and assembly of the singnaling complex. J Biol Chem. 2006; 281:35088-96.
12. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis.1997;18:2714–23.
13. Bednarski AE, Elgin SCR, Pakrasi HB. An inquiry into protein structure and genetic disease: introducing undergraduates to bioinformatics in a large introductory course. Cell Biol Educ. 2005;4:207–20.

14. Schymkowitz JWH, Rousseau F, Martins IC, Ferkinghoff-Borg J, Stricher F, Serrano L. Prediction of water and metal binding sites and their affinities by using the Fold-X force field. Proc Natl Acad Sci USA. 2005;102:10147–52.
15. Ng PC, Henikoff S. Accounting for human polymorphisms predicted to affect protein function. Genome Res. 2002;12:436–46.
16. Ferrer-Costa C, Orozco M, de la Cruz X. Sequence-based prediction of pathological mutations. Proteins. 2004;57:811–19
17. Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. Nucleic Acids Res. 2002;30:3894–3900.
18. Pupko T, Bell RE, Mayrose I, Glaser F, Ben-Tal N. Rate4Site: an algorithmic tool for the identification of functional regions in proteins by surface mapping of evolutionary determinants within their homologues. Bioinformatics. 2002;18 Suppl 1:S71–7.
19. Landau M, Mayrose I, Rosenberg Y, et al. ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. Nucleic Acids Res. 2005;33:W299–W302.
20. Kwasigroch JM, Gilis D, Dehouck Y, Rooman M. PoPMuSiC, rationally designing point mutations in protein structures. Bioinformatics.2002;18:1701–2.
21. Shen B, Vihinen M. RankViaContact: ranking and visualization of amino acid contacts. Bioinformatics. 2003;19:2161–2.
22. Lovell SC, Davis IW, Arendall WB 3rd, et al. Structure validation by Cα geometry: φ, ψ, and Cβ deviation. Proteins. 2003;50:437–50.
23. Word JM, Lovell SC, Richardson JS, Richardson DC. Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. J Mol Biol. 1999;285:1735–47.

Supplementary Figure 1. Additional data on patient A-II-3.

Panels A and B illustrate the pathohistology of colonic resection specimens from patient A-II-3 with multifocal ulcerations and intramural abscesses in colon mucosa (A), which consisted predominantly of neutrophil granulocytes indicating florid inflammation (B). Panel C shows phosphorylation of STAT3 (Tyr705) in EBV transformed B cells derived from a healthy donor (lanes 1 to 3) and patient A-II-3 (lanes 4 to 6) upon stimulation with IL10 and IFNα (positive control), respectively. FACS blots displayed in Panel D illustrate reconstitution of $IL10R2^{WT}$ expression in index patient A-II-3 cells upon lentiviral gene transfer. Panel E displays reconstitution of IL10-dependent STAT3 phosphorylation at the tyrosine 705 residue upon lentiviral transduction of wildtype IL10R2 in EBV transformed B cell line derived from patient A-II-3, as determined by ELISA. Panel E also shows data for a healthy donor. Results in panel E are representative of two independent experiments.

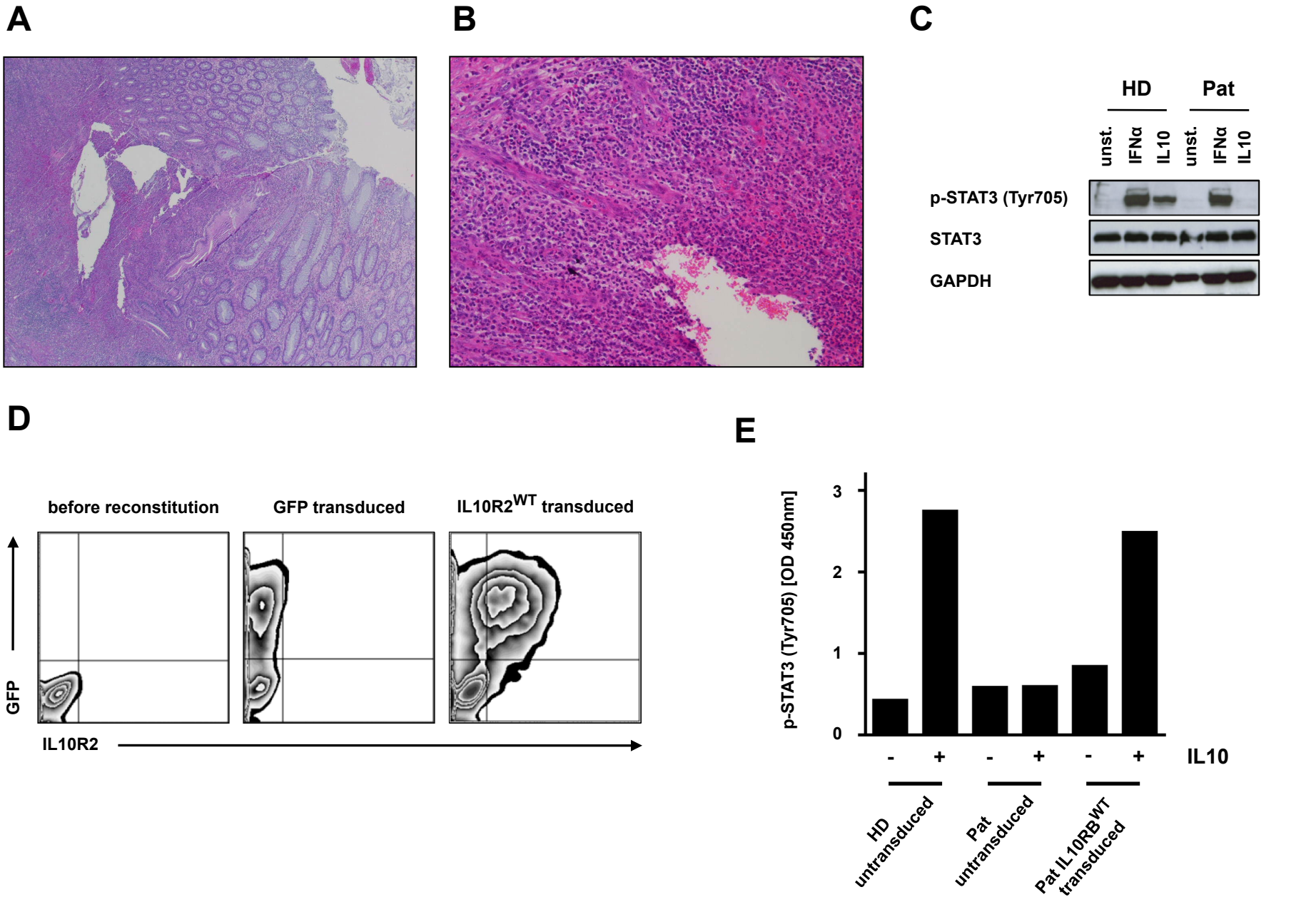
Supplementary Figure 2. Imaging, mutational analyses and functional data on *IL10RA*^{Thr84Ile} mutation.

Panel A illustrates the homozygous single nucleotide exchange leading to a nonsense mutation in exon 4 (c.G477A; p.Trp159X). Panel B illustrates the homozygous missense mutation in exon 4 of the *IL10RA* gene (c.G421A; p.Gly141Arg) identified in index patient B-II-5. Panel C illustrates mutational analysis of the *IL10RA* gene, showing a homozygous missense mutation in exon 3 of the *IL10RA* gene (c. C325T; p. Thr84Ile). Panel D shows data from a coloscopy in the additional, unrelated IL10R1-deficient patient illustrating multiple erosive lesions in the colon. Panels E and F show histopathological sections of this patient, revealing an erosive and ulcerative colitis with cryptic abscesses and neutrophilic infiltrates. Panel G shows normal expression levels of IL10R1 in patient B-II-5, as demonstrated by Western blot; GAPDH is used as a control. Panel H shows wildtype IL10R1, from PDB structure 1Y6K chain R. Residues Thr84, His40 and His41 are labeled, and nearby residues are shown. The figure shows the hydrogen bond between the oxygen at the tip of Thr84 and the peptide bond between His40 and His41. Panel I displays the predicted structure changes of IL10R1 with the substitution Thr84Ile. Residues Thr84, His40 and His41 are labeled, and nearby residues are shown. The hydrogen bond between position 84 and the background, shown in left panel, has been broken. Panel J displays a Western blot analysis revealing defective IL10-induced STAT3 phosphorylation. Panel K shows the increased TNFα secretion of this patient compared to a healthy donor, as assayed by ELISA in supernatants of PBMCs after stimulation with LPS alone or in co-stimulation with IL10 after 6 hours.

Supplementary Figure 3. Inflammatory cytokine arrays in IL10-deficient patients.

Panels A and B of this figure shows the raw data from the inflammatory cytokine array in patient B-II-5. The densitometric analysis of this array is shown in Figure 4E,F in the main document. Labels in the figure refer to the labels used in Figure 4E,F. As displayed in the figure, each cytokine analyzed is spotted twice on the array. HD, healthy donor; Pat, patient; LPS, lipopolysaccharide.

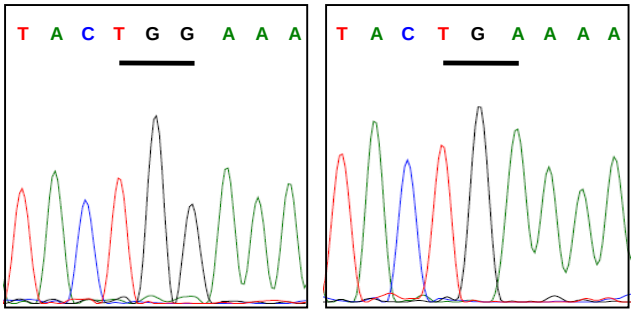
Panels C–F show data on the patient bearing *IL10RA*^{Ile84Thr} mutation. Similar to the patient with the *IL10RA*^{Gly141Arg} mutation, Panel C,D show a graphical representation of the increased secrection of proinflammatory cytokines upon LPS stimulation in the case of *IL10RA*^{Ile84Thr} mutation, which could not be counteracted by co-stimulation with exogenous IL10 in the patient. Numbering refers to the spots labeled in Supplementary Figure 3E,F. Panels E and F represent the respective source data from the inflammatory cytokine arrays analyzed in Panel C,D.



Supplementary Figure 1)

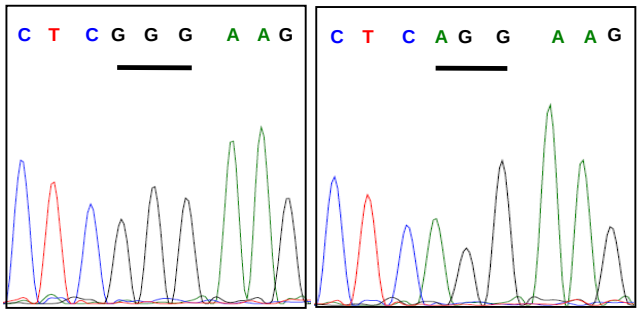
A

HD (*IL10RB*^{Trp159}) Pat (*IL10RB*^{Trp159X})



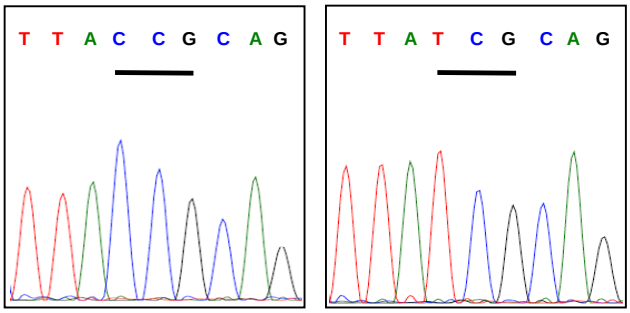
B

HD (*IL10RA*^{Gly141}) Pat (*IL10RA*^{Gly141Arg})

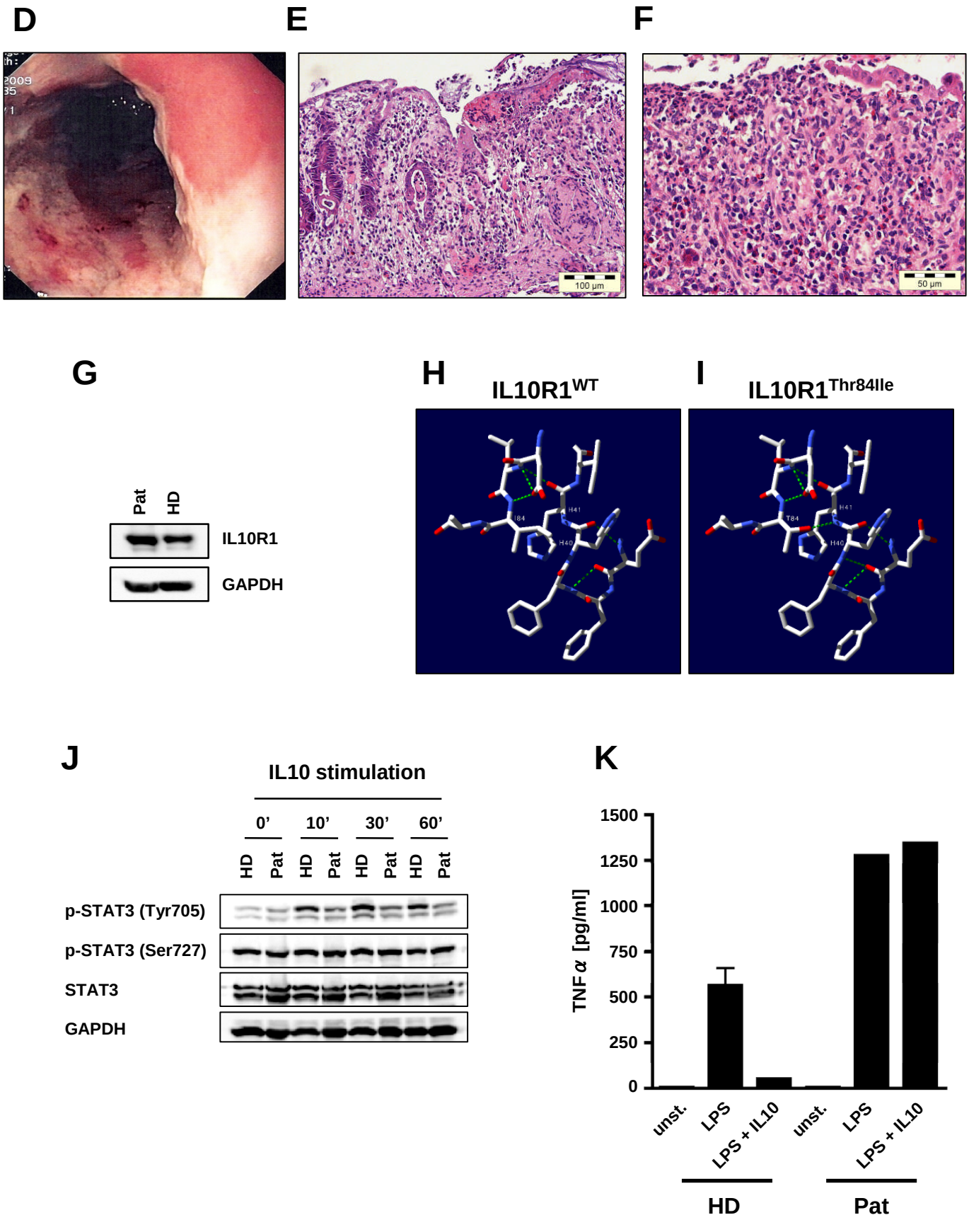


C

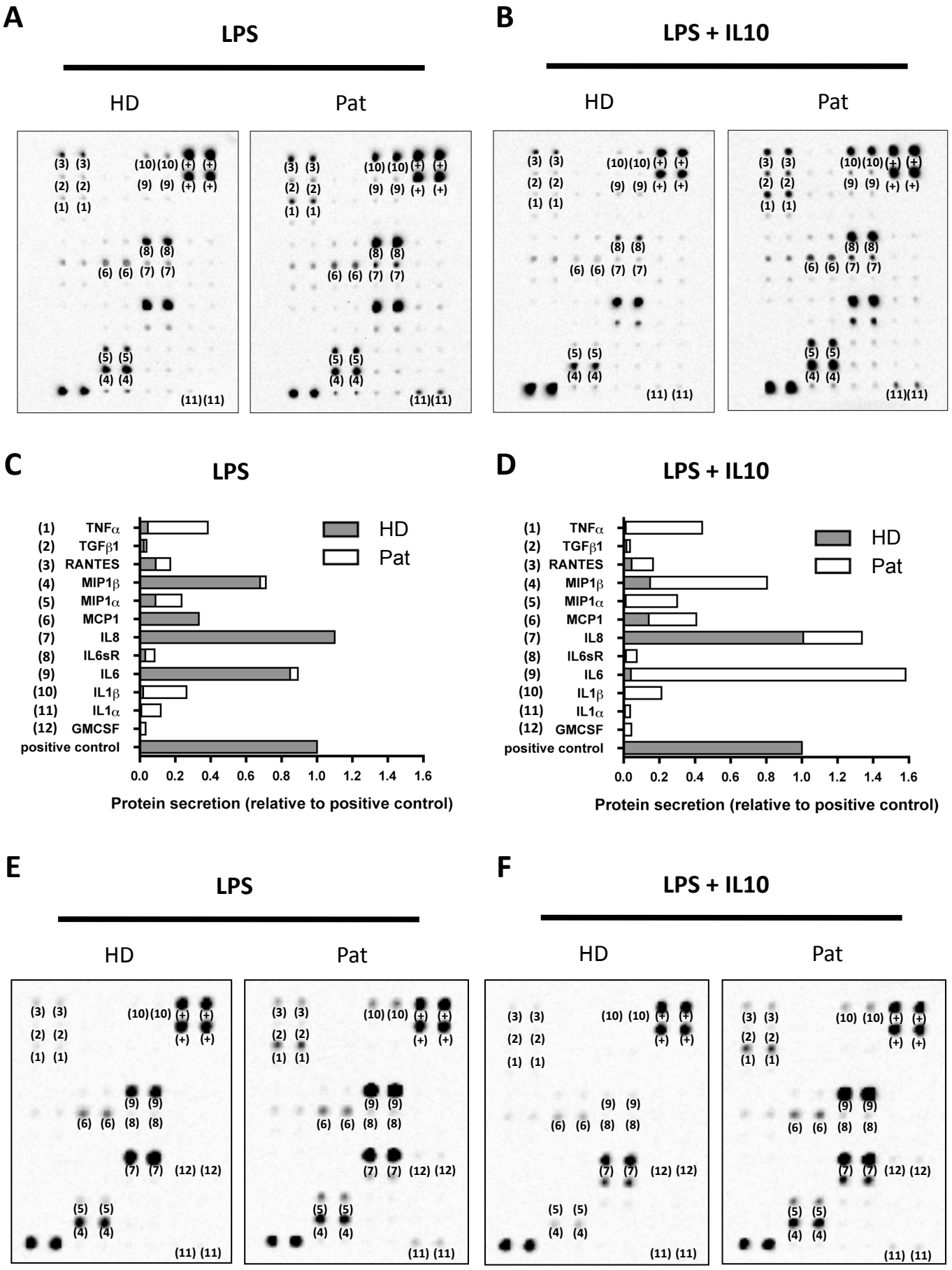
HD (*IL10RA*^{Thr84}) Pat (*IL10RA*^{Thr84Ile})



Supplementary Figure 2)



Supplementary Figure 2)



Supplementary Figure 3)