Published in final edited form as:
Sci Transl Med. 2016 November 02; 8(363): 363ra149. doi:10.1126/scitranslmed.aag1974.

Resolving *TYK2* locus genotype-to-phenotype differences in autoimmunity

Calliope A. Dendrou[1], **Adrian Cortes**[1,2], **Lydia Shipman**[1], **Hayley G. Evans**[1], **Kathrine E. Attfield**[3], **Luke Jostins**[2], **Thomas Barber**[1], **Gurman Kaur**[3], **Subita Balaram Kuttikkatte**[3], **Oliver A. Leach**[1], **Christiane Desel**[1], **Soren L. Faergeman**[1,4], **Jane Cheeseman**[5], **Matt J. Neville**[5,6], **Stephen Sawcer**[7], **Alastair Compston**[7], **Adam R. Johnson**[8], **Christine Everett**[8], **John I. Bell**[9], **Fredrik Karpe**[5,6], **Mark Ultsch**[8], **Charles Eigenbrot**[8], **Gil McVean**[2], and **Lars Fugger**[1,3,4,*]

[1]Oxford Centre for Neuroinflammation, Nuffield Department of Clinical Neurosciences, Division of Clinical Neurology, John Radcliffe Hospital, University of Oxford, Oxford OX3 9DS, UK

[2]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK

[3]MRC Human Immunology Unit, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, University of Oxford, Oxford OX3 9DS, UK

[4]Department of Clinical Medicine, Aarhus University Hospital, 8200 Aarhus N, Denmark

[5]Oxford Centre for Diabetes, Endocrinology, and Metabolism, University of Oxford, Oxford OX3 7LE, UK

[6]The National Institute for Health Research (NIHR) Oxford Biomedical Research Centre, OUH Trust, Churchill Hospital, Oxford OX3 7LE, UK

[7]Department of Clinical Neurosciences, University of Cambridge, Cambridge CB2 0QQ, UK

[8]Structural Biology and Biochemical Pharmacology, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA

[*]Correspondence to: lars.fugger@imm.ox.ac.uk.

Overline: Autoimmunity

Author contributions. C.A.D. contributed to the conception, coordination and design of the study, all experiments apart from the structural biology, and drafting and writing of the manuscript. A.Cortes performed all statistical genetics, epigenetic and next generation sequencing, and UK Biobank data analyses and contributed to the writing of the manuscript. L.S. contributed to the blood sample processing and analyses and the study conception and design, and L.S. and S.L.F. contributed to preliminary immunoblot analyses. H.G.E. generated the CRISPR/Cas9-modified cell lines, contributed to blood sample processing and performed the FACS sorting. K.E.A. and G.K. performed the animal immunizations, K.E.A. contributed to murine cell activation experiments and performed all murine perfusions, and C.D. contributed to flow cytometry of murine samples. L.J. contributed to the design and execution of genetic and statistical analyses. T.B. and O.A.L. contributed to blood sample processing. S.B.K. contributed to TaqMan SNP genotyping. J.C., M.J.N. and F.K. contributed to the coordination and sample obtainment from the Oxford BioBank, and M.J.N. performed genotyping for volunteer pre-selection. S.S. and A. Compston contributed MS genetic and clinical data. M.U., A.R.J., C. Everett and C. Eigenbrot designed and performed the crystallography and relevant biophysical analysis, and contributed to the writing of the final manuscript. J.I.B. contributed to the writing of the final manuscript. G.M. contributed to the conception and design of genetic and statistical analyses and to the writing of the manuscript. L.F. contributed to conception, design and coordination of the study, and drafting and writing of the manuscript.

Competing interests. L. Jostins is a consultant to Genomics plc. M. Ultsch, A. R. Johnson, C. Everett and C. Eigenbrot hold equities in Roche Holdings. J. I. Bell is on the board of Roche and holds equities in Roche Holdings. G. McVean is co-founder, holder of shares in and consultant to Genomics plc. The other authors declare no competing financial interests.

Data and materials availability. The X-ray crystal structure of the Ala-1057 JAK2 kinase domain at 2.66Å resolution has been deposited in the Protein Data Bank (PDB accession 5HEZ).

Europe PMC Funders Author Manuscripts

[9]Richard Doll Building, Roosevelt Drive, University of Oxford, Oxford OX3 7DG, UK

Abstract

Thousands of genetic variants have been identified that contribute to the development of complex diseases, but determining how to fully elucidate their biological consequences for translation into clinical benefit is challenging. Conflicting evidence regarding the functional impact of genetic variants in the tyrosine kinase 2 (*TYK2*) gene, which is differentially associated with common autoimmune diseases, currently obscures the potential of TYK2 as a therapeutic target. We aimed to resolve this conflict by performing genetic meta-analysis across disorders, subsequent molecular, cellular, *in vivo* and structural functional follow-up and epidemiological studies. Our data revealed a protective homozygous effect that defined a signaling optimum between autoimmunity and immunodeficiency and identified TYK2 as a potential drug target for multiple autoimmune disorders.

Introduction

Genome-wide association studies (GWAS) have helped to define the landscape of genetic pre-disposition to common polygenic diseases, including immune-mediated, metabolic, cardiovascular and neurological conditions. Elucidation of the biological consequences of disease-associated genetic variation can advance our understanding of the pathophysiological mechanisms that promote these diseases (1). This may allow for the identification of new therapeutic targets. The development of drugs against new targets can be complicated as insufficient efficacy and safety of new compounds have often led to clinical trial suspension (2), but retrospective analysis indicates that drug development programs supported by human genetic evidence are more likely to be successful (3). Therefore, it is conceivable that resolving the functional implications of disease-associated genetic variation can be used prospectively not only to identify potential targets, but to evaluate them through pre-clinical risk-benefit prediction prior to drug discovery and trialing (4).

We have investigated the tyrosine kinase 2 (*TYK2*) gene region based on previous GWAS reports of different, independent association signals within the region across multiple autoimmune disorders (5–11). We reasoned that such a complex association pattern could provide a framework for cross-comparisons that could be of additional value in dissecting the functional consequences of associated variants. Moreover, *TYK2* itself is an attractive immunological candidate for the gene affected by the associated variants as it encodes a non-receptor tyrosine kinase that is constitutively expressed across immune cell types. TYK2 has been reported to transduce signals downstream of the type I interferon (IFN), gp130, interleukin (IL)-10R2, IL-13Rα and IL-12Rβ1 cytokine receptor families, and thus has a pleiotropic role in host responses to infection and in tumor surveillance (12). The index single nucleotide polymorphisms (SNPs) that have been reported as representative of the association signals in the region include non-coding variants as well as missense SNPs such as rs34536443 and rs12720356, which are located directly in the *TYK2* gene (5–11). Although both these SNPs have been suggested to reduce the function of TYK2 in mediating cytokine signaling (13–15), this is inconsistent with rs34536443 conferring protection

Europe PMC Funders Author Manuscripts

against the conditions it is associated with, whilst rs12720356 confers protection against some diseases but risk for others. To resolve this conflict and to thereby uncover the predictive power gained from comprehensively studying the *TYK2* region, we have used a precision target analysis approach that includes (i) genetic, epigenetic and transcriptomic meta-analysis, (ii) genotype-dependent *in vitro* cytokine signaling assessments, (iii) the use of a humanized mouse model, (iv) health record association analysis, and (v) structural investigations.

Results

Joint analysis of genetic associations in the *TYK2* region across autoimmune diseases

As at least eight different index SNPs in the *TYK2* region have been reported to be autoimmune disease-associated (5–11) (table S1), we sought to evaluate whether these represent completely distinct associations and to determine whether single SNPs or more complex haplotypes drive these associations. We enriched our analysis for genetic data from conditions where non-coding variants in the region have been reported as associated and where comprehensive fine-mapping has not been previously performed. The SNP data utilized were obtained from 8,726 ankylosing spondylitis (5), 4,017 Crohn's disease (6), 16,691 multiple sclerosis (8), 2,814 psoriasis (9), and 3,871 ulcerative colitis (6) patients and 19,738 controls (6, 8), all of European ancestry, who were genotyped on the Immunochip. For fine-mapping we imputed untyped SNPs using data from the 1,000 Genomes Project to infer haplotypes (16, 17), and we then jointly analyzed the patient and control data by multinomial logistic regression (18). The strongest association signal is seen at rs34536443 ($P = 3.87 \times 10^{-21}$), which is highly likely to be the causative variant as this signal is explained by this single SNP alone (Fig. 1A and table S2). After controlling for this effect by conditional analysis, a secondary association was observed at rs9797854 ($P = 2.51 \times 10^{-12}$), which is the index SNP representing an associated haplotype that includes a total of 19 SNPs (Fig. 1A and table S2). Controlling for the first two signals revealed a tertiary association at rs12720356 ($P = 5.67 \times 10^{-9}$), which represents a haplotype that includes two other SNPs within the 90% credible set (Fig. 1A and table S2). Further conditioning indicated no additional associations in the *TYK2* region, demonstrating that for the diseases analyzed, these three associations in fact explain the previously reported signals represented by numerous different non-coding variants.

We next assessed the effect on protection against or risk for disease associated with carrying one or two copies of the minor allele at the index SNPs for each of the three association signals. For both rs9797854 and rs12720356 the allelic dosage effect fits an additive model. However, the rs9797854 minor allele only confers protection against multiple sclerosis, as we discovered no evidence of association with the other autoimmune diseases tested. The rs12720356 minor allele was found to confer a previously unreported risk for ankylosing spondylitis, risk for Crohn's disease and ulcerative colitis, and protection against psoriasis, but was not associated with multiple sclerosis at all (Fig. 1B and table S3A). In contrast, the rs34536443 minor allele is protective across all of the diseases tested, and surprisingly, its dosage effect shows a previously unrecognized significant departure from additivity ($P = 5.97 \times 10^{-4}$) (table S3B), with minor allele homozygosity conferring more than double the

protective effect of carrying one such allele (Fig. 1C and table 3C). Notably, for ulcerative colitis rs34536443-mediated protection is only observable in the homozygous state. Collectively, the findings of our meta-analysis, along with reports of association with other diseases (table S1), pinpoint rs34536443 as the only SNP identified genome-wide to date that protects against 10 different autoimmune conditions (Fig. 1C). This indicates a greater sharing of genetically determined pathophysiological mechanisms across conditions than has been previously recognized (19), and rs34536443 is therefore the prime variant for informing cross-disease therapeutic development.

Distinguishing the functional impact of the *TYK2* region association signals

Rs34536443 leads to the substitution of a highly conserved proline residue with an alanine at position 1104 (P1104A) in the kinase domain of the TYK2 protein (20) (fig. S1A). Given the specificity of this molecular effect and noting that rs34536443 was not found to correlate with *TYK2* gene expression differences when analyzing expression quantitative trait locus (eQTL) data sets from unstimulated and stimulated immune cell subsets (21–24) and tissues (25, 26) (table S4), we considered whether the secondary and tertiary associations have any influence on TYK2 or not. The SNPs in the 90% credible set that constitute the secondary signal haplotype are all located in the *SLC44A2* gene (Fig. 1A and table S2), which is a solute carrier family member that has been implicated as an autoantibody target in autoimmune hearing loss (27). These SNPs include missense variants that are unlikely to have a direct influence on *TYK2*, but also non-coding variants, two of which are potentially regulatory based on their co-localization with epigenetic marks of open chromatin and weak enhancers in immune cell types (table S2). Meta-analysis of eQTL data sets (21–26) provided no support for a correlation of the secondary signal with TYK2 transcript expression (table S4), although an effect for cell types or stimulatory conditions not included in the data sets analyzed cannot be ruled out. However, a lack of an impact on TYK2 is in keeping with the disease-specificity of this association, which is in sharp contrast to the association pattern of rs34536443.

The tertiary association signal index SNP, rs12720356, results in an I684S substitution in the TYK2 pseudokinase domain (fig. S1A). This domain is thought to regulate the kinase activity of TYK2 (28). As for rs34536443, there is some evidence of a correlation between the I864S substitution and reduced TYK2 activity in overexpression systems and a small number of lymphoblastoid cell lines upon stimulation with type I IFN, although it is unclear whether this might be a direct effect or could be related to reduced expression of TYK2 protein (13, 15). However, an overall similar functional impact of rs34536443 and rs12720356 is at odds with their cross-disease association pattern. TYK2 impairment as conferred by the minor alleles at each of the two SNPs initially seems plausible based on the protective effect of these alleles against psoriasis, rheumatoid arthritis, systemic lupus erythematosus and type 1 diabetes, but by extension this impairment would confer both protection against and risk for ankylosing spondylitis, Crohn’s disease and ulcerative colitis (Fig. 1C).

To begin to resolve this discrepancy between the genetic associations and the functional correlations, the rs34536443 and rs12720356 variants were introduced into HEK293T cells

by CRISPR-Cas9-mediated genomic modification (fig. S1, B and C). This allowed *TYK2* expression to be regulated by its endogenous promoter. The modified cells were likely to have a more uniform genomic background compared to lymphoblastoid cell lines, which were derived from different individuals. Upon stimulation with interferon-beta (IFNβ), TYK2 phosphorylation was reduced in an rs34536443-dependent fashion, but not by rs12720356 genotype, and neither SNP had an impact on the expression of TYK2 protein (fig. S1D and E). Given that the relative impact of TYK2 impairment depends on the action of the other Janus kinases (JAKs) that it functions in concert with (12), we also assayed the phosphorylation of the TYK2 and JAK substrate, signal transducer and activator of transcription (STAT). Reduced STAT phosphorylation (pSTAT) was observed by rs34536443 but not rs12720356 genotype; this reduction was consistent with an additive genetic effect, suggesting that the non-additive impact of rs34536443 genotype on disease risk may mainly relate to phenotypic differences downstream of pSTAT. Rs34536443 *C/C* minor allele homozygous cell lines showed >50% lower pSTAT compared to *G/G* major allele homozygous lines, although this was not as low as in TYK2 knockout lines (fig. S2). These data do not support an effect of rs12720356 on TYK2 function. However, as the kinase mediates signaling downstream of multiple cytokine receptor families (12), we next compared the effects of rs34536443 and rs12720356 on signaling induced by different cytokines using primary human immune cells.

The primary but not tertiary association signal influences TYK2-mediated cytokine signaling in human immune cells

To fully resolve which cytokine signaling pathways are specifically influenced by SNP genotype, we performed stimulations with at least one cytokine for each of the cytokine receptor families that activate TYK2 (12). We hypothesized that type I IFN, IL-10, IL-12 and IL-23 signaling were most likely to be affected as these signaling pathways are impaired in TYK2 immunodeficient patients, who completely lack TYK2 protein (29), but it was unclear if a less severe, SNP-dependent impairment of TYK2 function might differentially influence signal transduction downstream of these cytokines. A prior report has suggested that rs34536443 genotype might also affect IL-6, in addition to type I IFN and IL-10 signaling in T cells. However, this study assessed only a small number of rs34536443 major allele homozygotes and heterozygotes, and no minor allele homozygotes (14). The impact of TYK2 absence or impairment on IL-13 signaling has not been investigated.

For our analyses, human peripheral blood immune cells were obtained from the Oxford BioBank resource as healthy volunteers could be selected and recalled based on their genotype (fig. S3A). Thus, despite rs34536443 minor allele homozygotes having a frequency in the UK general population of only ~0.20%, we managed to analyze seven such individuals, representing a 22-fold enrichment for this genotype in our cohort (total $n = 162$ individuals). Using optimized stimulation conditions that allowed for a maximal response to cytokine (fig. S3, B and C), we assayed the phosphorylation of the different STAT molecules known to mediate signaling downstream of the different cytokines tested. We found ~80% lower pSTAT-positivity in rs34536443 minor compared to major allele homozygotes in response to type I IFN across all immune cell subsets, when assaying STAT1 and the non-canonical type I IFN transducer STAT3. (fig. S4 and Fig. 2, A and B). The influence of

genotype on pSTAT in the primary cells was generally consistent with a non-additive effect (fig. S4, Fig. 2, A and B and table S5). The differences in the percentage of pSTAT-positive cells were mirrored by differences in the total mean fluorescence intensity of pSTAT in stimulated compared to unstimulated cells (table S5 and S6). No genotype-dependent effects were found after IL-6, IL-10 or IL-13 administration (Fig. 2C-E, fig. S4A and table S6), indicating that the other JAKs can compensate for the impairment of TYK2 at the respective receptors for these cytokines.

Although the type I IFN, IL-6, and IL-10 receptors are constitutively expressed across immune cell subsets and the IL-13 receptor is found on monocytes directly *ex vivo*, detecting pSTAT in response to IL-12 and IL-23 necessitated cellular pre-activation, at least in part to facilitate up-regulation of the receptors that bind these cytokines on $CD4^+$ and $CD8^+$ T cells. Following pre-activation, there was approximately 70% less IL-12 and IL-23 signaling in $CD4^+$ memory T cells from rs34536443 *C/C* homozygotes as assayed by pSTAT-positivity (Fig. 2F, fig. S4A and table S6). Consistently, IL-12- and IL-23-dependent differences in $CD4^+$ T helper (T_H) type 1 and T_H17 signature cytokine production were detected, respectively (fig. S5A). No effect was found for T_H2 signature cytokine production (fig. S5A), suggesting that a prior report of an apparent TYK2-related difference in T_H2 cells under non-polarizing conditions may be indirectly due to T_H1 and/or T_H17 perturbations (14). Reduced IL-12 and IL-23 signaling was also observed in $CD8^+$ memory T cells (Fig. 2F and table S6). Notably, of the specific subset of genes whose expression was altered in a genotype-specific fashion upon IFNβ stimulation based on RNA-Seq analysis, there was impairment of IL-12Rβ2 transcript upregulation in $CD8^+$ T cells from rs34536443 minor compared to major allele homozygotes (fig. S5B). This indicates the potential for indirect as well as direct effects of rs34536443 on IL-12 signaling in this cellular compartment. In contrast, for all cytokines and cell types assayed, there were no significant correlations between rs12720356 genotype and pSTAT (tables S7 and S8). The other two SNPs, rs2304257 and rs5030390, that together with rs12720356 constitute the haplotype that drives the tertiary association signal in the *TYK2* region are non-coding. Rs2304257 is proximal to the *TYK2* promoter region (fig. S6), but as the tertiary association does not correlate with differences in TYK2 transcript expression and has no impact on pSTAT as a downstream measure of TYK2 expression and activity, we found no evidence of a functional effect of rs2304257 on the *TYK2* gene (table S9). However, the tertiary signal does correlate with the expression of the *ICAM4*, *ZGLP1* and *ICAM1* genes located in the *TYK2* region (21, 22, 24, 26). As other SNPs that are not autoimmune disease-associated contribute to the variation of *ICAM4* and *ZGLP1* expression, the correlation between the tertiary signal and the expression of these two genes is in fact irrelevant to disease (table S9). This reveals the genotype-dependent *ICAM1* expression difference in $CD14^+$ monocytes --- which could influence diapedesis of these cells or immunological synapse formation (30) --- as the only identified phenotypic effect solely explained by the tertiary association. This suggests rs5030390 rather than rs12720356 as the likely causative SNP as it localizes to regulatory elements proximal to the *ICAM1* promoter (31, 32) (fig. S6). Therefore, of the three genetic signals we assessed in the *TYK2* region, rs34536443 alone has a demonstrable impact on TYK2 function, specifically leading to impaired type I IFN, IL-12 and IL-23 signaling.

A humanized murine model of TYK2 P1104A

One or more of these cytokine signaling pathways have been pathophysiologically implicated through other genetic, cellular and clinical data in each of the autoimmune conditions that rs34536443 protects against (19, 33, 34), although in multiple sclerosis IFNβ is administered therapeutically (35). To address this apparent inconsistency, we generated humanized mice carrying the targeted, orthologous missense substitution, which in mice is at amino acid position 1124 (P1124A), rather than 1104. Tyk2 Ala-1124 homozygous mice had comparable immune profiles and Tyk2 transcript and protein expression relative to Tyk2 Pro-1124 homozygotes, although the phosphorylation of Tyk2 was impaired upon stimulation with IFNβ (fig. S7, A-C). As observed with the primary human immune cells, Tyk2 Ala-1124 homozygotes also displayed >50% lower pSTAT-positivity upon stimulation with type I IFN, IL-12 and IL-23 signaling across cell types (Fig. 3A-C). These differences in pSTAT can have downstream phenotypic implications: Tyk2 Ala-1124 compared to Pro-1124 homozygotes showed a 1.7-fold reduction in the capacity of IFNβ to block the IFNγ-induced up-regulation of major histocompatibility complex class II on peritoneal macrophages (fig. S7D), whilst lower IL-12 and IL-23 signaling leads to reduced T_H1 and T_H17 signature cytokine production *in vitro* (fig. S7, E-G).

Upon immunization with myelin oligodendrocyte glycoprotein peptide (MOG) in complete Freund's adjuvant (CFA) to induce the multiple sclerosis-like disease experimental autoimmune encephalomyelitis (EAE), heterozygous mice showed decreased disease incidence compared to Pro-1124 homozygotes, whilst Ala-1124 homozygosity led to a complete protection against EAE (Fig. 3D and fig. S8, A and B). Cytokine knockout (36) and cytokine receptor knockout (37) mouse studies suggest that the observed protection against disease is likely mediated by the diminished IL-23 and, to a lesser extent IL-12 signaling, particularly through an impact on $CD4^+$ T cells which are the main pathogenic effectors in the disease model used. Consistent with this, following immunization, peripheral T_H17 and T_H1 cell impairment was observed based on reduced signature cytokine production by draining lymph node cells from Ala-1124 homozygous mice, although total $CD4^+$ T cell numbers were unaffected (Fig. 3,E-H and fig. S8, C-I). A genotype-dependent difference in cytokine production, assayed directly *ex vivo*, was also detected in negative control mice immunized with adjuvant alone (Fig. 3, F-H). This indicates that the phenotypic difference is not entirely antigen-dependent --- in keeping with the association of rs34536443 across multiple autoimmune diseases in humans --- but is instead exaggerated as the immune response to the immunizing antigen develops. In the central nervous system, comparing Pro-1124 homozygous mice at peak disease and Ala-1124 homozygotes at an equivalent time point post immunization, the latter had a negligible number of infiltrating $CD4^+$ T cells, and any such cells did not show a pathogenic profile (Fig. 3, I-N and fig. S9, A-F).

In EAE, abrogated IFNβ signaling does not alter disease incidence but increases severity (38, 39). In the Tyk2 P1124A heterozygous mice that developed disease, severity was not increased relative to Pro-1124 homozygotes (fig. S9, G and H), indicating that a stronger Tyk2 impairment is required to mediate a detrimental enough reduction in IFNβ signaling. However, such an effect could not be seen in Ala-1124 homozygotes given the lack of EAE

development, as the impact of defective IL-23 and IL-12 signaling precedes any beneficial action of IFNβ.

TYK2 P1104A homozygosity defines an optimum between autoimmunity and immunodeficiency

A putative drug-dependent amelioration of human autoimmune disease by targeted TYK2 modulation could be counterbalanced by safety concerns associated with the decreased type I IFN, IL-12 and IL-23 signaling. Extrapolating from patients with genetically driven deficiencies of TYK2 (29) and other components of these cytokine pathways (40, 41), long-term therapeutic inhibition of the kinase could augment the risk of certain severe and recurrent infections. However, the Oxford BioBank rs34536443 minor allele homozygous individuals analyzed were all self-reported healthy and showed no immune profile perturbations apart from the impaired cytokine signaling (fig. S10). Moreover, interrogation of medical records from 249 minor allele homozygotes out of a total of 116,732 genotyped individuals of European ancestry from the UK Biobank resource, revealed no rs34536443-dependent effects on the majority of health measures, including mortality and malignancy, other than the association with autoimmune disease (table S10). Evidence suggestive of a protective effect against self-reported hypothyroidism/myxoedema ($P = 3.62 \times 10^{-4}$), which is commonly of autoimmune etiology, was also found (table S10C). The minor allele homozygotes also showed no increase in their hospitalization due to mycobacterial, bacterial, viral or fungal infections, noting that we had >80% power of detecting associations with a risk ratio of 6.0 or above (Fig. 4A).. This indicates that rs34536443 minor allele homozygosity may allow for the minimal amount of TYK2-mediated cytokine signaling needed to prevent immunodeficiency.

Unlike the rs34536443 minor allele homozygotes, all TYK2 knockout individuals reported to date have an immunodeficient phenotype --- perhaps not only because there is a complete lack of kinase function in these individuals --- but because the absence of the TYK2 protein has a further effect on signaling as cell surface cytokine receptor stability is lowered (29). This was observed with our TYK2 knockout cell lines as well with respect to the cell surface expression of the IFNAR1 chain of the type I IFN receptor. Conversely, the rs34536443 minor allele homozygous cell lines displayed no decrease in IFNAR1 cell surface expression (Fig. 4B and fig. S11A), which may explain the slightly higher pSTAT1-positivity detected upon stimulation in these cells relative to the knockout cell lines (fig. S2D). The primary immune cells obtained from the rs34536443 minor compared to major allele homozygous individuals also showed no reduction in cell surface type I IFN, IL-12 or IL-23 receptors, consistent with no genotype-dependent effect on TYK2 protein expression (Fig. 4C-E and fig. S11, B-F).

Thus, if the minimal difference in TYK2-mediated cytokine signaling when comparing TYK2 knockout cells to cells with the rs34536443 minor allele homozygous genotype primarily relates to the role of TYK2 protein in stabilizing cytokine receptors, this implies that the minor allele homozygosity drives a near complete loss of the signal transduction function of TYK2. To investigate how this effect is conferred at the molecular level, we determined the X-ray crystal structure of the Ala-1057 JAK2 kinase domain to 2.66Å

Europe PMC Funders Author Manuscripts

resolution, noting that the TYK2 and JAK kinase domains are structurally highly homologous (42). Superposition of this structure with that of the Pro-1057 JAK2 kinase domain (43) demonstrated an overall very similar arrangement of the principal elements of the protein kinase fold, consistent with the more than 80 JAK kinase structures in the Protein Data Bank, but with one exception. Electron density maps indicated an unambiguous and large change starting with the αFG helix and extending to the succeeding αG helix (Fig. 5A and fig. S12). The αFG helix includes the site of mutation and is unique to JAK kinases. The shift is such that the distance between corresponding Cα atoms in these helices is on the order of 15Å. The generation of this crystal structure required two elements: firstly, inclusion of the ATP-competitive inhibitor TG101209, a molecule which was found, as expected, at the ATP-binding site far from αFG and αG, and secondly, crystallization from a specific medium unrelated to that previously used for the Pro-1057 JAK2/TG101209 structure (43). An additional molecule of TG101209 was unexpectedly found contacting αG and also a neighboring protein molecule provided by the crystal lattice, a finding without precedent among JAK kinase crystal structures. This unintended arrangement prevented straightforward assignment of the structural changes to the Pro1057Ala mutation. Comprehensive screening of crystallization conditions did not identify a pair of similar conditions from which both Pro-1057 and Ala-1057 JAK2 complexes with TG101209 produced crystals, even extending to nucleation with micro-crystals in cross-seeding experiments. This distinctly different crystallization behavior of the Pro-1057 and Ala-1075 JAK2 proteins is consistent with the lack of crystal formation in exhaustive experiments using the Ala-1104 TYK2 protein, despite prior success with Pro-1104 TYK2 (44). We concluded that the relaxed conformational restraint characteristic of a change from proline to alanine produces a substantial expansion of the low energy protein conformational space of the kinase domain, but that the details of our new structure should be assessed with caution. Nonetheless, it is probable that such an effect on accessible conformations is intrinsic to our observed homozygous Ala-1104 TYK2-dependent cytokine signaling impairment, which defines an optimal degree of TYK2 function that protects against autoimmune disease without causing immunodeficiency (Fig. 5B and table S11).

Discussion

By resolving conflicting genetic associations and functional correlations in the *TYK2* gene region, our study has singled out a highly protective effect against autoimmunity. This arises from minor allele homozygosity at the rs34536443 SNP, which drives a near complete loss of TYK2 function and consequently impairs type I IFN, IL-12 and IL-23 signaling. The comparison of genetic associations in the region across multiple autoimmune conditions provided the necessary foundation for hypothesis generation regarding the relative role of the associated variants, with implications for drug development. The cross-disease association pattern alone indicated that, contrary to prior suggestions (13, 15), rs34536443 and rs12720356 cannot have fully analogous biological effects despite both being missense variants in the *TYK2* gene. Epigenetic and transcriptomic meta-analysis and genotype-specific cytokine signaling assessments confirmed this hypothesis by revealing that compared to the other associations we investigated, rs34536443 is distinguishable as the

Europe PMC Funders Author Manuscripts

only SNP with a demonstrable impact on *TYK2*. Therefore, TYK2 may be a potential target for drug-dependent inhibition for multiple common autoimmune disorders.

Through the elucidation of the relative influence of naturally occurring *TYK2* genetic variation on cytokine signaling, we have delineated a genotype-determined gradient of TYK2 activity that may be predictive of the dose-response relationship between drug-dependent targeting of the enzyme and therapeutic outcome. Given the relatively modest protective effect of rs34536443 heterozygosity, a partial drug-induced impairment of TYK2, as might be achieved with broad-spectrum kinase inhibitors, would be therapeutically inadequate. This is especially the case for ulcerative colitis where the minor allele of the SNP confers protection only in its homozygous state. Conversely, based on the consequences of TYK2 deficiency (29), the benefit imparted by a long-term drug-dependent loss of TYK2 or blockade of its binding to cytokine receptors would be offset by certain severe and recurring infections. Existing drugs that block cytokine signaling upstream of TYK2, such as antibodies that target IL-12/IL-23 in psoriasis, are showing therapeutic promise, but they have been linked to serious infections in some instances (45).

The precise mimicking of the impact of rs34536443 minor allele homozygosity could provide maximal efficacy with minimized risk, especially for the treatment of early-stage disease. However, although the unprecedented combination of genetic and functional follow-up data may help to prospectively predict therapeutic outcome of autoimmune disease treatment, a limitation of our study is that definitive demonstration of such an optimal effect will ultimately require the generation or identification and subsequent trialing of an appropriately specific pharmaceutical. This may be aided by additional structural analyses to further interrogate the molecular mechanism by which the TYK2 P1104A substitution exerts its effect and how this exactly relates to the conformational change we observed when comparing the Ala-1075 and Pro-1075 JAK2 kinase domain structures. It should also be noted that the precise degree to which drug-dependent manipulation of TYK2 activity may have some influence on susceptibility to infection remains to be determined, and this may vary by ethnicity- or geography-dependent variation in infectious disease prevalence and for emerging pathogens, some of which can subvert TYK2-mediated cytokine signaling responses (46).

Initiatives for the generation and collation of large clinically relevant datasets are now under way to accelerate progress towards precision medicine (47), but the use of these data to help improve healthcare provision to patients with common, complex polygenic diseases is a challenge. Our study demonstrates how the integration of cross-disease genetic data with molecular, cellular, structural and epidemiological information can help to provide pre-clinical predictions regarding (i) target identification, (ii) indications, (iii) dose responses, and (iv) side effects, to thereby facilitate drug development and aid the definition of a therapeutic optimum for these diseases.

Materials and Methods

Study design

The study was approved by the NRES Committee South Central, Oxford B (REC Ref. No.: 10/H0605/5;v.2), and fresh peripheral blood samples were obtained from white European volunteers (without self-reported autoimmune or other chronic immunological/inflammatory disorders) from the Oxford BioBank, which is a random, population-based cohort of healthy adults (n > 5,000 at the time of analysis). With informed consent, individuals were recalled for blood donation in a balanced, age- and sex-matched fashion based on pre-selected genotypes of interest. All murine experiments were approved by the University of Oxford Clinical Medicine Ethical Review Committee and licensed under the Animals Scientific Procedures Act of the UK Home Office.

Statistical analysis of disease associations

SNP data in the *TYK2* locus (chr19:10,016,000–10,928,000; HG19 coordinates) from 8,726 ankylosing spondylitis (5), 4,017 Crohn's disease (6), 16,691 multiple sclerosis (8), 2,814 psoriasis (9), and 3,871 ulcerative colitis (6) patients and 19,738 population controls (6, 8) genotyped on Immunochip were used to localize associations. SNPs with minor allele frequency < 1%, genotype missing rate > 5% in any one batch or > 1% overall, and not in Hardy-Weinberg equilibrium in any cohort ($P < 10^{-5}$) were removed. Samples with excess genotype missing rate (>10% of all SNPs or 2% of clean SNPs), abnormal rate of heterozygosity, duplicated, first to third degree relatives of samples already present, and/or with non-European ancestry were excluded. 291 SNPs and 55,857 samples passed quality control. For fine-mapping we inferred haplotypes using SHAPEIT and imputed untyped SNPs with IMPUTE version 2 using as a reference all samples from phase 3 of the 1000 Genomes Project (from IMPUTE2 website, released December 2013); untyped SNPs with an imputation information score <0.7 were removed and all imputed SNPs in the 90% credible sets for each association signal had an imputation information score of 0.971 or above. Association analysis was performed with multinomial logistic regression as implemented in the trinculo tool (18). For imputed SNPs, analysis was performed with dosages to account for genotype imputation uncertainty. For each SNP tested an additive effect was assumed and population structure was accounted for by including the first seven principal components as covariates. Conditional analysis for identification of independent effects was performed stepwise, by including previously identified effects as covariates (rs34536443 for signal 2; rs34536443 and rs9797854 for signal 3), until no independent genome-wide significant association ($P < 5 \times 10^{-8}$) was observed. The 90% credible sets were derived from the likelihood function of the multinomial logistic model (16). For each of the three signals, likelihoods for each SNP under the alternative model were normalized and posterior probabilities derived from the assumption of one causal SNP per signal. Likelihoods for signal two were estimated conditioning on signal one and likelihoods for signal three were calculated conditioning on signal one and two.

Epigenetic co-localizations and ATAC-Seq

We investigated regulatory consequences of associated SNPs by aligning credible sets with publicly available and in-house epigenetic modification data. We obtained cell subset-

Europe PMC Funders Author Manuscripts

specific chromatin segmentation maps for the locus from 111 reference epigenomes from the Roadmap Epigenomics Consortium, and 18 epigenomes from the Blueprint Epigenome Project (http://www.blueprint-epigenome.eu). In-house chromatin accessibility maps were generated using ATAC-Seq as previously described (48), to analyze monocytes from three donors. Cells were isolated and stimulated for 0, 2 and 8 h with LPS (Invivogen) in accordance with Fairfax *et al.* (23). 100 bp pair-end sequencing reads were generated as previously described (48). Reads were aligned to the UCSC HG19 reference genome using Bowtie2 with the parameters -X2000, -I0, --no-discordant and --no-mixed. Additional processing removed duplicate and non-uniquely aligned reads. Peak-calling was performed using MACS.

eQTL co-localizations

Publicly available eQTL datasets analyzed included the primary data and summary statistics indicated in tables S4 and S9. For each independent association signal we obtained 95% confidence sets (16). For each analysis, unconditional or conditional, the Bayes factor for a variant is proportional (under weak assumptions) to the posterior probability that it is the causal variant. The 95% confidence set consists of the most significant variants such that 0.95 of the posterior probability is accounted for. For co-localization of susceptibility and eQTL signals we computed Bayes factors comparing models where both signals are caused by the same causal variants against models where each signal is caused by distinct variants.

CRISPR-Cas9 modified cell lines

HEK293T cells (ATCC) were modified using the CRISPR-Cas9 system as previously described (49). For rs34536443 minor allele introduction the 5’-CACCGTGGGGGGGCTCTGGCTGG-3’, 5’-CTCTCACCGTGGGGGGGCTCTGG-3’, and 5’-AGCCAGGCCCGCAGCCCCACCGG-3’ guides and the 5’-ccttcggggtgaccctgtatgagctgctgacgcactgtgactccagccagagcgcccccacgGTGAGA GCCAGGCCCGCAGCCCCACCGGGAGTTTGCTAGAGCAATTAGAAAGGCATAGGCT G GGCCAGGC-3’ sense and 5’-GCCTGGCCCAGCCTATGCCTTTCTAATTGCTCTAGCAAA CTCCCGGTGGGGCTGCGGGCCTGGCTCTCACcgtgggggcgctctggctggagtcacagtgcgtcagc agc tcatacagggtcaccccgaagg-3’ anti-sense HDR templates were used, whilst the 5’-ATGCGTCAG ATGTCTGGTCC-3’ and 5’-GCCTCACAGCAGATTGTTCT-3’ primer pair were used for validation by Sanger sequencing. For rs12720356 the 5’-GCTGTGTCTTTACAGATATCA-3’ guide was used with the 5’-CAATTCCCCGCCATCCAGAGGGCAGAAGCAGGCAGGTTGC CCCAGAGCAGCTGTGTCTTTACAGatatcatggtgacagagtacgtggagcacggacccctggatgtgtggct gcg gagggagcggggccatgtgcccat-3’ sense and 5’-atgggcacatggccccgctccctccgcagccacacatccaggggt ccgtgctccacgtactctgtcaccatgatatCTGTAAAGACACAGCTGCTCTGGGGCAACCTGCCTG CT TCTGCCCTCTGGATGGCGGGGAATTG-3’ anti-sense templates for the major allele and the 5’-CAATTCCCCGCCATCCAGAGGGCAGAAGCAGGCAGGTTGCCCCAGAGCAGCTGT GTCTTTACAGatagcatggtgacagagtacgtggagcacggacccctggatgtgtggctgcggagggagcggggccatg tgc ccat-3’sense and 5’-

atgggcacatggccccgctccctccgcagccacacatccaggggtccgtgctccacgtactctgtcacca tgctatCTGTAAAGACACAGCTGCTCTGGGGCAACCTGCCTGCTTCTGCCCTCTGGAT GGCGGGGAATTG-3’ anti-sense HDR templates for the minor allele. The 5’-CTCTGCTGCATT CCGCATTT-3’ and 5’-CACAGGCCACACACCAGGTA-3’ primers were used for validation by Sanger sequencing. For knockout cell lines the genomic DNA in the region corresponding to the TYK2 N-terminal FERM domain was targeted as this is the domain with most naturally-occurring TYK2 knockout mutations (29). For these lines the 5’-ACTGTGGGAGCTGTCGAC CG-3’, 5’-GTGGGAGCTGTCGACCGAGG-3’, 5’-GCATGAGTTTGTGAATGACG-3’, and 5’-GAATGACGTGGCATCACTGT-3’ guides were used. PCR and Sanger sequencing confirmation was performed using the 5’-AGATCCCCAGAGATGCAAGC-3’ and 5’-GGAA CCTGCGGAAGACGTTC-3’ primers. Prior to stimulation of $1x10^6$ cells/ml with 1000 ng/ml IFNβ (Peprotech) for 15 min at 37°C, cells were serum-starved for 6 h.

Immunoblotting

Cells were lysed in 50 mM Tris/HCl pH 8.0, 10% (v/v) glycerol, 25 mM EDTA, 150 mM NaCl, 2 mM DTT, 0.5% NP-40 (Igepal CA-630), 25 mM sodium fluoride, 1 mM sodium orthovanadate supplemented with protease inhibitors (Roche) and 0.1% phosphatase inhibitor cocktail 2 (Sigma-Aldrich). Anti-actin (clone C4, MilliPore), and anti-N-terminal-TYK2 (clone 51, BD Biosciences) with anti-phospho-tyrosine (Y1054/Y1055) TYK2 (Cell Signaling Technology), or anti-N-terminal-TYK2 (H-135, Santa Cruz Biotechnology) with anti-phospho-tyrosine (Y1054/Y1055) TYK2 (Santa Cruz Biotechnology) were used for human and murine protein detection, respectively. Membranes were stained with IR-Dye® 680LT- and 800CW-conjugated secondary antibodies and visualized and quantified using an Odyssey Infra-Red Imaging System (LI-COR Biosciences).

Genotyping

Rs34536443, rs9797854, and rs12720356 were genotyped using TaqMan® SNP Genotyping Assays (Applied Biosystems), and mouse genotyping was performed using a custom-made TaqMan® SNP Genotyping Assay (Applied Biosystems) that enabled discrimination of alleles encoding Tyk2-Pro-1124 and Ala-1124.

Flow and mass cytometry

For cell-surface staining, cells were stained for 30 min and fixed with 1X BD CellFIX™ (BD Biosciences). Phosflow staining was performed according to the BD Biosciences protocols using BD Cytofix™ Fixation Buffer and BD Phosflow™ Perm Buffer III. Intracellular cytokine staining was performed using the BD Cytofix/Cytoperm™ Fixation/Permeablization kit.

Anti-human antibodies used were: anti-IFNAR1 (AbGent) along with Alexa-Fluor488 goat anti-rabbit IgG (Invitrogen); allophycocyanin (APC) anti-CD25 (clones M-A251 and 2A3), fluorescein isothiocyanate (FITC) anti-Lin1 (BD Biosciences); Alexa-Fluor700/APC/phycoerythrin (PE) anti-CD4 (clone RPA-T4), PE-Cy7 anti-CD11b (clone ICRF44), peridinin–chlorophyll–protein complex (PerCP)-Cy5.5 anti-CD14 (clone HCD14), Pacific Blue anti-CD16 (clone 3G8), Alexa-Fluor700 anti-CD27 (clone M-T271), APC anti-CD28

(clone CD28.2), Pacific Blue anti-CD45RA (clone HI100), APC anti-CD56 clone (HCD56), PerCP-Cy5.5 anti-CD57 (clone HNK-1), APC anti-CD123 (clone 6H6), APC-Cy7 anti-CD193 (5E8), APC anti-IFNγ (clone B27), APC anti-IgD (clone IA6-2), FITC/PE anti-IL-4 (clone MP4-25D2) (BioLegend); Alexa-Fluor488 anti-CD3 (clone OKT3), PE-Cy7 anti-CD4 (clone RPA-T4), Alexa-Fluor700/FITC/PE anti-CD8 (clone OKT8), Alexa-Fluor700 anti-CD11c (clone 3.9), FITC anti-CD15 (clone HI98), Alexa-Fluor488/Alexa-Fluor700/APC anti-CD19 (clone HIB19), eFluor450 anti-CD38 (clone HB7), PE-Cy7 anti-CD127 (clone eBioRDR5), PE anti-CD197 (clone 3D12), FITC/PE anti-IL-17 (clone N49-653) (eBioscience); FITC anti-IL-12Rβ1 (clone 69310), APC anti-IL-12Rβ2 (305719), PE anti-IL-23R (clone 218213) (R&D Systems).

Anti-mouse antibodies used were: PE anti-I-A/I-E (clone M5/114.15.2), PE anti-IL-4 (clone 11B11) (BD Biosciences); APC-eFluor780/FITC anti-CD11b (clone M1/70), APC anti-CD11c (clone N418), Alexa-Fluor700/eFluor450 anti-CD19 (clone eBio1D3), Alexa-Fluor700 anti-CD3e (clone 145-2c11), Alexa-Fluor700/PE anti-CD4 (clone RM4-5), APC/eFluor450/PE-Cy7 anti-CD8a (clone 53-6.7) APC anti-CD25 (PC61.5), eFluor450 anti-CD44 (IM7), FITC anti-CD49d (clone R1-2), FITC/PerCP-Cy5.5 anti-CD62L (clone MEL-14), PE anti-CD122 (clone TM-b1), eFluor660 anti-Eomes (clone Dan11mag), APC anti-IFNγ (clone XMG1.2), FITC anti-IL-17A (clone eBio17b7), eFluor450 anti-Ly-6C (clone HK1.4), PE-Cy7 anti-Ly6G (clone RB6-8C5), APC/PerCP-Cy5.5 anti-NK1.1 (clone PK136) (eBioscience). For phosflow staining, the following antibodies were used: Alexa-Fluor647 anti-pSTAT1 (clone 4a), Alexa-Fluor488 anti-pSTAT3 (clone 4/P-STAT3), Alexa-Fluor647 anti-pSTAT4 (clone 38/p-Stat4), Alexa-Fluor647 anti-pSTAT6 (clone 18/P-Stat6) (BD Biosciences).

Isotype control antibodies used were: APC/FITC/PE mouse IgG1, PE mouse IgG2b (R&D Systems), and PE rat IgG2a (BD Biosciences), rabbit polyclonal IgG (Abcam). Mass cytometry was performed using PBMCs and the Maxpar® Human Peripheral Blood Phenotyping Panel Kit (Fluidigm). Data obtained were analyzed using FlowJo (Tree Star). For FACS sorting, murine CNS mononuclear cells were first stained for 40 min at room temperature with 4 μg/ml of APC-labeled 1-A^b/myelin oligodendrocyte glycoprotein $(MOG)_{38-49}$ tetramer or negative control 1-A^b/hCLIP tetramer (NIH Tetramer Core Facility, Emory University). For a final 20 min cells were stained at 4°C with the tetramers, Fc block (BD Biosciences) and PE anti-CD4 (clone RM4-5), Alexa-Fluor488 anti-CD11b (clone M1/70), and eFluor450 anti-CD45.2 (clone 104) (eBioscience).

Human blood cell subset isolation and culturing

Peripheral blood mononuclear cells (PBMCs) were separated on Lymphoprep (Axis-Shield). Specific subsets ($CD3^+$ T cells, $CD4^+$ T cells, $CD8^+$ T cells, $CD19^+$ B cells and $CD14^+$ monocytes) were isolated using respective positive-selection MicroBeads (Miltenyi), and for RNA-Seq cells were serum-starved for 4 h and stimulated for 12 h with 400 ng/ml IFNβ. For cytokine stimulations, 400 ng/ml IFNα, IFNβ, IL-6, IL-10 or IL-13 (Peprotech) were used to stimulate PBMCs (1×10^6 cells/ml) for 15 min in RPMI at 37°C. For IL-12 stimulation, purified $CD3^+$ T cells were serum-starved for 4 h, activated for 72 h with 10 μg/ml phytohaemagglutinin (Sigma-Aldrich) in RPMI supplemented with 10% human AB serum,

2mM glutamine and penicillin/streptomycin (all from Sigma-Aldrich), then stimulated with 100 U/ml IL-2 (Peprotech) for 24 h, and serum-starved for 4 h prior to stimulation with 400 ng/ml IL-12 (Peprotech) for 15 min. For IL-23 stimulation, PBMCs were activated with Dynabeads® Human T-Activator CD3/CD28 (Invitrogen) according to the instructions but with a 1:2 bead-to-cell ratio for 5 days, serum-starved for 4 h and then stimulated with 400 ng/ml of IL-23 (R&D Systems) for 15 min.

RNA-Seq

Human polyA sequencing libraries were prepared according to the Illumina TruSeq protocol. Libraries were sequenced using Illumina HiSeq 4000 (75 bp paired-end reads). Reads were mapped to UCSC HG19 human using TopHat. Uniquely aligned reads were extracted and read duplicates removed using Picard (http://picard.sourceforge.net/). HT-Seq was used to quantify gene-expression using the gene annotations provided by Illumina's iGenomes Project. Differential expression was assessed with the R Bioconductor package DESeq2. Gene expression counts were normalized by library size and all other default parameters were used. Analysis was performed for each cell-type and stimulation condition. Evidence of association was quantified with the false-discovery rate routine in DESeq2 and Q-values < 0.05 were considered significant.

$CD4^+$ T-cell differentiation

$CD4^+$ naïve cells were isolated from human PBMCs and murine splenocytes using species-specific Naïve $CD4^+$ T Cell Isolation Kits (Miltenyi) and cultured for 5 days under polarizing conditions. Human cells were cultured with Dynabeads® Human T-Activator CD3/CD28 (Invitrogen) with 10 U/ml IL-2 (Peprotech) supplemented with the following for T_H1, T_H2 and T_H17 polarizations, respectively: 1 ng/ml IL-12 (Peprotech), 10 mg/ml anti-IL-4 (clone MP4-25D2, BioLegend); 4 μg/ml IL-4 (Peprotech), 1 μg/ml anti-IFNγ (clone B27, BioLegend), anti-IL-12p70 (clone 20C2, eBioscience); 10 ng/ml IL-1β (R&D Systems), 10 ng/ml IL-6 (Peprotech), 10 ng/ml IL-23 (R&D Systems), 1 ng/ml TGF-β1 (R&D Systems), 1 μg/ml anti-IL-4, 1 μg/ml anti-IFNγ. Murine cells were cultured with 5 μg/ml anti-CD3 (clone 145-2C11) and 1 μg/ml soluble anti-CD28 (37.51) (eBioscience) supplemented with the following for T_H1, T_H2 and T_H17 polarizations, respectively: 1 ng/ml IL-12 (Peprotech), 10 mg/ml anti-IL-4 (clone 11B11, eBioscience); 4 μg/ml IL-4 (Peprotech), 10 μg/ml anti-IFNγ (clone XMG1.2, eBioscience); 20 ng/ml IL-6 (Peprotech), 10 ng/ml IL-23 (R&D Systems), 1 ng/ml TGF-β1 (R&D Systems), 10 mg/ml anti-IL-4. For intracellular cytokine detection cells were plated in fresh media, stimulated with 10 ng/ml phorbol 12-myristate 13-acetate (PMA) and 1 μg/ml ionomycin, and treated with 10 μg/ml Brefeldin A (Sigma-Aldrich) and BD GolgiStop™ (BD Biosciences) for 5 h prior to staining.

Enzyme-linked immunosorbent assays

Cell culture supernatants were assayed for human or murine IFNγ, IL-4, IL-17A/F and murine GM-CSF using LEGEND MAX™ ELISA kits (BioLegend). Human plasma IgE was captured using anti-human IgE (clone G7-18, BD Biosciences). Human IgE full-length protein (Abcam) was the standard and biotinylated anti-human IgE (clone G7-26, BD Biosciences) and Europium-conjugated streptavidin (Perkin Elmer) were used for detection.

Europe PMC Funders Author Manuscripts

Murine plasma IgE was captured using anti-mouse IgE (clone R35-72), with purified mouse IgE as the standard and biotinylated anti-mouse IgE (clone R35-118) for detection (BD Biosciences).

Animals

Tyk2-Ala-1124-encoding targeting construct was generated by the Gene Recombineering Facility (Monash University). Gene targeting was performed in the embryonic stem (ES) cell line JM8F6. Successful homologous recombination between vector and endogenous *Tyk2* locus was detected by long-range PCR using the 5HR-F2 (5'-GGGGCTCC TGTCTACAGCTC-3') and 5HR-R2 (5'-TGTCGATCAGGATGATCTGG-3'), and the 3HR-F1 (5'-CGCGGGGATCTCA TGCTGGAGTTC-3') and 3HR-R1 (5'-CCCATGTGTGTCCCTTGA TCCTGC-3') primers and the LongAmp™ *Taq* PCR kit (New England Biolabs). Positive clones were further validated by Southern blotting using standard methods with probe amplification using the 5'-GGATCCGA ACATGGTCCCTTGGATGT-3' and 5'-GCTAGCGGCTTCACCTGTATCCCACT-3' primers. Correctly targeted ES cells were injected into albino C57BL6/J blastocysts. Confirmation of presence of transgenic *Tyk2* at the transcript level was performed by PCR and Sanger sequencing using the 5'-CTGCTGGACAACGACAGG-3' and 5'-ACACGC TGAACACGGAA-3' primers.

Real-time quantitative PCR

Murine and human glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) house-keeping gene and murine Tyk2 and human IFNAR1 transcripts were detected using TaqMan® Gene Expression Assays Mm99999915_g1, Hs02758991_g1, Mm00444469_m1 and Hs01066118_m1, respectively (Applied Biosystems). Relative transcript levels are expressed as 2^{-dCt}, where dCt = (cycle threshold for the transcript of interest) – (cycle threshold for the house-keeping gene transcript).

EAE

We immunized 8–12-week-old female mice subcutaneously in the flanks with 200 μg $MOG_{35–55}$ (Schafer) emulsified in incomplete Freund's adjuvant (Sigma-Aldrich) supplemented with 4 mg/ml *Mycobacterium tuberculosis* H37Ra (Difco; CFA), or with the pre-prepared Hooke™ MOG_{35-55}/CFA immunisation kit for all cellular analyses (Hooke Laboratories). We administered 200 ng *Bordetella pertussis* toxin (Sigma-Aldrich or Hooke Laboratories) intravenously on the day of immunisation and 48 h later. The following EAE scoring system was used: score 0 = no overt disease; score 1 = weak tail; score 2 = hind limb weakness, abnormal gait and/or impaired righting reflex; score 3 = partial hind limb paralysis; score 4 = complete hind limb paralysis.

Murine cell isolation and stimulation

Murine peritoneal cells were plated in DMEM/F-12 with GlutaMAX™ (Gibco™) and supplemented with 10% FCS; adherent cells were cultured for 24 h prior to stimulation with 10 ng/ml IFNβ (Pestka Biomedical Laboratories), 10 ng/ml IFNγ (Peprotech) or both for 18 h. CNS tissue obtained from perfused mice was digested in 1X HBSS (Sigma-Aldrich)

supplemented with 0.05% collagenase D (Roche), 0.1 μg/ml TLCK (Sigma-Aldrich) and 0.025 U/ml DNase I (ThermoFisher Scientific). Mononuclear cells were then isolated by discontinuous Percoll (GE Healthcare Life Sciences) gradient centrifugation. Spleens and inguinal lymph nodes were obtained by standard methods. For culturing cells were plated in RPMI supplemented with 10% FCS, 2 mM glutamine, penicillin/ streptomycin and with 1 mM sodium pyruvate, 0.1 mM non-essential amino acids and 0.5 μM beta-mercaptoethanol (GIBCO). For phosflow splenocytes were stimulated as for human PBMCs but with murine IFNβ (Pestka Biomedical Laboratories), IL-6, IL-10, IL-12 (Peprotech) or IL-23 (R&D Systems). For recall assays cells were stimulated for 72 h with 10 μg/ml MOG_{35-55} and restimulated with PMA and ionomycin as for $CD4^+$ T-cell differentiation.

Statistical analysis of HEK, human immune and murine data

Statistical tests were performed using GraphPad Prism and the R statistical software package. For analyses comparing two unpaired experimental groups Mann-Whitney *U* tests were used, whilst for paired analyses a Wilcoxon matched-pairs rank test was used, with a 5% significance level, and estimating one- or two-tailed *P*-values depending on prior assumptions regarding the directionality of the effect tested. For unpaired primary human pSTAT analyses a Bonferroni correction for multiple testing was used (taking $P = 0.05$ and considering eight independent hypotheses for four different cytokine types and two SNPs), the significance threshold estimated was $P = 0.00625$, noting that different cell subsets were not considered as independent tests as all are TYK2-positive and the SNPs are missense such that different subsets were analyzed to account for differences in cytokine receptor levels that might confound analyses of bulk populations. At this significance threshold, we obtained >80% power to detect the differences observed. For these analyses by rs34536443 and rs12720356 genotype, *P*-values provided are for non-additive and additive genetic models, respectively, noting that age and sex were not included as covariates as no age or sex effects were observed (all $P > 0.05$).

Structural analysis

Site-directed mutagenesis was applied to a previously reported JAK2 kinase domain construct (50) to produce a version bearing P1057A. DNA encoding the Ala-1057 JAK2 kinase domain was manufactured by GenScript. Protein was expressed and purified as previously described (50). Purified Ala-1057 JAK2/TG101209 complex (6 mg/mL in 250 mM NaCl, 25 mM bicine pH 8.5, 0.5 mM TCEP, and 1 mM inhibitor (from 50 mM DMSO)) was subjected to crystallization trials using screening solutions from Hampton Research and Qiagen. Optimized conditions used a reservoir with 1.9-2.0 M sodium malonate pH 7.0, 0.1 M Bis-tris propane pH 7.0, 10 mM $ZnCl_2$ in hanging drops made from 2 μL protein solution plus 2 μL reservoir at 18°C. Crystals were harvested and placed briefly in 100% reservoir as cryoprotectant, and then plunged into liquid nitrogen. Diffraction data were collected at 110K using APS beamline 21-IDG (wavelength = 0.97856 Å). Data were integrated and scaled using HKL2000 and elements of the CCP4 suite. Initial phases were obtained by molecular replacement using Phaser employing PDB 4JI9 (JAK2) as search probe. Refinement was performed using Refmac5, phenix.refine and BUSTER-TNT (final model 94% favored Ramachandran angles), and model manipulation and electron density inspection employed Coot.

Europe PMC Funders Author Manuscripts

UK Biobank data analysis

UK Biobank genotype data included 152,729 subjects. We excluded individuals of non-European ancestry (based on principal components analysis) and samples with abnormal heterozygosity rates and high missing genotype rates, leaving 116,732 individuals (male:female ratio = 61,339:55,214; mean age (±SEM) = 63.82 (±7.94) years). Data on hospital episode statistics (HES) contained in-patient data from 01 April 1997 onwards and included ICD-10-coded data from admissions, operations and procedures. We found no association between genotype and number or nature of hospitalisation events, aside from associations with autoimmune disease. In TYK2-deficient individuals the effect on immunodeficiency is 100% penetrant (29). In the UK Biobank, 1.2% of the cohort had at least one hospital episode recorded with relevant infections. With this prevalence and with 249 rs34536443 *C/C* individuals, we had > 80% power to detect an association with relative risk > 6.0 (using a recessive genetic model with a minor allele frequency of ~4%) and a case control ration of 10 (at alpha = 0.05). From the cancer registry, genetic data were available for 13,315 subjects and the most common cancer type (melanoma and other malignant neoplasms of the skin) has a prevalence of 3.4%. With this prevalence, we had > 80% power to detect an association (using a recessive genetic model as above) with a risk ratio of 2.621. Survival analysis (using the Cox proportional hazard regression model) with rs34536443 genotypes, and correcting for the effect of gender in life expectancy, resulted in no significant differences. In the UK Biobank, there were 2,146 death events registered, and we estimated power to detect a genetic association (recessive model) with the survSNP R package, assuming parameters obtained from the UK Biobank data. The study was powered to detect an association with reduced survival with a genotype hazard-ratio of 2.6 (considerably above the effect estimated for gender, 1.5). All data used were last updated June 2015.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the GAPC, IGAS, IMSGC, UKIBDGC, and WTCCC2 for data access. We thank the volunteers from the Oxford BioBank, NIHR Oxford Biomedical Research Centre, for their participation. The Oxford BioBank (www.oxfordbiobank.org.uk) is also part of the NIHR National Bioresource, which supported the recalling process of the volunteers. We also thank A. Smith and T. Sauka-Spengler [Weatherall Institute of Molecular Medicine (WIMM), University of Oxford] for help establishing the Southern blotting and ATAC-Seq methods, respectively; C. Monaco, P. Amjadi, and D. Ahern (Kennedy Institute of Rheumatology, University of Oxford) for mass cytometer access and technical support; and B. Davies, D. Biggs (Transgenic core, Wellcome Trust Centre for Human Genetics, University of Oxford) for transgenic mouse generation services. This study makes use of data generated by the Blueprint Consortium; a full list of the investigators who contributed to the generation of the data is available from www.blueprint-epigenome.eu and funding for the project was provided by the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement number 282510 BLUEPRINT. This research has also been conducted using the UK Biobank Resource.

We would also like to acknowledge the flow cytometry facility at the WIMM, which is supported by the MRC HIU; MRC MHU (MC_UU_12009); NIHR Oxford BRC and John Fell Fund (131/030 and 101/517), the EPA fund (CF182 and CF170) and by the WIMM Strategic Alliance awards G0902418 and MC_UU_12025. We thank the High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics (funded by Wellcome Trust grant reference 090532/Z/09/Z) for the generation of sequencing data. We thank Shamrock Structures, LLC for diffraction data collection. Use of the APS, an Office of Science User Facility operated for the U.S. Department of Energy (DOE), Office of Science by Argonne National Laboratory, was supported by the U.S. DOE under Contract

No. DE-AC02-06CH11357. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor (Grant 085P1000817).

Funding. Work in the authors' laboratories is supported by the UK and Danish Medical Research Councils, The Alan and Babette Sainsbury Charitable Fund, the Naomi Bramson Trust, the Clinical Neuroimmunology Fund, the Oxford Biomedical Research Centre, the Oak Foundation (L.F.), the Rosetrees Trust (L.F. and C.A.D.) and the Wellcome Trust (100308/Z/12/Z and 106130/Z/14/Z, L.F.; 100956/Z/13/Z, G.M.). L.S. was supported by a Christopher Welch Scholarship.

References and Notes

1. Manolio TA. Bringing genome-wide association findings into clinical use. Nat Rev Genet. 2013; 14:549–558. [PubMed: 23835440]
2. Hay M, Thomas DW, Craighead JL, Economides C, Rosenthal J. Clinical development success rates for investigational drugs. Nat Biotechnol. 2014; 32:40–51. [PubMed: 24406927]
3. Nelson MR, Tipney H, Painter JL, Shen J, Nicoletti P, Shen Y, Floratos A, Sham PC, Li MJ, Wang J, Cardon LR, et al. The support of human genetic evidence for approved drug indications. Nat Genet. 2015; 47:856–860. [PubMed: 26121088]
4. Plenge RM, Scolnick EM, Altshuler D. Validating therapeutic targets through human genetics. Nat Rev Drug Discov. 2013; 12:581–594. [PubMed: 23868113]
5. International Genetics of Ankylosing Spondylitis Consortium (IGAS). Cortes A, Hadler J, Pointon JP, Robinson PC, Karaderi T, Leo P, Cremin K, Pryce K, Harris J, Lee S, et al. Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. Nat Genet. 2013; 45:730–738. [PubMed: 23749187]
6. Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, Lee JC, Schumm LP, Sharma Y, Anderson CA, Essers J, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature. 2012; 491:119–124. [PubMed: 23128233]
7. Franke A, McGovern DP, Barrett JC, Wang K, Radford-Smith GL, Ahmad T, Lees CW, Balschun T, Lee J, Roberts R, Anderson CA, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet. 2010; 42:1118–1125. [PubMed: 21102463]
8. International Multiple Sclerosis Genetics Consortium (IMSGC). Beecham AH, Patsopoulos NA, Xifara DK, Davis MF, Kemppinen A, Cotsapas C, Shah TS, Spencer C, Booth D, Goris A, et al. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nat Genet. 2013; 45:1353–1360. [PubMed: 24076602]
9. Tsoi LC, Spain SL, Knight J, Ellinghaus E, Stuart PE, Capon F, Ding J, Li Y, Tejasvi T, Gudjonsson JE, Kang HM, et al. Identification of 15 new psoriasis susceptibility loci highlights the role of innate immunity. Nat Genet. 2012; 44:1341–1348. [PubMed: 23143594]
10. Diogo D, Bastarache L, Liao KP, Graham RR, Fulton RS, Greenberg JD, Eyre S, Bowes J, Cui J, Lee A, Pappas DA, et al. TYK2 protein-coding variants protect against rheumatoid arthritis and autoimmunity, with no evidence of pleiotropic effects on non-autoimmune complex traits. PLoS One. 2015; 10 e0122271.
11. Onengut-Gumuscu S, Chen WM, Burren O, Cooper NJ, Quinlan AR, Mychaleckyj JC, Farber E, Bonnie JK, Szpak M, Schofield E, Achuthan P, et al. Fine mapping of type 1 diabetes susceptibility loci and evidence for colocalization of causal variants with lymphoid gene enhancers. Nat Genet. 2015; 47:381–386. [PubMed: 25751624]
12. Schwartz DM, Bonelli M, Gadina M, O'Shea JJ. Type I/II cytokines, JAKs, and new strategies for treating autoimmune diseases. Nat Rev Rheumatol. 2016; 12:25–36. [PubMed: 26633291]
13. Tomasson MH, Xiang Z, Walgren R, Zhao Y, Kasai Y, Miner T, Ries RE, Lubman O, Fremont DH, McLellan MD, Payton JE, et al. Somatic mutations and germline sequence variants in the expressed tyrosine kinase genes of patients with *de novo* acute myeloid leukemia. Blood. 2008; 111:4797–4808. [PubMed: 18270328]
14. Couturier N, Bucciarelli F, Nurtdinov RN, Debouverie M, Lebrun-Frenay C, Defer G, Moreau T, Confavreux C, Vukusic S, Cournu-Rebeix I, Goertsches RH, et al. Tyrosine kinase 2 variant influences T lymphocyte polarization and multiple sclerosis susceptibility. Brain. 2011; 134:693–703. [PubMed: 21354972]

15. Li Z, Gakovic M, Ragimbeau J, Eloranta ML, Rönnblom L, Michel F, Pellegrini S. Two rare disease-associated Tyk2 variants are catalytically impaired but signaling competent. J Immunol. 2013; 190:2335–2344. [PubMed: 23359498]
16. Wellcome Trust Case Control Consortium. Maller JB, McVean G, Byrnes J, Vukcevic D, Palin K, Su Z, Howson JM, Auton A, Myers S, Morris A, et al. Bayesian refinement of association signals for 14 loci in 3 common diseases. Nat Genet. 2012; 44:1294–1301. [PubMed: 23104008]
17. 1000 Genomes Project Consortium. Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR. A global reference for human genetic variation. Nature. 2015; 526:68–74. [PubMed: 26432245]
18. Jostins L, McVean G. Trinculo: Bayesian and frequentist multinomial logistic regression for genome-wide association studies of multi-category phenotypes. Bioinformatics. 2016 pmid: 26873930.
19. Cho JH, Feldman M. Heterogeneity of autoimmune diseases: pathophysiological insights from genetics and implications for new therapies. Nat Med. 2015; 21:730–738. [PubMed: 26121193]
20. Ban M, Goris A, Lorentzen AR, Baker A, Mihalova T, Ingram G, Booth DR, Heard RN, Stewart GJ, Bogaert E, Dubois B, et al. Replication analysis identifies TYK2 as a multiple sclerosis susceptibility factor. Eur J Hum Genet. 2009; 17:1309–1313. [PubMed: 19293837]
21. Raj T, Rothamel K, Mostafavi S, Ye C, Lee MN, Replogle JM, Feng T, Lee M, Asinovski N, Frohlich I, Imboywa S, et al. Polarization of the effects of autoimmune and neurodegenerative risk alleles in leukocytes. Science. 2014; 344:519–523. [PubMed: 24786080]
22. Fairfax BP, Humburg P, Makino S, Naranbhai V, Wong D, Lau E, Jostins L, Plant K, Andrews R, McGee C, Knight JC. Innate immunity activity conditions the effect of regulatory variants upon monocyte gene expression. Science. 2014; 343:1246949. [PubMed: 24604202]
23. Lee MN, Ye C, Villani AC, Raj T, Li W, Eisenhaure TM, Imboywa SH, Chipendo PI, Ran FA, Slowikowski K, Ward LD, et al. Common genetic variants modulate pathogen-sensing responses in human dendritic cells. Science. 2014; 343:1246980. [PubMed: 24604203]
24. Fairfax BP, Makino S, Radhakrishnan J, Plant K, Leslie S, Dilthey A, Ellis P, Langford C, Vannberg FO, Knight JC. Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. Nat Genet. 2012; 44:502–510. [PubMed: 22446964]
25. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science. 2015; 348:648–660. [PubMed: 25954001]
26. Battle A, Mostafavi S, Zhu X, Potash JB, Weissman MM, McCormick C, Haudenschild CD, Beckman KB, Shi J, Mei R, Urban AE, et al. Characterizing the genetic basis of transcriptome diversity through RNA-sequencing of 922 individuals. Genome Res. 2014; 24:14–24. [PubMed: 24092820]
27. Nair TS, Kozma KE, Hoefling NL, Kommareddi PK, Ueda Y, Gong TW, Lomax MI, Lansford CD, Telian SA, Satar B, Arts HA, et al. Identification and characterization of choline transporter-like protein 2, an inner ear glycoprotein of 68 and 72 kDa that is the target of antibody-induced hearing loss. J Neurosci. 2004; 24:1772–1779. [PubMed: 14973250]
28. Lupardus PJ, Ultsch M, Wallweber H, Bir Kohli P, Johnson AR, Eigenbrot C. Structure of the pseudokinase-kinase domains from protein kinase TYK2 reveals a mechanism for Janus kinase (JAK) autoinhibition. Proc Natl Acad Sci U S A. 2014; 111:8025–8030. [PubMed: 24843152]
29. Kreins AY, Ciancanelli MJ, Okada S, Kong XF, Ramírez-Alejo N, Kilic SS, El Baghdadi J, Nonoyama S, Mahdaviani SA, Ailal F, Bousfiha A, et al. Human TYK2 deficiency: Mycobacterial and viral infections without hyper-IgE syndrome. J Exp Med. 2015; 212:1641–1662. [PubMed: 26304966]
30. Hogg N, Patzak I, Willenbrock F. The insider's guide to leukocyte integrin signalling and function. Nat Rev Immunol. 2011; 11:416–426. [PubMed: 21597477]
31. Roadmap Epigenomics Consortium. Integrative analysis of 111 reference human genomes. Nature. 2015; 518:317–330. [PubMed: 25693563]
32. Adams D, Altucci L, Antonarakis SE, Ballesteros J, Beck S, Bird A, Bock C, Boehm B, Campo E, Caricasole A, Dahl F, et al. BLUEPRINT to decode the epigenetic signature written in blood. Nat Biotechnol. 2012; 30:224–226. [PubMed: 22398613]

33. Teng MW, Bowman EP, McElwee JJ, Smyth MJ, Casanova JL, Cooper AM, Cua DJ. IL-12 and IL-23 cytokines: from discovery to targeted therapies for immune-mediated inflammatory diseases. Nat Med. 2015; 21:719–729. [PubMed: 26121196]
34. Ivashkiv LB, Donlin LT. Regulation of type I interferon responses. Nat Rev Immunol. 2014; 14:36–49. [PubMed: 24362405]
35. Calabresi PA, Kieseier BC, Arnold DL, Balcer LJ, Boyko A, Pelletier J, Liu S, Zhu Y, Seddighzadeh A, Hung S, Deykin A. ADVANCE Study Investigators, Pegylated interferon beta-1a for relapsing-remitting multiple sclerosis (ADVANCE): a randomised, phase 3, double-blind study. Lancet Neurol. 2014; 13:657–665. [PubMed: 24794721]
36. Cua DJ, Sherlock J, Chen Y, Murphy CA, Joyce B, Seymour B, Lucian L, To W, Kwan S, Churakova T, Zurawski S, et al. Interleukin-23 rather than interleukin-12 is the critical cytokine for autoimmune inflammation of the brain. Nature. 2003; 421:744–748. [PubMed: 12610626]
37. Awasthi A, Riol-Blanco L, Jäger A, Korn T, Pot C, Galileos G, Bettelli E, Kuchroo VK, Oukka M. Cutting edge: IL-23 receptor GFP reporter mice reveal distinct populations of IL-17-producing cells. J Immunol. 2009; 182:5904–5908. (2009). [PubMed: 19414740]
38. Prinz M, Schmidt H, Mildner A, Knobeloch KP, Hanisch UK, Raasch J, Merkler D, Detje C, Gutcher I, Mages J, Lang R, et al. Distinct and nonredundant *in vivo* functions of IFNAR on myeloid cells limit autoimmunity in the central nervous system. Immunity. 2008; 28:675–686. [PubMed: 18424188]
39. Axtell RC, de Jong BA, Boniface K, van der Voort LF, Bhat R, De Sarno P, Naves R, Han M, Zhong F, Castellanos JG, Mair R, et al. T helper type 1 and 17 cells determine efficacy of interferon-β in multiple sclerosis and experimental encephalomyelitis. Nat Med. 2010; 16:406–412. [PubMed: 20348925]
40. Altare F, Durandy A, Lammas D, Emile JF, Lamhamedi S, Le Deist F, Drysdale P, Jouanguy E, Döffinger R, Bernaudin F, Jeppsson O, et al. Impairment of mycobacterial immunity in human interleukin-12 receptor deficiency. Science. 1998; 280:1432–1435. [PubMed: 9603732]
41. Duncan CJ, Mohamad SM, Young DF, Skelton AJ, Leahy TR, Munday DC, Butler KM, Morfopoulou S, Brown JR, Hubank M, Connell J, et al. Human IFNAR2 deficiency: Lessons for antiviral immunity. Sci Transl Med. 2015; 7:307ra154.
42. Chrencik JE, Patny A, Leung IK, Korniski B, Emmons TL, Hall T, Weinberg RA, Gormley JA, Williams JM, Day JE, Hirsch JL, et al. Structural and thermodynamic characterization of the TYK2 and JAK3 kinase domains in complex with CP-690550 and CMP-6. J Mol Biol. 2010; 400:413–433. [PubMed: 20478313]
43. Siu M, Pastor R, Liu W, Barrett K, Berry M, Blair WS, Chang C, Chen JZ, Eigenbrot C, Ghilardi N, Gibbons P, et al. 2-Amino-[1,2,4]triazolo[1,5-a]pyridines as JAK2 inhibitors. Bioorg Med Chem Lett. 2013; 23:5014–5021. [PubMed: 23870430]
44. Tsui V, Gibbons P, Ultsch M, Mortara K, Chang C, Blair W, Pulk R, Stanley M, Starovasnik M, Williams D, Lamers M, et al. A new regulatory switch in a JAK protein kinase. Proteins. 2011; 79:393–401. [PubMed: 21117080]
45. Reich K, Langley RG, Papp KA, Ortonne JP, Unnebrink K, Kaul M, Valdes JM. A 52-week trial comparing briakinumab with methotrexate in patients with psoriasis. N Engl J Med. 2011; 365:1586–1596. [PubMed: 22029980]
46. Grant A, Ponia SS, Tripathi S, Balasubramanian V, Miorin L, Sourisseau M, Schwarz MC, Sanchez-Seco MP, Evans MJ, Best SM, Garcia-Sastre A. Zika virus targets human STAT2 to inhibit type I interferon signaling. Cell Host Microbe. 2016; 19:1–9. [PubMed: 26764588]
47. Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med. 2015; 372:793–795. [PubMed: 25635347]
48. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods. 2013; 10:1213–1218. [PubMed: 24097267]
49. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013; 8:2281–2308. [PubMed: 24157548]

50. Kulagowski JJ, Blair W, Bull RJ, Chang C, Deshmukh G, Dyke HJ, Eigenbrot C, Ghilardi N, Gibbons P, Harrison TK, Hewitt PR, et al. Identification of imidazo-pyrrolopyridines as a novel and potent JAK1 inhibitors. J Med Chem. 2012; 55:5901–5922. [PubMed: 22591402]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Single Sentence Summary

Resolving *TYK2* locus genotype-to-phenotype conflicts reveals an immune signaling optimum that may be targeted towards precision medicine for autoimmune diseases.

Europe PMC Funders Author Manuscripts

Accessible Summary

Determining the biological consequences of the thousands of genetic variants that contribute to common diseases for the purpose of improving healthcare provision is challenging. Genetic variants that influence autoimmune disease have been identified in the tyrosine kinase 2 (*TYK2*) gene, but conflicting evidence regarding their biological impact obscures the therapeutic potential of TYK2. By resolving this conflict, we have revealed a genetic effect that drives an optimal degree of immune signaling: low enough to be strongly protective against autoimmunity but high enough to prevent immunodeficiency. These findings indicate TYK2 as a potential drug target across multiple autoimmune conditions.

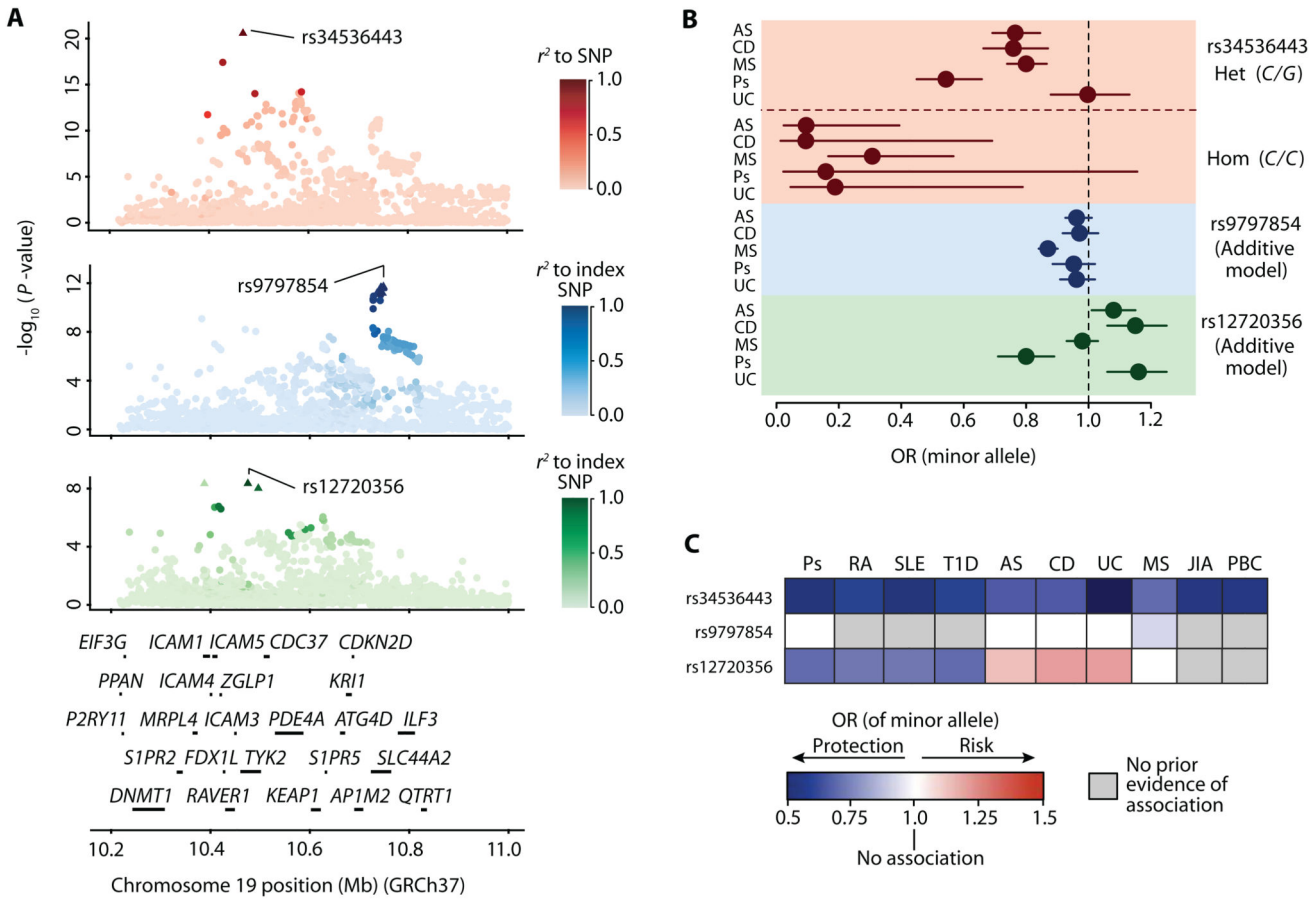

Fig. 1. Genetic associations in the *TYK2* gene region across autoimmune disorders.

(**A**) Joint association signal plots generated through multinomial logistic regression analysis. Primary association with the signal at rs34536443 (red). Secondary association with the rs9797854 index SNP (blue). Tertiary association with the rs12720356 index SNP. Scale bars indicate degree of LD (r^2) relative to each index SNP (green). Triangles indicate SNPs in the associated SNP cluster (90% credible set). The genes in the region are shown below the signal plots. (**B**) Odds ratios (ORs) for index SNPs for each of the five diseases and 95% confidence intervals (horizontal bars). For rs34536443, a non-additive effect was observed and ORs are shown separately for *C/G* heterozygosity (Het.) and *C/C* homozygosity (Hom.). For rs9797854 and rs12720356, the ORs fit an additive model. (**C**) Summary of associations of the minor alleles at rs34536443, rs9797854 and rs12720356 with psoriasis (Ps), rheumatoid arthritis (RA), systemic lupus erythematosus (SLE), type 1 diabetes (T1D), ankylosing spondylitis (AS), Crohn's disease (CD), ulcerative colitis (UC), multiple sclerosis (MS), juvenile idiopathic arthritis (JIA), and primary biliary cirrhosis (PBC). The rs34536443 association for UC is shown for the homozygous state; for all other associations the allelic OR is depicted.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

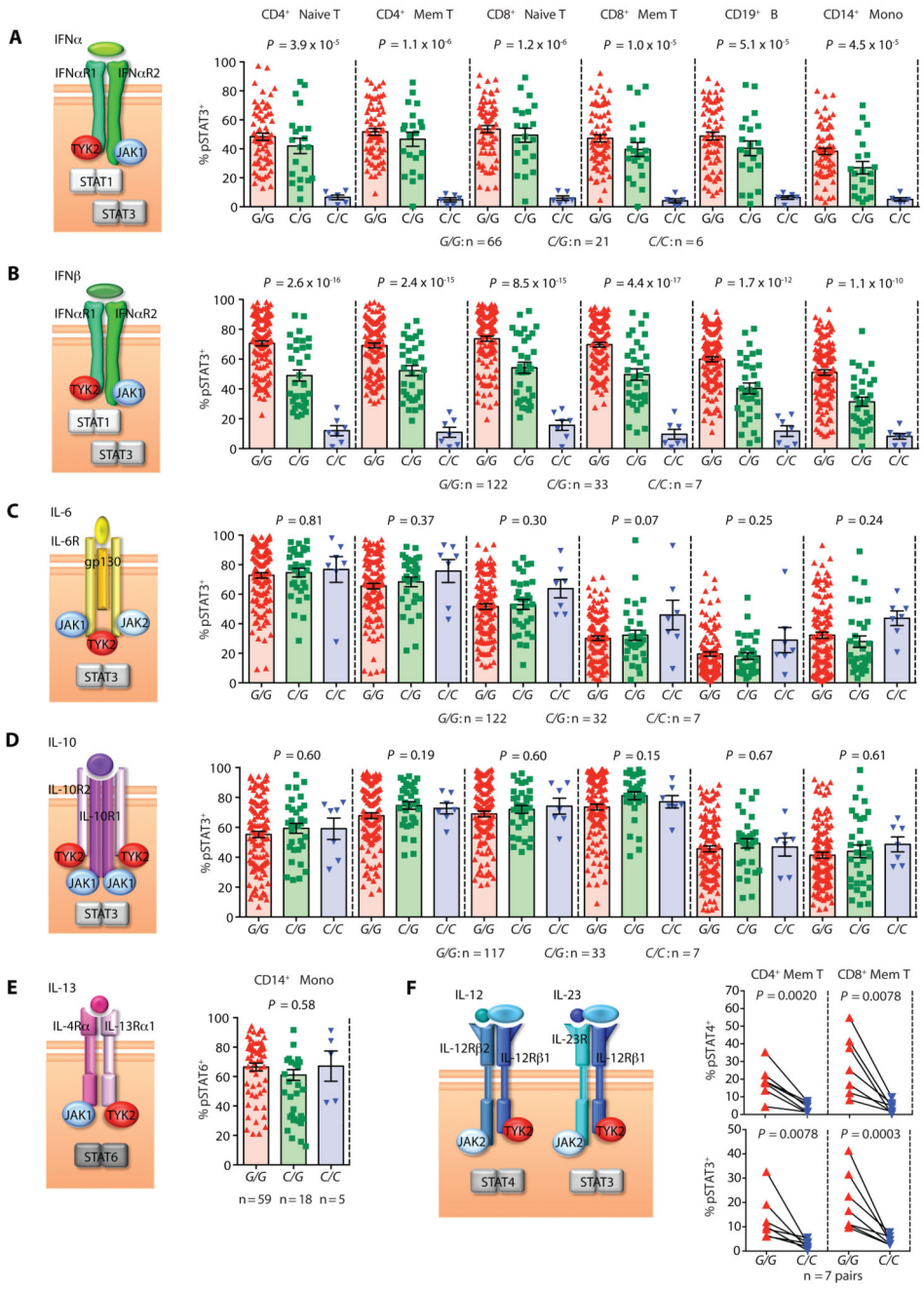

Fig. 2. Cytokine-induced STAT phosphorylation by rs34536443 genotype in primary human immune cells.

Percentage of pSTAT$^+$ cells by rs34536443 genotype for CD4$^+$ naïve T cells, CD4$^+$ memory (Mem) T cells, CD8$^+$ naïve T cells, CD8$^+$ memory T cells, CD19$^+$ B cells and CD14$^+$ monocytes (Mono) upon stimulation with (**A**) IFNα, (**B**) IFNβ, (**C**) IL-6 and **(D)** IL-10. CD14$^+$ monocytes were stimulated with (**E**) IL-13. CD4$^+$ memory T cells and CD8$^+$ memory T cells were stimulated with (**F**) IL-12 (top graph) and IL-23 (bottom graph). For each cytokine a schematic of its receptor chains and the TYK2, JAK and STAT molecules

that mediate its signaling are shown to the left of the respective graphs. For panels **A**-**E,** *P*-values provided are for a non-additive genetic model. For panels **E** and **F,** Wilcoxon matched-pairs rank tests were used. Error bars = ±SEM.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

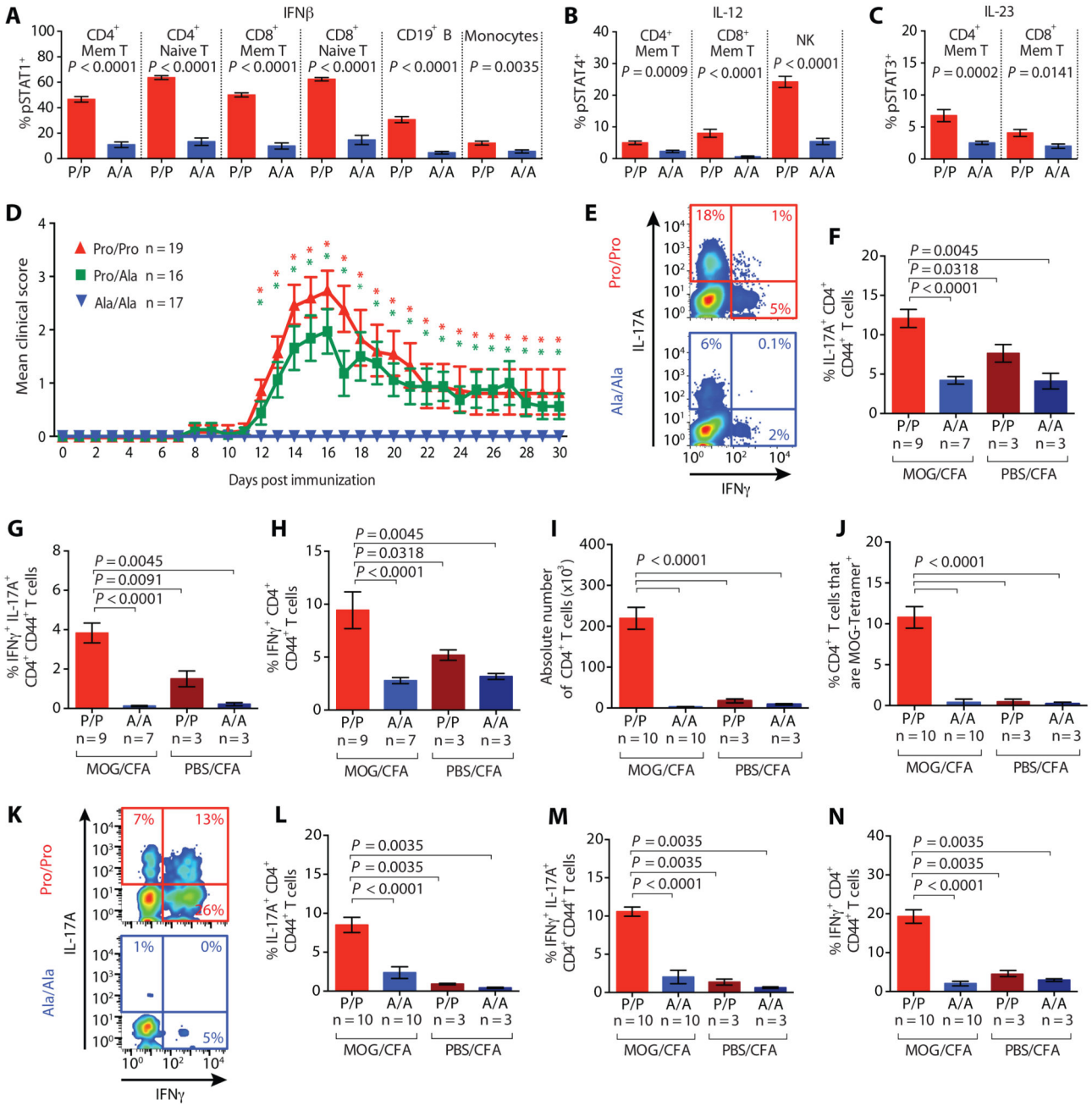

Fig. 3. Tyk2 Ala-1124 homozygous mice have impaired cytokine signaling and are protected against EAE.

Phosphorylation of STAT (pSTAT) in splenic immune cells from Tyk2 Pro-1124 (red, denoted by P/P; $n = 8$) and Ala-1124 (blue, denoted by A/A; $n = 8$) homozygous mice after stimulation with (**A**) IFNβ, (**B**) IL-12, and (**C**) IL-23. (**D**) EAE disease course shown as mean clinical score following immunization of Pro/Pro (red triangles), Pro/Ala (green squares) and Ala/Ala-1124 (blue triangles) mice. *denotes $P<0.05$ comparing Pro/Pro-1124 and Pro/Ala-1124 to Ala/Ala-1124 mice, respectively. (**E**) Representative intracellular

IL-17A and IFNγ staining, and proportion of (**F**) IL-17A$^+$, (**G**) IFNγ$^+$ IL-17A$^+$, and (**H**) IFNγ$^+$ CD4$^+$ memory T cells in draining lymph nodes at peak EAE (or equivalent time-point in mice without disease) following immunization with MOG/CFA or PBS/CFA (negative control). (**I**) Absolute CD4$^+$ T cell numbers in the central nervous system (CNS) at peak EAE. (**J**) Proportion of MOG-responsive CD4$^+$ T cells in the CNS at peak EAE. (**K**) Representative intracellular IL-17A and IFNγ staining, and proportion of (**L**) IL-17A$^+$, (**M**) IFNγ$^+$ IL-17A$^+$, and (**N**) IFNγ$^+$ CD4$^+$ memory T cells in the CNS at peak EAE following immunization with MOG/CFA or PBS/CFA. *P*-values were calculated using Mann-Whitney *U* tests. Error bars = ±SEM.

Europe PMC Funders Author Manuscripts

A

	rs34536443 genotype			
	G/G	*G/C*	*C/C*	Total
In UK Biobank	105,794 (90.63%)	10,689 (9.16%)	249 (0.21%)	116,732 (100%)
Mycobacterial	20 (86.96%)	3 (13.04%)	0 (0.00%)	23
Specific bacterial (e.g. *S. aureus*)	54 (90.00%)	5 (8.33%)	1 (1.67%)	60
Specific viral (e.g. HSV, VZV, viral encephalitis)	93 (96.88%)	3 (3.12%)	0 (0.00%)	96
Mucocutaneous candidiasis	46 (88.46%)	6 (11.54%)	0 (0.00%)	52
Total	213 (92.21%)	17 (7.36%)	1 (0.43%)	231

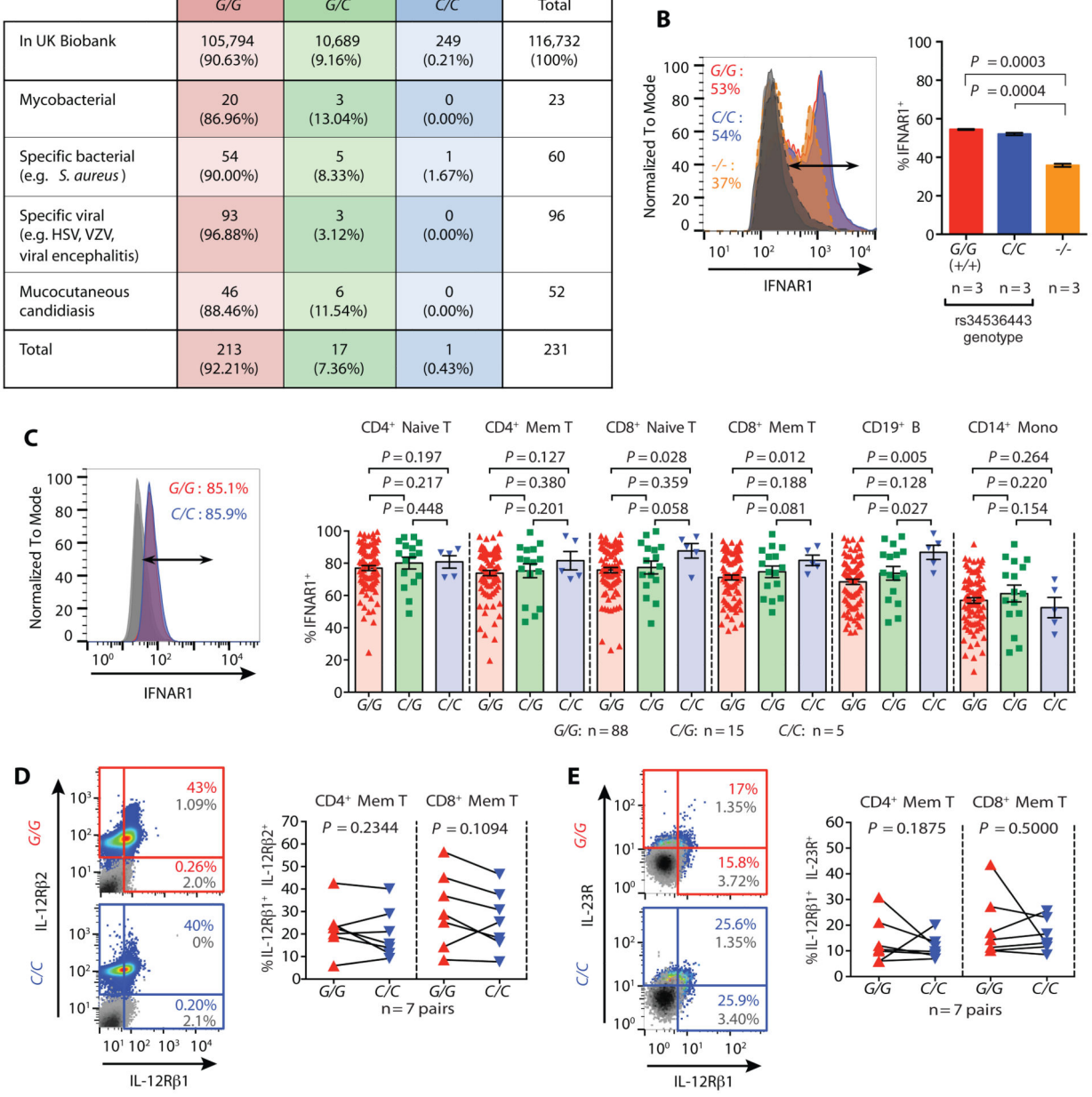

Fig. 4. Rs34536443 genotype does not lead to immunodeficiency based on UK Biobank health record analysis and cell surface cytokine receptor expression.

(**A**) Frequency of individuals in the UK Biobank hospitalized due to TYK2 immunodeficiency-relevant infections and categorized by rs34536443 genotype. (**B**) Cell surface type I IFN receptor chain 1 (IFNAR1) expression on HEK 293T cell lines genetically modified at the *TYK2* locus using CRISPR-Cas9. (**C**) Cell surface expression of IFNAR1 according to rs34536443 genotype. Representative surface expression shown for CD4$^+$ naïve T cells (left), and quantification of IFNAR1 expression across subsets (right).

(**D**) Cell surface expression of the IL-12 receptor chains (IL-12Rβ1 and IL-12Rβ2) on pre-activated T cells by rs34536443 genotype. Representative cytokine receptor expression shown for $CD4^+$ memory T cells (left), with corresponding quantification in $CD4^+$ and $CD8^+$ memory T cells (right). (**E**) Cell surface expression of the IL-23 receptor chains (IL-12Rβ1 and IL-23R) on pre-activated T cells according to rs34536443 genotype. Representative cytokine receptor expression shown for $CD8^+$ memory T cells (left), with corresponding quantification in $CD4^+$ and $CD8^+$ memory T cells (right). The grey histogram (panel **B**) and plots (panels **D** and **E**) show isotype control staining. For panels **B** and **C,** *P*-values were calculated using Mann-Whitney *U* tests. For panel **D,** *P*-values were calculated using Wilcoxon matched-pairs rank tests. Error bars = ±SEM.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

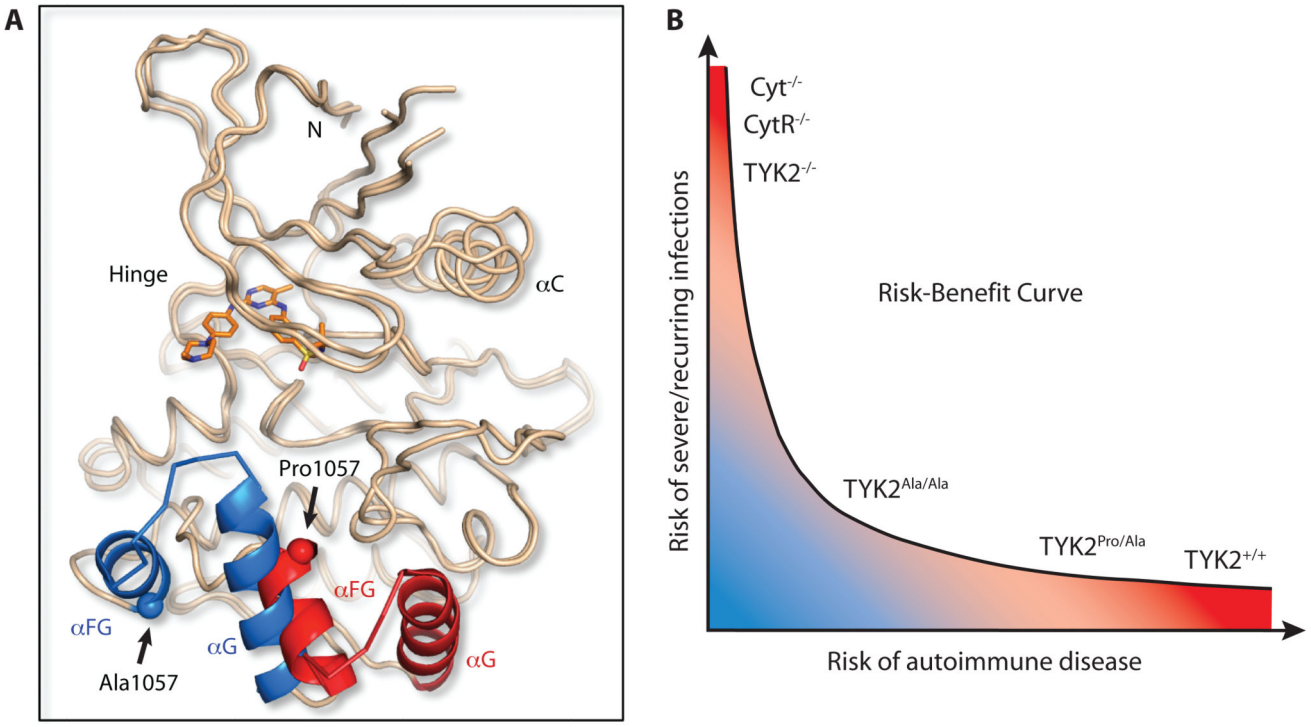

Fig. 5. JAK2 kinase domain Pro➔Ala substitution confers conformational change.

(**A**) Crystal structure of the TYK2-homologous JAK2-Ala-1057 kinase domain superimposed onto the JAK2-Pro-1057 kinase domain structure (PDB accession 4JI9). Both domains were complexed with the JAK2 inhibitor TG101209 (orange molecule) to stabilize the crystal structures. When the alanine residue is present at position 1057 a 15Å shift is observed in the αFG-αG helices (shown in blue) relative to their conformation when the proline residue is present (helices shown in red). (**B**) Relationship between genetically determined differences in TYK2-relevant cytokine signaling and risk for autoimmune disease versus severe and recurring infections. The kinase domain Pro➔Ala substitution likely confers a conformational change such that $TYK2^{Ala/Ala}$ mediates an optimal degree of cytokine signaling that may be low enough to prevent autoimmunity but high enough to prevent immunodeficiency. Cyt, cytokine; CytR, cytokine receptor.