Exome-wide association study reveals novel psoriasis susceptibility locus at *TNFSF15* and rare protective alleles in genes contributing to type I IFN signalling

Nick Dand[1,†], Sören Mucha[2,†], Lam C. Tsoi[3,4,5,†], Satveer K. Mahil[6], Philip E. Stuart[3], Andreas Arnold[7], Hansjörg Baurecht[8], A. David Burden[9], Kristina Callis Duffin[10], Vinod Chandran[11,12,13,14], Charles J. Curtis[15,16], Sayantan Das[5], David Ellinghaus[2], Eva Ellinghaus[2], Charlotta Enerback[17], Tõnu Esko[18], Dafna D. Gladman[11,13,14], Christopher E. M. Griffiths[19], Johann E. Gudjonsson[3], Per Hoffman[20,21], Georg Homuth[22], Ulrike Hüffmeier[23], Gerald G. Krueger[10], Matthias Laudes[24], Sang Hyuck Lee[15,16], Wolfgang Lieb[25], Henry W. Lim[26], Sabine Löhr[23], Ulrich Mrowietz[8], Martina Müller-Nurayid[27], Markus Nöthen[21], Annette Peters[27], Proton Rahman[28], André Reis[23], Nick J. Reynolds[29,30], Elke Rodriguez[8], Carsten O. Schmidt[31], Sarah L. Spain[1], Konstantin Strauch[27], Trilokraj Tejasvi[3], John J. Voorhees[3], Richard B. Warren[32], Michael Weichenthal[33], Stephan Weidinger[8], Matthew Zawistowski[5], Rajan P. Nair[3], Francesca Capon[1], Catherine H. Smith[6], Richard C. Trembath[1], Goncalo R. Abecasis[5], James T. Elder[3,34,‡], Andre Franke[2,‡], Michael A. Simpson[1,‡,*], Jonathan N. Barker[6,‡,*]

[1]Division of Genetics and Molecular Medicine, Faculty of Life Sciences & Medicine, King's College London, London, UK

[2]Institute of Clinical Molecular Biology, Christian-Albrechts-University of Kiel, Kiel, Germany

[3]Department of Dermatology, University of Michigan Medical School, Ann Arbor, Michigan, USA

[4]Department of Computational Medicine & Bioinformatics, University of Michigan Medical School, Ann Arbor, Michigan, USA

[5]Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, Michigan, USA

[6]St John's Institute of Dermatology, Faculty of Life Sciences & Medicine, King's College London, London, UK

[7]Clinic and Polyclinic of Dermatology, University Medicine Greifswald, Greifswald, Germany

[8]Department of Dermatology, Venereology and Allergy, University Hospital Schleswig-Holstein, Campus Kiel, Kiel, Germany

[9]Institute of Infection, Inflammation and Immunity, University of Glasgow, Glasgow, UK

[10]Department of Dermatology, University of Utah, Salt Lake City, Utah, USA

[11]Department of Medicine, University of Toronto, Toronto, Ontario, Canada

[12]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada

[13]Institute of Medical Science, University of Toronto, Toronto, Ontario, Canada

[14]Krembil Research Institute, University Health Network, Toronto, Ontario, Canada

[15]NIHR Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London, London, UK

[16]Social Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

[17]Department of Clinical and Experimental Medicine, Division of Cell Biology and Dermatology, Linköping University, Linköping, Sweden

[18]Estonian Biobank, Estonian Genome Center, University of Tartu, Tartu, Estonia

[19]Dermatology Centre, Salford Royal Hospital, University of Manchester, Manchester Academic Health Science Centre, Manchester, UK

[20]Genomics Research Group, Department of Biomedicine, University of Basel, Basel, Switzerland

[21]Institute of Human Genetics, University of Bonn, Bonn, Germany

[22]Department of Functional Genomics, Interfaculty Institute for Genetics and Functional Genomics, University Medicine and Ernst-Moritz-Arndt-University Greifswald, Greifswald, Germany

[23]Institute of Human Genetics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

[24]I. Department of Medicine, Christian-Albrechts-University of Kiel, Kiel, Germany

[25]Institute of Epidemiology and Biobank PopGen, Christian-Albrechts-University of Kiel, Kiel, Germany

[26]Department of Dermatology, Henry Ford Hospital, Detroit, Michigan, USA

[27]Institute of Genetic Epidemiology, Helmholtz Zentrum Munich, Neuherberg, Germany

[28]Memorial University, St. John's, Newfoundland, Canada

[29]Dermatological Sciences, Institute of Cellular Medicine, Newcastle University Medical School, Newcastle upon Tyne, UK

[30]Department of Dermatology, Royal Victoria Infirmary, Newcastle Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK

[31]Institute for Community Medicine, Study of Health in Pomerania/KEF, University Medicine Greifswald, Greifswald, Germany

[32]Dermatology Centre, Salford Road NHS Foundation Trust, The University of Manchester, Manchester Academic Health Science Centre, Manchester, UK

[33]Department of Dermatology, University Medical Center Schleswig-Holstein, Campus Kiel, Kiel, Germany

[34]Ann Arbor Veterans Hospital, Ann Arbor, Michigan, USA

* Corresponding authors:

Jonathan N. Barker
jonathan.barker@kcl.ac.uk
Phone: +44(0)20 7188 6412
Fax: +44(0)20 7188 8050

Michael A. Simpson
michael.simpson@kcl.ac.uk
Phone: +44(0)20 7848 8168
Fax: +44(0)20 7188 2585

† The authors wish it to be known that, in their opinion, the first three authors should be regarded as joint First Authors

‡ The authors wish it to be known that, in their opinion, the last four authors should be regarded as joint Senior Authors

Abstract

Psoriasis is a common inflammatory skin disorder for which multiple genetic susceptibility loci have been identified, but few resolved to specific functional variants. In this study we sought to identify common and rare psoriasis-associated gene-centric variation. Using exome arrays we genotyped four independent cohorts, totalling 11,861 psoriasis cases and 28,610 controls, aggregating the dataset through statistical meta-analysis. Single variant analysis detected a previously unreported risk locus at *TNFSF15* (rs6478108; $p = 1.50\times10^{-8}$, OR = 1.10), and association of common protein-altering variants at 11 loci previously implicated in psoriasis susceptibility. We validate previous reports of protective low-frequency protein-altering variants within *IFIH1* (encoding an innate antiviral receptor) and *TYK2* (encoding a Janus kinase), in each case establishing a further series of protective rare variants (minor allele frequency < 0.01) via gene-wide aggregation testing (*IFIH1*: $p_{burden} = 2.53\times10^{-7}$, OR = 0.707; *TYK2*: $p_{burden} = 6.17\times10^{-4}$, OR = 0.744). Both genes play significant roles in type I interferon (IFN) production and signalling. Several of the protective rare and low-frequency variants in *IFIH1* and *TYK2* disrupt conserved protein domains, highlighting potential mechanisms through which their effect may be exerted.

Introduction

Psoriasis is a common inflammatory hyperproliferative skin disorder with a significant genetic component to disease pathogenesis (1-3). It affects up to 2% of people worldwide, with affected individuals suffering high social and economic costs and increased morbidity and mortality (1, 4). Previous large-scale genome-wide association studies and meta-analyses have identified 63 loci that contribute to psoriasis susceptibility in populations of European origin (5-16). Recent studies have refined the understanding of the allelic architecture of psoriasis risk at several of these loci, including the major histocompatibility complex (MHC), through the detection of multiple independent secondary signals (8, 12, 17, 18). Many psoriasis risk loci harbour genes encoding components of disease-relevant biological processes, including innate and adaptive immune pathways and skin barrier function (2). Nevertheless, the precise molecular mechanisms through which the associated genetic variation confers psoriasis susceptibility remain uncertain for the majority of these signals (19). The identification of disease-associated protein-altering variation, including the effects of rare alleles, has the potential to illuminate the mechanisms underpinning the pathogenic process and to identify putative targets for therapeutic intervention. Notably in psoriasis, investigation of the protective common allele encoding a glutamine residue at position 381 of the interleukin-23 (IL-23) receptor has validated aberrant Th17 signalling as a key disease driver (20), consistent with the remarkable efficacy of therapeutics targeting this pathway (21).

Potential roles for common and low-frequency protein-altering variants in psoriasis susceptibility have been investigated in the Han Chinese (22, 23) and European (8) populations, but until now the contribution of rare protein-altering alleles to the disease architecture has not been systematically explored in any population. Here we present the most comprehensive investigation to date of protein-altering variation in psoriasis risk in the European population. The analysis encompasses four independent exome array association studies, referred to here as the UK, Estonia, Germany and Michigan studies. These were combined through meta-analysis to total 11,861 psoriasis cases and 28,610 controls (Supplementary Table 1). Our analysis focused on genetic variation outside of the MHC region, in which the psoriasis-associated HLA-C*06:02 allele and independent secondary

signals have been the subject of extensive investigation elsewhere (17, 18, 24-26). After quality control (QC), 167,587 single nucleotide variants (SNVs) that were successfully genotyped in each of the four cohorts were investigated. This set included 155,870 variants located within protein coding regions (and associated splice sites) of the genome and a further 11,717 non-coding SNVs including many tagging previously reported disease-associated SNVs. The allele frequency spectrum of the set of genotyped variants is skewed towards rare and low-frequency variants (Supplementary Table 3).

Results

Single marker association testing was performed in each of the four case-control cohorts using a linear mixed model with an empirically estimated relatedness matrix to control for population structure (27), and results were aggregated across studies via meta-analysis (Materials and Methods). Results for all variants achieving an association p-value $p < 1\times10^{-5}$ are summarised in Supplementary Table 5.

Single marker association tests uphold established psoriasis susceptibility loci

Of 67 previously reported independent psoriasis susceptibility signals across 63 loci, we were able to test for disease association at 24, either directly using the reported lead SNV or via a proxy ($r^2 > 0.8$ with the reported lead SNV). We observe significant disease associations at 23 (20 with genome-wide significance, $p < 5\times10^{-8}$, and three with $p < 10^{-4}$; Supplementary Table 6). We find no evidence of association at the recently reported 13q14.11 locus that encompasses *COG6* (10) (rs7993214: $p = 0.0589$; OR = 1.04). It should be noted that before QC the UK, Estonia, Germany and Michigan studies collectively included 7,885 psoriasis cases that were present in previously published analyses (6, 8, 12, 14) (Supplementary Table 1). As such, the associations that originate from these previous reports should not be considered independently replicated here.

Previously unreported genome-wide association at one locus

We detect genome-wide significant association at one further locus, mapping to *TNFSF15* at 9q32. An association is observed with the intronic variant rs6478108 ($p = 1.50\times10^{-8}$; OR = 1.10;

Supplementary Table 5). *TNFSF15*, encoding a member of the tumor necrosis factor superfamily of cytokines, is primarily expressed in endothelial cells and has previously been implicated in susceptibility to Crohn's disease (28). Although a previous study found nominal association between variants in this locus and psoriasis susceptibility in a Hungarian population (29), we establish 9q32 as a genome-wide significant susceptibility locus for the first time here.

Established association signals map to protein-altering variants at 11 susceptibility loci

Within psoriasis susceptibility loci, an ongoing challenge is to fine-map association signals to determine the underlying causal variants. Disease-associated protein-altering variants represent plausible candidates through which psoriasis risk is conferred. We searched for protein-altering SNVs with a consistent direction of effect across all four studies and exome-wide significant association ($p \leq 3.0\times10^{-7}$). We found 19 such variants within 11 different loci, all of which were previously reported susceptibility loci (Table 1; Supplementary Table 5). These observations extend the list of putative causal protein-altering alleles previously reported (8), most notably defining an additional candidate causal variant in *ERAP1* (rs30187: p.K528R; $p = 2.19\times10^{-11}$) that is predicted to be damaging by PolyPhen-2 and has a CADD score of 20.6 (Supplementary Table 5). Furthermore, conditional analysis indicates that this variant can account for the observed association of rs27432, the lead SNV in this locus reported by Tsoi *et al.* (8) (Supplementary Table 5).

Rare variant aggregation tests identify protective alleles for type I IFN genes

Despite the substantial sample size of the current study, evaluation of the contribution of individual rare and low-frequency variants to psoriasis susceptibility is limited by statistical power to detect association. We therefore performed a series of gene-based tests, aggregating variants with low minor allele frequency (MAF). At each MAF threshold (0.01 or 0.05) we performed a burden test to detect an excess of rare alleles in cases or controls, and a SKAT test, which is designed to detect scenarios in which the effects of the aggregated variants have different direction or magnitude (30) (Materials and Methods). This testing regime identified two genes, *IFIH1* and *TYK2*, with exome-wide significant

evidence of association ($p_{gene} < 2.5\times10^{-6}$; Table 2). Both *IFIH1* and *TYK2* are located in loci previously implicated in common variant studies of psoriasis risk (5).

In *IFIH1*, single marker association testing had identified two genome-wide significant psoriasis-associated protein-coding variants: the common rs1990760 (p.A946T; $MAF_{controls} = 0.38$; $p = 4.73\times10^{-18}$; OR = 0.86) and the low-frequency rs35667974 (p.I923V; $MAF_{controls} = 0.02$; $p = 1.10\times10^{-15}$; OR = 0.55). The latter contributes to the observed gene-based association of variants in *IFIH1* with MAF < 0.05 ($p_{SKAT} = 1.19\times10^{-20}$; $p_{burden} = 1.84\times10^{-19}$), although we also observe evidence of association with a MAF threshold of 0.01 ($p_{burden} = 2.53\times10^{-7}$; $p_{SKAT} = 6.02\times10^{-5}$). This association remains when conditioning on both rs35667974 and rs1990760 ($p_{conditional\text{-}burden} = 1.36\times10^{-8}$; $p_{conditional\text{-}SKAT} = 4.46\times10^{-6}$), or on either individually (Supplementary Table 8). Examination of allele frequencies of individual rare and low-frequency coding SNVs in *IFIH1* (Supplementary Table 10) reveals differences between cases and controls ($p < 0.05$) for six variants, each located within predicted functional domains of MDA5, the antiviral receptor encoded by this gene (Figure 1A). Notably, at each of these sites the minor allele is associated with decreased psoriasis risk, consistent with a previous study reporting two rare variant associations in this gene (31).

In *TYK2*, which encodes one of the Janus family of kinases (32), we detect a gene-level association across variants with MAF < 0.05 ($p_{SKAT} = 6.34\times10^{-41}$; $p_{burden} = 1.47\times10^{-39}$). As with *IFIH1*, a strong single-variant association contributes to the aggregated signal (rs34536443: $MAF_{cases} = 0.023$; $MAF_{controls} = 0.044$; $p = 1.72\times10^{-42}$; OR = 0.51). Nevertheless, the variants with MAF < 0.01 also display evidence for disease association ($p_{SKAT} = 2.82\times10^{-4}$; $p_{burden} = 6.17\times10^{-4}$). There is a complex linkage disequilibrium (LD) structure between individual variants that have been previously reported at this locus (see Discussion). Our data suggest that the observed aggregate rare variant association is independent of the single marker associations ($p_{conditional\text{-}SKAT} = 7.21\times10^{-5}$; $p_{conditional\text{-}burden} = 1.45\times10^{-4}$), although a suitable proxy to facilitate conditional analysis with the common disease-associated intronic SNV rs280519 was unavailable for this analysis (5). The low-frequency variant rs34536443 results in a substitution in TYK2's kinase domain (p.P1104A), as does the only rare variant that is nominally associated in a single marker test (rs35018800: $p = 0.0003$; OR = 0.68; Figure 1B; Supplementary Table 10).

Since *IFIH1* and *TYK2* are located in known psoriasis susceptibility loci, we further scrutinized genes in all previously reported psoriasis susceptibility loci (Online Methods). We observed suggestive evidence for aggregated rare variant association at four further genes (*IL23R*, *TNFAIP3*, *DDX58* and *STAT2*; Supplementary Table 9), the rare alleles displaying a protective effect in each case. Of these, we note that *DDX58* ($p_{burden} = 3.01\times10^{-5}$; $p_{SKAT} = 7.82\times10^{-5}$ for MAF < 0.05) encodes RIG-I, a paralog of the MDA5 receptor (encoded by *IFIH1*) with a closely related function (33). The most strongly associated single marker in the region (rs657454; $p = 2.16\times10^{-5}$; OR = 1.08; $MAF_{controls} = 0.38$) is not responsible for the observed association ($p_{conditional\text{-}burden} = 3.15\times10^{-5}$; $p_{conditional\text{-}SKAT} = 3.07\times10^{-5}$), although without a suitable proxy for conditional analysis we cannot fully rule out that the association is driven by the previously reported (8) common SNV rs11795343 (r^2 with rs657454 = 0.411). Furthermore *STAT2* ($p_{burden} = 3.80\times10^{-5}$; $p_{SKAT} = 9.48\times10^{-5}$ for MAF < 0.05), like *IFIH1, DDX58* and *TYK2*, also encodes an important component of the type I IFN signaling pathway.

Discussion

The systematic analysis of protein-altering variation reported here allows a thorough examination of the contribution of functional genetic mechanisms to psoriasis risk. For each locus in which we identified robustly associated single variants, the Supplementary Note provides a summary of evidence for functional involvement. For several loci (including those harbouring *IL23R*, *IL13* and *STAT2*), the most strongly associated functional variants remain those previously suggested by Tsoi *et al.* (8). Findings at other loci (1q21.3, 6q21, 16p11.2, 20q13.13) offered less clear interpretation but we did not find sufficient evidence to reject existing disease models involving candidate disease genes *LCE3B/C*, *TRAF3IP2*, *FBXL19* and *RNF114*, respectively (7, 34-36). We note the significant association of rs30187 in *ERAP1* ($p = 2.19\times10^{-11}$; OR = 0.89), a missense variant that can explain the association signal at the previously proposed causal variant rs27044 (8) and which, unlike the latter SNV, is predicted to be deleterious by both PolyPhen-2 and CADD (Supplementary Table 5). We also identified one significantly associated missense variant in the recently reported 19q13.33 locus (16):

rs602662 in the gene *FUT2* ($p = 3.29\times10^{-8}$; OR = 1.09). This gene encodes α-(1,2) fucosyltransferase, a Lewis antigen system enzyme that is central to determining an individual's secretor status (37) and is associated with protection from and susceptibility to certain viral, bacterial and fungal infections (38-40).

TYK2, as a Janus kinase, is widely expressed and facilitates a broad range of intracellular signaling processes (32). It provides another link between psoriasis-associated innate and adaptive immune pathways, having been shown, for example, both to mediate Th17 cell responses to IL-23 signaling and Th1 responses to IL-12 signaling, and to regulate type I interferon (IFN) signalling (41, 42). Our results highlight the complex LD structure at this locus, with protective associations at two previously identified independent missense variants which are each predicted to impair protein function (rs34536443 described earlier; rs12720356: $p = 1.39\times10^{-16}$; OR = 0.76; $MAF_{cases} = 0.068$; $MAF_{controls} = 0.083$) (5, 8). A third SNV, rs2304256 ($p = 2.88\times10^{-23}$; OR = 0.81; $MAF_{cases} = 0.243$; $MAF_{controls} = 0.280$), is in weak LD with both rs34536443 and rs12720356 ($r^2 = 0.107$ and 0.290, respectively) and its association disappears when conditioning on either SNV ($p_{conditional} = 0.0559$ and $p_{conditional} = 0.1063$, respectively). Our data included no proxy SNV for rs280519, another independent psoriasis signal with which rs2304256 is also in moderate LD ($r^2 = 0.357$). In other immune-mediated diseases, rs2304256 has been shown to represent a synthetic association due to neighbouring rarer variants including rs34536443 and rs12720356 (43, 44). It is notable that the observed association at *TYK2* under the rare variant aggregation tests is driven by two protective alleles which disrupt TYK2's kinase domain (Figure 1B). This may suggest that the catalytic activity of TYK2 helps to initiate and maintain the positive feedback loops that culminate in psoriatic inflammation. Indeed, our independently associated common SNV rs12720356 leads to the substitution p.I684S within the pseudokinase JAK-homology 2 domain (32, 42, 45), while conversely our likely synthetic association rs2304256 (p.V362F) impacts neither kinase domain.

IFIH1 encodes the innate antiviral receptor MDA5, which detects and binds to double-stranded RNA, promoting a pro-inflammatory type I interferon (IFN) response (46). Rare variants in *IFIH1*, including three of the six variants underlying our aggregation tests that exhibit nominal disease association, have previously been shown to be protective for type I diabetes (47), with evidence that

the associated rare alleles lead to a decrease in downstream IFNβ expression arising from impaired signal propagation (48). It is evident that this pathway is also relevant to psoriasis pathogenesis indicating potential shared mechanisms at this locus in these immune-mediated diseases. The most strongly associated rare variant is rs35667974, whose minor (C) allele exhibits a large protective effect ($p = 1.10\times10^{-15}$; $MAF_{cases} = 0.010$; $MAF_{controls} = 0.020$; OR = 0.55). This SNV is one of two independent rare variants at the *IFIH1* locus previously implicated by Li *et al.* in psoriasis susceptibility (31), the other (rs10930046) having not been tested in our study. Our single marker tests identified one further protein-altering variant with exome-wide significant disease association, but the association at the common variant rs1990760 (reported previously (10)) is lost when conditioning on rs35667974 ($p_{conditional} = 0.9619$), implying that it is a consequence of LD. It is also not predicted to be damaging (Supplementary Table 5), which suggests rs35667974 could represent the more likely functional variant at this locus.

Several of the variants we report here exhibit some degree of effect size heterogeneity in our meta-analysis (Supplementary Table 5). Notably, many variants display only modest evidence for association in the Estonia cohort, likely driven by the relatively small sample size for this study (Supplementary Table 1). However, each of the variants has consistent direction of effect across all four studies and are consistent with established psoriasis susceptibility signals, and as such represent robust associations.

The results of our meta-analysis contribute to our understanding of several mechanisms of psoriasis pathogenesis. However, it might have been anticipated that the large study size and exome-wide genotyping coverage would result in more novel biological insights than was borne out in practice. We therefore examined how completely the 167,587 variants in our study covered the autosomal protein-altering variants (outside the MHC region and predicted to impair protein function) that are observed in 33,370 European whole-exome- or whole-genome-sequenced samples in the ExAC reference dataset (Supplementary Table 11). Of 14,123 common variants ($MAF \geq 0.05$) in ExAC, 5,487 (38.9%) are absent from at least one version of the exome arrays used across our four studies. A further 1,564 (11.1%) were removed from the analysis during genotyping QC, meaning that 7,072 (50.1%) were tested in our analysis. A similar proportion of low-frequency

($0.01 \leq \text{MAF} < 0.05$) and rare ($0.001 \leq \text{MAF} < 0.01$) SNVs were tested. As expected, coverage of very rare variants with MAF below 0.001 was substantially sparser, the drop in coverage being more pronounced the lower the MAF (Supplementary Table 11).

To assess the impact of this incomplete coverage on our ability to map established psoriasis susceptibility signals to functional variants, we searched for all SNVs that are in moderate LD with a previously reported association ($r^2 > 0.2$) in 1000 Genomes European samples and predicted to impair protein function by at least one of SIFT, PolyPhen-2 and CADD (Supplementary Table 12). We found 23 such variants, of which 11 (47.8%, consistent with overall coverage) were not tested in our meta-analysis and are therefore potentially interesting candidate variants for future association testing. The 12 variants which were tested include 8 with strong evidence of association (and present in Table 1). The remaining four variants are not exome-wide significantly associated with psoriasis susceptibility, but none are in strong LD with the corresponding established signal (r^2 range 0.29-0.69; Supplementary Table 12).

We note that the rare and low-frequency variants found to be associated in this study display broadly protective effects on psoriasis risk. We cannot exclude that this is due to selection bias, since the exome array design is based largely on variants observed in whole exome sequencing studies of a range of complex traits, which do not include psoriasis (http://genome.sph.umich.edu/wiki/Exome_Chip_Design). This could limit the probability that the array includes rare variants associated with increased psoriasis risk, either individually or via gene-wide aggregation tests.

Previously established risk loci account for around 28% of the estimated heritability of psoriasis (16). Based on the method of So *et al*. (49) we find that the newly reported association at *TNFSF15* explains 0.23% of estimated heritability (50). Aggregated rare and low-frequency variants (MAF < 0.05) in *IFIH1* account for 0.47% of estimated heritability (0.17% after conditioning on previously reported associations); for *TYK2* we estimate 0.80% (0.06% after conditional analysis). While these figures do not substantially increase the cumulative proportion of heritability explained to date, they do highlight the possibility that some fraction of the residual unexplained heritability will be due both to many as yet unidentified psoriasis susceptibility loci and to rare variants at existing

loci. Further efforts to isolate such variants will require larger sample sizes and more comprehensive coverage of the full frequency spectrum of genetic variation.

In summary, we establish genome-wide significant psoriasis associations at the *TNFSF15* locus and identify a series of alleles at established psoriasis loci with plausible evidence for causality based on predicted effects on protein structure and function. Our investigation of alleles at the low end of the frequency spectrum with variant aggregation tests has expanded our understanding of the allelic architecture of psoriasis risk at the *IFIH1* and *TYK2* loci. The observation that rare alleles that disrupt conserved domains within each gene have protective effects is compatible with the hypothesis that the common ancestral alleles of *IFIH1* and *TYK2* contribute to a robust immune response to pathogens, but this comes at the expense of increased risk of immune-mediated disease. Our findings support a central role for type I IFN signalling in psoriasis pathogenesis, consistent with clinical observations that type I IFN therapy can induce or exacerbate psoriasis symptoms (51, 52). They also highlight putative therapeutic mechanisms; the efficacy of other janus kinase inhibitors (53-55) suggest that TYK2 in particular may be a fruitful drug target.

Materials and Methods

Study samples and genotyping

The meta-analysis includes four independent studies, referred to as the UK, Estonia, Germany and Michigan studies. In each study, all samples were collected from unrelated individuals of European ancestry after obtaining written informed consent. Enrolment of subjects in each study was approved by the ethics boards of the participating institutions, in accordance with Declaration of Helsinki principles. All cases had been diagnosed with psoriasis vulgaris by a dermatologist. DNA was isolated from blood using standard methods.

UK data. Psoriasis cases (n = 1,971) were recruited as previously described (8). Further cases (n = 960) were recruited from centres in the UK via the Biomarkers of Systemic Treatment Outcomes in Psoriasis (BSTOP) cohort study (www.kcl.ac.uk/lsm/research/divisions/gmm/departments/dermatology/Research/stru/groups/bstop/index.aspx) after research ethics approval (REC reference

11/H0802/7). Unselected population-based controls (n = 6,400) were obtained from the 1958 British Birth Cohort. Genotyping was performed using Illumina HumanExome-12 v1.1 BeadChip and Illumina HumanOmniExpressExome-8 v1.2 BeadChip for psoriasis cases, and Illumina HumanExome-12 v1.0 BeadChip for controls (Supplementary Table 1).

German data. All German psoriasis cases (n = 2,928) were recruited through local outpatient services at either the Department of Dermatology at Christian-Albrechts-University Kiel, or the Department of Dermatology and Allergy at the Technical University of Munich. The psoriasis cases were genotyped using Illumina HumanExome-12 v1.1, HumanCoreExome-12 v1.1B or HumanCoreExome-24 v1.0A BeadChips. German healthy control individuals (n = 15,966) were obtained from the PopGen biobank, the KORA S4 survey (an independent population-based sample from the general population living in the region of Augsburg, southern Germany), the Heinz-Nixdorf Recall (HNR) cohort, Bonn, and the SHIP and SHIP-TREND cohorts (56) (from the Study of Health in Pomerania, a prospective longitudinal population-based cohort study in West Pomerania). German controls were genotyped using Illumina HumanExome-12 v1.0, HumanCoreExome-24 v1.0A or HumanOmniExpressExome-8 v1.2A BeadChips (Supplementary Table 1).

Estonian data. All Estonian samples were provided by the population-based biobank of the Estonian Genome Center, University of Tartu. Subjects were recruited by general practitioners (GP) and physicians in the hospitals. Participants in the hospitals were randomly selected from individuals visiting GP offices or hospitals. Diagnosis of psoriasis on the basis of clinical symptoms was posed by a general practitioner and confirmed by a dermatologist (n = 1,459). At the time of recruitment, the controls (n = 3,167) did not report diagnosis of osteoarthritis, psoriasis, or autoimmune diseases. All Estonian samples were genotyped using Illumina HumanExome-12 v1.1 or HumanCoreExome-24 v1.0 BeadChips (Supplementary Table 1).

Michigan data. Psoriasis cases (n = 6,344) and unrelated, unaffected controls (n = 6,085) of European Caucasian descent were collected in North America and Sweden (Supplementary Table 1). The cohort was genotyped using the Affymetrix Axiom Biobank Plus Genotyping Array at the Affymetrix facility (Santa Clara, CA). In addition to the exome array content analysed in the present

study, the chip included genome-wide and customized content analysed as part of a concurrent GWAS meta-analysis (16).

Genotype calling and quality control

Initial genotype calling and QC was performed separately for each of the four studies. Subsequently a joint QC procedure was undertaken to ensure that consistent QC standards were adhered to (Supplementary Table 2).

UK data. Genotype calling was performed separately for the three different chips using Illumina's GenomeStudio Data Analysis software (samples clustered using GenTrain 2.0 algorithm). Sample QC was performed using PLINK (v1.07) (57) and R (58), with samples excluded based on call rate (< 0.95), suspected non-European ancestry, heterozygosity (±4 s.d. from the mean), array signal intensity (> 4 s.d. from the mean) and relatedness. SNVs were excluded due to call rate (< 0.99), deviations from Hardy-Weinberg equilibrium ($p < 0.0001$) and low GenomeStudio cluster separation score (< 0.4). We also excluded duplicate assays, tri-allelic variants and insertions/deletions from further analysis. zCall software (version 3) (59) was employed to improve genotype calling for samples and SNVs that passed the initial QC. Subsequently we excluded SNVs and samples having a revised call rate below 99% to give a total of 234,976 SNVs in 2,431 cases and 5,892 controls. Genotype intensity cluster plots were manually inspected for the 5,000 SNVs found to have the lowest p-values in a preliminary association test (see below). Where appropriate, genotypes were manually "rescued" using Evoker (version 2.3) (60).

German and Estonian data. We removed samples from the German and Estonian cohorts with high missingness (> 2%). SNVs were removed if they had low call rate (< 95%) or deviated from Hardy Weinberg equilibrium ($p < 0.0001$) across both cohorts combined. Triallelic variants, insertions, deletions and one of each pair of duplicated markers were excluded. Rare variant genotypes were called using the zCall algorithm after removing samples with a call rate < 95%. zCall was employed using default settings (59) for the German and Estonian cohorts separately.

Michigan data. We removed samples with high missingness (> 2%), and markers with low call rate (< 95%) or that deviated from Hardy Weinberg equilibrium ($p < 1\times10^{-6}$). Additional QC steps

and rare variant calling using zCall were performed in Kiel as described above for the German and Estonian datasets.

Joint quality control. All four datasets were filtered to exclude variants with call rate below 99% and samples with missingness above 1%. We took forward for subsequent analysis only those SNVs that were present in all four datasets. We excluded SNVs where alleles did not match between datasets or where the minor allele was ambiguous (that is, symmetrical SNVs with MAF > 0.45 and disagreement between datasets).

A subset of 32,403 independent SNVs was generated by excluding SNVs with MAF < 0.001 and SNVs within 250 kb of a previously-published psoriasis susceptibility locus or in regions of long-range LD as defined by Price *et al.* (61); and by using PLINK to perform LD-pruning (r^2 threshold = 0.2). Relationship inference was performed jointly across all samples based on this independent subset of SNVs, using KING (version 1.4) (62). For pairs of samples found to be related (second degree relative or closer; kinship coefficient > 0.0884), the sample with fewer missing genotypes was retained and the other excluded from further analysis. For each of the four datasets separately, principal component analysis (PCA) was performed based on the SNVs within the independent subset having MAF > 0.01 in that dataset (between 16,307 and 16,629 SNVs). In order to mitigate against population stratification we excluded PCA outliers from all four datasets (defined as samples lying > 6 standard deviations away from the mean for any of the first ten principal components) (Supplementary Fig. 1).

Following an initial round of association testing (described below), genotype intensity cluster plots for all non-MHC variants with single variant association p-value < 10^{-5} or included in a gene achieving a p-value < 10^{-5} in any aggregation test were manually inspected (and if necessary, removed or manually corrected) in all four datasets using Evoker (60). All analysis was subsequently repeated using these final datasets to give the results presented in this article. Cluster plots have been checked in the final datasets for all variants and genes reported here.

Linkage disequilibrium

All LD statistics reported in this work derive from 503 samples of European ancestry from 1000 Genomes (phase 3) (63). Estimates of r^2 and D' were calculated using PLINK.

Proxy markers for established psoriasis susceptibility variants

We curated established genome-wide significant psoriasis susceptibility variants from the literature (5-17, 35, 64) (Supplementary Table 6). Where established psoriasis variants were not present in our study, we identified tested SNVs within 500 kb with which they are in LD ($r^2 > 0.8$); of these, we used the SNV in strongest LD as a proxy for the established variant.

Single marker association testing

We used a linear mixed model (LMM) implemented in EMMAX (27) to test for association of single variants in each of the four studies. In each study population structure was controlled for using a genetic relatedness matrix derived from the set of 32,403 independent SNVs described above; to avoid confounding due to LD and known psoriasis association, we also estimated genomic inflation using the p-values of association for these SNVs. Evaluation of quantile-quantile (QQ) plots indicated that inflation was minimal (Supplementary Fig. 2), with median genomic control (λ_{GC}) values ranging from 1.005 to 1.048 across the four studies. We subsequently performed standard-error weighted fixed-effect meta-analysis using METAL (current version) (65) to obtain combined p-values. Since EMMAX does not guarantee the accuracy of effect size estimates for binary traits, we estimated odds ratios (ORs) separately. For this we used PLINK (v1.9; www.cog-genomics.org/plink/1.9/) (66) to perform logistic regression for each study with the first ten principal components as covariates, and the 'meta' package (67) in R for meta-analysis. We verified that the p-values generated under this method are consistent with our primary results generated by the LMM (Supplementary Fig. 3).

Single variants were only considered significantly associated with psoriasis susceptibility if their direction of effect was consistent across all four studies and p-value of association was below the exome-wide significance threshold of 3.0×10^{-7} (corresponding to 0.05/167,587 variants tested).

Where significantly associated protein-altering variants were identified in established psoriasis susceptibility loci, we assessed the degree to which each protein-altering variant corresponds to established association signal. This was done by estimating LD between the protein-altering variant and the known associated variant, and where a suitable proxy for the known variant existed in our data, by performing (bidirectional) conditional association testing with the protein-altering variant. Conditional association p-values were generated using EMMAX and METAL as above, with the genotype of the SNV to be conditioned on included as a fixed covariate in EMMAX and with the same genetic relatedness matrices as the unconditioned analysis.

Gene-based association testing

We prepared genotype data for gene-based association testing using EPACTS (v3.2.3; http://genome.sph.umich.edu/wiki/EPACTS) to annotate variants. We used RAREMETALWORKER (v4.13.5) (68) to generate score statistics and covariance information based on individual markers in each study; population structure was controlled for using a genetic relatedness matrix derived from the set of 32,403 independent SNVs described above. We subsequently used rareMETALS2 (v0.1; http://genome.sph.umich.edu/wiki/RareMETALS2) to perform combined gene-level meta-analysis, for each gene including all variants annotated as protein-altering (nonsynonymous, stop-gain and essential splice site) and having MAF below a fixed threshold. These combined tests comprised the GRANVIL (69) (burden) test and SKAT (30) (variance component) test, using MAF thresholds of both 0.01 and 0.05. To correct for exome-wide testing we used a Bonferroni-corrected threshold of $0.05/20{,}000 = 2.5\times10^{-6}$ to classify genes as significantly associated with psoriasis susceptibility.

Since RAREMETALWORKER and rareMETALS2 also provide single marker association test results we confirmed that meta-analysis p-values and effect sizes generated in this way are consistent with our primary results obtained as described above (Supplementary Fig. 4; Supplementary Table 4).

ORs were estimated for gene-based tests by collapsing all included rare variants across each gene into a single genotype, and performing logistic regression in PLINK and meta-analysis using the R ‘meta’ package as for single marker association testing (described above). Since both genes

achieving exome-wide significance fell within established psoriasis susceptibility loci for which exome-wide significant single variants were identified by our earlier analysis, we tested for gene-level association signal that could be attributed to rare variants independently of these known single variants. This was done by repeating the gene-level association tests and conditioning on the associated single variants, using the conditional analysis function implemented in rareMETALS2 (and excluding from the set of variants to be aggregated any associated single variants with sufficiently low MAF to be otherwise included).

We further investigated genes in all established psoriasis susceptibility loci. Our data included rare or low-frequency protein-altering variants in 412 genes located within 250 kb of a previously or newly reported single variant association. We checked these genes for GRANVIL and SKAT test p-values below a threshold of $0.05/412 = 1.214\times10^{-4}$.

Variant effect

We predicted variant effects using three *in silico* tools. We consider SIFT (70) scores below 0.05, PolyPhen-2 (71) estimated false-positive rate below 0.05 and scaled CADD (72) scores above 20 to indicate a predicted functional effect. For all variants, scores for all three prediction tools were generated via wANNOVAR (73). All amino acid substitutions described refer to the canonical protein sequence as defined by UniProt (74).

Exome array coverage

We collated variants included in the original exome array design from the online documentation (http://genome.sph.umich.edu/wiki/Exome_Chip_Design). Variants subsequently included on each of the genotyping arrays used were obtained from the relevant manufacturer (Illumina or Affymetrix; Supplementary Table 1).

To estimate the coverage of protein-altering variants by the genotyping arrays we downloaded annotated ExAC variants (release 0.3.1) (75). Biallelic SNVs were extracted which included an annotation of moderate or high impact to at least one protein-coding transcript. We further filtered these variants to those with non-zero alternative allele count in 33,370 European samples based on

variant calls for at least 10,000 chromosomes. After excluding SNVs on non-autosomal chromosomes and those within the MHC region this resulted in 1,655,908 SNVs, of which 14,123 were common (MAF ≥ 0.05), 9,957 were low-frequency ($0.01 \leq$ MAF < 0.05), 32,029 were rare but not very rare ($0.001 \leq$ MAF < 0.01). The majority (1,599,799) had MAF below 0.001.

To assess coverage of potential causal variants in established non-MHC psoriasis susceptibility loci we searched for variants in 1000 Genomes European samples that are in moderate LD ($r^2 \geq 0.2$) with a previously reported association, as described above. This resulted in 17,215 SNVs in total. To identify candidate exonic variants we extracted those which included a SIFT, PolyPhen or CADD annotation predicting impaired protein function.

Acknowledgements

This collaborative study was initiated and largely funded by the International Psoriasis Council. Funding is acknowledged from: a Medical Research Council (MRC) Stratified Medicine award (MR/L011808/1; to C.E.M.G., N.J.R, R.B.W., C.H.S., R.C.T., M.A.S. and J.N.B.) and The Psoriasis Association (RG2/10). S.K.M. was funded by a MRC Clinical Training Fellowship (MR/L001543/1). C.E.M.G. is a National Institute for Health Research (NIHR) Senior Investigator. UK psoriasis cases included samples collected by the Genetic Analysis of Psoriasis Consortium and Biomarkers of Systemic Treatment Outcomes study group; we acknowledge the enthusiastic collaboration and contributions of all of the clinical and research teams at participating centers in the U.K. and the Republic of Ireland, and patients. The research was supported by the NIHR Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College London, and by the NIHR Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of the NHS, NIHR, Department of Health or King's College London. We gratefully acknowledge capital equipment funding from the Maudsley Charity (grant 980) and Guy's and St Thomas's Charity (grant STR130505). We acknowledge use of the British 1958 Birth Cohort DNA collection, funded by the MRC grant G0000934 and the Wellcome Trust grant 068545/Z/02. This work was also supported by

the German Federal Ministry of Education and Research (BMBF) within the framework of the e:Med research and funding concept (sysINFLAME; grant 01ZX1306A). The project received infrastructure support from the DFG Excellence Cluster No. 306 “Inflammation at Interfaces” and the PopGen Biobank (Kiel, Germany; grant 01EY1103). The KORA research platform (KORA, Cooperative Research in the Region of Augsburg) was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by BMBF and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ. SHIP (Study of Health in Pomerania) is part of the Community Medicine Research net (CMR) of the University of Greifswald, Germany, which is funded by BMBF (grants 01ZZ9603, 01ZZ0103 and 01ZZ0403) and the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania, and the network ‘Greifswald Approach to Individualized Medicine (GANI_MED)’ funded by BMBF (grant 03IS2061A). Exome Chip data have been supported by BMBF (grant no. 03Z1CN22) and the Federal State of Mecklenburg-West Pomerania. Parts of the study were supported by the BMBF Metarthros grant (01EC1407A to A.R. and U.H.). A.F. receives an endowment professorship by the Foundation for Experimental Medicine (Zürich, Switzerland). This work was also supported by awards from the National Institutes of Health (R01AR042742, R01AR050511, R01AR054966, R01AR063611, R01AR065183 to J.T.E.), as well as a GAIN award from the Foundation for the National Institutes of Health (to G.R.A.). J.T.E. and T.T. are supported by the Ann Arbor Veterans Affairs Hospital. J.E.G. is supported by the Taubman Medical Research Institute as the Frances and Kenneth Eisenberg Emerging Scholar.

Conflict of Interest Statement

M.A.S. has a contract of service with Genomics plc.

References

1 Nestle, F.O., Kaplan, D.H. and Barker, J. (2009) Psoriasis. *N. Engl. J. Med.*, **361**, 496-509.

2 Mahil, S.K., Capon, F. and Barker, J.N. (2015) Genetics of Psoriasis. *Dermatol. Clin.*, **33**, 1-11.
3 Rahman, P. and Elder, J.T. (2005) Genetic epidemiology of psoriasis and psoriatic arthritis. *Ann. Rheum. Dis.*, **64 Suppl 2**, ii37-39.
4 Hawro, T., Zalewska, A., Hawro, M., Kaszuba, A., Krolikowska, M. and Maurer, M. (2015) Impact of psoriasis severity on family income and quality of life. *J. Eur. Acad. Dermatol. Venereol.*, **29**, 438-443.
5 Strange, A., Capon, F., Spencer, C.C., Knight, J., Weale, M.E., Allen, M.H., Barton, A., Band, G., Bellenguez, C., Bergboer, J.G. *et al.* (2010) A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. *Nat. Genet.*, **42**, 985-990.
6 Ellinghaus, E., Ellinghaus, D., Stuart, P.E., Nair, R.P., Debrus, S., Raelson, J.V., Belouchi, M., Fournier, H., Reinhard, C., Ding, J. *et al.* (2010) Genome-wide association study identifies a psoriasis susceptibility locus at TRAF3IP2. *Nat. Genet.*, **42**, 991-995.
7 Stuart, P.E., Nair, R.P., Ellinghaus, E., Ding, J., Tejasvi, T., Gudjonsson, J.E., Li, Y., Weidinger, S., Eberlein, B., Gieger, C. *et al.* (2010) Genome-wide association analysis identifies three psoriasis susceptibility loci. *Nat. Genet.*, **42**, 1000-1004.
8 Tsoi, L.C., Spain, S.L., Knight, J., Ellinghaus, E., Stuart, P.E., Capon, F., Ding, J., Li, Y., Tejasvi, T., Gudjonsson, J.E. *et al.* (2012) Identification of 15 new psoriasis susceptibility loci highlights the role of innate immunity. *Nat. Genet.*, **44**, 1341-1348.
9 Bowes, J., Budu-Aggrey, A., Huffmeier, U., Uebe, S., Steel, K., Hebert, H.L., Wallace, C., Massey, J., Bruce, I.N., Bluett, J. *et al.* (2015) Dense genotyping of immune-related susceptibility loci reveals new insights into the genetics of psoriatic arthritis. *Nat. Commun.*, **6**, 6046.
10 Yin, X., Low, H.Q., Wang, L., Li, Y., Ellinghaus, E., Han, J., Estivill, X., Sun, L., Zuo, X., Shen, C. *et al.* (2015) Genome-wide meta-analysis identifies multiple novel associations and ethnic heterogeneity of psoriasis susceptibility. *Nat. Commun.*, **6**, 6916.
11 Tsoi, L.C., Spain, S.L., Ellinghaus, E., Stuart, P.E., Capon, F., Knight, J., Tejasvi, T., Kang, H.M., Allen, M.H., Lambert, S. *et al.* (2015) Enhanced meta-analysis and replication studies identify five new psoriasis susceptibility loci. *Nat. Commun.*, **6**, 7001.
12 Stuart, P.E., Nair, R.P., Tsoi, L.C., Tejasvi, T., Das, S., Kang, H.M., Ellinghaus, E., Chandran, V., Callis-Duffin, K., Ike, R. *et al.* (2015) Genome-wide Association Analysis of Psoriatic Arthritis and Cutaneous Psoriasis Reveals Differences in Their Genetic Architecture. *Am. J. Hum. Genet.*, **97**, 816-836.
13 Capon, F., Bijlmakers, M.J., Wolf, N., Quaranta, M., Huffmeier, U., Allen, M., Timms, K., Abkevich, V., Gutin, A., Smith, R. *et al.* (2008) Identification of ZNF313/RNF114 as a novel psoriasis susceptibility gene. *Hum. Mol. Genet.*, **17**, 1938-1945.
14 Nair, R.P., Duffin, K.C., Helms, C., Ding, J., Stuart, P.E., Goldgar, D., Gudjonsson, J.E., Li, Y., Tejasvi, T., Feng, B.J. *et al.* (2009) Genome-wide scan reveals association of psoriasis with IL-23 and NF-kappaB pathways. *Nat. Genet.*, **41**, 199-204.
15 Ellinghaus, D., Ellinghaus, E., Nair, R.P., Stuart, P.E., Esko, T., Metspalu, A., Debrus, S., Raelson, J.V., Tejasvi, T., Belouchi, M. *et al.* (2012) Combined analysis of genome-wide association studies for Crohn disease and psoriasis identifies seven shared susceptibility loci. *Am. J. Hum. Genet.*, **90**, 636-647.
16 Tsoi, L.C., Stuart, P.E., Tian, C., Gudjonsson, J.E., Das, S., Zawistowski, M., Ellinghaus, E., Barker, J.N., Chandran, V., Dand, N. *et al.* (2017) Large scale meta-analysis characterizes genetic architecture for common psoriasis associated variants. *Nat. Commun.*, **8**, 15382.
17 Das, S., Stuart, P.E., Ding, J., Tejasvi, T., Li, Y., Tsoi, L.C., Chandran, V., Fischer, J., Helms, C., Duffin, K.C. *et al.* (2015) Fine mapping of eight psoriasis susceptibility loci. *Eur. J. Hum. Genet.*, **23**, 844-853.
18 Knight, J., Spain, S.L., Capon, F., Hayday, A., Nestle, F.O., Clop, A., Barker, J.N., Weale, M.E. and Trembath, R.C. (2012) Conditional analysis identifies three novel major histocompatibility complex loci associated with psoriasis. *Hum. Mol. Genet.*, **21**, 5185-5192.
19 Capon, F. and Barker, J.N. (2012) The quest for psoriasis susceptibility genes in the postgenome-wide association studies era: charting the road ahead. *Br. J. Dermatol.*, **166**, 1173-1175.

20 Di Meglio, P., Di Cesare, A., Laggner, U., Chu, C.C., Napolitano, L., Villanova, F., Tosi, I., Capon, F., Trembath, R.C., Peris, K. *et al.* (2011) The IL23R R381Q gene variant protects against immune-mediated diseases by impairing IL-23-induced Th17 effector response in humans. *PLOS ONE*, **6**, e17160.
21 Gandhi, M., Alwawi, E. and Gordon, K.B. (2010) Anti-p40 antibodies ustekinumab and briakinumab: blockade of interleukin-12 and interleukin-23 in the treatment of psoriasis. *Semin. Cutan. Med. Surg.*, **29**, 48-52.
22 Tang, H., Jin, X., Li, Y., Jiang, H., Tang, X., Yang, X., Cheng, H., Qiu, Y., Chen, G., Mei, J. *et al.* (2014) A large-scale screen for coding variants predisposing to psoriasis. *Nat. Genet.*, **46**, 45-50.
23 Zuo, X., Sun, L., Yin, X., Gao, J., Sheng, Y., Xu, J., Zhang, J., He, C., Qiu, Y., Wen, G. *et al.* (2015) Whole-exome SNP array identifies 15 new susceptibility loci for psoriasis. *Nat. Commun.*, **6**, 6793.
24 Nair, R.P., Stuart, P.E., Nistor, I., Hiremagalore, R., Chia, N.V., Jenisch, S., Weichenthal, M., Abecasis, G.R., Lim, H.W., Christophers, E. *et al.* (2006) Sequence and haplotype analysis supports HLA-C as the psoriasis susceptibility 1 gene. *Am. J. Hum. Genet.*, **78**, 827-851.
25 Feng, B.J., Sun, L.D., Soltani-Arabshahi, R., Bowcock, A.M., Nair, R.P., Stuart, P., Elder, J.T., Schrodi, S.J., Begovich, A.B., Abecasis, G.R. *et al.* (2009) Multiple Loci within the major histocompatibility complex confer risk of psoriasis. *PLOS Genet.*, **5**, e1000606.
26 Okada, Y., Han, B., Tsoi, L.C., Stuart, P.E., Ellinghaus, E., Tejasvi, T., Chandran, V., Pellett, F., Pollock, R., Bowcock, A.M. *et al.* (2014) Fine mapping major histocompatibility complex associations in psoriasis and its clinical subtypes. *Am. J. Hum. Genet.*, **95**, 162-172.
27 Kang, H.M., Sul, J.H., Service, S.K., Zaitlen, N.A., Kong, S.Y., Freimer, N.B., Sabatti, C. and Eskin, E. (2010) Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.*, **42**, 348-354.
28 Yamazaki, K., McGovern, D., Ragoussis, J., Paolucci, M., Butler, H., Jewell, D., Cardon, L., Takazoe, M., Tanaka, T., Ichimori, T. *et al.* (2005) Single nucleotide polymorphisms in TNFSF15 confer susceptibility to Crohn's disease. *Hum. Mol. Genet.*, **14**, 3499-3506.
29 Képíró, L., Széll, M., Kovács, L., Keszthelyi, P., Kemény, L. and Gyulai, R. (2014) Genetic risk and protective factors of TNFSF15 gene variants detected using single nucleotide polymorphisms in Hungarians with psoriasis and psoriatic arthritis. *Hum. Immunol.*, **75**, 159-162.
30 Wu, M.C., Lee, S., Cai, T., Li, Y., Boehnke, M. and Lin, X. (2011) Rare-variant association testing for sequencing data with the sequence kernel association test. *Am. J. Hum. Genet.*, **89**, 82-93.
31 Li, Y., Liao, W., Cargill, M., Chang, M., Matsunami, N., Feng, B.J., Poon, A., Callis-Duffin, K.P., Catanese, J.J., Bowcock, A.M. *et al.* (2010) Carriers of rare missense variants in IFIH1 are protected from psoriasis. *J. Invest. Dermatol.*, **130**, 2768-2772.
32 Babon, J.J., Lucet, I.S., Murphy, J.M., Nicola, N.A. and Varghese, L.N. (2014) The molecular regulation of Janus kinase (JAK) activation. *Biochem. J.*, **462**, 1-13.
33 Loo, Y.M. and Gale, M., Jr. (2011) Immune signaling by RIG-I-like receptors. *Immunity*, **34**, 680-692.
34 Hüffmeier, U., Uebe, S., Ekici, A.B., Bowes, J., Giardina, E., Korendowych, E., Juneblad, K., Apel, M., McManus, R., Ho, P. *et al.* (2010) Common variants at TRAF3IP2 are associated with susceptibility to psoriatic arthritis and psoriasis. *Nat. Genet.*, **42**, 996-999.
35 de Cid, R., Riveira-Munoz, E., Zeeuwen, P.L., Robarge, J., Liao, W., Dannhauser, E.N., Giardina, E., Stuart, P.E., Nair, R., Helms, C. *et al.* (2009) Deletion of the late cornified envelope LCE3B and LCE3C genes as a susceptibility factor for psoriasis. *Nat. Genet.*, **41**, 211-215.
36 Bijlmakers, M.J., Kanneganti, S.K., Barker, J.N., Trembath, R.C. and Capon, F. (2011) Functional analysis of the RNF114 psoriasis susceptibility gene implicates innate immune responses to double-stranded RNA in disease pathogenesis. *Hum. Mol. Genet.*, **20**, 3129-3137.
37 Serpa, J., Mendes, N., Reis, C.A., Santos Silva, L.F., Almeida, R., Le Pendu, J. and David, L. (2004) Two new FUT2 (fucosyltransferase 2 gene) missense polymorphisms, 739G-->A and 839T-->C, are partly responsible for non-secretor status in a Caucasian population from Northern Portugal. *Biochem. J.*, **383**, 469-474.
38 Carlsson, B., Kindberg, E., Buesa, J., Rydell, G.E., Lidon, M.F., Montava, R., Abu Mallouh, R., Grahn, A., Rodriguez-Diaz, J., Bellido, J. *et al.* (2009) The G428A nonsense mutation in FUT2

provides strong but not absolute protection against symptomatic GII.4 Norovirus infection. *PLOS ONE*, **4**, e5593.
39 Ikehara, Y., Nishihara, S., Yasutomi, H., Kitamura, T., Matsuo, K., Shimizu, N., Inada, K., Kodera, Y., Yamamura, Y., Narimatsu, H. *et al.* (2001) Polymorphisms of two fucosyltransferase genes (Lewis and Secretor genes) involving type I Lewis antigens are associated with the presence of anti-Helicobacter pylori IgG antibody. *Cancer Epidemiol. Biomarkers Prev.*, **10**, 971-977.
40 Thom, S.M., Blackwell, C.C., MacCallum, C.J., Weir, D.M., Brettle, R.P., Kinane, D.F. and Wray, D. (1989) Non-secretion of blood group antigens and susceptibility to infection by Candida species. *FEMS Microbiol. Immunol.*, **1**, 401-405.
41 Mahil, S.K., Capon, F. and Barker, J.N. (2015) Update on psoriasis immunopathogenesis and targeted immunotherapy. *Semin.Immunopathol.*, **38**, 11-27.
42 Sohn, S.J., Barrett, K., Van Abbema, A., Chang, C., Kohli, P.B., Kanda, H., Smith, J., Lai, Y., Zhou, A., Zhang, B. *et al.* (2013) A restricted role for TYK2 catalytic activity in human cytokine responses revealed by novel TYK2-selective inhibitors. *J. Immunol.*, **191**, 2205-2216.
43 Diogo, D., Bastarache, L., Liao, K.P., Graham, R.R., Fulton, R.S., Greenberg, J.D., Eyre, S., Bowes, J., Cui, J., Lee, A. *et al.* (2015) TYK2 protein-coding variants protect against rheumatoid arthritis and autoimmunity, with no evidence of major pleiotropic effects on non-autoimmune complex traits. *PLOS ONE*, **10**, e0122271.
44 Lopez-Isac, E., Campillo-Davo, D., Bossini-Castillo, L., Guerra, S.G., Assassi, S., Simeon, C.P., Carreira, P., Ortego-Centeno, N., Garcia de la Pena, P., Beretta, L. *et al.* (2016) Influence of TYK2 in systemic sclerosis susceptibility: a new locus in the IL-12 pathway. *Ann. Rheum. Dis.*, **75**, 1521-1526.
45 Li, Z., Gakovic, M., Ragimbeau, J., Eloranta, M.L., Ronnblom, L., Michel, F. and Pellegrini, S. (2013) Two rare disease-associated Tyk2 variants are catalytically impaired but signaling competent. *J. Immunol.*, **190**, 2335-2344.
46 Prens, E.P., Kant, M., van Dijk, G., van der Wel, L.I., Mourits, S. and van der Fits, L. (2008) IFN-alpha enhances poly-IC responses in human keratinocytes by inducing expression of cytosolic innate RNA receptors: relevance for psoriasis. *J. Invest. Dermatol.*, **128**, 932-938.
47 Nejentsev, S., Walker, N., Riches, D., Egholm, M. and Todd, J.A. (2009) Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes. *Science*, **324**, 387-389.
48 Looney, B.M., Xia, C.Q., Concannon, P., Ostrov, D.A. and Clare-Salzler, M.J. (2015) Effects of type 1 diabetes-associated IFIH1 polymorphisms on MDA5 function and expression. *Curr. Diab.Rep.*, **15**, 96.
49 So, H.C., Gui, A.H., Cherny, S.S. and Sham, P.C. (2011) Evaluating the heritability explained by known susceptibility variants: a survey of ten complex diseases. *Genet. Epidemiol.*, **35**, 310-317.
50 Grjibovski, A.M., Olsen, A.O., Magnus, P. and Harris, J.R. (2007) Psoriasis in Norwegian twins: contribution of genetic and environmental effects. *J. Eur. Acad. Dermatol. Venereol.*, **21**, 1337-1343.
51 Afshar, M., Martinez, A.D., Gallo, R.L. and Hata, T.R. (2013) Induction and exacerbation of psoriasis with Interferon-alpha therapy for hepatitis C: a review and analysis of 36 cases. *J. Eur. Acad. Dermatol. Venereol.*, **27**, 771-778.
52 Funk, J., Langeland, T., Schrumpf, E. and Hanssen, L.E. (1991) Psoriasis induced by interferon-alpha. *Br. J. Dermatol.*, **125**, 463-465.
53 Papp, K., Pariser, D., Catlin, M., Wierz, G., Ball, G., Akinlade, B., Zeiher, B. and Krueger, J.G. (2015) A phase 2a randomized, double-blind, placebo-controlled, sequential dose-escalation study to evaluate the efficacy and safety of ASP015K, a novel Janus kinase inhibitor, in patients with moderate-to-severe psoriasis. *Br. J. Dermatol.*, **173**, 767-776.
54 Papp, K.A., Krueger, J.G., Feldman, S.R., Langley, R.G., Thaci, D., Torii, H., Tyring, S., Wolk, R., Gardner, A., Mebus, C. *et al.* (2016) Tofacitinib, an oral Janus kinase inhibitor, for the treatment of chronic plaque psoriasis: Long-term efficacy and safety results from 2 randomized phase-III studies and 1 open-label long-term extension study. *J. Am. Acad. Dermatol.*, **74**, 841-850.
55 Bissonnette, R., Luchi, M., Fidelus-Gort, R., Jackson, S., Zhang, H., Flores, R., Newton, R., Scherle, P., Yeleswaram, S., Chen, X. *et al.* (2016) A randomized, double-blind, placebo-controlled, dose-escalation study of the safety and efficacy of INCB039110, an oral janus kinase 1 inhibitor, in patients with stable, chronic plaque psoriasis. *J. Dermatolog. Treat.*, **27**, 332-338.

56 Volzke, H., Alte, D., Schmidt, C.O., Radke, D., Lorbeer, R., Friedrich, N., Aumann, N., Lau, K., Piontek, M., Born, G. *et al.* (2011) Cohort profile: the study of health in Pomerania. *Int. J. Epidemiol.*, **40**, 294-307.
57 Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*, **81**, 559-575.
58 R Core Team (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
59 Goldstein, J.I., Crenshaw, A., Carey, J., Grant, G.B., Maguire, J., Fromer, M., O'Dushlaine, C., Moran, J.L., Chambert, K., Stevens, C. *et al.* (2012) zCall: a rare variant caller for array-based genotyping: Genetics and population analysis. *Bioinformatics*, **28**, 2543-2545.
60 Morris, J.A., Randall, J.C., Maller, J.B. and Barrett, J.C. (2010) Evoker: a visualization tool for genotype intensity data. *Bioinformatics*, **26**, 1786-1787.
61 Price, A.L., Weale, M.E., Patterson, N., Myers, S.R., Need, A.C., Shianna, K.V., Ge, D., Rotter, J.I., Torres, E., Taylor, K.D. *et al.* (2008) Long-range LD can confound genome scans in admixed populations. *Am. J. Hum. Genet.*, **83**, 132-135; author reply 135-139.
62 Manichaikul, A., Mychaleckyj, J.C., Rich, S.S., Daly, K., Sale, M. and Chen, W.M. (2010) Robust relationship inference in genome-wide association studies. *Bioinformatics*, **26**, 2867-2873.
63 1000 Genomes Project Consortium. (2015) A global reference for human genetic variation. *Nature*, **526**, 68-74.
64 Cargill, M., Schrodi, S.J., Chang, M., Garcia, V.E., Brandon, R., Callis, K.P., Matsunami, N., Ardlie, K.G., Civello, D., Catanese, J.J. *et al.* (2007) A large-scale genetic association study confirms IL12B and leads to the identification of IL23R as psoriasis-risk genes. *Am. J. Hum. Genet.*, **80**, 273-290.
65 Willer, C.J., Li, Y. and Abecasis, G.R. (2010) METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*, **26**, 2190-2191.
66 Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M. and Lee, J.J. (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*, **4**, 7.
67 Schwarzer, G. (2007) meta: An R package for meta-analysis. *R News*, **7**, 40-45.
68 Feng, S., Liu, D., Zhan, X., Wing, M.K. and Abecasis, G.R. (2014) RAREMETAL: fast and powerful meta-analysis for rare variants. *Bioinformatics*, **30**, 2828-2829.
69 Magi, R., Kumar, A. and Morris, A.P. (2011) Assessing the impact of missing genotype data in rare variant association analysis. *BMC Proc.*, **5 Suppl 9**, S107.
70 Kumar, P., Henikoff, S. and Ng, P.C. (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc*, **4**, 1073-1081.
71 Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S. and Sunyaev, S.R. (2010) A method and server for predicting damaging missense mutations. *Nat Methods*, **7**, 248-249.
72 Kircher, M., Witten, D.M., Jain, P., O'Roak, B.J., Cooper, G.M. and Shendure, J. (2014) A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.*, **46**, 310-315.
73 Chang, X. and Wang, K. (2012) wANNOVAR: annotating genetic variants for personal genomes via the web. *J. Med. Genet.*, **49**, 433-436.
74 The UniProt Consortium. (2015) UniProt: a hub for protein information. *Nucleic Acids Res*, **43**, D204-D212.
75 Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B. *et al.* (2016) Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, **536**, 285-291.

Legends to Figures

Figure 1 – Rare and low-frequency protein-altering variants in IFIH1 and TYK2

Frequency of alternative allele in cases and controls across all four studies for rare and low-frequency variants, displayed by protein consequence. ** designates exome-wide significant association ($p < 3.0\times10^{-7}$); * designates nominally significant association ($p < 0.05$). Common protein-altering variants that we report to be associated are marked by red triangles. Variant effect predictions (by SIFT, PolyPhen-2 and CADD) are red where a substitution is predicted to be damaging, white where it is not, and grey where no prediction was possible. SNV = single nucleotide variant; ESS = essential splice site; AA pos = amino acid position; PP = PolyPhen-2. (**a**) *IFIH1* variants (MDA5 protein): CARD = caspase activation recruitment domain; Hel = helicase domain. (**b**) *TYK2* variants: FERM = 4.1/ezrin/radixin/moesin domain; SH2-like = Src homology 2-like domain; JH2 = JAK-homology 2; Prot. kin. JH1 = protein kinase JAK-homology 1.

Tables

Table 1 – Exome-wide significant protein-altering variants

AAF = alternative allele frequency; OR = estimated odds ratio; variant effect prediction: S = predicted "Damaging" by SIFT; P = predicted "probably damaging" by PolyPhen-2; C = CADD score > 20.

Locus	Protein-altering SNV	hg19 position	Ref/Alt	Gene	Consequence	Variant effect prediction	AAF_{case}	AAF_{cont}	OR (95% CI)	Meta-analysis p-value	Corresponding established signal
1p31.3	rs11209026	67705958	G/A	*IL23R*	R381Q	P, C	0.046	0.062	0.719 (0.667 - 0.774)	2.00×10^{-18}	rs9988642
1q21.3	rs1332500	152692074	G/C	*C1orf68*	S26T		0.376	0.348	1.123 (1.086 - 1.161)	1.07×10^{-12}	rs6677595
	rs873775	152692472	A/C	*C1orf68*	T159P		0.376	0.348	1.123 (1.086 - 1.161)	1.11×10^{-12}	rs6677595
2q24.2	rs1990760	163124051	C/T	*IFIH1*	A946T		0.647	0.618	1.161 (1.122 - 1.200)	4.73×10^{-18}	rs1990760
	rs35667974	163124637	T/C	*IFIH1*	I923V	P	0.010	0.020	0.548 (0.473 - 0.634)	1.10×10^{-15}	rs1990760
5q15	rs27044	96118852	G/C	*ERAP1*	Q730E		0.698	0.725	0.869 (0.839 - 0.900)	1.28×10^{-13}	rs27432
	rs30187	96124330	T/C	*ERAP1*	K528R	P, C	0.633	0.663	0.886 (0.857 - 0.916)	2.19×10^{-11}	rs27432
5q31.1	rs20541	131995964	A/G	*IL13*	Q144R		0.814	0.783	1.170 (1.123 - 1.219)	3.59×10^{-13}	rs20541
6q21	rs33980500	111913262	C/T	*TRAF3IP2*	D19N	S, P, C	0.108	0.075	1.451 (1.374 - 1.533)	1.92×10^{-39}	rs33980500
	rs13190932	111913070	G/A	*TRAF3IP2*	R83W	S	0.084	0.060	1.404 (1.320 - 1.492)	1.34×10^{-28}	rs33980500
	rs458017	111696091	T/C	*REV3L*	Y1156C		0.085	0.064	1.313 (1.236 - 1.395)	1.37×10^{-19}	rs33980500
12q13.3	rs2066807	56740682	C/G	*STAT2*	M594I		0.053	0.073	0.729 (0.681 - 0.781)	1.56×10^{-17}	rs2066808
16p11.2	rs9938550	30999142	A/G	*HSD3B7*	T250A		0.599	0.626	0.882 (0.853 - 0.911)	4.74×10^{-13}	rs10782001

19p13.2	rs34536443	10463118	G/C	*TYK2*	P1104A	S, P, C	0.023	0.044	0.506 (0.458 - 0.558)	1.72×10^{-42}	rs34536443
	rs2304256	10475652	C/A	*TYK2*	V362F		0.243	0.280	0.814 (0.784 - 0.844)	2.88×10^{-23}	rs34536443 and rs12720356
	rs12720356	10469975	A/C	*TYK2*	I684S	S, P	0.068	0.083	0.763 (0.716 - 0.812)	1.39×10^{-16}	rs12720356
	rs1051738	10577843	C/A	*PDE4A*	A736E		0.166	0.184	0.879 (0.843 - 0.918)	2.02×10^{-7}	rs34536443
19q13.33	rs602662	49206985	G/A	*FUT2*	G258S	P	0.511	0.467	1.090 (1.056 - 1.126)	3.29×10^{-8}	rs281379
20q13.13	rs4647958	48600631	T/C	*SNAI1*	V118A		0.141	0.125	1.161 (1.107 - 1.217)	9.21×10^{-10}	rs495337

Table 2 – Exome-wide significant gene-based associations

MAF = minor allele frequency; OR = odds ratio estimated by collapsing test; n_{SNVs} = number of SNVs included in test (this may vary between unconditioned and conditional analysis since SNVs to be conditioned on which are sufficiently rare are not included in the test statistic for the conditional test); cMAF = cumulative minor allele frequency of SNVs included in test. Exome-wide significant p-values ($p_{gene} < 2.5\times10^{-6}$) are indicated in bold.

Locus	Gene	SNVs conditioned on	MAF < 0.01					MAF < 0.05				
			p_{burden}	OR (95% CI)	p_{SKAT}	n_{SNVs}	cMAF	p_{burden}	OR (95% CI)	p_{SKAT}	n_{SNVs}	cMAF
Unconditioned analysis												
2q24.2	*IFIH1*	-	**2.53×10^{-7}**	0.707 (0.626 - 0.799)	6.02×10^{-5}	24	0.0261	**1.84×10^{-19}**	0.620 (0.564 - 0.682)	**1.19×10^{-20}**	25	0.0461
19p13.2	*TYK2*	-	6.17×10^{-4}	0.744 (0.626 - 0.885)	2.82×10^{-4}	17	0.0115	**1.47×10^{-39}**	0.593 (0.549 - 0.641)	**6.34×10^{-41}**	19	0.0675
Conditional analysis												
2q24.2	*IFIH1*	rs35667974 and rs1990760	**1.36×10^{-8}**	0.687 (0.607 - 0.776)	4.46×10^{-6}	24	0.0261	**1.36×10^{-8}**	0.687 (0.607 - 0.776)	4.46×10^{-6}	24	0.0261
19p13.2	*TYK2*	rs34536443, rs2304256 and rs12720356	1.45×10^{-4}	0.728 (0.611 - 0.868)	7.21×10^{-5}	17	0.0115	1.85×10^{-5}	0.790 (0.701 - 0.890)	6.53×10^{-5}	18	0.0239

Abbreviations

IFN – Interferon

LD – Linkage disequilibrium

MAF – Minor allele frequency

MHC - Major histocompatibility complex

PCA – Principal component analysis

OR – Odds ratio

QC – Quality control

SNV – Single nucleotide variant

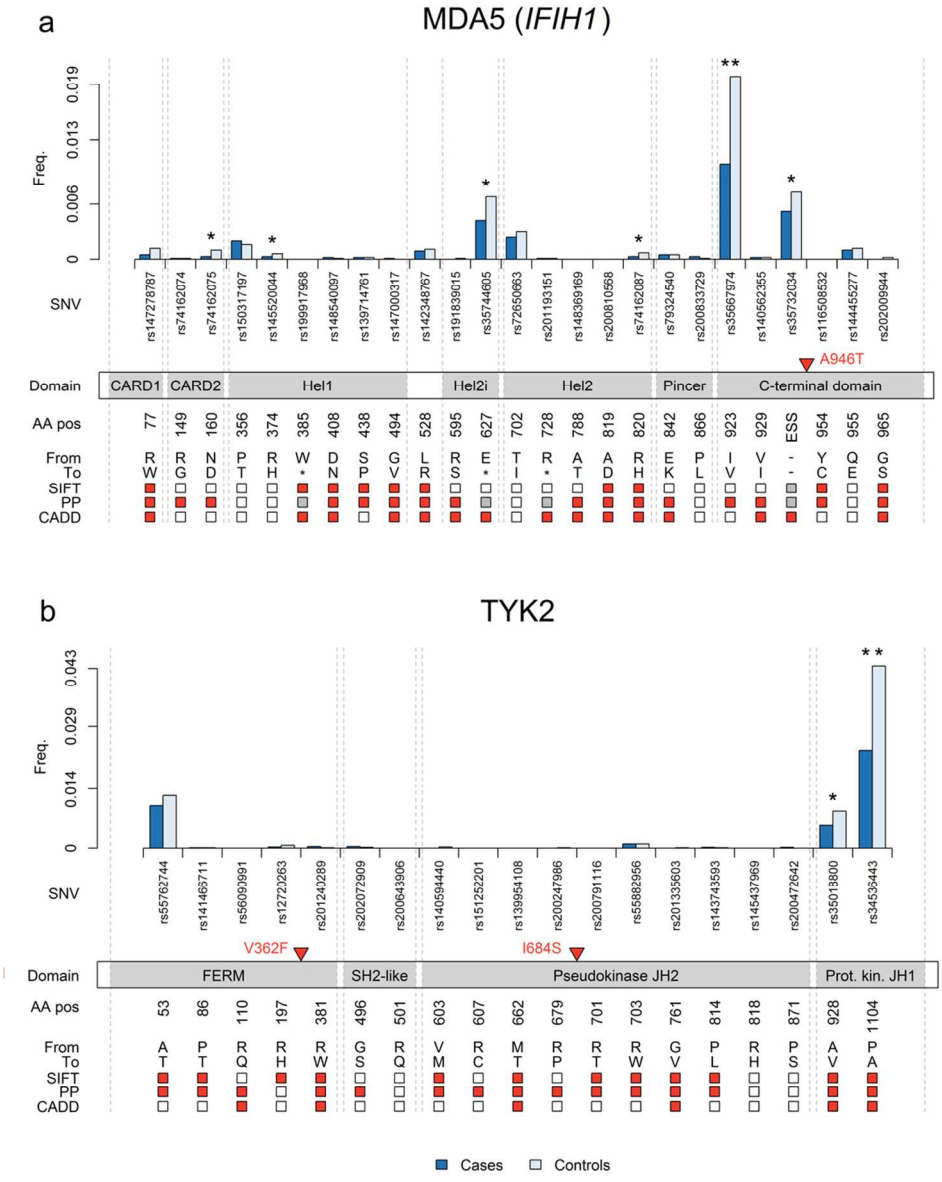

Figure 1 – Rare and low-frequency protein-altering variants in IFIH1 and TYK2

Frequency of alternative allele in cases and controls across all four studies for rare and low-frequency variants, displayed by protein consequence. ** designates exome-wide significant association ($p < 2.9\times10^{-7}$); * designates nominally significant association ($p < 0.05$). Common protein-altering variants that we report to be associated are marked by red triangles. Variant effect predictions (by SIFT, PolyPhen-2 and CADD) are red where a substitution is predicted to be damaging, white where it is not, and grey where no prediction was possible. SNV = single nucleotide variant; ESS = essential splice site; AA pos = amino acid position; PP = PolyPhen-2. (**a**) *IFIH1* variants (MDA5 protein): CARD = caspase activation recruitment domain; Hel = helicase domain. (**b**) *TYK2* variants: FERM = 4.1/ezrin/radixin/moesin domain; SH2-like = Src homology 2-like domain; JH2 = JAK-homology 2; Prot. kin. JH1 = protein kinase JAK-homology 1.

104x126mm (300 x 300 DPI)