The Functional Consequences of Autoimmune Variants in the *Tyrosine Kinase 2* Gene Region

Lydia Lambert, University College, University of Oxford

Thesis submitted for Doctor of Philosophy in Clinical Neurology, Trinity Term 2013

Abstract

The *tyrosine kinase 2* (*TYK2*) gene was first implicated in autoimmune disease in 2009 when a nonsynonymous single nucleotide polymorphism (nsSNP) in *TYK2* was reported to be associated with multiple sclerosis (MS). The immunological function of TYK2, as a kinase involved in signal transduction downstream of numerous different cytokine receptors, further strengthened the candidacy of the gene as an MS-relevant risk factor. More recently, this nsSNP has been associated with several other autoimmune conditions. In addition, another three different SNPs in the region have been found to be associated with a number of autoimmune diseases, sometimes in opposing directions. Considering the complex genetic association pattern that is emerging for the *TYK2* region across autoimmune conditions, it was hypothesised that this complexity reflects shared but also distinct pathogenic mechanisms, with the consequences of disease-associated SNPs being unlikely to all be restricted to genotype-dependent effects influencing TYK2. Therefore, the main aim of the work presented in this thesis has been to address this hypothesis by investigating the functional consequences of the disease-associated SNPs in the *TYK2* gene region.

Using an *in vitro* cell line system and primary human immune cells, obtained from a genotype-selectable donor cohort, protection at the MS-associated nsSNP was found to correlate with reduced TYK2-mediated signalling downstream of the type I interferon receptor. However, other cytokine signalling pathways were not affected, indicating a greater specificity to the function of TYK2 than has previously been appreciated. For the other SNPs in the region, substantial effects on TYK2 were not observed but immune cell subset-specific correlations with RNA-level expression of other genes in the region were identified. Thus, this is the first study to support the concept that a careful cross-comparative analysis of SNP association patterns in a single genomic region across multiple autoimmune diseases can have significant implications for enabling the delineation of pathways common or specific to different conditions, and this is of particular importance for drug repositioning strategies.

Table of contents

Total word count: approx. 47,000 words

Acknowledgements

Firstly, I would like to thank the Christopher Welch Trust for my scholarship, as my D.Phil study would not have been possible without their financial support. I also owe equally great thanks to my supervisor, Professor Lars Fugger, for giving me the opportunity to join his group and for his supervision throughout my D.Phil study.

I am immensely grateful for the supervision, support and friendship of Dr Calliope Dendrou, who deserves a special mention as she has taught me so much. I also thank all past and present members of the Fugger research group for creating a fun but productive working environment, and especially thank Dr Kate McCormick and Dr Adam Gregory for their guidance and friendship.

I would like to acknowledge our collaborators in the research team at the Oxford BioBank, Dr Matt Neville and Research Nurse Jane Cheeseman, and I also thank all of the volunteers that donated blood samples for use in my studies.

Last but by no means least, I thank my parents, Marcos and Rosemary, and my husband-to-be, Chris, for their support and patience.

Abbreviations

+pop/unstim	fold change in phospho-STAT MFI in the positive population relative to unstimulated cells
%Max	percentage of maximum: a normalised value accounting for the exact number of events captured for each sample by flow cytometry
%pSTAT+	percentage of cells gated as phospho-STAT-positive in stimulated cells relative to unstimulated cells
ACPA	anti-citrullinated protein antibody
AIRE	autoimmune regulator
AITD	autoimmune thyroid disease
ATP	adenosine triphosphate
APC	antigen-presenting cell (in the context of immune cells)
APC	allophycocyanin (in the context of flow cytometry antibody fluorochromes)
APS	ammonium persulphate
AS	ankylosing spondylitis
BBB	blood-brain barrier
BCR	B cell receptor
BSA	bovine serum albumin
CDC37	cell division cycle 37 protein
CD*x*	cluster of differentiation (where *x* is a number)
CDCV	common disease-common variant
cDNA	complementary deoxyribonucleic acid
CeD	celiac disease
CFA	complete Freund's adjuvant
CI	confidence interval
CNS	central nervous system
CO_2	carbon dioxide
CrD	Crohn's disease
CSF	cerebrospinal fluid
C_t	threshold cycle
DC	dendritic cell
dC_t	difference between the expression of the gene of interest and that of the house-keeping gene
DMEM	Dulbecco's Modified Eagle's Medium
DNA	deoxyribonucleic acid

EAE	experimental autoimmune encephalomyelitis
EBV	Epstein-Barr virus
EDTA	ethylenediaminetetraacetic acid
eQTL	expression quantitative trait locus
FAM	6-carboxyfluorescein
FCS	fetal calf serum
FDX1L	ferrodoxin-1 like protein
FERM	4.1 ezrin, radixin, moesin (domain of Janus kinases)
FIA	TYK2 protein with a phenylalanine residue at position 362, an isoleucine at 684 and an alanine at 1104
FIP	TYK2 protein with a phenylalanine residue at position 362, an isoleucine at 684 and a proline at 1104
FITC	fluorescein isothiocyanate
FoxP3	forkhead box P3
GA	glatiramer acetate
GFP	green fluorescent protein
GWAS	genome-wide association study
Het	individual heterozygous for the alleles at a single nucleotide polymorphism
HIES	hyper-immunoglobulin E syndrome
HLA	human leukocyte antigen
IBD	inflammatory bowel disease
ICAM	intercellular adhesion molecule
IFN-	interferon-
IFNAR	interferon (α, β, Ω) receptor subunit
Ig	immunoglobulin
IL-*x*	interleukin-*x* (where *x* is a number)
IL-*x*R	interleukin-*x* receptor (where *x* is a number)
IPEX	immunodysregulation, polyendocrinopathy, enteropathy, and X-linked syndrome
JAK	Janus kinase (human)
Jak	Janus kinase (mouse)
Jakininbs	Janus kinase inhibitors
JAKMIP-1	Janus kinase and microtubules interacting protein 1
JH	Janus kinase homology domain
JIA	juvenile idiopathic arthritis
kDa	kilo-dalton
KEAP1	kelch-like associated protein 1

KO	knockout
LCL	lymphoblastoid cell line
LD	linkage disequilibrium
LIF	leukemia inhibitory factor
MAF	minor allele frequency
MBP	myelin basic protien
mDC	myeloid dendritic cell
MEF	molecules of equivalent fluorochrome
MFI	mean fluorescence intensity
MGB	minor groove binder
MHC	Major Histocompatibility Complex
MinA	minor allele
MOG	myelin oligodendrocyte glycoprotein
MPD	myeloproliferative disease
MRI	magnetic resonance imaging
mRNA	messenger ribonucleic acid
MS	multiple sclerosis
n	number of individuals/replicates per group
NFκB	nuclear factor kappa B
NK	natural killer
ns	not significant
nsSNP	nonsynonymous single nucleotide polymorphism
OR	odds ratio
PBC	primary biliary cirrhosis
PBMCs	peripheral blood mononuclear cells
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
pDC	plasmacytoid dendritic cell
PDE4A	cAMP-dependent phosphodiesterase 4A
PE	phycoerythrin
PerCP-Cy5.5	peridinin-chlorophyll-protein complex-cyanine 5.5
PLP	proteolipid protein
pers. comm.	personal communication
PML	progressive multifocal leukoencephalopathy
PPMS	primary progressive multiple sclerosis
Prot	individual homozygous for the protective allele of a single nucleotide polymorphism

PS	psoriasis
pSTAT	phosphorylated signal transducer and activator of transcription molecule
pTYK2	phosphorylated tyrosine kinase 2
pY	phospho-tyrosine
RA	rheumatoid arthritis
RAVER1	ribonucleoprotein, PTB-binding protein 1
Risk	individual homozygous for the risk allele of a single nucleotide polymorphism
RNA	ribonucleic acid
rpm	rotations per minute
RPMI	Roswell Park Memorial Institute medium
RRMS	relapsing-remitting multiple sclerosis
rs118..	the rs11879191 SNP
rs127..	the rs12720356 SNP
rs230..	the rs2304256 SNP
rs345..	the rs34536443 SNP
RT-PCR	real-time quantitative polymerase chain reaction
S1PR5	sphingosine-1-phosphate receptor 5
SCID	severe combined immunodeficiency
SDS	sodium dodecyl sulphate
SEM	standard error of the mean
SH2	Src homology 2
SLE	systemic lupus erythematosus
SNP	single nucleotide polymorphism
SOCS	suppressor of cytokine signalling
SPMS	secondary progressive multiple sclerosis
STAT	signal transducer and activator of transcription
stim	stimulated
stim/unstim	fold change in phospho-STAT mean fluorescence intensity in stimulated relative to unstimulated cells
T1D	type 1 diabetes
TAMRA	tetramethylrhodamine
TEMED	tetramethylethylenediamine
TCR	T cell receptor
Tfh	follicular helper T cell
Th	helper T cell
T_m	melting temperature
TNF-α	tumour necrosis factor-α

TNFRSF1A	tumour necrosis factor superfamily member 1A
Tr1	type 1 regulatory T cell
Treg	regulatory T cell (CD25hi CD4+)
TYK2	human tyrosine kinase 2
Tyk2	mouse tyrosine kinase 2
UBC	ubiquitin C
UC	ulcerative colitis
unstim	unstimulated
UV	ultra-violet
VIA	TYK2 protein with a valine residue at position 362, an isoleucine at 684 and an alanine at 1104
VIP	TYK2 protein with a valine residue at position 362, an isoleucine at 684 and a proline at 1104
VO	vector only
VSP	TYK2 protein with a valine residue at position 362, a serine at 684 and a proline at 1104
WIMM	Weatherall Institute of molecular medicine
WTCCC	Wellcome Trust Case Control Consortium
WTCHG	Wellcome Trust Centre for Human Genetics, Oxford
*x*h	hours (where *x* is a number)
ZGLP1	zinc finger GATA-like protein 1

1. Introduction

1.1 Multiple sclerosis: clinical background

Multiple sclerosis (MS) is a common and severe disease of the central nervous system (CNS) characterised by demyelination, neurodegeneration and autoimmune inflammation. Although neurodegeneration has been identified as the correlate of clinical disability in MS (Ferguson et al., 1997) the precise events initiating disease remain to be elucidated. MS has a prevalence of approximately 1 in 1,000 individuals in Northern Europeans (Oksenberg and Baranzini, 2010) and affects about 2.5 million individuals worldwide; it is the most common cause of neurological disability in young adults in western society (Compston, 2005). The age of disease onset ranges from 20-40 years and MS more commonly affects females than males (Compston, 2005). It is estimated that MS shortens life expectancy by up to ten years and death occurs due to disease-associated complications in two thirds of patients (Compston and Coles, 2008).

Disease presentation is clinically heterogeneous with symptoms reflecting the site of CNS damage and can include sensory, motor and/or cognitive deficits. MS typically begins as a relapse-remitting disease (RRMS) that later develops into a progressive unrelenting condition (secondary-progressive MS; SPMS) in ~65% patients. A fifth of patients present with a progressive disease from the outset (PPMS) (Compston and Coles, 2008). In RRMS, the initial stages of disease include relapses, characterised by CNS inflammation, demyelination and variable axonal damage, with periods of complete recovery (remission). However over time, recovery is incomplete and extensive, and chronic neurodegeneration ensues. The exact triggers of initial disease and also disease progression are unknown. At present, the later stages of the disease remain refractory to treatment, but there are several therapeutic options for newly diagnosed patients presenting with RRMS.

1.1.1. Current therapies

Two decades ago there were no licensed treatments for MS, however there are now several approved disease-modifying therapies, and also more than 20 drugs at different stages of investigation in clinical trials with several in late-phase development (Haghikia et al., 2013). Interferon-β (IFN-β) is currently the first line treatment for RRMS and was originally used due to early views of MS having a viral aetiology. It is now believed that the therapeutic benefit of IFN-β is mediated by multiple, complex mechanisms (Dhib-Jalbut and Marks, 2010). Although IFN-β reduces relapse rates by ~35% and can reduce disease severity, around a third of patients are unresponsive to treatment (Compston and Coles, 2008). With a similar efficacy and proportion of non-responders, glatiramer acetate (GA) is a mixture of synthetic peptides of four amino acids and is believed to induce tolerance or anergy of myelin-reactive T cells (Schmied et al., 2003). More recently approved therapies include mitoxantrone, an inhibitor of DNA synthesis and repair, and natalizumab, a humanized antibody against the α4β1 integrin expressed on lymphocytes (Rice et al., 2005). The first oral therapy for MS, fingolimod, was approved in 2010 and modulates sphigosine-1-phosphate signalling, thus affecting lymphocyte migration from lymph nodes into the CNS (Pelletier and Hafler, 2012). Fingolimod, natalizumab and mitoxantrone have a greater efficacy compared to IFN-β and GA but are associated with severe side effects and so their use is restricted to more advanced or progressive MS (Compston and Coles, 2008). For example, natalizumab treatment is correlated with an increased risk of developing progressive multifocal leukoencephalopathy (PML), a typically fatal opportunistic viral infection (Rudick et al., 2006). More recently another oral therapy, teriflunomide, has been approved in the United States for treating MS, however the potential teratogenicity of this drug offsets its superior efficacy and may similarly limit its use (Oh and O'Connor, 2013).

Following promising results in phase III trials, two new oral therapies, BG12 and laquinimod were approved for MS in late 2012 (Methner and Zipp, 2013). In trials involving RRMS patients, these drugs both achieved significant reductions in annual relapse rate and improvements in magnetic resonance imaging (MRI) parameters, with only mild side effects (Comi et al., 2012; Fox et al., 2012; Gold et al., 2012). In a recent review, Methner and Zipp are optimistic that BG12 and laquinimod may offer therapeutic options far more effective than those previously available, as available data

suggests that they may have both immunomodulatory and also neuroprotective effects. Alemtuzumab may also achieve such dual effects and is currently awaiting approval for MS following phase III trial data reporting the drug to be superior to IFN-β, achieving a reduction in relapse rates and also reducing the accumulation of disability (Cohen et al., 2012; Coles et al., 2012).

Most of the currently available immunomodulatory therapies do not prevent the accumulation of irreversible CNS damage and disease progression. Similarly there are few options for treatment of PPMS or SPMS, highlighting the need for neuroprotective treatments. Amiloride has been demonstrated to reduce axonal damage and ameliorate disease in experimental autoimmune encephalomyelitis (EAE), a mouse model of MS (Friese et al., 2007; Vergo et al., 2011). A pilot study of amiloride treatment in PPMS reports significant reductions in MRI measures of neurodegeneration and these promising results are providing the rationale for further clinical trials (Arun et al., 2013). Amiloride is an oral therapy with an established safety profile having long been in clinical use for hypertension, and repositioning of this drug to the treatment of MS represents a cost-effective approach. Many novel neuroprotective agents are also being developed for the treatment of MS (Luessi et al., 2012).

In summary, the range of therapeutic options in MS is undoubtedly improving, however there still an unmet need for novel safe, cost-effective and efficacious treatments. New therapeutic strategies will need to include targeting of both the immune dysregulation and the neurodegenerative processes underpinning MS, potentially through combinatorial approaches. A major hindrance in addressing this unmet therapeutic demand in MS is that the immunopathogenesis and aetiology of the disease are not fully understood.

1.2 Immune tolerance and MS as an autoimmune disease

The immune system consists of a non-specific innate arm designed to rapidly respond to foreign insults, and a complementary adaptive arm constituting a repertoire of B and T cells with the capacity to recognise a hugely diverse range of specific antigens. The B and T cell antigen receptors (BCR and TCR, respectively) are generated by the random joining of different gene segments and given the stochastic nature of this process, there is an inherent propensity for the production of cells able to recognise not only foreign but also self antigens. In order to minimise the potential for these self-reactive cells to cause autoimmunity, the immune repertoire is shaped by central tolerance mechanisms to reduce the selection and development of autoreactive cells. The repertoire is then monitored throughout life via peripheral tolerance mechanisms to ensure immune homeostasis.

During development, positive and negative selection in the thymus ensures that T cells are able to recognise host major histocompatibility complex (MHC) molecules but do not express TCRs that react strongly to self-peptides. Autoreactive T cells are induced to undergo apoptosis or may differentiate into regulatory T cells (Tregs) (Hogquist et al., 2005). Central tolerance induction is induced by ectopic expression of tissue-specific antigens in the thymus under the control of the transcription factor autoimmune regulator (AIRE), and mutations in the *AIRE* gene have been identified as a cause of multi-organ autoimmunity in both mice and humans (Anderson et al., 2002; Peterson et al., 1998). The CNS is not a target of autoimmunity resulting from AIRE deficiency, however CNS antigens such as myelin basic protein (MBP) and myelin oligodendrocyte glycoprotein (MOG) are expressed in the thymus (Cassan and Liblau, 2007). The process of central tolerance refines the immune repertoire but does not guarantee total elimination of self-reactive T cells; indeed CNS autoantigen-specific T cells are found in peripheral blood from healthy controls as well as MS patients (Kerlero de Rosbo et al., 1993; Meinl et al., 1993). In addition some myelin epitopes are not expressed during development, for example only a splice variant of the myelin-specific proteolipid protein (PLP) is expressed in the thymus, meaning that the full length PLP expressed in the CNS is not tolerised against (Anderson et al., 2000; Klein et al., 2000). As central tolerance is incomplete, peripheral tolerance mechanisms are required to prevent any of these potentially pathogenic escapees

from responding inappropriately. These include T cell anergy, clonal ignorance, activation-induced cell death and Treg-mediated suppression (Mueller, 2010).

In MS, it is believed that T cells specific for CNS autoantigens manage to evade tolerance mechanisms and are triggered to become pathogenically active in the periphery. Despite being found in blood from both MS patients and healthy controls, myelin-specific autoreactive T cells are more activated in MS patients (Lovett-Racke et al., 1998). The precise mechanisms triggering activation are unknown, however there are several strong hypotheses mainly focusing on the role of microbial infection. Molecular mimicry describes the cross-reactivity of TCRs for pathogen and self-peptides proposing that T cell activation by the former may trigger inappropriate responses upon encountering the respective mimicked self-peptide. For example, a TCR cloned from an MS patient has been shown to recognise both MBP and also a peptide from Epstein-Barr virus (EBV) (Lang et al., 2002). Furthermore, a cross-reactive TCR recognising both a microbial peptide common to many bacteria and also a peptide from MBP is reported to induce MS-like disease in transgenic mice (Harkiolaki et al., 2009).

Bystander activation is an alternative and complementary explanation for the triggering of autoimmune disease, whereby a pro-inflammatory response leads to non-specific activation of self-reactive T cells. For example, there is evidence suggesting that B cells derived from MS patients display aberrant pro-inflammatory cytokine responses that mediate bystander activation of autoreactive T cells (Bar-Or et al., 2010).

1.3 The immunopathogenesis of MS

Immune cell infiltration into the CNS, demyelination and neurodegeneration are the pathological hallmarks of MS. The development of these inflammatory lesions is thought to occur when autoreactive CD4+ T cells specific for CNS antigens become activated in secondary lymphoid organs and cross the blood-brain barrier (BBB), infiltrating the CNS where they are re-activated by local antigen presentation. These infiltrating CD4+ T cells were believed to play a primary role in

initiating the cascades leading to CNS tissue damage, however more recent developments in our understanding of MS immunopathogenesis implicate a range of other T cell subsets, as well as humoral immunity and innate cells, including CNS-resident microglia. Not only are these cells present in MS lesions but there is also evidence that they have distinct contributions to disease pathogenesis.

1.3.1. CD4+ T cells

The strong human leukocyte antigen (HLA) class II genetic association with MS (discussed in more detail in section 1.5.1) is consistent with a prominent role for CD4+ T cells, which recognise peptides presented by this class of MHC molecules. Upon activation and depending on the prevailing cytokine milieu, CD4+ T cells differentiate into effectors of the T helper (Th) lineages such as Th1, Th2, Th17, or the more recently discovered T follicular helper (Tfh) lineage, or into suppressive Tregs.

MS was classically thought to result from skewing towards a Th1 (and away from Th2) cell phenotype, with elevated levels of the Th1 signature cytokine, IFN-γ, being reported in blood and cerebrospinal fluid (CSF) from MS patients (Olsson et al., 1990). Consistent with this, IFN-γ treatment exacerbates MS (Panitch et al., 1987) and adoptive transfer of autoreactive Th1 cells is sufficient to induce MS-like disease in mice (Stromnes and Goverman, 2006). The more recently described Th17 lineage is also involved in the pathogenesis of MS. Th17 cells are characterised by their expression of interleukin-17 (IL-17), which is highly expressed in active MS lesions (Tzartos et al., 2008) and has been shown to increase BBB permeability (Kebir et al., 2007). Additionally, a higher frequency of Th17 cells, rather than Th1 cells, is found in CSF from MS patients, with numbers peaking during relapse (Brucklacher-Waldert et al., 2009). Th17 cells from the blood and CSF of MS patients also show increased proliferation compared to Th1 cells and have higher expression of adhesion molecules and activation markers, as well as being less susceptible to suppression by Tregs (Brucklacher-Waldert et al., 2009). However, there is evidence from EAE studies that Th1 cells are required to infiltrate into the CNS before Th17 cells (O'Connor et al., 2008)

and that Th17 cells in the CNS can convert into Th1-like IFN-γ-producing cells, suggesting a more prominent role for Th1 cells in disease pathogenesis (Hirota et al., 2011).

Tfh cells are the most recently described CD4+ T cell effector lineage and are required for the formation of B cell germinal centres. Structures similar to germinal centres, termed ectopic lymphoid follicles, are found in the meninges of patients with SPMS (Serafini et al., 2004) and are believed to be sites of T and B cell activation and differentiation. Although the precise role of Tfh cells in MS is yet to be characterised, a recent study reported that there is an increased frequency of these cells in blood from patients with RRMS and SPMS compared to healthy controls and that Tfh cell frequency was correlated with disease progression (Christensen et al., 2013). Increased expression of Tfh markers was also found in CSF from these patients (Christensen et al., 2013), supporting a role for these cells in MS that requires further study.

As a key regulatory CD4+ T cell subset, the forkhead box P3 (FoxP3)+ Tregs play a major role in maintaining peripheral tolerance, with mutations in the *FOXP3* gene causing the severe inflammatory autoimmune condition immunodysregulation, polyendocrinopathy, enteropathy, and X-linked syndrome (IPEX) (Bennett et al., 2001; Gambineri et al., 2003; Wildin et al., 2001). The number of Tregs present in blood and CSF from patients with MS is comparable to healthy controls, however *in vitro* assays reveal an impaired ability to suppress myelin-specific effector T cell proliferation, which could be related to the reduced expression of FoxP3 seen in Tregs derived from MS patients (Huan et al., 2005; Viglietta et al., 2004). Tregs from MS patients also show reduced motility and impaired migration in a well-established *in vitro* brain endothelium model (Schneider-Hohendorf et al., 2010). An additional regulatory CD4+ T cell subset, the Treg type 1 (Tr1) cells also show impaired suppressive function in MS, producing less of the anti-inflammatory cytokine IL-10 than Tr1 cells derived from healthy controls (Martinez-Forero et al., 2008).

1.3.2. CD8+ T cells

In addition to the evident role for CD4+ T cells in MS, cytotoxic CD8+ T cells were also identified in MS lesions in the 1980s (Traugott et al., 1983). Interestingly, CD8+ effector T cells

greatly outnumber CD4+ T cells in active MS lesions and there is evidence that they are clonally expanded, suggesting that antigen stimulation has occurred within the CNS (Babbe et al., 2000). Further indicating the importance of CD8+ T cells in the disease, a role for viruses in disease aetiology has long been speculated and alleles of the MHC class I region have been associated with MS (discussed later in sections 1.4 and 1.5.1, respectively). CD8+ T cells are thought to be involved in CNS tissue damage through cytokine production as well as the secretion of cytotoxic granules containing granzyme and contact-mediated lysis via perforin (Friese and Fugger, 2005).

Autoreactive CD8+ T cells in the blood and CSF of MS patients produce IFN-γ in response to stimulation with CNS autoantigens, and these cells have also been shown to produce IL-17 in MS lesions (Tzartos et al., 2008; Wallstrom et al., 2000) and also directly associate with and damage axons (Medana et al., 2001). Further implicating these cells in disease pathogenesis, CD8+ T cells with an activated phenotype are found in the CSF and cortex of MS patients with early stage disease (Jilek et al., 2007; Lucchinetti et al., 2011). Moreover, directly incriminating CD8+ T cells in disease initiation, a transgenic mouse model expressing human MHC class I and a myelin-specific TCR derived from MS patient CD8+ T cells was shown to develop spontaneous MS-like disease (Friese et al., 2008).

In addition to effector cells, regulatory CD8+ T cell populations have also been identified that have the capacity to suppress autoreactive T cells. In contrast to regulatory CD4+ cell subsets, which have been extensively characterised, the phenotype and specific function of human regulatory CD8+ T cell subsets is less well defined due to most studies being performed in animal models. However, there is evidence of a role for regulatory CD8+ T cells in MS. CNS autoantigen-specific regulatory CD8+ T cells from both healthy subjects and MS patients have been shown to suppress CD4+ effector T cell proliferation and this suppressive ability is reported to correlate with disease remission and also be impaired during MS relapse (Baughman et al., 2011). In addition, GA treatment of MS has been shown to cause an expansion of regulatory CD8+ T cells, which are able to directly target and kill CD4+ T cells (Karandikar et al., 2002).

1.3.3. B cells

As previously mentioned, ectopic lymphoid follicles are found in the MS-affected brain, which contain proliferating B cells and plasma cells as well as dendritic cells (DCs) and T cells (Serafini et al., 2004). Early studies detected clonally expanded immunoglobulins (Igs) in CSF from MS patients (Delmotte, 1971; Kabat et al., 1942) and the presence of oligoclonal bands is a characteristic and diagnostic feature of MS. However, until fairly recently the production of the Igs that form oligoclonal bands by B cells present in the CSF of MS patients was only assumed and not empirically demonstrated (Obermeier et al., 2008). B cells, plasma cells and myelin-specific antibodies are also present in areas of active demyelination in MS patient tissue, and established brain lesions are associated with Ig deposition and complement-mediated demyelination (Breij et al., 2008; Esiri, 1977; Genain et al., 1999).

A key role for B cells in MS immunopathogenesis is further bolstered by the recent success of rituximab in clinical trials. This drug specifically depletes B cells as it targets the CD20 molecule expressed by these cells and MS patients receiving rituximab treatment showed rapid and significant improvements in both MRI and clinical measures (Hauser et al., 2008). As the drug does not deplete plasma cells, which produce antibodies, the trial data suggest the relevance of non-antibody-mediated functions of B cells, such as antigen presentation in activating T cells and also the modulation of T cell activity. Compared to healthy controls, B cells derived from MS patients show exaggerated responses to stimulation, producing more lymphotoxin and tumour necrosis factor-α (TNF-α), which drive pro-inflammatory responses by both CD4+ and CD8+ T cells (Bar-Or et al., 2010). In line with this, the B cell depletion achieved by rituximab treatment is associated with a concomitant reduction in T cell proliferation and cytokine production (Bar-Or et al., 2010).

1.3.4. Innate immune cells

In addition to B cells, innate immune cells including macrophages/microglia and DCs act as antigen-presenting cells (APCs) both in the periphery and in the CNS, and these cells also have distinct effector functions and can modulate T cell responses, thus contributing to MS pathogenesis.

Macrophages and CNS-resident microglia are considered as the main APCs in MS and are the most well studied innate immune cells in disease pathogenesis. Infiltrating macrophages and activated microglia are prominent in active MS lesions in both RRMS and progressive disease (Lucchinetti et al., 2000; Prineas and Wright, 1978). The effects of chronic microglial activation are not restricted to lesions but believed to induce widespread inflammatory changes in the MS-affected brain, as well as contributing significantly to neurodegeneration in progressive MS (Lassmann et al., 2012). As well as actively phagocytosing myelin debris, both macrophages and microglia secrete pro-inflammatory cytokines and reactive oxygen species, thus damaging neurons and the oligodendrocyte processes which form myelin (Gandhi et al., 2010). Macrophages and microglia also produce anti-inflammatory mediators and neurotrophic factors, thus promoting repair (Gandhi et al., 2010). The role of these cells in MS pathology is thus a complex interplay between beneficial and detrimental effects.

DCs are professional APCs that play a key role in T cell activation and differentiation. Plasmacytoid dendritic cells (pDCs) isolated from MS patients show impaired activation and immunoregulatory function (Stasiolek et al., 2006) and the pro-inflammatory function of myeloid DCs (mDCs) is also dysregulated in MS, with these cells showing hyperactivation and exaggerated production of inflammatory cytokines (Karni et al., 2006).

Innate cells other than APCs have also been implicated in MS. For example, blood neutrophil counts are elevated in RRMS (Debruyne et al., 1998) and granulocytes isolated from MS patients show increased expression of molecules involved in adhesion and migration, as well as enhanced effector activity (such as degranulation), which may contribute to MS pathogenesis by increasing inflammation and tissue injury (Naegele et al., 2012). Mast cells are also reported to accumulate in MS lesions (Olsson, 1974) and may play a key role in autoreactive T cell responses, as mice lacking mast cells develop a milder form of EAE (Gregory et al., 2005).

In addition, there is some evidence that innate-like lymphocytes play a role in disease. Gamma-delta (γδ) T cells have a limited TCR repertoire that recognises non-MHC restricted antigen. A high frequency of γδ T cells is found in the blood and CSF of MS patients, and within lesions these cells may cause direct damage and lysis of oligodendrocytes (Freedman et al., 1991; Selmaj et al., 1991; Stinissen et al., 1995). Natural killer (NK) cells are also present in MS lesions (Traugott, 1985) and *in vitro* studies demonstrate their cytotoxic activity towards oligodendrocytes (Saikali et al., 2007). However, there is also evidence to suggest that the expansion of CD95+ NK cells during MS remission may mediate a protective effect by actively suppressing autoreactive T cell responses (Takahashi et al., 2004). NKT cells may also play a protective role in MS, as cells isolated from patients show an immunoregulatory phenotype (Araki et al., 2003).

1.3.5. Summary: our understanding of MS

Evidently, MS is a complex disease and although our knowledge has greatly improved, the exact mechanisms underpinning disease initiation, pathogenesis and progression remain to be fully elucidated. Delineating the precise aetiology of MS would greatly progress our understanding of the disease, and identifying the contributory factors will inform the development of targeted and more efficacious therapies.

1.4 The multi-factorial aetiology of MS: environmental factors

It is widely believed that MS is a complex disease with no single cause: common polymorphisms at multiple genes determine disease predisposition, while environmental factors influence disease penetrance in genetically susceptible individuals. The first indication for a role of the environment in MS aetiology came from observed differences in the geographical distribution of disease, which could not be explained by genetics alone. There is a higher prevalence and incidence of MS in populations with increasing distance from the equator, suggesting a correlation between MS

and reduced sunlight exposure (Ascherio and Munger, 2007b). The change in risk observed in migration studies also further supports a role for environmental factors, with individuals relocating from geographical regions with low to high prevalence showing an increased risk of developing MS, and vice versa (Ascherio and Munger, 2007b; Compston and Coles, 2008).

The effect of latitude on MS prevalence is an established observation that has more recently been associated with exposure to sunlight and subsequent vitamin D production (Ascherio et al., 2010). A lack of sunlight and ultra-violet (UV) exposure has been reported as the most significant environmental risk factor for MS, having an estimated 20-fold greater effect than other such influences (Sloka et al., 2011). UV exposure is required for the conversion of dietary vitamin D into its biologically active form, which has many effects on the immune system (as reviewed by (Koch et al., 2013)). Disappointingly, high dose vitamin D supplementation has not achieved significant beneficial effects in trials with MS patients, with no effect relapse rate or MRI measures, however larger studies of longer duration may be required (James et al., 2013; Stein et al., 2011).

Salt is another dietary factor that has recently been implicated in EAE, with two recent publications demonstrating a role for sodium chloride in driving pathogenic Th17 responses (Kleinewietfeld et al., 2013; Wu et al., 2013). Although it is tempting to speculate that the high salt levels in the Western diet may have contributed to the recent increase in the prevalence of MS, the role of salt as an environmental risk factor in MS requires further investigation.

There is growing evidence for a role of commensal gut bacteria in MS susceptibility. Commensal microbiota were shown to be necessary for disease induction as mice housed in germ-free conditions were completely resistant to the development of myelin autoantigen-induced demyelinating disease (Berer et al., 2011). As this is a relatively new field, there are yet to be any large studies in individuals with MS, however a small study involving 15 patients reported that there were no overall differences in gut microbiota compared to healthy controls (Mowry et al., 2012). It is suggested that treatment with GA may alter bacterial populations in the gut (Mowry et al., 2012), however the extent to which this action mediates a therapeutic effect is unclear and requires further study. Although larger investigations are required to elucidate the role of the gut microbiome in MS, alterations in commensal microbiota have been reported in individuals with type 1 diabetes (T1D)

and inflammatory bowel disease (IBD) suggesting a more general role in determining susceptibility to autoimmunity (Pillai, 2013). This field may offer novel therapeutic options in MS; indeed the administration of a probiotic helminth is reportedly well tolerated in a pilot study involving patients with RRMS (Fleming et al., 2011). Treatment was also associated with increased anti-inflammatory cytokine levels in the blood as well as a reduction in the number of new lesions (Fleming et al., 2011), however the potential of probiotic dietary supplements in MS therapy requires much further study.

In addition to nutritional factors, epidemiological studies strongly implicate a role for infection in the aetiology of MS. Although many viruses have been reportedly associated with MS, the only consistent convincing evidence supports a role for EBV as an environmental risk factor. Around 95% of adults are seropositive for EBV and seronegativity is associated with a very low risk of MS, with a meta-analysis reporting an odds ratio of 0.06 compared to seropositive individuals (Ascherio and Munger, 2007a). Individuals that have suffered from infectious mononucleosis (caused by EBV infection) are reported to be twice as likely to develop MS as those that have not (Handel et al., 2010). Despite strong evidence for an epidemiological association between EBV and MS, this does not imply a causal role of the virus and the precise role of EBV in MS pathogenesis remains unclear. Potential mechanisms of infection triggering autoreactivity have been discussed in section 1.2.

Another environmental factor influencing MS risk is smoking, with smokers being 1.6 times more likely to develop MS than individuals that have never smoked (Hernan et al., 2001). Smoking is associated with increased brain atrophy and a greater lesion load in MS patients, and also specifically an increase in the number of gadolinium enhancing lesions (Zivadinov et al., 2009), suggesting that smoking increases permeability of the BBB.

There is undoubtedly a role for environmental factors in MS aetiology. Further study is required in larger cohorts and over longer time periods, however even strong associations do not necessarily indicate a causal role. Therefore, a concomitant investigation of the molecular mechanisms underpinning these associations is also necessary. In addition, the backdrop of genetic

susceptibility that these environmental factors are acting on must be considered in order to address the interplay between the environment and genetics in the context of disease development.

1.5 MS genetics

The familial clustering of MS was first noted over a century ago (Eichorst, 1896) and provided the initial evidence of a genetic contribution to the disease (Doolittle et al., 1990; Mackay, 1950; Robertson et al., 1996; Schapira et al., 1963), with the lifetime disease risk in siblings of an affected individual being six times that of the risk observed in the general population (Hemminki et al., 2009). Studies on adoptees (Ebers et al., 1995), half-siblings (Sadovnick et al., 1996) and spouses (Ebers et al., 2000) of individuals with MS further supported the existence of a genetic component. Twin studies also demonstrated the presence of pre-disposing genetic factors, with a consistently higher concordance rate for MS development being observed in monozygotic compared to dizygotic twins (Hansen et al., 2005; Hawkes and Macgregor, 2009; Willer et al., 2003). Taken together these investigations indicate that MS does not follow a Mendelian pattern of inheritance but that it is a polygenic disorder with a complex architecture of genetic risk (Sadovnick et al., 1997).

The common disease-common variant (CDCV) hypothesis states that for common conditions, like MS, a number of variants will contribute to disease development, each will contribute a moderate proportion of the overall disease risk and variants will be common and thus will be distributed throughout the population in which the disease manifests (Fisher, 1930). This model for common disease genetics was further amended to account for the potential existence of a few additional genetic factors with larger effect sizes that are maintained in the population despite their greater effect on disease development because they confer some other selective advantage. Hence, the relationship between the disease-associated genetic variants and the effect size on disease risk would be predicted to follow an L-shaped distribution (Wang et al., 2005).

1.5.1. The MHC: the strongest genetic effect on MS susceptibility

The MHC region is the most gene dense part of the genome; it is highly polymorphic and stretches for 2.5 megabases on chromosome six. It contains 230 known genes and pseudogenes (Horton et al., 2004), which fall into three classes. Class I includes the *HLA-A*, *-B* and *-C* genes, which encode membrane-bound proteins that constitute the alpha-chain of class I receptors that associate with β2 microglobulin; MHC class I molecules are found on all nucleated cells and platelets. Class II includes genes for *HLA-DRB1*, *-DRA*, *-DQB1*, *DQA1*, *DPB1* and *DPA1*, which encode MHC alpha- and beta-chains that are expressed mainly on APCs and some T cells (Baecher-Allan et al., 2006). Class III includes heat shock proteins, certain cytokines and complement proteins (Kumanovics et al., 2003). The highly polymorphic nature of the genes encoding MHC class I and II molecules is central to the role of these molecules in the discrimination of self and non-self by the immune system, which is critical in responses to infection as well as self-tolerance (discussed in section 1.2).

An association within the MHC region with MS was first identified in the 1970s and genetic variation in the HLA region remains the largest and most replicated genetic risk factor in MS (Compston et al., 1976; Jersild et al., 1973; Terasaki et al., 1976; Winchester et al., 1975). The *DRB1*15:01-DRB5*0101-DQB1*06:02* haplotype (*DR2* haplotype) was found to have the greatest effect on MS susceptibility (Olerup and Hillert, 1991) and further genetic analyses dissecting the contribution of the individual alleles in the haplotype have demonstrated that the *DRB1*15:01* allele is responsible for the disease association with an odds ratio (OR) of 3.1 (Etzensperger et al., 2008; Sawcer et al., 2011). Additional HLA class II alleles are also associated to a lesser extent with MS susceptibility; DRB1*03:01 and DRB1*13:03 have ORs of 1.26 and 2.4, respectively (Sawcer et al., 2011). In addition, the HLA class I allele *HLA-A2* is associated with protection against MS (OR = 0.73) (Sawcer et al., 2011).

Certain MHC haplotypes are also associated with other autoimmune diseases and although it is generally accepted that polymorphism influences the antigen-presenting ability of these molecules, functional assessment of disease-associated haplotypes remains incomplete. Some progress has been made using transgenic mice to model the effects of HLA alleles on MS susceptibility. In mice

expressing a TCR cloned from an MS patient, the protective effect of *HLA-A2* in a murine model of MS-like disease was shown to correlate with a reduced frequency of circulating CD8+ T cells that showed impaired proliferation and cytokine production in response to stimulation (Friese et al., 2008). Additionally, crystallography studies have revealed cross-reactivity of an MS patient-derived TCR with a peptide of myelin basic protein presented by HLA-DRB1*1501 and an EBV peptide complexed with HLA-DRB5*0101 (Lang et al., 2002). This implicates certain disease-associated HLA alleles in molecular mimicry and offers a functional link between HLA alleles and potential disease mechanisms. In addition to influencing MS susceptibility, recent evidence also correlates expression of the *DRB1*15:01* allele with an increased degree of demyelination, inflammation and axonal loss in MS spinal cord lesions (DeLuca et al., 2013), suggesting a link between a genetic risk factor and the extent of disease pathology.

Even a full elucidation of the functional effects of associated MHC alleles will not provide a complete explanation of disease aetiology as the MHC region associations explain only approximately 10% of MS heritability (Sawcer et al., 2011). Following the discovery of the MHC association, no other contributing genetic factors were consistently and significantly associated with MS due to the limited power of linkage studies and early association studies (Sawcer et al., 2005), although it was estimated that some 20-100 non-MHC variants were likely to contribute to disease risk (Sawcer et al., 2010). In 2007, the advent of large-scale genome-wide association studies (GWAS) provided an approach able to detect the smaller effect sizes of the non-MHC MS-associated common variants (Hafler et al., 2007).

1.5.2. GWAS in MS and progress from genes to function

In recent years, genome-wide investigation of disease-associated genetic variation has become possible due to several converging factors. Since the publication of the first complete genome sequence, there has been a concerted effort to catalogue the genetic variation across the genome through projects such as the International HapMap Project (Altshuler et al., 2010) and the 1000 Genomes Project (Abecasis et al., 2010). This required the concomitant development of

technologies enabling millions of single nucleotide polymorphisms (SNPs) to be genotyped for any single sample. Additionally, international multi-centre collaborations have created case-control cohorts large enough for high-powered GWAS to be performed. It has been estimated that for a sufficiently powered GWAS, at least 2000 cases and 2000 controls are required and SNPs are typically considered to have reached genome-wide significance when the *P*-value for the association is less than $5x10^{-8}$, with this significance threshold being set such that it accounts for the increased likelihood of detecting false positive signals when testing millions of SNPs across the entire genome (Oksenberg and Baranzini, 2010).

The first GWAS and meta-analyses identified some 20 non-MHC MS-associated loci (ANZgene, 2009; Baranzini et al., 2009; Burton et al., 2007; De Jager et al., 2009; Hafler et al., 2007; Patsopoulos et al., 2011; Sanna et al.). A larger GWAS completed in 2011 identified further novel loci, taking the total number to over 50 MS-associated regions (Sawcer et al., 2011). Notably, the loci showed significant enrichment for genes with known immunological function, with only two loci containing genes for pathways relating to neurodegeneration. The authors noted a particular enrichment of genes involved in the differentiation and signalling of T helper cells (Sawcer et al., 2011). By considering the associated regions collectively, gene networks highlighted the biological pathways that may be influenced by genetic variation, emphasising the need for functional studies to elucidate the precise molecular mechanisms underpinning the identified genetic associations.

The functional contribution of MS-associated genetic variation has been published for some replicated GWAS SNPs. The rs6897932 SNP in the IL-7 receptor gene (*IL7R*) was reported in the first MS GWAS, with the association being replicated and the functional consequences explored soon after (Hafler et al., 2007). Gregory and colleagues demonstrated that the common allele associated with MS risk causes increased skipping of the exon encoding the transmembrane domain of the IL-7 receptor, enhancing production of a soluble receptor isoform which may antagonise signalling (Gregory et al., 2007).

The potential translational impact of understanding of MS genetics has been debated, as the genetic contribution of non-HLA regions to overall MS susceptibility is relatively small. However, recent work has demonstrated that such associations may inform clinical practice. The risk allele of a

SNP in the TNF receptor superfamily member 1A (*TNFRSF1A)* gene region was shown to cause increased production of a soluble receptor isoform capable of blocking TNF signalling (Gregory et al., 2012). In line with this genetic effect, anti-TNF therapies cause a worsening of MS, but are beneficial in diseases that are not associated with genetic risk in the *TNFRSF1A* region highlighting the ability of functional work to provide an insight into therapeutically relevant disease mechanisms (Gregory et al., 2012). Thus, determining the functional consequences of the causal variants in other associated regions will progress our knowledge of the pathways involved in the pathogenesis of MS and also have potential clinical implications.

1.5.3. Limitations of GWAS

Although GWAS data have contributed to our understanding of the genetic underpinning of autoimmune disease in terms of the number of independent associations identified, a significant proportion of the genetic heritability of MS remains unexplained. The potential sources of this missing heritability that have not been captured by current arrays could include common variants of smaller effect size, rare variants (with a minor allele frequency of less than 1-5%), gene-gene interactions, gene-environment interactions, heritable epigenetic influences and other types of genetic variation apart from SNPs, including structural variants such as insertions, deletions and translocations (Manolio et al., 2009). Additional variants may also only be found through next generation sequencing technologies and the analysis of isolated populations or multiplex families (Albrechtsen et al., 2013; Baranzini et al., 2010; Nejentsev et al., 2009; Rivas et al., 2011).

Whilst GWAS have not discovered all the genetic variation that contributes to MS heritability, the study of identified associations can be informative but requires consideration of the inherent limitations of the GWAS approach. GWAS have been designed to take advantage of the haplotypic structure of the genome whereby regions of deoxyribonucleic acid (DNA) bound by recombination sites tend to be inherited as blocks. Therefore, an interrogation of genetic variation across the genome does not require typing of all variants, as typing SNPs that may serve to tag haplotypic variation will be sufficient to capture an underlying association with disease. Hence, due

to linkage disequilibrium (LD), SNPs that are identified as being disease-associated may not be causative but may simply markers of the truly causative genetic variants (Donnelly, 2008). Knowing which variant is truly causative may have implications for downstream hypotheses concerning the functional effects of the variation. Additionally, whilst candidate genes are typically assigned to each identified associated region, these regions often contain numerous genes and hypotheses regarding effects of genetic variants on gene function must be investigated with care, in particular when considering associated variants in non-coding regions. Thus, improved mapping of the variants responsible for identified association signals is required as an initial step towards determining their functional implications.

1.5.4. Fine-mapping associations identified through GWAS: the ImmunoChip

The custom Illumina 200 K Infinium ImmunoChip was designed as a novel cost-effective and efficient high-density genotyping array through a collaboration between the Wellcome Trust Case-Control Consortium (WTCCC) and leading researchers from across the autoimmune disease field, including groups investigating MS, T1D autoimmune thyroid disease (AITD), ankylosing spondylitis (AS), celiac disease (CeD), Crohn's disease (CrD), IgA deficiency, primary biliary cirrhosis (PBC), ulcerative colitis (UC), rheumatoid arthritis (RA), systemic lupus erythematosus (SLE) and psoriasis (PS). The ImmunoChip detects 196,524 polymorphisms in total, across 186 loci that reached genome-wide significance in GWAS of one or more immune-mediated diseases, as well as ~700 small insertions/deletions (Cortes and Brown, 2011). The ImmunoChip also includes some SNPs from non-immunological disease GWAS (e.g. ischemic stroke, Parkinson's disease) and all known SNPs from population-based sequencing projects such as the 1000 Genomes Project, thereby allowing for high resolution mapping (fine-mapping) of common and also some rare variants. Overall, the ImmunoChip has superior resolution and power to prior GWAS, as it is more likely to directly genotype and localise causal variants due to the increased density of detectable markers (Cortes and Brown, 2011).

Since 2011, ImmunoChip datasets have been published for several autoimmune diseases (see Table 1.1.). The most recently published ImmunoChip data is for MS and reports the discovery of 48 new non-MHC variants, bringing the total number of non-MHC MS-associated variants to 110 (IMSGC, 2013). It is estimated that these variants, along with the MHC, account for 20% of MS heritability (IMSGC, 2013). Although ImmunoChip data may not reveal a complete picture of disease heritability, the improved delineation of associated variants will promote the discovery of genotype-dependent functional effects and how they contribute to autoimmune disease pathogenesis. To date, a significant degree of overlap has been detected between the genetic architecture of MS and other autoimmune diseases (IMSGC, 2013; Parkes et al., 2013). Around 22% of all MS-associated signals were found to overlap with at least one other autoimmune disease, with IBD having the highest degree of overlap (9.1%) (IMSGC, 2013). This reinforces the concept that common autoimmune diseases share at least some pathogenic mechanisms and indicates that studying the functional consequences of disease-associated variants may prove informative for several different autoimmune conditions. Grouping of candidate genes from ImmunoChip loci into networks highlights pathways that may suggest the mechanisms underpinning the genetic association with disease (Figure 1.1.) (Cortes et al., 2013). Many candidate genes from shared associated loci are involved in T helper cell and cytokine signalling pathways, for example the gene encoding tyrosine kinase 2 (*TYK2*; Figure 1.1.).

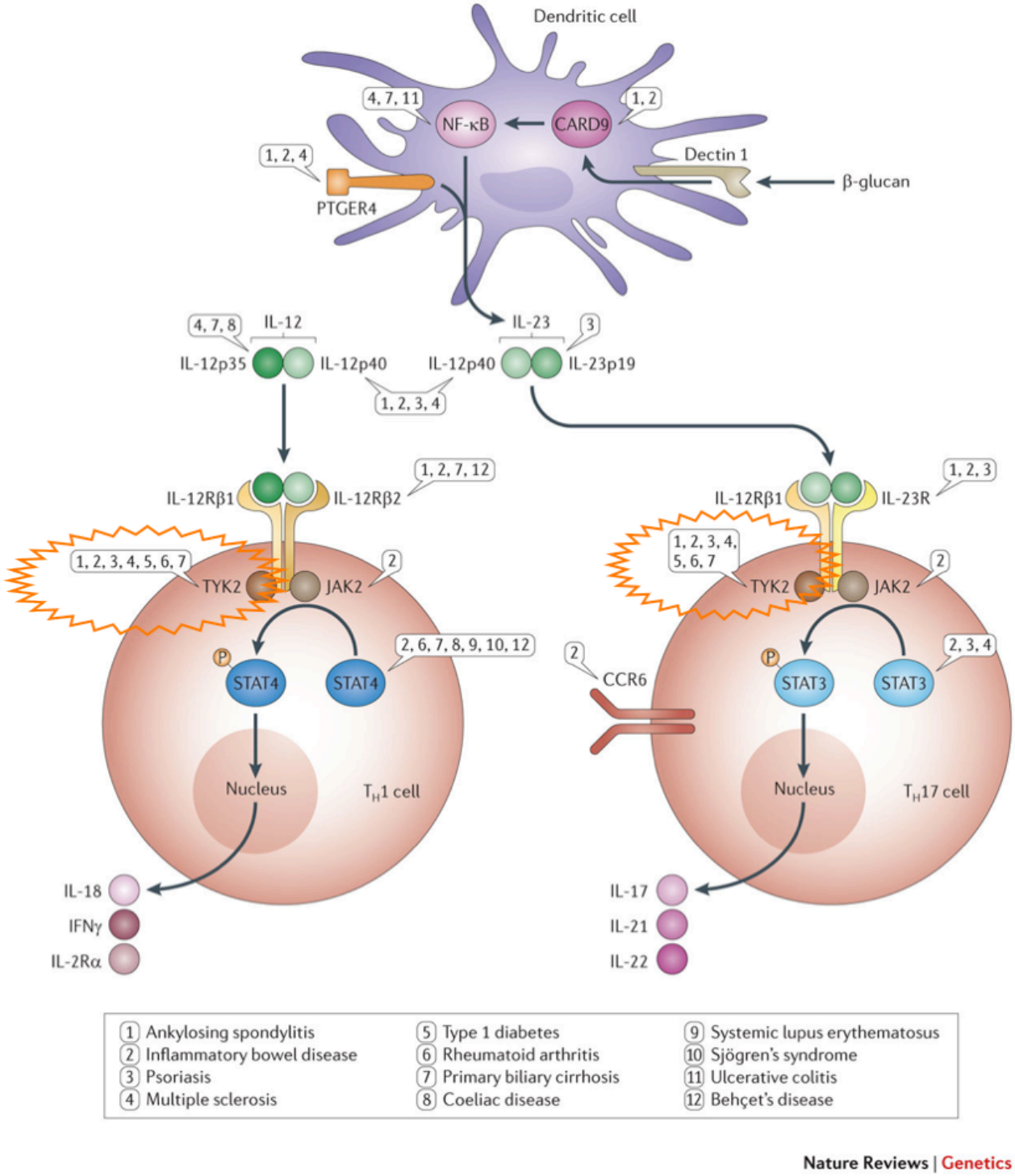


Figure 1.1. Published ImmunoChip data identify loci shared across multiple autoimmune diseases containing genes implicated in T helper cell differentiation and activation pathways
In line with data from prior GWAS, ImmunoChip data highlight a role for T helper cell signalling pathways across autoimmune diseases. Tyrosine kinase 2 (TYK2; highlighted in orange) plays a key role in cytokine signalling pathways and the *TYK2* gene region has been associated with multiple autoimmune diseases (see text; section 1.6). Diagram modified from (Cortes et al., 2013).

1.6 The *TYK2* region as an autoimmune disease locus

The *TYK2* gene on human chromosome 19 encodes non-receptor tyrosine kinase 2 (TYK2). Four SNPs in the *TYK2* gene region have now been associated with multiple autoimmune diseases (see Table 1.1), suggesting that variation in this region may have a more general role in maintaining the balance between self-tolerance and autoimmunity. *TYK2* was first identified as a potential candidate gene through a nonsynonymous SNP (nsSNP) scan in a combined analysis of MS, AS and AITD (Burton et al., 2007). The association with MS was subsequently replicated in a larger case-control cohort and across several populations (ANZgene, 2009; Ban et al., 2009; Johnson et al., 2009; Mero et al., 2009). Recently published ImmunoChip data have confirmed the association of the originally identified rs34536443 SNP with MS (IMSGC, 2013) and also with other autoimmune diseases (see Table 1.1). Notably, the rs34536443 SNP is the most associated non-MHC variant for PBC, PS, RA and juvenile idiopathic arthritis (JIA), having the greatest effect on protection against disease according to the reported ORs (Eyre et al., 2012; Hinks et al., 2013; Liu et al., 2012; Tsoi et al., 2012) (see Table 1.1).

GWAS have also identified another nsSNP in the *TYK2* gene, rs12720356, as an independent signal from the rs34536443 SNP that is associated with IBD (both CrD and UC; (Anderson et al., 2011; Franke et al., 2010; Strange et al., 2010)), and also PS (Tsoi et al., 2012). Interestingly, the minor allele of the rs12720356 SNP is protective against PS but is associated with risk in IBD (Anderson et al., 2011; Franke et al., 2010). Another independent signal in the region, rs11879191, is associated with IBD as well as AS (Cortes et al., 2013; Jostins et al., 2012). The rs11879191 SNP does not lie within an exon of the *TYK2* gene but is instead found in intron 1 of the neighbouring gene *CDC37* that encodes cell division cycle 37 protein. Although *TYK2* is reported as the candidate gene for the rs11879191 association signal (Cortes et al., 2013; Jostins et al., 2012), there is currently no functional evidence to support this.

At present, there are no published ImmunoChip data for T1D or SLE but previous GWAS for both conditions have indicated the existence of an association in the *TYK2* gene region. For T1D,

an nsSNP scan identified rs2304256 as being associated with the disease, with the minor allele being protective (Wallace et al., 2010). A recent review also cites a personal communication from the authors of the T1D ImmunoChip, revealing that this SNP association has now been replicated (Parkes et al., 2013). Conflicting results from several small GWAS (<2,000 cases) have been published regarding the association of rs2304256 with SLE (Cunninghame Graham et al., 2007; Hellquist et al., 2009; Jarvinen et al., 2010; Sigurdsson et al., 2005; Suarez-Gestal et al., 2009), however several meta-analyses report an association of rs2304256 with the disease (Cunninghame Graham et al., 2011; Lee et al., 2012; Suarez-Gestal et al., 2009). The publication of the ImmunoChip data for SLE will clarify this potential association with the *TYK2* region. An association has not been reported in the *TYK2* region following the publication of ImmunoChip datasets for AITD, atopic dermatitis and CeD (Cooper et al., 2012; Ellinghaus et al., 2013; Trynka et al., 2011).

Table 1.1. Autoimmune disease-associated SNPs in the *TYK2* gene region
Four SNPs in the *TYK2* gene region are reported to be associated with autoimmune disease in published data from GWAS and ImmunoChip. Minor allele frequency (MAF) refers to frequency in the control cohort from the dataset highlighted in bold. The odds ratio (OR) refers to the minor allele and the 95% confidence interval (95% CI) has been provided from studies that report this value in their publication. OR<1 indicates a protective effect on disease susceptibility; OR>1 indicates risk.
RA[a] – anti-citrullinated protein antibody (ACPA-) positive RA patients only
JIA[b] – oligoarticular and rheumatoid factor-negative polyarticular JIA only
AS[c] – reported SNP rs35164067 in complete LD (r^2 and D-prime = 1) with rs11879191.
SLE[d] – genome-wide significance only reported when performed combined analysis of all available data; no OR reported for combined analysis, so OR calculated from the cohort in the provided reference only.

SNP I.D.	Minor allele (MAF)	Associated disease	OR for minor allele (95% CI)	Dataset	Reference
rs34536443	C (0.05)	MS	0.76 (0.68-0.85)	GWAS	(Ban et al., 2009)
			0.77	GWAS	(Mero et al., 2009)
			0.75 (0.68-0.83)	GWAS	(Johnson et al., 2009)
			0.78	ImmunoChip	(IMSGC, 2013)
		PBC	0.52 (0.44-0.63)	ImmunoChip	(Liu et al., 2012)
		RA[a]	0.62	ImmunoChip	(Eyre et al., 2012)
		JIA[b]	0.56 (0.47-0.67)	**ImmunoChip**	(Hinks et al., 2013)
		PS	0.53	ImmunoChip	(Tsoi et al., 2012)
rs12720356	C (0.1)	CrD	1.12 (1.06-1.19)	GWAS	(Franke et al., 2010)
			1.22 (1.14-1.31)	GWAS	(Anderson et al., 2011)
		UC	1.17 (1.09-1.26)	GWAS	(Anderson et al., 2011)
		PS	0.71 (0.62-0.81)	GWAS	(Strange et al., 2010)
			0.8	**ImmunoChip**	(Tsoi et al., 2012)
rs11879191	A (0.2)	CrD	0.87 (0.84-0.91)	**ImmunoChip**	(Jostins et al., 2012)
		UC	0.89 (0.85-0.93)	ImmunoChip	(Jostins et al., 2012)
		AS[c]	0.88	ImmunoChip	(Cortes et al., 2013)
rs2304256	A (0.27)	T1D	0.86 (0.82-0.9)	GWAS	(Wallace et al., 2010)
			0.86	**ImmunoChip**	(Parkes et al., 2013)
		SLE[d]	0.79 (0.7-0.9)	GWAS	(Suarez-Gestal et al., 2009)

1.7 Structure of the Janus kinase TYK2

Three of the four identified SNPs in the *TYK2* gene region are nonsynonymous and thus may affect TYK2 protein function by changing its amino acid sequence (Figure 1.2.). TYK2 was the first Janus kinase (JAK) to be cloned (Firmbach-Kraft et al., 1990) and the family share seven JAK homology (JH) regions and four functional domains (Figure 1.2.).

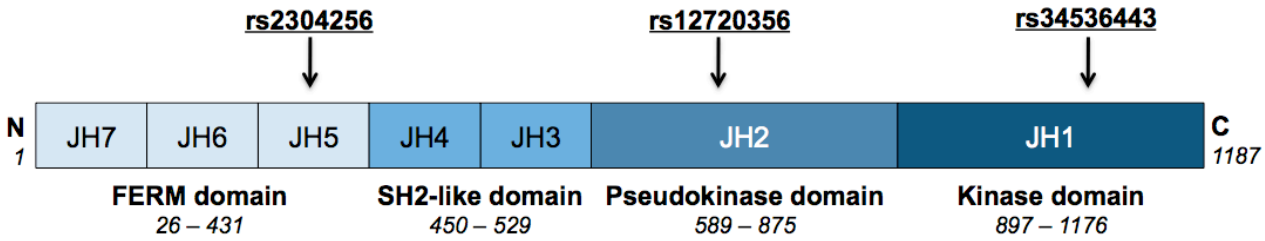


Figure 1.2. TYK2 protein structure and the location of nsSNP-encoded amino acid substitutions
A schematic of TYK2 protein structure from the amino acid (N) to carboxy (C) terminal with numbers indicating amino acid residues. The JAKs share seven JAK homology (JH) regions and four functional domains. The FERM and SH2 domains are involved in the interaction of JAKs with cytokine receptors, and additionally mutations of the FERM domain have been shown to affect kinase activity (see main text). The pseudokinase domain has been demonstrated to exert regulatory effects on the catalytic JH1 domain (see main text). The rs2304256, rs12720356 and rs34536443 SNPs have been reported to be associated with various autoimmune diseases and affect the amino acid sequence in different domains of TYK2 (indicated by arrows; see main text). The specific amino acid substitutions and their relation to disease are discussed in detail in Chapter 3. Diagram modified from (Levine and Gilliland, 2008).

The 4.1 ezrin, radixin, moesin (FERM) domain anchors JAKs to cytokine receptors, and was first demonstrated to be necessary and sufficient for JAK3 binding to the common gamma chain receptor subunit (Chen et al., 1997). Subsequently, mutations in the FERM domain have also been shown to abolish JAK3 activity in patients with severe combined immunodeficiency (SCID; (Zhou et al., 2001)). Additionally, a nuclear localization signal has been identified in the FERM domain of TYK2 (Ragimbeau et al., 2001) suggesting potential roles for this JAK in the nucleus, as discussed in section 1.9. The T1D- and SLE-associated rs2304256 SNP affects the amino acid sequence in the FERM domain and may thus alter the function of TYK2 (Figure 1.2.).

Although the JAK Src homology 2 (SH2)-like domain shows homology to typical SH2 domains, it lacks the conserved arginine residue found in the key phospho-tyrosine-binding position suggesting that it has an alternative function (Al-Lazikani et al., 2001). Consistent with this, there is some evidence for a structural role for the SH2-like domain of TYK2, as it has been shown to be required for maintenance of interferon (alpha, beta and omega) receptor 1 (IFNAR1) surface expression in a cell line (Ragimbeau et al., 2003).

The presence of two neighbouring tyrosine kinase domains is a unique feature of JAK enzymes that led to their family name deriving from the two-faced Roman god Janus (Wilks et al.,

1991). However, the JH2 domain was originally characterised as a pseudokinase domain, due to lacking typical kinase domain motifs (Wilks et al., 1991). The pseudokinase domain regulates the JH1 kinase activity (see below) and as the minor allele of the rs12720356 SNP introduces an amino acid change in this domain (Figure 1.2), it may affect TYK2 function as has been demonstrated for other mutations in this region. Certain mutations in the pseudokinase domain lead to hyperactivity of JAK2, which is known to cause myeloproliferative diseases (MPD) and account for nearly all cases of polycythemia vera (James et al., 2005). Artificially introducing an equivalent mutation in TYK2 similarly renders the molecule hyperactive (Staerk et al., 2005). Pathological hyperactivity resulting from a mutation in the Jak1 pseudokinase domain was also recently reported in an N-ethyl-N-nitrosourea (ENU) mutagenesis-derived mouse model with an SLE-like phenotype, and the equivalent mutations were found to cause acute lymphoblastic leukaemia in humans (Sabrautzki et al., 2013). However, not all pseudokinase domain mutations result in JAK kinase hyperactivity: a naturally occurring mutation in this domain in mice renders Tyk2 catalytically inactive (Shaw et al., 2003) and studies in cell lines have also identified pseudokinase domain mutations that result in hypoactivity of TYK2 (Velazquez et al., 1995; Yeh et al., 2000).

The molecular mechanism by which the pseudokinase domain exerts a regulatory effect on JH1 activity has remained elusive, but a recent report suggests that the JAK2 pseudokinase domain may in fact be a genuine kinase (Ungureanu et al., 2011). Phosphorylation of the serine-523 and tyrosine-570 residues of JAK2 had previously been recognised as having a regulatory role, however the kinase responsible for these phosphorylation events had not been identified. The authors demonstrate *in vitro* and in cells that it is the so-called "pseudokinase" domain of JAK2 that autophosphorylates these residues and that this is critical for maintaining inactivity in unstimulated cells (Ungureanu et al., 2011). Additionally, mutations associated with MPD were shown to prevent this regulatory phosphorylation, thus explaining the mechanism by which they cause constitutive kinase domain activation (Ungureanu et al., 2011).

So far, little is known about the structural interactions between domains, however high-resolution crystal structures of all four JAK family kinase domains have been published and have improved our understanding of the function of the catalytic domain (Boggon et al., 2005; Chrencik et

al., 2010; Korniski et al., 2010; Williams et al., 2009). For example, in contrast to previous studies of cytokine signalling in cell lines (Gauzzi et al., 1996) and to the function of the JAK2 and JAK3 catalytic domains, characterisation of the purified TYK2 kinase domain revealed that phosphorylation of both activation loop tyrosines at position 1054 and 1055 (pY1054 and pY1055, respectively) does not achieve greater catalytic activity than pY1054 alone (Korniski et al., 2010). However, it is noteworthy that the kinetic properties determined through this structural work may not reflect the function of the full-length protein, as there is evidence to suggest a striking (i.e. 1000-fold) difference in the potencies of several compounds at inhibiting the activity of full-length TYK2 compared to that of the purified kinase domain (Emmons et al., 2012).

There are no published crystal structures for the disease-associated nonsynonymous TYK2 variants, however there is evidence to suggest that the rs34536443 SNP may affect TYK2 function (further detail in Chapter 3). The minor allele of this nsSNP substitutes a highly conserved proline residue in the kinase domain (Figure 1.2.) and there is evidence to suggest that this has consequences for TYK2 signalling, as had been previously predicted by protein modeling (Couturier et al., 2011; Kaminker et al., 2007).

1.8 TYK2 function: cytokine signalling, non-canonical roles and disease

1.8.1. TYK2 deficiency

TYK2 was first characterised as being critical for type 1 IFN responsiveness in a fibroblast cell line found to be deficient for the protein (Velazquez et al., 1992) but it is now known to signal downstream of multiple cytokines, including IL-4, -6, -10, -12, -13 and -23, for example (Figure 1.3.). Thus, TYK2 plays a role in both the innate and adaptive system. TYK2 and the other members of the JAK enzyme family mediate signal transduction for cytokine receptors, which lack intrinsic enzyme activity. Cytokine-induced auto- and trans-phosphorylation of JAKs activates them to phosphorylate signal transducer and activator of transcription (STAT) proteins, which function as

both cytoplasmic signalling molecules and also transcription factors (Darnell et al., 1994) (Figure 1.3.).

Whilst members of the JAK family do have overlapping roles, their more unique functions are evident from the study of humans with specific deficiencies and from work with knockout (KO) mice. There have been two reported cases of human TYK2 deficiency, however full immunophenotyping has only been performed on cells from one of these patients (Kilic et al., 2012; Minegishi et al., 2006). Both individuals had homozygous deletions in *TYK2*, which caused frameshifts and thus the presence of a premature stop codon (Kilic et al., 2012; Minegishi et al., 2006). No TYK2 protein was detected in lymphocytes from either patient. Both TYK2-deficient patients were severely immunodeficient with enhanced susceptibility to various mycobacterial and viral infections (Kilic et al., 2012; Minegishi et al., 2006). Only one TYK2 deficient patient showed clinical features resembling hyper-immunoglobulin E syndrome (HIES), which is typically characterised by atopic dermatitis, recurrent staphylococcal infections of the skin and lungs and high serum IgE levels (Minegishi et al., 2006). However, this patient did not display the developmental defects characteristic of HIES and also presented with more pronounced immunodeficiency than that typically observed in classical HIES caused by mutations in STAT3 (Minegishi, 2009; Minegishi et al., 2006).

Correlating with the extensive clinical immunodeficiency, patient-derived cells displayed severely impaired type 1 IFN and IL-12 signalling, and also reduced IL-6, -10 and -23 signalling (Minegishi et al., 2006). In line with the defect in IL-12 signalling, patient cells showed impaired differentiation of Th1 cells (Minegishi et al., 2006). Consistent with this, *Tyk2*-knockout (*Tyk2*-KO) mice show impaired differentiation of Th1 and Th17 cells and also reduced, but not abolished, responses to type 1 IFNs, IL-12 and IL-23 (Ishizaki et al., 2011a; Karaghiosoff et al., 2000; Shaw et al., 2003; Shimoda et al., 2000). Mirroring the reduction in IL-10 signalling observed in cells derived from the TYK2 deficiency patient, cells from *Tyk2*-KO mice also display reduced responses to stimulation with IL-10 (Shaw et al., 2006), however IL-6 signalling is not impaired in *Tyk2*-KO mice (Shimoda et al., 2000) suggesting some species differences in the relative importance of TYK2 function across cytokine pathways.

Despite these apparent discrepancies in the degree to which specific cytokine signalling pathways are affected by TYK2 deficiency between species, the clinical presentation of TYK2 deficiency patients was consistent with findings in mice. For example, *Tyk2*-KO mice show impaired anti-viral responses (Karaghiosoff et al., 2000; Prchal-Murphy et al., 2012; Strobl et al., 2005). Intriguingly, the role of TYK2 in infection is evident not only from the increased susceptibility to microbes and viruses observed in the deficient patients and mice, but also by the targeting of TYK2 by viruses as an immune evasion mechanism (Taylor and Mossman, 2013). For example, human metapneumovirus reduces TYK2 expression at both the messenger ribonucleic acid (mRNA) and protein levels (Ren et al., 2011) and West Nile virus, Myxoma virus and EBV all inhibit TYK2 phosphorylation, thus preventing type 1 IFN-mediated anti-viral signalling (Geiger and Martin, 2006; Guo et al., 2005; Wang et al., 2009).

Mice deficient in Tyk2, either due to genetic manipulation or naturally occurring deleterious mutations, also show enhanced susceptibility to bacterial and parasitic infection (Aizu et al., 2006; Shaw et al., 2003), and show resistance to lipopolysaccharide-induced shock (Kamezaki et al., 2004; Karaghiosoff et al., 2003), emphasizing the role of Tyk2 not only in adaptive but also in innate immune responses. The phenotype of *Tyk2*-KO mice is less severe than that of mice deficient in the other JAKs, with no developmental abnormalities being reported (Karaghiosoff et al., 2000; Shimoda et al., 2000). Knockout of *Jak1* or *Jak2* results in perinatal or embryonic lethality, respectively, as these JAKs mediate signalling of growth factors in addition to cytokines (Parganas et al., 1998; Rodig et al., 1998). JAK3 deficiency is not lethal but causes SCID in both mice and humans, demonstrating its critical role in lymphocyte development and function (Nosaka et al., 1995).

In TYK2 deficiency, the respective signalling pathways are impaired, but cytokine responses are not completely abolished, suggesting some degree of compensation by other JAKs. The extent of redundancy between TYK2 and other JAKs remains unclear and there are evidently species-specific differences. Studies have suggested that TYK2 may play an inferior role in IFN-α signalling, requiring the presence of, and functioning downstream of, activated JAK1 (Muller et al., 1993). However, a full characterisation of the requirement for TYK2 by various cytokines has not been performed.

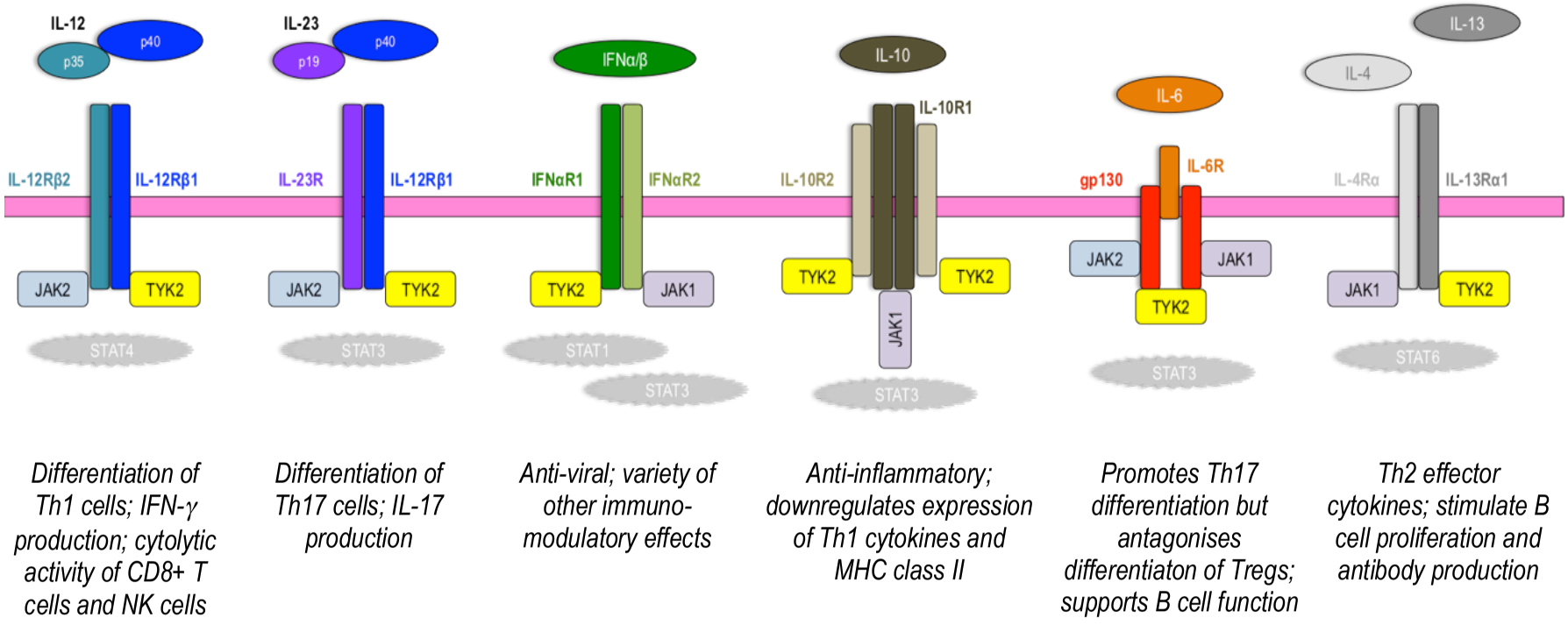


Figure 1.3. TYK2-mediated cytokine signalling pathways and their roles in the immune system
TYK2 is involved in the signal transduction of multiple cytokines with varied roles in the immune system. Binding of cytokine to its cognate receptor leads to activation of intra-cellular JAK-STAT signalling and downstream changes in gene expression. Diagram modified from (Watford and O'Shea, 2006).

1.8.2. TYK2 and mouse models of autoimmune disease

Prior to the genetic association of the *TYK2* region with human autoimmunity, functional studies in animal models had implicated Tyk2 in autoimmune disease. Mice deficient in Tyk2 show complete resistance to disease in models of collagen-induced arthritis, PS and colitis, correlating with defects in Th1 and Th17 cytokine production (Ishizaki et al., 2011a; Ishizaki et al., 2011b; Ortmann et al., 2001). Tyk2-deficient mice also show complete resistance to EAE induced by a truncated MOG peptide emulsified in complete Freund's adjuvant (CFA) (Oyamada et al., 2009; Spach et al., 2009). These mice show a complete lack of immune cell infiltration into the CNS and a reduced frequency of MOG-specific Th1 cells in the periphery, consistent with a reduction in IL-12-driven Th1 cell differentiation (Oyamada et al., 2009). Although this study highlights a critical role of Tyk2 in pathogenic CD4+ T cell responses, the use of CFA in EAE induction specifically skews towards a Th1 and Th17 immune response (Tigno-Aranjuez et al., 2009) and thus does not provide an ideal model for studying the potential role of Tyk2 in other immune cell subsets that may also be relevant in MS.

1.8.3. Regulation of TYK2 function

Given that TYK2 is implicated in infection and autoimmunity, it follows that its activity is heavily regulated not only by the action of its own pseudokinase domain but also through the action of other molecules, including phosphatases such as CD45 and protein tyrosine phosphatase 1B (Irie-Sasaki et al., 2001; Myers et al., 2001) and the suppressor of cytokine signalling (SOCS) proteins. SOCS proteins are rapidly induced by cytokine signalling to form a classical negative feedback loop (Shuai and Liu, 2003). SOCS3 has recently been shown to directly bind a specific motif in the kinase domain of JAK1, JAK2 and TYK2 (but not JAK3), having a direct inhibitory effect rather than competing with adenosine triphosphate (ATP) or acting as a pseudo-substrate (Babon et al., 2012). Similarly, SOCS1 has been shown to directly bind TYK2, specifically associating via pY1054 and

pY1055 and consequently inhibiting type 1 IFN-stimulated STAT phosphorylation as well as inducing degradation of IFNAR1 (Piganis et al., 2011). The kinase inhibitory region of Socs1 has also been shown to ameliorate EAE through effects on pathogenic CD4+ T cell responses (Jager et al., 2011), echoing the protective effect of *Tyk2*-KO in models of autoimmunity.

1.9 From canonical JAK-STAT signalling to other roles of TYK2

Apart from classical JAK-STAT signalling, TYK2 has been demonstrated to phosphorylate additional substrates upon cytokine-induced activation, for example TYK2 phosphorylates protein kinase D2, which results in the ubiquitination and degradation of IFNAR1 (Zheng et al., 2011). Cell line work has also shown that whilst STAT phosphorylation requires TYK2 and JAK activity, TYK2 alone is sufficient for type 1 IFN-induced activation of nuclear factor kappa-B (NFκB), although this requires further confirmation in primary cells (Hervas-Stubbs et al., 2011; Yang et al., 2005a)).

In addition to mediating signal transduction in its membrane-proximal location, TYK2 has been found to localise to the nucleus (Ragimbeau et al., 2001). A recent publication has demonstrated that type 1 IFN-induced TYK2 activity in the nucleus is associated with changes in histone methylation and acetylation in the promoter region of a typical type 1 IFN target gene (Ahmed et al., 2013). Further supporting a role for JAKs in the nucleus, JAK2 has been shown to phosphorylate histones (Dawson et al., 2009). However, the relative importance of the function of TYK2 at the plasma membrane as opposed to in the nucleus remains to be elucidated.

Several alternative TYK2 binding partners have also been identified. For example, TYK2 has been immunoprecipitated with the DNA repair enzyme Ku80, which was originally identified as the target of autoantibodies in SLE, although the functional consequence of this interaction is unknown (Adam et al., 2000; Reeves, 1992). TYK2 also interacts with the JAK and microtubule interacting protein, JAKMIP-1, which is expressed in activated immune cells, and intriguingly is found at high levels in neurons, providing the first indirect implication for a putative role of TYK2 in these cells (Libri et al., 2008; Steindler et al., 2004). Subsequent to these studies, TYK2 has been

shown to be upstream of STAT3 in the signalling pathway mediating amyloid-β-induced neuronal apoptosis and neurons from *Tyk2*-KO mice were resistant to cell death induced by this pathway (Wan et al., 2010). As well as an involvement in neurodegeneration, TYK2 may also play a role in negatively regulating CNS myelin repair as *Tyk2*-KO mice show enhanced repair compared to wild-type mice in the Theiler's murine encephalomyelitis virus model of MS-like disease (Bieber et al., 2010).

1.10 "JAKinibs" and TYK2 as a potential therapeutic target in autoimmune disease

The development of JAK inhibitors, or "JAKinibs", as a new therapeutic modality is currently an area of intense investigation (Kontzias et al., 2012). Since mutations in JAK2 were identified as a cause of MPDs, the JAKs became therapeutic targets with two JAKinibs having now been approved for clinical use (O'Shea et al., 2013b). The JAK1/2 inhibitor ruxolitinib is now an approved therapy for myelofibrosis and a JAK3/1 inhibitor (tofacitinib) has been approved for use in methotrexate-resistant arthritis (O'Shea et al., 2013a). Amongst others, these JAKinibs are showing promise in pre-clinical models and also trials in patients with a range of autoimmune diseases, including SLE, PS and IBD (as reviewed by (Kontzias et al., 2012)).

The full side effect profile of these drugs remains to be characterised and there is evidence from various trials that JAKinibs increase susceptibility to infection, however the extent to which this immunosuppression may also contribute to cancer risk remains to be determined. Both of the JAKinibs currently in clinical use are not highly selective for specific JAKs and the ongoing development of selective inhibitors may help to reduce unwanted adverse effects (Kontzias et al., 2012). Another benefit of increased drug specificity could be the more precise targeting of disease-relevant cytokine signalling pathways. For example, the use of tofacitinib to inhibit JAK3/1 has been reported to exacerbate EAE (Yoshida et al., 2012) whilst in contrast, selective inhibition of TYK2 causes a dose-dependent amelioration of EAE (Sareum, 2013). This is consistent with data from *Tyk2*-KO mice (Oyamada et al., 2009; Spach et al., 2009) and also the purported reduced kinase

activity of the protective TYK2 variant associated with MS and other autoimmune diseases (i.e. the variant encoded by the minor allele of the rs34536443 SNP) (Couturier et al., 2011).

The TYK2 kinase domain was the last of the JAKs to be crystallised, however the recent publication of a crystal structure (Chrencik et al., 2010; Korniski et al., 2010) has fuelled the development of TYK2-specific inhibitors. Liang and colleagues have recently reported their progress in generating potent and orally bioavailable TYK2-specific inhibitory compounds, and note that several patent applications have been made for similar work (Liang et al., 2013a; Liang et al., 2013b). Oral treatment of mice with the lead TYK2-selective compound was shown to inhibit IL-12 signalling *in vivo* (Liang et al., 2013b) and TYK2 inhibition is showing promise in pre-clinical trials in several models of autoimmune disease (Norman, 2012; Sareum, 2013). The translational potential of pharmacologically targeting TYK2 will be maximised through an improved understanding of the role of TYK2 signalling in different autoimmune diseases. As the *TYK2* gene region is associated with multiple autoimmune diseases, elucidating the functional consequences of these genetic variants will be valuable in fuelling such progress.

1.11 Aims of this D.Phil thesis

As detailed in Table 1.1, multiple autoimmune conditions are associated with each SNP in the *TYK2* region and in most cases each disease is associated with only one SNP, with the exception of PS and IBD for which there are two such associations. The fact that there are both distinct and overlapping associations across the four SNPs suggests that they exert different functional effects; if the consequences of each SNP was the same, then all SNPs would be associated across all of the diseases with the same direction of association. This would be inconsistent with the genetic data which shows that some diseases are only associated with one SNP in the region and also that the minor allele of the rs12720356 SNP is associated with risk for IBD but with protection against PS (Anderson et al., 2011; Franke et al., 2010; Strange et al., 2010; Tsoi et al., 2012). It is thus likely that each SNP has a distinct functional effect and although *TYK2* is identified as the candidate gene

for each of these SNPs, the molecular mechanisms underpinning these genetic associations may not be restricted to effects on TYK2.

The overall aim of the presented work was to investigate the functional effects of autoimmune disease-associated variation in the *TYK2* gene region. As three of the four focal variants are nsSNPs (rs2304256, rs12720356 and rs34536443), they may affect the kinase activity of TYK2 and thus influence cytokine signalling. Indeed there is some evidence to suggest that this may be the case for the rs34536443 SNP (Couturier et al., 2011), however this required further thorough study in a larger cohort. Following initial analyses of variant kinase activity and cytokine signalling in a cell line (Chapter 3), samples from a pre-selected cohort of genotyped individuals were used to investigate the effect of genotype on TYK2-mediated signalling in primary immune cell subsets (Chapters 4 and 5). In addition, these three nsSNPs and the intronic rs11879191 SNP (located in the *CDC37* gene) may exert regulatory effects on the expression of the *TYK2* gene, or other genes in the region. There is no existing evidence for these SNPs affecting *TYK2* expression, however some expression quantitative trait loci (eQTL) data suggests that the rs2304256, rs12720356 and rs11879191 SNPs may affect expression of genes in the region in certain cell types (Brown et al., 2013; Fairfax et al., 2012; Veyrieras et al., 2008; Wallace et al., 2012). Some of this work is confined to the use of cell lines, which may not reflect the regulation of gene expression in primary human immune cell subsets, and thus further investigation was warranted (Chapter 6).

Investigating the functional effects of autoimmune disease-associated variation in the *TYK2* gene region in primary human cells will improve our understanding of the molecular mechanisms underpinning shared genetics and also clarify potential disease-specific pathways, thus informing therapeutic development.

2. Materials and methods

2.1 Cell culture

U1A cells (a kind gift from Dr Ana Costa-Pereira, Imperial College London) were grown in 10% carbon dioxide (CO_2) at 37°C in Dulbecco's Modified Eagle Medium (DMEM; PAA Laboratories GmbH, UK) supplemented with 20% sterile filtered, heat-inactivated fetal calf serum (FCS; Sigma-Aldrich, UK) and penicillin/streptomycin (PAA Laboratories GmbH, UK; final concentrations of 1 U/ml and 1 μg/ml, respectively). HT-1080 cells (American Type Culture Collection, United States) were grown in 5% CO_2 at 37°C in the same media, but containing 10% FCS. Primary cells were grown in 5% CO_2 at 37°C in Roswell Park Memorial Institute medium (RPMI; Sigma-Aldrich, UK) supplemented with 10% FCS, 2mM glutamine (Sigma-Aldrich, UK) and penicillin/streptomycin. For T cells, β-mercaptoethanol (final concentration of 50 nM; GIBCO, Invitrogen, Life Technologies, UK) was also added to the culture media.

2.2 Cloning TYK2 constructs encoding nsSNP variants

In collaboration with Dr Christian Siebold (Wellcome Trust Centre for Human Genetics (WTCHG), University of Oxford), sequences were designed in order to express the wild-type *TYK2* gene (made by Eurofins MWG Operon, Germany), which was cloned into the pHLsec vector (a kind gift from Dr Christian Siebold) using the restriction sites *Nhe*I and *Kpn*I (both from New England Biolabs (NEB), UK) and the Rapid Ligation Kit (Roche Diagnostics, UK). Several rounds of site-directed mutagenesis were then performed using the QuikChange® II XL site-directed mutagenesis kit (Stratagene, UK) in order to generate separate constructs for expression of TYK2 variants with each of the following amino acid substitutions: valine-362-phenylalanine, isoleucine-684-serine and proline-1104-alanine. An additional construct with both of the latter substitutions was also generated (see Chapter 3, section 3.3.). Construct sequences were verified before further use (Weatherall

Institute of Molecular Medicine (WIMM) Sequencing Lab). See table below for mutagenic primers (designed according to manufacturer's guidelines) and thermal cycling parameters used.

Mutation	Mutagenic primers (5' to 3')
Proline-1104-alanine	GACTCGTCACAATCCGCTCCCACCAAGTTCT
	AGAACTTGGTGGGAGCGGATTGTGACGAGTC
Valine-362-phenylalanine	CGAAAGCCCACAAGGCATTCGGGCAGC
	GCTGCCCGAATGCCTTGTGGGCTTTCG
Isoleucine-684-serine	CGCGGTCCTGAGAACAGTATGGTGACCGAATAC
	GTATTCGGTCACCATACTGTTCTCAGGACCGCG

Segment	Cycles	Temperature (°C)	Time
1	1	95	1 minute
2	18	95	50 seconds
		60	50 seconds
		68	1 – 2.5 minutes per kilobase of plasmid length
3	1	68	7 minutes

2.3 Transfections and cell line stimulations

Cells were seeded at 0.2-0.8 x 10^5 in 24-well plates 24 hours (24h) prior to transfection using the lipid-based reagent TransIT®-2020 (Mirus Bio LCC, USA). Just prior to transfection, culture media was replaced with 500 µl fresh media. For each well, 0.5 µg DNA was mixed with 50 µl serum-free media (Opti-MEM I Red with GlutaMAX I, Invitrogen, Life Technologies, UK) prior to the addition of 1.5 µl TransIT®-2020 reagent. Additional fresh media was added 5h later for optimal cell growth and survival. Cells were monitored and fed before being harvested 48h post-transfection for stimulation and lysis. Cells were stimulated in 500 µl of fresh, warm media with 1000 U/ml IFN-β (PeproTech EC Ltd, UK) for 15 minutes at 37°C. Subsequently, the media was removed and cells were lysed according to the protocol below.

2.3 Protein lysates and western blotting

Cells were lysed in the following buffer, a stock of which was kept at room temperature: 20 mM HEPES pH 7.5, 100 mM sodium chloride, 0.05% Triton X-100 and 1mM dithiothreitol (all

Sigma-Aldrich, UK). Lysis buffer was supplemented with protease inhibitors (Roche Mini Complete from Roche Diagnostics, UK) immediately prior to use, along with 1 mM sodium orthovanadate (NEB, UK) and 0.1% phosphatase inhibitor cocktail 2 (Sigma-Aldrich, UK).

Following removal of culture media and a cold phosphate buffered saline (PBS) wash, 70 µl of filter-sterilised lysis buffer was added per well of adherent cells (cultured in 24-well plates). Plates were then shaken for 20 minutes on an orbital plate shaker (Thermo MiniMix, ThermoScientific, UK). Lysates were then stored at minus 20°C until use and aliquoted to avoid repeated freeze/thaw cycles. Protein concentration was determined using the Nanodrop ND-1000 (ThermoScientific, UK), which computes concentration by the absorbance at 280 nm. For western blotting, polyacrylamide gels and buffers were made in-house according to the table below. Lysates were diluted in 2X Laemmli buffer (details below) and boiled for 5 minutes at 95°C before loading onto the gel.

Component	**10% Resolving gel**	**Stacking gel**
MilliQ water	5.9 ml	6.2 ml
Acrylamide (Flowgen Bioscience)	5.0 ml	1.3 ml
Tris HCl pH 8.8 (Sigma-Aldrich, UK)	2.8 ml	2.5 ml
10% Sodium dodecyl sulphate (SDS; Sigma-Aldrich, UK)	150 µl	100 µl
10% Ammonium persulphate (APS; Sigma-Aldrich, UK)	150 µl	200 µl
Tetramethylethylenediamine (TEMED; Sigma-Aldrich, UK)	6 µl	12 µl

5X Running buffer	**10X Blotting buffer**
25 mM Tris (Sigma-Aldrich, UK) 250 mM Glycine (Fisher Scientific, UK) 0.1% SDS	250 mM Tris 1.9 M Glycine
2X Laemmli buffer	
4% SDS 20% UltraPure glycerol (Invitrogen, Life Technologies, UK) 125 mM Tris HCl pH 6.8 (Sigma-Aldrich, UK) 5% β-Mercaptoethanol 10% Bromophenol blue (Bio-Rad Laboratories, UK	
Blocking solution	
Odyssey Blocking buffer (LI-COR Biotechnology, UK) **OR** Bovine serum albumin (BSA) in TBST 1X TBS (10X = 200 mM Tris, 1.5 M NaCl) 0.1% Tween (Sigma-Aldrich, UK) 5% BSA (Sigma-Aldrich, UK)	

Gels were run and proteins blotted using the Mini Trans-Blot® system (Bio-Rad Laboratories, UK) and the See-Blue2 pre-stained protein marker (Invitrogen, Life Technologies, UK) was run as an indicator of protein size. Following blotting, the nitrocellulose membrane (GE Healthcare Life Sciences, UK) was immersed in blocking solution for one hour. The same blocking buffer was then used for primary antibody incubations. The table below lists the antibodies used. Following four 10-minute washes in PBS containing 0.1% tween, the membrane was incubated with the appropriate IR-Dye®-conjugated secondary antibody in the dark for 30 minutes in PBS containing 0.1% tween (Sigma-Aldrich, UK) and 0.01% SDS (Fisher BioReagents, UK). Following three 10-minute washes with PBS containing 0.1% tween and two washes with PBS alone, images of the membrane were obtained using the Odyssey Infra-Red Imaging System (Software version 2.0; LI-COR Biosciences, UK). Equal-sized gates were drawn around bands of interest using the software, which provided values for the fluorescence intensity of selected bands in acquired images.

Antigen	Species	Supplier	Conc^{n} used	Blocking solution
Rabbit IgG (secondary antibody; IR-Dye® 680LT conjugate)	Goat	LI-COR Biotechnology, UK	1:10,000	BSA
Mouse IgG (secondary antibody; IR-Dye® 800CW conjugate)	Goat	LI-COR Biotechnology, UK	1:10,000	BSA
Human actin	Mouse	MilliPore, UK	1:5000	BSA
Human pTYK2	Rabbit	NEB, UK	1:1000	BSA
Human pSTAT1	Mouse	B.D. Biosciences, UK	1:250	Odyssey
Human STAT1	Mouse	B.D. Biosciences, UK	1:500	BSA
Human TYK2	Mouse	BD Biosciences, UK	1:1000	BSA

2.4 Donor recruitment

Blood samples were donated by consenting healthy individuals under ethical approval from the National Research Ethics Service, Oxfordshire (REC B) for the project entitled “Studying the molecular mechanisms of action of genes with relevance to multiple sclerosis in healthy controls”; REC ref 10/H0605/5. The Vacuette® Safety Blood Collection set and heparin tubes (Greiner Bio-One, UK) were used by a qualified and experienced phlebotomist. Donors were recruited via posters advertising the study and also in collaboration with the Oxford BioBank (Professor Fredrik Karpe, Research Nurse Jane Cheeseman and Bioresource Co-ordinator Dr Matt Neville; Oxford Centre for

Diabetes, Endocrinology and Metabolism, University of Oxford (OCDEM)). Donors were of white European ethnicity and self-reported non-autoimmune/chronic inflammatory disease status.

All described flow cytometry work with donor blood samples was performed with the help and under the guidance of Dr. Calliope Dendrou. For RNA level data, donor sample processing was performed alongside Dr. Calliope Dendrou, Dr. Thomas Barber, Dr Adam Gregory and Mr Christos Krastev.

2.5 Genomic DNA extraction from human blood

For the non-Oxford BioBank donors, the QiaAmp DNA Blood midi kit (Qiagen, UK) was used to isolate DNA from 1-2 ml of blood, following the manufacturer's protocol. The quality and concentration of DNA obtained was determined using the Nanodrop ND-1000 (ThermoScientific, UK), with typical $A_{260/280}$ absorbance ratios of 1.7-2.1 and concentrations of >100 ng/µl. The Oxford BioBank stores frozen, dried-down genomic DNA from its donors.

2.6 SNP genotyping of human donor DNA

TaqMan® SNP assays (Applied Biosystems®, Life Technologies, UK) were used to genotype donor DNA according to the manufacturer's instructions. SNP genotyping assays were diluted to a 20X working stock with 1X TE buffer (10 mM Tris-HCl, 1 mM ethylenediaminetetraacetic acid (EDTA), pH8, DNase-free, sterile-filtered water), vortexed and aliquoted for storage at minus 20°C. Reactions were performed according to manufacturer's instructions using MicroAmp™ Optical 384-well Reaction Plates (DNA, RNA, RNase-free; Applied Biosystems®, Life Technologies, UK) and the Applied Biosystems® 7900HT Real-Time PCR System. On each plate that was run, two wells were allocated as no-template controls, containing DNase-free water instead of DNA. The section below provides details of the reaction components, thermal cycling parameters and SNP assays performed. All genotyping data were double scored by a

second operator to minimise error. For genotyping of Oxford BioBank donors, Dr Matt Neville used the same SNP assays according to the manufacturer's protocol for dried-down DNA.

SNP genotyping assays reaction components (µl per well)	
2.25	Template DNA (10 ng)
2.5	TaqMan® Universal PCR Master Mix (2X), No AmpErase® UNG (Applied Biosystems®, Life Technologies, UK)
0.25	20X working stock of SNP Genotyping Assay

SNP I.D.	SNP assay
rs34536443	C_60866522_10
rs2304256	C_25473911_10
rs12720356	C_34042925_10
rs11879191	C_1931097_10

Segment	Cycles	Temperature (°C)	Time
1 AmpliTaq Gold® enzyme activation	1	95	10 minutes
2 PCR	40	92	15 seconds
		60	1 minute

TaqMan® SNP Genotyping Assays contain two primers to specifically amplify the sequence of interest and two TaqMan® MGB probes, which have either the VIC® or FAM™ 5'-reporter dye. At the 3'-end, the non-fluorescent quencher is designed to provide greater sensitivity than dye quenchers, such as TAMRA®. According to the manufacturer, the 3'-minor groove binder (MGB) increases the melting temperature (T_m) of probes allowing them to be of shorter length. This enhances the efficiency of probe hybridization and allele discrimination by giving greater T_m differences between matched and mismatched probes. This minimises the chance of non-specific fluorescence emission.

2.7 Peripheral blood mononuclear cell (PBMC) isolation

Blood tubes were left rotating for 20 minutes to ensure thorough mixing with the anti-coagulant before subsequent dilution of blood at a ratio of 1:1 with sterile PBS (OXOID Ltd, UK) at room temperature. Diluted blood (25 ml) was then carefully layered onto 15 ml of Lymphoprep™ (Axis-Shield, UK). Density gradient centrifugation was performed at 800 x *g* for 20 minutes at room temperature with the brake set to "off" (Multifuge Heraeus centrifuge). Following centrifugation, the white layer of PBMCs was removed and washed twice with 1X ice-cold sterile PBS pH 7.4

supplemented with 1% heat-inactivated, filtered human male AB serum (PAA Laboratories GmbH, UK) and centrifuged at 300 x *g* for 10 min at 4°C using the "full" brake setting. PBMCs were subsequently kept ice-cold and the cell count and viability were estimated using trypan blue (Sigma-Aldrich, UK) on a haemocytometer. Typically, viability was >95% and cell counts were 1-2 x 10^6 cells/ml.

2.8 Isolation of polymorphonuclear granulocytes

Granulocytes were isolated from the erythrocyte/granulocyte fraction of three Lymphoprep™ gradients per donor by removal of the plasma and Lymphoprep phases. The remaining cell fraction was then lysed via resuspension in 50 ml of filter-sterilised lysis buffer containing 150 mM NH_4Cl, 1 mM $KHCO_3$ and 0.1 mM EDTA (pH 7.2-7.4; all Sigma-Aldrich, UK) and incubated at room temperature for 20 minutes. Cells were pelleted by centrifugation at 400 x *g* for 10 minutes, and the supernatant was discarded. The pellet was resuspended again in 20 ml 1 X RBC Lysis Buffer and incubated for a further 20 minutes at room temperature, following which the cells were pelleted by centrifugation at 400 x *g* for 10 minutes and supernatant was discarded, yielding polymorphonuclear leukocytes which were lysed for RNA extraction.

2.9 Recombinant human cytokines

The table below lists details of the recombinant human cytokines used.

Cytokine	Supplier	Catalogue number
IL-2	PeproTech EC Ltd, UK	200-02
IL-6	PeproTech EC Ltd, UK	200-06
IL-10	PeproTech EC Ltd, UK	200-10
IL-12 (p70)	PeproTech EC Ltd, UK	200-12
IL-13	Miltenyi Biotec, UK	130-093-953
IL-23 (p40+p19)	Miltenyi Biotec, UK	130-095-757
IFN-α	PeproTech EC Ltd, UK	200-02A
IFN-β	PeproTech EC Ltd, UK	300-02BC
IFN-γ	PeproTech EC Ltd, UK	300-02

2.10 Antibodies for flow cytometry

Antibody concentrations were determined from manufacturer's recommendations and volumes used per test were determined by titration. The table below lists details for all antibodies used.

Primary antibodies

Antigen (all human)	**Fluorochrome**	**Catalogue number**	**Clone** (or species and isotype for polyclonals)	**Supplier**
CD14	Peridinin-chlorophyll-protein complex-cyanine 5.5 (PerCP-Cy5.5)	325622	HCD14	BioLegend, UK
CD19	Alexa Fluor® 700	56-0199-42	HIB19	eBioscience, UK
CD19	Allophycocyanin (APC)	17-0199-73	HIB19	eBioscience, UK
CD4	APC	300514	RPA-T4	BioLegend, UK
CD4	Phycoerythrin (PE)	300508	RPA-T4	BioLegend, UK
CD4	Alexa Fluor® 700	300526	RPA-T4	BioLegend, UK
CD4	PE-Cy7	25-0049-42	RPA-T4	eBioscience, UK
CD45RA	Pacific Blue	304123	HI100	BioLegend, UK
CD8	Alexa Fluor® 700	56-0086-42	OKT8	eBioscience, UK
CD8	Fluorescein isothiocyanate (FITC)	11-0086-42	OKT8	eBioscience, UK
CD8	PE	12-0086-42	OKT8	eBioscience, UK
ICAM1	APC	322712	HCD54	BioLegend, UK
ICAM3	PE	330006	CBR-IC3/1	BioLegend, UK
ICAM4	Unconjugated	AF7179	Sheep polyclonal IgG	R&D Systems, UK
IFNAR1	Unconjugated	AP8550c	Rabbit polyclonal IgG	Abgent Europe, UK
IFN-γ	APC	506510	B27	BioLegend, UK
IL-10R	PE	556013	3F9	BD Biosciences, UK
IL-12 Rβ2	APC	FAB1959A	305719	R&D Systems, UK
IL-12Rβ1	FITC	FAB839F	69310	R&D Systems, UK
IL-13R	PE	FAB1462P	419718	R&D Systems, UK
IL-17	FITC	53-7179-42	N49-653	eBioscience, UK
IL-23R	PE	FAB14001P	218213	R&D Systems, UK
IL-4	PE	500808	MP4-25D2	BioLegend, UK
IL-6R	APC	562090	M5	BD Biosciences, UK
pSTAT1	Alexa Fluor® 647	612597	4a	BD Biosciences, UK
pSTAT2	FITC	IC2890F	Rabbit polyclonal IgG	R&D Systems, UK
pSTAT3	Alexa Fluor® 488	557814	4/P-STAT3	BD Biosciences, UK
pSTAT4	Alexa Fluor® 647	558137	38/p-Stat4	BD Biosciences, UK
pSTAT5	Alexa Fluor® 488	612598	47/Stat5(pY694)	BD Biosciences, UK
pSTAT6	Alexa Fluor® 647	612601	18/P-Stat6	BD Biosciences, UK
pTYK2	Unconjugated	9321S	Rabbit polyclonal IgG	NEB, UK
STAT1	PE	558537	1/Stat1	BD Biosciences, UK
STAT3	APC	560392	M59-50	BD Biosciences, UK

Secondary antibodies

Antibody	Fluorochrome	Cat. No.	Supplier
Goat anti-rabbit	Alexa Fluor® 488	A11008	Invitrogen, Life Technologies, UK
Goat anti-sheep	Alexa Fluor® 488	A11015	Invitrogen, Life Technologies, UK

Isotype controls

Antibody	Fluorochrome	Cat. No.	Supplier
Mouse IgG1	PE	IC002P	R&D Systems, UK
Mouse IgG1	FITC	551954	BD Biosciences, UK
Mouse IgG1	APC	IC002A	R&D Systems, UK
Mouse IgG2b	PE	IC0041P	R&D Systems, UK
Rat IgG2a	PE	551799	BD Biosciences, UK
Rabbit polyclonal	Unconjugated	ab37415	AbCam, UK
Sheep polyclonal IgG	Unconjugated	5-001-A	R&D Systems, UK
Rat IgG2a	Unconjugated	553927	BD Biosciences, UK

2.11 Primary human cell activations for RNA isolation and expression time courses

Monocytes were positively selected using CD14 microbeads (Miltenyi Biotec, UK) according to manufacturer's instructions. Isolated cells were then resuspended at 1 x 10^6 cells/ml in RPMI containing 10% FCS and cultured with 500 U/ml IFN-γ for 48h. Cells from the CD14-negative lymphocyte fraction were resuspended at 1 x 10^6 cells/ml in RPMI containing 10% FCS and cultured with Human T-Activator Dynabeads (Invitrogen, Life Technologies, UK) at a bead:cell ratio of 1:1 and 30 U/ml IL-2 for the indicated times.

Following stimulation, cells were washed and lysed in preparation for RNA extraction, as detailed in section 2.14.

2.12 Flow cytometry staining protocols

2.12.1. Whole blood surface staining

For surface staining, 1 ug/ml mouse serum (Sigma-Aldrich, UK) or appropriate unconjugated isotype control antibody was added to 100 µl whole blood and incubated at room temperature for 15

minutes. Fluorescently conjugated staining antibodies were then added and the sample incubated at 4°C for 30 minutes. Subsequently, 2 ml of 1X BD FACS Lysing Solution (10X stock; BD Biosciences, UK) was added and cells incubated at room temperature for 10 minutes. Cells were then pelleted by centrifugation at 300 x *g* for 5 minutes before being washed twice with cold PBS containing 1% human AB serum. Cells were then resuspended in 300 µl of 1X BD Cell Fix (BD Biosciences, UK).

2.12.2. PBMC surface staining

Following PBMC isolation, cells were counted and resuspended at 5 x 10^6 per ml in RPMI. For surface staining, 1 µg/ml mouse serum IgG (Sigma-Aldrich, UK) or appropriate unconjugated isotype control antibody (see table in section 2.10) was added to 200 µl of resuspended PBMCs and incubated at room temperature for 15 minutes. Fluorescently conjugated staining antibodies were then added and the sample incubated at 4°C for 30 minutes. Cells were then washed twice by centrifugation at 300 x *g* for 5 minutes with cold PBS containing 1% human AB serum. Cells were then resuspended in 300 µl of 1X BD Cell Fix (BD Biosciences, UK).

2.12.3. PBMC stimulation and phosflow staining

Following PBMC isolation, cells were counted and resuspended at 5 x 10^6 per ml in RPMI. 5x10^5 cells were stimulated for the indicated times and cytokine concentrations at 37°C before being washed with cold PBS containing 1% AB serum. The BD Phosflow protocol was then followed in order the stain both surface and intracellular antigens using BD Cytofix (BD Biosciences, UK), Phosflow wash buffer and BD Perm Buffer III (BD Biosciences, UK).

2.13 Flow cytometry data collection and analysis

Fixed samples were run on a CyAn™ ADP Analyser flow cytometer using the Summit software (both Beckman Coulter, UK) for data collection. Data files were analysed using FlowJo software (version 9.6.2 (Tree Star, Inc., Ashland, OR, U.S.A.), GraphPad Prism® (version 8 for Mac; http://www.graphpad.com/prism/Prism.htm) and Microsoft Excel® (version 12.3.6 for Mac).

To control for day-to-day and machine-to-machine variation, mean fluorescence intensity (MFI) values were normalised using fluorescent beads (FluoroSpheres 6-Peak from Dako UK Ltd, UK) that were run after setting the voltages on each day on a specific machine. For each relevant fluorochrome (FITC, PE or APC), the MFI values for each fluorescence peak were plotted against the number of molecules of equivalent fluorochrome (MEF) provided by the manufacturer. The slope of the best fit line was then used as a normalisation coefficient to convert MFI values across different days/machines into standardised MEF units.

Outliers were identified for exclusion using the ROUT method on Prism® with Q = 0.1% to minimise the false discovery rate (Motulsky and Brown, 2006). Values of 0% positivity for phospho-STAT were also excluded from analyses. Where there were significant effects of age or gender on STAT phosphorylation, data were normalised as appropriate in order to minimise the impact of such confounding factors.

2.14 RNA extraction & cDNA synthesis

RNA extraction from human PBMCs was performed using the Qiagen RNeasy mini kit, according to manufacturer's instructions for the spin protocol. Briefly, PBMCs were centrifuged and the cell pellet lysed before addition of ethanol and subsequent loading into the RNA-binding column. Following several washes, RNA was eluted and the concentration determined using the Nanodrop ND-1000 (ThermoScientific, UK) before storage at minus 80°C. The QuantiTect Reverse Transcription kit (Qiagen, UK) was used to synthesise complementary DNA (cDNA; typically from 500-1,000 ng of RNA were used per sample per reaction) which was then stored at minus 20°C for subsequent use in real-time quantitative polymerase chain reactions (qPCRs).

2.15 Real-time quantitative PCR

Primers and probes for qPCR were designed using the Primer3 program (version 0.4.0; http://frodo.wi.mit.edu/primer3/) and based on the human gene sequences available through Ensembl (http://www.ensembl.org/) for genome build GRCh37. According to the guidelines provided by Applied Biosystems® (Life Technologies, UK) the PCR products should be of 50-150 nucleotides in length. Primers and probes should have a guanine/cytosine (GC) content of 20-80% and successions of four or more identical nucleotides should be avoided. The primers should have a melting temperature of 58-60°C and the five nucleotides at the 3'-end must not contain more than two guanine and/or cytosine nucleotides. Probes should have a melting temperature of 68-70°C and must contain fewer guanine than cytosine nucleotides. Probes labeled with FAM™ and TAMRA® at the 5'- and 3'-ends, respectively, were purified by high performance liquid chromatography and ordered, along with primers, from Eurofins MWG Operon, Germany. Alternatively, assays were ordered from Applied Biosystems®. All assays were designed to detect mature mRNA and to amplify sequences containing exon-exon boundaries to avoid detection of genomic DNA or pre-mRNA. Assay optimisation involved titration of primers and probes for self-designed kits. For all assays, cDNA titrations were performed in order to plot a standard curve of the threshold cycle (C_t) values obtained for a range of cDNA concentrations, enabling calculation of assay efficiency using an online calculator (Agilent Technologies, USA; http://www.genomics.agilent.com/biocalculators/calcSlopeEfficiency.jsp). Assays were optimised to an efficiency of ≥90%. The table below shows the details of the Applied Biosystems® assays and primer-probe sets used.

	Product code
TYK2	Hs01105963_m1
ICAM1	Hs00164932_m1
ICAM4	Hs00169941_m1
ICAM5	Hs00959192_m1
FDX1L	Hs00365407_g1
ZGLP1	Hs00913460_g1
S1PR5	Hs00924881_m1
RAVER1	Hs00913477_m1
KEAP1	Hs00202227_m1

	Sequences (f / r – forward/reverse primer; p - probe)	**primer conc^{n} (nM)**	**probe conc^{n} (nM)**
CDC37	f - GAAGTGCTTCGATGTGAAGG r - TTAGAGTTGGGGACCCAGAG p - AAGATGGACCCCACCGACGC	200	200
PDE4A	f - CCAAATCACAGGGTTGAAAA r - TTCAGGTTCTCCAGTTCTTGG p - CCCCGATTTGGGGTGAAGACC	400	200
ICAM3	f - TTGGAGACGTCCCTATCAAAG r - GAGCCTGTACACGGTGATGT p - AGCTGGTGGCCAGTGGCATG	400	200
UBC	f - TCGTCACTTGACAATGCAGAT r - CATCTTCCAGCTGTTTTCCA p - TCCCTCCTGACCAGCAGAGGC	300	150

All assays were performed using TaqMan® Fast Universal PCR Master Mix (2X), No AmpErase® UNG (Applied Biosystems®, Life Technologies, UK). All cDNA samples were run in duplicate with a housekeeping gene assay for the same samples run on the same plate. Reactions were performs in MicroAmp^{TM} Optical 384-well Reaction Plates sealed with MicroAmp^{TM} optical adhesive film lids (both from Applied Biosystems®, Life Technologies, UK). The ABI PRISM® 7900HT Sequence Detection System with SDS v2.4 software (Applied Biosystems®, Life Technologies, UK) was used to perform the reactions. The PCR programme used is shown below.

Segment	**Cycles**	**Temperature (°C)**	**Time**
1	1	50	2 minutes
2	1	90	10 minutes
3	40	95	15 seconds
		60	1 minute

Transcript abundance in each cDNA sample was measured by the cycle number at which the detected fluorescence reached a threshold level (the C_t) that exceeded the background fluorescence. Transcripts that have a high abundance require fewer PCR cycles to reach the threshold level than transcripts present at a lower abundance, and hence the former have a lower C_t value. The expression level of each transcript was measured as delta C_t (dC_t), that is defined as the difference between the C_t of the transcript of interest and the C_t of the housekeeping gene (the *UBC* gene); the lower the dC_t value, the higher the transcript expression. Fold change in *TYK2* gene expression relative to the housekeeping gene was calculated as 2^{-dCt}. For time-courses, samples were normalised to the housekeeping gene and then to the unstimulated sample, calculated using 2^{-ddCt}. As RNA from two different cohorts was used for qPCR analyses by genotype, fold expression values were normalised to the mean expression of each gene in the common allele homozygote group.

qPCR data were analysed using GraphPad Prism® (version 8 for Mac; http://www.graphpad.com/prism/Prism.htm) and Microsoft Excel® (version 12.3.6 for Mac). Outliers were identified for exclusion using the ROUT method on Prism® with Q = 0.1% to minimise the false discovery rate (Motulsky and Brown, 2006).

2.16 Statistical analysis

All statistical tests were performed using GraphPad Prism® (version 8 for Mac; http://www.graphpad.com/prism/Prism.htm) and Microsoft Excel® (version 12.3.6 for Mac). Power calculations were performed using the R software (version 3.0.1 for Mac; http://www.r-project.org./; (Cohen, 1988; Team, 2008)). For cell line work, but not work with scarce and valuable donor samples, each experiment was performed a minimum of three independent times, with the following results Chapters featuring representative example figures and also indicating the number of replicates or donors in each data set.

2.16.1. Statistical analyses in Chapter 3: western blotting data from cell lines

Paired t-tests were used for analysis of western blotting experiments in which there were equal numbers of replicates and band fluorescence values were paired i.e. quantified from the same western blotting membrane. Where values were not all paired or where there an unequal number of replicates, an unpaired t-test was performed. In both cases, equal standard deviations were assumed, as there was no *a priori* reason to believe otherwise. To determine the degree of difference in variability between groups, an F test was performed to compare the variances across groups. The significance threshold was set at $P<0.05$.

2.16.2. Statistical analyses in Chapters 4 and 5: flow cytometry data from primary human immune cells

For correlations of antibody staining and age and also for assessment of the relationship between isotype control MFI or calibration bead normalisation factor and antibody MFI, linear regression analyses were performed with the *P*-value threshold set at 0.05. For comparison of flow cytometry data across cell subsets or by gender, unpaired t-tests were performed assuming equal standard deviations (as there was no *a priori* reason to believe otherwise) with a significance threshold set at $P<0.05$.

Regression analyses to test for correlations between rs12320356 or rs34536443 genotype and receptor expression or phospho-signalling in human immune cell subsets used a Bonferroni correction for multiple testing. For investigating receptor expression by genotype, the *P*-value threshold was defined as 0.05/2 = 0.025, considering previously published data which suggests a role for TYK2 in regulating the expression of 2 cytokine receptor subunits i.e. IFNAR1 and IL-10R2 (Gauzzi et al., 1997; Ragimbeau et al., 2003; Zheng et al., 2011). For phospho-signalling data, taking a *P*-value of 0.05 and considering three cytokine signalling pathways for two SNPs (based on previous literature; (Couturier et al., 2011)) the significance threshold was defined as 0.05/6 = 0.0083. Based on an average sample size of 64 individuals for IFN-β pSTAT signalling assays, the presented work had 74% power to detect the observed average r^2 of 0.0159 at this significance

threshold. Previously published data using a sample of 19 individuals (Couturier et al., 2011) had only 15.3% power to detect such differences, and thus the presented work has approximately 5-fold greater power.

2.16.3. Statistical analyses in Chapter 6: qPCR data from cell subsets and flow cytometry data for ICAMs

As for age and gender effects in flow cytometry data, linear regression and unpaired t-tests were performed, respectively. For comparing gene expression across immune cell subsets or for different stimulation conditions/timepoints, paired or unpaired t-tests were used depending on the nature of the data.

Regression analyses to test for correlations between SNP genotype and gene expression in human immune cell subsets used a Bonferroni correction for multiple testing. For lymphocytes, published eQTL data suggested an effect of two SNPs on the expression of three genes (see Table 6.3. in Chapter 6; (Brown et al., 2013; Veyrieras et al., 2008)) so a *P*-value significance threshold of 0.05/6 = 0.0083 was set. For qPCR data from lymphocytes, the sample size for the only significant trend was 142 and there was 73% power to detect the observed genotype-dependent correlation (r^2 of 0.07258) at this significance threshold. For monocytes and granulocytes, published eQTL data suggested an effect of 2 SNPs on the expression of 2 genes (see Table 6.3. in Chapter 6; (Fairfax et al., 2012; Wallace et al., 2012)) so a *P*-value significance threshold of 0.05/4 = 0.0125 was set. For qPCR data from monocytes and granulocytes, the average sample size for significant trends was 134 and there was 72.5% power to detect the observed average genotype-dependent correlation (r^2 of 0.06996) at this significance threshold.

For flow cytometric quantification of ICAM expression, linear regression was performed in a similar manner to earlier analyses to determine the best way to normalise data (i.e. to isotype control and/or calibration beads). To compare ICAM levels across genotype groups, linear regression analyses and unpaired t-tests were performed with a significance threshold of $P<0.05$.

3. Disease-associated amino acid substitutions in TYK2: *in vitro* assessment of variant kinase activity in a TYK2-deficient cell line

3.1 Introduction

As three of the four autoimmune disease-associated genetic variants in the *TYK2* region are nonsynonymous (see Chapter 1, Table 1.1.), these SNPs were hypothesised to be the causative variants accounting for each of the independent association signals. The amino acid substitutions within the TYK2 protein resulting from variation at these SNPs are shown in Table 3.1. and Figure 3.1. below. Given that the rs2304256, rs12720356, and rs34536443 SNPs encode amino acid variation in the FERM, pseudokinase and kinase domains, respectively (Figure 3.1.), they may influence the function of these domains. Thus, these nsSNPs may ultimately affect the ability of TYK2 to mediate cytokine-induced signalling by affecting: the binding of TYK2 to cytokine receptors, the pseudokinase-mediated regulation of the kinase domain, or the catalytic activity of the kinase domain, respectively.

Table 3.1. The three autoimmune disease-associated nsSNPs in *TYK2*: exonic location, alleles and encoded amino acids
Three nsSNPs within the *TYK2* gene are associated with autoimmune disease. The allele, encoded amino acid and position within the gene and TYK2 protein are shown. Data from NCBI RefSNP website (http://www.ncbi.nlm.nih.gov/snp).

SNP I.D.	Location in *TYK2* (exon)	SNP alleles as in the mRNA sequence (major / minor)	Encoded amino acid (major / minor)	Position of substitution in TYK2 protein
rs2304256	6	G / T	valine / phenylalanine (V / F)	362
rs12720356	13	T / G	isoleucine / serine (I / S)	684
rs34536443	21	C / G	proline / alanine (P / A)	1104

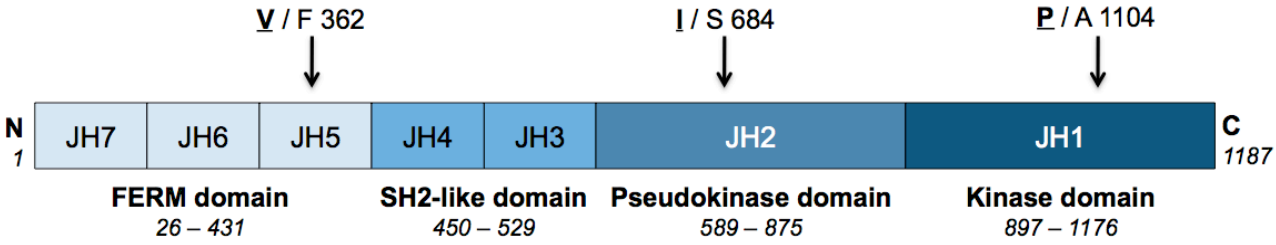


Figure 3.1. The location of autoimmune disease-associated amino acid substitutions in the TYK2 protein
The location of the amino acid residues varying by rs2304256, rs12720356 and rs34536443 nsSNP genotype are shown (from left to right). The amino acid encoded by the major allele is in bold and underlined. For a functional description of the depicted domains see Chapter 1, section 1.7. Diagram modified from (Levine and Gilliland, 2008).

In addition to their location in distinct domains of the TYK2 protein, a consideration of the extent to which the specific amino acids are conserved across the JAK family and across species provides a further indication of the potential functional relevance of these nsSNP-induced substitutions, based on the rationale that selective pressures will promote amino acid conservation in the interest of maintaining protein function. The PolyPhen tool (Adzhubei et al., 2010) can be used to predict whether or not amino acid substitutions are likely to be damaging to protein function, based on parameters such as conservation. Analyses using PolyPhen indicate that the amino acid substitutions caused by the minor alleles of both the rs12720356 and rs34536443 SNPs (I684S and P1104A, respectively) (Ban et al., 2009; Kaminker et al., 2007) are expected to have an impact on protein function, whereas the valine residue encoded by the major allele of the rs2304256 SNP is not highly conserved and thus substitution at this position (V362F) is not predicted to be damaging. An initial *in vitro* study of the effects of these variants on TYK2 phosphorylation is not totally consistent with these *in silico* predictions (Tomasson et al., 2008). This study suggests that the P1104A substitution (due to rs34536443) leads to reduced kinase activity, but that I684S (due to rs12720356) has no effect on TYK2 function, whereas V362F (due to rs2304256) was shown to increase TYK2 activity. However, a critical caveat of these findings is the qualitative rather than quantitative nature of the analysis, and thus the main aim of the work presented in this Chapter was to more accurately

measure the magnitude and direction of the effect of the different disease-associated substitutions on TYK2 activity.

To assess the impact of each of these nsSNP-dependent substitutions *in vitro*, transfection of the *TYK2*-null U1A (also termed 11,1) cell line was performed to exclude the effect other genomic or environmental influences. This cell line was originally characterised as a type 1 IFN-unresponsive mutant derivative of the HT-1080 fibroblast cell line (Pellegrini et al., 1989). U1A cells show impaired type I IFN signalling, that is, they are completely unresponsive to IFN-α and show only a minor response to IFN-β (Pellegrini et al., 1989). Further study identified the cause of the signalling defect through the use of serial genetic complementation, with the cloned cDNA capable of restoring responsiveness to type 1 IFNs being that encoding TYK2 (Velazquez et al., 1992), a molecule of unknown function at the time (Firmbach-Kraft et al., 1990). Thereafter the U1A cell line has been utilised in order to elucidate some of the basic biology of TYK2, including investigation of the role of the different functional domains of the protein and also the effect of mutations on kinase activity and on the regulation of cell surface IFNAR1 levels (Marijanovic et al., 2006; Ragimbeau et al., 2003; Ragimbeau et al., 2001; Staerk et al., 2005; Tomasson et al., 2008; Velazquez et al., 1995; Yeh et al., 2000).

Constructs expressing the different nsSNP-dependent TYK2 variants were generated by site-directed mutagenesis and transiently transfected into U1A cells, and type I IFN-induced phosphorylation of TYK2 and the substrate STAT1 were used as readouts of variant activity. The extent of cytokine-induced phosphorylation of TYK2 is a marker of catalytic activity as it is a requirement for the function of the kinase, and can be achieved by both autophosphorylation as well as JAK1-mediated trans-phosphorylation in the U1A cells transfected with TYK2. However, TYK2 mutants have been reported that are constitutively phosphorylated but are not catalytically active and cannot phosphorylate substrates such as the STAT molecules. This means that TYK2 phosphorylation alone is not a sufficient measure of catalytic activity (Yeh et al., 2000) and therefore phosphorylation of STAT1 was also quantified, noting that this reflects the action of both TYK2 and JAK1, which together mediate type 1 IFN signalling.

3.2 Chapter aim

The aim of the work presented in this Chapter was to investigate the effect of the three nsSNPs on the catalytic activity of TYK2 using a reductionist, cell line-based approach prior to assessing the role of these variants in the more complex context of primary human immune cell signalling. The nsSNPs result in amino acid substitutions in functionally important domains of the TYK2 protein and thus may influence kinase activity, however the distinct association pattern of these SNPs with different autoimmune diseases suggests that each variant is unlikely to have the same overall effect on TYK2 function. In addition, the discrepant findings from previously published studies warranted a more detailed analysis of their effects on TYK2 activity.

3.3 Results

The different variants of TYK2 are referred to using the amino acid code of the residues expressed at positions 362, 684 and 1104, respectively, as encoded by the three focal nsSNPs (Table 3.2.). In addition to constructs expressing the minor allele at each SNP, a construct expressing the minor allele of both the rs2304256 and rs34536443 SNPs was generated (encoding FIA TYK2), as there were individuals with this haplotype providing samples for studies in primary human immune cells (Chapters 4, 5 and 6). Given that individuals homozygous for the minor allele of the rs34536443 SNP are relatively rare, and some of these individuals have variation at the rs2304256 SNP in the cohort available for our donor selection, it was important to determine whether individuals with the FIA TYK2 variant would need to be considered as a separate group to those with VIA TYK2. Alternatively, these individuals could be pooled, thus increasing the total number of donors homozygous for the protective minor allele of the rs34536443 SNP and thus increasing our

power to detect potential genotype-dependent differences in TYK2 signalling in primary human immune cells. Constructs expressing other nsSNP combinations were not generated, as individuals expressing these haplotypes were not selected for subsequent studies.

Table 3.2. Amino acids encoded by different *TYK2* haplotypes and the respective abbreviations used in this Chapter

	Amino acid position			
	362	**684**	**1104**	**Abbreviation**
Most common haplotype (major allele at all 3 SNPs)	valine	isoleucine	proline	VIP
Minor allele at rs2304256 (major allele at rs12720356 and rs34536443)	phenylalanine	isoleucine	proline	FIP
Minor allele at rs12720356 (major allele at rs2304256 and rs34536443)	valine	serine	proline	VSP
Minor allele at rs34536443 (major allele at rs2304256 and rs12720356)	valine	isoleucine	alanine	VIA
Minor allele at rs2304256 and rs34536443 (major allele at rs12720356)	phenylalanine	isoleucine	alanine	FIA

3.3.1. IFN-β-induced phosphorylation of TYK2 and STAT1 in the U1A cell line transfected with a construct expressing the most common *TYK2* nsSNP haplotype

Having confirmed the absence of TYK2 protein in U1A cells (Figure 3.3.1.a), transfection of VIP TYK2 restored responsiveness to IFN-β, with cells showing a significant increase in levels of phosphorylated TYK2 and STAT1 after a 15 minute stimulation (pTYK2 and pSTAT1, respectively; P=0.0436 and 0.0035, respectively; Figures 3.3.1. and 3.3.2.). Consistent with published work, cells transfected with vector only (VO) showed a small pSTAT1 response to IFN-β stimulation mediated by JAK1, however this increased ~7-fold in cells expressing VIP TYK2 (P=0.0131; Figure 3.3.2.), demonstrating the significant contribution of TYK2 in IFN-β-induced STAT1 phosphorylation.

Fold increase in phosphorylation was calculated following a normalisation to the fluorescence signal for the unstimulated VIP-transfected cells for each separate experiment performed, and this normalisation has been used throughout this Chapter to account for any inter-experimental variation in fluorescence between different western blots. Replicate values include

multiple wells transfected at the same time point, and two or more separate transfections were performed for each experiment.

It is notable that the presence of VIP TYK2 was associated with a basal level of TYK2 and STAT1 phosphorylation in unstimulated cells, which was not seen in VO transfected cells (Figure 3.3.1. and 3.3.2.). This may be due to the overexpression of TYK2 in this transfection system, however given that IFN-β stimulation significantly increased levels of both pTYK2 and pSTAT1 relative to unstimulated cells (Figure 3.3.1. and 3.3.2.) the transfection system provides a readout of cytokine-induced TYK2 signalling that can be compared across different TYK2 variants.

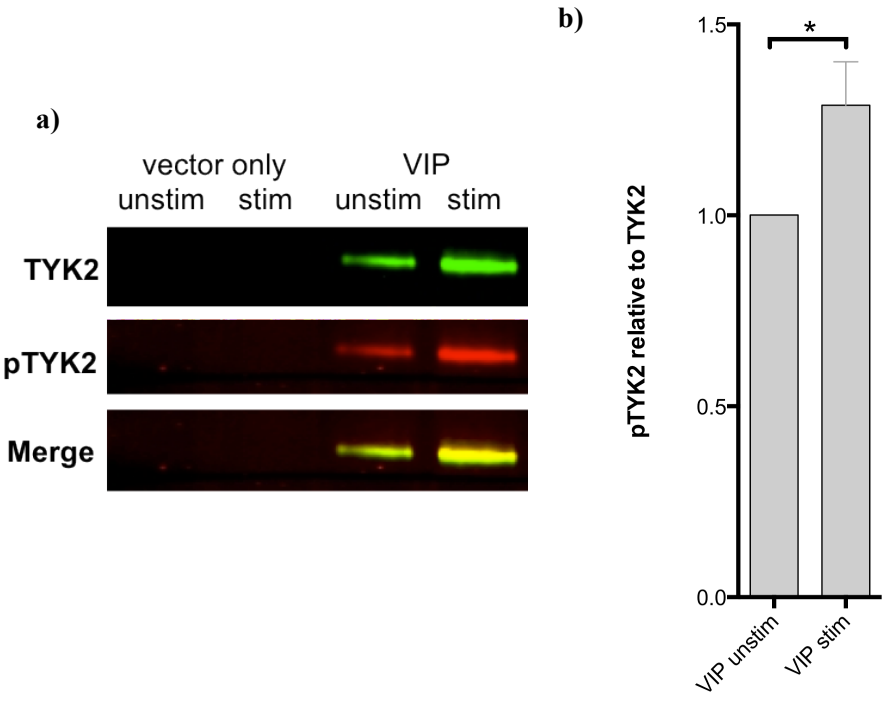


Figure 3.3.1. Expression of TYK2 protein and IFN-β-induced TYK2 phosphorylation in U1A cells transfected with VIP TYK2
U1A cells were transfected with vector only (VO) or a construct expressing the most common nsSNP haplotype of TYK2 (VIP; see main text). (a) Following a 15 minute stimulation with IFN-β, cells were lysed and samples prepared for western blotting, where each membrane was probed for both TYK2 and phosphorylated TYK2 (pTYK2), as shown in the representative blot. (b) The fluorescence of each band was quantified and the value of pTYK2 normalised to TYK2 was obtained. The mean ± standard error of the mean (SEM) values of pTYK2 normalised to TYK2 were plotted. Due to fluorescence variation across experiments, values from each individual membrane were normalised to the unstimulated condition (VIP unstim; n=7). VIP TYK2 transfected cells showed a significant increase in TYK2 phosphorylation in response to IFN-β stimulation (VIP stim; P=0.0436).

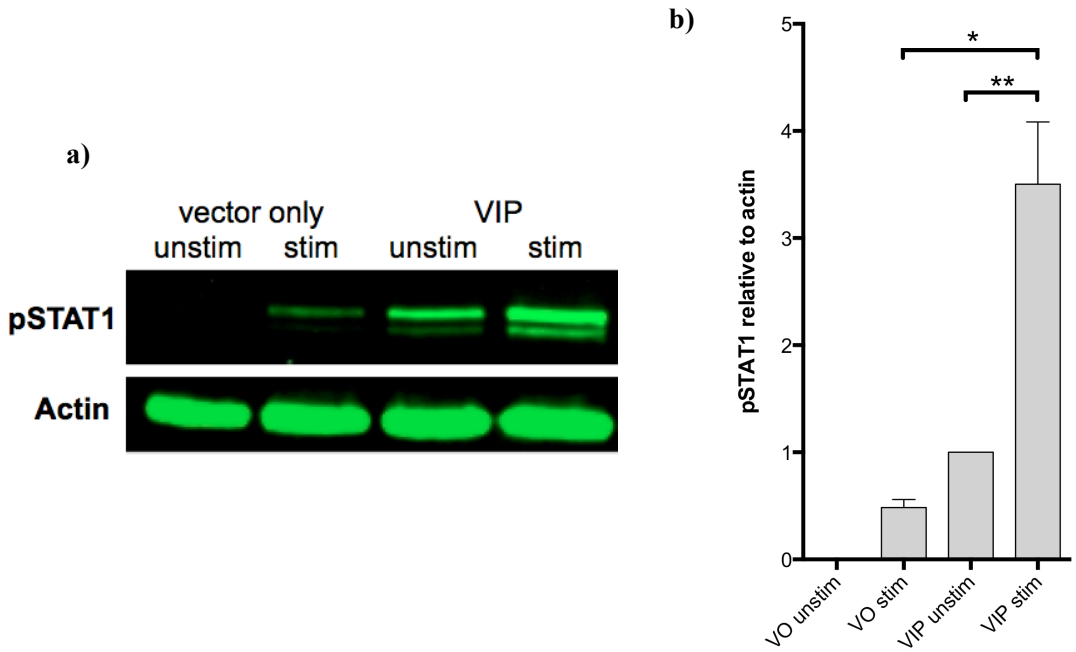


Figure 3.3.2. IFN-β-induced STAT1 phosphorylation in U1A cells transfected with VIP TYK2

U1A cells were transfected with vector only (VO) or a construct expressing the most common haplotype of TYK2 (VIP; see text). (a) Following a 15 minute stimulation with IFN-β, cells were lysed and samples prepared for western blotting, where each membrane was probed for both phosphorylated STAT1 (pSTAT1) and actin, as shown in the representative blot. (b) The fluorescence of each band was quantified and the mean ± SEM values of pSTAT1 normalised to actin were plotted. Compared to VO cells (n=3), VIP TYK2 transfected cells (n=7) show a significantly increased pSTAT1 response to IFN-β stimulation (P=0.0131). The levels of pSTAT1 were significantly increased in stimulated compared to unstimulated cells (P=0.0035).

3.3.2. The effect of the minor allele of rs2304256 on TYK2 signalling: IFN-β stimulation of U1A cells transfected with VIP compared to FIP TYK2

The IFN-β responsiveness of U1A cells transfected with the most common haplotype (resulting in VIP TYK2 expression) was compared to cells expressing the minor allele of the rs2304256 SNP, which encodes a phenylalanine rather than valine at position 362 (FIP TYK2). As transient transfection of these cells achieved overexpression of TYK2, it was important to determine whether TYK2 protein expression levels varied between the different constructs, despite the same amount of DNA being used. There was no significant difference between the TYK2 levels relative to actin in VIP compared to FIP transfected cells (P>0.05; Figure 3.3.3.a). Phosphorylation of TYK2 is required for catalytic activity (Korniski et al.) and may be mediated by TYK2 itself or another JAK

family member, however phosphorylation is not necessarily sufficient for kinase function, as some mutants have been reported that are phosphorylated yet catalytically inactive (Yeh et al., 2000). Thus, in addition to measuring the extent of TYK2 phosphorylation, the phosphorylation of the TYK2 substrate STAT1 was also quantified as a measure of the ability of different TYK2 variants to mediate signal transduction.

There is evidence from humans and mice that a lack of TYK2 protein causes a reduction in STAT1 protein levels (Karaghiosoff et al., 2000; Minegishi et al., 2006), thus is was important to first determine whether overexpression of TYK2 in this system led to differences in STAT1 levels between cells transfected with different TYK2 variants. There was no significant difference in STAT1 levels relative to actin between cells transfected with VO, VIP TYK2 or FIP TYK2 ($P>0.05$; Figure 3.3.4.)

Despite the significant difference in basal pTYK2 levels in unstimulated cells ($P=0.0396$; n=6), there was no difference in the extent of TYK2 phosphorylation in stimulated cells expressing FIP TYK2 compared to VIP TYK2 ($P>0.05$; Figure 3.3.3.b and 3.3.3.c). Also, there was neither a difference in background or stimulation-induced levels of pSTAT1 between the different constructs ($P>0.05$; Figure 3.3.4.b and 3.3.4.c), suggesting that FIP TYK2 is able to transduce IFN-β signalling to an extent indistinguishable from VIP TYK2, regardless of differences in basal pTYK2 levels in unstimulated cells.

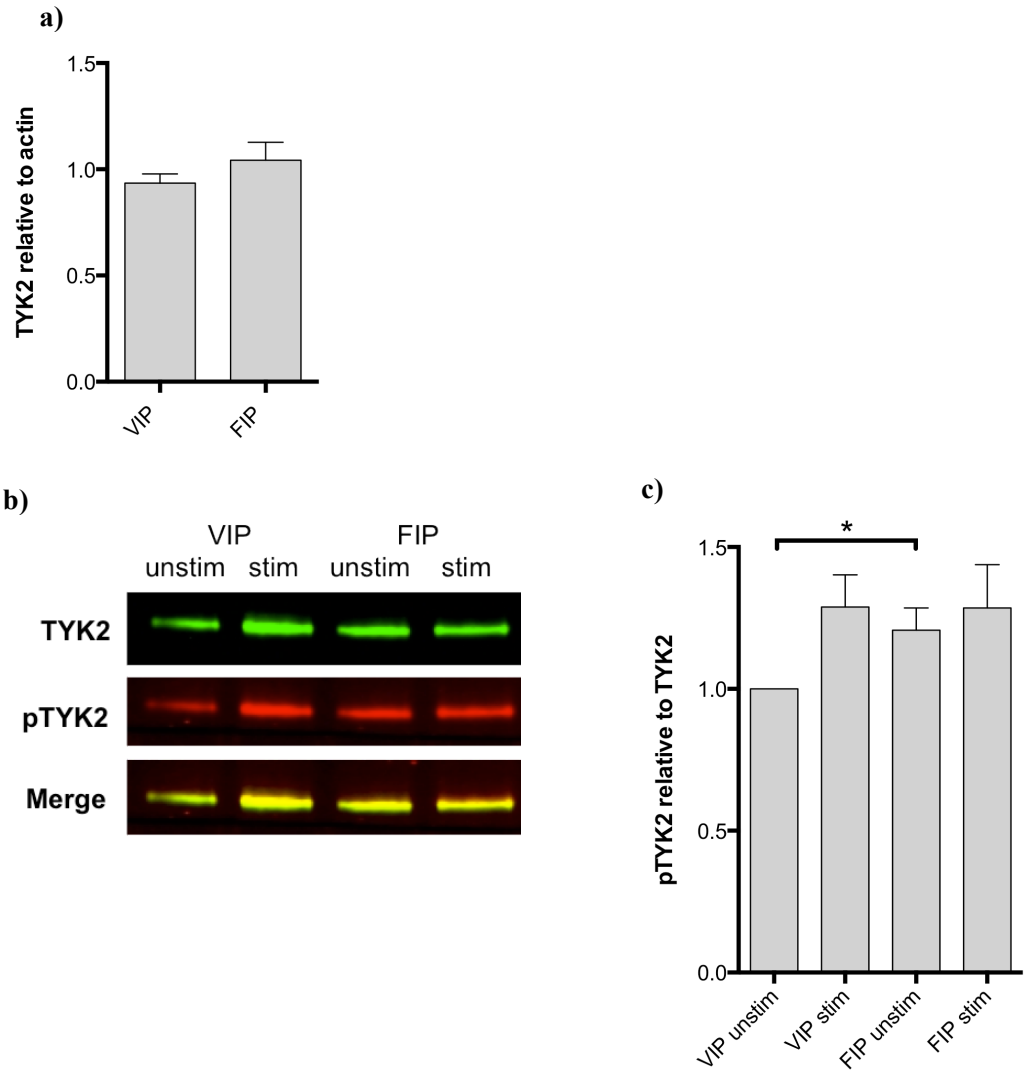


Figure 3.3.3. The effect of phenylalanine-362 on IFNβ-induced TYK2 phosphorylation
U1A cells were transfected with constructs expressing either the most common nsSNP-dependent haplotype of TYK2 (VIP) or the minor allele of rs2304256 (FIP). (a) TYK2 protein levels were quantified relative to the actin loading control and the mean ± SEM values are shown. There was no significant difference in the TYK2 protein levels between VIP and FIP TYK2 transfected cells (P>0.05; n=6). (b) Following a 15 minute stimulation with IFN-β, cells were lysed and samples prepared for western blotting, where each membrane was probed for both TYK2 and pTYK2, as shown in the representative blot. (c) The fluorescence of each band was quantified and the value of pTYK2 normalised to TYK2. The mean ± SEM values of pTYK2 normalised to TYK2 were plotted. Due to fluorescence variation across experiments, values from each individual membrane were normalised to VIP unstim. There was no significant difference in TYK2 phosphorylation between cells transfected with VIP and FIP (P>0.05; n=6). However, there was a significant increase in the background level of pTYK fluorescence in unstimulated cells when comparing FIP to VIP (P=0.0396).

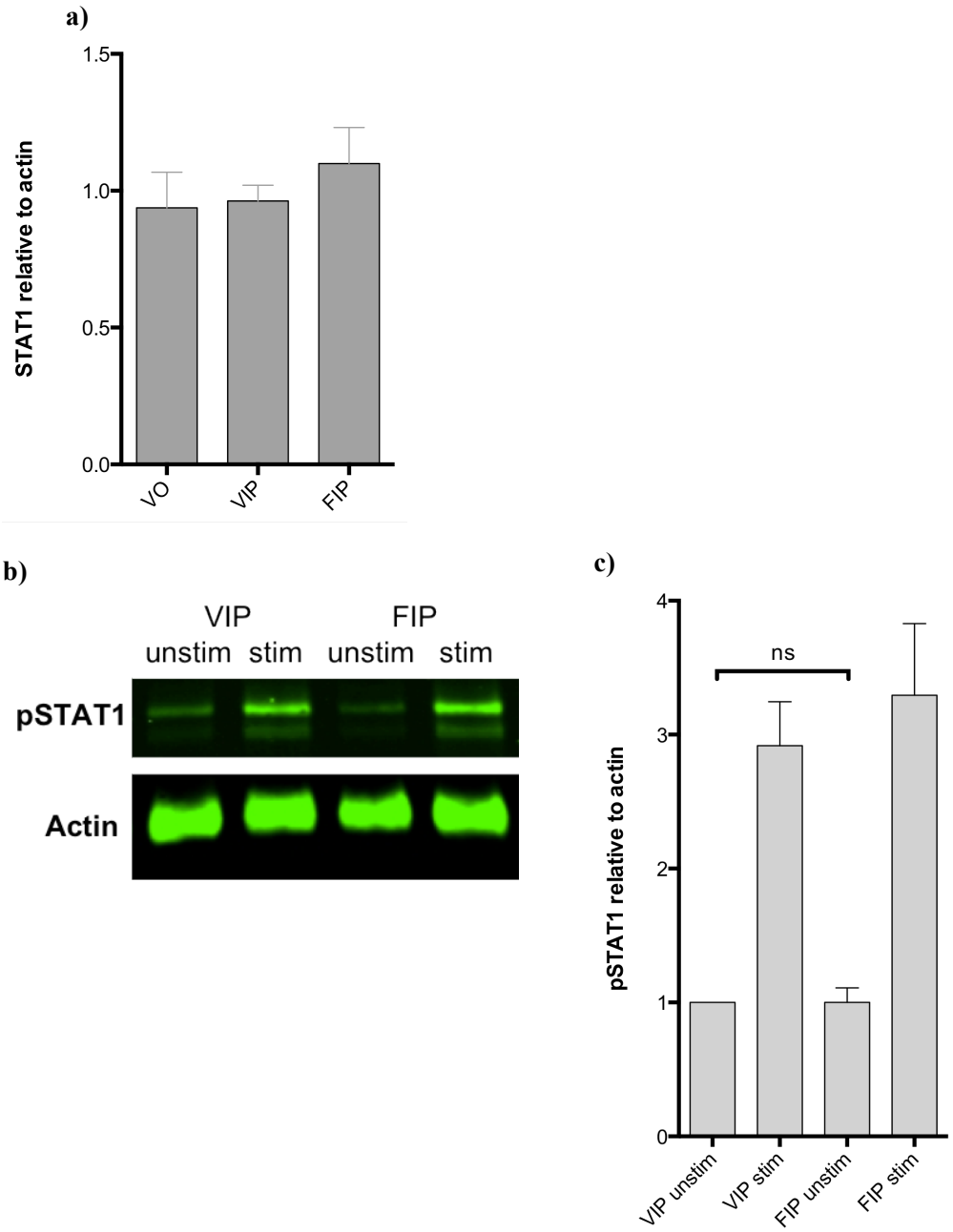


Figure 3.3.4. The effect of phenylalanine-362 on IFNβ-induced STAT1 phosphorylation
U1A cells were transfected with constructs expressing either the most common haplotype of TYK2 (VIP) or the minor allele of rs2304256 (FIP). (a) STAT1 protein levels were quantified relative to the actin loading control and the mean ± SEM values are shown. There was no significant difference in the STAT1 protein levels between VO, VIP TYK2 and FIP TYK2 transfected cells (n=3-6). (b) Following a 15 minute stimulation with IFN-β, cells were lysed and samples prepared for western blotting, where each membrane was probed for both phosphorylated STAT1 (pSTAT1) and actin, as shown in the representative blot. (c) The fluorescence of each band was quantified and the mean ± SEM values of pSTAT1 normalised to actin were plotted. There was no significant difference in basal pSTAT1 in unstimulated cells or in IFN-β-induced STAT1 phosphorylation between cells transfected with VIP compared to FIP TYK2 (P>0.05; n=6).

3.3.3. The effect of the minor allele of rs12720356 on TYK2 signalling: IFN-β stimulation of U1A cells transfected with VIP compared to VSP TYK2

Transfected cells were not selected for equal TYK2 expression and although the mean TYK2 protein level relative to actin was not significantly different (P>0.05), there was significantly more variation in TYK2 levels for cells transfected VSP TYK2 compared to VIP TYK2 (P<0.0001; Figure 3.3.7.a), despite the same amount of DNA being used across transfections. In addition, there were significantly higher levels of STAT1 protein relative to the actin loading control in cells transfected with VSP TYK2 compared to VIP TYK2 (P=0.0439; n=9-11; Figure 3.3.8.a).

There was a significant difference in both basal and IFN-β-induced pTYK2 levels for cells transfected with VSP TYK2 compared to VIP TYK2 (P=0.0048 and 0.0182, respectively; n=5; Figure 3.3.7.b and 3.3.7.c). However, this reduction in TYK2 phosphorylation was not sufficient to lead to a significantly reduced phosphorylation of STAT1 in unstimulated or stimulated cells (Figure 3.3.8.b and 3.3.8.c) suggesting that similarly to FIP TYK2, the VSP variant does not display reduced signalling competence. However unlike cells transfected with FIP, cells transfected with VSP TYK2 displayed significantly more variable and significantly increased levels of TYK2 and STAT1, respectively, thus complicating interpretation of the phosphorylation data. The increased variation in TYK2 protein levels may have contributed to the observed increase in STAT1 levels, and the latter may have influenced the pSTAT1 readout of TYK2 activity. In this case, it would have been ideal to normalise pSTAT1 levels to STAT1 for each sample by probing for both on the same membrane, as was performed for pTYK2 and TYK2, where appropriate antibodies were available for simultaneous quantification.

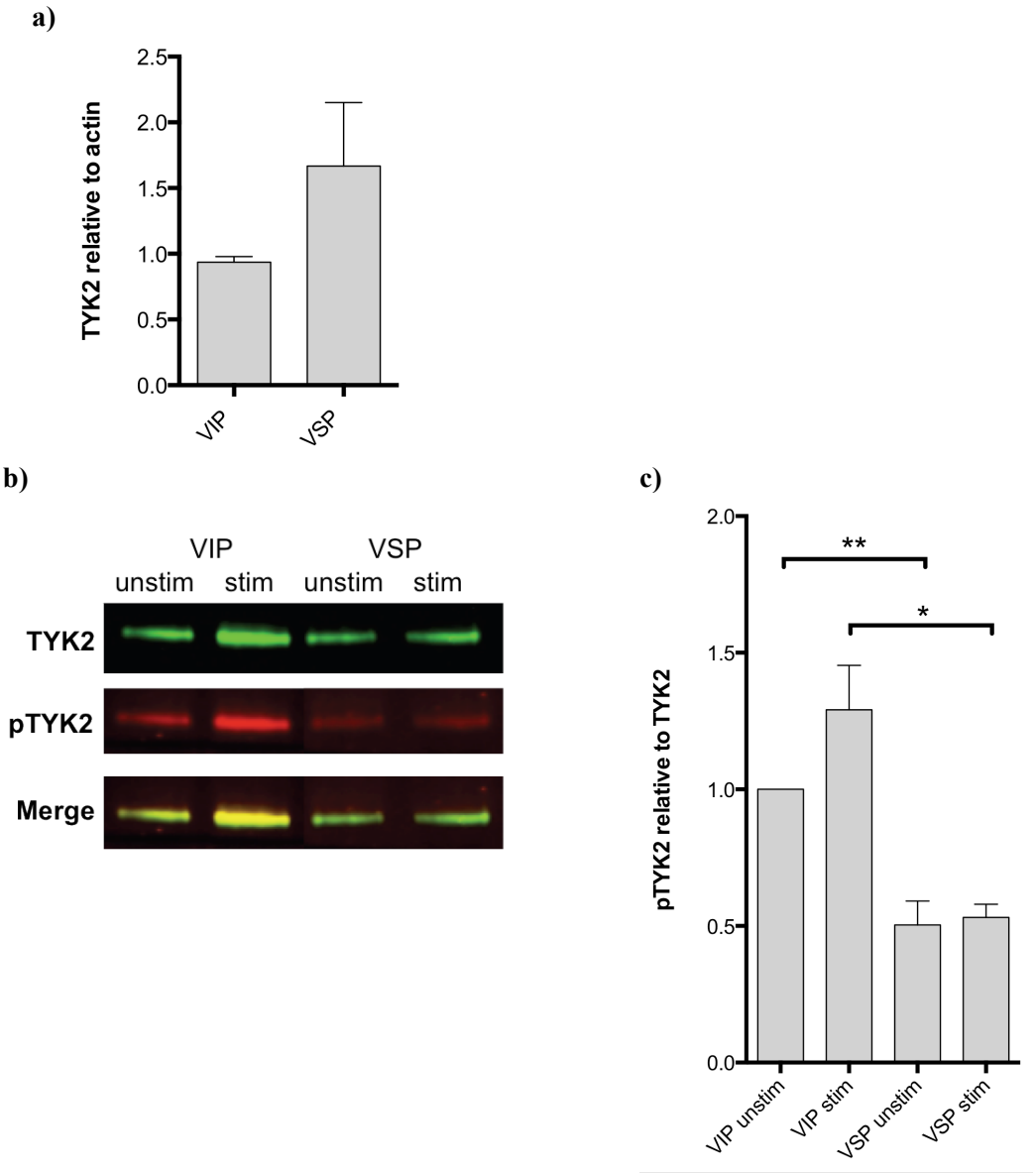


Figure 3.3.7. The effect of serine-684 on IFNβ-induced TYK2 phosphorylation

U1A cells were transfected with constructs expressing either the most common haplotype of TYK2 (VIP) or the minor allele of rs12720356 (VSP). (a) TYK2 protein levels were quantified relative to the actin loading control and the mean ± SEM values are shown. There was no significant difference in the mean TYK2 protein levels (P>0.05), however there was significantly more variability in TYK2 protein levels for cells transfected with VSP TYK2 compared to VIP TYK2 (P<0.0001; n=6). (b) Following a 15 minute stimulation with IFN-β, cells were lysed and samples prepared for western blotting, where each membrane was probed for both TYK2 and phosphorylated TYK2 (pTYK2), as shown in the representative blot. (c) The fluorescence of each band was quantified and the value of pTYK2 normalised to TYK2. The mean ± SEM values of pTYK2 normalised to TYK2 were plotted. Due to fluorescence variation across experiments, values from each individual membrane were normalised to VIP unstim. There was a significant reduction in both basal and IFN-β-induced pTYK2 levels in cells expressing VSP TYK2 compared to VIP (P=0.0048 and 0.0182, respectively; n=5).

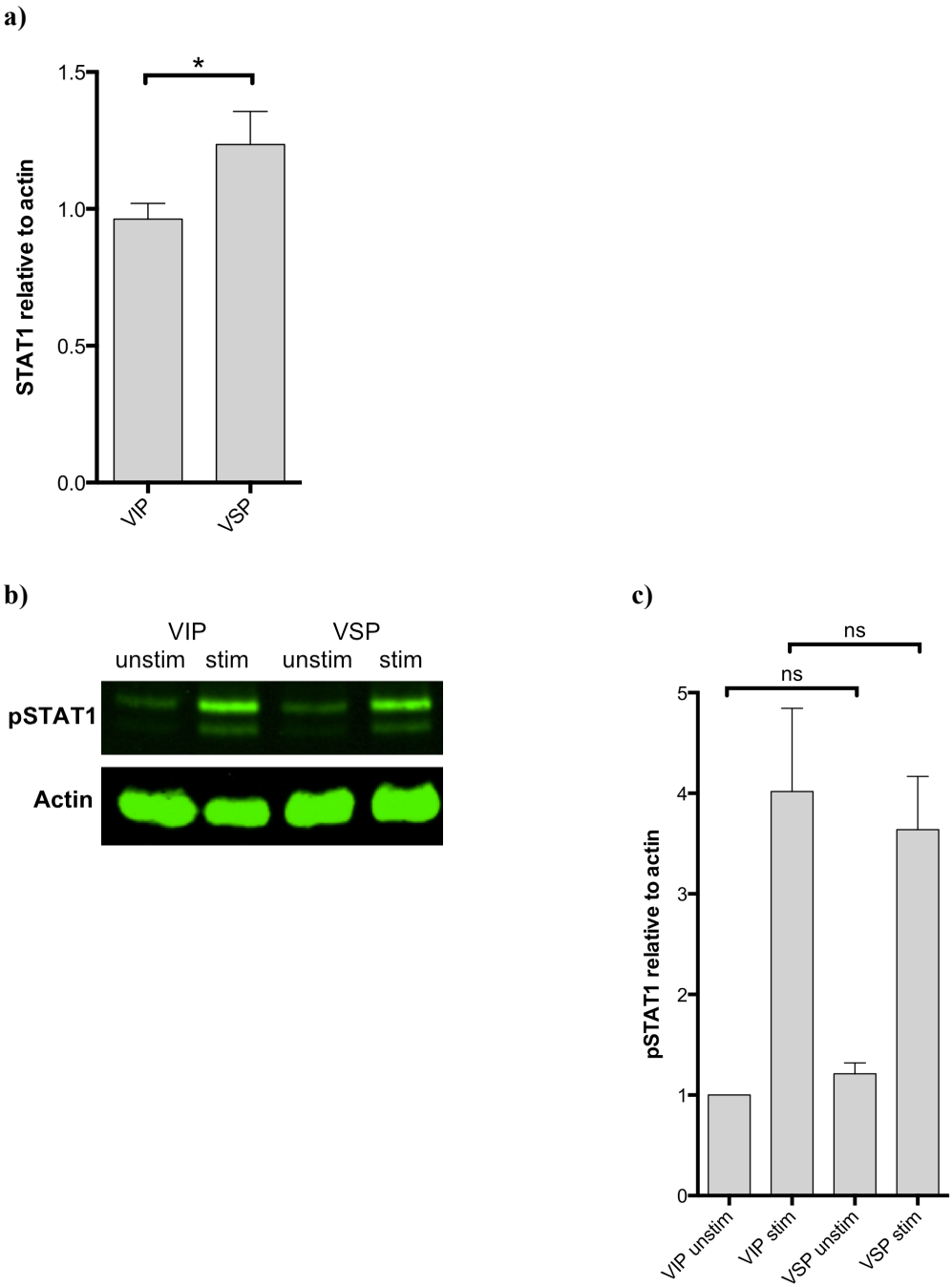


Figure 3.3.8. The effect of serine-684 on IFNβ-induced STAT1 phosphorylation
U1A cells were transfected with constructs expressing either the most common haplotype of TYK2 (VIP) or the minor allele of rs12720356 (VSP). (a) STAT1 protein levels were quantified relative to the actin loading control and the mean ± SEM values are shown. There were significantly higher levels of STAT1 protein relative to the actin loading control in cells transfected with VSP TYK2 compared to VIP TYK2 (*P*=0.0439; n=9-11). (b) Following a 15 minute stimulation with IFN-β, cells were lysed and samples prepared for western blotting, where each membrane was probed for both phosphorylated STAT1 (pSTAT1) and actin, as shown in the representative blot. (c) The fluorescence of each band was quantified and the mean ± SEM values of pSTAT1 normalised to actin were plotted. Due to fluorescence variation across experiments, values from each individual membrane were normalised to VIP unstim. There was no significant difference in basal or IFN-β-induced STAT1 phosphorylation between cells transfected with VIP and VSP TYK2 (n=5).

3.3.4. The effect of the minor allele of rs34536443 on TYK2 signalling: IFN-β stimulation of U1A cells expressing VIP compared to VIA TYK2

The minor allele of rs34536443 encodes an alanine at position 1104 in the TYK2 protein and cells were transfected with constructs expressing either this minor allele alone (i.e. VIA TYK2), or along with the minor allele of rs2304256 (i.e. FIA TYK2) as both nsSNP haplotypes were found in donors selected to provide samples for primary human cell work (Chapters 4, 5 and 6).

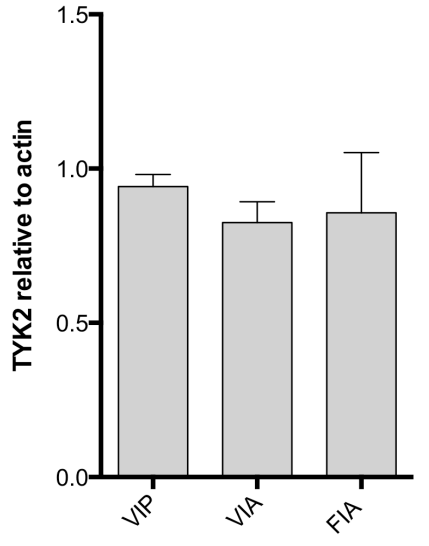
There was no significant difference in the mean TYK2 protein levels relative to actin between cells expressing VIA and FIA TYK2, compared to VIP ($P>0.05$; Figure 3.3.9.a), however there was significantly more variation in TYK2 protein levels for cells transfected with FIA TYK2 ($P<0.0001$; n=3-6). It is highly unlikely that this variation accounts for the highly significant reduction in IFN-β-stimulated phosphorylation of TYK2 observed both for cells transfected with VIA and FIA TYK2 (both $P<0.0001$; n=3-6; Figure 3.3.9.b and 3.3.9.c). The levels of basal pTYK2 in unstimulated cells were also significantly reduced both for cells transfected with VIA and FIA TYK2 (both $P<0.0001$; n=3-6; Figure 3.3.9.b and 3.3.9.c).

The significant reduction in both basal and stimulated pTYK2 levels was reflected in significant reductions in the respective pSTAT1 levels. For VIA, pSTAT1 levels were significantly reduced in both the unstimulated and stimulated cells relative to cells transfected with VIP TYK2 ($P=0.0002$ and $P=0.0124$, respectively; n=3-6; Figure 3.3.10.b and 3.3.10.c). Similarly for FIA, pSTAT1 levels were significantly reduced in both the unstimulated and stimulated cells relative to cells transfected with VIP TYK2 ($P<0.0001$ and $P=0.017$, respectively; n=3-6; Figure 3.3.10.b and 3.3.10.c). There was no significant difference in STAT1 protein levels relative to actin for cells transfected with VIA TYK2 compared to VIP TYK2 ($P>0.05$; Figure 3.3.10.a). As previously demonstrated, the variant encoded by the minor allele of the rs2304256 SNP does not affect STAT1 protein levels (FIP TYK2; $P>0.05$; Figure 3.3.4.a).

Unlike FIP and VSP TYK2, both basal and IFN-β-induced phosphorylation of both TYK2 and STAT1 was significantly reduced by both of the variants with an alanine at position 1104 in the TYK2 protein. The level of pTYK2 and pSTAT1 in both unstimulated and stimulated cells did not

significantly differ between cells transfected with VIA and FIA TYK2 (P>0.05), suggesting that the functional effect of these two nsSNP haplotypes is equivalent and thus that individuals with variation at the rs2304256 SNP could be included in investigations of the effect of alanine-1104 on TYK2 signalling in primary human immune cells (Chapters 4 and 5). Additionally, these data reinforce the lack of an effect of the minor allele of the rs2304256 SNP on TYK2 function, as there is no apparent difference between signalling in cells transfected with FIP compared to VIP, or FIA compared to VIA.

a)



b)

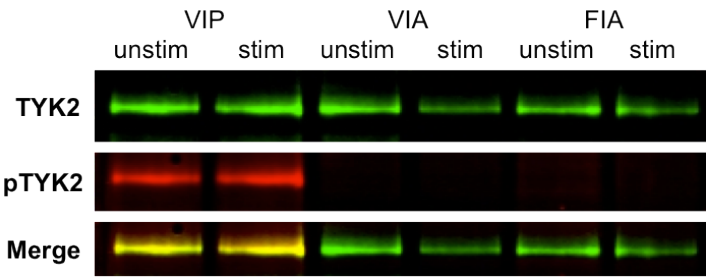


figure continued on next page...

c)

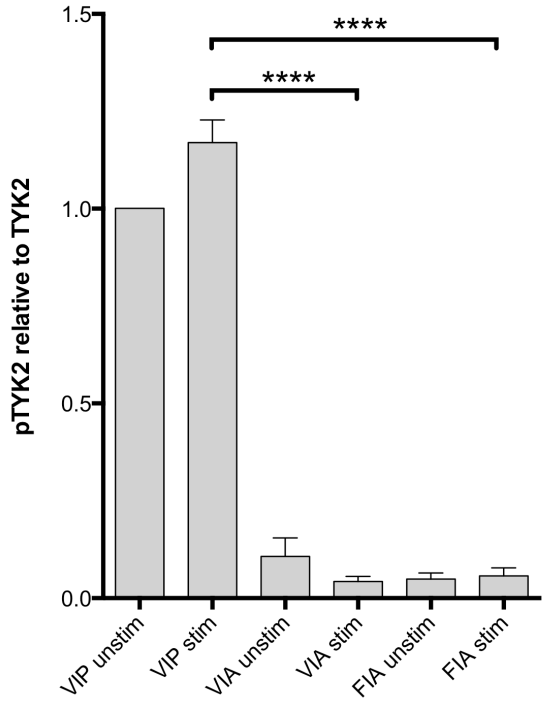


Figure 3.3.9. The effect of alanine-1104 and the FIA haplotype on IFNβ-induced TYK2 phosphorylation

U1A cells were transfected with constructs expressing either the most common haplotype of TYK2 (VIP), the minor allele of rs34536443 (VIA) or the minor alleles of both rs2304256 and rs34536443 (FIA). (a) TYK2 protein levels were quantified relative to the actin loading control and the mean ± SEM values are shown. There was no significant difference in the mean TYK2 protein levels between cells transfected with the different TYK2 constructs ($P>0.05$), however there was significantly more variability in TYK2 protein levels for cells transfected with FIA TYK2 compared to VIP TYK2 ($P<0.0001$; n=6). (b) Following a 15 minute stimulation with IFN-β, cells were lysed and samples prepared for western blotting, where each membrane was probed for both TYK2 and phosphorylated TYK2 (pTYK2), as shown in the representative blot. (c) The fluorescence of each band was quantified and the value of pTYK2 normalised to TYK2. The mean ± SEM values of pTYK2 normalised to TYK2 were plotted. Due to fluorescence variation across experiments, values from each individual membrane were normalised to VIP unstim. Asterisks representing significant differences between groups are only shown for the stimulated condition. Compared to VIP TYK2, there was significantly reduced IFN-β-induced pTYK2 in cells expressing VIA TYK2 ($P<0.0001$; n=6) and also FIA TYK2 ($P<0.0001$; n=3). Additionally, basal pTYK2 levels were significantly reduced in cells transfected with VIA and FIA TYK2 compared to VIP (both $P<0.0001$; n=6 and 3, respectively).

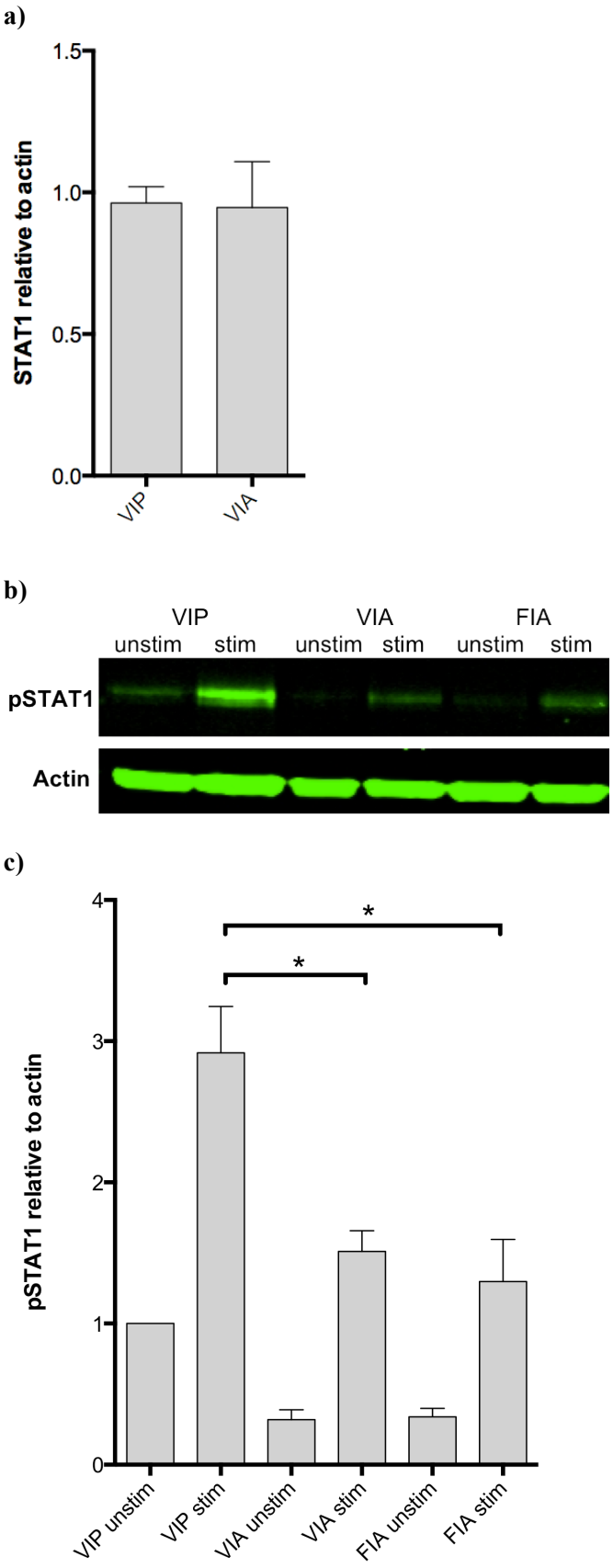


Figure 3.3.10. The effect of alanine-1104 and the FIA haplotype on IFNβ-induced STAT1 phosphorylation

U1A cells were transfected with constructs expressing either the most common haplotype of TYK2 (VIP), the minor allele of rs34536443 (VIA) or both the minor allele of rs2304256 and rs34536443

(FIA). (a) STAT1 protein levels were quantified relative to the actin loading control and the mean ± SEM values are shown. There was no significant difference in the STAT1 protein levels between cells transfected with VIA TYK2 compared to VIP TYK2 (n=6). (b) Following a 15 minute stimulation with IFN-β, cells were lysed and samples prepared for western blotting, where each membrane was probed for both phosphorylated STAT1 (pSTAT1) and actin, as shown in the representative blot. (c) The fluorescence of each band was quantified and the mean ± SEM values of pSTAT1 normalised to actin were plotted. Due to fluorescence variation across experiments, values from each individual membrane were normalised to VIP unstim. Asterisks representing significant differences between groups are only shown for the stimulated condition. Compared to VIP TYK2, there was significantly reduced IFN-β-induced pSTAT1 in cells expressing VIA TYK2 (P=0.0124; n=6) and also FIA TYK2 (P<0.017; n=3). Additionally, basal pSTAT1 levels were significantly reduced in cells transfected with VIA and FIA TYK2 compared to VIP (P=0.0002 and P<0.0001, respectively; n=6 and 3, respectively).

3.4 Discussion

U1A cells were transiently transfected with constructs expressing variants of *TYK2* that are encoded by three autoimmune disease-associated nsSNPs. The phosphorylation of TYK2 and STAT1 in response to IFN-β stimulation was quantified by western blotting.

For the FIP TYK2 variant encoded by the minor allele of the rs2304256 SNP, there was a significant increase in basal levels of pTYK2 in unstimulated cells but there was no difference in the extent of stimulation-induced TYK2 phosphorylation compared to the protein encoded by the most common haplotype, VIP TYK2. Only one paper has previously investigated the FIP variant, initially finding an increase in stimulation-induced phosphorylation of TYK2, however they could not replicate this finding (Tomasson et al., 2008). Moreover, the Tomasson *et al.* study did not perform quantification of any observed qualitative differences. Consistent with the lack of a difference in the extent of TYK2 phosphorylation, there was also no difference in pSTAT1 levels upon stimulation of cells expressing FIP compared to VIP TYK2. In contrast to the observed difference in basal pTYK2 levels, cells transfected with FIP TYK2 did not show a difference in basal STAT1 phosphorylation, suggesting that the basal difference in TYK2 phosphorylation is not a reflection of the molecule's catalytic activity and that FIP TYK2 signalling is not impaired relative to that of VIP TYK2. Given that there was no difference in pTYK2 or pSTAT1 for FIP compared to VIP upon stimulation, it was

deemed unlikely that rs2304256-dependent differences in TYK2-mediated phosphorylation would be observed with human primary immune cells, leading to the hypothesis that this SNP may instead influence gene expression (Chapter 6).

In contrast to FIP TYK2, VSP TYK2 (encoded by the minor allele of the rs12720356 SNP) displayed reduced phosphorylation in both the unstimulated and stimulated conditions. Tomasson and colleagues had previously noted reduced TYK2 protein levels in cells expressing this variant but found no difference in TYK2 phosphorylation compared to VIP TYK2, although this was not formally quantified (Tomasson et al., 2008). In the work presented in this Chapter, the observed difference in TYK2 phosphorylation was not reflected by a change in STAT1 phosphorylation, suggesting that signalling is mediated by VSP TYK2 to the same extent as by VIP TYK2, however analyses of this variant were complicated by significant variation in TYK2 and STAT1 expression levels.

The presented data demonstrate a significant reduction in phosphorylation of both TYK2 and STAT1 for cells expressing the alanine-1104 variant of TYK2, thus providing strong evidence that the minor allele of the rs34536443 SNP results in impaired TYK2 signalling. Consistent with the work presented in this Chapter, Couturier and colleagues report defective IFN-β signalling in expanded T cells derived from individuals heterozygous for the rs34536443 SNP (Couturier et al., 2011). A recent paper by Li *et al.* also investigated the activity of the VIA TYK2 variant in stably transfected cell lines and EBV-transformed cell lines, reporting reduced activity of VIA TYK2 in cell-free assays (Li et al., 2013). However, inconsistent with their own findings, as well as the data presented in this Chapter and the findings of Couturier *et al.*, Li and colleagues show no reduction in TYK2 or STAT1 phosphorylation in cell-based stimulation assays, however phosphorylation signals were not quantified (Li et al., 2013). The discrepancies between the two assay systems used could suggest that, in their cell-based assays, JAK1 can compensate for the apparent reduction in TYK2 catalytic activity observed in the cell-free assay, thereby restoring type 1 IFN signalling in cells. An analogous discrepancy between the two assay systems was also observed for the VSP TYK2 variant (Li et al., 2013). The disparity between the findings of Li and colleagues compared to the data presented in this Chapter and the findings of Couturier *et al.* may relate to the use of IFN-α in the

former and IFN-β in the latter, as although are both type 1 IFNs using the same receptor chains, there is evidence to suggest that IFN-β interacts more strongly than IFN-α with IFNAR1 (de Weerd et al., 2013), which is directly bound to TYK2 (Colamonici et al., 1994) (JAK1 binds IFNAR2; (Domanski et al., 1997)). The IFN-β-IFNAR1 complex has been demonstrated to be capable of signalling in the absence of IFNAR2 (de Weerd et al., 2013), and although the precise role of TYK2 in this instance is yet to be elucidated, it is possible that the relative importance of TYK2 may vary downstream of IFN-β compared to IFN-α.

As TYK2 mediates signal transduction for a variety of cytokine receptors in primary human immune cells and always acts together with another JAK (or JAKs), SNP-dependent effects on TYK2-mediated signalling will relate to their direct effect on TYK2 catalytic activity and also the relative importance of TYK2 activity downstream of each cytokine receptor. Collectively considering the presented data and also previously published work, there is strong evidence that the VIA TYK2 variant encoded by the minor allele of the rs34536443 SNP is catalytically impaired, warranting further investigation to determine the relevance of this reduced activity across TYK2-mediated cytokine pathways in primary human immune cells. Overall there is no consensus regarding the degree to which rs12720356 genotype influences TYK2 activity, as studies of VSP TYK2 have involved the use of overexpression systems in which, for example, STAT1 levels are also altered. Thus, the inconsistent findings from previous studies may not reflect the effect that this variant may have on TYK2-mediated signalling in primary human immune cells that naturally express TYK2. In contrast, there is no evidence to suggest that the rs2304256 genotype affects TYK2 signalling either in the context of the major or minor allele of the rs34536443 SNP, and thus the functional effect of this SNP may not directly be on TYK2 activity. This means that individuals with variation at rs2304256 can be selected for the studying the effect of rs34536443 genotype on TYK2 signalling in primary human immune cells, increasing the number of available donors carrying the minor allele of the latter SNP for better powered investigation of potential genotype-dependent functional effects.

4. Characterisation of TYK2-mediated cytokine signalling pathways in primary human immune cell subsets

4.1. Introduction

For two of the three nsSNPs investigated, there was some evidence from analyses in the U1A cell line (Chapter 3) that the minor allele was associated with reduced TYK2 function. Thus, these nsSNPs, rs34536443 and rs12720356, were prioritised for further assessment of their impact on TYK2-mediated cytokine signalling in primary human immune cells. Although the cell line work provided an initial indication that nsSNP genotype may affect TYK2 activity, the over-expression of TYK2 and the observed upregulation of STAT1 in the transfection system may not reflect what would be seen under physiological conditions, and thus genotype-dependent findings required validation in a primary cell setting. This is particularly the case for the rs12720356 SNP, where differences in TYK2 phosphorylation but not in phosphorylation of the TYK2 substrate STAT1, were observed. Moreover, type 1 IFN signalling is the only pathway that can be investigated using U1A cells and TYK2 is in fact involved in the signalling pathways of multiple additional cytokines (Figure 4.1) (Watford and O'Shea, 2006). Thus, potential genotype-dependent differences in TYK2 activity across pathways required investigation in primary human immune cells, which express the cytokine receptors for TYK2-mediated signalling pathways including IL-6, IL-10, IL-12, IL-13 and IL-23 (Figure 4.1.1.).

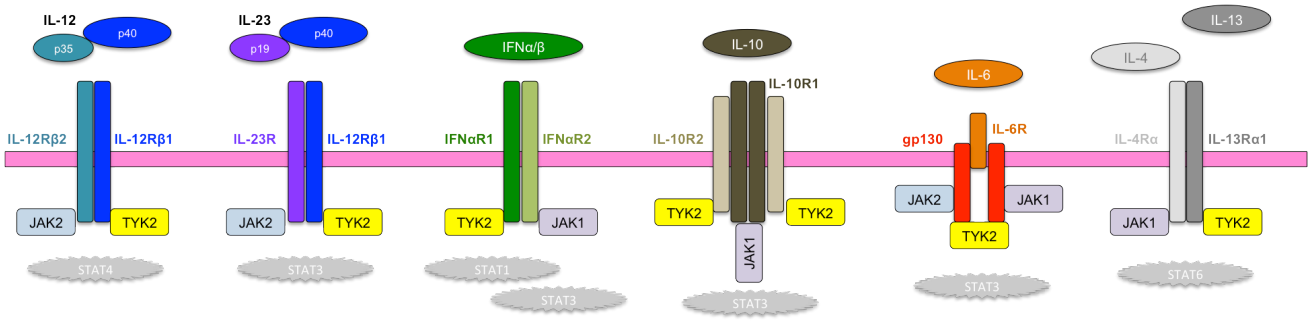


Figure 4.1. TYK2-mediated cytokine signalling pathways
TYK2 is involved in the signal transduction of multiple cytokines with varied roles in the immune system. Binding of cytokine to its cognate receptor leads to activation of intracellular JAK-STAT signalling and downstream changes in gene expression. Diagram modified from (Watford and O'Shea, 2006).

Along with other the other JAK family members, TYK2 mediates signal transduction via STAT phosphorylation downstream of the cytokine-cytokine receptor interaction. Using STAT phosphorylation (pSTAT) as a readout will thus reflect the activity of both TYK2 and the JAK (or JAKs) mediating signal transduction at a particular receptor. By assessing STAT rather than TYK2 phosphorylation, the cytokine pathways in which genotype-dependent differences in TYK2 activity have the greatest impact on signalling will be determined. Thus, a cytokine pathway in which pSTAT levels correlate with *TYK2* nsSNP genotype will reflect the inability of JAK activity to fully compensate for reduced TYK2 function, indicating the relative functional importance of TYK2 in the particular signalling cascade. Therefore, the cytokine pathways that are most relevant to the autoimmune diseases with which the *TYK2* nsSNPs are associated will be distinguished.

The disease-associated genetic variants to be investigated have small effects on overall disease risk (see ORs in Table 1.1. of Chapter 1) and are thus likely to have relatively subtle functional effects. Thus, the use of primary human cells for assessment of potential genotype-dependent differences requires a consideration of the factors that may contribute to the variation in the pSTAT readout used and that may consequently reduce the ability to detect differences attributable to the nsSNPs of interest. These variables may include inter-individual variation due to other genetic variants or environmental factors that may influence cytokine signalling, and experimental variation, for example due to differences in flow cytometer function through time.

4.2 Chapter aim

The aim of the work presented in this chapter was to characterise selected representative TYK2-mediated cytokine pathways in primary human immune cell subsets, in order for an informed subsequent study of TYK2 signalling by nsSNP genotype (Chapter 5). This characterisation involved assessing expression of IFNAR1, IL-6R, IL-10R and IL-13R across primary human immune cell subsets directly *ex vivo*, as cell surface receptor expression dictates cytokine-induced TYK2-mediated signalling. Moreover, studies using cell lines have suggested that TYK2 may play a role in regulating cell surface expression of IFNAR1 and IL-10R2 (Gauzzi et al., 1997; Ragimbeau et al., 2003; Zheng et al., 2011). Thus, any nsSNP-dependent changes in TYK2 activity could alter cytokine receptor levels and potentially affect STAT phosphorylation as a consequence. Therefore, rs34536443- and rs12720356-dependent differences in cell surface receptor expression were investigated by flow cytometry to help determine whether the pSTAT readout of TYK2 activity was likely to be a direct measure of TYK2/JAK activity, or whether it may be a more complex readout reflecting prior *in vivo* TYK2-dependent effects on receptor level as well as assay-induced TYK2 activity.

In addition to considering variation in receptor levels, the level of cytokine and the stimulation time required for maximal responsiveness were also investigated to ensure a subsequent optimal analysis of *TYK2* nsSNP-dependent signalling. Additionally, given that the use of primary human cells to study genotype-dependent effects may be further complicated by inter-individual variation (at other genetic variants that are relevant to TYK2-mediated cytokine signalling pathways or by different environmental influences), the overall study design with respect to donor selection was considered to maximise the potential to detect pSTAT variation attributable to *TYK2* nsSNP genotype.

4.3. Results

4.3.1. Cytokine receptor expression in TYK2-mediated pathways across immune cell subsets directly *ex vivo*: optimising the approach to quantification

Cytokine receptor expression level is a key component governing the dynamics of cytokine responses, as has been demonstrated for cytokines such as IL-2 and IL-7, for example (Busse et al., 2010; Cotari et al., 2013; Hofer et al., 2012). Therefore, an assessment of genotype-dependent differences in signalling across individuals requires consideration of variation in cell surface receptors levels within and across samples. TYK2-dependent signalling downstream of representative type I short-chain (IL-13), long-chain (IL-6), and type II (IFN-β; IL-10) cytokines (Leonard, 2001) was assessed in Chapter 5. These pathways were selected as their receptors are known to be expressed on immune cells directly *ex vivo*, without the need for *in vitro* cell activation to upregulate receptor expression, which could mask nsSNP-dependent effects by promoting unphysiological, stimulation-induced biases in signalling.

Within any given individual, cytokine receptor expression may vary with respect to both the number of cells expressing the receptor and also the receptor level on a per cell basis, which may be reflected in apparent differences in cytokine-induced STAT phosphorylation between individuals. In this way, variation in receptor expression could promote pSTAT variation, thus confounding the identification of relatively subtle potential nsSNP-dependent differences in STAT signalling. Detailed characterisation of the relationship between receptor expression and cytokine responsiveness in human peripheral blood mononuclear cells (PBMCs) requires the ability to consider receptor level and the pSTAT response in a single test, such that the two can be directly correlated and the proportion of pSTAT variation attributable to receptor variation can be determined (Cotari et al., 2013). However, co-immunostaining for the cytokine receptors and pSTAT could not be performed, as the methanol-based permeabilisation reagent required to enable intracellular anti-pSTAT staining (Krutzik and Nolan, 2003) was not found to be conducive to staining with any of the antibodies against the cytokine receptors of interest (data not shown).

In the absence of the ability to directly correlate pSTAT and receptor levels, an alternative approach to account for cytokine receptor variation across PBMC samples was needed. As previously demonstrated for IL-2RA (Dendrou et al., 2009b), cytokine receptor expression may vary across PBMC subsets and thus receptor expression was characterised across immune cell subpopulations, including naïve and memory CD4+ and CD8+ T cells, B cells, and monocytes (Figure 4.3.1.) by multi-parameter flow cytometry. Considering TYK2-mediated cytokine pathways across cell subsets could help to minimise receptor-dependent variation in pSTAT measures, as an alternative to direct determination of the relationship between receptor and pSTAT levels, which was not possible.

It is noteworthy, that apart from pSTAT variation due to differential receptor expression across subsets, it was not hypothesised that nsSNP-dependent differences in pSTAT levels would be found in only certain cell subpopulations, as subset-specific differences in the kinase activity of TYK2 have not been previously described. As TYK2 is ubiquitously expressed in immune cells and the genetic variants of interest are nonsynonymous, potential functional effects will be evident across the immune cell subsets studied.

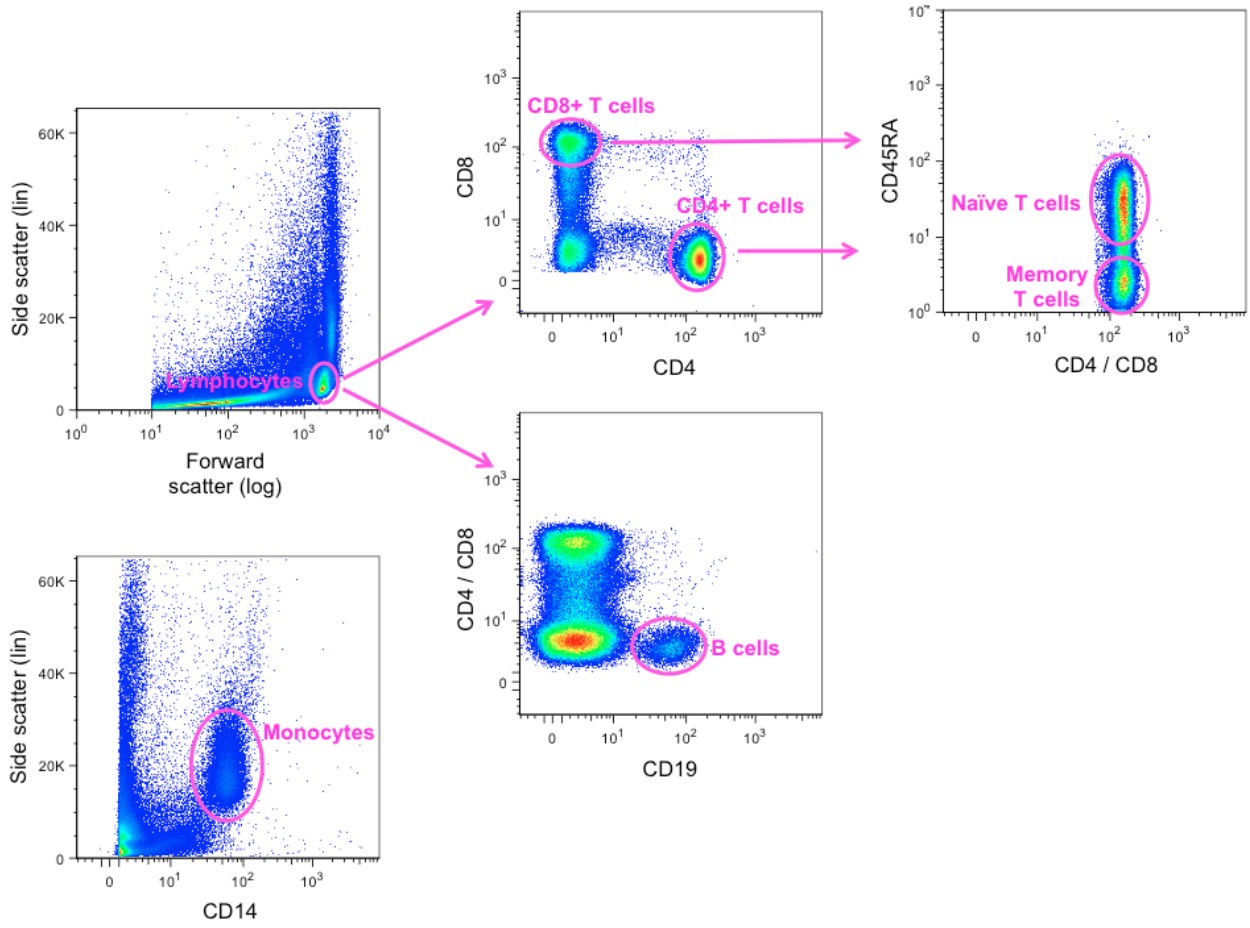


Figure 4.3.1. Flow cytometry gating strategy for stained human PBMCs
PBMCs were stained using markers of the main immune cell subsets. Lymphocytes were first gated by their forward and side scatter profile, and then divided into CD4+, CD8+ or CD19+ subsets. CD4+ and CD8+ T cells we further sub-divided based on their expression of CD45RA, a marker of naïve T cells. Monocytes were gated based on their higher side scatter relative to lymphocytes and on expression of CD14.

In order to accurately quantify cell subset-specific cytokine receptor levels, it was important to consider potential sources of variation, including inter-individual variation (e.g. due to unspecific staining) and also experimental variation due to differences in flow cytometer function and voltage settings through time. Isotype controls can account for unspecific antibody staining which may arise from the binding of antibodies to Fc receptors expressed on certain immune cell types, for example, rather than binding to the specific protein that they have been raised against. Therefore, isotype control antibody staining was used as an indicator of the background fluorescence levels, and appropriate unconjugated isotype Ig was also used prior to staining to aid Fc receptor blocking (see Chapter 2, section 2.12.1. and 2.12.2.). Whilst isotype control staining may reflect variation in

fluorescence readouts due to day-to-day changes in flow cytometer settings and function to some extent, it has previously been demonstrated that the use of calibration beads may better reflect this variation (Dendrou et al., 2009a; Ferreira et al., 2013; Perfetto et al., 2012; Schwartz et al., 1996). The range of the bead fluorescence intensities is more similar to the MFI of the marker of interest, in contrast to the isotype control staining which only has a low-level fluorescence intensity. Normalisation to the calibration beads involves calculating the normalisation coefficient, which is the slope of the linear relationship between the mean fluorescence intensity (MFI) of the calibration bead fluorescence peaks and the molecules of equivalent fluorochrome (MEF) values of each bead population (which are provided by the manufacturer).

In order to determine how to best normalise receptor expression data acquired by flow cytometry, the extent of variation accounted for by the isotype control staining and the calibration beads was explored. For all cell subsets and cytokine receptors of interest, the MFI of the isotype control was plotted against the MFI of the receptor staining to determine the extent to which unspecific antibody binding accounted for variation in the data and whether this relationship was significant (Table 4.1.). Unspecific binding was found to significantly contribute to variation in receptor staining and was highest for the rabbit polyclonal IgG antibody (r^2>0.8 such that unspecific binding explains >80% of receptor level variation), and lowest for the mouse IgG1 antibody (r^2 = 0.06-0.23), with the IgG2 antibodies showing an intermediate binding non-specificity (r^2 = 0.18-0.48) as expected (Burton, 1985; Unkeless et al., 1981).

Although isotype antibody staining alone is traditionally used as a control, the use of calibration beads has been demonstrated as a more accurate method for accounting for variation in machine function and voltage settings (Dendrou et al., 2009a; Ferreira et al., 2013; Perfetto et al., 2012; Schwartz et al., 1996). Thus, the extent to which staining variation was reflected by the calibration beads on each experimental day was also determined. As shown in Table 4.1. for all cytokine receptors studied, a significant proportion of fluorescence variation was accounted for by this coefficient, ranging from 21-47% of the staining fluorescence. For IFNAR1, the isotype control accounts for a larger proportion of variation relative to the beads, which is more likely to reflect the increased level of unspecific binding; for antibody isotypes known to have low unspecific binding

(i.e. for IL-6R and IL-10R), the calibration beads accounted for a larger proportion of the variation in receptor staining than the isotype. This indicates that the beads are likely to better capture variation in flow cytometer function than isotype control antibodies.

In summary, the analysis in Table 4.1. indicates that both the isotype control staining and calibration beads account for significant proportions of variability in the data, due to unspecific binding and flow cytometer settings/function, respectively. Thus, both were used for normalisation of cytokine receptor MFI values to enable expression to be compared across cell subsets in donor samples over time.

Table 4.1. Correlating receptor staining with the MFI of the isotype control and the normalisation coefficient calculated using the calibration beads
The MFI of the appropriate isotype control antibodies was plotted against the MFI of the staining for each receptor, as was the normalisation coefficient calculated using calibration beads on each experimental day. The r^2 of the observed linear relationships and the extent of significance (*P*-value) are shown for each receptor, for memory CD4+ T cells (which are representative of the other investigated lymphocyte subsets) and monocytes. As IL-13R is only expressed by monocytes, this is the only subset shown.

			isotype MFI vs R MFI		beads vs R MFI	
Receptor	**Isotype**	**Cell subset**	r^2	***P*-value**	r^2	***P*-value**
IFNAR1	rabbit polyclonal IgG	CD4+ memory	0.84	<0.0001	0.21	0.0001
		Monocytes	0.88	<0.0001	0.27	<0.0001
IL-6R	mouse IgG1	CD4+ memory	0.06	0.03	0.29	<0.0001
		Monocytes	0.23	<0.0001	0.21	<0.0001
IL-10R	rat IgG2a	CD4+ memory	0.18	<0.0001	0.47	<0.0001
		Monocytes	0.22	<0.0001	0.37	<0.0001
IL-13R	mouse IgG2b	Monocytes	0.48	<0.0001	0.4	<0.0001

4.3.2. Cytokine receptor expression in TYK2-mediated pathways across immune cell subsets directly *ex vivo*: the effects of age and sex

Following consideration of the most accurate way to normalise receptor expression over time, the potential confounding effects of donor age and sex were investigated, as receptor level dependence on both these variables has been previously described for some cytokine receptors (Dendrou et al., 2009a). Any outliers were excluded as detailed in Chapter 2, section 2.13. For all cell subsets and all cytokine receptors studied, there were no correlations between receptor expression

and age (all $P>0.05$), noting that the age of donors in the selected cohort ranged from 25-63 years with an average age of 46 years. With respect to sex, the donor male-to-female ratio in the total cohort was 43:51. Similarly to the lack of correlations with age, there was predominantly no effect of sex, apart from a significant difference for IL-6R expression on monocytes where on average males had a higher expression than females ($P=0.0383$; Figure 4.3.2.); this difference was therefore accounted for in subsequent analyses.

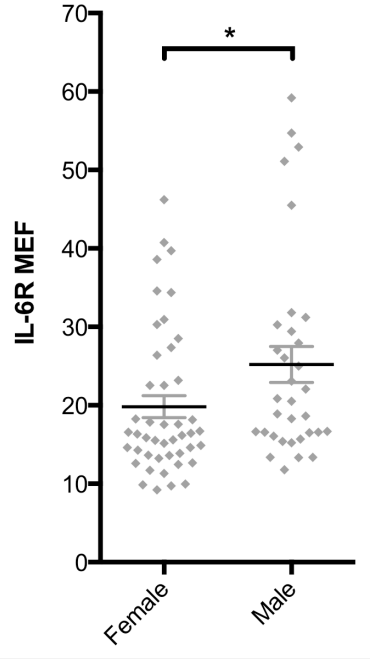


Figure 4.3.2. IL-6R expression in monocytes in samples from female and male donors
CD14+ monocytes: expression of IL-6R relative to isotype control staining and normalised to calibration beads. Significant effect of sex: males had significantly increased receptor expression on monocytes compared to females (unpaired t-test $P=0.0383$).

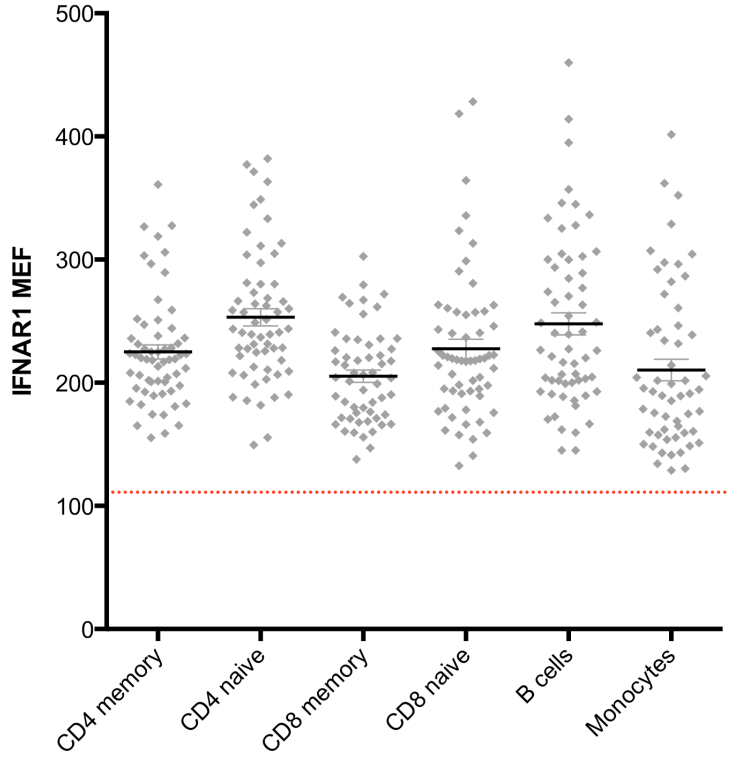
4.3.3. Quantifying cytokine receptor expression in TYK2-mediated pathways across immune cell subsets directly *ex vivo*

Based on the previous analyses, quantification of cytokine receptor expression across immune cell subsets involved normalising the receptor staining MFI to the isotype control MFI and also the calibration bead normalisation coefficient, thus giving a value for receptor staining as an MEF. Additionally, IL-6R expression by monocytes was normalised by sex. If cells were negative for a particular receptor, the resultant MEF value would thus be less than or equal to the calibration bead normalisation coefficient i.e. when receptor MFI/isotype MFI =1, the bead normalisation coefficient value represents the MEF when there is no cell surface receptor expression. The lowest

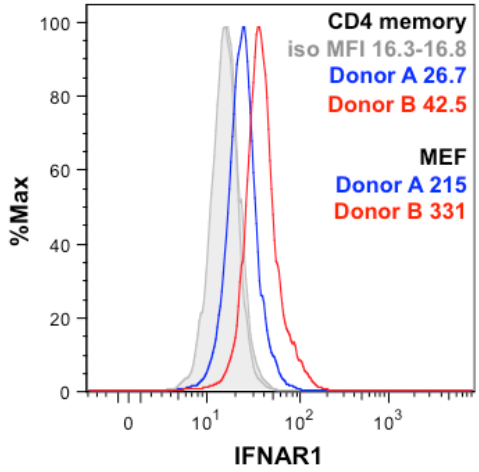
calibration bead normalisation coefficient value observed throughout the experimental period is indicated by a red dotted line in Figures 4.3.3., 4.3.4. and 4.3.5., which show IFNAR1, IL-6R and IL-10R expression across immune cell subsets, respectively. The Table in part c of each of these figures summarises the extent of variation in receptor expression between cell subsets.

All cell subsets from the individuals studied were positive for IFNAR1 (Figure 4.3.3.a), with a similar range of MEF values, however the mean receptor expression varied across cell subsets. For example, memory CD4+ T cells expressed significantly less IFNAR1 than naïve CD4+ T cells (P=0.0022; Figure 4.3.3.).

a)



b)



c)

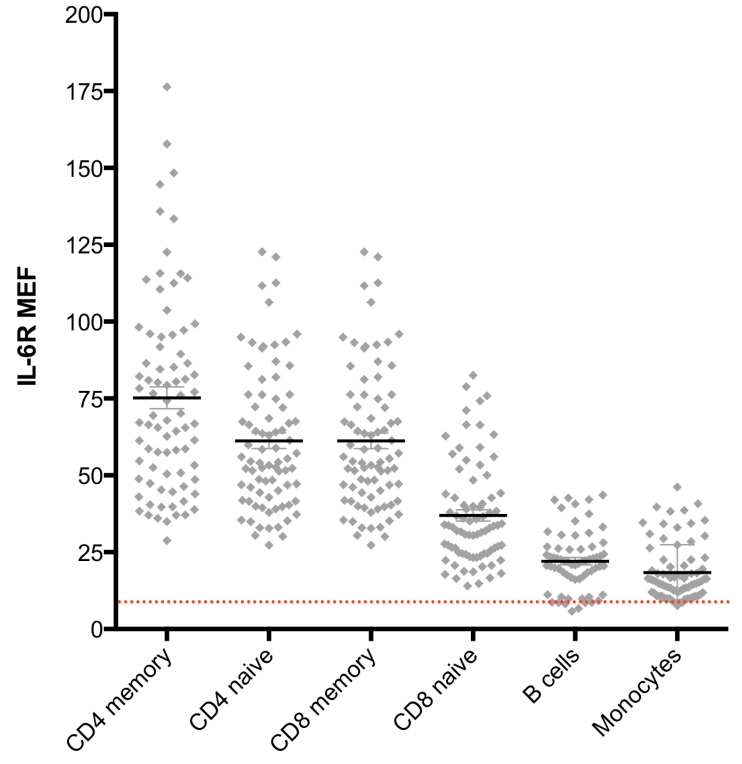
	n	Comparing IFNAR1 expression between cell subsets (*P*-value)				
		CD4+ naive	CD8+ memory	CD8+ naive	B cells	Monocytes
CD4+ memory	60	0.0022	0.0114	ns	0.0326	ns
CD4+ naive	60		<0.0001	0.0035	ns	0.0002
CD8+ memory	56			ns	<0.0001	ns
CD8+ naïve	60				0.0273	ns
B cells	60					0.0033
Monocytes	56					

Figure 4.3.3. Expression of IFNAR1 across primary human immune cell subsets
(a) IFNAR1 staining was normalised to the isotype control MFI and the calibration beads, and thus presented as an MEF value (y-axis) across immune cell subsets. The red dotted line indicates the lower limit of the calibration bead normalisation coefficient i.e. the MEF value that represents no receptor staining positivity. Mean normalisation coefficient = 141 and range 106.6-200. Graph shows mean ± SEM. (b) A representative histogram showing IFNAR1 staining for two individuals (red and blue lines) relative to the isotype control (filled grey line). The y-axis %Max value refers to a normalised value calculated by the FlowJo software to account for the exact number of events captured for each sample. (c) A summary table of the differences in IFNAR1 expression across cell subsets. *P*-values are from unpaired t-tests. ns = not significant ($P>0.05$). n = number per group.

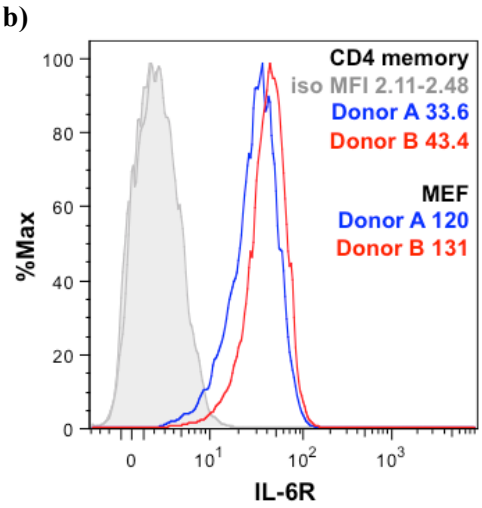
In contrast to IFNAR1, expression of IL-6R was highly variable across cell subsets with B cells and monocytes expressing significantly lower levels than T cell subsets, with some donors being negative for IL-6R in these cell subsets (Figure 4.3.4.). For both CD4+ and CD8+ T cells,

memory subsets expressed significantly more IL-6R than naïve cells (P=0.0016 and P<0.0001, respectively; Figure 4.3.4.).

a)



b)



c)

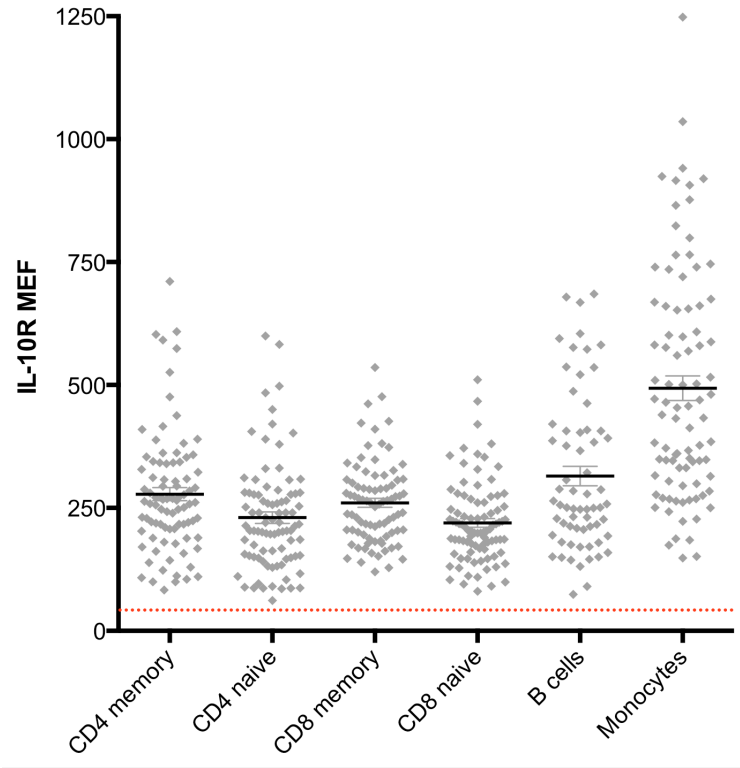
	n	Comparing IL-6R expression between cell subsets (*P*-value)				
		CD4+ naive	CD8+ memory	CD8+ naive	B cells	Monocytes
CD4+ memory	80	0.0016	0.0016	<0.0001	<0.0001	<0.0001
CD4+ naive	80		ns	<0.0001	<0.0001	<0.0001
CD8+ memory	80			<0.0001	<0.0001	<0.0001
CD8+ naïve	80				<0.0001	<0.0001
B cells	65					ns
Monocytes	76					

Figure 4.3.4. Expression of IL-6R across primary human immune cell subsets
(a) IL-6R staining was normalised to the isotype control MFI and the calibration beads, and thus presented as an MEF value (y-axis) across immune cell subsets. The red dotted line indicates the lower limit of the calibration bead normalisation coefficient i.e. the MEF value that represents no receptor staining positivity. Mean normalisation coefficient = 11.4 and range 6.4-33.4. Graph shows mean ± SEM. (b) A representative histogram showing IL-6R staining for two individuals (red and blue lines) relative to the isotype control (filled grey line). (c) A summary table of the differences in IL-6R expression across cell subsets. *P*-values are from unpaired t-tests. ns = not significant (P>0.05). n = number per group.

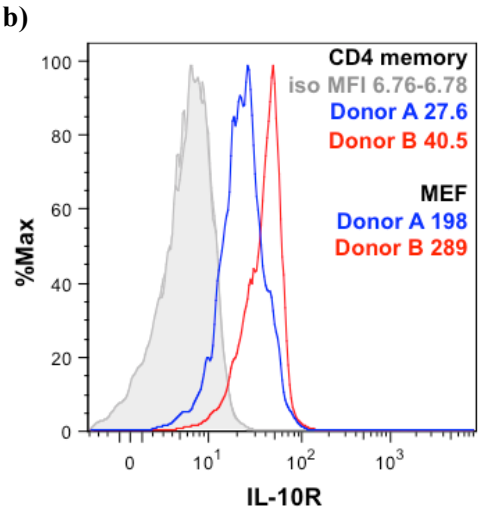
IL-10R expression was significantly higher in monocytes compared to all lymphocyte cell subsets (P<0.0001 for all; Figure 4.3.5.). For both CD4+ and CD8+ T cells, the memory subset

expressed significantly higher IL-10R than naïve cells (P=0.007 and P=0.0016, respectively). In contrast to IFNAR1, IL-6R and IL-10R, expression of IL-13R was only detectable on CD14-positive monocytes (Figure 4.3.6.).

a)



b)



c)

		Comparing IL-10R expression between cell subsets (*P*-value)				
	n	**CD4+ naive**	**CD8+ memory**	**CD8+ naive**	**B cells**	**Monocytes**
CD4+ memory	87	0.007	ns	0.0003	ns	<0.0001
CD4+ naive	88		0.0456	ns	0.0001	<0.0001
CD8+ memory	83			0.0016	0.007	<0.0001
CD8+ naïve	89				<0.0001	<0.0001
B cells	64					<0.0001
Monocytes	88					

Figure 4.3.5. Expression of IL-10R across primary human immune cell subsets
(a) IL-10R staining was normalised to the isotype control MFI and the calibration beads, and thus presented as an MEF value (y-axis) across immune cell subsets. The red dotted line indicates the lower limit of the calibration bead normalisation coefficient i.e. the MEF value that represents no receptor staining positivity. Mean normalisation coefficient = 75.9 and range 45.7-120.8. Graph shows mean ± SEM. (b) A representative histogram showing IL-10R staining for two individuals (red and blue lines) relative to the isotype control (filled grey line). (c) A summary table of the differences in IL-10R expression across cell subsets. *P*-values are from unpaired t-tests. ns = not significant (P>0.05). n = number per group.

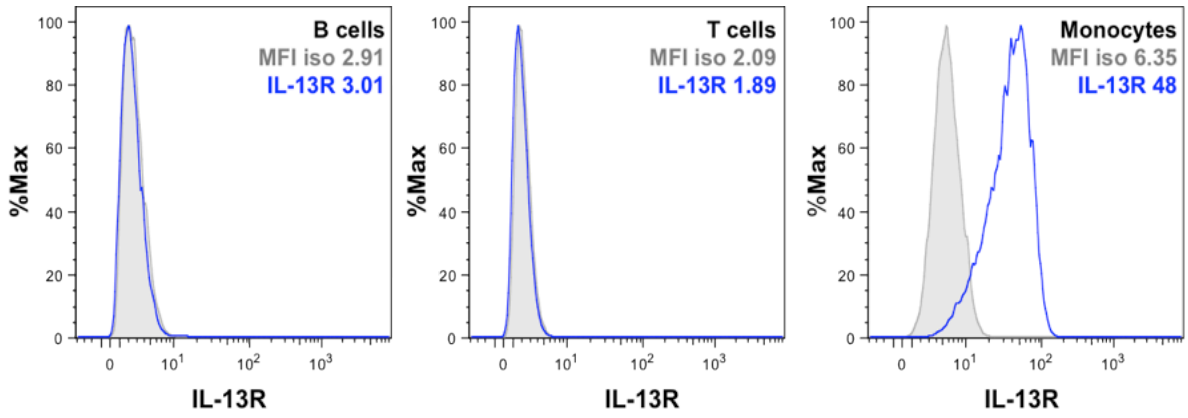


Figure 4.3.6. IL-13R is expressed in monocytes but not in T and B cells
Representative staining histograms are shown for expression of IL-13R (blue line) relative to the isotype control (filled grey line) in B cells, T cells and monocytes. The y-axis %Max value refers to a normalised value calculated by the FlowJo software to account for the exact number of events captured for each sample.

4.3.4. Investigating genotype-dependent differences: study design and cytokine receptor expression

A key aim of this thesis was to elucidate the functional consequences of the rs34536443 and rs12720356 SNPs on TYK2-mediated cytokine signalling, and therefore the existence of any genotype-dependent differences in cytokine receptor level were also investigated, to determine whether putative genotype-dependent differences in TYK2 signalling were likely to be directly due to the effect of the nsSNPs on TYK2 activity rather than to any effects on receptor expression.

The total sample size required was based on the lowest r^2 value (~0.2) observed for genotype-to-phenotype correlations from previous studies of autoimmune disease-associated SNPs (Dendrou et al., 2009b; Gregory et al., 2012) with ORs in the same range as those for the *TYK2* nsSNPs. For this r^2 value and using a conservative significance threshold incorporating multiple testing considerations (P=0.025 for receptor level analyses and P=0.0083 for pSTAT analyses; see Chapter 2, section 2.16.2.), and aiming for 80% study power, the estimated sample size needed was ~45 individuals to assess genotype-dependent differences in receptor levels and ~55 individuals to assess pSTAT level variation by genotype. The total sample size in the presented work was 94.

Given the low MAF of these two SNPs (0.05 and 0.10, respectively) (Hinks et al., 2013; Tsoi et al., 2012), it was necessary to pre-select donors from a large cohort in collaboration with the

Oxford BioBank. If donors were selected randomly from the general population, a much larger sample size would be required in order to obtain enough individuals with the rarer genotypes for an accurate assessment of potential genotype-to-phenotype correlations in primary immune cells. Comparing the genotype frequencies that would be expected for the rs34536443 and rs12720356 SNPs in a healthy control population to the frequencies observed in the selected cohort, pre-selection enabled enrichment of individuals homozygous for the minor allele by 12-fold and 4-fold, respectively (highlighted blue in Table 4.2.). Moreover, as the chance of a randomly selected individual being homozygous for the minor allele of either the rs12720356 or rs34536443 SNP is 1% and 0.25%, respectively, a cohort recruited by random selection may not have enabled study of these genotypes in our total sample size of 94.

All selected donors were of the same ethnicity (white European) and were all of self-reported non-autoimmune disease status, thus avoiding potential confounding effects on genotype-dependent phenotypic differences that could be due to the presence of an inflammatory condition or due to the action of immunomodulatory treatments.

Table 4.2. Pre-selection of donors from a genotyped cohort for investigations of TYK2 function by nsSNP genotype
Donors were selected from a large cohort of individuals genotyped for rs34536443 and rs12720356. Expected genotype frequencies were calculated from cited allele frequencies (IMSGC, 2013; Tsoi et al., 2012) and assuming Hardy-Weinberg equilibrium. Observed frequency refers to values calculated from the total cohort of donors recruited to provide samples. Risk = homozygous for the risk allele. Het = heterozygote. Prot = homozygous for the protective allele. For the rs12720356 SNP, the direction of disease association refers to IBD (PS is associated in the opposite direction; see Chapter 1, Table 1.1.).

SNP (abbreviation)	**MAF**	**Expected frequency (%)**			**Observed frequency (%)**			**Fold enrichment**		
		Risk	**Het**	**Prot**	**Risk**	**Het**	**Prot**	**Risk**	**Het**	**Prot**
rs34536443 (rs345..)	0.05	90.3	9.60	0.25	78.00	18.90	3.10	0.86	1.97	12.40
rs12720356 (rs127..)	0.10	1.00	18.00	81.00	4.40	16.30	79.4	4.40	0.91	0.98

Table 4.3. summarises the age and sex of individuals in each genotype group of the selected cohort. Individuals homozygous for the major alleles at both SNPs had a haplotype permitting their use in both the rs345..Risk and rs127..Prot genotype groups. There was no significant difference in

age across groups for the rs34536443 genotype, however the mean age was significantly lower in individuals homozygous for the protective allele of the rs12720356 SNP compared to heterozygous individuals (P=0.0045). Based on previous data, there was no effect of age on cytokine receptor expression across cell subsets. The number of males and females was not equal for each genotype group (Table 4.3.) and thus receptor expression data were only normalised by sex where determined to be necessary (i.e. monocyte IL-6R).

Table 4.3. The age and sex of individuals forming each genotype group
The mean age and age range are shown to the nearest year for each genotype group. There were unequal numbers of males and females in each genotype group.

Genotype	Mean age (age range) in year	Number of males:females
rs345..Risk / rs127..Prot	45.1 (28-56)	22:24
rs345..Het	47 (25-60)	11:12
rs345..Prot	41 (28-52)	2:3
rs127..Het	51 (41-63)	5:8
rs127..Risk	45 (37-53)	3:4

Cytokine receptor expression data were plotted by donor genotype following normalisation as previously determined (sections 4.3.1. and 4.3.2.). For all cell subsets, there were no significant trends by rs34536443 or rs12720356 genotype for IFNAR1, IL-6R, IL-10R and IL-13R (P>0.025 for all; Table 4.4.). This suggests that any potential differences in phospho-signalling by genotype would not be due to altered receptor expression by immune cells, and that the readout of STAT phosphorylation was thus be more likely to directly reflect the activity of TYK2.

Table 4.4. Receptor expression across cell subsets: the effect of rs34536443 and rs12720356 genotype

For each receptor and cell subset, the number of individuals in each analysis is provided (n). There were no significant trends by rs34536443 or rs12720356 genotype for IFNAR1, IL-6R, IL-10R and IL-13R across immune cell subsets. ns = not significant (P>0.025).

		rs345.. genotype				rs127.. genotype			
		n per group				n per group			
		Risk	Het	Prot	*P*-value	Risk	Het	Prot	*P*-value
IFNAR1	CD4+ memory	29	9	3	ns	6	13	29	ns
	CD4+ naive	29	10	3	ns	6	12	29	ns
	CD8+ memory	29	8	3	ns	6	10	29	ns
	CD8+ naïve	29	10	3	ns	7	13	29	ns
	B cells	29	9	3	ns	7	13	29	ns
	Monocytes	28	7	3	ns	7	11	28	ns
IL-6R	CD4+ memory	38	17	5	ns	7	13	38	ns
	CD4+ naive	38	17	5	ns	7	13	38	ns
	CD8+ memory	38	15	5	ns	6	12	38	ns
	CD8+ naïve	38	17	5	ns	7	13	38	ns
	B cells	31	11	3	ns	7	13	31	ns
	Monocytes	37	17	4	ns	7	13	37	ns
IL-10R	CD4+ memory	42	22	4	ns	7	12	42	ns
	CD4+ naive	42	22	4	ns	7	12	42	ns
	CD8+ memory	42	20	5	ns	6	11	42	ns
	CD8+ naïve	43	22	5	ns	7	13	43	ns
	B cells	31	11	3	ns	7	12	31	ns
	Monocytes	42	22	4	ns	7	12	42	ns
IL-13R	Monocytes	21	16	4	ns	-	-	-	-

4.3.5. Measuring STAT phosphorylation by flow cytometry: a readout of TYK2 activity

Fluorescently conjugated antibodies against phosphorylated STAT molecules were used to quantify TYK2-mediated cytokine signalling by flow cytometry, in samples simultaneously stained for the immune cell subset markers used previously for characterisation of receptor expression. For the selected cytokine pathways (IFN-β, IL-6, IL-10 and IL-13), the main STAT molecule(s) known to mediate signalling were investigated. STAT3 is the main mediator of intracellular signalling downstream of IL-6 and IL-10, whereas STAT6 mediates IL-13 signalling (Kisseleva et al., 2002). Although STAT1 and STAT2 are considered to be the main mediators of type 1 IFN signalling, STAT3 and STAT5 are also phosphorylated in this pathway (as reviewed by (Platanias, 2005)).

Previous work has demonstrated reduced IFN-β-induced pSTAT2 signalling in expanded T cells derived from individuals heterozygous for the minor allele of the rs34536443 SNP compared to those homozygous for the major allele, but no apparent effect on phosphorylation of STAT1 was observed (Couturier et al., 2011). However, these analyses were performed with a small sample size (n=19), and phosphorylation was assayed by western blotting whereas the presented work aimed to measure genotype-dependent differences in pSTAT levels by flow cytometry, which is a quantitative approach with greater sensitivity. Therefore, anti-pSTAT antibodies were assessed for suitability for use in flow cytometric analyses.

Phospho-specific antibodies are available for use in flow cytometry that bind pSTAT1 and 3-6, with such binding being observed upon cytokine-induced phosphorylation (Figure 4.3.7.). In contrast, there is no widely used antibody against pSTAT2 available specifically for use in flow cytometry, however a polyclonal antibody (R&D Systems, UK) that has one published use for this application (Thacker et al., 2010) was tested. Although robust IFN-β-induced phosphorylation of STAT1, STAT3 and STAT5 was consistently observed, there was no detectable positive pSTAT2 staining in cells stimulated under the same conditions (Figure 4.3.7.). In the absence of a robust antibody for flow cytometric quantification of cytokine-induced pSTAT2, phosphorylation of the other main mediator of IFN-β-induced signalling, STAT1 was quantified. As pSTAT1 levels were not reported to correlate with rs34536443 genotype in previous work (Couturier et al., 2011), phosphorylation of STAT3 was also assessed in response to IFN-β stimulation.

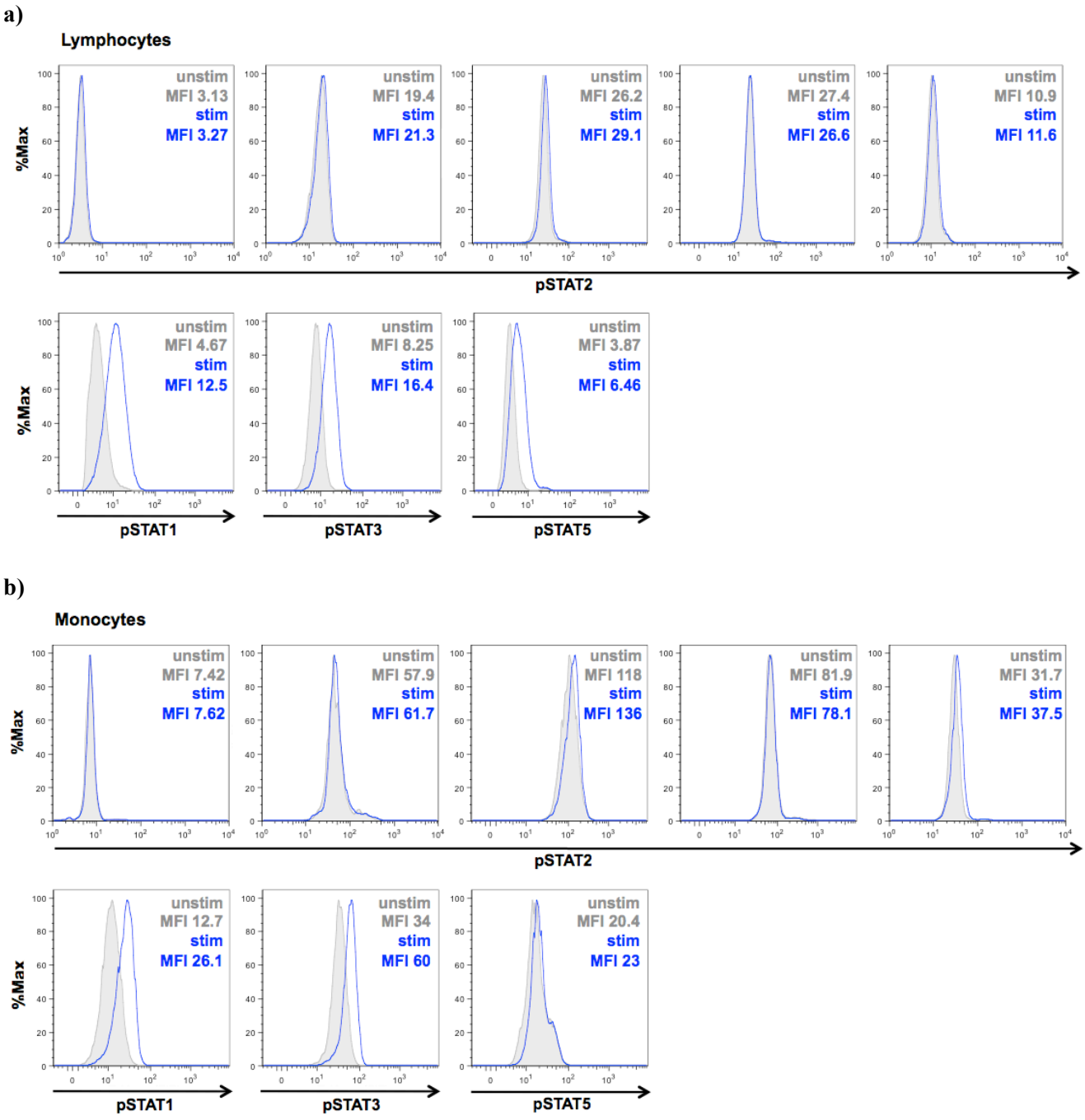


Figure 4.3.7. IFN-β-induced phosphorylation of different STAT molecules
Following a 15 minute stimulation with 100 ng/ml IFN-β, PBMCs were stained for intracellular pSTAT levels. Representative histograms are shown for pSTAT1, pSTAT3 and pSTAT5, and pSTAT2 staining from five different individuals is shown. (a) In contrast to the robust signal for pSTATs 1, 3 and 5 observed in stimulated (blue line) compared to unstimulated (grey line, filled) lymphocytes there was no observed shift in the MFI of pSTAT2 (n=5). (b) Similarly, there was no robust IFN-β-induced increase in pSTAT2 levels in monocytes. The monocyte pSTAT5 response was also minimal compared to that of pSTAT1 and pSTAT3, consistent with previously published data (van Boxel-Dezaire et al., 2010). The y-axis %Max value refers to a normalised value calculated by the FlowJo software to account for the exact number of events captured for each sample.

Having established the pSTAT staining to be used and characterised the variation in cytokine receptor levels, the optimal stimulation conditions required to achieve a maximal STAT response were considered, in terms of both cytokine concentration and the duration of stimulation required. A suboptimal cytokine stimulation, whereby only some of the receptor-positive cells respond, could introduce further variation in pSTAT levels throughout the experiment, thus limiting the ability to detect potential genotype-dependent differences in TYK2-mediated signalling.

Although it is not possible to directly correlate cytokine receptor and pSTAT levels, the previous characterisation of receptor expression demonstrated variation across immune cell subsets (section 4.3.3.). Thus, the analysis of specific subsets could be used as an approach to reduce the variation in receptor levels for any single pSTAT measurement. The cytokine concentration required for maximal STAT phosphorylation across cell subsets was investigated for IFN-β, IL-6 and IL-10 and the proportion of cells that were pSTAT-positive (pSTAT+; for gating see Figure 4.3.9.) had reached a plateau between 100-1000 ng/ml of cytokine, indicating saturation across all cell subsets at this range of concentrations (Figure 4.3.8.a-h). A plateau in pSTAT responsiveness was also reached in monocytes at similar concentrations of IL-13 (Figure 4.3.8.i), noting that only this cell subset expresses IL-13R directly *ex vivo*. At a chosen high concentration of 400 ng/ml of cytokine, the peak pSTAT response to stimulation with IFN-β, IL-6 and IL-10 was reached at 10-15 minutes (Figure 4.3.8.j), in line with previously published studies of cytokine-induced phospho-signalling (Montag and Lotze, 2006). Thus in subsequent work, stimulation of PBMCs was performed using a cytokine concentration of 400 ng/ml for 15 minutes.

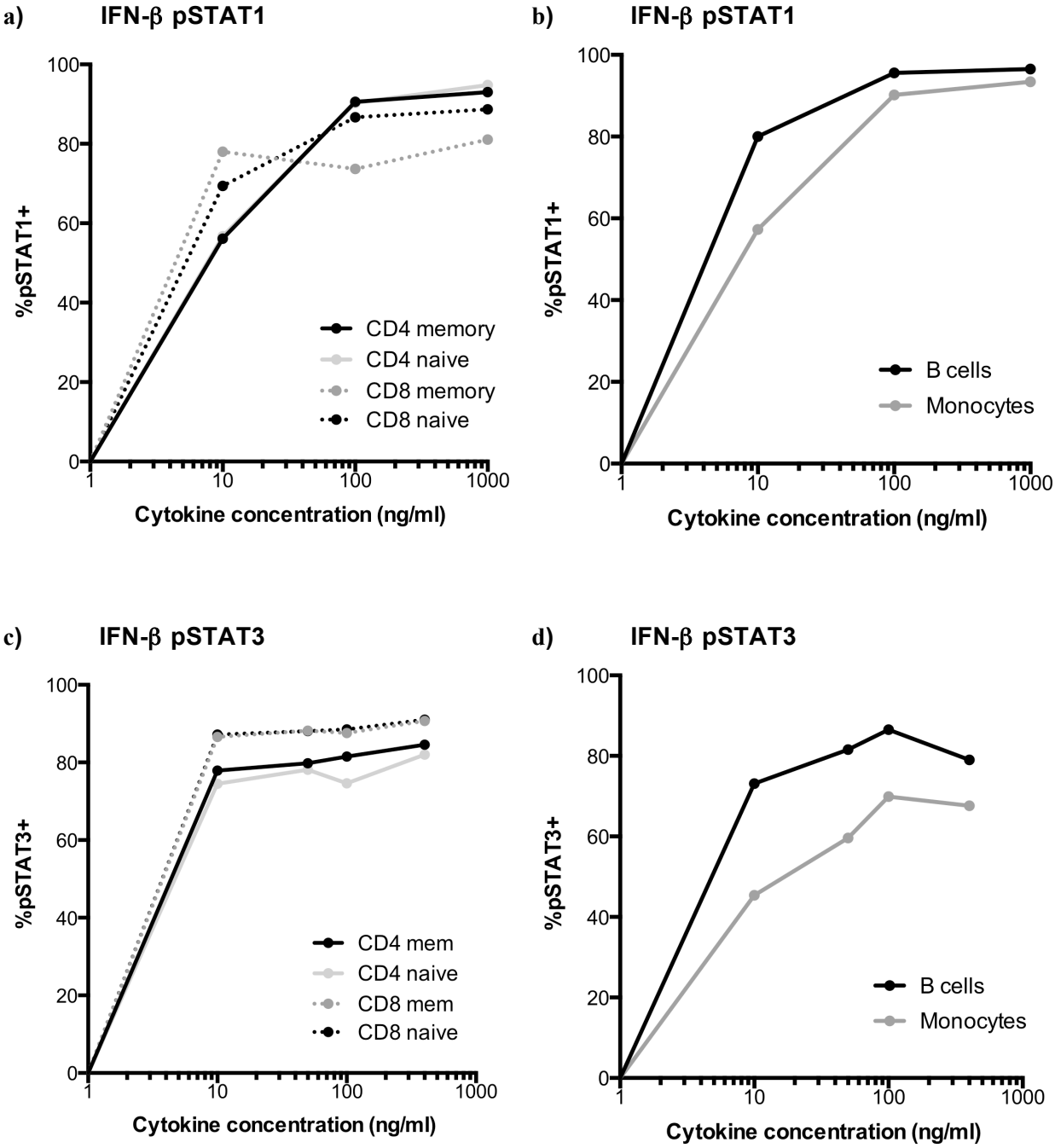


figure continued on next page…

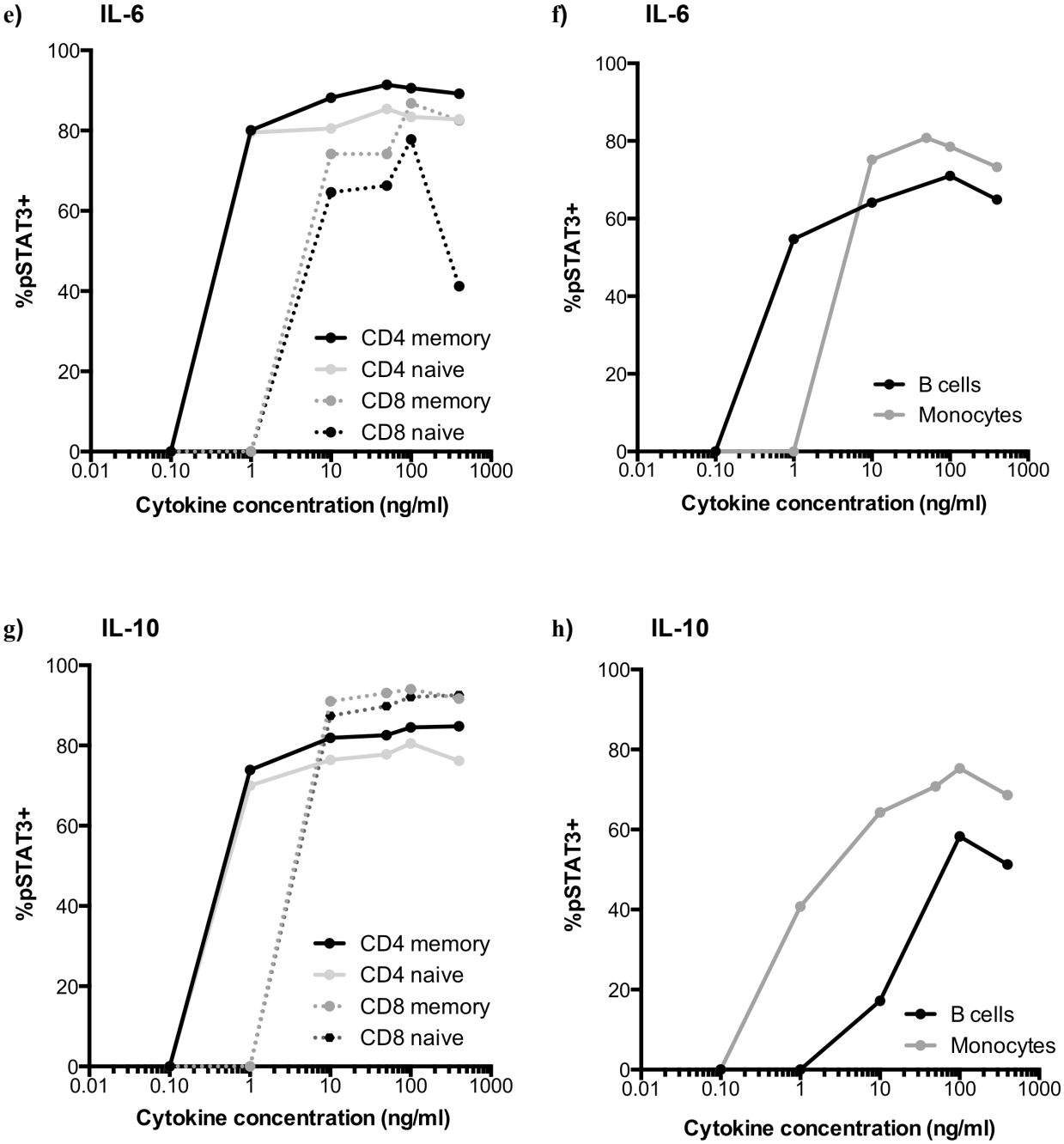


figure continued on next page…

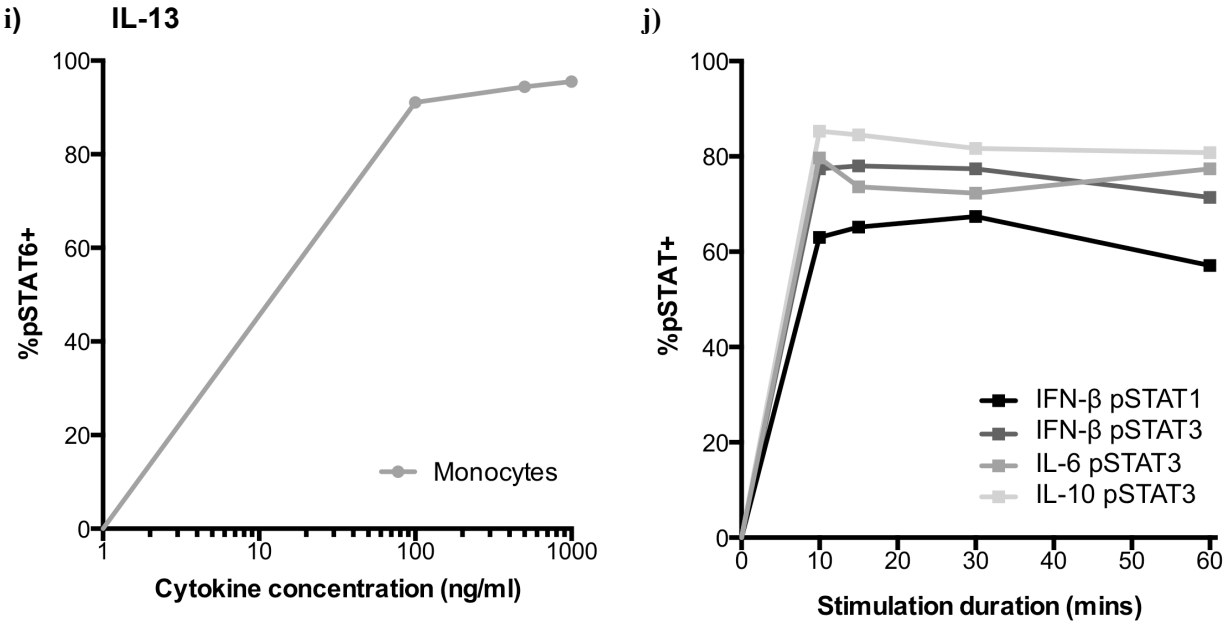


Figure 4.3.8. Optimising cytokine stimulation of PBMCs to achieve maximal pSTAT response
Representative data from cytokine titration and stimulation timecourse experiments. (a-i) Cytokine titrations were performed for IFN-β, IL-6, IL-10 and IL-13, across cell subsets known to express the respective receptors. The proportion of cells becoming pSTAT+ (relative to unstimulated cells; see Figure 4.3.9. for gating) reached a plateau indicating saturation across all cell subsets at concentrations between 100-1000 ng/ml. (j) A timecourse was performed for IFN-β, IL-6 and IL-10 to determine the duration of stimulation achieving maximal pSTAT response. In line with previously published studies of cytokine-induced phospho-signalling (Montag and Lotze, 2006), the optimal timepoint was between 10-15 minutes for all cell subsets and cytokines. Representative data from CD4+ T cells is shown, except for IL-13 signalling which was only detectable in monocytes.

For the data presented in Figure 4.3.8., cytokine-induced STAT phosphorylation is measured as the proportion of cells becoming pSTAT+ in stimulated cells relative to unstimulated cells. These pSTAT+ values were obtained from overlaid histograms for the staining of pSTAT in stimulated and unstimulated samples, with the pSTAT+ population being defined as the population with pSTAT fluorescence intensity values above the intersection of the unstimulated and stimulated pSTAT histograms (Figure 4.3.9.a). As well as measuring the proportion of cells becoming pSTAT+, the assessment of genotype-dependent differences in pSTAT signalling will also include measurement of the MFI of the pSTAT+ population relative to unstimulated cells (+pop/unstim). Whilst these measures are inter-related they do, however, allow an investigation of whether genotype-dependent differences in TYK2 activity mainly influence the likelihood of a cell passing an activation threshold

to become pSTAT+ (Figure 4.3.9.b), whether they influence the overall level of pSTAT obtained on a per cell basis (Figure 4.3.9.c) or both. The fold change in pSTAT MFI of the total population upon stimulation was also assessed relative to unstimulated cells (stim/unstim), as this is a measure that does not rely on gating of pSTAT+ cells and thus cannot be influenced by any putative gating biases.

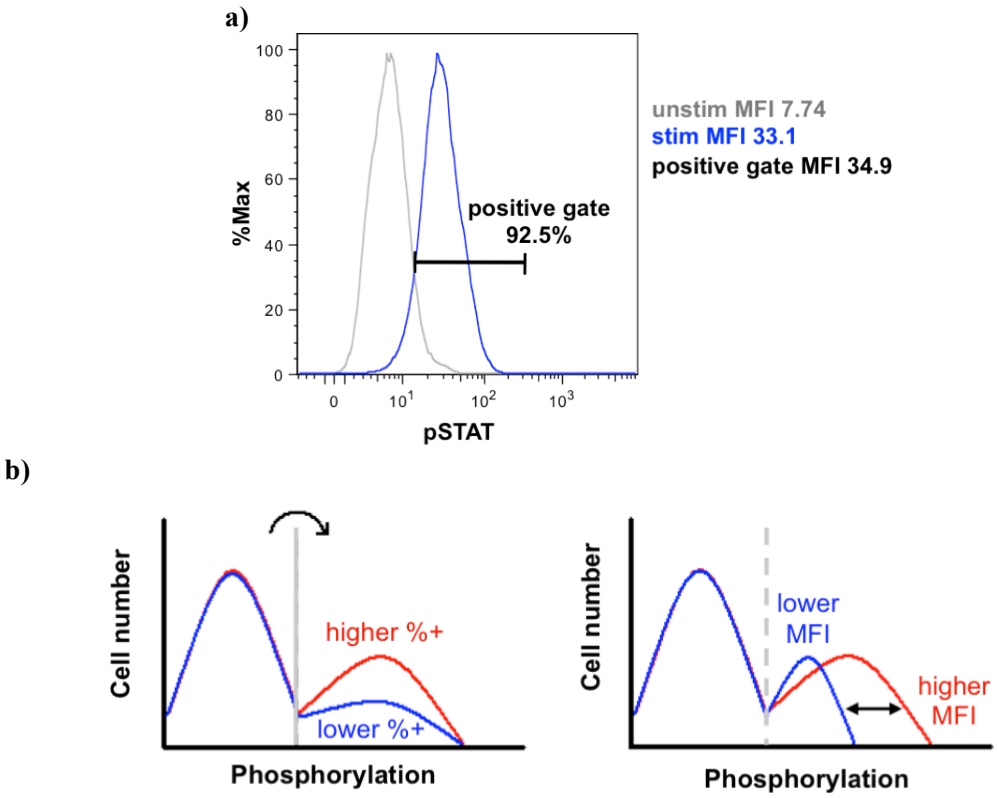


Figure 4.3.9. STAT phosphorylation gating examples
(a) The pSTAT staining histograms of the unstimulated sample (unstim; grey line) and stimulated sample (stim; blue line) from a single donor were overlaid and the pSTAT+ population defined as the population with x-axes values gated from the intersection of the histograms to the upper limit of the stimulated population i.e. in this representative example 92.5% of the cells are pSTAT+. The MFI values of the stimulated population and the pSTAT+ population are also shown and were normalised to the pSTAT MFI of unstimulated cells to provide further measures of TYK2-mediated cytokine signalling. In this case, the stimulated cells show a 4.3-fold increase in pSTAT upon stimulation and the pSTAT+ cells show a 4.5-fold increase relative to unstimulated cells. The y-axis %Max value refers to a normalised value calculated by the FlowJo software to account for the exact number of events captured for each sample. (b) Using these inter-related but distinct measures of STAT phosphorylation will help to delineate whether potential genotype-dependent changes in TYK2 activity influence both or either the probability of cells within a population becoming pSTAT+ (left panel) and/or the level (MFI) of pSTAT achieved upon stimulation (right panel).

As pSTAT levels in stimulated cells are normalised to the unstimulated condition, the latter provides a staining control, and thus there was no need for an isotype control as was necessary for analyses of receptor staining. For quantification of receptor levels, the isotype control staining did not completely account for experimental variation due to changes in flow cytometer function (Table 4.1.) and therefore normalisation to the calibration bead coefficient was required. To determine whether such a normalisation would also be necessary for measurement of STAT phosphorylation, the degree to which the pSTAT levels of the unstimulated cells could reflect such experimental variation was determined.

The relationship between the calibration bead normalisation coefficients and the pSTAT MFI for both the unstimulated and stimulated cells was highly significant (Table 4.5.). The proportion of variability in pSTAT MFI reflected by the variation in the bead normalisation coefficient values being ~11% for pSTAT1 (fluorescence detected in the APC channel and stimulated by the red 642 nm flow cytometer laser), and ~55% for pSTAT3 (fluorescence detected in the FITC channel and stimulated by the blue 488 nm flow cytometer laser). It is noteworthy that the r^2 values obtained for the correlation between the calibration bead normalisation coefficient values and the unstimulated or stimulated pSTAT MFIs are similar, indicating that the experimental variation captured by the beads equally affected the pSTAT MFI measurement under both conditions. Therefore, there was no significant relationship between the calibration bead normalisation coefficients and value calculated for the fold pSTAT change in stimulated relative to unstimulated cells (Table 4.5.), indicating that normalising the pSTAT measures to the bead coefficients was unnecessary and would be redundant given the performed normalisation between the stimulated and unstimulated conditions.

Table 4.5. Determining whether to normalise fold-change in pSTAT MFI to the calibration beads

The normalisation coefficient calculated using the calibration beads was plotted against the MFI of unstimulated cells or stimulated cells, or the stimulation-induced fold change in pSTAT MFI (stim/unstim). The r^2 of the observed linear relationships and the extent of significance (*P*-value) are shown for IFN-β, IL-6 and IL-10 signalling in memory CD4+ T cells which are representative of all immune cell subsets studied.

	unstim MFI vs beads		stim MFI vs beads		stim/unstim vs beads	
Stimulation & readout	r^2	***P*-value**	r^2	***P*-value**	r^2	***P*-value**
IFN-β pSTAT1	0.1054	0.0018	0.1298	0.0004	0.0016	0.7047
IFN-β pSTAT3	0.5542	<0.0001	0.5644	<0.0001	0.0054	0.4895
IL-6 pSTAT3	0.5250	<0.0001	0.5610	<0.0001	0.0073	0.4756
IL-10 pSTAT3	0.5326	<0.0001	0.6083	<0.0001	0.00004	0.9471

4.4 Discussion

The data presented in this chapter were generated with the purpose of determining the optimal study design to enable the identification of any potential differences in TYK2-mediated signalling that are due to rs12720356 and rs34536443 nsSNP genotype (Chapter 5). This required a consideration of the variables that influence cytokine responsiveness of PBMC subsets, including:

- variables that cannot be controlled but can be accounted for – such as experimental variation, that can be minimised by appropriate data normalisation; and variation in receptor levels that could be segregated, and thus reduced, by considering distinct immune cell subsets.
- unknown inter-individual variation – such as differences in genetic background and environmental influences whose effects can be reduced (relative to putative *TYK2* nsSNP-dependent effects) by using an appropriate sample size and enriching for the genotype groups of interest.
- variables that can be controlled – such as the cytokine concentration and the duration of stimulation.

Having determined how to analyse receptor expression across individuals through time and minimise experimental variation (sections 4.3.1. and 4.3.2.), there were found to be no correlations

between receptor expression and genotype, for either the rs12720356 or rs34536443 SNP, across all immune cell subsets considered (section 4.3.4.). This indicates that any observed correlation between nsSNP genotype and pSTAT levels was more likely to reflect altered TYK2 activity, rather than differential receptor expression. Although not all PBMC subsets were being studied (e.g. NK cells, DCs), it was hypothesised that the effects of *TYK2* nsSNP genotype would be seen across cell types and not be subset-specific, apart from where receptor expression is restricted (e.g. IL-13R on monocytes; Figure 4.3.6.). By using markers of the most abundant immune cell subsets, this hypothesis was addressed and, as discussed, receptor level variation minimised.

The IFN-β, IL-6, IL-10 and IL-13 signalling pathways were selected as they represent the main cytokine sub-families (Leonard, 2001) and also the receptors for these cytokine pathways are expressed on PBMCs directly *ex vivo* (section 4.3.3.) without the need for cell activation or expansion. Activation of cells is required for the study of IL-12 and IL-23 signalling, as the shared IL-12Rβ1 subunit is not expressed on resting T cells (Fahey et al., 2007). Such activation-induced upregulation of receptor expression may introduce additional variation in phospho-signalling, making the detection of potential genotype-dependent differences more difficult. Therefore, optimisation of stimulation conditions to achieve robust receptor upregulation across individuals is ongoing, with a view to investigating IL-12 and IL-23 signalling by *TYK2* nsSNP genotype, as these pathways are also relevant in the context of autoimmune disease (see Chapter 5, section 5.4.).

Following the characterisation of receptor expression for the selected TYK2-mediated signalling pathways, the optimal stimulation conditions were determined (section 4.3.5) in order to detect peak responsiveness to IFN-β, IL-6, IL-10 and IL-13 in subsequent studies of cytokine-induced STAT phosphorylation by nsSNP genotype (Chapter 5). By using three inter-related but distinct measures of pSTAT, the potential influence of *TYK2* nsSNP genotype on cytokine signalling was assessed in terms of the proportion of cells responding to stimulation as well as the extent of STAT phosphorylation on a per cell basis (section 4.3.5.).

In summary, the work presented in this Chapter has determined the experimental methods and approach to data analysis, and has considered the overall study design with respect to donor selection to enrich for the minor alleles of the rs12720356 and rs34536443 SNPs, as well as

calculating the estimated sample size required to achieve 80% study power (section 4.3.4.). Thus, a cohort of 94 donors that have been pre-selected by genotype provided samples for investigation of potential nsSNP-dependent effects on TYK2-mediated cytokine-induced phospho-signalling in the following Chapter.

5. Genotype-to-phenotype correlations: TYK2-mediated cytokine signalling in primary human immune cell subsets by rs12720356 and rs34536443 genotype

5.1 Introduction

The rs12720356 and rs34536443 nsSNPs are associated with multiple autoimmune diseases, and although the precise molecular mechanisms underpinning these associations remain to be fully elucidated, there is evidence to suggest that these variants affect TYK2 activity (Chapter 3 and (Couturier et al., 2011; Li et al., 2013)). TYK2 mediates signalling for a variety of cytokines and as demonstrated in the previous Chapter for the signalling pathways selected for investigation (IFN-β, IL-6, IL-10 and IL-13), there is no correlation between nsSNP genotype and receptor expression in human PBMC subsets directly *ex vivo* (Chapter 4, section 4.3.4.). This indicates that potential genotype-dependent effects on signalling are more likely directly reflect TYK2 activity. Having optimised the conditions for cytokine stimulation and the approach to phospho-signalling data analysis, the potential consequences of the rs12720356 and rs34536443 nsSNPs on TYK2-mediated cytokine signalling were investigated.

TYK2 always mediates cytokine-induced signalling alongside another JAK (or JAKs) and the extent of compensation between JAK family members (with respect to how differences in the level or activity of any single JAK family member influences a specific cytokine signalling pathway) is not well characterised. A complete lack of TYK2 protein has been shown to cause significant defects in multiple cytokine signalling pathways, resulting in a severely immunodeficient phenotype in KO mice and in the single reported case of human TYK2 deficiency that has been fully characterised (Karaghiosoff et al., 2000; Minegishi et al., 2006; Shimoda et al., 2000). However, given that the minor alleles of the rs12720356 and rs34536443 SNPs are found in healthy controls (as these are common nsSNPs), their functional consequences are likely to involve more subtle

differences in TYK2 activity, noting that neither of these SNPs causes a substantial absence of TYK2 protein (Chapter 3; (Couturier et al., 2011; Li et al., 2013)).

Although there are no published studies of TYK2 signalling in primary human immune cells by rs12720356 genotype, one study has reported reduced IFN-β signalling in expanded T cells from individuals heterozygous for the minor allele of rs34536443 compared to individuals homozygous for the common allele (Couturier et al., 2011). This study also reported genotype-dependent effects on IL-6- and IL-10-induced gene expression, suggesting a broad effect of rs34536443 genotype across cytokine signalling pathways. However, a more thorough investigation of potential genotype-dependent effects on TYK2-mediated signalling was warranted as these data are from expanded T cells, that have been exposed to *in vitro* culturing conditions for a prolonged period and thus may not be representative of cytokine signalling in primary human immune cells. As the receptors for the aforementioned cytokines are expressed by PBMC subsets directly *ex vivo* (Chapter 4, section 4.3.3.), cell expansion is not necessary for investigation of these signalling pathways and may have introduced additional variation in the aforementioned study (Couturier et al., 2011). Furthermore, as mentioned in Chapter 4, phosphorylation was assayed by western blotting (Couturier et al., 2011) whereas the presented work aimed to measure genotype-dependent differences in cytokine-induced phospho-signalling by flow cytometry, which is a quantitative approach with greater sensitivity. The presented approach also used a proximal readout of signalling (i.e. pSTAT levels) rather than quantifying gene expression induced further downstream of cytokine stimulation, as had been performed for previous work with only a small sample size (n=19) (Couturier et al., 2011).

For the presented flow cytometric analyses of TYK2-mediated cytokine signalling, the cohort of donors providing samples was selected by genotype from a healthy control population. This achieved a manifold enrichment of donors homozygous for the minor allele of either the rs34536443 or the rs12720356 SNP compared to the genotype frequencies that would be expected via random selection of individuals (Chapter 4, section 4.3.4.). Notably, to our knowledge the presented work is the first to investigate cytokine signalling in cells derived from individuals homozygous for the minor allele of the rs34536443 SNP and the first to explore the potential effects of rs12720356 genotype on primary human immune cell signalling. The total sample size (n=94) used for the

presented analyses of phospho-signalling by genotype exceeded that estimated to achieve 80% study power (Chapter 4, section 4.3.4.) and was also greater than previously published investigations of the functional effects of rs34536443 genotype (n=19) (Couturier et al., 2011).

5.2 Chapter aim

The aim of this chapter was to investigate potential nsSNP genotype-dependent effects on TYK2-mediated cytokine signalling in human immune cell subsets directly *ex vivo*. Using the previously determined experimental conditions and approach to data analysis (from Chapter 4), flow cytometry was used to quantify cytokine-induced STAT phosphorylation as a proximal readout of TYK2 activity in the IFN-β, IL-6, IL-10 and IL-13 signalling pathways in samples from donors pre-selected by rs12720356 and rs34536443 genotype.

5.3 Results

As described previously, the rs34536443 and rs12720356 SNPs will be abbreviated in figures and data tables to rs345.. and rs127.., respectively. For the rs12720356 SNP, the disease association (Risk/Het/Prot) refers to IBD (PS is associated in the opposite direction; see Chapter 1, Table 1.1.). In order to correct for multiple testing, the *P*-value threshold for significance of genotype-dependent correlations was set at 0.0083 (see Chapter 2, section 2.16.2). For the effects of age and sex the significance threshold used was $P<0.05$. Any outliers were excluded from the analyses as detailed in Chapter 2, section 2.13.

5.3.1. Investigating genotype-dependent variation in cytokine-induced STAT phosphorylation in immune cell subsets directly *ex vivo*: the effects of age and sex

Before considering potential correlations between nsSNP genotype and TYK2-mediated cytokine signalling it was important to consider the effects of age and sex, which could introduce variation into the data. With a view to normalising signalling data to minimise the influence of these potentially confounding factors, statistical analyses were performed for all immune cell subsets and cytokine signalling pathways under investigation. Three different (although not entirely independent) measures of STAT phosphorylation were considered (i.e. %pSTAT+, stim/unstim and +pop/unstim; see Chapter 4, Figure 4.3.9.) and Table 5.1. summarises the significant findings. Interestingly, although IL-6R expression by monocytes was found to vary by sex in the previous Chapter (Figure 4.3.2.), no such trend was found in IL-6-induced pSTAT3 levels, suggesting that the sex-dependent effect on IL-6R expression was not substantial enough to translate into a sex-dependent effect on phospho-signalling. For significant trends (Table 5.1.), the respective pSTAT data were normalised to donor age or sex prior to data analysis by genotype as required.

Table 5.1. Significant effects of age and/or sex on cytokine-induced pSTAT levels
For all cytokine pathways and cell subsets under investigation, statistical analyses were performed to determine whether there were any correlations between pSTAT levels and age or sex. Results of the statistical analyses are only shown for the specific cytokine pathways and cell subsets where there was at least one significant effect of age or sex on a measure of pSTAT. The number of donors per analysis (n) is indicated for significant effects. Based on these analyses, cytokine signalling data were normalised as appropriate to minimise the variation introduced by these potentially confounding factors. ns = not significant (P>0.05).

			***P*-values (n)**		
Signalling pathway	**Variable**	**Cell subset**	**%pSTAT+**	**stim/unstim**	**+pop/unstim**
IFN-β pSTAT3	Age	B cells	0.0291 (63)	0.0016 (64)	0.0031 (63)
	Sex	CD4+ naive	0.0435 (90)	ns	ns
IL-6 pSTAT3	Age	CD8+ naive	0.0333 (88)	0.0144 (89)	0.0073 (88)
	Age	Monocytes	0.0235 (58)	ns	ns
	Sex	CD4+ naive	0.0104 (82)	ns	ns
IL-10 pSTAT3	Sex	B cells	0.0298 (62)	ns	ns

5.3.2. IFN-β signalling by genotype in immune cell subsets directly *ex vivo*: STAT3 phosphorylation

Based on the cell line work presented in Chapter 3 and previously published work in expanded T cells (Couturier et al., 2011), the effect of rs34536443 and rs12720356 genotype was considered in the context of the IFN-β signalling in primary human immune cell subsets directly *ex vivo*.

Table 5.2. provides a summary of *P*-values and the number of individuals per group for analyses of IFN-β-induced pSTAT3 signalling by rs34536443 and rs12720356 genotype. For naïve and memory subsets of both CD4+ and CD8+ T cells, IFN-β-induced pSTAT3 levels were significantly reduced in individuals homozygous for the protective allele of the rs34536443 SNP, using all three measures of STAT phosphorylation (Table 5.2.; Figures 5.3.1.-5.3.4.) with genotype accounting for 9.79-36.27% of the variation in the data (Table 5.2.). For B cells and monocytes, the same trend was seen by rs34536443 genotype however significance was not reached for all three pSTAT3 quantification methods (Table 5.2.; Figures 5.3.5. and 5.3.6.), potentially due to the lower number of individuals homozygous for the protective allele of the rs34536443 SNP in these analyses compared to the staining for other cell subsets.

IFN-β-induced pSTAT3 signalling was only significantly correlated with rs12720356 genotype in memory CD4+ T cells (Figure 5.3.1.a; P=0.0073, r^2=0.114), however there were non-significant step-wise trends for IFN-β-induced pSTAT3 signalling by rs12720356 genotype in other T cell subsets (Table 5.2.; Figures 5.3.1.-5.3.4.) but not B cells and monocytes (Table 5.2.; Figures 5.3.5. and 5.3.6.). Collectively, this may suggest that the rs12720356 SNP has a more subtle influence on TYK2 function than the rs34536443 SNP, and thus a larger sample size may be required to reveal such an effect. Based on the disease association pattern of the rs12720356 SNP, however, it was hypothesised that this SNP was likely to exert additional effects on genes other than *TYK2* (see Chapter 6).

Table 5.2. Summary of statistical analyses for IFN-β-induced pSTAT3 signalling by genotype
IFN-β-induced pSTAT3 signalling in different immune cell subsets was quantified using three different measures: the proportion of cells gated as pSTAT-positive (%pSTAT+), the fold change in pSTAT level in stimulated compared to unstimulated cells (stim/unstim) and the fold change in pSTAT in the pSTAT-positive population compared to unstimulated cells (+pop/unstim). Data were normalised by age or sex as appropriate based on earlier analyses. The *P*-values for linear regression analyses of pSTAT levels by rs34536443 and rs12720356 genotype and the number of individuals per genotype group (n per group) are shown. Using a Bonferroni correction for multiple testing, the *P*-value threshold for significance of genotype-dependent correlations was set at 0.0083; ns = not significant ($P>0.0083$). For *P*-values meeting the threshold, r^2 values are shown, indicating the proportion of variation in the data that is accounted for by genotype. Cells highlighted in blue indicate genotype groups for which there were three or fewer individuals.

		rs345.. genotype				rs127.. genotype			
		n per group				n per group			
Cell subset	**Gating**	**Risk**	**Het**	**Prot**	***P*-value (r^2)**	**Risk**	**Het**	**Prot**	***P*-value (r^2)**
CD4+ memory	%pSTAT+	42	22	5	<0.0001 (0.2663)	7	13	42	0.0073 (0.114)
	stim/unstim	44	22	5	0.0014 (0.1376)	7	13	44	ns
	+pop/un	43	22	5	0.0016 (0.1373)	7	13	43	ns
CD4+ naive	%pSTAT+	43	22	5	0.0064 (0.0979)	7	13	43	ns
	stim/unstim	44	22	5	0.0025 (0.1253)	7	13	44	ns
	+pop/un	43	22	4	0.003 (0.124)	7	13	43	ns
CD8+ memory	%pSTAT+	44	21	5	<0.0001 (0.3627)	7	12	44	ns
	stim/unstim	44	21	5	<0.0001 (0.2028)	6	12	44	ns
	+pop/un	44	21	5	0.0007 (0.1551)	7	12	44	ns
CD8+ naïve	%pSTAT+	44	22	5	<0.0001 (0.2475)	7	13	44	ns
	stim/unstim	44	22	5	0.0053 (0.1075)	7	13	44	ns
	+pop/un	44	22	5	0.003 (0.1209)	7	13	44	ns
B cells	%pSTAT+	30	11	1	ns	7	13	30	ns
	stim/unstim	31	11	2	0.0013 (0.22)	7	13	31	ns
	+pop/un	31	11	1	0.0055 (0.1731)	7	13	31	ns
Monocytes	%pSTAT+	45	19	3	ns	7	13	45	ns
	stim/unstim	46	22	5	<0.0001 (0.1998)	7	13	46	ns
	+pop/un	45	19	3	ns	7	13	45	ns

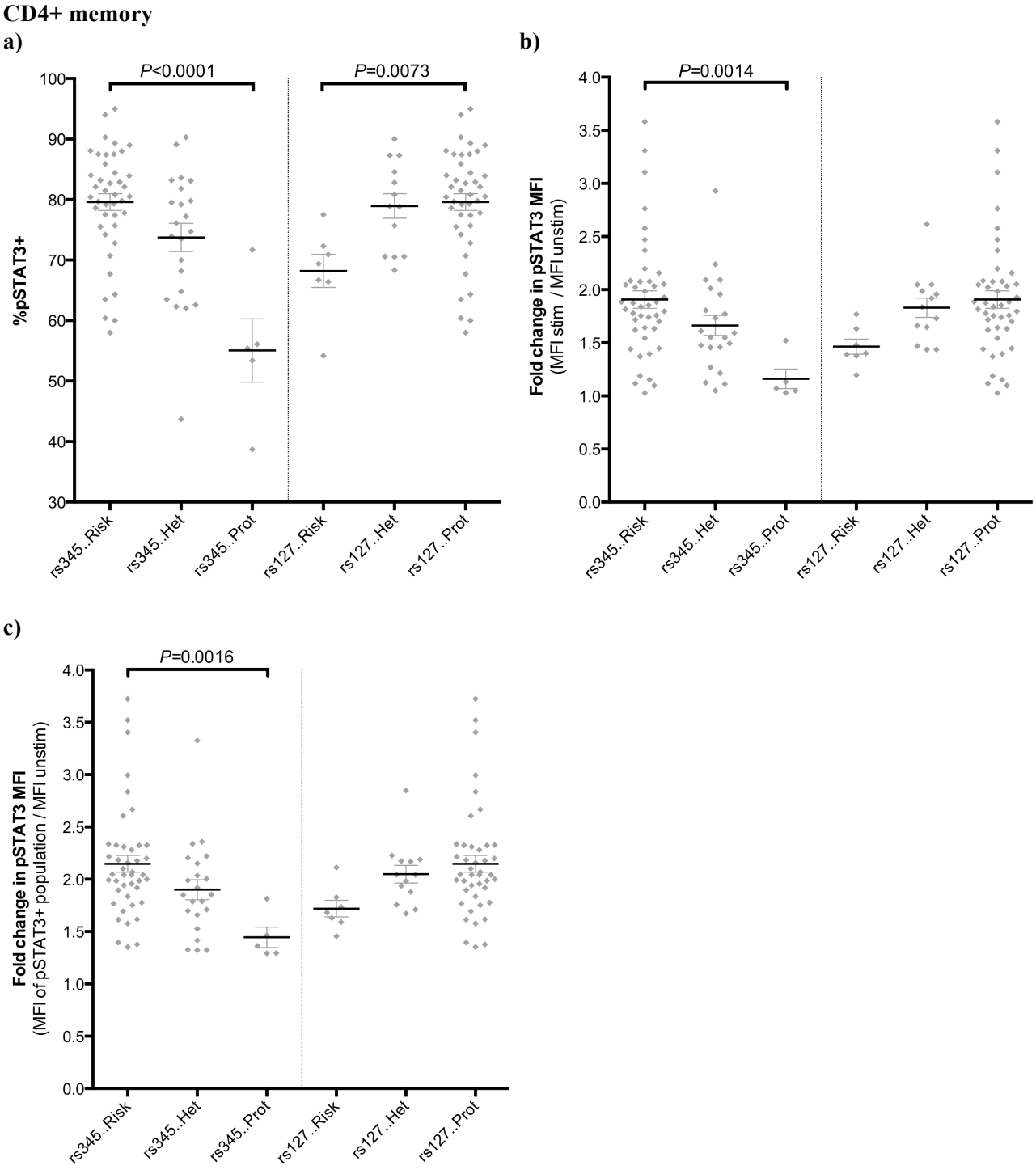


Figure 5.3.1. IFN-β-induced pSTAT3 signalling by genotype in memory CD4+ T cells
Following a 15 minute stimulation with IFN-β, PBMCs were stained for cell subset markers and intra-cellular pSTAT3 levels for analysis of signalling by rs34536443 and rs12720356 genotype. *P*-values calculated by linear regression analyses are shown where reaching the significance threshold set to account for multiple testing ($P<0.0083$). For a summary of *P*-values and numbers per group for these analyses see Table 5.2. (a) There was a significant trend in the proportion of cells becoming pSTAT3-positive (%pSTAT3+) correlating with rs34536443 genotype ($P<0.0001$, $r^2=0.2663$) and also rs12720356 genotype ($P=0.0073$, $r^2=0.114$). (b) The fold change in pSTAT3 in IFN-β-stimulated compared to unstimulated cells correlated with rs34536443 genotype ($P=0.0014$, $r^2=0.1376$) but not rs12720356 genotype. (c) The fold change in pSTAT3 levels observed in the pSTAT-positive population also correlated with rs34536443 genotype ($P=0.0016$, $r^2=0.1373$) but not rs12720356 genotype. Graphs show mean ± SEM.

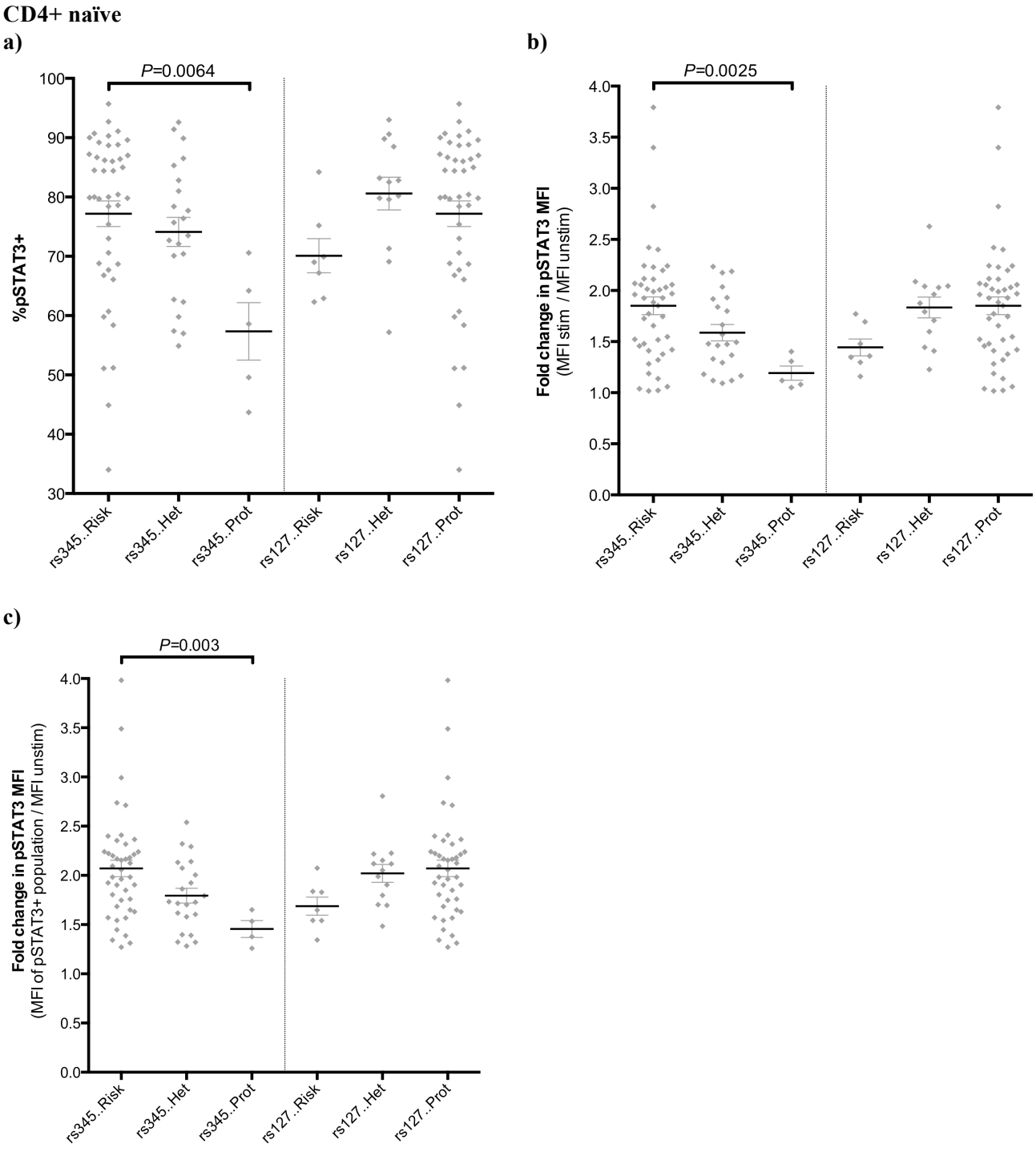


Figure 5.3.2. IFN-β-induced pSTAT3 signalling by genotype in naive CD4+ T cells
Following a 15 minute stimulation with IFN-β, PBMCs were stained for cell subset markers and intra-cellular pSTAT3 levels for analysis of signalling by rs34536443 and rs12720356 genotype. *P*-values calculated by linear regression analyses are shown where reaching the significance threshold set to account for multiple testing ($P<0.0083$). For a summary of *P*-values and numbers per group for these analyses see Table 5.2. (a) There was a significant trend in the proportion of cells becoming pSTAT3-positive (%pSTAT3+) correlating with rs34536443 genotype ($P=0.0064$, $r^2=0.0979$) but not rs12720356 genotype. (b) The fold change in pSTAT3 in IFN-β-stimulated compared to unstimulated cells correlated with rs34536443 genotype ($P=0.0025$, $r^2=0.1253$) but not rs12720356 genotype. (c) The fold change in pSTAT3 levels observed in the pSTAT-positive population also correlated with rs34536443 genotype ($P=0.003$, $r^2=0.124$) but not rs12720356 genotype. Graphs show mean ± SEM.

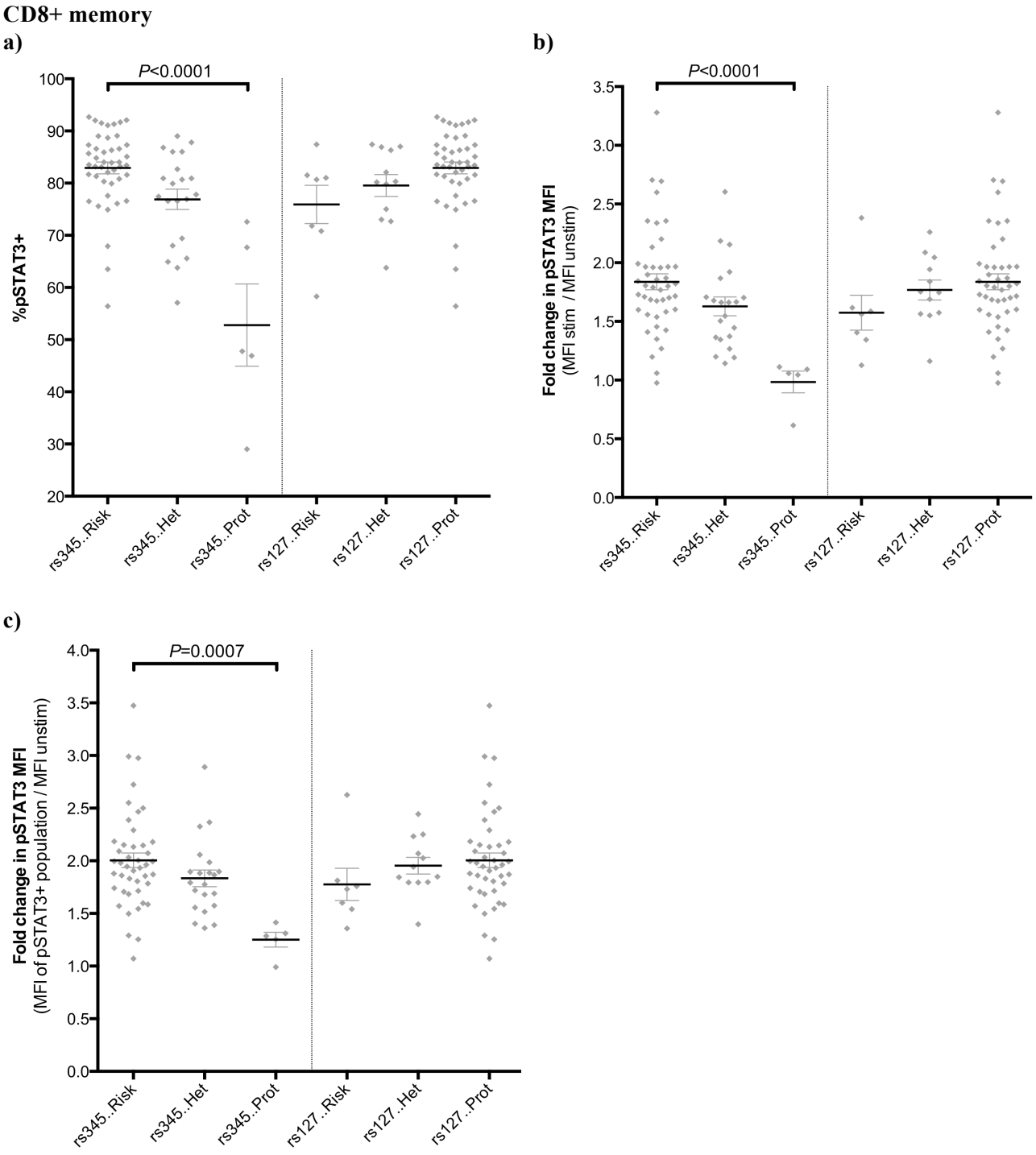


Figure 5.3.3. IFN-β-induced pSTAT3 signalling by genotype in memory CD8+ T cells
Following a 15 minute stimulation with IFN-β, PBMCs were stained for cell subset markers and intra-cellular pSTAT3 levels for analysis of signalling by rs34536443 and rs12720356 genotype. *P*-values calculated by linear regression analyses are shown where reaching the significance threshold set to account for multiple testing ($P<0.0083$). For a summary of *P*-values and numbers per group for these analyses see Table 5.2. (a) There was a significant trend in the proportion of cells becoming pSTAT3-positive (%pSTAT3+) correlating with rs34536443 genotype ($P<0.0001$, $r^2=0.3627$) but not rs12720356 genotype. (b) The fold change in pSTAT3 in IFN-β-stimulated compared to unstimulated cells correlated with rs34536443 genotype ($P<0.0001$, $r^2=0.2028$) but not rs12720356 genotype. (c) The fold change in pSTAT3 levels observed in the pSTAT-positive population also correlated with rs34536443 genotype ($P=0.0007$, $r^2=0.1551$) but not rs12720356 genotype. Graphs show mean ± SEM.

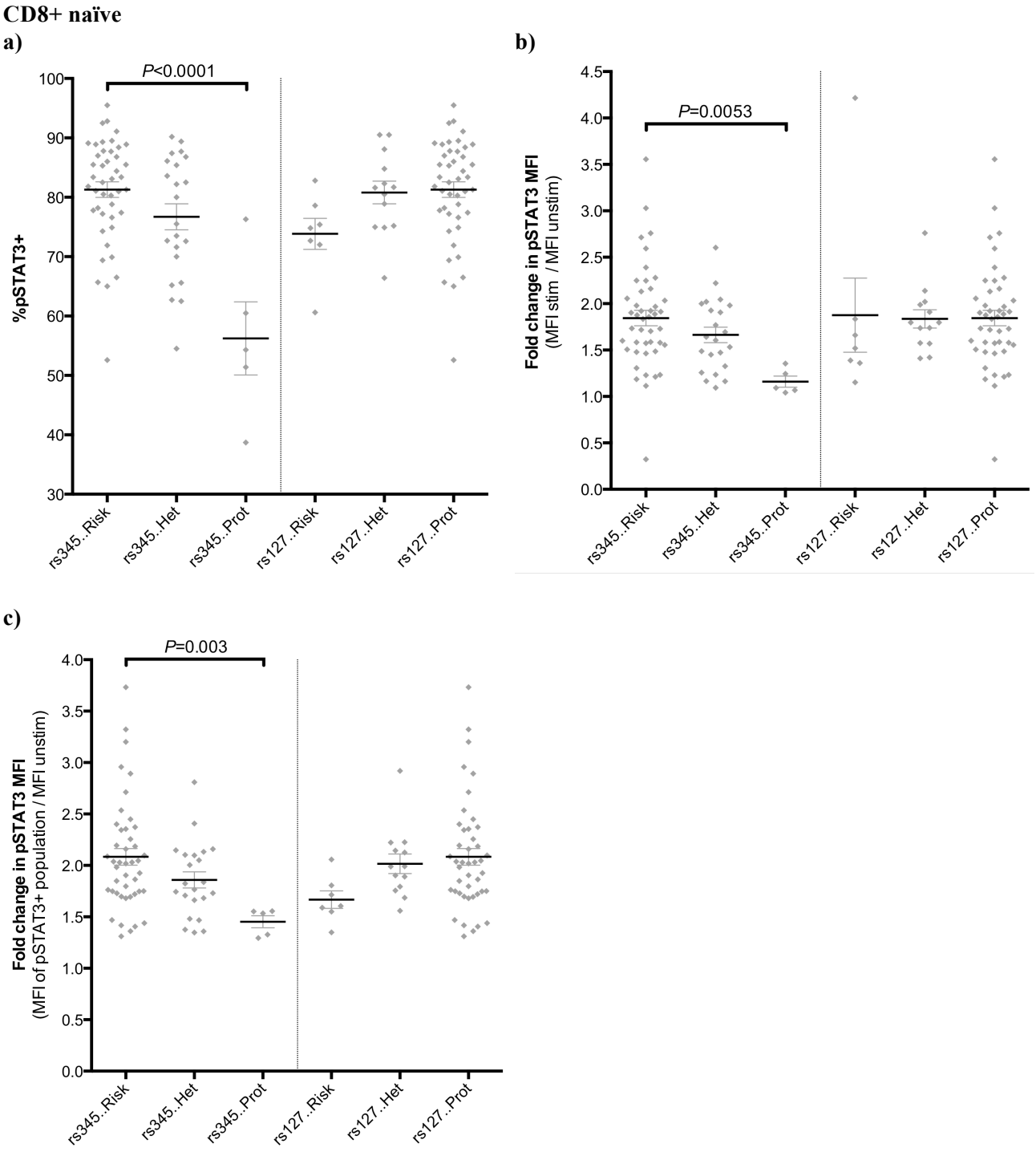


Figure 5.3.4. IFN-β-induced pSTAT3 signalling by genotype in naive CD8+ T cells
Following a 15 minute stimulation with IFN-β, PBMCs were stained for cell subset markers and intra-cellular pSTAT3 levels for analysis of signalling by rs34536443 and rs12720356 genotype. *P*-values calculated by linear regression analyses are shown where reaching the significance threshold set to account for multiple testing ($P<0.0083$). For a summary of *P*-values and numbers per group for these analyses see Table 5.2. (a) There was a significant trend in the proportion of cells becoming pSTAT3-positive (%pSTAT3+) correlating with rs34536443 genotype ($P<0.0001$, $r^2=0.2475$) but not rs12720356 genotype. (b) The fold change in pSTAT3 in IFN-β-stimulated compared to unstimulated cells correlated with rs34536443 genotype ($P=0.0053$, $r^2=0.1075$) but not rs12720356 genotype. (c) The fold change in pSTAT3 levels observed in the pSTAT-positive population also correlated with rs34536443 genotype ($P=0.003$, $r^2=0.1209$) but not rs12720356 genotype. Graphs show mean ± SEM.

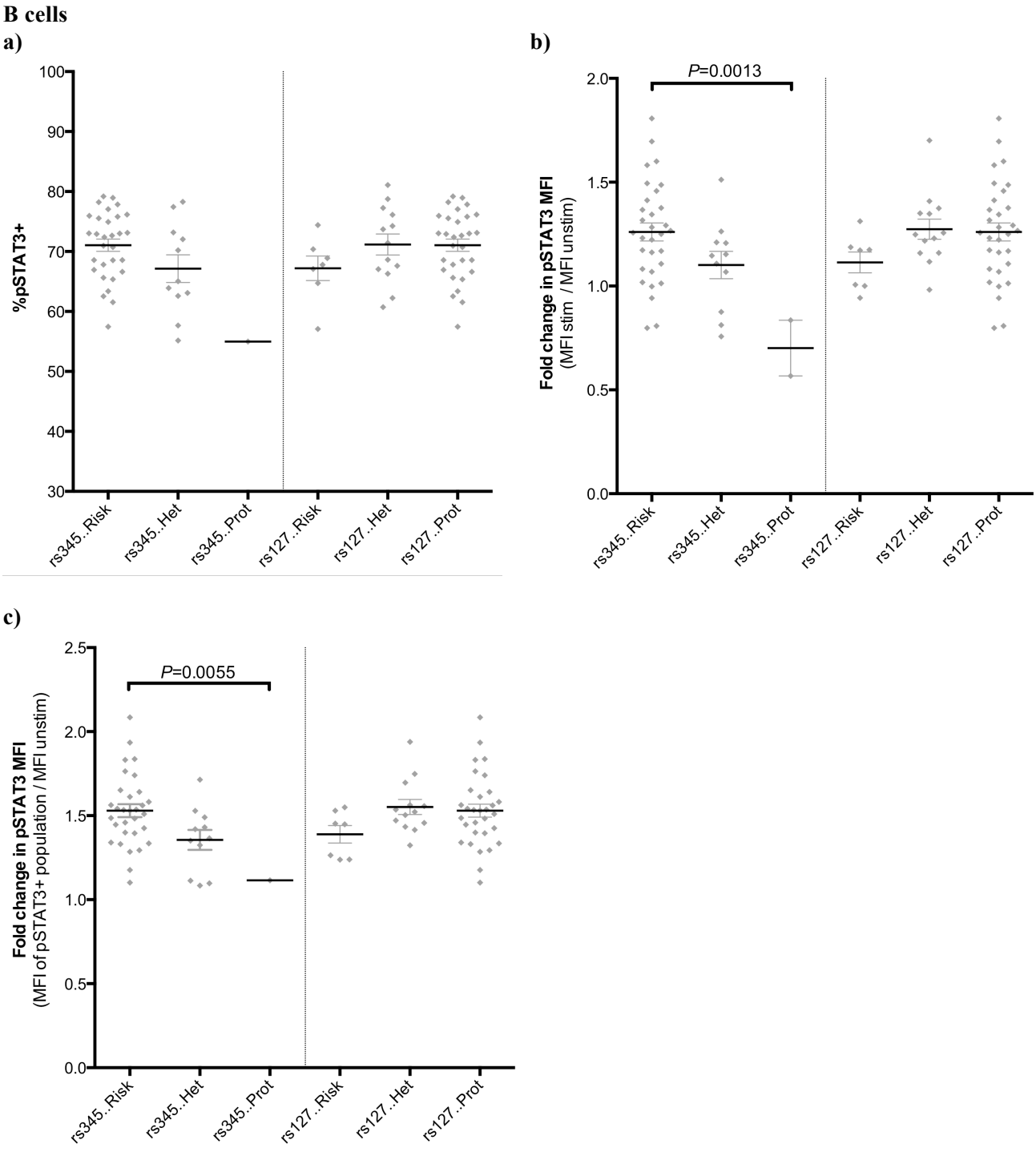


Figure 5.3.5. IFN-β-induced pSTAT3 signalling by genotype in B cells

Following a 15 minute stimulation with IFN-β, PBMCs were stained for cell subset markers and intra-cellular pSTAT3 levels for analysis of signalling by rs34536443 and rs12720356 genotype. *P*-values calculated by linear regression analyses are shown where reaching the significance threshold set to account for multiple testing ($P<0.0083$). For a summary of *P*-values and numbers per group for these analyses see Table 5.2. (a) There was no significant trend in the proportion of cells becoming pSTAT3-positive (%pSTAT3+) correlating with either rs34536443 or rs12720356 genotype (b) The fold change in pSTAT3 in IFN-β-stimulated compared to unstimulated cells correlated with rs34536443 genotype ($P=0.0013$, $r^2=0.22$) but not rs12720356 genotype. (c) The fold change in pSTAT3 levels observed in the pSTAT-positive population also correlated with rs34536443 genotype ($P=0.0055$, $r^2=0.1731$) but not rs12720356 genotype. Graphs show mean ± SEM.

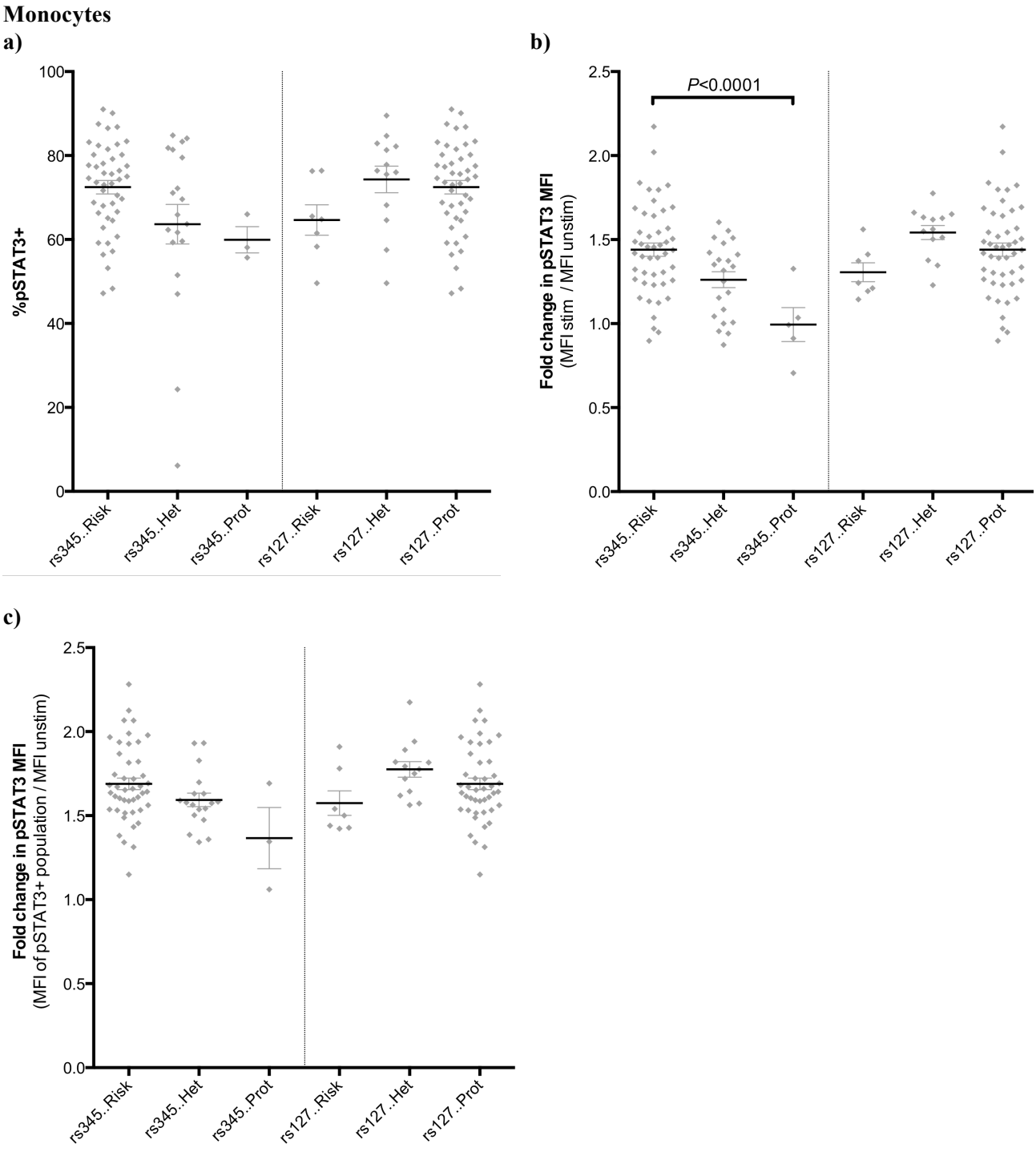


Figure 5.3.6. IFN-β-induced pSTAT3 signalling by genotype in monocytes
Following a 15 minute stimulation with IFN-β, PBMCs were stained for cell subset markers and intra-cellular pSTAT3 levels for analysis of signalling by rs34536443 and rs12720356 genotype. *P*-values calculated by linear regression analyses are shown where reaching the significance threshold set to account for multiple testing ($P<0.0083$). For a summary of *P*-values and numbers per group for these analyses see Table 5.2. (a) There was no significant trend in the proportion of cells becoming pSTAT3-positive (%pSTAT3+) correlating with either rs34536443 or rs12720356 genotype (b) The fold change in pSTAT3 in IFN-β-stimulated compared to unstimulated cells correlated with rs34536443 genotype ($P<0.0001$, $r^2=0.1998$) but not rs12720356 genotype. (c) There was no genotype-dependent correlation with the fold change in pSTAT3 levels observed in the pSTAT-positive population. Graphs show mean ± SEM.

5.3.3. IFN-β signalling by genotype in immune cell subsets directly *ex vivo*: STAT1 phosphorylation

Previous work has demonstrated a correlation between rs34536443 genotype and IFN-β-induced phosphorylation of STAT2, but not that of pSTAT1 (Couturier et al., 2011). However this was in a relatively small sample size (n=19 compared to n=94 here), and given the absence of a robust pSTAT2 antibody for flow cytometry, IFN-β-induced pSTAT1 levels were quantified in addition to pSTAT3 in the presented work. Table 5.3. summarises the *P*-values and the number of individuals per group for analyses of IFN-β-induced pSTAT1 levels by rs34536443 and rs12720356 genotype. There were no significant correlations between IFN-β-induced pSTAT1 signalling and rs12720356 genotype (Table 5.3.). In contrast, for naïve and memory subsets of both CD4+ and CD8+ T cells stimulation-induced pSTAT1 signalling was found to significantly correlate with rs34536443 genotype (Table 5.3.; Figures 5.3.7.-5.3.10.) for all three pSTAT quantification methods with genotype accounting for 10.38-15.98% of the variability in the data (Table 5.3.). However, for pSTAT1 signalling in B cells and monocytes (Table 5.3.; Figure 5.3.11. and 5.3.12.), there were suggestive trends (that did not meet the significance threshold) towards a reduction in IFN-β-induced pSTAT1 levels in individuals carrying the minor protective allele of the rs34536443 SNP; a larger sample size will be required to determine whether these trends will reach statistical significance.

Unlike for the pSTAT3 staining, the numbers of individuals in the rs34536443 protective homozygous group were analogous for the pSTAT1 staining across cell subsets. Therefore, collectively the pSTAT3 and pSTAT1 data by rs34536443 genotype may suggest a more pronounced effect of genotype-dependent IFN-β-induced differences in TYK2 activity in T cells compared to B cells and monocytes.

Table 5.3. Summary of statistical analyses for IFN-β-induced pSTAT1 signalling by genotype
IFN-β-induced pSTAT1 signalling in different immune cell subsets was quantified using three different measures: the proportion of cells gated as pSTAT-positive (%pSTAT+), the fold change in pSTAT level in stimulated compared to unstimulated cells (stim/unstim) and the fold change in pSTAT in the pSTAT-positive population compared to unstimulated cells (+pop/unstim). Data were normalised by age or sex as appropriate based on earlier analyses. The *P*-values for linear regression analyses of pSTAT levels by rs34536443 and rs12720356 genotype and the number of individuals per genotype group (n per group) are shown. Using a Bonferroni correction for multiple testing, the *P*-value threshold for significance of genotype-dependent correlations was set at 0.0083; ns = not significant ($P>0.0083$). For *P*-values meeting the threshold, r^2 values are shown, indicating the proportion of variation in the data that is accounted for by genotype.

		rs345.. genotype				rs127.. genotype			
		n per group				n per group			
Cell subset	**Gating**	**Risk**	**Het**	**Prot**	***P*-value (r^2)**	**Risk**	**Het**	**Prot**	***P*-value (r^2)**
CD4+ memory	%pSTAT+	45	22	4	0.0022 (0.1276)	7	13	45	ns
	stim/unstim	45	22	5	0.0014 (0.137)	7	13	45	ns
	+pop/un	45	22	4	0.0062 (0.1038)	7	13	45	ns
CD4+ naive	%pSTAT+	45	22	4	0.001 (0.1463)	7	13	45	ns
	stim/unstim	45	22	5	0.001 (0.1449)	7	13	45	ns
	+pop/un	45	23	4	0.0055 (0.1066)	7	13	45	ns
CD8+ memory	%pSTAT+	44	20	4	0.0009 (0.1553)	7	13	44	ns
	stim/unstim	44	20	5	0.0007 (0.1598)	7	13	44	ns
	+pop/un	44	20	4	0.0064 (0.1074)	7	13	44	ns
CD8+ naïve	%pSTAT+	46	22	4	0.001 (0.1445)	7	13	46	ns
	stim/unstim	46	23	5	0.0005 (0.1553)	7	13	46	ns
	+pop/un	46	22	4	0.0022 (0.1258)	7	13	46	ns
B cells	%pSTAT+	43	20	4	ns	7	13	43	ns
	stim/unstim	44	21	5	ns	7	13	44	ns
	+pop/un	42	20	4	ns	7	13	42	ns
Monocytes	%pSTAT+	41	20	4	ns	7	13	41	ns
	stim/unstim	43	21	4	ns	7	13	43	ns
	+pop/un	41	20	4	ns	7	13	41	ns

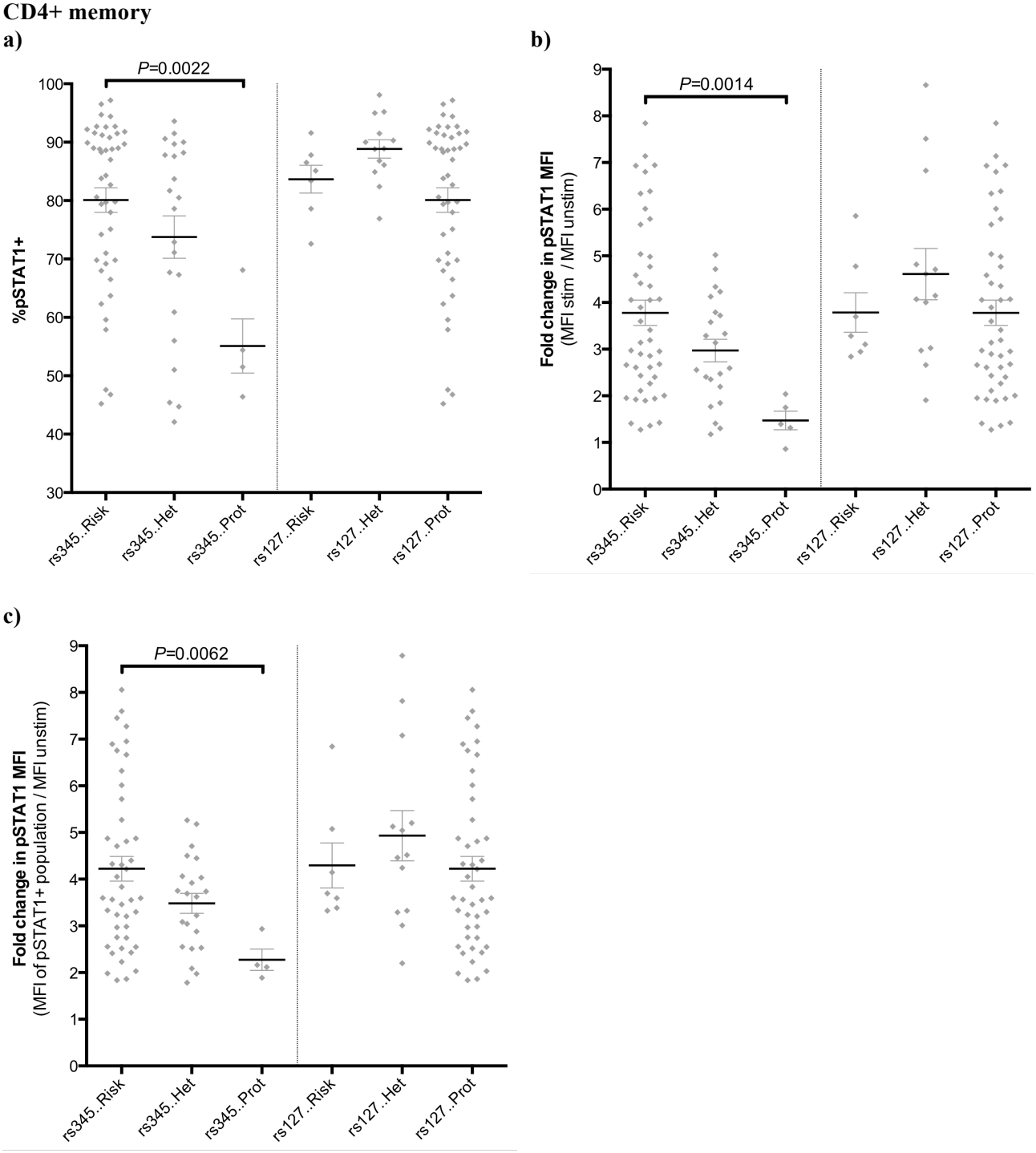


Figure 5.3.7. IFN-β-induced pSTAT1 signalling by genotype in memory CD4+ T cells
Following a 15 minute stimulation with IFN-β, PBMCs were stained for cell subset markers and intra-cellular pSTAT1 levels for analysis of signalling by rs34536443 and rs12720356 genotype. *P*-values calculated by linear regression analyses are shown where reaching the significance threshold set to account for multiple testing ($P<0.0083$). For a summary of *P*-values and numbers per group for these analyses see Table 5.3. (a) There was a significant trend in the proportion of cells becoming pSTAT1-positive (%pSTAT1+) correlating with rs34536443 genotype ($P=0.0022$, $r^2=0.1276$) but not rs12720356 genotype. (b) The fold change in pSTAT1 in IFN-β-stimulated compared to unstimulated cells correlated with rs34536443 genotype ($P=0.0014$, $r^2=0.137$) but not rs12720356 genotype. (c) The fold change in pSTAT1 levels observed in the pSTAT-positive population also correlated with rs34536443 genotype ($P=0.0062$, $r^2=0.1038$) but not rs12720356 genotype. Graphs show mean ± SEM.

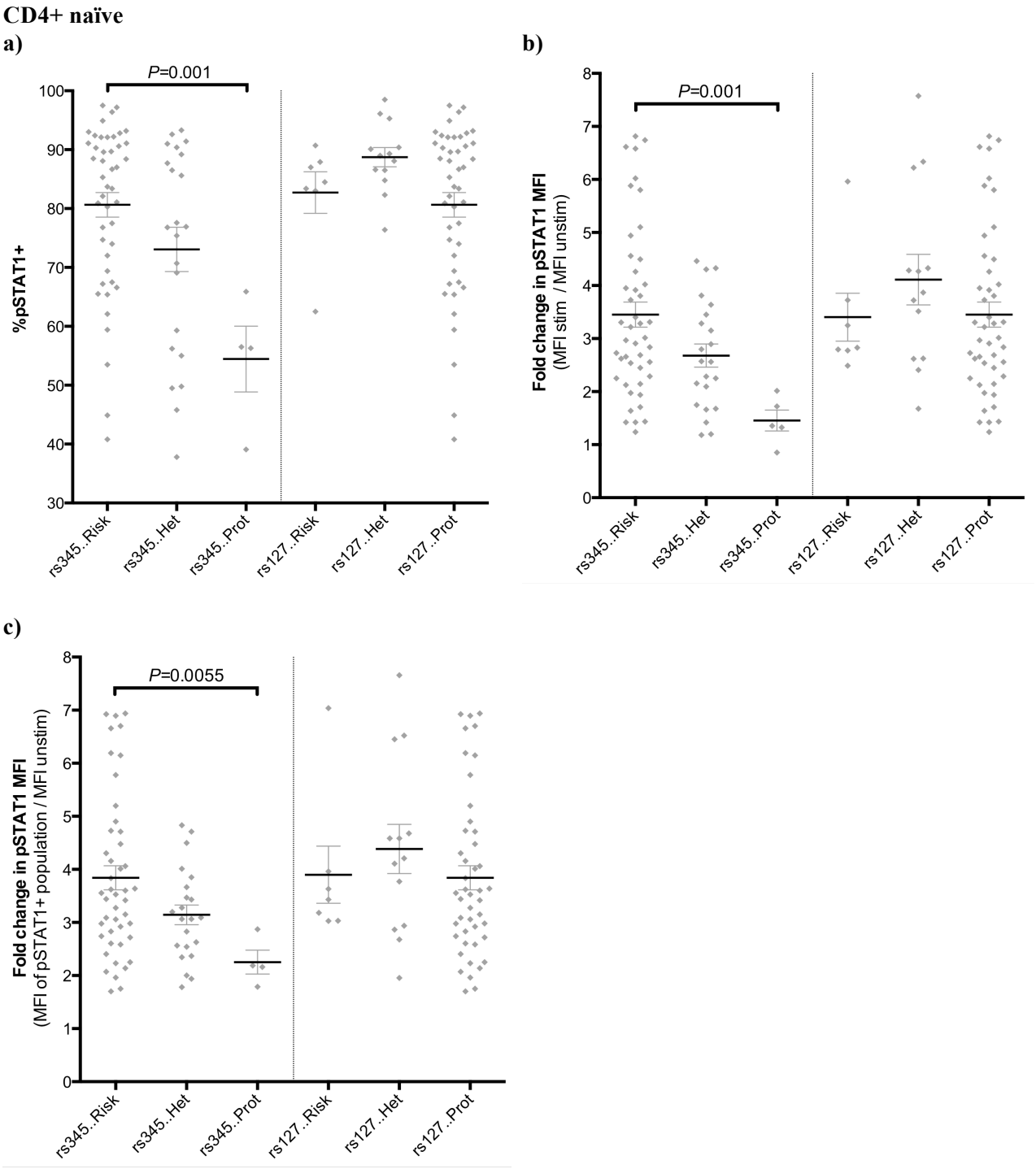


Figure 5.3.8. IFN-β-induced pSTAT1 signalling by genotype in naive CD4+ T cells
Following a 15 minute stimulation with IFN-β, PBMCs were stained for cell subset markers and intra-cellular pSTAT1 levels for analysis of signalling by rs34536443 and rs12720356 genotype. *P*-values calculated by linear regression analyses are shown where reaching the significance threshold set to account for multiple testing ($P<0.0083$). For a summary of *P*-values and numbers per group for these analyses see Table 5.3. (a) There was a significant trend in the proportion of cells becoming pSTAT1-positive (%pSTAT1+) correlating with rs34536443 genotype ($P=0.001$, $r^2=0.1463$) but not rs12720356 genotype. (b) The fold change in pSTAT1 in IFN-β-stimulated compared to unstimulated cells correlated with rs34536443 genotype ($P=0.001$, $r^2=0.1449$) but not rs12720356 genotype. (c) The fold change in pSTAT1 levels observed in the pSTAT-positive population also correlated with rs34536443 genotype ($P=0.0055$, $r^2=0.1066$) but not rs12720356 genotype. Graphs show mean ± SEM.

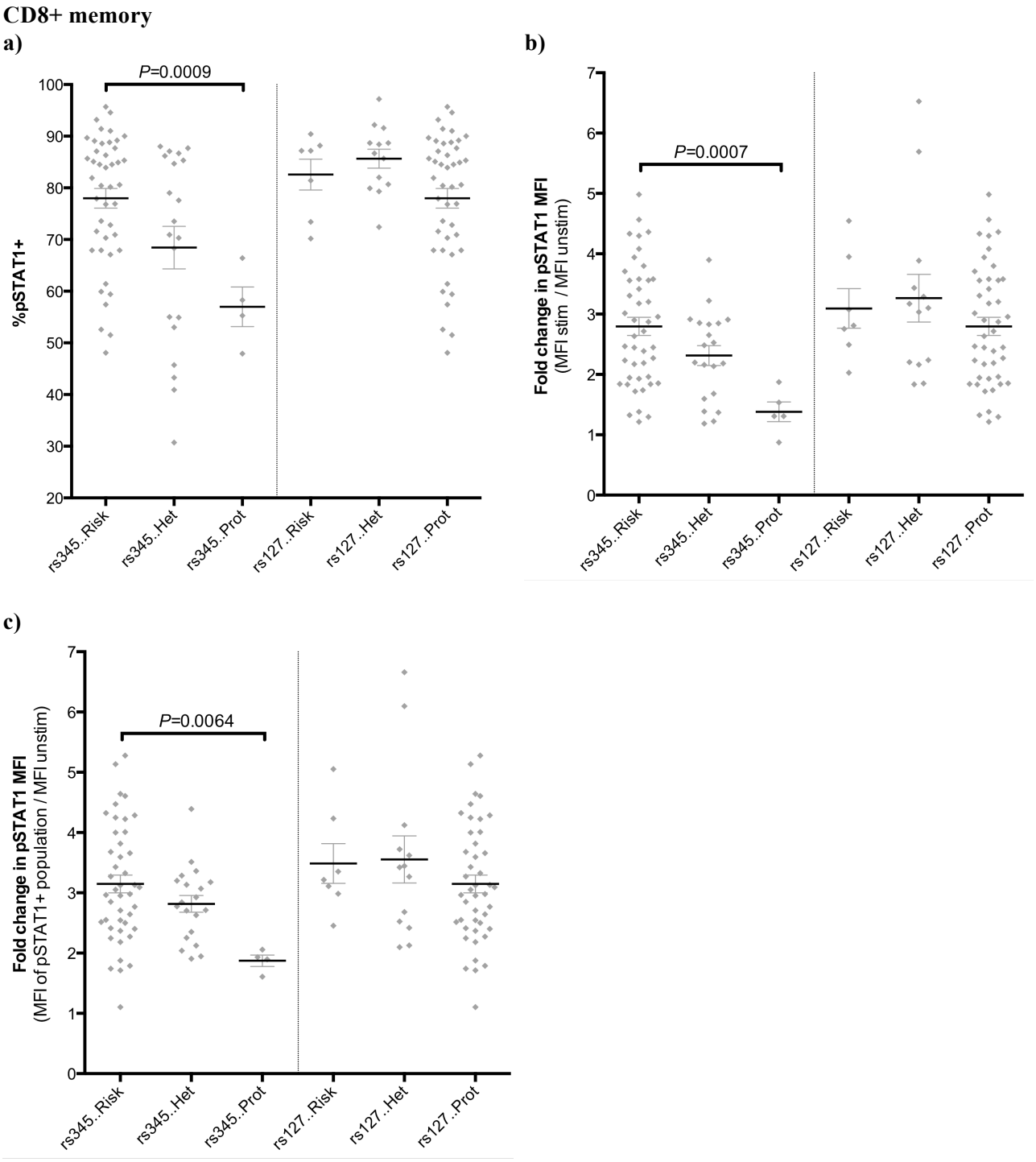


Figure 5.3.9. IFN-β-induced pSTAT1 signalling by genotype in memory CD8+ T cells
Following a 15 minute stimulation with IFN-β, PBMCs were stained for cell subset markers and intra-cellular pSTAT1 levels for analysis of signalling by rs34536443 and rs12720356 genotype. *P*-values calculated by linear regression analyses are shown where reaching the significance threshold set to account for multiple testing ($P<0.0083$). For a summary of *P*-values and numbers per group for these analyses see Table 5.3. (a) There was a significant trend in the proportion of cells becoming pSTAT1-positive (%pSTAT1+) correlating with rs34536443 genotype ($P=0.0009$, $r^2=0.1553$) but not rs12720356 genotype. (b) The fold change in pSTAT1 in IFN-β-stimulated compared to unstimulated cells correlated with rs34536443 genotype ($P=0.0007$, $r^2=0.1598$) but not rs12720356 genotype. (c) The fold change in pSTAT1 levels observed in the pSTAT-positive population also correlated with rs34536443 genotype ($P=0.0064$, $r^2=0.1074$) but not rs12720356 genotype. Graphs show mean ± SEM.

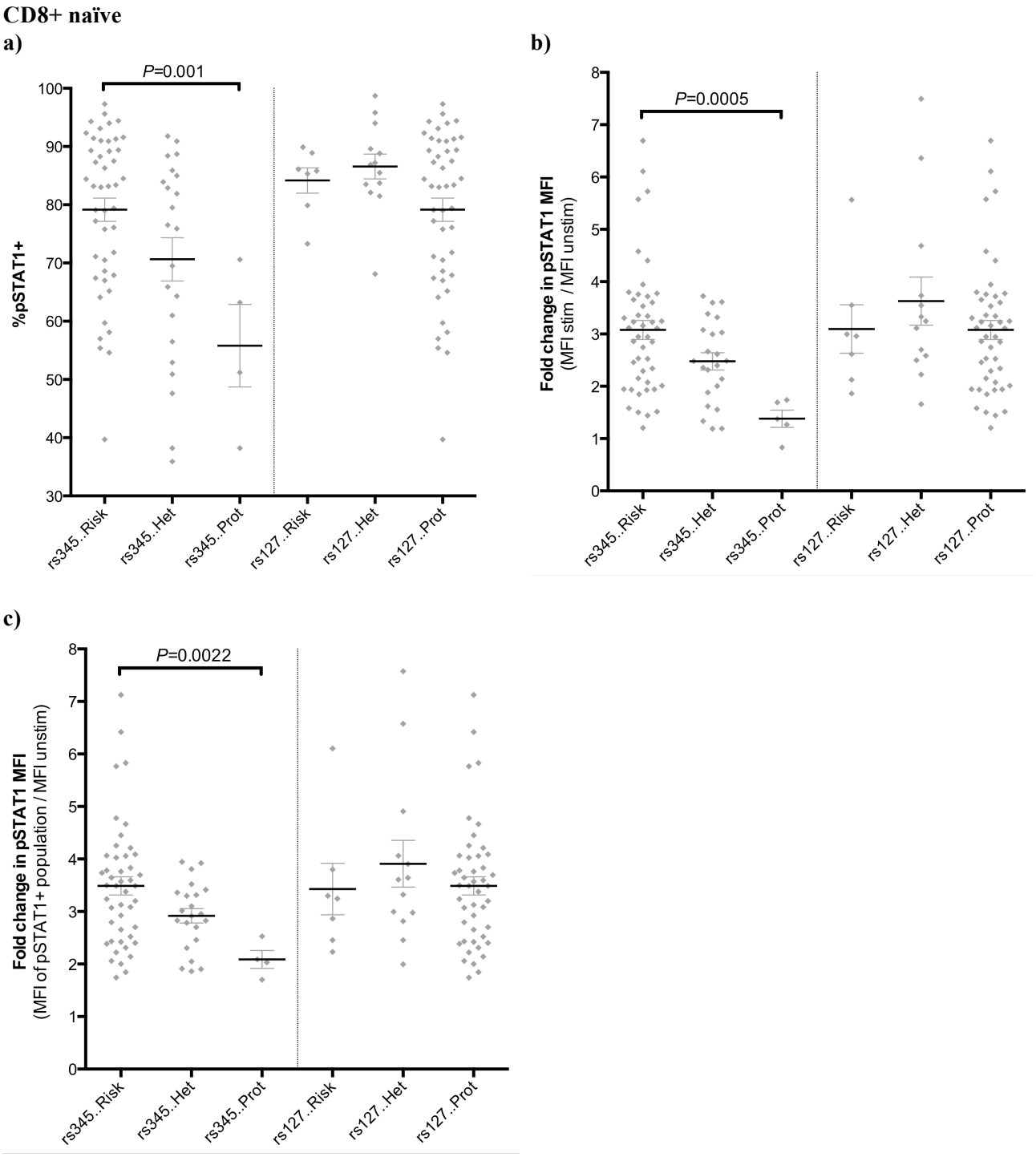


Figure 5.3.10. IFN-β-induced pSTAT1 signalling by genotype in naive CD8+ T cells
Following a 15 minute stimulation with IFN-β, PBMCs were stained for cell subset markers and intra-cellular pSTAT1 levels for analysis of signalling by rs34536443 and rs12720356 genotype. *P*-values calculated by linear regression analyses are shown where reaching the significance threshold set to account for multiple testing ($P<0.0083$). For a summary of *P*-values and numbers per group for these analyses see Table 5.3. (a) There was a significant trend in the proportion of cells becoming pSTAT1-positive (%pSTAT1+) correlating with rs34536443 genotype ($P=0.001$, $r^2=0.1445$) but not rs12720356 genotype. (b) The fold change in pSTAT1 in IFN-β-stimulated compared to unstimulated cells correlated with rs34536443 genotype ($P=0.0005$, $r^2=0.1553$) but not rs12720356 genotype. (c) The fold change in pSTAT1 levels observed in the pSTAT-positive population also correlated with rs34536443 genotype ($P=0.0022$, $r^2=0.1258$) but not rs12720356 genotype. Graphs show mean ± SEM.

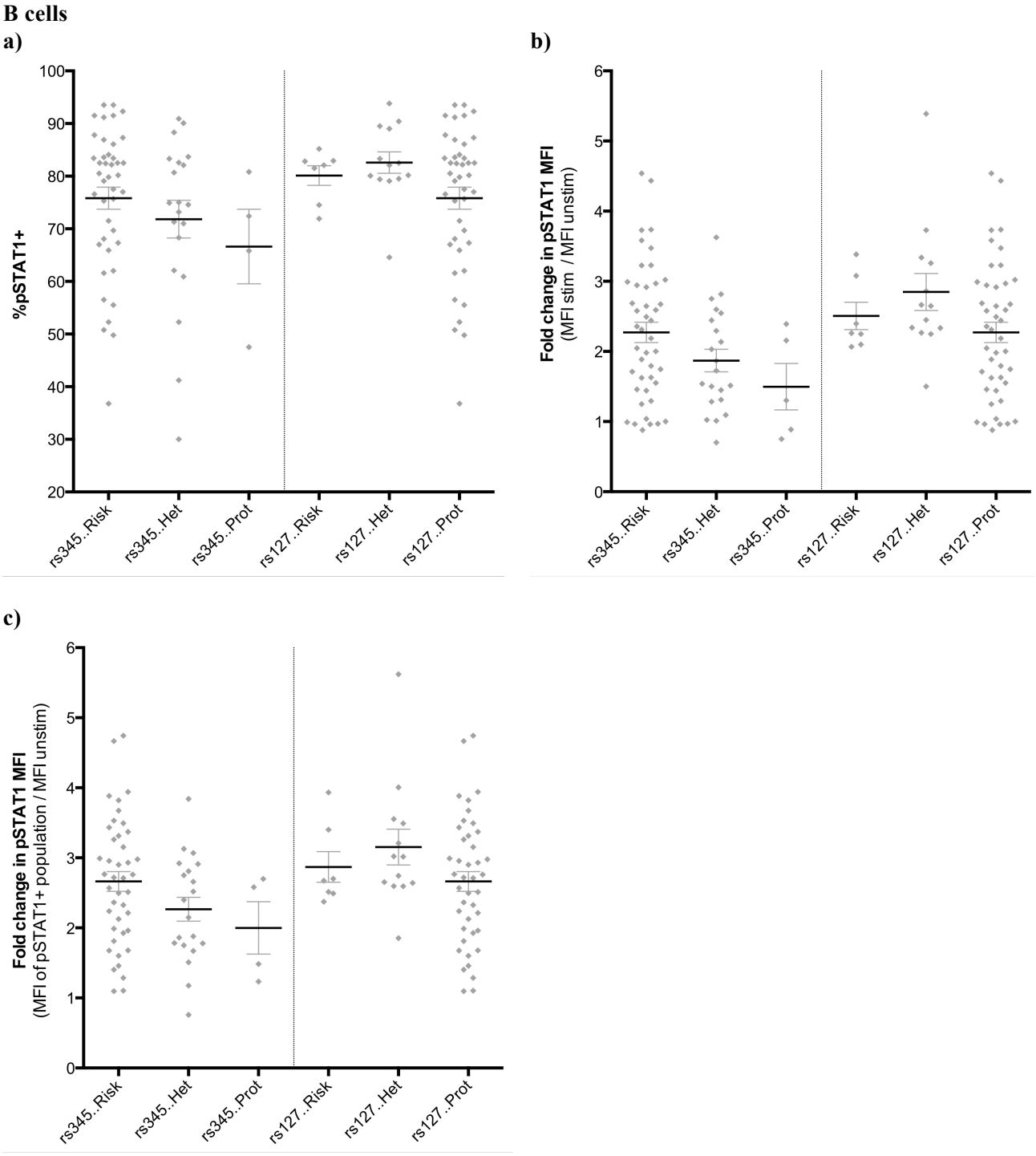


Figure 5.3.11. IFN-β-induced pSTAT1 signalling by genotype in B cells
Following a 15 minute stimulation with IFN-β, PBMCs were stained for cell subset markers and intra-cellular pSTAT1 levels for analysis of signalling by rs34536443 and rs12720356 genotype. *P*-values calculated by linear regression analyses are shown where reaching the significance threshold set to account for multiple testing ($P<0.0083$). For a summary of *P*-values and numbers per group for these analyses see Table 5.3. (a) There was no significant trends in the proportion of cells becoming pSTAT1-positive (%pSTAT1+) correlating with either rs34536443 or rs12720356 genotype. (b) The fold change in pSTAT1 in IFN-β-stimulated compared to unstimulated cells was also found not to correlate with genotype. (c) The fold change in pSTAT1 levels observed in the pSTAT-positive population was not found to correlate with genotype. Graphs show mean ± SEM.

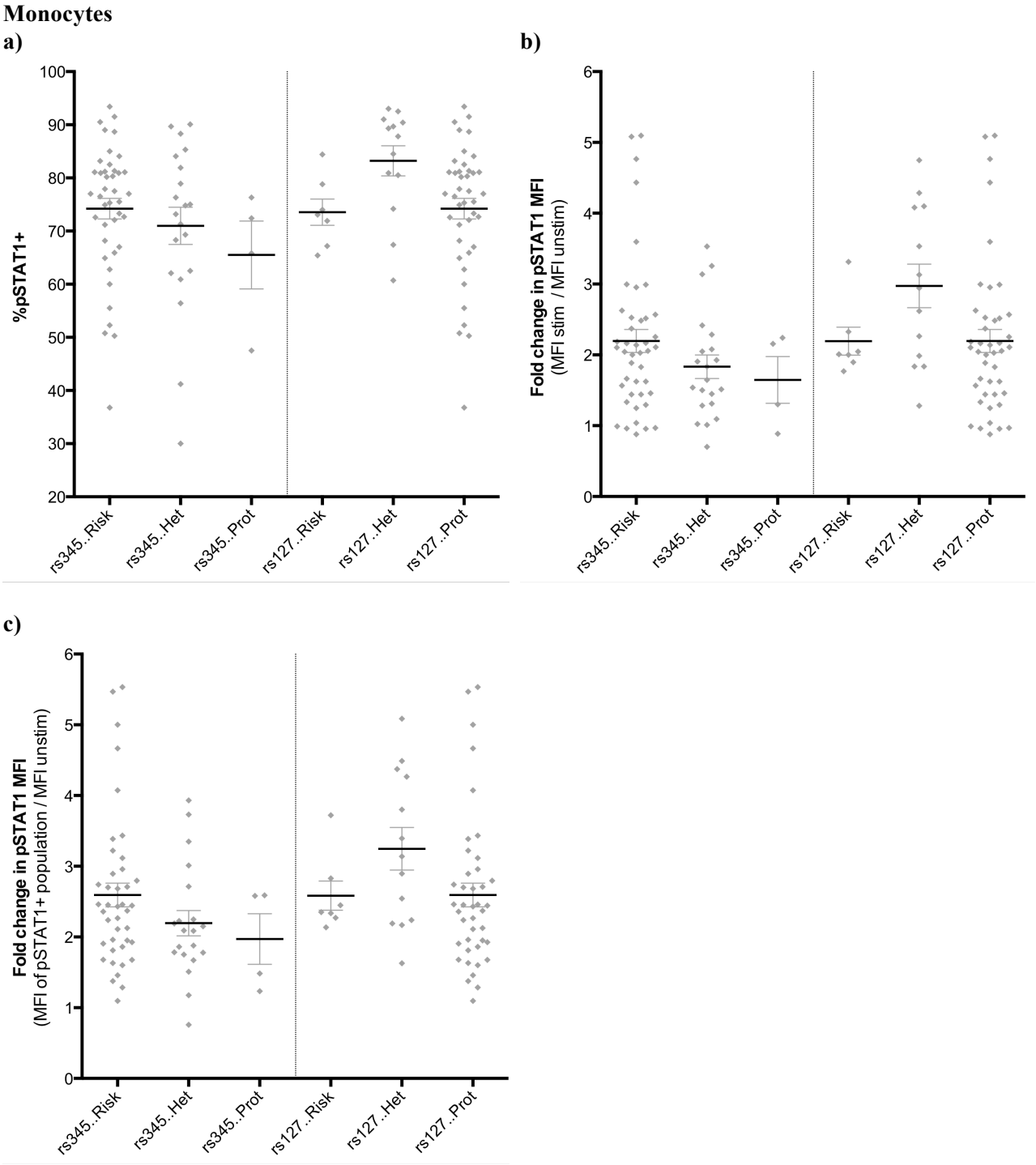


Figure 5.3.12. . IFN-β-induced pSTAT1 signalling by genotype in monocytes
Following a 15 minute stimulation with IFN-β, PBMCs were stained for cell subset markers and intra-cellular pSTAT1 levels for analysis of signalling by rs34536443 and rs12720356 genotype. *P*-values calculated by linear regression analyses are shown where reaching the significance threshold set to account for multiple testing ($P<0.0083$). For a summary of *P*-values and numbers per group for these analyses see Table 5.3. (a) There was no significant trends in the proportion of cells becoming pSTAT1-positive (%pSTAT1+) correlating with either rs34536443 or rs12720356 genotype. (b) The fold change in pSTAT1 in IFN-β-stimulated compared to unstimulated cells was also found not to correlate with genotype. (c) The fold change in pSTAT1 levels observed in the pSTAT-positive population was not found to correlate with genotype. Graphs show mean ± SEM.

5.3.4. IL-6 signalling by genotype in immune cell subsets directly *ex vivo*: STAT3 phosphorylation

In our cohort of individuals, there were no significant correlations between IL-6-induced pSTAT3 signalling and either rs34536443 or rs12720356 genotype across the immune cell subsets studied and for all measures of STAT3 phosphorylation (summarised in Table 5.4.). This indicates that any genotype-dependent effects on TYK2 function have no impact on IL-6 signalling, potentially due to the ability of the other JAK family kinases acting at the IL-6R to compensate for reduced TYK2 activity.

Table 5.4. Summary of statistical analyses for IL-6-induced pSTAT3 signalling by genotype
IL-6-induced pSTAT3 signalling in different immune cell subsets was quantified using three different measures: the proportion of cells gated as pSTAT-positive (%pSTAT+), the fold change in pSTAT level in stimulated compared to unstimulated cells (stim/unstim) and the fold change in pSTAT in the pSTAT-positive population compared to unstimulated cells (+pop/unstim). Data were normalised by age or sex as appropriate based on earlier analyses. The number of individuals per genotype group for linear regression analyses (n per group) is shown. Using a Bonferroni correction for multiple testing, the *P*-value threshold for significance of genotype-dependent correlations was set at 0.0083; ns = not significant (P>0.0083). Cells highlighted in blue indicate genotype groups for which there three or fewer individuals.

		rs345.. genotype				rs127.. genotype			
		n per group				n per group			
Cell subset	**Gating**	**Risk**	**Het**	**Prot**	***P*-value**	**Risk**	**Het**	**Prot**	***P*-value**
CD4+ memory	%pSTAT+	42	22	5	ns	7	13	42	ns
	stim/unstim	43	22	5	ns	7	13	43	ns
	+pop/un	42	22	5	ns	7	13	42	ns
CD4+ naive	%pSTAT+	41	20	3	ns	7	13	41	ns
	stim/unstim	43	22	5	ns	7	13	43	ns
	+pop/un	42	20	5	ns	7	13	42	ns
CD8+ memory	%pSTAT+	39	17	4	ns	6	12	39	ns
	stim/unstim	41	19	4	ns	6	12	41	ns
	+pop/un	38	17	4	ns	6	12	38	ns
CD8+ naïve	%pSTAT+	41	22	5	ns	7	13	41	ns
	stim/unstim	42	22	5	ns	7	13	42	ns
	+pop/un	40	22	5	ns	7	13	40	ns
B cells	%pSTAT+	23	8	2	ns	6	10	23	ns
	stim/unstim	31	11	3	ns	7	13	31	ns
	+pop/un	23	8	2	ns	6	10	23	ns
Monocytes	%pSTAT+	28	12	4	ns	4	10	28	ns
	stim/unstim	44	22	5	ns	7	13	44	ns
	+pop/un	27	11	4	ns	4	10	12	ns

5.3.5. IL-10 signalling by genotype in immune cell subsets directly *ex vivo*: STAT3 phosphorylation

As summarised in Table 5.5., there were only two significant correlations between nsSNP genotype and IL-10-induced pSTAT3 signalling. The proportion of memory CD8+ T cells becoming pSTAT-positive (%pSTAT+) upon IL-10 stimulation was found to correlate with rs12720356 genotype, with significantly reduced signalling observed in individuals carrying the minor risk allele (P=0.0035; Table 5.5.). In B cells, there was a significant reduction in the fold change in pSTAT3 level observed in stimulated compared to unstimulated cells (stim/unstim; P=0.006) in individuals carrying the minor protective allele of the rs34536443 SNP (P=0.006; Table 5.5.). A larger sample size will be required to determine if the statistical significance of these two correlations remains. However, overall the data suggest that genotype-dependent effects on TYK2 activity do not have a substantial impact on IL-10 signalling.

Table 5.5. Summary of statistical analyses for IL-10-induced pSTAT3 signalling by genotype
IL-10-induced pSTAT3 signalling in different immune cell subsets was quantified using three different measures: the proportion of cells gated as pSTAT-positive (%pSTAT+), the fold change in pSTAT level in stimulated compared to unstimulated cells (stim/unstim) and the fold change in pSTAT in the pSTAT-positive population compared to unstimulated cells (+pop/unstim). Data were normalised by age or sex as appropriate based on earlier analyses. The number of individuals per genotype group for linear regression analyses (n per group) is shown. Using a Bonferroni correction for multiple testing, the P-value threshold for significance of genotype-dependent correlations was set at 0.0083; ns = not significant (P>0.0083). For P-values meeting the threshold, r^2 values are shown, indicating the proportion of variation in the data that is accounted for by genotype. Cells highlighted in blue indicate genotype groups for which there were three or fewer individuals.

		rs345.. genotype				rs127.. genotype			
		n per group				n per group			
Cell subset	Gating	Risk	Het	Prot	P-value (r^2)	Risk	Het	Prot	P-value (r^2)
CD4+ memory	%pSTAT+	43	22	5	ns	7	13	43	ns
	stim/unstim	44	22	5	ns	7	13	44	ns
	+pop/un	42	22	5	ns	6	13	42	ns
CD4+ naive	%pSTAT+	40	20	5	ns	7	13	40	ns
	stim/unstim	44	22	5	ns	7	13	44	ns
	+pop/un	44	21	5	ns	7	13	44	ns
CD8+ memory	%pSTAT+	42	21	5	ns	7	12	42	0.0035 (0.1353)
	stim/unstim	42	21	4	ns	7	12	42	ns
	+pop/un	42	21	4	ns	7	12	42	ns
CD8+ naïve	%pSTAT+	43	22	5	ns	7	13	43	ns
	stim/unstim	44	22	4	ns	7	13	44	ns
	+pop/un	44	22	5	ns	7	13	44	ns
B cells	%pSTAT+	31	10	2	ns	6	12	31	ns
	stim/unstim	30	11	3	0.006 (0.1313)	7	13	30	ns
	+pop/un	30	11	2	ns	7	13	30	ns
Monocytes	%pSTAT+	42	20	4	ns	6	12	42	ns
	stim/unstim	44	22	5	ns	7	13	44	ns
	+pop/un	40	20	4	ns	6	12	40	ns

5.3.5. IL-13 signalling directly *ex vivo*: pSTAT6 data

As IL-13R expression is restricted to monocytes (Chapter 4, Figure 4.3.6.), IL-13-induced pSTAT6 signalling was only studied in this cell subset. No significant correlations with rs34536443 genotype were identified (Table 5.6.). Given this absence of correlations with the rs34536443 SNP and the fact that rs12720356 genotype was not found to strongly influence the other cytokine signalling pathways investigated, an effect of the latter SNP on IL-13-induced STAT6 phosphorylation was not investigated.

Table 5.6. Summary of statistical analyses for IL-13-induced pSTAT6 signalling in monocytes by genotype

IL-13-induced pSTAT6 signalling in monocytes was quantified using three different measures: the proportion of cells gated as pSTAT-positive (%pSTAT+), the fold change in pSTAT level in stimulated compared to unstimulated cells (stim/unstim) and the fold change in pSTAT in the pSTAT-positive population compared to unstimulated cells (+pop/unstim). Data were normalised by age or sex as appropriate based on earlier analyses. The number of individuals per genotype group for linear regression analyses (n per group) is shown. Using a Bonferroni correction for multiple testing, the *P*-value threshold for significance of genotype-dependent correlations was set at 0.0083; ns = not significant (*P*>0.0083).

		rs345.. genotype			
		n per group			
Cell subset	**Gating**	**Risk**	**Het**	**Prot**	***P*-value**
Monocytes	%pSTAT+	22	17	4	ns
	stim/unstim	20	15	4	ns
	+pop/un	20	16	4	ns

5.4 Discussion

The aim of the work presented in this chapter was to investigate the effects of nsSNP genotype on TYK2-mediated cytokine signalling in primary human immune cells, with a view to better characterising genotype-dependent differences in the broader context of autoimmune disease-relevant immunological pathways.

5.4.1. The effect of rs12720356 genotype on immune cell signalling and future investigation of this SNP

Both the rs34536443 and rs12720356 SNPs were found to have effects on TYK2 phosphorylation in cell lines (Chapter 3), however unlike the rs34536443 SNP, the effects of the rs12720356 SNP on pTYK2 levels did not correlate with changes in STAT1 phosphorylation. Therefore it was necessary to further investigate the consequences of rs12720356 genotype on TYK2-mediated cytokine signalling in primary human immune cells, to ascertain whether an effect of this SNP on TYK2 did in fact translate into a downstream effect on STAT-mediated signalling. However, unlike the rs34536443 SNP, there were no consistent effects of rs12720356 genotype across all cell subsets and all measures of pSTAT investigated, although two correlations did reach statistical significance. IFN-β-induced STAT3 phosphorylation in memory CD4+ T cells and IL-10 signalling in memory CD8+ T cells were both found to be significantly reduced in individuals carrying the minor allele of the rs12720356 SNP. There were no significant trends by genotype for IFN-β-induced pSTAT1 or IL-6-induced pSTAT3 signalling, however there were trends towards reduced IFN-β-induced pSTAT3 levels in other T cell subsets that did not reach significance, as was the case in memory CD4+ T cells, for example. This could suggest a more subtle effect of rs12720356 genotype on TYK2 signalling, and the analysis of a larger cohort of individuals is currently underway in order to ascertain whether the rs12720356 SNP does indeed influence TYK2-mediated signal transduction.

For both the rs12720356 and rs34536443 SNPs, the minor allele is associated with protection against PS (Strange et al., 2010; Tsoi et al., 2012) and thus a shared functional consequence of reduced IFN-β signalling would be consistent with this association. However, the rs12720356 SNP is also associated in the opposite direction with IBD, for which the minor allele is associated with risk (Anderson et al., 2011; Franke et al., 2010). The rs34536443 SNP is not associated with IBD and, other than PS, there are no other diseases that are associated with both the rs34536443 and rs12720356 SNPs. Considering these disparate disease associations, it is likely that these nsSNPs have distinct functional consequences. The lack of consistent significant correlations between rs12720356 genotype and IFN-β signalling across cell subsets with the current donor cohort further suggests that the rs12720356 SNP is likely to have additional disease-relevant functional consequences, which may be mediated via effects on genes other than *TYK2* (as was investigated in Chapter 6).

5.4.2. The effect of rs34536443 genotype on immune cell signalling

In contrast to the rs12720356 SNP, significant correlations between rs34536443 genotype and IFN-β signalling were consistently observed across cell subsets. The minor allele of the rs34536443 SNP was associated with significantly reduced IFN-β signalling in T cells, in terms of both pSTAT1 and pSTAT3. This was the case for all three measures of STAT phosphorylation, indicating that IFN-β signalling was reduced both in terms of the proportion of cells becoming pSTAT-positive and also the levels of pSTAT on a per cell basis. There were some significant correlations between rs34536443 genotype and IFN-β signalling in B cells and monocytes, and the observed trends not reaching significance may be better resolved in a larger sample size (recruitment is ongoing). For IFN-β signalling in B cells and monocytes, a significant genotype-dependent effect was observed for at least one measure of pSTAT3, however for pSTAT1 all three of the measures were not significantly correlated with genotype, which could reflect limitations of the antibodies used or perhaps suggest a cell subset-specific difference in the extent of STAT usage. Interestingly, in *Tyk2*-KO mice, type 1 IFN-induced phosphorylation of STAT3 is more drastically reduced than that

of STAT1 and STAT2 (Karaghiosoff et al., 2000). Further study is required to investigate the downstream consequences of these genotype-dependent changes in STAT phosphorylation on gene expression and cell function.

To our knowledge, the presented work was the first to investigate IFN-β signalling in primary immune cell subsets from individuals homozygous for the minor allele of the rs34536443 SNP, with previously published work involving study of only heterozygous individuals (Couturier et al., 2011). Additionally, the presented work has a relatively larger sample of individuals heterozygous for the minor allele of the rs34536443 SNP (average across analyses n=21) compared to previous work (n=9; (Couturier et al., 2011)) as well as a larger total sample size, thus increasing the overall statistical power of the presented analyses (see Chapter 2, section 2.16.2.). Overall, the step-wise trend in IFN-β-induced pSTAT levels across risk homozygous, heterozygous and protective homozygous individuals is consistent with a gene dosage effect as indicated by the association of the rs34536443 SNP with disease, with allelic risk following a conventional multiplicative model such that individuals homozygous for the protective allele are more protected than individuals carrying just one copy of the protective allele (heterozygotes). The importance of investigating signalling using all three genotype groups, to assess whether potential genotype-dependent phenotypic differences are consistent with the disease association pattern for these variants, is further emphasized when considering discrepancies between the presented findings and previous work. In the data presented in this Chapter, there was no significant correlation between rs34536443 genotype and IL-6-induced STAT3 phosphorylation, however a previous study has reported a reduction in IL-6-induced gene expression in individuals heterozygous for the rs34536443 SNP compared to risk homozygous individuals (Couturier et al., 2011). In the presented work, there was a similar trend towards reduced IL-6 signalling if considering rs34536443 risk allele homozygotes and heterozygotes only, however statistical analyses across the three genotype groups demonstrated no significant effect of the rs34536443 SNP on proximal IL-6 signalling across cell subsets; there was no significant step-wise trend in the pSTAT readout mirroring the pattern of disease risk (as was observed for IFN-β signalling). In line with the observed lack of an effect of genotype on IL-6 signalling, this pathway is not impaired in *Tyk2*-KO mice (Shimoda et al., 2000).

Previously, reduced IL-10-induced gene expression has been reported in expanded T cells derived from individuals heterozygous for the protective minor allele of the rs34536443 SNP (Couturier et al., 2011), however given the broadly anti-inflammatory role of IL-10, it is unclear how such a reduction in IL-10 signalling would be associated with protection against autoimmune disease. Indeed in the presented work, there was no significant effect of rs34536443 genotype on STAT3 signalling induced by IL-10 stimulation across the cell subsets studied. In addition, no effect of rs34536443 genotype was observed for IL-13-induced pSTAT6 levels in monocytes, which is consistent with the protective role of IL-13 in models of MS-like disease (Cash et al., 1994; Ochoa-Reparaz et al., 2008).

The significant genotype-dependent effect on IFN-β signalling and lack of an effect on that of IL-6, IL-10 and IL-13 cannot simply be explained by differential usage of certain JAK or STAT molecules at each respective receptor. Type I IFNs, IL-10 and IL-13 all mediate signalling via TYK2 and JAK1 (Watford and O'Shea, 2006), however considering the presented findings it seems that the ability of JAK1 to compensate for reduced TYK2 activity varies across these pathways, thereby also indicating that TYK2 is less important in mediating signal transduction downstream of certain cytokines. The pathway-specific effect of genotype is also not STAT-specific as STAT3 phosphorylation is significantly reduced in a genotype-dependent fashion in the context of IFN-β signalling, but not in response to IL-6 and IL-10 stimulation. It is possible that the differential relative importance of TYK2 activity across cytokine pathways relates to de-phosphorylation: for example, protein tyrosine phosphatase epsilon selectively dephosphorylates STAT3 when it is phosphorylated downstream of stimulation with IL-6 and IL-10, but not with IFN-β (Tanuma et al., 2001).

5.4.3. The restricted functional effects of rs34536443 genotype suggest a more pathway-specific role for TYK2: IFN-β signalling

Overall, the presented findings suggest a more restricted role of TYK2 in cytokine signalling than would be predicted based on previous work. For example, somewhat more general signalling

defects have been observed across multiple cytokine pathways in humans and mice deficient in TYK2 (Karaghiosoff et al., 2000; Minegishi et al., 2006; Shimoda et al., 2000). However, as only a single TYK2-deficient patient has been fully characterised (Minegishi et al., 2006), the effects of TYK2 deficiency in humans have not been confirmed. In addition, the complete absence of TYK2 protein may have more severe implications than the presence of a TYK2 variant with reduced activity, for example in terms of the regulation of receptor levels (Gauzzi et al., 1997; Ragimbeau et al., 2003; Zheng et al., 2011), with this having further implications on cytokine signalling. Notably, however, the study of *Tyk2*-KO mice does suggest that IL-6 signalling is not impaired even in the absence of Tyk2 (Shimoda et al., 2000). This is in contrast to previously published work on the effects of rs34536443 genotype in which IL-6-induced signalling was affected in expanded human T cells (Couturier et al., 2011). However, our data from a larger cohort that includes individuals homozygous for the minor protective allele of the rs34536443 SNP, demonstrates no genotype-dependent effects on IL-6, IL-10 or IL-13-induced signalling in primary human immune cells directly *ex vivo*.

In the presented work, the effects of rs34536443 genotype were restricted to IFN-β signalling, and future work will involve an investigation of potential genotype-dependent effects on IFN-α signalling. Whilst these type 1 IFNs share the same receptor signalling complex, IFN-β interacts more strongly than IFN-α with IFNAR1 (de Weerd et al., 2013; Kalie et al., 2008), to which TYK2 binds (Colamonici et al., 1994), whereas JAK1 interacts with the IFNAR2 chain (Domanski et al., 1997)). IFN-β has also been demonstrated to be more potent than IFN-α at inducing gene expression (Leaman et al., 2003). Interestingly, a recent paper has demonstrated that the IFN-β-IFNAR1 complex is capable of mediating signalling in the absence of IFNAR2 (de Weerd et al., 2013), suggesting that the relative importance of TYK2 may differ between type 1 IFNs. Consistent with this, recently published work characterising the inhibitory effects of a TYK2-selective compound demonstrated no effect on IFN-α signalling in primary human fibroblasts (Sohn et al., 2013), although the effect on IFN-β signalling was not investigated. Additionally, both overlapping and distinct gene expression patterns induced by IFN-α and IFN-β have been identified further indicating that these cytokines have some non-redundant functions (Jaitin et al., 2006; Pappas et al.,

2009). Therefore, it is possible that the rs34536443 SNP will not affect both IFN-α and IFN-β signalling in the same way, and this will have implications for our understanding of the effects of rs34536443 genotype in the context of the associated autoimmune diseases.

5.4.4. Type 1 IFNs in autoimmune disease

Since the anti-viral properties of IFNs were first discovered over 50 years ago (Isaacs and Lindenmann, 1957), the complexity of their signalling and their role in regulating both innate and adaptive immunity has become increasingly appreciated (de Weerd and Nguyen, 2012). There are three main IFN families: type 1 IFNs include 13 IFN-α sub-types, a single IFN-β and the less well characterised IFN-ε, -κ and -ω; IFN-γ is the only member of the type 2 IFN family, and the type 3 IFN-λ family is the most recently described (de Weerd and Nguyen, 2012). Type 1 IFNs have the most extensive range of biological activities and have both beneficial and detrimental effects, for example, although type 1 IFNs mediate anti-viral effects, IFN-β signalling is toxic and lethal in the context of certain bacterial infections (Decker et al., 2005). As previously discussed, IFN-α and IFN-β have overlapping but also distinct roles. A CNS-specific protective role for IFN-β has been demonstrated in models of virus infection (Sandberg et al., 1994) and of autoimmune disease (EAE) (Teige et al., 2003), and notably IFN-β is prescribed as a first-line disease-modifying therapy for MS (Haghikia et al., 2013). However, this is inconsistent with the genotype-dependent effect of the rs34536443 SNP, whereby reduced IFN-β signalling correlates with protection against MS, although for other autoimmune conditions a predominantly protective effect of type 1 IFN signalling has not been described.

There are several possible explanations that would reconcile these discrepancies. Firstly, reduced IFN-β signalling may protect against autoimmune disease development, but may have an opposing effect with respect to MS progression subsequent to onset, potentially through an effect on the neuronal rather than immune compartment (Dhib-Jalbut and Marks, 2010; Sandberg et al., 1994; Teige et al., 2003). Secondly, if the rs34536443 SNP is found to also influence IFN-α-induced

signalling in future work, it may be that this effect specifically confers protection against autoimmunity, whereas the effect on IFN-β is less relevant. Thirdly, with respect to the effect of the rs34536443 SNP in the context of MS, it is possible that the protective influence of the minor allele is predominantly due to its influence on the TYK2-mediated signalling induced by other cytokines such as IL-12 or IL-23, and these pathways will be investigated in future work.

With respect to how reduced type 1 IFN signalling may promote protection against the development of autoimmune disease, this may relate to reduced anti-viral immune responses, consistent with the increased viral susceptibility observed in TYK2 deficiency (Karaghiosoff et al., 2000; Minegishi et al., 2006; Prchal-Murphy et al., 2012; Strobl et al., 2005). The mounting of a weaker response to viral infection could reduce the chance of viral-triggering events that are implicated in the initiation of autoimmune disease, such as bystander activation and molecular mimicry (as discussed in Chapter 1, section 1.2). It has recently been shown that type 1 IFNs are anti-viral in acute infection but, somewhat paradoxically, promote viral persistence in chronic infection (Teijaro et al., 2013; Wilson et al., 2013). Thus, another hypothesis is that a genotype-dependent reduction in type 1 IFN signalling may in fact be involved in diminishing chronic viral persistence, therefore reducing the chance of autoreactive T cell activation. Further investigation is required to determine the mechanisms linking reduced type 1 IFN signalling to protection against autoimmune disease in the context of effects on viral immunity.

Elevated production of type 1 IFNs is characteristic of an anti-viral immune response but is also associated with many autoimmune diseases, where it may be a general marker of inflammation and/or be directly involved in promoting the inflammatory response. A distinctive pattern of type 1 IFN-dependent gene expression, termed the IFN signature, is most prominent in SLE where it correlates with disease severity (Baechler et al., 2003) and is likely to be pathogenic, as anti-IFN-α therapy is proving beneficial (Merrill et al., 2011). The association of the rs34536443 SNP with SLE has not yet been assessed in a large case-control cohort, however the SNP is associated with PS and RA, which have both been described as having an IFN signature and thus the association of a protective variant that reduces type 1 IFN signalling would be consistent with this (van der Fits et al., 2004; van der Pouw Kraan et al., 2007).

In the case of MS, patients that are unresponsive to IFN-β therapy have been reported to have increased levels of endogenous type 1 IFNs prior to treatment (Axtell et al., 2010) consistent with previous studies that had suggested the presence a type 1 IFN signature in immune cells from IFN-β non-responders (Comabella et al., 2009). This may indicate that at least in a subset of MS patients, type 1 IFN may have a detrimental effect as is postulated for several other autoimmune conditions. One hypothesis to explain these findings proposes that type 1 IFN signalling has a protective effect in MS patients with a more Th1-driven disease, but a detrimental effect in patients with more Th17-driven disease (Axtell et al., 2013). This hypothesis is based on observations from adoptive transfer experiments performed in the context of EAE (Axtell et al., 2010). In line with this, MS patients identified as unresponsive to IFN-β therapy have been shown to have elevated serum IL-17F prior to commencing treatment (Axtell et al., 2010) suggesting that their disease was more Th17-driven.

Whilst this model may provide an explanation as to how a genetic variant can mediate protection against MS via reducing IFN-β signalling, the extent to which this model may apply to other autoimmune diseases is unclear. For example, the Th17 axis is heavily implicated in diseases that are not associated with the rs34536443 SNP such as IBD and AS, where a genetically-driven reduction in IFN-β signalling would also be predicted to be protective based on the model for MS. However, other cytokines such as IL-12 and IL-23 are known to be relevant to Th1 and Th17 cell-driven pathology, respectively. As TYK2 is associated with the receptor complex for both of these cytokines, further investigation is required to determine whether there is an effect of rs34536443 genotype on IL-12 and/or IL-23 signalling. Further delineating the effects of this SNP will be informative in determining the relevance of certain cytokine pathways across the associated autoimmune diseases. Moreover, as rs34536443 genotype-dependent effects on IFN-β signalling were observed not only for CD4+ T cells but also for CD8+ T cells, B cells and monocytes, the relative importance of changes in cytokine-induced signalling in these cell types also remains to be determined. For example, type 1 IFNs enhance the survival and proliferation of activated CD4+ and CD8+ T cells (Marrack et al., 1999) and also of mature B cells (Morikawa et al., 1987; Su and David,

1999), which may contribute to autoimmune disease pathogenesis and be influenced by the effects of rs34536443 genotype on IFN-β signalling.

5.4.5. Future work: investigating rs34536443 genotype-dependent effects on other cytokine pathways and on immune cell signalling *in vivo*

Studies of IL-12 and IL-23 signalling by genotype have not been performed for the presented work but such investigations are planned and optimisation is ongoing in order to upregulate the IL-12Rβ1 subunit shared by the IL-12 and IL-23 receptors, which is not expressed on resting peripheral T cells (Fahey et al., 2007). Interestingly, recently published work reports inhibition of IL-12, but not IL-23, signalling by a TYK2-selective inhibitor in expanded human CD4+ T cells (Sohn et al., 2013), suggesting that TYK2 function may be more important for signalling downstream of the former cytokine. Interestingly, whilst both IL-12 and IL-23 are heavily implicated in autoimmunity, both genetic associations at several loci and also drug efficacies suggest a degree of disease-specificity to the precise relevance of each cytokine, which to some extent parallels the disease association pattern of the rs34536443 SNP. For example, it is intriguing to speculate that the rs34536443 SNP may influence IL-12 signalling but not that of IL-23, noting that this SNP is not associated with IBD or AS. Both of these diseases, however, show associations with the rs11209026 nsSNP in *IL23R*, with the minor allele conferring protection (Cortes et al., 2013; Jostins et al., 2012) by reducing IL-23R-mediated signalling (Di Meglio et al., 2011; Di Meglio et al., 2013; Pidasheva et al., 2011).

Whilst a lack of a putative effect of the rs34536443 SNP on IL-23 signalling is consistent with the lack of an association with IBD and AS, it seems paradoxical when considering that IL-23 signalling is involved in the development of Th17 cells and that these cells are also implicated in diseases that *are* associated with the rs34536443 SNP. However, although ustekinumab, a drug that blocks both IL-12 and IL-23 signalling, is efficacious for CrD (Sandborn et al., 2008; Sandborn et al., 2012), this drug showed no clinical benefit in trials for MS (Segal et al., 2008). Conversely, secukinumab, which blocks IL-17 (the signature cytokine produced by Th17 cells) has shown efficacy in MS (Havrdová, 2012) but unexpectedly not in CrD (Hueber et al., 2012). This raises the

possibility that IL-23 signalling does not solely act to generate Th17 cells in the context of autoimmune disease. Indeed, the pathogenic relevance of IL-17 and IFN-γ double-positive cells in IBD has been recognised, for example, with these cells requiring IL-23 for their development (Ahern et al., 2010). Thus, elucidating the exact nature of the cytokine pathways most affected by rs34536443 genotype, and then further investigating their downstream consequences, will facilitate a deeper insight into the pathophysiological mechanisms that are shared by or specific to different autoimmune diseases.

Analysis of genotype-dependent differences in gene expression patterns of different immune cell subsets will enable further delineation of the downstream effects of rs34536443 genotype. In addition, the impact of this nonsynonymous variation will be investigated *in vivo* using a transgenic mouse model that has been generated, which carries the orthologous amino acid substitution. In particular, the implications of this substitution will be studied in the context of TCR-MHC class I and II humanised mouse models (Friese et al., 2008; Gregersen et al., 2006). Moreover, further functional analyses will also be performed in order to determine whether the potential non-canonical roles of TYK2, such as in type 1 IFN-induced NFκB signalling (Yang et al., 2005a) and its action in the nucleus (Ahmed et al., 2013), are also relevant to autoimmune disease.

6. Genotype-to-phenotype correlations: investigating the effects autoimmune disease-associated SNPs on gene expression in the *TYK2* region

6.1 Introduction

The disease association pattern of SNPs in the *TYK2* region (Table 6.1 and Figure 6.1.), and the data from functional studies presented in the previous Chapters collectively suggest that not all of the disease-associated SNPs in the region are likely to influence TYK2 activity. Based on the results of Chapters 3 and 5, the minor allele of the rs34536443 SNP correlates with reduced cytokine-induced TYK2 kinase activity across immune cell subsets. However, this SNP is not associated with IBD, which suggests that the variants that *are* associated with IBD are likely to have distinct disease-associated functional consequences outside of effects on TYK2. The rs12720356 and rs11879191 SNPs are both associated with IBD, and there is evidence that the former may correlate with some genotype-dependent differences on TYK2-mediated signalling, although these differences are not as consistent as for the rs34536443 SNP (see Chapter 5) and considering the above rationale, it is hypothesised that the rs12720356 SNP will have additional biological effects. The rs11879191 SNP is located in intron 1 of the gene neighbouring *TYK2*, which encodes cell division cycle 37 (*CDC37*), and may thus influence expression of the *CDC37* gene or other genes in the region, especially given that physical proximity to a specific gene is not a robust indicator of whether or not that gene is affected by the SNP. Similarly, the rs34536443 SNP has not been associated with T1D, and variation at the T1D-associated rs2304256 SNP was not found to correlate with genotype-dependent effects on TYK2 signalling in the previously presented cell line work (see Chapter 3). This suggests that although this SNP alters the amino acid sequence of TYK2, its disease-associated functional effects may not be on TYK2.

Table 6.1. Autoimmune disease-associated SNPs in the *TYK2* gene region
GWAS and ImmunoChip have identified four SNPs in the *TYK2* gene region that are associated with autoimmune disease. For details of minor allele frequency and odds ratios see Chapter 1, Table 1.1.

SNP	Associated disease	References
rs34536443	MS	(Ban et al., 2009; IMSGC, 2013; Johnson et al., 2009; Mero et al., 2009)
	PBC	(Liu et al., 2012)
	RA	(Eyre et al., 2012)
	JIA	(Hinks et al., 2013)
	PS	(Tsoi et al., 2012)
rs12720356	CrD	(Anderson et al., 2011; Franke et al., 2010)
	UC	(Anderson et al., 2011)
	PS	(Strange et al., 2010; Tsoi et al., 2012)
rs11879191	CrD	(Jostins et al., 2012)
	UC	(Jostins et al., 2012)
	AS	(Cortes et al., 2013)
rs2304256	T1D	(Parkes et al., 2013; Wallace et al.)
	SLE	(Suarez-Gestal et al., 2009)

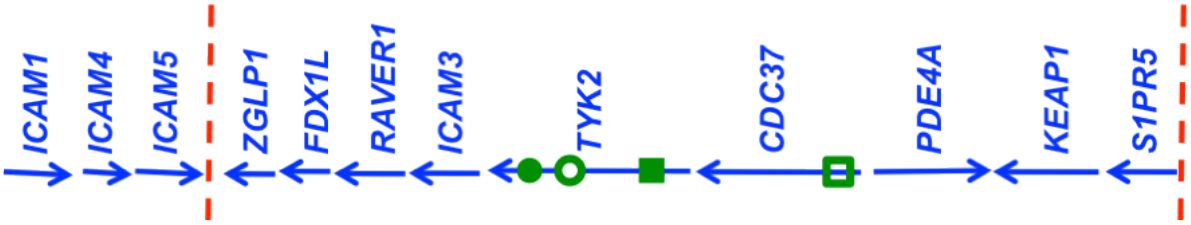


Figure 6.1. Genes in the *TYK2* region on chromosome 19p13 and the location of the focal SNPs
Diagrammatic representation of genes (blue lines) in the *TYK2* region, with arrowheads indicating the direction of transcription. Red dashed lines indicate the boundaries of the association region based on recombination rate information from the HapMap phase II recombination map (Sawcer et al., 2011). The location of the four autoimmune disease-associated SNPs of interest is indicated by the green symbols: the rs34536443 SNP is located in exon 21 of the *TYK2* gene (filled circle), the rs12720356 SNP is located in exon 13 (empty circle) and the rs2304256 SNP (filled square) is in exon 6. The rs11879191 SNP (empty square) is located in intron 1 of the gene encoding cell division cycle 37 (*CDC37*). *ICAM*: intercellular adhesion molecule. *ZGLP1*: zinc finger GATA-like protein 1. *RAVER1*: ribonucleoprotein, PTB-binding 1. *PDE4A*: cAMP-specific phosphodiesterase 4A. *KEAP1*: kelch-like ECH-associated protein 1. *S1PR5*: sphingosine-1-phosphate receptor 5. Diagram modified from (Sawcer et al., 2011).

If the rs12720356, rs11879191, and rs2304256 SNPs are capable of influencing genes other than *TYK2*, it was hypothesised that they (or SNPs in high LD with them; Table 6.2.) may be located in regulatory elements, thereby having an impact on gene expression. Notably, none of the focal variants or SNPs in high LD with them (Figure 6.1. and Table 6.2.) are located proximal to intron-

exon boundaries and thus they are unlikely to affect splicing. There is at least some evidence that these SNPs may be located within regulatory elements, based on publicly available data from the ENCODE Project Consortium (Bernstein et al., 2012; Birney et al., 2007) that has aimed to identify and characterise functional elements within the human genome in a high-throughput manner (http://genome.ucsc.edu/ENCODE/). For example, the rs11879191 SNP co-localises with a DNase hypersensitivity site in data from 125 ENCODE cell lines and also with certain histone lysine methylation marks, which are both characteristic of regulatory elements. In addition, a CREB-binding protein (CBP) interaction site co-localises with the rs11879191 SNP in the Jurkat T cell line (Hollenhorst et al., 2009), further suggesting a potential regulatory element at this site, as CBP is involved in co-activating many transcription factors (Vo and Goodman, 2001).

Table 6.2. SNPs that are in high LD with the focal disease-associated SNPs
The focal disease-associated SNPs may in fact be markers of the actual causal variants, which would be in high LD. Using information provided by our collaborator Professor Gilean McVean (WTCHG, University of Oxford), data from recent ImmunoChip publications (Cortes et al., 2013; Hinks et al., 2013) and an online tool (http://www.broadinstitute.org/mpg/snap/ldsearch.php), SNPs in high LD (r^2>0.9) with the focal variants were identified.

Focal SNP	Other SNPs in high LD (r^2>0.9)
rs2304256	rs35251378, rs11085725, rs34725611
rs12720356	-
rs34536443	rs74956615
rs11879191	rs35164067

Data from several expression quantitative trait loci (eQTLs) studies further support the hypothesis that the rs12720356, rs11879191, and rs2304256 SNPs may exert regulatory effects on gene expression in the *TYK2* region (Table 6.3.). For example, data from lymphoblastoid cell lines (LCLs) correlate alleles of the rs11879191 SNP with changes in the expression of the *CDC37* gene (Veyrieras et al., 2008) and also correlate rs12720356 genotype with expression differences for the *FDX1L* and *PDE4A* genes (Brown et al., 2013). Notably, these eQTL data were obtained from LCLs and thus they must be considered with caution due to potential cell line-specific artefacts (Choy et al., 2008; Plagnol et al., 2008). Some eQTL data are available from primary human monocytes

(Fairfax et al., 2012), with rs12720356 genotype having been correlated with ICAM4 RNA-level expression differences, noting that only two individuals homozygous for the minor allele were analysed, as donors were not selected by genotype in this study. From this same data set, rs2304256 genotype was correlated with ICAM3 expression in monocytes (Fairfax et al., 2012), however, the majority of other immune cell subsets have not been considered to date.

Table 6.3. Allele-specific gene expression correlating with SNPs in the *TYK2* region
Three of the four autoimmune disease-associated SNPs in the *TYK2* gene region are significantly correlated with changes in the expression of particular genes in certain cell types. There are no eQTL data for the rs34536443 SNP.

SNP I.D.	Gene affected	Cell type	Reference
rs12720356	*ICAM4*	Primary human monocytes	(Fairfax et al., 2012)
	FDX1L	Lymphoblastoid cell lines (LCLs)	(Brown et al., 2013)
	PDE4A		
rs2304256	*ICAM3*	Primary human monocytes	(Fairfax et al., 2012; Wallace et al., 2012)
rs11879191	*CDC37*	LCLs	(Veyrieras et al., 2008)

6.2 Chapter aim

The aim of this chapter was to investigate whether the autoimmune disease-associated SNPs in the *TYK2* region influence the expression levels of *TYK2* and/or other genes in the region (Figure 6.1.), particularly noting that other genes, such as *ICAM1* and *ICAM3*, are also immunologically-relevant candidates. RNA-level expression analyses were performed by genotype using primary human lymphocytes, monocytes and granulocytes, following a characterisation of the expression pattern of the genes of interest. The only gene in the region that was not investigated was *ICAM5* as it is known to only be expressed in neurons (Mori et al., 1987). In addition, protein-level expression was assessed in parallel for gene products expressed at the cell surface (ICAM1, ICAM3 and ICAM4), using primary human PBMCs.

6.3 Results

As indicated in Chapters 4 and 5, donors were pre-selected by genotype and recalled for studies of cytokine receptor expression and phospho-signalling, and thus immune cell subsets were also obtained for RNA-level analyses. In addition, samples that were collected from a cohort used for previous studies in the laboratory (Gregory et al., 2012) were also included to increase the sample size. Comparing the genotype frequencies that would be expected for these SNPs in a healthy control population to the frequencies observed in the selected cohort, pre-selection enabled enrichment of individuals homozygous for the minor allele of the focal SNPs by between 4- and 12-fold (highlighted blue in Table 6.4.).

Table 6.4. Pre-selection of donors from a genotyped cohort for function investigations of SNPs in the *TYK2* gene region
Donors were pre-selected from a large cohort of genotyped individuals. Expected genotype frequencies were calculated from cited minor allele frequencies (MAF; (IMSGC, 2013; Jostins et al., 2012; Parkes et al., 2013; Tsoi et al., 2012) and assuming Hardy-Weinberg equilibrium. Observed frequency refers to values calculated from the total cohort of donors recruited to provide samples. Risk = homozygous for the risk allele. Het = heterozygote. Prot = homozygous for the protective allele. For the rs12720356 SNP, the direction of disease association refers to IBD (PS is associated in the opposite direction; see Chapter 1, Table 1.1.).

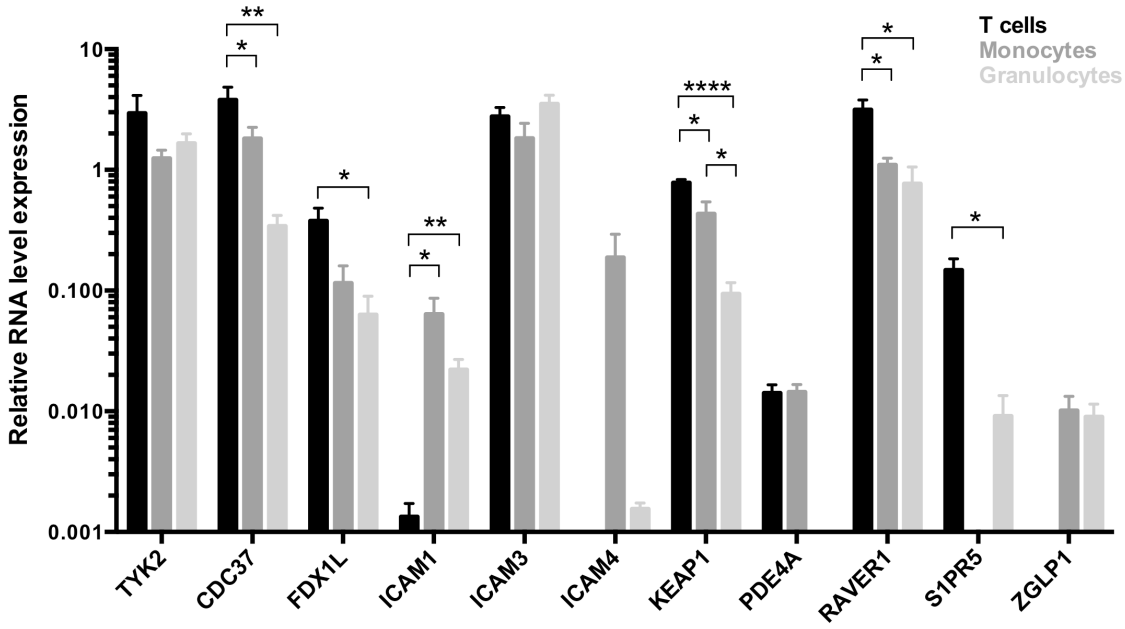
SNP (abbreviation)	MAF	Expected frequency			Frequency in cohort used			Fold enrichment		
		Risk	Het	Prot	Risk	Het	Prot	Risk	Het	Prot
rs34536443 (rs345..)	0.05	90.25	9.50	0.25	78.00	18.90	3.10	0.86	1.97	12.40
rs12720356 (rs127..)	0.10	1.00	18.00	81.00	4.40	16.30	79.40	4.40	0.91	0.98
rs2304256 (rs230..)	0.27	53.29	39.42	7.29	22.80	38.60	38.60	0.72	0.98	5.29
rs11879191 (rs118..)	0.20	64.00	32.00	4.00	34.40	42.70	22.90	0.54	1.33	5.73

6.3.1. Expression of genes in the *TYK2* region in resting and activated primary immune cell subsets

The expression of genes in the *TYK2* region was characterised through a preliminary analysis, in order to identify the most appropriate cell type and activation state in which to investigate potential genotype-dependent correlations. RNA-level expression was investigated in the most

abundant PBMC subsets under resting conditions: T cells, monocytes and granulocytes, noting that the former two subsets have been most strongly implicated in the autoimmune diseases with which the *TYK2* region SNPs are associated (Table 6.1.).

In certain cell subsets, the RNA-level expression of particular genes was undetectable, for example ICAM4 and ZGLP1 in T cells, S1PR5 in monocytes and PDE4A in granulocytes. Expression of TYK2 did not significantly differ between cell subsets (Figure 6.3.1.). T cells expressed significantly more KEAP1 and RAVER1 than monocytes (P=0.0287 and 0.0209, respectively; Figure 6.3.1.) and granulocytes (P<0.0001 and 0.0148, respectively; Figure 6.3.1.). Expression of CDC37, FDX1L and S1PR5 was also significantly higher in T cells relative to granulocytes (P=0.0016, 0.0261 and 0.0081, respectively; Figure 6.3.1.). ICAM1 expression in T cells was significantly lower than in monocytes and granulocytes (P=0.0335 and 0.0039, respectively; Figure 6.3.1.). Monocytes expressed significantly higher levels of both CDC37 and KEAP1 than granulocytes (P=0.0162 and 0.0256, respectively; Figure 6.3.1.)



Gene	T cells vs monocytes	T cells vs granulocytes	Monocytes vs granulocytes
TYK2	ns	ns	ns
CDC37	ns	0.0016 (T>G)	0.0162 (M>G)
FDX1L	ns	0.0261 (T>G)	ns
ICAM1	0.0335 (T<M)	0.0039 (T<G)	ns
ICAM3	ns	ns	ns
ICAM4	-	-	ns
KEAP1	0.0287 (T>M)	<0.0001 (T>G)	0.0256 (M>G)
PDE4A	ns	-	-
RAVER1	0.0209 (T>M)	0.0148 (T>G)	ns
S1PR5	-	0.0081 (T>G)	-
ZGLP1	-	-	ns

Figure 6.3.1. RNA-level expression of genes in the TYK2 region in resting primary human immune cell subsets
The expression of genes of interest was quantified by qPCR relative to expression of the housekeeping gene ubiquitin C (*UBC*) in resting T cells, monocytes and granulocytes (n=3-4). Expression levels were compared between cell subsets and significant *P*-values are shown in the table, where a dash indicates that gene expression was undetectable in one of the subsets under comparison. ns = not significant (*P*>0.05). For significant differences between cell subsets, the relationship is shown in brackets, where M = monocytes, T = T cells and G = granulocytes.

In order to investigate whether expression of the genes of interest was influenced by cellular activation state, stimulations were performed for the most likely disease-relevant subsets, T cells and monocytes. The stimulants used were selected to be subset-specific and the time-points shown are those for which there was a change in expression of at least one gene in the region compared to resting cells and where the cells remained viable. Stimulation of T cells with anti-CD3/CD28

Dynabeads® and 30 U/ml IL-2 did not significantly alter expression levels of CDC37, FDX1L, ICAM1, PDE4A and RAVER1 (P>0.05 for all; Table 6.5.). Compared to unstimulated T cells, levels of ICAM3 and S1PR5 were significantly reduced at all stimulation timepoints investigated (Table 6.5.), however levels of KEAP1 were significantly increased after stimulation for 24h and 72h (P=0.0439 and 0.0446, respectively; Table 6.5.). Expression of TYK2 was significantly reduced in T cells after stimulation for 24h and 72h (P=0.0012 and 0.0035, respectively; Table 6.5.).

Expression of TYK2, CDC37, FDX1L, ICAM1, KEAP1, PDE4A and RAVER1 was not significantly different between unstimulated monocytes and those stimulated for 48h with 500 U/ml IFN-γ (P>0.05 for all; Table 6.6.). In contrast, stimulated monocytes expressed significantly lower levels of ICAM3, ICAM4 and ZGLP1 compared to unstimulated cells (P=0.0002, <0.0001 and 0.0229, respectively; Table 6.6.).

As the expression of genes of interest was detected in at least one resting cell subset, and across the genes of interest and the subsets considered expression was predominantly unaltered by stimulation, subsequent qPCR analyses by genotype were performed on cDNA samples from unstimulated cell subsets.

Table 6.5 Investigating the effect of T cell activation on expression of the genes of interest
The expression of genes of interest was quantified by qPCR relative to expression of the housekeeping gene *UBC* in T cells stimulated for 24h, 72h or 120h (n=4) with Human T Activator Dynabeads® at a bead:cell ratio of 1:1 and 30 U/ml IL-2. Expression was normalised to the unstimulated condition (0h). The *P*-values calculated from t-test analyses are shown and for significant differences the arrow indicates whether expression was increased (↑) or decreased (↓) upon stimulation. ns = not significant (P>0.05).

Gene	0h vs 24h	0h vs 72h	0h vs 120h
TYK2	ns	0.0012 ↓	0.0035 ↓
CDC37	ns	ns	ns
FDX1L	ns	ns	ns
ICAM1	ns	ns	ns
ICAM3	0.0006 ↓	0.0009 ↓	0.003 ↓
KEAP1	0.0439 ↑	0.0446 ↑	ns
PDE4A	ns	ns	ns
RAVER1	ns	ns	ns
S1PR5	0.0002 ↓	0.0088 ↓	0.0079 ↓

Gene	Stim vs unstim
TYK2	ns
CDC37	ns
FDX1L	ns
ICAM1	ns
ICAM3	0.0002 ↓
ICAM4	<0.0001 ↓
KEAP1	ns
PDE4A	ns
RAVER1	ns
ZGLP1	0.0229 ↓

Table 6.6. Investigating the effect of monocyte stimulation on expression of the genes of interest
The expression of genes of interest was quantified by qPCR relative to expression of the housekeeping gene *UBC* in monocytes stimulated (stim) for 48h with 500 U/ml IFN-γ (n=4). Expression was normalised to the unstimulated condition (unstim). The *P*-values calculated from t-test analyses are shown and for significant differences the arrow indicates whether expression was increased (↑) or decreased (↓) upon stimulation. ns = not significant (*P*>0.05).

6.3.2. The effect of age and sex on expression of genes in the *TYK2* region in resting immune cell subsets

Prior to analyses of gene expression by SNP genotype, it was important to consider whether there were significant trends in the data when sorted by donor age or sex that could confound potential genotype-dependent trends. Age was found to contribute to a small (3.8%-16.97%) but significant proportion of the variation in expression of several genes in specific cell subsets (Table 6.7.), whereas a significant effect of sex was found only for expression of CDC37 in lymphocytes (*P*=0.0116; Table 6.7.). For the specific assays and cell subsets demonstrating significant trends by age or sex (Table 6.7.), the data were normalised as appropriate prior to considering the existence of potential genotype-dependent effects on gene expression.

Table 6.7. Identifying potential confounding factors: significant trends in gene expression by donor age and sex
For the qPCR assays and cell subsets shown, there were significant trends between gene expression and donor age or sex. Data were normalised as appropriate for subsequent analyses of gene expression by genotype.

Gene	Cell subset	Variable	*P*-value (r^2 for age only)
TYK2	Granulocytes	Age	0.0025 (0.07933)
CDC37	Lymphocytes	Sex	0.0116
	Granulocytes	Age	0.0379 (0.03857)
FDX1L	Granulocytes	Age	0.0004 (0.1052)
KEAP1	Granulocytes	Age	0.0278 (0.04364)
	Monocytes	Age	<0.0001 (0.1697)
S1PR5	Lymphocytes	Age	0.0146 (0.04372)
ZGLP1	Monocytes	Age	0.011 (0.04342)

6.3.3. Investigating expression of genes in the *TYK2* region in resting immune cell subsets by SNP genotype

The *P*-value threshold for significance of genotype-dependent trends in gene expression was corrected for multiple testing; for lymphocytes the corrected significance threshold was set at *P*=0.0083; for monocytes and granulocytes the corrected significance threshold was set at *P*=0.0125 (for details see Chapter 2, section 2.16.3.). In the following figures, only *P*-values reaching the defined significance thresholds are shown and the presented data are from the total donor cohort segregated by genotype at each SNP (rather than by haplotypes defined by multiple SNPs).

There were no significant genotype-dependent correlations in gene expression for the rs34536443 SNP, indicating that the functional consequences of this variant are likely to be restricted to the previously identified effects on TYK2-mediated signalling (Chapters 3 and 5), which were not due to changes in the expression of the *TYK2* gene at the RNA level (Figure 6.3.2.). Indeed there were no significant correlations between TYK2 expression and genotype studied for the three other focal SNPs in any of the cell subsets studied (Figure 6.3.2.), suggesting that their functional effects are mediated via effects on other genes. There were also no genotype-dependent correlations across the investigated cell subsets for expression of FDX1L (Figure 6.3.3.), KEAP1 (Figure 6.3.4.), S1PR5 (Figure 6.3.5.) and ZGLP1 (Figure 6.3.6.).

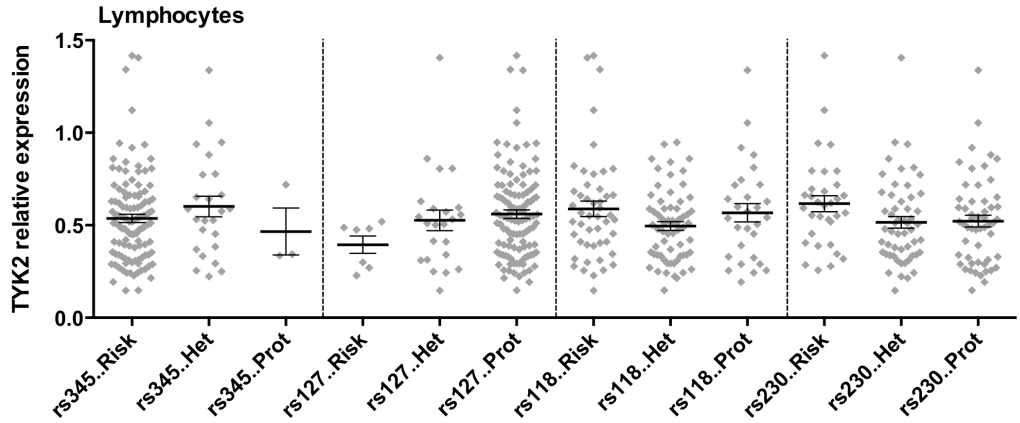
Table 6.8. summarises the significant genotype-dependent correlations identified in the presented RNA-level expression analyses. For the rs12720356 SNP, five significant genotype-dependent correlations were identified. In monocytes, there were significant correlations between rs12720356 genotype and expression of CDC37 (*P*=0.0021, r^2=0.06277; Figure 6.3.7.b), RAVER1 (*P*=0.0067, r^2=0.04956; Figure 6.3.8.b), PDE4A (*P*=0.0035, r^2=0.05801; Figure 6.3.9.b) and ICAM3 (*P*=0.0012, r^2=0.04379); Figure 6.3.10.b), with expression being highest for all genes in individuals homozygous for the protective allele. In granulocytes, ICAM4 expression was found to correlate with rs12720356 genotype (*P*=0.0011, r^2=0.1072, Figure 6.3.11.c) with expression being highest in individuals homozygous for the risk allele of the rs12720356 SNP.

For the rs11879191 SNP, genotype-dependent correlations were identified for expression of ICAM4 in monocytes (P=0.0005, r^2=0.08421; Figure 6.3.11.b) and ICAM1 in granulocytes (P=0.0019, r^2=0.09986; Figure 6.3.12.c), with expression being highest in individuals homozygous for the risk allele. For the rs2304256 SNP, genotype-dependent correlations were identified for expression of PDE4A in lymphocytes (P=0.0011, r^2=0.07258; Figure 6.3.9.a) and ICAM3 in monocytes (P=0.0049, r^2=0.05429; Figure 6.3.10.b), with expression being highest in individuals homozygous for the risk allele.

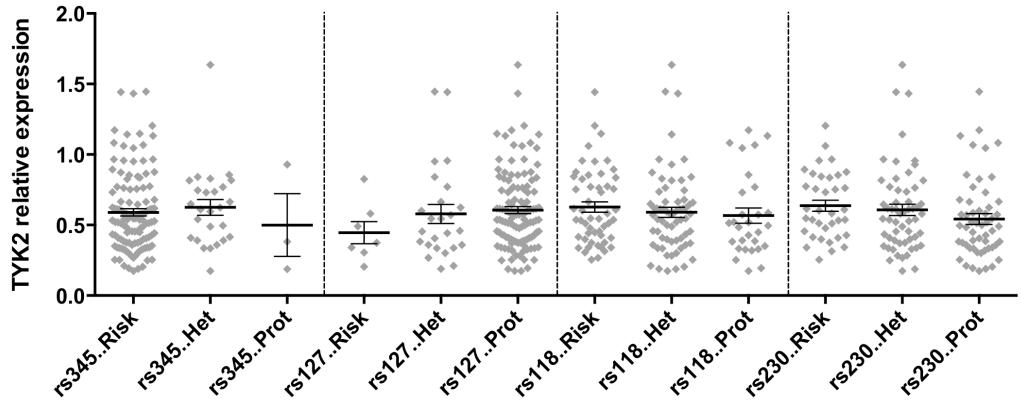
Table 6.8. Summary of correlations between SNP genotype and gene expression
A summary of the identified significant correlations between SNP genotype and RNA-level expression for the genes of interest in the investigated cell subsets (see accompanying Figures). *P*-values reaching the defined significance thresholds are shown, with the respective r^2 values.

SNP I.D.	Cell subset	Gene of interest	Correlation *P*-value (r^2)	Figure
rs12720356	Monocytes	*CDC37*	0.0021 (0.06277)	6.3.7.b
		RAVER1	0.0067 (0.04956)	6.3.8.b
		PDE4A	0.0035 (0.05801)	6.3.9.b
		ICAM3	0.0012 (0.04379)	6.3.10.b
	Granulocytes	*ICAM4*	0.0011 (0.1072)	6.3.11.c
rs11879191	Monocytes	*ICAM4*	0.0005 (0.08421)	6.3.11.b
	Granulocytes	*ICAM1*	0.0019 (0.09986)	6.3.12.c
rs2304256	Lymphocytes	*PDE4A*	0.0011 (0.07258)	6.3.9.a
	Monocytes	*ICAM3*	0.0049 (0.05429)	6.3.10.b

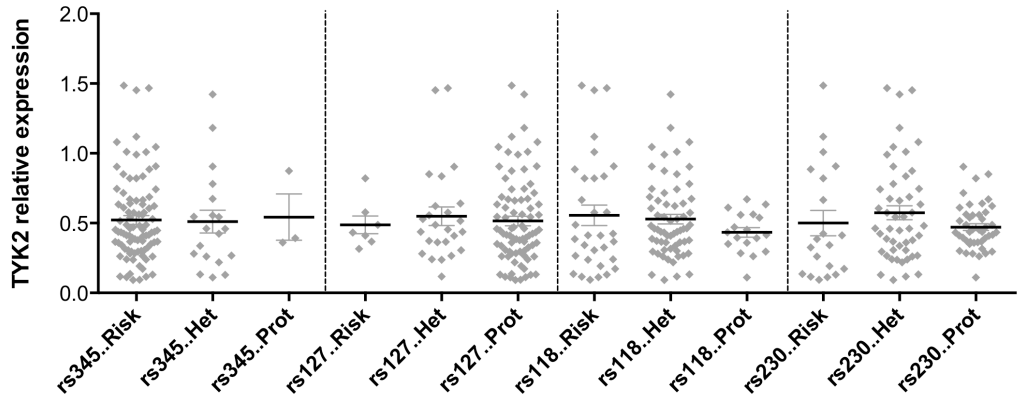
a)



b)



c)

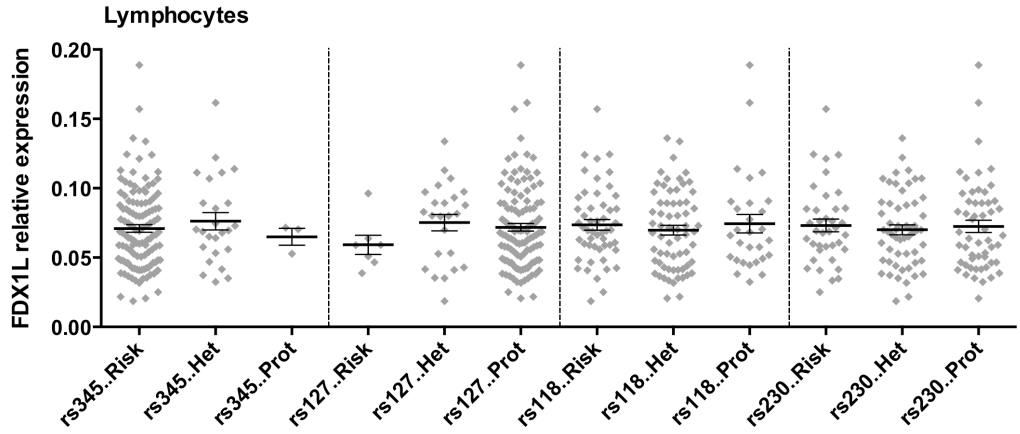


d)

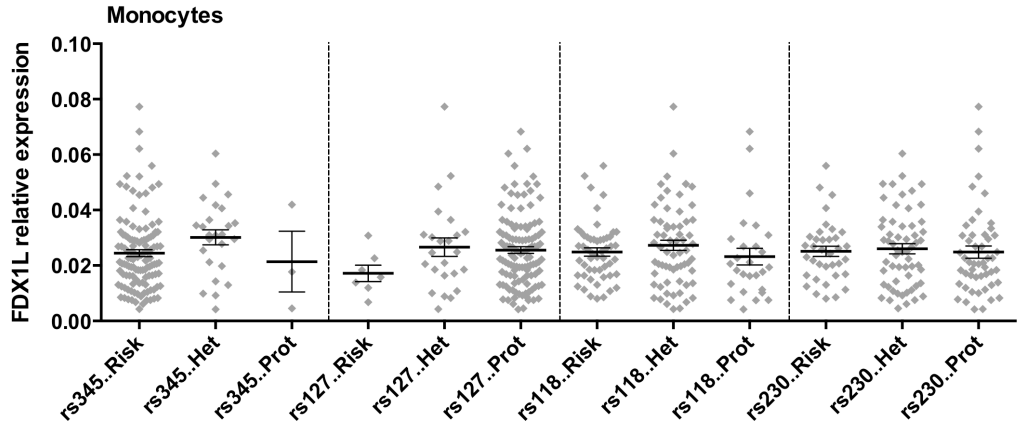
	rs345.Risk	rs345..Het	rs345..Prot	rs127..Risk	rs127..Het	rs127..Prot	rs118..Risk	rs118..Het	rs118..Prot	rs230..Risk	rs230..Het	rs230..Prot
Lymphocytes	115	25	3	7	21	113	48	64	29	33	54	54
Monocytes	120	26	3	7	25	117	52	65	29	36	57	54
Granulocytes	91	19	3	7	25	81	32	63	18	20	49	45

Figure 6.3.2. Expression of TYK2 in immune cell subsets by genotype
TYK2 expression was quantified by qPCR relative to expression of the housekeeping gene *UBC* in lymphocytes (a), monocytes (b) and granulocytes (c) from genotyped individuals. The number of individuals per genotype group is shown for each cell subset (d). In all cell subsets, TYK2 expression was not found to correlate with genotype. Graphs show mean ± SEM.

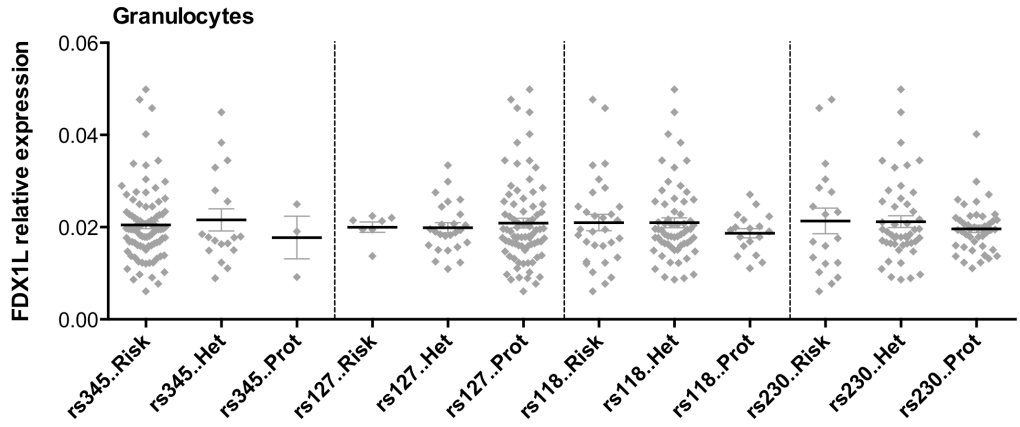
a)



b)



c)

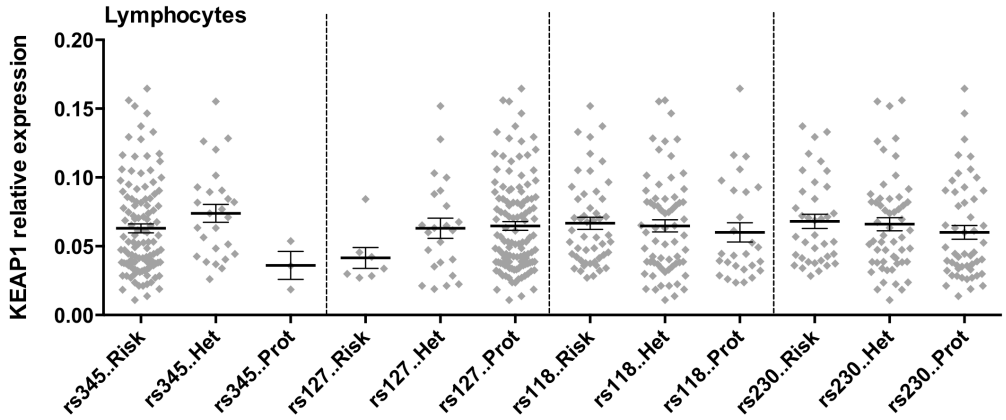


d)

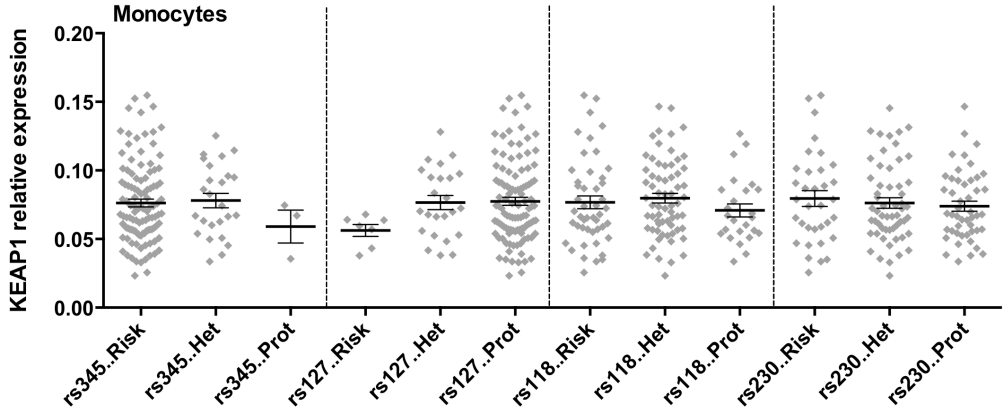
	rs345..Risk	rs345..Het	rs345..Prot	rs127..Risk	rs127..Het	rs127..Prot	rs118..Risk	rs118..Het	rs118..Prot	rs230..Risk	rs230..Het	rs230..Prot
Lymphocytes	119	25	3	7	24	117	51	65	29	38	56	54
Monocytes	119	25	3	7	24	116	52	65	27	36	58	51
Granulocytes	90	18	3	7	25	79	31	62	18	19	48	44

Figure 6.3.3. Expression of FDX1L in immune cell subsets by genotype
FDX1L expression was quantified by qPCR relative to expression of the housekeeping gene *UBC* in lymphocytes (a), monocytes (b) and granulocytes (c) from genotyped individuals. The number of individuals per genotype group is shown for each cell subset (d). In all cell subsets, FDX1L expression was not found to correlate with genotype. Graphs show mean ± SEM.

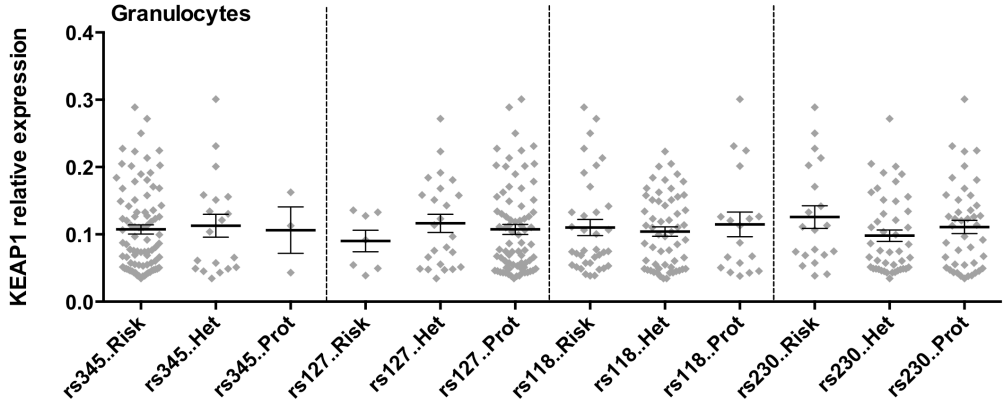
a)



b)



c)

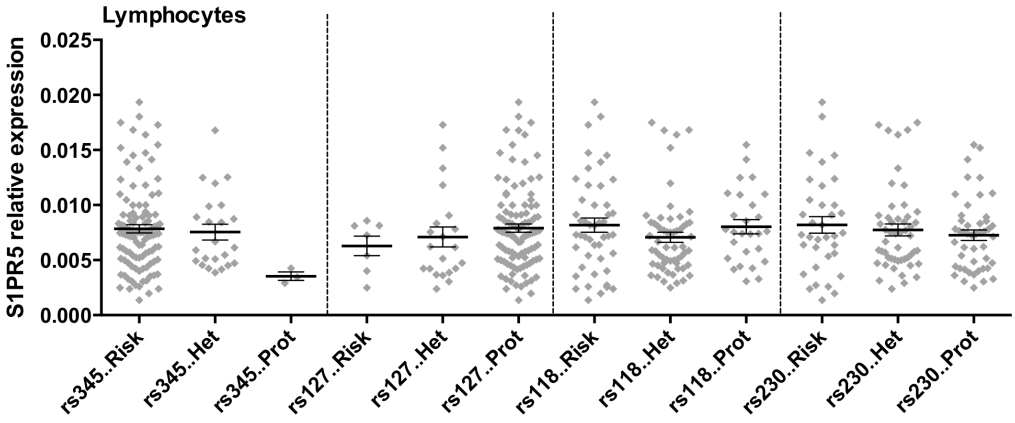


d)

	rs345..Risk	rs345..Het	rs345..Prot	rs127..Risk	rs127..Het	rs127..Prot	rs118..Risk	rs118..Het	rs118..Prot	rs230..Risk	rs230..Het	rs230..Prot
Lymphocytes	114	26	3	7	23	120	50	66	27	37	55	51
Monocytes	110	25	3	7	24	107	47	65	25	35	56	48
Granulocytes	86	19	3	7	24	77	36	57	18	20	45	43

Figure 6.3.4. Expression of KEAP1 in immune cell subsets by genotype
KEAP1 expression was quantified by qPCR relative to expression of the housekeeping gene *UBC* in lymphocytes (a), monocytes (b) and granulocytes (c) from genotyped individuals. The number of individuals per genotype group is shown for each cell subset (d). In all cell subsets, KEAP1 expression was not found to correlate with genotype. Graphs show mean ± SEM.

a)

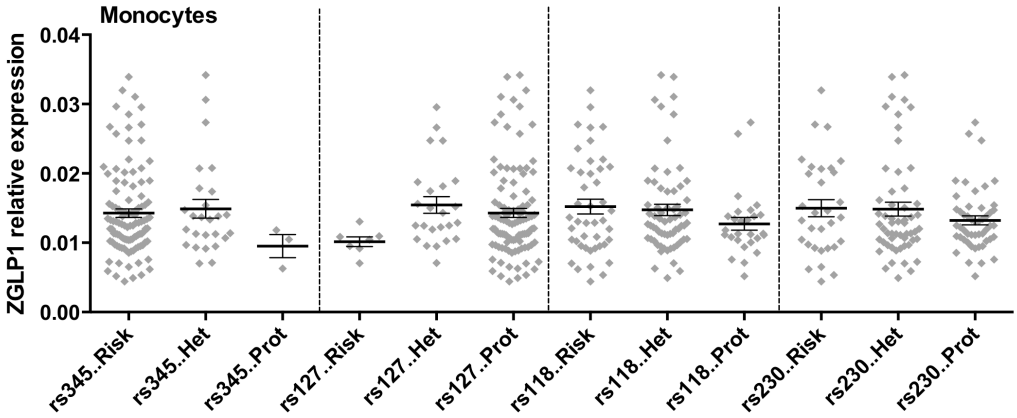


b)

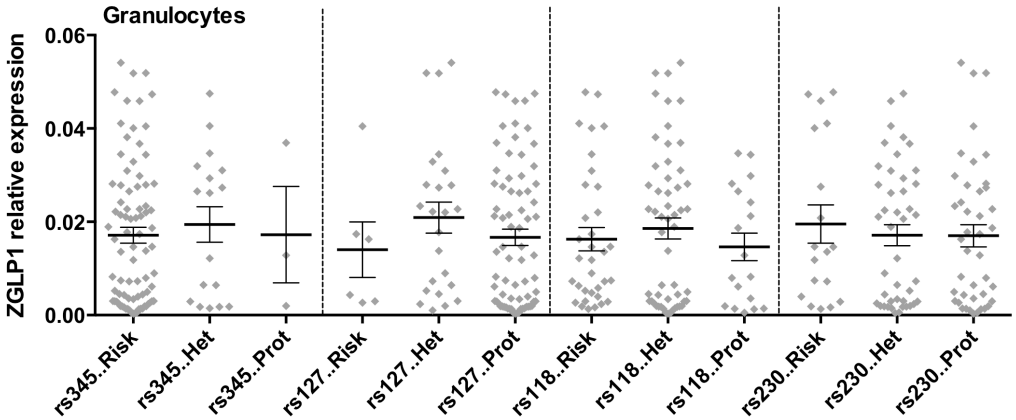
	rs345..Risk	rs345..Het	rs345..Prot	rs127..Risk	rs127..Het	rs127..Prot	rs118..Risk	rs118..Het	rs118..Prot	rs230..Risk	rs230..Het	rs230..Prot
Lymphocytes	116	23	3	7	21	106	47	59	27	35	51	47

Figure 6.3.5. Expression of S1PR5 in lymphocytes by genotype
S1PR5 expression was quantified by qPCR relative to expression of the housekeeping gene *UBC* in lymphocytes (a) from genotyped individuals. The number of individuals per genotype group is shown (b). S1PR5 expression was not found to correlate with genotype in lymphocytes. Graphs show mean ± SEM.

a)



b)

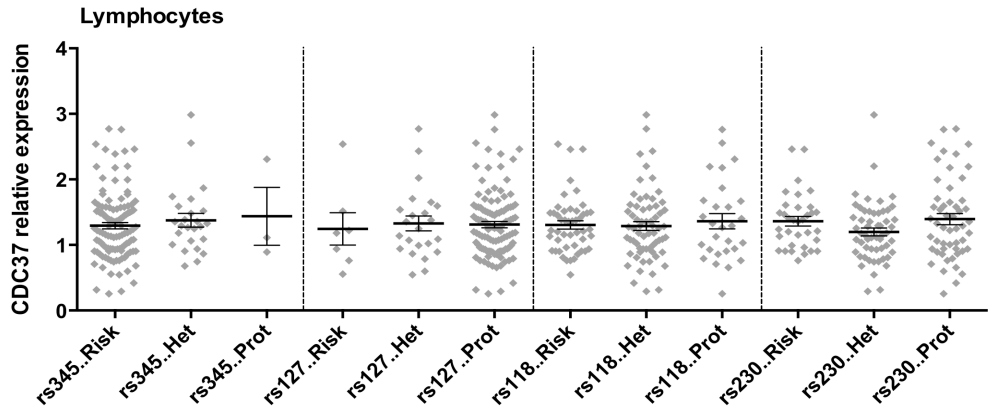


c)

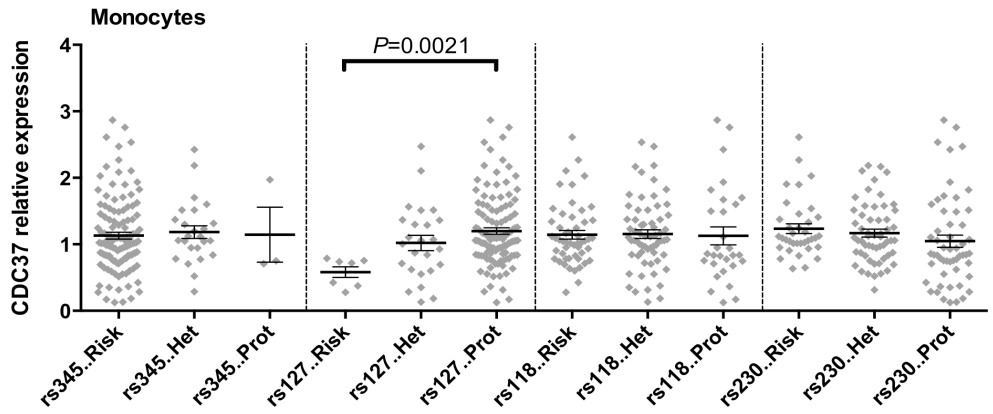
	rs345..Risk	rs345..Het	rs345..Prot	rs127..Risk	rs127..Het	rs127..Prot	rs118..Risk	rs118..Het	rs118..Prot	rs230..Risk	rs230..Het	rs230..Prot
Monocytes	108	26	3	7	24	106	44	65	27	32	56	50
Granulocytes	83	17	3	6	24	73	34	54	18	18	43	42

Figure 6.3.6. Expression of ZGLP1 in immune cell subsets by genotype
ZGLP1 expression was quantified by qPCR relative to expression of the housekeeping gene *UBC* in monocytes (a) and granulocytes (b) from genotyped individuals. The number of individuals per genotype group is shown for each cell subset (c). In all cell subsets, ZGLP1 expression was not found to correlate with genotype. Graphs show mean ± SEM.

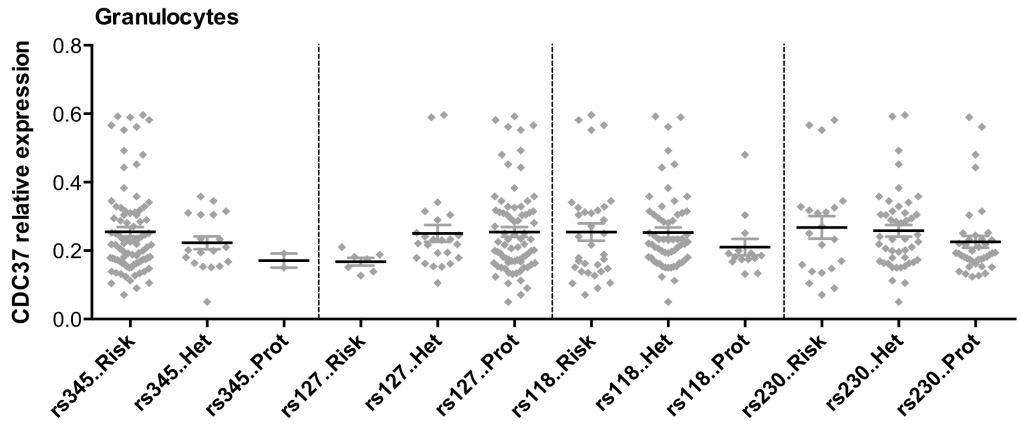
a)



b)



c)

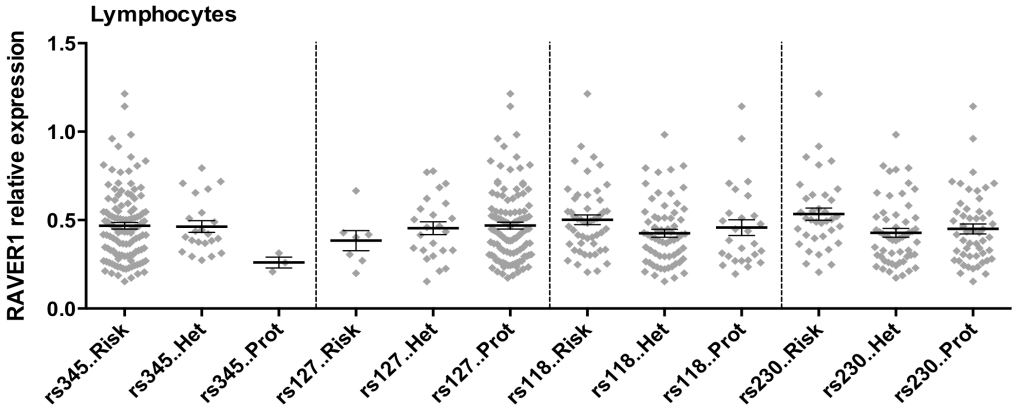


d)

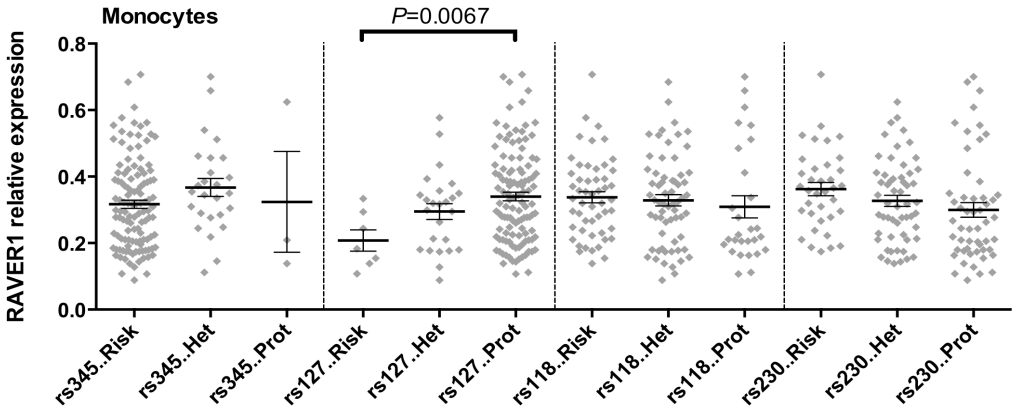
	rs345..Risk	rs345..Het	rs345..Prot	rs127..Risk	rs127..Het	rs127..Prot	rs118..Risk	rs118..Het	rs118..Prot	rs230..Risk	rs230..Het	rs230..Prot
Lymphocytes	115	25	3	7	23	114	48	65	28	34	55	53
Monocytes	120	25	3	7	25	116	52	64	29	54	56	54
Granulocytes	86	19	2	7	24	77	33	60	14	21	46	40

Figure 6.3.7. Expression of CDC37 in immune cell subsets by genotype
CDC37 expression was quantified by qPCR relative to expression of the housekeeping gene *UBC* in lymphocytes (a), monocytes (b) and granulocytes (c) from genotyped individuals. The number of individuals per genotype group is shown for each cell subset (d). In lymphocytes and granulocytes, CDC37 expression was not found to correlate with genotype, however there was significant correlation between CDC37 expression and rs12720356 genotype in monocytes (P=0.0021, r^2=0.06277). Graphs show mean ± SEM.

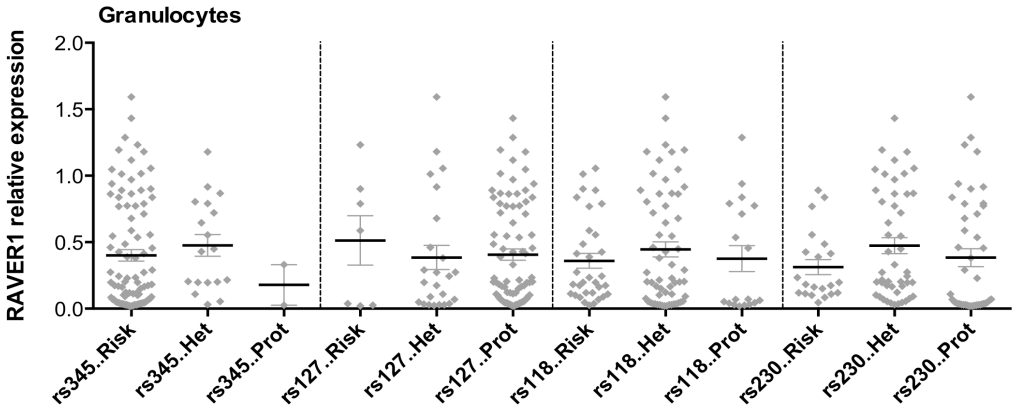
a)



b)



c)

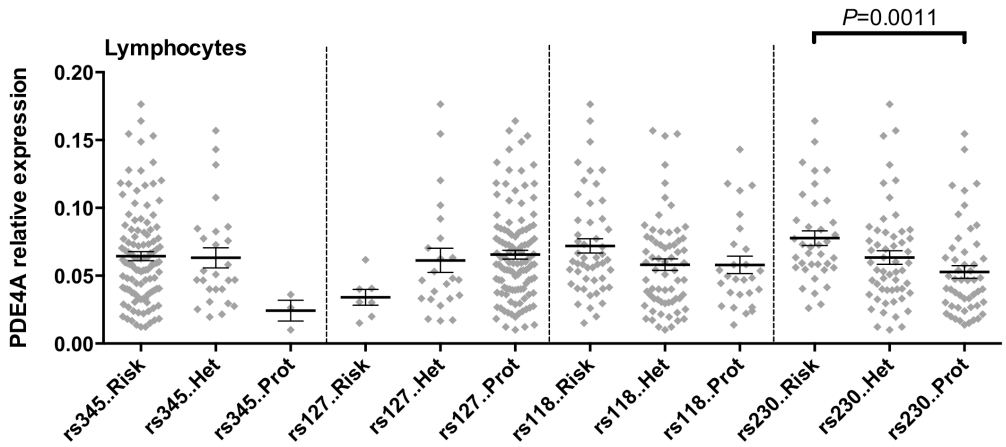


d)

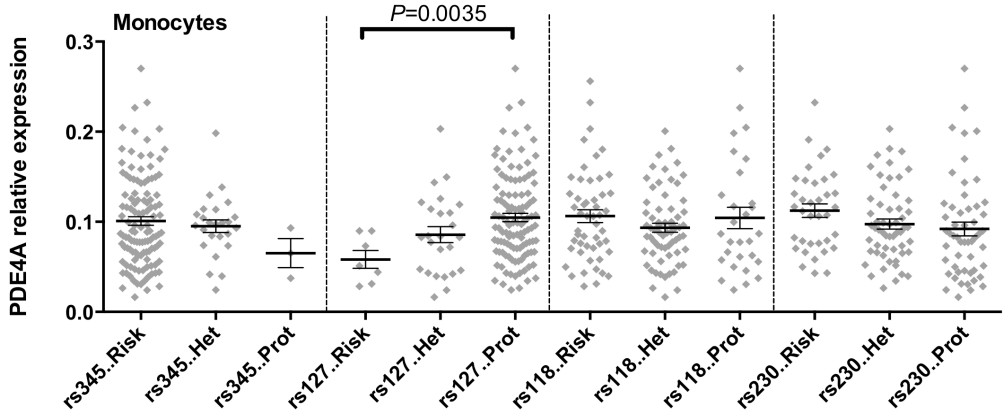
	rs345..Risk	rs345..Het	rs345..Prot	rs127..Risk	rs127..Het	rs127..Prot	rs118..Risk	rs118..Het	rs118..Prot	rs230..Risk	rs230..Het	rs230..Prot
Lymphocytes	119	22	3	7	24	114	51	64	27	36	55	52
Monocytes	118	26	3	7	25	115	52	64	28	36	56	53
Granulocytes	91	18	2	7	25	80	33	60	18	21	47	43

Figure 6.3.8. Expression of RAVER1 in immune cell subsets by genotype
RAVER1 expression was quantified by qPCR relative to expression of the housekeeping gene *UBC* in lymphocytes (a), monocytes (b and e) and granulocytes (c) from genotyped individuals. The number of individuals per genotype group is shown for each cell subset (d). In T cells and granulocytes, RAVER1 expression was not found to correlate with genotype, however in monocytes there was significant correlation between RAVER1 expression and rs12720356 genotype (P=0.0067, r^2=0.04956). Graphs show mean ± SEM.

a)



b)

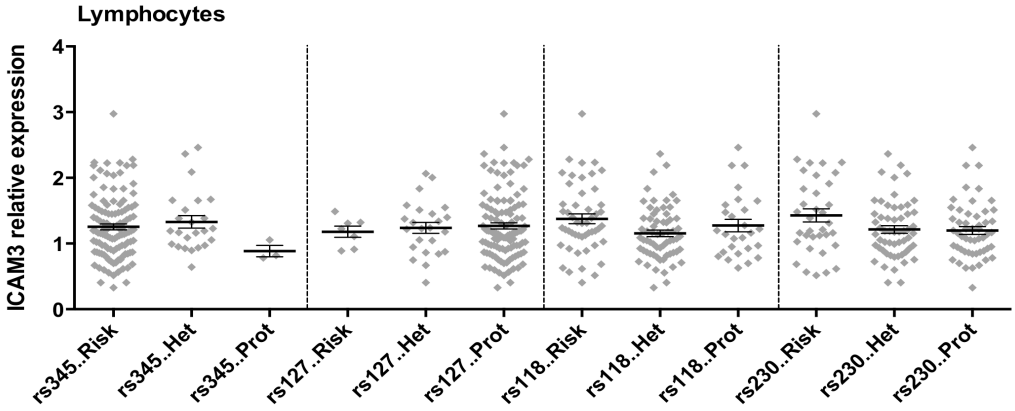


c)

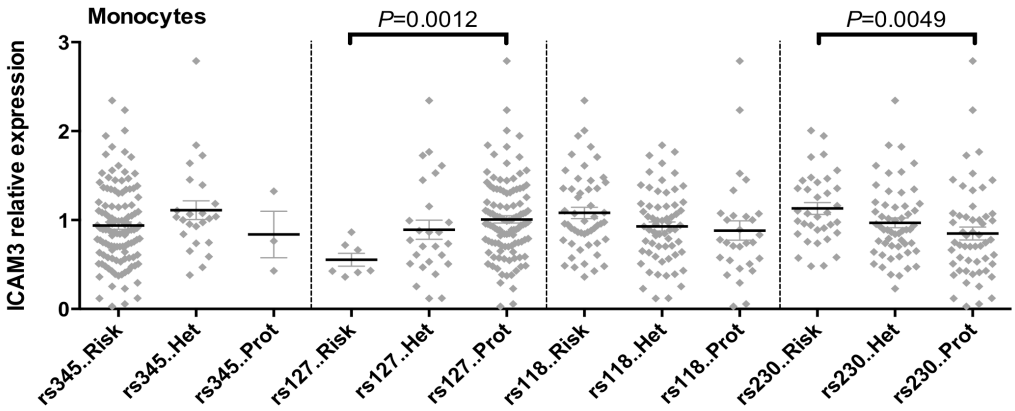
	rs345..Risk	rs345..Het	rs345..Prot	rs127..Risk	rs127..Het	rs127..Prot	rs118..Risk	rs118..Het	rs118..Prot	rs230..Risk	rs230..Het	rs230..Prot
Lymphocytes	114	26	3	7	23	113	50	66	27	37	55	51
Monocytes	117	25	3	7	25	113	51	65	29	36	56	53

Figure 6.3.9. Expression of PDE4A in immune cell subsets by genotype
PDE4A expression was quantified by qPCR relative to expression of the housekeeping gene *UBC* in lymphocytes (a) and monocytes (b) from genotyped individuals. The number of individuals per genotype group is shown for each cell subset (c). PDE4A expression was found to correlate with rs2304256 genotype in lymphocytes (P=0.0011, r^2=0.07258) and with rs12720356 genotype in monocytes (P=0.0035, r^2=0.05801). Graphs show mean ± SEM.

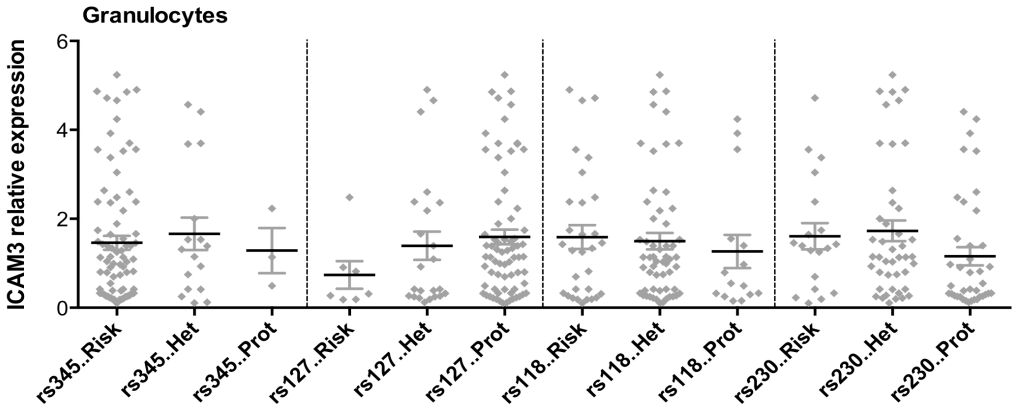
a)



b)



c)

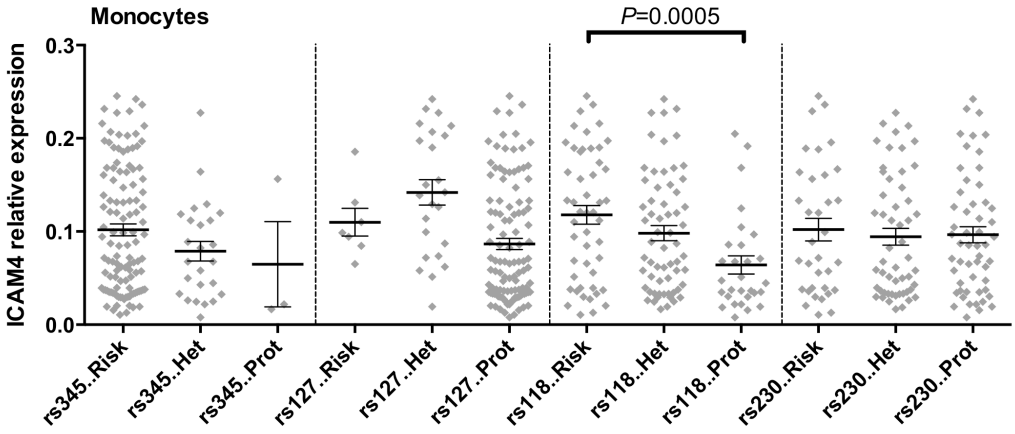


d)

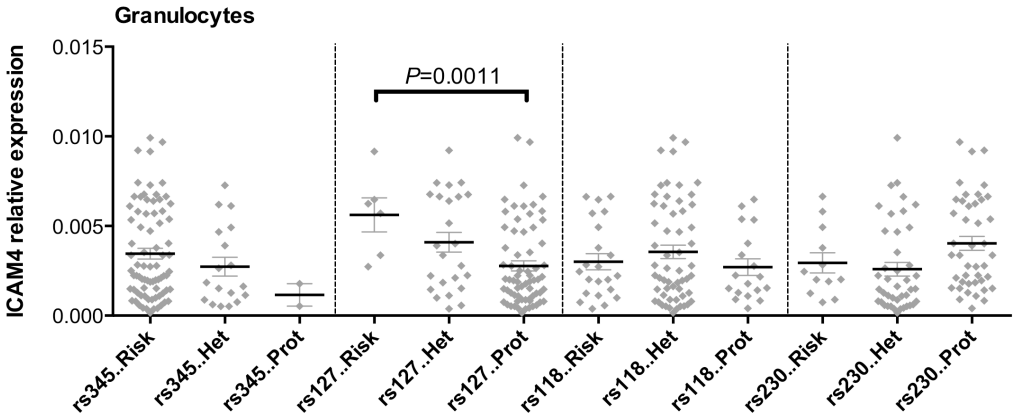
	rs345..Risk	rs345..Het	rs345..Prot	rs127..Risk	rs127..Het	rs127..Prot	rs118..Risk	rs118..Het	rs118..Prot	rs230..Risk	rs230..Het	rs230..Prot
Lymphocytes	117	24	3	7	24	114	50	65	27	35	56	52
Monocytes	119	24	3	7	26	113	51	63	29	35	55	54
Granulocytes	82	17	3	7	23	72	30	57	15	19	44	39

Figure 6.3.10. Expression of ICAM3 in immune cell subsets by genotype
ICAM3 expression was quantified by qPCR relative to expression of the housekeeping gene *UBC* in lymphocytes (a), monocytes (b) and granulocytes (c) from genotyped individuals. The number of individuals per genotype group is shown (d). In lymphocytes and granulocytes, ICAM3 expression did not correlate with genotype. However, in monocytes there was significant correlation between ICAM3 expression and genotype for rs12720356 (P=0.0012, r^2=0.04379) and rs2304256 (P=0.0049, r^2=0.05429). Graphs show mean ± SEM.

a)



b)

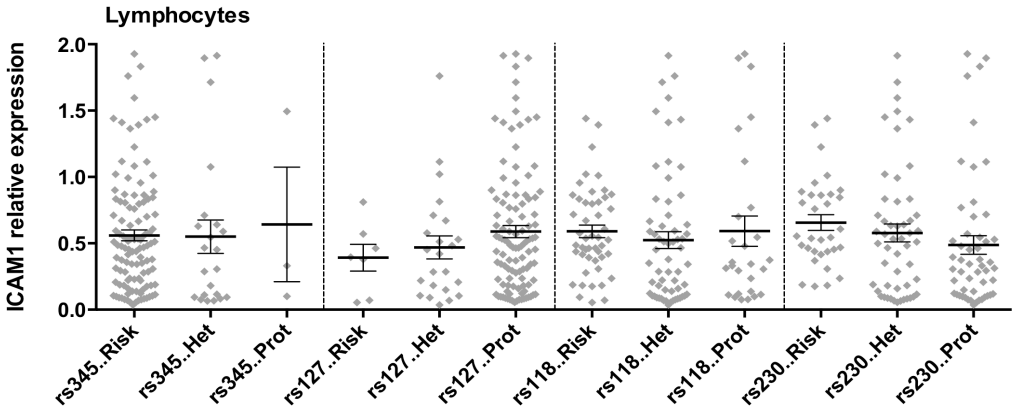


c)

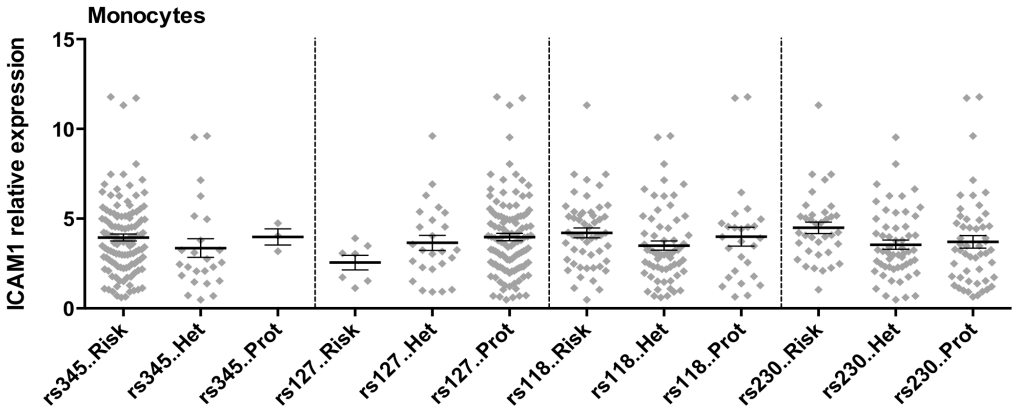
	rs345..Risk	rs345..Het	rs345..Prot	rs127..Risk	rs127..Het	rs127..Prot	rs118..Risk	rs118..Het	rs118..Prot	rs230..Risk	rs230..Het	rs230..Prot
Monocytes	115	25	3	7	24	112	49	62	29	34	53	54
Granulocytes	77	18	2	6	24	67	22	57	18	12	42	43

Figure 6.3.11. Expression of ICAM4 in immune cell subsets by genotype
ICAM4 expression was quantified by qPCR relative to expression of the housekeeping gene *UBC* in monocytes (a) and granulocytes (b) from genotyped individuals. The number of individuals per genotype group is shown for each cell subset (c). ICAM4 expression was found to correlate with rs11879191 genotype in monocytes (P=0.0005, r^2=0.08421) and rs12720356 genotype in granulocytes (P=0.0011, r^2=0.1072). Graphs show mean ± SEM.

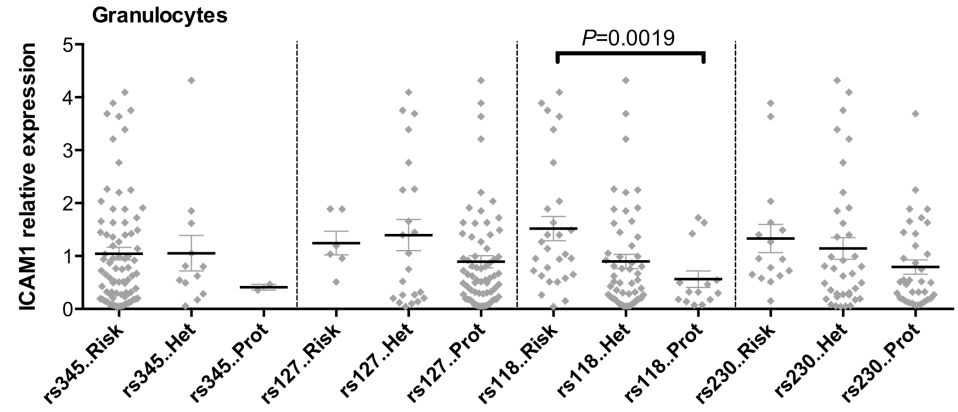
a)



b)



c)



d)

	rs345..Risk	rs345..Het	rs345..Prot	rs127..Risk	rs127..Het	rs127..Prot	rs118..Risk	rs118..Het	rs118..Prot	rs230..Risk	rs230..Het	rs230..Prot
Lymphocytes	112	24	3	7	23	108	48	61	27	32	52	52
Monocytes	119	24	3	7	35	114	52	64	27	36	56	52
Granulocytes	80	12	2	6	22	66	29	52	14	16	37	36

Figure 6.3.12. Expression of ICAM1 in immune cell subsets by genotype
ICAM1 expression was quantified by qPCR relative to expression of the housekeeping gene *UBC* in lymphocytes (a), monocytes (b) and granulocytes (c) from genotyped individuals. The number of individuals per genotype group is shown for each cell subset (d). ICAM1 expression was found to correlate with rs11879191 genotype in granulocytes (P=0.0019, r^2=0.09986). Graphs show mean ± SEM.

6.3.4. Correlating RNA-level expression data with protein levels for ICAM1, ICAM3 and ICAM4

Preliminary analyses were performed (and further investigation is ongoing) in order to investigate whether the genotype-dependent correlations identified at the RNA level corresponded to protein-level differences. The ICAMs were selected for this preliminary protein-level work due to their known immunological relevance, their surface expression on immune cells and the availability of antibodies for flow cytometry. Staining protocols had previously been optimised for PBMCs rather than whole blood, and thus flow cytometric expression analyses were performed for monocytes, as a cell type representative of the myeloid lineage. Having obtained the RNA-level correlations for expression of ICAM4 and ICAM1 in granulocytes (Figures 6.3.11.c and 6.3.12.c, respectively) protein-level data are now being acquired specifically for this cell type.

In order to accurately quantify ICAM expression by flow cytometry, it was important to first determine how to best normalise the data to minimise variability (for details see Chapter 4, section 4.3.1.). For ICAM3, the bead normalisation co-efficient was significantly correlated with antibody staining ($P<0.0001$; Table 6.9.) reflecting 54.5% of the variability in the data, however the isotype control MFI did not significantly correlate with ICAM3 staining ($P=0.2676$; Table 6.9.). This suggested that the calibration beads, but not the isotype control, should be used for data normalisation in this case, similarly to previous work quantifying cell surface expression of CD25 (Dendrou et al., 2009a). In contrast, both the calibration beads and isotype control staining were found to reflect a significant proportion of variation in staining for both ICAM1 (both $P<0.0001$; Table 6.9.) and ICAM4 ($P<0.0001$ and 0.0033, respectively; Table 6.9.) and thus both were used for normalisation. There were no significant effects of age or sex on ICAM expression levels. ICAM4 was undetectable in lymphocytes (data not shown) and ICAM1 and ICAM3 staining was significantly higher on monocytes compared to lymphocytes (both $P<0.0001$). For representative ICAM staining on monocytes see Figure 6.3.13.

Table 6.9. Correlating ICAM antibody staining with the MFI of the isotype control and the normalisation co-efficient calculated using the calibration beads

The MFI of the appropriate isotype control antibodies was plotted against the MFI of the staining for each ICAM, as was the normalisation co-efficient calculated using calibration beads on each experimental day. The r^2 of the observed linear relationships and the extent of significance (*P*-value) are shown for ICAM1, ICAM3 and ICAM4. All MFI values refer to the monocyte population.

		isotype MFI vs ICAM MFI		beads vs ICAM MFI	
	Isotype	r^2	***P*-value**	r^2	***P*-value**
ICAM1	mouse IgG1	0.1287	<0.0001	0.3838	<0.0001
ICAM3	mouse IgG1	0.0119	0.2676	0.5450	<0.0001
ICAM4	polyclonal sheep IgG	0.5118	<0.0001	0.1079	0.0033

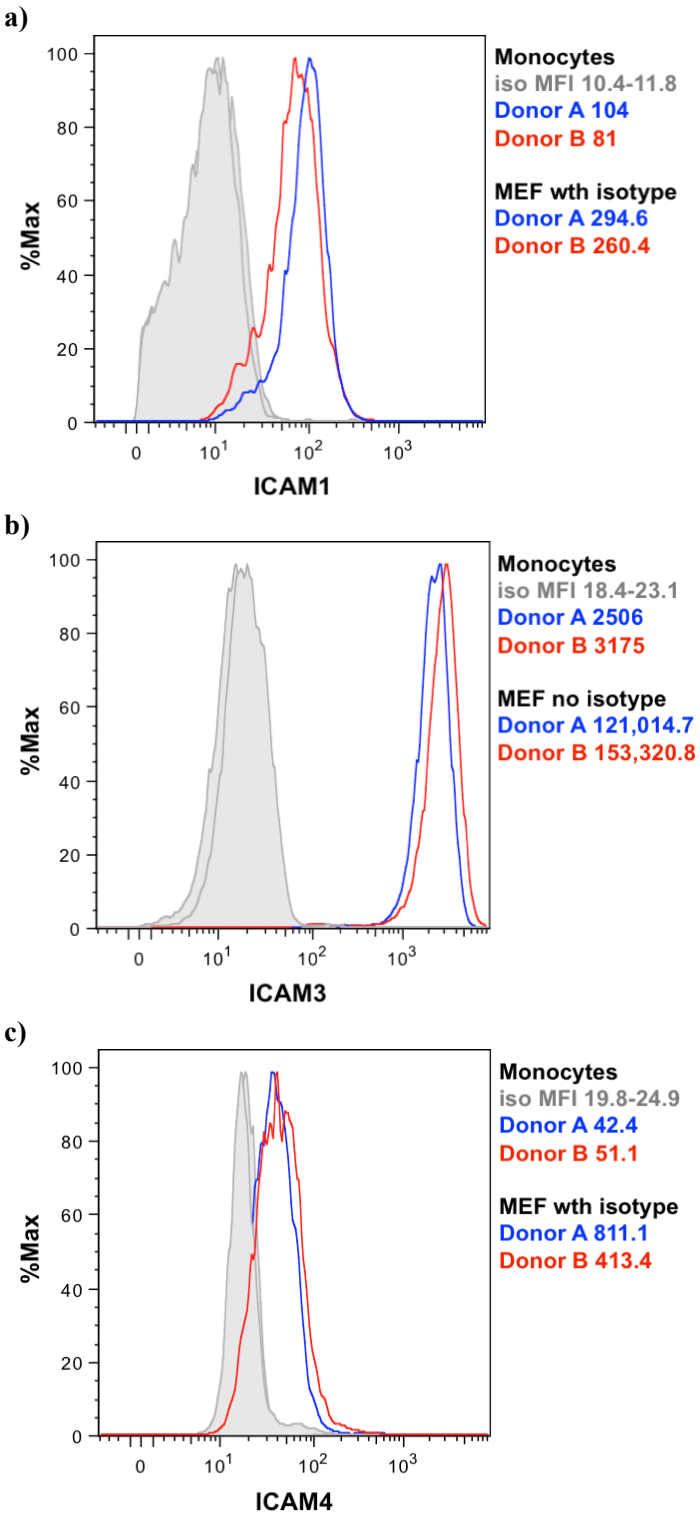


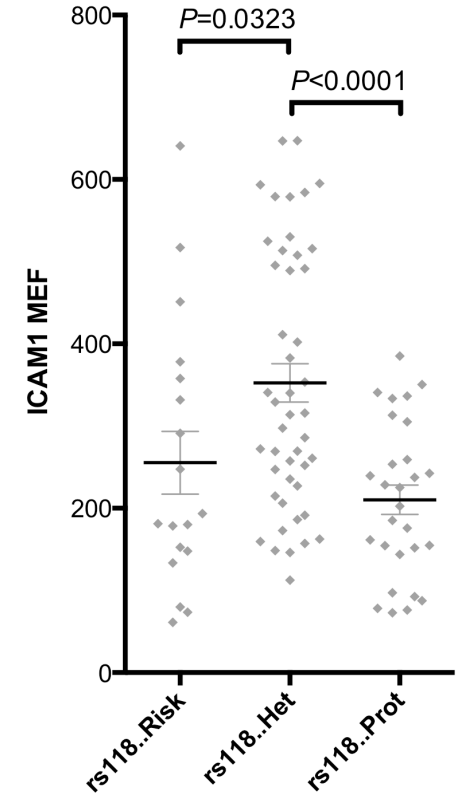
Figure 6.3.13. Representative staining histograms of ICAM expression on CD14+ monocytes
Representative histogram showing ICAM staining on monocytes from two individuals (red and blue lines) relative to the isotype control (filled grey line), with the respective MFI values provided and MEF values representing normalised data as detailed. The y-axis %Max value refers to a normalised value calculated by the FlowJo software to account for the exact number of events captured for each sample.

Expression of ICAM1 in granulocytes was found to correlate with rs11879191 genotype at the RNA level, however in protein-level analyses there was no significant step-wise genotype-dependent correlation observed for ICAM1 expression on monocytes (Figure 6.3.14); in fact, ICAM1 expression was found to be significantly increased in rs11879191 heterozygous individuals compared to both protective and risk allele homozygotes ($P<0.0001$ and $P=0.0323$, respectively; Figure 6.3.14.). The lack of a significant step-wise genotype-dependent correlation for this subset is consistent with the RNA-level data, and an analysis of cell surface ICAM1 levels on granulocytes is now required.

For ICAM3, RNA-level expression was significantly correlated with the genotype of both the rs12720356 and rs2304256 SNPs (Figure 6.3.10.b), with increased expression correlating with protection and risk at each SNP, respectively. At the protein level, whilst no significant correlation was observed for either SNP ($P>0.05$; Figure 6.3.15.; noting the smaller total sample size compared to RNA-level analyses), a weak step-wise trend was observed. A larger sample size will be required in order to determine whether the RNA-level differences translate into a subtle but significant effect on monocyte ICAM3 protein levels.

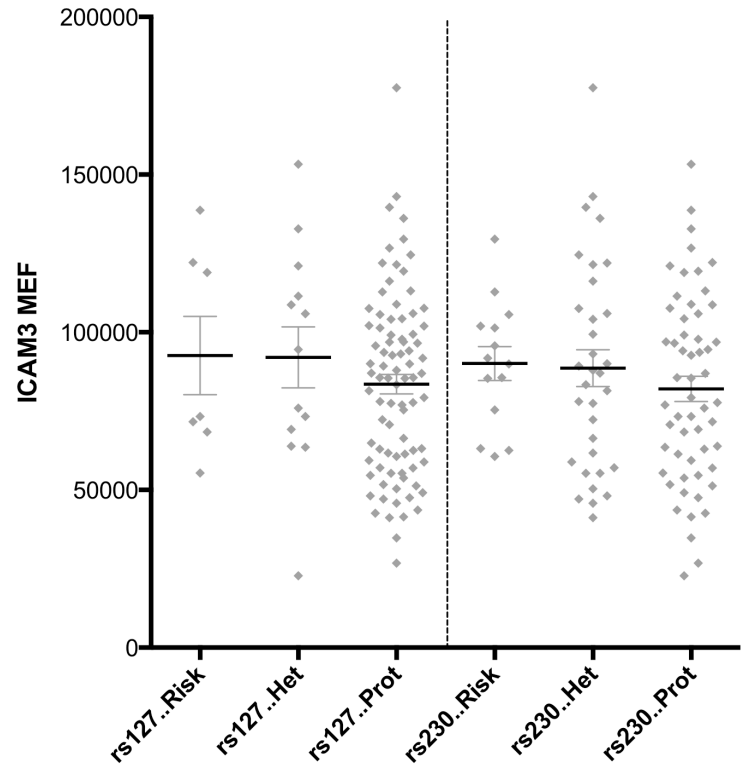
For ICAM4, significant RNA-level correlations were obtained with rs11879191 genotype in monocytes (Figure 6.3.11.b) and rs12720356 genotype in granulocytes (Figure 6.3.11.c), with increased expression correlating with risk in both cases. In the current sample size, a statistically significant step-wise correlation was not observed across the three genotype groups for the rs11879191 SNP and protein-level expression of ICAM4 on monocytes (Figure 6.3.16.). However, monocytes from individuals carrying two copes of the risk allele of the rs11879191 SNP were found to have significantly higher cell surface expression of ICAM4 compared to both heterozygotes and protective homozygotes ($P=0.0087$ and 0.0111, respectively; Figure 6.3.16.). The analysis of a larger sample size will be required to determine whether ICAM4 protein-level expression on monocytes does indeed mirror the RNA-level genotype-dependent effect. Consistent with the absence of a step-wise correlation between RNA-level expression of ICAM4 and rs12720356 genotype in monocytes, there was no such correlation at the protein level, with the heterozygotes having significantly lower cell-surface expression of ICAM4 than both risk and protective homozygotes ($P=0.0004$ and 0.0176,

respectively; Figure 6.3.16). Given the observed significant correlation detected at the RNA level between rs12720356 genotype and ICAM4 levels on granulocytes, data for this subset are now being acquired.



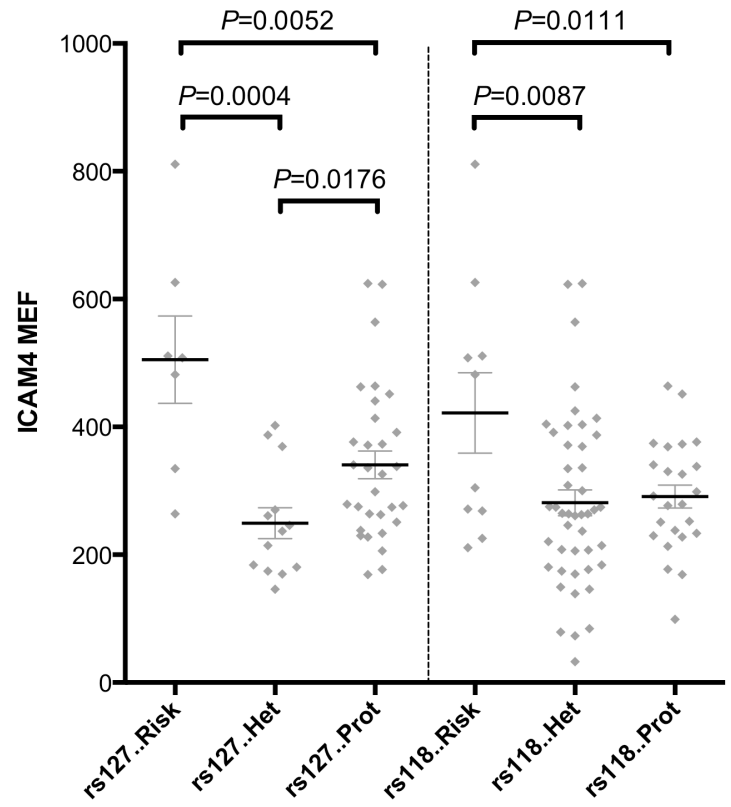
n	18	46	28

Figure 6.3.14. Protein-level expression of ICAM1 on monocytes by genotype
Expression of ICAM1 on monocytes was quantified by flow cytometry. There was no significant step-wise trend by genotype, however there was significantly increased expression of ICAM1 in heterozygous individuals compared to those homozygous for the protective allele of the rs11879191 SNP (P<0.0001). However, ICAM1 levels were also significantly higher in heterozygous individuals than those homozygous for the risk allele (P=0.0323), which is inconsistent with the RNA-level data and also the predicted pattern of allelic risk. Graphs show mean ± SEM and tables show the number of samples (n) per genotype group.



n	7	13	85	14	33	57

Figure 6.3.15. Protein-level expression of ICAM3 on monocytes by genotype
Expression of ICAM3 on monocytes was quantified by flow cytometry. ICAM3 levels were not found to significantly correlate with rs12720356 or rs2304256 genotype in data from the total cohort. Graphs show mean ± SEM and tables show the number of samples (n) per genotype group.



n	7	13	31	24	45	10

Figure 6.3.16. Protein-level expression of ICAM4 on monocytes by rs12720356 and rs11879191 genotype
Expression of ICAM4 on monocytes was quantified by flow cytometry. There were no significant step-wise trends by genotype. For the rs12720356 SNP, there was significantly lower expression of ICAM4 in heterozygous individuals compared to both risk and protective homozygotes (*P*=0.0004 and 0.0176, respectively). For the rs11879191 SNP, there was significantly increased expression of ICAM4 in individuals homozygous for the risk allele of rs11879191 compared to both heterozygous and protective homozygous individuals (*P*=0.0087 and 0.0111, respectively). Graphs show mean ± SEM and tables show the number of samples (n) per genotype group.

6.4 Discussion

The main aim of the work presented in this chapter was to investigate whether expression of genes in the *TYK2* region correlated with the genotype of the four focal disease-associated SNPs. It was first important to characterise expression of the genes of interest across immune cell subsets and upon activation in order to determine the most appropriate samples in which to initially investigate potential effects of SNP genotype. Expression of these genes was detected in resting cell subsets and was predominantly either unaffected or reduced upon stimulation, with the exception of KEAP1 for which expression was upregulated in activated T cells. However, KEAP1 was expressed at relatively high levels in resting T cells (compared to granulocytes) and thus for all qPCR assays subsequent expression analyses by genotype were performed using cDNA samples from unstimulated cells.

6.4.1. Identified correlations between SNP genotype and gene expression

Based on the findings presented in Chapter 5, which identified significant effects of rs34536443 genotype on IFN-β signalling, it was hypothesised that rs34536443 SNP genotype would not correlate with gene expression differences, as indeed was the case. However, based on the lack of a consistent effect on TYK2-mediated signalling and the published evidence of a regulatory role in terms of gene expression, it was hypothesised that correlations between genotype and gene expression would be observed for the rs12720356, rs11879191 and rs2304256 SNPs. The data presented in this Chapter provide further evidence that these three independent association signals influence the regulation of gene expression in the *TYK2* gene region, as summarised in Table 6.10.

Table 6.10. Summary of correlations between SNP genotype and gene expression
A summary of the correlations between SNP genotype and gene expression, indicating the cell subset in which the effect was identified and also whether increased expression (↑) was associated with disease risk or protection. A dash indicates the lack of a significant correlation. As has been the convention, disease association for the rs12720356 SNP refers to IBD, with PS being associated in the opposite direction (Chapter 1, Table 1.1.). M = monocytes, T = T cells and G = granulocytes.

	rs127..	rs118..	rs230..
ICAM4	↑G = risk	↑M = risk	-
ICAM1	-	↑G = risk	-
ICAM3	↑M = prot	-	↑M = risk
CDC37	↑M = prot	-	-
PDE4A	↑M = prot	-	↑L = risk
RAVER1	↑M = prot	-	-

Whilst both the rs11879191 and rs12720356 SNPs are associated with IBD, only the former is also associated with AS, and only the latter with PS (albeit in the opposite direction to IBD), leading to the hypothesis that these variants would exert distinct functional effects mediated either through regulation of the expression of different genes, and/or the differential regulation of gene expression in different cell subsets. As the rs2304256 SNP is not associated with IBD, AS or PS it is also likely to have distinct functional effects that are T1D-specific. Although the identified correlations between SNP and gene expression do not provide definitive evidence of a causal functional effect, considering the function of SNP-correlated genes may suggest potential disease-relevant pathways for further study at the level of protein expression, and also in terms of cellular function. Considering the identified correlations, expression of ICAM1 and ICAM3 is relatively restricted to immune cells compared to the more ubiquitously expressed CDC37 and RAVER1, and ICAM4. Thus, genotype-dependent differences in expression of ICAM1 and ICAM3 may be more relevant in the context of autoimmune disease than effects on other genes in the region, given their prominent immunological function.

The presented work identified a correlation between ICAM4 expression on monocytes and rs11879191 genotype, with increased ICAM4 correlating with risk for IBD and AS. A correlation was also observed between ICAM4 expression in granulocytes and rs12720356 SNP genotype, with increased expression correlating with IBD risk but protection against PS. As ICAM4 is mainly expressed by erythrocytes (Wu et al., 2009) and as there are no published data regarding its function

in monocytes and granulocytes, a putative disease-relevant mechanism underpinning the observed correlations between SNP genotype and ICAM4 expression remains to be identified.

In contrast to ICAM4, there is a greater understanding of the function of the other genes for which genotype-dependent correlations were identified in the immune cell subsets studied, however any specific role of these gene products in autoimmune disease is not well characterised. Increased ICAM1 expression in granulocytes correlated with risk for IBD and AS at the rs11879191 SNP, and increased ICAM1 expression has been shown to promote adhesion and migration of neutrophils across vascular endothelium (Yang et al., 2005b). As neutrophil migration and activation are features of pathology in both IBD and AS (Alzoghaibi, 2005; Yazici et al., 2004), this may be a potential mechanism linking the association of the rs11879191 SNP genotype with disease and further investigation is required to obtain further evidence for an effect of this SNP on ICAM1 protein levels.

Increased expression of ICAM3 in monocytes correlated with the risk allele of the rs2304256 SNP and thus with risk for T1D, consistent with previously published eQTL data (Fairfax et al., 2012; Wallace et al., 2012). RNA-level expression of ICAM3 in monocytes was also found to correlate with rs12720356 genotype, with higher expression correlating with protection against IBD and risk for PS. ICAM3 expression by monocytes plays a role promoting leukocyte recruitment and entry into tissues (Estecha et al., 2012) as well as contributing to pro-inflammatory responses by monocytes with respect to the synthesis and secretion of chemokines and prostaglandins (Kessel et al., 1998; Pontsler et al., 2002). However, the putative mechanism by which altered ICAM3 levels may be relevant in the context of autoimmune diseases is unknown. It is also unclear whether increased ICAM3 may have opposite consequences on disease susceptibility for IBD compared to PS, or whether the opposing direction of association of the rs12720356 SNP with these diseases is due to additional distinct effects. Future work will include increasing the sample size for investigation of these genotype-dependent correlations at the RNA level and also for protein-level analyses.

In the presented work, increased expression of CDC37, RAVER1 and PDE4A in monocytes correlated with the major allele of the rs12720356 SNP and thus correlated with protection against

IBD but risk for PS. PDE4A expression in lymphocytes was found to correlate with rs2304256 genotype, with increased expression correlating with T1D risk. It is unclear how altered expression of these genes can correlate with both risk and protection in different autoimmune diseases, or indeed whether all of the observed genotype-dependent gene expression differences are truly disease-relevant. This is particularly the case when considering that only the general functions of these genes have been described in fundamental cellular processes, and thus the specific relevance of the expression of these genes in the context of autoimmune disease remains to be elucidated. CDC37 is a co-chaperone that promotes interaction of kinases with heat-shock protein 90 (Hsp90) to stabilise and promote their activity (Stepanova et al., 1996). RAVER1 is a modulator of alternative splicing (Gromak et al., 2003) and also binds to cytoskeletal components, which may relate to a role in distributing mRNA to sub-cellular compartments (Lee et al., 2009). PDE4A is a cAMP-specific phosphodiesterase isoform that is ubiquitously expressed and along with other PDE4 isoforms regulates multiple cellular signalling pathways. Notably, PDE4 inhibitors do have anti-inflammatory effects and are in clinical trials for multiple autoimmune diseases (Kumar et al., 2013), however they are also associated with diverse side effects.

6.4.2. Published eQTLs not replicated in the presented work

As detailed in Table 6.3, previous studies have reported correlations between rs12720356 genotype and expression of PDE4A and FDX1L that were not detected in the presented data. This is likely to be related to the use of LCLs in the published work (Brown et al., 2013). The regulation of gene expression in these EBV-transformed cell lines is known to be subject to artificial effects arising from their generation and culturing conditions (Choy et al., 2008; Plagnol et al., 2008), and thus data from LCLs may not be representative of the more physiological context of the primary immune cell subsets investigated in this Chapter. Similarly, a correlation between rs11879191 genotype and CDC37 expression in LCLs has previously been reported (Veyrieras et al., 2008) but no such correlation was observed in the presented work.

In primary human monocytes, expression of ICAM4 has previously been correlated with rs12720356 genotype (Fairfax et al., 2012) however no such correlation was detected in the presented data, which did however include samples from a greater number of individuals homozygous for the minor allele (here n=7 compared to n=2 in previous work (Fairfax et al., 2012)). In line with the reported eQTL, there was a significant reduction of ICAM4 expression in monocytes from individuals homozygous for the common protective allele compared to heterozygous individuals (P=0.0002), but there was no significant step-wise trend across all three genotype groups (as would be predicted by the pattern of the association of this SNP with disease).

6.4.3. Conclusion and future work

In summary, the presented RNA-level data support the hypothesis that the different disease-associated SNPs in the *TYK2* region do not all influence *TYK2* and that they are each likely to exert distinct functional effects, in keeping with their complex association pattern across multiple autoimmune conditions. Based on these data and those from previous Chapters, the main effect of the rs34536443 SNP is on TYK2 function. Thus, for diseases that are not associated with the rs34536443 SNP, it was predicted that the disease-associated SNPs in the *TYK2* region would likely affect genes other than *TYK2*. Consistent with this, there were no genotype-dependent correlations observed for RNA-level expression of TYK2, whereas the rs12720356, rs11879191 and rs2304256 SNPs were each correlated with expression of other genes in the region.

Further investigation is required in order to determine the extent to which the identified expression correlations are truly disease-relevant. On-going work is being carried out to fully confirm the observed correlations by increasing the sample size for both RNA- and protein-level analyses. The acquisition of a larger cohort will allow for the analysis of gene expression not only by individual SNPs but also by haplotypes defined by collective consideration of the four focal SNPs in the *TYK2* region. Analyses will also be performed to determine whether the disease-associated SNPs best explain the variation in gene expression; the identification of non-associated SNPs with a larger impact on gene expression may suggest that correlations with the disease-associated variants are not

in fact disease-relevant, which will help to further dissect the genetically-driven functional effects that *do* influence disease development. For example, in the presented data, RNA-level expression of ICAM3 in monocytes was found to correlate with rs2304256 genotype, which is consistent with previous data (Fairfax et al., 2012). However the published study identifies the rs2304240 SNP to be more significantly correlated with ICAM3 expression than the rs2304256 SNP (P=4.29x10^{-23} compared to P=0.000116, respectively) suggesting that the former is more likely to be driving this effect. As yet, the rs2304240 SNP has not been reported to be associated with autoimmune disease, and thus altered expression of ICAM3 in monocytes may not be disease-relevant.

In addition, further gene expression data will be obtained using other methodologies (such as microarrays and/or transcriptome sequencing technologies) to further validate the presented qPCR-based findings and to explore the potential for long-range effects of the SNPs on gene expression regulation. A recently published study reports a correlation between rs12720356 genotype and the gene encoding cyclin-dependent kinase inhibitor 2D (*CDKN2D;* (Westra et al., 2013)), which is located outside of the disease-associated *TYK2* gene region defined through GWAS (as shown schematically in Figure 6.1.). A *trans*-eQTL was also reported for this SNP, whereby rs12720356 genotype was found to correlate with expression of the gene encoding interferon alpha-inducible protein 27 (*IFI27*) located on chromosome 14 (Westra et al., 2013). The study by Westra *et al.* used whole blood samples from genotyped individuals, whereas previous studies have demonstrated the existence of distinct cell subset-specific eQTLs (Fairfax et al., 2012). Thus, future work will involve performing further expression analyses in immune cell sub-populations to enhance the likelihood of detecting such subset-specific differences that may be overlooked in studies of whole blood samples.

Collectively considering the presented data and those from previous eQTL studies, there is evidence suggesting a regulatory function for three of the four disease-associated variants in the *TYK2* region (or for variants in high LD with the reported SNPs). Future work will involve using chromatin conformation capture technologies to detect the presence of a physical interaction between associated variants and the promoter of genes whose expression correlates with variant genotype (van de Werken et al., 2012). This will provide additional confirmation and also a molecular basis for the genotype-dependent correlations identified at the RNA and protein level. Finally, the functional

implications of confirmed SNP-dependent differences in gene expression will be assessed at the cellular and systemic level, in order to delineate how the genotype-dependent pathways influenced by disease-associated SNPs in the *TYK2* region influence autoimmunity.

7. General conclusions

The overall aim of the work presented in this thesis was to investigate the functional effects of four distinct autoimmune disease associations in the *TYK2* gene region. As detailed in Chapter 1 (Table 1.1.), each of the reported focal SNPs is associated with multiple conditions and most diseases are associated with just one SNP, with the exception of IBD and PS (each of which are associated with two SNPs). Based on this complex pattern of disease association, it was hypothesised that each of the independent genetic associations in the *TYK2* region would be likely to have distinct functional consequences that would not be restricted to an effect on *TYK2* itself, despite *TYK2* being cited as the candidate gene for all of the identified association signals. Indeed, in the presented work, a substantial effect on TYK2 was only demonstrated for a single disease-associated SNP.

Three of the four reported disease-associated SNPs in the *TYK2* gene region are nonsynonymous and influence the amino acid sequence of known functional domains in the TYK2 protein. In the presented work, the minor allele of the rs34536443 nsSNP, that results in the amino acid substitution of a highly conserved residue in the TYK2 kinase domain, was correlated with reduced IFN-β signalling in both cell lines (Chapter 3) and primary human immune cells (Chapter 5). There was no evidence to suggest an effect of the rs2304256 nsSNP on TYK2 function in cell lines (Chapter 3), consistent with the respective amino acid substitution not being located at a residue that is conserved across the JAK family and across species. Although there was some indication that the rs12720356 nsSNP-dependent substitution in the pseudokinase domain influences TYK2 activity using the transfection system (Chapter 3), there was little support for such an effect in genotype-dependent analyses of cytokine signalling in primary human cells (Chapter 5). Thus, for the three nsSNPs under investigation there was only strong evidence that rs34536443 genotype has consequences on TYK2-mediated signalling. Importantly, the effects of this SNP were found to be restricted to the IFN-β signalling, as there was no effect of rs34536443 genotype on STAT phosphorylation induced by IL-6, IL-10 or IL-13 stimulation. This has implications with respect to the genetically-driven involvement of specific cytokine signalling pathways in different autoimmune

conditions. Additionally, these findings also indicate a previously unappreciated specificity in the function of TYK2, which not only represents progress in our general understanding of immune signalling but is also of particular importance in terms of the therapeutic targeting of TYK2 by selective compounds for the purpose of immunomodulation.

In order to further examine the functional consequences of disease-associated variation in the *TYK2* region, RNA-level expression analyses were performed by genotype in primary human immune cell subsets for all four focal SNPs. As hypothesised, no correlation was observed for the rs34536443 SNP, but a number of significant correlations were observed between rs2304256, rs12720356 and rs11879191 genotype and the RNA-level expression of several genes in the *TYK2* region. Notably, expression of TYK2 was not found to correlate with SNP genotype for any of these variants. This is consistent with hypothesis that the rs34536443 SNP affects TYK2 whereas the other disease-associated SNPs in the region influence other genes, given that they are not associated with the same conditions as the rs34536443 SNP (with the sole exception of PS which is associated with both the rs34536443 and rs12720356 SNPs). Further work is required to validate the observed RNA-level correlations at the protein level and to ascertain their relevance to autoimmune disease mechanisms.

The concept that a careful consideration of the genetic association pattern of a specific genomic region across diseases can inform the degree of sharing of pathophysiological pathways between autoimmune diseases is of particular importance. Several meta-analyses have estimated the extent of genetic overlap between autoimmune diseases (Parkes et al., 2013), however assigning overlap based loosely on the existence of a genetic association in a certain region is likely to over-estimate the extent to which disease mechanisms are truly shared. This is especially the case for SNPs associated with diseases in the opposite direction, for example, the rs12720356 SNP is associated with both IBD and PS, but the minor allele confers risk for the former condition and protection against the latter (Anderson et al., 2011; Franke et al., 2010; Strange et al., 2010; Tsoi et al., 2012). This potentially suggests that the same genotype-dependent pathway could have different consequences for two different autoimmune conditions. Moreover, such bioinformatic pathway analyses are based on assumptions regarding the likely candidate gene in a given associated region

and do not typically rely on functional data. Thus, whilst this approach has a role may be informative to a certain extent in identifying putative common disease-relevant pathways from the wealth of recently identified disease-associated genetic variation, its limitations must be recognised, particularly as it has been suggested as a useful approach for directing drug repositioning strategies across autoimmune diseases (Sanseau et al., 2012). A recent proof-of-principle study has indicated that the pathways affected by common autoimmune disease-associated genetic variants can be of clinical relevance with respect to therapeutic targeting, despite their small individual effects on overall disease risk. Genetic risk for MS at the rs1800693 SNP in the *TNFRSF1A* gene has been found to drive an endogenous TNF antagonist, in keeping with the clinical experience with TNF antagonistic drugs that have been found to exacerbate the disease (Gregory et al., 2012). Conversely, the rs1800693 SNP is also associated with AS (Cortes et al., 2013; Evans et al., 2011), but in the opposite direction to MS, and TNF antagonist administration in AS has been shown to be efficacious (Braun et al., 2011; van der Heijde et al., 2011). This example demonstrates how careful consideration of a genetic association across conditions could have significant implications that could not be predicted without accounting for both the directionality and the functional effects of a disease-associated variant.

In the case of TYK2, selective inhibitors are under development (Liang et al., 2013a; Liang et al., 2013b), and promising results having been obtained from the use of these compounds in various pre-clinical models of autoimmune disease (Norman, 2012; Sareum, 2013). Such compounds would mimic (and magnify) the protective effect of reduced TYK2-mediated signalling that occurs in carriers of the minor allele of the rs34536443 SNP, and could therefore be considered for testing in conditions that this SNP is associated with. Whether such inhibitors will be useful if repositioned for the treatment of diseases that are not associated with the rs34536443 SNP, such as IBD, T1D and AS, remains to be determined. Potential therapeutic targets other than TYK2 may be identified for these diseases pending further investigation of the other associated SNPs (rs12720356, rs11879191 and rs2304256) in the region.

In summary, although *TYK2* is an attractive candidate gene for the four independent autoimmune disease associations in the gene region due to its immunological relevance, the work

presented in this thesis supports the paradigm that for genomic regions in which multiple disease-associated signals are detected, a thorough cross-comparative genetic and functional investigation of individual SNPs is necessary in order to delineate the genetically-driven pathways that are relevant to different diseases. The functional dissection of the effects of the associated SNPs in the region provides evidence that the *TYK2* gene may be causal for only one of these association signals, with the other SNPs correlating with expression differences of genes other than *TYK2*. Hence, future work will aim to further ascertain the shared and distinct mechanisms underpinning the multiple autoimmune disease associations in the *TYK2* gene region.

8. References

Abecasis, G.R., Altshuler, D., Auton, A., Brooks, L.D., Durbin, R.M., Gibbs, R.A., Hurles, M.E., and McVean, G.A. (2010). A map of human genome variation from population-scale sequencing. Nature *467*, 1061-1073.

Adam, L., Bandyopadhyay, D., and Kumar, R. (2000). Interferon-alpha signaling promotes nucleus-to-cytoplasmic redistribution of p95Vav, and formation of a multisubunit complex involving Vav, Ku80, and Tyk2. Biochem Biophys Res Commun *267*, 692-696.

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.

Ahern, P.P., Schiering, C., Buonocore, S., McGeachy, M.J., Cua, D.J., Maloy, K.J., and Powrie, F. (2010). Interleukin-23 drives intestinal inflammation through direct activity on T cells. Immunity *33*, 279-288.

Ahmed, C.M., Noon-Song, E.N., Kemppainen, K., Pascalli, M.P., and Johnson, H.M. (2013). Type I IFN receptor controls activated TYK2 in the nucleus: implications for EAE therapy. J Neuroimmunol *254*, 101-109.

Aizu, K., Li, W., Yajima, T., Arai, T., Shimoda, K., Nimura, Y., and Yoshikai, Y. (2006). An important role of Tyk2 in APC function of dendritic cells for priming CD8+ T cells producing IFN-gamma. Eur J Immunol *36*, 3060-3070.

Al-Lazikani, B., Sheinerman, F.B., and Honig, B. (2001). Combining multiple structure and sequence alignments to improve sequence detection and alignment: application to the SH2 domains of Janus kinases. Proc Natl Acad Sci U S A *98*, 14796-14801.

Albrechtsen, A., Grarup, N., Li, Y., Sparso, T., Tian, G., Cao, H., Jiang, T., Kim, S.Y., Korneliussen, T., Li, Q.*, et al.* (2013). Exome sequencing-driven discovery of coding polymorphisms associated with common metabolic phenotypes. Diabetologia *56*, 298-310.

Altshuler, D.M., Gibbs, R.A., Peltonen, L., Dermitzakis, E., Schaffner, S.F., Yu, F., Bonnen, P.E., de Bakker, P.I., Deloukas, P., Gabriel, S.B.*, et al.* (2010). Integrating common and rare genetic variation in diverse human populations. Nature *467*, 52-58.

Alzoghaibi, M.A. (2005). Neutrophil expression and infiltration into Crohn's intestine. Saudi J Gastroenterol *11*, 63-72.

Anderson, A.C., Nicholson, L.B., Legge, K.L., Turchin, V., Zaghouani, H., and Kuchroo, V.K. (2000). High frequency of autoreactive myelin proteolipid protein-specific T cells in the periphery of naive mice: mechanisms of selection of the self-reactive repertoire. J Exp Med *191*, 761-770.

Anderson, C.A., Boucher, G., Lees, C.W., Franke, A., D'Amato, M., Taylor, K.D., Lee, J.C., Goyette, P., Imielinski, M., Latiano, A.*, et al.* (2011). Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. Nat Genet *43*, 246-252.

Anderson, M.S., Venanzi, E.S., Klein, L., Chen, Z., Berzins, S.P., Turley, S.J., von Boehmer, H., Bronson, R., Dierich, A., Benoist, C., and Mathis, D. (2002). Projection of an immunological self shadow within the thymus by the aire protein. Science *298*, 1395-1401.

ANZgene, A.a.N.Z.M.S.G.C. (2009). Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. Nat Genet *41*, 824-828.

Araki, M., Kondo, T., Gumperz, J.E., Brenner, M.B., Miyake, S., and Yamamura, T. (2003). Th2 bias of CD4+ NKT cells derived from multiple sclerosis in remission. Int Immunol *15*, 279-288.

Arun, T., Tomassini, V., Sbardella, E., de Ruiter, M.B., Matthews, L., Leite, M.I., Gelineau-Morel, R., Cavey, A., Vergo, S., Craner, M.*, et al.* (2013). Targeting ASIC1 in primary progressive multiple sclerosis: evidence of neuroprotection with amiloride. Brain *136*, 106-115.

Ascherio, A., and Munger, K.L. (2007a). Environmental risk factors for multiple sclerosis. Part I: the role of infection. Ann Neurol *61*, 288-299.

Ascherio, A., and Munger, K.L. (2007b). Environmental risk factors for multiple sclerosis. Part II: Noninfectious factors. Ann Neurol *61*, 504-513.

Ascherio, A., Munger, K.L., and Simon, K.C. (2010). Vitamin D and multiple sclerosis. Lancet Neurol *9*, 599-612.

Axtell, R.C., de Jong, B.A., Boniface, K., van der Voort, L.F., Bhat, R., De Sarno, P., Naves, R., Han, M., Zhong, F., Castellanos, J.G.*, et al.* (2010). T helper type 1 and 17 cells determine efficacy of interferon-beta in multiple sclerosis and experimental encephalomyelitis. Nat Med *16*, 406-412.

Axtell, R.C., Raman, C., and Steinman, L. (2013). Type I interferons: beneficial in Th1 and detrimental in Th17 autoimmunity. Clin Rev Allergy Immunol *44*, 114-120.

Babbe, H., Roers, A., Waisman, A., Lassmann, H., Goebels, N., Hohlfeld, R., Friese, M., Schroder, R., Deckert, M., Schmidt, S.*, et al.* (2000). Clonal expansions of CD8(+) T cells dominate the T cell infiltrate in active multiple sclerosis lesions as shown by micromanipulation and single cell polymerase chain reaction. J Exp Med *192*, 393-404.

Babon, J.J., Kershaw, N.J., Murphy, J.M., Varghese, L.N., Laktyushin, A., Young, S.N., Lucet, I.S., Norton, R.S., and Nicola, N.A. (2012). Suppression of cytokine signaling by SOCS3: characterization of the mode of inhibition and the basis of its specificity. Immunity *36*, 239-250.

Baecher-Allan, C., Wolf, E., and Hafler, D.A. (2006). MHC class II expression identifies functionally distinct human regulatory T cells. J Immunol *176*, 4622-4631.

Baechler, E.C., Batliwalla, F.M., Karypis, G., Gaffney, P.M., Ortmann, W.A., Espe, K.J., Shark, K.B., Grande, W.J., Hughes, K.M., Kapur, V.*, et al.* (2003). Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. Proc Natl Acad Sci U S A *100*, 2610-2615.

Ban, M., Goris, A., Lorentzen, A.R., Baker, A., Mihalova, T., Ingram, G., Booth, D.R., Heard, R.N., Stewart, G.J., Bogaert, E.*, et al.* (2009). Replication analysis identifies TYK2 as a multiple sclerosis susceptibility factor. Eur J Hum Genet *17*, 1309-1313.

Bar-Or, A., Fawaz, L., Fan, B., Darlington, P.J., Rieger, A., Ghorayeb, C., Calabresi, P.A., Waubant, E., Hauser, S.L., Zhang, J., and Smith, C.H. (2010). Abnormal B-cell cytokine responses a trigger of T-cell-mediated disease in MS? Ann Neurol *67*, 452-461.

Baranzini, S.E., Mudge, J., van Velkinburgh, J.C., Khankhanian, P., Khrebtukova, I., Miller, N.A., Zhang, L., Farmer, A.D., Bell, C.J., Kim, R.W.*, et al.* (2010). Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis. Nature *464*, 1351-1356.

Baranzini, S.E., Wang, J., Gibson, R.A., Galwey, N., Naegelin, Y., Barkhof, F., Radue, E.W., Lindberg, R.L., Uitdehaag, B.M., Johnson, M.R.*, et al.* (2009). Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. Hum Mol Genet *18*, 767-778.

Baughman, E.J., Mendoza, J.P., Ortega, S.B., Ayers, C.L., Greenberg, B.M., Frohman, E.M., and Karandikar, N.J. (2011). Neuroantigen-specific CD8+ regulatory T-cell function is deficient during acute exacerbation of multiple sclerosis. J Autoimmun *36*, 115-124.

Bennett, C.L., Christie, J., Ramsdell, F., Brunkow, M.E., Ferguson, P.J., Whitesell, L., Kelly, T.E., Saulsbury, F.T., Chance, P.F., and Ochs, H.D. (2001). The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. Nat Genet *27*, 20-21.

Berer, K., Mues, M., Koutrolos, M., Rasbi, Z.A., Boziki, M., Johner, C., Wekerle, H., and Krishnamoorthy, G. (2011). Commensal microbiota and myelin autoantigen cooperate to trigger autoimmune demyelination. Nature *479*, 538-541.

Bernstein, B.E., Birney, E., Dunham, I., Green, E.D., Gunter, C., and Snyder, M. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57-74.

Bieber, A.J., Suwansrinon, K., Kerkvliet, J., Zhang, W., Pease, L.R., and Rodriguez, M. (2010). Allelic variation in the Tyk2 and EGF genes as potential genetic determinants of CNS repair. Proc Natl Acad Sci U S A *107*, 792-797.

Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigo, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E.*, et al.* (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799-816.

Boggon, T.J., Li, Y., Manley, P.W., and Eck, M.J. (2005). Crystal structure of the Jak3 kinase domain in complex with a staurosporine analog. Blood *106*, 996-1002.

Braun, J., van den Berg, R., Baraliakos, X., Boehm, H., Burgos-Vargas, R., Collantes-Estevez, E., Dagfinrud, H., Dijkmans, B., Dougados, M., Emery, P.*, et al.* (2011). 2010 update of the ASAS/EULAR recommendations for the management of ankylosing spondylitis. Ann Rheum Dis *70*, 896-904.

Breij, E.C., Brink, B.P., Veerhuis, R., van den Berg, C., Vloet, R., Yan, R., Dijkstra, C.D., van der Valk, P., and Bo, L. (2008). Homogeneity of active demyelinating lesions in established multiple sclerosis. Ann Neurol *63*, 16-25.

Brown, C.D., Mangravite, L.M., and Engelhardt, B.E. (2013). Integrative Modeling of eQTLs and Cis-Regulatory Elements Suggests Mechanisms Underlying Cell Type Specificity of eQTLs. PLoS Genet *9*, e1003649.

Brucklacher-Waldert, V., Stuerner, K., Kolster, M., Wolthausen, J., and Tolosa, E. (2009). Phenotypical and functional characterization of T helper 17 cells in multiple sclerosis. Brain *132*, 3329-3341.

Burton, D.R. (1985). Immunoglobulin G: functional sites. Mol Immunol *22*, 161-206.

Burton, P.R., Clayton, D.G., Cardon, L.R., Craddock, N., Deloukas, P., Duncanson, A., Kwiatkowski, D.P., McCarthy, M.I., Ouwehand, W.H., Samani, N.J.*, et al.* (2007). Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. Nat Genet *39*, 1329-1337.

Busse, D., de la Rosa, M., Hobiger, K., Thurley, K., Flossdorf, M., Scheffold, A., and Hofer, T. (2010). Competing feedback loops shape IL-2 signaling between helper and regulatory T lymphocytes in cellular microenvironments. Proc Natl Acad Sci U S A *107*, 3058-3063.

Cash, E., Minty, A., Ferrara, P., Caput, D., Fradelizi, D., and Rott, O. (1994). Macrophage-inactivating IL-13 suppresses experimental autoimmune encephalomyelitis in rats. J Immunol *153*, 4258-4267.

Cassan, C., and Liblau, R.S. (2007). Immune tolerance and control of CNS autoimmunity: from animal models to MS patients. J Neurochem *100*, 883-892.

Chen, M., Cheng, A., Chen, Y.Q., Hymel, A., Hanson, E.P., Kimmel, L., Minami, Y., Taniguchi, T., Changelian, P.S., and O'Shea, J.J. (1997). The amino terminus of JAK3 is necessary and sufficient for binding to the common gamma chain and confers the ability to transmit interleukin 2-mediated signals. Proc Natl Acad Sci U S A *94*, 6910-6915.

Choy, E., Yelensky, R., Bonakdar, S., Plenge, R.M., Saxena, R., De Jager, P.L., Shaw, S.Y., Wolfish, C.S., Slavik, J.M., Cotsapas, C.*, et al.* (2008). Genetic analysis of human traits in vitro: drug response and gene expression in lymphoblastoid cell lines. PLoS Genet *4*, e1000287.

Chrencik, J.E., Patny, A., Leung, I.K., Korniski, B., Emmons, T.L., Hall, T., Weinberg, R.A., Gormley, J.A., Williams, J.M., Day, J.E.*, et al.* (2010). Structural and thermodynamic characterization of the TYK2 and JAK3 kinase domains in complex with CP-690550 and CMP-6. J Mol Biol *400*, 413-433.

Christensen, J.R., Bornsen, L., Ratzer, R., Piehl, F., Khademi, M., Olsson, T., Sorensen, P.S., and Sellebjerg, F. (2013). Systemic inflammation in progressive multiple sclerosis involves follicular T-helper, Th17- and activated B-cells and correlates with progression. PLoS One *8*, e57820.

Cohen, J. (1988). Statistical power analysis for the behavioural sciences, 2nd edn (Lawrence Erlbaum Associates).

Cohen, J.A., Coles, A.J., Arnold, D.L., Confavreux, C., Fox, E.J., Hartung, H.P., Havrdova, E., Selmaj, K.W., Weiner, H.L., Fisher, E.*, et al.* (2012). Alemtuzumab versus interferon beta 1a as first-line treatment for patients with relapsing-remitting multiple sclerosis: a randomised controlled phase 3 trial. Lancet *380*, 1819-1828.

Colamonici, O., Yan, H., Domanski, P., Handa, R., Smalley, D., Mullersman, J., Witte, M., Krishnan, K., and Krolewski, J. (1994). Direct binding to and tyrosine phosphorylation of the alpha subunit of the type I interferon receptor by p135tyk2 tyrosine kinase. Mol Cell Biol *14*, 8133-8142.

Coles, A.J., Twyman, C.L., Arnold, D.L., Cohen, J.A., Confavreux, C., Fox, E.J., Hartung, H.P., Havrdova, E., Selmaj, K.W., Weiner, H.L.*, et al.* (2012). Alemtuzumab for patients with relapsing multiple sclerosis after disease-modifying therapy: a randomised controlled phase 3 trial. Lancet *380*, 1829-1839.

Comabella, M., Lunemann, J.D., Rio, J., Sanchez, A., Lopez, C., Julia, E., Fernandez, M., Nonell, L., Camina-Tato, M., Deisenhammer, F.*, et al.* (2009). A type I interferon signature in monocytes is associated with poor response to interferon-beta in multiple sclerosis. Brain *132*, 3353-3365.

Comi, G., Jeffery, D., Kappos, L., Montalban, X., Boyko, A., Rocca, M.A., and Filippi, M. (2012). Placebo-controlled trial of oral laquinimod for multiple sclerosis. N Engl J Med *366*, 1000-1009.

Compston, A., and Coles, A. (2008). Multiple sclerosis. Lancet *372*, 1502-1517.

Compston, A., McDonald, I., Noseworthy, J., Lassmann, H., Miller, D., Smith, K., Wekerle, H., Confavreux, C. (2005). McAlpine's Multiple Sclerosis, Fourth edn (London: Churchill Livingstone).

Compston, D.A., Batchelor, J.R., and McDonald, W.I. (1976). B-lymphocyte alloantigens associated with multiple sclerosis. Lancet *2*, 1261-1265.

Cooper, J.D., Simmonds, M.J., Walker, N.M., Burren, O., Brand, O.J., Guo, H., Wallace, C., Stevens, H., Coleman, G., Franklyn, J.A.*, et al.* (2012). Seven newly identified loci for autoimmune thyroid disease. Hum Mol Genet *21*, 5202-5208.

Cortes, A., and Brown, M.A. (2011). Promise and pitfalls of the Immunochip. Arthritis Res Ther *13*, 101.

Cortes, A., Hadler, J., Pointon, J.P., Robinson, P.C., Karaderi, T., Leo, P., Cremin, K., Pryce, K., Harris, J., Lee, S.*, et al.* (2013). Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. Nat Genet *45*, 730-738.

Cotari, J.W., Voisinne, G., Dar, O.E., Karabacak, V., and Altan-Bonnet, G. (2013). Cell-to-cell variability analysis dissects the plasticity of signaling of common gamma chain cytokines in T cells. Sci Signal *6*, ra17.

Couturier, N., Bucciarelli, F., Nurtdinov, R.N., Debouverie, M., Lebrun-Frenay, C., Defer, G., Moreau, T., Confavreux, C., Vukusic, S., Cournu-Rebeix, I.*, et al.* (2011). Tyrosine kinase 2 variant influences T lymphocyte polarization and multiple sclerosis susceptibility. Brain *134*, 693-703.

Cunninghame Graham, D.S., Akil, M., and Vyse, T.J. (2007). Association of polymorphisms across the tyrosine kinase gene, TYK2 in UK SLE families. Rheumatology (Oxford) *46*, 927-930.

Cunninghame Graham, D.S., Morris, D.L., Bhangale, T.R., Criswell, L.A., Syvanen, A.C., Ronnblom, L., Behrens, T.W., Graham, R.R., and Vyse, T.J. (2011). Association of NCF2, IKZF1, IRF8, IFIH1, and TYK2 with systemic lupus erythematosus. PLoS Genet *7*, e1002341.

Darnell, J.E., Jr., Kerr, I.M., and Stark, G.R. (1994). Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. Science *264*, 1415-1421.

Dawson, M.A., Bannister, A.J., Gottgens, B., Foster, S.D., Bartke, T., Green, A.R., and Kouzarides, T. (2009). JAK2 phosphorylates histone H3Y41 and excludes HP1alpha from chromatin. Nature *461*, 819-822.

De Jager, P.L., Jia, X., Wang, J., de Bakker, P.I., Ottoboni, L., Aggarwal, N.T., Piccio, L., Raychaudhuri, S., Tran, D., Aubin, C.*, et al.* (2009). Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. Nat Genet *41*, 776-782.

de Weerd, N.A., and Nguyen, T. (2012). The interferons and their receptors--distribution and regulation. Immunol Cell Biol *90*, 483-491.

de Weerd, N.A., Vivian, J.P., Nguyen, T.K., Mangan, N.E., Gould, J.A., Braniff, S.J., Zaker-Tabrizi, L., Fung, K.Y., Forster, S.C., Beddoe, T.*, et al.* (2013). Structural basis of a unique interferon-beta signaling axis mediated via the receptor IFNAR1. Nat Immunol.

Debruyne, J., Philippe, J., Dereuck, J., Willems, A., and Leroux-Roels, G. (1998). Relapse markers in multiple sclerosis: are in vitro cytokine production changes reflected by circulatory T-cell phenotype alterations? Mult Scler *4*, 193-197.

Decker, T., Muller, M., and Stockinger, S. (2005). The yin and yang of type I interferon activity in bacterial infection. Nat Rev Immunol *5*, 675-687.

Delmotte, P. (1971). Gel isoelectric focusing of cerebrospinal fluid proteins: a potential diagnostic tool. Z Klin Chem Klin Biochem *9*, 334-336.

DeLuca, G.C., Alterman, R., Martin, J.L., Mittal, A., Blundell, S., Bird, S., Beale, H., Hong, L.S., and Esiri, M.M. (2013). Casting light on multiple sclerosis heterogeneity: the role of HLA-DRB1 on spinal cord pathology. Brain *136*, 1025-1034.

Dendrou, C.A., Fung, E., Esposito, L., Todd, J.A., Wicker, L.S., and Plagnol, V. (2009a). Fluorescence intensity normalisation: correcting for time effects in large-scale flow cytometric analysis. Adv Bioinformatics, 476106.

Dendrou, C.A., Plagnol, V., Fung, E., Yang, J.H., Downes, K., Cooper, J.D., Nutland, S., Coleman, G., Himsworth, M., Hardy, M.*, et al.* (2009b). Cell-specific protein phenotypes for the autoimmune locus IL2RA using a genotype-selectable human bioresource. Nat Genet *41*, 1011-1015.

Dhib-Jalbut, S., and Marks, S. (2010). Interferon-beta mechanisms of action in multiple sclerosis. Neurology *74 Suppl 1*, S17-24.

Di Meglio, P., Di Cesare, A., Laggner, U., Chu, C.C., Napolitano, L., Villanova, F., Tosi, I., Capon, F., Trembath, R.C., Peris, K., and Nestle, F.O. (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.

Di Meglio, P., Villanova, F., Napolitano, L., Tosi, I., Terranova Barberio, M., Mak, R.K., Nutland, S., Smith, C.H., Barker, J.N., Todd, J.A., and Nestle, F.O. (2013). The IL23R A/Gln381 Allele Promotes IL-23 Unresponsiveness in Human Memory T-Helper 17 Cells and Impairs Th17 Responses in Psoriasis Patients. J Invest Dermatol *133*, 2381-2389.

Domanski, P., Fish, E., Nadeau, O.W., Witte, M., Platanias, L.C., Yan, H., Krolewski, J., Pitha, P., and Colamonici, O.R. (1997). A region of the beta subunit of the interferon alpha receptor different from box 1 interacts with Jak1 and is sufficient to activate the Jak-Stat pathway and induce an antiviral state. J Biol Chem *272*, 26388-26393.

Donnelly, P. (2008). Progress and challenges in genome-wide association studies in humans. Nature *456*, 728-731.

Doolittle, T.H., Myers, R.H., Lehrich, J.R., Birnbaum, G., Sheremata, W., Franklin, G.M., Nelson, L.M., and Hauser, S.L. (1990). Multiple sclerosis sibling pairs: clustered onset and familial predisposition. Neurology *40*, 1546-1552.

Ebers, G.C., Sadovnick, A.D., and Risch, N.J. (1995). A genetic basis for familial aggregation in multiple sclerosis. Canadian Collaborative Study Group. Nature *377*, 150-151.

Ebers, G.C., Yee, I.M., Sadovnick, A.D., and Duquette, P. (2000). Conjugal multiple sclerosis: population-based prevalence and recurrence risks in offspring. Canadian Collaborative Study Group. Ann Neurol *48*, 927-931.

Eichorst (1896). Uber infantile und hereditare multiple sclerose. Virch Arch *146*, 173-193.

Ellinghaus, D., Baurecht, H., Esparza-Gordillo, J., Rodriguez, E., Matanovic, A., Marenholz, I., Hubner, N., Schaarschmidt, H., Novak, N., Michel, S.*, et al.* (2013). High-density genotyping study identifies four new susceptibility loci for atopic dermatitis. Nat Genet *45*, 808-812.

Emmons, T.L., Wrightstone, A.D., Baima, E.T., Brown, S., Sommers, C.D., Hirsch, J.L., Pegg, L.E., Weinberg, R.A., Fischer, H.D., Wittwer, A.J., and Tomasselli, A.G. (2012). Differential kinetics and inhibition of purified recombinant tyrosine kinase 2 (TYK-2) and its catalytic domain JH-1. Protein Pept Lett *19*, 485-491.

Esiri, M.M. (1977). Immunoglobulin-containing cells in multiple-sclerosis plaques. Lancet *2*, 478.

Estecha, A., Aguilera-Montilla, N., Sanchez-Mateos, P., and Puig-Kroger, A. (2012). RUNX3 regulates intercellular adhesion molecule 3 (ICAM-3) expression during macrophage differentiation and monocyte extravasation. PLoS One *7*, e33313.

Etzensperger, R., McMahon, R.M., Jones, E.Y., and Fugger, L. (2008). Dissection of the multiple sclerosis associated DR2 haplotype. J Autoimmun *31*, 201-207.

Evans, D.M., Spencer, C.C., Pointon, J.J., Su, Z., Harvey, D., Kochan, G., Oppermann, U., Dilthey, A., Pirinen, M., Stone, M.A.*, et al.* (2011). Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. Nat Genet *43*, 761-767.

Eyre, S., Bowes, J., Diogo, D., Lee, A., Barton, A., Martin, P., Zhernakova, A., Stahl, E., Viatte, S., McAllister, K.*, et al.* (2012). High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. Nat Genet *44*, 1336-1340.

Fahey, A.J., Robins, R.A., and Constantinescu, C.S. (2007). Reciprocal effects of IFN-beta and IL-12 on STAT4 activation and cytokine induction in T cells. J Leukoc Biol *81*, 1562-1567.

Fairfax, B.P., Makino, S., Radhakrishnan, J., Plant, K., Leslie, S., Dilthey, A., Ellis, P., Langford, C., Vannberg, F.O., and Knight, J.C. (2012). Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. Nat Genet *44*, 502-510.

Ferguson, B., Matyszak, M.K., Esiri, M.M., and Perry, V.H. (1997). Axonal damage in acute multiple sclerosis lesions. Brain *120 (Pt 3)*, 393-399.

Ferreira, R.C., Freitag, D.F., Cutler, A.J., Howson, J.M., Rainbow, D.B., Smyth, D.J., Kaptoge, S., Clarke, P., Boreham, C., Coulson, R.M.*, et al.* (2013). Functional IL6R 358Ala allele impairs classical IL-6 receptor signaling and influences risk of diverse inflammatory diseases. PLoS Genet *9*, e1003444.

Firmbach-Kraft, I., Byers, M., Shows, T., Dalla-Favera, R., and Krolewski, J.J. (1990). tyk2, prototype of a novel class of non-receptor tyrosine kinase genes. Oncogene *5*, 1329-1336.

Fisher, R.A. (1930). The Genetical Theory of Natural Selection. Oxford University Press, Oxford.

Fleming, J.O., Isaak, A., Lee, J.E., Luzzio, C.C., Carrithers, M.D., Cook, T.D., Field, A.S., Boland, J., and Fabry, Z. (2011). Probiotic helminth administration in relapsing-remitting multiple sclerosis: a phase 1 study. Mult Scler *17*, 743-754.

Fox, R.J., Miller, D.H., Phillips, J.T., Hutchinson, M., Havrdova, E., Kita, M., Yang, M., Raghupathi, K., Novas, M., Sweetser, M.T.*, et al.* (2012). Placebo-controlled phase 3 study of oral BG-12 or glatiramer in multiple sclerosis. N Engl J Med *367*, 1087-1097.

Franke, A., McGovern, D.P., Barrett, J.C., Wang, K., Radford-Smith, G.L., Ahmad, T., Lees, C.W., Balschun, T., Lee, J., Roberts, R.*, et al.* (2010). Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet *42*, 1118-1125.

Freedman, M.S., Ruijs, T.C., Selin, L.K., and Antel, J.P. (1991). Peripheral blood gamma-delta T cells lyse fresh human brain-derived oligodendrocytes. Ann Neurol *30*, 794-800.

Friese, M.A., Craner, M.J., Etzensperger, R., Vergo, S., Wemmie, J.A., Welsh, M.J., Vincent, A., and Fugger, L. (2007). Acid-sensing ion channel-1 contributes to axonal degeneration in autoimmune inflammation of the central nervous system. Nat Med *13*, 1483-1489.

Friese, M.A., and Fugger, L. (2005). Autoreactive CD8+ T cells in multiple sclerosis: a new target for therapy? Brain *128*, 1747-1763.

Friese, M.A., Jakobsen, K.B., Friis, L., Etzensperger, R., Craner, M.J., McMahon, R.M., Jensen, L.T., Huygelen, V., Jones, E.Y., Bell, J.I., and Fugger, L. (2008). Opposing effects of HLA class I molecules in tuning autoreactive CD8+ T cells in multiple sclerosis. Nat Med *14*, 1227-1235.

Gambineri, E., Torgerson, T.R., and Ochs, H.D. (2003). Immune dysregulation, polyendocrinopathy, enteropathy, and X-linked inheritance (IPEX), a syndrome of systemic autoimmunity caused by mutations of FOXP3, a critical regulator of T-cell homeostasis. Curr Opin Rheumatol *15*, 430-435.

Gandhi, R., Laroni, A., and Weiner, H.L. (2010). Role of the innate immune system in the pathogenesis of multiple sclerosis. J Neuroimmunol *221*, 7-14.

Gauzzi, M.C., Barbieri, G., Richter, M.F., Uze, G., Ling, L., Fellous, M., and Pellegrini, S. (1997). The amino-terminal region of Tyk2 sustains the level of interferon alpha receptor 1, a component of the interferon alpha/beta receptor. Proc Natl Acad Sci U S A *94*, 11839-11844.

Gauzzi, M.C., Velazquez, L., McKendry, R., Mogensen, K.E., Fellous, M., and Pellegrini, S. (1996). Interferon-alpha-dependent activation of Tyk2 requires phosphorylation of positive regulatory tyrosines by another kinase. J Biol Chem *271*, 20494-20500.

Geiger, T.R., and Martin, J.M. (2006). The Epstein-Barr virus-encoded LMP-1 oncoprotein negatively affects Tyk2 phosphorylation and interferon signaling in human B cells. J Virol *80*, 11638-11650.

Genain, C.P., Cannella, B., Hauser, S.L., and Raine, C.S. (1999). Identification of autoantibodies associated with myelin damage in multiple sclerosis. Nat Med *5*, 170-175.

Gold, R., Kappos, L., Arnold, D.L., Bar-Or, A., Giovannoni, G., Selmaj, K., Tornatore, C., Sweetser, M.T., Yang, M., Sheikh, S.I., and Dawson, K.T. (2012). Placebo-controlled phase 3 study of oral BG-12 for relapsing multiple sclerosis. N Engl J Med *367*, 1098-1107.

Gregersen, J.W., Kranc, K.R., Ke, X., Svendsen, P., Madsen, L.S., Thomsen, A.R., Cardon, L.R., Bell, J.I., and Fugger, L. (2006). Functional epistasis on a common MHC haplotype associated with multiple sclerosis. Nature *443*, 574-577.

Gregory, A.P., Dendrou, C.A., Attfield, K.E., Haghikia, A., Xifara, D.K., Butter, F., Poschmann, G., Kaur, G., Lambert, L., Leach, O.A.*, et al.* (2012). TNF receptor 1 genetic risk mirrors outcome of anti-TNF therapy in multiple sclerosis. Nature *488*, 508-511.

Gregory, G.D., Robbie-Ryan, M., Secor, V.H., Sabatino, J.J., Jr., and Brown, M.A. (2005). Mast cells are required for optimal autoreactive T cell responses in a murine model of multiple sclerosis. Eur J Immunol *35*, 3478-3486.

Gregory, S.G., Schmidt, S., Seth, P., Oksenberg, J.R., Hart, J., Prokop, A., Caillier, S.J., Ban, M., Goris, A., Barcellos, L.F.*, et al.* (2007). Interleukin 7 receptor alpha chain (IL7R) shows allelic and functional association with multiple sclerosis. Nat Genet *39*, 1083-1091.

Gromak, N., Rideau, A., Southby, J., Scadden, A.D., Gooding, C., Huttelmaier, S., Singer, R.H., and Smith, C.W. (2003). The PTB interacting protein raver1 regulates alpha-tropomyosin alternative splicing. EMBO J *22*, 6356-6364.

Guo, J.T., Hayashi, J., and Seeger, C. (2005). West Nile virus inhibits the signal transduction pathway of alpha interferon. J Virol *79*, 1343-1350.

Hafler, D.A., Compston, A., Sawcer, S., Lander, E.S., Daly, M.J., De Jager, P.L., de Bakker, P.I., Gabriel, S.B., Mirel, D.B., Ivinson, A.J.*, et al.* (2007). Risk alleles for multiple sclerosis identified by a genomewide study. N Engl J Med *357*, 851-862.

Haghikia, A., Hohlfeld, R., Gold, R., and Fugger, L. (2013). Therapies for multiple sclerosis: translational achievements and outstanding needs. Trends Mol Med *19*, 309-319.

Handel, A.E., Williamson, A.J., Disanto, G., Handunnetthi, L., Giovannoni, G., and Ramagopalan, S.V. (2010). An updated meta-analysis of risk of multiple sclerosis following infectious mononucleosis. PLoS One *5*.

Hansen, T., Skytthe, A., Stenager, E., Petersen, H.C., Bronnum-Hansen, H., and Kyvik, K.O. (2005). Concordance for multiple sclerosis in Danish twins: an update of a nationwide study. Mult Scler *11*, 504-510.

Harkiolaki, M., Holmes, S.L., Svendsen, P., Gregersen, J.W., Jensen, L.T., McMahon, R., Friese, M.A., van Boxel, G., Etzensperger, R., Tzartos, J.S.*, et al.* (2009). T cell-mediated autoimmune disease due to low-affinity crossreactivity to common microbial peptides. Immunity *30*, 348-357.

Hauser, S.L., Waubant, E., Arnold, D.L., Vollmer, T., Antel, J., Fox, R.J., Bar-Or, A., Panzara, M., Sarkar, N., Agarwal, S.*, et al.* (2008). B-cell depletion with rituximab in relapsing-remitting multiple sclerosis. N Engl J Med *358*, 676-688.

Havrdová (2012). Positive proof of concept of AIN457, an antibody against interleukin-17A, in relapsing-remitting multiple sclerosis. In ECTRIMS.

Hawkes, C.H., and Macgregor, A.J. (2009). Twin studies and the heritability of MS: a conclusion. Mult Scler *15*, 661-667.

Hellquist, A., Jarvinen, T.M., Koskenmies, S., Zucchelli, M., Orsmark-Pietras, C., Berglind, L., Panelius, J., Hasan, T., Julkunen, H., D'Amato, M.*, et al.* (2009). Evidence for genetic association and interaction between the TYK2 and IRF5 genes in systemic lupus erythematosus. J Rheumatol *36*, 1631-1638.

Hemminki, K., Li, X., Sundquist, J., Hillert, J., and Sundquist, K. (2009). Risk for multiple sclerosis in relatives and spouses of patients diagnosed with autoimmune and related conditions. Neurogenetics *10*, 5-11.

Hernan, M.A., Olek, M.J., and Ascherio, A. (2001). Cigarette smoking and incidence of multiple sclerosis. Am J Epidemiol *154*, 69-74.

Hervas-Stubbs, S., Perez-Gracia, J.L., Rouzaut, A., Sanmamed, M.F., Le Bon, A., and Melero, I. (2011). Direct effects of type I interferons on cells of the immune system. Clin Cancer Res *17*, 2619-2627.

Hinks, A., Cobb, J., Marion, M.C., Prahalad, S., Sudman, M., Bowes, J., Martin, P., Comeau, M.E., Sajuthi, S., Andrews, R.*, et al.* (2013). Dense genotyping of immune-related disease regions identifies 14 new susceptibility loci for juvenile idiopathic arthritis. Nat Genet *45*, 664-669.

Hirota, K., Duarte, J.H., Veldhoen, M., Hornsby, E., Li, Y., Cua, D.J., Ahlfors, H., Wilhelm, C., Tolaini, M., Menzel, U.*, et al.* (2011). Fate mapping of IL-17-producing T cells in inflammatory responses. Nat Immunol *12*, 255-263.

Hofer, T., Krichevsky, O., and Altan-Bonnet, G. (2012). Competition for IL-2 between Regulatory and Effector T Cells to Chisel Immune Responses. Front Immunol *3*, 268.

Hogquist, K.A., Baldwin, T.A., and Jameson, S.C. (2005). Central tolerance: learning self-control in the thymus. Nat Rev Immunol *5*, 772-782.

Hollenhorst, P.C., Chandler, K.J., Poulsen, R.L., Johnson, W.E., Speck, N.A., and Graves, B.J. (2009). DNA specificity determinants associate with distinct transcription factor functions. PLoS Genet *5*, e1000778.

Horton, R., Wilming, L., Rand, V., Lovering, R.C., Bruford, E.A., Khodiyar, V.K., Lush, M.J., Povey, S., Talbot, C.C., Jr., Wright, M.W.*, et al.* (2004). Gene map of the extended human MHC. Nat Rev Genet *5*, 889-899.

Huan, J., Culbertson, N., Spencer, L., Bartholomew, R., Burrows, G.G., Chou, Y.K., Bourdette, D., Ziegler, S.F., Offner, H., and Vandenbark, A.A. (2005). Decreased FOXP3 levels in multiple sclerosis patients. J Neurosci Res *81*, 45-52.

Hueber, W., Sands, B.E., Lewitzky, S., Vandemeulebroecke, M., Reinisch, W., Higgins, P.D., Wehkamp, J., Feagan, B.G., Yao, M.D., Karczewski, M.*, et al.* (2012). Secukinumab, a human anti-IL-17A monoclonal antibody, for moderate to severe Crohn's disease: unexpected results of a randomised, double-blind placebo-controlled trial. Gut *61*, 1693-1700.

IMSGC (2013). Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nature Genetics.

Irie-Sasaki, J., Sasaki, T., Matsumoto, W., Opavsky, A., Cheng, M., Welstead, G., Griffiths, E., Krawczyk, C., Richardson, C.D., Aitken, K.*, et al.* (2001). CD45 is a JAK phosphatase and negatively regulates cytokine receptor signalling. Nature *409*, 349-354.

Isaacs, A., and Lindenmann, J. (1957). Virus interference. I. The interferon. Proc R Soc Lond B Biol Sci *147*, 258-267.

Ishizaki, M., Akimoto, T., Muromoto, R., Yokoyama, M., Ohshiro, Y., Sekine, Y., Maeda, H., Shimoda, K., Oritani, K., and Matsuda, T. (2011a). Involvement of Tyrosine Kinase-2 in Both the IL-12/Th1 and IL-23/Th17 Axes In Vivo. J Immunol.

Ishizaki, M., Muromoto, R., Akimoto, T., Ohshiro, Y., Takahashi, M., Sekine, Y., Maeda, H., Shimoda, K., Oritani, K., and Matsuda, T. (2011b). Tyk2 deficiency protects joints against destruction in anti-type II collagen antibody-induced arthritis in mice. Int Immunol.

Jager, L.D., Dabelic, R., Waiboci, L.W., Lau, K., Haider, M.S., Ahmed, C.M., Larkin, J., 3rd, David, S., and Johnson, H.M. (2011). The kinase inhibitory region of SOCS-1 is sufficient to inhibit T-helper 17 and other immune functions in experimental allergic encephalomyelitis. J Neuroimmunol *232*, 108-118.

Jaitin, D.A., Roisman, L.C., Jaks, E., Gavutis, M., Piehler, J., Van der Heyden, J., Uze, G., and Schreiber, G. (2006). Inquiring into the differential action of interferons (IFNs): an IFN-alpha2 mutant with enhanced affinity to IFNAR1 is functionally similar to IFN-beta. Mol Cell Biol *26*, 1888-1897.

James, C., Ugo, V., Le Couedic, J.P., Staerk, J., Delhommeau, F., Lacout, C., Garcon, L., Raslova, H., Berger, R., Bennaceur-Griscelli, A.*, et al.* (2005). A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. Nature *434*, 1144-1148.

James, E., Dobson, R., Kuhle, J., Baker, D., Giovannoni, G., and Ramagopalan, S.V. (2013). The effect of vitamin D-related interventions on multiple sclerosis relapses: a meta-analysis. Mult Scler.

Jarvinen, T.M., Hellquist, A., Koskenmies, S., Einarsdottir, E., Koskinen, L.L., Jeskanen, L., Berglind, L., Panelius, J., Hasan, T., Ranki, A.*, et al.* (2010). Tyrosine kinase 2 and interferon regulatory factor 5 polymorphisms are associated with discoid and subacute cutaneous lupus erythematosus. Exp Dermatol *19*, 123-131.

Jersild, C., Svejgaard, A., Fog, T., and Ammitzboll, T. (1973). HL-A antigens and diseases. I. Multiple sclerosis. Tissue Antigens *3*, 243-250.

Jilek, S., Schluep, M., Rossetti, A.O., Guignard, L., Le Goff, G., Pantaleo, G., and Du Pasquier, R.A. (2007). CSF enrichment of highly differentiated CD8+ T cells in early multiple sclerosis. Clin Immunol *123*, 105-113.

Johnson, B.A., Wang, J., Taylor, E.M., Caillier, S.J., Herbert, J., Khan, O.A., Cross, A.H., De Jager, P.L., Gourraud, P.A., Cree, B.C.*, et al.* (2009). Multiple sclerosis susceptibility alleles in African Americans. Genes Immun.

Jostins, L., Ripke, S., Weersma, R.K., Duerr, R.H., McGovern, D.P., Hui, K.Y., Lee, J.C., Schumm, L.P., Sharma, Y., Anderson, C.A.*, et al.* (2012). Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature *491*, 119-124.

Kabat, E.A., Moore, D.H., and Landow, H. (1942). An Electrophoretic Study of the Protein Components in Cerebrospinal Fluid and Their Relationship to the Serum Proteins. J Clin Invest *21*, 571-577.

Kalie, E., Jaitin, D.A., Podoplelova, Y., Piehler, J., and Schreiber, G. (2008). The stability of the ternary interferon-receptor complex rather than the affinity to the individual subunits dictates differential biological activities. J Biol Chem *283*, 32925-32936.

Kamezaki, K., Shimoda, K., Numata, A., Matsuda, T., Nakayama, K., and Harada, M. (2004). The role of Tyk2, Stat1 and Stat4 in LPS-induced endotoxin signals. Int Immunol *16*, 1173-1179.

Kaminker, J.S., Zhang, Y., Waugh, A., Haverty, P.M., Peters, B., Sebisanovic, D., Stinson, J., Forrest, W.F., Bazan, J.F., Seshagiri, S., and Zhang, Z. (2007). Distinguishing cancer-associated missense mutations from common polymorphisms. Cancer Res *67*, 465-473.

Karaghiosoff, M., Neubauer, H., Lassnig, C., Kovarik, P., Schindler, H., Pircher, H., McCoy, B., Bogdan, C., Decker, T., Brem, G.*, et al.* (2000). Partial impairment of cytokine responses in Tyk2-deficient mice. Immunity *13*, 549-560.

Karaghiosoff, M., Steinborn, R., Kovarik, P., Kriegshauser, G., Baccarini, M., Donabauer, B., Reichart, U., Kolbe, T., Bogdan, C., Leanderson, T.*, et al.* (2003). Central role for type I interferons and Tyk2 in lipopolysaccharide-induced endotoxin shock. Nat Immunol *4*, 471-477.

Karandikar, N.J., Crawford, M.P., Yan, X., Ratts, R.B., Brenchley, J.M., Ambrozak, D.R., Lovett-Racke, A.E., Frohman, E.M., Stastny, P., Douek, D.C.*, et al.* (2002). Glatiramer acetate (Copaxone) therapy induces CD8(+) T cell responses in patients with multiple sclerosis. J Clin Invest *109*, 641-649.

Karni, A., Abraham, M., Monsonego, A., Cai, G., Freeman, G.J., Hafler, D., Khoury, S.J., and Weiner, H.L. (2006). Innate immunity in multiple sclerosis: myeloid dendritic cells in secondary progressive multiple sclerosis are activated and drive a proinflammatory immune response. J Immunol *177*, 4196-4202.

Kebir, H., Kreymborg, K., Ifergan, I., Dodelet-Devillers, A., Cayrol, R., Bernard, M., Giuliani, F., Arbour, N., Becher, B., and Prat, A. (2007). Human TH17 lymphocytes promote blood-brain barrier disruption and central nervous system inflammation. Nat Med *13*, 1173-1175.

Kerlero de Rosbo, N., Milo, R., Lees, M.B., Burger, D., Bernard, C.C., and Ben-Nun, A. (1993). Reactivity to myelin antigens in multiple sclerosis. Peripheral blood lymphocytes respond predominantly to myelin oligodendrocyte glycoprotein. J Clin Invest *92*, 2602-2608.

Kessel, J.M., Hayflick, J., Weyrich, A.S., Hoffman, P.A., Gallatin, M., McIntyre, T.M., Prescott, S.M., and Zimmerman, G.A. (1998). Coengagement of ICAM-3 and Fc receptors induces chemokine secretion and spreading by myeloid leukocytes. J Immunol *160*, 5579-5587.

Kilic, S.S., Hacimustafaoglu, M., Boisson-Dupuis, S., Kreins, A.Y., Grant, A.V., Abel, L., and Casanova, J.L. (2012). A patient with tyrosine kinase 2 deficiency without hyper-IgE syndrome. J Pediatr *160*, 1055-1057.

Kisseleva, T., Bhattacharya, S., Braunstein, J., and Schindler, C.W. (2002). Signaling through the JAK/STAT pathway, recent advances and future challenges. Gene *285*, 1-24.

Klein, L., Klugmann, M., Nave, K.A., Tuohy, V.K., and Kyewski, B. (2000). Shaping of the autoreactive T-cell repertoire by a splice variant of self protein expressed in thymic epithelial cells. Nat Med *6*, 56-61.

Kleinewietfeld, M., Manzel, A., Titze, J., Kvakan, H., Yosef, N., Linker, R.A., Muller, D.N., and Hafler, D.A. (2013). Sodium chloride drives autoimmune disease by the induction of pathogenic TH17 cells. Nature *496*, 518-522.

Koch, M.W., Metz, L.M., Agrawal, S.M., and Yong, V.W. (2013). Environmental factors and their regulation of immunity in multiple sclerosis. J Neurol Sci *324*, 10-16.

Kontzias, A., Kotlyar, A., Laurence, A., Changelian, P., and O'Shea, J.J. (2012). Jakinibs: a new class of kinase inhibitors in cancer and autoimmune disease. Curr Opin Pharmacol *12*, 464-470.

Korniski, B., Wittwer, A.J., Emmons, T.L., Hall, T., Brown, S., Wrightstone, A.D., Hirsch, J.L., Gormley, J.A., Weinberg, R.A., Leone, J.W.*, et al.* (2010). Expression, purification, and characterization of TYK-2 kinase domain, a member of the Janus kinase family. Biochem Biophys Res Commun *396*, 543-548.

Krutzik, P.O., and Nolan, G.P. (2003). Intracellular phospho-protein staining techniques for flow cytometry: monitoring single cell signaling events. Cytometry A *55*, 61-70.

Kumanovics, A., Takada, T., and Lindahl, K.F. (2003). Genomic organization of the mammalian MHC. Annu Rev Immunol *21*, 629-657.

Kumar, N., Goldminz, A.M., Kim, N., and Gottlieb, A.B. (2013). Phosphodiesterase 4-targeted treatments for autoimmune diseases. BMC Med *11*, 96.

Lang, H.L., Jacobsen, H., Ikemizu, S., Andersson, C., Harlos, K., Madsen, L., Hjorth, P., Sondergaard, L., Svejgaard, A., Wucherpfennig, K.*, et al.* (2002). A functional and structural basis for TCR cross-reactivity in multiple sclerosis. Nat Immunol *3*, 940-943.

Lassmann, H., van Horssen, J., and Mahad, D. (2012). Progressive multiple sclerosis: pathology and pathogenesis. Nat Rev Neurol *8*, 647-656.

Leaman, D.W., Chawla-Sarkar, M., Jacobs, B., Vyas, K., Sun, Y., Ozdemir, A., Yi, T., Williams, B.R., and Borden, E.C. (2003). Novel growth and death related interferon-stimulated genes (ISGs) in melanoma: greater potency of IFN-beta compared with IFN-alpha2. J Interferon Cytokine Res *23*, 745-756.

Lee, J.H., Rangarajan, E.S., Yogesha, S.D., and Izard, T. (2009). Raver1 interactions with vinculin and RNA suggest a feed-forward pathway in directing mRNA to focal adhesions. Structure *17*, 833-842.

Lee, Y.H., Choi, S.J., Ji, J.D., and Song, G.G. (2012). Associations between PXK and TYK2 polymorphisms and systemic lupus erythematosus: a meta-analysis. Inflamm Res *61*, 949-954.

Leonard, W.J. (2001). Cytokines and immunodeficiency diseases. Nat Rev Immunol *1*, 200-208.

Levine, R.L., and Gilliland, D.G. (2008). Myeloproliferative disorders. Blood *112*, 2190-2198.

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.

Liang, J., Tsui, V., Van Abbema, A., Bao, L., Barrett, K., Beresini, M., Berezhkovskiy, L., Blair, W.S., Chang, C., Driscoll, J.*, et al.* (2013a). Lead identification of novel and selective TYK2 inhibitors. Eur J Med Chem *67C*, 175-187.

Liang, J., van Abbema, A., Balazs, M., Barrett, K., Berezhkovsky, L., Blair, W., Chang, C., Delarosa, D., Devoss, J., Driscoll, J.*, et al.* (2013b). Lead Optimization of a 4-Aminopyridine Benzamide Scaffold To Identify Potent, Selective, and Orally Bioavailable TYK2 Inhibitors. J Med Chem *56*, 4521-4536.

Libri, V., Schulte, D., van Stijn, A., Ragimbeau, J., Rogge, L., and Pellegrini, S. (2008). Jakmip1 is expressed upon T cell differentiation and has an inhibitory function in cytotoxic T lymphocytes. J Immunol *181*, 5847-5856.

Liu, J.Z., Almarri, M.A., Gaffney, D.J., Mells, G.F., Jostins, L., Cordell, H.J., Ducker, S.J., Day, D.B., Heneghan, M.A., Neuberger, J.M.*, et al.* (2012). Dense fine-mapping study identifies new susceptibility loci for primary biliary cirrhosis. Nat Genet *44*, 1137-1141.

Lovett-Racke, A.E., Trotter, J.L., Lauber, J., Perrin, P.J., June, C.H., and Racke, M.K. (1998). Decreased dependence of myelin basic protein-reactive T cells on CD28-mediated costimulation in multiple sclerosis patients. A marker of activated/memory T cells. J Clin Invest *101*, 725-730.

Lucchinetti, C., Bruck, W., Parisi, J., Scheithauer, B., Rodriguez, M., and Lassmann, H. (2000). Heterogeneity of multiple sclerosis lesions: implications for the pathogenesis of demyelination. Ann Neurol *47*, 707-717.

Lucchinetti, C.F., Popescu, B.F., Bunyan, R.F., Moll, N.M., Roemer, S.F., Lassmann, H., Bruck, W., Parisi, J.E., Scheithauer, B.W., Giannini, C.*, et al.* (2011). Inflammatory cortical demyelination in early multiple sclerosis. N Engl J Med *365*, 2188-2197.

Luessi, F., Siffrin, V., and Zipp, F. (2012). Neurodegeneration in multiple sclerosis: novel treatment strategies. Expert Rev Neurother *12*, 1061-1076; quiz 1077.

Mackay, R.P. (1950). Familial occurrence of multiple sclerosis and its implications. Arch Neurol Psychiatry *64*, 155-157.

Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A.*, et al.* (2009). Finding the missing heritability of complex diseases. Nature *461*, 747-753.

Marijanovic, Z., Ragimbeau, J., Kumar, K.G., Fuchs, S.Y., and Pellegrini, S. (2006). TYK2 activity promotes ligand-induced IFNAR1 proteolysis. Biochem J *397*, 31-38.

Marrack, P., Kappler, J., and Mitchell, T. (1999). Type I interferons keep activated T cells alive. J Exp Med *189*, 521-530.

Martinez-Forero, I., Garcia-Munoz, R., Martinez-Pasamar, S., Inoges, S., Lopez-Diaz de Cerio, A., Palacios, R., Sepulcre, J., Moreno, B., Gonzalez, Z., Fernandez-Diez, B.*, et al.* (2008). IL-10 suppressor activity and ex vivo Tr1 cell function are impaired in multiple sclerosis. Eur J Immunol *38*, 576-586.

Medana, I., Martinic, M.A., Wekerle, H., and Neumann, H. (2001). Transection of major histocompatibility complex class I-induced neurites by cytotoxic T lymphocytes. Am J Pathol *159*, 809-815.

Meinl, E., Weber, F., Drexler, K., Morelle, C., Ott, M., Saruhan-Direskeneli, G., Goebels, N., Ertl, B., Jechart, G., Giegerich, G., and et al. (1993). Myelin basic protein-specific T lymphocyte repertoire in multiple sclerosis. Complexity of the response and dominance of nested epitopes due to recruitment of multiple T cell clones. J Clin Invest *92*, 2633-2643.

Mero, I.L., Lorentzen, A.R., Ban, M., Smestad, C., Celius, E.G., Aarseth, J.H., Myhr, K.M., Link, J., Hillert, J., Olsson, T.*, et al.* (2009). A rare variant of the TYK2 gene is confirmed to be associated with multiple sclerosis. Eur J Hum Genet.

Merrill, J.T., Wallace, D.J., Petri, M., Kirou, K.A., Yao, Y., White, W.I., Robbie, G., Levin, R., Berney, S.M., Chindalore, V.*, et al.* (2011). Safety profile and clinical activity of sifalimumab, a fully human anti-interferon alpha monoclonal antibody, in systemic lupus erythematosus: a phase I, multicentre, double-blind randomised study. Ann Rheum Dis *70*, 1905-1913.

Methner, A., and Zipp, F. (2013). Multiple sclerosis in 2012: Novel therapeutic options and drug targets in MS. Nat Rev Neurol *9*, 72-73.

Minegishi, Y. (2009). Hyper-IgE syndrome. Curr Opin Immunol *21*, 487-492.

Minegishi, Y., Saito, M., Morio, T., Watanabe, K., Agematsu, K., Tsuchiya, S., Takada, H., Hara, T., Kawamura, N., Ariga, T.*, et al.* (2006). Human tyrosine kinase 2 deficiency reveals its requisite roles in multiple cytokine signals involved in innate and acquired immunity. Immunity *25*, 745-755.

Montag, D.T., and Lotze, M.T. (2006). Rapid flow cytometric measurement of cytokine-induced phosphorylation pathways [CIPP] in human peripheral blood leukocytes. Clin Immunol *121*, 215-226.

Mori, K., Fujita, S.C., Watanabe, Y., Obata, K., and Hayaishi, O. (1987). Telencephalon-specific antigen identified by monoclonal antibody. Proc Natl Acad Sci U S A *84*, 3921-3925.

Morikawa, K., Kubagawa, H., Suzuki, T., and Cooper, M.D. (1987). Recombinant interferon-alpha, -beta, and -gamma enhance the proliferative response of human B cells. J Immunol *139*, 761-766.

Motulsky, H.J., and Brown, R.E. (2006). Detecting outliers when fitting data with nonlinear regression - a new method based on robust nonlinear regression and the false discovery rate. BMC Bioinformatics *7*, 123.

Mowry, E., Waubant, E., Chehoud, C., DeSantis, T., Kuczynski, J., and Warrington, J. (2012). Gut bacterial populations in multiple sclerosis and in health. Neurology *78*.

Mueller, D.L. (2010). Mechanisms maintaining peripheral tolerance. Nat Immunol *11*, 21-27.

Muller, M., Briscoe, J., Laxton, C., Guschin, D., Ziemiecki, A., Silvennoinen, O., Harpur, A.G., Barbieri, G., Witthuhn, B.A., Schindler, C., and et al. (1993). The protein tyrosine kinase JAK1 complements defects in interferon-alpha/beta and -gamma signal transduction. Nature *366*, 129-135.

Myers, M.P., Andersen, J.N., Cheng, A., Tremblay, M.L., Horvath, C.M., Parisien, J.P., Salmeen, A., Barford, D., and Tonks, N.K. (2001). TYK2 and JAK2 are substrates of protein-tyrosine phosphatase 1B. J Biol Chem *276*, 47771-47774.

Naegele, M., Tillack, K., Reinhardt, S., Schippling, S., Martin, R., and Sospedra, M. (2012). Neutrophils in multiple sclerosis are characterized by a primed phenotype. J Neuroimmunol *242*, 60-71.

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.

Norman, P. (2012). Selective JAK1 inhibitor and selective Tyk2 inhibitor patents. Expert Opin Ther Pat *22*, 1233-1249.

Nosaka, T., van Deursen, J.M., Tripp, R.A., Thierfelder, W.E., Witthuhn, B.A., McMickle, A.P., Doherty, P.C., Grosveld, G.C., and Ihle, J.N. (1995). Defective lymphoid development in mice lacking Jak3. Science *270*, 800-802.

O'Connor, R.A., Prendergast, C.T., Sabatos, C.A., Lau, C.W., Leech, M.D., Wraith, D.C., and Anderton, S.M. (2008). Cutting edge: Th1 cells facilitate the entry of Th17 cells to the central nervous system during experimental autoimmune encephalomyelitis. J Immunol *181*, 3750-3754.

O'Shea, J.J., Holland, S.M., and Staudt, L.M. (2013a). JAKs and STATs in immunity, immunodeficiency, and cancer. N Engl J Med *368*, 161-170.

O'Shea, J.J., Kontzias, A., Yamaoka, K., Tanaka, Y., and Laurence, A. (2013b). Janus kinase inhibitors in autoimmune diseases. Ann Rheum Dis *72 Suppl 2*, ii111-115.

Obermeier, B., Mentele, R., Malotka, J., Kellermann, J., Kumpfel, T., Wekerle, H., Lottspeich, F., Hohlfeld, R., and Dornmair, K. (2008). Matching of oligoclonal immunoglobulin transcriptomes and proteomes of cerebrospinal fluid in multiple sclerosis. Nat Med *14*, 688-693.

Ochoa-Reparaz, J., Rynda, A., Ascon, M.A., Yang, X., Kochetkova, I., Riccardi, C., Callis, G., Trunkle, T., and Pascual, D.W. (2008). IL-13 production by regulatory T cells protects against experimental autoimmune encephalomyelitis independently of autoantigen. J Immunol *181*, 954-968.

Oh, J., and O'Connor, P.W. (2013). An update of teriflunomide for treatment of multiple sclerosis. Ther Clin Risk Manag *9*, 177-190.

Oksenberg, J.R., and Baranzini, S.E. (2010). Multiple sclerosis genetics-is the glass half full, or half empty? Nat Rev Neurol.

Olerup, O., and Hillert, J. (1991). HLA class II-associated genetic susceptibility in multiple sclerosis: a critical evaluation. Tissue Antigens *38*, 1-15.

Olsson, T., Zhi, W.W., Hojeberg, B., Kostulas, V., Jiang, Y.P., Anderson, G., Ekre, H.P., and Link, H. (1990). Autoreactive T lymphocytes in multiple sclerosis determined by antigen-induced secretion of interferon-gamma. J Clin Invest *86*, 981-985.

Olsson, Y. (1974). Mast cells in plaques of multiple sclerosis. Acta Neurol Scand *50*, 611-618.

Ortmann, R., Smeltz, R., Yap, G., Sher, A., and Shevach, E.M. (2001). A heritable defect in IL-12 signaling in B10.Q/J mice. I. In vitro analysis. J Immunol *166*, 5712-5719.

Oyamada, A., Ikebe, H., Itsumi, M., Saiwai, H., Okada, S., Shimoda, K., Iwakura, Y., Nakayama, K.I., Iwamoto, Y., Yoshikai, Y., and Yamada, H. (2009). Tyrosine kinase 2 plays critical roles in the pathogenic CD4 T cell responses for the development of experimental autoimmune encephalomyelitis. J Immunol *183*, 7539-7546.

Panitch, H.S., Hirsch, R.L., Schindler, J., and Johnson, K.P. (1987). Treatment of multiple sclerosis with gamma interferon: exacerbations associated with activation of the immune system. Neurology *37*, 1097-1102.

Pappas, D.J., Coppola, G., Gabatto, P.A., Gao, F., Geschwind, D.H., Oksenberg, J.R., and Baranzini, S.E. (2009). Longitudinal system-based analysis of transcriptional responses to type I interferons. Physiol Genomics *38*, 362-371.

Parganas, E., Wang, D., Stravopodis, D., Topham, D.J., Marine, J.C., Teglund, S., Vanin, E.F., Bodner, S., Colamonici, O.R., van Deursen, J.M.*, et al.* (1998). Jak2 is essential for signaling through a variety of cytokine receptors. Cell *93*, 385-395.

Parkes, M., Cortes, A., van Heel, D.A., and Brown, M.A. (2013). Genetic insights into common pathways and complex relationships among immune-mediated diseases. Nat Rev Genet.

Patsopoulos, N.A., Esposito, F., Reischl, J., Lehr, S., Bauer, D., Heubach, J., Sandbrink, R., Pohl, C., Edan, G., Kappos, L.*, et al.* (2011). Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. Ann Neurol *70*, 897-912.

Pellegrini, S., John, J., Shearer, M., Kerr, I.M., and Stark, G.R. (1989). Use of a selectable marker regulated by alpha interferon to obtain mutations in the signaling pathway. Mol Cell Biol *9*, 4605-4612.

Pelletier, D., and Hafler, D.A. (2012). Fingolimod for multiple sclerosis. N Engl J Med *366*, 339-347.

Perfetto, S.P., Ambrozak, D., Nguyen, R., Chattopadhyay, P.K., and Roederer, M. (2012). Quality assurance for polychromatic flow cytometry using a suite of calibration beads. Nat Protoc *7*, 2067-2079.

Peterson, P., Nagamine, K., Scott, H., Heino, M., Kudoh, J., Shimizu, N., Antonarakis, S.E., and Krohn, K.J. (1998). APECED: a monogenic autoimmune disease providing new clues to self-tolerance. Immunol Today *19*, 384-386.

Pidasheva, S., Trifari, S., Phillips, A., Hackney, J.A., Ma, Y., Smith, A., Sohn, S.J., Spits, H., Little, R.D., Behrens, T.W.*, et al.* (2011). Functional studies on the IBD susceptibility gene IL23R implicate reduced receptor function in the protective genetic variant R381Q. PLoS One *6*, e25038.

Piganis, R.A., De Weerd, N.A., Gould, J.A., Schindler, C.W., Mansell, A., Nicholson, S.E., and Hertzog, P.J. (2011). Suppressor of cytokine signaling (SOCS) 1 inhibits type I interferon (IFN) signaling via the interferon alpha receptor (IFNAR1)-associated tyrosine kinase Tyk2. J Biol Chem *286*, 33811-33818.

Pillai, S. (2013). Rethinking mechanisms of autoimmune pathogenesis. J Autoimmun.

Plagnol, V., Uz, E., Wallace, C., Stevens, H., Clayton, D., Ozcelik, T., and Todd, J.A. (2008). Extreme clonality in lymphoblastoid cell lines with implications for allele specific expression analyses. PLoS One *3*, e2966.

Platanias, L.C. (2005). Mechanisms of type-I- and type-II-interferon-mediated signalling. Nat Rev Immunol *5*, 375-386.

Pontsler, A.V., St Hilaire, A., Marathe, G.K., Zimmerman, G.A., and McIntyre, T.M. (2002). Cyclooxygenase-2 is induced in monocytes by peroxisome proliferator activated receptor gamma and oxidized alkyl phospholipids from oxidized low density lipoprotein. J Biol Chem *277*, 13029-13036.

Prchal-Murphy, M., Semper, C., Lassnig, C., Wallner, B., Gausterer, C., Teppner-Klymiuk, I., Kobolak, J., Muller, S., Kolbe, T., Karaghiosoff, M.*, et al.* (2012). TYK2 kinase activity is required for functional type I interferon responses in vivo. PLoS One *7*, e39141.

Prineas, J.W., and Wright, R.G. (1978). Macrophages, lymphocytes, and plasma cells in the perivascular compartment in chronic multiple sclerosis. Lab Invest *38*, 409-421.

Ragimbeau, J., Dondi, E., Alcover, A., Eid, P., Uze, G., and Pellegrini, S. (2003). The tyrosine kinase Tyk2 controls IFNAR1 cell surface expression. EMBO J *22*, 537-547.

Ragimbeau, J., Dondi, E., Vasserot, A., Romero, P., Uze, G., and Pellegrini, S. (2001). The receptor interaction region of Tyk2 contains a motif required for its nuclear localization. J Biol Chem *276*, 30812-30818.

Reeves, W.H. (1992). Antibodies to the p70/p80 (Ku) antigens in systemic lupus erythematosus. Rheum Dis Clin North Am *18*, 391-414.

Ren, J., Kolli, D., Liu, T., Xu, R., Garofalo, R.P., Casola, A., and Bao, X. (2011). Human metapneumovirus inhibits IFN-beta signaling by downregulating Jak1 and Tyk2 cellular levels. PLoS One *6*, e24496.

Rice, G.P., Hartung, H.P., and Calabresi, P.A. (2005). Anti-alpha4 integrin therapy for multiple sclerosis: mechanisms and rationale. Neurology *64*, 1336-1342.

Rivas, M.A., Beaudoin, M., Gardet, A., Stevens, C., Sharma, Y., Zhang, C.K., Boucher, G., Ripke, S., Ellinghaus, D., Burtt, N.*, et al.* (2011). Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease. Nat Genet *43*, 1066-1073.

Robertson, N.P., Clayton, D., Fraser, M., Deans, J., and Compston, D.A. (1996). Clinical concordance in sibling pairs with multiple sclerosis. Neurology *47*, 347-352.

Rodig, S.J., Meraz, M.A., White, J.M., Lampe, P.A., Riley, J.K., Arthur, C.D., King, K.L., Sheehan, K.C., Yin, L., Pennica, D.*, et al.* (1998). Disruption of the Jak1 gene demonstrates obligatory and nonredundant roles of the Jaks in cytokine-induced biologic responses. Cell *93*, 373-383.

Rudick, R.A., Stuart, W.H., Calabresi, P.A., Confavreux, C., Galetta, S.L., Radue, E.W., Lublin, F.D., Weinstock-Guttman, B., Wynn, D.R., Lynn, F.*, et al.* (2006). Natalizumab plus interferon beta-1a for relapsing multiple sclerosis. N Engl J Med *354*, 911-923.

Sabrautzki, S., Janas, E., Lorenz-Depiereux, B., Calzada-Wack, J., Aguilar-Pimentel, J.A., Rathkolb, B., Adler, T., Cohrs, C., Hans, W., Diener, S.*, et al.* (2013). An ENU Mutagenesis-Derived Mouse Model with a Dominant Jak1 Mutation Resembling Phenotypes of Systemic Autoimmune Disease. Am J Pathol.

Sadovnick, A.D., Dyment, D., and Ebers, G.C. (1997). Genetic epidemiology of multiple sclerosis. Epidemiol Rev *19*, 99-106.

Sadovnick, A.D., Ebers, G.C., Dyment, D.A., and Risch, N.J. (1996). Evidence for genetic basis of multiple sclerosis. The Canadian Collaborative Study Group. Lancet *347*, 1728-1730.

Saikali, P., Antel, J.P., Newcombe, J., Chen, Z., Freedman, M., Blain, M., Cayrol, R., Prat, A., Hall, J.A., and Arbour, N. (2007). NKG2D-mediated cytotoxicity toward oligodendrocytes suggests a mechanism for tissue injury in multiple sclerosis. J Neurosci *27*, 1220-1228.

Sandberg, K., Eloranta, M.L., and Campbell, I.L. (1994). Expression of alpha/beta interferons (IFN-alpha/beta) and their relationship to IFN-alpha/beta-induced genes in lymphocytic choriomeningitis. J Virol *68*, 7358-7366.

Sandborn, W.J., Feagan, B.G., Fedorak, R.N., Scherl, E., Fleisher, M.R., Katz, S., Johanns, J., Blank, M., and Rutgeerts, P. (2008). A randomized trial of Ustekinumab, a human interleukin-12/23 monoclonal antibody, in patients with moderate-to-severe Crohn's disease. Gastroenterology *135*, 1130-1141.

Sandborn, W.J., Gasink, C., Gao, L.L., Blank, M.A., Johanns, J., Guzzo, C., Sands, B.E., Hanauer, S.B., Targan, S., Rutgeerts, P.*, et al.* (2012). Ustekinumab induction and maintenance therapy in refractory Crohn's disease. N Engl J Med *367*, 1519-1528.

Sanna, S., Pitzalis, M., Zoledziewska, M., Zara, I., Sidore, C., Murru, R., Whalen, M.B., Busonero, F., Maschio, A., Costa, G.*, et al.* (2010). Variants within the immunoregulatory CBLB gene are associated with multiple sclerosis. Nat Genet.

Sanseau, P., Agarwal, P., Barnes, M.R., Pastinen, T., Richards, J.B., Cardon, L.R., and Mooser, V. (2012). Use of genome-wide association studies for drug repositioning. Nat Biotechnol *30*, 317-320.

Sareum (2013). TYK2 kinase programme.

Sawcer, S., Ban, M., Maranian, M., Yeo, T.W., Compston, A., Kirby, A., Daly, M.J., De Jager, P.L., Walsh, E., Lander, E.S.*, et al.* (2005). A high-density screen for linkage in multiple sclerosis. Am J Hum Genet *77*, 454-467.

Sawcer, S., Ban, M., Wason, J., and Dudbridge, F. (2010). What role for genetics in the prediction of multiple sclerosis? Ann Neurol *67*, 3-10.

Sawcer, S., Hellenthal, G., Pirinen, M., Spencer, C.C., Patsopoulos, N.A., Moutsianas, L., Dilthey, A., Su, Z., Freeman, C., Hunt, S.E.*, et al.* (2011). Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature *476*, 214-219.

Schapira, K., Poskanzer, D.C., and Miller, H. (1963). Familial and conjugal multiple sclerosis. Brain *86*, 315-332.

Schmied, M., Duda, P.W., Krieger, J.I., Trollmo, C., and Hafler, D.A. (2003). In vitro evidence that subcutaneous administration of glatiramer acetate induces hyporesponsive T cells in patients with multiple sclerosis. Clin Immunol *106*, 163-174.

Schneider-Hohendorf, T., Stenner, M.P., Weidenfeller, C., Zozulya, A.L., Simon, O.J., Schwab, N., and Wiendl, H. (2010). Regulatory T cells exhibit enhanced migratory characteristics, a feature impaired in patients with multiple sclerosis. Eur J Immunol *40*, 3581-3590.

Schwartz, A., Fernandez Repollet, E., Vogt, R., and Gratama, J.W. (1996). Standardizing flow cytometry: construction of a standardized fluorescence calibration plot using matching spectral calibrators. Cytometry *26*, 22-31.

Segal, B.M., Constantinescu, C.S., Raychaudhuri, A., Kim, L., Fidelus-Gort, R., and Kasper, L.H. (2008). Repeated subcutaneous injections of IL12/23 p40 neutralising antibody, ustekinumab, in patients with relapsing-remitting multiple sclerosis: a phase II, double-blind, placebo-controlled, randomised, dose-ranging study. Lancet Neurol *7*, 796-804.

Selmaj, K., Brosnan, C.F., and Raine, C.S. (1991). Colocalization of lymphocytes bearing gamma delta T-cell receptor and heat shock protein hsp65+ oligodendrocytes in multiple sclerosis. Proc Natl Acad Sci U S A *88*, 6452-6456.

Serafini, B., Rosicarelli, B., Magliozzi, R., Stigliano, E., and Aloisi, F. (2004). Detection of ectopic B-cell follicles with germinal centers in the meninges of patients with secondary progressive multiple sclerosis. Brain Pathol *14*, 164-174.

Shaw, M.H., Boyartchuk, V., Wong, S., Karaghiosoff, M., Ragimbeau, J., Pellegrini, S., Muller, M., Dietrich, W.F., and Yap, G.S. (2003). A natural mutation in the Tyk2 pseudokinase domain underlies altered susceptibility of B10.Q/J mice to infection and autoimmunity. Proc Natl Acad Sci U S A *100*, 11594-11599.

Shaw, M.H., Freeman, G.J., Scott, M.F., Fox, B.A., Bzik, D.J., Belkaid, Y., and Yap, G.S. (2006). Tyk2 negatively regulates adaptive Th1 immunity by mediating IL-10 signaling and promoting IFN-gamma-dependent IL-10 reactivation. J Immunol *176*, 7263-7271.

Shimoda, K., Kato, K., Aoki, K., Matsuda, T., Miyamoto, A., Shibamori, M., Yamashita, M., Numata, A., Takase, K., Kobayashi, S.*, et al.* (2000). Tyk2 plays a restricted role in IFN alpha signaling, although it is required for IL-12-mediated T cell function. Immunity *13*, 561-571.

Shuai, K., and Liu, B. (2003). Regulation of JAK-STAT signalling in the immune system. Nat Rev Immunol *3*, 900-911.

Sigurdsson, S., Nordmark, G., Goring, H.H., Lindroos, K., Wiman, A.C., Sturfelt, G., Jonsen, A., Rantapaa-Dahlqvist, S., Moller, B., Kere, J.*, et al.* (2005). Polymorphisms in the tyrosine kinase 2 and interferon regulatory factor 5 genes are associated with systemic lupus erythematosus. Am J Hum Genet *76*, 528-537.

Sloka, S., Silva, C., Pryse-Phillips, W., Patten, S., Metz, L., and Yong, V.W. (2011). A quantitative analysis of suspected environmental causes of MS. Can J Neurol Sci *38*, 98-105.

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.

Spach, K.M., Noubade, R., McElvany, B., Hickey, W.F., Blankenhorn, E.P., and Teuscher, C. (2009). A single nucleotide polymorphism in Tyk2 controls susceptibility to experimental allergic encephalomyelitis. J Immunol *182*, 7776-7783.

Staerk, J., Kallin, A., Demoulin, J.B., Vainchenker, W., and Constantinescu, S.N. (2005). JAK1 and Tyk2 activation by the homologous polycythemia vera JAK2 V617F mutation: cross-talk with IGF1 receptor. J Biol Chem *280*, 41893-41899.

Stasiolek, M., Bayas, A., Kruse, N., Wieczarkowiecz, A., Toyka, K.V., Gold, R., and Selmaj, K. (2006). Impaired maturation and altered regulatory function of plasmacytoid dendritic cells in multiple sclerosis. Brain *129*, 1293-1305.

Stein, M.S., Liu, Y., Gray, O.M., Baker, J.E., Kolbe, S.C., Ditchfield, M.R., Egan, G.F., Mitchell, P.J., Harrison, L.C., Butzkueven, H., and Kilpatrick, T.J. (2011). A randomized trial of high-dose vitamin D2 in relapsing-remitting multiple sclerosis. Neurology *77*, 1611-1618.

Steindler, C., Li, Z., Algarte, M., Alcover, A., Libri, V., Ragimbeau, J., and Pellegrini, S. (2004). Jamip1 (marlin-1) defines a family of proteins interacting with janus kinases and microtubules. J Biol Chem *279*, 43168-43177.

Stepanova, L., Leng, X., Parker, S.B., and Harper, J.W. (1996). Mammalian p50Cdc37 is a protein kinase-targeting subunit of Hsp90 that binds and stabilizes Cdk4. Genes Dev *10*, 1491-1502.

Stinissen, P., Vandevyver, C., Medaer, R., Vandegaer, L., Nies, J., Tuyls, L., Hafler, D.A., Raus, J., and Zhang, J. (1995). Increased frequency of gamma delta T cells in cerebrospinal fluid and peripheral blood of patients with multiple sclerosis. Reactivity, cytotoxicity, and T cell receptor V gene rearrangements. J Immunol *154*, 4883-4894.

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.

Strobl, B., Bubic, I., Bruns, U., Steinborn, R., Lajko, R., Kolbe, T., Karaghiosoff, M., Kalinke, U., Jonjic, S., and Muller, M. (2005). Novel functions of tyrosine kinase 2 in the antiviral defense against murine cytomegalovirus. J Immunol *175*, 4000-4008.

Stromnes, I.M., and Goverman, J.M. (2006). Passive induction of experimental allergic encephalomyelitis. Nat Protoc *1*, 1952-1960.

Su, L., and David, M. (1999). Inhibition of B cell receptor-mediated apoptosis by IFN. J Immunol *162*, 6317-6321.

Suarez-Gestal, M., Calaza, M., Endreffy, E., Pullmann, R., Ordi-Ros, J., Domenico Sebastiani, G., Ruzickova, S., Jose Santos, M., Papasteriades, C., Marchini, M.*, et al.* (2009). Replication of recently identified systemic lupus erythematosus genetic associations: a case-control study. Arthritis Res Ther *11*, R69.

Takahashi, K., Aranami, T., Endoh, M., Miyake, S., and Yamamura, T. (2004). The regulatory role of natural killer cells in multiple sclerosis. Brain *127*, 1917-1927.

Tanuma, N., Shima, H., Nakamura, K., and Kikuchi, K. (2001). Protein tyrosine phosphatase epsilonC selectively inhibits interleukin-6- and interleukin- 10-induced JAK-STAT signaling. Blood *98*, 3030-3034.

Taylor, K.E., and Mossman, K.L. (2013). Recent advances in understanding viral evasion of type I interferon. Immunology *138*, 190-197.

Team, R.D.C. (2008). R: A language and environment for statistical computing. (R Foundation for Statistical Computing).

Teige, I., Treschow, A., Teige, A., Mattsson, R., Navikas, V., Leanderson, T., Holmdahl, R., and Issazadeh-Navikas, S. (2003). IFN-beta gene deletion leads to augmented and chronic demyelinating experimental autoimmune encephalomyelitis. J Immunol *170*, 4776-4784.

Teijaro, J.R., Ng, C., Lee, A.M., Sullivan, B.M., Sheehan, K.C., Welch, M., Schreiber, R.D., de la Torre, J.C., and Oldstone, M.B. (2013). Persistent LCMV infection is controlled by blockade of type I interferon signaling. Science *340*, 207-211.

Terasaki, P.I., Park, M.S., Opelz, G., and Ting, A. (1976). Multiple sclerosis and high incidence of a B lymphocyte antigen. Science *193*, 1245-1247.

Thacker, S.G., Berthier, C.C., Mattinzoli, D., Rastaldi, M.P., Kretzler, M., and Kaplan, M.J. (2010). The detrimental effects of IFN-alpha on vasculogenesis in lupus are mediated by repression of IL-1 pathways: potential role in atherogenesis and renal vascular rarefaction. J Immunol *185*, 4457-4469.

Tigno-Aranjuez, J.T., Jaini, R., Tuohy, V.K., Lehmann, P.V., and Tary-Lehmann, M. (2009). Encephalitogenicity of complete Freund's adjuvant relative to CpG is linked to induction of Th17 cells. J Immunol *183*, 5654-5661.

Tomasson, M.H., Xiang, Z., Walgren, R., Zhao, Y., Kasai, Y., Miner, T., Ries, R.E., Lubman, O., Fremont, D.H., McLellan, M.D.*, et al.* (2008). Somatic mutations and germline sequence variants in the expressed tyrosine kinase genes of patients with de novo acute myeloid leukemia. Blood *111*, 4797-4808.

Traugott, U. (1985). Characterization and distribution of lymphocyte subpopulations in multiple sclerosis plaques versus autoimmune demyelinating lesions. Springer Semin Immunopathol *8*, 71-95.

Traugott, U., Reinherz, E.L., and Raine, C.S. (1983). Multiple sclerosis: distribution of T cell subsets within active chronic lesions. Science *219*, 308-310.

Trynka, G., Hunt, K.A., Bockett, N.A., Romanos, J., Mistry, V., Szperl, A., Bakker, S.F., Bardella, M.T., Bhaw-Rosun, L., Castillejo, G.*, et al.* (2011). Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. Nat Genet *43*, 1193-1201.

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.

Tzartos, J.S., Friese, M.A., Craner, M.J., Palace, J., Newcombe, J., Esiri, M.M., and Fugger, L. (2008). Interleukin-17 production in central nervous system-infiltrating T cells and glial cells is associated with active disease in multiple sclerosis. Am J Pathol *172*, 146-155.

Ungureanu, D., Wu, J., Pekkala, T., Niranjan, Y., Young, C., Jensen, O.N., Xu, C.F., Neubert, T.A., Skoda, R.C., Hubbard, S.R., and Silvennoinen, O. (2011). The pseudokinase domain of JAK2 is a dual-specificity protein kinase that negatively regulates cytokine signaling. Nat Struct Mol Biol *18*, 971-976.

Unkeless, J.C., Fleit, H., and Mellman, I.S. (1981). Structural Aspects and Heterogeneity of Immunoglobulin Fc Receptors. Adv Immunol *31*, 247-270.

van Boxel-Dezaire, A.H., Zula, J.A., Xu, Y., Ransohoff, R.M., Jacobberger, J.W., and Stark, G.R. (2010). Major differences in the responses of primary human leukocyte subsets to IFN-beta. J Immunol *185*, 5888-5899.

van de Werken, H.J., Landan, G., Holwerda, S.J., Hoichman, M., Klous, P., Chachik, R., Splinter, E., Valdes-Quezada, C., Oz, Y., Bouwman, B.A.*, et al.* (2012). Robust 4C-seq data analysis to screen for regulatory DNA interactions. Nat Methods *9*, 969-972.

van der Fits, L., van der Wel, L.I., Laman, J.D., Prens, E.P., and Verschuren, M.C. (2004). In psoriasis lesional skin the type I interferon signaling pathway is activated, whereas interferon-alpha sensitivity is unaltered. J Invest Dermatol *122*, 51-60.

van der Heijde, D., Sieper, J., Maksymowych, W.P., Dougados, M., Burgos-Vargas, R., Landewe, R., Rudwaleit, M., and Braun, J. (2011). 2010 Update of the international ASAS recommendations for the use of anti-TNF agents in patients with axial spondyloarthritis. Ann Rheum Dis *70*, 905-908.

van der Pouw Kraan, T.C., Wijbrandts, C.A., van Baarsen, L.G., Voskuyl, A.E., Rustenburg, F., Baggen, J.M., Ibrahim, S.M., Fero, M., Dijkmans, B.A., Tak, P.P., and Verweij, C.L. (2007). Rheumatoid arthritis subtypes identified by genomic profiling of peripheral blood cells: assignment of a type I interferon signature in a subpopulation of patients. Ann Rheum Dis *66*, 1008-1014.

Velazquez, L., Fellous, M., Stark, G.R., and Pellegrini, S. (1992). A protein tyrosine kinase in the interferon alpha/beta signaling pathway. Cell *70*, 313-322.

Velazquez, L., Mogensen, K.E., Barbieri, G., Fellous, M., Uze, G., and Pellegrini, S. (1995). Distinct domains of the protein tyrosine kinase tyk2 required for binding of interferon-alpha/beta and for signal transduction. J Biol Chem *270*, 3327-3334.

Vergo, S., Craner, M.J., Etzensperger, R., Attfield, K., Friese, M.A., Newcombe, J., Esiri, M., and Fugger, L. (2011). Acid-sensing ion channel 1 is involved in both axonal injury and demyelination in multiple sclerosis and its animal model. Brain *134*, 571-584.

Veyrieras, J.B., Kudaravalli, S., Kim, S.Y., Dermitzakis, E.T., Gilad, Y., Stephens, M., and Pritchard, J.K. (2008). High-resolution mapping of expression-QTLs yields insight into human gene regulation. PLoS Genet *4*, e1000214.

Viglietta, V., Baecher-Allan, C., Weiner, H.L., and Hafler, D.A. (2004). Loss of functional suppression by CD4+CD25+ regulatory T cells in patients with multiple sclerosis. J Exp Med *199*, 971-979.

Vo, N., and Goodman, R.H. (2001). CREB-binding protein and p300 in transcriptional regulation. J Biol Chem *276*, 13505-13508.

Wallace, C., Rotival, M., Cooper, J.D., Rice, C.M., Yang, J.H., McNeill, M., Smyth, D.J., Niblett, D., Cambien, F., Tiret, L.*, et al.* (2012). Statistical colocalization of monocyte gene expression and genetic risk variants for type 1 diabetes. Hum Mol Genet *21*, 2815-2824.

Wallace, C., Smyth, D.J., Maisuria-Armer, M., Walker, N.M., Todd, J.A., and Clayton, D.G. (2010). The imprinted DLK1-MEG3 gene region on chromosome 14q32.2 alters susceptibility to type 1 diabetes. Nat Genet *42*, 68-71.

Wallstrom, E., Khademi, M., Andersson, M., and Olsson, T. (2000). Increased numbers of mononuclear cells from blood and CSF expressing interferon-gamma mRNA in multiple sclerosis are from both the CD4+ and the CD8+ subsets. Eur J Neurol *7*, 71-76.

Wan, J., Fu, A.K., Ip, F.C., Ng, H.K., Hugon, J., Page, G., Wang, J.H., Lai, K.O., Wu, Z., and Ip, N.Y. (2010). Tyk2/STAT3 signaling mediates beta-amyloid-induced neuronal cell death: implications in Alzheimer's disease. J Neurosci *30*, 6873-6881.

Wang, F., Barrett, J.W., Shao, Q., Gao, X., Dekaban, G.A., and McFadden, G. (2009). Myxoma virus selectively disrupts type I interferon signaling in primary human fibroblasts by blocking the activation of the Janus kinase Tyk2. Virology *387*, 136-146.

Wang, W.Y., Barratt, B.J., Clayton, D.G., and Todd, J.A. (2005). Genome-wide association studies: theoretical and practical concerns. Nat Rev Genet *6*, 109-118.

Watford, W.T., and O'Shea, J.J. (2006). Human tyk2 kinase deficiency: another primary immunodeficiency syndrome. Immunity *25*, 695-697.

Westra, H.J., Peters, M.J., Esko, T., Yaghootkar, H., Schurmann, C., Kettunen, J., Christiansen, M.W., Fairfax, B.P., Schramm, K., Powell, J.E.*, et al.* (2013). Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat Genet.

Wildin, R.S., Ramsdell, F., Peake, J., Faravelli, F., Casanova, J.L., Buist, N., Levy-Lahad, E., Mazzella, M., Goulet, O., Perroni, L.*, et al.* (2001). X-linked neonatal diabetes mellitus, enteropathy and endocrinopathy syndrome is the human equivalent of mouse scurfy. Nat Genet *27*, 18-20.

Wilks, A.F., Harpur, A.G., Kurban, R.R., Ralph, S.J., Zurcher, G., and Ziemiecki, A. (1991). Two novel protein-tyrosine kinases, each with a second phosphotransferase-related catalytic domain, define a new class of protein kinase. Mol Cell Biol *11*, 2057-2065.

Willer, C.J., Dyment, D.A., Risch, N.J., Sadovnick, A.D., and Ebers, G.C. (2003). Twin concordance and sibling recurrence rates in multiple sclerosis. Proc Natl Acad Sci U S A *100*, 12877-12882.

Williams, N.K., Bamert, R.S., Patel, O., Wang, C., Walden, P.M., Wilks, A.F., Fantino, E., Rossjohn, J., and Lucet, I.S. (2009). Dissecting specificity in the Janus kinases: the structures of JAK-specific inhibitors complexed to the JAK1 and JAK2 protein tyrosine kinase domains. J Mol Biol *387*, 219-232.

Wilson, E.B., Yamada, D.H., Elsaesser, H., Herskovitz, J., Deng, J., Cheng, G., Aronow, B.J., Karp, C.L., and Brooks, D.G. (2013). Blockade of chronic type I interferon signaling to control persistent LCMV infection. Science *340*, 202-207.

Winchester, R., Ebers, G., Fu, S.M., Espinosa, L., Zabriskie, J., and Kunkel, H.G. (1975). B-cell alloantigen Ag 7a in multiple sclerosis. Lancet *2*, 814.

Wu, C., Orozco, C., Boyer, J., Leglise, M., Goodale, J., Batalov, S., Hodge, C.L., Haase, J., Janes, J., Huss, J.W., 3rd, and Su, A.I. (2009). BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. Genome Biol *10*, R130.

Wu, C., Yosef, N., Thalhamer, T., Zhu, C., Xiao, S., Kishi, Y., Regev, A., and Kuchroo, V.K. (2013). Induction of pathogenic TH17 cells by inducible salt-sensing kinase SGK1. Nature *496*, 513-517.

Yang, C.H., Murti, A., Valentine, W.J., Du, Z., and Pfeffer, L.M. (2005a). Interferon alpha activates NF-kappaB in JAK1-deficient cells through a TYK2-dependent pathway. J Biol Chem *280*, 25849-25853.

Yang, L., Froio, R.M., Sciuto, T.E., Dvorak, A.M., Alon, R., and Luscinskas, F.W. (2005b). ICAM-1 regulates neutrophil adhesion and transcellular migration of TNF-alpha-activated vascular endothelium under flow. Blood *106*, 584-592.

Yazici, C., Kose, K., Calis, M., Kuzuguden, S., and Kirnap, M. (2004). Protein oxidation status in patients with ankylosing spondylitis. Rheumatology (Oxford) *43*, 1235-1239.

Yeh, T.C., Dondi, E., Uze, G., and Pellegrini, S. (2000). A dual role for the kinase-like domain of the tyrosine kinase Tyk2 in interferon-alpha signaling. Proc Natl Acad Sci U S A *97*, 8991-8996.

Yoshida, H., Kimura, A., Fukaya, T., Sekiya, T., Morita, R., Shichita, T., Inoue, H., and Yoshimura, A. (2012). Low dose CP-690,550 (tofacitinib), a pan-JAK inhibitor, accelerates the onset of experimental autoimmune encephalomyelitis by potentiating Th17 differentiation. Biochem Biophys Res Commun *418*, 234-240.

Zheng, H., Qian, J., Baker, D.P., and Fuchs, S.Y. (2011). Tyrosine phosphorylation of protein kinase D2 mediates ligand-inducible elimination of the Type 1 interferon receptor. J Biol Chem *286*, 35733-35741.

Zhou, Y.J., Chen, M., Cusack, N.A., Kimmel, L.H., Magnuson, K.S., Boyd, J.G., Lin, W., Roberts, J.L., Lengi, A., Buckley, R.H.*, et al.* (2001). Unexpected effects of FERM domain mutations on catalytic activity of Jak3: structural implication for Janus kinases. Mol Cell *8*, 959-969.

Zivadinov, R., Weinstock-Guttman, B., Hashmi, K., Abdelrahman, N., Stosic, M., Dwyer, M., Hussein, S., Durfee, J., and Ramanathan, M. (2009). Smoking is associated with increased lesion volumes and brain atrophy in multiple sclerosis. Neurology *73*, 504-510.