Pathogenic mechanism of chronic inflammatory diseases : linking genotype, immune cell function and pathology in Spondyloarthritis

Surya Koturan

Université de Paris
Ecole doctorale BioSPC
ED 562 – Immunologie
Laboratoire d'Immunoregulation, Institut Pasteur

Pathogenic mechanism of chronic inflammatory diseases: Linking genotype, immune cell function and pathology in Spondyloarthritis

Par Surya KOTURAN

Thèse de doctorat d'Immunologie

Dirigée par Dr Elisabetta Bianchi

Présentée et soutenue publiquement à Paris le 25/11/2020

Devant un jury composé de :

Rapporteurs :	Pr Henri-Jean Garchon, Université de Versailles Dr Matteo Vecellio, University of Oxford
Examinateurs :	Dr Claire Vandiedonck, Université de Paris Dr Hussein Al-Mossawi, University of Oxford
Directeur de thèse : Co-encadrant de thèse:	Dr Elisabetta Bianchi, Institut Pasteur Dr Lars Rogge, Institut Pasteur

Abstract

Spondyloarthritis (SpA) is a family of related chronic inflammatory diseases (CID), and is characterised by inflammation of the spine and sacroiliac joints, the hallmarks of the axial form of this disease (AxSpA). Genome-wide association studies (GWAS) have identified genetic loci associated to SpA disease susceptibility. GWAS also have provided evidence of a key role for immune signalling pathways in the pathogenesis of SpA, as many of the identified loci map to immune genes. In particular, GWAS have suggested a role for the interleukin-23/interleukin-17 (IL-23/IL-17) axis in the pathogenesis of several CIDs. However, for most associated loci, the mechanism by which they affect pathogenesis and the immune cell populations in which they act are still not known. The goal of this project is to understand the molecular mechanisms of the disease pathogenesis by studying immune cell function and response to biologics in SpA.

To understand the role of susceptibility loci in CIDs we designed an nCounter® gene expression panel (NanoString Technologies): the Autoimmune Discovery Consortium Panel. This panel was created utilising the reported loci from GWAS in 9 CIDs, with the addition of several immune genes, such as cytokine and chemokine genes. We used this panel to determine the expression pattern of "GWAS genes" in immune cell populations isolated from 50 axial SpA (AxSpA) patients. The cell-type-specific gene expression profiles were then correlated with the patient genotype to identify expression quantitative trait loci (eQTLs). Gene expression analysis showed that ~80% of genes associated with CIDs were expressed in T cell subsets ($CD4^+$ and $CD8^+$, under resting and activated states), implicating their role in disease. The eQTL study revealed a genomic region on chromosome 11, including SNPs that affect the expression of cathepsin W (*CTSW*), a lysosomal peptidase implicated in cytotoxic activity in $CD8^+$ T cells. Furthermore, we demonstrated that transcription factor NFATC2 (Nuclear factor of activated T-cells, cytoplasmic 2) binds in an allele-specific manner to the eQTL rs12225345, supporting its role in *CTSW* regulation in T cells. These data illustrate how allele-specific binding of transcription factors could contribute to the regulation of disease-associated genes, and may play a role in the pathogenesis of the disease.

In the second part of this project, we studied the molecular mechanism of AxSpA pathogenesis and the impact of biologics (TNF inhibitors, TNFi) on immune responses of AxSpA patients.
To study the role for the IL-23/IL-17 axis in the pathogenesis of AxSpA, we have characterised the immune cells that produce IL-17A in AxSpA patients. We compared the IL-17 production capacity of cell subsets of the innate (MAIT, γδT, and neutrophils) and adaptive ($CD4^+$ and $CD8^+$ T-cells) arms of the immune system. We identified MAIT cells as the main producers of IL-17 in AxSpA. We also observed that both innate and adaptive lymphocytes express genes belonging to the IL-23/IL-17 pathway and genes previously associated with SpA susceptibility. To understand the mechanism of action of TNF blockade in axSpA patients, we investigated the effect of TNFi on immune responses to microbial and pathway-related stimuli. We demonstrated that anti-TNF therapy induced profound changes in patients' innate immune responses, but has minor effects on Th1/Th17 immunity. Additionally, we observed that TNFi affect the NFκB transcriptional network – an important regulatory network for innate immune response genes. We also reported that TNFi steers monocytes/macrophages towards an M2-like profile, which may be an important factor in the regulation of inflammatory responses.

Keywords: chronic inflammatory diseases, spondyloarthritis, innate and adaptive immune cells, cell-type-specific gene expression, eQTLs

Résumé

La spondylarthrite (SpA) est une famille de maladies inflammatoires chroniques apparentées (CID), et se caractérise par une inflammation de la colonne vertébrale et des articulations sacro-iliaques, caractéristiques de la forme axiale de cette maladie (AxSpA). Des études d'association pangénomique (GWAS) ont identifié des locus génétiques associés à la susceptibilité à la maladie de la SpA. GWAS a également fourni la preuve d'un rôle clé pour les voies de signalisation immunitaire dans la pathogenèse de la SpA, car de nombreux loci identifiés sont associés à des gènes immunitaires. En particulier, les GWAS ont suggéré un rôle pour l'axe interleukine-23 / interleukine-17 (IL-23 / IL-17) dans la pathogenèse de plusieurs CID. Cependant, pour la plupart des loci associés, le mécanisme par lequel ils affectent la pathogenèse et les populations de cellules immunitaires dans lesquelles ils agissent ne sont toujours pas connus. L'objectif de ce projet est de comprendre les mécanismes moléculaires de la pathogenèse de la maladie en étudiant la fonction des cellules immunitaires et la réponse aux traitement biologique dans la SpA.

Pour comprendre le rôle des loci de susceptibilité dans les CID, nous avons conçu un panel d'expression génique nCounter® (NanoString Technologies): le panel Autoimmune Discovery Consortium. Ce panel a été créé en utilisant les loci rapportés de GWAS dans 9 CID, avec l'ajout de plusieurs gènes immunitaires, tels que les gènes de cytokine et de chimiokine. Nous avons utilisé ce panel pour déterminer le modèle d'expression des «gènes GWAS» dans les populations de cellules immunitaires isolées de 50 patients SpA axiaux (AxSpA). Les profils d'expression génique spécifiques au type de cellule ont ensuite été corrélés avec le génotype du patient pour identifier les locus de caractères quantitatifs d'expression (eQTL). L'analyse de l'expression génique a montré qu'environ 80% des gènes associés aux CID étaient exprimés dans des sous-ensembles de cellules T ($CD4^+$ et $CD8^+$, à l'état de repos et activé), ce qui implique leur rôle dans la maladie. L'étude eQTL a révélé un cluster sur le chromosome 11, y compris des SNP qui affectent l'expression de la *CTSW* (cathepsin W), une peptidase lysosomale impliquée dans l'activité cytotoxique des cellules T $CD8^+$. En outre, nous avons démontré que le facteur de transcription NFATC2 (Nuclear factor of activated T-cells, cytoplasmic 2) se lie de manière allélique spécifique à l'eQTL rs12225345, soutenant son rôle dans la régulation CTSW dans les lymphocytes T. Ces données illustrent comment la liaison allélique spécifique des facteurs de transcription pourrait contribuer à la régulation des gènes associés à la maladie et jouer un rôle dans la pathogenèse de la maladie.

Dans la deuxième partie de ce projet, nous avons étudié le mécanisme moléculaire de la pathogenèse AxSpA et l'impact des traitement biologiques (inhibiteurs du TNF, TNFi) sur les réponses immunitaires des patients axSpA. Pour étudier le rôle de l'axe IL-23 / IL-17 dans la pathogenèse de AxSpA, nous avons caractérisé les cellules immunitaires qui produisent l'IL-17A chez les patients AxSpA. Nous avons comparé la capacité de production d'IL-17 de sous-ensembles cellulaires des bras innés (MAIT, γδT et polynucléaires neutrophiles) et adaptatifs (cellules T $CD4^+$ et $CD8^+$) du système immunitaire. Nous avons identifié les cellules MAIT comme les principaux producteurs d'IL-17 dans AxSpA. Nous avons également observé que les lymphocytes innés et adaptatifs expriment des gènes appartenant à la voie IL-23 / IL-17 et des gènes précédemment associés à la sensibilité SpA. Pour comprendre le mécanisme d'action du blocage du TNF chez les patients AxSpA, nous avons étudié l'effet du TNFi sur les réponses immunitaires aux stimuli microbiens et liés à la voie. Nous avons démontré que la thérapie anti-TNF induisait des changements profonds dans la réponse immunitaire innée des patients, mais avait des effets mineurs sur l'immunité Th1 / Th17. De plus, nous avons observé que le TNFi affecte le réseau transcriptionnel du NFκB - un réseau de régulation important pour les gènes de réponse immunitaire innée. Nous avons également signalé que le TNFi oriente les monocytes / macrophages vers un profil de type M2, ce qui peut être un facteur important dans la régulation des réponses inflammatoires.

Mot clefs : maladies chroniques inflammatoire, spondyloarthrites, cellules immunitaires innées et adaptatives, expression génetique, eQTLs

TABLE OF CONTENTS

To my parents, Anitha & Saji

Acknowledgements

I would like to express my sincere gratitude to my supervisors Dr Elisabetta Bianchi and Dr Lars Rogge for their support over the last four years. Lars, thanks for giving me a chance to come work in your lab despite the initial hurdles with funding and bureaucracy. Elisabetta, I appreciate your tenacity and genuine interest in science. Your guidance has been invaluable throughout this project and in writing the manuscript. Thank you for pushing me to do better and for being patient with me.

I am grateful to the Milieu Intérieur LabEx consortium for funding my PhD project over the last 3.5 years, and also to the department of Immunology at IP for the additional financial support.

I would like to thank the members of my jury for taking the time to read and evaluate this work. Thanks to Dr Henri-Jean Garchon and Dr Matteo Vecellio for accepting to be my rapporteurs. I appreciate the time you dedicated to read and correct my thesis thoroughly and for your comments and suggestions. Thanks to Dr Claire Vandiedonck and Dr Hussein Al-Mossawi for being part of my jury and the time you dedicated to discuss my work.

I thank all the present and past lab members of Team Imreg for integrating me into the lab when I was new and for showing me the ropes. Thanks to Dr Hanane Yahia-Cherbal, who has been instrumental in this project, especially in the final year of my thesis. Your encouragement during this project, both scientifically and emotionally means a lot to me. I have learnt a lot from you and I will miss working with you, the fun lunch breaks and me trying to help you cope with your Coca-Cola addiction! I thank Nicolas Rosine for being an amazing lab partner from the beginning. I will not forget the time we worked together in the CIH: the many hours spent in the cell sorting facility with Charlie Brown (cell sorter), late night “McDo dinners” in the lab, listening to me complain about the failed experiments and French administration, and good banter. Your advice and support have been very important to me and for this project! Thanks to Claire Leloup for your support and for all your help in the lab, with the cytometry experiments and for teaching me about important parts of the French food and culture: caramel au beurre salé, Foire aux vins et Beaujolais nouveau! Thanks to Dr Silvia Menegatti who has been a vital part of the SpA cohort recruitment and the anti-TNF study in our lab. Thanks Tharshana, for helping with the last few experiments and your help from the CB UTechS days- I wish you a lot of success in your PhD! Special thanks to Marie-Luce Kop for all her help since my first week at Pasteur. Marie-Luce, I admire your grit to power through bureaucracy to get matters done in time.

Thanks to many members of the ImmunoPasteur department for helpful discussions and fun moments during departmental retreats, seminars and happy hours B.C. (Before Covid19). Merci à tous pour vos amitiés, votre gentillesse et soutien moral chaque jour!

I am grateful to Pascal Campagne, the bioinformatician who worked on a very important part of this project. I appreciate all your help, the time you took out of your busy schedule to explain and answer my questions! I also thank Vincent Guillemot for his tips and tricks in statistics and R. Vincent, your R-tutorials during lockdown have been very interesting and I will keep practicing.

Special thanks to our collaborators at the Rheumatological service at Hôpital Cochin: Dr Maxime Dougados, Dr Corinne Miceli-Richard and Nathalie Menagé for organising the patient recruitment. Special thanks to Corinne for your feedback and answering my rheumatology-related questions whenever I came to you. I would also like to thank all the patients and donors who participated in this study and without whom this work would not have been possible.

I have been lucky enough to meet some incredible people along the way, and make good friends for life. I thank my amazing group of friends for always being there for me, despite the distance and time zone difference. You have always believed in me and have cheered me on through thick and thin. Thank you Gayathri, Sonali and Shanthy for being my constant sounding board. Thanks to HMG2015 crew: Jen, Ben, Annique and Lance- you inspire me, motivate me and always push me to keep going. Mariano and Anna, thank you for listening, our apéros and conversations over whisky have helped me a lot, especially in the past couple of months. Big thanks to the rest of my friends at Pasteur- Laura, Livia, Maria, Adria, Crispin & Borja! I am so grateful that I got to know you all. I have learnt so much from each and every one of you, and I will cherish the memories I have with you. Gracias…grazie…obrigada.

Thomas, ondanks dat jij graag de wijde wereld in trekt, ben ik blij dat je een paar jaar geleden in Parijs bent geland. Je bent een ongelooflijke steun geweest tijdens mijn doctoraat en het coronavirus lock down. Bedankt dat je de afgelopen zes maanden zo begripvol en geduldig met me bent geweest. Ik ben blij dat onze wegen elkaar kruisten en jij deel uitmaakt van mijn leven. Hartelijk dank aan je familie in Antwerpen voor hun vriendelijke woorden en steun, en jij ook Vincent!

Last but not least, the most important acknowledgements go to my loving parents. Amma and Acha, thank you for always supporting me and believing in me. I wouldn't be here if it weren't for you and the sacrifices you made. I am forever grateful for everything.

An immunologist and a cardiologist are kidnapped.
The kidnappers threaten to shoot one of them, but promise to spare whoever has made the greater contribution to humanity. The cardiologist says, "Well, I've identified drugs that have saved the lives of millions of people."
Impressed, the kidnappers turn to the immunologist. "What have you done?" they ask.
The immunologist says, "The thing is, the immune system is very complicated …" And the cardiologist says, "Just shoot me now."

Ed Yong (The Atlantic, August 5, 2020)

INTRODUCTION

Chapter 1

CHRONIC INFLAMMATORY DISEASES

1.1 Chronic inflammatory diseases (CID): An overview

Chronic inflammatory diseases (CIDs) are disorders that feature dysregulation of normal immune response resulting in persistent inflammation. During CIDs the immune system either overreacts (autoinflammation) or fails to distinguish between self from non-self (autoimmunity), or it induces an exaggerated response to antigens (allergies) (Janeway et al., 2001). CIDs encompass a spectrum of chronic and polygenic immune disorders featuring inflammation in the target organs and tissues- blood, bones, skin, GI tract etc., as depicted in **Figure 1**. The prevalence rate of CIDs in the Western population is estimated to be 5% - 7% (El-Gabalawy, Guenther & Bernstein, 2010). Some examples of CIDs are systemic lupus erythematosus (SLE), rheumatoid arthritis (RA), inflammatory bowel disease (IBD), psoriasis (Pso), type 1 diabetes (T1D), ankylosing spondylitis (AS) and multiple sclerosis (MS). Clinically, the pathophysiology of these diseases are distinct. However, immune dysregulation through imbalance in cytokine secretion resulting in prolonged inflammation is a shared signature in their pathology (Kuek, Hazleman, & Östör, 2007).

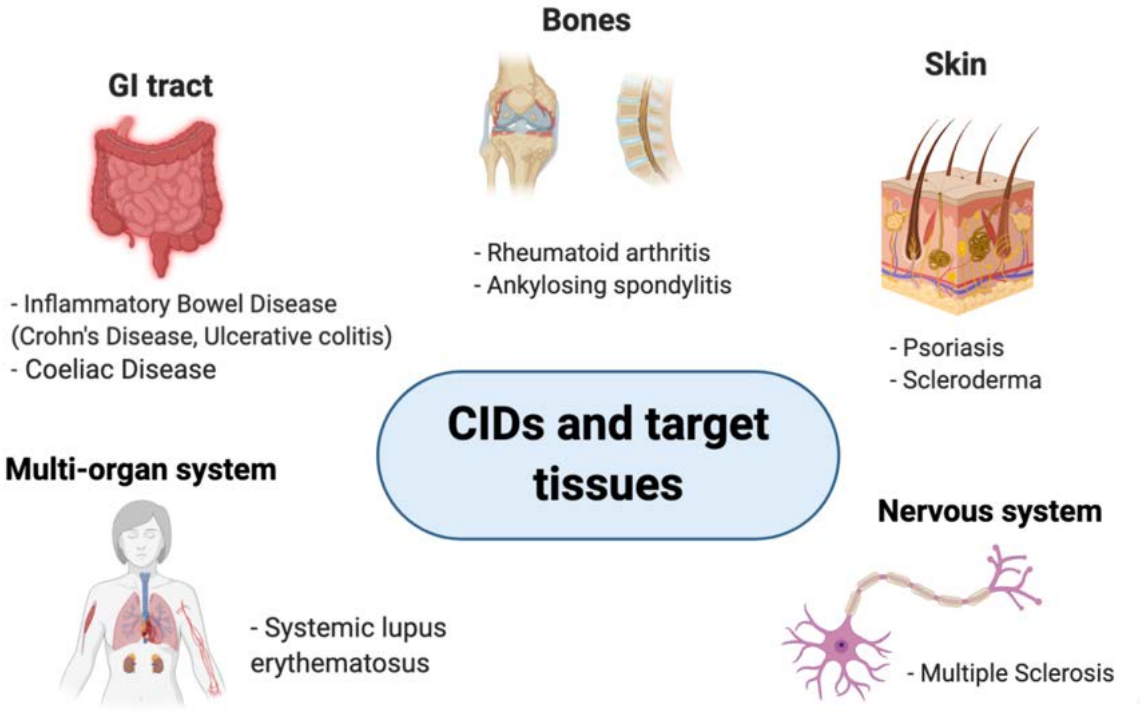


Figure 1: Chronic inflammatory diseases and target tissues and organs. CIDs are a spectrum of chronic and complex inflammatory disorders with localised inflammation and other systemic effects.

Epidemiologically, the incidence rate of immune disorders has been on the rise since the 1950s (Bach, 2002). Around 1 in 30 individuals have a type of inflammatory or autoimmune disease, thus making them a major global health problem (Wandstrat & Wakeland, 2001). Numerous factors can be attributed to the rise in CIDs, namely, environment, interaction between genetics and environment, lifestyle and socio-economic factors **(Figure 2)**. These factors alone and/or combined can perturb interactive biological networks that maintain immune homeostasis. Due to the multifactorial nature of these diseases, it is particularly challenging to understand their pathophysiology.

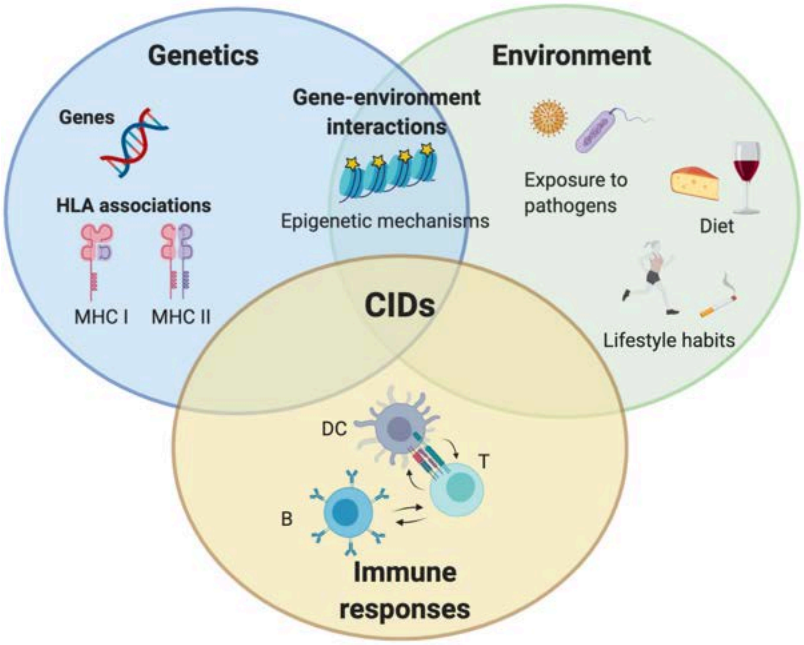


Figure 2: Contributing factors to chronic inflammatory diseases. CIDs can be triggered as a result of host genetics, environmental and lifestyle elements, combined gene-environment interactions and the resulting immune regulation.

1.2 Genetic susceptibility to CIDs

The development of CIDs has a strong hereditary component. Heritability (h^2) is defined as the proportion of total phenotypic variation for a given trait (V_P) in a population that can be attributed to the observed genetic variance (V_G) among individuals. It explains how much of the phenotypic variation is due to the genetic factors. h^2 is the proportion of genetic variance from additive genetic factors, $h^2 = V_A / V_P$, where V_A is the genetic variance explained by the sum total of the genetic effects and V_P is the observed phenotypic variance (Wray &

Visscher, 2008). The heritability in genetic diseases is classically assessed using family based studies to calculate sibling recurrence risk and twin concordance rates.

The sibling recurrence risk (λ_s) is the ratio of the disease risk in siblings of affected individuals to the disease incidence in the general population, whereas twin concordance rate refers to the proportion of siblings of the affected twins that also have the disease. Most CIDs have a sibling recurrence risk score between 6 and 20 and concordance rates between 25% - 50% in monozygotic (MZ) twins and 2% - 12% in dizygotic (DZ) twins (Goris & Liston, 2012). For example, the computed relative risk in RA from λ_s is 5 – 10, MZ concordance rate was 12% - 15% and the DZ concordance rate was between 2% - 3% (Wandstrat & Wakeland, 2001). Additionally, co-occurrence of multiple CIDs in individuals along with clustering of these diseases in families indicate the strong genetic influences that contribute to susceptibility to these diseases.

1.2.1 Genetic studies in CIDs: Pre-GWAS era

Before the advent of sequencing technologies and big data analysis in genomics, genes associated with disease predisposition were identified through candidate-gene approaches and family-based linkage studies. However, for CIDs, these approaches did not yield much progress in identifying novel risk loci. Candidate gene studies required prior knowledge of the disease mechanism to select genes and to test genetic variants for association with the disease. With limited understanding of the disease biology, identifying new candidate genes was difficult. Additional drawbacks included testing one or few polymorphisms at a time and the lack of statistical power to detect significant associations. Despite these limitations, candidate-gene studies were still able to identify some prominent susceptibility genes in immune diseases, like *CTLA4* (cytotoxic T-lymphocyte-associated protein 4) associated with Grave's disease (GD) and T1D (Ricaño-Ponce & Wijmenga, 2013). *CTLA4* association to GD and T1D were replicated later on by several association studies and *CTLA4* is now a well-established locus for other CIDs like RA (Gregersen et al., 2009) and Coeliac's disease (CeD) (Trynka et al., 2011).

Unlike candidate-gene studies, linkage analysis is conducted in a hypothesis-free manner without requiring prior knowledge of the disease pathology. Family-based linkage studies work by a transmission model to explain the inheritance pattern of a phenotype (a complex trait, e.g. an immune disease like IBD) and the genotype observed in a pedigree (Lander & Schork, 1994). This approach was successful in monogenic diseases where linkage can be established within a single pedigree. Due to the polygenic nature of complex immune

diseases, linkage analysis had limited success in identifying disease-associated genes. However, using this method Hugot et al. mapped variants in the *NOD2,* a gene associated with CD susceptibility (Hugot et al., 1996). *NOD2* was later identified as a causal gene for IBD by the same team (Hugot et al., 2001). Genetic approaches to study immune diseases, prior to GWAS, identified associations of several CIDs with the human leukocyte antigen (HLA) alleles within the major histocompatibility complex (MHC) on chromosome 6 (Kumar, Wijmenga & Xavier, 2014).

1.2.2 Genetic studies in CIDs: GWAS era

GWAS is a large-scale analysis of genetic variants across the human genome that is used to study the relationship between single nucleotide polymorphisms (SNPs) with a phenotypic trait within a population (Hardy & Singleton, 2009). Conceptually, in the first stage of case-control GWAS, cohorts of ethnically matched patients and healthy controls are recruited (ideally). In the second stage, genomic DNA isolated from the individuals is hybridised on SNP arrays with 300,000 to millions of markers with fluorescently labelled alleles. In the third stage, the fluorescent signals corresponding to the alleles of SNPs are used to calculate the allele frequencies in the cases and in the controls. Variants associated with the disease will be found at a higher frequency in cases than in controls as shown in **Figure 3**. A statistical test is carried out to assess how likely a variant is to be associated with a trait and the results are represented in p-values. The p-value in GWAS indicates the probability that the allele is likely to be associated with the trait (Bush & Moore, 2012). The results can be represented as a “Manhattan plot” as seen in **Figure 3**. The association is also represented as odds ratios (OR), which is the ratio between the odds of individuals having the disease associated with a specific allele and the odds of individuals who have the disease but do not carry the same associated allele.

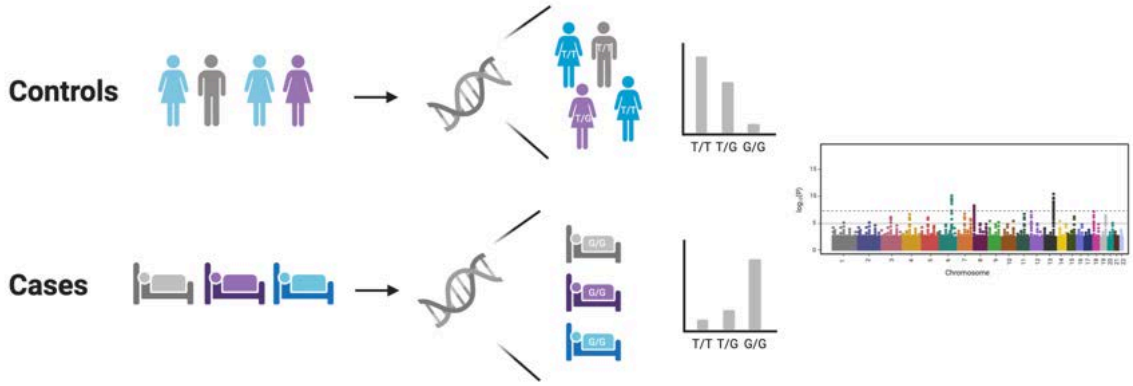


Figure 3: Genome-wide association study methodology. GWAS surveys the entire genome for SNPs that occur more frequently in individuals with the disease (cases) compared to the healthy individuals (controls). When the frequency of a SNP between cases and control is significantly different, that SNP is associated with disease risk.

Unlike family-based linkage studies, GWAS is done in a hypothesis-free manner without the bias from prior knowledge about the disease pathology. Therefore, through GWAS, several novel loci can be identified that may give new insights into disease pathogenesis. Association studies have been successful in identifying hundreds of loci linked to immune disease susceptibility owing to the linkage disequilibrium (LD) structure in the genome. In multifactorial diseases, GWAS is effective in identifying common variants with modest effect sizes, i.e. a modest proportion of individuals in a population carry the allele associated with a phenotype, which makes the variant moderately penetrant (Lobo, 2008). However, these variants with modest effect sizes do not fully account for the observed disease susceptibility. The discrepancy may be due to the rare variants with minor allele frequency (MAF) <1% and low to modest effect sizes, that are challenging to detect using GWAS (McCarthy et al., 2008). Due to this drawback the effect of rare variants is often discounted in association studies. Other limitations of GWAS include population stratification bias and sample size bias resulting in false-positive and false-negative results. Moreover, overcoming the burden of multiple testing requires a large sample size, in order to reach the genome-wide significance threshold (p-value < 10^{-8}), and rigorous replication studies are necessary to validate the findings (Flint, 2013). However, the advantages of GWAS outweigh the disadvantages especially in measuring the disease risk in large case-control cohorts of polygenic disorders.

Through immune-related GWAS, common variants in the non-MHC region with odds ratios (OR) ranging between 1.04 and 3.99 were identified (Ricaño-Ponce & Wijmenga, 2013). GWAS and replication studies conducted in individual CIDs have successfully identified and replicated numerous loci, such as 137 loci in IBD (De Lange et al., 2017), 56 loci in Pso (Tsoi et al., 2017), and 52 loci in MS (Beecham et al., 2013). A combined meta-analysis study in

CIDs (ankylosing spondylitis, Crohn's disease, psoriasis, primary sclerosing cholangitis and ulcerative colitis), collectively revealed a total of 156 loci including 27 new associations (Ellinghaus et al., 2016). There are a plethora of public data repositories that maintain GWAS summary statistics and data from published studies worldwide- for e.g., GWAS catalog, Immunobase consortium and Open Targets Genetics.

However, most GWAS associations explain less than 50% of the heritability of a disease, and the rest remains unaccounted for. For complex diseases, heritability assessed through GWAS is small and the remaining unexplained heritability is termed as "missing heritability". Rare variants in the genome that are missed during GWAS may be a source for the missing heritability in autoimmune diseases. However, exon sequencing studies of GWAS loci have revealed that rare exonic variants have minor impact on immune disease risk (Hunt et al., 2013). Another source could be rare variants in the intronic region or from the gene-gene interactions and epigenetic regulations in polygenic diseases (Ye, Gillespie & Rodriguez, 2018). A plausible reason for missing out variants in GWAS could be due to the population bias from SNP-genotyping arrays. Genotyping arrays were often designed with polymorphisms that occur in the individuals of European ancestry, under-representing the variants from non-European populations. (Bush & Moore, 2012). This urges the need to extend the genome-wide investigations in selected cohorts with extensive population data and disease cohorts of appropriate sample sizes. A recent initiative from the UK Biobank aims to achieve these goals with 500,000 participants of the British population with diverse genetic ancestry between the ages of 40-69. Using the available genetic data, association studies for over 4000 different phenotypes were conducted (Bycroft et al., 2017). Furthermore, a large phenotype-wide association study (PheWAS) on the UK Biobank data under the leadership of Benjamin Neale's team was conducted with two rounds of results. These results were made publicly available to enable replication studies in independent disease cohorts and have been a useful resource for complex disease genetics research (http://www.nealelab.is/uk-biobank/).

1.2.3 Shared loci and shared pathogenesis in CIDs: findings from GWAS

The existence of shared pathogenetic factors among CIDs has been well-established from the co-occurrence pattern of the diseases, HLA haplotypes and through GWAS findings. Co-occurrence of chronic inflammatory immune disorders has been reported in patients with one form of CIDs- for example, occurrence of CeD in patients after the onset of T1D has been previously reported (Barera et al., 2002). Similarly, patients with AS, a prominent form of SpA have shown to exhibit extra-articular manifestations such as Pso, IBD and anterior uveitis (El Maghraoui, 2011).

The knowledge from genetic and functional studies have elucidated the role of the immune system in CIDs and emphasizes how dysregulation of the immune system and disrupted homeostasis can lead to immune disorders. In the interest of achieving replication of the association studies and the fine-mapping of the GWAS loci, a SNP microarray called "Immunochip" was developed by a collaborative effort led by the Wellcome Trust Case-Control Consortium (Cortes & Brown, 2011). The chip was designed exclusively for 11 autoimmune and inflammatory diseases from experts of RA, AS, SLE, T1D, autoimmune thyroid disease, CeD, MS, ulcerative colitis (UC), Crohn's disease (CD), and Pso.

A highlight of Immunochip studies is that the many reported genetic loci are shared among several diseases, as shown, as an example, in **Figure 4** for loci identified in chromosome 6. Almost all of the CIDs share the genetic association to the HLA locus. There are also hundreds of non-MHC loci identified in this manner that demonstrate shared associations to several diseases, which suggested similar effector pathways in regulating the immune responses observed in autoimmune and inflammatory diseases. Although the concomitance of these CIDs in patients and the shared genetic loci among the co-occurring diseases give an insight into the shared immune pathways at play, the disease mechanism and biology still remain incompletely understood.

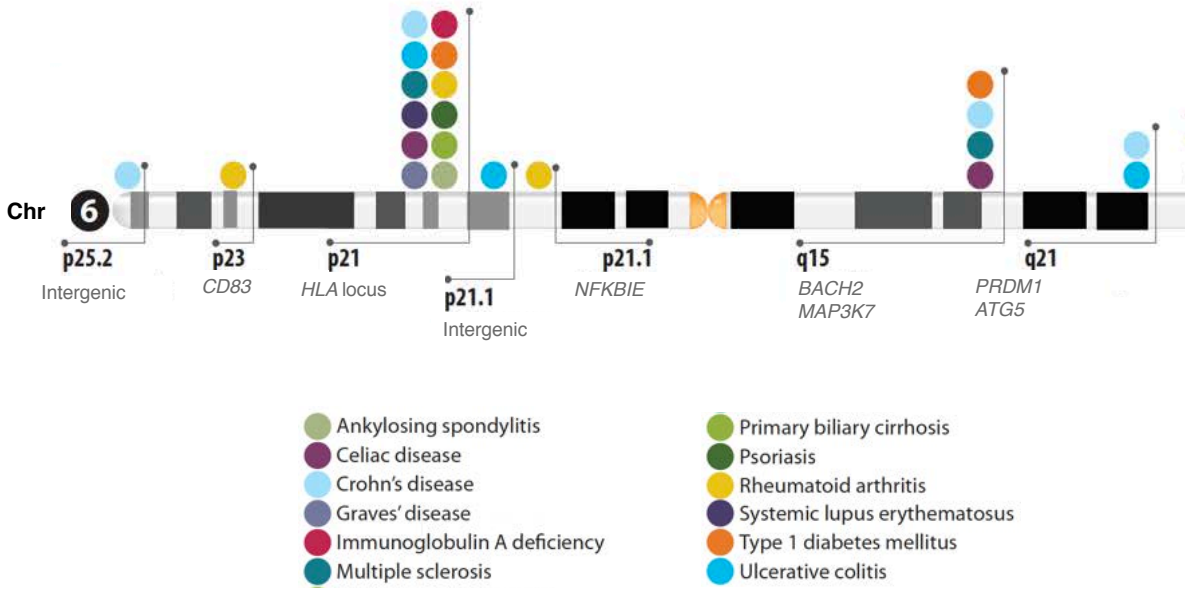


Figure 4: Loci associated with chronic inflammatory diseases on chromosome 6. Susceptibility loci reported on chromosome 6 associated to different immune disorders identified through GWAS. The figure is adapted from Ricaño-Ponce & Wijmenga, 2013.

Therefore, the shared genetic background observed in CIDs may be important in providing information about the disease mechanism and the underlying biology. However due to the limited resolution of the GWAS approach, it is difficult to determine whether the shared genetic variants are true associations or if they are detected as a result of the close proximity of the SNPs in the locus.

Shared genetic loci from Immunochip studies have highlighted the pleiotropic nature of the disease-associated genes and SNPs. The overlap in the shared loci can chiefly indicate three things: (i.) the SNP is correlated and concordant- i.e., the same SNP confers disease risk for more than one disease, (ii.) the SNP is correlated and discordant- the SNP can be associated with increased risk for one disease and has a protective effect for another disease and (iii.) the SNP is not correlated- i.e., the shared locus shows different haplotypes (Parkes et al., 2013). By careful analysis of GWAS loci in different chronic inflammatory and autoimmune diseases, Zhernakova et al. assessed the common immune pathways that are detected in the Immunochip consortium in a comprehensive review (Zhernakova et al., 2009). (**Table 1**). Similarly, another article analysed the overlap in the shared loci detected from Immunochip studies for 6 immune diseases- AS, CeD, IBD, Pso, RA and T1D (Parkes et al., 2013a). The findings from both articles identified shared disease-associated genes and the immune pathways they are involved in antigen presentation, T-cell differentiation, the interleukin-23 pathway, transcription factors that are essential for immune cell function and differentiation (NF-κB), and other innate immune response pathways. The findings from these combined analyses highlight the potential role of the shared loci in disease pathogenesis by modulating immune signalling as summarised in **Table 1**. The authors also analysed genes that are unique to one disease, such as genes of chemokines and cytokines that could point to new pathways that are relevant for disease biology (Zhernakova et al., 2009; Parkes et al., 2013).

Pathway/genes	Shared candidate genes	Associated diseases
Aminopeptidase	*ERAP1, ERAP2*	AS, Pso, IBD
IL-2 and IL-21	*IL2, IL21, IL2RA, IL2RB*	IBD, CeD, RA, T1D & MS
IL-23 and T_h1	*IL23R, IL12B, IL12A, TYK2, JAK2, STAT3,STAT4, IL27 & CCR6*	Pso, RA, MS & IBD
Tyrosine phosphatases	*PTPN2, PTPN22*	IBD, CeD, RA, T1D & SLE
Transcription factors	*NFKB1*	MS, Pso, IBD
	IRF4	Pso, CeD
	TNFAIP3	IBD, CeD, Pso, RA, T1D & SLE
Cytokines	*IL10*	IBD, T1D & SLE
	IL18RAP	IBD, CeD & T1D
Others	*BACH2*	AS, IBD, CeD, T1D & MS
	CARD9	AS and IBD
	TAGAP	IBD, CeD, Pso, RA, MS &T1D
	ZMIZ1	IBD, Pso, CeD & MS
	PTGER4	AS, IBD & MS

Table 1: Some genes and pathways that are associated with two or more CIDs. Examples of some shared candidate genes and the implicated pathways based on the analysis conducted by Zhernakova et al., 2009 and Parkes at al., 2013. The lead positional candidate gene is mentioned, as well as the pathway linked and the shared associated diseases

1.2.4 GWAS loci affect the function of immune signalling pathway genes

The MHC loci were the first reported and strongest genetic association with most CIDs. The MHC region is very gene-dense with extensive LD pattern and hundreds of immune-related genes. Diseases exhibiting classical autoimmune phenotypes such as SLE, T1D, CeD etc. are associated with HLA class II alleles, whereas diseases with prominent autoinflammatory phenotypes are generally associated with specific HLA class I alleles, for example HLA-B27 for AS, HLA-Cw6 for Pso and HLA-B51 with Behçet's disease. Further analysis of MHC associations to CIDs has been challenging due to the strength of these associations and the tight LD structure in this locus. However with dense SNP-typing and HLA-imputation methods several independent MHC loci have been reported in CIDs (Parkes et al., 2013).

Immune-GWAS have revealed a number of additional genes that are vital for T-cell function, mainly for differentiation of peripheral T-cells into subsets of T_h1 cells, T_h17 cells and regulatory T cells (T_{reg} cells). Identification of susceptibility genes like *IL10*, *IL12*, *IL18RAP*, and *STAT4* have suggested the involvement of T_h1 cells in disease pathology (Moss et al., 2004; Raphael, Nalawade, Eagar, & Forsthuber, 2015). Genes associated with the function of T_h17 cells, like STAT3, *IL23R* and *IL21*, are also susceptibility genes for CD, RA and Pso, suggesting their mechanistic role in disease (Patel & Kuchroo, 2015). Consistently, the sites of inflammation in these diseases, have been found infiltrated with T_h17 cells that mediate the inflammatory responses (Maddur et al., 2012). Also, diminished T_{reg} activity observed in autoimmune diseases is possibly resulting from polymorphisms in genes like *IL2* and *IL2RA*, that regulate T_{reg} function (Brusko, Putnam & Bluestone, 2009). Disease susceptibility genes involved in T-cell activation, like *CTLA4*, *TAGAP*, and *ICOSLG* have linked the involvement of activated T-cells in the disease pathology of T1D and CeD. Hence, disrupted T-cell function due to genetic variants from the susceptibility loci could have a critical role in the pathogenetic mechanisms of some CIDs.

PTPN2 is a protein tyrosine phosphatase gene associated to T1D and CD. *PTPN2* may play a direct role in the destruction of beta cells in T1D and it is suggested to modulate innate immune responses in CD (Sharp et al., 2015). Shared adapter protein coding gene *SH2B3*, associated with T1D and CeD, has been shown to mediate T-cell receptor and TNF-alpha signalling (Zhernakova et al., 2009). In diseases like SLE, the identified susceptibility loci were genes in the B-cell signalling pathway like *TNFSF5,* which is crucial for B-cell activation and proliferation; and *BANK1* (B-cell scaffold protein with ankyrin repeats 1) that manages to the autoantibody production observed in lupus.

Additionally, the involvement of these disease-associated genes in several immune signalling pathways suggests the interplay between innate and adaptive immune system. For example, *NOD2*, a susceptibility gene for CD, is crucial for maintaining intestinal epithelial barrier function and to produce anti-microbial peptides to protect the epithelium. Genes associated to autophagy, such as *ATG16L1* (autophagy 16 related-like 1) and *IRGM* (immunity-related GTPase M), support the role of intracellular processing of bacterial particles in CD (Levine & Deretic, 2007). A cytokine gene associated to TNF signalling pathway, *TNFSF15,* is linked to IBD (CD & UC), and has been reported to control both T_h1 and T_h17 cellular functions (Takedatsu et al., 2008). *TNFAIP3* encodes for a ubiquitin-modifying enzyme, A20, which is vital in regulating TNF-induced NF-κB activation through TLRs. Therefore, polymorphisms in *TNFAIP3* could affect the production of functional A20, disrupting signalling downstream of TLR that can then lead to long-term inflammation (Zhernakova et

al., 2009). These associated genes are also suggestive of the role of pathogen exposure in the onset of certain CIDs. Therefore GWAS associated genes in immune disorders point to a strong genetic link between dysregulation in the immune signalling pathways and disease.

Chapter 2

MOVING FORWARD FROM GWAS TO CANDIDATE CAUSAL VARIANTS

2.1 Challenges in post-GWAS era

Association studies have paved the way to remarkable progress in studying the genetic architecture of complex diseases in an unbiased manner. Although GWAS have revolutionised the field of complex disease genetics, they do not immediately reveal a clear picture of the mechanisms that lead to disease. It is challenging to establish a functional link between GWAS loci and disease, as ~90% of the disease-associated SNPs are situated in the non-coding regions of the genome, such as enhancers and promoters (Hrdlickova et al., 2014).

GWAS associations denote susceptibility loci in the genome that convey either disease risk or protection. A susceptibility locus is defined as the region in the genome that contains the genetic variant associated to a disease. In current practice, the identified SNP is linked to the nearest gene, however this gene may not explain the disease mechanism and the underlying biology since the susceptibility loci often contain multiple genes. Therefore, there is a need to develop approaches to identify the causal gene, and to test the effects of the genetic variants on the function of the nearest genes.

An additional difficulty is given by the fact that the associated SNPs might have a context-specific activity or can only be of relevance in a specific cell type or tissue. Approaches to identify relevant cell types by studying their transcriptional profiles to detect the expression of disease-associated genes have shown promising results. One example is the suggested role of dendritic cells in CD based on the enrichment of IBD loci in these cells from Gene Ontology analysis (Jostins et al., 2012). However, many attempts to explore the functional role of susceptibility loci in disease mechanism are not done in the appropriate cell types, and it is therefore difficult to ascertain their biological relevance. The challenging aspects in establishing causal links from GWAS variants are depicted in **Figure 5**.

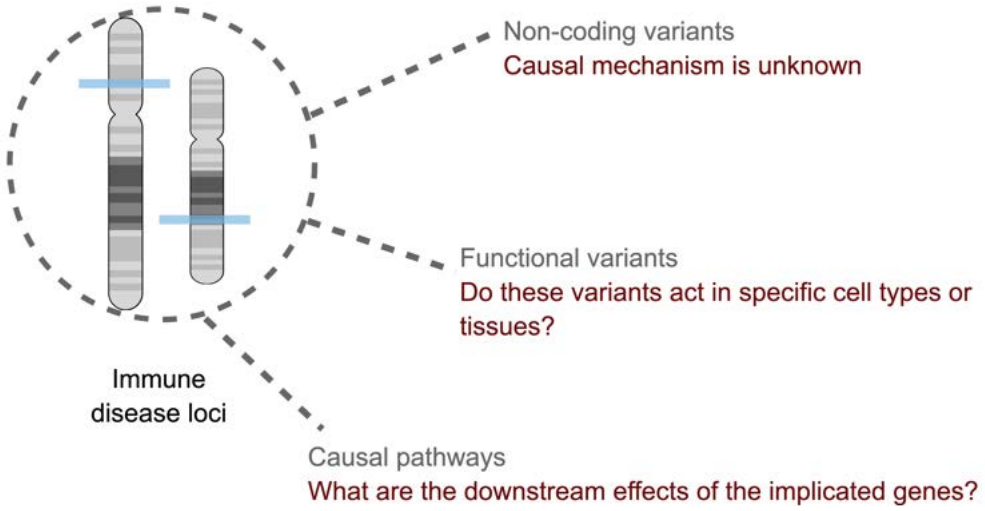


Figure 5: Challenges in translating GWAS variants to function. Modified from the seminar titled "Immune disease variants converge on regulation of CD4 T cell activation" at the CSHL Systems immunology meeting, Trynka G, 2019.

Linkage disequilibrium (LD) is a property of adjacent SNPs that determines the degree to which an allele of one SNP is inherited with an allele of another SNP in the same locus within a population (Bush & Moore, 2012). The high density genotyping arrays used for association studies utilise the LD pattern of the SNPs in the locus to select representative SNPs- termed "tag SNPs". In GWAS, the genotyped tag SNPs and LD pattern are used to extrapolate the disease association. However, LD patterns are population dependent. Therefore, tag SNPs selected for one population may not show the same inheritance pattern in another population. The tag SNP could directly influence the biological mechanism that give rise to the associated disease. In this scenario the tag SNP is referred to as the functional SNP. Alternatively, the genotyped tag SNP is not the functional SNP, but it may be in LD with another SNP in the locus that affects the phenotype (Hirschhorn & Daly, 2005). Hence it is crucial to recall that a significant SNP association from GWAS should not be assumed as being causal without precise functional studies. Methods to conduct functional studies on the GWAS loci are ongoing with the help of fine mapping, SNP prioritisation by combined genetic and bioinformatic approaches and by *in vitro* and *in vivo* experimental methods to verify the predicted mechanisms of likely causal genes (Edwards et al., 2013).

2.2 Fine-mapping efforts to link GWAS variants to causal implications

There are numerous pipelines that have been proposed to conduct the post-GWAS analysis for the functional assessment of non-coding variants. Most methodologies employ fine-mapping- which is a general term used for the refining process of the GWAS variants using statistical, bioinformatic and functional methods to sift out the SNPs that are less likely to be involved in a causal mechanism. The typical analysis pipeline for fine-mapping proceeds in the following steps: genotype imputation to enlist the associated variants, examining the LD structure of the genomic loci highlighted by the SNPs, shortlisting the independent SNP associations using regression models, merging the data from multiple studies (meta-analyses) to increase the resolution and finally, integrating SNP information from functional annotation database and gene expression data to assign causality. The outline of this general strategy to prioritise GWAS SNPs to determine regions of interest is described in the schematic in **Figure 6** (Schaid, Chen & Larson, 2018). A brief overview of the processes involved in fine-mapping is given below:

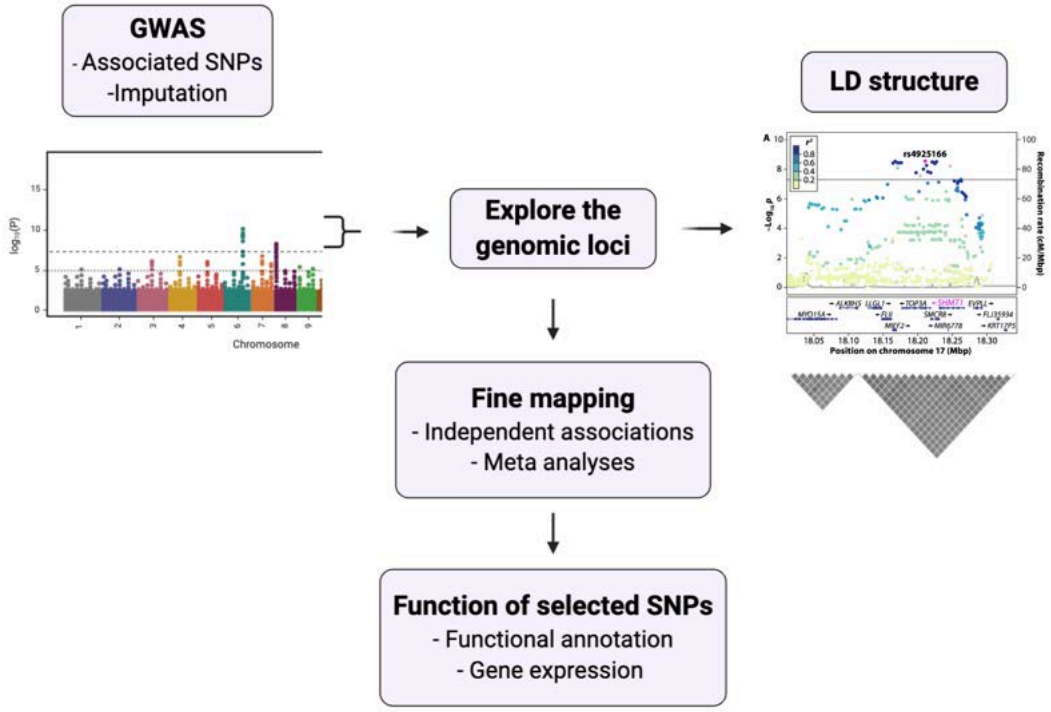


Figure 6: Post-GWAS analysis pipeline for fine-mapping associated variants. Genomic loci indicated by the associated variants from GWAS are examined further for LD patterns. If strong LD is established, then independent associations and replicated variants from multiple studies are shortlisted for more analysis using gene expression data to assess the likely function of the selected SNPs. (Modified from Schaid, Chen & Larson, 2018).

2.2.1 Genotype imputation

All the SNPs located in a genomic region are not included in the SNP genotyping arrays giving rise to intermittent missing genotypes. Imputation is a method used to predict the genotypes that are not directly genotyped. These predicted genotypes can then be used to boost the power of association studies and facilitate meta-analysis. Imputation also improves the resolution of an associated region and enables identification of a causal SNP (Marchini & Howie, 2010).

2.2.2 Identify independent associations

Multiple SNPs in a region can show marginal association to a particular trait while analysing one SNP at a time. The LD pattern among SNPs in a genomic locus can show several non-causal SNPs to be correlated with one causal SNP. When SNPs are analysed jointly, the local LD among the SNPs is accounted for and as a result only the causal variant should show association to the trait. However, if there are multiple causal SNPs and they show a correlation with each other and to several other non-causal SNPs, identifying the true source of the signal can be challenging. Hence in some fine-mapping analyses, a conditional approach is used to detect secondary association signals at a given locus. In this method, the effect from the primary SNP associated to a trait is adjusted to verify if the secondary signals show an association to the trait independent of the primary SNP (Yang et al., 2012).

2.2.3 Integrative analysis combining multi-cohort studies

Fine-mapping resolution can be improved by pooling individual-level data from multiple studies in complex diseases. (Pasaniuc & Price, 2017). Although it can be challenging to attain individual-level data from multi-cohort and multi-centric studies due to privacy concerns, summary statistics from GWAS can be used as they are publicly available for most studies. Summary association statistics report the estimated effect sizes and standard errors for each SNP analysed in a GWAS. Summary statistics from different GWAS conducted for a specific disease can be used for a meta-analysis, where association statistics of each study are analysed jointly to increase power to detect SNP associations of small effect sizes.

2.2.4 Integrating functional annotation

Fine-mapping accuracy can be improved by integrating functional annotations that assigns a biological function to sequence information. This approach is useful in assessing the likely functional role of SNPs and can help in the prioritisation process for follow-up functional

studies. There are a number of publicly available resources with large-scale information about genomic regions, protein coding and non-coding regions, like ENCODE, FANTOM5 and Roadmap Epigenomics Project (Pasaniuc & Price, 2017). Combining different layers of data from tissues and cell types along with the corresponding annotation gives an insight into the biological context in which certain SNPs act. This approach is widely accepted based on the background provided by early research showing that disease-associated genomic loci are enriched in regulatory marks in specific cell types, suggesting their potential function in pathogenesis (Schaub et al, 2012), as described in detail in chapter 3.

2.2.5 Integrating gene expression data with GWAS

The majority of the genetic variants identified through GWAS are located in the non-coding regions of the genome. However several GWAS variants are situated within regulatory regions with enhancers, promoters, silencers and insulators. Additionally, these trait-associated variants are more likely to be expression quantitative trait loci (eQTLs), which are genomic regions that influence the expression levels of a nearby or distant gene (Nicolae et al., 2010), suggesting that the altered gene expression affects the associated disease phenotype. Statistical approaches to combine eQTL and GWAS data facilitates the identification of susceptibility genes and potential causal pathways involved in disease mechanisms.

Research teams focused on translating GWAS associations to disease mechanisms have developed fine-mapping algorithms that aid in identification of candidate causal variants. For example, the Probabilistic Identification of Causal SNPs (PICS) algorithm was developed by Farh et al. to enhance the selection of loci associated to autoimmune and inflammatory disease variants. PICS computes the probability of a SNP being causal by utilising the haplotype structure and the association pattern observed at a locus (Farh et al., 2015). This algorithm used summary statistics from Immunochip studies and other GWAS data for 21 diseases. In this study the authors integrated fine-mapped variants from PICS with transcription and *cis*-regulatory element information (RNA-seq, ENCODE ChIP-seq, histone modification profiles and DNase hypersensitivity profiles) from primary immune cells. With this approach they reported that nearly 90% of the causal variants are in fact in the non-coding region and importantly, ~60% of them are situated in immune-cell enhancers.

2.3 eQTL studies unravel susceptibility genes in CIDs

As demonstrated by Farh et al. the majority of the disease-associated SNPs identified through GWAS are located in the non-coding regions of the genome, suggesting their role in gene regulation. Studying the effect of these SNPs on gene expression in a specific cell type or tissue could explain the regulatory mechanism of the genetic variants implicated in a disease. eQTL mapping is a method used to identify genetic variants that influence the expression levels of a gene. An eQTL is a locus with a SNP that is correlated with the variation in the expression levels of a gene (Gilad, Rifkin & Pritchard, 2008). A genomic locus with eQTLs may indicate the presence of gene regulatory elements like transcription factors binding sites (TFBS). With eQTL mapping, SNPs identified through GWAS can be prioritised by their regulatory effects on gene expression in a specific cell type or tissue, and this can then be used to investigate their downstream functional role in disease.

Conventionally, the eQTL analysis is proximity-based, where you define an arbitrary distance around the eQTL SNP to detect the affected genes (eGenes). If the eGene is located within a given window of genomic distance (e.g. <1 Mb) of the SNP, the local effect is referred to as a *cis*-eQTL (**Figure 7a**). On the contrary, if the eGene is located at a greater distance (e.g. >5 Mb or on a different chromosome), the observed effect on the gene is indirect, and is called a *trans*-eQTL effect (**Figure 7b**).

eQTL mapping is a powerful method for prioritisation of disease susceptibility genes from GWAS loci. *Cis*-eQTLs tend to have large effect sizes, and can therefore be detected in a modest sample size. *Cis*-eQTLs are often situated near the transcription start sites (TSS) of genes or within gene bodies (Brown, Mangravite, & Engelhardt, 2013). As a general observation, the effect size of the eQTL SNP increases with the decrease in the distance between the eQTL SNP and TSS. Due to their location near the TSS, *cis*-eQTLs have been shown to modify transcription factor binding sites (TFBS) or other *cis*-regulatory elements (CREs) like DNase I hypersensitivity sites (DHSs) (Westra & Franke, 2014). Unlike *cis*-eQTLs, *trans*-eQTLs show small effect sizes and therefore large samples sizes are required to detect them. *Trans*-eQTLs, however, provide meaningful insights into disease pathogenesis by identifying multiple eGenes and gene networks that may be involved in the disease mechanism. Due to the sample size limitation in most eQTL analyses, to perform a *trans*-eQTL analysis one must rely on meta-analysis to increase the sample size and to boost statistical power.

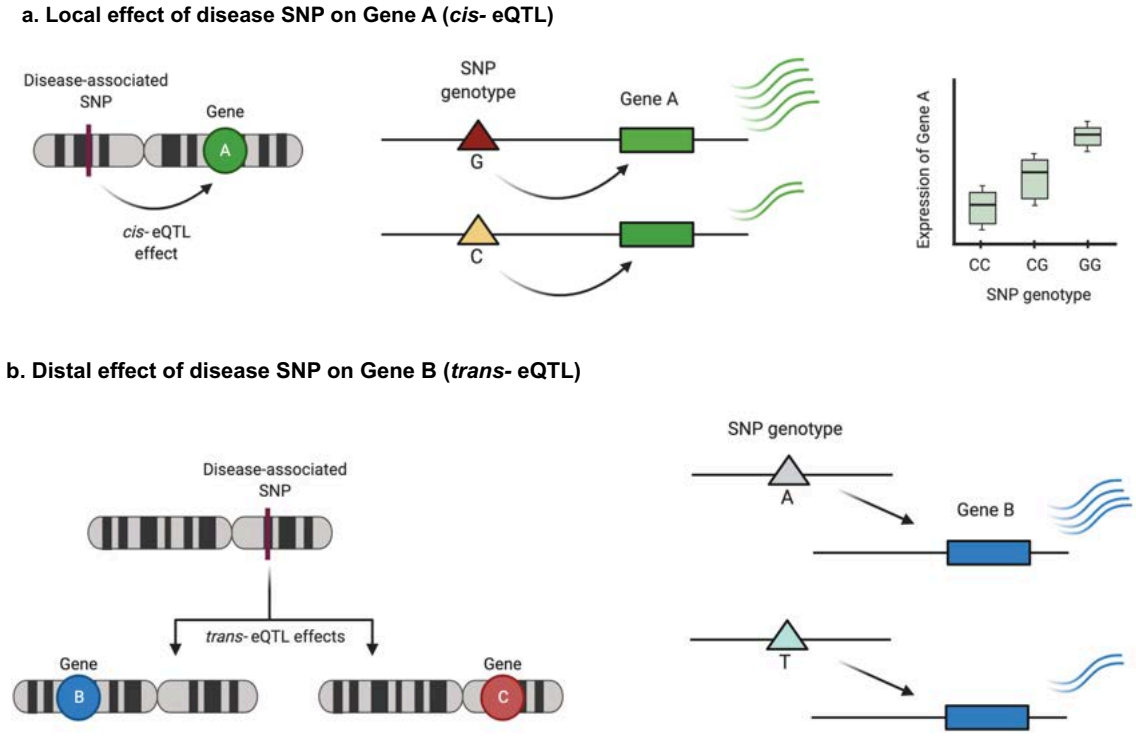


Figure 7: Effect of *cis*- and *trans*- acting SNPs in gene expression. (a.) The disease SNP shows a local effect on Gene A, as both the SNP and the gene are located on the same chromosome within an arbitrary genomic distance (<1Mb). (b) Distal effect from a disease-associated SNP on the genes B and C, which are located at a genomic distance >5Mb. The gene expression levels are altered by the presence of a SNP that either increases or decreases the levels of mRNA transcripts. Modified from Cheung & Spielman, 2009; Võsa et al., 2018.

The eQTL approach has enabled prioritisation of susceptibility genes in immune diseases. Early GWAS and replication studies identified numerous eQTLs, for e.g., in CD out of 71 detected loci, 39 were eQTLs, and in T1D 32 out of 53 loci were eQTLs (Ricaño-Ponce & Wijmenga, 2013). Since then, hundreds of eQTL analyses have been conducted in cardiometabolic diseases, autoimmune and inflammatory disorders. One prominent example is the eQTLGen Consortium: a recent initiative in whole blood eQTL studies, which identified ~17,000 *cis*-eQTL and ~6000 *trans*- eQTL analyses from 37 datasets including over 31,000 individuals (Võsa et al., 2018).

2.3.1 eQTLs are context-specific

Large-scale epigenomic studies like the Roadmap Epigenomics project (Kundaje et al., 2015) have identified *cis*-regulatory functions for ~20% of the non-coding regions and showed an enrichment of cell-type specific *cis*-regulatory associations, demonstrating that SNPs affect gene expression in a context-specific manner, and the importance of performing eQTL studies in functionally pertinent cell types and tissues. Early eQTL studies were

predominantly done in peripheral blood mononuclear cells (PBMCs) in unstimulated conditions, and in tissues. An initiative from the Genotype-Tissue Expression (GTEx) consortium enabled eQTL studies across 53 human tissues in ~1000 individuals, and also demonstrated the tissue-specific *cis*-regulatory effects of GWAS variants (Lonsdale et al., 2013; Melé et al., 2015). The functional relevance of detected tissue eQTLs were assessed by cross-referencing them with GWAS associations from the Wellcome Trust Case Control Consortium (WTCCC) studies in several complex diseases. With this approach, they found that eQTLs detected in whole blood and lymphoblastoid cell lines in GTEx also included SNPs that are associated to immune diseases like CD, RA and T1D.

Later on, eQTL studies interrogating the inter-individual variation in the human immune system were performed mainly in cell lines as reviewed by Zhernakova et al., 2017 and more recently in selected cell types like monocytes (Fairfax et al., 2014) and dendritic cells (Lee et al., 2014). For example, Fairfax et al. conducted an eQTL study in human primary monocytes activated with innate stimuli (IFNγ and LPS) for 2h or 24h (Fairfax et al., 2012). This study found that ~50% of the detected *cis*-eQTLs from IFNγ stimulated monocytes, overlapped with the GWAS-associated loci. The authors also reported that several eQTLs identified were dependent on the duration of stimulation. Using the newly identified stimulus-specific eQTLs, they explored their role in disease risk from previously reported GWAS loci. One of the associations they reported is in the *CARD9* locus, where a constitutively active eQTL SNP (rs4078099) was present, that was previously tagged from GWAS in IBD and AS (Evans et al., 2011), and was also shown to be associated with risk allele for CD (Jostins et al., 2012). Fairfax et al. detected a second stimulus-induced eQTL (rs36119806) in *CARD9,* also associated to CD risk, from IFNγ-stimulated monocytes. The risk alleles of both the SNPs were correlated with reduced *CARD9* expression levels. In the context of CD, *CARD9* is a mediator of innate and adaptive immune responses in the gut, and it is associated with IFNγ production in response to bacterial infection in the gut (Sokol et al., 2013). Low levels of *CARD9* may be linked to disruption in the maintenance of mucosal immune responses to clear out pathogens, thereby exacerbating the disease. This study elegantly shows that cell-type-dependent eQTLs which overlap with GWAS loci can explain the role of risk alleles in disease mechanism.

In order to study the role of regulatory variants in immune cell types, the DICE (database of immune cell expression, expression quantitative trait loci and epigenomics) project was established by Schmiedel et al. in 2018. The cell-type-specific gene expression patterns and the potential regulatory function of disease-associated SNPs were interrogated in 13 different immune cell types from 91 donors, under 2 activation states (Schmiedel et al.,

2018). The cell-type-specific gene expression data and genotyping data from the donors were used to discover 16,989 unique *cis*-associations and 6,298 *trans*-associations. The authors also reported SNPs that show cell-type-dependent effects, for e.g., the RA-associated variant, rs3093026, was associated with the expression of the chemokine receptor gene *CCR6* specifically in T_h1/T_h17 and T_h17 cell subsets. Therefore, studying cell-type-specific effects of genetic variants plays a vital role in linking GWAS variants to cell type function, which could in turn explain disease pathogenesis.

2.4 Integrating GWAS and eQTL data: TWAS

Disease-associated variants from GWAS are enriched for eQTLs (Nicolae et al., 2010). Therefore, susceptibility genes and SNPs from GWAS can be probed for causal implications in disease using eQTL data. This approach is called transcriptome-wide association analysis (TWAS), and it evaluates the association between the expression of each gene and a trait or disease of interest (Pasaniuc & Price, 2017). Since the number of genes is lower than the number of genetic variants, using gene expression rather than genotypes for association (like in GWAS) reduces the burden from multiple testing. However, such an approach is impractical as it would require gene expression data from hundreds of cases and controls and from multiple tissues and cell types. Alternatively, transcriptomic reference data can be used to impute gene expression pattern based on genotypes, without having the need to perform large-scale RNA-sequencing experiments. When performing TWAS, the information from GWAS and eQTL database is used to predict the expected gene expression in cases and controls. Gene expression can be predicted based on genotypes due to their high heritability and as most of its heritability comes from the nearest variants to the genes (Cano-Gamez & Trynka, 2020). TWAS also utilises eQTL data with the tissue or cell-type-specific gene expression and the genotypes of all the included donors as a training set to build gene expression predictors. The predictors take into account the SNPs that are located near a susceptibility gene to determine its expression levels and are used to impute the gene expression values for the genotypes of the individuals included in GWAS. The imputed gene expression values are then tested for association with the disease to obtain a set of genes that either positively or negatively affect it (Cano-Gamez & Trynka, 2020).

TWAS can enable the prioritisation of candidate causal genes in complex diseases. With the help of eQTL studies from multiple tissues and cell types with appropriate sample sizes these approaches can advance the functional gene prioritisation process that is important for determining the most likely candidate causal genes in a disease. However SNP prioritisation methods are mainly based on observational and theoretical analysis. Therefore it is important to integrate the regulatory information in the region with prioritised SNPs and carry out downstream experimental validation to establish causality.

Chapter 3

APPROACHES TO STUDY THE ROLE OF REGULATORY VARIANTS IN DISEASE

3.1 Regulatory elements and gene expression

Regulatory elements are important to initiate or curb gene expression in specific cell types in response to various developmental cues. These elements are present within the non-coding regions of the genome. An important class of regulatory molecules that affect gene expression are transcription factors (TFs) that can act as an activators or repressors depending on the biological context. TFs have the ability to bind to cis-regulatory elements that are situated in the transcription start sites (TSS) like promoter. They can also interact with distal regulatory components located far away from the TSS, such as enhancers, silencers and insulators. TFs enable the correct docking of RNA polymerase II which is important for the formation of the transcription-initiation complex (Rojano et al., 2019). Key regulatory elements that are involved in the transcriptional process are:

- **Promoters** are DNA sequences situated in the 5' region of the genes. They activate transcription through the action of TFs that enable the binding of RNA polymerase II to a consensus sequence forming the transcription-initiation complex.
- **Enhancers** are short DNA sequences that are bound by DNA-binding proteins called activators, located several base pairs away from the TSS. Enhancers interact with activators and form DNA loops that bring them closer to the promoter region, and increase gene transcription by interacting with RNA polymerase II.
- **Silencers** are short sequences of DNA that bind to repressors (DNA-binding proteins) to decrease transcription. They are either located close to the TSS or far away with the ability to form DNA loops that enable them to interact with promoter regions.
- **Insulators** are regulatory sequences that inhibit interactions between chromatin domains to regulate chromatin states. Transcriptional repressors like CCCTC-binding factor (CTCF) can act as an insulator by binding to its target sequence and blocking the interactions between promoters and enhancers.

The regulatory elements have an important role in governing transcriptional activity and gene expression. They can act as a transcriptional rheostat to synthesise the target gene transcripts based on the context. Gene expression is a very dynamic process that involves various events at the transcriptional level that can be affected by genetic variants, such as transcription factor binding to promoter and enhancer regions, chromatin regulation and interaction with other regulatory elements, alternative splicing and post-translational

modifications. Regulatory variants found in both genic and intergenic sites can disrupt these dynamic processes, affecting gene expression (Knight, 2014), and GWAS variants in the non-coding region of the genome are enriched in regulatory regions (Hrdlickova et al., 2014). These regulatory variants have been shown to affect transcription factor binding sites (TFBS), enhancers and DNA methylation, and can modulate the expression of downstream target genes as eQTLs (Glinos, Soskic, & Trynka, 2017; Rojano et al., 2019) (**Figure 8**).

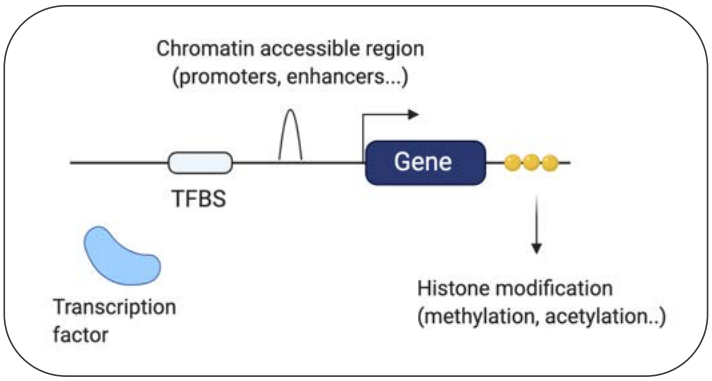


Figure 8: Epigenomic approach to study the role of disease variants. Immune disease variants are functionally annotated using epigenomics. This approach ascertains the epigenetic landscape of disease associated genomic loci to identify regulatory elements that enable better understanding of the effects of disease SNPs in target gene regulation. Adapted from Glinos, Soskic, & Trynka, 2017.

3.2 Epigenetic changes and chromatin accessibility

Putative regulatory regions are determined by studying the epigenetic modifications in the genomic locus. The basic unit of chromatin is the nucleosome, which consists of a histone octamer structure resulting from ~147bp DNA encircled around two molecules of the four core histones (H2A, H2B, H3 and H4) (Barth & Imhof, 2010). The physical interaction of DNA and histone can lead to the assembly of individual nucleosomes into several higher-order chromatin structures which can hinder DNA accessibility for the transcriptional machinery (Talbert & Henikoff, 2010). However, formation of chromatin loops bring about regulatory elements such as promoters and enhancers in close proximity and can enable transcription. Through several DNA-binding proteins or transcription factors (TFs) (Zhang et al., 2016).

Epigenetic modifications are heritable chemical or physical changes in chromatin-mainly histone modifications and DNA methylation. Histones are important players in epigenetic regulation of gene expression, and are subjected to post-translational modifications such as methylation, acetylation, ubiquitination, phosphorylation, that can be assessed using

chromatin immunoprecipitation (Kimura, 2013). For example, in actively transcribed genes, histone marks like trimethylation of Lysine 4 on Histone 3 (H3K4me3) is a signature of gene promoters, monomethylated H3K4 (H3K4me1) is enriched at distal regulatory elements like enhancers, and acetylated H3K27 (H3K27ac) is a hallmark of transcription start sites and enhancers (**Figure 9**) (Kuznetsova et al., 2020). Modifications of histones can create binding sites that are recognized by transcription factors (TFs), which upon binding recruit RNA polymerase II (Turner, 2005) and initiate transcription. The formation and removal of histone modifications occur in a coordinated manner with nucleosome turnover events that alter the regulatory landscape dynamically at each cell cycle (Chory et al., 2019). The histone modifications can thereby regulate chromatin accessibility and activation of regulatory elements like enhancers and promoters.

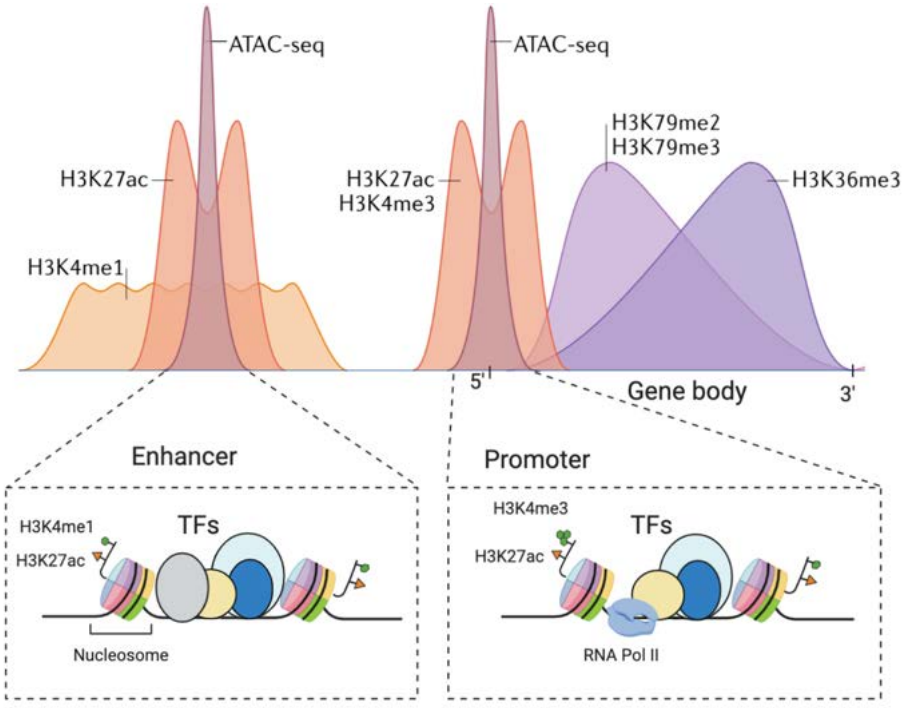


Figure 9: Epigenetic landscape of regulatory regions. Graphic representation of chromatin modifications and transcription factors in active promoters and enhancers. Active enhancers are characterised by the presence of H3K4 monomethylation (H3K4me1) and H3K27 acetylation as marked in the figure. Active promoters have H3K4me3 and H3K27ac. The ATAC-seq peaks represent open chromatin regions. Transcription factor binding occurs in regions outside the nucleosome, characterized by ATAC- seq signal enrichment as shown in the peaks. Adapted from Kuznetsova et al. 2020.

Chromatin accessibility (CA) refers to the extent to which the nuclear macromolecules can physically interact with chromatinized DNA (Klemm, Shipony, & Greenleaf, 2019). The accessibility remodelling is an important step during which TFs compete with nucleosomes to

gain access to their sites on the nucleosomal DNA (Lambert A et al., 2018). During nucleosome turnover, regions can remain exposed and accessible for small periods of time, during which the TFs can gain access to the histone-bound DNA. Hence, TFs play a crucial role in chromatin remodelling machinery by providing DNA-sequence specificity.

Chromatin accessibility is a direct consequence of epigenetic changes and organisation of nucleosomes that enable DNA-sequence specific TF binding. It can therefore be used to ascertain the regulatory potential at a genomic locus. The chromatin accessibility can be quantified through DNase I hypersensitivity assay to identify sites (DNase I hypersensitivity sites, DHSs) sensitive to cleavage by DNase I enzyme. Increased number in DHSs is a direct measurement of chromatin accessible regions. Another technique that is widely used to assess chromatin accessibility is via assay for transposase- accessible chromatin using sequencing (ATAC- seq), where the open chromatin regions are fragmented by transposase. These fragments are then sequenced to retrieve a stack of mapped reads, which is interpreted as the signature of a chromatin accessible region, illustrated in **Figure 9**.

3.2.1 Methods to functionally annotate the epigenomic landscape

High-throughput sequencing technologies have achieved functional annotation of the regulatory features in the epigenome. Several international consortia aim to produce data to annotate the regulatory elements in the non-coding genome, using different experimental approaches, such as Chromatin Immunoprecipitation Sequencing (ChIP-seq), chromosome conformation capture methods, DNase I hypersensitivity assays, DNase Sequencing (DNase-seq) and RNA Sequencing (RNA-seq). There are also tools available to process and utilise the information on TFBS, enhancer and promoter sites and interactions, DNA methylation sites, introns, splice sites, among many others. Some of the international projects that maintain databases for these data are **ENCODE (The ENCyclopedia of DNA Elements) and Roadmap Epigenomics Project**. ENCODE consists of epigenomic maps of several human cells and tissues. ENCODE provides information about chromatin state, TFBS and RNA transcripts. These data can be used to identify functional DNA elements and regulatory SNPs. Roadmap Epigenomics Project shows an overview of the human epigenome by including DNA methylation data, histone modifications, chromatin accessibility and mRNA transcripts in cell types and tissues.

Table 2: List of online data resources and tools in regulatory variant analysis. Some of the online tools used to assess the functional SNPs based on the regulatory data in the genomic loci.

Name	Description	URL
ENCODE	Encyclopedia of DNA elements Project	https://www.encodeproject.org
FANTOM	Functional Annotation of the Mammalian Genome project	http://fantom.gsc.riken.jp/5/
UCSC Genome Browser	Interactive visualisation of genomes	https://genome.ucsc.edu/
Roadmap Epigenomics Project	NIH Roadmap Epigenomics Mapping Consortium	http://www.roadmapepigenomics.org
International Human Epigenome Consortium	International Human Epigenome Consortium Data Portal	http://ihec-epigenomes.org/outcomes/ihec-data-portal/
BLUEPRINT	European hematopoietic epigenome project	http://www.blueprint-epigenome.eu
JASPAR	Transcription factor binding profile database	http://jaspar.genereg.net
PROMO	Transcription factor binding site analysis	http://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3

These data can be accessed through the University of California Santa Cruz (UCSC) Genome Browser. Based on the experimental data generated from these consortia, the regulatory variants are annotated and classified into database such as **RegulomeDB** and **HaploReg**. Some commonly used online data resources and tools for identifying putative functional variants are given in the **Table 2**.

A more recent addition to the list of resources available for regulatory variant analysis is the Open Targets Genetics platform (https://genetics.opentargets.org/). This platform provides a comprehensive catalogue of associations between traits/diseases, SNPs and genes. Open Targets Genetics is a compendium of association studies for complex disorders focusing on multi-trait analysis and variant-gene-trait association from the UK Biobank and the GWAS Catalog. This new resource has been very useful to gather information by gene or SNP or by complex disease/trait. Gene search generates the list of associated traits and variants identified in the gene locus. Similarly, SNP queries list information about the assigned nearest genes in the indicated SNP locus, eQTL effects of the SNP, distance of the SNP to the TSS and other potential regulatory information from published studies. Phenome-wide association studies (PheWAS), with SNP-phenotype associations (Pendergrass et al., 2011) and summary statistics from the UK Biobank and the GWAS Catalog are also listed for genetic variants, with level of significance, odds ratio and β-value (effect size) and direction of the

association (a β+ indicates risk, and a β- value denotes protective effect). This repository has been extremely useful to prioritize relevant SNPs for further functional analysis in independent disease cohorts.

3.3 Linking regulatory landscape with disease

The majority of the known GWAS associations overlap with a functional region, or are in LD with another SNP overlapping a functional region (Schaub et al., 2012). This effective approach to translate the role of regulatory variants in disease pathogenesis was established in a seminal paper focusing on post-GWAS analysis conducted on the ENCODE data, in order to link regulatory information with GWAS loci, and to identify functional SNPs potentially important in disease (Schaub et al., 2012). In this approach, Schaub et al. determined whether the lead variant from GWAS is a functional SNP or if it is in strong LD with a functional SNP. Next, they assessed whether the GWAS lead SNP or another SNP in strong LD with the lead SNP affect gene expression.

For example, as illustrated in **Figure 10**, SNP1 (the lead GWAS SNP) is in LD with SNP6 (functional SNP), which coincides with a TFBS. SNP3 is an eQTL SNP in LD with the lead SNP. However, neither SNP1 nor SNP3 are located in a functional region overlapping DHSs or TFBSs. This would suggest that the SNP6 is more likely to be the functional SNP associated with the biological phenotype (Schaub et al., 2012). Although a predicted TF binding motif is located at SNP2, a variant in LD with SNP1, this sequence does not overlap with DHS or ChIP-seq peaks, making it less likely to be functional. Similarly, SNP4 overlaps with a regulatory element, however it is not in LD with the lead variant and so it is less likely to be a functional variant for the disease.

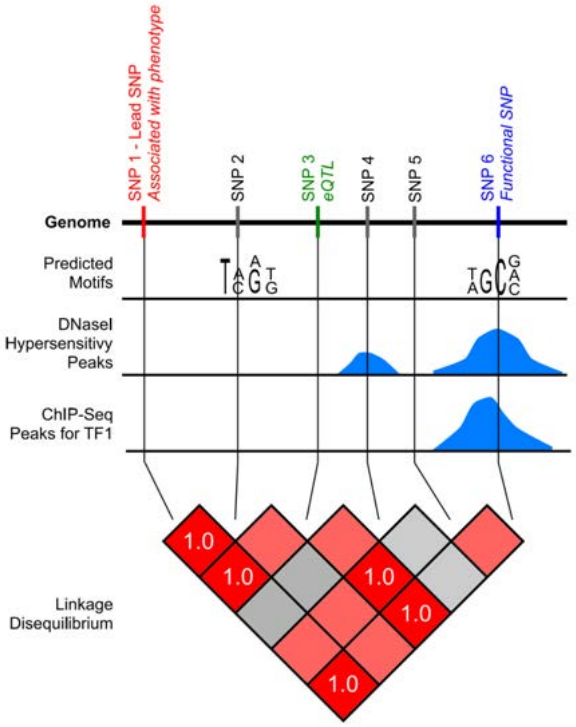


Figure 10: Overview of the functional SNP approach. This figure shows the schematic representation of the method used to identify functional SNPs. Regulatory data in the locus shows predicted motifs, DNase I hypersensitivity sites, ChIP-seq peaks and TF binding sites from ChIP-seq data. SNP1 is the lead SNP that is associated with a phenotype in GWAS. SNP2 is another variant in the locus in LD with the lead SNP. SNP3 is an eQTL SNP in the locus. SNP6 is highly likely to be functional as it is in LD with the lead SNP and the eQTL SNP (r^2>0.8), and it coincides with DNase I hypersensitivity peak with a TF binding site. This schematic is adapted from Schaub et al., 2012.

Using a similar approach for SNP prioritisation, Maurano et al. also reported that more than 90% of the disease-associated variants present in the non-coding region showed enrichment in regulatory features (like DHSs), that disrupt TFBS (Maurano et al., 2012). Therefore, functional annotation of GWAS loci has been successful though combining regulatory evidence from multiple sources. Using this method, several lead SNPs have been annotated to assess their functional relevance in disease association. For example, rs7163757, was a reported lead SNP for type 2 diabetes, that overlaps with a ChIP-seq peak and DHS peak, identifying it as a functional SNP. Additionally, DNA footprinting analysis revealed that the functional SNP overlaps with a predicted NFAT binding site (Schaub et al., 2012). This was one of the first integrative studies that combined GWAS, gene expression and evidence of regulatory activity to prioritise and identify functional loci that could explain the underlying biological mechanism of disease. Furthermore, Maurano et al. showed that GWAS SNPs associated to autoimmune diseases were found in DHSs in immune cells and some of them were reported to alter binding sites for several TFs involved in the IRF9 network associated with type I interferon induction, which may mediate autoimmune and inflammatory disorders (Maurano et al., 2012).

Another disease-specific example using this approach was reported in the UK IBD consortium paper, where a non-coding SNP (rs12212067) located in the intronic region of the gene *FOXO3A* was associated with reduced severity CD and RA (Lee et al., 2013). In this study, the authors assessed the influence of this variant in CD prognosis through the observed LD pattern in the locus and by testing for allele-specific expression. They also showed that FOXO3-regulates production of cytokine TGFβ1 in monocytes by measuring allele-specific binding of FOXO3 to the *TGFβ1* promoter (Lee et al., 2013).

3.3.1 Regulatory marks aid in identifying disease-relevant cell types

The context-specific nature of the epigenetic regulatory landscape makes it challenging to interpret the consequence of functional variants. The regulatory variant effects can occur only in a certain cell type, and/or under a certain activation state. Due to this diversity in the gene regulatory mechanism it is important to study the effect of SNPs in appropriate cell types and tissues to understand their functional mechanism (Kellis et al., 2014). The overlapping of SNP position with a regulatory feature such as histone marks is also an effective way to shortlist disease-related cell types (Trynka et al., 2013; Farh et al., 2015). This approach was used to explore the selective enrichment of GWAS variants within the regulatory regions in specific cell types. For example, Maurano et al. reported a significant enrichment of GWAS variants in DHSs in T_h17 and T_h1 immune cell subtypes for CD which is concordant with the implicated disease pathology (Maurano et al., 2012). Moreover, Farh et al. showed that histone marks signifying transcription initiation (H3K3me3) are enriched with RA and T1D associated variants in $CD4^+$ T cells (Farh et al., 2015). Similarly, an enrichment of GWAS SNPs was observed in a region bearing marks of active transcription in regulatory T cells (T_{reg}), which may be implicated in triggering autoimmune diseases including IBD, CeD and RA (Trynka & Raychaudhuri, 2013). These studies have provided an important link in complex disease genetics by implicating a role for disease-associated SNPs in immune cell function, activation and differentiation (Calderon et al., 2019; Soskic et al., 2019)

3.4 GWAS variants affect immune cell functions

Genetic studies in immune disorders have identified several loci associated with immunological parameters like immune cell counts, ratio between different immune cell populations and cytokine secretion in response to stimuli (Brodin et al., 2015). For instance, the risk allele of a SNP associated to T1D was associated with increased number of $CD4^+$ T

cells (Ferreira et al., 2010). Similarly, another study reported an overlap of SNPs associated to asthma and eosinophil counts emphasising their correlation (Astle et al., 2016). Hu et al. demonstrated that the genes in proximity to a locus associated to RA were enriched in effector memory $CD4^+$ T cells (Hu et al., 2014). Upon further analysis the authors reported a non-coding variant that correlates with proliferation capacity of effector memory $CD4^+$ T cells.

Another aspect of genetic studies in immune diseases is to understand how SNPs affect the effector mechanism of immune cells by regulating the secretion of cytokines and chemokines. Like gene expression, cytokine levels in the blood is highly heritable as shown by Brodin et al. in 2015. Studies from Mihai Netea's group have characterised cellular response to bacterial and fungal stimulations have identified cytokine-QTLs (cQTLs) for inflammatory cytokine levels such as IL-10, IL-6, IL-8 and TNFα (Li et al., 2016). cQTL associations are relevant in terms of autoimmune diseases, as many of the associated cytokines are targets for biologic therapies, for e.g., the usage of anti-TNFα for IBD, SpA, Ps & RA (Garcês, Demengeot & Benito-Garcia, 2013). Another article from by Li et al. worked with whole blood, PBMCs and monocyte-derived macrophages and examined the immune response to bacterial, fungal, viral, and non-microbial stimuli in 500 individuals (Li et al., 2016). The study identified monocyte-specific cQTLs associated with infectious disease susceptibility and T-cell-specific cQTLs being associated with autoimmune diseases (Li et al., 2016). Furthermore, studies conducted with standardised whole blood cultures by the Milieu Intérieur consortium from 1000 healthy donors with flow cytometric analysis and gene expression assays revealed the complex immune response signatures to a wide range of physiologically appropriate stimuli (Duffy et al., 2014; Urrutia et al., 2016), and the genetic influences at play in transcriptional variation in human immune responses (Piasecka et al., 2018). All these findings highlight how the genetic control of cytokines through cQTLs can affect immune response regulation.

3.5 Experimental validation of functional variants to assign causal relationships

The study of regulatory variants from GWAS associations has enabled the identification of target genes and genomic regions of interest for disease pathogenesis. Although these variants are implied in regulating gene expression, the lack of mechanistic insights that functionally link a specific risk variant to the underlying disease pathogenesis is a major limitation. Integrating GWAS variants with the regulatory information of a given locus, using eQTL mapping, ATAC-seq, and chromatin conformation capture assays, we can

postulate the function of a variant to an associated phenotype. The hypotheses generated from this approach can then be experimentally validated in an appropriate cell type or disease (Osgood & Knight, 2018).

Reporter gene assays are effective in determining the functional consequence of SNPs by testing allele-specific expression. This method was used by Vecellio et al. to validate a possible regulatory region in the *RUNX3* locus with a variant associated to ankylosing spondylitis (Vecellio et al., 2016). In this study, using reporter gene assays, the risk allele was correlated with decreased binding of the TF IRF4 and the decreased *RUNX3* expression in $CD8^+$ T cells, implicating a potential role for them in AS pathology.

Using the CRISPR/Cas9 method point mutations or a knockout can be introduced to study the downstream effects of the implicated variant and/or gene. This approach has been successfully employed in a study in Parkinson's disease, where an associated risk variant was identified in the enhancer region, which affected the expression of a disease-associated gene *SNCA* (α-synuclein). Using this approach, the authors established a causal relationship between the risk allele of the SNP and altered TF binding causing increased *SNCA* expression, that leads to disease (Soldner et al., 2016). Therefore, integrative approaches using regulatory data and experimental methods like reporter gene assays and gene-editing technologies can be useful to establish causal relationships.

Chapter 4

SPONDYLOARTHRITIS

4.1 Overview of spondyloarthropathies

Spondyloarthritis (SpA) is a collective term for a spectrum of inter-related chronic inflammatory rheumatic diseases that share genetic, clinical and pathophysiological features (Joachim Sieper & Poddubnyy, 2016). The spondyloarthropathies are categorised into two disease subsets: either axial SpA (AxSpA) or peripheral SpA, based on the dominant clinical and radiological manifestation (Rudwaleit et al., 2011) as shown in **Figure 11**. The subsets of SpA include psoriatic arthritis (PsA), spondylitis associated IBD, reactive arthritis (ReA) undifferentiated SpA (uSpA) and ankylosing spondylitis (AS) (Costantino, Breban & Garchon, 2018). The clinical characteristics of AxSpA are the involvement of sacroiliac joints and spine (axial skeleton), and peripheral SpA largely presents features such as enthesitis, dactylitis, and arthritis that prominently affect the lower extremities. SpA may also have extra-articular clinical manifestations including psoriasis, anterior uveitis and IBD (Joachim Sieper & Poddubnyy, 2016).

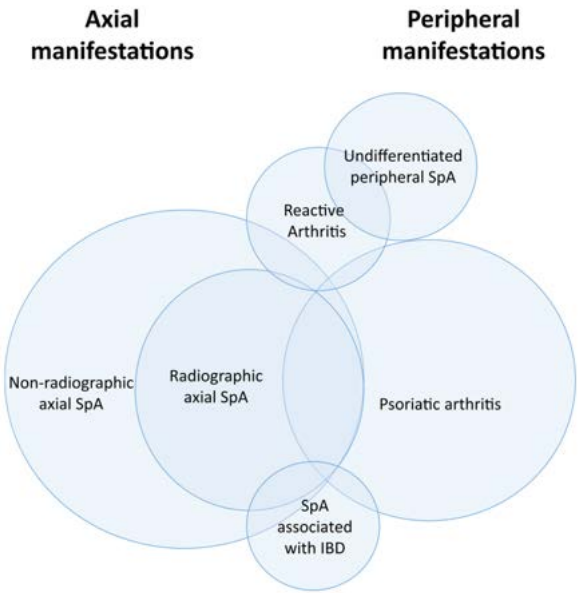


Figure 11: Schematic representing the SpA disease subsets and overlap between various spondyloarthropathies. The two main forms of SpA, axial and peripheral, with the corresponding disease subsets and shared features. Adapted from Proft & Poddubnyy, 2018.

4.2 Classification criteria

The criteria for disease classification for SpA are based on (i.) inflammatory back pain (IBP), (ii.) imaging evidence and (iii.) HLA-B27 status. IBP is a key clinical symptom of SpA that reflects the extent of inflammation in the sacroiliac joints, spine and spinal entheses. However, the diagnostic accuracy and specificity in using IBP for classification is limited. On the other hand, imaging of the sites of SpA-associated inflammation in the sacroiliac joints and spine has been shown to improve disease classification of SpA patients (Akgul & Ozgocmen, 2011). The conventional radiographic methods used to assess sacroiliitis could not aid diagnosis at early stages of SpA, as the structural changes manifest later in the disease. Magnetic resonance imaging (MRI) enables clinicians to capture active inflammation in the spine and sacroiliac joints also in early axial form of SpA. Specifically, MRI-based assessment can identify osteitis (bone marrow oedema), enthesitis, synovitis and ankylosis (Akgul & Ozgocmen, 2011). Considering this, the new Assessment of SpondyloArthritis International Society (ASAS) integrated an “imaging arm” and a “clinical arm” into the classification criteria for identifying axial SpA (Rudwaleit et al., 2011). In the clinical arm, HLA-B27 positivity is an important factor to support early diagnosis of SpA, since about 8% of the general population are HLA-B27 positive and around 90% of patients with SpA (AS) are HLA-B27 positive (Akassou & Bakri, 2018).

The ASAS classification for axial SpA includes patients under 45 years of age with persistent inflammatory back pain lasting over a period of 3 months. Patients should also have evidence of sacroiliitis on X-ray or MRI along with an additional SpA feature, or HLA-B27 positivity and at least two additional extra-articular features (Pso, uveitis or IBD) (Rudwaleit et al., 2011) (**Figure 12**).

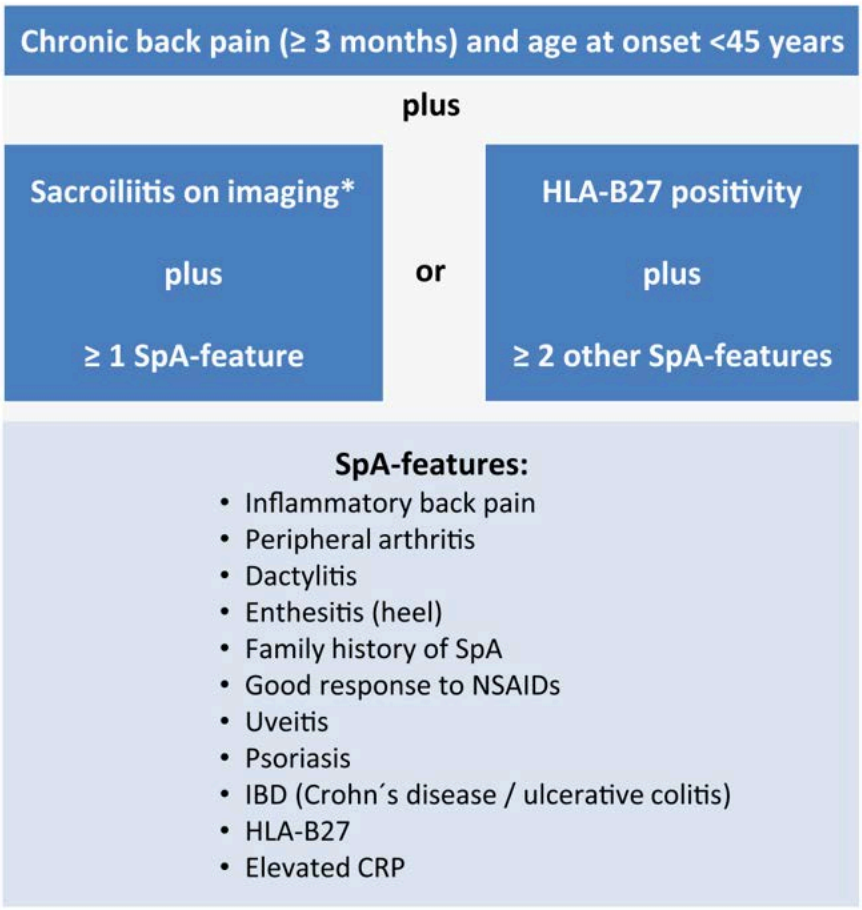


Figure 12: Assessment in Spondyloarthritis International Society (ASAS) classification criteria for axial spondyloarthritis (SpA). *Sacroiliitis on imaging refers to definite radiographic sacroiliitis according to the modified New York criteria or active sacroiliitis on magnetic resonance imaging according to the ASAS consensus definition. CRP, C-reactive protein; HLA-B27, human leucocyte antigen-B27; IBD, inflammatory bowel disease; NSAIDs, nonsteroidal anti-inflammatory drugs. Adapted from Proft & Poddubnyy, 2018.

4.3 Ankylosing spondylitis

Ankylosing spondylitis (AS) is the prototypic form of SpA with radiographic changes in the sacroiliac joints (radiographic axial SpA) (Dougados & Baeten, 2011). Persistent inflammation at target sites can lead to fibrosis and bone calcification. The more advanced cases of AS present osteoproliferation, bony fusions of vertebral joints and formation of syndesmophytes (Taurog, Chhabra & Colbert, 2016) that affect flexibility of the spine by forming long bony column often referred to as “bamboo spine” (Sieper et al., 2002) (**Figure 13**). Furthermore, nearly 25% of AS patients present evidence of musculoskeletal features such as peripheral enthesitis. AS patients may also develop extra-articular manifestations, for example, 5-10% of AS patients have IBD, nearly 10% show psoriasis, approximately 25% have uveitis and around 70% of cases have subclinical ileitis (Thomas & Brown, 2010). Some of the major accompanying co-morbidities of AS are cardiovascular and gastrointestinal complications, contributing to prominent socio-economic impact for young adults.

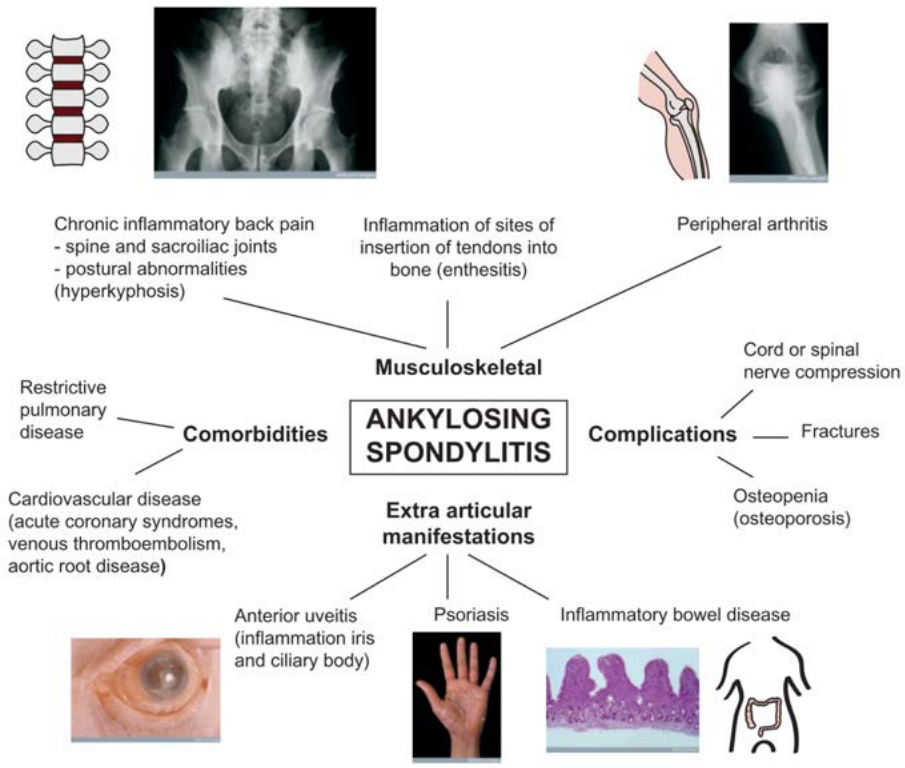


Figure 13: Diagrammatic representation of the musculoskeletal and extra-articular manifestations of AS together with comorbidities and major complications. AS is characterized by musculoskeletal complications such as inflammation and osteoproliferation, followed by bone fusion of the spine and sacroiliac joints. It also presents anterior uveitis, psoriasis and IBD as extra-articular manifestations. Comorbidities involving pulmonary diseases and cardiovascular diseases along with other complications such as osteopenia are often observed. Adapted from Osgood & Knight, 2018.

AS has a worldwide prevalence between 0.1 – 1.4% (Dean et al., 2014), with a higher prevalence rate in Europe (0.25%) and North America (0.20%) (Sieper et al., 2006). AS also shows higher prevalence rate in males compared to females with an approximate male to female ratio of 3:1 (Ngo, Steyn, & Mccombe, 2014; Stolwijk et al., 2016). Men also tend to be more severely affected than women. Commonly observed age of onset for AS in multiple cohorts is in the second and third decade of life, with a determined mean age of onset of 28.3 years according to a study conducted by a German rheumatological database sampling over 8000 subjects (Sieper et al., 2002). AS is a highly heritable form of arthropathy with a strong genetic association to HLA-B27 (Brown, 2011). The prevalence rate of AS correlates with the HLA-B27 status in specific populations. More than 80% of the Caucasian European AS cases are HLA-B27 positive. However, only around 6% of HLA-B27 carriers actually develop AS (Matthew A. Brown, 2011).

4.3.1 HLA-B27 and AS

The association of HLA-B*27 was discovered in the 70's by Caffrey and James when the polygenic nature of AS was poorly understood (Caffrey & James, 1973). Although HLA is the main genetic factor associated to AS, the exact mechanism and role of HLA-B27 is not clear. This association was detected and replicated across diverse populations suggesting that this is in fact a true disease-association. Being a highly polymorphic region, there are many subtypes that have also been associated to AS based on the population ancestry. HLA-B*27:05 is the most frequent ancestral allele which is present in nearly all populations. Other associated subtypes that are derived from B*27:05 are population-dependent, namely, B*27:02 in the Caucasian and Mediterranean populations, B*27:06 in south-east Asia, B*27:04 and B*27:07 are the major subtypes in Asians (Khan et al., 2007; Li & Brown, 2017). Additionally, Immunochip studies in AS identified many other HLA-B alleles that were linked to either increased disease risk (B*13:02, B*40:01, B*40:02, B*47:01, B*51:01) or protection (B*07:02 and B*57:01). Independent variants in the MHC class II genes were also associated to AS susceptibility - for example, HLA-DRB1, HLA-DPA1, HLA-DPB1 specifically in the French and Spanish populations (Reveille, 2014). Non-HLA related genes in the MHC locus have also been associated with AS. However, these associations are difficult to interpret due to strong LD between variants in the non-HLA genes (MICA, TNF, etc.) with the MHC locus as discussed by Costantino, Breban & Garchon (2018).

4.3.2 Pathogenetic mechanisms in AS

The aetio-pathological mechanisms in AS are not fully understood. The HLA-B27 association along with the genetic associations from GWAS have provided insights into the underlying disease biology. However, many functional aspects of these genetic loci are yet to be studied. The actual mechanism of HLA-B27 in AS disease biology is not known. Classically there are three main hypotheses that discuss the potential pathogenetic mechanisms of HLA-B27 in AS (**Figure 14**):

4.3.2.1 The arthritogenic peptide theory

MHC class I molecules (HLA-B27) bind peptides derived from intracellular proteins and present them to $CD8^+$ cytotoxic T cells (CTLs). The heavy chains of HLA-B27 are synthesised in the endoplasmic reticulum (ER) where they form a peptide-loading complex with β2-microglobulin, which enables the cell surface presentation of antigenic peptides and self-antigens to CTLs. Considering this role of HLA-B27, Parham and Benjamin proposed that in AS, an HLA-B27-restricted CTL response is initiated by the endogenous arthritogenic peptides, that may resemble microbial peptides (Benjamin & Parham, 1990). There are several studies supporting this theory from the 90s. For example, a study conducted in T cells reported that HLA-B27-restricted $CD8^+$ T cells recognised both infected and uninfected target cells that were present in the peripheral blood and synovial fluid from the joints of ReA and AS patients (Hermann et al., 1993). There has also been some evidence in identifying potential pathogen-specific peptides in the synovial fluid of ReA patients, for example *Yersinia*- and *Chlamydia*-derived peptides were reported in this study (Hermann et al., 1993). However, the autoantigens that cross-react with autoreactive T cells have not been identified yet. Therefore it has not been possible to identify a suitable arthritogenic peptide from disease-relevant tissue sites that proves the HLA-B27-restricted and CTL-mediated cross-reactivity (Mchugh & Bowness, 2012). Moreover, there are some contradictory evidence from studies conducted in $CD8\alpha^{-/-}$ or B27-depleted transgenic rat models, in which the disease developed and persisted, suggesting that the disease is not entirely $CD8^+$ T cell dependent (May et al., 2003; Taurog et al., 2009).

4.3.2.2 Unfolded protein response and the ER stress response

HLA-B27 has a tendency to form misfolded homodimers in the ER that lead to cytosolic protein degradation. Under normal circumstances, the heavy chain of HLA-B27 is internalised from the cell surface and processed by lysosomes. But a proportion of the HLA-B27 heavy chains dislocates from the ER membrane soon after synthesis before binding to β2-microglobulin and gets degraded by cytosolic proteasomes. The accumulated misfolded proteins can create a stress response in the ER and initiate the unfolded protein response (UPR) mechanism and/or autophagy. The ER-associated degradation (ERAD) pathway and unfolded protein response (UPR) mechanisms are used to discard misfolded and non-assembled ER proteins (Colbert et al., 2009). Evidence of this mechanism has been detected in bone-marrow derived macrophages from HLA-B27 transgenic rats, where HLA-B27 misfolding observed after cytokine stimulation correlated with increased IL-23 production (Turner et al., 2007; DeLay et al., 2009). Although these data have indicated that AS/SpA could be the result of incorrectly folded HLA-B27 accumulating in the ER, contributing to ER stress and consequent inflammatory responses, there is no direct evidence in human SpA for this mechanism. Furthermore, a study performed in B27-transgenic rats by introducing copies of human β2-microglobulin gene reduced B27 misfolding and alleviated colitis, but exacerbated arthritic features (Tran et al., 2006). Studies in AS human tissues did not show convincing evidence for ER stress and UPR. However, a report found signs of autophagy in the gastrointestinal tract of AS patients with increased IL-23 production and secretion by gut-derived mononuclear cells (Ciccia et al., 2014). This study suggests that autophagy related pathways might be implicated in human AS. However, most of the autophagy linked genes identified in GWAS are associated with CD, and are not shared with AS, raising question about the importance of this pathway in AS pathogenesis.

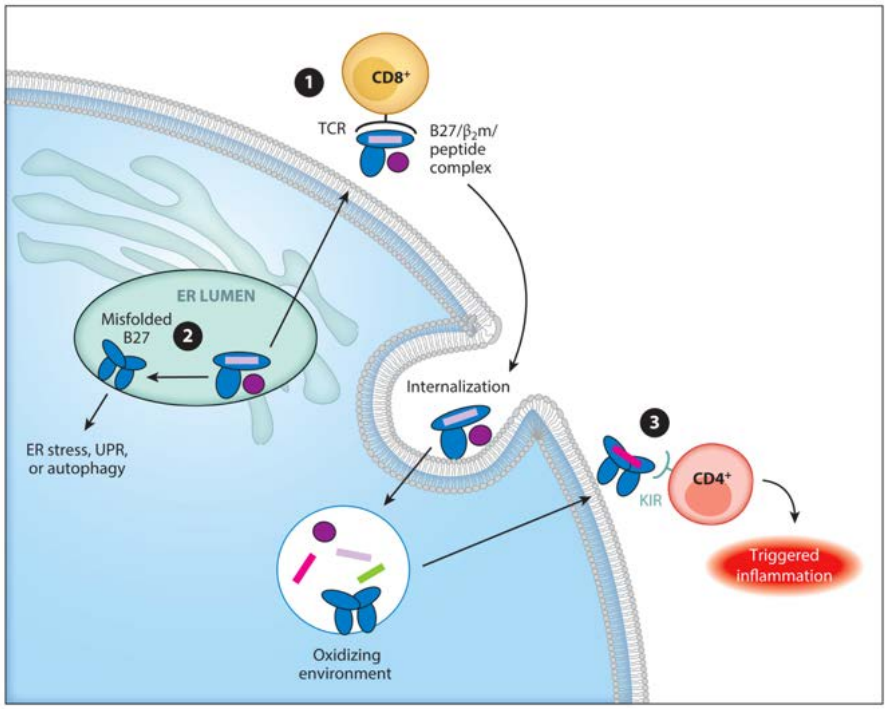


Figure 14: Hypotheses explaining pathogenic mechanisms in AS. 1. Arthritogenic peptide hypothesis: Self- and/or pathogen-derived peptides are selected and presented by properly folded forms of HLA-B27 for recognition by the autoreactive CD8+ T cells. **2. B27 misfolding hypothesis:** Misfolding of B27 within the ER causes ER stress, unfolded protein response (UPR), or autophagy, which has downstream effects on cellular function. **3. Cell surface B27 free heavy chain expression and immune recognition hypothesis:** B27 free heavy chains including dimers are expressed at the cell surface, where they are recognized by cells bearing killer immunoglobulin–like receptors (KIR) and/or leukocyte immunoglobulin–like receptors to trigger inflammation. Adapted from Bowness, 2015.

4.3.2.3 Cell surface B27 free heavy chain expression and immune recognition hypothesis

HLA-B27 homodimers can be expressed on the cell surface after the β2-microglobulin detaches from the HLA heavy chain. These homodimers can bind to killer immunoglobulin-like receptors (KIR3DL2) and leukocyte immunoglobulin-like receptors (LILR) on the surface of NK cells and certain subsets of T cells ($CD4^+$ T cells, $CD8^+$ T cells and γδT cells) (Bowness, 2015). Antigen presenting cells that express HLA-B27 homodimers can therefore interact with KIRs expressed on NK cells (Chan et al., 2005) and T cells (Wong-Baeza et al., 2013) and induce pro-inflammatory effects (Bowness, 2015). HLA-B27 expressing APCs have also been reported to enhance IL-17 production from $KIR3DL2^+$ $CD4^+$ T cells in SpA patients, linking HLA-B27 with IL-17 secretion (Bowness et al., 2011). Moreover, KIR3DL2/B27 interaction have shown to enhance the proliferation and survival of T_h17 cells in AS (Taurog et al., 2016).

4.4 Genetic epidemiology of AS

Family-based studies gave the first indication that genetic factors contribute to disease predisposition in AS. The recurrence risk of AS in first-degree relative or sibling of patients was estimated to be 80 (Brown et al., 2000). Furthermore, disease recurrence patterns from twin studies in AS estimated heritability to be >90% (Brown et al., 1997; Robinson & Brown, 2014). However, it was not clear whether the high recurrence rate in families came from shared genetics or environmental factors until the discovery of HLA-B27 association to AS susceptibility in 1973 (Brewerton et al., 1973; Caffrey & James, 1973). The MHC locus on the chromosome 6 is one the most gene dense regions in the genome with several genes associated to immunological function. This region is by far the strongest association to AS susceptibility through MHC class I molecule HLA-B27, which contributes to 20.5% of the estimated heritability (Ellinghaus et al., 2016). Although 96% of AS patients are HLA-B27 positive, only a small fraction of HLA-B27 positive patients develop AS. HLA-B27 remains the strongest known genetic association between the MHC antigen and a disease, but it alone is not sufficient to develop the disease, especially since monozygotic twins with AS show a concordance rate of ~75% and dizygotic twins only around 15% (Breban, 2006). Hence there must be other factors besides the MHC region that contribute to disease susceptibility.

4.4.1 Associations outside the MHC region: GWAS results in AS

Since the discovery of HLA-B27, over 40 loci have been associated to AS susceptibility, which include both MHC and non-MHC genetic factors as shown in **Figure 15**. Some of these were identified through early candidate-gene based approaches. However, the majority of these loci were identified through GWAS. The first GWAS in AS was conducted by the Wellcome Trust Case Control Consortium and the Australo-Anglo-American Spondylitis consortium in 2007, and identified 2 non-MHC susceptibility loci - *ERAP1* and *IL23R* (Burton et al., 2007). Later, the International Genetics of Ankylosing Spondylitis Consortium (IGAS) conducted a GWAS using the Immunochip whole-genome microarray and identified 13 associated loci (Cortes et al., 2013). A more recent study by Ellinghaus et al. combined the Immunochip data from 5 chronic inflammatory diseases (AS, Crohn's disease, psoriasis, primary sclerosing cholangitis, and ulcerative colitis) to increase the statistical power (Ellinghaus et al., 2016), and reported 117 variants associated to AS, 17 of which were novel associations.

Despite its limitations, GWAS was able to uncover some important susceptibility loci in AS that indicated biological mechanisms and pathways that could be important in its pathology. The main findings from GWAS in AS pointed to the involvement of aminopeptidase genes, genes of the tumour necrosis factor (TNF) and IL-23/IL-17 pathways.

Aminopeptidase genes such as *ERAP1* and *ERAP2* encode enzymes in the endoplasmic reticulum (ER) that trim and process peptides before they are presented on the MHC class I molecules (Brown, Kenna, & Wordsworth, 2015). ERAP1 association with AS susceptibility has been demonstrated in AS disease models by Evans et al., who reported abnormal or reduced peptide trimming due to SNPs in *ERAP1* that disrupt antigen presentation by HLA class I molecules (Evans et al., 2011). *ERAP1* activity has also been linked to the expression of HLA-B27 free heavy chain on the cell surface implied in the T_h17 responses in AS (Chen et al., 2016). Additionally, there are other aminopeptidase genes that are implicated in AS, such as *LNPEP* and *NPEPPS*.

Another pathway associated with AS pathology through GWAS is the IL-23/IL-17 axis. IL-23 is a pro-inflammatory cytokine that is important in the activation of like the T_h17 subset of T cells. Many variants in genes of this pathway have been detected through GWAS including: *IL12B*, tyrosine kinase 2 (*TYK2*), *CARD9*, *PTGER4*, *JAK2*, nuclear factor kappa B subunit 1 (*NFKB1*), *IL27* and IL23R (Li & Brown, 2017). Members of this pathway, like *TYK2*, play an important role in the signal transduction of several cytokines, including IL-10, IL-12, IL-23, IL-6, IFN-α and IFN-β (Dendrou et al., 2016). Genes of this pathway show pleiotropic effects in many CIDs. For example, *IL23R* locus is associated to AS-related disorders, like Pso and IBD. Variants in *TYK2* have been reported in AS, Pso, IBD, T1D, MS and RA. There are also other shared genetic associations between these diseases, for e.g., *IL12B*, *IL23R*, *JAK2* and *IL27* are shared susceptibility loci for AS and IBD. Although these loci are implicated in many CIDs, suggesting a shared pathology through IL-23R signalling, the exact biological mechanism is yet to be understood.

The TNF signalling pathway has also been highlighted in AS pathology through GWAS. SNPs were identified in TNF signalling pathway genes, such as *TNFRSF1A*, *TRADD* and *TNFSF15*. TNF is a key player in initiating and maintaining inflammatory responses and therefore is an important target for biologic treatments. TNF inhibitors are widely used in managing several forms of SpA and other CIDs, to treat inflammation.

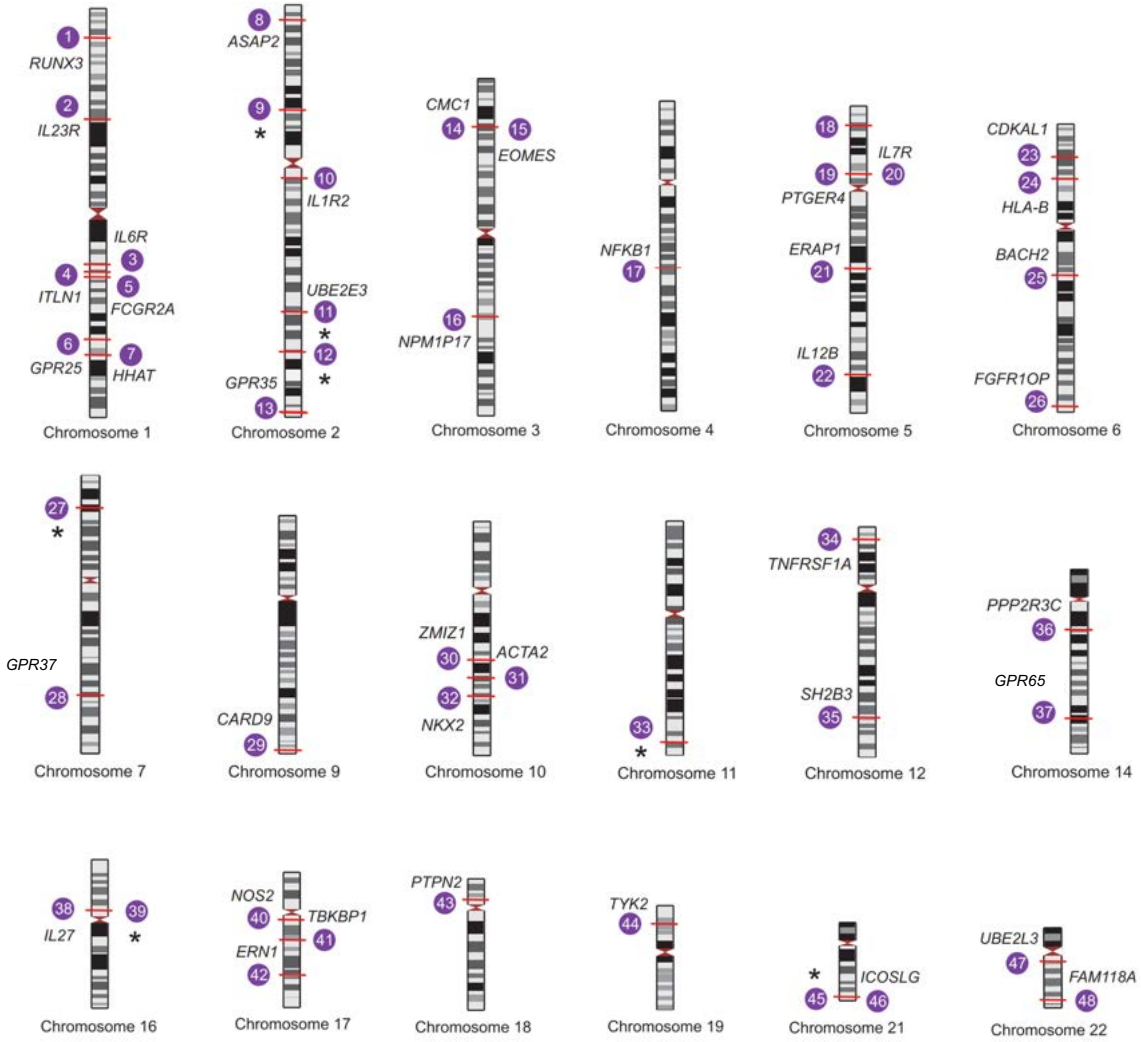


Figure 15: Overview of AS GWAS loci. Loci associated with AS at genome-wide significance reported by Ellinghaus et al. are plotted (numbered 1–48). The implicated gene and chromosomal position (marked in red) are shown for each locus. Asterix indicates loci where there is no reported gene. Adapted from Osgood & Knight, 2018.

To date, only 27.82% of the overall heritability in AS has been explained. The majority of the heritability component comes from the association with HLA-B27 (20.44%) (**Figure 16**). All the other associations from candidate gene studies and GWAS add up to nearly 7% (Ellinghaus et al., 2016). The rest 72% of the missing heritability could be attributed to rare variants, gene-gene and gene-environment interactions as it is the case for many complex diseases (Eichler et al., 2010). Solving the missing heritability puzzle in AS requires further study integrating methods to assess protein interactions, epistasis and the effects of structural variants (deletions, duplications and inversions) and rare variants (O'Rielly, Uddin, & Rahman, 2016).

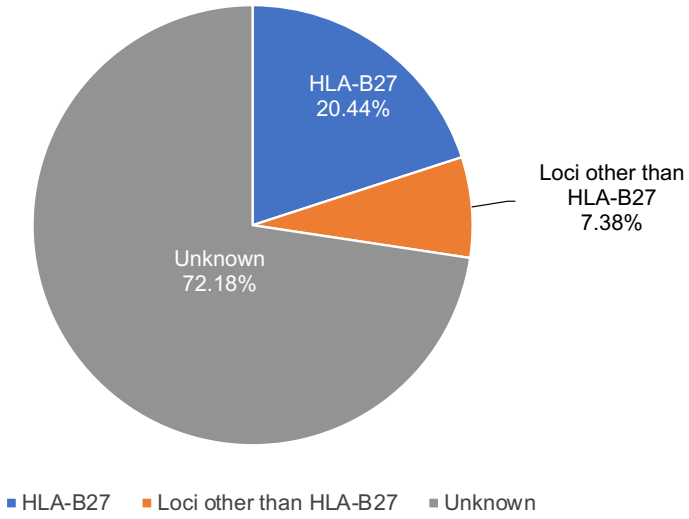


Figure 16: Heritability explained in AS. 27.82% of the heritability in AS is attributed to HLA-B27 and other non-MHC loci and the rest 72% remain unaccounted for. Figure adapted from Ellinghaus et al., 2016 and Costantino, Breban & Garchon, 2018.

4.5 The IL-23/IL-17 pathway in spondyloarthritis

The IL-17 superfamily of cytokines consists of IL-17A, IL-17B, IL-17C, IL-17D, IL-17E and IL-17F. IL-17A is the most common and widely studied member of this family and it is often referred to as IL-17 in the literature. It was initially described in $CD4^+$ T cells of the T_h17 phenotype but later identified in other cell types, including $CD8^+$ T cells, γδT cells**,** natural killer T cells (NKT cells) and mucosal associated invariant T (MAIT) cells, as reviewed by Taams et al., 2018. Mast cells, neutrophils and innate lymphoid cells have also been reported to produce IL-17A (Cua & Tato, 2010). IL-17A signalling has been implicated in many inflammatory diseases including RA and SpA (Lubberts, 2015). IL-17 can regulate transcriptional activation and consequent production of pro-inflammatory cytokines and chemokines such as TNF, IL-1, CCL2 and CXCL8. It has also been shown to increase the secretion of granulocyte colony-stimulating factor (G-CSF) and granulocyte-macrophage colony-stimulating factor (GM-CSF) in T cells and macrophages (Taams et al., 2018). Moreover, in the context of inflammatory arthritis, IL-17A has been reported to induce matrix metalloproteinases (MMPs) in target cells, which drive the degradation of extra-cellular matrix at the inflamed synovial joints (Chabaud et al., 2000).

Interleukin 23 (IL-23), has been shown to regulate the differentiation of $CD4^+$ T_h17 cells and IL-17 production (Aggarwal et al., 2003; McGeachy et al., 2009). Furthermore, Sherlock et al. reported that the IL-23R expressing T cells that reside in entheses can induce SpA under the influence of IL-23 present in the milieu (Sherlock et al., 2012). Turner et al. had shown that HLA-B27 misfolding could result in ER stress and unfolded protein response (Turner et al., 2005); furthermore, DeLay et al. showed that ER stress was correlated with IL-23 production in HLA-B27 transgenic rat models (DeLay et al., 2009). The misfolded protein was also reported to induce UPR that can consequently increase IL-23 production from the APC (Colbert et al., 2010). These results suggest a possible link between IL-23 and ER stress induced UPR hypothesis in SpA. Additionally, several variants identified through association studies in the *IL23R* and the *IL-12B* locus, which encodes for the IL-12p40 subunit shared between IL-12 and IL-23, have been associated with susceptibility to SpA-related disorders (PsA, AS, Pso and IBD) (Cortes et al., 2013). The summary of the polymorphisms identified in the IL-23/IL-17 pathway in AS and PsA is shown in the **Figure 17**. Therefore in order to determine the role of IL-23/IL-17 pathway in SpA, it is important to investigate the immune cells that are involved in the production of these cytokines and how they behave in the context of disease.

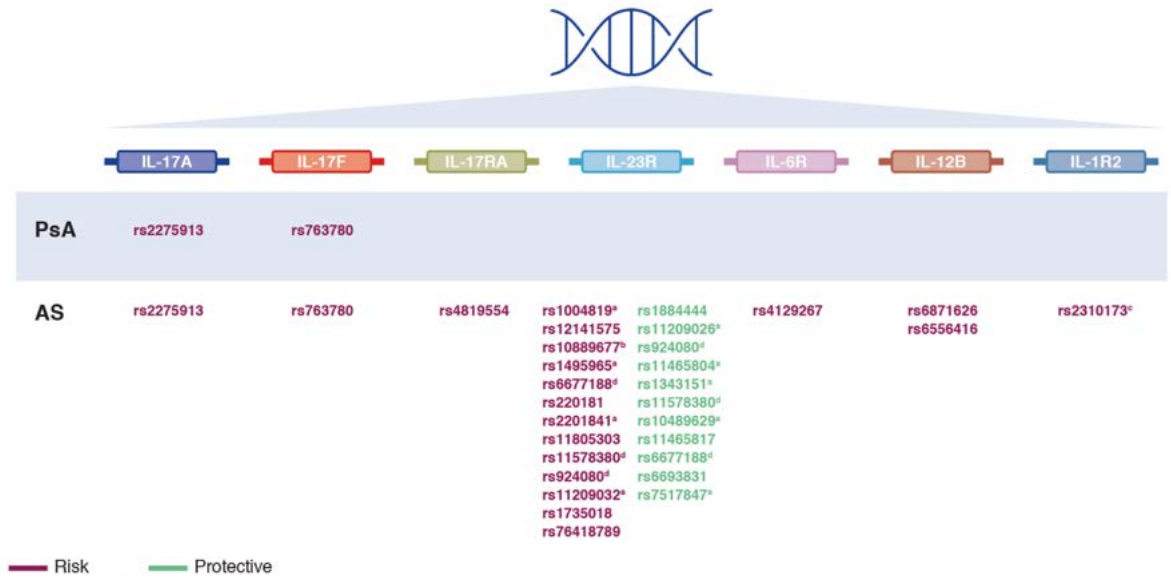


Figure 17: Single nucleotide polymorphisms identified in the IL-23/IL-17 signalling pathway that have been linked to axial SpA and PsA. [a]Significant association shown in European but not Asian populations; [b]No risk associated with this SNP shown in certain studies; [c]No risk associated with this SNP shown in certain studies207; [d]SNP can be associated with risk or protection depending on the specific mutation. AS is ankylosing spondylitis and PsA is psoriatic arthritis. Adapted from McGonagle et al., 2019.

4.6 Innate and adaptive immune cells implicated in AS pathology

Although the precise pathological mechanisms of AS are not fully understood, several different immune cell populations may be implicated in the pathogenesis of the disease.

4.6.1 Innate immune cells

Natural Killer (NK) cells comprise 1 – 6% of the peripheral blood mononuclear cells. NK cells are important components of innate immunity and are involved in immunosurveillance and defence against pathogens. They are characterised by the expression of surface markers CD56 and CD16 and the lack of CD3. NK cells are further divided into two groups based on CD56 expression: $CD56_{dim}$ and $CD56_{bright}$. 90% of the circulating NK cells are $CD56_{dim}$ NK cells that express perforin and killer immunoglobulin-like receptors (KIRs), whereas $CD56_{bright}$ NK subset are localised to the secondary lymphoid tissues (Caligiuri, 2008). Impaired function of NK cells has been associated with immune mediated disorders such as AS, SLE, RA and MS (Fogel, Yokoyama, & French, 2013). AS patients have shown an increased percentage of $CD56_{dim}$ $CD16^{+}$ NK cells with high expression of the NK-inhibitory receptor *CEACAM1* (carcino-embryonic antigen-cell adhesion molecule) (Azuz-Lieberman et al., 2005). This study observed that *CEACAM1* is induced by IL-8 and stromal cell derived factor 1 (*SDF-1*), which are present in elevated levels in the sera of AS patients, suggesting a potential role for *CEACAM1* expressing NK cells in AS pathology.

HLA molecules have been reported to interact with NK cell inhibitory receptors, for e.g., KIR3DL1 has been shown to recognise HLA-B27. KIR3DL1 interacts with HLA-B27 and suppress cytolytic potential of NK cells and T cells (Rezaiemanesh et al., 2018). Additionally, GWAS have reported genetic polymorphisms in KIR genes (*KIR2DL1*, *KIR3DL1*, *KIR2DS5*, *KIR3DS1*, and *KIR2DL5*) associated to AS susceptibility in different populations, which emphasises their potential role in disease mechanism.

Monocytes are precursors to antigen presenting cells, including dendritic cells and macrophages, and are characterised by the expression of surface markers CD14 and CD16. Microarray data from peripheral blood monocytes from SpA patients showed an increase in *MNDA* (myeloid nuclear differentiation antigen) which is a marker for inflammatory monocyte/macrophage (Gu et al., 2002). Another study in SpA patient-derived monocytes also found increased expression of toll-like receptors 2 and 4 (*TLR2* and *TLR4*), which were reduced by TNFα-blockade (De Rycke et al., 2005). Monocyte-derived macrophages have

also been implicated in AS disease pathogenesis through UPR from HLA-B27 misfolding, which eventually results in the production of pro-inflammatory cytokines such as TNF, IL-1α, IL-6 and IL-23 (Colbert, Tran & Layh-Schmitt, 2014). There is additional evidence in the literature that reported elevated frequency of classical monocytes ($CD14^+$ $CD16^-$) and reduced frequency of non-classical monocytes ($CD14^-$ $CD16^+$) from the peripheral blood of AS patients (Surdacki et al., 2014; Conrad et al., 2015).

Furthermore, SNPs associated to AS are enriched in open chromatin regions in immune cells including monocytes, suggesting a potential role of these cells in disease biology (Farh et al., 2015). eQTL studies in monocytes have revealed that innate immune pathways are affected through variants in the regulatory region in a stimulus-dependent manner (Fairfax et al., 2014). For example, polymorphisms in *IL7R* associated with pre-disposition to AS were found to be eQTLs in blood-monocytes activated with an innate stimulus (Al-Mossawi et al., 2019). The same study also reported an enrichment of $IL7R^+$ monocytes in the synovial fluid from AS patients giving further emphasis on the potential role of monocytes in AS pathology. Another study explored the role of micro-RNA (miRNA) in SpA associated pathology by determining the expression profile of miRNA in monocytes derived from SpA patients (Fogel et al., 2019). The authors identified an enrichment of deregulated miRNAs, where miRNA are either degraded or not translated during gene regulation, in pathways associated to monocyte/macrophage differentiation and polarization, and in pathways that control the expression and secretion of pro-inflammatory cytokines. These studies highlight the potential function of monocyte in AS pathogenesis, and could be useful for developing targeted therapeutic interventions.

4.6.2 Innate-like T cells

γδT cells are “unconventional T cells” that present features of adaptive T cells such as antigen recognition through T cell receptors and orchestrate T cell effector functions, but act in an innate-like manner. γδT cells express heterodimeric T-cell receptors (TCRs) composed of γ and δ chains, and are mainly found in the tissues but they are also present in peripheral blood in smaller proportions (0.5- 5%). γδT cells expressing IL-17 were detected in peripheral blood from Pso patients, in whom the lower percentage of the variable domain- γ in (Vγ)9Vδ2 T cells showed a direct correlation to increased cell migration of the to the inflamed skin (Laggner et al., 2011).

Increased frequency of γδT cells that express IL-17A were reported in patients with active AS, ReA and enthesitis, compared to healthy donors (Kenna et al., 2012). Similarly, enrichment of IL-17A$^+$ γδT cells was observed in the synovial fluid from patients with other forms of SpA (ReA, PsA and uSpA) implicating their role in the tissue-specific features of SpA (Taams et al., 2018).

Mucosal-associated invariant T cells (MAIT cells) are unique innate-like T cells that are characterised by expression of an invariant T cell receptor (TCR) α-chain (Vα7.2–Jα33 in humans (Toubal et al., 2019). MAIT cells also show high expression of CD161 and MHC class I- related gene protein (MR1) restriction, as shown by Oliver Lantz's team in 2003 (Treiner et al., 2003). MAIT cells reportedly express pro-inflammatory cytokines such as IL-17, IFNγ and TNF and also the cytolytic molecule granzyme B upon stimulation with phorbol myristate acetate and ionomycin (PMA-Ionomycin) (Dusseaux et al., 2011). A potential role of IL-17 producing MAIT cells has been proposed in diseases of the SpA family. For example, patients with AS exhibited an increased frequency of IL-17$^+$ MAIT cells in peripheral blood compared to healthy controls (Gracey et al., 2016; Hayashi et al., 2016). Another group reported low frequencies of IL-17-producing CD8$^+$CD161$^+$Vα7.2$^+$ MAIT cells together with IL-17 producing CD8$^+$ T cells in synovial fluid derived from PsA patients (Menon et al., 2014). Additionally, IL-17$^+$ CD8$^+$ MAIT cells have been detected in psoriasis patients from skin and blood samples. These studies provide insights into the variability in MAIT cell frequency in blood and target tissues and the increased activation of MAIT cells that might be important in disease mechanism.

4.6.3 Adaptive immune cells

CD4$^+$ T cells of the T$_h$17 phenotype produce IL-17 along with other proinflammatory cytokines like IL-6, IL-22, and IFNγ. HLA-B27 homodimers can bind to KIR3DL2 and promote the differentiation of KIR3DL2$^+$ CD4$^+$ T cells that produce IL-17A (Bowness et al., 2011). IL-17 producing CD4$^+$ T cells have been detected at higher frequencies in the peripheral blood of patients with AS and RA, compared to that of healthy donors as reviewed by Shen, Goodall & Gaston, 2009. In AS patients, IL17A$^+$ CD4$^+$ T cell percentages in peripheral blood correlated with disease activity (Al-Mossawi et al., 2019). The presence of IL-17A$^+$ CD4$^+$ T cells has also been reported in synovial fluid from patients with PsA, ReA and AS (Al-Mossawi et al., 2017; Shen, Goodall & Hill Gaston, 2010). These additional findings support the potential involvement of T$_h$17 cells in AS.

CD8$^+$ T cells are involved in the adaptive arm of the immune system that differentiate into effector and cytotoxic T lymphocyte subsets. Effector CD8$^+$ T cells secrete pro-inflammatory cytokines like TNFα, IFNγ and IL-17. IL-17$^+$ CD8$^+$ T cells have been described in inflammatory diseases (Srenathan, Steel & Taams, 2016). The first report of such a phenotype was through *IL17* mRNA expression detection in CD8$^+$ T cell clones derived from skin lesions of patients with Pso. This phenotype was also described in other inflammatory diseases, such as MS and PsA. In AS patients with an advanced form of the disease, high frequencies of IL-17$^+$ CD8$^+$ T cells were detected in the peripheral blood (Taams et al., 2018). Additionally, an increased number of IL-17 producing CD8$^+$ T cells has been reported in the synovial fluid from patients with AS and PsA, when compared with the cell numbers in peripheral blood (Menon et al., 2014). This phenotype of CD8$^+$ T cells might be involved in the HLA class I associated mechanism of AS pathogenesis.

Furthermore, SNPs identified in the aminopeptidase genes involved in MHC class I peptide trimming – *ERAP1* and *ERAP2*, and *RUNX3*, a vital transcription factor for CD8$^+$ T cell differentiation have also been associated with PsA, AS and Pso. The location of the SNPs associated to AS were enriched in epigenetic marks of active transcription in immune cells, including CD8$^+$ T cells (Li et al., 2017). Vecellio et al. demonstrated that variants in the *RUNX3* locus that coincided with epigenetic marks for enhancer region (H3K4Me1) affected binding of the transcription factor IRF4 and *RUNX3* expression, suggesting the role of CD8$^+$ T cells in AS (Vecellio et al., 2016).

Cytotoxic activity is another possible mechanism for CD8$^+$ T cell associated pathogenicity in AS. The cytotoxic activity of CD8$^+$ T cells is mediated through perforins and granzymes, and the activation of the Fas/FasL pathway. Altered cytotoxic activity in AS patients has been reported recently by Gracey et al., who observed a reduced expression of cytotoxicity-associated genes in whole blood samples of AS patients (Gracey et al., 2020). The reduced levels of granzyme and perforin was correlated with reduced CD8$^+$ T cell frequency in blood and an increased CD8$^+$ T cell frequency in the synovial fluid from AS patients. This study demonstrated the enrichment of cytotoxic CD8$^+$ T cells in the site of inflammation and underlined their importance in the joint inflammation observed in AS.

4.7 Overview of the therapeutic strategies in diseases of the SpA family

The ASAS/EULAR recommendations state that the first-line of treatment for patients diagnosed with active axial SpA are nonsteroidal anti-inflammatory drugs (NSAIDs). NSAIDs, including selective cyclooxygenase-2 (COX-2) antagonists, are effective in reducing inflammation, pain and stiffness exhibited by AxSpA patients. They are administered either continuously or on demand, based on the appearance of flares in patients. According to reported clinical studies, nearly 10% of the patients achieve partial remission with NSAIDs. Patients show better response rate to NSAIDs if the diagnosis and treatment starts early in the course of their disease (Sieper et al., 2015). Disease-modifying anti-rheumatic drugs and glucocorticoids (DMARDs) are a class of drugs that include methotrexate and sulfasalazine compounds that have been shown to inhibit radiographic disease progression, but do not aid in the management of the axial manifestations of SpA. However, DMARDs might help in treating peripheral form of SpA (Joachim Sieper & Poddubnyy, 2016). Systemic glucocorticoid therapy is not recommended in the axial disease, as they are less efficient in AxSpA compared to other inflammatory diseases.

Patients who show poor response to NSAIDs or exhibit other contraindications or intolerance for NSAIDs have the option of adopting biological DMARDs (bDMARDs) such as TNFα-inhibitors (TNFi) or monoclonal antibody against IL-17A (Secukinumab) (Van Der Heijde et al., 2017). TNFi are considered the first-line of bDMARDs in AxSpA and PsA. There are five TNFi that are currently available: infliximab (IFX), etanercept (ETN), adalimumab (ADA), golimumab (GOL), and certolizumab pegol (CZP). TNFi have brought about remarkable progress in AxSpA patients' treatment, with high efficacy despite the presence of radiographic modifications in patients. TNFi, specifically monoclonal antibodies, have also shown to improve peripheral arthritis, enthesitis, and extra-articular manifestations like uveitis, Pso and IBD. TNFi improve functional and spinal mobility, along with substantial reduction of C-reactive protein levels (CRP), and increase the overall quality of life and productivity in patients (Braun et al., 2018). Anti-IL-17A (Secukinumab) is effective in patients presenting active AS (radiographic form of AxSpa) and PsA, reducing the common signs and symptoms and by slowing the radiographic progression (Blair, 2019). IL-17 inhibition has now been introduced in clinical practice and shows similar levels of clinical efficacy as TNFi across most clinical manifestations, including peripheral arthritis, axial disease, dactylitis and enthesitis (McGonagle et al., 2019).

However, a major challenge lies ahead in teasing apart the biological mechanisms behind the success of IL-17A inhibitors. The IL-23/IL-17 axis has been previously implicated in SpA/AS pathology, with evidence showing that IL-23 enhances IL-17A production (Sherlock et al., 2012). Therefore, IL-23 inhibition was expected to show similar results as IL-17A inhibitors. On the contrary, IL-12/IL-23 inhibition with ustekinumab in AxSpA was discontinued due to poor efficacy, although it was effective in treating patients with Pso (Papp et al., 2008). These opposing results in the efficacy of IL-17A and IL-23 inhibitors in AS might suggest that IL-17A in the context of AS is produced and maintained during the inflammatory response in an IL-23-independent manner (McGonagle et al., 2019).

GOALS OF THE PROJECT

PART I | Linking GWAS variants in the susceptibility loci with nearest gene function

GWAS have successfully identified loci associated to several chronic inflammatory diseases over the past decade. There is evidence of considerable overlap in the loci linked to susceptibility to a wide range of immune disease. A susceptibility locus consists of the top significant variant from the GWAS that is associated to a trait/disease and the gene that is in close proximity to the variant. However, for the majority of the associations, there is no mechanistic evidence that links the SNPs and gene in the susceptibility locus to the disease. Attempts to assign causal relationships for the GWAS variants have been successful for a fraction of the variants. Translating GWAS variants to function is challenging as ~90% of the GWAS variants lie in the non-coding region of the genome. Furthermore, it is not well-established whether the susceptibility genes have a function in cell types that are relevant in disease biology.

The first part of my thesis tackles these issues by investigating the following questions in an inflammatory disease model: axial spondyloarthritis (AxSpA)

1. Are genes in the susceptibility loci of CIDs expressed in immune cells? In which immune cell populations are they expressed?
2. Do variants associated to CIDs affect the nearest genes in the susceptibility loci?
3. How do variants associated to CIDs affect the nearest gene function in immune cells?

PART II | Innate and adaptive IL-17A producing cells in axial spondyloarthritis

The genetic association studies in AxSpA revealed several variants in genes of the IL-23/IL-17 axis- *IL23R*, *IL12B*, *IL6R*, etc., suggesting a possible role in disease pathogenesis. IL-17A-blockade has been used successfully in AxSpA over the past 6 years, however the cell populations that serve as the main source of IL-17A remain uncharacterised. Surprisingly, the failure of anti-IL-23 in AxSpA suggests a biological mechanism that leads to IL-17A-mediated inflammation in AxSpA, independent of IL-23. This emphasises the importance of identifying cell populations producing IL-17A in AxSpA, and the mechanisms of IL-23 independent induction of this cytokine.

The second part of my thesis addresses the question: What are the IL-17A producing immune cell populations in AxSpA?

The submitted manuscript is appended in **Chapter 9** and the **Annex II**.

PART III | Mechanism of action of TNF-blockers in AxSpA and immunological correlates of responsiveness to TNFi.

TNF inhibitors (TNFi) are highly effective in reducing inflammation and associated clinical symptoms in AxSpA. However, 30-40% of the patients respond poorly to anti-TNF therapy. There is very little known about how TNFis affect immune responses in patients, and there are no methods currently in clinical practice that can identify the patients who will respond to TNFi before they undergo the therapy. Hence there is a need to improve our understanding of how the molecular pathways affected by TNFis and to develop strategies to guide the choice of therapies for patients affected by AxSpA.

Part III discusses the results from the collaborative study investigating 1) the mechanism of action of TNF-blockers in AxSpA and 2) the immunological correlates of responsiveness to TNFi. The manuscript currently in revision is appended in **Chapter 10** and the **Annex II**.

Chapter 5

MATERIALS AND METHODS

5.1 Patient cohort description

This project was conducted as a part of the collaborative effort between the Unité Mixte Recherche at Institut Pasteur and the Service de Rhumatologie B, headed by Pr Maxime Dougados, at Hôpital Cochin, Paris. A cohort of patients with axial spondyloarthritis (AxSpA) was recruited from our collaborators. Whole blood samples of 50-60 mL volume were collected from each patient in Vacutainers with Lithium-Heparin anti-coagulating agent (BD Vacutainer® Safety-Lok™, Becton-Dickenson, Franklin Lakes, NJ, USA).

This study was performed by fulfilling the current Good Clinical Practice Guidelines and with the approval of the Ethical Review Committee (Comité de Protection des Personnes Ile-de-France III). All the participants included in the study gave their written informed consent prior to sample collection. The patients recruited met with either modified New York criteria for ankylosing spondylitis or the Amor criteria for spondyloarthritis (Amor, Dougados, & Mijiyawa, 1990; Linden, Valkenburg, & Cats, 1984). Patients exhibiting inflammatory back pain symptoms in the buttock, lumbar or thoracic spine over 3 months and other symptoms suggestive of spondyloarthritis were included in the study. The age range of patients was >18 years and <70 years.

Patients who showed other spinal diseases (e.g. discarthosis), history of any biological treatments or history of other conditions that might interfere with the criteria mentioned in the informed consent, such as, psychological disorders, substance addiction, etc. were excluded from the study. Patients were prescribed corticosteroids (prednisone) only at a lower dose (<10mg) per day for at least weeks prior to the blood sample collection.

Records of patient characteristics including age, sex, disease duration, HLA-B27 positivity, ongoing treatments (e.g. analgesics, NSAIDs, DMARDs including biologics, physiotherapy), co-morbidities, and other main clinical features of SpA (acute anterior uveitis, psoriasis, inflammatory bowel disease, enthesitis, peripheral articular involvement) were collected on a Case Record Form during each visit. Erythrocyte Sedimentation Rate (ESR) and C-Reactive Protein (CRP) were collected at the baseline. CRP measurements were quantified from sera collected on the first visit using the high-sensitivity CRP (hs-CRP) test. Disease activity score was calculated as per the criteria dictated in the Ankylosing Spondyltitis Disease Activity Score (ASDAS) (Machado et al., 2011). Cholesterol (HDL, LDL) levels and blood count were collected during visits, along with systematic radiological evaluation to monitor disease progression by X-rays and MRI of the spine and pelvis.

5.2 Blood sample collection and processing

5.2.1 Whole blood TruCulture stimulation system

TruCulture systems were manufactured in accordance with EN ISO 13485 (Medical Device Directive) standards, at EDI GmbH (Reutlingen, Germany), a subsidiary of Myriad RBM (Austin, Texas). The same manufacturing batch of TruCulture tubes has been used for all the samples processed in this study. TruCulture tubes were stored at -20°C and thawed at room temperature (RT) 30 minutes before usage, or thawed at room temperature for 10 minutes and pre-warmed for 10 additional minutes at 37°C.

Venous blood was collected from patients in sterile Lithium-Heparin vacutainers (BD Vacutainer®, Becton-Dickenson, Franklin Lakes, NJ, USA). Seven vacutainer tubes, containing 7 mL of blood, were collected for all AxSpA patients. Of the total 50 mL of the collected blood, 21 mL were used for the TruCulture assays study. Only three TruCulture stimulation conditions were utilized for this project (Part I). The rest of the conditions were used for the anti-TNF study described in Part III. The rest of the blood was used for isolating peripheral blood mononuclear cells (PBMCs). All TruCulture tubes and tubes for sample collection were identified with cryogenic labels prior to start of sample processing. Each tube was identified by patient ID (serial number), time point, stimulus pre-loaded in the TruCulture assay and date of sample collection.

Within 30 minutes of blood sample collection, 1mL of the whole blood was distributed using an electronic pipette distributor into the each of pre-warmed TruCulture tubes with an appropriate stimulus. The samples were mixed well with the medium and stimulus present in the tube by inverting the them several times. The tubes were then incubated using a bench-top heating block at 37°C (± 1°C), for 22h (± 15 min). After the incubation period, a filter system with a plunger was inserted to separate the sedimented cell pellet and the supernatant. Supernatant was harvested and 3 aliquots of 400 µL were prepared with an electronic adjustable tip spacing multichannel equalizer pipette and stored at -80°C until the time of use. Cell pellets were re-suspended in 2 mL of TRIzol™ LS reagent (Sigma) and mixed well with a vortex mixer for 5 minutes and stored at -80°C until the time of RNA extraction.

The stimuli used for this system were lipopolysaccharide (LPS at 10ng/ml) and staphylococcus enterotoxin B (SEB at 0.4 µg/ml) stimulation and the negative control (∅) as per the description in (Duffy et al., 2014). We utilised 25 patient samples in the three conditions

to validate the gene expression panel (Autoimmune disease consortium panel) we designed with NanoString.

5.2.2 Experimental work flow for PBMC isolation and cell culture

5.2.2.1 PBMC isolation from patient blood samples

Venous blood samples were collected from patients in sterile Lithium-Heparin tubes (BD Vacutainer®, Becton-Dickenson, Franklin Lakes, NJ, USA) for peripheral blood mononuclear cells (PBMC) isolation. PBMCs were separated from whole blood diluted in 1:3 volumes with room temperature Phosphate Buffered Saline (PBS, Fisher Scientific) using Lymphocyte separation medium (Eurobio, France) and density gradient centrifugation. The layer of PBMCs was collected and was transferred into a new Falcon® tube and then centrifuged with cold PBS to remove platelets. Cells to be counted were re-suspended in a known volume of PBS. The cells were then counted on a Neubauer chamber slide using 0.4% Trypan Blue, using the formula given below:

$$\text{Total \# of cells} = \text{Mean cell count} \times \text{dilution factor} \times \text{volume (ml)} \times 10^4$$

5.2.2.2 Healthy donor samples

Human mononuclear cells were isolated by Lymphocyte separation medium by density gradient centrifugation (Eurobio, France) from buffy coat samples (EFS Hôpital Saint-Louis).

5.2.2.3 Cell purification by magnetic separation: monocytes and NK cells

The isolated PBMCs were labelled with human CD14 or CD56 (MicroBeads (Miltenyi Biotec), 20μL beads for 10^7 cells suspended in separation buffer (1x PBS + 0.5% FCS + 2mM EDTA) of appropriate volume (80μL for 10^7 cells) and were incubated for 15 minutes at 4°C on a rotor. After the incubation period the magnetic separation was carried out using the MS columns following the kit protocol. The positive fraction was counted using the same method described before and then cultured in 48-well plates. The purity of monocytes was over 97% as verified by flow cytometry (LSR II, BD Biosciences). The $CD14^-$ fraction was used to isolate the T cell subsets with fluorescence-activated cell sorting (FACS).

5.2.2.4 Fluorescence-activated cell sorting (FACS) for T cells

The counted $CD14^-$ fraction was stained in the dark at 4°C for 20 minutes in FACS buffer (1x PBS + 1%FCS) with anti-human CD3-APC (Clone BW264/56, Miltenyi Biotec®), anti-human CD4 APC-Vio770 (Clone M-T466, Miltenyi Biotech®) and anti-human CD8- PerCP Vio 700 (BW135/80 Miltenyi Biotec®). For the experiments on healthy blood donors, we performed first a CD3 positive magnetic separation using anti-CD3 monoclonal antibodies coated beads (Miltenyi Biotec®). The CD3 positive fraction was labelled with CD3-BV711 (BD Horizon) to sort the T cell subsets. T cells were purified from healthy donor samples using the respective markers: CD4-BV750 (clone: SK3, BD Horizon™), CD8- PerCP-Vio700 (Miltenyi Biotec®), γδT cells (TCRgd- PE (clone: 11F2), TCRd2- PE (clone: 123R2)) and MAIT cells (Va7.2- APC (clone: 3C10), Biolegend® and CD161-BV421 (clone: HP3G10), Sony Biotechnology®). Cells were purified on a FACS Aria II (BD Biosciences, San Jose, CA) at the Center for Translational Science (CRT)/Cytometry Biomarkers Unit of Technology and Service (CB UTechS) at Pasteur Institute in Paris. The purity of each population was verified with post-sort analysis. Cells were collected into 5mL FACS tubes coated with medium.

5.2.2.5 Cell culture conditions

The purified cells were pelleted at 3500 rpm for 5 minutes. For the unstimulated sample, cells were immediately lysed in 350µL of Buffer RLT plus (Qiagen RNeasy Micro Kit, Valencia, CA). The stimulated samples were cultured in 1ml of pre-warmed medium Roswell Park Memorial Institute (RPMI) 1640 (Gibco Life Technologies, Oslo, Norway) + 5% FCS (HyClone, Fisher Scientific) + Penicillin/Streptomycin and the respective stimuli in a 5% CO_2 incubator for later time points. $CD14^+$ Monocytes were cultured in 48-well plates (500,000 cells/400 or 500 µL, respectively) in pre-warmed RPMI 1640 medium (Invitrogen) without FCS and antibiotics, and LPS from *Escherichia coli* (LPS, Invivogen) at 20ng/ml concentration for 16h at 37°C. T cells were cultured with anti-CD3/anti-CD28 human T activator Dynabeads (Invitrogen) at a bead:T cell ratio of 1:4 at 37°C. Cell lysates were stored at -80°C until RNA extraction. Cells isolated from healthy donor samples were stimulated as per the following conditions: NK cells activated with IL-18 (50ng/mL), monocytes stimulated with LPS from *Escherichia coli* (LPS, Invivogen, at 20ng/mL), T- cells (CD4, CD8 and γδT cells) were activated with IL-2 (20U/mL) and TCR-stimulation (Human T Cell TransAct™, Miltenyi Biotec®). MAIT cells were activated with a combination of cytokines (IL-2 (20U/mL) + IL-1β (10ng/mL) + IL-23 (10ng/mL) and the tetramer MR1/5-OPRU (0.2nM, NIH Tetramer Core Facility).

5.3 Gene expression Analysis

5.3.1 RNA extraction

TruCulture RNA samples were extracted in a randomised manner to avoid batch effects. Cell pellets in TRIzol LS were thawed on ice for 30 minutes. The thawed samples were mixed vigorously with a vortex mixer for 5 minutes at 2000 rpm to dislodge the frozen pellet and to resuspend it and to lyse it homogenously. The phenol-chloroform method (Chomczynski & Sacchi, 1987) was used to proceed with the RNA extraction. To the 600 µL of TRIzol lysate, 150 µL of chloroform (Merck) was added and mixed thoroughly using a vortex mixer. The samples were centrifuged for 15 minutes at 14000g at 4°C to enable phase separation. After centrifugation, the upper aqueous phase was transferred to a clean Eppendorf tube and 300 µL of isopropyl alcohol (Sigma) was added to precipitate the RNA. The RNA was then washed with 70% ethanol twice and then dissolved in 30 µL RNase-free water. Aliquots of 4 µL were prepared for quality control measures to avoid repeated freeze-thaw cycles. All the extracted RNA samples were stored at -80°C until use.

Total RNA from cell populations $CD4^+$ T cells, $CD8^+$ T cells, γδT cells, MAIT cells, NK cells and monocytes from patients and controls, was isolated using RNeasy Micro Kit or RNeasy Mini Kit (Qiagen RNeasy Micro Kit, Valencia, CA), based on the cell numbers obtained. Extraction was performed following the protocol provided by the manufacturer. RNA concentration was estimated using Qubit RNA HS Assay Kit (Life Technologies, Grand Island, New York, USA) according to the manufacturer's instructions.

5.3.2 RNA quantification and quality control

RNA quantification was done using Qubit RNA HS Assay Kit (Life Technologies, Grand Island, New York, USA) according to the manufacturer's instructions on the Qubit 2.0 fluorometer. RNA quality was assessed using RNA 6000 Nano Kit on the Agilent BioAnalyzer 2100 system (Agilent Technologies, Palo Alto, CA). The RNA integrity number (RIN) was determined using the LabChip System software. The quality threshold was set at RIN > 8.

5.3.3 NanoString AID panel design

We designed a gene expression assay in collaboration with NanoString Technologies called Autoimmune Discovery Consortium Panel (AID panel) with 755 genes. This panel comprises 518 genes associated to 9 different autoimmune and autoinflammatory diseases, namely, MS, RA, SLE, T1D, AS, Pso, CD, UC and IBD (gene list included in **Annex I**). These

genes were selected after reviewing summary statistics of published GWA studies curated by Open Targets Genetics platform- previously maintained by the ImmunoBase (https://genetics.opentargets.org/immunobase). All the disease-associated genes in the panel have been reported to be the nearest gene in the susceptibility locus for an immune disease. We also included 237 genes related to immune response and 15 internal reference genes.

5.3.4 Gene Expression Analysis with NanoString nCounter® Technology

NanoString nCounter® technology utilises direct molecular barcoding to detect target messenger ribonucleic acid (mRNA) molecules in a sample. The technology is multiplexed, and allows simultaneous detection and quantification of up to 800 mRNA molecules without the amplification step required in conventional gene expression technologies. The probe pair consists of a biotinylated Capture probe at the 3' end, and a Reporter probe, which carries the barcode on its 5' end. The Capture probe has a 35-50-base sequence that is complementary to a target mRNA molecule, along with a short common sequence coupled to the biotin. The Reporter probe consists of a second 35-50-base complementary sequence to a the target mRNA coupled to a colour-coded tag that emits the detection signal. The colour code is made up of six positions and each position can be one of four colours, enabling a combination of thousands of barcode combinations. Unique pairs of Capture probe and Reporter probe combinations are created in this manner to measure mRNA transcripts for each gene of interest (Geiss et al., 2008).

The hybridisation step was carried out in 12-tube PCR strips. RNA samples were diluted with RNase-free water to have a total concentration of 100ng in 5µL. A mix of hybridisation buffer and Reporter probes were added to the diluted RNA, followed by the addition of Capture probes. The contents of the tubes were mixed gently and transferred to a thermocycler set at 65°C for 22h. During the hybridisation reaction a tripartite structure with the target mRNA bound to the specific Reporter and Capture probes is formed. After the hybridisation step, the samples were transferred to the Prep Station for further processing steps to remove excess and unbound probes by affinity purification. This step is entirely automated to avoid variabilities between users. The remaining tripartite complexes are then immobilised and aligned onto a cartridge coated with streptavidin. The target mRNA of interest is identified by the colour code generated by the fluorescent barcode on the Reporter probe. The expression level is measured by counting the number of barcodes for each target mRNA as per the technology described by Geiss et al., 2008. Cartridges were scanned using the nCounter® Digital analyser at the highest resolution of 555 fields of view (FOV) per flow cell to obtain a Reporter Code Count (RCC) data set.

nCounter® assays were performed at the CB UTechs technology core facility of the Pasteur Institute (Paris, France).

5.3.5 Gene expression data analysis

Each sample was analysed in a separate multiplexed reaction including in each, eight negative probes and six serial concentrations of positive control probes. Negative control analysis was performed to determine the background level for each sample. We generated raw data (RCC files) from whole blood TruCulture samples stimulated with LPS and SEB and the negative control (n=25), $CD8^+$ T cells (n=49), $CD4^+$ T cells (n=55) and monocytes (n=45) under resting and activated states using the NanoString AID panel. Quality control and data normalisation was carried out as per the established pipeline provided by NanoString using nSolver Analysis software (version 3.0) using three steps described here: (1.) positive control normalisation, (2.) negative control normalisation and (3.) internal reference (housekeeping) gene normalisation. (1.) The positive control normalisation was used to correct for technical variation by calculating the geometric mean of the positive probe counts of each sample and then by calculating the average geometric mean of the positive probes across all the samples. A scaling factor was calculated by dividing the average geometric mean by the geometric mean for each sample. The computed scaling factor was then multiplied to the corresponding positive and negative controls and the target gene counts to correct for technical variation. (2.) Negative control normalisation, a background threshold was calculated by taking the mean of the negative controls + 2 standard deviations (SD) and then this value was subtracted from the gene counts. (3.) The variation in mRNA input was corrected using the same method as the positive control normalisation, but the geometric means were calculated for the selected housekeeping genes.

Housekeeping genes were selected using geNorm algorithm (Vandesompele et al., 2002) built-in to the nSolver analysis software. Using this method we selected the following housekeeping genes: *POLR2A*, *SDHA* and *TBP* for whole blood data, *GUSB*, *RPL19* and *SDHA* for $CD4^+$ and $CD8^+$ T cell data set; and *OAZ1*, *HPRT1* and *PPIA* for monocyte data set. Probes with low counts were determined with respect to the background level, defined as the mean of the negative controls + 2*SD. All the 755 genes including low count genes were used for the eQTL analysis. The selected housekeeping genes for innate and adaptive cell type data from healthy donors data were the following: *G6PD*, *GUSB*, *OAZ1* and *RPL19*. MAIT cells were normalised separately using *ALAS1*, *HPRT1*, *PP1A* and *EEF1G*.

5.4 Genotyping data

All patient samples were genotyped using the Illumina Global Screening Array (GSA), in collaboration with the genotyping platform at the Centre National de Recherche en Génomique (CNG) in Ivry. The genotyping data analysis was done in collaboration with the Center of Bioinformatics, Biostatistics and Integrative Biology (C3BI) at the Pasteur Institute. Genotyping data were obtained were subjected to standard quality control steps and all the samples passed the initial quality control done by the CNG. There were two sibling pairs in the cohort which were excluded There were no sex inconsistencies between the genotyping data and the clinical information. Imputation was performed using IMPUTE2 (2.3.2) (Howie, Donnelly, & Marchini, 2009). and the 1000G Phase 3 reference panel, based on 250 kb windows and a buffer region of 250 kb. We removed imputed SNPs with information metric < 0.8 as well as imputed SNPs with > 5% missing data (individual genotype probabilities < 0.8 were considered missing data).

5.4.1 SNP selection process for functional analysis

We compiled a list of SNPs based on the results from the eQTL analysis from the T cell data set ($CD4^+$ and $CD8^+$ T cells under resting and activated states). All the eQTLs indicated a region on chromosome 11 around the gene *CTSW*. We queried if the eQTLs from our analysis were also SNPs from association studies in immune diseases, specifically in ankylosing spondylitis. This step was done by consulting the online resources: GWAS Catalog (https://www.ebi.ac.uk/gwas/) and Open Targets Genetics (https://genetics.opentargets.org/). Linkage disequilibrium (LD) between eQTLs and GWAS variants was determined using the LDpair and LDmatrix settings from LDlink interactive tool (https://ldlink.nci.nih.gov/?tab=home) in the European population (EUR: CEU, TSI, FIN, GBR and IBS), indicated in R^2. The LD threshold was set at $R^2 > 0.8$.

To explore the regulatory information of the given locus. The SNP positions were marked using the SNP database dbSNP 151 track available on the UCSC Genome Browser version GRCh37/hg19 (http://genome.ucsc.edu/cgi-bin/hgGateway). The SNP locations were overlapped with data tracks for regulatory element marks- histone marks, DNase hypersensitivity sites using the “track hubs” option embedded in the genome browser. When the SNPs overlapped with histone marks indicating open chromatin conformation, we checked if there were also transcription factors binding sites (TFBS) in the given position using a combination of three different online tools. Firstly, we used HaploReg v4.1 (https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php) to assess if there were

published data on TFBS and altered motifs by giving the variant information. Secondly, within the UCSC genome browser we plugged in the JASPAR TFBS track with profiles of transcription factor predictions from the JASPAR CORE collection. Thirdly, we utilised Promo v3 to predict the possible TF binding sites by providing the sequence of the locus 10 bases upstream and downstream from the location of the SNP of interest. We compiled the three list of TFs based on the SNP of interest and cross-referenced the expression of these factors in immune cell types from public RNA-Seq profiles curated in DICE eQTLs (https://dice-database.org/) (Schmiedel et al., 2019) and the Human Blood Atlas (part of the Human Protein Atlas) (https://www.proteinatlas.org/humanproteome/blood). Based on the TF list and the expression pattern in cell types we tested the TFs for allele-specific binding pattern in relevant immune cell subsets.

5.4.2 eQTL analysis

The normalised gene expression from the cell type-specific data set and genotyping data were used to carry out expression quantitative trait locus (eQTL) mapping, using the R package MatrixEQTL (Shabalin, 2012). The aim was to correlate existing polymorphisms in regions surrounding SNPs of interest (i.e., variants mentioned in the literature as associated with AS) with expression level of disease-associated genes from the AID panel. Sex, HLA-B27 status, cohort and ethnicity (European vs non-European background) were included as covariates in the analysis. We selected the variants post-imputation based on the loci associated to AS and analysed 68442 SNPs for genotype-gene expression correlation. Expression data were log-transformed.

Two types of associations were investigated: (i) associations between SNPs and expression levels of genes that were at a genomic distance < 250kb from SNPs (henceforth, cis-eQTL); (ii) associations between SNPs and expression levels of genes that were at a genomic distance > 250kb from SNPs (henceforth, trans-eQTL). P-values were further adjusted to account for multiple testing using permutations. The observed effect size is denoted as Z- score, which is a standardised estimate (i.e., the slope of the regression divided by the level of variation observed in both variables, gene expression and genotypes). We calculated the estimates of effect size S_{cis} and S_{trans} based on the permuted data. In this method, rows of gene expression tables were permuted in order to simulate the distribution of statistic parameters $S_{cis\text{-}max}$ and $S_{trans\text{-}max}$ in under the null hypothesis (no genuine association between SNPs and gene expression). These parameters are standardised estimates of the maximum effect size (of SNPs on gene expression) across the genome in both types, cis- and trans-associations. We ran regressions on the data (gene expression as a function of

genotypes) and recorded the estimates (observed values) $S_{cis-obs}$ and similarly, $S_{trans-obs}$. Then, the distribution of the highest effect size across the genome $S_{cis-max}$ and $S_{trans-max}$ were obtained with the permuted data. This process was repeated over a large number of permutations (i > 10,000) to generate a simulation of the empirical distribution of the S_{max} estimates. This distribution was then used to compute adjusted P-values as: the number of times where $S_{cis-max}$ is higher than $S_{cis-obs}$ in *cis*-associations, divided by the total number of permutations. The same process was performed for computing $S_{trans-max}$ and $S_{trans-obs}$ in the *trans*- analysis. This method was applied to each cell type data by stimulation condition (resting and activated).

5.5 Affinity capture assay for DNA binding proteins

Oligonucleotides were designed with 10 bases upstream and downstream of the location of the two SNPs of interest to test for preferential binding. The oligonucleotides were synthesised and purchased from Eurofins Biotech (**Table 3**). The reverse complement of the oligonucleotides were tagged with biotin. Both the forward strand and the biotinylated reverse complement strand were annealed in TE plus 100mM NaCl buffer at 95°C to the form the biotinylated double-stranded (ds) oligonucleotides.

Total cell extracts were prepared from $CD8^+$ T cells (5 $x10^6$ cells per reaction) lysed in extraction buffer (20 mM Hepes/NaOH pH 7.9, 0.35 M NaCl, 0.1% Nonidet P-40, 1 mM MgCl2, 0.5 mM EDTA, 20% glycerol and proteinase inhibitors), using a vortex mixer, followed by high speed centrifugation at 4°C for 5 minutes. The lysate was removed and the salt content was adjusted to 0.35M using the dilution buffer (20mM HEPES, 20% glycerol and proteinase inhibitor), followed by the addition of the respective biotinylated ds-oligonucleotides (2 pmol) and 2μg poly(dI-dC). The lysates were then incubated for 45 minutes at 4°C to enable binding. After the incubation, 20μL Dynabeads™ M-280 Streptavidin (Thermo Fisher Scientific) pre-incubated with 2% BSA, was added to pull down the bound protein and then incubated for 15 min at 4°C. The samples were then washed 4 times with ice-cold wash buffer (25 mM Hepes/NaOH pH 7.9, 0.15 M NaCl, 0.1% Nonidet P-40) and once with ice-cold PBS. The bound complex was then eluted in 2X sample buffer and denatured at 95°C for 5 minutes.

The protein samples were separated on SDS gels by Western blot and transferred to PVDF membrane (Amersham). The membranes were probed with anti-NFATC2 (4G6-G5, Santa Cruz biotechnologies; concentration:1μg/mL) and anti-RBPJK (E7, Santa Cruz

biotechnologies; concentration:1µg/mL). Proteins were visualised using ECL (SuperSignal™ West Pico PLUS Chemiluminescent Substrate (Thermo Fisher Scientific) on the ImageQuant LAS 4000mini chemiluminescence analyser (Fuji) at an appropriate exposure.

Table 3: Oligonucleotides used for high affinity DNA capture assay

rs658524_ref	Oligo 1	Forward primer	GTTTCTCCAGCCCCACCCTCACACCGTCCT
		Reverse primer	AGGACGGTGTGAGGGTGGGGCTGGAGAAAC
rs658524_alt	Oligo 2	Forward primer	GTTTCTCCAGCCCCGCCCTCACACCGTCCT
		Reverse primer	AGGACGGTGTGAGGGCGGGGCTGGAGAAAC
rs12225345_ref	Oligo 3	Forward primer	TCCAACCTGGGTGGAAAAAAAAAAGAAAGA
		Reverse primer	TCTTTCTTTTTTTTTCCCACCCAGGTTGGA
rs12225345_alt	Oligo 4	Forward primer	TCCAACCTGGGTGGGAAAAAAAAAGAAAGA
		Reverse primer	TCTTTCTTTTTTTTTCCCACCCAGGTTGGA

5.6 Generation of reporter gene constructs for regulatory SNP rs12225345

To measure the activity of NFATC2 binding on transcription in the presence of regulatory SNPs in the *CTSW* promoter region, we cloned a fragment of the promoter region containing the reference and alternative allele of rs12225345. The sequence was doubled to create a long oligonucleotide of 104 bp with two sites of the SNPs to increase the sensitivity of the assay as shown in table 4. The oligos were cloned upstream of the luciferase gene into the pGL4.23 vector (Promega) with a minimal promoter, using the restriction sites BglII and NheI. Correct cloning was verified by sequencing.

Table 4: Primers used for reporter gene assay

rs12225345_ref	Long (2 sites for SNP)	Forward primer	5'-CTAGCGCCACTGCACTCCAACCTGGGTGGAAAAAAAAAAGAAAGA AAGAAAGAAGCCACTGCACTCCAACCTGGGTGGAAAAAAAAAGAAAG AAAGAAAGAAA- 3'
		Reverse primer	5'-GATCTTTCTTTCTTTCTTTCTTTTTTTTTTCCACCCAGGTTGGAGTG CAGTGGCTTCTTTCTTTCTTTCTTTTTTTTTTCCACCCAGGTTGGAGTGC AGTGGCG- 3'
rs12225345_alt	Long (2 sites for SNP)	Forward primer	5'-CTAGCGCCACTGCACTCCAACCTGGGTGGGAAAAAAAAAGAAAGAA AGAAAGAAGCCACTGCACTCCAACCTGGGTGGGAAAAAAAAAGAAAGA AAGAAAGAAA-3'
		Reverse primer	5'-GATCTTTCTTTCTTTCTTTCTTTTTTTTTCCCACCCAGGTTGGAGTGC AGTGGCTTCTTTCTTTCTTTCTTTTTTTTTCCCACCCAGGTTGGAGTGCA GTGGCG-3'

5.7 Transfection with Lipofectamine™

Kidney epithelial 293 T cells were cultured in DMEM with 10% serum. Cells were transfected using Lipofectamine™ 2000 according to manufacturer's instruction (Life Technologies). Cells were transfected in a 24 well plate with 100 ng of pGL4.23 vector of interest and 5 ng of pRL-TK vector with Renilla luciferase to normalize for transfection efficiency. Where indicated, 0.5 µg of the vector expressing NFATC2 (pEF6/His) was added. To obtain equal amounts of DNA in all samples, the appropriate amount of pBluescript vector was added to the transfected DNA. The Luciferase activity was measured 24h after transfection using Dual-Luciferase Reporter Assay System kits (Promega).

5.8 Statistical Analysis and Data Visualisation

Statistical analyses were performed using GraphPad Prism version 7 (GraphPad Software, San Diego, CA) and R Studio (v1.1.463). Unless otherwise indicated, horizontal bars represent the mean. Bar graphs, pie charts and box plots were generated using R packages ggplot2 (v3.3) and ggpubr (v0.4). Upset plots were generated using the R package UpSetR (v1.4). Principal Component Analysis (PCA), hierarchical clustering, paired and unpaired t tests were performed with Qlucore Omics Explorer version 3.0 (Qlucore, Lund, Sweden). Before applying PCA and hierarchical clustering, mRNA expression levels were log-transformed, mean-centred and scaled to unit variance.

5.9 Protein de-glycosylation assay

$CD8^+$ T cells were lysed in radioimmunoprecipitation (RIPA) buffer. Protein quantification was done using Bradford's assay. The lysate was then mixed with glycoprotein denaturing buffer (New England Biolabs, Ipswich,MA) and incubated at 100 °C for 10 min. 1µg of the target glycoprotein was incubated for protein de-glycosylation reaction using Endo H (Promega) and recombinant PNGase F (Promega), at 37°C, for 1h30min.

Chapter 6

RESULTS: PART I

6.1 Patient cohort description

This project was done in collaboration with the Service du Rhumatologie B team at Hôpital Cochin (Paris) directed by Pr Maxime Dougados. A cohort of 68 patients with axial spondyloarthritis (AxSpA) were enrolled in this study. Blood was collected from the patients to perform whole blood TruCulture assays and to isolate immune cell populations. In parallel to this, DNA samples were isolated to genotype the patients. The TruCulture assays were tested and standardised with the first 12 patients recruited in the study in the work done by Menegatti et al (in revision). The workflow of the patient recruitment and experimental set up is given in **Figure 18**. After quality control measures, we used 53 patients for further analysis.

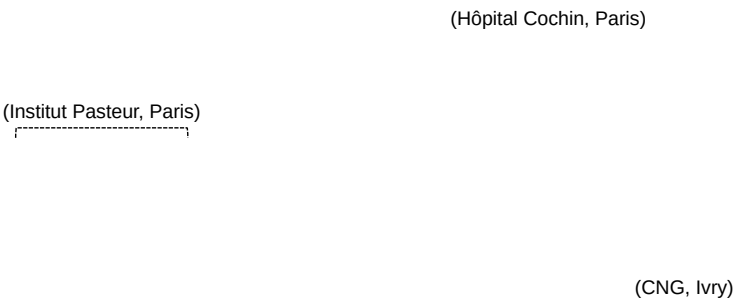


Figure 18: Axial spondyloarthritis (AxSpA) patients enrolled in the study. A cohort of patients with AxSpA were included in the study. Peripheral blood samples were obtained for all patients (n = 68) at baseline before the onset of any biologic treatment.

6.1.2 Patient Characteristics

The clinical characteristics of the patients included in the study are summarised in **Table 5**. The average age was 39±13 years (ranging from 18 to 69 years), and 74% of the subjects were male, consistently with the observed prevalence of SpA in males (Gran, Husby, & Hordvik, 1985; W. Lee et al., 2007; Ngo et al., 2014). 78% of the overall patients (n=53) were HLA-B27 positive; 72% of female and 80% of male patients were HLA-B27 positive. 49% of the patients enrolled had a familial history of spondyloarthritis and related disorders (**Figure 19**).

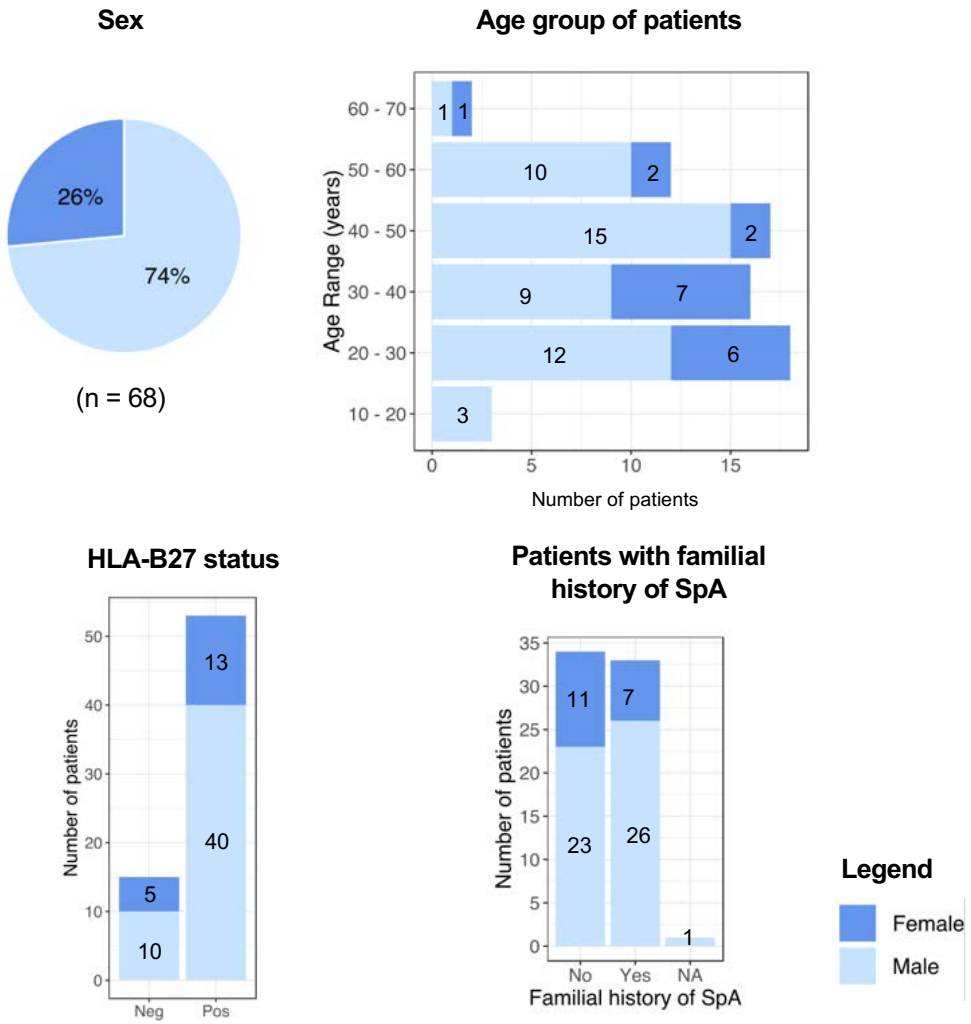


Figure 19: Demographic characteristics of SpA patient cohort recruited for the study. Each data point represents one patient at baseline (D0). Pie chart shows the proportion of males and females in the cohort. Bar chart graphs indicate the number of women and men in the respective age groups, HLA-B27 status in men and women and the number of male and female patients with familial history of SpA.

All the included patients were naïve to any biological treatment. The majority of the patients presented with radiographic features, either X-ray sacroiliitis (63%) and/or MRI sacroiliitis (88%). 47% of the patients showed elevated levels of CRP at baseline (>6 mg/L) (see **Figure 20**). Patients also presented extra-articular manifestations such as: 32% also had acute anterior uveitis (n=22), 19% had psoriasis (n=13), 4% had SpA-associated inflammatory bowel disease (n=1) and nearly half of the patients (47%) also presented with enthesitis (**Figure 20**). Prior to recruitment, 82% of the patients were treated with non-steroidal anti-inflammatory drugs (NSAIDs), 7% with DMARDs (methotrexate, n=4; sulphasalazine, n=1) and 3% with corticosteroids (**Figure 20**). Patients who were prescribed corticosteroids were

only admitted to the study if a lower dosage (<10 mg of prednisone per day) was maintained for at least 4 weeks prior to the recruitment.

Variable	Results
Number of patients	**68**
Age (years, mean ± SD)	39±13
Gender (% Female)	26.5%
Disease duration (years, mean ± SD)	6±10
HLA-B27 positive (%)	78 %
ASAS criteria (%)	100%
Radiological sacroiliitis (mNY) (%)	63%
Abnormal CRP (%)	47%
MRI – SI joints (%)	88%
Familial history of SpA (%)	49%
Current symptoms or past history of:	%
- Peripheral arthritis (%)	18%
- Uveitis (%)	32%
- Psoriasis (%)	19%
- IBD (%)	4%
- Enthesitis (%)	47%
Therapies before biologic treatment:	
- DMARDs (%)	7%
- Corticosteroids (%)	3%
- NSAIDs (%)	82%

Table 5: Clinical and demographic data of spondyloarthritis patients included in the study at baseline. Values are reported as means ± SD, or proportions (%) of total patients.

Disease activity was evaluated with the Bath AS Disease Activity Index (BASDAI) (Garrett et al., 1994) and the Ankylosing Spondylitis Disease Activity Score (ASDAS) (MacHado et al., 2011; Machado et al., 2015; Rudwaleit et al., 2009) (**Figure 20**).

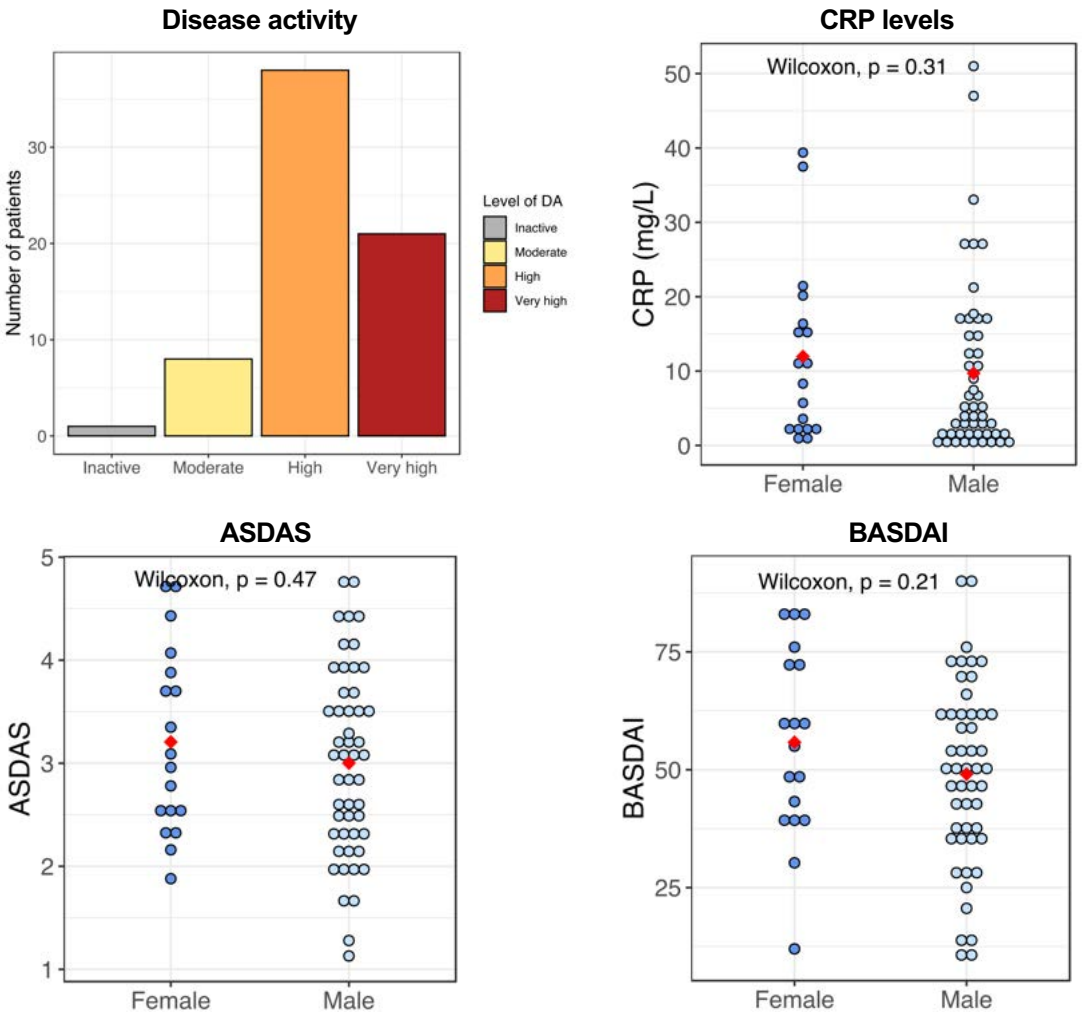


Figure 20: Clinical characteristics of the patients to assess disease status and progression. Bar plot indicates the frequency of patients with very high, high, moderate or low disease activity evaluated using the ASDAS score. Dot plots show the CRP levels measured in mg/L and disease activity scores- ASDAS and BASDAI in men and women present in the cohort. Mean CRP level and disease activity scores are marked by the red dot in the plot. P-values were calculated using Wilcoxon rank sum test between the men and women.

6.2 Expression of disease susceptibility genes identified in GWAS

To study the role of genes in the susceptibility loci reported by GWAS in immune diseases, we first asked whether these genes are expressed in immune cell populations from AxSpA patients. To this end, we designed a gene expression assay, the Autoimmune Discovery Consortium Panel, in collaboration with NanoString. For this panel, we selected 518 genes associated in GWAS studies to 9 different autoimmune and autoinflammatory diseases: MS, RA, SLE, T1D, AS, Pso, CD, UC and IBD (**Figure 21**). These genes were selected after reviewing summary statistics of published GWA studies curated by Open Targets Genetics platform, previously maintained by the ImmunoBase (https://genetics.opentargets.org/). All the genes reported in these studies were the gene nearest to the disease-associated SNP identified in the GWAS. We also included in the panel 237 genes related to immune responses, and 15 internal reference genes, for a total of 755 genes.

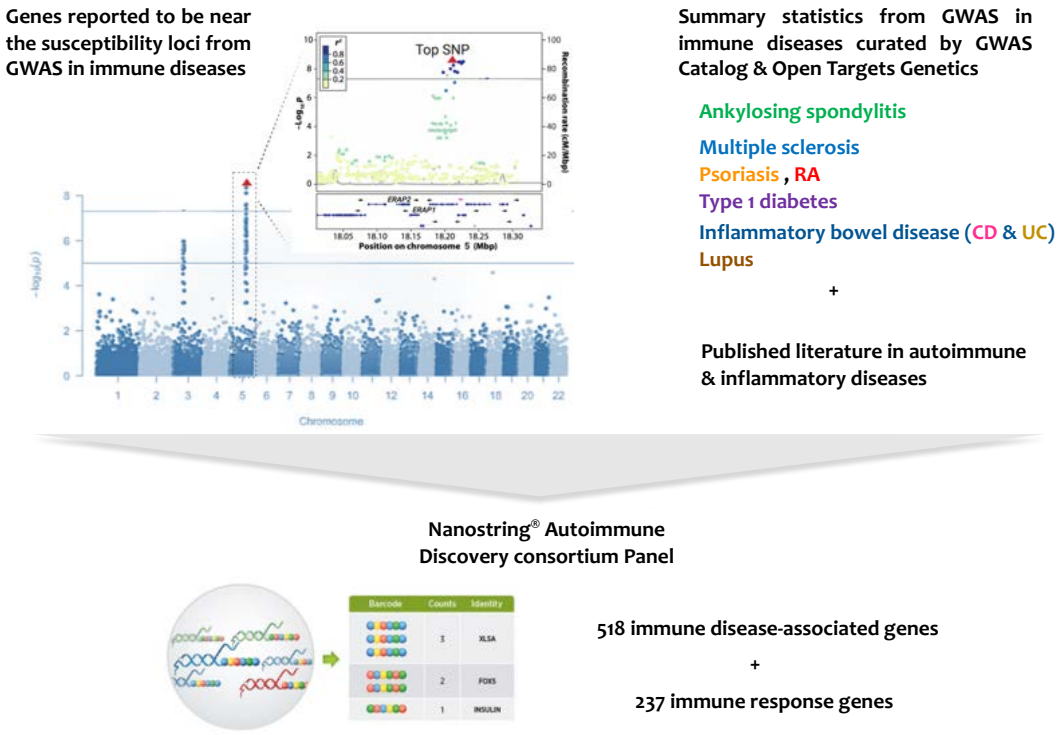


Figure 21: Design of the NanoString Autoimmune Discovery Consortium Panel gene expression assay. Employing the nearest-gene concept, from reported genetic loci associated to 9 different immune-mediated diseases, we designed with NanoString a gene expression assay for 755 genes.

Using this panel, we performed gene expression analysis on whole blood samples and cell populations isolated from patients. The TruCulture assays from 25 patients at baseline were used to explore the expression pattern of genes associated to diseases in whole blood under innate and adaptive stimulation conditions as described by Duffy et al. (2014). (**Table 6**). This tested the NanoString gene expression panel we designed to address our question.

Table 6: Innate and Adaptive Immune Stimuli Used in TruCulture Assays

Stimulus	Concentration	Supplier	Sensor or Receptor
Null (∅)	-	NA	-
LPS-EB (hi)	10 ng/mL		TLR4
Enterotoxin SEB	0.4 µg/mL	Bernhard Nocht Institute	TCR

From the overall gene expression data, we observed that 87% of the genes in the panel (n=658) were expressed in whole blood, while the remaining 13% of the genes were below the background threshold. In particular, nearly 80% of the GWAS genes (n=409) were detected in whole blood (**Figure 22**).

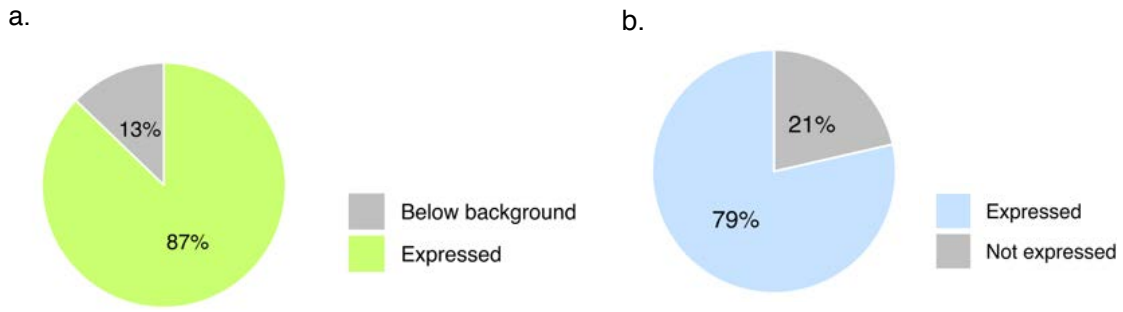


Figure 22: Proportion of genes detected in TruCulture whole blood assays using the new NanoString panel. Percentage of genes detected from the whole panel is depicted in **a**. The percentage of disease-associated genes detected out of the 518 genes included in the panel is showed in **b**.

Since nearly 80% of the genes associated with CIDs can be detected in blood samples, it implies that these genes are important in immune cell function. Therefore, it supports the hypothesis that dysregulation in the immune cell function plays an important role in disease mechanism.

6.2.1 Expression profile of GWAS genes in whole blood

In line with observation in **Figure 22**, we asked whether the detected genes participate in the innate or the adaptive arm of the immune response. To address this question, we compared the expression pattern of GWAS genes in response to innate and adaptive stimuli in whole blood cultures from 25 AxSpA patients. We performed a principal component analysis (PCA) to assess the data structure. PCA is a dimensionality reduction method that simplifies the complexity in a data set while retaining trends and patterns in the data (Lever, Krzywinski, & Altman, 2017). As shown in **Figure 23a**, the main driving factor of the clusters in the PCA is the stimulus condition. The other variables, such as sex (**Figure 23b**) and HLA-B27 status (**Figure 23c**), did not have an observable effect on the clustering of the samples. This analysis also demonstrated that there were no outliers, as the different stimulation groups defined homogeneous clusters. The maximum variance (PC1 = 58%) is attributed to cell activation versus non stimulated conditions (Null) (**Figure 23a**).

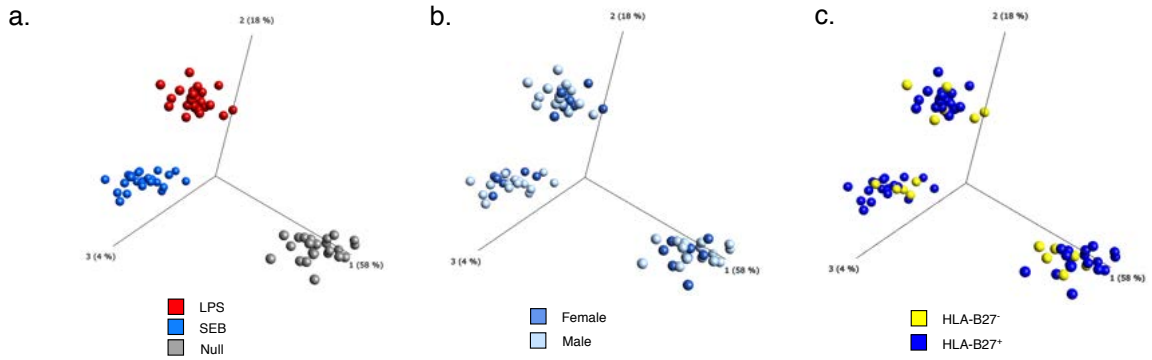
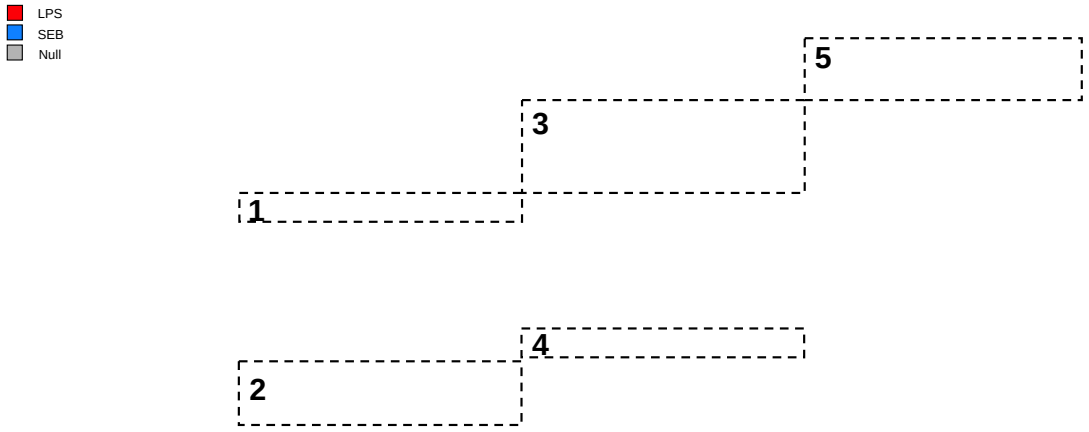


Figure 23: Gene expression analysis of whole blood TruCulture assays. Principal component analysis (PCA) on the 25 SpA patients under (a.) LPS (red), SEB (blue) and Null (grey) stimulation conditions. In (b.) samples were coloured by sex and in (c.) by HLA-B27 status. Gene expression values are log2 transformed, centred to a mean value of zero and scaled to unit variance.

By performing an unsupervised hierarchical clustering analysis, we observed that several disease-associated genes were modulated selectively by the innate stimulus (156 genes) or the adaptive stimulus (133 genes) (**Figure 24**). We also observed a cluster of 70 genes relatively upregulated in the unstimulated (Null) sample. This context-dependent induction following treatment with an innate stimulus (LPS) and adaptive stimulus (SEB) supports the involvement of the disease-associated genes in both innate and adaptive immune functions. Therefore, we wanted to investigate the potential role of disease-associated genes in innate and adaptive immune cell populations- such as monocytes and T cells.



Defined clusters:

1 & 2: Genes upregulated in the SEB stimulated group

3 & 4: Genes upregulated in the LPS stimulated group

5: Genes upregulated in the NULL

Figure 24: Gene expression profile in whole blood cultures following innate and adaptive stimulation. Heat map representation generated by agglomerative hierarchical clustering of disease-associated genes above background, after LPS (red) and SEB (blue) stimulation conditions, or left unstimulated (NULL in grey) from 25 SpA patients. In the heatmap, the columns represent patient samples and the rows represent genes. The group of genes upregulated by a stimulus marked by black dotted boxes. Gene expression values are $\log_2$ transformed, centred to a mean value of zero and scaled to unit variance. Bar to the left represents the scale for the level of gene expression- red indicates high levels of expression, light blue indicates low levels of expression and yellow indicates no relative change in the expression value.

6.2.2 Expression pattern of disease-associated genes in innate and adaptive cell populations in patients

To gain insight into the functional role of disease-associated genes in innate and adaptive immune responses, we performed a gene expression analysis in representative cell populations purified from AxSpA patients. We compared the gene expression patterns of the panel genes in $CD4^+$ and $CD8^+$ T cells and monocytes under resting and activated states. The initial data structure was demonstrated by conducting PCA in the three data set to determine the quality of the data for further analysis and to assess the extent to which the gene expression is affected by stimulation.

6.2.2.1 Cell-type-specific gene expression analysis in T cells

PCA performed on the $CD4^+$ T cell data indicated that the main factor driving gene expression variance was the stimulation condition, which contributed 71% of the variance (**Figure 25a**). We did not observe any significant effects from sex (**Figure 25b**) and HLA-B27 status (**Figure 25c**) as presented in PCA plots below.

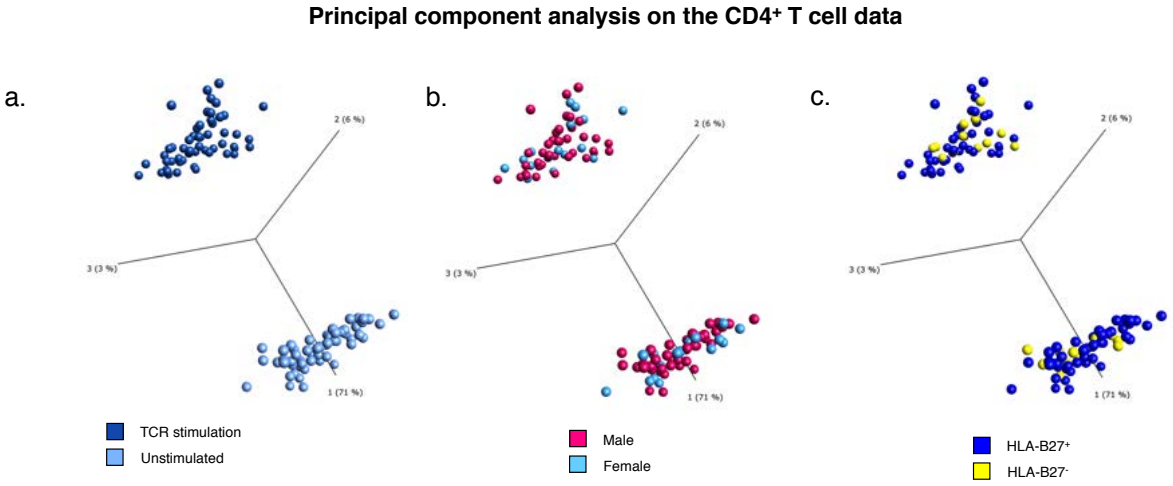


Figure 25: Data structure of the gene expression data from $CD4^+$ T cells. Principal component analysis (PCA) conducted in $CD4^+$ T cell data set from 55 SpA patients. (a.) Data points coloured by stimulation conditions- TCR stimulation (16h) in dark blue and unstimulated in light blue. (b.) Data points coloured by sex of the patients- males in pink and females in sky blue (c.) Data points coloured by HLA-B27 status of the patients- HLA-$B27^+$ patients in blue and HLA-$B27^-$ patients in yellow. Gene expression data are $\log_2$ transformed, centred to a mean value of zero and scaled to unit variance.

We detected 618 genes (~82%) from the total 755 genes in the panel that were above the background threshold (**Figure 26a**). Nearly 70% of the genes that were above the background were disease-associated GWAS genes (**Figure 26b**). Hierarchical clustering analysis performed in the CD4$^+$ T cell data grouped the data by stimulation (**Figure 27**). 57% of the genes we detected were upregulated after TCR activation in CD4$^+$ T cells (**Figure 27**).

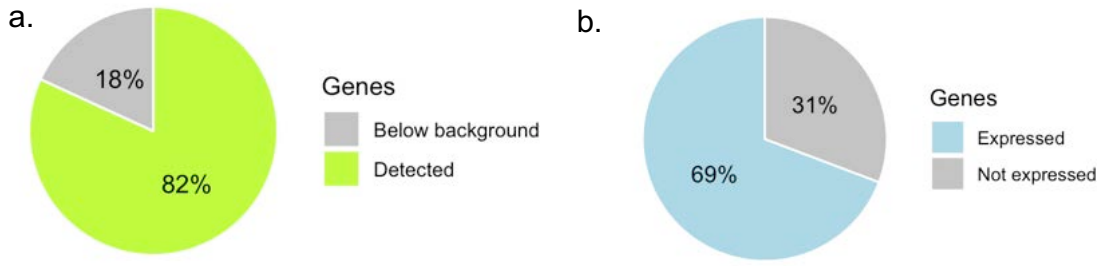


Figure 26: Proportion of genes detected from the NanoString panel. Percentage of genes detected in the CD4+ T cell data is coloured in lime green. Genes that fall fellow background are in grey (a.). The percentage of disease-associated genes detected out of 618 genes is shown in (b.).

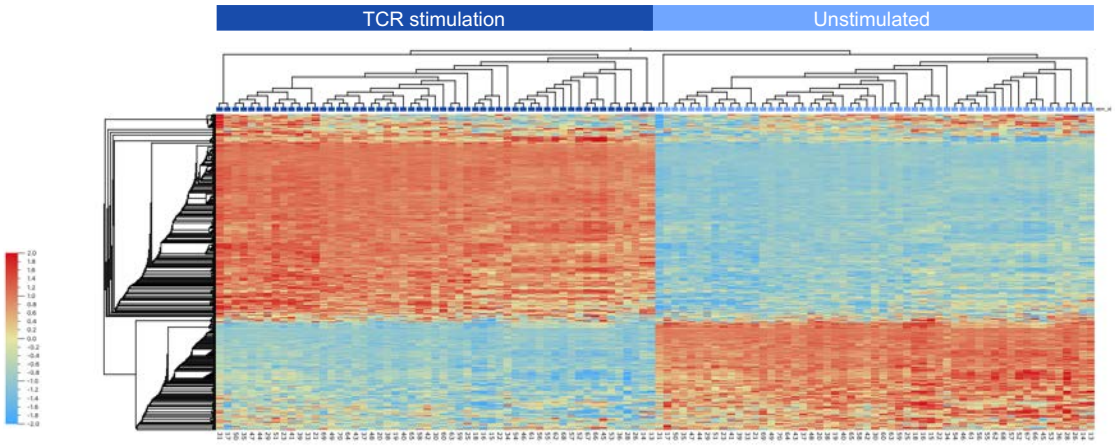


Figure 27: Gene expression pattern in CD4$^+$ T cells upon TCR activation. Heatmap shows the unsupervised clustering of genes (in rows) and samples (in columns). The samples are grouped by the stimulation condition: TCR stimulation (16h) in dark blue and the unstimulated group in light blue. Genes with expression values below the background threshold were removed. Gene expression data are $\log_2$ transformed, centred to a mean value of zero and scaled to unit variance. Bar to the left represents the scale for the level of gene expression- red indicates higher levels of expression, light blue indicates lower levels of expression and yellow indicates no relative change in the expression value

Similarly, in the CD8$^+$ T cell data, PCA shows that the stimulation condition is the main factor driving gene expression variance, with 65% of the variance captured by PC1 (**Figure 28a**). We did not observe any significant effects from sex (**Figure 28b**) and HLA-B27 status (**Figure 28c**) as shown in the PCA plots below.

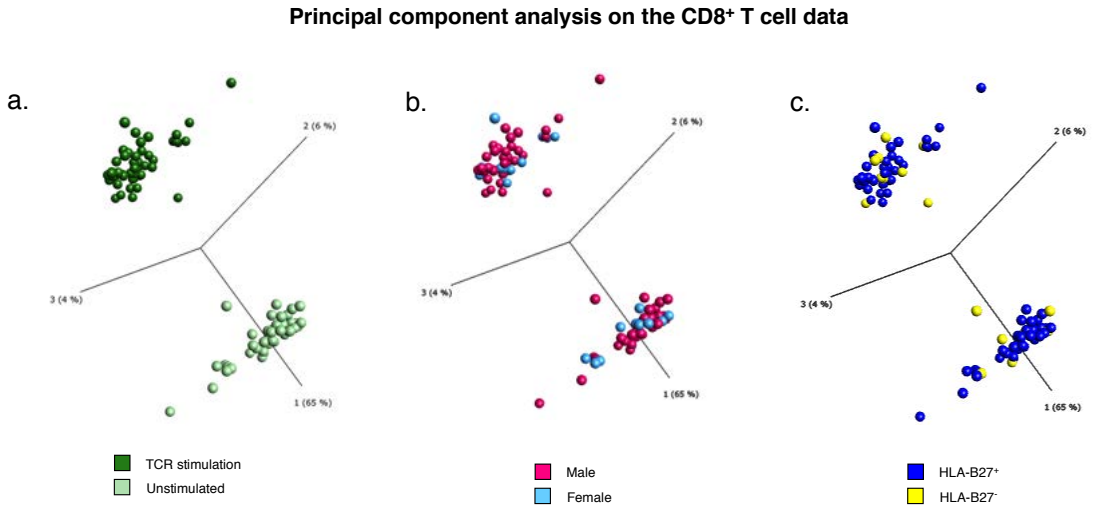


Figure 28: Data structure of the gene expression data from CD8⁺ T cells. Principal component analysis (PCA) conducted in CD8⁺ T cell data set from 49 SpA patients. Data points coloured in (a.) by stimulation conditions- TCR stimulation (16h) in dark green and unstimulated in light green, (b.) by sex of the patients- males in pink and females in sky blue and (c.) by HLA-B27 status of the patients- HLA-B27 positive patients in blue and HLA-B27 negative patients in yellow. Gene expression data are log_2 transformed, centred to a mean value of zero and scaled to unit variance

624 genes (~83%) from the total 755 genes in the panel were above the background threshold (**Figure 29a**), and out of the 624 genes, 69% were specifically disease-associated genes (**Figure 29b**).

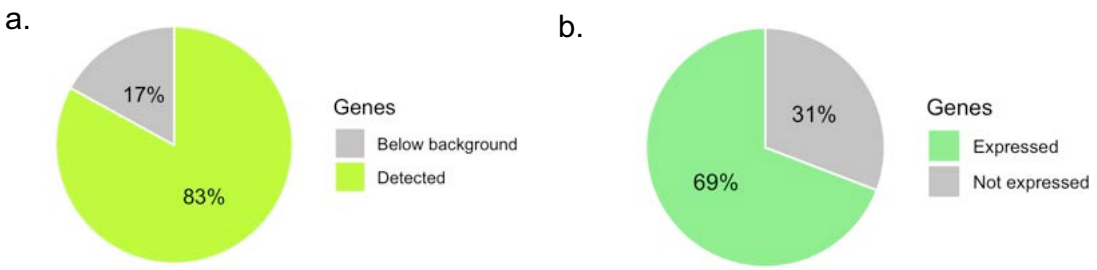


Figure 29: Proportion of genes detected from the NanoString panel. Percentage of genes detected in the CD8⁺ T cell data is coloured in lime green. Genes that fall fellow background is in grey is shown in (a). The percentage of disease associated genes detected out of 624 genes is shown in (b.).

We performed a hierarchical clustering analysis in the CD8⁺ T cell data, which grouped the samples by stimulation condition. 53% of the genes we detected in CD8⁺ T cells were induced after TCR activation (**Figure 30**).

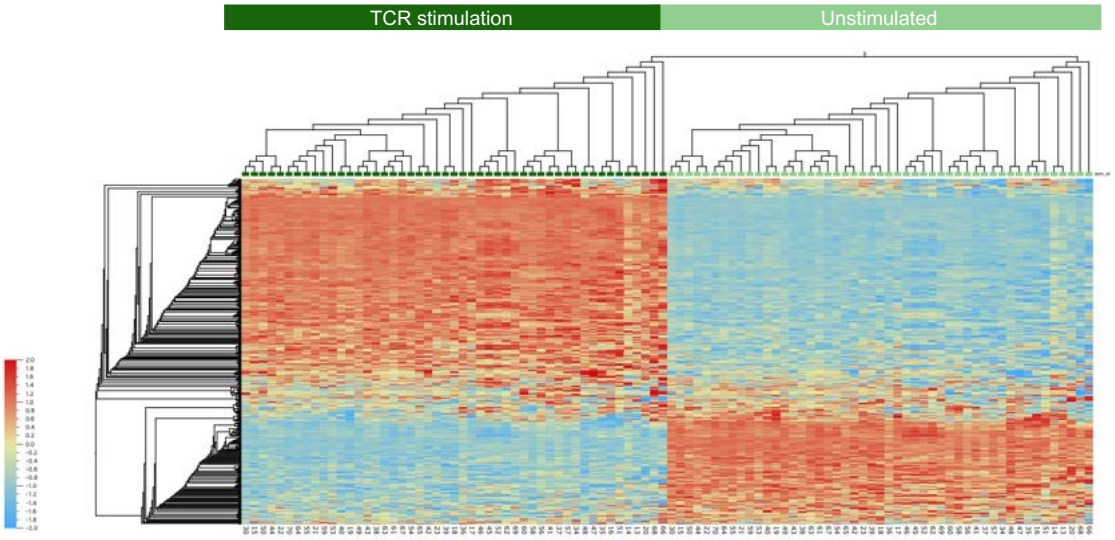


Figure 30: Gene expression pattern in CD8⁺ T cells upon TCR activation. Heatmap shows the unsupervised clustering in of genes (in rows) and samples (in columns). The samples clustered by the stimulation condition- TCR stimulation (16h) in dark green and the unstimulated group in light green. Genes with expression values below the background threshold were removed. Gene expression data are log_2 transformed, centred to a mean value of zero and scaled to unit variance. Bar to the left represents the scale for the level of gene expression- red indicates higher levels of expression, light blue indicates lower levels of expression and yellow indicates no relative change in the expression value

As a quality control for cell stimulation, we also checked the expression of cytokine genes in unstimulated and activated T cells (**Figure 31**). Since most of the genes associated with disease were induced upon TCR activation (16h), they may have a potential role in regulating T cells function.

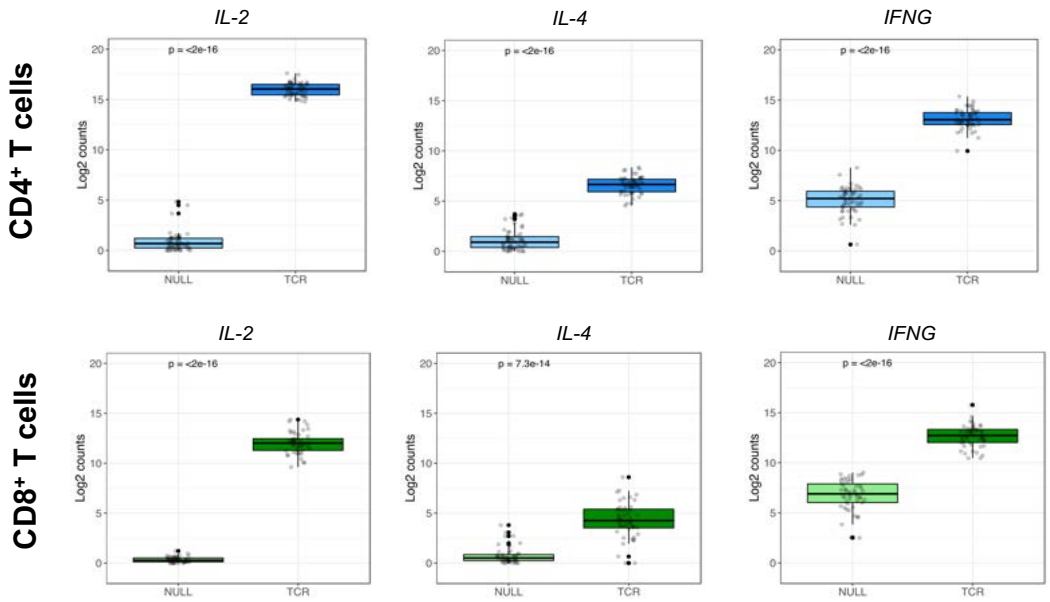


Figure 31: Expression of TCR stimulation induced genes in T cells. Box plots show the expression levels of genes in the unstimulated (NULL) and stimulated for 16h (TCR) CD4⁺ (in blue) and CD8⁺ T-cells (plots in green). Gene expression data are log_2 transformed. A paired T-test was performed between the two groups and the corresponding p-value is indicated within the graph.

6.2.2.2 Cell-type specific gene expression analysis in monocytes

We performed a PCA in the monocyte dataset and observed that cell stimulation with LPS contributed the largest part of the variance (PC1 = 38%) (data not shown). However, for some donors (n=15), the separation between unstimulated and stimulated samples was not clear, possibly due to cell activation of the unstimulated samples during the preparation procedure. We explored this variability further in the following sections. We did not observe any significant effects from sex and HLA-B27 status (data not shown). After removing technical outliers from the data, we further analysed 30 samples. We observed that nearly 87% of the genes from the panel were detected in this dataset as illustrated in the pie chart below (**Figure 32a**). The rest of the genes were below the background threshold. 67% of the genes that were above background were disease-associated genes (**Figure 32b**).

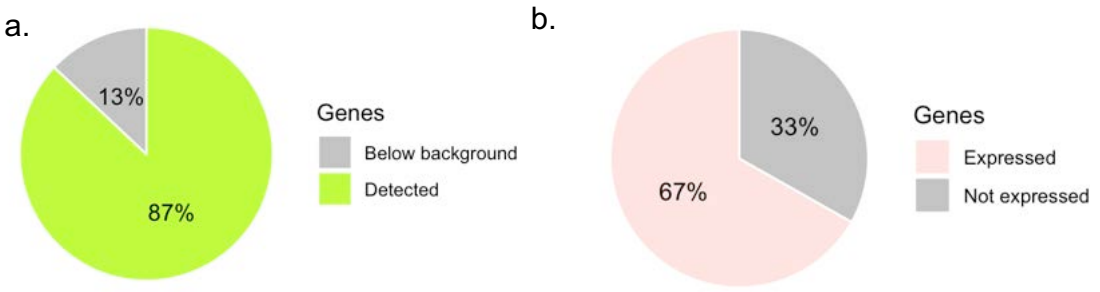


Figure 32: Proportion of genes detected from the NanoString panel. Percentage of genes detected in the $CD8^+$ T cell data is coloured in lime green. Genes that fall fellow background is in grey is shown in (a). The percentage of disease associated genes detected out of 657 genes is shown in (b).

The hierarchical clustering performed on the rest of the 30 patients showed that the clustering is driven by the stimulation condition (**Figure 33**). To have an overview of the effect of the LPS stimulation, we analysed the levels of cytokine genes- *IL1A*, *IL1B* and *IL6R*, that are typically induced by LPS in monocytes as shown in boxplots in **Figure 34**. The difference in gene expression between stimulated and unstimulated group maintained the direction of the induction pattern with statistical significance ($p < 0.05$) (**Figure 34**).

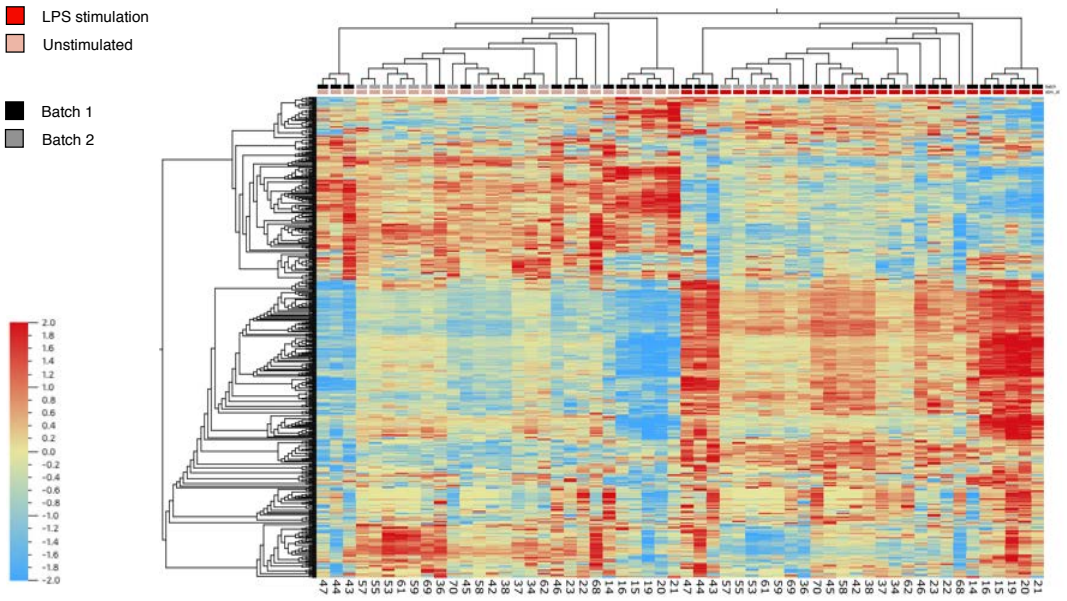


Figure 33: Gene expression pattern in revised monocyte data. Heatmap shows the unsupervised clustering in of genes (in rows) and samples (in columns). The samples coloured by the stimulation condition- LPS stimulation in red and the unstimulated group in light pink. The batch of the CD14 magnetic beads are indicated in black and grey. Genes with expression values below the background threshold were removed. Gene expression data are $\log_2$ transformed, centred to a mean value of zero and scaled to unit variance. Bar to the left represents the scale for the level of gene expression- red indicates higher levels of expression, light blue indicates lower levels of expression and yellow indicates no relative change in the expression value

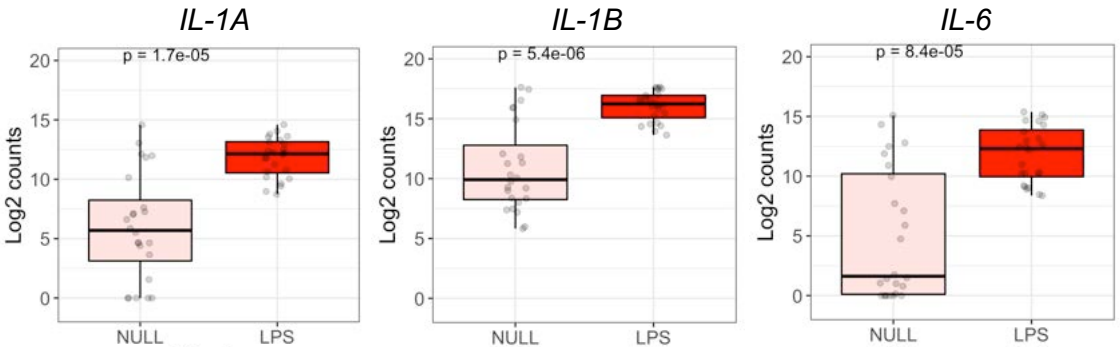


Figure 34: Expression of LPS-induced cytokine genes in monocytes. Box plots show the expression levels of genes in the unstimulated (NULL) and stimulated (LPS) monocytes. Gene expression data are $\log_2$ transformed. A paired T-test was performed between the two groups and the corresponding p-value is indicated within the graph.

6.2.3 Analysis of disease-associated gene modules in isolated immune cell types from AS patients

We characterised the expression pattern of disease-associated genes in $CD4^+$, $CD8^+$ T cells and monocytes isolated from patients with AS. Firstly, we identified which of the genes associated with ankylosing spondylitis and its extra-articular conditions like Pso and IBD are induced upon stimulation in monocytes and T cells. Then we conducted the same analysis for genes associated with other chronic inflammatory disorders, like T1D and MS, which are not associated with SpA. The number of genes assigned to each disease module is given in the table below (**Table 7**) and the number of overlapping genes between AS and other diseases are given in the Venn diagrams in **Figure 35a & b**.

Table 7: Number of genes assigned per each disease module

Chronic inflammatory disease	Number of genes in the panel
Ankylosing spondylitis	43
Psoriasis	48
Inflammatory bowel disease	253
Type 1 Diabetes	36
Multiple sclerosis	104

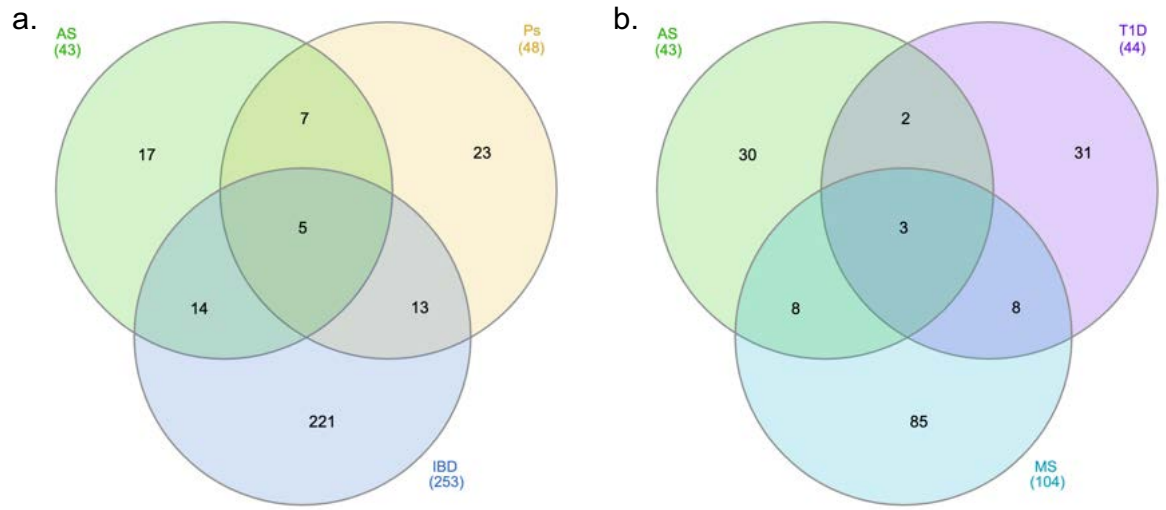
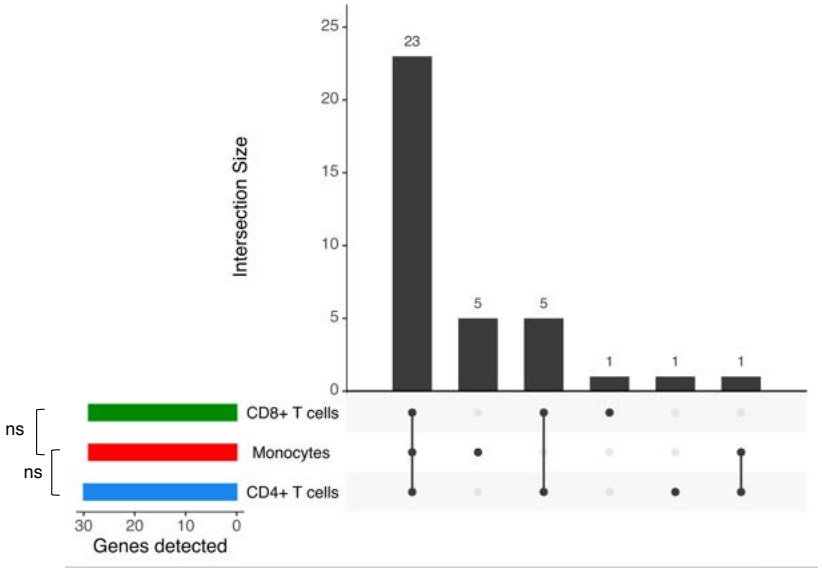


Figure 35: Venn diagram showing the overlap in the number of genes in disease modules. The number of shared genes among AS and its extra-articular manifestations are shown in (a). The overlap of genes between AS and unrelated CIDs (T1D and MS) are shown in (b).

6.2.3.1 Gene expression analysis of the ankylosing spondylitis associated gene module

We detected expression of 36 genes out of the 43 genes present in the ankylosing spondylitis module in the stimulated purified cell populations. We did not observe a significant difference between the number of genes expressed between the different cell populations (**Figure 36**). Twenty-three of the genes were shared among T cells and monocytes. Only a few genes from the detected list were distinctive to a cell type, i.e., five genes were uniquely induced in monocytes, and another five were specific to T cells. Only one gene was uniquely induced in each $CD4^+$ and $CD8^+$ T cells.



Test statistic between number of genes expressed between monocytes and CD4 T cells

```
2-sample test for equality of
proportions with continuity correction

data:  c(29, 30) out of c(43, 43)
X-squared = 0, df = 1, p-value = 1
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.2426226  0.1961110
sample estimates:
   prop 1    prop 2
0.6744186 0.6976744
```

Test statistic between number of genes expressed between monocytes and CD8 T cells

```
2-sample test for equality of proportions
without continuity correction

data:  c(29, 29) out of c(43, 43)
X-squared = 4.5078e-31, df = 1, p-value = 1
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.1980722  0.1980722
sample estimates:
   prop 1    prop 2
0.6744186 0.6744186
```

Figure 36: Profile of ankylosing spondylitis gene module. Upset plots indicate the number of genes detected in the three stimulated cell types. The bottom panel of the upset plots shows the number of genes detected in each cell type. The dots represent the cell types where the genes are detected. Line connecting the dots shows the number of genes shared between cell types. The bar plot on the top panel shows the number of genes shared among the cell types. The statistics from two-proportions test between the number of genes detected in monocytes and in T cells are shown below the figure.

By performing unsupervised hierarchical clustering, we analysed the levels of the gene expression as shown in **Figure 37**, and we observed that the cell type primarily drives the clustering of the samples. The majority of the genes in the AS module had relatively higher expression in LPS-stimulated monocytes compared to the activated T cells ($CD8^+$ T cells shown in the **Figure 37**). We observed a cluster of genes with higher expression in T cells consisting of transcription factor genes that are important for regulating T cell functions, such as, T-bet (*TBX21*) and *RUNX3* (**Figure 38**). A study published in 2017 demonstrated that AS patients showed higher T-bet expression in NK and $CD8^+$ T cells compared to healthy individuals (Lau et al., 2017). Another group evaluated the functional mechanism of *RUNX3* association with AS susceptibility (Evans et al., 2011), by exploring the allele-specific effects on transcription factor IRF4 recruitment and *RUNX3* expression in primary $CD8^+$ T cells from AS patients (Vecellio et al., 2016). We noted higher levels in $CD8^+$ T cells of *EOMES*, which encodes an essential protein for the differentiation of effector $CD8^+$ T cells as previously described (Pearce et al., 2003) (**Figure 38**).

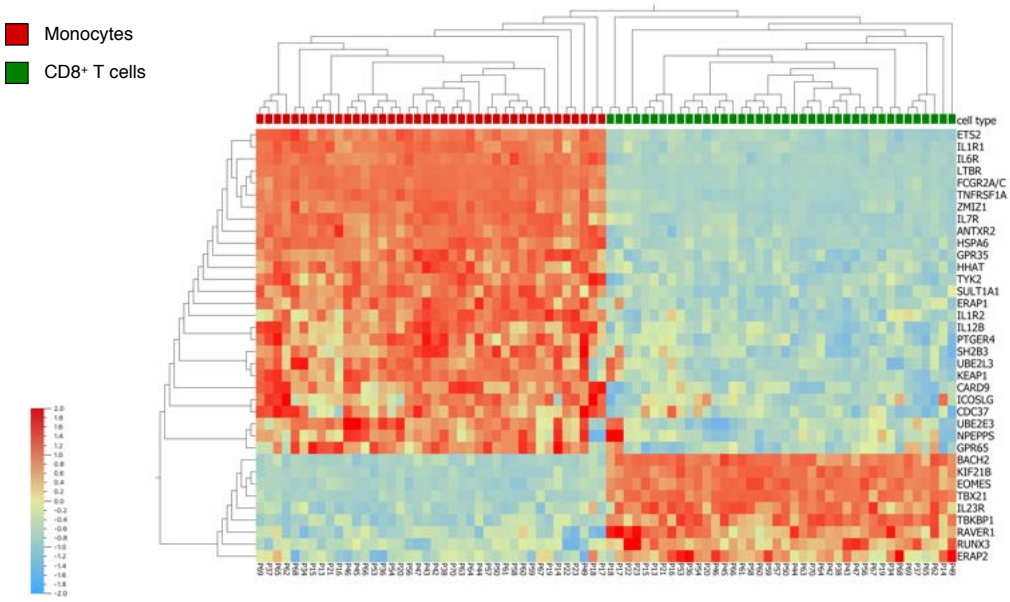


Figure 37: Expression of AS-linked gene module. Gene expression pattern of the disease module for AS in activated monocytes and CD8+ T cells is shown in the heatmap ordered by hierarchical clustering. Gene expression data are log_2 transformed, centred to a mean value of zero and scaled to unit variance. Bar to the left represents the scale for the level of gene expression- red indicates higher levels of expression, light blue indicates lower levels of expression and yellow indicates no relative change in the expression value.

Other genes with preferential expression in T cells included *KIF21B,* which encodes a kinesin-like protein involved in Golgi-to-ER retrograde transport, *RAVER1* encoding a ribonucleoprotein, *CDKAL1,* a methylthiotransferase that controls the quality of protein translation, and *TBKBP1,* which encode an adaptor protein that binds to the NF-κB factors, as

seen in the heatmap in **Figure 37**. We also observed that aminopeptidases essential for peptide trimming during the generation of HLA class I-binding peptides were expressed in both monocytes and T cells, with higher expression of *ERAP1* in monocytes (**Figure 38**) and of *ERAP2* in T cells, respectively.

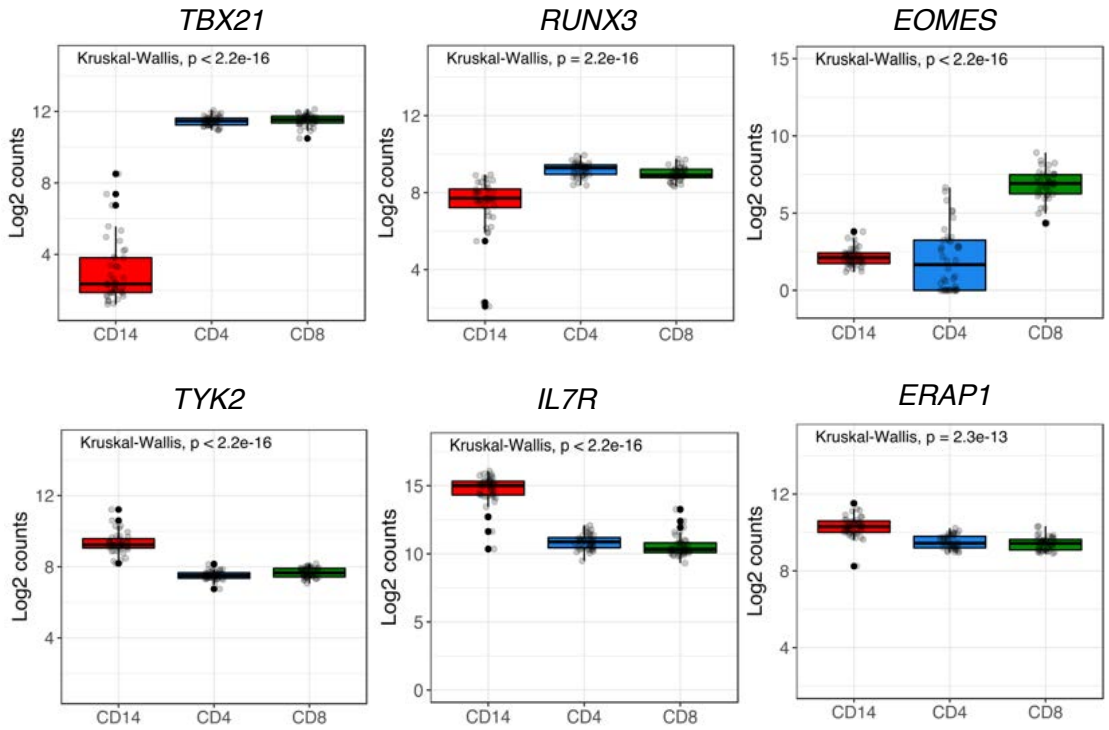


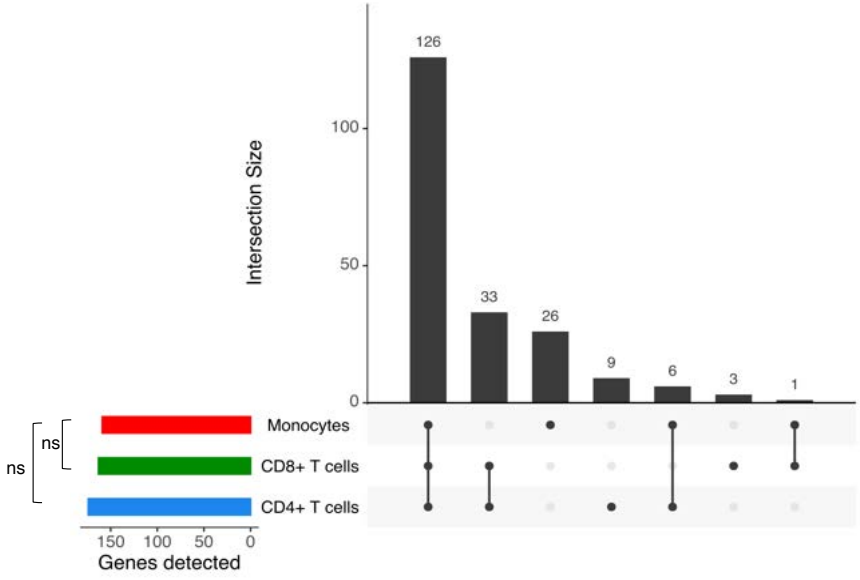
Figure 38: Expression of AS module genes in activated monocytes and T cells. Box plots show the expression levels of genes in the LPS activated $CD14^+$ monocytes (in red) and TCR stimulated $CD4^+$ (T cells (dark blue) and $CD8^+$ T-cells dark green). Gene expression data are log_2 transformed. A non-parametric multiple group comparison (Kruskal-Wallis) was performed and the corresponding p-value is indicated within the graph.

TYK2 (tyrosine kinase 2), a gene associated with several CIDs is detected at higher levels in LPS activated monocytes than in activated T cells (**Figure 38**). Similarly, we saw that the interleukin-7 receptor (IL7R) was more expressed in monocytes following LPS stimulation, compared to stimulated T cells (**Figure 38**), which concurs with the findings from Al-Mossawi et al. (2019). Interleukin-7 (IL-7) is a crucial regulator of T-cell survival and proliferation, whose functional effects are controlled by the soluble form of the IL-7R which is induced in myeloid cells. Therefore, polymorphisms in the IL-7R could impact the interaction between IL-7 and IL-7R, and affect the downstream inflammatory response.

6.2.3.2 Gene expression analysis of genes associated with AS-related diseases

Inflammatory bowel disease

Inflammatory Bowel Disease (IBD) is another extra-articular manifestation of ankylosing spondylitis (AS). 5 - 10% of AS cases are related to IBD; and there is an association between gut inflammation in AS and IBD (Rudwaleit & Baeten, 2006). Using the NanoString panel, we assessed the expression pattern of genes associated with IBD (Crohn's disease and ulcerative colitis) in monocytes and T cells. We detected expression in immune cells of 204 genes out of the 253 genes in the panel associated with IBD.



Test statistic between number of genes expressed between monocytes and CD4 T cells

```
2-sample test for equality of proportions
with continuity correction

data:  c(159, 174) out of c(253, 253)
X-squared = 1.7215, df = 1, p-value = 0.1895
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.14573992  0.02716285
sample estimates:
   prop 1    prop 2
0.6284585 0.6877470
```

Test statistic between number of genes expressed between monocytes and CD8 T cells

```
2-sample test for equality of proportions with
continuity correction

data:  c(159, 163) out of c(253, 253)
X-squared = 0.076863, df = 1, p-value = 0.7816
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.10357954  0.07195898
sample estimates:
   prop 1    prop 2
0.6284585 0.6442688
```

Figure 39: Profile of IBD gene module. Upset plots indicate the number of genes induced among the three stimulated cell types. The bottom panel of the upset plots shows the number of genes detected in each cell type. The dots represent the cell types where the genes are detected. Line connecting the dots shows the number of genes shared between cell types. The bar plot on the top panel shows the number of genes shared among the cell types and the calculations are given below the figure. The statistics from two-proportions test between the number of genes detected in monocytes and in T cells are shown below the figure.

Expression of 174 genes was identified in $CD4^+$ T cells, 163 genes in the $CD8^+$ T cells and 159 genes in monocytes. Similar to AS and Pso, most genes (126) were shared among the three activated immune cell subsets (**Figure 39**). A set of 26 genes were uniquely expressed in monocytes. To visualize the levels of gene expression, we did a hierarchical clustering (**Figure 40**). We noted a stronger expression of IBD-associated genes in monocytes, which could be due to the fact that LPS is a stronger stimulation compared to the TCR activation.

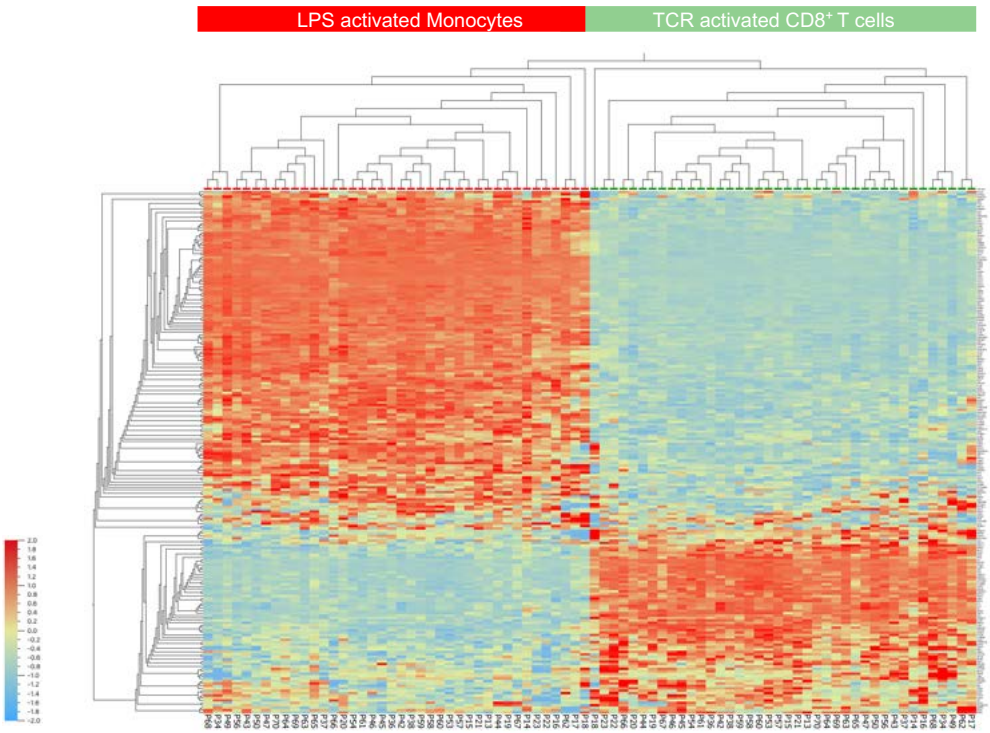


Figure 40: Expression of IBD gene module. Gene expression pattern of the disease module for IBD T1D in activated monocytes and CD8+ T cells is shown in the heatmap ordered by hierarchical clustering. Gene expression data are $\log_2$ transformed, centred to a mean value of zero and scaled to unit variance. Bar to the left represents the scale for the level of gene expression- red indicates higher levels of expression, light blue indicates lower levels of expression and yellow indicates no relative change in the expression value.

We observed several upregulated IBD susceptibility genes in monocytes such as: pattern recognition receptor *NOD2*, pro-inflammatory cytokine genes like tumor necrosis factor superfamily 15 (*TNFSF15*) and genes of cytokine members of IL-6 family, such as Oncostatin M (*OSM*, **Figure 41**). *NOD2* expression is restricted to monocytes, where it can drive the innate inflammatory responses by activating NF-κB and MAPK (mitogen-activated protein kinases) and caspase-1 pathways (Ogura et al., 2001).

TNFSF15 encodes the TNF superfamily cytokine, *TL1A,* which has been associated to IBD susceptibility by signalling through the death receptor 3 (DR3) and activating NF-κB to co-stimulate IFN-γ production in T cells (Prehn et al., 2007). Increased expression of *TL1A* was observed in monocytes and monocyte-derived DCs following bacterial stimulation. *TL1A* was shown to have a role in the inflammation of the gut mucosa and ileum, which is a prominent feature of active IBD (Bamias et al., 2003). However, an eQTL study in 2018 demonstrated that reduced expression of *TNFSF15* from peripheral blood monocytes is associated with IBD risk (Richard et al., 2018). The study showed that the protective allele for IBD at *TNFSF15* was strongly associated with increased mRNA levels of *TNFSF15* in monocytes in both IBD patients and healthy donors. High levels of oncostatin-M (*OSM*) were previously described in the inflamed intestinal mucosa of IBD patients. *OSM* is also a predictive factor for response to anti-TNF therapy in IBD (West et al., 2017) (**see Figure 41**).

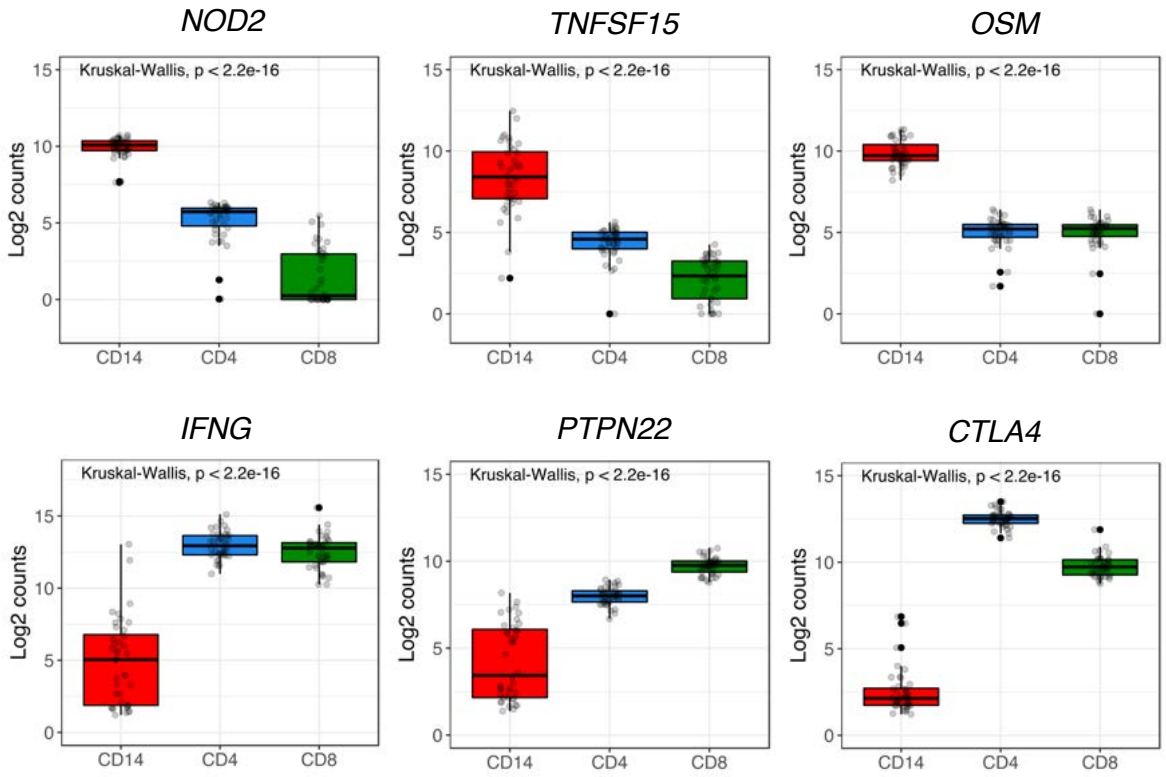


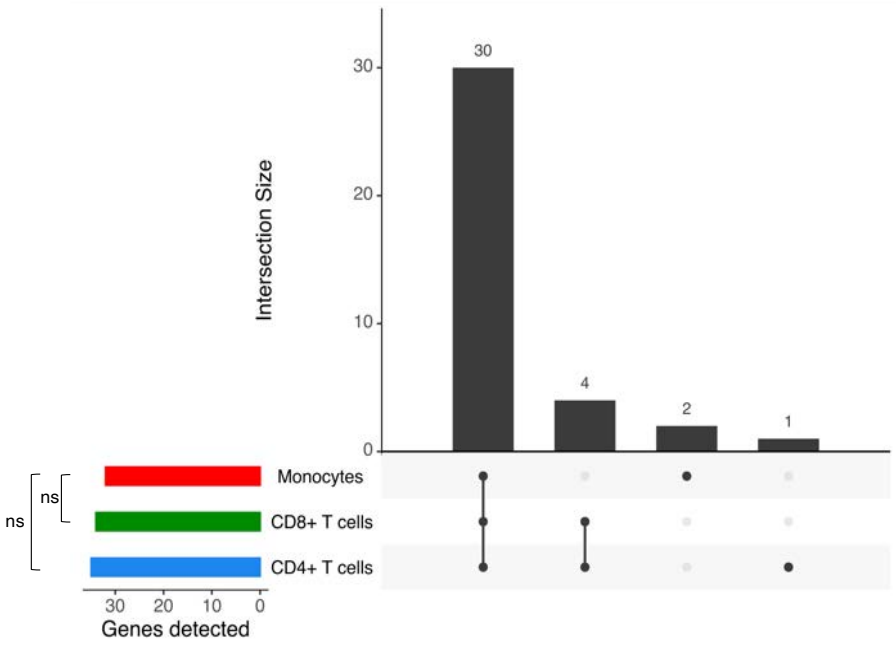
Figure 41: Expression of IBD module genes in activated monocytes and T cells. Box plots show the expression levels of genes in the LPS activated $CD14^+$ monocytes (in red) and TCR stimulated $CD4^+$ T cells (dark blue) and $CD8^+$ T cells (dark green). Gene expression data are log_2 transformed. A non-parametric multiple group comparison (Kruskal-Wallis) was performed and the corresponding p-value is indicated within the graph.

Among the genes that were relatively increased in T cells were the proinflammatory cytokine gene *IFNG*, protein tyrosine phosphatase non-receptor type 22 (*PTPN22*) and cytotoxic T-lymphocyte-associated protein 4 (*CTLA4*) (**Figure 41**). IFNG was identified as a major cytokine in IBD pathogenesis (Langer et al., 2019). Prehn et al. previously reported that production of *TNFSF15* from monocytes stimulated through FcγR induces *IFNG* in $CD4^+$ T-cells (Prehn et al., 2007). We noticed an increase in *PTPN22* expression in activated T cells (**Figure 41**), which is a protein tyrosine phospahatase (PTP) that regulates T-cell activation and effector function (Brownlie, Zamoyska, & Salmond, 2018). *PTPN22* also influences inflammatory signalling in lymphocytes by affecting cytokine secretion and autophagy, which is an essential pathway for clearing out the invading pathogens. In IBD, protein tyrosine phosphatases (PTPs) also have a role in controlling the intestinal epithelial barrier function. Therefore, dysfunction of PTPs can lead to exaggerated immune responses and consequent chronic intestinal inflammation featured in IBD (Spalinger et al., 2015).

We observed high expression levels of *CTLA4* in $CD4^+$ and $CD8^+$ T cells (**Figure 41**). In IBD patients, *CTLA4* was shown to have an essential role in T_{reg} suppressive function during intestinal inflammation (Read et al., 2006). It is possible that SNPs in *CTLA4* may disrupt the regulatory T cell function and exacerbate IBD-associated colitis.

Psoriasis

Psoriasis (Pso) is one of the extra-articular manifestations of ankylosing spondylitis (AS). 5-25% psoriasis cases go on to develop AS and SpA-associated inflammatory spinal disease (Machado et al., 2013). In our expression data, we detected 37 genes from the 48 genes assigned to Pso from the panel. We observed that similar to AS, the Pso genes (30) were shared genes among the three cell types (**Figure 42**). The hierarchical clustering revealed genes affected by both LPS stimulation in monocytes and TCR-activation in T cells (**Figure 43**).



Test statistic between number of genes expressed between monocytes and CD4 T cells

```
2-sample test for equality of proportions
with continuity correction

data:  c(32, 35) out of c(48, 48)
X-squared = 0.19763, df = 1, p-value = 0.6566
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.2666067  0.1416067
sample estimates:
   prop 1    prop 2
0.6666667 0.7291667
```

Test statistic between number of genes expressed between monocytes and CD8 T cells

```
2-sample test for equality of proportions with
continuity correction

data:  c(32, 34) out of c(48, 48)
X-squared = 0.048485, df = 1, p-value = 0.8257
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.2477528  0.1644194
sample estimates:
   prop 1    prop 2
0.6666667 0.7083333
```

Figure 42: Profile of psoriasis gene module. Upset plots indicate the number of genes induced in the three stimulated cell types. The bottom panel of the upset plots shows the number of genes detected in each cell type. The dots represent the cell types where the genes are detected. Line connecting the dots shows the number of genes shared between cell types. The bar plot on the top panel shows the number of genes shared among the cell types. The statistics from two-proportions test between the number of genes detected in monocytes and in T cells are shown below the figure.

We observed upregulated levels of transcription factor genes like *IRF4* in T cells (**Figure 44**). We also saw that monocytes had relatively higher expression of *TRAF3IP2*, which encodes for an adaptor protein involved in the induction of IL-17-dependent inflammatory responses (Qian et al., 2007). This observation supports our previous findings from the analysis in AS module that the disease-associated genes suggest a role for the interaction between innate and adaptive cells in disease pathogenesis.

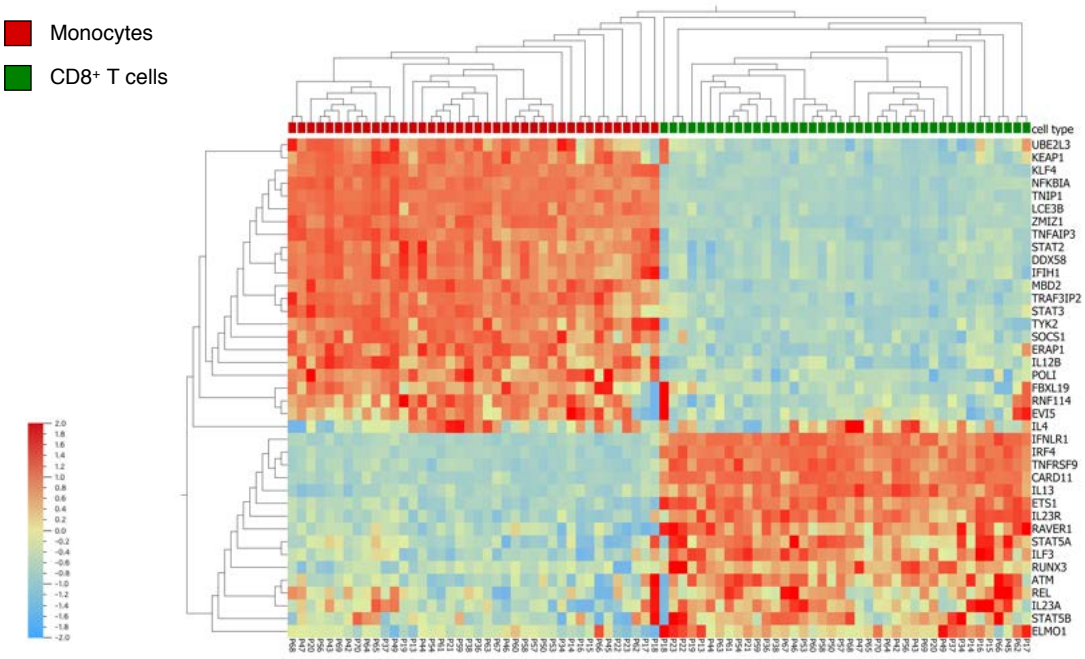


Figure 43: Expression of psoriasis gene module. Gene expression pattern of the disease module for Pso in activated monocytes and CD8+ T cells is shown in the heatmap ordered by hierarchical clustering. Gene expression data are $\log_2$ transformed, centred to a mean value of zero and scaled to unit variance. Bar to the left represents the scale for the level of gene expression- red indicates higher levels of expression, light blue indicates lower levels of expression and yellow indicates no relative change in the expression value.

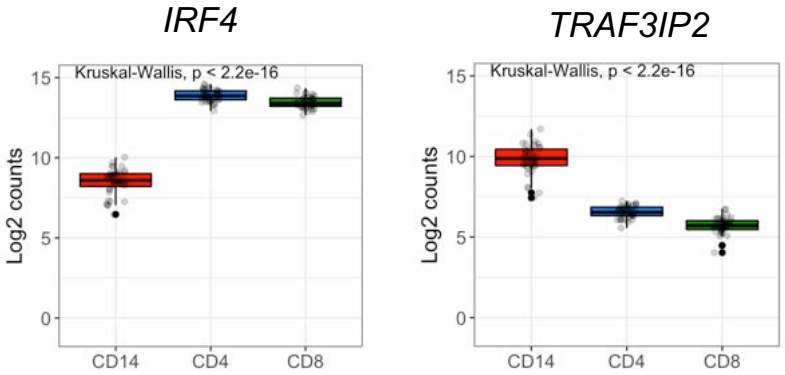
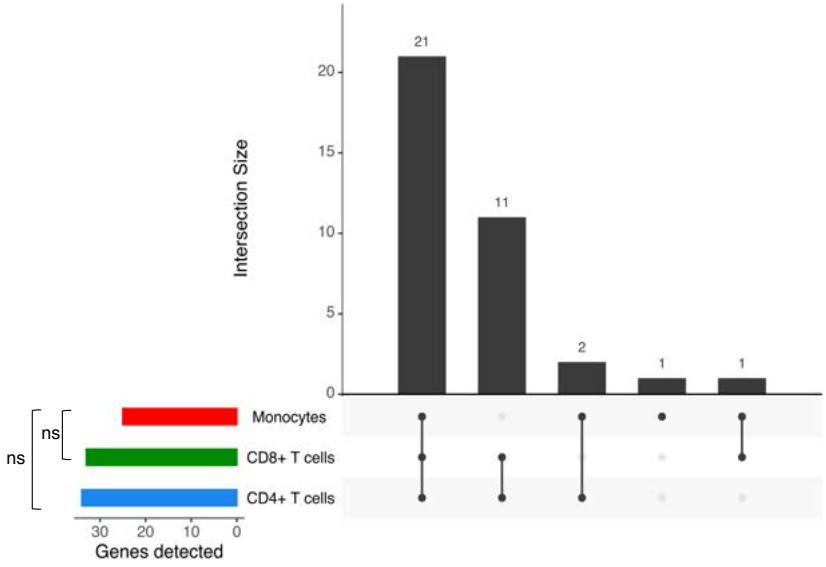


Figure 44: Expression of Pso module genes in activated monocytes and T cells. Box plots show the expression levels of genes in the LPS activated $CD14^+$ monocytes (in red) and TCR stimulated $CD4^+$ T cells (dark blue) and $CD8^+$ T cells (dark green). Gene expression data are $\log_2$ transformed. A non-parametric multiple group comparison (Kruskal-Wallis) was performed and the corresponding p-value is indicated within the graph.

6.2.3.3 Gene expression analysis of genes associated with other CIDs

Type 1 Diabetes

Type 1 Diabetes (T1D) is a chronic metabolic disorder with autoimmune destruction of pancreatic islet beta cells that produce insulin. We investigated the disease gene module (36 genes) from T1D to compare the expression pattern of genes from unrelated inflammatory and autoimmune diseases to that of AS-associated genes. The UpSet analysis showed that nearly all of the genes were expressed in $CD4^+$ T cells (34 genes) and $CD8^+$ T cells (33 genes), and we noted 25 genes that were detected in monocytes (**Figure 45 & 46**).



Test statistic between number of genes expressed between monocytes and CD4 T cells

```
2-sample test for equality of proportions
with continuity correction

data:  c(25, 34) out of c(44, 44)
X-squared = 3.2916, df = 1, p-value = 0.06963
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.418984102  0.009893193
sample estimates:
   prop 1    prop 2
0.5681818 0.7727273
```

Test statistic between number of genes expressed between monocytes and CD8 T cells

```
2-sample test for equality of proportions with
continuity correction

data:  c(25, 33) out of c(44, 44)
X-squared = 2.4782, df = 1, p-value = 0.1154
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.39894314  0.03530677
sample estimates:
   prop 1    prop 2
0.5681818 0.7500000
```

Figure 45: Profile of T1D gene module. Upset plots indicate the number of genes induced in three stimulated cell types. The bottom panel of the upset plots shows the number of genes detected in each cell type. The dots represent the cell types where the genes are detected. Line connecting the dots shows the number of genes shared between cell types. The bar plot on the top panel shows the number of genes shared among the cell types and the calculations are given below the figure. The statistics from two-proportions test between the number of genes detected in monocytes and in T cells are shown below the figure.

Among the genes with higher expression in monocytes we observed *IL10*, an adaptor protein that encodes for phosphatase for JAK2 (*SH2B3*) and *NRP1* (Neuropilin-1), which codes for a peptide that is highly linked to diabetic neuropathy (Bondeva & Wolf, 2015) (**Figure**

47). CD14^{+} CD16^{+} monocytes from peripheral blood have been reported to produce cytokine IL-10 upon activation (Skrzeczyńska-Moncznik et al., 2008). However, in the context of T1D, a recent study showed that type I interferons selectively inhibit IL-10 signalling in effector T-cells in T1D (Iglesias et al., 2018).

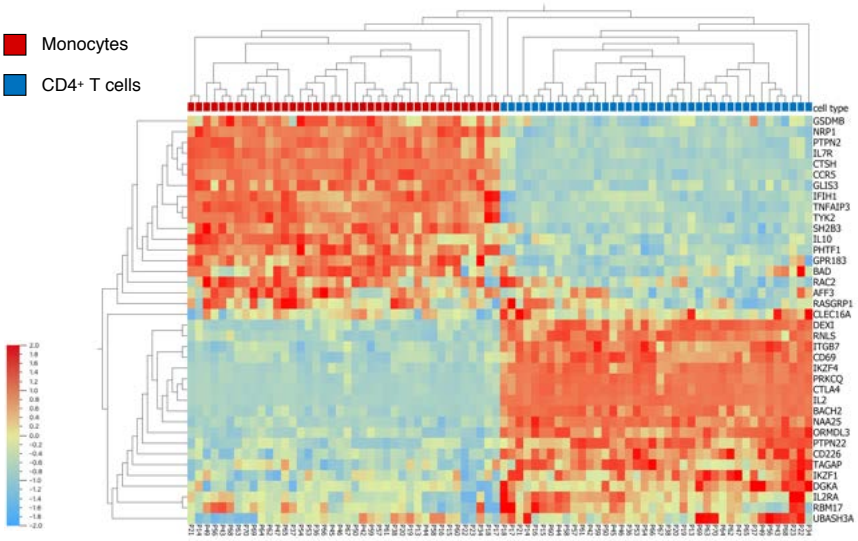


Figure 46: Expression of T1D associated gene module. Gene expression pattern of the disease module for T1D in activated monocytes and CD4+ T cells is shown in the heatmap ordered by hierarchical clustering. Gene expression data are $\log_2$ transformed, centred to a mean value of zero and scaled to unit variance. Bar to the left represents the scale for the level of gene expression- red indicates higher levels of expression, light blue indicates lower levels of expression and yellow indicates no relative change in the expression value.

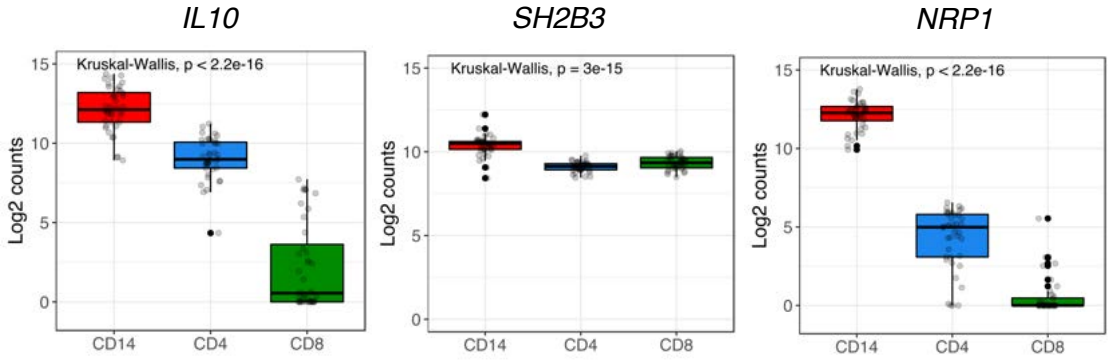


Figure 47: Expression of T1D module genes in activated monocytes. Box plots show the expression levels of genes in the LPS activated CD14^{+} monocytes (in red) and TCR stimulated CD4^{+} T cells (dark blue) and CD8^{+} T-cells (dark green). Gene expression data are $\log_2$ transformed. A non-parametric multiple group comparison (Kruskal-Wallis) was performed and the corresponding p-value is indicated within the graph.

Among the genes with higher expression in T cells was the ORMDL Sphingolipid Biosynthesis Regulator 3 (*ORMDL3*), a crucial regulator of ER-stress and UPR through activation of the transcription factor 6 (*ATF6*) (**Figure 48**). Dysfunction in this gene is shown to affect the function of pancreatic beta-cells and the development of insulin resistance in T1D (Eizirik DL et al., 2008). *TAGAP*, a member of the Rho GTPase-activator superfamily, was reported to be upregulated in activated $CD4^+$ and $CD8^+$ T cells (Schmiedel B et al. 2018). *TAGAP* is also associated with RA susceptibility, however, its biological function and relevance are not understood well in either T1D or RA.

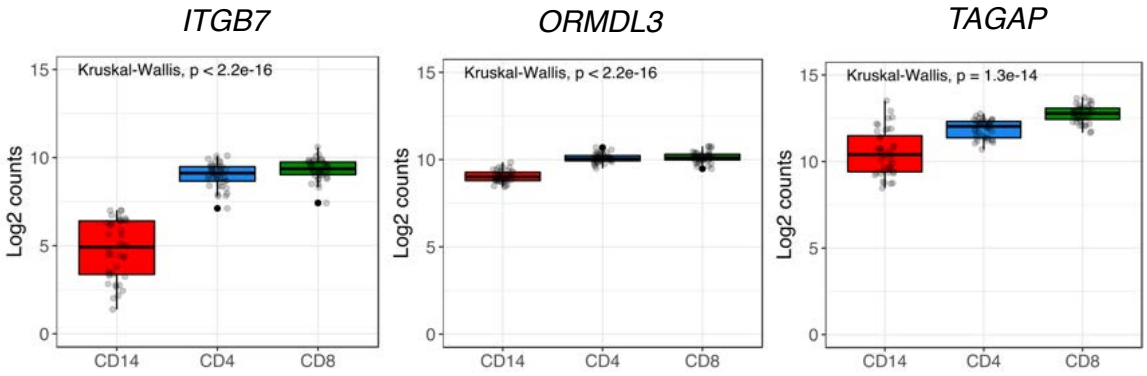
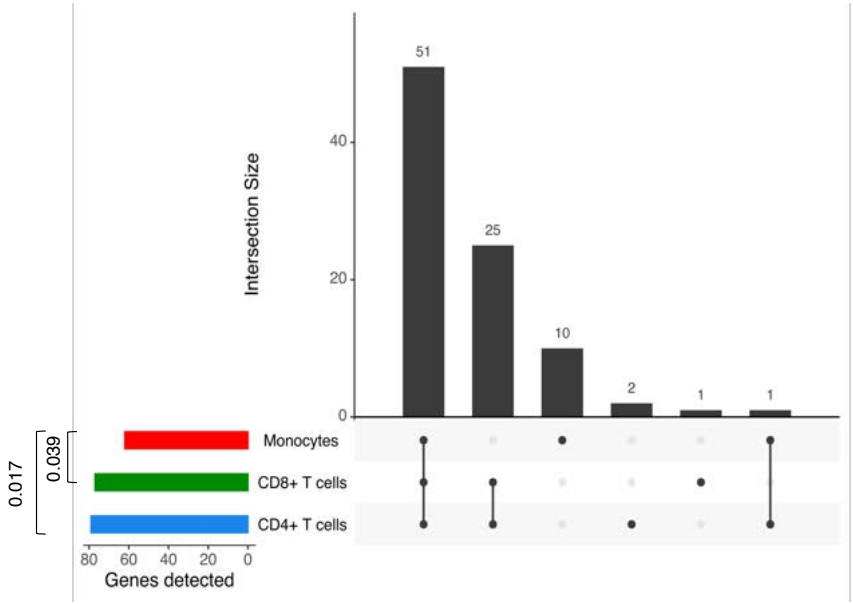


Figure 48: Expression of T1D module genes in activated T cells. Box plots show the expression levels of genes in the LPS activated $CD14^+$ monocytes (in red) and TCR stimulated $CD4^+$ T cells (dark blue) and $CD8^+$ T-cells dark green). Gene expression data are log_2 transformed. A non-parametric multiple group comparison (Kruskal-Wallis) was performed and the corresponding p-value is indicated within the graph.

Multiple Sclerosis

Multiple sclerosis (MS) is a common autoimmune disease that targets myelin in the central nervous system (CNS). The expression of genes in the MS gene module was analysed in activated monocytes and T cells. We detected 90 genes from the 104 genes of the module: 86% of these were expressed in $CD4^+$ (79 genes) and $CD8^+$ T cells (77 genes). We detected 62 genes in the activated monocytes (**Figure 49**). The overall gene signature shows that a large number of MS-associated genes could be of functional relevance in T cells.



Test statistic between number of genes expressed between monocytes and CD4 T cells

```
2-sample test for equality of proportions
with continuity correction

data:  c(62, 79) out of c(104, 104)
X-squared = 5.6365, df = 1, p-value = 0.01759
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.29812684 -0.02879624
sample estimates:
   prop 1    prop 2
0.5961538 0.7596154
```

Test statistic between number of genes expressed between monocytes and CD8 T cells

```
2-sample test for equality of proportions with
continuity correction

data:  c(62, 77) out of c(104, 104)
X-squared = 4.2507, df = 1, p-value = 0.03924
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.280308187 -0.008153351
sample estimates:
   prop 1    prop 2
0.5961538 0.7403846
```

Figure 49: Profile of MS gene module. Upset plots indicate the number of genes induced in three stimulated cell types. The bottom panel of the upset plots shows the number of genes detected in each cell type. The dots represent the cell types where the genes are detected. Line connecting the dots shows the number of genes shared between cell types. The bar plot on the top panel shows the number of genes shared among the cell types and the calculations are given below the figure. The statistics from two-proportions test between the number of genes detected in monocytes and in T cells are shown below the figure.

Hierarchical clustering analysis performed on the MS-module gene set indicates that there are cell type-specific gene clusters (**Figure 50**). Genes that encode for TNF-receptor superfamily (*CD40*), NF-κB inhibitor Zeta (*NFKBIZ*) and members of the glycosyl hydrolase 2 family, Mannosidase beta (*MANBA*), were upregulated in the activated monocytes (**Figure 51**). *CD40* is constitutively expressed in B-cells and DCs. However, CD40 was also expressed in T cells, monocytes and macrophages upon cell activation (Aarts et al., 2017). *NFKBIZ* encodes for IκBζ, which is a repressor of NF-κB signalling and may influence IL6 production and Th17 differentiation. *MANBA* encodes a glycosyl hydrolase localised to the lysosome. The *MANBA* gene is located adjacent to the *NFKB1* gene, which may also be influenced by SNP in the *MANBA* locus (Hussman et al., 2016).

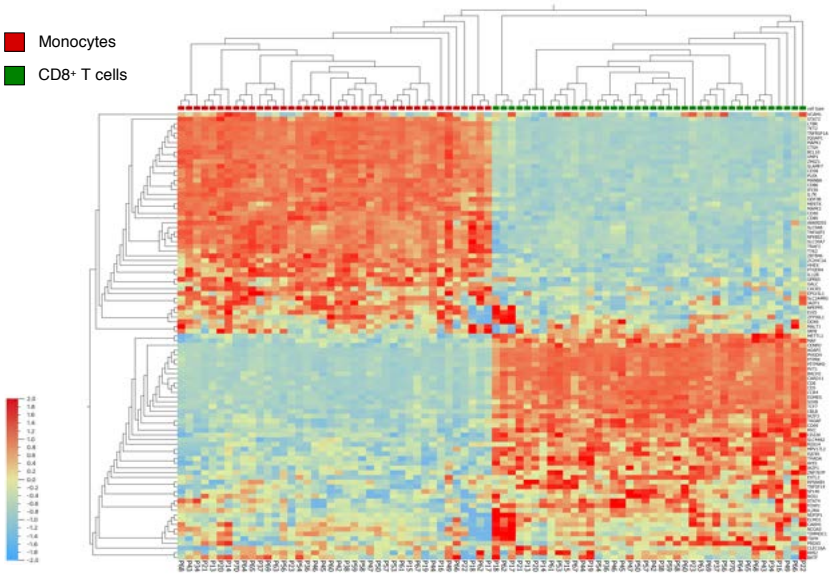


Figure 50: Expression of MS associated gene module. Gene expression pattern of the disease module for MS in activated monocytes and CD8+ T cells is shown in the heatmap ordered by hierarchical clustering. Gene expression data are log_2 transformed, centred to a mean value of zero and scaled to unit variance. Bar to the left represents the scale for the level of gene expression- red indicates higher levels of expression, light blue indicates lower levels of expression and yellow indicates no relative change in the expression value.

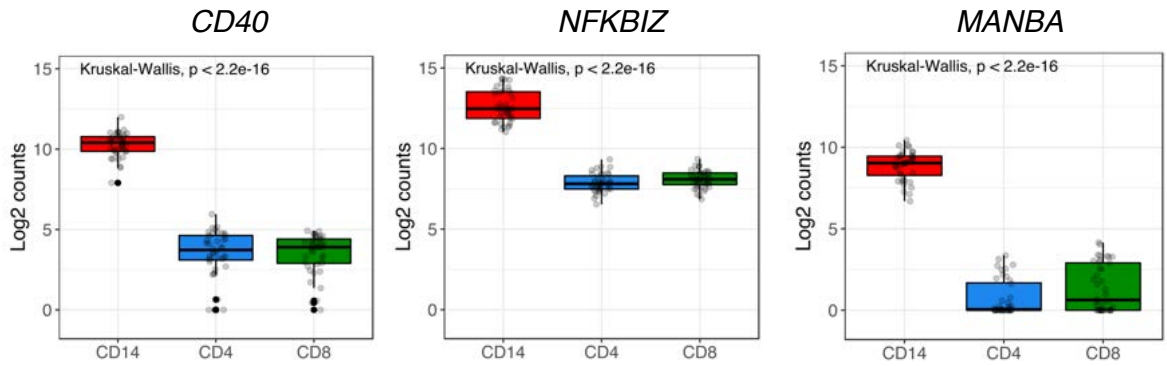


Figure 51: Expression of MS module genes in activated monocytes. Box plots show the expression levels of genes in the LPS activated $CD14^+$ monocytes (in red) and TCR stimulated $CD4^+$ T cells (dark blue) and $CD8^+$ T-cells (dark green). Gene expression data are log_2 transformed. A non-parametric multiple group comparison (Kruskal-Wallis) was performed and the corresponding p-value is indicated within the graph.

Among the genes with higher expression in T cells were: *SOX8*, a transcription factor essential for neural crest development (International Multiple Sclerosis Genetics Consortium, 2013), and *RGS14,* a scaffolding protein that integrates G protein and Ras/ERK/MAP kinase signalling pathways implicated in controlling synaptic plasticity (Lee SE et al., 2010) (**Figure 52**). *TNFSF14* is another member of the tumor necrosis factor (TNF) ligand family, that can act as a costimulatory signal that activates the NF-KB pathway to induce pro-inflammatory cytokine genes (Cheung TC et al., 2009).

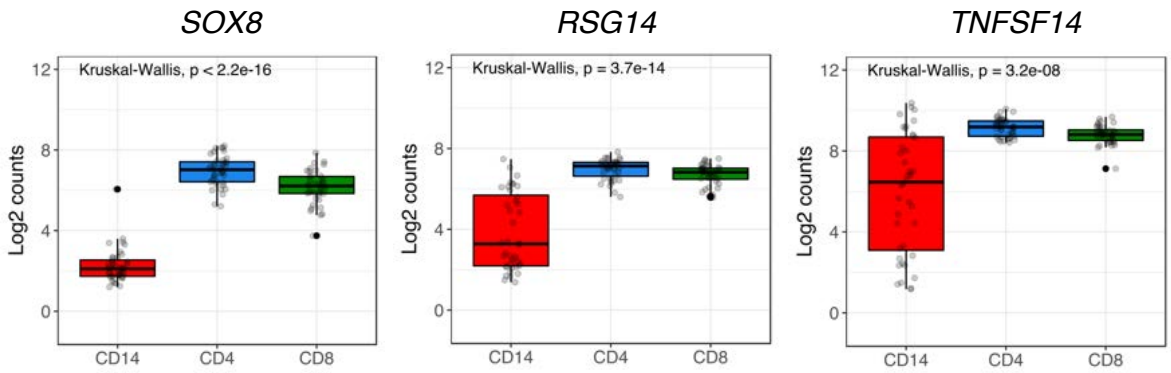


Figure 52: Expression of MS module genes in activated T cells. Box plots show the expression levels of genes in the LPS activated $CD14^+$ monocytes (in red) and TCR stimulated $CD4^+$ T cells (dark blue) and $CD8^+$ T-cells dark green). Gene expression data are log_2 transformed. A non-parametric multiple group comparison (Kruskal-Wallis) was performed and the corresponding p-value is indicated within the graph.

The UpSet analysis and the detected gene expression patterns support the notion that both the adaptive and innate arms of the immune system are involved in orchestrating the immune response in chronic inflammatory conditions. We observed no significant difference in the numbers of genes expressed in T cells and in monocytes in the case of genes associated with AS (p-value = 1), and its extra-articular manifestations. We observed a trend for genes associated with T1D and MS towards larger expression in T cells than in monocytes (p values = 0.069 for T1D and 0.017 for MS), suggesting an important involvement of T cell populations in these diseases as discussed it is discussed in the literature.

This analysis demonstrated that the expression pattern of GWAS-genes is cell-type-dependent. Therefore, to understand the functional role of disease susceptibility genes, it is crucial to characterise them in the appropriate cell type and possibly in the appropriate activation state, as the polymorphisms reported in these genes, therefore may conditionally alter gene function and the subsequent immune response.

6.2.4 Genes of the IL-23/IL-17 axis in chronic inflammatory diseases

From genetic association studies, we observed a substantial overlap among the susceptibility genes for AS, Pso and IBD, suggesting the involvement of common pathways in these diseases. The IL-23/IL-17 axis and Th17-mediated inflammation have been strongly implicated in the pathogenic mechanisms of these diseases by several clinical, experimental and genetic studies (Gaffen et al., 2014; Schön & Erpenbeck, 2018; Sieper, Poddubnyy, & Miossec, 2019). However, the immune cells in which these genes function in the context of chronic inflammatory diseases are not fully understood. Using our NanoString panel we also explored the expression pattern of the genes in the IL-23/IL-17 pathway that were identified as disease-associated through GWAS (**Figure 53**). Gene sets were obtained from the molecular signature database (MSigDB) for the IL-17 signalling pathway (standard name: BIOCARTA_IL17_PATHWAY; systematic name: M19422) and for the IL-23 signalling pathway (standard name: PID_IL23_PATHWAY; systematic name: M196).

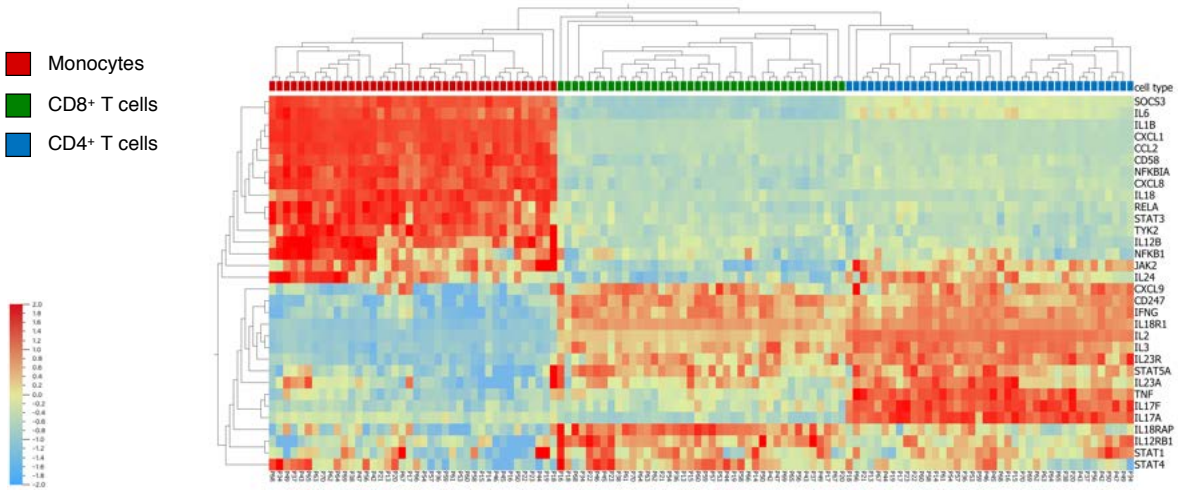


Figure 53: Expression of genes from IL-23/IL-17 pathway in activated innate and adaptive immune cells. Gene expression pattern of genes from the IL-23/IL-17 pathway implicated in AS, Pso and IBD pathogenesis is shown in the heatmap ordered by hierarchical clustering. Gene expression data are log2 transformed, centred to a mean value of zero and scaled to unit variance. Bar to the left represents the scale for the level of gene expression- red indicates higher levels of expression, light blue indicates lower levels of expression and yellow indicates no relative change in the expression value.

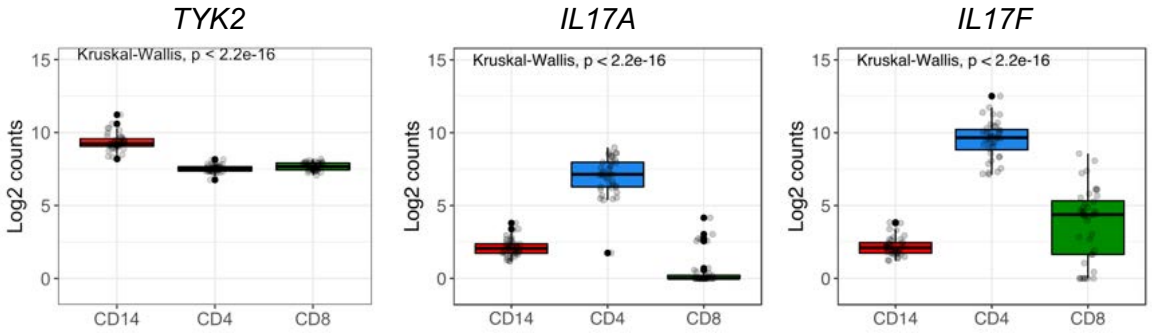


Figure 54: IL-23/IL-17 pathway genes in activated monocytes and T cells. Box plots show the expression levels of genes in the LPS activated $CD14^+$ monocytes (in red) and TCR stimulated $CD4^+$ T cells (dark blue) and $CD8^+$ T cells (dark green). Gene expression data are log_2 transformed. A non-parametric multiple group comparison (Kruskal-Wallis) was performed and the corresponding p-value is indicated within the graph.

We noted that genes involved in Th17 cells and adaptive immunity induction, such as *TYK2* was expressed in activated monocytes (**Figure 54**). Role of *TYK2* in IL-12-mediated pathogenic IFNγ production and IL-23-induced pathogenic IL-17A production was previously shown by Ishizaki et al. (2011) in $Tyk^{-/-}$ mice models (Ishizaki et al., 2011). We observed the cytokine genes that are crucial in mediating Th17 response- *IL17A* and *IL17F* genes were expressed mainly in activated $CD4^+$ T cells (**Figure 54**).

6.2.5 Expression pattern of CID associated genes in innate and adaptive immune cell subsets from healthy donors

Given the cell type-specific gene expression pattern we observed in patients with an inflammatory disease background, we enquired whether the CID-associated genes show a similar expression pattern in healthy controls. We wanted to understand whether the expression of CID-associated genes is influenced by the inflammatory environment observed in patients with AxSpA. We isolated representative cell populations of the innate immune system ($CD56^+$ NK cells and $CD14^+$ monocytes, and adaptive system ($CD4^+$ and $CD8^+$ T cells) from healthy controls (n=3). We also included the innate-like γδT cells. We noted that the main factor driving gene expression variance was the cell type (PC1= 52%) (**Figure 55**).

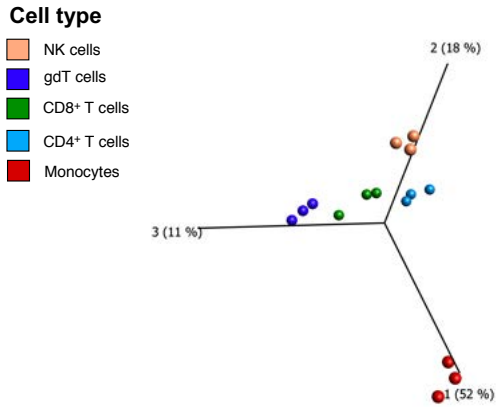


Figure 55: Structure of the gene expression data from innate and adaptive cell populations from healthy donors. Principal component analysis (PCA) conducted in 5 cell populations from 3 healthy controls. Data points are coloured (a.) by cell type as indicated in the key on the left, (b.) by stimulation- IL-18 stimulation (20h) in dark blue, IL-2+TCR stimulation (6 days) in light blue and LPS stimulation (20h) in red. Gene expression data are log_2 transformed, centred to a mean value of zero and scaled to unit variance

We observed that nearly 89% of the genes (671) in our panel was above the background threshold. Hierarchical clustering analysis showed a higher expression of a large number of genes in monocytes stimulated with LPS (**Figure 56a**). From these 671 genes, 453 were CID- associated genes, that showed cell-type-specific gene induction. We noted a group of genes that are expressed in relatively higher levels specifically in LPS-activated monocytes compared to the other activated cell populations. We performed a multi group comparison to across the five cell types, and observed 298 differentially expressed genes ($p< 0.006$, FDR = 1%) (**Figure 56a**). Several CID-linked genes were expressed at relatively higher levels in the innate cell populations- LPS-activated monocytes and then in IL-18 activated NK cells.

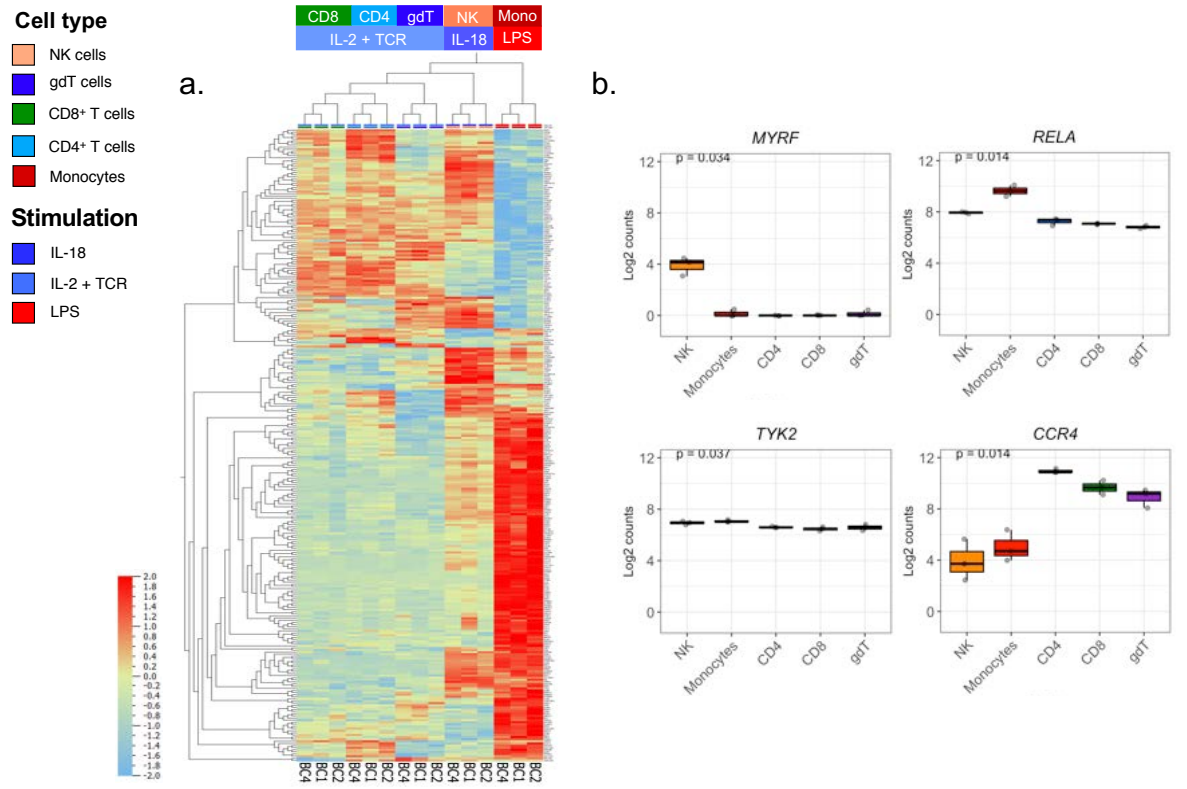


Figure 56: Differential gene expression analysis in disease-associated gene expression analysis in activated innate and adaptive immune cells. (a.) Heatmap shows the hierarchical clustering of differentially expressed genes (in rows) and samples (in columns) across five cell populations are shown, with a p-value = 0.006 and FDR at 1%. The cell type and stimulation conditions are indicated in the key on the left. condition- TCR stimulation (16h) in dark green and the unstimulated group in light green. Genes with expression values below the background threshold were removed. Gene expression data are $\log_2$ transformed, centred to a mean value of zero and scaled to unit variance. Bar to the left represents the scale for the level of gene expression- red indicates higher levels of expression, light blue indicates lower levels of expression and yellow indicates no relative change in the expression value. (b.) Plots showing levels of differentially expressed genes in cell populations. A non-parametric multiple group comparison (Kruskal-Wallis) was performed and the corresponding p-value is indicated within the graph.

We saw a distinct gene expression pattern for disease-associated genes in innate and adaptive immune cells. For example, susceptibility genes for CD were expressed at higher levels in the innate immune cells, such as *MYRF* (myelin regulatory factor) (Liu et al., 2015) in activated NK cells, and the NFκB subunit *RELA* (Liu et al., 2015; Ellinghaus et al., 2016) in LPS-activated monocytes (**Figure 56b**). We observed the C-C chemokine receptor 4 (*CCR4*) associated to MS (Beecham et al., 2013; Patsopoulos et al., 2019), expressed at higher levels in the activated T cells subsets compared to activated monocytes and activated NK cells (**Figure 56b**). Tyrosine kinase 2 (*TYK2*) (**Figure 56b**)- a GWAS gene linked to susceptibility of multiple CIDs was induced in both innate an adaptive cell populations, concordant with the observation from AS patient-derived cell type data (see **Figure 38**).

6.2.5.1 Ankylosing spondylitis susceptibility genes in innate and adaptive cell subsets

We analysed the expression signature in innate and adaptive cell subsets of genes in the 43 loci associated to AS. There was an overall enrichment of susceptibility genes in activated monocytes and NK cells, compared to activated T cells as seen in the hierarchical clustering analysis in **Figure 57a**. The associated genes are differentially expressed across the cell types, for e.g., *EOMES*, *RUNX3*, *IL7R* and Prostaglandin E2 receptor EP4 subtype (*PTGER4*), are upregulated in NK and monocytes and downregulated in the T cells ($p<0.05$) (**Figure 57b**), consistently with the signature observed in the hierarchical clustering.

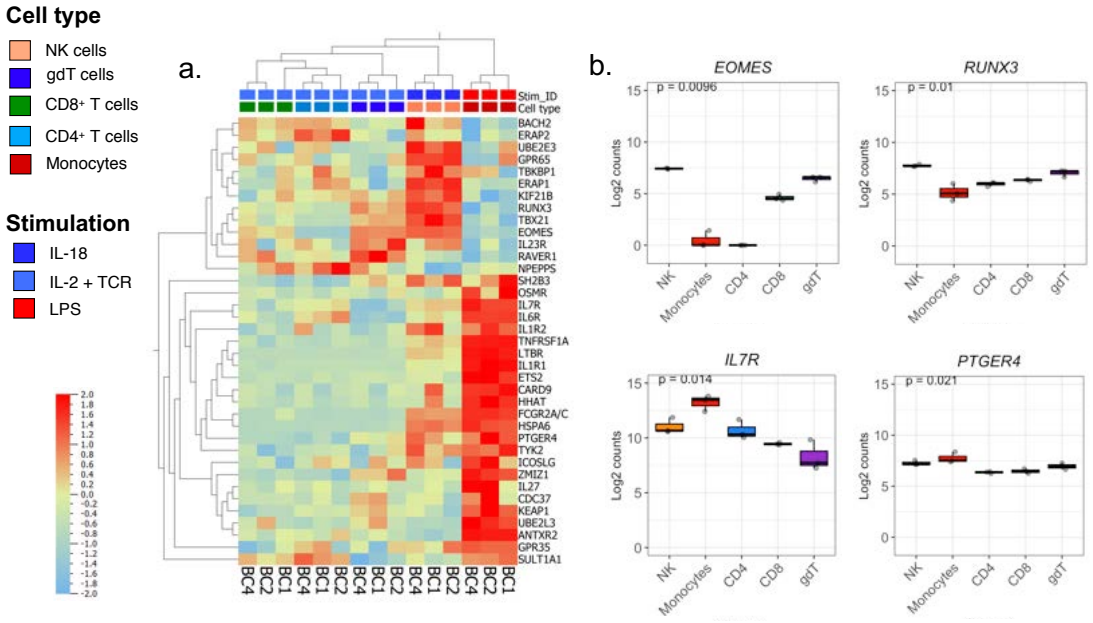


Figure 57: Expression of AS-linked gene module in healthy donor cells. (a.) Gene expression pattern of 37 detected genes from the disease module for AS in activated monocytes, NK cells, $CD4^+$ T cells, $CD8^+$ T cells and γδT cells ordered by hierarchical clustering. Gene expression data are log_2 transformed, centred to a mean value of zero and scaled to unit variance. Bar to the left represents the scale for the level of gene expression- red indicates higher levels of expression, light blue indicates lower levels of expression and yellow indicates no relative change in the expression value. (b.) Levels of genes associated to AS across the five stimulated cell types. A non-parametric multiple group comparison (Kruskal-Wallis) was performed and the corresponding p-value is indicated within the graph.

The expression pattern of AS module genes in NK cells and monocytes suggest that these genes might have a role in the innate immune response. Interestingly, we observed that anti-TNF therapy in AxSpA has a more pronounced impact on the innate immunity pathways than the T_h1/T_h17 immune responses (Menegatti et al, in revision). Innate immune cells act as the "first responders" during an inflammatory event. Therefore, the stronger expression of the susceptibility genes in innate cells could imply that these genes are important in orchestrating the innate immunity signature in AxSpA.

6.2.5.2 IL-23/IL-17 axis genes in innate and adaptive immune cells in healthy donors

We studied the expression of genes of the IL-23/IL-17 pathway in the healthy donor cell type data. Gene set was obtained from the molecular signature database (MSigDB) for the IL-17 signalling pathway and for the IL-23 signalling pathway. Although the IL-23/IL-17 pathway genes implicate the involvement of Th17 immunity, we observed that most of these genes are also induced in the innate immune cells: activated monocytes and NK cells, indicating that they are implicated in additional pathways, besides Th17 cell induction (**Figure 58**).

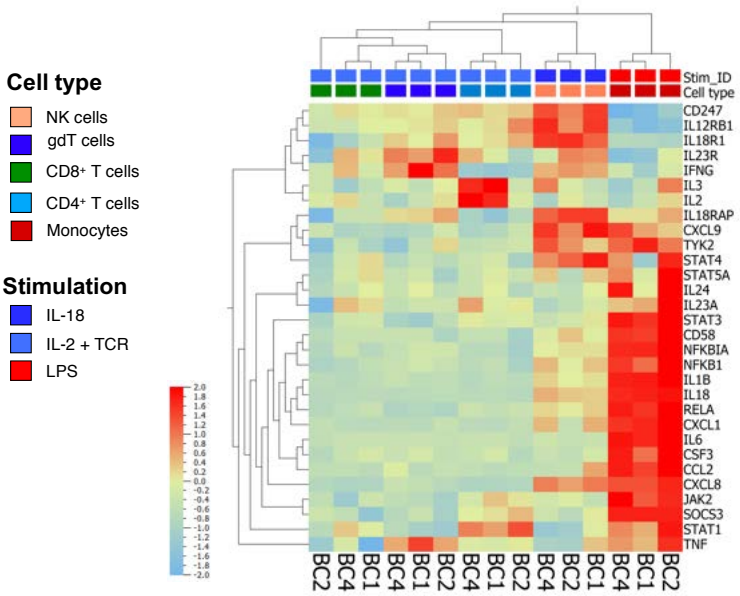


Figure 58: Expression of genes from IL-23/IL-17 pathway in activated innate and adaptive immune cells from healthy controls. Gene expression pattern of genes from the IL-23/IL-17 pathway implicated in AS, Pso and IBD pathogenesis is shown in the heatmap ordered by hierarchical clustering. Gene expression data are $\log_2$ transformed, centred to a mean value of zero and scaled to unit variance. Bar to the left represents the scale for the level of gene expression- red indicates higher levels of expression, light blue indicates lower levels of expression and yellow indicates no relative change in the expression value.

Cytokine genes like *IL17A* and *IL17F* were not detected in stimulated cells from healthy donors. However, we observed that these genes were upregulated in activated $CD4^+$ T cells from AxSpA patients (**Figure 54**). The context-specific induction of IL-23/IL-17 pathway associated genes in innate immune cells could be essential to initiate the Th17-mediated response by the adaptive immune cells.

6.2.5.3 Expression of disease associated genes in MAIT cells from healthy donors

We also analysed the expression pattern of susceptibility genes associated to AS in MAIT cells from healthy donors. The isolated MAIT cells were activated with a combination of cytokines (IL-2 (20U/mL) + IL-1β (10ng/mL) + IL-23 (10ng/mL) and the tetramer MR1/5-OPRU (0.2nM, NIH Tetramer Core Facility). Most of the genes were below the background or detected at very low levels. However, genes such as *TNFRSF1A*, *IL7R* and *ERAP2* were relatively higher in resting MAIT cells (**Figure 59a**), while *IL23R* was expressed upon cell stimulation. Similarly, genes associated to Pso (**Figure 59b**) and IBD (**Figure 59c**) showed elevated levels in resting MAITs. The observed gene induction pattern may suggest that the conditions in which MAIT cells were cultured did not reflect the physiological setting observed in patients that is required to trigger these genes.

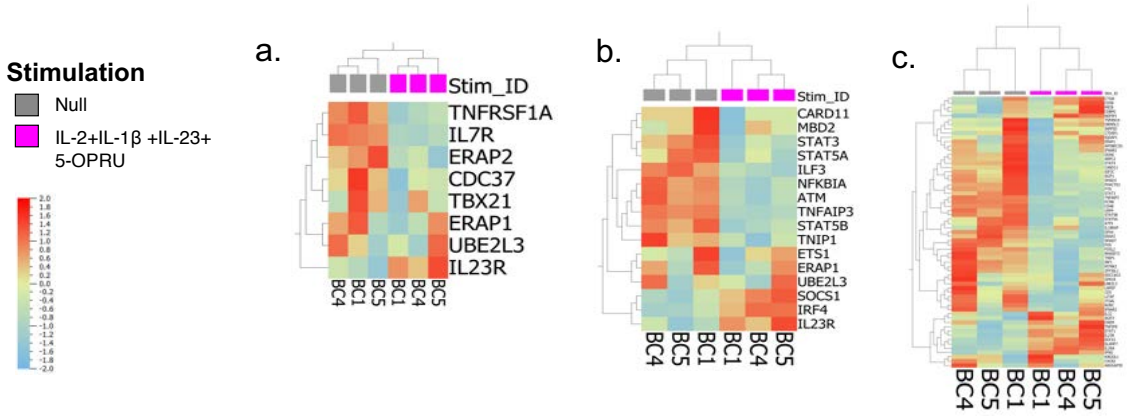


Figure 59: Expression of disease associated genes module in MAIT cells from healthy donors. (a.) Gene expression pattern of detected genes associated to (a.) ankylosing spondylitis, (b.) psoriasis and (c.) IBD, in MAIT cells ordered by hierarchical clustering. Gene expression data are $\log_2$ transformed, centred to a mean value of zero and scaled to unit variance. Bar to the left represents the scale for the level of gene expression- red indicates higher levels of expression, light blue indicates lower levels of expression and yellow indicates no relative change in the expression value.

The expression pattern of CID-associated genes in immune cells in AxSpA patients and healthy controls implies a functional role for these genes in the immune system. The associated genes we studied are the nearest genes with respect to a disease-associated variants. Therefore, we wanted to explore if the susceptibility variants had a role in regulating the nearest genes, using the cell-type-specificity of the genes. To answer this question, we performed an expression-quantitative trait locus (eQTL) mapping study, between the SNP genotype data from patients and the corresponding cell type gene expression data. With this approach, we wanted to identify SNPs that affected the expression levels of the genes in the susceptibility locus in immune cell types.

6.3 eQTL analysis in immune cell populations

To address the second question in the Part I of this project, to identify whether variants associated to CIDs affect the nearest genes in the susceptibility loci, we performed an eQTL mapping in the T cell data. It was shown that context-dependent eQTLs that affect eGenes can help in prioritising putative causal genes and SNPs (Zhernakova et al., 2017). Therefore to elaborate the potential role of SNPs in regulating genes in the susceptibility loci, we performed an eQTL analysis. The eQTL analysis was performed using gene expression data from $CD4^+$ and $CD8^+$ T cells purified from AxSpA patients. The genotyping data of the SNPs from the Illumina Infinium Global Screening Array was included. However, the genotyping array did not include all the SNPs associated with AS. To overcome this shortcoming, we first assessed the genomic loci associated with AS, and imputed the missing SNPs in close proximity with these loci to yield a set of 68442 SNPs for further analysis. Based on the positions of both genes and polymorphisms, we carried out the analysis by separating *cis*-effects (i.e., <250 kb) and *trans*-effects (i.e., >250 kb). The analysis was conducted in resting and activated cell data separately, and detected several significant *cis*- associations. *Trans*-eQTL analysis did not yield significant results. We observed that the majority of the *cis*-hits detected in the data were in resting cell data as given below (**Table 8**).

Table 8: The number of detected cis-eQTLs from the T cell data

	CD4+ T cells		CD8+ T cells	
eGene	**Resting**	**Activated**	**Resting**	**Activated**
FADS2	52	13	4	3
CTSW	15	-	12	1

6.3.1 *Cis*-eQTL analysis in resting and activated $CD4^+$ T cells

We observed 67 *cis*-eQTLs in the unstimulated dataset that correlated with the expression of the eQTL-genes (eGenes) *FADS2 (*fatty acid desaturase 2) and *CTSW* (cathepsin W), located on chromosome 11. The top significant *cis*-eQTL hits are shown in the Manhattan plot in **Figure 60a**, and the top 10 significant hits are listed in **Table 9** along with the test statistic (z-score >5) and adjusted p-values (<0.05). The eQTL effect of two of the top significant hits on the expression of eGenes from the resting $CD4^+$ T cells data is shown in the boxplots in **Figure 60b**.

The majority of the individuals have the alternative allele (G) for rs658524 (allele frequency of 0.81 in the European population, from the 1000 Genome (Auton et al., 2015), which is correlated with high expression of *CTSW* in resting CD4$^+$ T cells. The eQTL SNP rs968567 has an allele frequency of 0.85 for the reference allele ("C") in the European population (1000 Genomes), which is correlated with reduced levels of *FADS2*.

Table 9: The number of detected *cis*-eQTLs in the resting CD4$^+$ T cell data

SNPs	Position (hg19)	eGene	Test statistic	Adjusted p-value
rs968567	61595564	*FADS2*	8.816	0.0001305
rs61897793	61599347	*FADS2*	8.816	0.0001305
rs658524	65647260	*CTSW*	8.612	0.0001305
rs10791830	65661291	*CTSW*	-8.415	0.0001305
rs7943728	61547068	*FADS2*	8.368	0.0001305
rs61896141	61556039	*FADS2*	8.368	0.0001305
rs693824	65654783	*CTSW*	-8.177	0.0001305
rs2308217	65655259	*CTSW*	8.177	0.0001305
rs641018	65655393	*CTSW*	-8.177	0.0001305
rs656980	65656282	*CTSW*	-8.177	0.0001305

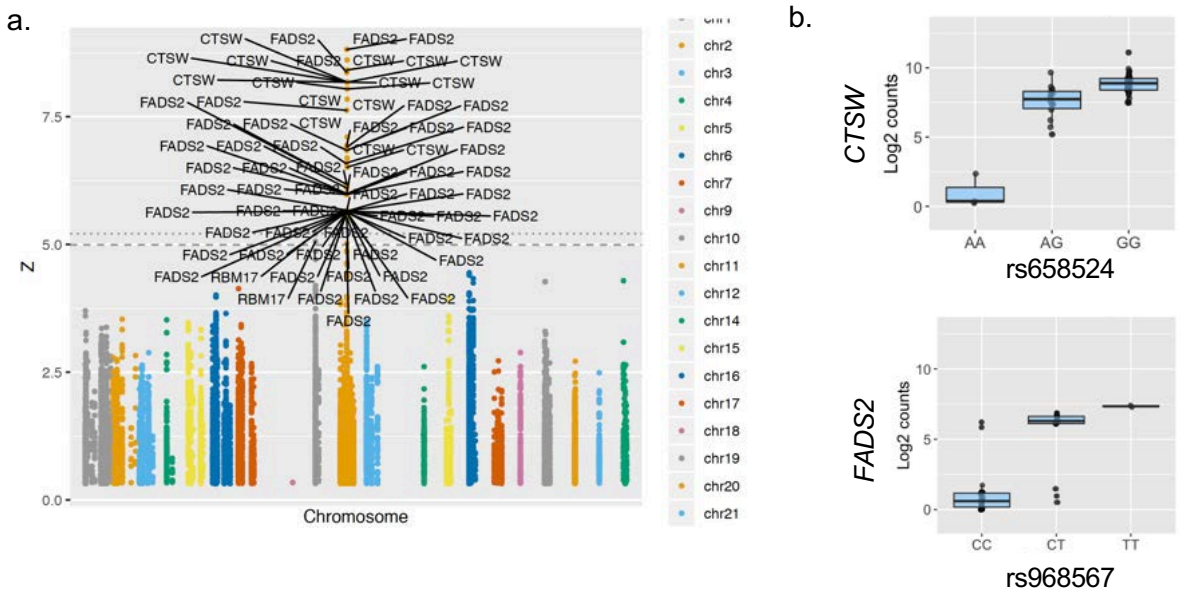


Figure 60: cis-eQTLs identified from resting CD4$^+$ T cells. (a) Manhattan plot showing the significant cis-eQTL hits. The x-axis shows all the chromosomes and the coloured dots denote the SNPs tested. The horizontal lines on the graph (dotted and dashed) represent the adjusted-P < 0.05 and adjusted-P < 0.1.(b.) Box plots showing the levels of eGene expression for CTSW and FADS2. The normalised gene expression (log_2 transformed counts) is on the y-axis and the SNP genotype group is on the x-axis.

In the TCR-activated data set, we observed only 13 significant *cis*-eQTLs, which correlated with the expression of only one eGene - *FADS2*. **Table 10** shows the SNPs linked to *FADS2* expression with the position of the SNP and the corresponding test statistic and p-value. The SNPs detected above the level of significance threshold are indicated in the Manhattan plot in **Figure 61a**. The reference allele of SNP rs174534 (A) (frequency of 0.65 in the European population, 1000 Genomes) correlated with lower expression levels of *FADS2* (**Figure 61b**). A similar pattern in the direction of the allelic effect was observed for rs968567, as shown in the box plot in **Figure 61b**.

Table 10: The number of detected *cis*-eQTLs in the TCR-stimulated CD4⁺ T cell data

SNPs	Position (hg19)	eGene	Test statistic	Adjusted p-value
rs968567	61595564	*FADS2*	6.06	0.003119
rs61897793	61599347	*FADS2*	6.06	0.003119
rs174582	61607168	*FADS2*	5.871	0.005718
rs7943728	61547068	*FADS2*	5.753	0.007927
rs61896141	61556039	*FADS2*	5.753	0.007927
rs108499	61547237	*FADS2*	5.521	0.01962
rs174534	61549458	*FADS2*	5.521	0.01962
rs174538	61560081	*FADS2*	5.521	0.01962
rs174559	61581656	*FADS2*	5.424	0.02534
rs174585	61611694	*FADS2*	5.321	0.03405

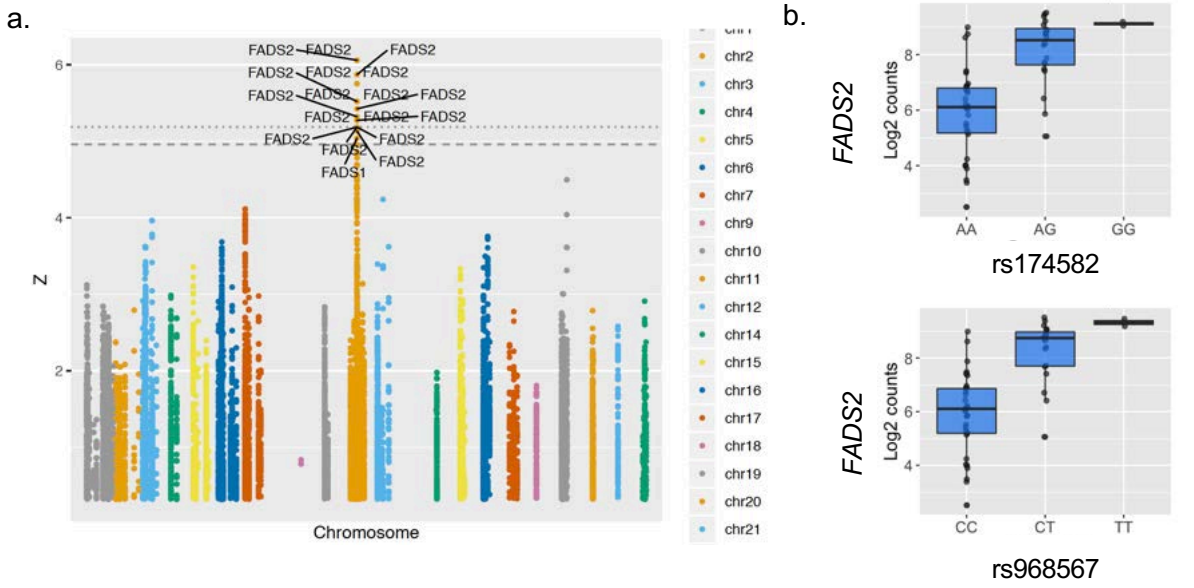


Figure 61: cis-eQTLs identified in TCR-activated CD4⁺ T cells. (a) Manhattan plot showing the significant cis-eQTLs. The x-axis shows all the chromosomes and the coloured dots denote the SNPs tested. The horizontal lines on the graph (dotted and dashed) represent the adjusted-P < 0.05 and adjusted-P < 0.1. (b) Box plots showing the levels of eGene FADS2. The normalised gene expression $\log_2$ transformed counts are on the y-axis and the SNP genotype group is on the x-axis.

6.3.2 *Cis*-eQTL analysis in resting and activated CD8⁺ T cells

In the unstimulated dataset we detected 16 *cis*-eQTLs that correlated with the expression of the genes *FADS2* and *CTSW* located on chromosome 11 (Manhattan plot in **Figure 62a**). A list of the top 10 significant SNPs is listed in the **Table 11**. The eQTL effect on the expression of *FADS2* and *CTSW* is shown in the boxplots in **Figure 62b**. The reference allele of rs10791830 (C) (frequency: 0.81) is correlated with high CTSW expression (**Figure 62b**). The reference allele for the *FADS2* SNP rs108499 (ref: C, AF= 0.67, European population, 1000 Genomes) correlates with lower eGenes expression (**Figure 62b**).

Table 11: The number of detected *cis*-eQTLs in the resting CD8⁺ T cell data

SNPs	Position (hg19)	eGene	Test statistic	Adjusted p-value
rs658524	65647260	*CTSW*	6.548	0.0019
rs10791830	65661291	*CTSW*	-6.004	0.009299
rs568617	65653242	*CTSW*	-5.967	0.009999
rs694994	65653309	*CTSW*	-5.967	0.009999
rs677931	65654665	*CTSW*	-5.967	0.009999
rs693824	65654783	*CTSW*	-5.933	0.0114
rs641018	65655393	*CTSW*	-5.933	0.0114
rs656980	65656282	*CTSW*	-5.933	0.0114
rs507062	65661656	*CTSW*	-5.933	0.0114
rs588114	65662752	*CTSW*	-5.933	0.0114

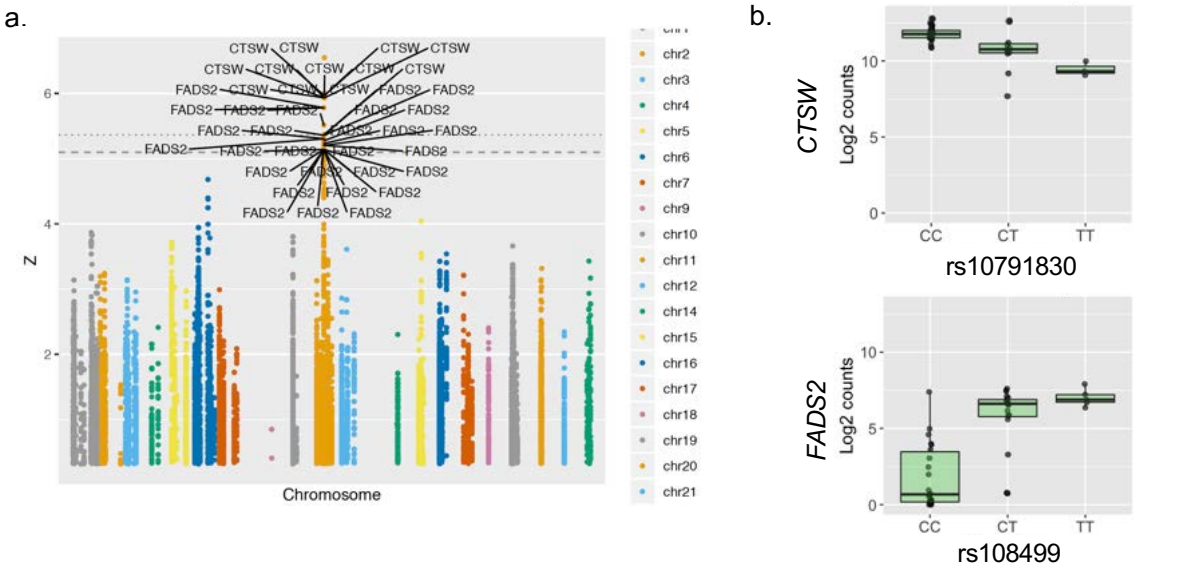


Figure 62: *cis*-eQTLs identified from resting CD8⁺ T cells. (a) Manhattan plot showing the significant cis-eQTL hits. The x-axis shows all the chromosomes and the coloured dots denote the SNPs tested. The horizontal lines on the graph (dotted and dashed) represent the adjusted-P < 0.05 and adjusted-P < 0.1. (b.) Box plots showing the levels of eGenes CTSW and FADS2. The normalised gene expression $\log_2$ transformed counts are on the y-axis and the SNP genotype group is on the x-axis.

In the TCR activated CD8 data we observed only 4 significant *cis*-eQTLs, and these correlated with the expression of eGenes *FADS2* (rs108499, rs174534 & rs174538) and *CTSW* (rs658524) with a z-score >5 and adjusted p-values p<0.05 (**Table 12** & **Figure 63a**). eQTLs rs658524 and rs174534 in **Figure 63b** showed the same allelic effect direction for the gene expression levels of *CTSW* and *FADS2*. Some of the detected eQTLs in $CD8^+$ T cell data were also shared in resting and/or activated $CD4^+$ T cells. These findings suggest that these eQTLs are potentially regulatory and could control the eGenes in a similar manner in both $CD4^+$ and $CD8^+$ T cells.

Table 12: The number of detected *cis*-eQTLs in the TCR-stimulated $CD8^+$ T cell data

SNPs	Position (hg19)	eGene	Test statistic	Adjusted p-value
rs658524	65647260	*CTSW*	5.868	0.0116
rs108499	61547237	*FADS2*	5.458	0.0336
rs174534	61549458	*FADS2*	5.458	0.0336
rs174538	61560081	*FADS2*	5.458	0.0336

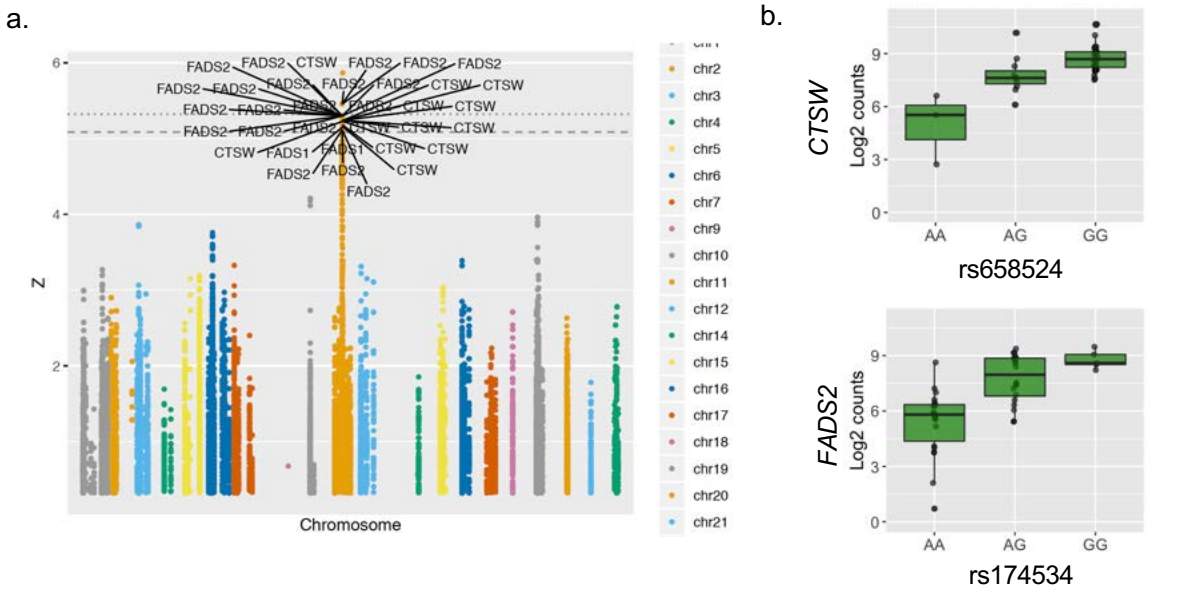


Figure 63: *cis*-eQTLs identified from TCR-activated $CD8^+$ T cells. (a) Manhattan plot showing the significant cis-eQTL hits. The x-axis shows all the chromosomes and the coloured dots denote the SNPs tested. The horizontal lines on the graph (dotted and dashed) represent the adjusted-P < 0.05 and adjusted-P < 0.1. (b) Box plots showing the levels of eGenes CTSW and FADS2. The normalised gene expression $\log_2$ transformed counts are on the y-axis and the SNP genotype group is on the x-axis.

This analysis highlighted a region on chromosome 11 as a region of interest for T-cell function in the context of AS, since several of the SNPs identified as eQTLs had been associated with immune and inflammatory diseases through GWAS. For example, rs10791830, an eQTL for *CTSW* in resting $CD4^+$ and $CD8^+$ T cells, was previously associated to IBD and Crohn's disease (De Lange et al., 2017), and most recently with AS susceptibility based on the round 2 analysis of the UK Biobank data (http://www.nealelab.is/uk-biobank/).

6.3.3 Cell-type specific expression of *FADS2* and *CTSW*

To investigate the cell-type-specific expression of *FADS2* and *CTSW*, we surveyed published studies and gene expression databases to assess the expression pattern of these eGenes in immune cell subsets. From the DICE database we observed that *FADS2* was expressed at high levels in NK cells ($CD16^{+}CD56^{+}$), followed by classical and non-classical monocytes, and B-cells (**Figure 64**) (Schmiedel et al., 2018). *FADS2* was expressed at relatively low levels in T-cell subsets as seen in the box plot below (**Figure 64**).

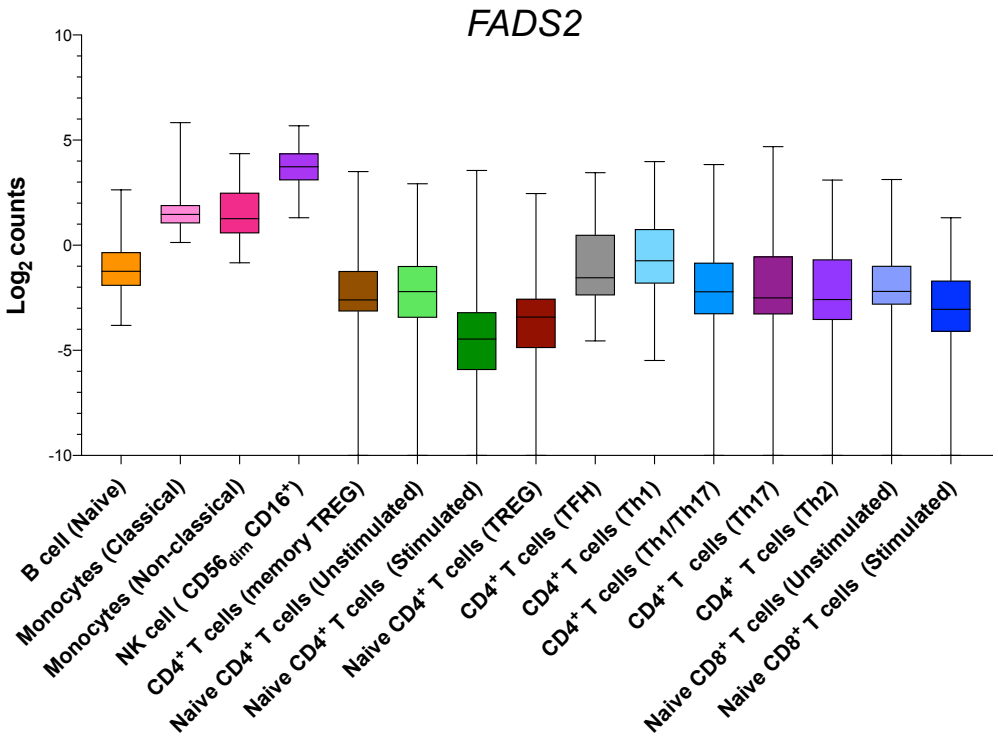


Figure 64: Expression levels of eGene FADS2 in immune cell subsets. The gene expression of FADS2 as measured by Schmeidel et al. using bulk RNA-seq from 15 immune cell subsets from 91 donors (https://dice-database.org/genes/FADS2). The expression is shown in the y-axis as log2 transformed normalised transcripts per million, labelled as "Log2 counts".

The *FADS* cluster locus is a polymorphic region consisting of fatty acid desaturase genes 1,2 and 3 (*FADS1*, *FADS2* & *FADS3*), which have been previously reported in genetic studies concerning immune (Ellinghaus et al., 2016) and cardiometabolic traits (Chen et al., 2019). A recent integrative gene regulatory network analysis reported allele specificity in transcription factor binding for *FADS2* regulation (Van Der Wijst et al., 2018). This study reported the regulatory role of eQTL rs968567, an RA-associated GWAS variant, for *FADS2* expression, by identifying allele-specific binding to this sequence of a sterol binding factor (SREBF2), using ENCODE ChIP-seq data. Additionally, eQTLs for *FADS2* were previously

detected in monocytes (Zeller et al., 2010) and *FADS2* expression has been well-characterised in tissue studies from GTEx data (Reynolds et al., 2018).

On the other hand, we noted that *CTSW* is expressed at high levels in CTLs CD8$^+$ T cells and NK cells (**Figure 65**), suggesting its role in cytotoxicity-associated activity. As the eQTLs detected for *CTSW* were in CD8$^+$ T cells, and since *CTSW* is predominantly expressed in CTLs, we explored the role of the eQTLs in regulating *CTSW* expression in CD8+ T cells.

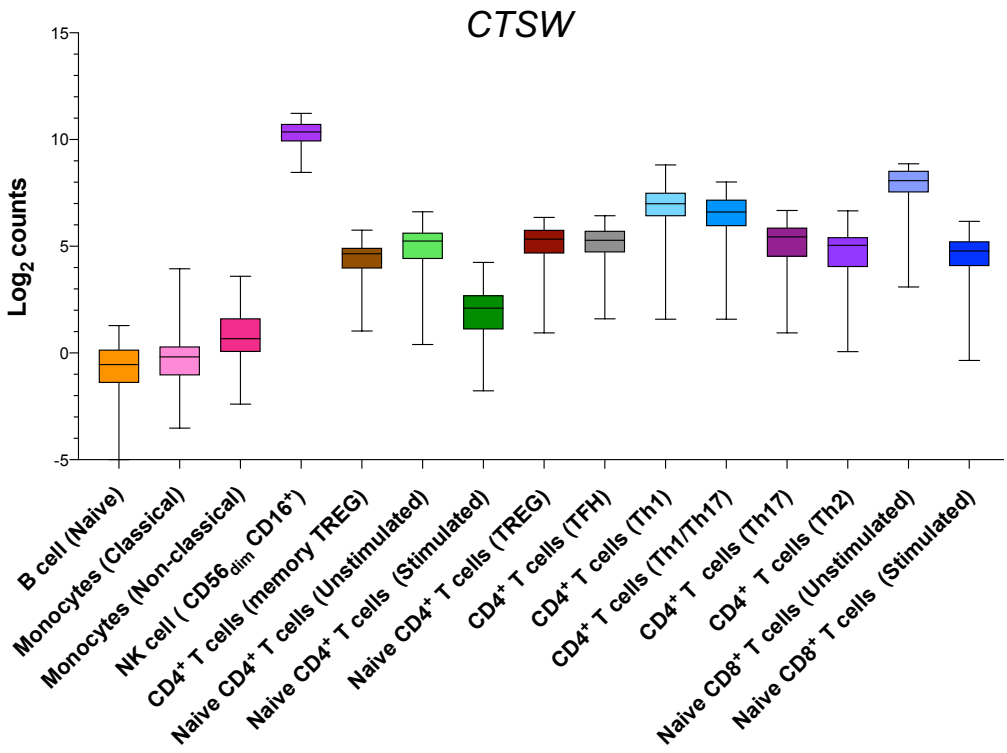


Figure 65: Expression levels of eGene CTSW in immune cell subsets. The gene expression of FADS2 as presented by Schmeidel et al. using bulk RNA-seq from 15 immune cell subsets from 91 donors (https://dice-database.org/genes/CTSW). The expression is shown in the y-axis as log2 transformed normalised transcripts per million labelled as “Log2 counts”.

6.3.4 Cysteine cathepsins in the immune system

Cathepsins are proteases that are classified into three categories by the amino acid residues present in their active sites, namely, aspartic proteases, serine protease and cysteine proteases. These enzymes require an acidic environment to be active. Due to their selective activity in low pH, they were initially thought to be intracellular enzymes involved in proteolytic activities in the endosomes and lysosomes (Turk, 2001). The role of cathepsins is not confined to the endo-lysosomal system as they have been later detected in the nucleus, cytoplasm and the extra-cellular and extra-lysosomal space (Brix et al., 2008).

Most cysteine cathepsins show ubiquitous expression in human tissues - for example, B, C, F, H, L, O, and X cathepsins, whereas cathepsins K, S, V and W show tissue- and cell type-restricted expression suggesting a specific role. Cysteine cathepsins are involved in several innate and adaptive immune responses due to their constitutive expression in immune cells (**Figure 66**). For example, cathepsins B,F, L and H are found in macrophages; cathepsin C has been reported in CTLs, macrophages, granulocytes and mast cells. Similarly, cathepsin K is expressed at high levels in osteoclasts, epithelial cells and synovial fibroblasts in RA joints. Cathepsin S is present in antigen presenting cells (DCs, B-cells and macrophages) and is implicated in MHC class II antigen presentation (Vidak et al., 2019). Cathepsin W is predominantly expressed in cytotoxic cells like $CD8^+$ T cells and NK cells and is implicated in cytotoxicity (Brown et al., 1998; Wex et al., 2003).

Cysteine cathepsins are involved in a wide range of cellular and physiological functions such as ion channel activity, innate immunity, complement activation, apoptosis, peptide synthesis and vesicular trafficking (Brix et al., 2008; Repnik et al., 2012). Cysteine cathepsins are known to mediate autophagy, which is an essential cellular process that enables lysosomal degradation of misfolded protein. Dysregulated autophagic pathway has been associated with diseases such as rheumatoid arthritis, dermatitis, osteoporosis, IBD, cardiovascular disease and cancer. (Patel et al., 2018).

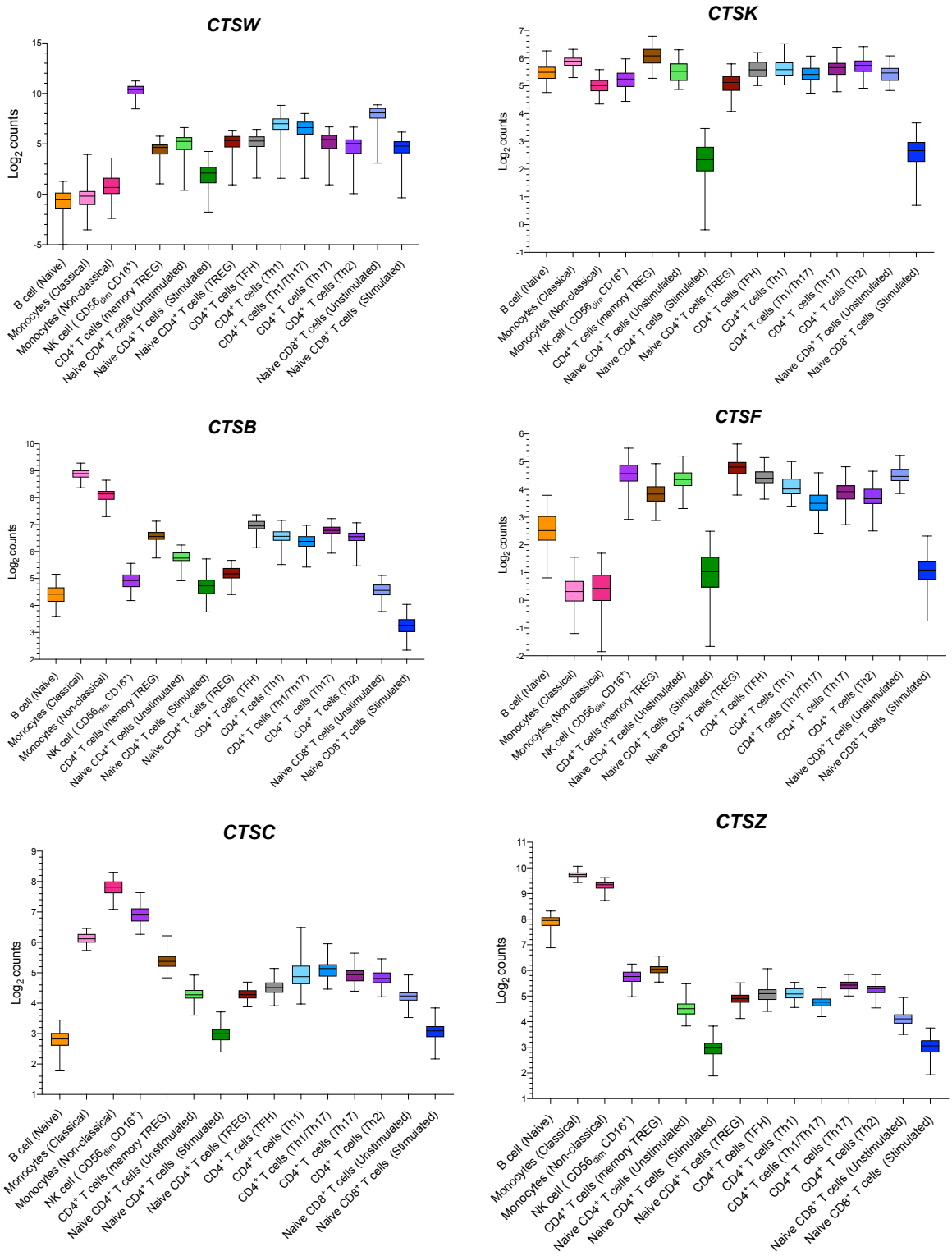


Figure 66: Expression levels of cathepsins in immune cell types. The expression data is acquired from the DICE database (Schmiedel et al., 2018).

6.3.5 *CTSW* eQTLs are lead variants from GWAS

Previous studies have established that trait-associated SNPs are more likely to be eQTLs (Nicolae et al., 2010). We therefore assessed if the eQTLs we identified for *CTSW* are also lead variants from GWAS in AS and other CIDs. We searched for lead variants in the CTSW locus associated to CIDs in the GWAS Catalog (https://www.ebi.ac.uk/gwas/). We also consulted an updated online resource to verify the associated lead SNPs with CIDs from the UK Biobank PheWAS analysis (http://www.nealelab.is/uk-biobank/).

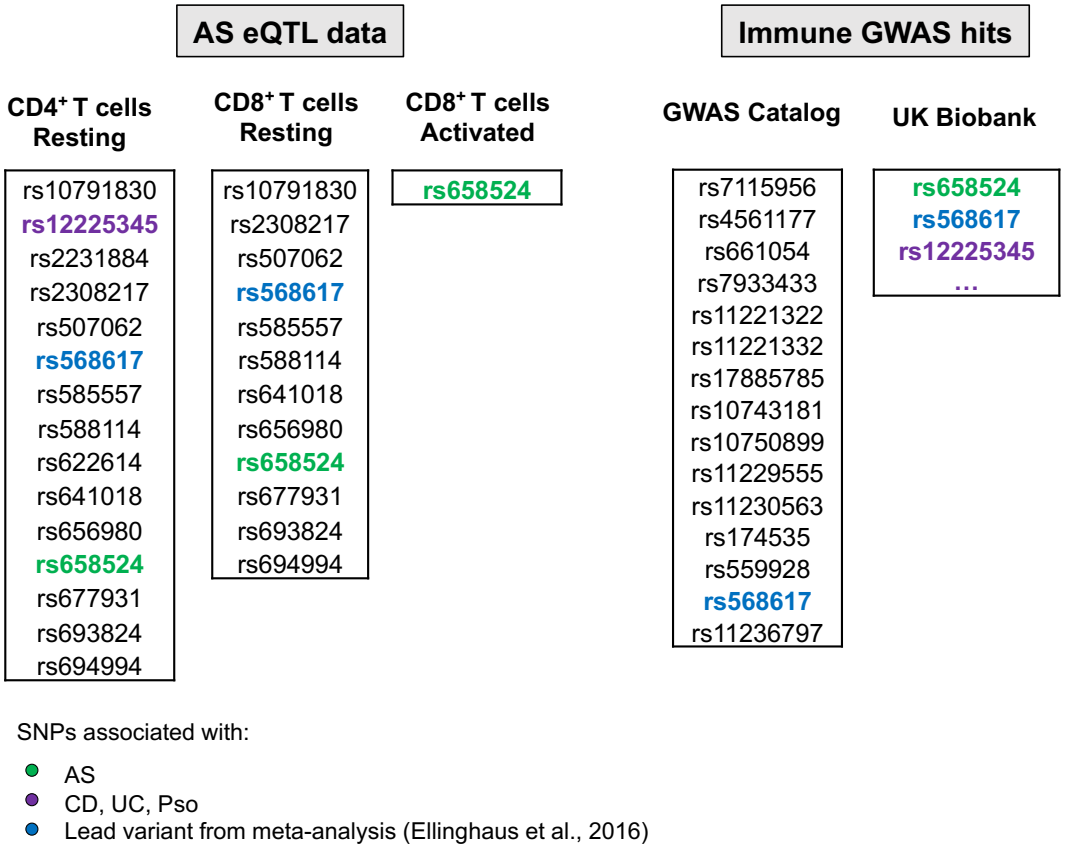


Figure 67: List of eQTL SNPs from our AS data that are lead variants reported in immune GWAS. The eQTLs that are also a lead SNP associated with AS is indicated in green, with IBD (CD & UC) and Pso is in purple and the lead variant from the meta-analysis study is in dark blue.

We identified three SNPs - rs658524, rs568617, and rs12225345 - that were associated with CID susceptibility from GWAS (**Figure 67**). The SNP rs12225345 was reported to be associated with IBD (De Lange et al., 2017) and Pso (UK Biobank PheWAS, http://www.nealelab.is/uk-biobank/). SNPs rs658524 and rs568617 were initially detected as GWAS variants for CD (Liu et al., 2015) and IBD (De Lange et al., 2017). These two SNPs

were later identified as risk variants for AS in the UK Biobank round 2 PheWAS analysis (http://www.nealelab.is/uk-biobank/).

We assessed the LD structure of the 15 unique and significant eQTL SNPs in the *CTSW* locus using LDlink (https://ldlink.nci.nih.gov/?tab=ldmatrix) to calculate the pairwise LD between SNPs. The haplotypes were extracted from the 1000 Genomes European population. We observed a correlation coefficient, $r^2 > 0.8$, showing that the SNPs we tested are in strong LD, as shown in the correlation matrix in **Figure 68**.

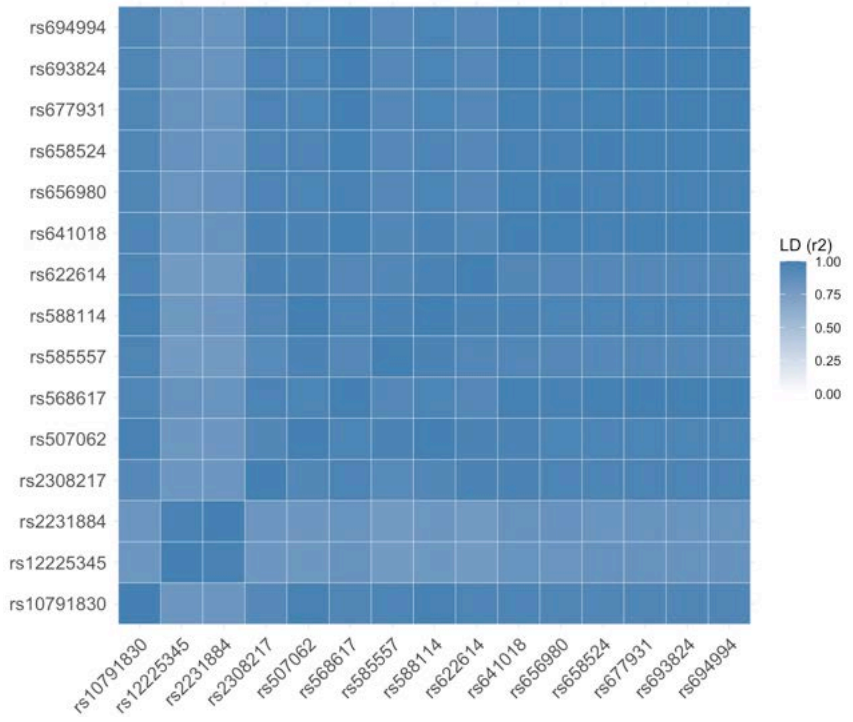


Figure 68: Pairwise LD matrix of the SNPs. LD correlation matrix for the variants in the CTSW locus that are eQTLs and lead GWAS SNPs. LD values were computed from LDlink in the European population from 1000 Genomes.

The regional plot indicating the location of the SNPs in the *CTSW* locus is shown in **Figure 69**. The plot indicates the position of the 15 SNPs we tested in the locus with respect to the gene location. The adjusted p-values from our eQTL analysis are shown on the y-axis. The SNPs are clustered near the gene body of *CTSW*. The tight linkage of the SNPs (red dots) to the top SNP rs658524 (purple diamond) is indicated by the r^2 values. LD values were computed from the 1000 Genomes (March 2014 release) European population using LocusZoom (Pruim et al., 2011).

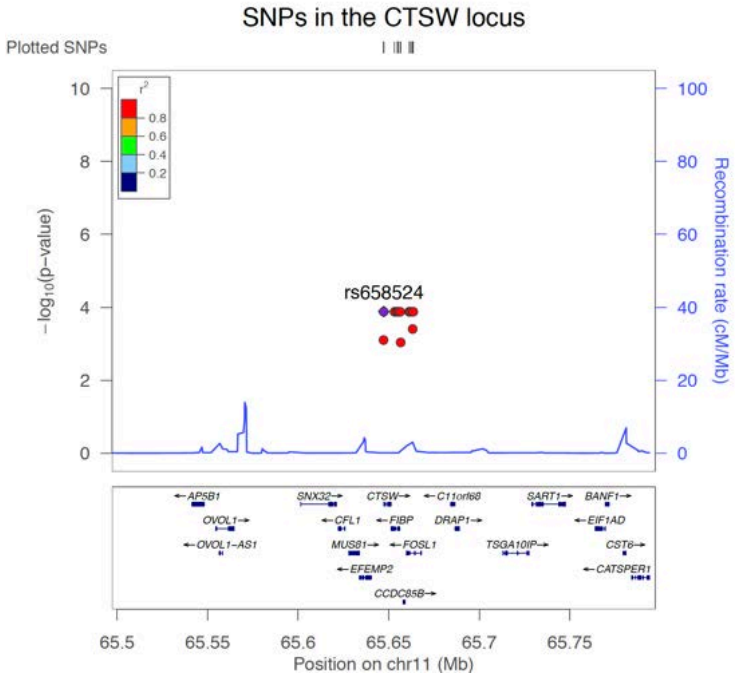


Figure 69. Regional plot for the significant eQTLs over the CTSW locus. The SNPs are coloured by the strength of the LD association (r^2) as shown by the scale on the left. The adjusted p-value from the eQTL analysis is shown in the y-axis as $-\log_{10}$ scale. The genomic region on the chr11 where the SNPs are located with respect to CTSW gene is indicated in the bottom panel of the LocusZoom plot.

The strong LD among the eQTL SNPs makes it difficult to pin-point the functional SNP that regulates CTSW expression. To gain biologically meaningful insights from our eQTL data, we integrated in our study the analysis of epigenetic regulatory marks of the locus by examining the publicly available chromatin conformation data at the SNP location. Using the dbSNP data track of the UCSC genome browser, we located the SNPs in the *CTSW* locus. The eQTL SNPs from our analysis are marked in fig. 70 with red rectangle boxes. We added the Roadmap Epigenomics data track available on UCSC genome browser (GRCh37/hg19) for regulatory element information. This includes DNase hypersensitivity sites (DHS) and histone marks from $CD4^+$ and $CD8^+$ T cells, as shown in **Figure 70** (http://genome.ucsc.edu/cgi-bin/hgGateway) (Kundaje et al., 2015).

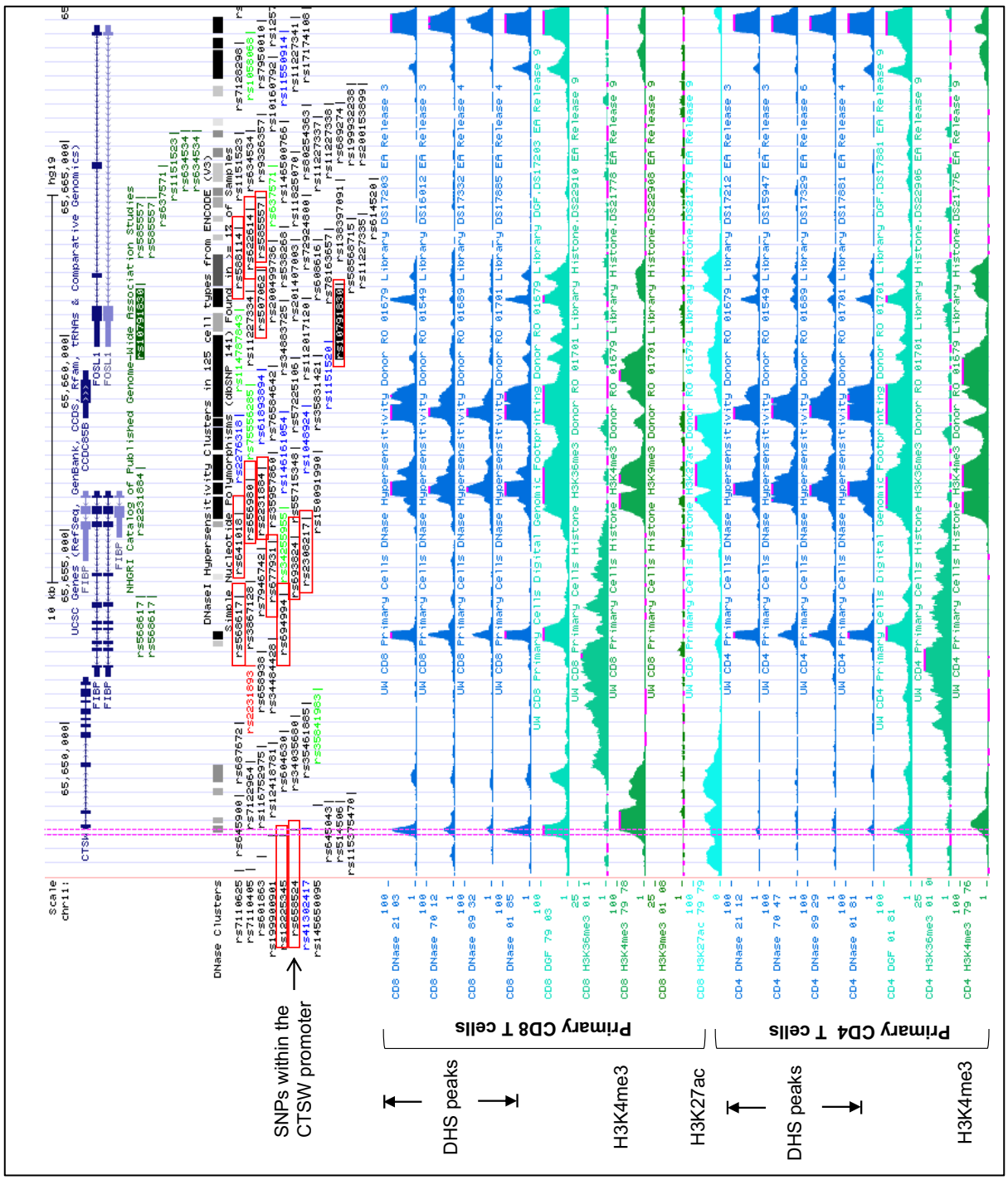
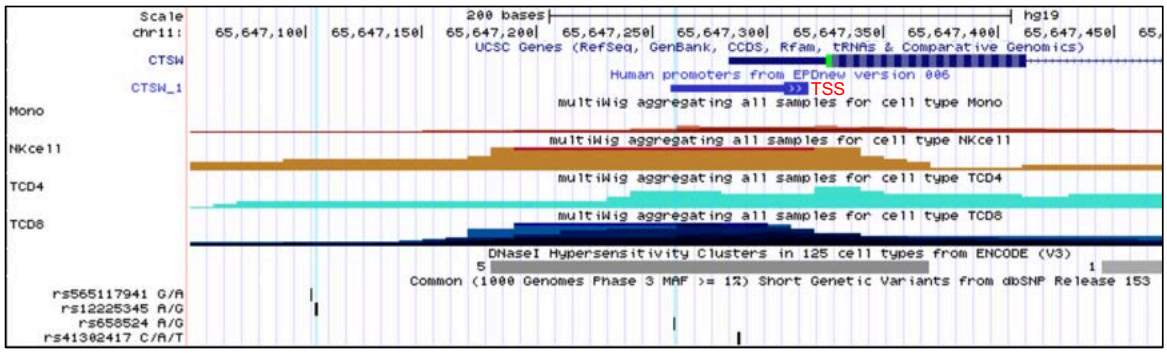


Figure 70: Regulatory information at the *CTSW* locus on Chr11. The genomic coordinates are extracted from the UCSC genome browser (hg19). The top most panel shows the position of the gene transcripts in the locus (UCSC, RefSeq). The eQTL SNPs in LD are indicated by red boxes. *CTSW* promoter SNPs are indicated in magenta dotted lines. The regulatory data tracks were added-on from Roadmap Epigenomics Project. The DHS peaks are marked in dark blue. The dark green peaks indicate the H3K4me3 peaks, and H3K27ac peaks are shown in turquoise.

We assessed the co-occurrence of DHSs and methylation marks with the SNP positions (**Figure 70**). From this analysis, we observed that two eQTL SNPs (rs658524 & rs12225345, marked in magenta dotted lines), are in the *CTSW* promoter region and coincided with DHS peaks in $CD8^+$ primary T cells. Additionally, these SNPs also overlapped with marks of active promoters and enhancers (H3K4me3 and H3K27ac) in $CD8^+$ T cells (**Figure 70**). This evidence strongly suggests a possible regulatory mechanism for these eQTLs in $CD8^+$ T cells.

DHS peaks overlapping these SNPs in $CD4^+$ T cells were less pronounced, compatibly with lower expression of CTSW in these cells, compared to $CD8^+$ T cells (**Figure 70**). However, there were some H3K4me3 peaks in the CD4 cells. This observation may indicate a potential regulatory activity for these SNPs in regulating *CTSW* in $CD4^+$ T cells also.

rs658524 is located 48bp upstream of the transcription start site (TSS) and rs12225345 at 203bp upstream of the TSS for *CTSW* (**Figure 71**). The SNP location with respect to the TSS and overlapping chromatin accessibility peaks from immune cell subsets are indicated in **Figure 71**. The chromatic accessibility information was obtained from the UCSC data track for ATAC-seq data generated by Calderon et al. from immune cell subsets (Calderon et al., 2019).



	Distance from *CTSW* TSS
rs658524	- 48 bp
rs12225345	- 203 bp

Figure 71: Location of the SNPs reported at the *CTSW* promoter, as distance from the TSS. SNP positions are marked by pale blue lines. The TSS location for *CTSW* obtained from the Eukaryotic promoter database (EPD) is indicated by the blue box with white arrows showing the direction of transcription. The ATAC-seq peaks from monocytes, NK, $CD4^+$ and $CD8^+$ T cells are shown.

The UK Biobank PheWAS analysis reported the "G" allele (AF= 0.81) of rs658524 as correlated with AS risk. We noted that patients with the GG genotype showed elevated levels of *CTSW,* as shown in the box plot in **Figure 72a**. In our resting CD4$^+$ T cell eQTL data, we observed a significant eQTL effect from rs12225345 correlated with low levels of *CTSW* expression in the G allelic group, and a similar trend in the activated CD8$^+$ T cell data (**Figure 72a**). These two SNP are also in LD with each other (r^2 > 0.8) (**Figure 72b**): the G allele of rs658524 that correlates with increased *CTSW* expression, is linked with the A allele of rs12225345. The high levels of *CTSW* in patients with the corresponding SNP genotype may indicate a potential functional link between these SNPs and *CTSW* regulation in CD8$^+$ T cells.

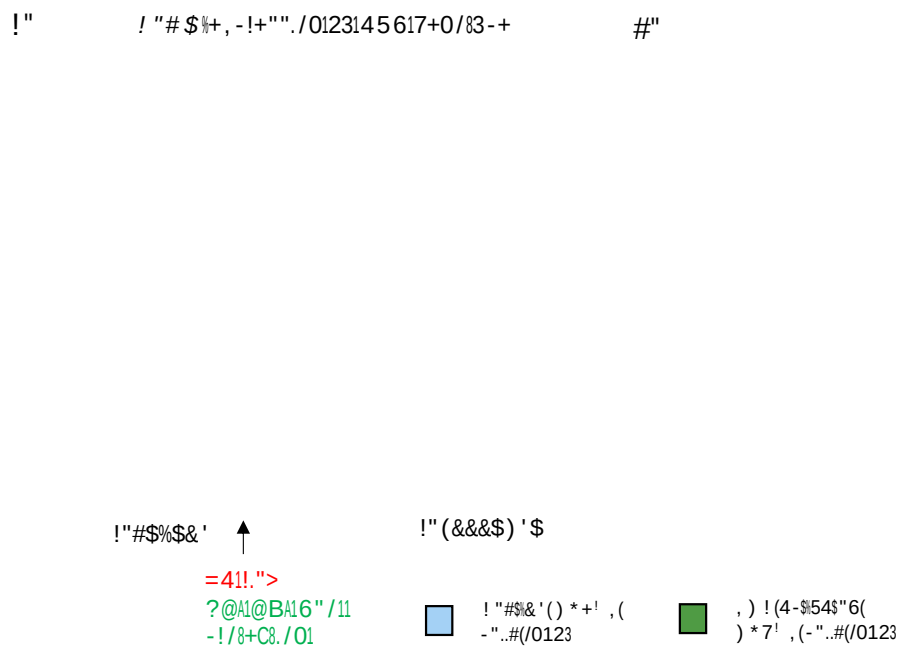


Figure 72: Allelic direction for the eQTL effect of two promoter SNPs on CTSW expression. (a.) Box plots show the levels of CTSW mRNA in resting CD4$^+$ T cells (light blue) and stimulated CD8$^+$ T cells (dark green). The genotype groups of the SNP are shown on the x-axis. The y-axis represents the normalised and log_2 transformed expression data. (b.) The LD calculation between rs658524 and rs12225345 extracted from LDlink using 1000 Genomes European population haplotypes. Shown are the Pearson correlation coefficient (r^2) and the standardized linkage disequilibrium coefficient (D'), along with the correlation pattern between the alleles of the variants.

6.4 *CTSW* eQTLs are located at transcription factor binding sites

Given the correlation we observed with the SNP genotype and *CTSW* expression levels, we asked if rs658524 and rs12225345 overlap with predicted transcription factor binding sites. Firstly, we consulted HaploRegv4 (Ward & Kellis, 2012) to retrieve the predicted or validated TF sites in the SNP location from published studies. Secondly, we utilised Promo3 - an online bioinformatics tool, to predict TF binding motifs in the sequence of the reference and alternative alleles for the SNPs (Farré et al., 2003; Messeguer et al., 2002). Thirdly, we identified predicted TF motifs in a given genomic locus from JASPAR dB CORE 2020 (Fornes et al., 2020) track on UCSC genome browser. We compiled the results from these three approaches to refine the list of possible TFs for further analysis. In particular, we selected TFs that are predicted to show allele-specific binding (**Table 13**).

Table 13: TFs that are predicted to show allele-specific binding to rs658524 and rs12225345

rs658524	rs12225345
Egr-1	IRF1
IRF1, IRF2	NFATC1, NFATC2, NFATC3
KLF9, KLF10	STAT1, STAT4
ZNF740	RBPJK

We analysed the expression profiles of these TFs in immune cells using the DICE database (https://dice-database.org/) and the Human Blood Atlas (https://www.proteinatlas.org/humanproteome/blood), as an additional measure to select relevant TF that are expressed in the same cell type where CTSW is expressed (**Figure 73**).

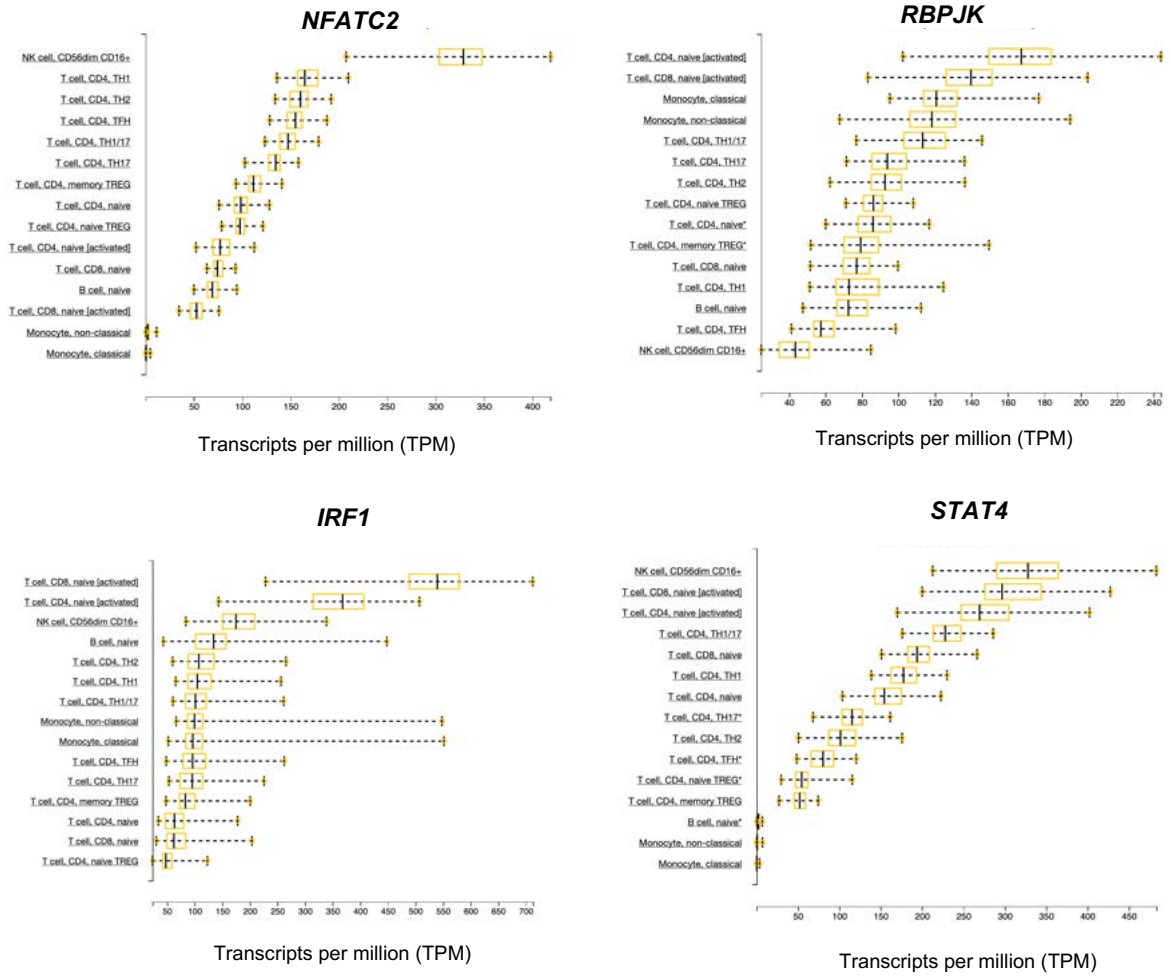


Figure 73: Gene expression profiles of TFs in immune cell subsets. The TF gene expression from RNA-seq data from the DICE database is shown as boxplots. The levels of expression are depicted as transcripts per million (TPM).

We observed that the TFs such as *NFATC2*, *RBPJK, IRF1* and *STAT4* were expressed at higher levels in either resting and activated $CD8^+$ and $CD4^+$ T cells, suggesting they may be involved in *CTSW* regulation (see **Figure 73**). *NFATC2* also showed high level of expression in NK cells. Therefore, SNPs that disrupt the binding of these TF could play a role in regulating the transcription of *CTSW* in these cell types. To assess this hypothesis further we tested whether any of these TF bind to rs658524 and rs12225345 in an allele-preferential fashion.

6.4.1 NFATC2 shows allele-preferential binding at *CTSW* genomic sequences in vitro

We tested the binding specificity of the predicted TFs using high-affinity DNA capture assay. The biotinylated oligos with either the reference or the alternative alleles of SNPs rs658524 (A/G) and rs12225345 (A/G) were pre-incubated with the cell extract from resting and stimulated $CD8^+$ T cells. The bound proteins were then tested for the presence of various TFs using Western blot analysis (**Figure 74**).

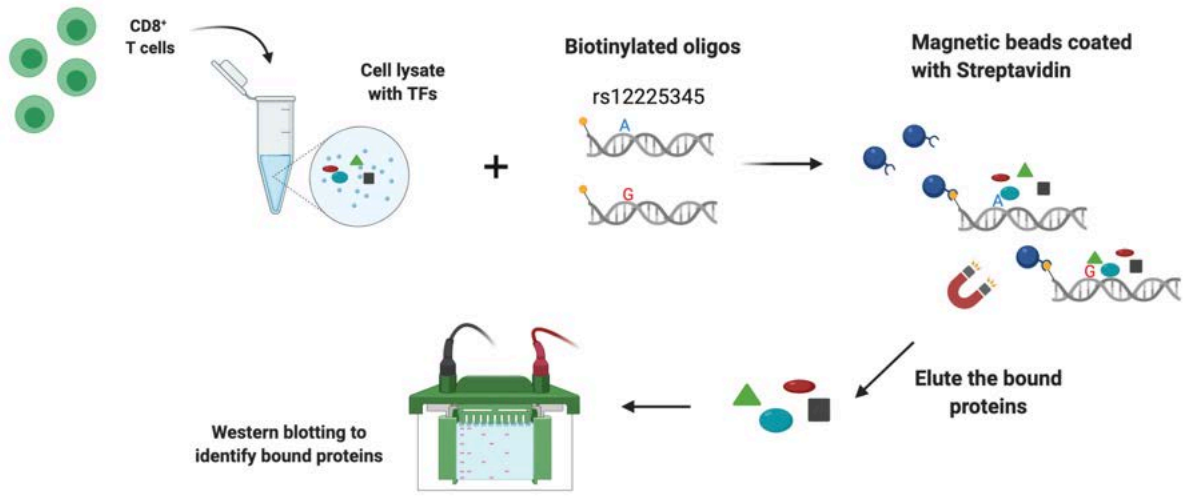


Figure 74: Work flow for high-affinity DNA capture assay. Lysates from unstimulated or PMA/ionomycin stimulated $CD8^+$ T cells were incubated with biotinylated oligos, followed by incubation with streptavidin beads. Proteins bounds to the DNA sequences are eluted, separated by SDS-PAGE and analysed by Western blotting. Membranes are probed with antibodies directed towards the TFs of interest.

From the Western blot analysis, we observed allele-preferential binding of NFATC2 (nuclear factor of activated T-cells, cytoplasmic 2) to rs12225345 in resting $CD8^+$ T cells (**Figure 75a**). The predicted consensus sequence for NFATC2 from the JASPAR CORE 2020 database (http://jaspar.genereg.net/), together with the genomic sequence upstream of the *CTSW* promoter encompassing rs12225345, is shown in the bottom panel of **Figure 75a**. The binding of NFATC2 is reduced in the presence of the alternative allele of rs12225345 (G), as shown in **Figure 75a**. Binding of NFATC1 to the same sequence was much weaker (**Figure 75b**).

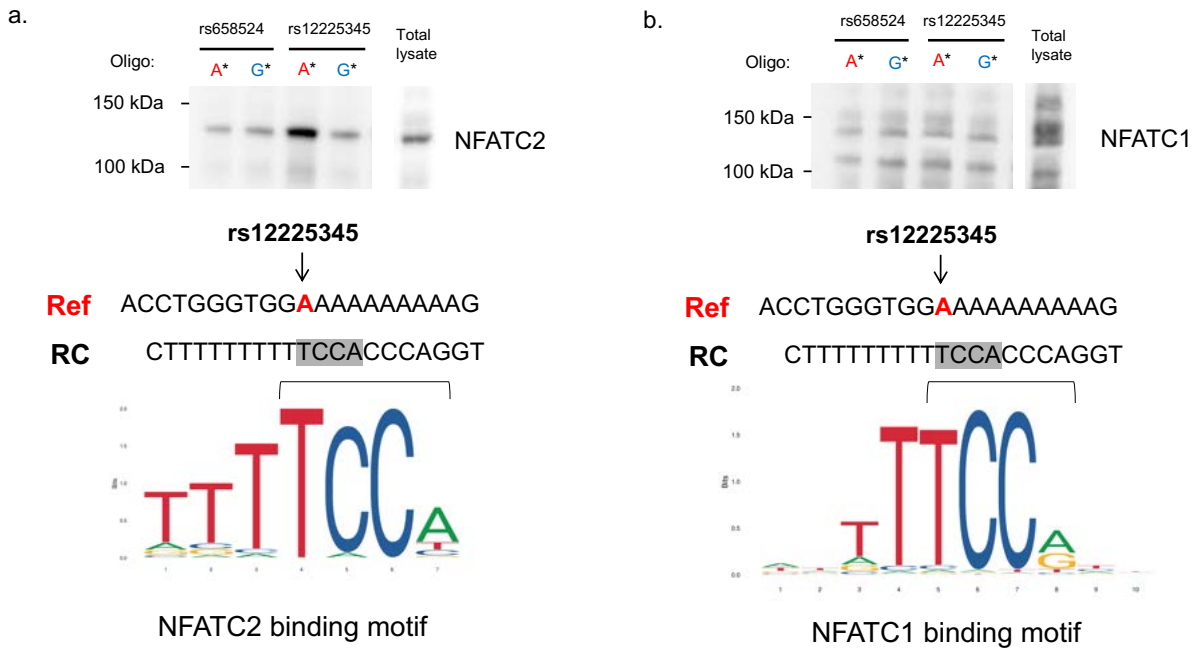


Figure 75: Allele-specific binding of NFATC2 in unstimulated CD8$^+$ T cells. DNA affinity capture assay was performed in lysates from CD8+ T cells with biotinylated oligos containing the sequences of the CTSW promoter encompassing SNPs rs658524 and rs12225345. Western blots from the DNA-affinity capture assay was probed with (a.) anti-NFATC2 antibody and (b.) anti-NFATC2 antibody are shown. Binding motif for NFATC2 and NFATC1 are shown as per the JASPAR database, and the sequence of the CTSW promoter region indicating the alternative allele of rs12225345. "RC" denotes the reverse complement sequence. The matching sequence with the predicted motif is highlighted in grey, and is indicated in the sequence logo with square brackets.

We then performed a competitive DNA affinity capture assay to confirm the allele-specificity of NFATC2 binding (lanes 5 to 10 in **Figure 76**). Binding to the biotinylated oligo of the reference allele (A*) was competed by decreasing amounts (5x, 2x and 1x) of the non-biotinylated oligos of alleles A or G. In the competition assay, we observed that the A allele (lanes 5, 7 & 9 of **Figure 76**) was more efficient in competing binding compared to the non-biotinylated G allele (lanes 6, 8 & 10 in **Figure 76**).

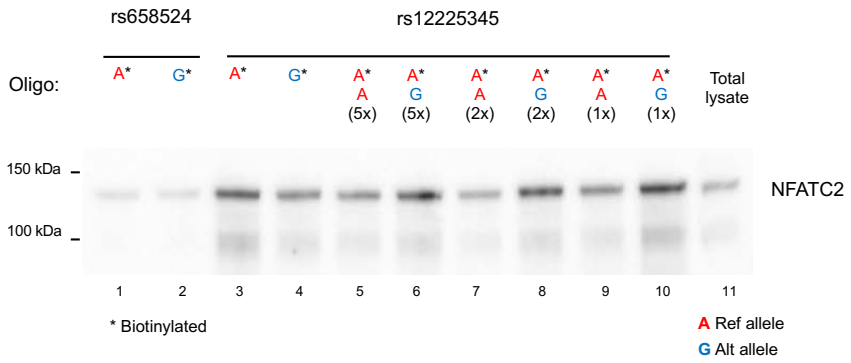


Figure 76: DNA affinity capture assay on stimulated CD8⁺ T cells. (a.) DNA affinity capture assay was performed in CD8+ T cells lysates with biotinylated oligos of the sequences from the CTSW promoter containing the SNPs rs658524 or rs12225345. Western blot from the DNA-affinity capture assay was probed with anti-NFATC2 antibody. Lanes 1 and 2 are the reference and alternative alleles of the SNP rs658524, and lanes 3 and 4 correspond to the reference and alternative alleles for rs12225345. Lanes 5 -10 show the binding of NFATC2 from competition assay between biotinylated (marked with *) and non-biotinylated oligos. Lane 11 is the total lysate from the stimulated CD8+ T cells (PMA (100ng/ml) + Ionomycin (1µg/mL) + IFNγ (10ng/mL) for 30min).

These results support NFATC2 preferential binding to the reference allele (A) of rs12225345. Presence of the alternative allele (G) may result in weaker binding of NFATC2 affecting gene transcription, which is consistent with lower levels of *CTSW* expression in CD8⁺ T cells carrying this allele (see **Figure 75a**).

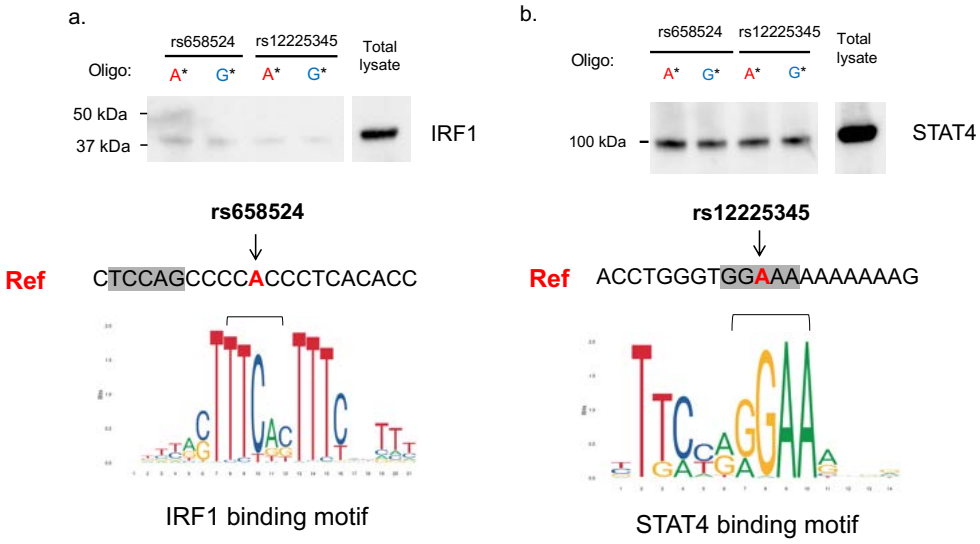


Figure 77: Transcription factor binding pattern for predicted sites in unstimulated CD8⁺ T cells. DNA affinity capture assay was performed in lysates from CD8+ T cells with biotinylated oligos containing the sequences of the CTSW promoter encompassing SNPs rs658524 and rs12225345. Western blots from the DNA-affinity capture assay was probed with (a.) anti-IRF1 antibody and (b.) anti-STAT4 antibody are shown. Binding motif for IRF1 and STAT4 are shown as per the JASPAR database, and the sequence of the *CTSW* promoter region indicating the alternative allele of rs658524, the matching sequence with the predicted motif is highlighted in grey, and is indicated in the sequence logo with square brackets.

On the other hand, we did not observe allele-specificity for other TFs predicted for rs658524, such as IRF1 (**Figure 77a**) and for rs12225345, STAT4 (**Figure 77b**) in unstimulated CD8$^+$ T cells. Binding of IRF1 was very weak in the presence of either alleles of rs658524 (**Figure 77a**). STAT4 bound everywhere irrespective of the allele present for rs12225345 (**Figure 77b**), suggesting that the SNPs is not adequate to alter the binding potential of the TF. The lower panel of **Figure 77b** shows the consensus sequence of the binding motif for STAT4 (JASPAR database).

Interestingly, we noted that the transcriptional regulator of the Notch signalling pathway, RBPJK (Recombination Signal Binding Protein For Immunoglobulin Kappa J Region), binds preferentially to the alternative allele (G) of rs12225345 (**Figure 78a**). The predicted consensus sequence for RBPJK is shown in **Figure 78b**, as presented in JASPAR CORE 2020 database (http://jaspar.genereg.net/). The alternative allele (G) of rs12225345 was correlated with low expression levels of *CTSW* (See **Figure 72a**). The selective binding of RBPJK to the alternative allele may suggest a possible repressive mechanism through Notch signalling pathway in low levels of *CTSW* expression in activated CD8$^+$ T cells.

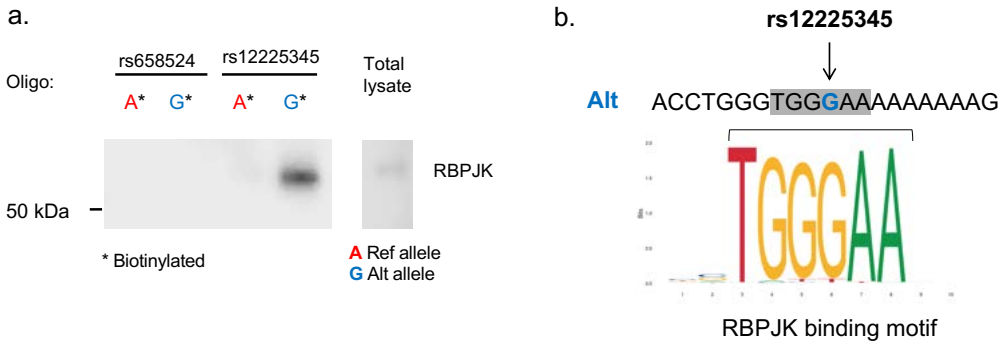


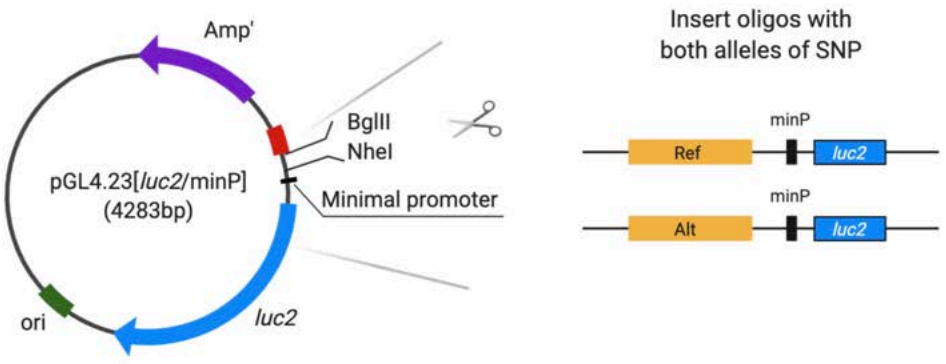
Figure 78: Allele-specific binding of RBPJK in stimulated CD8+ T cells. (a.) DNA affinity capture assay was performed in lysates from CD8+ T cells with biotinylated oligos containing the sequences of the CTSW promoter encompassing SNPs rs658524 and rs12225345. Western blot from the DNA-affinity capture assay was probed with anti-RBPJK antibody. (b.) Binding motif for RBPJK as shown in the JASPAR database, and the sequence of the CTSW promoter region indicating the alternative allele of rs12225345.

The allele-preferential binding of TF supports the regulatory role of the variant rs12225345 located within the *CTSW* promoter region. Based on the selective binding affinity of NFATC2 in the presence of the reference allele, we hypothesise that NFATC2 is important in the regulation of *CTSW* expression in CD8$^+$ T cells. We explored this hypothesis further using reporter gene assays to validate the role of these putative SNPs in NFATC2-mediated activation of *CTSW* transcriptional activity.

6.4.2 *Cis*-eQTLs regulate *CTSW* expression through NFATC2-mediated transcription activation

To test the role NFATC2 binding to rs12225345 in transcription regulation *in vitro*, we performed a dual luciferase reporter gene assay in the human kidney epithelial cell line HEK293T. We constructed a luciferase reporter plasmid by cloning either the reference or the alternate allele into a luciferase expressing plasmid with a minimal promoter (pGL4.23[*luc2*/minP]) as illustrated in **Figure 79a**. The constructs with reference and alternative alleles were then transfected into HEK293T cells, with or without the vector expressing the transcription factor NFATC2, using the Lipofectamine™ 2000 protocol (**Figure 79b**). The luciferase ratio of the experimental reporter (Firefly) to the control reporter (Renilla) was then used to measure the transcriptional activity in the presence of each allele.

a.



b.

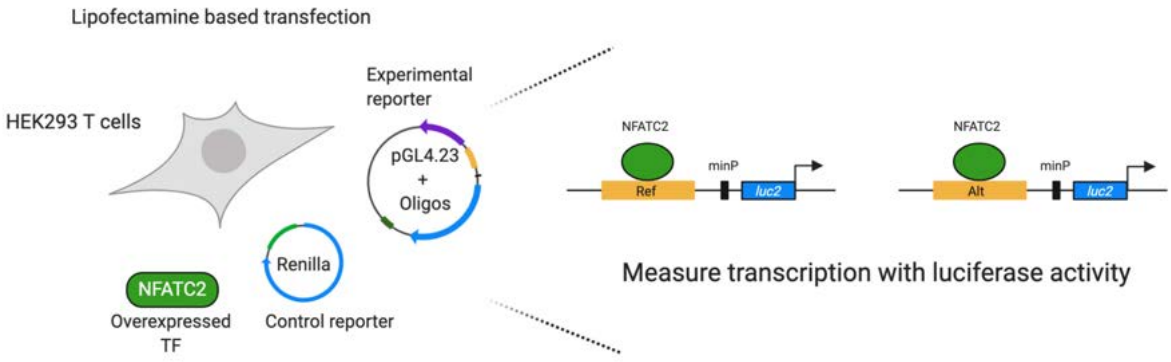


Figure 79: Dual reporter gene assay to test the eQTL-dependent transcriptional activity. (a.) pGL4.23 vector with the minimal promoter used to clone the CTSW promoter construct with the SNP rs12225345 (A/G). (b.) The firefly luciferase reporter containing the SNP alleles, the control reporter with Renilla luciferase and the NFATC2 vector were co-transfected into HEK293 T cells using Lipofectamine 2000. The luciferase activity measured was used to assess transcription levels.

Presence of *CTSW* promoter sequences increased luciferase transcription over the background (**Figure 80** column 1). Overexpression of NFATC2 showed a further increase in transcriptional activity, significantly more pronounced in the presence of the "A" allele of rs12225345, compared to the allele "G" as shown in **Figure 80**. This difference in transcriptional activity is consistent with the findings from our eQTL study, showing increased *CTSW* expression in the presence of the "A" allele (See **Figure 72a**), We also observed that in the absence of NFATC2, the transcriptional activity did not show allele-dependency compared to the vector without the promoter constructs, as indicated by the dark grey and light grey bars in **Figure 80**. The reporter gene assay results support our hypothesis that NFATC2 regulates the transcription activity in the *CTSW* promoter region, in an allele-dependent manner.

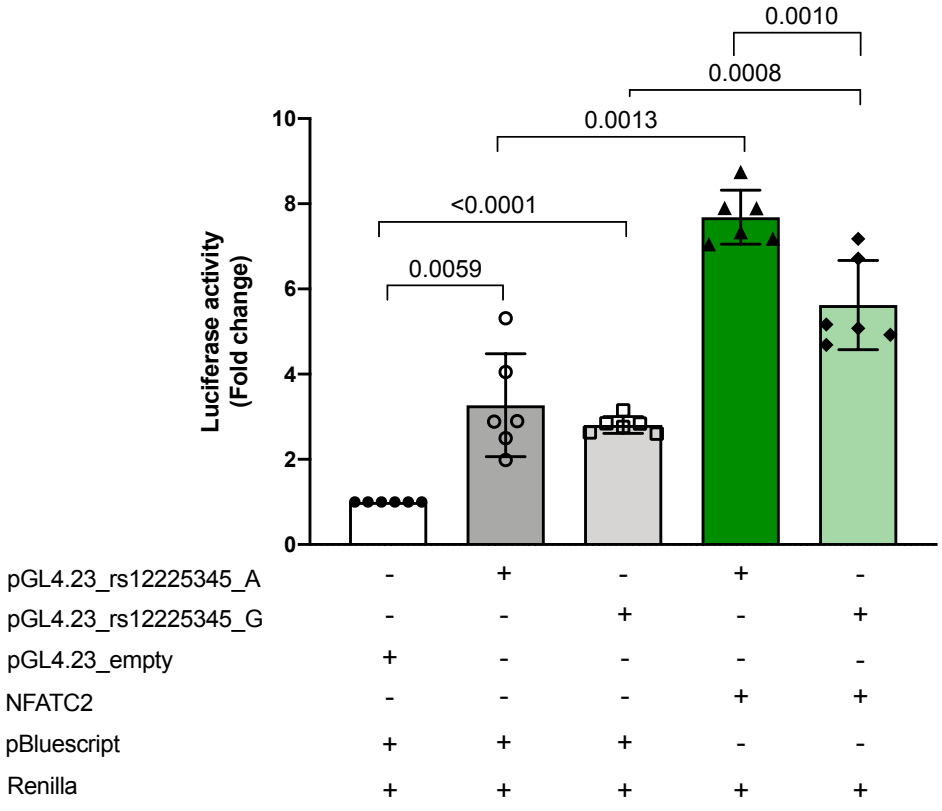


Figure 80: NFATC2 increases transcription in an allele-dependent manner at the CTSW promoter. Luciferase reporter gene assays were carried out in the HEK293 T cell line. Cells were transfected with the CTSW promoter construct containing the SNP rs12225345 (A/G), with or without NFATC2, as indicated in the panel below the graph. Firefly luciferase activity was measured and normalised to the control reporter Renilla luciferase activity in each sample, and is represented as fold changes over the empty vector alone. The average and standard deviation of 6 independent experiments are shown. T-test was performed between different conditions as indicated (p-values).

6.4.3 Cathepsin W is sensitive to Endo H and PNGase F

Cell-type-specific *CTSW* expression in CTLs and NK cells led us to investigate the cellular compartments where cathepsin W is localized to understand its functional relevance in cytotoxic cells. Post-translational modification of cathepsin W takes place through the protein trafficking pathway. Therefore, cathepsin W (CatW) may be present in different vesicular compartments at various stages of its maturation. To improve our understanding of the localisation of cathepsin W, we performed an enzymatic treatment using total lysates from $CD8^+$ T cells with the glycosidases Endo H and PNGase F.

Protein trafficking can be monitored by studying the glycosylation of proteins. Glycosidases are enzymes that hydrolyse glycosidic structures in proteins. Endoglycosidase H (Endo H) cleaves high mannose structures of a glycoprotein in the endoplasmic reticulum (ER). When protein structures are resistant to Endo H it can be inferred that the protein has progressed to the other vesicular compartments and has acquired complex glycan structures as part of post-translational modification. Peptide-N-Glycosidase F (PNGase F) is another enzyme that can cleave both high mannose structures and the hybrid complex sugar molecules (asparagine-linked oligosaccharides) attached to the protein in the Golgi compartment (**Figure 81**).

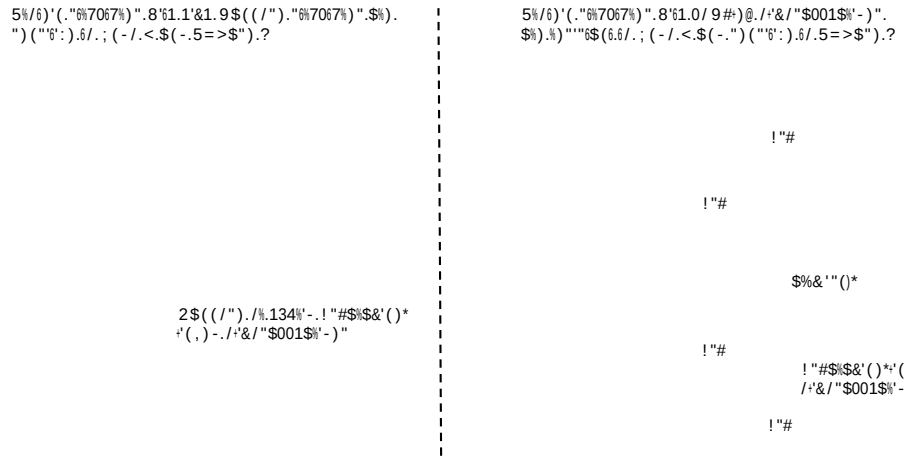


Figure 81: Enzymatic action of glycosidases Endo H and PNGase F (Structures adapted from QA-Bio)

Glycoproteins can also have high mannose structures with added phosphate residues in the Golgi (mannose-6-phosphates, M6P) that are sensitive to Endo H (**Figure 82**). Therefore, sensitivity and resistance towards Endo H and PNGase F can help us determine the stage and location of the protein processing.

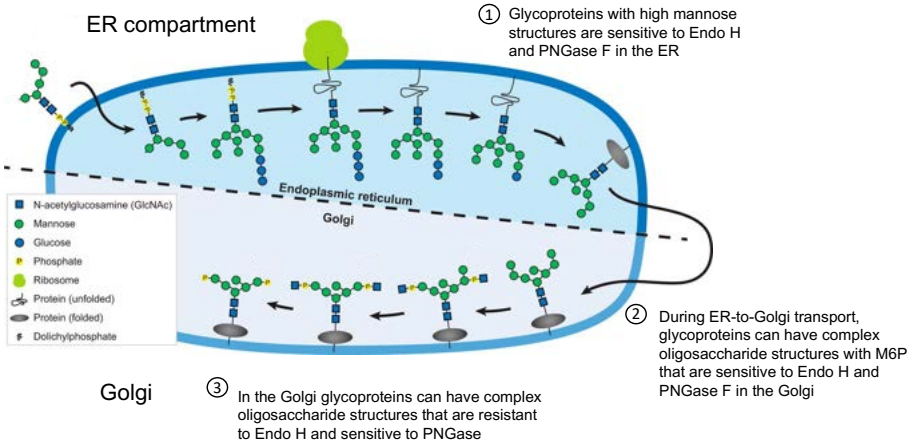


Figure 82: Protein-processing and trafficking through mannose-6-phosphate pathway in the ER and Golgi. (1.) High mannose and asparagine-linked structures of the glycoproteins in the ER make them sensitive to Endo H and PNGase F. (2.) When glycoproteins move from ER to Golgi, they acquire M6P conformation, and retain their sensitivity to Endo H. (3.) Upon reaching the Golgi, most glycoproteins add on complex oligosaccharide structures that are resistant to Endo H, but still remain PNGase F sensitive. (Adapted from Caval et al., 2019)

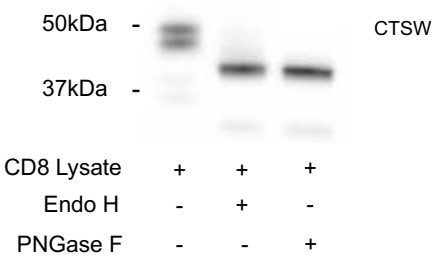


Figure 83: Deglycosylation of cathepsin W. Total lysates from CD8+ T cells were examined in the absence or presence of Endo H and PNGase F. Decrease in the CTSW band from 50 to 37 kDa was noted.

We observed that CatW is sensitive to both the enzymes (**Figure 83**). The high mannose-type structure from CatW is cleaved by Endo H and PNGase F, compared to the untreated total cell lysate. CatW shows two bands in untreated total lysate close to 50 kDa, which may represent two different glycosylation product, since glycosidase treatment results in a single band with a lower molecular weight (37 kDa). A similar observation regarding murine cathepsin W deglycosylation was reported in lysates from IL2/ConA-activated blasts by Ondr & Pham (2004). The sensitivity of CatW to glycosidases may suggest that it may still be in the early stages of protein processing within the ER compartment.

Chapter 7

DISCUSSION & CONCLUSION: PART I

In the first part of this project, we have addressed the molecular mechanisms of the disease pathogenesis by studying immune cell function in AxSpA. We linked genetic variants in susceptibility loci of CIDs to immune cell function using gene expression analysis and eQTL mapping. We also demonstrated the mechanism by which variants in the susceptibility loci regulate the nearest genes in the locus.

7.1 Quantifying expression of GWAS genes using NanoString

The NanoString nCounter® technology proved to be a robust and reproducible method to answer our specific question about the expression pattern of genes in the CID susceptibility loci. Methods to study the whole transcriptome such as RNA-seq have the advantages of identification of rare transcripts, splice variants and non-coding RNAs to characterise complex diseases. However, the experimental procedures involved (library preparation, conversion of mRNA to cDNA by reverse transcription and amplification) can introduce variability in the data. NanoString technologies focuses on targeted transcriptomics by directly measuring a transcript of interest, and it's a highly quantitative, reproducible technique, which can be applied to small amount of starting material. We utilised this approach for the specific task of assessing the expression profiles of GWAS genes in blood and immune cell populations.

7.2 GWAS genes are involved in the interplay between innate and adaptive immunity

The whole blood cultures using the TruCulture system were used to challenge the patient immune system in an *ex vivo* setting with stimulation conditions that represent the *in vivo* inflammatory conditions. The TruCulture system is a highly standardised *ex vivo* assay that can be used to capture the physiological cellular interactions without the variability that arise from experimental processes (blood draw, cell isolation, cell culture conditions, etc.).

As the numerous genetic loci associated with CIDs implicate dysregulation of the immune system, we looked at the expression pattern in whole blood cultures of genes in the susceptibility loci. Expression analysis of the candidate genes showed that around 80% of the GWAS genes can be detected in blood cells, supporting their role in the immune system function. The whole blood gene expression data from patients with AxSpA revealed that disease associated genes are activated under innate or adaptive stimulation. The stimulation-dependent expression pattern of GWAS gene clusters shows that they participate in both innate and adaptive immune system. Therefore, we suggest that these genes have a

functional role in cell populations of the innate and adaptive immunity. To test this hypothesis, we analysed the expression of GWAS genes in T cells and monocytes derived from AxSpA patients. Our analysis revealed that genes associated to CIDs are expressed in both innate and adaptive immune cell subsets when subjected to physiologically appropriate stimulation conditions. Gene expression pattern shows that most GWAS genes are expressed in monocytes after a microbial stimulus and in T cells activated through TCR signalling in AxSpA patients.

During an immune response, cells from innate immune compartment serve as the first line of defence against infection. Innate immune cells when challenged can express genes that aid in fighting against the infection directly through anti-microbial peptides or indirectly by attracting and activating the adaptive immune system through secretion of cytokines and chemokines. The fact that we observe nearly 70% of the genes in the susceptibility loci in activated innate and adaptive immune cells, reveals that the GWAS genes can be activated by stimulation conditions that occur during an inflammatory response. The activation of these genes may imply that they have a role in the coordinated immune response by the innate and adaptive cells.

7.3 Expression of disease-associated genes and upon stimulation in distinct immune cell populations

We analysed whether disease-associated genes are also cell-type-specific, i.e., to assess whether genes associated to a disease are expressed more in a specific cell type, and whether that may be of biological relevance. We did not observe a significant difference between the number of disease-associated genes detected in T cells and monocytes for AS, Pso and IBD disease modules in the upset analysis. We observed a significant difference in the number of genes detected between T cells and monocytes in MS gene module, with most genes detected in T cells. The difference in the significance among the gene modules could be due to the fact that the number of genes per disease module was not evenly distributed and this is a possible limitation in this analysis (**Table 7**).

GWAS genes associated with AS, Pso and IBD could be activated in the presence of an innate stimulus in monocytes and in T-cells activated via TCR-signalling. This suggests that susceptibility genes may be involved in the regulation of immune responses mediated by innate and adaptive cells in CIDs. The context-specific gene expression in autoimmune diseases has been discussed by the Fairfax lab for *IL7R* association in AS (Al-Mossawi et al.,

2019). This work elucidated the possible function of monocytes in AS, as it showed activation of *IL7R* (AS GWAS gene) in LPS-activated monocytes. Prior to the IL7R study, another publication from Bowness group had reported the increased expression of *GPR65*, a GWAS gene for AS, in GM-CSF$^+$ IL17A$^+$ T cell subsets (CD4$^+$, CD8$^+$, γδT cells) and NK cell in AS patients (Al-Mossawi et al., 2017). These studies and the expression pattern we observed for GWAS genes in innate and adaptive cells suggest their potential involvement in the immune events in AxSpA.

T1D gene module also showed a trend that indicated most disease-associated genes were detected in CD4$^+$ T cells and monocytes (**Figure 45**). This is interesting as studies have shown that immune cell types including autoreactive CD4$^+$ T cells and APCs have effector roles in pancreatic β-cell destruction- a main feature of T1D (Coppieters et al., 2012; Sarikonda et al., 2014). We also noted a similar pattern in MS module, where the number of GWAS genes expressed in T cells is significantly higher than in monocytes (**Figure 49**). The expression levels showed that GWAS genes from MS can be involved in both T cells and monocytes. It is important to note that the disease-associated gene signatures were measured in patients with AxSpA- an inflammatory disease with overlapping immune features with many CIDs. However, the functional relevance of GWAS genes associated to T1D, MS and other CIDs in immune cells can be assessed by using appropriate immune cell types or tissues derived from representative disease cohorts. This analysis however provides information about in which cell type and under which activation condition the disease-associated genes can be measured.

GWAS genes and variants may present a range of functional consequences in disease pathogenesis. A variant that affects the function of a gene which may be vital in the development of one or several diseases can have the opposite effect in another disease. An example is the role of *TYK2* variant rs34536443 in autoimmunity and infectious diseases. This SNP has been reported to impair cytokine signalling in immune cells (Li et al., 2013), and present protection against Pso, RA, SLE and T1D, while it confers disease risk for AS, CD and UC (Dendrou et al., 2016). Furthermore, *TYK2* knock-out can increase the susceptibility to mycobacterial infections as previously shown by Kreins et al. (2015). These results have demonstrated that disease-associations can have multiple and/or disparate effects. In such situations, the analysis of individuals with inborn errors of immunity can provide important information about the function of GWAS genes. This approach has been particularly useful for the analysis of *TYK2* variants as reported by (Boisson-Dupuis et al., 2018). Analysis of immune cells from *IL23R*- or *IL12RB2*-deficient individuals in this study has revealed the non-

redundant roles of these genes in vivo and provided detailed information about the signalling downstream of these receptors.

The context-dependent expression pattern of GWAS genes showed that these genes potentially have a function in T cells and/or monocytes which may be related to the disease pathogenesis. For example, GWAS gene *TBX21*- which encodes T-bet, a T-cell specific transcription factor, showed a significantly higher level in $CD4^+$ and $CD8^+$ T cells. However, we also noted that *ERAP1*- an aminopeptidase that trims HLA class I-binding precursors to enable antigen presentation, is expressed in both monocytes and T cells. The trimmed peptides are presented by MHC class I molecules on the cell surface of APCs that are then recognised by surface receptors of $CD8^+$ T cells and NK cells. Interestingly, aminopeptidases (*ERAP1* & *ERAP2*) have been shown to regulate cytotoxic T cell (Zervoudi et al., 2013) and NK cell responses (Cifaldi et al., 2015).

7.4 The GWAS genes may be involved in the effector mechanism of immune cells in patients with inflammatory diseases

To assess whether chronic inflammation in patients affects the expression of GWAS genes, we analysed the gene expression in immune cells isolated from healthy donors. We observed the same stimulation-dependent gene expression pattern, with majority of the genes upregulated in the innate immune cells (monocytes and NK cells). However, it is important to note that the overall number of GWAS genes detected and the range of gene expression was much lower in healthy donors. This might suggest that the stimulation conditions were not sufficient to activate the susceptibility genes. It may also suggest that the baseline inflammatory environment required to activate these genes may be more prominent in patients than in healthy donors. However, the gene expression analysis was done in patient and healthy donor data separately due to the differences in the RNA amounts in the two groups.

The association of IL-23/IL-17 pathway in CIDs, like AS, IBD and Pso was identified through GWAS. These findings indicated a role of Th17-driven immunity in the pathogenetic mechanisms of these diseases. We observed that the genes in the IL-23/IL-17 pathway are not only expressed in adaptive T-cells, but also in innate cells - both in controls and AxSpA patients, consistent with the literature in CID pathophysiology (Schön & Erpenbeck, 2018). This supports the previous observation that there is an interplay between the innate and adaptive immune cells to mediate the immune responses.

With the observed interplay between innate and adaptive immune cells, we investigated whether GWAS genes were also expressed in innate-like cells. γδT cells and MAIT cells are "inbetweener" populations that have the properties of adaptive cells, but mediate innate-like responses. The expression pattern of GWAS genes in immune cells from healthy donors showed that these genes can be activated with the appropriate immune stimulus, that is concordant with the findings from the patient data. We observed a group of genes that are significantly upregulated in activated monocytes compared to activated NK cells and T cell subsets. However, this observation has an added caveat that the effect of LPS stimulation may be stronger compared to other stimuli. The analysis on MAIT cells were performed separately due to the technical issue of low cell numbers and low RNA quantity. GWAS genes were detected more in resting MAIT cells than the activated cells from healthy donors. MAIT cells being innate-like might already be in a challenged state, which could explain why the susceptibility genes are activated in them without additional stimulation.

7.5 Variants in the susceptibility loci affect the expression of *FADS2* and *CTSW*

To analyse whether GWAS genes are affected by genetic variants that are in the susceptibility loci we performed an eQTL analysis. Cell-type-specific eQTLs are like signposts that indicate a region of interest that may be linked to a biological mechanism involved in the disease pathogenesis. eQTLs are useful to identify candidate genes/variants from a GWAS locus that are most likely to have a functional role. Cell-type-specific eQTL-Genes (eGenes) also give a clue about possible biological pathways that are perturbed in patients.

eQTL studies in the literature are done using gene expression data from cells or tissues derived from healthy donors such as the tissue-specific eQTLs in the GTEx project (Ardlie et al., 2015) and recently, the eQTL analysis from immune cell-types conducted by the DICE study (Schmiedel et al., 2018). Such studies have shown genotype-gene expression correlations in representative cell types derived from healthy donors in activated cell states. Due to this context-dependent nature of eQTLs, they are often detected in cells isolated from healthy donors under stimulation. These stimulation conditions may resemble the pathological/physiological settings observed in patients, and thus eQTLs detected in this manner may be more informative in terms of disease biology. However, these signatures might not resemble the actual genotype-gene expression associations in patients. eQTL studies conducted in patient cohorts can therefore give a better insight into the effects of variants in expression of genes associated to the disease. We performed eQTL mapping from

gene expression data from T cells isolated from AxSpA patients, under both resting and activated states. We detected more eQTLs from resting T cells from AxSpA patients than the activated T cells. This observation could be attributed to the fact that the immune cells from CID patients may already be in an activated state, due to the higher baseline in patients compared to healthy controls. Another possibility for this observation could be the temporal effects on stimulus-dependent eQTLs. Fairfax et al. reported time-dependent and stimulation-dependent gain or loss of eQTL effects in LPS-stimulated monocytes from healthy donors (Fairfax et al., 2014). One example is the effect of rs2275888 on *IFNB1* gene, where the eQTL effect was not observed for the SNP in *IFNB1* expression in resting monocytes. Interestingly, after 2h of LPS stimulation, *IFNB1* showed differences in expression levels by SNP genotype, which was lost after 24h of LPS stimulation (Fairfax et al., 2014). In our study, the patient cells may already be in an activated state due to the chronic inflammation. When we top-up cells, that are already in a challenged state, with a stimulus for a longer period (>>2h), we may lose the eQTL effect over time, and that could be why we did not detect many eQTLs in stimulated cells.

The eQTL analysis revealed many SNPs in a region on chromosome 11 acting in *cis*-that could be of interest in T cell function for AxSpA. Our eQTL analysis yielded several significant SNPs in close proximity to one another affecting the expression of genes *FADS2* and *CTSW* located on chromosome 11. We did not observe significant *trans*-eQTL associations, possibly given the limited sample size of our cohort.

FADS2 is one of the fatty acid desaturases in a highly polymorphic region called the "FADS" cluster. Several SNPs in this locus have been previously associated with blood-related traits such as high sensitivity C-reactive protein (hs-CRP) (Roke et al., 2013), levels of metabolites from the ω-3 and ω-6 fatty acid pathways in the serum and immune cell counts (UKBB v2, 2018), and cardiovascular diseases. The SNPs in this region are also strongly linked to CIDs such as IBD, Pso & AS (Ellinghaus et al., 2016). The pleiotropic effect of SNPs in the FADS cluster make it challenging to define concrete associations from variant-to-gene function-to-disease. However, a recent work from Franke group showed an integrative approach studying the regulatory role of SNPs in genomic loci like the FADS cluster for linking GWAS variants to function (Van Der Wijst et al., 2018). In this approach, the authors elucidated regulatory gene networks that control the expression of *FADS2* by combining information about allele-specific TF binding and enhancer and promoter elements, with an identified *cis*-eQTL rs968567 from GTEx data, which is a GWAS variant linked to RA susceptibility originally reported by Stahl et al. (2010). Van der Wijst et al. discussed the allele-

specificity of sterol binding transcription factor SREBF2 using the ENCODE ChIP-seq data analysis in a previous paper from the same group (Zhernakova et al., 2017).

To understand the function of *FADS2*, we explored its expression pattern in immune cells from the DICE database (Schmiedel et al., 2018). *FADS2* expression is cell-type-specific, with highest levels seen in NK cells, followed by classical and non-classical monocytes, suggesting a potential functional role in these cell types (**Figure 64**). However, we observed that in the AxSpA patient data, the expression of *FADS2* was similar in both activated monocytes and T cells. Moreover, the SNP from our eQTL analysis, rs968567, was already shown to affect the *FADS2* promoter activity by allele-dependent binding of the transcription factor ELK1 (a member of the ETS domain transcription factor family), using reporter assays in HepG2 and HeLa cell lines (Lattka et al., 2010). Therefore we focused our analysis on the SNPs that regulate the expression of *CTSW,* since there's no data about the regulation of *CTSW* expression, and because this gene is expressed specifically in immune cells, suggesting a role for cathepsin W in immune function.

Due to the limitations in the sample size and statistical power in our eQTL analysis, we did not detect long-range effects that were reported through association studies and meta-analyses. We did not detect previously reported significant and robust associations linked to AS, e.g., the non-MHC locus *ERAP1* (Burton et al., 2007; Evans et al., 2011). SNPs reported in *ERAP1* have shown strong association to AS susceptibility and were replicated by several large independent cohorts with substantial statistical power, as reviewed by Costantino, Breban & Garchon (2018). A study performed in 2015 showed a strong correlation between expression of *ERAP1* and polymorphisms in *ERAP1* in monocyte-derived dendritic cells (Costantino et al., 2015). *ERAP1* is an aminopeptidase implicated in AS susceptibility through its peptide trimming function in the ER and antigen presentation by MHC class I molecules. As our eQTL analysis was not performed in monocytes or APCs, we may have missed replicating this association reported by Costantino et al. (2015). Therefore it would be interesting to extend the eQTL analysis to monocytes and other APCs, preferably in a larger cohort to improve the resolution of the signals detected.

7.6 Functional annotation of GWAS SNPs that are eQTLs is important to make variant-to-gene function associations

The *CTSW* eQTLs we identified were localised to the region on the chromosome 11 that has been previously reported in autoimmune disease susceptibility. Therefore, we wanted to see if our eQTLs were also reported GWAS variants, that may be functional in regulating *CTSW* expression.

Integrating eQTL data from relevant tissues and cell types along with the GWAS variant information can help in selecting SNPs that are more likely to have a functional role in the phenotype. Using this approach we identified the three *CTSW* eQTLs - rs658524, rs568617, and rs12225345 - as GWAS variants associated with AS, IBD and Pso susceptibility from the 15 significant eQTLs. The lead variant and *CTSW* eQTL rs568617 was reported in the meta-analysis study conducted by Ellinghaus et al. (2016) was linked to IBD susceptibility. rs658524 and rs122253445 were found in the UKBB PheWAS data, where rs658524 was a risk variant for AS, and rs12225345 was associated with Pso. This concurs with the premise reported by Nicolae et al. that nearly half of the SNPs associated with complex traits (autoimmune diseases, cancers and neuropsychiatric disorders) are eQTLs in at least one cell type (Nicolae et al., 2010). GWAS-associated variants can therefore affect the phenotype by altering the amount of mRNA synthesised, which may have a direct or indirect effect on the underlying biology associated with the disease. The regional LD in the locus highlighted in our eQTL analysis showed that all the eQTLs are in strong LD with each other and that they are co-inherited. Due to the strong LD structure it is complicated to determine the functional consequence without empirical validation.

Interestingly, the “G” allele of rs658524 was correlated with high *CTSW* expression in T cells from AxSpA patients. The UKBB PheWAS analysis reported the “G” allele of rs658524 as being associated with AS risk, and with a protective effect in Pso, CD & UC. The “G” allele of rs658524 was correlated with the “A” allele of rs12225345, which also corresponds to the genotype group with high *CTSW* expression. These findings further indicated a possible functional importance for these SNPs. Pleiotropic effect of SNPs with discordant associations have been previously reported in the literature. For instance, a follow-up study on IBD GWAS associated regions in an independent case-control group identified variants associated with IBD that show discordant effects (Rivas et al., 2011). The authors reported a rare missense variant rs2476601 in *PTPN22*, associated with risk of T1D, RA and Vitiligo, and protective against CD (Rivas et al., 2011). Furthermore, discordant effects of SNPs were also identified for the variants in *TYK2* that showed increased susceptibility to mycobacterial infections and

protection against autoimmune diseases, as discussed by Dendrou et al. (2016). Conceptually, the discordant effect of a SNP in two diseases may be due to the fact the associated SNP might be perturbing the function of a specific cell type and a related biological pathway that can lead to one disease. The SNP perturbation can also have a concomitant effect on another cell type or pathway that can lead to a protective effect against another disease. Another possibility is that, the disruption in the biological mechanism from the SNP in one disease may be biologically linked to having less of another disease.

To prioritize SNPs that are most likely to have a functional role in *CTSW* regulation, we analysed the overlap of the eQTL variants that we have identified, which were also GWAS SNPs at cis-regulatory elements (CREs). We observed that SNPs rs658524 and rs12225345 were in the non-coding region upstream of *CTSW* TSS, and in strong LD with each other (r^2 > 0.8) (**Figure 68**). Public epigenetic regulatory data from $CD8^+$ T cells showed that promoter and enhancer marks (H3K4me3 and H3K27ac) overlapped with the location of the two SNPs (**Figure 70**). This analysis suggests that these SNPs may be important for the regulation of *CTSW* in $CD8^+$ T cells. We noticed that the SNPs also coincided with chromatin accessible regions in $CD8^+$ T cells from the ATAC-seq data in immune cell populations (Calderon et al., 2019) (**Figure 71**), further emphasising the importance of these SNPs in *CTSW* regulation in $CD8^+$ T cells. It is however, important to note that the chromatin accessibility data showed that these regions are also accessible in NK cells (**Figure 71**), suggesting a similar regulatory mechanism for these SNPs in *CTSW* expression in NK cells.

Furthermore, we cannot exclude the possibility that there could be other variants in the locus that have causal implications. The region on the chromosome 11 has other genes like *FIBP* and *FOSL1,* which are expressed in specific immune cells. The SNPs in this region may also have a causal role for these genes. For instance the lead variant rs568617 from the meta-analysis by Ellinghaus et al. is located within *FIBP*, and it may be important in regulating *FIBP* expression in cell types where they are expressed that could imply causal relationships. FIBP is expressed in monocytes, NK cells and T cells based on the DICE database (not shown). Similarly, there are variants within *FOSL1* that may be influencing its expression in immune cells where the gene is prominently expressed- $CD4^+$ T cells, $CD8^+$ T cells and monocytes (DICE database; plot not shown).

Using a similar approach Vecellio et al. identified the functional role of rs4648889 in the regulation of *RUNX3* (GWAS gene for AS and PsA) by allele-specific binding of the TF IRF4 in T-cells (Vecellio et al., 2016). The authors also conducted a conditional SNP analysis between a lead GWAS SNP associated with AS (rs6600247) and SNP in LD with the lead

variant (rs4648889). Through the conditional analysis and the regulatory information at the SNP location (DHSs, ChIP-seq peaks for TF binding and histone modification) they identified rs4648889 as the most likely variant with a functional implication in AS.

7.7 Functional SNPs alter transcriptional factor binding activity to regulate *CTSW* expression

Genomic sequence-dependent TF binding at cis-regulatory elements is a mechanism that drives cell-type-specific gene expression (Liu & Tjian, 2018). From the eukaryotic promoter database we observed that rs12225345 and rs658524 are located 48 bp and 203 bp upstream of the *CTSW* TSS, respectively. Since 147 nucleotide make up a single turn around a nucleosome, the SNPs may be close to one another on the surface of the nucleosome. Due to their close proximity, the TFs binding to these SNPs may interact and cooperate with other TFs. We predicted binding sites for several TFs such as IRFs at rs658524 site, and NFAT and IRF4 are known to cooperate with each other (Hogan et al., 2003). Such interactions may be taking place during the regulation of *CTSW* expression.

The utilization of three online tools- HaploRegv4, Promo3 and JASPAR database ensured the reliability of the TF list we generated. Testing for allele-specific binding pattern, we identified NFATC2 as an important TF in *CTSW* regulation that binds in the presence of "A" allele of the SNP rs12225345. Using high affinity DNA capture assays, we demonstrated the preferential binding of NFATC2 by allele for rs12225345 in resting $CD8^+$ T cells, in comparison to other predicted TFs (**Figure 75**). The competitive binding assay validated that the "A" allele of rs12225345 is more effective in binding the TF than the allele "G" (**Figure 76**). These results provide substantial evidence that shows the effect of eQTL variants in regulating nearest gene expression implicated in disease mechanism. NFATC2 is a well-known transcriptional regulator associated with T-cell development and activation. NFAT family members have specific roles in $CD4^+$ T cells of Th1, Th17, T_{reg} and T_{fh} as reviewed by Hermann-Kleiter & Baier (2010). The differential role of NFATC1 and NFATC2 was reported in $CD8^+$ T cell differentiation and CTL function during acute viral infection recently by Xu, Keller, & Martinez (2019). The authors showed that despite the similarity in DNA binding motifs for NFATC1 and NFATC2, the former was essential in the differentiation of effector $CD8^+$ T cells, whereas the latter is essential to promote CTLs. Interestingly, a study published in 2017, had shown that NFATC1 contributed to $CD8^+$ T cell-mediated cytotoxicity, with ChIP-seq data showing NFATC1 binding for genes that are associated with CTL activity (Klein-Hessling et al., 2017). In our data, we observed weak binding for NFATC1 for *CTSW* regulation in $CD8^+$

T cells. However, a possible role for NFATC1-mediated *CTSW* regulation and CTL activity cannot be ruled out.

An interesting finding was that, while NFATC2 preferentially bound to the "A" allele of the SNP rs12225345, RBPJK (Recombination Signal Binding Protein For Immunoglobulin Kappa J Region) showed preferential binding to the alternate "G" allele of rs12225345 (**Figure 78**). RBPJK (CSL/CBF-1) is a DNA-binding transcription factor in the Notch signalling pathway. Notch-mediated gene transcription is initiated by the conversion of RBPJK from a co-repressor to a co-activator that occurs during the notch signalling pathway activation. Notch signalling is activated when the membrane bound Notch protein is cleaved to release the Notch intracellular domain (NICD). RBPJK acts as a repressor or activator based on the availability of NICD in the nucleus. In the absence of NICD, the target gene transcription is hindered by co-repressor complexes assembled on RBPJK. The cleaved NICD translocates to the nucleus and forms the "Notch transcription complex" with RBPJK, and replaces the co-repressor protein complex by recruiting co-activator protein Mastermind (MAM). We are currently studying whether RBPJK has a activator or repressor properties in this context. However, the allele "G" of rs12225345 where RBPJK binds is correlated with low expression levels of *CTSW*. These findings are consistent with reduced NFATC2-regulated transcription, and possibly, the recruitment of a repressor complex by RBPJK. It is also interesting to note that both NFATC2 and RBPJK, belong to the class Rel homology region (RHR) factors.

In the literature, *CTSW* has been reported to have a biphasic expression pattern after TCR activation, starting with high expression levels at 0h, followed by a dip after 6h of stimulation (Stoeckle et al., 2009). The *CTSW* expression increases again after 24-48h, reaching peak levels at 72h. NFATC2 activity is induced after TCR activation, hence there may be some constitutive TFs that maintain the *CTSW* high expression at unstimulated state of the cells. Therefore, the increase in *CTSW* expression from 24-48h may be NFATC2 dependent. When $CD8^+$ T cell are activated and in the presence of the "G" allele, RBPJK might be involved in regulating *CTSW* transcription in a repressive manner that results in decreased *CTSW* levels. *RBPJK* expression has been detected in activated $CD4^+$ and $CD8^+$ T cells from the DICE database (Schmiedel et al., 2018), consistent with a role for Notch-mediated signalling in activated T cells.

NFATC2 has been shown to inhibit Notch signalling by competing with RBPJK (CSL) for DNA binding with Notch transcriptional complex in osteoblasts (Zanotti, Smerdel-Ramoya, & Canalis, 2013). Notch signalling is known to play an important role in T cell differentiation and function. RBPJK has been shown to promote differentiation of pathogenic Th17 cells

through IL-23R signalling (Meyer zu Horste et al., 2016). The authors elucidated the dual role of RBPJK as a transcriptional regulator in Th17 cells by inducing IL-23R and supressing the production of anti-inflammatory cytokine IL-10. A similar mechanism may be occurring in $CD8^+$ T cells during their effector function. The essential role of Notch signalling in effector memory $CD8^+$ T cells in airway inflammation was reported by Okamoto et al., where the authors observed the activation of Notch 1 in TCR-activated $CD8^+$ T_{eff} cells (Okamoto et al., 2008). Another study showed that Notch 1 signalling is important to mediate CTL activity by regulation of expression of cytotoxic molecules (perforins and granzyme B) through RBPJK-mediated transcription activation (Cho et al., 2009). This study also reported that Notch signalling regulates *EOMES* expression- a key TF in CTL development through RBPJK binding at *EOMES* promoter region, showing a direct evidence in Notch 1 regulated *EOMES* transcription and CTL activity. In our data, the RBPJK binding to promoter region of *CTSW* was enabled by the presence of the "G" allele of rs12225345.

Our work demonstrated the effects of allele-dependent NFATC2 binding in transcriptional activation using dual reporter gene assay. We report an increase in luciferase activity in the presence of *CTSW* promoter sequences and NFATC2 for the "A" allele compared to "G" allele of rs12225345. Considering these results it would be interesting to understand how *CTSW* is involved in $CD8^+$ T cell function, and whether it may have a potential role in pathogenesis of CIDs such as AS and IBD. rs12225345 has shown concordant effects on CD and PsA- another form of AxSpA (UKBB, round 2 results). However, discordant effect of rs12225345 has not been reported in the literature for AS and IBD. A recent report showed that AxSpA patients have less number of $CD8^+$ intraepithelial lymphocytes (IELs) in gut compared to healthy donors, whereas an increase in number of IELs were reported in IBD patients compared to healthy donors (Regner et al., 2018). There are clinical cases of AxSpA and quiescent IBD and AxSpA and concomitantly active IBD, drawing attention to the gut-joint axis as reviewed by Gracey et al. (2020)

7.8 Possible implication of $CD8^+$ T cells in AS.

Cytotoxic T lymphocytes (CTLs) synthesize lytic proteins like perforins and granzymes, which are stored in lytic granules and released upon CTL triggering (Stinchcombe & Griffiths, 2007) as illustrated in **Figure 84**. CTL-mediated responses have been documented in several autoimmune and inflammatory diseases as discussed by Gravano & Hoyer, 2013. For example, patients with SLE exhibit elevated levels of $CD8^+$ T cells that express perforin and granzyme B, which were correlated with high disease activity (Blanco et al., 2005).

Despite the strong genetic association with HLA-B27 and other $CD8^+$ T cell function related genes (*EOMES, TBX21, RUNX3, NPEPPS, ERAP1* and *ERAP2*) to AS susceptibility, the specific role of $CD8^+$ T cells in AS is yet to be defined. MHC class I restricted CTLs have been reported to present self and non-self peptides in AS patients (Hermann et al., 1993; Fiorillo et al., 2000). Moreover, increase in circulating $CD8^+$ $CD28^-$ T cells that exhibit a dual cytotoxic and regulatory role has been observed in peripheral blood from AS patients (Schirmer et al., 2001). Studies from Robert Inman's group have observed the expansion of $CD8^+$ T cells in AS patients (Faham et al., 2017) and also described a unique $CD8^+$ T cell subset prevalent in synovial fluid (SF) from AS patients that expresses several integrins (β7, CD103, CD29 and CD49a) (Qaiyum et al., 2019). The $CD8^+$ T cells from SF of AS patients reported by Qaiyum et al. also showed a cytotoxic profile with high levels of *GZMB* and *PRF1*, and regulatory potential with increased *IL-10* and *TNFAIP3*. To the contrary, a recent study from the same group reported a reduction in cytotoxic potential of $CD8^+$ T cells with low expression of granzyme B and perforin using gene, protein and cellular data from whole blood (Gracey, Yao, et al., 2020). These studies suggest the existence of altered cytotoxicity and a possible role of $CD8^+$ T cells in AS pathogenesis.

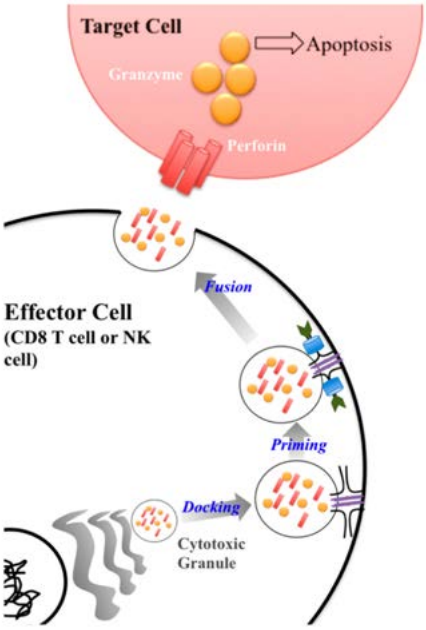


Figure 84: Granule-mediated cytotoxicity in CTLs. Cytotoxic granules (granzymes and perforins) are produced by the effector cells and are released into the immunological synapse via exocytosis. Perforin forms a pore on the plasma membrane of the target cell allowing granzyme to be delivered and consequently induce cell death. Adapted from Willenbring & Johnson, 2017.

7.9 Role of cysteine cathepsins in cytotoxic T lymphocytes

Cathepsins have been shown to have an important function during the synthesis and activation of granzyme B. In the secretory pathway, granzyme B is in the pro-granzyme B form (zymogen) in the ER, with an N-terminal peptide that prevents it from getting activated. When pro-granzyme B reaches the Golgi, it moves on to the lysosomes using the mannose-6-phosphate secretory pathway (Stinchcombe & Griffiths, 2007). In the lysosome, cysteine protease cathepsin C cleaves the N-terminal dipeptide of the pro-granzyme B and activates it (D'Angelo et al., 2010). The authors also showed a similar function for cathepsin H in converting pro-granzyme B into its active form in vitro.

Cysteine cathepsins have also been implicated in the protection of cytotoxic cells from the activity of their lytic granules. The acidic environment in the granules prevents the granzymes and perforins from getting activated. However, it remains unclear how the CTLs remain protected from perforins and granules when they are released into the lytic synapse. Cathepsin B (CatB) has shown to play a role in the self-protection mechanism of cytotoxic cells during exocytosis of granules (Balaji et al., 2002). This study demonstrated that CatB inhibition leads to perforin-dependent death of CTLs, suggesting that CatB might prevent self-destruction of CTLs when the granules are released into the cytolytic synapse, as depicted in **Figure 85**. CatB is present in the membrane of the secretory granules containing perforins and granzymes. Upon exocytosis, the membrane of the lytic granules fuses with the plasma membrane and release the granules into the lytic synapse. The CatB present on the granular membrane fuses with the plasma membrane of the CTLs and might act on the cytolytic molecules by cleaving and deactivating them.

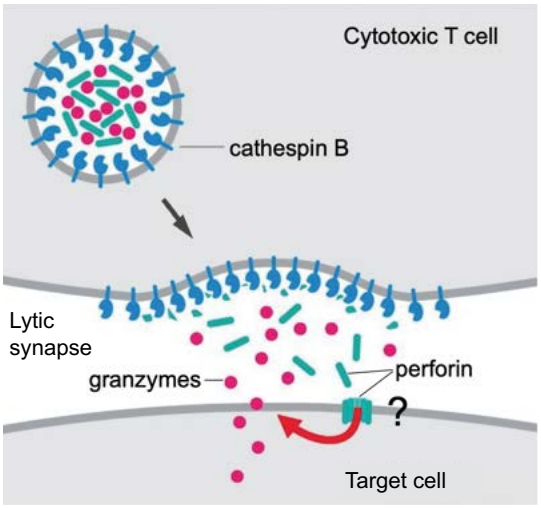


Figure 85: Perforin and granzyme secretion from cytotoxic cell in the presence of cathepsin B. Adapted from (Griffiths, 2013).

However, another study reported that CTLs in CatB deficient mice were able to survive and maintain their cytotoxic potential and effector function *in vitro* and *in vivo* (Baran et al., 2006). Therefore the lack of CatB may be compensated by different cysteine cathepsins present in granular compartments. Of note, CatW is predominantly expressed in cytotoxic cells compared to CatB and has a higher expression in NK cells compared to $CD8^+$ CTLs (Stoeckle et al., 2009). However, the enzymatic properties and the function of CatW in cytotoxicity are still unclear.

7.9.1 Cathepsin W in cytotoxic cells and cytotoxicity

CatW was initially identified as a cysteine endopeptidase of the papain family that was mainly expressed in $CD8^+$ T cells (Linnevers, Smeekens, & Brömme, 1997). Brown et al. referred to this enzyme as "lymphopain" due to its restricted expression in cytotoxic T cells and NK cells (Brown et al., 1998), suggesting a role in cytotoxic cell-mediated apoptosis. Phylogenetic analyses showed that cathepsin F and W share similarities in the amino acid sequences compared to other peptidases (Brinkworth et al., 2000).

Wex et al. explored the functional role of CatW by using cytotoxicity assays with the NK cell line NK92 against the erythropoietic cell line K562 (Wex et al., 2003). This study used anti-sense mediated inhibition to downregulate CatW and observed impaired cytotoxic activity in NK92 cells, which was correlated with low levels of CatW. Moreover, they observed that CatW protein levels decreased over time from cell lysates during the cytotoxic assay, compatibly with a release during the cytotoxic attack. Interestingly, Wex et al. also observed that overexpressed CatW was localised to the ER compartment of NK92 cells using immunofluorescence experiments. These data suggest that CatW was not directed to the target K562 cells during the cytotoxic attack by the NK92, indicating a role within the cytotoxic cell. Stoeckle et al. reported that siRNA knockdown of CatW in $CD8^+$ T cells co-incubated with target T2 cells loaded with CMV-specific peptides did not perturb target cell killing using a chromium release assay (Stoeckle et al., 2009). Furthermore, the authors did not observe an effect on IFNγ production after CatW knockdown, suggesting that, although CatW is released during target cell killing from human CTLs, it is not a requirement to initiate cytotoxicity (Stoeckle et al., 2009). This may also suggest that, in the absence of CatW there may be other cathepsins that can compensate for its function. Cathepsin F and W have shown phylogenetic similarities as previously reported by Brinkworth et al. (2000); CatF is also expressed in T cells and NK cells (**Figure 66**). CatW expressing cells were also detected in the gastrointestinal

tissue of patients with IBD (CD & UC) and autoimmune gastritis as reported by Buhling et al. in 2002.

Another paper discussed the potential involvement of CatW in the endo-lysosomal pathway. A study published in 2015 described the role of CatW in influenza A virus (IAV) replication and reported that CatW is required to release IAV from the late endosome (Edinger et al., 2015). The authors observed that the knockdown of CatW led to the accumulation of viral nucleoprotein in the endosome, preventing translocation of the viral particles into the nucleus. This finding indicates that CatW is present in the endosomal pathway. Together, the earlier hypothesis about the role of CatW in cytotoxic granule formation and exocytosis, and the results from Edinger et al. suggest a role for CatW in the vesicular transport pathway within the cell.

7.9.2 Cathepsin W may be involved in autophagy

Cathepsins have also been implicated in the autophagy process. Autophagy is a constitutive cellular mechanism for compartmentalisation and lysosomal degradation of various cytoplasmic components- misfolded proteins, damaged organelles, invading pathogens, etc. Autophagic responses are vital intrinsic cellular processes that protect cells from internal and external stresses. Autophagy-related genes (ATG) have been linked to membrane-trafficking pathways involved in phagocytosis, lysosomal delivery and exocytosis of secretory granules (Ma et al., 2013; Levine & Kroemer, 2019). Autophagy also has an important role in cellular immunity - especially in the differentiation and activation of immune cells, such as APCs and T lymphocytes as reviewed by Ma et al. (2013). Several genetic polymorphisms in ATG are also associated with CIDs like SLE (*ATG5*) (Harley et al., 2008) and CD (*ATG16L1* and *NOD2*) (Fritz et al., 2011). Autophagy is suggested to influence the expression of IL-23 in driving the gut inflammation in AS patients (Ciccia et al., 2014).

Cathepsins have been implicated in autophagic processes - impaired *CTSB* and *CTSD* can cause saposin C deficiency in fibroblasts, leading to the accumulation of autophagosomes that is a feature Gaucher disease - of a lysosomal storage disorder (Tatti et al., 2013). Another function for autophagy in secretory lysosome exocytosis was reported in osteoclasts. Bone resorption mechanism requires secretory lysosome fusion to the specialised osteoclast cell membrane during which cathepsin K is released for bone matrix degradation (DeSelm et al., 2011). It would be interesting to explore whether cathepsin W may have a role in the vesicle trafficking observed in CTLs.

7.9.3 Cathepsin W may be localized to the compartments of vesicular transport pathways

We observed that CatW is sensitive to both glycosidases Endo H and PNGase F, suggesting that CatW possesses high mannose structures (**Figure 83**). Glycoproteins in the ER can acquire high mannose structures, as they pass through the stages of glycosylation. Upon reaching the Golgi, glycoproteins tend to lose their high mannose structures. However, there is a group of proteins leaving the ER, that retain a high mannose structures, which becomes phosphorylated in the Golgi (M6P). These proteins are targeted towards lysosomes and vesicular compartments that express M6P-receptors (M6P-R). Therefore, the M6Ps act as recognition markers that aid in the transportation of these enzymes to the endosomal/lysosomal compartments through recognition by the M6P-R as reviewed by Gary-Bobo et al. (2007). Therefore, in the case of CatW, the glycosidase sensitivity may indicate three things: (1.) CatW is an ER protein and has a site-specific function within the ER compartment, (2.) CatW may have a phosphorylated high mannose conformation, that may be targeted to endo-lysosomal compartments, and (3.) some CatW may be in a simple high mannose structure and be localized in the ER, and another pool of CatW may adopt M6P structures in the Golgi as it moves along the protein trafficking pathway. The second and third possibilities, are compatible with a localization of CatW in lysosomes and exocytosis vesicles that may perform a potential CTL-related function.

Our results complement the previous work from Wex et al. (2001) and Ondr & Pham (2004), who reported that CatW is sensitive to Endo H, suggesting localisation of CatW to ER. Moreover, Stoeckle et al. proposed that CatW is localized to the ER and the Golgi apparatus based on immunofluorescence microscopy data (Stoeckle et al., 2009). Edinger et al. (2015) showed that cathepsin W is required for release of influenza A virus (IAV) from the late endosome for viral replication. This study also observed that CatW was co-localised with calnexin (ER-resident protein), and with LAMP1 (Lysosomal-associated membrane protein 1, or CD107a) - a marker for late endosome in A549 cells. These studies may have captured the different stages of maturation of CatW in the vesicular compartments, which suggests that CatW may be processed and transported out of the cell through the trafficking pathway. The preprocathespsin W (immature form) may start out in the ER, which is then translocated to the Trans-Golgi network (TGN) to be parcelled out to the endosome/lysosome. The acidic pH in the endosomal/lysosomal compartment could enable the maturation of CatW. The mature form of CatW may then aid in protein degradation and autophagy related events. CatW may also have a site-specific function within the organelles during the processing, transport and release of lytic molecules from cytotoxic cells (**Figure 86**). These results along with the

suggested role of CatW in cytotoxicity may indicate its potential function in the release of lytic granules, however this hypothesis requires further study.

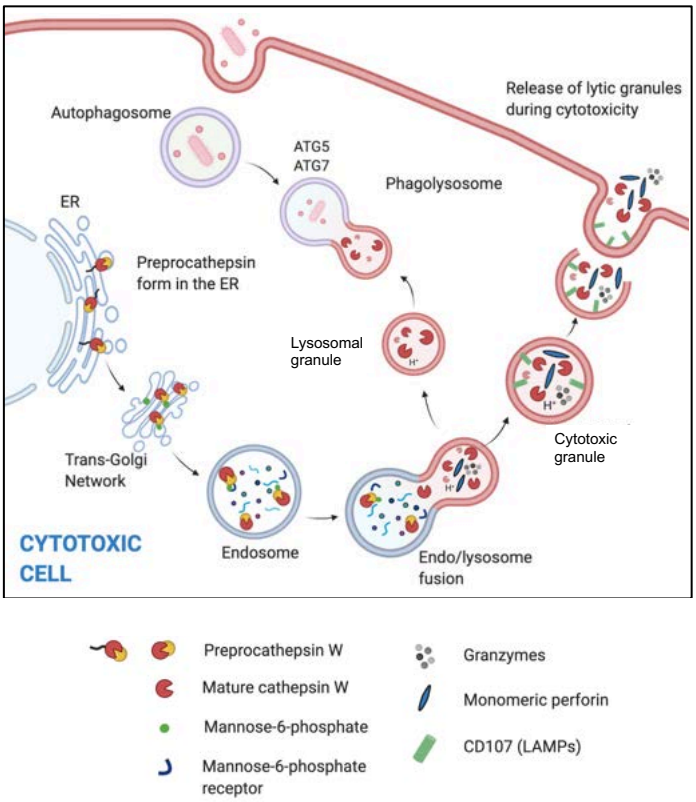


Figure 86: Schematic explaining the hypothetical role of cathepsin W in membrane trafficking pathways in cytotoxic cells (Adapted from Levine & Kroemer, 2019; Yadati et al., 2020)**.**

Conclusion

The genetic associations in CIDs have implicated the immune system in the pathogenesis of these diseases through the identification of several immune function-related genes. However, the evidence linking these susceptibility genes to immune function, and the immune cell types in which they are expressed have been incompletely explored. We report that the disease-susceptibility genes are expressed in innate and adaptive immune cells that may be relevant to the disease biology of AxSpA. With the eQTL analysis in T cells, we showed that disease-associated variants: rs658524 and rs12225345 can affect *CTSW* expression - a nearest gene in the susceptibility locus for AS, Pso, CD and UC. We further demonstrated a regulatory role for rs12225345 in *CTSW* expression by altering TF binding motif for NFATC2. This work allows us to better understand the mechanism by which variants associated to diseases can moderate expression of genes in the susceptibility locus, which could be important in the mechanism of CIDs.

ANNEX I

List of abbreviations

AS	Ankylosing spondylitis
ASAS	Assessment of SpondyloArthritis international Society
ATAC	Assay for transposase- accessible chromatin
ATG	Autophagy-related genes
AxSpA	Axial Spondyloarthritis
CatW	Cathepsin W
CD	Crohn's disease
CeD	Coeliac's disease
CID	Chronic inflammatory diseases
cQTL	cytokine quantitative trait loci
CRE	cis-regulatory elements
CTLA4	Cytotoxic T-lymphocyte-associated protein 4
CTSW	Cathepsin W (transcript)
DMARDs	Disease-modifying anti-rheumatic drugs
DNA	Deoxyribonucleic acid
Endo H	Endo-ß-N-acetylglucosaminidase H
eQTLs	Expression quantitative trait loci
ER	Endoplasmic reticulum
EULAR	European League Against Rheumatism
GD	Grave's disease
GWAS	Genome-wide association studies
HLA	Human leukocyte antigen
IBD	Inflammatory bowel disease
IRF1	Interferon Regulatory Factor 1
LAMP1	Lysosomal-associated membrane protein 1
M6P	Mannose-6-Phosphates
Mb	Megabase
MHC	Major histocompatibility complex
MS	Multiple sclerosis
NFATC1	Nuclear factor-activated T cells c1
NFATC2	Nuclear factor-activated T cells c2
NFκB	Nuclear Factor kappa-light-chain-enhancer of activated B cells
NICD	Notch intracellular domain
NOD2	Nucleotide-binding oligomerization domain-containing protein 2
NSAIDs	Non-steroidal anti-inflammatory drugs
PheWAS	Phenotype-wide association study
PNGase F	Peptide-N4-(N-acetyl-beta-glucosaminyl) asparagine amidase
Pso	Psoriasis
RA	Rheumatoid arthritis
RBPJK	Recombination Signal Binding Protein For Immunoglobulin Kappa J Region
RNA	Ribonucleic acid
SLE	Systemic lupus erythematosus
SNP	Single nucleotide polymorphism
STAT4	Signal Transducer And Activator Of Transcription 4
T1D	Type 1 diabetes
TF	Transcription factor
TFBS	Transcription factor binding sites

TGN	Trans-Golgi network
TNFi	TNF inhibitors
TSS	Transcription start sites
TWAS	Transcriptome-wide association studies
UC	Ulcerative colitis

List of figures

List of tables

List of genes in the AID panel

Gene	Alias / Previous Symbol
AAMP	Angio Associated Migratory Cell Protein
ABHD6	Abhydrolase Domain Containing 6, Acylglycerol Lipase
ACKR2	CMKBR9,CCBP2;chemokine binding protein 2
ACOXL	acyl-Coenzyme A oxidase-like
ACSL6	FACL6;fatty-acid-Coenzyme A ligase, long-chain 6
ADA	Adenosine Deaminase
ADAM30	a disintegrin and metalloproteinase domain 30
ADCY3	Adenylate Cyclase 3
ADCY7	Adenylate Cyclase 7
AFF3	LAF4;lymphoid nuclear protein related to AF4,AF4/FMR2 family, member 3
AGAP2	CENTG1;centaurin, gamma 1
AHI1	Abelson helper integration site
AHR	Aryl Hydrocarbon Receptor
AHSA2	AHSA2;AHA1, activator of heat shock 90kDa protein ATPase homolog 2 (yeast),activator of HSP90 ATPase homolog 2
AIRE	APECED;autoimmune regulator (autoimmune polyendocrinopathy candidiasis ectodermal dystrophy)
AMIGO3	Adhesion Molecule With Ig Like Domain 3
ANKRD55	Ankyrin Repeat Domain 55
ANTXR2	ANTXR Cell Adhesion Molecule 2
APEH	D3F15S2,DNF15S2,D3S48E;N-acylaminoacyl-peptide hydrolase
APOBEC3G	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3G
ARG1	arginase, liver
ARHGAP30	Rho GTPase Activating Protein 30
ARID5B	AT rich interactive domain 5B (MRF1-like)
ARPC2	actin related protein 2/3 complex, subunit 2 (34 kD),actin related protein 2/3 complex subunit 2, 34kDa
ATF4	TXREB;activating transcription factor 4 (tax-responsive enhancer element B67)
ATG16L1	APG16L,ATG16L;APG16 autophagy 16-like (S. cerevisiae),ATG16 autophagy related 16-like (S. cerevisiae),ATG16 autophagy related 16-like 1 (S. cerevisiae)
ATG5	APG5L;APG5 (autophagy 5, S. cerevisiae)-like,APG5 autophagy 5-like (S. cerevisiae),ATG5 autophagy related 5 homolog (S. cerevisiae)
ATM	ATA,ATDC,ATC,ATD;ataxia telangiectasia mutated (includes complementation groups A, C and D),ataxia telangiectasia mutated
B2M	Beta-2-Microglobulin
B3GNT2	B3GNT1;UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 1
BABAM2	BRE;brain and reproductive organ-expressed (TNFRSF1A modulator)
BACH2	BTB and CNC homology 1, basic leucine zipper transcription factor 2
BAD	BCL2 Associated Agonist Of Cell Deat
BANK1	B Cell Scaffold Protein With Ankyrin Repeats 1
BATF	basic leucine zipper transcription factor, ATF-like

BATF3	basic leucine zipper transcription factor, ATF-like 3
BCL10	BCL10 Immune Signaling Adaptor
BCL3	D19S37,BCL4
BCL6	ZNF51;zinc finger protein 51
BID	BH3 Interacting Domain Death Agonist
BLK	B lymphoid tyrosine kinase
BLNK	B Cell Linker
BORCS5	LOH12CR1;loss of heterozygosity, 12, chromosomal region 1
BSN	ZNF231;bassoon (presynaptic cytomatrix protein)
BTK	AGMX1,IMD1;Bruton agammaglobulinemia tyrosine kinase
BTLA	B And T Lymphocyte Associated
BTNL2	butyrophilin-like 2 (MHC class II associated),butyrophilin-like 2
C1QBP	HABP1;complement component 1, q subcomponent binding protein
C1QTNF6	C1q and tumor necrosis factor related protein 6
C1orf106	C1orf106;chromosome 1 open reading frame 106
C1orf53	Chromosome 1 Open Reading Frame 53
C5orf30	Chromosome 5 Open Reading Frame 30
CALM3	calmodulin 3 (phosphorylase kinase, delta)
CARD11	caspase recruitment domain family, member 11
CARD14	PSORS2;psoriasis susceptibility 2,caspase recruitment domain family, member 14
CARD9	caspase recruitment domain family, member 9
CARM1	Coactivator Associated Arginine Methyltransferase 1
CASP1	IL1BC;caspase 1, apoptosis-related cysteine protease (interleukin 1, beta, convertase),caspase 1, apoptosis-related cysteine peptidase (interleukin 1, beta, convertase),caspase 1, apoptosis-related cysteine peptidase
CASP10	caspase 10, apoptosis-related cysteine protease,caspase 10, apoptosis-related cysteine peptidase
CASP2	NEDD2;neural precursor cell expressed, developmentally down-regulated 2,caspase 2, apoptosis-related cysteine peptidase
CASP3	caspase 3, apoptosis-related cysteine protease,caspase 3, apoptosis-related cysteine peptidase
CASP8	caspase 8, apoptosis-related cysteine protease,caspase 8, apoptosis-related cysteine peptidase
CAVIN1	PTRF;polymerase I and transcript release factor
CBLB	Cas-Br-M (murine) ectropic retroviral transforming sequence b,Cas-Br-M (murine) ecotropic retroviral transforming sequence b,Cbl proto-oncogene B, E3 ubiquitin protein ligase
CCDC116	Coiled-Coil Domain Containing 116
CCDC88B	CCDC88;coiled-coil domain containing 88
CCL11	SCYA11;small inducible cytokine subfamily A (Cys-Cys), member 11 (eotaxin),chemokine (C-C motif) ligand 11
CCL13	SCYA13;small inducible cytokine subfamily A (Cys-Cys), member 13,chemokine (C-C motif) ligand 13
CCL18	SCYA18;small inducible cytokine subfamily A (Cys-Cys), member 18, pulmonary and activation-regulated,chemokine (C-C motif) ligand 18 (pulmonary and activation-regulated),chemokine (C-C motif) ligand 18
CCL19	SCYA19;small inducible cytokine subfamily A (Cys-Cys), member 19,chemokine (C-C motif) ligand 19
CCL2	SCYA2;small inducible cytokine A2 (monocyte chemotactic protein 1, homologous to mouse Sig-je),chemokine (C-C motif) ligand 2

CCL20	SCYA20;small inducible cytokine subfamily A (Cys-Cys), member 20,chemokine (C-C motif) ligand 20
CCL21	SCYA21;small inducible cytokine subfamily A (Cys-Cys), member 21,chemokine (C-C motif) ligand 21
CCL22	SCYA22;small inducible cytokine subfamily A (Cys-Cys), member 22,chemokine (C-C motif) ligand 22
CCL23	SCYA23;small inducible cytokine subfamily A (Cys-Cys), member 23,chemokine (C-C motif) ligand 23
CCL24	SCYA24;small inducible cytokine subfamily A (Cys-Cys), member 24,chemokine (C-C motif) ligand 24
CCL3	SCYA3;small inducible cytokine A3 (homologous to mouse Mip-1a),chemokine (C-C motif) ligand 3
CCL4	LAG1,SCYA4;small inducible cytokine A4 (homologous to mouse Mip-1b),chemokine (C-C motif) ligand 4
CCL5	D17S136E,SCYA5;small inducible cytokine A5 (RANTES),chemokine (C-C motif) ligand 5
CCL7	SCYA6,SCYA7;small inducible cytokine A7 (monocyte chemotactic protein 3),chemokine (C-C motif) ligand 7
CCL8	SCYA8;small inducible cytokine subfamily A (Cys-Cys), member 8 (monocyte chemotactic protein 2),chemokine (C-C motif) ligand 8
CCNY	C10orf9;chromosome 10 open reading frame 9
CCR1	SCYAR1,CMKBR1;chemokine (C-C motif) receptor 1
CCR2	CMKBR2;chemokine (C-C motif) receptor 2
CCR4	chemokine (C-C motif) receptor 4
CCR5	CMKBR5;chemokine (C-C motif) receptor 5,chemokine (C-C motif) receptor 5 (gene/pseudogene)
CCR6	STRL22;chemokine (C-C motif) receptor 6
CCR7	CMKBR7,EBI1;chemokine (C-C motif) receptor 7
CD19	CD19 antigen
CD209	CD209 antigen
CD226	CD226 antigen
CD244	natural killer cell receptor 2B4,CD244 natural killer cell receptor 2B4,CD244 molecule, natural killer cell receptor 2B4
CD247	CD3Z;CD3z antigen, zeta polypeptide (TiT3 complex),CD247 antigen
CD27	TNFRSF7;tumor necrosis factor receptor superfamily, member 7
CD274	PDCD1LG1;programmed cell death 1 ligand 1,CD274 antigen
CD28	CD28 antigen (Tp44)
CD40	TNFRSF5;tumor necrosis factor receptor superfamily, member 5,CD40 molecule, TNF receptor superfamily member 5
CD40LG	HIGM1,IMD3,TNFSF5;tumor necrosis factor (ligand) superfamily, member 5 (hyper-IgM syndrome)
CD44	MIC4,MDU2,MDU3;CD44 antigen (homing function and Indian blood group system)
CD45R0	CD45
CD45RA	CD45
CD45RB	CD45
CD48	BCM1;CD48 antigen (B-cell membrane protein),CD48 molecule
CD5	LEU1;CD5 antigen (p56-62)
CD58	LFA3;CD58 antigen, (lymphocyte function-associated antigen 3)
CD6	CD6 antigen
CD69	CD69 antigen (p60, early T-cell activation antigen)

CD80	CD28LG,CD28LG1;CD80 antigen (CD28 antigen ligand 1, B7-1 antigen),CD80 molecule
CD83	CD83 antigen (activated B lymphocytes, immunoglobulin superfamily),CD83 molecule
CD86	CD28LG2;CD86 antigen (CD28 antigen ligand 2, B7-2 antigen)
CDC37	CDC37 (cell division cycle 37, S. cerevisiae, homolog),CDC37 cell division cycle 37 homolog (S. cerevisiae),cell division cycle 37 homolog (S. cerevisiae)
CDK2	Cyclin Dependent Kinase 2
CDK6	Cyclin Dependent Kinase 6
CDKAL1	CDK5 regulatory subunit associated protein 1-like 1
CEBPB	TCF5;CCAAT/enhancer binding protein (C/EBP), beta
CEBPG	CCAAT/enhancer binding protein (C/EBP), gamma
CENPO	Centromere Protein O
CEP250	CEP2;centrosomal protein 2,centrosomal protein 250kDa
CEP57	centrosomal protein 57kDa
CFLAR	CASP8AP1;CASP8 and FADD-like apoptosis regulator
CHUK	TCF16
CIITA	MHC2TA;MHC class II transactivator,class II, major histocompatibility complex, transactivator
CISD1	C10orf70,ZCD1;chromosome 10 open reading frame 70,zinc finger, CDGSH-type domain 1
CISH	cytokine inducible SH2-containing protein
CLEC16A	KIAA0350;KIAA0350,C-type lectin domain family 16, member A,C-type lectin domain family 16 member A
CLEC4A	CLECSF6;C-type (calcium dependent, carbohydrate-recognition domain) lectin, superfamily member 6,C-type lectin domain family 4, member A
CLEC4E	CLECSF9;C-type (calcium dependent, carbohydrate-recognition domain) lectin, superfamily member 9,C-type lectin domain family 4, member E
CLEC5A	CLECSF5;C-type (calcium dependent, carbohydrate-recognition domain) lectin, superfamily member 5,C-type lectin domain family 5, member A,C-type lectin domain family 5 member A
CLEC6A	CLECSF10;C-type (calcium dependent, carbohydrate-recognition domain) lectin, superfamily member 10,C-type lectin domain family 6, member A,C-type lectin domain family 6 member A
CLEC7A	CLECSF12;C-type (calcium dependent, carbohydrate-recognition domain) lectin, superfamily member 12,C-type lectin domain family 7, member A,C-type lectin domain family 7 member A
COG6	Component Of Oligomeric Golgi Complex 6
CPEB4	Cytoplasmic Polyadenylation Element Binding Protein 4
CREM	CAMP Responsive Element Modulator
CRTC3	CREB Regulated Transcription Coactivator 3
CSF1	colony stimulating factor 1 (macrophage)
CSF1R	FMS;McDonough feline sarcoma viral (v-fms) oncogene homolog
CSF2	colony stimulating factor 2 (granulocyte-macrophage)
CSF2RB	IL3RB;colony stimulating factor 2 receptor, beta, low-affinity (granulocyte-macrophage)
CSF3	GCSF,G-CSF,C17orf33;chromosome 17 open reading frame 33,colony stimulating factor 3 (granulocyte)
CSF3R	CD114;colony stimulating factor 3 receptor (granulocyte)

CSK	c-src tyrosine kinase,CSK, non-receptor tyrosine kinase
CTDSP1	CTD (carboxy-terminal domain, RNA polymerase II, polypeptide A) small phosphatase 1
CTLA4	CELIAC3,IDDM12;celiac disease 3,insulin-dependent diabetes mellitus 12
CTNNB1	CTNNB;catenin (cadherin-associated protein), beta 1 (88kD),catenin (cadherin-associated protein), beta 1, 88kDa,catenin (cadherin-associated protein), beta 1
CTSH	CPSB
CTSW	cathepsin W (lymphopain)
CTSZ	cathepsin Z
CX3CR1	GPR13,CMKBRL1;chemokine (C-X3-C) receptor 1,chemokine (C-X3-C motif) receptor 1
CXCL1	MGSA,GRO1,FSP;GRO1 oncogene (melanoma growth stimulating activity, alpha),fibroblast secretory protein,chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)
CXCL10	INP10,SCYB10;small inducible cytokine subfamily B (Cys-X-Cys), member 10,chemokine (C-X-C motif) ligand 10
CXCL11	SCYB9B,SCYB11;small inducible cytokine subfamily B (Cys-X-Cys), member 11,chemokine (C-X-C motif) ligand 11
CXCL13	SCYB13;small inducible cytokine B subfamily (Cys-X-Cys motif), member 13 (B-cell chemoattractant),chemokine (C-X-C motif) ligand 13
CXCL2	GRO2;GRO2 oncogene,chemokine (C-X-C motif) ligand 2
CXCL3	GRO3;GRO3 oncogene,chemokine (C-X-C motif) ligand 3
CXCL5	SCYB5;small inducible cytokine subfamily B (Cys-X-Cys), member 5 (epithelial-derived neutrophil-activating peptide 78),chemokine (C-X-C motif) ligand 5
CXCL8	IL8;interleukin 8,chemokine (C-X-C motif) ligand 8
CXCL9	CMK,MIG;monokine induced by gamma interferon,chemokine (C-X-C motif) ligand 9
CXCR1	CMKAR1,IL8RA;interleukin 8 receptor, alpha,chemokine (C-X-C motif) receptor 1
CXCR2	IL8RB;interleukin 8 receptor, beta,chemokine (C-X-C motif) receptor 2
CXCR3	GPR9;G protein-coupled receptor 9,chemokine (C-X-C motif) receptor 3
CXCR4	chemokine (C-X-C motif), receptor 4 (fusin),chemokine (C-X-C motif) receptor 4
CXCR5	BLR1;Burkitt lymphoma receptor 1, GTP-binding protein,Burkitt lymphoma receptor 1, GTP binding protein (chemokine (C-X-C motif) receptor 5),chemokine (C-X-C motif) receptor 5
CXCR6	chemokine (C-X-C motif) receptor 6
CXorf21	Chromosome X Open Reading Frame 21
DAP	Death Associated Protein
DBP	D site of albumin promoter (albumin D-box) binding protein
DCLRE1B	DNA cross-link repair 1B (PSO2 homolog, S. cerevisiae)
DDAH1	Dimethylarginine Dimethylaminohydrolase 1
DDX58	DEAD (Asp-Glu-Ala-Asp) box polypeptide 58
DDX6	HLR2;DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 6 (RNA helicase, 54kD),DEAD (Asp-Glu-Ala-Asp) box polypeptide 6,DEAD (Asp-Glu-Ala-Asp) box helicase 6
DENND1B	FAM31B,C1orf218;family with sequence similarity 31, member B,chromosome 1 open reading frame 218,DENN/MADD domain containing 1B
DEXI	Dexi homolog (mouse)
DGKA	DAGK,DAGK1;diacylglycerol kinase, alpha (80kD),diacylglycerol kinase, alpha 80kDa

DHCR7	SLOS;Smith-Lemli-Opitz syndrome
DLD	LAD,GCSL;dihydrolipoamide dehydrogenase (E3 component of pyruvate dehydrogenase complex, 2-oxo-glutarate complex, branched chain keto acid dehydrogenase complex)
DLK1	delta-like homolog (Drosophila),delta-like 1 homolog (Drosophila)
DNASE1L3	deoxyribonuclease I like 3
DNMT3A	DNA (cytosine-5-)-methyltransferase 3 alpha
DOK3	Docking Protein 3
DPP4	CD26,ADCP2;dipeptidylpeptidase IV (CD26, adenosine deaminase complexing protein 2),adenosine deaminase complexing protein 2,dipeptidyl-peptidase 4
DUSP4	Dual Specificity Phosphatase 4
EBI3	Epstein-Barr Virus Induced 3
EDEM3	C1orf22;chromosome 1 open reading frame 22,ER degradation enhancer, mannosidase alpha-like 3
EFR3B	KIAA0953;KIAA0953,EFR3 homolog B (S. cerevisiae)
EGR1	Early Growth Response 1
EGR2	KROX20;early growth response 2 (Krox-20 homolog, Drosophila)
EIF3C	EIF3S8;eukaryotic translation initiation factor 3, subunit 8, 110kDa,eukaryotic translation initiation factor 3, subunit C
ELMO1	engulfment and cell motility 1 (ced-12 homolog, C. elegans)
ELP1	DYS,IKBKAP;dysautonomia (Riley-Day syndrome, hereditary sensory autonomic neuropathy type III),inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase complex-associated protein
EMSY	C11orf30;chromosome 11 open reading frame 30
ENTPD1	CD39
EOMES	eomesodermin (Xenopus laevis) homolog
EPO	Erythropoietin
EPS15L1	epidermal growth factor receptor pathway substrate 15-like 1
ERAP1	Endoplasmic Reticulum Aminopeptidase 1
ERAP2	Endoplasmic Reticulum Aminopeptidase 2
ERBB3	LCCS2;lethal congenital contracture syndrome 2,v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 3
ERRFI1	ERBB Receptor Feedback Inhibitor 1
ESRRA	ESRL1
ETS1	EWSR2;v-ets avian erythroblastosis virus E26 oncogene homolog 1
ETS2	v-ets erythroblastosis virus E26 oncogene homolog 2 (avian),v-ets avian erythroblastosis virus E26 oncogene homolog 2
ETV7	ets variant gene 7 (TEL2 oncogene)
EVI5	Ecotropic Viral Integration Site 5
EXOC2	SEC5L1;SEC5-like 1 (S. cerevisiae)
EXTL2	exostoses (multiple)-like 2
F11R	JAM1;junctional adhesion molecule 1
FADD	Fas (TNFRSF6)-associated via death domain
FADS1	LLCDL1
FADS2	LLCDL2
FADS3	LLCDL3
FAM213A	C10orf58;chromosome 10 open reading frame 58,family with sequence similarity 213, member A

FAM213B	C1orf93;chromosome 1 open reading frame 93,family with sequence similarity 213, member B
FAM98B	family with sequence similarity 98, member B
FAS	FAS1,APT1,TNFRSF6;tumor necrosis factor receptor superfamily, member 6,Fas (TNF receptor superfamily, member 6)
FASLG	APT1LG1,TNFSF6;tumor necrosis factor (ligand) superfamily, member 6,Fas ligand (TNF superfamily, member 6)
FBXL19	F-Box And Leucine Rich Repeat Protein 19
FCER1A	FCE1A;Fc fragment of IgE, high affinity I, receptor for; alpha polypeptide
FCER1G	Fc fragment of IgE, high affinity I, receptor for; gamma polypeptide
FCGR1A/B	Fc fragment of IgG, high affinity Ia, receptor for (CD64),Fc fragment of IgG, high affinity Ia, receptor (CD64)
FCGR2A/C	Fc fragment of IgG, low affinity IIc, receptor for (CD32),Fc fragment of IgG, low affinity IIc, receptor for (CD32) (gene/pseudogene)
FCGR2B	FCG2,FCGR2;Fc fragment of IgG, low affinity IIb, receptor for (CD32),Fc fragment of IgG, low affinity IIb, receptor (CD32)
FCGR3A/B	FCGR3,FCG3;Fc fragment of IgG, low affinity IIIa, receptor for (CD16),Fc fragment of IgG, low affinity IIIa, receptor (CD16a)
FCGRT	Fc fragment of IgG, receptor, transporter, alpha
FCMR	FAIM3;Fas apoptotic inhibitory molecule 3
FCRL1	Fc receptor-like 1
FIGNL1	fidgetin-like 1
FKBP5	FK506-binding protein 5
FLI1	Friend leukemia virus integration 1
FLRT1	Fibronectin Leucine Rich Transmembrane Protein 1
FOS	v-fos FBJ murine osteosarcoma viral oncogene homolog,FBJ murine osteosarcoma viral oncogene homolog
FOSL1	FOS like antigen 1
FOSL2	FOS like antigen 2
FOXO3	FKHRL1,FOXO3A
FOXP1	Forkhead Box P1
FOXP3	IPEX;immune dysregulation, polyendocrinopathy, enteropathy, X-linked
FUT2	SE;fucosyltransferase 2 (secretor status included)
FYN	FYN oncogene related to SRC, FGR, YES
GALC	galactosylceramidase (Krabbe disease)
GART	PRGS,PGFT
GATA3	GATA-binding protein 3
GBP1	guanylate binding protein 1, interferon-inducible, 67kDa,guanylate binding protein 1, interferon-inducible
GBP5	guanylate binding protein 5
GFI1	ZNF163;growth factor independent 1
GLIS3	ZNF515;zinc finger protein 515
GMPPB	GDP-Mannose Pyrophosphorylase B
GNA12	guanine nucleotide binding protein (G protein) alpha 12
GNLY	LAG2
GPR18	G Protein-Coupled Receptor 18
GPR183	EBI2;Epstein-Barr virus induced gene 2 (lymphocyte-specific G protein-coupled receptor)
GPR25	G Protein-Coupled Receptor 25

GPR35	G Protein-Coupled Receptor 35
GPR65	G Protein-Coupled Receptor 65
GPX1	Glutathione Peroxidase 1
GPX4	glutathione peroxidase 4 (phospholipid hydroperoxidase)
GSDMA	GSDM,GSDM1;gasdermin,gasdermin 1
GSDMB	GSDML;gasdermin-like
GZMA	HFSP,CTLA3;granzyme A (granzyme 1, cytotoxic T-lymphocyte-associated serine esterase 3)
GZMB	CTLA1,CSPB;granzyme B (granzyme 2, cytotoxic T-lymphocyte-associated serine esterase 1)
GZMK	granzyme K (serine protease, granzyme 3; tryptase II),granzyme K (granzyme 3; tryptase II)
HCK	hemopoietic cell kinase
HHAT	Hedgehog Acyltransferase
HHEX	PRHX
HNF4A	TCF14,MODY,MODY1
HSPA6	heat shock 70kD protein 6 (HSP70B'),heat shock 70kDa protein 6 (HSP70B')
ICAM1	Intercellular Adhesion Molecule 1
ICAM2	Intercellular Adhesion Molecule 2
ICAM3	Intercellular Adhesion Molecule 3
ICAM4	intercellular adhesion molecule 4
ICAM5	TLCN
ICOS	Inducible T Cell Costimulator
ICOSLG	Inducible T Cell Costimulator ligand
IDO1	IDO,INDO;indoleamine-pyrrole 2,3 dioxygenase
IFI16	Interferon Gamma Inducible Protein 16
IFI30	interferon gamma inducible protein 30
IFI35	interferon-induced protein 35
IFIH1	interferon induced with helicase C domain 1,interferon induced, with helicase C domain 1
IFIT2	IFI54,G10P2;interferon-induced protein with tetratricopeptide repeats 2
IFITM1	IFI17;interferon induced transmembrane protein 1 (9-27)
IFNA1/13	Interferon Alpha 1/13
IFNA2	Interferon Alpha 2
IFNAR1	IFNAR;interferon (alpha, beta and omega) receptor 1
IFNAR2	IFNABR;interferon (alpha, beta and omega) receptor 2
IFNB1	IFNB;interferon, beta 1, fibroblast
IFNG	Interferon Gamma
IFNGR1	Interferon Gamma Receptor 1
IFNGR2	IFNGT1;interferon gamma receptor 2 (interferon gamma transducer 1)
IFNL1	IL29;interleukin 29,interleukin 29 (interferon, lambda 1)
IFNL2/3	IL28B;interleukin 28B,interleukin 28B (interferon, lambda 3)
IFNLR1	IL28RA;interleukin 28 receptor, alpha,interleukin 28 receptor, alpha (interferon, lambda receptor),interferon, lambda receptor 1
IGFBP1	IBP1;insulin-like growth factor binding protein 1
IGFBP3	insulin-like growth factor binding protein 3

IKBKB	inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase beta
IKBKE	inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase epsilon
IKBKG	IP2,IP1;incontinentia pigmenti,inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase gamma
IKZF1	ZNFN1A1;zinc finger protein, subfamily 1A, 1 (Ikaros),IKAROS family zinc finger 1 (Ikaros)
IKZF2	ZNFN1A2;zinc finger protein, subfamily 1A, 2 (Helios),IKAROS family zinc finger 2 (Helios)
IKZF3	ZNFN1A3;zinc finger protein, subfamily 1A, 3 (Aiolos),IKAROS family zinc finger 3 (Aiolos)
IKZF4	ZNFN1A4;zinc finger protein, subfamily 1A, 4 (Eos),IKAROS family zinc finger 4 (Eos)
IL10	interleukin 10
IL10RA	IL10R;interleukin 10 receptor, alpha
IL10RB	CRFB4,D21S58,D21S66;interleukin 10 receptor, beta
IL11RA	interleukin 11 receptor, alpha
IL12A	NKSF1;interleukin 12A (natural killer cell stimulatory factor 1, cytotoxic lymphocyte maturation factor 1, p35)
IL12B	NKSF2;interleukin 12B (natural killer cell stimulatory factor 2, cytotoxic lymphocyte maturation factor 2, p40)
IL12RB1	IL12RB;interleukin 12 receptor, beta 1
IL12RB2	interleukin 12 receptor, beta 2
IL13	interleukin 13
IL13RA1	interleukin 13 receptor, alpha 1
IL15	interleukin 15
IL15RA	interleukin 15 receptor, alpha
IL16	interleukin 16 (lymphocyte chemoattractant factor)
IL17A	CTLA8,IL17;interleukin 17 (cytotoxic T-lymphocyte-associated serine esterase 8)
IL17B	interleukin 17B
IL17F	interleukin 17F
IL18	interleukin 18 (interferon-gamma-inducing factor)
IL18R1	interleukin 18 receptor 1
IL18RAP	Interleukin 18 Receptor Accessory Protein
IL1A	interleukin 1A
IL1B	interleukin 1B
IL1R1	IL1R,IL1RA;interleukin 1 receptor, type I
IL1R2	IL1RB;interleukin 1 receptor, type II
IL1RAP	Interleukin 1 Receptor Accessory Protein
IL1RL1	Interleukin 1 Receptor Like 1
IL1RN	Interleukin 1 Receptor Antagonist
IL2	Interleukin 2
IL20RB	FNDC6;fibronectin type III domain containing 6,interleukin 20 receptor beta,interleukin 20 receptor beta subunit
IL21	Interleukin 21
IL21R	Interleukin 21 recepter
IL22	Interleukin 22
IL23A	interleukin 23, alpha subunit p19
IL23R	Interleukin 23 receptor

IL24	Interleukin 24
IL27	IL30;interleukin 30
IL2RA	IL2R,IDDM10;insulin-dependent diabetes mellitus 10,interleukin 2 receptor, alpha
IL2RB	IL15RB;interleukin 15 receptor, beta,interleukin 2 receptor, beta
IL2RG	SCIDX1,IMD4,CIDX;severe combined immunodeficiency,combined immunodeficiency, X-linked,interleukin 2 receptor, gamma
IL3	interleukin 3 (colony-stimulating factor, multiple)
IL31	Interleukin 31
IL32	Interleukin 32
IL33	C9orf26;chromosome 9 open reading frame 26 (NF-HEV)
IL4	Interleukin 4
IL4R	Interleukin 4 receptor
IL5	interleukin 5 (colony-stimulating factor, eosinophil)
IL6	IFNB2;interleukin 6 (interferon, beta 2)
IL6R	Interleukin 6 receptor
IL6ST	interleukin 6 signal transducer (gp130, oncostatin M receptor)
IL7	Interleukin 7
IL7R	Interleukin 7 receptor
IL9	Interleukin 9
ILF3	interleukin enhancer binding factor 3, 90kD,interleukin enhancer binding factor 3, 90kDa
INPP5B	inositol polyphosphate-5-phosphatase, 75kD,inositol polyphosphate-5-phosphatase, 75kDa
INPP5D	inositol polyphosphate-5-phosphatase, 145kD,inositol polyphosphate-5-phosphatase, 145kDa
INPP5E	JBTS1;Joubert syndrome 1,inositol polyphosphate-5-phosphatase, 72 kDa
IP6K1	IHPK1;inositol hexaphosphate kinase 1
IP6K2	IHPK2;inositol hexaphosphate kinase 2
IPMK	Inositol Polyphosphate Multikinase
IQCB1	IQ calmodulin-binding motif containing 1
IQGAP1	IQ Motif Containing GTPase Activating Protein 1
IRAK1	Interleukin 1 Receptor Associated Kinase 1
IRAK2	Interleukin 1 Receptor Associated Kinase 2
IRAK3	interleukin-1 receptor-associated kinase 3
IRAK4	Interleukin 1 Receptor Associated Kinase 4
IRF1	Interferon Regulatory Factor 1
IRF3	Interferon Regulatory Factor 2
IRF4	Interferon Regulatory Factor 4
IRF5	Interferon Regulatory Factor 5
IRF7	Interferon Regulatory Factor 7
IRF8	ICSBP1;interferon consensus sequence binding protein 1
ITGA2B	GP2B;integrin, alpha 2b (platelet glycoprotein IIb of IIb/IIIa complex, antigen CD41B),integrin, alpha 2b (platelet glycoprotein IIb of IIb/IIIa complex, antigen CD41)
ITGA4	CD49D;integrin, alpha 4 (antigen CD49D, alpha 4 subunit of VLA-4 receptor)
ITGA5	FNRA;integrin, alpha 5 (fibronectin receptor, alpha polypeptide)

ITGA6	integrin, alpha 6
ITGAE	integrin, alpha E (antigen CD103, human mucosal lymphocyte antigen 1; alpha polypeptide)
ITGAL	CD11A;integrin, alpha L (antigen CD11A (p180), lymphocyte function-associated antigen 1; alpha polypeptide)
ITGAM	CR3A,CD11B;integrin, alpha M (complement component receptor 3, alpha; also known as CD11b (p170), macrophage antigen alpha polypeptide),integrin, alpha M (complement component 3 receptor 3 subunit)
ITGAX	CD11C;integrin, alpha X (antigen CD11C (p150), alpha polypeptide),integrin, alpha X (complement component 3 receptor 4 subunit)
ITGB1	FNRB,MSK12,MDF2;integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29 includes MDF2, MSK12)
ITGB2	CD18,MFI7;integrin, beta 2 (antigen CD18 (p95), lymphocyte function-associated antigen 1; macrophage antigen 1 (mac-1) beta subunit),integrin, beta 2 (complement component 3 receptor 3 and 4 subunit)
ITGB7	integrin, beta 7
ITGB8	integrin, beta 8
ITIH4	ITIHL1;inter-alpha (globulin) inhibitor H4 (plasma Kallikrein-sensitive glycoprotein)
ITPKA	inositol 1,4,5-trisphosphate 3-kinase A
JAK1	Janus Kinase 1
JAK2	Janus Kinase 2
JAK3	Janus Kinase 3
JAZF1	JAZF Zinc Finger 1
JRKL	jerky (mouse) homolog-like,
JUN	v-jun avian sarcoma virus 17 oncogene homolog,jun oncogene
JUNB	JunB Proto-Oncogene, AP-1 Transcription Factor Subunit
KCNJ2	potassium inwardly-rectifying channel, subfamily J, member 2
KEAP1	Kelch Like ECH Associated Protein 1
KIAA1841	
KIF21B	Kinesin Family Member 21B
KIR2DL1	killer cell immunoglobulin-like receptor, two domains, long cytoplasmic tail, 1
KLF4	Kruppel-like factor 4 (gut)
LAG3	lymphocyte-activation gene 3
LAMP3	lysosomal-associated membrane protein 3
LAT	linker for activation of T cells
LBH	limb bud and heart development homolog (mouse)
LCE3B	Late Cornified Envelope 3B
LCE3D	SPRL6B,SPRL6A;small proline rich-like (epidermal differentiation complex) 6B
LCK	lymphocyte-specific protein tyrosine kinase
LCP2	lymphocyte cytosolic protein 2
LEF1	lymphoid enhancer-binding factor 1
LGALS3	LGALS2;lectin, galactoside-binding, soluble, 3
LGALS9	lectin, galactoside-binding, soluble, 9
LIF	leukemia inhibitory factor
LILRB4	leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 4
LIME1	Lck Interacting Transmembrane Adaptor 1
LITAF	lipopolysaccharide-induced TNF factor

LNPEP	leucyl/cystinyl aminopeptidase
LPXN	Leupaxin
LSP1	Lymphocyte Specific Protein 1
LST1	Leukocyte Specific Transcript 1
LTA	TNFB;lymphotoxin alpha (TNF superfamily, member 1)
LTB	TNFC
LTB4R	P2RY7,GPR16,CMKRL1
LTB4R2	Leukotriene B4 Receptor 2
LTBR	Lymphotoxin Beta Receptor
LYST	CHS1;Chediak-Higashi syndrome 1
MAF	v-maf avian musculoaponeurotic fibrosarcoma oncogene homolog
MALT1	MLT;mucosa associated lymphoid tissue lymphoma translocation gene 1
MAML2	mastermind (Drosophila)-like 2,mastermind-like 2 (Drosophila)
MAMSTR	MEF2 Activating Motif And SAP Domain Containing Transcriptional Regulator
MANBA	mannosidase, beta A, lysosomal
MAP3K14	Mitogen-Activated Protein Kinase Kinase Kinase 14
MAP3K7	TAK1
MAP3K8	COT,ESTF
MAP4K1	Mitogen-Activated Protein Kinase Kinase Kinase Kinase 1
MAP4K2	Mitogen-Activated Protein Kinase Kinase Kinase Kinase 2
MAP4K4	Mitogen-Activated Protein Kinase Kinase Kinase Kinase 4
MAPK1	PRKM2,PRKM1
MAPK11	PRKM11
MAPK14	CSPB1,CSBP1,CSBP2
MAPK3	PRKM3
MAPK8	PRKM8
MAPKAPK2	MAPK Activated Protein Kinase 2
MARCO	Macrophage Receptor With Collagenous Structure
MBD2	Methyl-CpG Binding Domain Protein 2
MBP	Myelin Basic Protein
MECP2	RTT,MRX16,MRX79;mental retardation, X-linked 16,mental retardation, X-linked 79,Rett syndrome,methyl CpG binding protein 2 (Rett syndrome),methyl CpG binding protein 2
MED1	TRIP2,PPARGBP,PPARBP;PPAR binding protein
MERTK	c-mer proto-oncogene tyrosine kinase
METTL1	C12orf1;methyltransferase-like 1
MICB	MHC Class I Polypeptide-Related Sequence B
MIF	GLIF;macrophage migration inhibitory factor (glycosylation-inhibiting factor)
MLH3	mutL (E. coli) homolog 3,mutL homolog 3 (E. coli)
MLX	TCFL4;transcription factor-like 4,MAX-like protein X
MMEL1	MMEL2;membrane metallo-endopeptidase-like 2
MMP3	STMY1,STMY;matrix metalloproteinase 3 (stromelysin 1, progelatinase),stromelysin 1
MMP9	CLG4B;matrix metalloproteinase 9 (gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase)

MPV17L2	MPV17 Mitochondrial Inner Membrane Protein Like 2
MR1	HLALS;major histocompatibility complex, class I-like sequence
MST1	D3F15S2,HGFL,DNF15S2;hepatocyte growth factor-like
MST1R	RON,PTK8;PTK8 protein tyrosine kinase 8
MSTO1	misato homolog 1 (Drosophila)
MTF1	Metal Regulatory Transcription Factor 1
MTMR3	Myotubularin Related Protein 3
MUC1	PUM,MCKD1;mucin 1, transmembrane,medullary cystic kidney disease 1 (autosomal dominant)
MX1	myxovirus (influenza) resistance 1
MYC	v-myc avian myelocytomatosis viral oncogene homolog
MYD88	myeloid differentiation primary response gene (88)
MYRF	C11orf9;chromosome 11 open reading frame 9
NAA25	C12orf30;chromosome 12 open reading frame 30
NADSYN1	NAD Synthetase 1
NCF2	neutrophil cytosolic factor 2 (65kD, chronic granulomatous disease, autosomal 2)
NCOA5	Nuclear Receptor Coactivator 5
NDFIP1	Nedd4 Family Interacting Protein 1
NDUFAF1	NADH dehydrogenase (ubiquinone) complex I, assembly factor 1
NFATC1	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 1
NFATC2	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 2
NFATC3	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 3
NFIL3	IL3BP1
NFKB1	nuclear factor of kappa light polypeptide gene enhancer in B-cells 1
NFKB2	nuclear factor of kappa light polypeptide gene enhancer in B-cells 2 (p49/p100)
NFKBIA	NFKBI;nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha
NFKBIE	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, epsilon
NFKBIL1	NFKBIL;nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor-like 1
NFKBIZ	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, zeta
NKX2-3	NKX2C;NK-2 (Drosophila) homolog C,NK2 transcription factor related, locus 3 (Drosophila)
NLRP2	NALP2;NACHT, leucine rich repeat and PYD containing 2
NLRP3	C1orf7,CIAS1;cold autoinflammatory syndrome 1
NOD1	CARD4;caspase recruitment domain family, member 4
NOD2	IBD1,CARD15;caspase recruitment domain family, member 15
NOS2	NOS2A;nitric oxide synthase 2A (inducible, hepatocytes),nitric oxide synthase 2, inducible
NOTCH1	TAN1;Notch (Drosophila) homolog 1 (translocation-associated),Notch homolog 1, translocation-associated (Drosophila)
NOTCH2	Notch (Drosophila) homolog 2,Notch homolog 2 (Drosophila)
NPEPPS	Aminopeptidase Puromycin Sensitive
NRP1	Neuropilin 1
NUPR1	Nuclear Protein 1, Transcriptional Regulator
NUSAP1	Nucleolar And Spindle Associated Protein 1

NXPE1	FAM55A;family with sequence similarity 55, member A,neurexophilin and PC-esterase domain family, member 1
NXPE4	neurexophilin and PC-esterase domain family, member 4
ODF3B	
OR5B21	olfactory receptor, family 5, subfamily B, member 21
ORMDL3	ORM1 (S. cerevisiae)-like 3,ORM1-like 3 (S. cerevisiae)
OSM	Oncostatin M
OSMR	Oncostatin M Receptor
P2RY10	purinergic receptor P2Y, G-protein coupled, 10,purinergic receptor P2Y10
PADI4	PADI5;peptidyl arginine deiminase, type V,peptidyl arginine deiminase, type IV
PARK7	Parkinson disease (autosomal recessive, early onset) 7,parkinson protein 7
PAX5	paired box gene 5 (B-cell lineage specific activator protein),paired box gene 5 (B-cell lineage specific activator)
PDCD1	SLEB2;systemic lupus erythematosus susceptibility 2
PDCD1LG2	Programmed Cell Death 1 Ligand 2
PDCD2	Programmed Cell Death 2
PDLIM4	PDZ And LIM Domain 4
PF4	platelet factor 4
PFKFB4	6-Phosphofructo-2-Kinase/Fructose-2,6-Biphosphatase 4
PHACTR2	C6orf56;chromosome 6 open reading frame 56
PHGDH	Phosphoglycerate Dehydrogenase
PHRF1	RNF221
PHTF1	PHTF
PIGR	Polymeric Immunoglobulin Receptor
PITPNM2	phosphatidylinositol transfer protein, membrane associated 2
PLAU	Plasminogen Activator, Urokinase
PLAUR	Plasminogen Activator, Urokinase Receptor
PLCG2	phospholipase C, gamma 2 (phosphatidylinositol-specific)
PLCH2	PLCL4;phospholipase C-like 4,phospholipase C, eta 2
PLCL1	PLCE;phospholipase C, epsilon,phospholipase C-like 1,phospholipase C like 1
PLCL2	PLCE2;phospholipase C, epsilon 2,phospholipase C-like 2
PLD2	Phospholipase D2
PLEK	Pleckstrin
PLEKHG5	pleckstrin homology domain containing, family G (with RhoGef domain) member 5
PLTP	Phospholipid Transfer Protein
PMPCA	INPP5E;inositol polyphosphate-5-phosphatase, 72 kD,peptidase (mitochondrial processing) alpha
PNKD	paroxysmal nonkinesiogenic dyskinesia
PNMT	PENT
POLI	RAD3OB,RAD30B;polymerase (DNA directed) iota,polymerase (DNA) iota
POU3F1	OTF6;POU domain class 3, transcription factor 1
PPARG	peroxisome proliferative activated receptor, gamma,peroxisome proliferator-activated receptor gamma
PPIL4	peptidylprolyl isomerase (cyclophilin)-like 4
PRDM1	BLIMP1;PR domain containing 1, with ZNF domain,PR domain 1

PRDX5	Peroxiredoxin 5
PRF1	perforin 1 (pore forming protein)
PRKCB	PRKCB2,PKCB,PRKCB1;protein kinase C, beta 1,protein kinase C, beta
PRKCD	protein kinase C, delta
PRKCH	PRKCL;protein kinase C, eta
PRKCQ	protein kinase C, theta
PRM3	Protamine 3
PROCR	protein C receptor, endothelial
PRSS53	Serine Protease 53
PTGDR2	GPR44;G protein-coupled receptor 44
PTGER4	prostaglandin E receptor 4 (subtype EP4)
PTGS2	prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)
PTK2	PTK2 protein tyrosine kinase 2
PTPN11	NS1;Noonan syndrome 1
PTPN2	PTPT
PTPN22	PTPN8;protein tyrosine phosphatase, non-receptor type 8,protein tyrosine phosphatase, non-receptor type 22 (lymphoid)
PTPN6	Protein Tyrosine Phosphatase Non-Receptor Type 6
PTPRK	Protein Tyrosine Phosphatase Receptor Type K
PVT1	pvt-1 (murine) oncogene homolog, MYC activator,Pvt1 oncogene homolog (mouse)
PXK	PX domain containing serine/threonine kinase
RABEP2	Ras Interacting Protein 1
RAC2	ras-related C3 botulinum toxin substrate 2 (rho family, small GTP binding protein Rac2)
RAD51B	RAD51L1;RAD51 (S. cerevisiae)-like 1,RAD51-like 1 (S. cerevisiae),RAD51 homolog B (S. cerevisiae)
RAF1	v-raf-1 murine leukemia viral oncogene homolog 1
RASGRP1	RAS guanyl releasing protein 1 (calcium and DAG-regulated)
RASIP1	
RASSF5	Ras association (RalGDS/AF-6) domain family member 5
RAVER1	ribonucleoprotein, PTB-binding 1
RBM17	RNA Binding Motif Protein 17
RBPJ	IGKJRB1,RBPSUH;recombining binding protein suppressor of hairless (Drosophila)
RCAN1	DSCR1;Down syndrome critical region gene 1
REL	v-rel avian reticuloendotheliosis viral oncogene homolog
RELA	NFKB3;nuclear factor of kappa light polypeptide gene enhancer in B-cells 3
RELB	v-rel avian reticuloendotheliosis viral oncogene homolog B (nuclear factor of kappa light polypeptide gene enhancer in B-cells 3)
REV3L	REV3 (yeast homolog)-like, catalytic subunit of DNA polymerase zeta,REV3-like,REV3-like, polymerase (DNA directed),
RGS1	IER1;regulator of G-protein signalling 1,regulator of G-protein signaling 1
RGS14	regulator of G-protein signalling 14,regulator of G-protein signaling 14
RIMBP3	RIMS Binding Protein 3
RIPK1	receptor (TNFRSF)-interacting serine-threonine kinase 1
RIPK2	receptor-interacting serine-threonine kinase 2

RIT1	RIT;Ric (Drosophila)-like, expressed in many tissues
RMI2	C16orf75;chromosome 16 open reading frame 75,RMI2, RecQ mediated genome instability 2, homolog (S. cerevisiae)
RNASET2	Ribonuclease T2
RNF114	ZNF313;zinc finger protein 313
RNF186	Ring Finger Protein 186
RNLS	C10orf59;chromosome 10 open reading frame 59
RORA	RAR-related orphan receptor A
RORC	RAR-related orphan receptor C
RPS6KA2	ribosomal protein S6 kinase, 90kD, polypeptide 2
RPS6KA4	ribosomal protein S6 kinase, 90kD, polypeptide 4
RPS6KB1	STK14A;ribosomal protein S6 kinase, 70kD, polypeptide 1
RTEL1	C20orf41;chromosome 20 open reading frame 41
RTKN2	PLEKHK1;pleckstrin homology domain containing, family K member 1
RUNX1	AML1,CBFA2;acute myeloid leukemia 1,runt-related transcription factor 1
RUNX3	CBFA3;runt-related transcription factor 3
S100A8	CAGA,CFAG;S100 calcium-binding protein A8 (calgranulin A),S100 calcium binding protein A8 (calgranulin A)
S100A9	CAGB,CFAG;S100 calcium-binding protein A9 (calgranulin B),S100 calcium binding protein A9 (calgranulin B)
S1PR1	EDG1;endothelial differentiation, sphingolipid G-protein-coupled receptor, 1
SBNO2	KIAA0963;KIAA0963,strawberry notch homolog 2 (Drosophila)
SCAMP3	C1orf3
SDCCAG3	ENTR1, Endosome Associated Trafficking Regulator 1
SEC16A	KIAA0310;KIAA0310,SEC16 homolog A (S. cerevisiae)
SELE	ELAM1,ELAM;endothelial adhesion molecule 1
SELL	LYAM1,LNHR;lymphocyte adhesion molecule 1
SELPLG	Selectin P Ligand
SERINC3	TDE1;tumor differentially expressed 1
SERPING1	serpin peptidase inhibitor, clade G (C1 inhibitor), member 1
SH2B1	SH2B Adaptor Protein 1
SH2B3	SH2B Adaptor Protein 3
SIGIRR	single immunoglobulin and toll-interleukin 1 receptor (TIR) domain
SKP1	SKP1A;S-phase kinase-associated protein 1A (p19A)
SLAMF1	SLAM;signaling lymphocytic activation molecule
SLAMF7	SLAM Family Member 7
SLAMF8	SLAM Family Member 8
SLC10A4	solute carrier family 10, member 4
SLC11A1	LSH,NRAMP,NRAMP1;solute carrier family 11 (proton-coupled divalent metal ion transporter), member 1
SLC15A2	solute carrier family 15 (H+/peptide transporter), member 2,solute carrier family 15 (oligopeptide transporter), member 2
SLC15A4	solute carrier family 15, member 4,solute carrier family 15 (oligopeptide transporter), member 4
SLC22A5	CDSP;solute carrier family 22 (organic cation/carnitine transporter), member 5
SLC2A4RG	SLC2A4 Regulator
SLC30A7	solute carrier family 30 (zinc transporter), member 7

SLC44A2	solute carrier family 44 (choline transporter), member 2
SLC45A1	DNB5;deleted in neuroblastoma 5,solute carrier family 45, member 1
SLC9A8	solute carrier family 9 (sodium/hydrogen exchanger), isoform 8
SMAD3	MADH3;MAD, mothers against decapentaplegic homolog 3 (Drosophila),SMAD, mothers against DPP homolog 3 (Drosophila)
SMAD5	MADH5;MAD, mothers against decapentaplegic homolog 5 (Drosophila),SMAD, mothers against DPP homolog 5 (Drosophila)
SMAD7	MADH8,MADH7;MAD, mothers against decapentaplegic homolog 7 (Drosophila),SMAD, mothers against DPP homolog 7 (Drosophila)
SMG7	C1orf16;chromosome 1 open reading frame 16,smg-7 homolog, nonsense mediated mRNA decay factor (C. elegans)
SMIM20	C4orf52;chromosome 4 open reading frame 52
SMURF1	SMAD Specific E3 Ubiquitin Protein Ligase 1
SNAPC4	small nuclear RNA activating complex, polypeptide 4, 190kD,small nuclear RNA activating complex, polypeptide 4, 190kDa
SNX32	SNX6B;sorting nexin 6B
SOCS1	Suppressor Of Cytokine Signaling 1
SOCS3	Suppressor Of Cytokine Signaling 3
SOX8	SRY (sex determining region Y)-box 8,SRY box 8
SP140	SP140 Nuclear Body Protein
SPHK2	Sphingosine Kinase 2
SPP1	BNSP,OPN;osteopontin,bone sialoprotein I
SPRED2	Sprouty Related EVH1 Domain Containing 2
SPRY4	sprouty homolog 4 (Drosophila)
STAT1	signal transducer and activator of transcription 1, 91kD,signal transducer and activator of transcription 1, 91kDa
STAT2	signal transducer and activator of transcription 2, 113kD,signal transducer and activator of transcription 2, 113kDa
STAT3	signal transducer and activator of transcription 3 (acute-phase response factor)
STAT4	Signal Transducer And Activator Of Transcription 4
STAT5A	Signal Transducer And Activator Of Transcription 5A
STAT5B	Signal Transducer And Activator Of Transcription 5B
STAT6	signal transducer and activator of transcription 6, interleukin-4 induced
STMN3	stathmin-like 3
SULT1A1	STP,STP1;sulfotransferase family, cytosolic, 1A, phenol-preferring, member 1
SULT1A2	STP2;sulfotransferase family, cytosolic, 1A, phenol-preferring, member 2
SYK	spleen tyrosine kinase
SYNGR1	Synaptogyrin 1
TAB1	MAP3K7IP1;mitogen-activated protein kinase kinase kinase 7 interacting protein 1,TGF-beta activated kinase 1/MAP3K7 binding protein 1
TAGAP	T-cell activation GTPase activating protein
TBK1	TANK-binding kinase 1
TBKBP1	TBK1 Binding Protein 1
TBX21	T-Box Transcription Factor 21
TCF7	transcription factor 7 (T-cell specific, HMG-box)
TEC	Tec Protein Tyrosine Kinase
TET2	KIAA1546;KIAA1546,tet oncogene family member 2
TGFB1	TGFB,DPD1;transforming growth factor, beta 1

TGFBI	CSD3,LCD1,CSD1,CSD2;transforming growth factor, beta-induced, 68kD,transforming growth factor, beta-induced, 68kDa
TGFBR1	MSSE,ESS1;multiple self-healing squamous epithelioma,transforming growth factor beta receptor I
TGFBR2	MFS2;transforming growth factor, beta receptor II (70/80kDa),transforming growth factor beta receptor II
THADA	thyroid adenoma associated
THEMIS	C6orf207,C6orf190,TSEPA;chromosome 6 open reading frame 207,chromosome 6 open reading frame 190,thymocyte selection pathway associated
TICAM1	toll-like receptor adaptor molecule 1
TIGIT	VSIG9,VSTM3;V-set and immunoglobulin domain containing 9,V-set and transmembrane domain containing 3,T cell immunoreceptor with Ig and ITIM domains
TIMMDC1	C3orf1;chromosome 3 open reading frame 1
TIRAP	Toll-interleukin 1 receptor (TIR) domain-containing adaptor protein
TLE3	transducin-like enhancer of split 3 (E(sp1) homolog, Drosophila)
TLR1	toll-like receptor 1
TLR2	toll-like receptor 2
TLR3	toll-like receptor 3
TLR4	toll-like receptor 4
TLR5	SLEB1;systemic lupus erythematosus susceptibility 1,toll-like receptor 5
TLR7	toll-like receptor 7
TLR8	toll-like receptor 8
TLR9	toll-like receptor 9
TMBIM1	Transmembrane BAX Inhibitor Motif Containing 1
TMEM50B	C21orf4;chromosome 21 open reading frame 4
TNC	HXB,DFNA56;hexabrachion (tenascin C, cytotactin),deafness, autosomal dominant 56
TNF	TNFA;tumor necrosis factor (TNF superfamily, member 2)
TNFAIP3	tumor necrosis factor, alpha-induced protein 3
TNFAIP6	tumor necrosis factor, alpha-induced protein 6
TNFRSF10C	tumor necrosis factor receptor superfamily, member 10c, decoy without an intracellular domain
TNFRSF11A	PDB2,LOH18CR1;tumor necrosis factor receptor superfamily, member 11a, activator of NFKB,Paget disease of bone 2,loss of heterozygosity, 18, chromosomal region 1,tumor necrosis factor receptor superfamily, member 11a, NFKB activator
TNFRSF13B	tumor necrosis factor receptor superfamily, member 13B
TNFRSF13C	tumor necrosis factor receptor superfamily, member 13C
TNFRSF14	tumor necrosis factor receptor superfamily, member 14 (herpesvirus entry mediator),tumor necrosis factor receptor superfamily, member 14
TNFRSF17	BCMA;tumor necrosis factor receptor superfamily, member 17
TNFRSF18	tumor necrosis factor receptor superfamily, member 18
TNFRSF1A	TNFR1;tumor necrosis factor receptor superfamily, member 1A
TNFRSF1B	TNFR2;tumor necrosis factor receptor superfamily, member 1B
TNFRSF4	TXGP1L;tumor necrosis factor receptor superfamily, member 4
TNFRSF6B	tumor necrosis factor receptor superfamily, member 6b, decoy
TNFRSF8	CD30,D1S166E;tumor necrosis factor receptor superfamily, member 8
TNFRSF9	ILA;tumor necrosis factor receptor superfamily, member 9

TNFSF10	tumor necrosis factor (ligand) superfamily, member 10
TNFSF11	tumor necrosis factor (ligand) superfamily, member 11
TNFSF12	tumor necrosis factor (ligand) superfamily, member 12
TNFSF13B	TNFSF20;tumor necrosis factor (ligand) superfamily, member 13b
TNFSF14	tumor necrosis factor (ligand) superfamily, member 14
TNFSF15	tumor necrosis factor (ligand) superfamily, member 15
TNFSF18	tumor necrosis factor (ligand) superfamily, member 18
TNFSF4	TXGP1;tax-transcriptionally activated glycoprotein 1, 34kD,tumor necrosis factor (ligand) superfamily, member 4
TNFSF8	CD30LG;tumor necrosis factor (ligand) superfamily, member 8
TNIP1	TNFAIP3 Interacting Protein 1
TNNI2	AMCD2B;troponin I, skeletal, fast,arthrogryposis multiplex congenita, distal, type 2B,troponin I type 2 (skeletal, fast)
TNP2	transition protein 2 (during histone to protamine replacement)
TNPO3	LGMD1F;limb girdle muscular dystrophy 1F (autosomal dominant)
TOLLIP	Toll Interacting Protein
TPD52	Tumor Protein D52
TRADD	TNFRSF1A-associated via death domain
TRAF1	TNF Receptor Associated Factor 1
TRAF2	TNF Receptor Associated Factor 2
TRAF3	TNF Receptor Associated Factor 3
TRAF3IP2	C6orf4,C6orf5,C6orf6,C6orf2;chromosome 6 open reading frame 5,chromosome 6 open reading frame 2
TRAF4	TNF Receptor Associated Factor 4
TRAF5	TNF Receptor Associated Factor 5
TRAF6	TNF receptor-associated factor 6, E3 ubiquitin protein ligase
TRIB1	tribbles homolog 1 (Drosophila)
TRPT1	TRNA Phosphotransferase 1
TSFM	Ts Translation Elongation Factor, Mitochondrial
TSPAN14	TM4SF14;transmembrane 4 superfamily member 14
TSPAN33	Tetraspanin 33
TTYH3	tweety homolog 3 (Drosophila)
TUBD1	tubulin, delta 1
TXK	PTK4;PTK4 protein tyrosine kinase 4
TXNDC11	Thioredoxin Domain Containing 11
TYK2	Tyrosine Kinase 2
UBA7	UBE1L;ubiquitin-activating enzyme E1-like,ubiquitin-like modifier activating enzyme 7
UBASH3A	Ubiquitin Associated And SH3 Domain Containing A
UBE2D1	SFT;stimulator of Fe transport,ubiquitin-conjugating enzyme E2D 1 (UBC4/5 homolog, yeast),ubiquitin-conjugating enzyme E2D 1
UBE2E3	ubiquitin-conjugating enzyme E2E 3 (homologous to yeast UBC4/5),ubiquitin-conjugating enzyme E2E 3 (UBC4/5 homolog, yeast),ubiquitin-conjugating enzyme E2E 3
UBE2L3	ubiquitin-conjugating enzyme E2L 3
UBQLN4	C1orf6;chromosome 1 open reading frame 6

UCN	Urocortin
UCN2	Urocortin 2
UHRF1BP1	C6orf107;chromosome 6 open reading frame 107
UQCC1	C20orf44,UQCC;chromosome 20 open reading frame 44,ubiquinol-cytochrome c reductase complex chaperone
USF1	Upstream Transcription Factor 1
USP4	UNP;ubiquitin specific peptidase 4 (proto-oncogene)
VAMP3	Vesicle Associated Membrane Protein 3
VCAM1	Vascular Cell Adhesion Molecule 1
VDR	vitamin D (1,25- dihydroxyvitamin D3) receptor
VMP1	TMEM49;transmembrane protein 49
VSIR	C10orf54;chromosome 10 open reading frame 54
WDFY4	C10orf64;chromosome 10 open reading frame 64
WWOX	WW domain-containing oxidoreductase
XBP1	XBP2
YDJC	YdjC homolog (bacterial)
ZAP70	SRK;zeta-chain (TCR) associated protein kinase (70 kD),zeta-chain (TCR) associated protein kinase 70kDa,zeta chain of T cell receptor associated protein kinase 70kDa,zeta chain of T cell receptor associated protein kinase 70
ZBTB16	ZNF145;zinc finger protein 145 (Kruppel-like, expressed in promyelocytic leukemia)
ZBTB46	ZNF340,BTBD4;BTB (POZ) domain containing 4
ZC2HC1A	C8orf70,FAM164A;chromosome 8 open reading frame 70,family with sequence similarity 164, member A
ZEB1	TCF8,PPCD3;transcription factor 8 (represses interleukin 2 expression),posterior polymorphous corneal dystrophy 3
ZFP36L1	BRF1;zinc finger protein, C3H type, 36-like 1,zinc finger protein 36, C3H type-like 1
ZFP90	zinc finger protein 90 homolog (mouse)
ZGPAT	KIAA1847;KIAA1847,zinc finger CCCH-type with G-patch domain
ZMIZ1	RAI17;retinoic acid induced 17
ZNF767P	ZNF767;zinc finger family member 767
ZNF831	C20orf174;chromosome 20 open reading frame 174

Chapter 8

BIBLIOGRAPHY: PART I

Aarts, S. A. B. M., Seijkens, T. T. P., van Dorst, K. J. F., Dijkstra, C. D., Kooij, G., & Lutgens, E. (2017). The CD40-CD40L dyad in experimental autoimmune encephalomyelitis and multiple sclerosis. *Frontiers in Immunology*, *8*(DEC). https://doi.org/10.3389/fimmu.2017.01791

Aggarwal, S., Ghilardi, N., Xie, M. H., De Sauvage, F. J., & Gurney, A. L. (2003). Interleukin-23 promotes a distinct CD4 T cell activation state characterized by the production of interleukin-17. *Journal of Biological Chemistry*, *278*(3), 1910–1914. https://doi.org/10.1074/jbc.M207577200

Akassou, A., & Bakri, Y. (2018). Does HLA-B27 status influence ankylosing spondylitis phenotype? *Clinical Medicine Insights: Arthritis and Musculoskeletal Disorders*, *11*(Figure 1). https://doi.org/10.1177/1179544117751627

Akgul, O., & Ozgocmen, S. (2011). Classification criteria for spondyloarthropathies. *World Journal of Orthopaedics*, *2*(12), 107–115. https://doi.org/10.5312/wjo.v2.i12.107

Al-Mossawi, H., Yager, N., Taylor, C. A., Lau, E., Danielli, S., de Wit, J., … Fairfax, B. P. (2019). Context-specific regulation of surface and soluble IL7R expression by an autoimmune risk allele. *Nature Communications*, *10*(1), 1–11. https://doi.org/10.1038/s41467-019-12393-1

Al-Mossawi, M. H., Chen, L., Fang, H., Ridley, A., de Wit, J., Yager, N., … Bowness, P. (2017). Unique transcriptome signatures and GM-CSF expression in lymphocytes from patients with spondyloarthritis. *Nature Communications*, *8*(1), 1510. https://doi.org/10.1038/s41467-017-01771-2

Amor, B., Dougados, M., & Mijiyawa, M. (1990). [Criteria of the classification of spondylarthropathies]. *Revue Du Rhumatisme et Des Maladies Osteo-Articulaires*, *57*(2), 85–89. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2181618

Ardlie, K. G., DeLuca, D. S., Segrè, A. V., Sullivan, T. J., Young, T. R., Gelfand, E. T., … Lockhart. (2015). The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science*, *348*(6235), 648–660. https://doi.org/10.1126/science.1262110

Astle, W. J., Elding, H., Jiang, T., Allen, D., Ruklisa, D., Mann, A. L., … Sauerwein, H. P. (2016). The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. *Cell*, *167*(5), 1415-1429.e19. https://doi.org/10.1016/j.cell.2016.10.042

Auton, A., Abecasis, G. R., Altshuler, D. M., Durbin, R. M., Bentley, D. R., Chakravarti, A., … Schloss, J. A. (2015, September 30). A global reference for human genetic variation. *Nature*. Nature Publishing Group. https://doi.org/10.1038/nature15393

Azuz-Lieberman, N., Markel, G., Mizrahi, S., Gazit, R., Hanna, J., Achdout, H., … Mandelboim, O. (2005). The involvement of NK cells in ankylosing spondylitis. *International Immunology*, *17*(7), 837–845. https://doi.org/10.1093/intimm/dxh270

Balaji, K. N., Schaschke, N., Machleidt, W., Catalfamo, M., & Henkart, P. A. (2002). Surface cathepsin B protects cytotoxic lymphocytes from self-destruction after degranulation. *Journal of Experimental Medicine*, *196*(4), 493–503. https://doi.org/10.1084/jem.20011836

Bamias, G., Martin, C., Marini, M., Hoang, S., Mishina, M., Ross, W. G., … Cominelli, F. (2003). Expression, Localization, and Functional Activity of TL1A, a Novel Th1-Polarizing Cytokine in Inflammatory Bowel Disease. *The Journal of Immunology*, *171*(9), 4868–4874. https://doi.org/10.4049/jimmunol.171.9.4868

Baran, K., Ciccone, A., Peters, C., Yagita, H., Bird, P. I., Villadangos, J. A., & Trapani, J. A. (2006). Cytotoxic T lymphocytes from cathepsin B-deficient mice survive normally in vitro and in vivo after encountering and killing target cells. *Journal of Biological Chemistry*, *281*(41), 30485–30491. https://doi.org/10.1074/jbc.M602007200

Barera, G., Bonfanti, R., Viscardi, M., Bazzigaluppi, E., Calori, G., Meschi, F., … Chiumello, G. (2002). Occurrence of celiac disease after onset of type 1 diabetes: A 6-year prospective longitudinal study. *Pediatrics*, *109*(5), 833–838. https://doi.org/10.1542/peds.109.5.833

Barth, T. K., & Imhof, A. (2010). Fast signals and slow marks: the dynamics of histone modifications. *Trends in Biochemical Sciences*, *35*(11), 618–626. https://doi.org/10.1016/j.tibs.2010.05.006

Beecham, A. H., Patsopoulos, N. A., Xifara, D. K., Davis, M. F., Kemppinen, A., Cotsapas, C., … McCauley, J. L. (2013a). Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nature Genetics*, *45*(11), 1353–1362. https://doi.org/10.1038/ng.2770

Beecham, A. H., Patsopoulos, N. A., Xifara, D. K., Davis, M. F., Kemppinen, A., Cotsapas, C., … McCauley, J. L. (2013b). Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nature Genetics*, *45*(11), 1353–1362. https://doi.org/10.1038/ng.2770

Benjamin, R., & Parham, P. (1990). Guilt by association: HLA-B27 and ankylosing spondylitis. *Trends in Immunology*, *11*(C), 137–142.

Blair, H. A. (2019, March 1). Secukinumab: A Review in Ankylosing Spondylitis. *Drugs*. Springer International Publishing. https://doi.org/10.1007/s40265-019-01075-3

Blanco, P., Pitard, V., Viallard, J. F., Taupin, J. L., Pellegrin, J. L., & Moreau, J. F. (2005). Increase in activated CD8+ T lymphocytes expressing perforin and granzyme B correlates with disease activity in patients with systemic lupus erythematosus. *Arthritis and Rheumatism*, *52*(1), 201–211. https://doi.org/10.1002/art.20745

Boisson-Dupuis, S., Ramirez-Alejo, N., Li, Z., Patin, E., Rao, G., Kerner, G., … Casanova, J. L. (2018). Tuberculosis and impaired IL-23–dependent IFN- immunity in humans homozygous for a common TYK2 missense variant. *Science Immunology*, *3*(30). https://doi.org/10.1126/sciimmunol.aau8714

Bondeva, T., & Wolf, G. (2015). Role of Neuropilin-1 in Diabetic Nephropathy. *Journal of Clinical Medicine*, *4*(6), 1293–1311. https://doi.org/10.3390/jcm4061293

Bowness, P. (2015). HLA-B27. https://doi.org/10.1146/annurev-immunol-032414-112110

Bowness, P., Ridley, A., Shaw, J., Chan, A. T., Wong-Baeza, I., Fleming, M., … Kollnberger, S. (2011). Th17 Cells Expressing KIR3DL2 $^{+}$ and Responsive to HLA-B27 Homodimers Are Increased in Ankylosing Spondylitis. *The Journal of Immunology*, *186*(4), 2672–2680. https://doi.org/10.4049/jimmunol.1002653

Braun, J., Deodhar, A., Landewé, R., Baraliakos, X., Miceli-Richard, C., Sieper, J., … Van Der Heijde, D. (2018). Impact of baseline C-reactive protein levels on the response to secukinumab in ankylosing spondylitis: 3-year pooled data from two phase III studies. *RMD Open*, *4*(2), 749. https://doi.org/10.1136/rmdopen-2018-000749

Breban, M. (2006). Genetics of spondyloarthritis. *Best Practice and Research: Clinical Rheumatology*, *20*(3), 593–599. https://doi.org/10.1016/j.berh.2006.03.002

Brewerton, D. A., Hart, F. D., Nicholls, A., Caffrey, M., James, D. C. O., & Sturrock, R. D. (1973). ANKYLOSING SPONDYLITIS AND HL-A 27. *The Lancet*, *301*(7809), 904–907. https://doi.org/10.1016/S0140-6736(73)91360-3

Brinkworth, R. I., Tort, J. F., Brindley, P. J., & Dalton, J. P. (2000). Phylogenetic relationships and theoretical model of human cathepsin W (lymphopain), a cysteine proteinase from cytotoxic T lymphocytes. *International Journal of Biochemistry and Cell Biology*, *32*(3), 373–384. https://doi.org/10.1016/S1357-2725(99)00129-6

Brix, K., Dunkhorst, A., Mayer, K., & Jordans, S. (2008). Cysteine cathepsins: Cellular roadmap to different functions. *Biochimie*, *90*(2), 194–207. https://doi.org/10.1016/j.biochi.2007.07.024

Brodin, P., Jojic, V., Gao, T., Bhattacharya, S., Angel, C. J. L., Furman, D., … Davis, M. M. (2015). Variation in the human immune system is largely driven by non-heritable influences. *Cell*, *160*(1–2), 37–47. https://doi.org/10.1016/j.cell.2014.12.020

Brown, C. D., Mangravite, L. M., & Engelhardt, B. E. (2013). Integrative Modeling of eQTLs and Cis-Regulatory Elements Suggests Mechanisms Underlying Cell Type Specificity of eQTLs. *PLoS Genetics*, *9*(8). https://doi.org/10.1371/journal.pgen.1003649

Brown, J., Matutes, E., Singleton, A., Price, C., Molgaard, H., Buttle, D., & Enver, T. (1998). Lymphopain, a cytotoxic T and natural killer cell-associated cysteine proteinase. *Leukemia*, *12*(11), 1771–1781. https://doi.org/10.1038/sj.leu.2401164

Brown, M. A., Laval, S. H., Brophy, S., & Calin, A. (2000). Recurrence risk modelling of the genetic susceptibility to ankylosing spondylitis. *Annals of the Rheumatic Diseases*, *59*(11), 883–886. https://doi.org/10.1136/ard.59.11.883

Brown, Matthew A. (2011). Progress in the genetics of ankylosing spondylitis. *Briefings in Functional Genomics*, *10*(5), 249–257. https://doi.org/10.1093/bfgp/elr023

Brown, Matthew A., Kennedy, L. G., MacGregor, A. J., Darke, C., Duncan, E., Shatford, J. L., … Wordsworth, P. (1997). Susceptibility to ankylosing spondylitis in twins: The role of genes, HLA, and the environment. *Arthritis and Rheumatism*, *40*(10), 1823–1828. https://doi.org/10.1002/art.1780401015

Brown, Matthew A, Kenna, T., & Wordsworth, B. P. (2015). Genetics of ankylosing spondylitis—insights into pathogenesis. *Nature Reviews Rheumatology*, *12*(2), 1–11. https://doi.org/10.1038/nrrheum.2015.133

Brownlie, R. J., Zamoyska, R., & Salmond, R. J. (2018, July 1). Regulation of autoimmune and anti-tumour T-cell responses by PTPN22. *Immunology*. Blackwell Publishing Ltd. https://doi.org/10.1111/imm.12919

Brusko, T. m., Putnam, A. l., & Bluestone, J. a. (2009). Erratum: Human regulatory T cells: Role in autoimmune disease and therapeutic opportunities (Immunological Reviews (2008) 223 (371390)). *Immunological Reviews*, *229*(1), 388. https://doi.org/10.1111/j.1600-065X.2009.00788.x

Buhling, F., Kellner, U., Guenther, D., Kahl, S., Brömme, D., Weber, E., … Wex, T. (2002). Characterization of novel anti-cathepsin W antibodies and cellular distribution of cathepsin W in the gastrointestinal tract. *Biological Chemistry*, *383*(7–8), 1285–1289. https://doi.org/10.1515/BC.2002.144

Burton, P. R., Clayton, D. G., Cardon, L. R., Craddock, N., Deloukas, P., Duncanson, A., … Brown, M. J. M. A. M. J. M. A. (2007). Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nature Genetics*, *39*(11), 1329–1337. https://doi.org/10.1038/ng.2007.17

Bush, W. S., & Moore, J. H. (2012). Chapter 11: Genome-Wide Association Studies. *PLoS Computational Biology*, *8*(12). https://doi.org/10.1371/journal.pcbi.1002822

Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., … Marchini, J. (2017). Genome-wide genetic data on ~500,000 UK Biobank participants. *BioRxiv*, 166298. https://doi.org/10.1101/166298

Caffrey, M. F. P., & James, D. C. O. (1973). Human lymphocyte antigen association in ankylosing spondylitis. *Nature*, *242*(5393), 121. https://doi.org/10.1038/242121a0

Calderon, D., Nguyen, M. L. T., Mezger, A., Kathiria, A., Müller, F., Nguyen, V., … Pritchard, J. K. (2019). Landscape of stimulation-responsive chromatin across diverse human immune cells. *Nature Genetics*, *51*(10), 1494–1505. https://doi.org/10.1038/s41588-019-0505-9

Caligiuri, M. A. (2008). Human natural killer cells. *Blood*, *112*(3), 461–469. https://doi.org/10.1182/blood-2007-09-077438

Cano-Gamez, E., & Trynka, G. (2020). From GWAS to Function : Using Functional Genomics to Identify the Mechanisms Underlying Complex Diseases, *11*(May), 1–21. https://doi.org/10.3389/fgene.2020.00424

Chabaud, M., Garnero, P., Dayer, J. M., Guerne, P. A., Fossiez, F., & Miossec, P. (2000). Contribution of interleukin 17 to synovium matrix destruction in rheumatoid arthritis. *Cytokine*, *12*(7), 1092–1099. https://doi.org/10.1006/cyto.2000.0681

Chan, A. T., Kollnberger, S. D., Wedderburn, L. R., & Bowness, P. (2005). Expansion and enhanced survival of natural killer cells expressing the killer immunoglobulin-like receptor KIR3DL2 in spondylarthritis. *Arthritis and Rheumatism*, *52*(11), 3586–3595. https://doi.org/10.1002/art.21395

Charles A Janeway, J., Travers, P., Walport, M., & Shlomchik, M. J. (2001). Autoimmune responses are directed against self antigens. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK27155/

Chen, L., Ridley, A., Hammitzsch, A., Al-Mossawi, M. H., Bunting, H., Georgiadis, D., … Bowness, P. (2016). Silencing or inhibition of endoplasmic reticulum aminopeptidase 1 (ERAP1) suppresses free heavy chain expression and Th17 responses in ankylosing spondylitis. *Annals of the Rheumatic Diseases*, *75*(5), 916–923. https://doi.org/10.1136/annrheumdis-2014-206996

Chen, Y., Estampador, A. C., Keller, M., Poveda, A., Dalla-Riva, J., Johansson, I., … Varga, T. V. (2019). The combined effects of FADS gene variation and dietary fats in obesity-related traits in a population from the far north of Sweden: the GLACIER Study. *International Journal of Obesity*, *43*(4), 808–820. https://doi.org/10.1038/s41366-018-0112-3

Cheung, V. G., & Spielman, R. S. (2009). Genetics of human gene expression: Mapping DNA variants that influence gene expression. *Nature Reviews Genetics*, *10*(9), 595–604. https://doi.org/10.1038/nrg2630

Cho, O. H., Shin, H. M., Miele, L., Golde, T. E., Fauq, A., Minter, L. M., & Osborne, B. A. (2009). Notch Regulates Cytolytic Effector Function in CD8 + T Cells . *The Journal of Immunology*, *182*(6), 3380–3389. https://doi.org/10.4049/jimmunol.0802598

CHOMZYNSKI, P. (1987). Single-Step Method of RNA Isolation by Acid Guanidinium Thiocyanate–Phenol–Chloroform Extraction. *Analytical Biochemistry*, *162*(1), 156–159. https://doi.org/10.1006/abio.1987.9999

Chory, E., Calarco, J., Hathaway, N., Bell, O., Neel, D., & Crabtree, G. (2019). Nucleosome Turnover Regulates Histone Methylation Patterns over the Genome. Molecular Cell, 73(1), 61-72.e3. doi: 10.1016/j.molcel.2018.10.028

Ciccia, F., Accardo-Palumbo, A., Rizzo, A., Guggino, G., Raimondo, S., Giardina, A., … Triolo, G. (2014). Evidence that autophagy, but not the unfolded protein response, regulates the expression of IL-23 in the gut of patients with ankylosing spondylitis and subclinical gut inflammation. *Annals of the Rheumatic Diseases*, *73*(8), 1566–1574. https://doi.org/10.1136/annrheumdis-2012-202925

Ciccia, F., Alessandro, R., Rizzo, A., Accardo-Palumbo, A., Raimondo, S., Raiata, F., … Triolo, G. (2014). Macrophage phenotype in the subclinical gut inflammation of patients with ankylosing spondylitis. *Rheumatology (United Kingdom)*, *53*(1), 104–113. https://doi.org/10.1093/rheumatology/ket323

Cifaldi, L., Romania, P., Falco, M., Lorenzi, S., Meazza, R., Petrini, S., … Fruci, D. (2015). ERAP1 regulates natural killer cell function by controlling the engagement of inhibitory receptors. *Cancer Research*, *75*(5), 824–834. https://doi.org/10.1158/0008-5472.CAN-14-1643

Colbert, R. A., Delay, M. L., Klenk, E. I., & Layh-Schmitt, G. (2010). From HLA-B27 to spondyloarthritis: A journey through the ER. *Immunological Reviews*. https://doi.org/10.1111/j.0105-2896.2009.00865.x

Colbert, R. A., DeLay, M. L., Layh-Schmitt, G., & Sowders, D. P. (2009). HLA-B27 misfolding and spondyloarthropathies. *Prion*, *3*(1), 15–26. https://doi.org/10.4161/pri.3.1.8072

Colbert, R. A., Tran, T. M., & Layh-Schmitt, G. (2014). HLA-B27 misfolding and ankylosing spondylitis. *Molecular Immunology*, *57*(1), 44–51. https://doi.org/10.1016/j.molimm.2013.07.013

Conrad, K., Wu, P., Sieper, J., & Syrbe, U. (2015). In vivo pre-activation of monocytes in patients with axial spondyloarthritis. *Arthritis Research and Therapy*, *17*(1), 1–12. https://doi.org/10.1186/s13075-015-0694-2

Coppieters, K. T., Dotta, F., Amirian, N., Campbell, P. D., Kay, T. W. H., Atkinson, M. A., … von Herrath, M. G. (2012). Demonstration of islet-autoreactive CD8 T cells in insulitic lesions from recent onset and long-term type 1 diabetes patients. *Journal of Experimental Medicine*, *209*(1), 51–60. https://doi.org/10.1084/jem.20111187

Cortes, A., & Brown, M. A. (2011). Promise and pitfalls of the Immunochip. *Arthritis Research and Therapy*, *13*(1), 2010–2012. https://doi.org/10.1186/ar3204

Cortes, A., Hadler, J., Pointon, J. P., Robinson, P. C., Karaderi, T., Leo, P., … Brown, M. A. (2013). Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. *Nature Genetics*, *45*(7), 730–738. https://doi.org/10.1038/ng.2667

Costantino, F. (2015). ERAP1 Gene Expression Is Influenced by Nonsynonymous Polymorphisms Associated With Predisposition to Spondyloarthritis. Arthritis & Rheumatology, 67(6), 1525–1534. Retrieved from https://doi.org/10.1002/art.39072

Costantino, F., Breban, M., & Garchon, H.-J. (2018). Genetics and Functional Genomics of Spondyloarthritis. *Frontiers in Immunology*, *9*(December), 1–11. https://doi.org/10.3389/fimmu.2018.02933

Cua, D. J., & Tato, C. M. (2010). Innate IL-17-producing cells: the sentinels of the immune system. *Nat Rev Immunol*, *10*(7), 479–489. https://doi.org/10.1038/nri2800

D'Angelo, M. E., Bird, P. I., Peters, C., Reinheckel, T., Trapaniand, J. A., & Sutton, V. R. (2010). Cathepsin H is an additional convertase of pro-granzyme B. *Journal of Biological Chemistry*, *285*(27), 20514–20519. https://doi.org/10.1074/jbc.M109.094573

David Ellinghaus , Luke Jostins, Sarah L Spain, Adrian Cortes, Jörn Bethune, Buhm Han, Yu Rang Park, Soumya Raychaudhuri, Jennie G Pouget, Matthias Hübenthal, Trine Folseraas, Yunpeng Wang, Tonu Esko, Andres Metspalu, Harm-Jan Westra, Lude Franke, Tune H, M.

A. B. & A. F. (2016). Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. *Nature Genetics*, *48*(May), 510–518.

De Lange, K. M., Moutsianas, L., Lee, J. C., Lamb, C. A., Luo, Y., Kennedy, N. A., … Barrett, J. C. (2017). Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nature Genetics*, *49*(2), 256–261. https://doi.org/10.1038/ng.3760

De Rycke, L., Vandooren, B., Kruithof, E., De Keyser, F., Veys, E. M., & Baeten, D. (2005). Tumor necrosis factor α blockade treatment down-modulates the increased systemic and local expression of Toll-like receptor 2 and Toll-like receptor 4 in spondylarthropathy. *Arthritis and Rheumatism*, *52*(7), 2146–2158. https://doi.org/10.1002/art.21155

Dean, L. E., Jones, G. T., Macdonald, A. G., Downham, C., Sturrock, R. D., & Macfarlane, G. J. (2014). Global prevalence of ankylosing spondylitis. *Rheumatology (United Kingdom)*, *53*(4), 650–657. https://doi.org/10.1093/rheumatology/ket387

DeLay, M. L., Turner, M. J., Klenk, E. I., Smith, J. A., Sowders, D. P., & Colbert, R. A. (2009a). HLA-B27 misfolding and the unfolded protein response augment interleukin-23 production and are associated with Th17 activation in transgenic rats. *Arthritis and Rheumatism*, *60*(9), 2633–2643. https://doi.org/10.1002/art.24763

DeLay, M. L., Turner, M. J., Klenk, E. I., Smith, J. A., Sowders, D. P., & Colbert, R. A. (2009b). HLA-B27 misfolding and the unfolded protein response augment interleukin-23 production and are associated with Th17 activation in transgenic rats. *Arthritis and Rheumatism*, *60*(9), 2633–2643. https://doi.org/10.1002/art.24763

Dendrou, C. A., Cortes, A., Shipman, L., Evans, H. G., Attfield, K. E., Jostins, L., … Fugger, L. (2016). Resolving TYK2 locus genotype-To-phenotype differences in autoimmunity. *Science Translational Medicine*, *8*(363), 363ra149. https://doi.org/10.1126/scitranslmed.aag1974

DeSelm, C. J., Miller, B. C., Zou, W., Beatty, W. L., van Meel, H., Takahata, Y., … Virgin, H. W. (2011). Autophagy proteins regulate the secretory component of osteoclastic bone resorption. *Developmental Cell*, *21*(5), 966–974. https://doi.org/10.1016/j.devcel.2011.08.016

Dougados, M., & Baeten, D. (2011). Spondyloarthritis. *The Lancet*, *377*(9783), 2127–2137. https://doi.org/10.1016/S0140-6736(11)60071-8

Duffy, D., Rouilly, V., Libri, V., Hasan, M., Beitz, B., David, M., … Albert, M. L. (2014). Functional analysis via standardized whole-blood stimulation systems defines the boundaries of a healthy immune response to complex stimuli. *Immunity*, *40*(3), 436–450. https://doi.org/10.1016/j.immuni.2014.03.002

Dusseaux, M., Martin, E., Serriari, N., Péguillet, I., Premel, V., Louis, D., … Lantz, O. (2011). Human MAIT cells are xenobiotic-resistant, tissue-targeted, CD161 hi IL-17-secreting T cells. *Blood*, *117*(4), 1250–1259. https://doi.org/10.1182/blood-2010-08-303339

Edinger, T. O., Pohl, M. O., Yángüez, E., & Stertz, S. (2015). Cathepsin W is required for escape of influenza A virus from late endosomes. *MBio*, *6*(3), 1–12. https://doi.org/10.1128/mBio.00297-15

Edwards, S. L., Beesley, J., French, J. D., & Dunning, M. (2013). Beyond GWASs: Illuminating the dark road from association to function. *American Journal of Human Genetics*, *93*(5), 779–797. https://doi.org/10.1016/j.ajhg.2013.10.012

Eichler, E. E., Flint, J., Gibson, G., Kong, A., Leal, S. M., Moore, J. H., & Nadeau, J. H. (2010). Missing heritability and strategies for finding the underlying causes of complex disease. *Nature Reviews. Genetics*, *11*(6), 446–450. https://doi.org/10.1038/nrg2809

El-Gabalawy, H., Guenther, L. C., & Bernstein, C. N. (2010). Epidemiology of immune-mediated inflammatory diseases: Incidence, prevalence, natural history, and comorbidities. *Journal of Rheumatology*, *37*(SUPPL. 85), 2–10. https://doi.org/10.3899/jrheum.091461

El Maghraoui, A. (2011). Extra-articular manifestations of ankylosing spondylitis: Prevalence, characteristics and therapeutic implications. *European Journal of Internal Medicine*, *22*(6), 554–560. https://doi.org/10.1016/j.ejim.2011.06.006

Ellinghaus, D., Jostins, L., Spain, S. L., Cortes, A., Bethune, J., Han, B., … Franke, A. (2016). Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. *Nature Genetics*, *48*(5), 510–518. https://doi.org/10.1038/ng.3528

Evans, D. M., Spencer, C. C. A., Pointon, J. J., Su, Z., Harvey, D., Kochan, G., … Donnelly, P. (2011). Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nature Genetics*, *43*(8), 761–767. https://doi.org/10.1038/ng.873

Faham, M., Carlton, V., Moorhead, M., Zheng, J., Klinger, M., Pepin, F., … Inman, R. D. (2017). Discovery of T Cell Receptor β Motifs Specific to HLA–B27–Positive Ankylosing Spondylitis by Deep Repertoire Sequence Analysis. *Arthritis and Rheumatology*, *69*(4), 774–784. https://doi.org/10.1002/art.40028

Fairfax, B. P., Humburg, P., Makino, S., Naranbhai, V., Wong, D., Lau, E., … Knight, J. C. (2014). Innate Immune Activity Conditions the Effect of Regulatory Variants upon Monocyte Gene Expression. *Science*, *343*(6175), 1246949. https://doi.org/10.1126/science.1246949

Fairfax, B. P., Makino, S., Radhakrishnan, J., Plant, K., Leslie, S., Dilthey, A., … Knight, J. C. (2012). Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nature Genetics*, *44*(5), 502–510. https://doi.org/10.1038/ng.2205

Farh, K. K. H., Marson, A., Zhu, J., Kleinewietfeld, M., Housley, W. J., Beik, S., … Bernstein, B. E. (2015). Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature*, *518*(7539), 337–343. https://doi.org/10.1038/nature13835

Farré, D., Roset, R., Huerta, M., Adsuara, J. E., Roselló, L., Albà, M. M., & Messeguer, X. (2003). Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN. *Nucleic Acids Research*, *31*(13), 3651–3653. https://doi.org/10.1093/nar/gkg605

Ferreira, M. A. R., Mangino, M., Brumme, C. J., Zhao, Z. Z., Medland, S. E., Wright, M. J., … Martin, N. G. (2010). Quantitative Trait Loci for CD4:CD8 Lymphocyte Ratio Are Associated with Risk of Type 1 Diabetes and HIV-1 Immune Control. *American Journal of Human Genetics*, *86*(1), 88–92. https://doi.org/10.1016/j.ajhg.2009.12.008

Fiorillo, M. T., Maragno, M., Butler, R., Dupuis, M. L., & Sorrentino, R. (2000). CD8+ T-cell autoreactivity to an HLA-B27-restricted self-epitope correlates with ankylosing spondylitis. *Journal of Clinical Investigation*, *106*(1), 47–53. https://doi.org/10.1172/JCI9295

Flint, J. (2013). Gwas. *Current Biology*, *23*(7), R265–R266. https://doi.org/10.1016/j.cub.2013.01.040

Fogel, L. A., Yokoyama, W. M., & French, A. R. (2013, July 11). Natural killer cells in human autoimmune disorders. *Arthritis Research and Therapy*. BioMed Central. https://doi.org/10.1186/ar4232

Fogel, O., Bugge Tinggaard, A., Fagny, M., Sigrist, N., Roche, E., Leclere, L., … Tost, J. (2019). Deregulation of microRNA expression in monocytes and CD4+ T lymphocytes from patients with axial spondyloarthritis. *Arthritis Research and Therapy*, *21*(1), 1–14. https://doi.org/10.1186/s13075-019-1829-7

Fornes, O., Castro-Mondragon, J. A., Khan, A., Van Der Lee, R., Zhang, X., Richmond, P. A., … Mathelier, A. (2020). JASPAR 2020: Update of the open-Access database of transcription factor binding profiles. *Nucleic Acids Research*, *48*(D1), D87–D92. https://doi.org/10.1093/nar/gkz1001

Fritz, T., Niederreiter, L., Adolph, T., Blumberg, R. S., & Kaser, A. (2011, November 1). Crohn's disease: NOD2, autophagy and ER stress converge. *Gut*. BMJ Publishing Group. https://doi.org/10.1136/gut.2009.206466

Gaffen, S. L., Jain, R., Garg, A. V., & Cua, D. J. (2014). The IL-23-IL-17 immune axis: From mechanisms to therapeutic testing. *Nature Reviews Immunology*, *14*(9), 585–600. https://doi.org/10.1038/nri3707

Garcês, S., Demengeot, J., & Benito-Garcia, E. (2013). The immunogenicity of anti-TNF therapy in immune-mediated inflammatory diseases: A systematic review of the literature with a meta-analysis. *Annals of the Rheumatic Diseases*, *72*(12), 1947–1955.

https://doi.org/10.1136/annrheumdis-2012-202220
Garrett, S., Jenkinson, T., Kennedy, L. G., Whitelock, H., Gaisford, P., & Calin, A. (1994). A new approach to defining disease status in ankylosing spondylitis: The bath ankylosing spondylitis disease activity index. *Journal of Rheumatology*, *21*(12), 2286–2291. Retrieved from https://europepmc.org/article/med/7699630

Gary-Bobo, M., Nirde, P., Jeanjean, A., Morere, A., & Garcia, M. (2007). Mannose 6-Phosphate Receptor Targeting and its Applications in Human Diseases. *Current Medicinal Chemistry*, *14*(28), 2945–2953. https://doi.org/10.2174/092986707782794005
Geiss, G. K., Bumgarner, R. E., Birditt, B., Dahl, T., Dowidar, N., Dunaway, D. L., … Dimitrov, K. (2008). Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nature Biotechnology*, *26*(3), 317–325. https://doi.org/10.1038/nbt1385
Gilad, Y., Rifkin, S. a., & Pritchard, J. K. (2008). Revealing the architecture of gene regulation: the promise of eQTL studies. *Trends in Genetics*, *24*(8), 408–415. https://doi.org/10.1016/j.tig.2008.06.001
Glinos, D. A., Soskic, B., & Trynka, G. (2017). Immunogenomic approaches to understand the function of immune disease variants. *Immunology*, *152*(4), 527–535. https://doi.org/10.1111/imm.12796
Goris, A., & Liston, A. (2012). The immunogenetic architecture of autoimmune disease. *Cold Spring Harbor Perspectives in Biology*, *4*(3). https://doi.org/10.1101/cshperspect.a007260
Gracey, E., Qaiyum, Z., Almaghlouth, I., Lawson, D., Karki, S., Avvaru, N., … Inman, R. D. (2016). IL-7 primes IL-17 in mucosal-associated invariant T (MAIT) cells, which contribute to the Th17-axis in ankylosing spondylitis. *Annals of the Rheumatic Diseases*, *75*(12), 2124–2132. https://doi.org/10.1136/annrheumdis-2015-208902
Gracey, E., Vereecke, L., McGovern, D., Fröhling, M., Schett, G., Danese, S., … Elewaut, D. (2020, August 1). Revisiting the gut–joint axis: links between gut inflammation and spondyloarthritis. *Nature Reviews Rheumatology*. Nature Research. https://doi.org/10.1038/s41584-020-0454-9
Gracey, E., Yao, Y., Qaiyum, Z., Lim, M., Tang, M., & Inman, R. D. (2020). Altered Cytotoxicity Profile of CD8+ T Cells in Ankylosing Spondylitis. *Arthritis and Rheumatology*, *72*(3), 428–434. https://doi.org/10.1002/art.41129
Gran, J. T., Husby, G., & Hordvik, M. (1985). Prevalence of ankylosing spondylitis in males and females in a young middle-aged population of Tromsø, northern Norway. *Ann Rheum Dis*, (44), 359–367.
Gravano, D. M., & Hoyer, K. K. (2013). Promotion and prevention of autoimmune disease by CD8+ T cells. *Journal of Autoimmunity*, *45*, 68–79. https://doi.org/10.1016/j.jaut.2013.06.004
Gregersen, P. K., Amos, C. I., Lee, A. T., Lu, Y., Remmers, E. F., Kastner, D. L., … Siminovitch, K. A. (2009). REL, encoding a member of the NF-B family of transcription factors, is a newly defined risk locus for rheumatoid arthritis. *Nature Genetics*, *41*(7), 820–823. https://doi.org/10.1038/ng.395
Griffiths, G. M. (2013). Open questions: Missing pieces from the immunological jigsaw puzzle. *BMC Biology*, *11*, 10–11. https://doi.org/10.1186/1741-7007-11-10
Gu, J., Märker-Hermann, E., Baeten, D., Tsai, W., Galdman, D., Maksymowych, W., … Yu, D. T. Y. (2002). A 588-gene microarray analysis of the peripheral blood mononuclear cells of spondyloarthropathy patients. *Rheumatology*, *41*(7), 759–766. https://doi.org/10.1093/rheumatology/41.7.759
Hardy, J., & Singleton, A. (2009). Genomewide association studies and human disease. *The New England Journal of Medicine*, *360*(17), 1759–1768. https://doi.org/10.1056/NEJMra0808700
Harley, J. B., Alarcón-Riquelme, M. E., Criswell, L. A., Jacob, C. O., Kimberly, R. P., Moser, K. L., … Kelly, J. A. (2008). Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. *Nature Genetics*, *40*(2), 204–210. https://doi.org/10.1038/ng.81
Hayashi, E., Chiba, A., Tada, K., Haga, K., Kitagaichi, M., Nakajima, S., … Miyake, S. (2016). Involvement of mucosal-associated invariant T cells in ankylosing spondylitis. *Journal of Rheumatology*, *43*(9), 1695–1703. https://doi.org/10.3899/jrheum.151133

Hermann-Kleiter, N., & Baier, G. (2010, April 15). NFAT pulls the strings during CD4+ T helper cell effector functions. *Blood*. American Society of Hematology. https://doi.org/10.1182/blood-2009-10-233585

Hermann, E., Meyer zum Büschenfelde, K. H., Fleischer, B., & Yu, D. T. Y. (1993). HLA-B27-restricted CD8 T cells derived from synovial fluids of patients with reactive arthritis and ankylosing spondylitis. *The Lancet*, *342*(8872), 646–650. https://doi.org/10.1016/0140-6736(93)91760-J

Hirschhorn, J. N., & Daly, M. J. (2005). Genome-wide association studies for common diseases and complex traits. *Nature Reviews Genetics*, *6*(2), 95–108. https://doi.org/10.1038/nrg1521

Hogan, P. G., Chen, L., Nardone, J., & Rao, A. (2003, September 15). Transcriptional regulation by calcium, calcineurin, and NFAT. *Genes and Development*. Cold Spring Harbor Laboratory Press. https://doi.org/10.1101/gad.1102703

Howie, B. N., Donnelly, P., & Marchini, J. (2009). A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. *PLoS Genetics*, *5*(6), e1000529. https://doi.org/10.1371/journal.pgen.1000529

Hrdlickova, B., de Almeida, R. C., Borek, Z., & Withoff, S. (2014). Genetic variation in the non-coding genome: Involvement of micro-RNAs and long non-coding RNAs in disease. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, *1842*(10), 1910–1922. https://doi.org/10.1016/j.bbadis.2014.03.011

Hu, X., Kim, H., Raj, T., Brennan, P. J., Trynka, G., Teslovich, N., … Raychaudhuri, S. (2014). Regulation of Gene Expression in Autoimmune Disease Loci and the Genetic Basis of Proliferation in CD4+ Effector Memory T Cells. *PLoS Genetics*, *10*(6). https://doi.org/10.1371/journal.pgen.1004404

Hugot, J. P., Chamaillard, M., Zouali, H., Lesage, S., Cézard, J. P., Belaiche, J., … Thomas, G. (2001). Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature*, *411*(6837), 599–603. https://doi.org/10.1038/35079107

Hugot, J. P., Laurent-Puig, P., Gower-Rousseau, C., Olson, J. M., Lee, J. C., Beaugerie, L., … Thomas, G. (1996). Mapping of a susceptibility locus for Crohn's disease on chromosome 16. *Nature*, *379*(6568), 821–823. https://doi.org/10.1038/379821a0

Hunt, K. A., Mistry, V., Bockett, N. A., Ahmad, T., Ban, M., Barker, J. N., … Van Heel, D. A. (2013). Negligible impact of rare autoimmune-locus coding-region variants on missing heritability. *Nature*, *498*(7453), 232–235. https://doi.org/10.1038/nature12170

Hussman, J. P., Beecham, A. H., Schmidt, M., Martin, E. R., McCauley, J. L., Vance, J. M., … Pericak-Vance, M. A. (2016). GWAS analysis implicates NF-κB-mediated induction of inflammatory T cells in multiple sclerosis. *Genes and Immunity*, *17*(5), 305–312. https://doi.org/10.1038/gene.2016.23

Iglesias, M., Arun, A., Chicco, M., Lam, B., Talbot, C. C., Ivanova, V., … Raimondi, G. (2018). Type-I interferons inhibit interleukin-10 signaling and favor Type 1 diabetes development in nonobese diabetic mice. *Frontiers in Immunology*, *9*(JUL), 1–14. https://doi.org/10.3389/fimmu.2018.01565

Ishizaki, M., Akimoto, T., Muromoto, R., Yokoyama, M., Ohshiro, Y., Sekine, Y., … Matsuda, T. (2011). Involvement of Tyrosine Kinase-2 in Both the IL-12/Th1 and IL-23/Th17 Axes In Vivo. *The Journal of Immunology*, *187*(1), 181–189. https://doi.org/10.4049/jimmunol.1003244

J Sieper, J Braun, M Rudwaleit, A Boonen, A. Z. (2002). Ankylosing spondylitis: an overview. *Annals of the Rheumatic Diseases*, *61*. https://doi.org/10.1136/ard.61.suppl_3.iii8

Jean François Bach. (2002). Susceptibility To autoimmune and allergic D Iseases. *New England Journal of Medicine*, *347*(12), 911–920. https://doi.org/10.1056/NEJMra020100

Kellis, M., Wold, B., Snyder, M. P., Bernstein, B. E., Kundaje, A., Marinov, G. K., … Hardison, R. C. (2014). Defining functional DNA elements in the human genome. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(17), 6131–6138. https://doi.org/10.1073/pnas.1318948111

Kenna, T. J., Davidson, S. I., Duan, R., Bradbury, L. A., McFarlane, J., Smith, M., … Brown, M. A. (2012). Enrichment of circulating interleukin-17-secreting interleukin-23 receptor-positive

γ/δ T cells in patients with active ankylosing spondylitis. *Arthritis and Rheumatism*, *64*(5), 1420–1429. https://doi.org/10.1002/art.33507

Khan, M. A., Mathieu, A., Sorrentino, R., & Akkoc, N. (2007). The pathogenetic role of HLA-B27 and its subtypes. *Autoimmunity Reviews*, *6*(3), 183–189. https://doi.org/10.1016/j.autrev.2006.11.003

Kimura, H. (2013). Histone modifications for human epigenome analysis. *Journal of Human Genetics*, *58*(7), 439–445. https://doi.org/10.1038/jhg.2013.66

Klein-Hessling, S., Muhammad, K., Klein, M., Pusch, T., Rudolf, R., Flöter, J., … Serfling, E. (2017). NFATc1 controls the cytotoxicity of CD8+ T cells. *Nature Communications*, *8*(1). https://doi.org/10.1038/s41467-017-00612-6

Klemm, S. L., Shipony, Z., & Greenleaf, W. J. (n.d.). Chromatin accessibility and the regulatory epigenome. https://doi.org/10.1038/s41576-018-0089-8

Knight, J. C. (2014). Approaches for establishing the function of regulatory genetic variants involved in disease. *Genome Medicine*, *6*(10), 1–15. https://doi.org/10.1186/s13073-014-0092-4

Kreins, A. Y., Ciancanelli, M. J., Okada, S., Kong, X. F., Ramírez-Alejo, N., Kilic, S. S., … Boisson-Dupuis, S. (2015). Human TYK2 deficiency: Mycobacterial and viral infections without hyper-IgE syndrome. *Journal of Experimental Medicine*, *212*(10), 1641–1662. https://doi.org/10.1084/jem.20140280

Kuek, A., Hazleman, B. L., & Östör, A. J. K. (2007, April). Immune-mediated inflammatory diseases (IMIDs) and biologic therapy: A medical revolution. *Postgraduate Medical Journal*. BMJ Publishing Group. https://doi.org/10.1136/pgmj.2006.052688

Kumar, V., Wijmenga, C., & Xavier, R. J. (2014). Genetics of immune-mediated disorders: From genome-wide association to molecular mechanism. *Current Opinion in Immunology*, *31*, 51–57. https://doi.org/10.1016/j.coi.2014.09.007

Kuznetsova, T., Prange, K. H. M., Glass, C. K., & de Winther, M. P. J. (2020). Transcriptional and epigenetic regulation of macrophages in atherosclerosis. *Nature Reviews Cardiology*, *17*(4), 216–228. https://doi.org/10.1038/s41569-019-0265-3

Laggner, U., Di Meglio, P., Perera, G. K., Hundhausen, C., Lacy, K. E., Ali, N., … Nestle, F. O. (2011). Identification of a Novel Proinflammatory Human Skin-Homing Vγ9Vδ2 T Cell Subset with a Potential Role in Psoriasis. *The Journal of Immunology*, *187*(5), 2783–2793. https://doi.org/10.4049/jimmunol.1100804

Lambert, S., Jolma, A., Campitelli, L., Das, P., Yin, Y., & Albu, M. et al. (2018). The Human Transcription Factors. Cell, 175(2), 598-599. doi: 10.1016/j.cell.2018.09.045

Lander, E., & Schork, N. J. (1994). Genetic dissection of complex traits. *Science*, *265*(4), 2037–2048.

Langer, V., Britzen-laurent, N., Stürzl, M., Langer, V., Vivi, E., Regensburger, D., … Stürzl, M. (2019). IFN- g drives inflammatory bowel disease pathogenesis through VE-cadherin – directed vascular barrier disruption Graphical abstract Find the latest version : IFN- γ drives inflammatory bowel disease pathogenesis through VE-cadherin – directed vascular bar, *129*(11), 4691–4707.

Lattka, E., Eggers, S., Moeller, G., Heim, K., Weber, M., Mehta, D., … Adamski, J. (2010). A common FADS2 promoter polymorphism increases promoter activity and facilitates binding of transcription factor ELK1. *Journal of Lipid Research*, *51*(1), 182–191. https://doi.org/10.1194/jlr.M900289-JLR200

Lau, M. C., Keith, P., Costello, M. E., Bradbury, L. A., Hollis, K. A., Thomas, R., … Kenna, T. J. (2017). Genetic association of ankylosing spondylitis with TBX21 influences T-bet and pro-inflammatory cytokine expression in humans and SKG mice as a model of spondyloarthritis. *Annals of the Rheumatic Diseases*, *76*(1), 261–269. https://doi.org/10.1136/annrheumdis-2015-208677

Lee, J. C., Espéli, M., Anderson, C. A., Linterman, M. A., Pocock, J. M., Williams, N. J., … Smith, K. G. C. (2013). Human SNP links differential outcomes in inflammatory and infectious disease to a FOXO3-regulated pathway. *Cell*, *155*(1), 57–69. https://doi.org/10.1016/j.cell.2013.08.034

Lee, M. N., Ye, C., Villani, A.-C., Raj, T., Li, W., Eisenhaure, T. M., … Hacohen, N. (2014). Common genetic variants modulate pathogen-sensing responses in human dendritic cells.

Science (New York, N.Y.), *343*(6175), 1246980. https://doi.org/10.1126/science.1246980
Lee, W., Reveille, J. D., Jr, J. C. D., Learch, T. J., Ward, M. M., & Weisman, M. H. (2007). Are there gender differences in severity of ankylosing spondylitis? Results from the PSOAS cohort. *Ann Rheum Dis*, *66*, 633–638. https://doi.org/10.1136/ard.2006.060293
Lever, J., Krzywinski, M., & Altman, N. (2017). Points of Significance: Principal component analysis. *Nature Methods*, *14*(7), 641–642. https://doi.org/10.1038/nmeth.4346
Levine, B., & Deretic, V. (2007). Unveiling the roles of autophagy in innate and adaptive immunity. *Nature Reviews Immunology*, *7*(10), 767–777. https://doi.org/10.1038/nri2161
Levine, B., & Kroemer, G. (2019). Leading Edge Review Biological Functions of Autophagy Genes: A Disease Perspective. https://doi.org/10.1016/j.cell.2018.09.048
Li, Y., Oosting, M., Deelen, P., Ricaño-Ponce, I., Smeekens, S., Jaeger, M., … Netea, M. G. (2016). Inter-individual variability and genetic influences on cytokine responses to bacteria and fungi. *Nature Medicine*, *22*(8), 952–960. https://doi.org/10.1038/nm.4139
Li, Y., Oosting, M., Smeekens, S. P., Jaeger, M., Aguirre-Gamboa, R., Le, K. T. T., … Netea, M. G. (2016). A Functional Genomics Approach to Understand Variation in Cytokine Production in Humans. *Cell*, *167*(4), 1099-1110.e14. https://doi.org/10.1016/j.cell.2016.10.017
Li, Z, Haynes, K., Pennisi, D. J., Anderson, L. K., Song, X., Thomas, G. P., … Brown, M. A. (2017). Epigenetic and gene expression analysis of ankylosing spondylitis-associated loci implicate immune cells and the gut in the disease pathogenesis. *Genes and Immunity*, *18*(3), 135–143. https://doi.org/10.1038/gene.2017.11
Li, Zhi, Gakovic, M., Ragimbeau, J., Eloranta, M.-L., Rönnblom, L., Michel, F., & Pellegrini, S. (2013). Two Rare Disease-Associated Tyk2 Variants Are Catalytically Impaired but Signaling Competent. *The Journal of Immunology*, *190*(5), 2335–2344. https://doi.org/10.4049/jimmunol.1203118
Li, Zhixiu, & Brown, M. A. (2017). Progress of genome-wide association studies of ankylosing spondylitis. *Clinical and Translational Immunology*, *6*(12). https://doi.org/10.1038/cti.2017.49
Linden, S. Van Der, Valkenburg, H. A., & Cats, A. (1984). Evaluation of Diagnostic Criteria for Ankylosing Spondylitis. *Arthritis & Rheumatism*, *27*(4), 361–368. https://doi.org/10.1002/art.1780270401
Linnevers, C., Smeekens, S. P., & Brömme, D. (1997). Human cathepsin W, a putative cysteine protease predominantly expressed in CD8+ T-lymphocytes. *FEBS Letters*, *405*(3), 253–259. https://doi.org/10.1016/S0014-5793(97)00118-X
Liu, J. Z., Van Sommeren, S., Huang, H., Ng, S. C., Alberts, R., Takahashi, A., … Weersma, R. K. (2015). Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nature Genetics*, *47*(9), 979–986. https://doi.org/10.1038/ng.3359
Liu, Z., & Tjian, R. (2018, April 1). Visualizing transcription factor dynamics in living cells. *Journal of Cell Biology*. Rockefeller University Press. https://doi.org/10.1083/jcb.201710038
Lobo, I. (2008). Multifactorial Inheritance and Genetic Disease. *Nature Education*, *1*(1), 5. Retrieved from https://www.nature.com/scitable/topicpage/multifactorial-inheritance-and-genetic-disease-919/
Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., … Moore, H. F. (2013). The Genotype-Tissue Expression (GTEx) project. *Nature Genetics*, *45*(6), 580–585. https://doi.org/10.1038/ng.2653
Lubberts, E. (2015). The IL-23-IL-17 axis in inflammatory arthritis. *Nat Rev Rheumatol*, *11*(7), 415–429. https://doi.org/10.1038/nrrheum.2015.53
Luke Jostins, Ripke, S., Weersma, R. K., Duerr, R. H., McGovern, D. P., Hui, K. Y., … Kaida Ning, Isabelle Cleynen, Emilie Theatre, Sarah L. Spain, Soumya Raychaudhuri, Philippe Goyette, Zhi Wei, Clara Abraham, Jean-Paul Achkar, Tariq Ahmad, Leila Amininejad, Ashwin N. Ananthakrishnan, V. A. (2012). Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature*, *490*(7422), 119–124. https://doi.org/10.1038/nature11582
Ma, Y., Galluzzi, L., Zitvogel, L., & Kroemer, G. (2013, August 22). Autophagy and cellular immune responses. *Immunity*. Cell Press. https://doi.org/10.1016/j.immuni.2013.07.017
Machado, P., Landewé, R., Braun, J., Baraliakos, X., Hermann, K. G. A., Hsu, B., … Van Der

Heijde, D. (2013). Ankylosing spondylitis patients with and without psoriasis do not differ in disease phenotype. *Annals of the Rheumatic Diseases*, *72*(6), 1104–1107. https://doi.org/10.1136/annrheumdis-2012-202922

MacHado, P., Landewé, R., Lie, E., Kvien, T. K., Braun, J., Baker, D., & Van Der Heijde, D. (2011). Ankylosing Spondylitis Disease Activity Score (ASDAS): Defining cut-off values for disease activity states and improvement scores. *Annals of the Rheumatic Diseases*, *70*(1), 47–53. https://doi.org/10.1136/ard.2010.138594

Machado, P., Navarro-Compán, V., Landewé, R., Gaalen, F. A., Roux, C., & Heijde, D. (2015). Brief Report: Calculating the Ankylosing Spondylitis Disease Activity Score If the Conventional C-Reactive Protein Level Is Below the Limit of Detection or If High-Sensitivity C-Reactive Protein Is Used: An Analysis in the DESIR Cohort. *Arthritis & Rheumatology*, *67*(2), 408–413. https://doi.org/10.1002/art.38921

Maddur, M. S., Miossec, P., Kaveri, S. V., & Bayry, J. (2012, July 1). Th17 cells: Biology, pathogenesis of autoimmune and inflammatory diseases, and therapeutic strategies. *American Journal of Pathology*. Elsevier. https://doi.org/10.1016/j.ajpath.2012.03.044

Marchini, J., & Howie, B. (2010). Genotype imputation for genome-wide association studies. *Nature Reviews Genetics*, *11*(7), 499–511. https://doi.org/10.1038/nrg2796

Maurano, M. T., Humbert, R., Rynes, E., Thurman, R. E., Haugen, E., Wang, H., … Stamatoyannopoulos, J. A. (2012). Systematic localization of common disease-associated variation in regulatory DNA. *Science*, *337*(6099), 1190–1195. https://doi.org/10.1126/science.1222794

May, E., Dorris, M. L., Satumtira, N., Iqbal, I., Rehman, M. I., Lightfoot, E., & Taurog, J. D. (2003). CD8αβ T Cells Are Not Essential to the Pathogenesis of Arthritis or Colitis in HLA-B27 Transgenic Rats. *The Journal of Immunology*, *170*(2), 1099–1105. https://doi.org/10.4049/jimmunol.170.2.1099

McCarthy, M. I., Abecasis, G. R., Cardon, L. R., Goldstein, D. B., Little, J., Ioannidis, J. P. a, & Hirschhorn, J. N. (2008). Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nature Reviews. Genetics*, *9*(5), 356–369. https://doi.org/10.1038/nrg2344

McGeachy, M. J., Chen, Y., Tato, C. M., Laurence, A., Joyce-Shaikh, B., Blumenschein, W. M., … Cua, D. J. (2009). The interleukin 23 receptor is essential for the terminal differentiation of interleukin 17-producing effector T helper cells in vivo. *Nature Immunology*, *10*(3), 314–324. https://doi.org/10.1038/ni.1698

McGonagle, D. G., McInnes, I. B., Kirkham, B. W., Sherlock, J., & Moots, R. (2019). The role of IL-17A in axial spondyloarthritis and psoriatic arthritis: Recent advances and controversies. *Annals of the Rheumatic Diseases*, *78*(9), 1167–1178. https://doi.org/10.1136/annrheumdis-2019-215356

Mchugh, K., & Bowness, P. (2012). The link between HLA-B27 and SpA-new ideas on an old problem. *Rheumatology (United Kingdom)*, *51*(9), 1529–1539. https://doi.org/10.1093/rheumatology/kes061

Melé, M., Ferreira, P. G., Reverter, F., DeLuca, D. S., Monlong, J., Sammeth, M., … Guigó, R. (2015). The human transcriptome across tissues and individuals. *Science*, *348*(6235), 660–665. https://doi.org/10.1126/science.aaa0355

Menon, B., Gullick, N. J., Walter, G. J., Rajasekhar, M., Garrood, T., Evans, H. G., … Kirkham, B. W. (2014). Interleukin-17+CD8+ T cells are enriched in the joints of patients with psoriatic arthritis and correlate with disease activity and joint damage progression. *Arthritis and Rheumatology*, *66*(5), 1272–1281. https://doi.org/10.1002/art.38376

Messeguer, X., Escudero, R., Farré, D., Núñez, O., Martínez, J., & Albà, M. M. (2002). PROMO: Detection of known transcription regulatory elements using species-tailored searches. *Bioinformatics*, *18*(2), 333–334. https://doi.org/10.1093/bioinformatics/18.2.333

Meyer zu Horste, G., Wu, C., Wang, C., Cong, L., Pawlak, M., Lee, Y., … Kuchroo, V. K. (2016). RBPJ Controls Development of Pathogenic Th17 Cells by Regulating IL-23 Receptor Expression. *Cell Reports*, *16*(2), 392–404. https://doi.org/10.1016/j.celrep.2016.05.088

Moss, R. B., Moll, T., El-Kalay, M., Kohne, C., Hoo, W. S., Encinas, J., & Carlo, D. J. (2004, December). Th1/Th2 cells in inflammatory disease states: Therapeutic implications. *Expert Opinion on Biological Therapy*. Expert Opin Biol Ther. https://doi.org/10.1517/14712598.4.12.1887

Ngo, S. T., Steyn, F. J., & Mccombe, P. A. (2014). Frontiers in Neuroendocrinology Gender differences in autoimmune disease. *Frontiers in Neuroendocrinology*, *35*(3), 347–369. https://doi.org/10.1016/j.yfrne.2014.04.004

Nicolae, D. L., Gamazon, E., Zhang, W., Duan, S., Eileen Dolan, M., & Cox, N. J. (2010). Trait-associated SNPs are more likely to be eQTLs: Annotation to enhance discovery from GWAS. *PLoS Genetics*, *6*(4). https://doi.org/10.1371/journal.pgen.1000888

O'Rielly, D. D., Uddin, M., & Rahman, P. (2016). Ankylosing spondylitis. *Current Opinion in Rheumatology*, *28*(4), 337–345. https://doi.org/10.1097/BOR.0000000000000297

Ogura, Y., Inohara, N., Benito, A., Chen, F. F., Yamaoka, S., & Núñez, G. (2001). Nod2, a Nod1/Apaf-1 Family Member That Is Restricted to Monocytes and Activates NF-κB. *Journal of Biological Chemistry*, *276*(7), 4812–4818. https://doi.org/10.1074/jbc.M008072200

Okamoto, M., Takeda, K., Joetham, A., Ohnishi, H., Matsuda, H., Swasey, C. H., … Gelfand, E. W. (2008). Essential role of Notch signaling in effector memory CD8+ T cell-mediated airway hyperresponsiveness and inflammation. *Journal of Experimental Medicine*, *205*(5), 1087–1097. https://doi.org/10.1084/jem.20072200

Ondr, J. K., & Pham, C. T. N. (2004). Characterization of murine cathepsin W and its role in cell-mediated cytotoxity. *Journal of Biological Chemistry*, *279*(26), 27525–27533. https://doi.org/10.1074/jbc.M400304200

Osgood, J. A., & Knight, J. C. (2018). Translating GWAS in rheumatic disease: approaches to establishing mechanism and function for genetic associations with ankylosing spondylitis. *Briefings in Functional Genomics*, *17*(August), 1–11. https://doi.org/10.1093/bfgp/ely015

Papp, K. A., Langley, R. G., Lebwohl, M., Krueger, G. G., Szapary, P., Yeilding, N., … Reich, K. (2008). Efficacy and safety of ustekinumab, a human interleukin-12/23 monoclonal antibody, in patients with psoriasis: 52-week results from a randomised, double-blind, placebo-controlled trial (PHOENIX 2). *The Lancet*, *371*(9625), 1675–1684. https://doi.org/10.1016/S0140-6736(08)60726-6

Parkes, M., Cortes, A., Van Heel, D. A., & Brown, M. A. (2013a). Genetic insights into common pathways and complex relationships among immune-mediated diseases. *Nature Reviews Genetics*, *14*(9), 661–673. https://doi.org/10.1038/nrg3502

Parkes, M., Cortes, A., Van Heel, D. A., & Brown, M. A. (2013b). Genetic insights into common pathways and complex relationships among immune-mediated diseases. *Nature Reviews Genetics*, *14*(9), 661–673. https://doi.org/10.1038/nrg3502

Pasaniuc, B., & Price, A. L. (2017). Dissecting the genetics of complex traits using summary association statistics. *Nature Reviews Genetics*, *18*(2), 117–127. https://doi.org/10.1038/nrg.2016.142

Patel, D. D., & Kuchroo, V. K. (2015). Review Th17 Cell Pathway in Human Immunity: Lessons from Genetics and Therapeutic Interventions. https://doi.org/10.1016/j.immuni.2015.12.003

Patel, S., Homaei, A., El-Seedi, H. R., & Akhtar, N. (2018). Cathepsins: Proteases that are vital for survival but can also be fatal. *Biomedicine and Pharmacotherapy*, *105*(April), 526–532. https://doi.org/10.1016/j.biopha.2018.05.148

Patsopoulos, N. A., Baranzini, S. E., Santaniello, A., Shoostari, P., Cotsapas, C., Wong, G., … De Jager, P. L. (2019). Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. *Science*, *365*(6460). https://doi.org/10.1126/science.aav7188

Pearce, E. L., Mullen, A. C., Martins, G. A., Krawczyk, C. M., Hutchins, A. S., Zediak, V. P., … Reiner, S. L. (2003). Control of Effector CD8+ T Cell Function by the Transcription Factor Eomesodermin. *Science*, *302*(5647), 1041–1043. https://doi.org/10.1126/science.1090148

Pendergrass, S. A., Brown-Gentry, K., Dudek, S. M., Torstenson, E. S., Ambite, J. L., Avery, C. L., … Ritchie, M. D. (2011). The use of phenome-wide association studies (PheWAS) for exploration of novel genotype-phenotype relationships and pleiotropy discovery. *Genetic Epidemiology*, *35*(5), 410–422. https://doi.org/10.1002/gepi.20589

Piasecka, B., Duffy, D., Urrutia, A., Quach, H., Patin, E., Posseme, C., … Quintana-Murci, L.

(2018). Distinctive roles of age, sex, and genetics in shaping transcriptional variation of human immune responses to microbial challenges. *Proceedings of the National Academy of Sciences*, *115*(3), E488–E497. https://doi.org/10.1073/pnas.1714765115

Prehn, J. L., Thomas, L. S., Landers, C. J., Yu, Q. T., Michelsen, K. S., & Targan, S. R. (2007). The T Cell Costimulator TL1A Is Induced by FcγR Signaling in Human Monocytes and Dendritic Cells. *The Journal of Immunology*, *178*(7), 4033–4038. https://doi.org/10.4049/jimmunol.178.7.4033

Proft, F., & Poddubnyy, D. (2018). Ankylosing spondylitis and axial spondyloarthritis: recent insights and impact of new classification criteria. *Therapeutic Advances in Musculoskeletal Disease*, *10*(5–6), 129–139. https://doi.org/10.1177/1759720X18773726

Pruim, R. J., Welch, R. P., Sanna, S., Teslovich, T. M., Chines, P. S., Gliedt, T. P., … Frishman, D. (2011). LocusZoom: Regional visualization of genome-wide association scan results. *Bioinformatics*, *27*(13), 2336–2337. https://doi.org/10.1093/bioinformatics/btq419

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., … Sham, P. C. (2007). REPORT PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics Am. J. Hum. Genet*, *8181*, 559–575. https://doi.org/10.1086/519795

Qaiyum, Z., Gracey, E., Yao, Y. C., & Inman, R. D. (2019). Integrin and transcriptomic profiles identify a distinctive synovial CD8+ T cell subpopulation in spondyloarthritis. *Annals of the Rheumatic Diseases*, *78*(11), 1566–1575. https://doi.org/10.1136/annrheumdis-2019-215349

Qian, Y., Liu, C., Hartupee, J., Altuntas, C. Z., Fatih Gulen, M., Jane-Wit, D., … Li, X. (2007). The adaptor Act1 is required for interleukin 17-dependent signaling associated with autoimmune and inflammatory disease. https://doi.org/10.1038/ni1439

Raphael, I., Nalawade, S., Eagar, T. N., & Forsthuber, T. G. (2015). T cell subsets and their signature cytokines in autoimmune and inflammatory diseases. *Cytokine*, *74*(1), 5–17. https://doi.org/10.1016/j.cyto.2014.09.011

Read, S., Greenwald, R., Izcue, A., Robinson, N., Mandelbrot, D., Francisco, L., … Powrie, F. (2006). Blockade of CTLA-4 on CD4 + CD25 + Regulatory T Cells Abrogates Their Function In Vivo . *The Journal of Immunology*, *177*(7), 4376–4383. https://doi.org/10.4049/jimmunol.177.7.4376

Regner, E. H., Ohri, N., Stahly, A., Gerich, M. E., Fennimore, B. P., Ir, D., … Kuhn, K. A. (2018). Functional intraepithelial lymphocyte changes in inflammatory bowel disease and spondyloarthritis have disease specific correlations with intestinal microbiota. *Arthritis Research and Therapy*, *20*(1), 149. https://doi.org/10.1186/s13075-018-1639-3

Repnik, U., Stoka, V., Turk, V., & Turk, B. (2012). Lysosomes and lysosomal cathepsins in cell death. *Biochimica et Biophysica Acta - Proteins and Proteomics*, *1824*(1), 22–33. https://doi.org/10.1016/j.bbapap.2011.08.016

Reveille, J. D. (2014). An update on the contribution of the MHC to as susceptibility. *Clinical Rheumatology*, *33*(6), 749–757. https://doi.org/10.1007/s10067-014-2662-7

Reynolds, L. M., Howard, T. D., Ruczinski, I., Kanchan, K., Seeds, M. C., Mathias, R. A., & Chilton, F. H. (2018). Tissue-specific impact of FADS cluster variants on FADS1 and FADS2 gene expression. *PLoS ONE*, *13*(3), 1–15. https://doi.org/10.1371/journal.pone.0194610

Rezaiemanesh, A., Abdolmaleki, M., Abdolmohammadi, K., Aghaei, H., Pakdel, F. D., Fatahi, Y., … Nicknam, M. H. (2018). Immune cells involved in the pathogenesis of ankylosing spondylitis. *Biomedicine and Pharmacotherapy*, *100*(November 2017), 198–204. https://doi.org/10.1016/j.biopha.2018.01.108

Ricã No-Ponce, I., & Wijmenga, C. (2013). Mapping of Immune-Mediated Disease Genes. *Annu. Rev. Genomics Hum. Genet*, *14*, 325–353. https://doi.org/10.1146/annurev-genom-091212-153450

Richard, A. C., Peters, J. E., Savinykh, N., Lee, J. C., Hawley, E. T., Meylan, F., … Smith, K. G. C. (2018). Reduced monocyte and macrophage TNFSF15/TL1A expression is associated with susceptibility to inflammatory bowel disease. *PLoS Genetics*, *14*(9), e1007458. https://doi.org/10.1371/journal.pgen.1007458

Rivas, M. A., Beaudoin, M., Gardet, A., Stevens, C., Sharma, Y., Zhang, C. K., … Daly, M. J. (2011). Deep resequencing of GWAS loci identifies independent rare variants associated

with inflammatory bowel disease. *Nature Genetics*, *43*(11), 1066–1073. https://doi.org/10.1038/ng.952

Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., … Kellis, M. (2015). Integrative analysis of 111 reference human epigenomes. *Nature*, *518*(7539), 317–329. https://doi.org/10.1038/nature14248

Robinson, P. C., & Brown, M. A. (2014). Genetics of ankylosing spondylitis. *Molecular Immunology*. Elsevier Ltd. https://doi.org/10.1016/j.molimm.2013.06.013

Rojano, E., Seoane, P., Ranea, J. A. G., & Perkins, J. R. (2019). Regulatory variants: From detection to predicting impact. *Briefings in Bioinformatics*, *20*(5), 1639–1654. https://doi.org/10.1093/bib/bby039

Roke, K., Ralston, J. C., Abdelmagid, S., Nielsen, D. E., Badawi, A., El-Sohemy, A., … Mutch, D. M. (2013). Variation in the FADS1/2 gene cluster alters plasma n-6 PUFA and is weakly associated with hsCRP levels in healthy young adults. *Prostaglandins Leukotrienes and Essential Fatty Acids*, *89*(4), 257–263. https://doi.org/10.1016/j.plefa.2013.06.003

Rudwaleit, M., Van Der Heijde, D., Landewé, R., Akkoc, N., Brandt, J., Chou, C. T., … Sieper, J. (2011). The Assessment of SpondyloArthritis international Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. *Annals of the Rheumatic Diseases*, *70*(1), 25–31. https://doi.org/10.1136/ard.2010.133645

Rudwaleit, M., van der Heijde, D., Landewe, R., Listing, J., Akkoc, N., Brandt, J., … Sieper, J. (2009). The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. *Annals of the Rheumatic Diseases*, *68*(6), 777–783. https://doi.org/10.1136/ard.2009.108233

Sarikonda, G., Pettus, J., Phatak, S., Sachithanantham, S., Miller, J. F., Wesley, J. D., … Von Herrath, M. (2014). CD8 T-cell reactivity to islet antigens is unique to type 1 while CD4 T-cell reactivity exists in both type 1 and type 2 diabetes. *Journal of Autoimmunity*, *50*, 77–82. https://doi.org/10.1016/j.jaut.2013.12.003

Schaid, D. J., Chen, W., & Larson, N. B. (2018). From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nature Reviews Genetics*, *19*(8), 491–504. https://doi.org/10.1038/s41576-018-0016-z

Schaub, M. A., Boyle, A. P., Kundaje, A., Batzoglou, S., & Snyder, M. (2012). Linking disease associations with regulatory information in the human genome. *Genome Research*, *22*(9), 1748–1759. https://doi.org/10.1101/gr.136127.111

Schirmer, M., Goldberger, C., Duftner, C., Clausen, J., & Falkenbach, A. (2001). Enrichment of CD8+ CD28- cytotoxic T cells in circulating lymphocytes of patients with ankylosing spondylitis. Retrieved from http://europepmc.org/abstract/PMC/PMC3273168

Schmiedel, B. J., Singh, D., Madrigal, A., Valdovino-Gonzalez, A. G., White, B. M., Zapardiel-Gonzalo, J., … Vijayanand, P. (2018). Impact of Genetic Polymorphisms on Human Immune Cell Gene Expression. *Cell*, 1–15. https://doi.org/10.1016/j.cell.2018.10.022

Schön, M. P., & Erpenbeck, L. (2018). The interleukin-23/interleukin-17 axis links adaptive and innate immunity in psoriasis. *Frontiers in Immunology*, *9*(JUN), 1–13. https://doi.org/10.3389/fimmu.2018.01323

Shabalin, A. A. (2012). Matrix eQTL: Ultra fast eQTL analysis via large matrix operations. *Bioinformatics*, *28*(10), 1353–1358. https://doi.org/10.1093/bioinformatics/bts163

Sharp, R. C., Abdulrahim, M., Naser, E. S., & Naser, S. A. (2015). Genetic variations of PTPN2 and PTPN22: Role in the pathogenesis of Type 1 diabetes and Crohn's disease. *Frontiers in Cellular and Infection Microbiology*. Frontiers Media S.A. https://doi.org/10.3389/fcimb.2015.00095

Shen, H., Goodall, J. C., & Hill Gaston, J. S. (2009). Frequency and phenotype of peripheral blood Th17 cells in ankylosing spondylitis and rheumatoid arthritis. *Arthritis and Rheumatism*, *60*(6), 1647–1656. https://doi.org/10.1002/art.24568

Shen, H., Goodall, J. C., & Hill Gaston, J. S. (2010). Frequency and phenotype of T helper 17 cells in peripheral blood and synovial fluid of patients with reactive arthritis. *Journal of Rheumatology*, *37*(10), 2096–2099. https://doi.org/10.3899/jrheum.100146

Sherlock, J. P., Joyce-Shaikh, B., Turner, S. P., Chao, C.-C., Sathe, M., Grein, J., … Cua, D. J. (2012). IL-23 induces spondyloarthropathy by acting on ROR-γt+ CD3+CD4-CD8- entheseal resident T cells. *Nature Medicine*, *18*(7), 1069–1076.

https://doi.org/10.1038/nm.2817
Sieper, J, Braun, J., Dougados, M., & Baeten, D. (2015). Axial spondyloarthritis. *Nature Reviews Disease Primers*, *1*(6), 1–16. https://doi.org/10.1016/j.coi.2006.09.008
Sieper, Joachim, & Poddubnyy, D. (2016). New evidence on the management of spondyloarthritis. *Nature Reviews Rheumatology*. https://doi.org/10.1038/nrrheum.2016.42

Sieper, Joachim, Poddubnyy, D., & Miossec, P. (2019). The IL-23–IL-17 pathway as a therapeutic target in axial spondyloarthritis. *Nature Reviews Rheumatology*, *15*(12), 747–757. https://doi.org/10.1038/s41584-019-0294-7
Sieper, Joachim, Rudwaleit, M., Khan, M. A., & Braun, J. (2006). Concepts and epidemiology of spondyloarthritis. *Best Practice and Research: Clinical Rheumatology*, *20*(3), 401–417. https://doi.org/10.1016/j.berh.2006.02.001
Skrzeczyńska-Moncznik, J., Bzowska, M., Loseke, S., Grage-Griebenow, E., Zembala, M., & Pryjma, J. (2008). Peripheral blood CD14high CD16+ monocytes are main producers of IL-10. *Scandinavian Journal of Immunology*, *67*(2), 152–159. https://doi.org/10.1111/j.1365-3083.2007.02051.x
Sokol, H., Conway, K. L., Zhang, M., Choi, M., Morin, B., Cao, Z., … Xavier, R. J. (2013). Card9 mediates intestinal epithelial cell restitution, t-helper 17 responses, and control of bacterial infection in mice. *Gastroenterology*, *145*(3), 591-601.e3. https://doi.org/10.1053/j.gastro.2013.05.047
Soldner, F., Stelzer, Y., Shivalila, C. S., Abraham, B. J., Latourelle, J. C., Barrasa, M. I., … Jaenisch, R. (2016). Parkinson-associated risk variant in distal enhancer of α-synuclein modulates target gene expression. *Nature*, *533*(7601), 95–99. https://doi.org/10.1038/nature17939
Soskic, B., Cano-Gamez, E., Smyth, D. J., Rowan, W. C., Nakic, N., Esparza-Gordillo, J., … Trynka, G. (2019). Chromatin activity at GWAS loci identifies T cell states driving complex immune diseases. *Nature Genetics*, *51*(10), 1486–1493. https://doi.org/10.1038/s41588-019-0493-9
Spalinger, M. R., McCole, D. F., Rogler, G., & Scharl, M. (2015). Role of protein tyrosine phosphatases in regulating the immune system: Implications for chronic intestinal inflammation. *Inflammatory Bowel Diseases*, *21*(3), 645–655. https://doi.org/10.1097/MIB.0000000000000297
Srenathan, U., Steel, K., & Taams, L. S. (2016). IL-17+ CD8+ T cells: Differentiation, phenotype and role in inflammatory disease. *Immunology Letters*, *178*, 20–26. https://doi.org/10.1016/j.imlet.2016.05.001
Stahl, E. A., Raychaudhuri, S., Remmers, E. F., Xie, G., Eyre, S., Thomson, B. P., … Plenge, R. M. (2010). Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nature Genetics*, *42*(6), 508–514. https://doi.org/10.1038/ng.582
Stinchcombe, J. C., & Griffiths, G. M. (2007). Secretory Mechanisms in Cell-Mediated Cytotoxicity. *Annual Review of Cell and Developmental Biology*, *23*(1), 495–517. https://doi.org/10.1146/annurev.cellbio.23.090506.123521
Stoeckle, C., Gouttefangeas, C., Hammer, M., Weber, E., Melms, A., & Tolosa, E. (2009). Cathepsin W expressed exclusively in CD8+ T cells and NK cells, is secreted during target cell killing but is not essential for cytotoxicity in human CTLs. *Experimental Hematology*, *37*(2), 266–275. https://doi.org/10.1016/j.exphem.2008.10.011
Stolwijk, C., van Onna, M., Boonen, A., & van Tubergen, A. (2016). Global Prevalence of Spondyloarthritis: A Systematic Review and Meta-Regression Analysis. *Arthritis Care and Research*, *68*(9), 1320–1331. https://doi.org/10.1002/acr.22831
Surdacki, A., Sulicka, J., Korkosz, M., Mikoałjczyk, T., Telesińska-Jasiówka, D., Klimek, E., … Grodzicki, T. K. (2014). Blood monocyte heterogeneity and markers of endothelial activation in ankylosing spondylitis. *Journal of Rheumatology*, *41*(3), 481–489. https://doi.org/10.3899/jrheum.130803
Taams, L. S., Steel, K. J. A., Srenathan, U., Burns, L. A., & Kirkham, B. W. (2018). IL-17 in the immunopathogenesis of spondyloarthritis. *Nature Reviews Rheumatology*, *14*(8), 453–466. https://doi.org/10.1038/s41584-018-0044-2
Takedatsu, H., Michelsen, K. S., Wei, B., Landers, C. J., Thomas, L. S., Dhall, D., … Targan, S. R. (2008). TL1A (TNFSF15) Regulates the Development of Chronic Colitis by Modulating

Both T-Helper 1 and T-Helper 17 Activation. *Gastroenterology*, *135*(2), 552–567. https://doi.org/10.1053/j.gastro.2008.04.037

Talbert, P. B., & Henikoff, S. (2010, April 3). Histone variants ancient wrap artists of the epigenome. *Nature Reviews Molecular Cell Biology*. Nature Publishing Group. https://doi.org/10.1038/nrm2861

Tatti, M., Motta, M., Di Bartolomeo, S., Cianfanelli, V., & Salvioli, R. (2013). Cathepsin-mediated regulation of autophagy in saposin C deficiency. *Autophagy*. Taylor and Francis Inc. https://doi.org/10.4161/auto.22557

Taurog, J. D., Chhabra, A., & Colbert, R. A. (2016). Ankylosing spondylitis and axial spondyloarthritis. *New England Journal of Medicine*, *374*(26), 2563–2574. https://doi.org/10.1056/NEJMra1406182

Taurog, J. D., Dorris, M. L., Satumtira, N., Tran, T. M., Sharma, R., Dressel, R., … Reichardt, H. M. (2009). Spondylarthritis in HLA-B27/human β2-microglobulin- transgenic rats is not prevented by lack of CD8. *Arthritis and Rheumatism*, *60*(7), 1977–1984. https://doi.org/10.1002/art.24599

Thomas, G. P., & Brown, M. A. (2010). Genetics and genomics of ankylosing spondylitis. *Immunological Reviews*, *233*, 350–373. https://doi.org/10.1016/B978-0-12-374934-5.00029-5

Toubal, A., Nel, I., Lotersztajn, S., & Lehuen, A. (2019). Mucosal-associated invariant T cells and disease. *Nature Reviews Immunology*, *19*(10), 643–657. https://doi.org/10.1038/s41577-019-0191-y

Tran, T. M., Dorris, M. L., Satumtira, N., Richardson, J. A., Hammer, R. E., Shang, J., & Taurog, J. D. (2006). Additional human β2-microglobulin curbs HLA-B27 misfolding and promotes arthritis and spondylitis without colitis in male HLA-B27-transgenic rats. *Arthritis and Rheumatism*, *54*(4), 1317–1327. https://doi.org/10.1002/art.21740

Treiner, E., Duban, L., Bahram, S., Radosavljevic, M., Wanner, V., Tilloy, F., … Lantz, O. (2003). Erratum: Selection of evolutionarily conserved mucosal-associated invariant T cells by MR1 (Nature (2003) 422 (164-169)). *Nature*, *423*(6943), 1018.

Trynka, G., Hunt, K. A., Bockett, N. A., Romanos, J., Mistry, V., Szperl, A., … Van Heel, D. A. (2011). Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. *Nature Genetics*, *43*(12), 1193–1201. https://doi.org/10.1038/ng.998

Trynka, G., & Raychaudhuri, S. (2013). Using chromatin marks to interpret and localize genetic associations to complex human traits and diseases. *Current Opinion in Genetics and Development*, *23*(6), 635–641. https://doi.org/10.1016/j.gde.2013.10.009

Trynka, G., Sandor, C., Han, B., Xu, H., Stranger, B. E., Liu, X. S., & Raychaudhuri, S. (2013). Chromatin marks identify critical cell types for fine mapping complex trait variants. *Nature Genetics*, *45*(2), 124–130. https://doi.org/10.1038/ng.2504

Tsoi, L. C., Stuart, P. E., Tian, C., Gudjonsson, J. E., Das, S., Zawistowski, M., … Elder, J. T. (2017). Large scale meta-analysis characterizes genetic architecture for common psoriasis associated variants. *Nature Communications*, *8*(May), 1–8. https://doi.org/10.1038/ncomms15382

Turk, V. (2001). NEW EMBO MEMBERS' REVIEW: Lysosomal cysteine proteases: facts and opportunities. *The EMBO Journal*, *20*(17), 4629–4633. https://doi.org/10.1093/emboj/20.17.4629

Turner, B. M. (2005, February 20). Reading signals on the nucleosome with a new nomenclature for modified histones. *Nature Structural and Molecular Biology*. Nat Struct Mol Biol. https://doi.org/10.1038/nsmb0205-110

Turner, M. J., DeLay, M. L., Bai, S., Klenk, E., & Colbert, R. A. (2007). HLA-B27 up-regulation causes accumulation of misfolded heavy chains and correlates with the magnitude of the unfolded protein response in transgenic rats: Implications for the pathogenesis of spondylarthritis-like disease. *Arthritis and Rheumatism*, *56*(1), 215–223. https://doi.org/10.1002/art.22295

Turner, M. J., Sowders, D. P., DeLay, M. L., Mohapatra, R., Bai, S., Smith, J. A., … Colbert, R. A. (2005). HLA-B27 Misfolding in Transgenic Rats Is Associated with Activation of the Unfolded Protein Response. *The Journal of Immunology*, *175*(4), 2438–2448.

https://doi.org/10.4049/jimmunol.175.4.2438
Urrutia, A., Duffy, D., Rouilly, V., Posseme, C., Djebali, R., Illanes, G., … Albert, M. L. (2016). Standardized Whole-Blood Transcriptional Profiling Enables the Deconvolution of Complex Induced Immune Responses. *Cell Reports*, *16*(10), 2777–2791. https://doi.org/10.1016/j.celrep.2016.08.011

Van Der Heijde, D., Ramiro, S., Landewé, R., Baraliakos, X., Van Den Bosch, F., Sepriano, A., … Braun, J. (2017). 2016 update of the ASAS-EULAR management recommendations for axial spondyloarthritis. *Annals of the Rheumatic Diseases*, *76*(6), 978–991. https://doi.org/10.1136/annrheumdis-2016-210770
Van Der Wijst, M. G. P., De Vries, D. H., Brugge, H., Westra, H. J., & Franke, L. (2018). An integrative approach for building personalized gene regulatory networks for precision medicine. *Genome Medicine*, *10*(1), 1–15. https://doi.org/10.1186/s13073-018-0608-4
Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., & Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*, *3*(7), 1–12. https://doi.org/10.1186/gb-2002-3-7-research0034
Vecellio, M., Roberts, A. R., Cohen, C. J., Cortes, A., Knight, J. C., Bowness, P., & Wordsworth, B. P. (2016). The genetic association of RUNX3 with ankylosing spondylitis can be explained by allele-specific effects on IRF4 recruitment that alter gene expression. *Annals of the Rheumatic Diseases*, *75*(8), 1534–1540. https://doi.org/10.1136/annrheumdis-2015-207490
Vidak, E., Javoršek, U., Vizovišek, M., & Turk, B. (2019). Cysteine Cathepsins and their Extracellular Roles: Shaping the Microenvironment. *Cells*, *8*(3), 264. https://doi.org/10.3390/cells8030264
Võsa, U., Claringbould, A., Westra, H.-J., Jan Bonder, M., Deelen, P., Zeng, B., … Franke, L. (2018). Unraveling the polygenic architecture of complex traits using blood eQTL meta-analysis. *BioRxiv*, *18*, 10. https://doi.org/10.1101/447367
Wandstrat, A., & Wakeland, E. (2001). The genetics of complex autoimmune diseases: Non-MHC susceptibility genes. *Nature Immunology*, *2*(9), 802–809. https://doi.org/10.1038/ni0901-802
Ward, L. D., & Kellis, M. (2012). HaploReg: A resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Research*, *40*(D1), 930–934. https://doi.org/10.1093/nar/gkr917
West, N. R., Hegazy, A. N., Owens, B. M. J., Bullers, S. J., Linggi, B., Buonocore, S., … Uhlig, H. (2017). Oncostatin M drives intestinal inflammation and predicts response to tumor necrosis factor–neutralizing therapy in patients with inflammatory bowel disease. *Nature Medicine*, *23*(5), 579–589. https://doi.org/10.1038/nm.4307
Westra, H. J., & Franke, L. (2014). From genome to function by studying eQTLs. *Biochimica et Biophysica Acta - Molecular Basis of Disease*. https://doi.org/10.1016/j.bbadis.2014.04.024
Wex, T., Wex, H., Hartig, R., Wilhelmsen, S., & Malfertheiner, P. (2003). Functional involvement of cathepsin W in the cytotoxic activity of NK-92 cells. *FEBS Letters*, *552*(2–3), 115–119. https://doi.org/10.1016/S0014-5793(03)00895-0
Willenbring, R. C., & Johnson, A. J. (2017). Finding a balance between protection and pathology: The dual role of perforin in human disease. *International Journal of Molecular Sciences*, *18*(8), 1–18. https://doi.org/10.3390/ijms18081608
Wong-Baeza, I., Ridley, A., Shaw, J., Hatano, H., Rysnik, O., McHugh, K., … Kollnberger, S. (2013). KIR3DL2 Binds to HLA-B27 Dimers and Free H Chains More Strongly than Other HLA Class I and Promotes the Expansion of T Cells in Ankylosing Spondylitis. *The Journal of Immunology*, *190*(7), 3216–3224. https://doi.org/10.4049/jimmunol.1202926
Wray, B. N. R., Queensland, P. D., & Peter, M. (2008). Estimating Trait Heritability Genetic variation in a population can result from a variety of things . What are the ways we, *0*, 12–14.
Xu, T., Keller, A., & Martinez, G. J. (2019). NFAT1 and NFAT2 Differentially Regulate CTL Differentiation Upon Acute Viral Infection. *Frontiers in Immunology*, *10*(FEB), 184. https://doi.org/10.3389/fimmu.2019.00184
Yang, J., Ferreira, T., Morris, A. P., Medland, S. E., Madden, P. A. F., Heath, A. C., … Visscher,

P. M. (2012). Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nature Genetics*, *44*(4), 369–375. https://doi.org/10.1038/ng.2213

Ye, J., Gillespie, K. M., & Rodriguez, S. (2018). Unravelling the roles of susceptibility loci for autoimmune diseases in the post-GWAS era. *Genes*, *9*(8), 1–15. https://doi.org/10.3390/genes9080377

Zanotti, S., Smerdel-Ramoya, A., & Canalis, E. (2013). Nuclear factor of activated T-cells (NFAT)c2 inhibits notch receptor signaling in osteoblasts. *Journal of Biological Chemistry*, *288*(1), 624–632. https://doi.org/10.1074/jbc.M112.340455

Zeller, T., Wild, P., Szymczak, S., Rotival, M., Schillert, A., Castagne, R., … Cambien, F. (2010). Genetics and beyond - the transcriptome of human monocytes and disease susceptibility. *PLoS ONE*, *5*(5). https://doi.org/10.1371/journal.pone.0010693

Zervoudi, E., Saridakis, E., Birtley, J. R., Seregin, S. S., Reeves, E., Kokkala, P., … Stratikos, E. (2013). Rationally designed inhibitor targeting antigentrimming aminopeptidases enhances antigen presentation and cytotoxic T-cell responses. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(49), 19890–19895. https://doi.org/10.1073/pnas.1309781110

Zhang, K., Li, N., Ainsworth, R., & Wang, W. (2016). Systematic identification of protein combinations mediating chromatin looping. Nature Communications, 7(1). doi: 10.1038/ncomms12249

Zhernakova, A., Van Diemen, C. C., & Wijmenga, C. (2009). Detecting shared pathogenesis from the shared genetics of immune-related diseases. *Nature Reviews Genetics*, *10*(1), 43–55. https://doi.org/10.1038/nrg2489

Zhernakova, D. V., Deelen, P., Vermaat, M., Van Iterson, M., Van Galen, M., Arindrarto, W., … Franke, L. (2017). Identification of context-dependent expression quantitative trait loci in whole blood. *Nature Genetics*, *49*(1), 139–145. https://doi.org/10.1038/ng.3737

Chapter 9

RESULTS: PART II

Characterization of peripheral and entheseal IL-17 secreting Mucosal Associated Invariant T (MAIT) cells in Axial Spondyloarthritis

Nicolas Rosine ǂ, Hannah Rowe ǂ, Surya Koturan ǂ, et al. (submitted)

ǂ Equal contribution

Manuscript I

Characterization of peripheral and entheseal IL-17 secreting Mucosal Associated Invariant T (MAIT) cells in Axial Spondyloarthritis

Summary

The IL-23/IL-17 axis has been implicated in the pathogenesis of axSpA through association studies. However, the success of anti-IL-17A therapy and the failure of anti-IL-23 raised the possibility of IL-23-independent IL-17 production in SpA, and made us wonder what are the IL-17 producing cells in axSpA. In the third part of this project we have characterised the immune cells that produce IL-17A from 18 axSpA patients. We compared the IL-17 production capacity of five cell populations of the innate (MAIT, γδT, and neutrophils) and adaptive ($CD4^+$ and $CD8^+$ T-cells) arms of the immune system.

From analysing the secreted protein data, we identified MAIT cells as the main producers of IL-17 in axSpA, compared to CD4+ T, CD8+ T and γδ T cells. In our data, neutrophils did not produce IL-17A, although they have been suggested in the literature as the main producers of IL-17A in AS by Appel et al. We also generated gene expression data for *IL-17A* using nCounter® technology, that complements our protein data with MAIT cells showing the highest level of IL17A expression in these cells, followed by $CD4^+$ T cells. We also noted a significantly higher expression of *IL-17F* and *IL-23R* in MAIT cells. Additionally, we studied the stimulation conditions that induce *IL17F* in MAIT cells, and observed that TCR activation associated with IL-7 or with IL-18 induced a high expression of *IL17F* by MAIT cells.

Our results emphasize the importance of MAIT cells in the SpA pathogenesis. This premise is further supported by the tissue data collected in collaboration with the McGonagle's team at University of Leeds, who could detect the presence of MAIT cells in the entheseal samples from healthy donors. Our data suggest the involvement of MAIT cells in the inflammatory events in enthesitis - a signature of SpA pathogenesis.

It is also important to note that both innate and adaptive lymphocytes express genes belonging to the IL-23/IL-17 pathway, and genes previously associated with axSpA susceptibility. The detailed manuscript, figures and supplementary information is provided in the following pages.

Manuscript draft

Characterization of peripheral and entheseal IL-17 secreting Mucosal Associated Invariant T (MAIT) cells in Axial Spondyloarthritis

Nicolas Rosine[1,2ł], Hannah Rowe[3ł], Surya Koturan[1,2ł], Hanane Yahia-Cherbal[1], Claire Leloup[1], Abdulla Watad[3], Francis Berenbaum[4,5], Jeremie Sellam[4,5], Maxime Dougados[6,7,8], Vishu Kumar Aimanianda Bopaiah[9], Charlie Bridgewood[3], Elisabetta Bianchi[1,8], Lars Rogge[1,8†], Dennis G McGonagle[3†], Corinne Miceli-Richard[1,7,8†].

[1] Institut Pasteur, Immunoregulation Unit, Department of Immunology, Paris, France
[2] Paris Diderot Université, Paris 7, Sorbonne Paris Cité, Paris, France
[3] University of Leeds Institute of Rheumatic and Musculoskeletal Medicine, Leeds, UK
[4] Sorbonne Université, Service de Rhumatologie, Hôpital Saint-Antoine, AP-HP, 75012, Paris, France
[5] Centre de Recherche Saint-Antoine, INSERM UMR_S 938, Paris, France
[6] INSERM (U1153) Clinical epidemiology and biostatistics, PRES Sorbonne Paris-Cité, Paris, France
[7] Paris Descartes Université, Service de Rhumatologie, Hôpital Cochin, AP-HP, 75014, Paris, France
[8] Unité Mixte AP-HP/ Institut Pasteur, Institut Pasteur, Immunoregulation Unit, Paris, France
[9]Institut Pasteur, Molecular mycology Unit, Paris France

NR, HR and SK contributed equally

LR, CMR and DGM contributed equally

Key words: IL-17A, IL-17F, IL-17 producing cells, axial spondyloarthritis, MAITs

Correspondence to:
Corinne Miceli-Richard, Institut Pasteur, Immunoregulation Unit, Department of Immunology, 25 rue du Docteur Roux, 75015 Paris, France
E-mail: corinne.miceli@pasteur

Dennis G McGonagle, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Leeds, united Kingdom
E-mail: d.g.mcgonagle@leeds.ac.uk

Abstract

Objectives: The failure of anti-IL-23 in Axial Spondyloarthritis (AxSpA) raised questions about the nature of IL-17-producing cells involved in AxSpA pathogenesis. We undertook this study to identify and characterize the major IL-17A-producing cell populations in AxSpA.

Methods: We compared IL-17A production capacity of 5 sorted cell populations from 18 AxSpA patients: neutrophils, MAIT, γδ T cells, CD4+T and CD8+T cells after stimulation by PMA+A23187+β1,3 glucan. IL-17A expression and production were assessed with ultra-sensitive technology (SimoA technology for protein production and Nanostring Technology for gene expression). We assessed the presence of MAIT in normal human enthesis soft tissue (EST) and adjacent perientheseal bone (PEB) (n=5) by immunophenotyping.

Results: On a per cell basis, we observed that MAIT produced the highest amount of IL-17A compared to CD4+T ($p < 0.01$), CD8+T ($p < 0.0001$) and γδ T cells ($p < 0.0001$). Neutrophils did not produce IL-17A. The results were confirmed by gene expression analysis. Significantly higher *IL23R* and *IL17F* expression were observed in MAIT compared to other cell types. Stimulation of MAIT with αCD3/CD28 in the presence of IL-23 did not induce *IL17A* expression. MAIT cells express high levels of IL-18 receptor subunits and stimulation with αCD3/CD28 associated to IL-7 or IL-18 induced a strong expression of *IL17F*. MAITs were present in both EST and PEB from normal human enthesis and were characterized by an immunoregulatory phenotype based on the expression of JAK1, STAT4 and TGFb1.

Conclusion: Innate T cells could play an important role in IL-17A production in AxSpA. MAIT cells displayed the highest production capacity for IL-17A compared to CD4+T, γδT, and CD8+T. Addition of IL-23 did not increase *IL17A* and *IL17F* expression by these cells. The combination of IL-7 and IL-18 on a background of CD3/CD28 stimulation induced high levels of IL-17F. The presence of MAIT cells in normal human enthesis suggest their key role in axial SpA pathogenesis.

Key messages

What is already known about this subject?

The success of anti-IL-17A therapy in Axial Spondyloarthritis has confirmed that IL-17A is a critical cytokine in the pathophysiology of the disease. However, the main IL-17 producing cell populations in AxSpA have not been clearly identified yet.

What does this study add?

MAIT have the highest IL-17A production capacity in AxSpA compared to CD4 +T, γδT and CD8+T cells and are present in human enthesis. These cells are also able to express high levels of *IL-17F* after IL-7 and IL-18 stimulation combined with CD3/CD28. Addition of IL-23 to CD3/CD28 stimulation does not induce the expression of *IL17A* or *IL17F* by MAIT.
Neutrophils do not produce IL-17A in AxSpA after a strong stimulation with β1,3 glucan. MAIT cells are present in normal enthesis and transcriptomic analysis confirmed expression on *IL-17A* and *IL-17F*.

How might this impact on clinical practice or future developments?

Further studies on MAIT cells might help to describe the IL-23-independent IL17 expression and to understand the failure of anti-IL-23. Expression of IL-17F by MAIT cells and their presence in normal human enthesis gives arguments in favour of the effectiveness of anti-IL-17A and anti-IL-17F in axial disease.

Introduction

Spondyloarthritis (SpA) is one of the most common chronic inflammatory rheumatic conditions, with prevalence in the general population ranging from 0.5% to 1.9%[1]. In addition to disabling rheumatologic manifestations, some patients with SpA develop severe extra-articular manifestations such as chronic inflammatory bowel disease (IBD), uveitis and psoriasis (Pso). The therapeutic challenge in 2020 and for the coming years will be to identify new therapeutics that are effective on the whole spectrum of clinical manifestations observed in SpA. In fact, SpA mainly affects young adults and the functional consequences of a poorly controlled disease lead to an alteration in their quality of life and professional capacity, which has a direct impact on healthcare costs. The appearance of biologics in the rheumatology therapeutic arsenal has revolutionized the treatment of chronic inflammatory diseases such as SpA. The first biologics available in this indication were those targeting the Tumor Necrosis Factor (TNF). However, clinical studies showed that 30 to 40% of patients did not respond to or acquired resistance to anti-TNF[2]. It was therefore important to continue to expand our therapeutic arsenal.

With the development of technologies available for biomedical research, the in-depth analysis of patient samples has made it possible to demonstrate the role of the IL-23/IL-17 axis in the physiopathology of the disease[3]. The importance of this axis was first supported by the discovery of genetic associations between several gene variants involved in IL-23/IL-17 pathway with SpA (*IL-23Rc, IL-12B, IL-6Rc*, etc.). As IL-17 is the terminal cytokine of this pathophysiological pathway, the development of new treatments has initially focused on blocking this cytokine. In fact, clinical trials have demonstrated the efficacy of biologics targeting IL-17 (secukinumab SCK and ixekizumab IXE) in the axial (AxSpA) [4, 5] and peripheral forms of SpA (psoriatic arthritis - PsA) [6-9].

More than ten years ago, scientific work has shown that IL-23 plays a crucial role in the production of IL-17 by maintaining the differentiation state of Th17 lymphocytes, the main known cellular source of IL-17 secretion[10]. IL-23 is a heterodimer composed of one IL12p40 and one IL23p19 subunit. Whereas the genetic associations of variants of the IL-17/IL-23 pathway had been demonstrated in the axial form of the disease (ankylosing spondylitis - AS)[2], surprisingly, clinical trials targeting IL-23, either via IL23p19 (risankisumab) or IL-12p40 (ustekinumab USK), failed in the axial forms of the disease[11,12] while their use in the peripheral forms was effective[13,14]. These results suggest a distinct pathophysiology between the different forms of the disease and underline our incomplete understanding of the mechanisms of action of these molecules during the course of the disease. One of the hypotheses put forward to explain this discordant effect of IL-23 targeting on axial and peripheral forms of SpA is that other IL-17 source cells may be independent of IL-23[15-17]. Assuming this, therapeutic strategies targeting IL-23 would not be effective enough in preventing the secretion of the effector cytokine (IL-17) and would be ineffective in treating the axial form in which these cells would be involved.

Several IL-17A producing cells have been described. Among those, CD4+ Th17 cells have been considered for several years as the main IL-17 producing cell type. Their increased prevalence has been reported in the peripheral blood of AxSpA patients[18]. Nevertheless, alternative sources of IL-17 have been further reported, in particular innate T lymphocytes: γδTs and MAITs. Their increased frequency in AxSpA has been reported by *Kenna et al.* for γδT[19] and *Gracey al.* for MAITs[17], suggesting their involvement in disease pathogeny. Neutrophils have been also described as IL-17A producing cells. *Appel et al.* suggested that they might be the main producers of IL-17A at the facet joint level in AS[20]. Nevertheless, the ability of neutrophils to produce IL-17A is very controversial[21].

The recent failure of Rizankimumab in AxSpA, an anti-IL-23p19 yet effective in psoriasis and psoriatic arthritis, questions the predominant role of Th17[11] in the context of AxSpA. The ineffectiveness of blocking IL-23, which is essential to stabilize the phenotype of these cells, suggests that other cell types

can produce IL-17, at least partially independently from IL-23. According to the literature, MAIT, γδT, and neutrophils represent potential candidates in that respect.

The objective of the present study was to compare the respective IL-17 production capacity of different cell subsets belonging to the adaptive (CD4+ and CD8+ T-cells) and innate (MAIT, γδT, and neutrophils) immune system in AxSpA and to study the conditions associated with their best profiles of *IL17* expression.

Material and Methods

Patients and samples

Blood samples from 18 patients with a clinical diagnosis of axial Spondyloarthritis (AxSpA) fulfilling the Assessment of SpondyloArthritis international Society criteria[22] were included in two tertiary centers (Cochin and Saint-Antoine Hospitals – Paris – France).

Human interspinous process and matched peripheral blood were obtained from 5 non-AxSpA patients who underwent elective spinal surgery for either decompression or scoliosis correction using methods previously reported[23]. All patients provided written informed consent before enrollment in the study as approved by the French Ethics Committee and by the north West-Greater Manchester West Research Ethics Committee. The clinical characteristics of the AxSpA patients and of the non-AxSpA patients who underwent spinal surgery are summarized in supplementary table 1 (**Annex II**).

Samples from AxSpA patients

Cell sorting

Neutrophils were isolated by negative magnetic cell sorting using MACSxpress Whole Blood Human Neutrophils isolation kit (Milteny Biotec®) according to the manufacturer's instructions.
For T cells, PBMCs were isolated from blood using lymphocyte separation medium (Eurobio®). After isolation, PBMCs were labeled with CD3 BUV395 (BD Biosciences®), CD4 VioBright FITC (Milteny Biotec®), CD8 PerCP vio700 (Milteny Biotec®), TCR Vδ PE (Milteny Biotec®), TCR Vδ2 PE (Milteny Biotec®), TCR Vα7.2 APC (BioLegend®), CD161 BV421 (Sony Biotecnology®). PBMCs subsets were isolated using a BD FACS Aria II according to the gating strategy presented in **Supplementary Figure 1** (**Annex II**).
For the experiments on healthy blood donors (n=3), we performed first a CD3 positive magnetic separation using anti-CD3 monoclonal antibodies coated beads (Milteny Biotec®) then the CD3 positive fraction was labeled with CD3 APC Vio770 (Milteny Biotec®) CD4 VioBright FITC (Milteny Biotec®), Vα7.2 APC (BioLegend®), CD161 BV421 (Sony Biotechnology®) CCR6 BV786 (BD Biosciences®) and CD3 subsets were isolated using a second gating strategy presented in **Supplementary Figure 2** (**Annex II**).

Cell stimulations and cultures

Neutrophils were plated in 48 well plates ($5x10^6$ cells/well) and cultured in complete media alone (Roswell Park Memorial Institute (1640 medium RPMI Invitrogen®)). Sorted T cells were plated in 48 well plates ($1x10^6$ cells/well) or 96 well plates ($2x10^5$ cells/well) according to the number of cells

obtained after sorting, and cultured in complete media (Roswell Park Memorial Institute (1640 medium RPMI Gibco life technologies®)) with 5% heat inactivated fetal calf serum (Hyclone, Fischer Scientific®) and penicillin/streptomycin.

Where applicable, cells were stimulated with Phorbol 12-mristate 13-acetate (50ng/ml) (Merck®), A23187 (5µM) (Merck®), IL-7 (20ng/ml) (Milteny Biotec®), IL-18 (50ng/ml) (R and D System®) IL-23 (20ng/ml) (Milteny Biotec®), Dynabeads Human T-activator CD3/CD28 (Thermo Fisher Scientific®).

β1,3 glucan (50µg/ml) was isolated from the alkali-insoluble (AI) fraction of the *A. fumigatus* mycelial cell-wall as described earlier[9,10].

Gene expression analysis

Gene expression profiles from AxSpA patients sorted cells were assessed using the nCounter® Autoimmune Discovery Consortium codeset (NanoString Technologies®). The resultant Reporter code count (RCC) data were imported in nSolver Analysis Software® for quality control and normalization. Transcriptional profiling of entheseal and blood MAITs was performed on non-AxSpA patients. MAITs cells were isolated from EST, PEB and peripheral blood from patients subjected to spinal surgery. Basal expression of cytokines, chemokines, growth factors, signalling molecules, tissue residency markers was assessed.

See supplementary materials for details (RNA preparation and gene expression analysis- **Annex II**).

Protein expression analyses

IL-17A concentrations (in fg/ml) in cell culture surpernatants from AxSpA patients were determined with Simoa IL-17A 2.0 Reagent kit (Quanterix corp. Lexington, MA 02421) using HD-1 Analyzer (Quanterix ®) (see supplementary Materials). The LOD was calculated by the blank+3SDs. When IL-17A was undetectable, we replaced by LOD value for presentation and analysis purpose. The lowest LOD was 7 fg/ml.

Intracellular TNF and IL-17 cytokine expression with and without stimulation with PMA/ionomycin was assessed in the presence of Golgi Plug in PEB derived MAIT cells from non-SpA patients subjected to spinal surgery.

Samples from non-AxSpA

Isolation of primary cells from entheses and match blood

Entheseal samples were separated into entheseal soft tissue (EST) and perientheseal bone (PEB) and both were enzymatically digested as previously described[23]. For both cell preparations, blood and

entheseal cells, density gradient separation (Lymphoprep) was conducted in order to obtain PBMCs and entheseal mononuclear cells (EMCs), respectively, as previously described[24].

Immunophenotyping and sorting of entheseal and peripheral blood MAIT cells

EMCs or PBMCs were stained with zombie aqua (live/dead discrimination), anti-CD45 (to exclude non-leucocytes), CD3 (T-cell inclusion). MAIT cells were identified by CD161+ and Vα7.2 TCR+. All flow cytometry was conducted on the Cytoflex LX (Beckman Coulter) and subsequent analysis completed using the CytExpert Acquisition and Analysis Software (V.2.3) and FlowJo software (Tree Star, USA). A full list of antibodies, clones and fluorophores can be found in online **Supplementary Table 3** (**Annex II**).

Statistical analysis

GraphPad Prism software was used for statistical analyses (GraphPad Software, La Jolla, California, USA).

For all graphs, non-significant results ($p>0.05$) were indicated, *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p>0.0001$. Error bars represent the SE of the mean (SEM).

Qlucore omics explorer 3.5 was used to generate heatmaps with unsupervised hierarchical clustering after applying to the data log transformation, mean centering, and unit variance scaling.

Results

Genes associated with AS susceptibility are differentially expressed in innate and adaptive T cell populations isolated from peripheral blood of axSpA patients

More than 40 loci have been significantly associated with AS susceptibility[2]. Among those, some are involved in IL-23/IL-17 expression pathways such as *IL-1R, IL-6R, TYK2, RUNX3*, and *IL-12B*. To better decipher the role of MAITs, CD4+, CD8+ and γδT-cells in AS pathogeny, we specifically analyzed the expression profiles of genes significantly associated to AS susceptibility, including those belonging to the IL-17/IL-23 pathways. The expression of 36 genes from the total of 45 AS-associated genes panel was detected in four T cell populations isolated from 9 axSpA patients and stimulated with PMA (50ng/ml) + A23187(5μM) + β-glucan (50μg/ml). The expression pattern observed after hierarchical clustering showed a clear distinction between the innate and adaptive T cell groups as shown in the heatmap (**Figure 1A**). Gene clusters consisting of members of the MHC class I mediated antigen processing & presentation - *NPEPPS* and *UBE2L3* - were expressed in CD4+ and CD8+ T cells. We observed genes that were expressed at relatively higher levels in a specific cell type, e.g., *PTGER4* in CD4+ T cells and *TYK2* in CD8+ T cells. MAIT cells express high levels of *IL23R* and the G protein-coupled receptors *GPR35* and *GPR65*. We also noted cell type-specific expression of several IL-23/IL-17 pathway genes (**Figure 1B**). *IL17F* was expressed about one log higher in CD4+ and MAIT cells, while *IL23R* was expressed two logs higher in MAIT and one log higher in γδT cells compared to CD4+ and CD8+ T cells. *NFKB1*, *RELA* and *NFKBIA* are preferentially expressed in CD4+ T cells, while several genes encoding cytokines and their receptors (*IL23A*, *IL23R, IL12RB1, IL18R1, IL18RAP, TNF* and *IFNG*) were expressed at higher levels in "innate" MAIT and γδT cells when compared to "adaptive" CD4+ and CD8+ T cells. *IL1R1*, *TYK2* and *RUNX3* were expressed at high levels in CD8+ T cells. Nevertheless, many other genes, not belonging to IL-17/IL-23 pathway, participated in cell clustering suggesting that those different cell types were involved in AS susceptibility beyond their relative role in IL-17/IL-23 pathway.

Figure 1

A

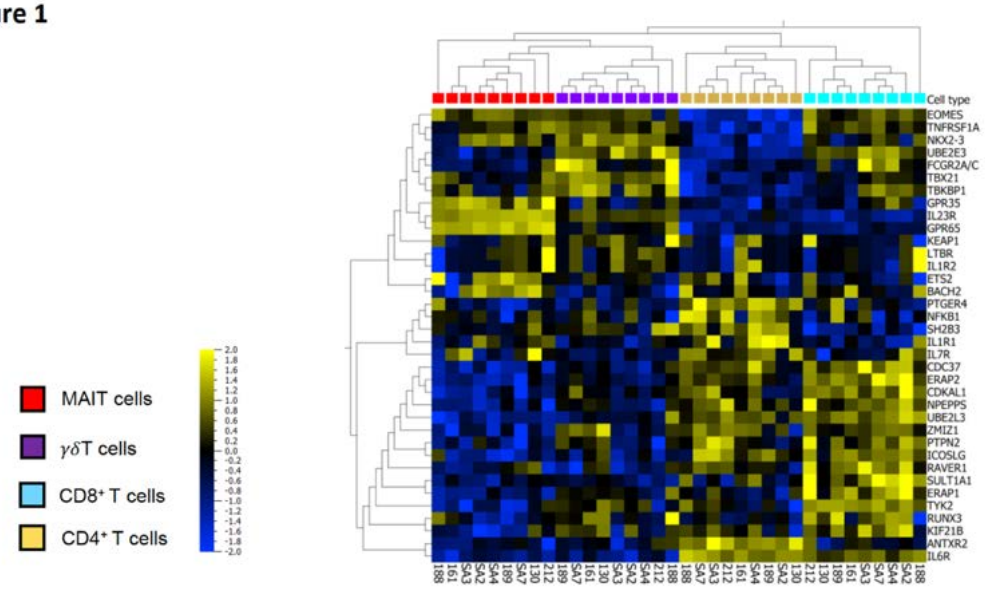


Figure 1A: Heatmap showing the expression of genes associated with AS in T cell subpopulations isolated from peripheral blood of axSpA patients. The expression of 36 genes genetically linked to AS in T cells from axSpA patients after 2h of stimulation by PMA (50ng/ml) + A23187(5µM) + β-glucan (50µg/ml). Cell populations are colour-coded as shown by the legend on the left. Columns in the heatmap represent patient samples (n=9) and the rows represent genes. The heatmap is ordered by hierarchical clustering. Gene expression data are $\log_2$ transformed, centred to a mean value of zero and scaled to unit variance. The bar on the left denotes the scale for the gene expression levels- yellow indicates higher- and blue lower levels of expression.

B

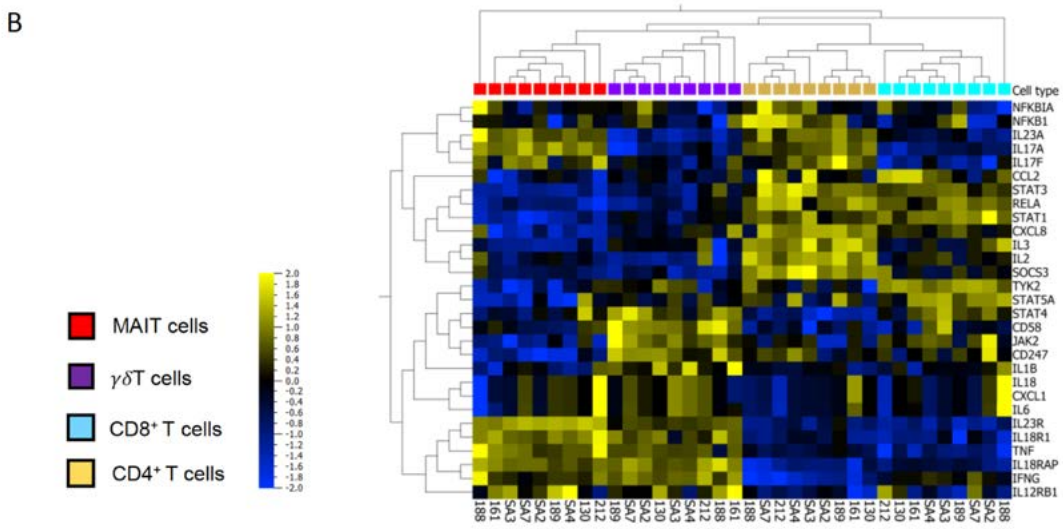


Figure 1B: The heatmap depicts the gene expression pattern of IL-23/IL-17 pathway genes in T cell populations isolated from axSpA patients as above. Genes associated with the IL-23 and IL-17 pathways were obtained from the molecular signature database (MSigDB) were selected. Shown are the mRNA levels of 29 genes after 2h of stimulation by PMA (50ng/ml) + A23187(5µM) + β-glucan (50µg/ml). Cell populations are colour-coded as shown by the legend on the left. Columns in the heatmap represent patient samples (n=9) and the rows represent genes. The heatmap is ordered by hierarchical clustering. Gene expression data are $\log_2$ transformed, centred to a mean value of zero and scaled to unit variance. The bar on the left denotes the scale for the gene expression levels- yellow indicates higher- and blue lower levels of expression.

Peripheral blood derived MAITs cells have a high potential of IL-17A and IL-17F secretion.

To better define the relative role of MAITs, CD4+, CD8+ and γδT cells in IL-17 expression, we sorted these cells, together with neutrophils which have been controversially suspected to secrete IL-17A [11]. Cell sorting was performed on peripheral blood from 18 AxSpA patients (a representative cytometry panel with the gating strategy is shown in **supplementary Figure 1- Annex II**). We compared IL-17A secretion on a per cell basis. We divided the IL-17A concentration by the number of sorted cells and normalized by the volume in which they were stimulated. Since not all cells from a population are likely to produce the cytokine, we expressed the result in femtograms (fg) per 1000 cells. This showed a high level of IL-17A production by MAIT cells, significantly higher than CD4+T ($p < 0.01$), γδT ($p < 0.0001$), CD8+T ($p < 0.0001$) and neutrophils ($p < 0.0001$). Although lower than in MAIT (mean = 478.60 fg / 1000 cells), the IL-17A production capacity by CD4+T was high (mean = 128.65 fg / 1000 cells). γδT cells produced a small amount of IL-17A with a mean of 13.71 fg / 1000 cells, and in the same range than IL-17A produced by CD8+T cells (mean = 4.66 fg / 1000 cells). βglucan is the main component of *Aspergillus Fumigatus hyphae*. Different teams have demonstrated a potential role of the hyphae of these fungi on the production of IL-17A by human neutrophils [25]. Despite of a strong β-glucan-associated stimulation, most neutrophils' samples did not exceed the detection limit (**Figure 2A**). Gene expression analysis confirmed protein data, MAIT cells displaying the highest level of *IL17A* expression followed by CD4+T cells. We noted low levels of *IL17A* expression in γδT and CD8+T cells. For neutrophils, *IL17A* was undetectable, results which were consistent with protein data (**Figure 2B**).

Figure 2

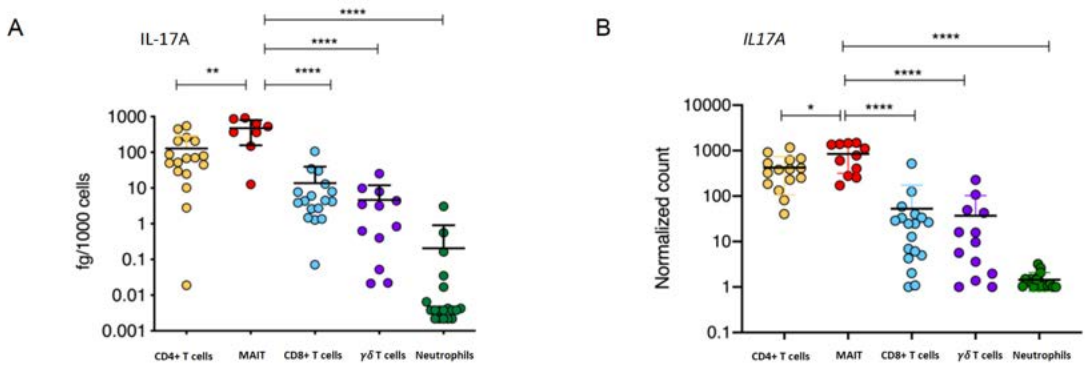


Figure 2: IL-17A protein production and *IL17A* gene expression by sorted cells from Axial SpA patients. (A) Left, IL-17A protein levels presented as femtograms per 1000 cells in sorted CD4+ T cells (in yellow), MAIT cells (in red), CD8+T cells (in blue), $\gamma\delta$ T cells (in purple) and neutrophils (in green) after 18h of stimulation by PMA (50ng/ml) + A23187(µM) + βglucan (50µg/ml). P-values were calculated using a Mann-Whitney T test. Horizontal bars indicate the median and SD. Differences are considered significant for P-values < 0.05. ****, P< 0.0001, ***, P<0.001, **, P<0.01, *, P<0.05, ns; not significant. (B) Right, Transcripts levels of *IL17A* presented as normalized counts in sorted CD4+ T cells (in yellow), MAIT cells (in red), CD8+T cells (in blue), $\gamma\delta$ T cells (in purple) and neutrophils (in green) after 2h of stimulation by PMA (50ng/ml) + A23187(µM) + βglucan (50µg/ml). P-values were calculated using a Mann-Whitney T test. Horizontal bars indicate the median and SD. Differences are considered significant for P-values < 0.05. ****, P< 0.0001, ***, P<0.001, **, P<0.01, *, P<0.05, ns; not significant

Next, we expanded our analysis to *IL17F*, *IL23R* and IFNγ expression. We observed that the expression level of *IL17F* displayed the same ranking as *IL17A* for the 5 cell populations (**Figure 3A**). Nevertheless, there was no significant difference between CD4+ and MAIT. We also observed a lower level of expression of *IL17F* compared with that of *IL17A* in all cell subsets. Regarding *IL23R* (**Figure 3B**), MAIT had the highest level of expression, followed by γδT cells and CD4+T cells. As expected, CD8+ T-cells and neutrophils did not express significant levels of *IL23R*. Contrarily to IL-17, the expression level of *IFNγ* was high in all T cell populations (**Figure 3C**) suggesting cell-specific expression profiles for *IL17*.

Collectively, these data indicate that MAIT cells represent the main producer of IL-17A at a per-cell level and express high levels of *IL17F* and *IL23R* compared to the other IL-17A-producing cells in AxSpA.

Figure 3

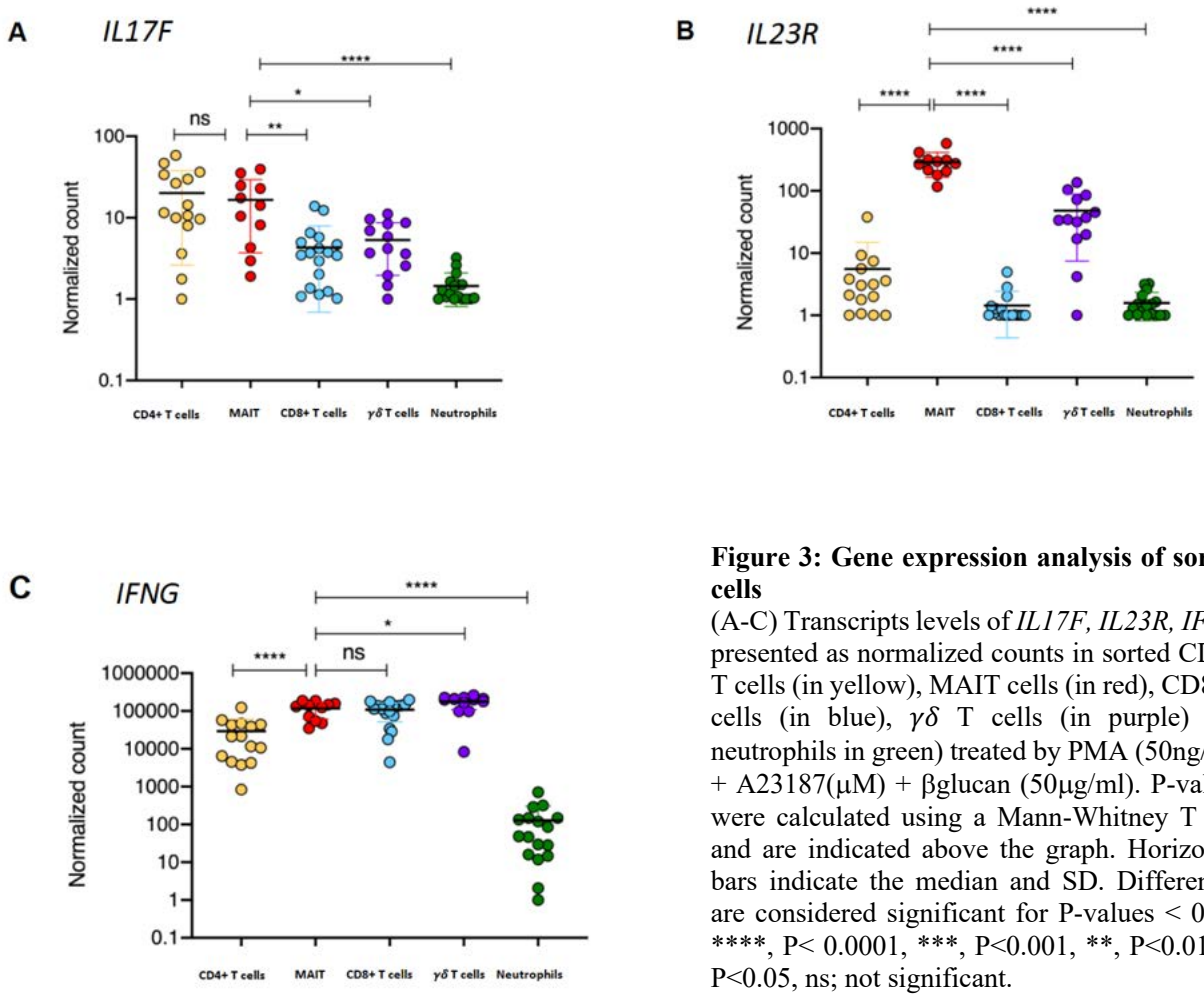


Figure 3: Gene expression analysis of sorted cells

(A-C) Transcripts levels of *IL17F, IL23R, IFNG* presented as normalized counts in sorted CD4+ T cells (in yellow), MAIT cells (in red), CD8+T cells (in blue), $\gamma\delta$ T cells (in purple) and neutrophils in green) treated by PMA (50ng/ml) + A23187(μM) + βglucan (50μg/ml). P-values were calculated using a Mann-Whitney T test and are indicated above the graph. Horizontal bars indicate the median and SD. Differences are considered significant for P-values < 0.05. ****, P< 0.0001, ***, P<0.001, **, P<0.01, *, P<0.05, ns; not significant.

Peripheral blood derived MAIT cells express IL-17F independently of IL-23 stimulation.

We further aimed to assessed which stimulation conditions were able to induce IL-17A and IL-17F expression by MAIT cells. We stimulated sorted MAIT (a representative flow cytometry gating strategy is shown on **Supplementary Figure 2- Annex II**) for 36 hours with anti CD3/CD28 alone or in combination with cytokines (IL-23, IL-7, IL-18 or both IL-7 and IL-18). We used CD4+CCR6+ T cells ("Th17-like" T cells) as control (**Figure 4**). We observed that a stimulation by CD3/CD28 associated with IL-7 or with IL-18 induced a high expression of *IL17F* by MAIT cells (**Figure 4B**). The combination of both 2 cytokines (IL-7 and IL-18) with CD3/CD28 even increased this level. Contrarily, we observed that addition of IL-23 to CD3/CD28 stimulation did not affect *IL17F* expression by MAIT cells (fig 3b). The expression of *IFNγ* was also remarkably high after a stimulation by CD3/CD28 associated with IL-7 or IL-18 (**Figure 4C**). We observed that CD3/CD28 stimulation alone or combined with IL-23 was not able to induce *IL17A* expression by MAIT cells. In comparison, CD4+CCR6+ T cells expressed high level of *IL17A* and *IL17F* after CD3/CD28 stimulation and CD3/CD28 stimulation with IL-23 (**Figure 4A**).

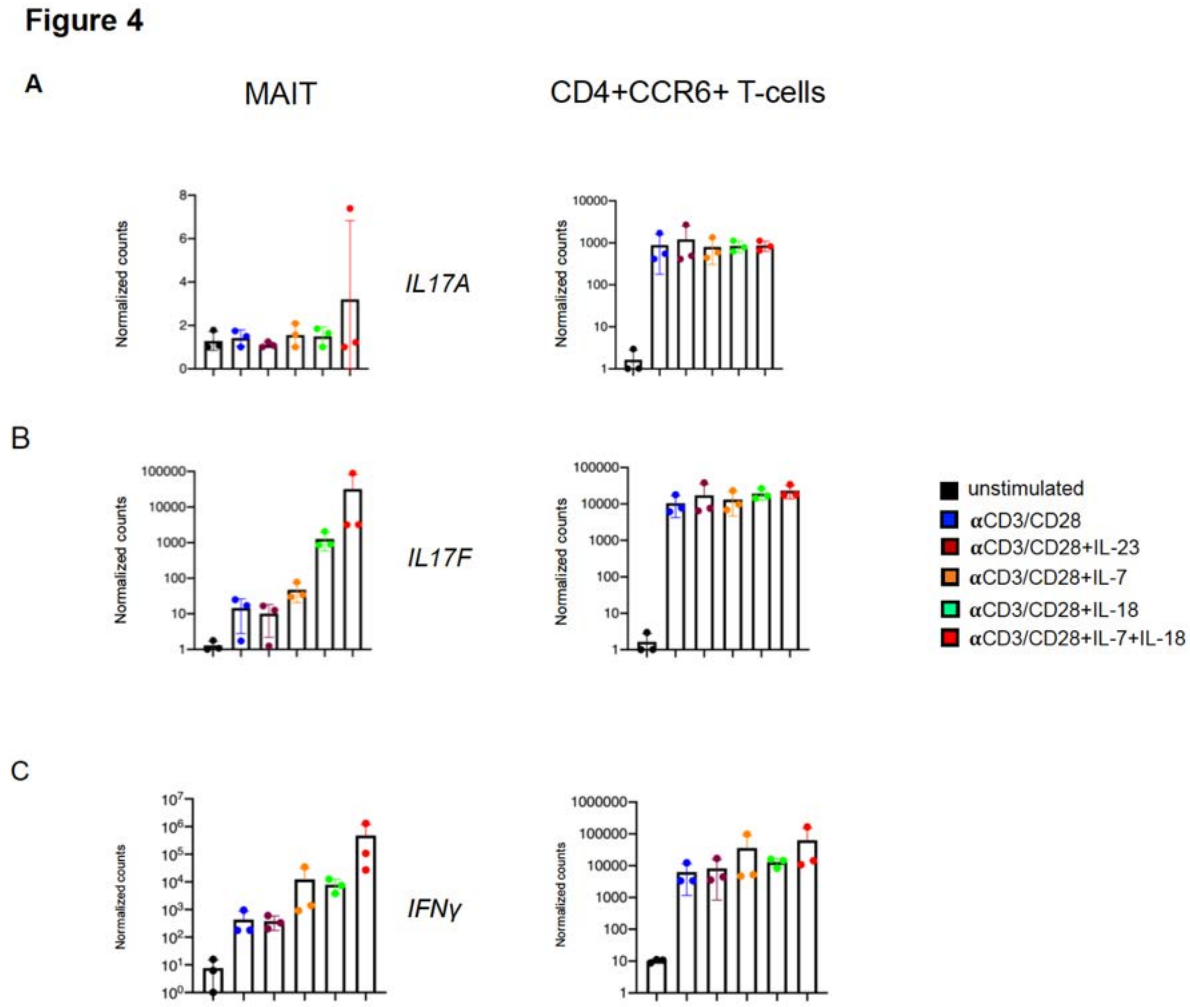


Figure 4: Stimulation conditions for IL-17 expression in sorted MAIT cells and CD4+CCR6+ T cells. (A-C) Transcripts levels of *IL17A, IL17F, IFNG* presented as normalized counts in sorted MAIT cells (left column) and CD4+CCR6+ T cells (right column) from donors (n=3) after 36 hours of stimulation in 6 different conditions: unstimulated (black), CD3/CD28 stimulation (blue), CD3/CD28 stimulation + IL-23 (20ng/ml) (burgundy), CD3/CD28 stimulation + IL-7(20ng/ml) (orange), CD3/CD28 stimulation + IL-18 (50ng/ml) (green), CD3/CD28 stimulation + IL-7 (20ng/ml)+ IL-18 (50ng/ml) (red).

MAIT cell are present in human entheses

Considering that the hallmark of SpA pathogenesis is an entheses inflammation, we further aimed to assess the presence of MAIT cells within human entheses. Human interspinous process and matched peripheral blood were obtained from 5 non-AxSpA and entheseal samples were separated into entheseal soft tissue (EST) and perientheseal bone (PEB). MAIT cells were identified as CD3 + cells expressing both CD161 and TCR Vα7.2. The gating strategy is shown on **supplementary Figure 4** (**Annex II**). MAIT cells were identified within human entheses as shown **Figure 5A**. Within both EST and PEB, MAIT cells mainly expressed the CD69 resident memory marker compared to MAIT cells from the peripheral blood which were mainly CD45RA+ corresponding to a naïve/circulating phenotype (**Figure 5A**).

Figure 5

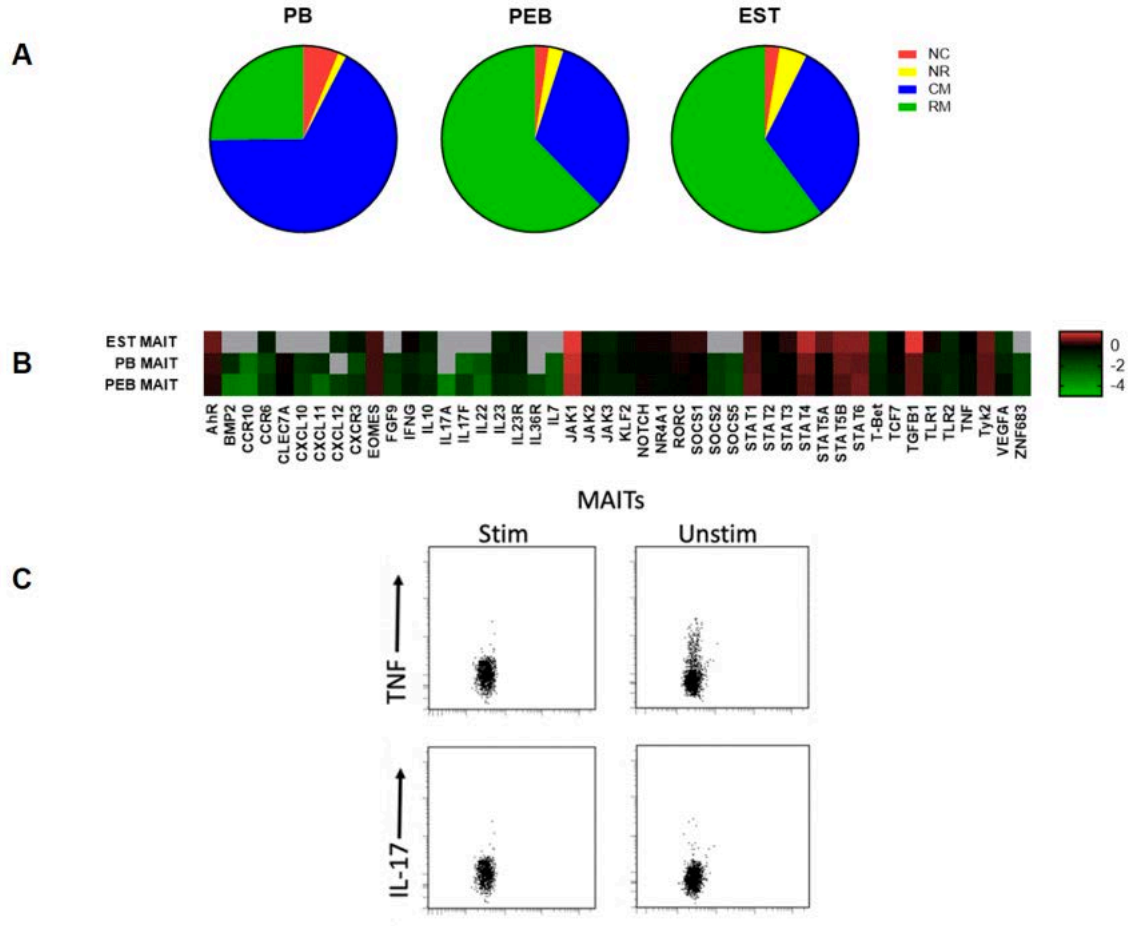


Figure 5: Entheseal MAIT cells Transcriptional profiling and pro-inflammatory cytokines induction. MAIT cells were sorted from Entheseal Soft Tissue (EST), Peri-entheseal Bone (PEB) and Peripheral Blood (PB) defined by their expression of Vα7.2 TCR and CD161. Cells expressing tissue resident (R)/memory (M) markers were identified by CD69+ and naïve (N)/circulating (C) cells by CD45RA+. Results displayed as mean from n=5 (A). Basal expression of cytokines, chemokines, growth factors, signalling molecules, tissue residency markers was assessed. Colour coding refers to differentially expressed genes where values less than −1 indicate low expression and values greater than 1 indicate higher expression, those with grey boxes indicates no values, values displayed are log 10ΔCt relative to HPRT (n=7) (B). Intracellular TNF and IL-17 cytokine expression with and without stimulation with PMA/ionomycin in the presence of Golgi Plug in PEB derived MAIT cells (C).

Transcriptional profiling of entheseal MAIT cells compared to peripheral blood supports their immunomodulatory status

The comparison of blood- and entheseal-derived (from both PEB and EST) MAIT cell transcripts (Vα7.2+ CD161+) suggested that EST MAIT cells were more immunomodulatory due to higher JAK1, STAT4 and TGFβ1 transcript expression (**Figure 5B**). Contrastingly, transcripts indicative of circulating T-cells, such as KLF2 and T-bet showed higher transcript expression in blood T-cells (**Figure 5B**). Entheseal MAIT cells also showed higher expression of growth factors and molecules

associated with tissue repair and homeostasis, such as VEGFA and IL-10 when compared to matched peripheral blood (**Figure 5B**).

Significance between Blood and EST MAIT's in CCR6, JAK2, STAT4 and TGFβ1 following non-parametric testing with Dunn's post-hoc testing p=0.023, 0.036, 0.046 and 0.046, respectively. Significance between Blood and PEB MAIT's in CXCR3 and TCF7 following non-parametric testing with Dunn's post-hoc testing p=0.050 and 0.045 respectively (**Figure 2A**)

Discussion

This work reports for the first time the ability of MAITs to express IL-17A as well as IL-17F according to various cell stimulation conditions. These MAITs were present in the spinal entheses of healthy subjects where they mainly showed a profile of resident memory cells.

Before studying more precisely the role of MAITs in the secretion of IL-17, we analysed the gene expression profiles of 45 genes whose polymorphisms were significantly associated with the predisposition to AS in the 4 T lymphocyte populations sufficiently represented in peripheral blood: CD4+ and CD8+ T lymphocytes, γδ T-lymphocytes and MAITs. We observed the expression of 80% of these genes in these 4 cell types with a differential expression from one cell population to another leading to cell clustering induced by these susceptibility genes. These new findings suggest that the functional consequences of polymorphisms carrying the susceptibility genes for AS should be studied in the corresponding cell subsets. A previous work from our lab [26] had already shown that some susceptibility SNPs could modify the function of CD4+ T cells. Similar mechanisms may also affect innate T cells since they expressed some of these genes at remarkably high levels.

The susceptibility genes differentially expressed in T-cell subsets are involved in numerous cellular functions that go well beyond their role in the IL23/IL17 pathway. Genes related to this latter pathway were evenly distributed across all cell clusters and not restricted to one specific cell type, suggesting their potential joint contribution in IL17-dependent pathophysiology of AS. Gene expression data using the NanoString technology showed that MAIT cells were able to express both *IL17A* and *IL17F* in stimulation condition, as shown for CD4+ T-cells. To better decipher the relative contribution of MAITs in IL-17A/F expression compared with CD4+ T-cells, we assessed IL-17A secretion in cell culture supernatant under cell stimulation. On a per-cell basis, MAIT cells were the main producers of IL-17A in AxSpA compared to CD4+T cells, CD8+T cells and γδT cells. This was demonstrated both at the mRNA and at the protein levels. For the latter, the analysis on per cell basis allowed us to precisely characterize the production capacity of each cell type which is generally challenging for small cell subsets. Similarly, MAIT cells were also able to express at least as much IL-17F than CD4+ T cells. MAITs strongly express *IL23R* and are able to express IL-17 in a context of IL-23 stimulation but our

work shows that other cytokines, alone or in combination, such as the IL-7/IL-18 combination are able to induce a strong expression of *IL-17* RNA. IL-18 is a member of the IL-1 family. As for IL-1, IL-18 is cleaved by caspase 1 leading to its active form. IL-18 forms a complex with a receptor consisting of the IL18Ra (ubiquitous) and IL-18Rb (expressed by lymphocytes and dendritic cells) subunit that induces the NFKB pathway via MyD88/IRAK/TRAF6. Associated with IL-12 and IL-15, IL-18 induces IFNγ secretion by Th1[28]. This combination of cytokines seems to have a similar effect on MAITs. A recent publication also showed the co-expression of IL-17F following this stimulation. Association with IL-7 appeared to potentiate this IL-17F expression. IL-7 is a key cytokine of the adaptive immune system, especially for the development of T-cells and dendritic cells but also for the expansion and survival of immature B-cells. Stromal and epithelial cells are the main producers of this cytokine but Ciccia et al. also showed that Paneth's intestinal cells could produce it. IL-7 is produced consistently and independently of stimulation but at low levels under physiological conditions. Elevated levels of IL-7 have been observed in sacroiliac joint fluid in patients with SpA. Detecting an elevated level of these cytokines at the enthesis level would support our hypothesis that MAITs represent the main source of IL17 in axial SpA.

The contribution to IL-17A production of γδT and CD8+T cells was minor, but the cell specific expression of genes with a polymorphism associated to the disease still suggest their involvement in disease pathogeny through other mechanism of action. Interestingly, neutrophils have been previously described as IL-17 producing cells. This was not confirmed in our work, even using a strong stimulation. Our choice to use the combination PMA+A23187+β1,3 glucan was guided by the objective of maximizing the potential of each cell population for the production of IL17A. Aspergillus hyphae had been used by *Taylor et al.* [25] to stimulate PBMC to induce IL-17A production by neutrophils but these results have not been reproduced. Despite a similar approach, our results rather support the data published by *Tamassia et al.* [21] with two other robust techniques (Nanostring technology and SimoA technology), that neutrophils do not substantially contribute to IL-17 production.

In the peripheral form of the disease, several publications suggest that MAITs are not the only IL-17 producing cells[17]. Both iNKT and γδT-cells are increased in the synovial fluid of patients with SpA and contribute to IL-17 expression but through an IL-23-dependent mechanism[29]. *Kenna et al.* observed a high expression of IL23R on the surface of γδT-cells[19] but this cell population did not seem to be the main source of Il-17 contrary to what we observed in the axial form of the disease. In this study, they defined Th17 as the major source of IL17A. ILC3 would also be able to produce IL-17A but *Blijdorp et al.*[30] have recently shown that these cells produced IL-22 rather than IL17A in the joints of patient with peripheral SpA, results in accordance with anti-IL-23 efficacy of in this indication.

Recently IL-17F got interest with the approval of Bimekizumab in psoriasis [27] and promising results in AxSpA. Bimekizumab blocks IL-17A and IL-17F that share 50% of homology. We demonstrated that MAITs were able to express high levels of *IL-17F* after stimulation by IL-18 and to a lesser extent by IL-7. The combination of IL-7 and IL-18 was the best inducer of *IL-17F*. Despite the high level of *IL23R* expression, IL-23 stimulation did not induce expression of either *IL17A* or *IL17F*. This suggests that IL-23 alone is not sufficient to trigger IL-17 production by these cells.

In conclusion, this study shows that both innate and adaptive lymphocytes express a range of genes belonging to the IL-23/IL-17 pathway and previously associated with AS susceptibility. These data reinforce the hypothesis of the major involvement of these cell subsets in the pathogeny of SpA. On a per cell basis, MAIT cells were shown to be the main producers of IL-17A in AxSpA compared to CD4+T cells, CD8+T cells and γδT cells. Moreover, IL-7/IL-18 combination were able to strongly induce *IL-17F* expression by MAITs cells were also able to express high levels of *IL17F,* independently from IL-23 signalling. MAITs cells were identified within entheseal tissues, showing a regulatory profile. Further studies are needed to assess whether similar findings are observed in the entheseal tissues from AxSpA patients and to what extent an *in-situ* production of IL-7 and Il-18 could induce IL-17F expression.

References

1. M. Dougados, D. Baeten, Spondyloarthritis. Lancet 377, 2127 (Jun 18, 2011).
2. J. Sieper, J. Braun, M. Dougados, D. Baeten, Axial spondyloarthritis. Nature reviews. Disease primers 1, 15013 (Jul 9, 2015).
3. J. A. Smith, R. A. Colbert, Review: The interleukin-23/interleukin-17 axis in spondyloarthritis pathogenesis: Th17 and beyond. Arthritis & rheumatology 66, 23
4. Baeten, D., Baraliakos, X., Braun, J., Sieper, J., Emery, P., van der Heijde, D., McInnes, I., van Laar, J.M., Landewe, R., Wordsworth, P., et al. (2013a). Anti-interleukin-17A monoclonal antibody secukinumab in treatment of ankylosing spondylitis: a randomised, double-blind, placebo-controlled trial. Lancet 382, 1705-1713.
5. van der Heijde D, Cheng-Chung Wei J, Dougados M, Mease P, Deodhar A, Maksymowych WP, Van den Bosch F, Sieper J, Tomita T, Landewé R, Zhao F, Krishnan E, Adams DH, Pangallo B, Carlier H; COAST-V study group. Ixekizumab, an interleukin-17A antagonist in the treatment of ankylosing spondylitis or radiographic axial spondyloarthritis in patients previously untreated with biological disease-modifying anti-rheumatic drugs (COAST-V): 16 week results of a phase 3 randomised, double-blind, active-controlled and placebo-controlled trial.
6. Mease, P.J., McInnes, I.B., Kirkham, B., Kavanaugh, A., Rahman, P., van der Heijde, D., Landewe, R., Nash, P., Pricop, L., Yuan, J., et al. (2015). Secukinumab Inhibition of Interleukin-17A in Patients with Psoriatic Arthritis. The New England journal of medicine 373, 1329-1339.
7. McInnes IB, Mease PJ, Kirkham B, Kavanaugh A, Ritchlin CT, Rahman P, van der Heijde D, Landewé R, Conaghan PG, Gottlieb AB, Richards H, Pricop L, Ligozio G, Patekar M, Mpofu S; FUTURE 2 Study Group. Secukinumab, a human anti-interleukin-17A monoclonal antibody, in patients with psoriatic arthritis (FUTURE 2): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet. 2015 Sep 19;386(9999):1137-46.
8. Mease PJ, van der Heijde D, Ritchlin CT, Okada M, Cuchacovich RS, Shuler CL, Lin CY, Braun DK, Lee CH, Gladman DD; SPIRIT-P1 Study Group. Ixekizumab, an interleukin-17A specific monoclonal antibody, for the treatment of biologic-naïve patients with active psoriatic arthritis: results from the 24-week randomised, double-blind, placebo-controlled and active (adalimumab)-controlled period of the phase III trial SPIRIT-P1. Ann Rheum Dis. 2017 Jan;76(1):79-87.
9. Nash P, Kirkham B, Okada M, Rahman P, Combe B, Burmester GR, Adams DH, Kerr L, Lee C, Shuler CL, Genovese M; SPIRIT-P2 Study Group. Ixekizumab for the treatment of patients with active psoriatic arthritis and an inadequate response to tumour necrosis factor inhibitors: results from the 24-week randomised, double-blind, placebo-controlled period of the SPIRIT-P2 phase 3 trial. Lancet. 2017 Jun 10;389(10086):2317-2327.
10. McGeachy, M.J., Chen, Y., Tato, C.M., Laurence, A., Joyce-Shaikh, B., Blumenschein, W.M., McClanahan, T.K., O'Shea, J.J., and Cua, D.J. (2009). The interleukin 23 receptor is essential for the terminal differentiation of interleukin 17-producing effector T helper cells in vivo. Nature immunology 10, 314-324.
11. Baeten D, Østergaard M, Wei JC, Sieper J, Järvinen P, Tam LS, Salvarani C, Kim TH, Solinger A, Datsenko Y, Pamulapati C, Visvanathan S, Hall DB, Aslanyan S, Scholl P, Padula SJ. Risankizumab, an IL-23 inhibitor, for ankylosing spondylitis: results of a randomised, double-blind, placebo-controlled, proof-of-concept, dose-finding phase 2 study. Ann Rheum Dis. 2018 Sep;77(9):1295-1302.
12. Deodhar A, Gensler LS, Sieper J, Clark M, Calderon C, Wang Y, Zhou Y, Leu JH, Campbell K, Sweet K, Harrison DD, Hsia EC, van der Heijde D. Three Multicenter, Randomized, Double-Blind, Placebo-Controlled Studies Evaluating the Efficacy and Safety of Ustekinumab in Axial Spondyloarthritis. Arthritis Rheumatol. 2019 Feb;71(2):258-270.

13. Ritchlin C, Rahman P, Kavanaugh A, McInnes IB, Puig L, Li S, Wang Y, Shen YK, Doyle MK, Mendelsohn AM, Gottlieb AB; PSUMMIT 2 Study Group. Efficacy and safety of the anti-IL-12/23 p40 monoclonal antibody, ustekinumab, in patients with active psoriatic arthritis despite conventional non-biological and biological anti-tumour necrosis factor therapy: 6-month and 1-year results of the phase 3, multicentre, double-blind, placebo-controlled, randomised PSUMMIT 2 trial.
14. McInnes IB, Kavanaugh A, Gottlieb AB, Puig L, Rahman P, Ritchlin C, Brodmerkel C, Li S, Wang Y, Mendelsohn AM, Doyle MK; PSUMMIT 1 Study Group. Efficacy and safety of ustekinumab in patients with active psoriatic arthritis: 1 year results of the phase 3, multicentre, double-blind, placebo-controlled PSUMMIT 1 trial. Lancet. 2013 Aug 31;382(9894):780-9.
15. Reinhardt A, Yevsa T, Worbs T, Lienenklaus S, Sandrock I, Oberdörfer L, Korn T, Weiss S, Förster R, Prinz I. Interleukin-23-Dependent γ/δ T Cells Produce Interleukin-17 and Accumulate in the Enthesis, Aortic Valve, and Ciliary Body in Mice. Arthritis Rheumatol. 2016 Oct;68(10):2476-86.
16. Cuthbert RJ, Watad A, Fragkakis EM, Dunsmuir R, Loughenbury P, Khan A, Millner PA, Davison A, Marzo-Ortega H, Newton D, Bridgewood C, McGonagle DG. Evidence that tissue resident human enthesis γδT-cells can produce IL-17A independently of IL-23R transcript expression. Ann Rheum Dis. 2019 Nov;78(11):1559-1565.
17. Gracey E, Qaiyum Z, Almaghlouth I, Lawson D, Karki S, Avvaru N, Zhang Z, Yao Y, Ranganathan V, Baglaenko Y, Inman RD. IL-7 primes IL-17 in mucosal-associated invariant T (MAIT) cells, which contribute to the Th17-axis in ankylosing spondylitis. Ann Rheum Dis. 2016 Dec;75(12):2124-2132.
18. Jandus C, Bioley G, Rivals J-P, et al. Increased numbers of circulating polyfunctional Th17 memory cells in patients with seronegative spondylarthritides. Arthritis Rheum 2008;58:2307–17. doi:10.1002/art.23655
19. Kenna TJ, Davidson SI, Duan R, et al. Enrichment of circulating interleukin-17-secreting interleukin-23 receptor-positive γ/δ T cells in patients with active ankylosing spondylitis. Arthritis Rheum 2012;64:1420–9. doi:10.1002/art.33507
20. Appel H, Maier R, Wu P, et al. Analysis of IL-17+cells in facet joints of patients with spondyloarthritis suggests that the innate immune pathway might be of greater relevance than the Th17-mediated adaptive immune response. Arthritis Research & Therapy 2011;13:R95. doi:10.1186/ar3370
21. Tamassia N, Arruda-Silva F, Calzetti F, et al. A Reappraisal on the Potential Ability of Human Neutrophils to Express and Produce IL-17 Family Members In Vitro: Failure to Reproducibly Detect It. Frontiers in Immunology 2018;9. doi:10.3389/fimmu.2018.00795
22. Rudwaleit M, van der Heijde D, Landewé R, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. Ann Rheum Dis 2009;68:777–83.
23. Cuthbert RJ, Watad A, Fragkakis EM, et al. Evidence that tissue resident human enthesis γδT-cells can produce IL-17A independently of IL-23R transcript expression. Ann Rheum Dis. 2019;78(11):1559-1565. doi:10.1136/annrheumdis-2019-215210
24. Bridgewood C, Russell T, Weedon H, et al. The novel cytokine Metrnl/IL-41 is elevated in Psoriatic Arthritis synovium and inducible from both entheseal and synovial fibroblasts. Clin Immunol. 2019;208:108253. doi:10.1016/j.clim.2019.108253
25. Taylor PR, Roy S, Leal SM, et al. Activation of neutrophils by autocrine IL-17A-IL-17RC interactions during fungal infection is regulated by IL-6, IL-23, RORγt and dectin-2. Nat Immunol 2014;15:143–51. doi:10.1038/ni.2797
26. Coffre M, Roumier M, Rybczynska M, et al. Combinatorial control of Th17 and Th1 cell functions by genetic variations in genes associated with the interleukin-23 signaling pathway in spondyloarthritis. Arthritis Rheum 2013;65:1510–21. doi:10.1002/art.37936
27. Papp KA, Merola JF, Gottlieb AB, et al. Dual neutralization of both interleukin 17A and interleukin 17F with bimekizumab in patients with psoriasis: Results from BE ABLE 1, a 12-week randomized, double-blinded, placebo-controlled phase 2b trial. J Am Acad Dermatol 2018;79:277-286.e10. doi:10.1016/j.jaad.2018.03.037

28. Ussher JE, Bilton M, Attwod E, et al. CD161++ CD8+ T cells, including the MAIT cell subset, are specifically activated by IL-12+IL-18 in a TCR-independent manner. *Eur J Immunol*. 2014;44(1):195-203. doi:10.1002/eji.201343509
29. Venken K, Jacques P, Mortier C, et al. RORγt inhibition selectively targets IL-17 producing iNKT and γδ-T cells enriched in Spondyloarthritis patients. *Nat Commun*. 2019;10(1):9. Published 2019 Jan 2. doi:10.1038/s41467-018-07911-6
30. Blijdorp ICJ, Menegatti S, Mens LJJV, et al. IL-22- and GM-CSF-expressing but not IL-17A-expressing group 3 innate lymphoid cells are expanded in the inflamed spondyloarthritis joint. Arthritis & Rheumatology;0. doi:10.1002/art.40736

RESULTS: PART III

Immune response profiling of spondyloarthritis patients reveals signaling networks mediating TNF-blocker function in vivo

Menegatti , et al. (submitted)

Manuscript II

Immune response profiling of spondyloarthritis patients reveals signalling networks mediating TNF-blocker function in vivo

Summary

Anti-TNF therapy has transformed the treatment strategies for several CIDs, including axSpA. However, the effectiveness of TNF-blockers and the response rate in patients are highly variable, as nearly 40% of axSpA patients are poor responders (Menegatti, Bianchi, & Rogge, 2019). The biological mechanism behind the treatment failure and the effect of anti-TNF therapy in patients' immune response remain uncharacterised. The third part of my project addressed this matter by analysing immune response in cell cultures from axSpA patients induced microbial stimuli and stimuli that activate specific signalling pathways. We performed whole-blood stimulation assays using TruCulture system with designated stimuli, from 80 axSpA patients before and after anti-TNF therapy, and measured the cytokine and chemokine profiles and monitored gene expression using NanoString nCounter® assays.

We stimulated blood samples from 12 patients with a range of microbial stimuli or signalling agonists, and we measured the levels of 31 secreted molecules. We observed that TNF-inhibitors (TNFi) affect induced immune responses towards microbes and stimuli targeting specific immune receptors. In particular, TNFi induced specific changes in patients' immune responses that are mostly detected in the challenged immune system, and not in the resting state.

We demonstrated that the effect of TNF-blockers is detectable after a single dose and remained stable over a period of three months after initiation of therapy. Pro-inflammatory molecules such as- MIP-1β, IL-1Ra and IL-8 were reduced in response to many stimuli, which indicated the influence TNF blockers have towards many intracellular pathways. TNFi directed its effects towards the innate immune responses of the patients, with minor effects on the Th1/Th17 immunity, as we did not see major effects on secreted IL-6, IFN-γ and IL-17 proteins. The gene expression data was consistent with the secreted protein analysis, as there were no effects of the treatment on genes *IL17A*, *IFNG* and *IL6.*

The analysis on expression pattern of immune-related genes at day 7 (D7) and after three months (D90) of anti-TNF therapy revealed that a number of genes of the NF-κB transcriptional network were affected. We report that TNFi target a core immune response gene signature that includes NF-κB genes (*NFKB1*, *RELA*, *NFKB2* & *RELB*) and NF-κB

targets (*IL1A*, *IL1B* & *CCL20*). Additionally, we observed a strong downregulation of the cyclooxygenase-2 (COX2) gene *PTGS2* - an enzyme involved in the biosynthetic pathway of prostaglandin synthesis, and *PTGER4*, which encodes the PGE2 receptor EP4. Signalling through EP4 upregulates *IL-23R* expression promoting human Th17 cell development, and suppresses disease progression in an experimental mouse model of autoimmune encephalomyelitis as reported by Yao et al. (2009). Of note, *PTGER4* has been associated with SpA susceptibility, as have been *NFKB1* and *CARD9*, which were also strongly downregulated by TNFi. Collectively, these data provide evidence that TNFi target the expression of genes closely linked to SpA pathogenesis. Our findings suggest that TNFi target several immune cell pathways that cooperate to control inflammation.

To study the effect of TNFi on the immune networks we compared the immune responses at D0 and D7 using Quantitative Set Analysis for Gene Expression (QuSAGE) for several gene modules. Group of genes of the “NF-κB transcription factors” and “NF-κB target genes” modules were observed to be downregulated by TNF-blockers. While the gene set showed a reduction in most immune events following TNF inhibition, we report an increased activity in the cytotoxic gene module. Cytotoxic granules are essential in the effector function of CTLs and NK cells. We observed an increase in the pathway activity after therapy for the genes linked to target cell lysis and “NK cells” module, such as granzymes (*GZMA* and *GZMB*), perforin (*PRF1*) and granulysin (*GNLY*), as well as killer cell lectin receptors (*KLRD1*, *KLRF1*).

Another key finding from our analysis is the decrease in the genes from the “M1-like monocyte” and the increase in “M2-like monocyte” gene module. This data indicate that the TNFi may affect monocyte/macrophage polarisation, specifically towards the activation of M2-phenotype macrophages. Our data also showed that many characteristic surface markers of the regulatory M2-macrophages, such as mannose receptor *MRC1*, and scavenger receptors *MSR1* and *CD163* are upregulated at D7 after the anti-TNF therapy, implicating their role in inflammation resolution in axSpA patients.

Through this study we show that TNFi target multiple immune cell pathways that cooperate to resolve inflammation. We propose that immune response profiling provides new insight into the biology of TNF-blocker action in patients and can identify signalling pathways associated with therapeutic responses to biologic therapies. The detailed manuscript, with figures and supplementary data of this work that is currently under revision with Annals of Rheumatic Diseases is given in the following pages.

Manuscript draft

Immune response profiling of spondyloarthritis patients reveals signaling networks mediating TNF-blocker function *in vivo*

Silvia Menegatti[1,2,†], Vincent Guillemot[3], Eleonora Latis[1,2], Hanane Yahia-Cherbal[1,2], Daniela Mittermüller[1], Vincent Rouilly[4], Elena Mascia[1], Nicolas Rosine[1,2], Surya Koturan[1,2], Gael A. Millot[3], Claire Leloup[1], Darragh Duffy[5], Aude Gleizes[6,7], Salima Hacein-Bey-Abina[6,7], Milieu Intérieur Consortium[‡], Jérémie Sellam[8,9], Francis Berenbaum[8,9], Corinne Miceli-Richard[1,10,11], Maxime Dougados[10,11,12], Elisabetta Bianchi[1,11], Lars Rogge[1,11*]

[1] Immunoregulation Unit, Institut Pasteur, Department of Immunology, Paris, France.

[2] Université Paris Diderot, Sorbonne Paris Cité, Paris, France.

[3] Bioinformatics and Biostatistics Hub – Département de Biologie Computationelle, Institut Pasteur, USR 3756 IP CNRS – Paris, France.

[4] DATACTIX, Paris, France.

[5] Institut Pasteur, Translational Immunology Lab, Department of Immunology, Paris, France.

[6] Clinical Immunology Laboratory, Groupe Hospitalier Universitaire Paris-Sud, Hôpital Kremlin-Bicêtre, Assistance Publique-Hôpitaux de Paris, Le-Kremlin-Bicêtre, France.

[7] UTCBS CNRS UMR 8258, INSERM U1267, Faculté de Pharmacie de Paris, Université de Paris, Paris, France.

[8] Sorbonne Université, Service de Rhumatologie, Hôpital Saint-Antoine, AP-HP, 75012, Paris, France.

[9] Centre de Recherche Saint-Antoine, INSERM UMR_S 938, Paris, France.

[10] Paris Descartes University, Rheumatology Department, Cochin Hospital, AP-HP, Paris, France.

[11] Unité Mixte AP-HP/Institut Pasteur, Institut Pasteur, Paris, France.

[12] INSERM U1153 Clinical Epidemiology and Biostatistics, PRES Sorbonne Paris-Cité, Paris, France.

[†] Present address: INSERM U932, Institut Curie, PSL Research University, Paris, France.

[‡] A list of members of the Milieu Intérieur Consortium is given in the Acknowledgements

* Corresponding author: Lars Rogge, PhD
Institut Pasteur, Immunoregulation Unit, Department of Immunology
25, rue du Dr. Roux, 75724 Paris, France
Phone: +33-1-40613822, mail: lars.rogge@pasteur.fr
ORCID ID: 0000-0003-1262-9204

Word count: 2998

ABSTRACT

Objectives

Anti-tumor necrosis factor (TNF) therapy has revolutionized treatment of several chronic inflammatory diseases, including spondyloarthritis (SpA). However, TNF-inhibitors (TNFi) are not effective in all patients and the biological basis for treatment failure remains unknown. We have analyzed induced immune responses to define the mechanism of action of TNF-blockers in SpA and to identify immunological correlates of responsiveness to TNFi.

Methods

Immune responses to microbial and pathway-specific stimuli were analyzed in peripheral blood samples from 80 axial SpA patients before and after TNFi treatment, using highly standardized whole-blood stimulation assays. Cytokines and chemokines were measured in a CLIA-certified laboratory, and gene expression was monitored using nCounter assays.

Results

Anti-TNF therapy induced profound changes in patients' innate immune responses. TNFi action was selective, and had only minor effects on Th1/Th17 immunity. Modular transcriptional repertoire analysis identified prostaglandin E_2 synthesis and signaling, leukocyte recirculation, macrophage polarization, dectin and IL-1 signaling, as well as the NF-kB transcription factor family as key pathways targeted by TNF-blockers *in vivo*. Analysis of induced immune responses before treatment initiation revealed that expression of molecules associated with leukocyte adhesion and invasion, chemotaxis and IL-1 signaling are correlated with therapeutic responses to anti-TNF.

Conclusions

We show that TNFi target multiple immune cell pathways that cooperate to resolve inflammation. We propose that immune response profiling provides new insight into the biology of TNF-blocker action in patients and can identify signaling pathways associated with therapeutic responses to biologic therapies.

KEYWORDS

Spondyloarthritis, anti-TNF therapy, immune response profiling, mechanisms of biologic drug action, therapeutic responses

KEY MESSAGES

What is already known about this subject?

- Anti-TNF therapy has revolutionized treatment of many chronic inflammatory diseases, including spondyloarthritis (SpA) and rheumatoid arthritis. However, TNF-inhibitors (TNFi) are not effective in 30-40% of patients. The immunosuppressive effects of TNF-blockers therefore expose a substantial fraction of patients to side-effects, in particular infections, without clinical benefit. Despite the extensive use of TNFi for many years, the biological basis for treatment failure remains unknown.

What did this study add?

- We demonstrate that anti-TNF therapy induces profound changes in patients' innate immune responses, but does not affect Th1/Th17 immunity.
- Modular transcriptional repertoire analysis showed that prostaglandin E_2 synthesis and signaling, leukocyte recirculation, macrophage polarization, dectin and IL-1 signaling, as well as the NF-kB transcription factor family are key pathways targeted by TNF-blockers *in vivo*.
- To investigate the concept that the immune status of patients before treatment initiation will define their response to TNFi treatment, we have searched for immunological transcripts that correlate with clinical efficacy of TNF-blockers in stimulated immune cells. We found that high expression of molecules associated with leukocyte adhesion and invasion, chemotaxis and IL-1 signaling is correlated with favorable outcome of anti-TNF therapy.

How might this study impact on clinical practice or future developments?

- We have established a robust pipeline to monitor immune responses in patients that can be translated into a clinical setting. We show that immune response profiling can identify signaling pathways associated with therapeutic responses to TNFi. Further studies will assess whether this approach can be used to develop molecular biomarkers to help stratify patients to the most appropriate therapy.

INTRODUCTION

Chronic inflammatory diseases (CID) are challenging illnesses that often strike at a young age and cause lifelong morbidity, representing a considerable burden for the affected individuals and for society. Spondyloarthritis (SpA) is a family of related inflammatory disorders with common pathologic and genetic features.[1-3] Clinical manifestations include spinal (axial) inflammation, peripheral arthritis, enthesitis and extra-articular features such as uveitis, psoriasis and inflammatory bowel disease.[4]

Anti-TNF therapy has proven effective to reduce inflammation and clinical symptoms in SpA, however, little is known about how TNF inhibitors (TNFi) affect immune responses in patients, and TNFi have been associated with infectious complications,[5] including *M. tuberculosis* reactivation.[6-8]

Furthermore, the high rate of non-responsiveness (30-40%) to TNFi exposes a substantial fraction of patients to side effects without clinical benefit, and it is still not possible to determine which patients will respond to TNFi before treatment initiation.[9-11] The recent introduction of antibodies blocking IL-17A has expanded the therapeutic options for axial SpA (axSpA), as well as psoriasis and psoriatic arthritis.[12, 13] It is therefore important to develop tools to guide treatment decisions for patients affected by SpA and other CID, to optimize clinical care and contain health care costs.

Here, we investigated the global impact of TNFi on immune responses to microbial or pathway-specific stimuli, with the goal to enhance our understanding of the molecular mechanism of action of TNF-blockers in SpA patients and to identify immunological correlates of responsiveness to TNFi.

METHODS

Patients

Peripheral blood samples were obtained from 80 biologic-naïve patients fulfilling ASAS criteria for axSpA,[14, 15] attending the Rheumatology Departments of Cochin or Saint-Antoine Hospitals (Paris, France). This study fulfills the current Good Clinical Practice Guidelines and a clinical protocol has been accepted by regulatory committees: Comité de Protection des Personnes Ile de France III; Référence CPP: n° AT-100), Institut Pasteur (Projet de recherché clinique n° 2011-32, CCTIRS (DGRI CCTIRS MG/CP°2012.035), and CNiL (Décision DR-2013-080). A written informed consent has been obtained from each subject.

Patients' demographics, HLA-B27 status, information regarding symptoms, ongoing treatments, co-morbidities and other main clinical features of spondyloarthritis were recorded on a Case Record Form before and 3 months (D90) after initiation of anti-TNF therapy (see **Table 1** and **online supplementary table 1**).

Primary responsiveness to anti-TNF therapy was based on the Ankylosing Spondylitis Disease Activity Score (ASDAS).[16] The "improvement score" was calculated as: ASDAS at baseline (D0) - ASDAS at D90). Patients achieving a delta ASDAS < 1.1 were classified as non-responders.[16]

Whole-Blood TruCulture Stimulation was performed with TruCulture assays (Myriad RBM, Texas).[17] Multi-analyte profiling of culture supernatants was performed with Luminex xMAP technology (Myriad-RBM, Austin, TX, USA), gene expression analysis with nCounter Technology (NanoString), with the Human Immunology v2 Gene Expression CodeSet.[18, 19]

Purification of monocytes and *in vitro* cell stimulation

To generate *in vitro* derived macrophages, monocytes were isolated from healthy donors and cultured with M-CSF in presence or absence of TNFi. Cells were polarized towards M1 with LPS (20 ng/mL, Invivogen) and IFN-□ (20 ng/ml, Milteny), or towards M2 with IL-4 and IL-13 (20 ng/ml, Miltenyi).

Data analysis

Quantitative set analysis of gene expression was performed using the R QuSage package.[20] Differential gene expression was analyzed using the LIMMA package,[21] principal component analysis and hierarchical clustering were performed with Qlucore Omics Explorer (Qlucore).

Methods are described in detail in the online supplementary material.

RESULTS

TNFi affect immune responses to microbes and stimuli targeting specific immune receptors

We analyzed immune responses in axSpA patients with indications for TNFi treatment (**Table 1**), using whole blood ("TruCulture") assays[17] (**figure 1A**). We stimulated blood samples from 12 patients with a range of microbial stimuli or signaling agonists, and we measured the levels of 31 secreted molecules (**online supplementary table 3** and **4, online supplementary figure 1A**). Three months (D90) after TNFi initiation, the induction of many pro-inflammatory cytokines and chemokines (such as MIP-1β, IL-1Ra and IL-8) was reduced in response to various stimuli, indicating that TNFi target intracellular pathways shared by a broad range of immune activators (**figure 1B**). In contrast, TNFi had no major effects on IL-6, IFN-γ and IL-17 (**online supplementary figure 1D**), although the Th17 pathway is suggested to be of key importance in SpA pathophysiology.[22]

Only few secreted proteins increased after TNFi therapy. Among these was IL-10 following stimulation with gardiquimod **(figure 1B**), a selective ligand for TLR7.

These results show that TNFi induce selective changes in patients' immune responses, mostly detected in the challenged immune system, and not in the resting state (**online supplementary figure 1D**).

The effects of TNFi are detected after a single injection and remain stable over time

To determine the early effects of TNFi, we analyzed 17 consecutive axSpA patients 7 days after initiation of TNFi therapy (**online supplementary figure 1B**). Secretion of pro-inflammatory mediators was already affected after a single TNFi injection (**figure 1C, D and G**) and over a broad range of stimuli (**online supplementary figure 2A**). Production of IL-6, IL-17 and IFN-□ was largely unaffected (**figure 1E** and **F**).

The reduction in pro-inflammatory mediators was maintained at D90 (**online supplementary figure 2B** and **C**), demonstrating that the effects of TNFi on immune responses remain stable over time.

TNF-blockers affect key transcriptional networks of innate immune responses

To gain insight into the mechanisms by which TNFi affect immune responses, we analyzed the expression of immune-related genes before and at D7 and D90 after TNFi treatment. TNF blockade profoundly altered the transcription of a large number of genes (**figure 2A**).

The majority of genes differentially expressed after therapy were shared by different stimulation conditions, revealing a "core immune response signature" targeted by TNFi (**figure 2B**), which included NF-kB genes, such as *NFKB1*, *RELA, NFKB2* and *RELB*, and NF-kB targets, such as *IL1A, IL1B* and *CCL20* (**figure 2C, 2D, S3A** and **online supplementary figure 3B**). In particular, TNFi strongly downmodulated expression of *PTGS2*, encoding cyclooxygenase (COX-2), the key enzyme in prostaglandin E_2 (PGE_2) biosynthesis, and *PTGER4* encoding the PGE_2 receptor EP4 (**figure 2D**). TNFi-induced downmodulation of *PTGS2* and *PTGER4* did not depend on the NSAID index at baseline (**online supplementary figure 4**). Consistent with our analysis of secreted proteins (**figure 1D**), *IL17A*, *IFNG* and *IL6* were largely unaffected (**online supplementary figure 3A**).

The analysis of patients stratified into responders and non-responders showed that the majority of differentially expressed genes are common to both groups, although a number of genes are uniquely affected in each patient subset (**online supplementary table 6** and **online supplementary figures 5 and 6**).

The effects of TNFi also on gene expression could be measured after a single injection and remained stable over time (**online supplementary figure 7A**).

To determine if changes in cell populations accounted for these effects, we analyzed cell counts at D0 and D90. While leukocyte and monocyte counts remained stable, we observed a modest decrease of neutrophils and increase of lymphocyte counts after TNFi therapy (**online supplementary figure 7B**).

Modular transcriptional repertoire analysis reveals multiple mechanisms of TNFi action *in vivo*

The observation that TNFi affected several molecules in the same signaling pathway prompted us to further define the effects of TNFi on immune networks. We compared immune responses at D0 and D7 using Quantitative Set Analysis for Gene Expression (QuSAGE)[20] (**online supplementary table 5**). The modules "NF-κB transcription factors" and "NF-κB target genes" were among those most strongly downregulated by TNFi (**figure 3A-C** and **online supplementary table 7**), followed by the "IL-1/IL-1R" module (**figure 3A** and **B**). Inspection of the individual genes in this module showed downregulation of *IL1A*, *IL1B*, *IRAK2*, *IL1R1*, and *IL1RN*, as well as a substantial increase of *SIGIRR*, after TNF blockade (**figure 3D**).

TNFi therapy also reduced the activity of the "dectin" module (**figure 3A and B** and **online supplementary figure 8A**), which groups C-type lectin receptors (CLRs) for *C. albicans* and other fungi such as Dectin-2 (encoded by *CLEC6A*), or Mincle (encoded by *CLEC4E*) and associated signaling molecules, such as *CARD9*, a molecule involved in antifungal immunity that mediates signals from CLRs to the NF-κB pathway via BCL10.[23]

While gene set activities for most gene modules were reduced by TNFi, we observed increased activity at D7 of the "cytotoxic molecules" module and of the M2-like monocytes gene module, while the overall activity of the module "M1-like monocytes" was reduced after TNFi, indicating that TNF-blockers may affect monocyte/macrophage polarization (**figure 3**).

In particular, we observed an upregulation of the genes encoding surface markers characteristic of regulatory macrophages, such as the mannose receptor *MRC1,* the scavenger receptors *MSR1* and *CD163*, the decoy receptor *IL1R2*, and of *IL10* (**figure 3G** and **online supplementary figure 8B**).

Analogous results were obtained at D90 after initiation of TNFi (**online supplementary figure 8C**), indicating the multiple immune pathways that mediate TNFi function in SpA patients.

Many of the genes affected by TNFi are expressed in monocytes and macrophages, which prompted us to investigate the roles of these cells in the response to TNFi. We stimulated monocytes from SpA patients with LPS in the presence or absence of etanercept (Eta), and measured transcript levels before and at different time points after stimulation (**online supplementary figure 9**). Several of the genes downregulated by etanercept were direct NF-κB target genes, such *NFKBIA*, *TNFAIP3*, *TNFAIP6*, or *IL1A* (**online supplementary figure 9**).

TNFi skew macrophage polarization towards an M2 phenotype *in vitro*

We then asked whether TNFi affect also macrophage gene expression. As the analysis of tissues is rarely performed in axSpA[24] we investigated the effects of two TNFi, etanercept and adalimumab, on *in vitro* differentiated macrophages (**figure 4A**). Although the effects of adalimumab on gene expression were stronger in our system, a core of 56 genes was regulated by both TNFi (**figure 4B-E**).

We noted strong downregulation of M1-macrophages genes such as *IL18* (**figure 4C, D** and **E**), while expression of genes associated with M2-macrophages, such *MRC1*, *MSR1* and *CLEC7A* was significantly increased (**figure 4E**).

TNFi also strongly downmodulated *PTGS2* expression in stimulated M1 macrophages (**figure 4E**), and affected the mRNA levels of chemokines and their receptors: the expression of *CCL19*, *CCL4*, and *CCL3* was downregulated, while *CCL13* and *CCL24* were upregulated by TNFi (**figure 4C, D** and **E**). These data are consistent with our results for TNFi treatment *in vivo* and suggest that TNFi may affect leukocyte recruitment to inflamed joints.

Finally, we confirmed a significant downregulation of NF-κB pathway genes (**figure 4C, D** and **F**). These data further support the notion that TNFi affect immune responses by acting on multiple inflammatory pathways and that phagocytic cells are important targets of these effects (**figure 4F**).

Immune gene expression associated with therapeutic responses to anti-TNF therapy

Finally, we investigated the correlation between therapeutic responses to TNFi and stimulated immune responses in 80 axSpA patients, before initiation of anti-TNF therapy. Response to therapy was calculated as the delta ASDAS "improvement score" (ASDAS D0 – ASDAS D90).[16, 25] 50 patients (62.5%) had either a major or a clinically important improvement ("responders", delta ASDAS ≥ 1.1), while 30 (37.5%) were non-responders

(**Table 1 and online supplementary table 1**). The analysis of whole-blood cultures stimulated with LPS or SEB revealed that 55 genes were differentially expressed between responders and non-responders (**Table 2** and **figure 5A**).

To explore if different types of anti-TNF drugs could have an impact on therapeutic responses to TNFi, we compared differential gene expression between responders and non-responders treated with soluble TNFR2 (n=53) to those treated with monoclonal antibodies (n=27). We found a good correlation (R=0.901) for the 55 genes differentially expressed. These data indicate that the type of TNF-blockers does not have a major effect on the genes significantly associated with therapeutic responses before treatment (**online supplementary figure 10B**).

A search of the DICE database[26] showed expression of these genes in different immune cells, including activated T cells, Treg, Th17 and NK cells (**figure 5B**). Notably, 29 of the genes were expressed specifically in resting classical or non-classical monocytes (**figure 5B**). These data suggest that several immune cell populations contribute to determine the efficacy of anti-TNF therapy in SpA patients.

Among the 55 differentially expressed genes, 15 regulate key steps of leukocyte migration and invasion: these include *PLAU* and *PLAUR*, the integrin subunits *ITGB1*, *ITGA5*, *ITGAX*, and *ITGA6*, and the CD2 ligand *CD58* (**figure 5B, 5C** and **Table 2**). The importance of leukocyte recirculation as a determinant of therapeutic responses to TNFi is supported by the observation that several genes encoding chemokines and their receptors, such as *CCL20*, *IL8, CXCL1, CXCL2* and *CXCR1* are expressed at higher levels in cultures from SpA patients responding to TNFi than in non-responders, while *CXCL9* is expressed at higher levels in non-responders (**figure 5B-C, Table 2 and online supplementary figure 10**). Expression of the receptors for the pro-inflammatory cytokines TNF (*TNFRSF1B*), IL-6 (*IL6R*) and IL-1 (*IL1R1*, *IL1R2* and *IL1RAP*) was also substantially higher in responders than in non-responders, as was expression of the IL-1R-associated kinases *IRAK1* and *IRAK3,* and of *NLRP3*, which controls caspase-1-dependent processing of pro-IL-1 and IL-18. These data indicate that the activation status of the IL-1 signaling pathway may influence responsiveness to TNFi. We also noted substantially higher expression in responders of *CLEC5A* (MDL-1, myeloid DAP12-associating lectin-1), an important mediator of autoimmune inflammation in experimental arthritis models[27] (**figure 5C** and **Table 2**).

DISCUSSION

To investigate immune responses in SpA patients, we have used highly standardized and robust assays that may be directly translated into a clinical setting. “TruCulture” assays were designed to preserve physiological cellular interactions and capture immune cell activity without introducing sample collection and manipulation variables.[28] We chose to analyse responses in whole blood, because tissue biopsies cannot be performed routinely in axSpA.

Most of the effects of TNFi could be observed only in stimulated cultures, supporting the notion that TNFi act on activated immune cells, rather than in homeostatic conditions. This may explain the relatively modest changes in gene expression in response to TNFi detected in a recent study of unstimulated PBMCs from axSpA patients.[29]

Our modular transcriptional repertoire analysis of the stimulation cultures[20] established a hierarchy of signaling pathways affected by anti-TNF therapy, with potential clinical implications.

We found a strong decrease of pro-inflammatory molecules produced primarily by innate immune cells, pointing to the importance of these cells in SpA pathogenesis. The decreased activity of the NF-κB module underlines the major role of these factors in mediating TNF-blocker functions. However, TNF-blockade had only minor effects on the expression and secretion of IL-6, contrary to what observed in RA patients.[30] These data suggest that this cytokine may be more relevant to RA, but less to SpA pathogenesis, consistent with the limited therapeutic efficacy of IL-6-blockade in SpA.[31]

We observed downregulation of the classical, M1-like module and an increase of the non-classically activated, M2-like monocyte gene module activity, consistent with the finding that TNFi can expand a cell population with a M2 macrophage-like appearance in *vivo* and *in vitro*.[32, 33] Analysis of the effects of TNFi *in vitro* provided direct evidence that TNFi act directly on macrophage polarization. These results are consistent with a previous study performed with *in vitro* differentiated macrophages from RA patients.[34] M2 macrophages, characterized by expression of IL-10, high-levels of scavenger and mannose receptors, *IL1R2* and *IL1RN*, are implicated in the resolution of inflammation and orchestrate tissue repair and remodeling[35, 36]. Polarization of monocytes/macrophages towards a M2-like profile may be an additional mechanism by which TNF-blockers act on the immune system to regulate inflammatory responses[37] and could also explain the increased risk of opportunistic infections observed for patients treated with TNFi, in particular *M. tuberculosis*.[38]

TNFi strongly downregulated expression of *PTGS2,* the key enzyme in prostaglandin E_2 (PGE_2) biosynthesis and target of non-steroidal anti-inflammatory drugs, the first-line treatment of SpA. PGE_2 is an important early mediator of enthesitis, the hallmark of SpA[39] and COX-2 inhibition may be an important mechanism of TNFi therapeutic action in this disease. PGE_2 induces vasodilation, which may facilitate neutrophil recruitment into the entheseal compartment[39]. We also found that expression of the PGE_2 receptor *PTGER4* (EP4) was downregulated by TNFi. Signaling through EP4 upregulates IL-23R expression promoting human Th17 cell development,[40] and suppresses disease progression in an experimental mouse model of autoimmune encephalomyelitis.[41] Of note, *PTGER4* has been associated with SpA susceptibility, as have been *NFKB1* and *CARD9,*[42] also strongly downregulated by TNFi. Collectively, these data provide evidence that TNFi target the expression of genes closely linked to SpA pathogenesis.

Our findings suggest that TNFi target several immune cell pathways that cooperate to control inflammation. Targeting PGE_2biosynthesis via *PTGS2* downregulation is of particular relevance for enthesitis, a critical early pathogenic feature of spondyloarthritis, while shifting the balance of macrophages from a pro-inflammatory phenotype to a pro-resolving phenotype is important for the resolution of synovitis. MDL-1/CLEC5A was among the most strongly downregulated molecule after TNFi therapy. Dengue virus-mediated activation of MDL-1/CLEC5A can trigger potent induction of TNF, IL-6 and IL-1 and NLRP3 inflammasome activation and shock.[43, 44] MDL-1/CLEC5A is also expressed in synovial tissue from RA patients and MDL-1/CLEC5A blockade reduced tissue inflammation and bone erosion in experimental arthritis models.[27] Reduction of MDL-1/CLEC5A expression by TNFi may result in inhibition of bone erosion and inflammatory cytokine production in SpA.

The involvement of multiple pathways in TNF-blocker functions could also explain the difficulties in identifying a genetic marker for treatment response to TNFi.[45] We could not identify a single gene whose expression correlates with responsiveness to TNFi, but rather a set of genes. A limitation is that our study focused on a pre-defined panel with 594 immune-

related genes. In addition, this study was performed in patients from France and and should be replicated in patients from different genetic and environmental backgrounds. Genome-wide studies may be necessary to identify unique molecular biomarkers. Nevertheless, our data indicate that high-level expression of molecules associated with leukocyte invasion and migration as well as IL-1 signaling in stimulated immune cells predisposes to favorable outcome of anti-TNF therapy in the consistent cohort of 80 patients recruited for this study.

In conclusion, we suggest that immune response profiling of patients is a powerful approach to define the mechanism of action of biologic drugs and may be a useful strategy to establish objective criteria guiding treatment decisions.

Contributors

S.M., E.B., and L.R. designed the study, analyzed data, interpreted results and wrote the manuscript; S.M., E.L., E.M., H.Y., D.M., C.L., N.R. and S.K. performed experiments; A.G. and S.H.-B.-A. analyzed drug levels and anti-drug antibiodies in serum samples; S.M., V.G., V.R., G.A.M., E.B. and L.R. performed bioinformatics data analysis; D.D. provided data from the Milieu Intérieur cohort. J.S. and F.B. provided patient samples and clinical data. C.M.R. and M.D. had overall medical oversight, provided patient samples and clinical data, performed clinical data analysis and revised the manuscript; all authors approved the manuscript.

Competing interests

The authors declare no competing financial interests.

Funding

S.M. was a scholar of the Pasteur-Paris University (PPU) International PhD program and supported by a grant from the Fondation pour la Recherche Médicale. E.L. was supported by a fellowship from the Université Paris Diderot. This study was supported by grants from Institut Pasteur, the French Government's Investissement d'Avenir Program, Laboratoire d'Excellence "Milieu Intérieur" (ANR-10-LABX-69-01), FOREUM Foundation for Research in Rheumatology, the Fondation Arthritis, MSD Avenir (Project iCARE-SpA), a Bourse Passerelle from Pfizer and a Sanofi Innovation Award Europe.

Acknowledgments

We thank Nathalie Menagé (Cochin Hospital, Paris) for patient-sample collection, Samuel T. LaBrie (Myriad RBM, Inc.) for multi-analyte profiling, the Center for Translational Science (CRT)/Cytometry Biomarkers Unit of Technology and Service (CB UTechS) at Institut Pasteur for support in conducting this study. Parts of this study have been presented at the Annual

European Congress of Rheumatology (EULAR) in June 2019 (Meeting Abstract: SP0054; "Deconvolution of the immune response").

‡ The Milieu Intérieur Consortium

The Milieu Intérieur Consortium¶ is composed of the following team leaders: Laurent Abel (Hôpital Necker), Andres Alcover, Hugues Aschard, Kalla Astrom (Lund University), Philippe Bousso, Pierre Bruhns, Ana Cumano, Caroline Demangel, Ludovic Deriano, James Di Santo, Françoise Dromer, Gérard Eberl, Jost Enninga, Jacques Fellay (EPFL, Lausanne), Ivo Gomperts-Boneca, Milena Hasan, Serge Hercberg (Université Paris 13, Paris), Olivier Lantz (Institut Curie, Paris), Hugo Mouquet, Etienne Patin, Sandra Pellegrini, Stanislas Pol (Hôpital Côchin, Paris), Antonio Rausell (INSERM UMR 1163 – Institut Imagine, Paris), Lars Rogge, Anavaj Sakuntabhai, Olivier Schwartz, Benno Schwikowski, Spencer Shorte, Frédéric Tangy, Antoine Toubert (Hôpital Saint-Louis, Paris), Mathilde Trouvier (Université Paris 13, Paris), Marie-Noëlle Ungeheuer, Darragh Duffy§, Matthew L. Albert (In Sitro)§, Lluis Quintana-Murci§

¶ unless otherwise indicated, partners are located at Institut Pasteur, Paris, France

§ co-coordinators of the Milieu Intérieur Consortium

REFERENCES

1. Brown MA, Kenna T, Wordsworth BP. Genetics of ankylosing spondylitis-insights into pathogenesis. Nat Rev Rheumatol. 2016 Feb; 12(2):81-91.
2. Dougados M, Baeten D. Spondyloarthritis. Lancet. 2011 Jun 18; 377(9783):2127-2137.
3. Sieper J, Braun J, Dougados M*, et al.* Axial spondyloarthritis. Nat Rev Dis Primers. 2015 Jul 9; 1:15013.
4. Taurog JD, Chhabra A, Colbert RA. Ankylosing Spondylitis and Axial Spondyloarthritis. N Engl J Med. 2016 Jun 30; 374(26):2563-2574.
5. Salmon-Ceron D, Tubach F, Lortholary O*, et al.* Drug-specific risk of non-tuberculosis opportunistic infections in patients receiving anti-TNF therapy reported to the 3-year prospective French RATIO registry. Ann Rheum Dis. 2011 Apr; 70(4):616-623.
6. Bongartz T, Sutton AJ, Sweeting MJ*, et al.* Anti-TNF antibody therapy in rheumatoid arthritis and the risk of serious infections and malignancies: systematic review and meta-analysis of rare harmful effects in randomized controlled trials. JAMA. 2006 May 17; 295(19):2275-2285.
7. Keane J, Gershon S, Wise RP*, et al.* Tuberculosis associated with infliximab, a tumor necrosis factor alpha-neutralizing agent. N Engl J Med. 2001 Oct 11; 345(15):1098-1104.
8. Tubach F, Salmon D, Ravaud P*, et al.* Risk of tuberculosis is higher with anti-tumor necrosis factor monoclonal antibody therapy than with soluble tumor necrosis factor receptor therapy: The three-year prospective French Research Axed on Tolerance of Biotherapies registry. Arthritis Rheum. 2009 Jul; 60(7):1884-1894.
9. Ermann J, Rao DA, Teslovich NC*, et al.* Immune cell profiling to guide therapeutic decisions in rheumatic diseases. Nat Rev Rheumatol. 2015 Sep; 11(9):541-551.

10. Reveille JD. Biomarkers for diagnosis, monitoring of progression, and treatment responses in ankylosing spondylitis and axial spondyloarthritis. Clin Rheumatol. 2015 Jun; 34(6):1009-1018.
11. Menegatti S, Bianchi E, Rogge L. Anti-TNF Therapy in Spondyloarthritis and Related Diseases, Impact on the Immune System and Prediction of Treatment Responses. Front Immunol. 2019; 10:382.
12. Baeten D, Sieper J, Braun J*, et al.* Secukinumab, an Interleukin-17A Inhibitor, in Ankylosing Spondylitis. N Engl J Med. 2015 Dec 24; 373(26):2534-2548.
13. McInnes IB, Mease PJ, Kirkham B*, et al.* Secukinumab, a human anti-interleukin-17A monoclonal antibody, in patients with psoriatic arthritis (FUTURE 2): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet. 2015 Sep 19; 386(9999):1137-1146.
14. Rudwaleit M, Landewe R, van der Heijde D*, et al.* The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part I): classification of paper patients by expert opinion including uncertainty appraisal. Ann Rheum Dis. 2009 Jun; 68(6):770-776.
15. Rudwaleit M, van der Heijde D, Landewe R*, et al.* The Assessment of SpondyloArthritis International Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. Ann Rheum Dis. 2011 Jan; 70(1):25-31.
16. Machado P, Landewe R, Lie E*, et al.* Ankylosing Spondylitis Disease Activity Score (ASDAS): defining cut-off values for disease activity states and improvement scores. Ann Rheum Dis. 2011 Jan; 70(1):47-53.
17. Duffy D, Rouilly V, Libri V*, et al.* Functional Analysis via Standardized Whole-Blood Stimulation Systems Defines the Boundaries of a Healthy Immune Response to Complex Stimuli. Immunity. 2014 Mar 20; 40(3):436-450.
18. Latis E, Michonneau D, Leloup C*, et al.* Cellular and molecular profiling of T-cell subsets at the onset of human acute GVHD. Blood Adv. 2020 Aug 25; 4(16):3927-3942.
19. Yahia-Cherbal H, Rybczynska M, Lovecchio D*, et al.* NFAT primes the human RORC locus for RORgammat expression in CD4(+) T cells. Nat Commun. 2019 Oct 16; 10(1):4698.
20. Yaari G, Bolen CR, Thakar J*, et al.* Quantitative set analysis for gene expression: a method to quantify gene set differential expression including gene-gene correlations. Nucleic Acids Res. 2013 Oct; 41(18):e170.
21. Ritchie ME, Phipson B, Wu D*, et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015 Apr 20; 43(7):e47.
22. Sherlock JP, Joyce-Shaikh B, Turner SP*, et al.* IL-23 induces spondyloarthropathy by acting on ROR-gammat(+) CD3(+)CD4(-)CD8(-) entheseal resident T cells. Nat Med. 2012; 18(7):1069-1076.
23. Netea MG, Joosten LA, van der Meer JW*, et al.* Immune defence against Candida fungal infections. Nat Rev Immunol. 2015 Oct; 15(10):630-642.
24. Bridgewood C, Watad A, Russell T*, et al.* Identification of myeloid cells in the human enthesis as the main source of local IL-23 production. Ann Rheum Dis. 2019 Jul; 78(7):929-933.
25. Machado PM, Landewe R, Heijde DV*, et al.* Ankylosing Spondylitis Disease Activity Score (ASDAS): 2018 update of the nomenclature for disease activity states. Ann Rheum Dis. 2018 Oct; 77(10):1539-1540.

26. Schmiedel BJ, Singh D, Madrigal A*, et al.* Impact of Genetic Polymorphisms on Human Immune Cell Gene Expression. Cell. 2018 Nov 29; 175(6):1701-1715 e1716.
27. Joyce-Shaikh B, Bigler ME, Chao CC*, et al.* Myeloid DAP12-associating lectin (MDL)-1 regulates synovial inflammation and bone erosion associated with autoimmune arthritis. J Exp Med. 2010 Mar 15; 207(3):579-589.
28. Duffy D, Rouilly V, Braudeau C*, et al.* Standardized whole blood stimulation improves immunomonitoring of induced immune responses in multi-center study. Clin Immunol. 2017 Oct; 183:325-335.
29. Wang XB, Ellis JJ, Pennisi DJ*, et al.* Transcriptome analysis of ankylosing spondylitis patients before and after TNF-alpha inhibitor therapy reveals the pathways affected. Genes Immun. 2017 Sep; 18(3):184-190.
30. Charles P, Elliott MJ, Davis D*, et al.* Regulation of cytokines, cytokine inhibitors, and acute-phase proteins following anti-TNF-alpha therapy in rheumatoid arthritis. J Immunol. 1999 Aug 1; 163(3):1521-1528.
31. Sieper J, Porter-Brown B, Thompson L*, et al.* Assessment of short-term symptomatic efficacy of tocilizumab in ankylosing spondylitis: results of randomised, placebo-controlled trials. Ann Rheum Dis. 2014 Jan; 73(1):95-100.
32. Vos AC, Wildenberg ME, Arijs I*, et al.* Regulatory macrophages induced by infliximab are involved in healing in vivo and in vitro. Inflamm Bowel Dis. 2012 Mar; 18(3):401-408.
33. Vos AC, Wildenberg ME, Duijvestein M*, et al.* Anti-tumor necrosis factor-alpha antibodies induce regulatory macrophages in an Fc region-dependent manner. Gastroenterology. 2011 Jan; 140(1):221-230.
34. Degboe Y, Rauwel B, Baron M*, et al.* Polarization of Rheumatoid Macrophages by TNF Targeting Through an IL-10/STAT3 Mechanism. Front Immunol. 2019; 10:3.
35. Udalova IA, Mantovani A, Feldmann M. Macrophage heterogeneity in the context of rheumatoid arthritis. Nat Rev Rheumatol. 2016 Aug; 12(8):472-485.
36. Schett G, Neurath MF. Resolution of chronic inflammatory disease: universal and tissue-specific concepts. Nat Commun. 2018 Aug 15; 9(1):3261.
37. Kratochvill F, Neale G, Haverkamp JM*, et al.* TNF Counterbalances the Emergence of M2 Tumor Macrophages. Cell Rep. 2015 Sep 22; 12(11):1902-1914.
38. Marino S, Cilfone NA, Mattila JT*, et al.* Macrophage polarization drives granuloma outcome during Mycobacterium tuberculosis infection. Infect Immun. 2015 Jan; 83(1):324-338.
39. Schett G, Lories RJ, D'Agostino MA*, et al.* Enthesitis: from pathophysiology to treatment. Nat Rev Rheumatol. 2017 Nov 21; 13(12):731-741.
40. Boniface K, Bak-Jensen KS, Li Y*, et al.* Prostaglandin E2 regulates Th17 cell differentiation and function through cyclic AMP and EP2/EP4 receptor signaling. J Exp Med. 2009 Mar 16; 206(3):535-548.
41. Yao C, Sakata D, Esaki Y*, et al.* Prostaglandin E2-EP4 signaling promotes immune inflammation through Th1 cell differentiation and Th17 cell expansion. Nat Med. 2009 Jun; 15(6):633-640.
42. Ellinghaus D, Jostins L, Spain SL*, et al.* Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. Nat Genet. 2016 May; 48(5):510-518.

43. Cheung R, Shen F, Phillips JH*, et al.* Activation of MDL-1 (CLEC5A) on immature myeloid cells triggers lethal shock in mice. J Clin Invest. 2011 Nov; 121(11):4446-4461.
44. Wu MF, Chen ST, Yang AH*, et al.* CLEC5A is critical for dengue virus-induced inflammasome activation in human macrophages. Blood. 2013 Jan 3; 121(1):95-106.
45. Sieberts SK, Zhu F, Garcia-Garcia J*, et al.* Crowdsourced assessment of common genetic contribution to predicting anti-TNF treatment response in rheumatoid arthritis. Nat Commun. 2016 Aug 23; 7:12460.

Table 1. Clinical characteristics of the 80 axSpA patients included in the study

Characteristic	SpA (*n* = 80)
Female n (%)	25 (31%)
Median (IQR) age at sampling (years)	37 (19-64)
Median (IQR) disease duration (years)	2 (0-33)
HLA-B27 positive n (%)	63 (79%)
Current smokers n (%)	40 (50%)
Median (IQR) CRP (mg/l) at baseline	6.06 (0.09-62)
Median (IQR) BASDAI at baseline	49.80 (9.40-90)
Median (IQR) ASDAS at baseline	3.05 (1.13-4.79)
Axial involvement n (%)	80 (100%)
Axial and enthesial involvement n (%)	38 (47.5%)
Radiological sacroiliitis n (%)	48 (60%)
MRI sacroiliitis n (%)	63 (79%)
TNF blocker	
Soluble TNF receptor Etanercept n (%)	53 (66.25%)
Monoclonal antibody Adalimumab n (%)	13 (16.25%)
Monoclonal antibody Golimumab n (%)	13 (16.25%)
Monoclonal antibody Infliximab n (%)	1 (1.25%)
Extra-articular manifestations	
Psoriasis n (%)	16 (20%)
Uveitis n (%)	26 (33%)
IBD (%)	3 (4%)
Response at D90	
Median (IQR) C-reactive protein (mg/l) at D90	1.95 (0-51.80)
Median (IQR) BASDAI at D90	23.50 (0-78)
Median (IQR) ASDAS at D90	1.44 (0.64-3.45)
Patients with major ASDAS improvement n (%)	20 (25%)
Patients with clinically important improvement ASDAS n (%)	30 (37.5%)
Non-Responder ASDAS n (%)	30 (37.5%)
Non-Responder ASDAS treated with Etanercept n (%)	22 (73.33%)* (41.5%)**
Non-Responder ASDAS treated with Adalimumab n (%)	5 (16.67%)* (38.5%)**
Non-Responder ASDAS treated with Golimumab n (%)	3 (10%)* (23.1%)**
Non-Responder ASDAS treated with Infliximab n (%)	0 (0 %)
Non-Responder BASDAI50 n (%)	52 (65%)

Median and interquartile range (IQR) or percentages are shown; CRP, C-reactive protein; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; ASDAS, Ankylosing Spondylitis Disease Activity Score; IBD, inflammatory bowel disease
*percentage of total Non-Responders ** percentage of patients treated with the indicated drug

Table 2. Genes differentially expressed between responders and non-responders to TNFi

Gene ID	Log Fold-change	P-value	Adjusted P-value (R/NR)
*PLAUR*_LPS	0.4816	2.86E-06	0.0023
*ITGB1*_LPS	0.2860	5.29E-06	0.0023
*CD14*_LPS	0.5704	1.78E-05	0.0041
*CCL20*_LPS	0.6264	2.04E-05	0.0041
*IL1R1*_LPS	0.7803	2.48E-05	0.0041
*IRAK1*_LPS	0.2964	3.41E-05	0.0041
*IRAK3*_LPS	0.3977	3.49E-05	0.0041
*CLEC5A*_LPS	0.7180	3.8E-05	0.0041
*ITGA5*_LPS	0.2684	0.0001	0.0066
*LTB4R*_LPS	0.5985	0.0001	0.0069
*LTA*_LPS	-0.3366	0.0001	0.0074
*BST1*_LPS	0.5186	0.0001	0.0077
*IL1RAP*_LPS	0.4707	0.0001	0.0083
*CD58*_LPS	0.2690	0.0001	0.0083
*CEBPB*_LPS	0.2989	0.0001	0.0083
*IL8*_LPS	0.5694	0.0002	0.0083
*IFNGR1*_LPS	0.3022	0.0002	0.0097
*IL1R2*_LPS	0.4411	0.0003	0.0121
*CXCL9*_LPS	-2.0206	0.0003	0.0121
*TNFRSF1B*_LPS	0.3157	0.0003	0.0121
*IL6R*_LPS	0.3360	0.0003	0.0121
*NLRP3*_LPS	0.3896	0.0003	0.0121
*CTNNB1*_LPS	0.1495	0.0003	0.0121
*FCGRT*_LPS	0.3159	0.0003	0.0121
*ITGAX*_LPS	0.3600	0.0003	0.0121
*IFNG*_LPS	-1.4398	0.0005	0.0180
*CXCL1*_LPS	0.4515	0.0006	0.0180
*FCGR2A*_LPS	0.2634	0.0006	0.0180
*ITGA6*_SEB	-0.2569	0.0006	0.0180
*PRKCD*_LPS	0.3330	0.0006	0.0187
*ZEB1*_LPS	0.3487	0.0007	0.0201
*CLEC7A*_LPS	0.3795	0.0007	0.0201
*PECAM1*_LPS	0.4050	0.0008	0.0218
*IRAK1*_SEB	0.1988	0.0009	0.0231
*APP*_LPS	0.1938	0.0010	0.0237
*FCER1G*_LPS	0.2902	0.0011	0.0255
*ICAM5*_SEB	0.5363	0.0011	0.0257
*IL8*_SEB	0.3880	0.0011	0.0257
*PLAUR*_SEB	0.3067	0.0012	0.0270
*IL7R*_SEB	-0.1991	0.0012	0.0270
*IGF2R*_LPS	0.2310	0.0013	0.0270
*IKZF3*_LPS	-0.1544	0.0013	0.0276
*TNFRSF8*_LPS	0.3647	0.0014	0.0276
*NFIL3*_LPS	0.2830	0.0015	0.0290
*LIF*_LPS	1.0229	0.0015	0.0292
*MBP*_LPS	0.2114	0.0016	0.0296
*TP53*_LPS	-0.1846	0.0016	0.0296
*CXCL2*_LPS	0.4914	0.0020	0.0371
*CXCR4*_LPS	0.2833	0.0022	0.0398
*ATG7*_LPS	0.2486	0.0024	0.0412
*CRADD*_SEB	0.3238	0.0025	0.0435
*PLAU*_LPS	0.4759	0.0027	0.0452
*SPP1*_SEB	0.4451	0.0028	0.0452
*SKI*_LPS	0.1760	0.0028	0.0452
*CXCR1*_LPS	0.6786	0.0029	0.0452
*TLR2*_LPS	0.2718	0.0031	0.0471
*MAP4K4*_LPS	0.2504	0.0031	0.0471
*DUSP4*_LPS	0.4570	0.0031	0.0471

Figures and figure legends

Figure 1

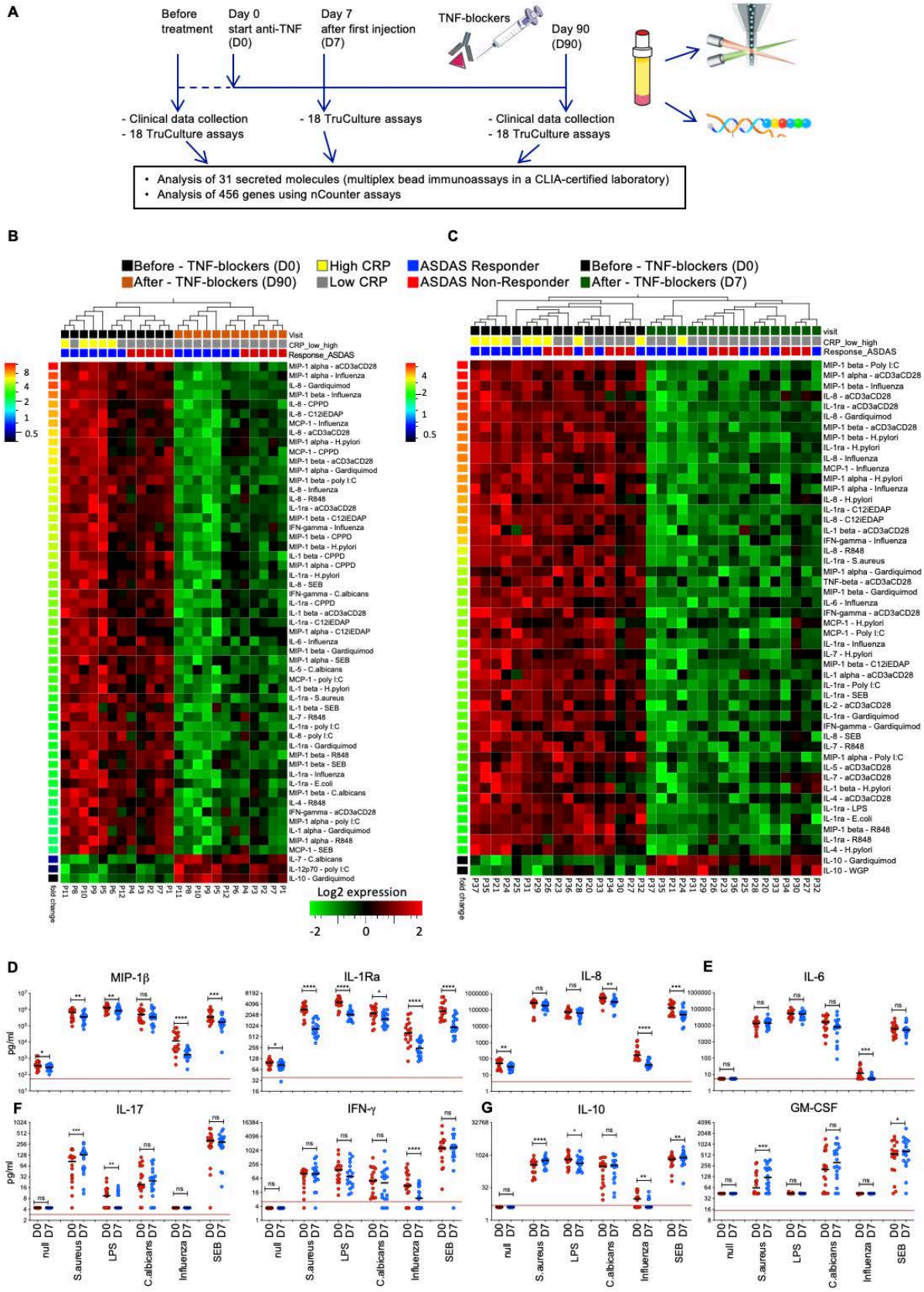


Figure 1. An immunologic signature of anti-TNF therapy

(**A**) Study design. Blood samples were collected from axSpA patients prior to (D0), 7 days (D7, for a subset of patients), and 3 months (D90) after beginning TNFi treatment. Clinical efficacy was monitored at D90 according to the current standard of care. (**B**) The levels of 31 secreted molecules in response to 18 different immune stimuli were compared in samples from 12 patients at D0 (black rectangles) and D90 (orange rectangles). Patients with CRP-levels > 6 mg/l are marked with yellow rectangles, while CRP-levels < 6 mg/l are indicated with grey rectangles. Patients responding to anti-TNF therapy (delta ASDAS ≥ 1.1) are marked in blue and non-responders (delta ASDAS < 1.1) are marked in red. The heatmap shows the levels of differentially secreted proteins (paired t-test, FDR ≤ 0.05, fold-change ≥ 2, red indicates higher and green lower levels of protein secretion). Analyte-stimulus combinations were ranked by decreasing fold-change (color-code bar, top left); patient IDs are indicated below the heatmaps. (**C**) The same analysis as in (**B**) was performed for additional 17 axSpA patients, sampled at D0 (blue rectangles) and D7 (green rectangles). (**D**-**G**) Levels of proteins identified in (**C**), for 5 representative stimuli and the unstimulated (null) condition, in 17 axSpA patients at D0 (red) and D7 (blue). Red lines indicate the least detectable dose (LDD) for each assay. *P*-values were calculated using a Wilcoxon matched-pairs test (SpA patients D0 versus D7) *: P<0.05; **: P<0.01; ***: P<0.001; ****: P<0.0001; ns: not significant. Horizontal black bars indicate the median. Y-axes are log2 scales

Figure 2

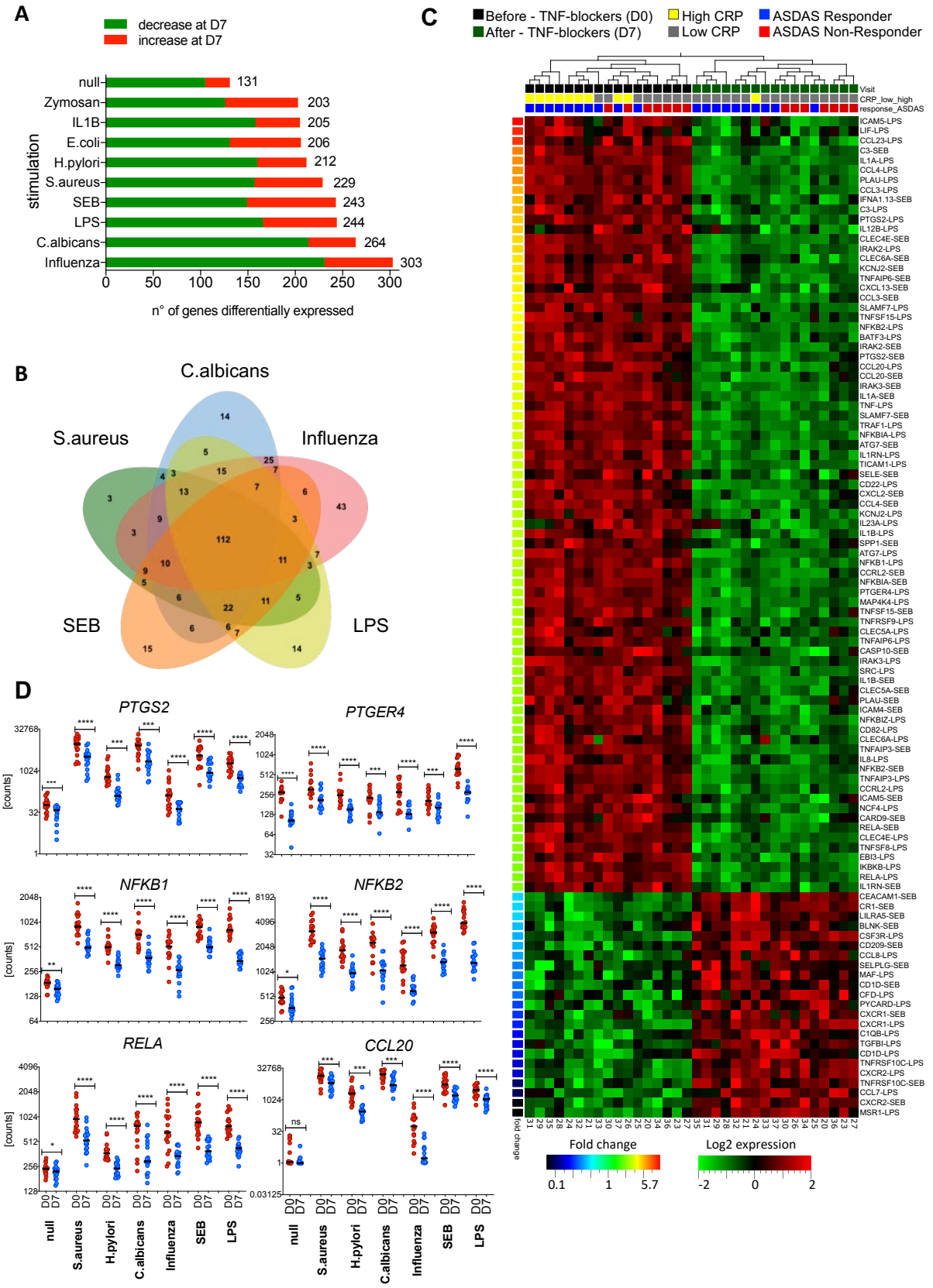

A
decrease at D7
increase at D7
null 131
Zymosan 203
IL1B 205
E.coli 206
H.pylori 212
S.aureus 229
SEB 243
LPS 244
C.albicans 264
Influenza 303
stimulation
0 50 100 150 200 250 300
n° of genes differentially expressed
B
C.albicans
S.aureus
Influenza
SEB
LPS
14
5
25
7
15
4
3
13
7
6
43
3
9
3
112
11
7
3
9
10
5
6
22
11
5
6
6
7
15
14
C
Before - TNF-blockers (D0)
After - TNF-blockers (D7)
High CRP
Low CRP
ASDAS Responder
ASDAS Non-Responder
Visit
CRP_low_high
response_ASDAS
ICAM5-LPS
LIF-LPS
CCL23-LPS
C3-SEB
IL1A-LPS
CCL4-LPS
PLAU-LPS
CCL3-LPS
IFNA1.13-SEB
C3-LPS
PTGS2-LPS
IL12B-LPS
CLEC4E-SEB
IRAK2-LPS
CLEC6A-SEB
KCNJ2-SEB
TNFAIP6-SEB
CXCL13-SEB
CCL3-SEB
SLAMF7-LPS
TNFSF15-LPS
NFKB2-LPS
BATF3-LPS
IRAK2-SEB
PTGS2-SEB
CCL20-LPS
CCL20-SEB
IRAK3-SEB
IL1A-SEB
TNF-LPS
SLAMF7-SEB
TRAF1-LPS
NFKBIA-LPS
ATG7-SEB
IL1RN-LPS
TICAM1-LPS
SELE-SEB
CD22-LPS
CXCL2-SEB
CCL4-SEB
KCNJ2-LPS
IL23A-LPS
IL1B-LPS
SPP1-SEB
ATG7-LPS
NFKB1-LPS
CCRL2-SEB
NFKBIA-SEB
PTGER4-LPS
MAP4K4-LPS
TNFSF15-SEB
TNFRSF9-LPS
CLEC5A-LPS
TNFAIP6-LPS
CASP10-SEB
IRAK3-LPS
SRC-LPS
IL1B-SEB
CLEC5A-SEB
PLAU-SEB
ICAM4-SEB
NFKBIZ-LPS
CD82-LPS
CLEC6A-LPS
TNFAIP3-SEB
IL8-LPS
NFKB2-SEB
TNFAIP3-LPS
CCRL2-LPS
ICAM5-SEB
NCF4-LPS
CARD9-SEB
RELA-SEB
CLEC4E-LPS
TNFSF8-LPS
EBI3-LPS
IKBKB-LPS
RELA-LPS
IL1RN-SEB
CEACAM1-SEB
CR1-SEB
LILRA5-SEB
BLNK-SEB
CSF3R-LPS
CD209-SEB
CCL8-LPS
SELPLG-SEB
MAF-LPS
CD1D-SEB
CFD-LPS
PYCARD-LPS
CXCR1-SEB
CXCR1-LPS
C1QB-LPS
TGFBI-LPS
CD1D-LPS
TNFRSF10C-LPS
CXCR2-LPS
TNFRSF10C-SEB
CCL7-LPS
CXCR2-SEB
MSR1-LPS
fold change
Fold change
0.1 1 5.7
Log2 expression
-2 0 2
D
PTGS2
PTGER4
NFKB1
NFKB2
RELA
CCL20
[counts]
null
S.aureus
H.pylori
C.albicans
Influenza
SEB
LPS

Figure 2. TNF-blockers strongly affect key transcriptional networks of innate immune responses

(**A**) Number of genes differentially expressed in 10 different TruCulture stimulation assays performed at D0 and D7 (17 patients, paired t-test, FDR ≤ 0.05). (**B**) Venn diagram of the genes differentially expressed as in (**A**), in 5 representative stimulation conditions. (**C**) Heatmap showing the genes most affected by TNFi (D0, black rectangles *versus* D7, green) in LPS and SEB stimulation conditions. Paired t-test, FDR ≤ 0.005 and fold-difference threshold of ≥ 2. Gene-stimulus combinations were ranked by decreasing fold-change (color code bottom left bar). (**D**) Expression levels of NF-κB family members for the unstimulated TruCulture assay and 5 representative stimuli at D0 (red) and D7 (blue) after initiation of TNFi therapy. *P*-values were determined using a Wilcoxon matched-pairs test (D0 *versus* D7, *: $P<0.05$; **: $P<0.01$; ***: $P<0.001$; ****: $P<0.0001$; ns: not significant, n = 17). Horizontal black bars indicate the median.

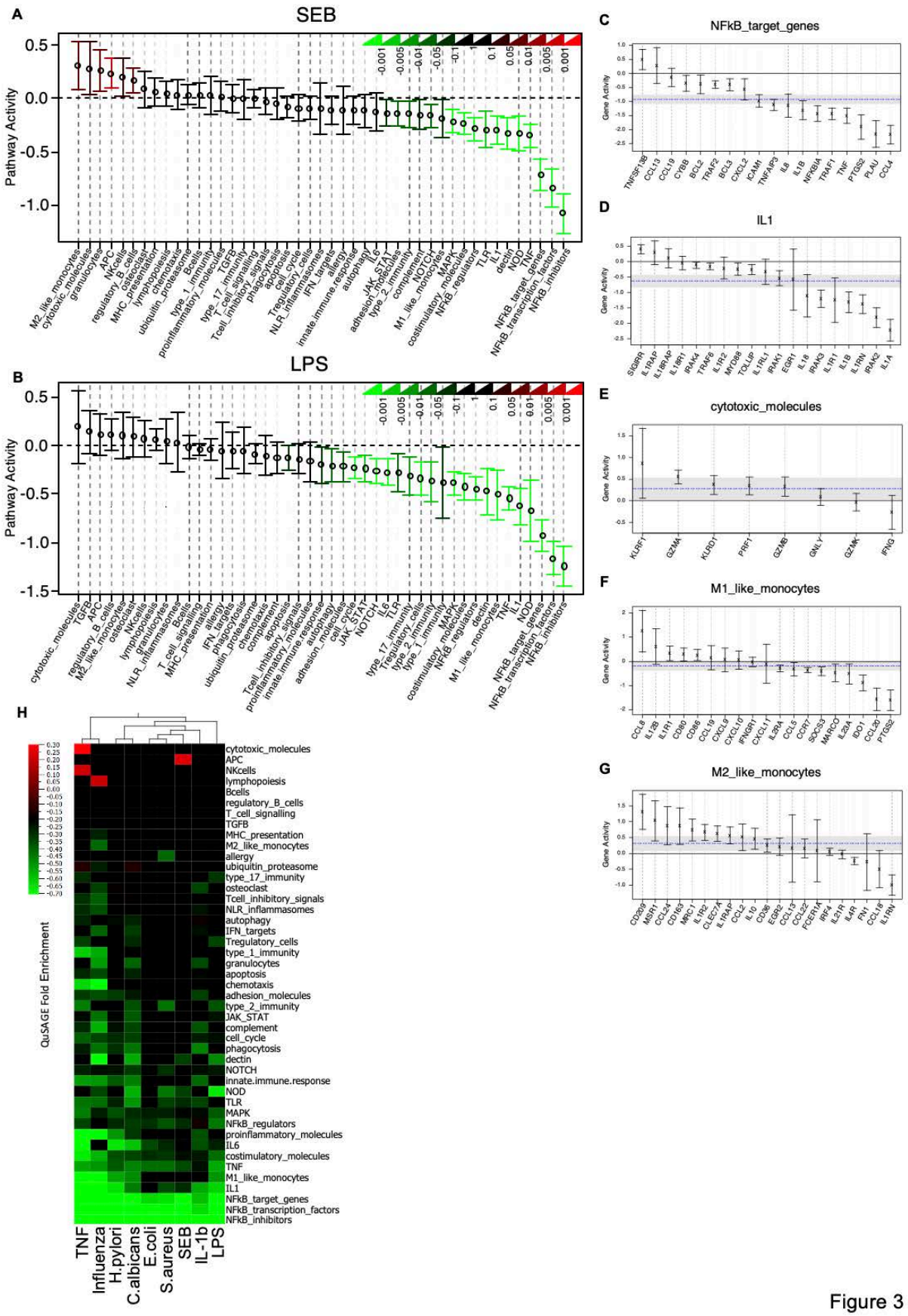

A
SEB
Pathway Activity
B
LPS
C
NFkB_target_genes
D
IL1
E
cytotoxic_molecules
F
M1_like_monocytes
G
M2_like_monocytes
Gene Activity
H
QuSAGE Fold Enrichment
TNF
Influenza
H.pylori
C.albicans
E.coli
S.aureus
SEB
IL-1b
LPS

Figure 3

Figure 3. Modular transcriptional repertoire analysis reveals multiple mechanisms of TNF-blocker action in SpA

(**A, B**) Effect of anti-TNF therapy on the activity of 45 gene modules (**online supplementary table 5**) generated from 456 immune-related genes. Whole-blood cultures were stimulated with SEB (**A**) or LPS (**B**). For each gene module, the mean activity fold-change and 95% confidence interval are plotted and color-coded according to their FDR-corrected *P*-values (means compared to fold-change zero). Confidence intervals overlapping the horizontal dotted line indicate statistically significant increased or decreased module activity at D7 as compared to D0. (**C-G**) Detailed gene activity in five representative modules with decreased (**C, D, E**, LPS-stimulation) or increased (**F, G**, SEB-stimulation) pathway activity after anti-TNF therapy. The cultures were stimulated with LPS and SEB, respectively. Represented are the mean fold-change and 95% confidence interval for individual genes in each module. The horizontal dashed blue line and the grey band indicate the mean differential expression of all genes in the module at D7 versus D0, and the 95% confidence interval. (**H**) QuSAGE fold-enrichment of gene set activity in 9 different stimulated cultures at D7 *versus* D0. For each module, the mean fold-change is color-coded to indicate increased (red) or decreased (green) module activity. Only changes reaching a significance threshold of FDR ≤ 0.01 are represented.

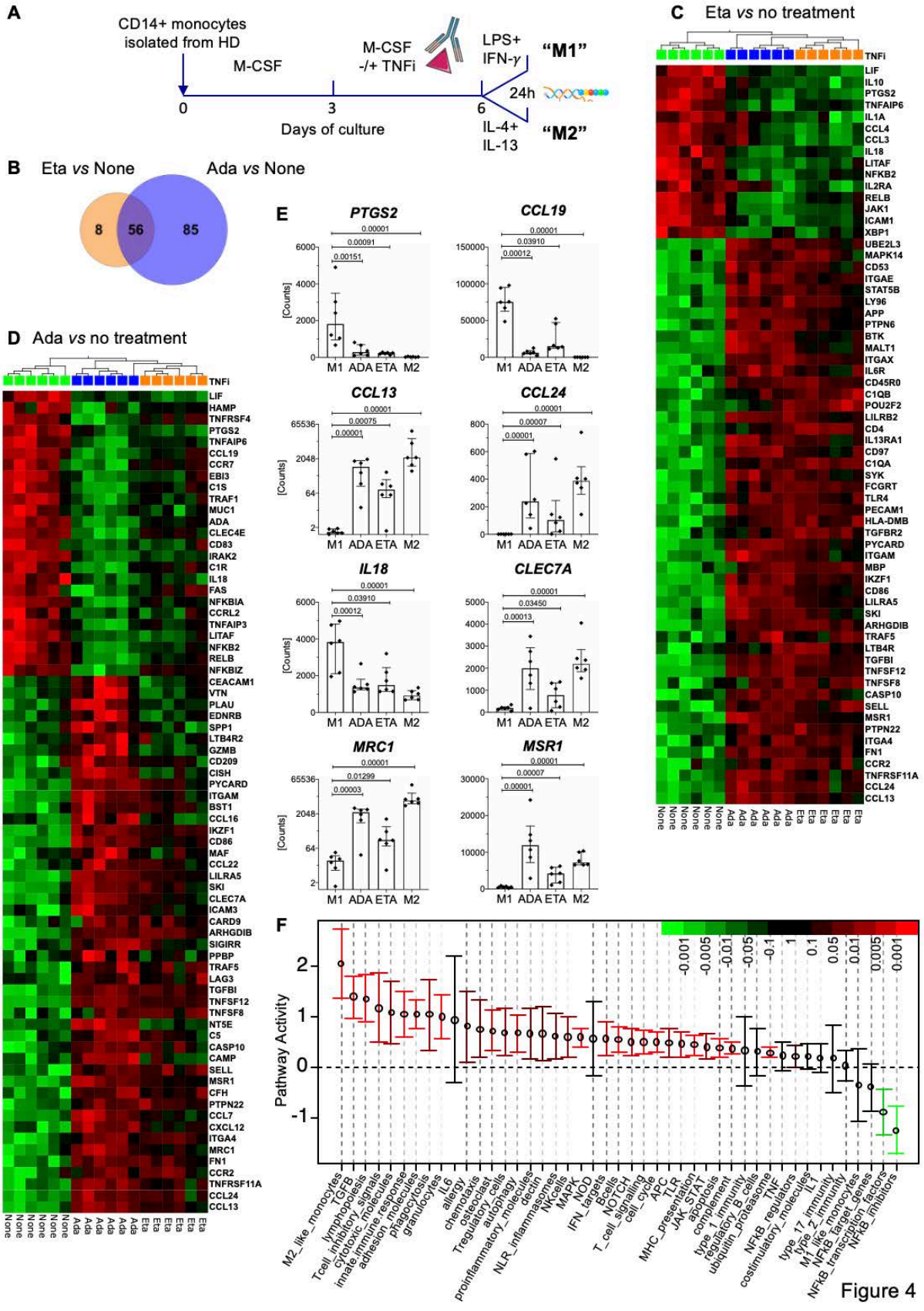

A
CD14+ monocytes isolated from HD
M-CSF
M-CSF -/+ TNFi
LPS+ IFN-γ
"M1"
24h
IL-4+ IL-13
"M2"
0
3
6
Days of culture
B Eta vs None
Ada vs None
8
56
85
C Eta vs no treatment
D Ada vs no treatment
E
PTGS2
CCL19
CCL13
CCL24
IL18
CLEC7A
MRC1
MSR1
[Counts]
M1 ADA ETA M2
F
Pathway Activity
M2_like_monocytes
TGFB
lymphopoiesis
Tcell_inhibitory_signals
cytotoxic_molecules
innate_immune_response
adhesion_molecules
Phagocytosis
granulocytes
IL6
allergy
chemotaxis
osteoclast
Treg
autophagy
cells
proinflammatory_molecules
dectin
NLR_inflammasomes
Nkcells
MAPK
NOD
IFN_targets
Bcells
NOTCH
T_cell_signalling
cell_cycle
APC
TLR
MHC_presentation
JAK_STAT
apoptosis
complement
type_1_immunity
regulatory_B_cells
ubiquitin_proteasome
NFkB_regulators
TNF
costimulatory_molecules
IL1
type_17_immunity
type_2_immunity
M1_like_monocytes
NFkB_target_genes
NFkB_transcription_factors
NFkB_inhibitors

Figure 4

Figure 4. TNFi have largely overlapping effects on *in vitro* differentiated M1-type macrophages

(**A**) Study design. CD14+ cells isolated from healthy donors were differentiated *in vitro* into macrophages in the presence or absence of etanercept or adalimumab. TNFi were added at day 3 and macrophages were polarized to the M1 subset in the presence or absence of etanercept (Eta) or adalimumab (Ada). Gene expression was analyzed with the nCounter Human Immunology v2 panel and with LIMMA, (paired sample adjusted *P*-value threshold 0.01). (**B**) Venn Diagram showing the overlap of genes affected by Eta or Ada. Analysis of paired samples with LIMMA, adjusted *P*-value threshold 0.01). (**C**, **D**) Heatmaps showing the genes most affected by Eta (orange rectangles) versus no treatment (green rectangles) (**C**) and Ada (blue rectangles) versus no treatment (**D**) in macrophages stimulated for 24h with LPS and IFN-g ("M1" polarization). (**C**) Paired t-test, Eta versus no treatment, adjusted p-value threshold 0.01. Included also gene expression levels for Ada-treated samples for the same genes. (**D**) Paired t-test, Ada versus no treatment and fold-change threshold of ≥ 2. Included also gene expression levels for Eta-treated samples for the same genes. Samples were ordered by hierarchical clustering and genes were ranked by decreasing fold-change. (**E**) Shown are the mRNA levels of 8 selected genes from (**C**) and (**D**) in untreated M1-polarized macrophages (M1), M1 macrophages treated with Ada, M1 macrophages treated with Eta or untreated M2-polarized and macrophages (M2). Symbols represent individual data points, boxes the median and whiskers the interquartile range. Adjusted *P*-values are those of the LIMMA analysis. (**F**) Effect of Ada on the activity of 45 gene modules (**online supplementary table 5**) as in **Figure 3**. For each gene module, the mean activity fold-change and 95% confidence interval are plotted and color-coded according to their FDR-corrected *P*-values compared to zero. Red and green bars indicate statistically significant increased or decreased module activity, respectively, in M1 polarized macrophages treated with Ada versus no treatment.

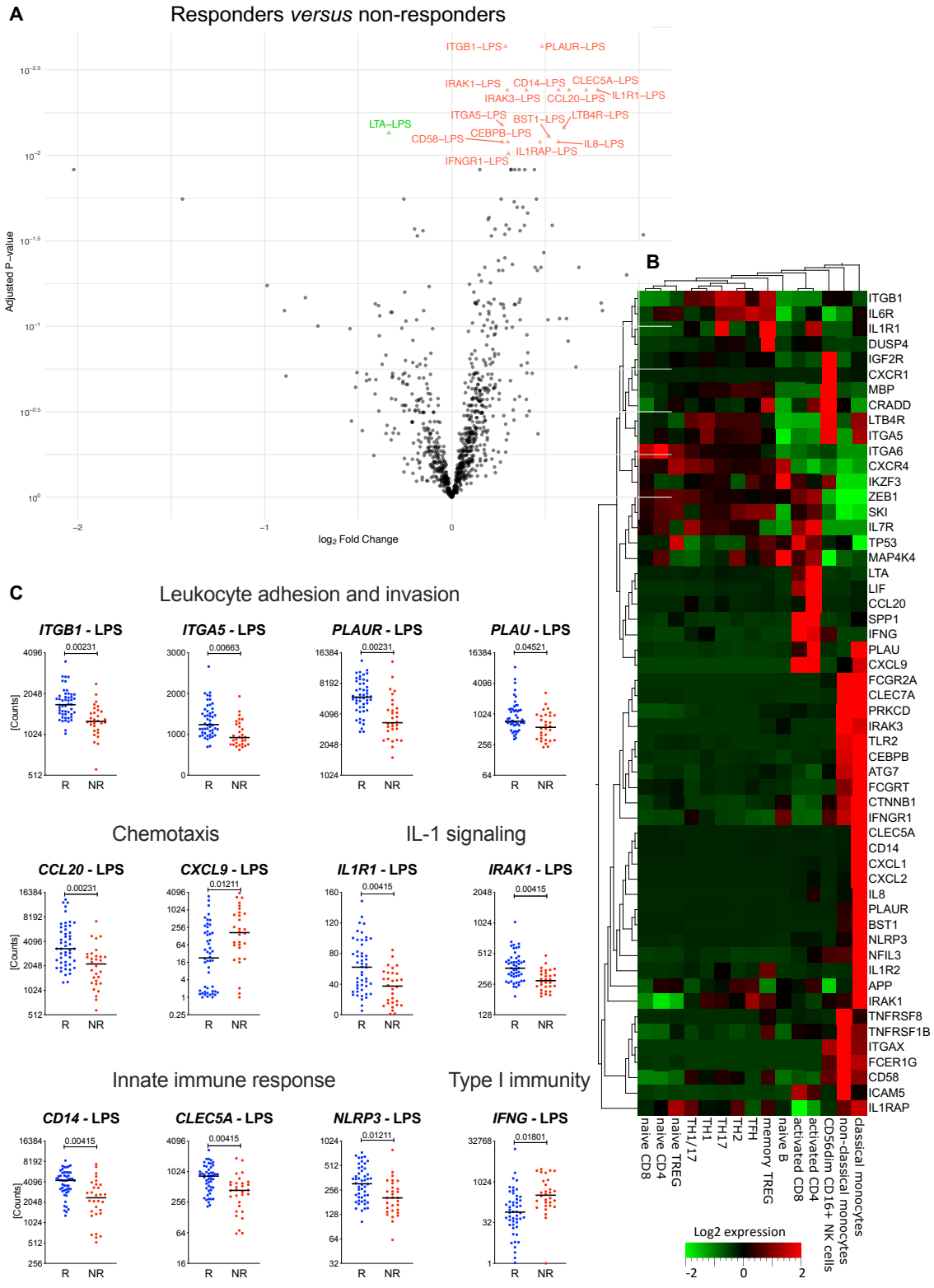


Figure 5

Figure 5. Immune gene expression associated with therapeutic responses to anti-TNF therapy

(**A**) Volcano plot representation of genes differentially expressed between 50 SpA patients responding to anti-TNF therapy and 30 non-responders in whole-blood cultures stimulated with LPS or SEB before initiation of therapy; red triangles: genes higher in responders; green triangle: higher in non-responders (LIMMA analysis, adjusted p-value < 0.05). Expression levels and fold-change values of the 58 gene-stimulus combinations (corresponding to 55 genes) that are the most differentially expressed between responders and non-responders are reported in **Table 2**. (**B**). The heatmap shows the expression levels of the differentially expressed genes in different immune cell subpopulations. Gene expression data were extracted from the DICE database (http://dice-database.org/). (**C**) The expression levels of selected gene-stimulus combinations correlated with treatment response are plotted before treatment initiation (D0). Patients with major or clinically important improvement of disease activity were grouped together as responders and are represented in blue (R, blue, *n*=50). Non-responders are represented in red (NR, red, *n*=30). The horizontal black line represents the median. Statistical significance was tested using LIMMA analysis (responders *versus* non-responders) and adjusted *P*-values are indicated above the graph.

ANNEX II

Supplementary information (Manuscript I)

Characterization of peripheral and entheseal IL-17 secreting Mucosal Associated Invariant T (MAIT) cells in Axial Spondyloarthritis

Supplementary figures and legends from manuscript 1 (Part II results)

Figure S1

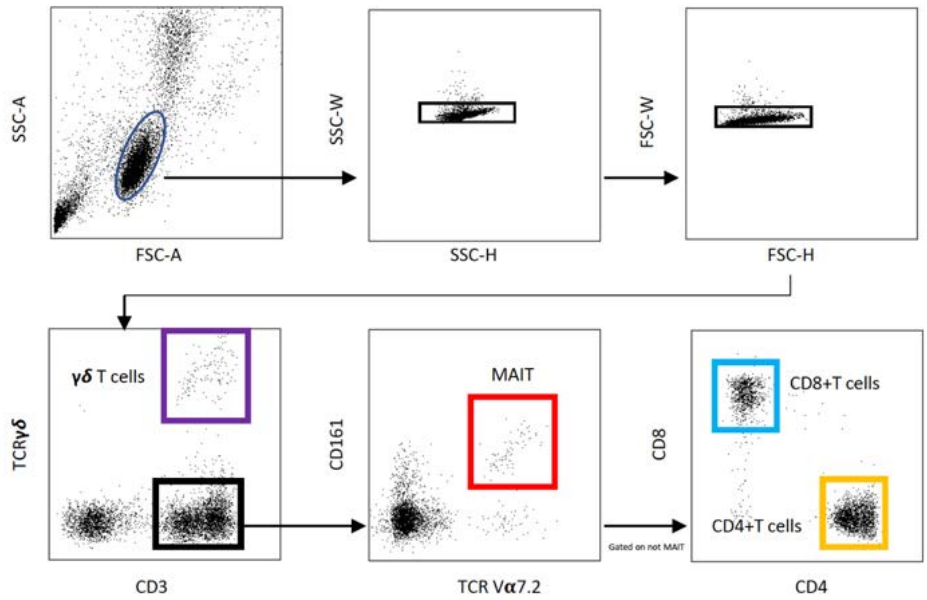


Figure S1: Flow cytometry gating strategy 1. Flow cytometry gating strategy for the isolation of MAIT, **γδ** T cells, CD4+ and CD8+ T cells.

Figure S2

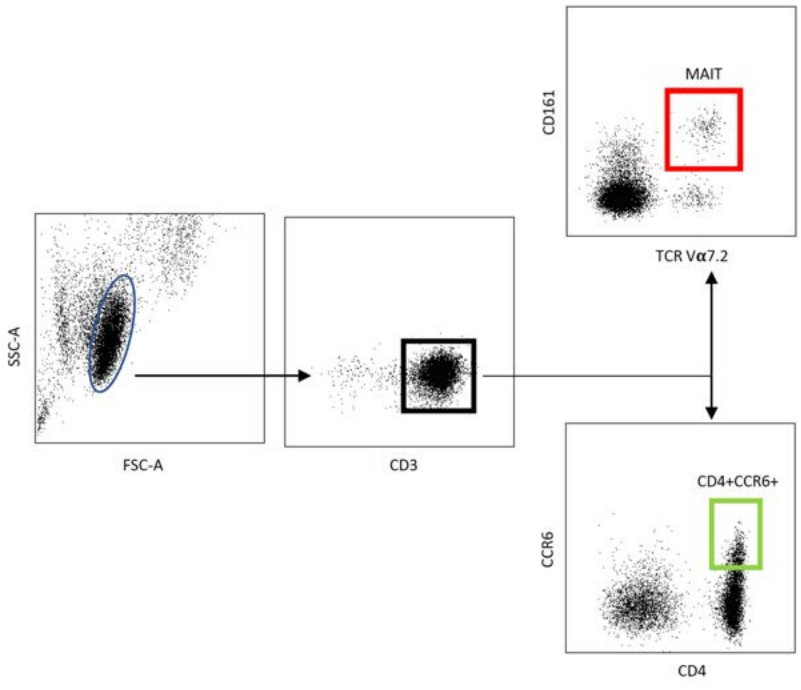


Figure S2: Flow cytometry gating strategy 2. Flow cytometry gating strategy, after CD3 positive magnetic sorting, of MAIT and CD4+ CCR6+T cells.

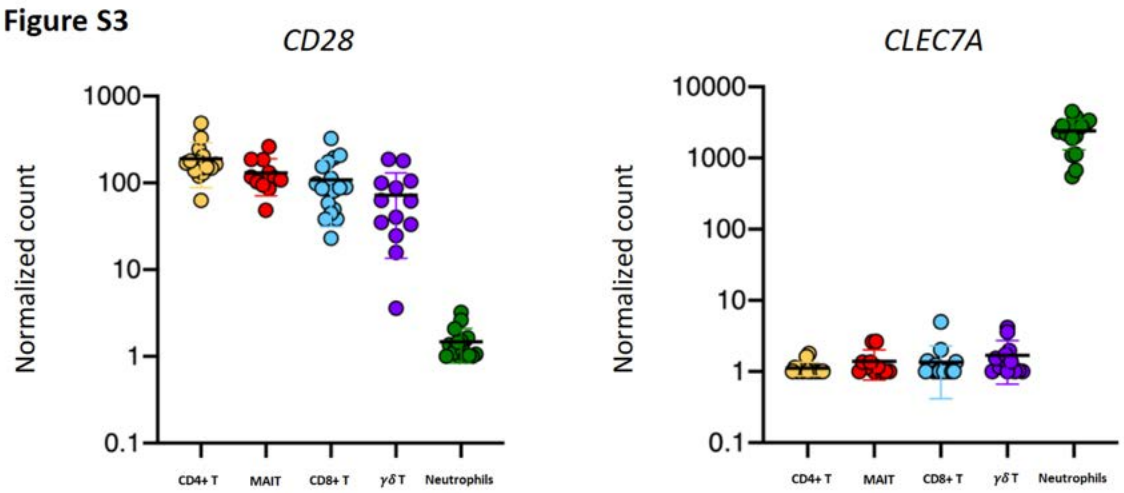


Figure S3: Transcripts levels of *CD28* and *CLEC7A* presented as normalized counts in sorted CD4+ T cells (in yellow), MAIT cells (in red), CD8+T cells (in blue), $\gamma\delta$ T cells (in purple) and neutrophils (in green) after 2h of stimulation by PMA (50ng/ml) + A23187(µM) + βglucan (50µg/ml).

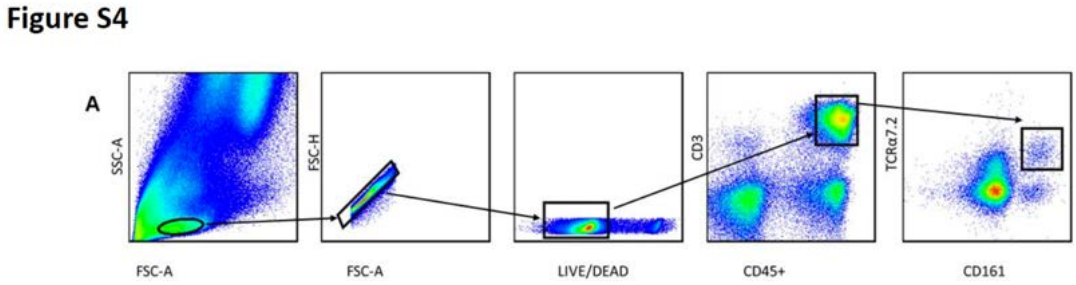


Figure S4: Flow cytometry gating strategy for phenotypic identification of MAIT cells in entheseal tissues and peripheral blood cells. Doublet excluded EMCs or PBMCs were stained with zombie aqua (live/dead discrimination), anti-CD45 (to exclude non-leucocytes), CD3 (T-cell inclusion), CD161 and Vα7.2 TCR.

Supplementary Materials

Patients characteristics

	AxSpA patients	Non-AxSpA patients with spinal surgery
Age (years)	38.5±14.79	42.6±27.4
Gender (M/F)	10/8	1/4
Disease duration (years)	9.26±10.43	NA
HLAB27 (%)	76,47%	NA
CRP (mg/l)	10.79±13.49	NA
BASDAI (0-10)	41.33±14.93	NA
BASFI (0-10)	28.8±19.9	NA
ASDAS-CRP	2.7±0.79	NA

Supplementary Table 1

RNA extraction (AxSpA patients)

Total RNA from sorted Neutrophils, CD4+, CD8+, MAIT, γδ T cells were isolated using RNeasy Micro Kit or Mini Kit (Qiagen RNeasy Micro Kit, Valencia, CA) following the protocol provided by the manufacturer. RNA concentration was estimated using Qubit RNA HS Assay Kit (Life Technologies, Grand Island, New York, USA) according to the manufacturer's instructions.

RNA quality was assessed on a selection of random samples across the cohorts using the RNA 6000 Nano Kit on an Agilent BioAnalyzer 2100 system (Agilent Technologies, Palo Alto, CA). The RNA integrity number (RIN) was determined using the LabChip System software. RNA aliquots were stored at -80°C until use.

Gene expression analysis (AxSpA patients)

Total RNAs were diluted with RNase-free water at 5 ng/µl into 0.2ml tubes of the 12-strip provided by NanoString. 25 or 50ng (5µl total volume) of total RNA from each sample were analyzed according to manufacturer's instructions.

Hybridization reactions were performed in 12-tube PCR strips. First 5µl of each sample were added. Next, a mix containing hybridization buffer and Reporter probes was added. Finally, the Capture probes were added and PCR strips were quickly transferred to a thermocycler set at 65°C. Samples were hybridized at 65°C for 22 hours. Cartridges were read on a nCounter Digital Analyzer at the highest resolution (555 fields-of-view (FOV) collected per flow cell) to yield a Reporter Code Count (RCC) data set.

Data QC and normalisation

We generated raw data (RCC files) using purified cell populations from 18 patients for various stimulation conditions, using the Nanostring® AS panel. Quality control and data normalization was carried out as per the established pipeline provided by Nanostring® using nSolver Analysis software (version 3.0) using three steps described here: (1.) positive control normalization, (2.) negative control normalization and (3.) internal reference (housekeeping) gene normalization. (1.) The positive control normalization was used to correct for technical variation by calculating the geometric mean of the positive probe counts of each sample and then by calculating the average geometric mean of the positive probes across all the samples. A scaling factor was calculated by dividing the average geometric mean by the geometric mean for each sample. The computed scaling factor was then multiplied to the corresponding positive and negative controls and the target gene counts to correct for technical variation. (2.) Negative control normalization, a background threshold was calculated by taking the mean of the negative controls + 2 standard deviations (SD) and then this value was subtracted from the gene counts. (3.) The variation in mRNA input was corrected using the same method as the positive control normalization and housekeeping normalization. Housekeeping genes were selected by comparing the most stable genes across the panel according to the standard of deviation and coefficient of variation across all the samples. Using this method, we selected the following housekeeping genes: GAPDH, PPIA and RASSF. Probes with low counts were determined with respect to the background level, defined as the mean of the negative controls + 2*Standard deviation.

Gene expression analysis on AS disease module

Genes associated with AS were assigned using public GWAS data and annotations from Immunobase consortium. We then assessed the expression levels of these genes in different public repositories for gene expression data (RNA-seq and microarray). The genes that showed low levels of expression were excluded from the panel. Following these steps we originally assigned 43 associated genes to AS, which was later modified based on the list of susceptibility loci summarized by Osgood & Knight, 2018, to give 45 associated genes in the panel. After generating the gene expression data from 5 cell types isolated from SpA patients, QC and normalization were applied using the standard Nanostring nSolver data analysis pipeline. We then determined genes that fall below the background in our data and excluded them from further analysis, resulting in 36 associated genes.

Supplementary Table 2

Complete list of TaqMan Assays used for analysis of gene expression on non-SpA patients.

Gene Symbol	Gene Name	Assay ID
AHR	aryl hydrocarbon receptor	HS00169233_m1
BMP2	bone morphogenetic protein 2	HS00154192_m1
CCR10	C-C motif chemokine receptor 10	HS00706455_s1
CCR6	C-C motif chemokine receptor 6	HS00171121_m1
CLEC7A	C-type lectin domain family 7 member A / dectin-1	Hs01902549_s1
CXCL10	C-X-C motif chemokine ligand 10	HS01124251_g1
CXCL11	C-X-C motif chemokine ligand 11	HS04187682_g1
CXCL12	C-X-C motif chemokine ligand 12	HS00171022_m1
CXCR3	C-X-C motif chemokine receptor 3	HS00171041_m1
KLF2	Krupple like factor 2	Hs00360439_g1
NR4A1	Nuclear Receptor Subfamily 4 Group A Member 1	Hs00374226_m1
T-Bet	T-bet	HS00203436_m1
FGF7	Fibroblast growth factor 7	Hs00940253_m1
FGF9	Fibroblast growth factor 9	Hs00181829_m1
FOXP3	Forkhead Box P3	Hs1085834_m1
HPRT1	hypoxanthine phosphoribosyl transferase 1	HS99999909_m1
IFNG	interferon gamma	HS00989291_m1
IL6	Interleukin 6	Hs00174131_m1
IL10	interleukin 10	HS00961622_m1
IL17A	interleukin 17A	HS00174383_m1
IL17F	interleukin 17F	HS00369400_m1
IL22	interleukin 22	HS01574152_g1
IL23R	interleukin 23 receptor	HS00332759_m1
IL36R	IL-36 receptor	Hs00543916_m1
JAK1	Janus kinase 1	HS01026983_m1
JAK2	Janus kinase 2	HS01078136_m1
JAK3	Janus kinase 3	HS00169663_m1
RORA	RAR related orphan receptor A	Hs00536545_m1
RORC	RAR related orphan receptor C	HS01076122_m1
SOCS1	suppressor of cytokine signalling 1	HS00705164_s1
SOCS2	suppressor of cytokine signalling 2	HS00919620_m1
SOCS5	suppressor of cytokine signalling 5	HS00367107_m1
STAT1	signal transducer and activator of transcription 1	HS01013996_m1
STAT2	signal transducer and activator of transcription 2	HS01013123_m1
STAT3	signal transducer and activator of transcription 3	HS00374280_m1

Single-molecule array (Simoa) IL-17A digital ELISA

IL-17A concentrations (in fg/ml) were determined with Simoa technology (Quanterix Simoa™ IL-17A Reagent Kit, Lexington, MA, USA) according to the manufacturer's instructions. The working dilutions were 1:4 for all samples, in working volumes of 80 µL. Each plate was run with a control sample to normalize the concentrations. Cytokine concentrations were interpolated from standard curves. The

lower limit of detection was 7 fg/mL.

Supplementary Table 3

List of antibodies used for flow cytometry (FC) and activation assays (AA) on non-SpA patients

Target	Colour	Clone	Manufacturer	Application
CD3	n/a	OKT3	Invitrogen	AA
CD3	BUV395	UCHT1	BD Biosciences	FC
CD45	RF710	HI30	Tonbo Biosciences	FC
CD161	BV510	HP-3G10	Biolegend	FC
TCR Vα7.2	BV785	3C10	Biolegend	FC
CD45RA	FITC	HI100	Biolegend	FC
CD69	APC	FN50	Biolegend	FC
IL17A	PE	eBio64DEC17	ThermoFisher	FC
TNFα	PE-CF594	MAb11	BD Biosciences	FC

Supplementary information (Manuscript II)

Immune response profiling of spondyloarthritis patients reveals signaling networks mediating TNF-blocker function *in vivo*

Silvia Menegatti[1,2,†], Vincent Guillemot[3], Eleonora Latis[1,2], Hanane Yahia-Cherbal[1,2], Daniela Mittermüller[1], Vincent Rouilly[4], Elena Mascia[1], Nicolas Rosine[1,2], Surya Koturan[1,2], Gael A. Millot[3], Claire Leloup[1], Darragh Duffy[5], Aude Gleizes[6,7], Salima Hacein-Bey-Abina[6,7], Milieu Intérieur Consortium[‡], Jérémie Sellam[6,7], Francis Berenbaum[6,7], Corinne Miceli-Richard[1,8,9], Maxime Dougados[8,9,10], Elisabetta Bianchi[1,9], Lars Rogge[1,9*]

Supplementary Methods

Patients

Peripheral blood samples were obtained from 80 consecutive patients with a definitive diagnosis of axial spondyloarthritis (axSpA) attending the Rheumatology Department of Cochin Hospital or the Rheumatology Department of Saint-Antoine Hospital (Paris, France). This study fulfills the current Good Clinical Practice Guidelines and a clinical protocol to analyze peripheral blood from SpA patients before and after therapy with TNF-blockers has been accepted by ethical committees (Comité de Protection des Personnes Ile de France III; Référence CPP: n° AT-100) and Institut Pasteur (Projet de recherché clinique n° 2011-32). The project has been approved by the “comité consultatif sur le traitement de l’information en matière de recherche dans le domaine de la santé (CCTIRS, Référence DGRI CCTIRS MG/CP°2012.035), as well as the “Commission Nationale de l’Information et des Libertés” (CNiL; Project “du genotype à la physiopathologie dans les spondylarthropathies, analyse de l’axe IL-23/Th17 chez les patients traités par un anti-TNF”; Décision DR-2013-080). A written informed consent, in compliance with the applicable regulatory and ethical requirements, has been obtained from each subject. All patients met assessment of spondyloarthritis international society (ASAS) criteria for axSpA.[1, 2] Blood was collected from each participant at days 0, 7 and/or 90 after initiation of anti-TNF therapy.

Inclusion criteria

- Patients aged over 18 and under 60 years
- Compliance with criteria established by the "Assessment of SpondyloArthritis international Society" (ASAS, http://www.asas-group.org/)

Exclusion criteria:

- Other spinal disease clearly defined (e.g. discarthrosis);
- History of any biotherapy;
- It is possible to include patients that have received corticosteroid treatment, with the condition that the therapy is stable for at least 4 weeks at the moment of inclusion, and with a dose inferior to 10 mg prednisone.
- patient with active IBD or ongoing uveitis
- patients with psoriatic involvement more than 10% of the skin surface.
- Pregnancy
- History or current disorders which might interfere with the validity of the informed consent and/or prevent an optimal compliance of the patient to the cohort (e.g. alcoholism, psychological disorders).
- No affiliation with a social security scheme
- Person deprived of liberty by judicial or administrative decision, person subjected to a legal protection measure

The first 12 patients were recruited and analyzed during 2015. Recruitment of the subsequent patients was between 2016 and 2018. Patients' demographics, HLA-B27 status, information regarding evaluation of symptoms (including duration of morning stiffness, pain or swelling in peripheral joints and back pain), ongoing treatments (e.g. analgesics, NSAIDs, DMARDs, physiotherapy), co-morbidities with a specific check-list including in particular cardiovascular and malignant diseases, and other main clinical features of spondyloarthritis (e.g. acute anterior uveitis, psoriasis, inflammatory bowel disease, enthesitis, peripheral articular involvement) were recorded on a Case Record Form before and 3 months after initiation of anti-TNF therapy (see **Table S1**). Axial, peripheral or enthesial presentation was clinically assessed.

The Ankylosing Spondylitis Disease Activity Score (ASDAS), the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI), erythrocyte sedimentation rate, C-reactive protein, cholesterol (HDL, LDL) and complete blood count were collected before and 3 months after initiation of anti-TNF therapy. C-Reactive Protein (CRP) levels were measured using the high-sensitivity test (hs-CRP test). Radiological evaluation (including plain X-rays and MRI of the spine and the pelvis) was collected systematically for each patient at baseline and at different times after the beginning of the biotherapy.

Definition of Disease Activity and Response to anti-TNF therapy

The criteria for determining disease activity and primary responsiveness to anti-TNF therapy based on the Ankylosing Spondylitis Disease Activity Score (ASDAS) have been described

previously.[3, 4] ASDAS-CRP was calculated at baseline (ASDAS D0) and 3 months after initiation of anti-TNF therapy (ASDAS D90). To assess the clinical response to anti-TNF therapy the "improvement score" (delta ASDAS = ASDAS D0 - ASDAS D90) was calculated. Delta ASDAS ≥2 defines a major improvement (responders), delta ASDAS ≥1.1 defines a clinically important improvement (partial responders) and patients achieving a delta ASDAS < 1.1 were classified as non-responders.[3, 4]

Whole-Blood TruCulture Stimulation

TruCulture tubes (Myriad RBM, Texas) are whole-blood stimulation systems consisting in syringe-based medical devices containing the indicated stimulus resuspended in 2 ml of buffered media.[5] Control tubes with no stimulants to assess background levels of genes and mediators of interest were included for each patient at each time point. TruCulture systems were manufactured in accordance with EN ISO 13485 (Medical Device Directive) standards, at EDI GmbH (Reutlingen, Germany), a subsidiary of Myriad RBM (Austin, TX, USA). All TruCulture tubes used in this study were prepared in the same batch, using the same lot of stimuli, and stored at -20°C until use. We performed whole blood stimulation experiments exactly as described previously.[5]

Multi-analyte Profiling

Supernatants from whole-blood stimulation systems were analyzed with Luminex xMAP technology by Myriad-RBM (Austin, TX, USA) as described.[5]

RNA Extraction

Total RNA was extracted from TruCulture cell pellets lysed in Trizol LS and stored at -80°C. Tubes containing cell lysate were thawed on ice 30 minutes before processing, vortexed twice for 5 min at 2000 rpm to complete thawing and RNA release and centrifuged (3000 x g for 5 min at 4°C) to pellet the cellular debris generated during the Trizol lysis. Total RNA was isolated according to a protocol provided by the supplier (Sigma-Aldrich).

RNA Quality Assessment

RNA concentration was estimated using Qubit RNA HS Assay Kit (Life Technologies, USA) according to the protocol provided by the manufacturer. RNA quality was assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies). The RNA Integrity Number (RIN) was determined using the LabChip System software and all samples with a RIN > 6 were processed for gene expression analysis.

Gene Expression Analysis with nCounter Technology

The nCounter system, a hybridization-based multiplexed assay, was used for the digital counting of transcripts using protocols provided by the supplier (NanoString). Briefly, 100 ng of total RNA from each sample was hybridized according to manufacturer's instructions with the Human Immunology v2 Gene Expression CodeSet, which contains 594 endogenous gene probes, 8 negative control probes (NEG A to NEG H) and 6 positive control probes (POS A to POS F) designed against six *in vitro* transcribed RNA targets at a range of concentrations (from

128fM to 0.125fM). Data collection was carried out in the nCounter Digital Analyzer at the highest standard data resolution (555 fields of view (FOV) collected per flow cell).

We used in total three different batches of the nCounter XT formulation. To correct for a potential batch effect, the expression level of 24 randomly selected RNA samples was measured with the three batches to calculate the calibration factor.

Quality Control of the NanoString Data

Each sample was analyzed in a multiplexed reaction including eight negative probes and six serial concentrations of positive control probes. Quality control consisted of checking the field of view counted (flag if < 0.75), binding density (flag if not in 0.05 − 2.75 range), linearity of positive controls (flag if R^2 < 0.9), and limit of detection for positive controls (flag if 0.5fM positive control < 2 standard deviation (SD) above the mean of the negative controls). Negative control analysis was performed to determine the background for each sample. Of note, we excluded three negative control probes (NEG B, NEG F, NEG H), for which we observed variable expression probably due to cross-reaction with bacterial nucleic acid present in two of the TruCulture stimulation systems (*S. aureus* and SEB). nSolver analysis software (version 3.0, NanoString) and R Software (version 3.3.3), NanoStringQCPro (version 1.12.0), NormqPCR (version 1.26.0) packages) were used for quality control and data normalization.

Normalization of the NanoString Data

A first step of normalization using the internal positive controls permitted correction of potential sources of variation associated with the technical platform (e.g. hybridization, purification, or binding efficiency). To do so, the geometric mean of the positive probe counts was calculated for each sample. The scaling factor for a sample was defined as: (average of all the sample geometric means) / (geometric mean of the considered sample). For each sample, we multiplied all gene counts by the corresponding scaling factor. Next, the background noise, defined as the mean + 2 SD across the five negative probe counts, was subtracted from each gene in a sample. Finally, to normalize for differences in RNA input we used the same method as in the positive control normalization, except that geometric means were calculated over three housekeeping genes (EEF1G, HPRT1 and TBP). These genes were selected using geNorm method [6], an established approach for identification of stable housekeeping genes, from the 15 candidate genes included in the CodeSet. The impact of anti-TNF treatment on the expression level of these housekeeping genes was also evaluated and none of them were affected by TNFi in patient samples.

Gene Filtering

The Human Immunology v2 gene CodeSet contained a total of 594 probes (15 correspond to housekeeping genes), of which 456 were included in downstream analysis after removing probes mapping to multiple genes or aligning to polymorphic regions with greater than two SNPs (9 probes) and probes with low counts (114 probes). Probes mapping to multiple locations and aligning to polymorphic regions with more than two SNPs were excluded from the analysis as described.[7]

We estimated the background level for each sample as the mean plus 2 standard deviations of the five negative probes counts, excluding NEG B, NEG F and NEG H for which we observed

significant differences in counts between conditions as previously explained. We defined as 30 counts the highest background level across all the genes in the different stimulations. In order to easily identify genes that were low in high proportions in a given condition, we calculated for each gene in each condition the percentage of samples with expression below the background (30 counts). We removed 114 genes which expression was below the background level in more than 80% of samples in one condition. A condition was considered a given stimulus at a given time point before or after anti-TNF treatment (D0, D7, D90).

Design of gene modules

We generated 45 gene modules by grouping genes included in the immunology_v2 panel according to the Molecular Signatures Database (MSigDB) annotation (http://software.broadinstitute.org/gsea/msigdb)[8] and manual curation from published literature (see **Table S4**). Each gene module contains a minimum of three genes, and the same gene can be included in different modules.

Quantitative set analysis of gene expression

We used quantitative set analysis of gene expression (QuSAGE) to identify differences in gene modules by quantifying gene-module activity using a probability density function.[9] The analysis was performed using R Bioconductor package v2.6.1. As compared to other gene set enrichment analysis methods, QuSAGE improves power by accounting for inter-gene correlations and quantifies gene-module activity with a complete probability density function (PDF). From this PDF, P values and confidence intervals can be easily extracted.

To generate heatmaps representing QuSAGE fold-enrichment of gene sets in the different stimulated cultures, only changes reaching a significance threshold of FDR ≤ 0.01 were represented. When this threshold was not reached for a given module in a specific culture, the value of 0 was assigned to the fold-change, to reflect no statistically significant change.

Venn diagram

The Venn diagram was generated using the web application jvenn (http://genoweb.toulouse.inra.fr:8091/app/example.html).

Purification of PBMCs and *in vitro* cell stimulation

Peripheral blood mononucleated cells (PBMCs) were isolated from fresh blood samples by gradient separation on Ficoll density gradient centrifugation (Lymphocyte separation medium, Eurobio, France) as described previously.[10] Monocytes were purified by magnetic cell sorting using anti-CD14 monoclonal antibody (mAb)-coated beads as recommended by the manufacturer (Miltenyi Biotec). The purity of monocytes was over 97% as verified by flow cytometry (LSR II, BD Biosciences). CD14+ cells were plated in 48-well plates at a final concentration of 1×10^6 PBMCs per ml and cultured for different times in pre-warmed Roswell Park Memorial Institute (RPMI) 1640 medium (Invitrogen) not supplemented with fetal calf serum, nor antibiotics. Untreated cells were immediately lysed in RLT buffer (Qiagen) with 1% β-mercaptoethanol to form the naïve subset and snap frozen for RNA extraction at a later date. All the rest of the monocytes were incubated or not with the soluble receptor etanercept (gift from Rheumatology Hardy B Unit of Cochin Hospital (Paris, France)) at a

concentration of 10 µg/ml for 10 minutes at 37°C [11] prior to the stimulation for various times with lipopolysaccharide (LPS, 20 ng/mL) from Escherichia coli (LPS, Invivogen). Cells were harvested after 15, 30, 60, 120 and 240 minutes of stimulation for analysis of mRNA expression. Cultured monocytes were lysed directly in RLT buffer (Qiagen) with 1% β-mercaptoethanol and homogenized by pipetting. mRNA was isolated using a RNeasy Micro kit (Qiagen) and analyzed with the nCounter Human Immunology v2 Gene Expression CodeSet.

Culture of Monocyte-Derived Macrophages

Monocytes were isolated from peripheral blood of six healthy donors using CD14 microbeads (Miltenyi Biotec) and cultured for 3 days in RPMI-Glutamax medium (Gibco) supplemented with antibiotics (penicillin and streptomycin) and 10% FCS in presence of 50 ng/ml M-CSF (Miltenyi Biotec). Monocyte-derived macrophages were subsequently cultured for three additional days in RPMI with M-CSF in presence or absence of etanercept or adalimumab (gifts from Rheumatology Hardy B Unit of Cochin Hospital (Paris, France)) at a concentration of 10 µg/ml, and then polarized for 24h towards the M1 subset with LPS (20 ng/mL, Invivogen) and IFN-γ (20 ng/ml, Milteny Biotec), or towards the M2 subset with IL-4 and IL-13 (both 20 ng/ml, Miltenyi Biotec). M1- and M2-macrophages were lysed in RLT buffer (Qiagen) with 1% β-mercaptoethanol and homogenized by pipetting. mRNA was isolated using a RNeasy Micro kit (Qiagen) and analyzed with the nCounter Human Immunology v2 Gene Expression CodeSet as described above.

Gene expression analysis for correlation to therapeutic responses

Using baseline clinical parameters (collected before the initiation of anti-TNF therapy), and baseline (D0) NanoString gene expression for LPS and SEB stimulations, differential gene expression analysis was performed to correlate therapeutic responses to TNFi in 80 axSpA patients, according to the delta ASDAS score.

Prior to the differential expression analysis, the NanoString gene expression dataset composed from LPS and SEB stimulations was filtered based on level of expression and pattern of expression. Lowly expressed genes were discarded when their normalized median count was below 30 counts in LPS and SEB stimulation conditions at D0 (R Software v3.3.3, dplyr v0.7.4).

We analyzed differential gene expression between the stimulation cultures from the 50 responders and 30 non-responders using the LIMMA package.[12] with an FDR correction for multiple testing. Age, sex, smoking history, B27 status, comorbidities and type of TNF inhibitor were included as covariates in the analysis. Genes were considered as differentially expressed when their adjusted p-values were lower than 0.05. The differentially expressed genes are reported in **Table 2** with their log Fold-Change, P-values and adjusted P-values.

Statistical analysis

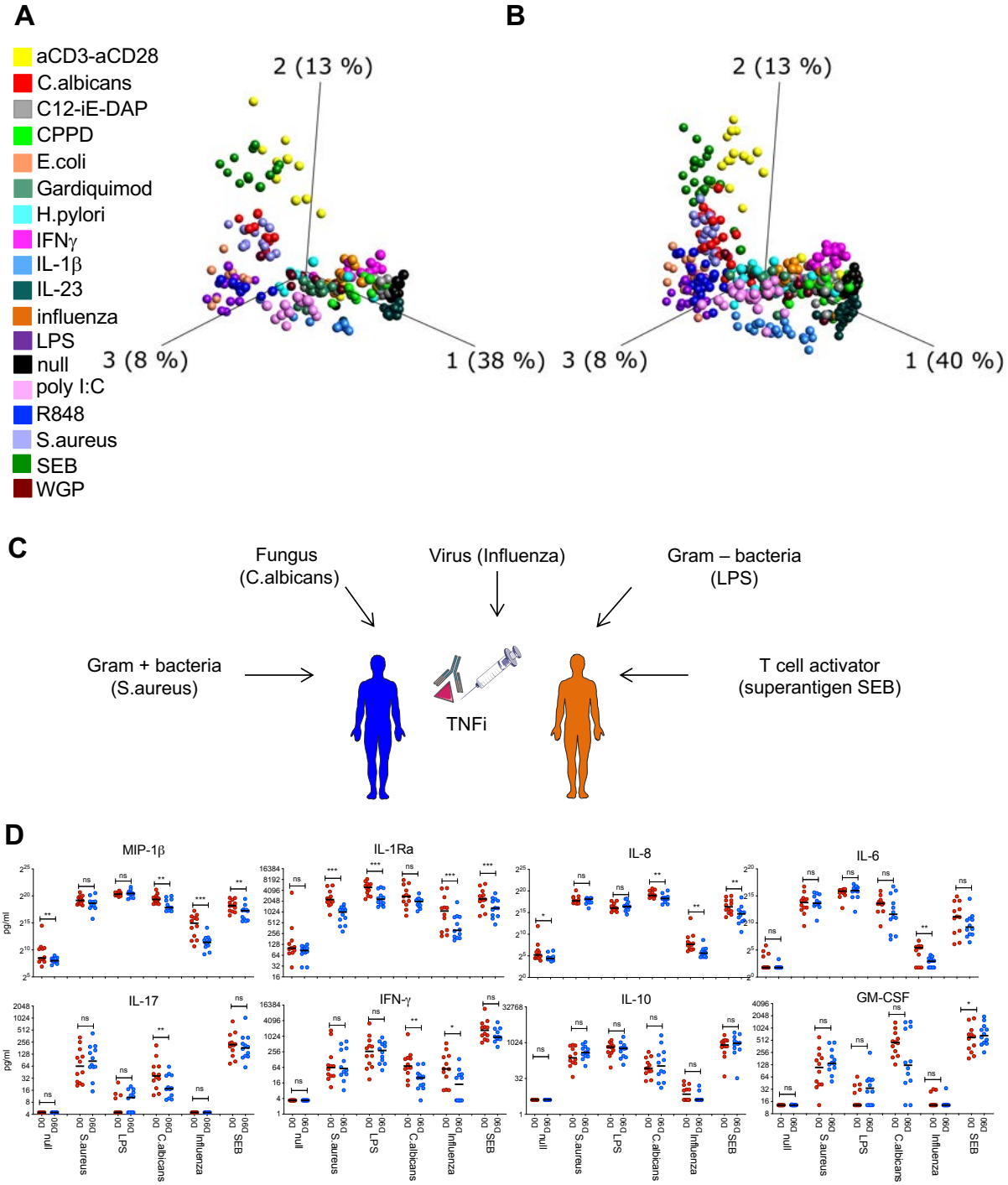
Unless otherwise indicated, horizontal bars represent the median. Statistical tests were two-sided and are specified in figure legends. Differences were considered to be significant when $P < 0.05$. Multiple testing corrections were applied where appropriate. Dot-plot graphs were compiled with GraphPad Prism v.7.0.

Principal component analysis (PCA) and agglomerative hierarchical clustering were performed with Qlucore Omics Explorer, version 3.6 (Qlucore). Before applying PCA and agglomerative hierarchical clustering, the variables (proteins or mRNA expression levels) were log-transformed, mean-centered per donor, and scaled to unit variance.

REFERENCES

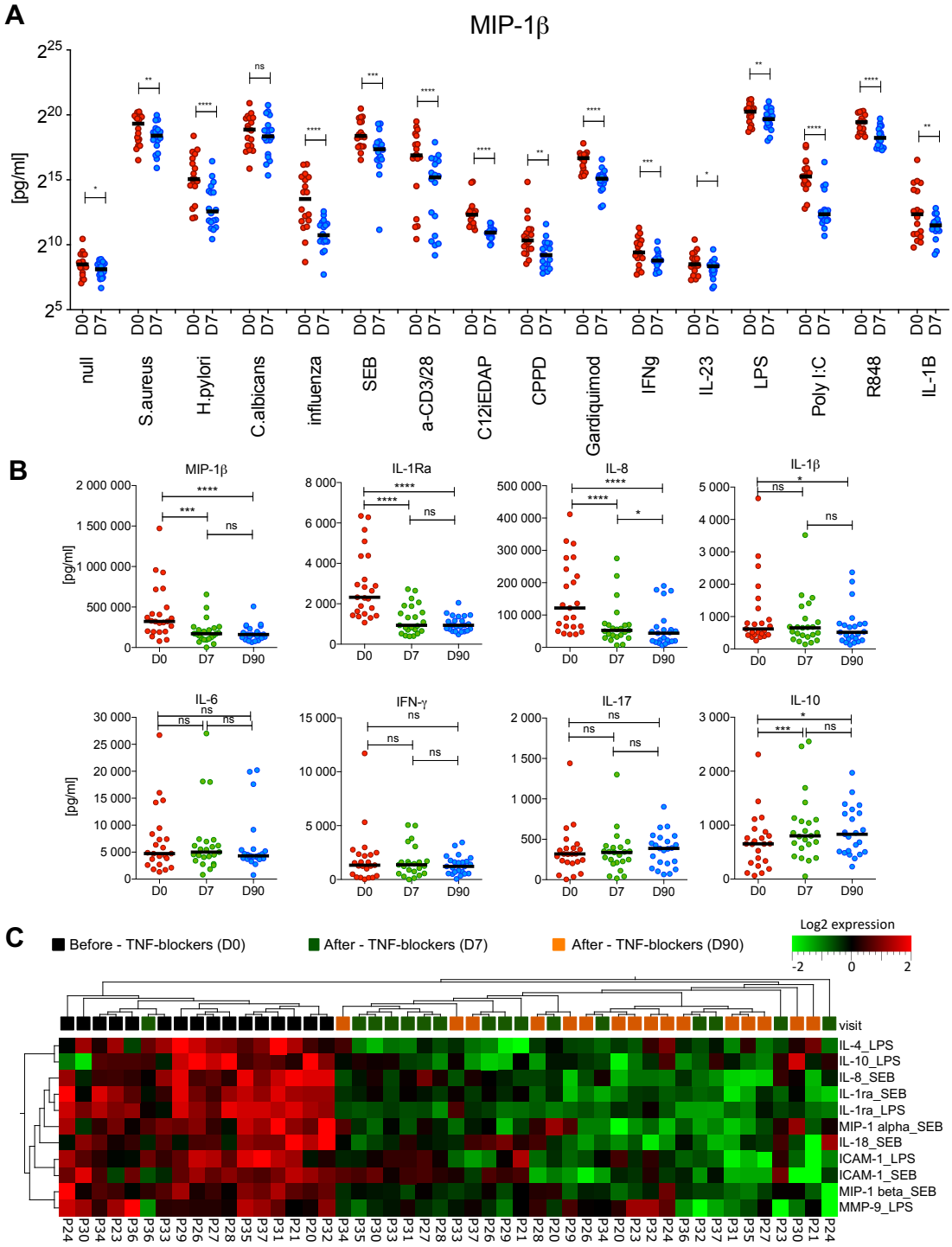
1. Rudwaleit M, Landewe R, van der Heijde D*, et al.* The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part I): classification of paper patients by expert opinion including uncertainty appraisal. Ann Rheum Dis. 2009 Jun; 68(6):770-776.
2. Rudwaleit M, van der Heijde D, Landewe R*, et al.* The Assessment of SpondyloArthritis International Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. Ann Rheum Dis. 2011 Jan; 70(1):25-31.
3. Machado P, Landewe R, Lie E*, et al.* Ankylosing Spondylitis Disease Activity Score (ASDAS): defining cut-off values for disease activity states and improvement scores. Ann Rheum Dis. 2011 Jan; 70(1):47-53.
4. Machado PM, Landewe R, Heijde DV*, et al.* Ankylosing Spondylitis Disease Activity Score (ASDAS): 2018 update of the nomenclature for disease activity states. Ann Rheum Dis. 2018 Oct; 77(10):1539-1540.
5. Duffy D, Rouilly V, Libri V*, et al.* Functional Analysis via Standardized Whole-Blood Stimulation Systems Defines the Boundaries of a Healthy Immune Response to Complex Stimuli. Immunity. 2014 Mar 20; 40(3):436-450.
6. Vandesompele J, De Preter K, Pattyn F*, et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. 2002 Jun 18; 3(7):RESEARCH0034.
7. Urrutia A, Duffy D, Rouilly V*, et al.* Standardized Whole-Blood Transcriptional Profiling Enables the Deconvolution of Complex Induced Immune Responses. Cell Rep. 2016 Sep 06; 16(10):2777-2791.
8. Liberzon A, Birger C, Thorvaldsdottir H*, et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. 2015 Dec 23; 1(6):417-425.
9. Yaari G, Bolen CR, Thakar J*, et al.* Quantitative set analysis for gene expression: a method to quantify gene set differential expression including gene-gene correlations. Nucleic Acids Res. 2013 Oct; 41(18):e170.
10. Coffre M, Roumier M, Rybczynska M*, et al.* Combinatorial control of Th17 and Th1 cell functions by genetic variations in genes associated with the interleukin-23 signaling pathway in spondyloarthritis. Arthritis Rheum. 2013 Jun; 65(6):1510-1521.
11. Mitoma H, Horiuchi T, Tsukamoto H*, et al.* Mechanisms for cytotoxic effects of anti-tumor necrosis factor agents on transmembrane tumor necrosis factor alpha-expressing cells: comparison among infliximab, etanercept, and adalimumab. Arthritis Rheum. 2008 May; 58(5):1248-1257.
12. Ritchie ME, Phipson B, Wu D*, et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015 Apr 20; 43(7):e47.
13. Covert MW, Leung TH, Gaston JE*, et al.* Achieving stability of lipopolysaccharide-induced NF-kappaB activation. Science. 2005 Sep 16; 309(5742):1854-1857.

Supplementary Figure 1



Supplementary Figure 1. Effects of different stimuli on protein signatures.
(**A**) Principal component analysis (PCA) was performed on the secreted protein data obtained from 12 patients before initiation of anti-TNF therapy (D0), measured in 18 different whole blood stimulations. Each filled circle represents a stimulated sample. Although the samples cluster by stimulation, some stimuli largely overlap, reflecting the activation of common signaling pathways. Values for each of the 31 analytes were centered to mean = zero and scaled to unit variance. (**B**) PCA was performed on the secreted protein data obtained from additional 17 patients at D0. The overall PCA structure of this cohort is similar to the one in (A). (**C**) Shown are the representative stimuli selected for further analysis of patient profiles before and after initiation of anti-TNF therapy: S. aureus (a gram-positive bacteria), C. albicans (a yeast), influenza virus, Lipopolysaccharide (LPS) and Staphylococcal enterotoxin B (SEB), a superantigen triggering T cell activation. (**D**) Plots (as in **Fig. 1**) indicate the levels of differentially secreted proteins for 6 representative stimuli and the unstimulated (null) condition, in 12 patients before (D0, in red) and 90 days after (D90, in blue) initiation of anti-TNF therapy (identified as described in **Fig. 1B**).

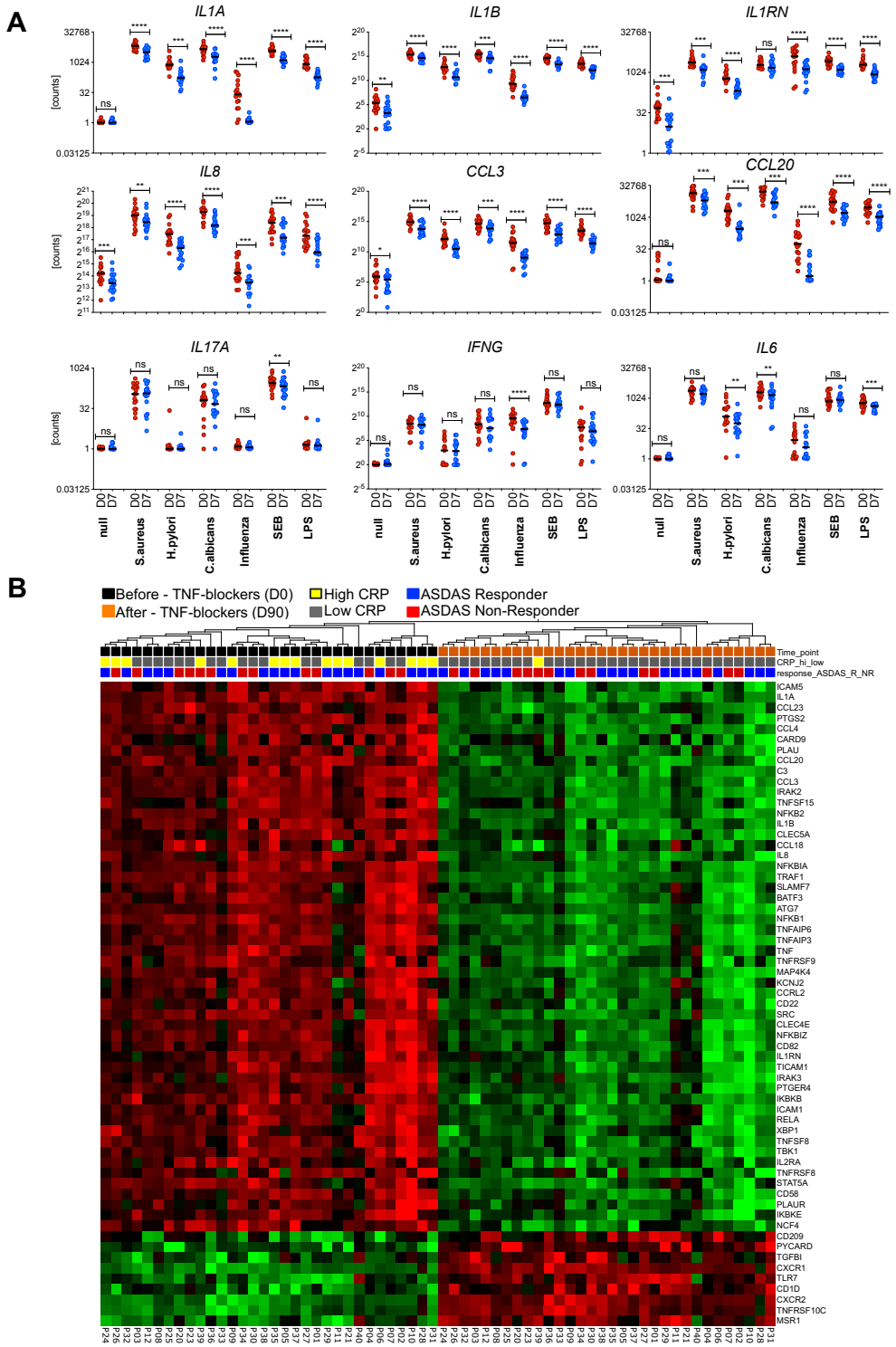
Supplementary Figure 2



Supplementary Figure 2. Effects of different stimuli on protein signatures before and at different time points after anti-TNF treatment.

(**A**) Quantification of MIP-1β in TruCulture assay supernatants from 17 patients, at D0 (red) and D7 days (blue). The stimuli present in the TruCulture assays are indicated below the x-axis. (**B**) Quantification of proteins in supernatants of TruCulture assays stimulated with SEB from patients at D0, D7 and D90 after initiation of anti-TNF therapy. Horizontal bars indicate the median. Significance was determined using a Wilcoxon matched-pairs test (SpA patients before *versus* after treatment) and *P*-values are indicated above the graph (*: *P*<0.05; **: *P*<0.01; ***: *P*<0.001; ****: P<0.0001; ns: not significant). (**C**) The levels of 31 secreted molecules in response to LPS and SEB were compared in samples from 17 patients at D0 (black rectangles), D7 (green rectangles) and D90 (orange rectangles). The heatmap shows the levels of differentially secreted proteins (paired t-test, FDR ≤ 0.01, red indicates higher and green lower levels of protein secretion).

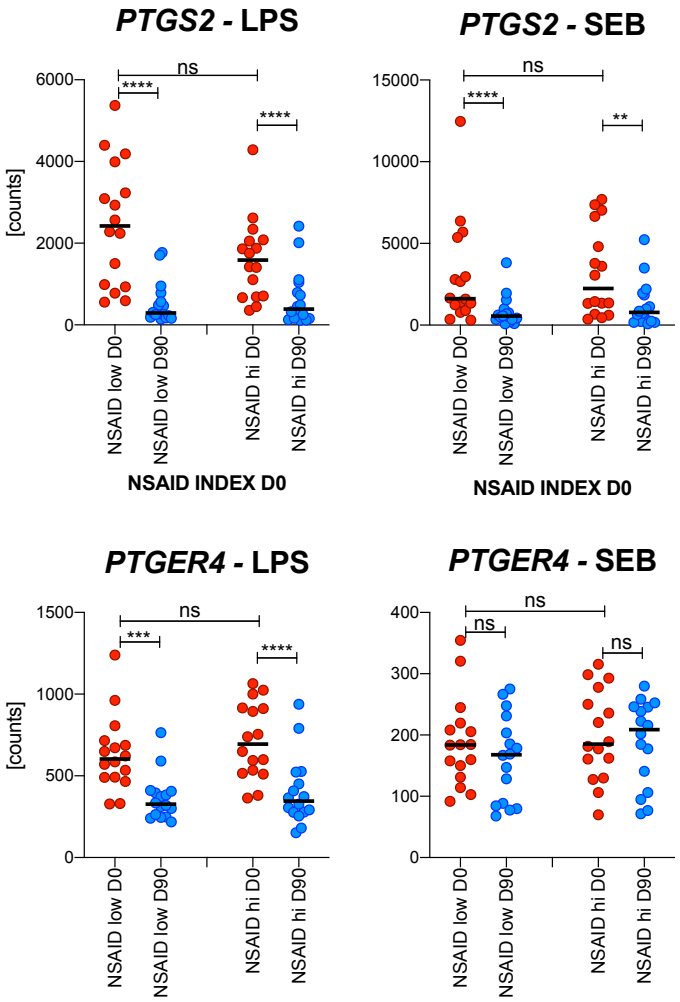
Supplementary Figure 3



Supplementary Figure 3. TNF-blockers strongly affect key regulators of innate immune responses.
(**A**) Plots indicate expression level of genes encoding molecules with pro-inflammatory properties and of *IL17A, IFNG* and *IL6* for the unstimulated condition and 5 representative stimuli, in samples before (D0, red) and 7 days after (D7, blue) initiation of anti-TNF therapy. Stimuli present in the TruCulture assays are indicated below the x-axis (n = 17, FDR ≤ 0.05, as in Fig. 2). (**B**) Heatmap of differentially expressed genes, comparing samples from 32 patients before (D0, black rectangles) and 90 days (D90, orange rectangles) after initiation of anti-TNF therapy. A paired t-test with false-discovery rate FDR ≤ 0.01 and a fold-change threshold of ≥ 2 identified 61 genes (ranked by decreasing fold-change). Red indicates high-level, and green low level of gene expression, respectively. Data are normalized and log2 transformed.

Supplementary Figure 4

A

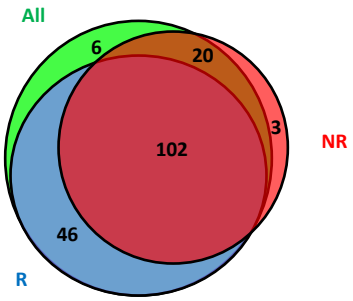


B

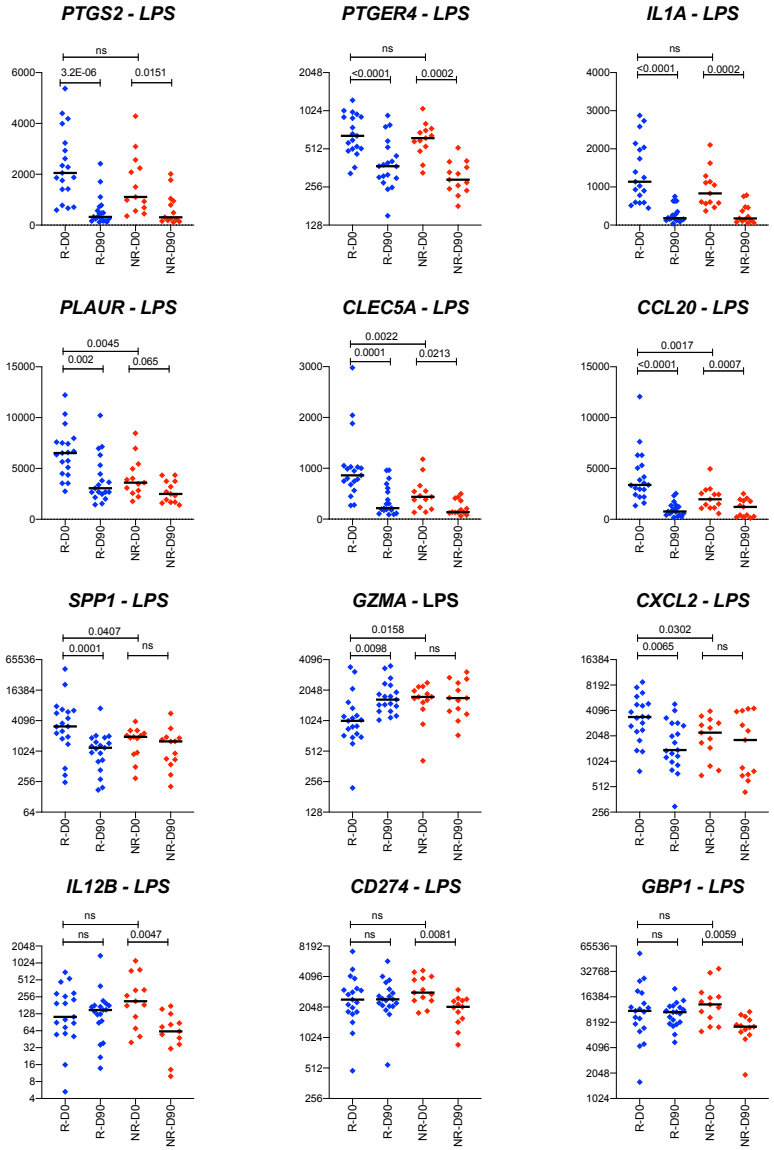
Patient_id	NSAID_D0	NSAID_D0_high/low
P01	11	low
P02	0	low
P03	150	hi
P04	84	hi
P05	177	hi
P06	89	hi
P07	100	hi
P08	46	low
P09	50	low
P10	50	low
P11	38	low
P12	50	low
P20	100	hi
P21	150	hi
P23	33,33	low
P24	0	low
P25	14,5	low
P26	83,33	hi
P27	50	low
P28	29	low
P29	100	hi
P30	11,11	low
P31	100	hi
P32	83,88	hi
P33	66,67	hi
P34	57	low
P35	100	hi
P36	6,33	low
P37	4,93	low
P38	100	hi
P39	110	hi
P40	76,34	hi

Supplementary Figure 4. The NSAID index was determined at baseline for the 32 patients for which gene expression data were available before (D0) and after (D90) TNFi treatment, and stratified patients according to the NSAID index (cut-off, median, **B**). *PTGS2* and *PTGER4* expression levels at D0 and D90 were plotted for the two groups of patients. Horizontal bars represent the median, and *P*-values are indicated above the graph (**: *P*<0.01; ***: *P*<0.001; ****: P<0.0001; ns: not significant).

Supplementary Figure 5



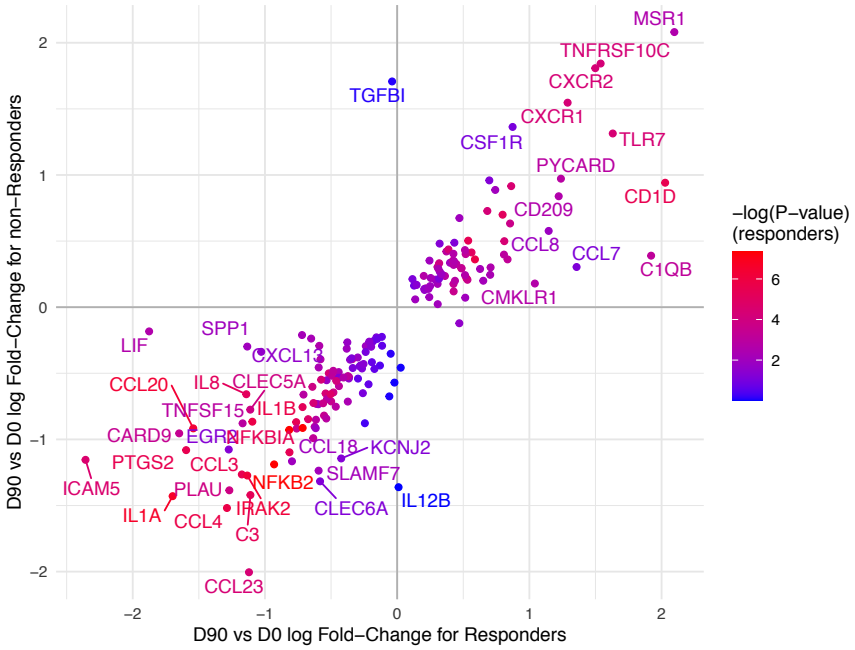
Supplementary figure 5. A. Gene expression data were analyzed in LPS-stimulated Truculture samples from 32 patients. Limma analysis was performed to compare gene expression at D0 versus D90 in 32 patients (all), or selectively in patients classified as Responders (R) or Non-responders (NR), according to ASDAS criteria. The Venn diagram shows the distribution of genes differentially expressed (adjusted p-value <0.05) in the indicated patient populations. The large majority (102) of differentially expressed genes was shared by all patient populations. Analysis of differentially expressed genes in NR patients alone identified 3 genes with significant changes between D0 and D90 specifically in these patients (*CD274*, *GBP1*, and *IL12B*).



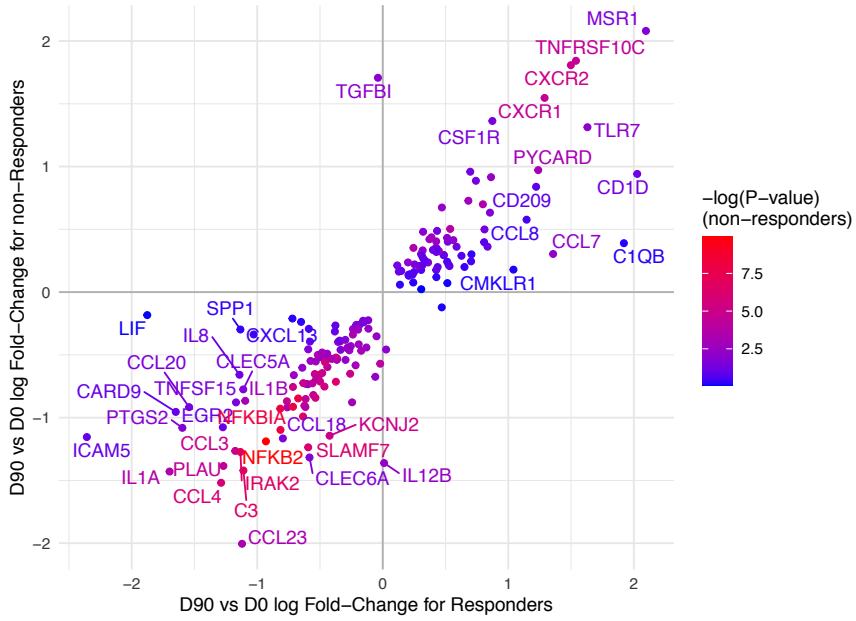
B. Selected genes from **A**. See also Supplementary Table 6.

Supplementary Figure 6

A



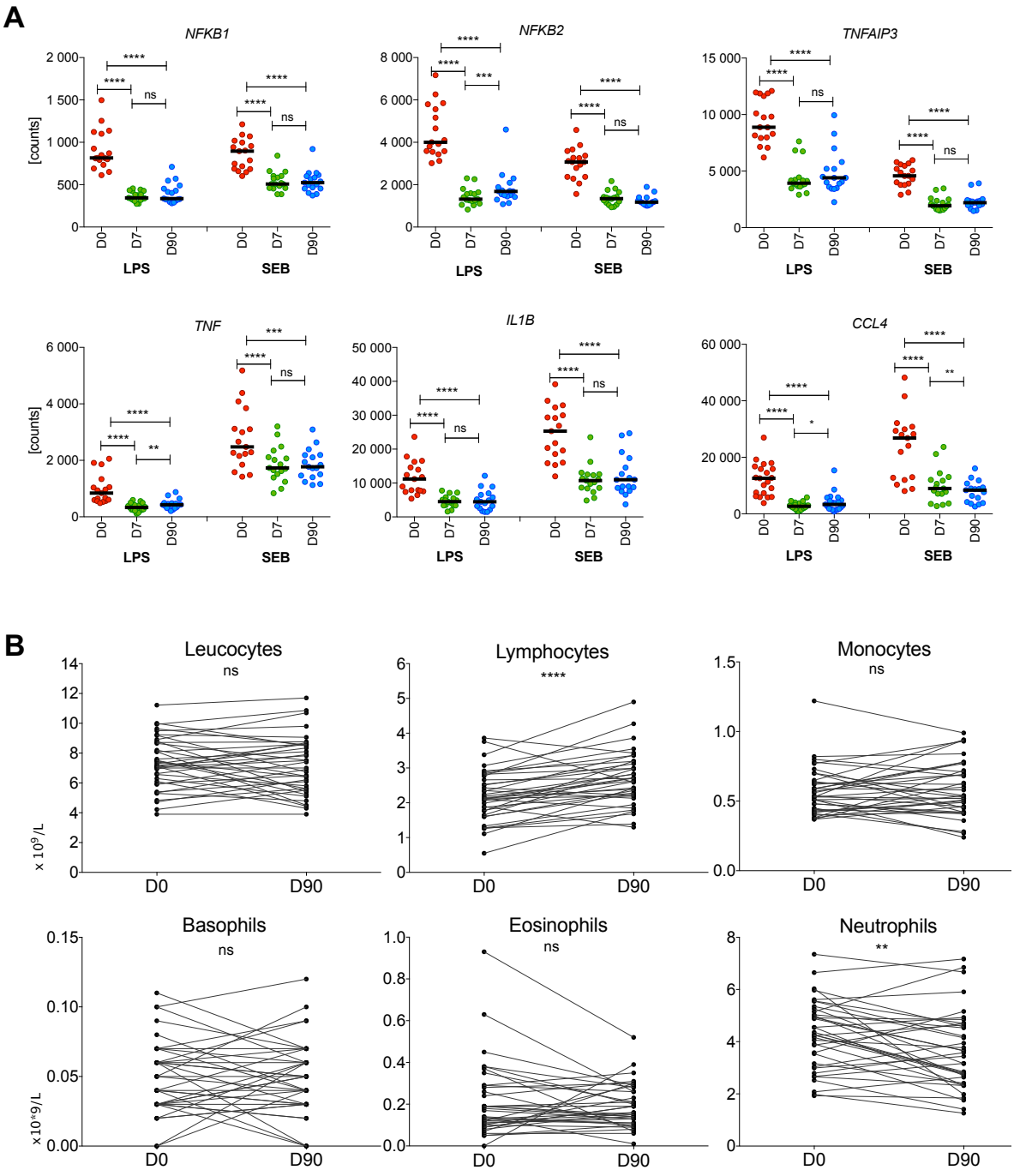
B



Supplementary figure 6. Differential gene expression between D0 (before treatment initiation) and D90 after treatment initiation.

Differential gene expression before and 90 days after TNFi treatment was calculated for Responders and Non-Responders (adjusted p-value <0.01, see **online supplementary table 6**), and Fold-changes of the differentially expressed genes were plotted for both populations. The labels identify the genes with log Fold-Change > 1 or < 1. The colors indicate the value of the adjusted p-value for each gene in Responders (**A**) and Non-responders (**B**)

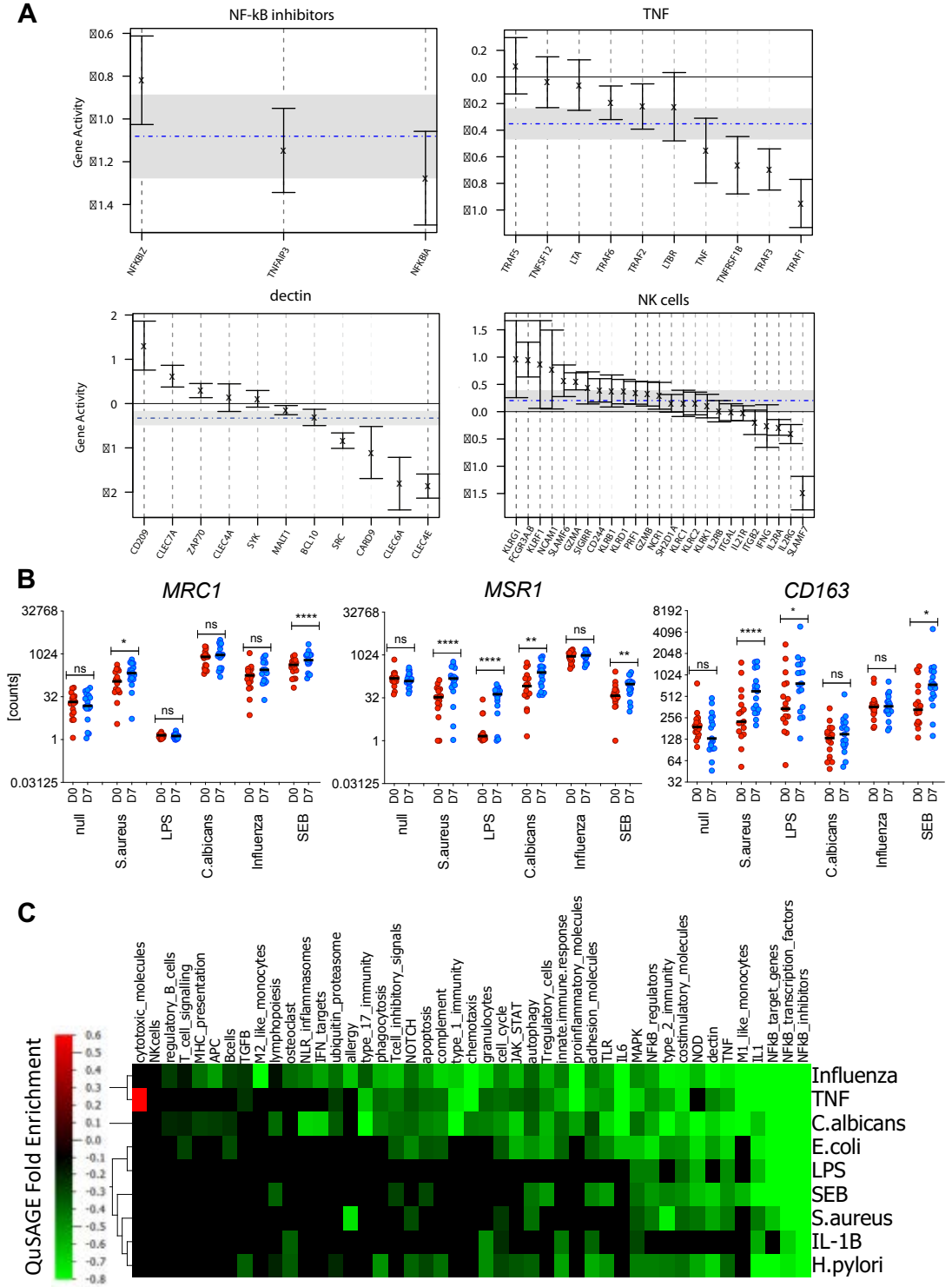
Supplementary Figure 7



Supplementary Figure 7. The effects of TNF-blockers on immune responses can be detected after a single injection and remain stable over time.

(**A**) Plots indicate gene expression levels of immune genes from stimulation cultures containing LPS or SEB performed before (D0, in red), 7 days (D7, in green) and 90 days (D90, in blue) after initiation of anti-TNF therapy (17 patients). (**B**) Complete blood cell counts (Coulter counter) in 37 axSpA patients at D0 and D90 after initiation of anti-TNF therapy. Significance was determined using a Wilcoxon matched-pair test (values before *versus* after treatment). *P*-values are indicated above the graph (*: P<0.05; **: P<0.01; ***: P<0.001; ****: P<0.0001; ns: not significant). We noted a modest decrease (1.23-fold) of neutrophil counts and a 1.24-fold increase of lymphocyte counts after TNF therapy.

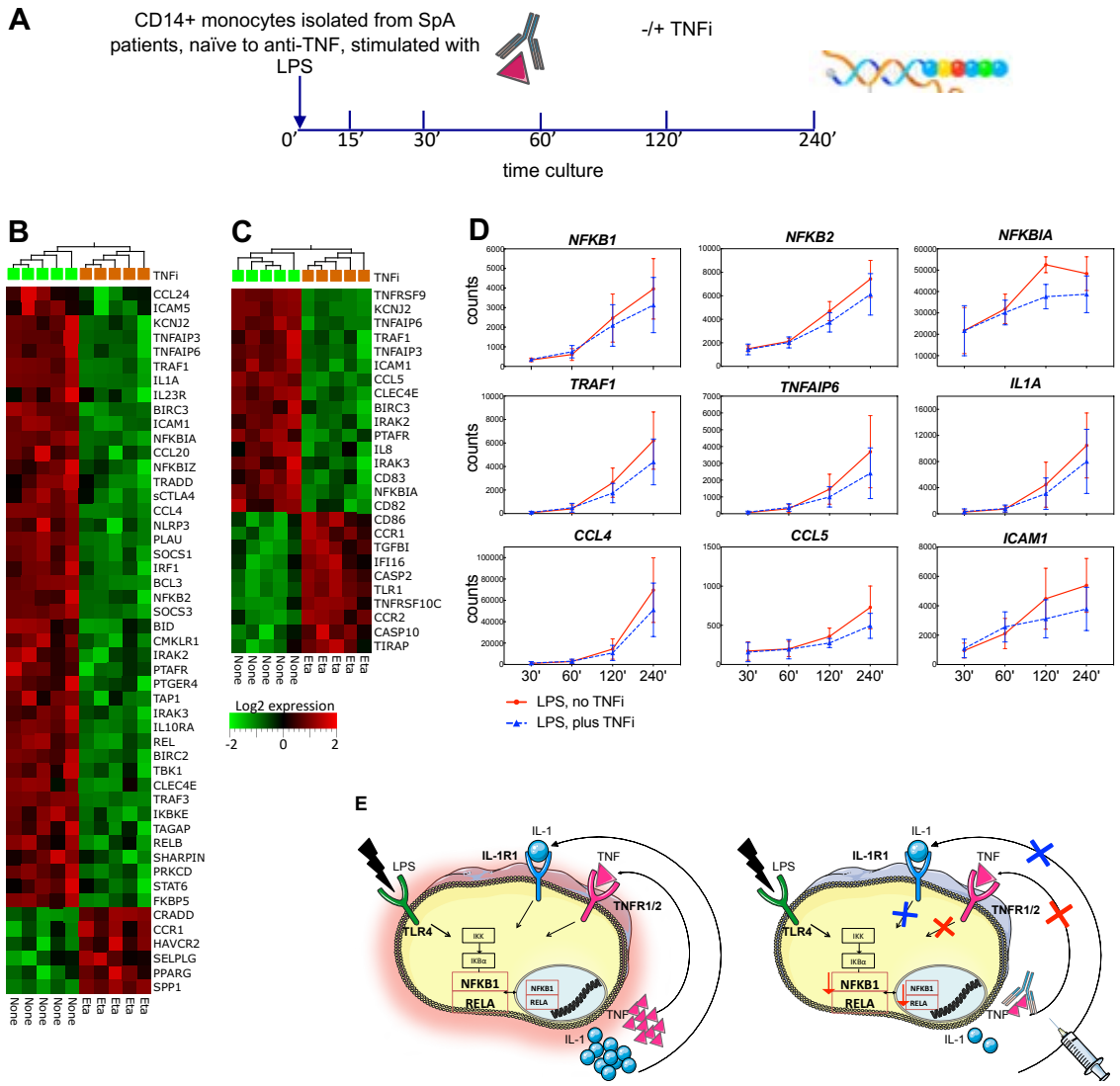
Supplementary Figure 8



Supplementary Figure 8. Modular transcriptional framework to assess the signaling pathways affected by TNF-blockers in stimulated immune cells.

(**A**) Fold changes in gene activity in modules before and 7 days after initiation of anti-TNF therapy (D7 *versus* D0) for SEB stimulated samples. Represented are the mean fold-change and 95% confidence interval for individual genes in each module. Gene activity = 0 signifies no change. The horizontal dashed blue line and the grey band indicate the mean differential expression of genes in the module at D7, compared to D0, and the 95% confidence interval, respectively. (**B**) Plots indicate expression levels of M2-like monocyte-related genes for the null and 5 representative stimuli in Truculture assays from 17 patients before (D0, in red) and 7 days (D7, in blue) after initiation of anti-TNF therapy. (**C**) Heatmap representing QuSAGE fold-enrichment of gene sets in 9 different stimulated cultures from 12 SpA patients, at D90 after initiation of anti-TNF therapy *versus* D0. For each module, the mean fold-change is represented and color-coded to indicate increased (red) or decreased (green) module activity.

Supplementary Figure 9

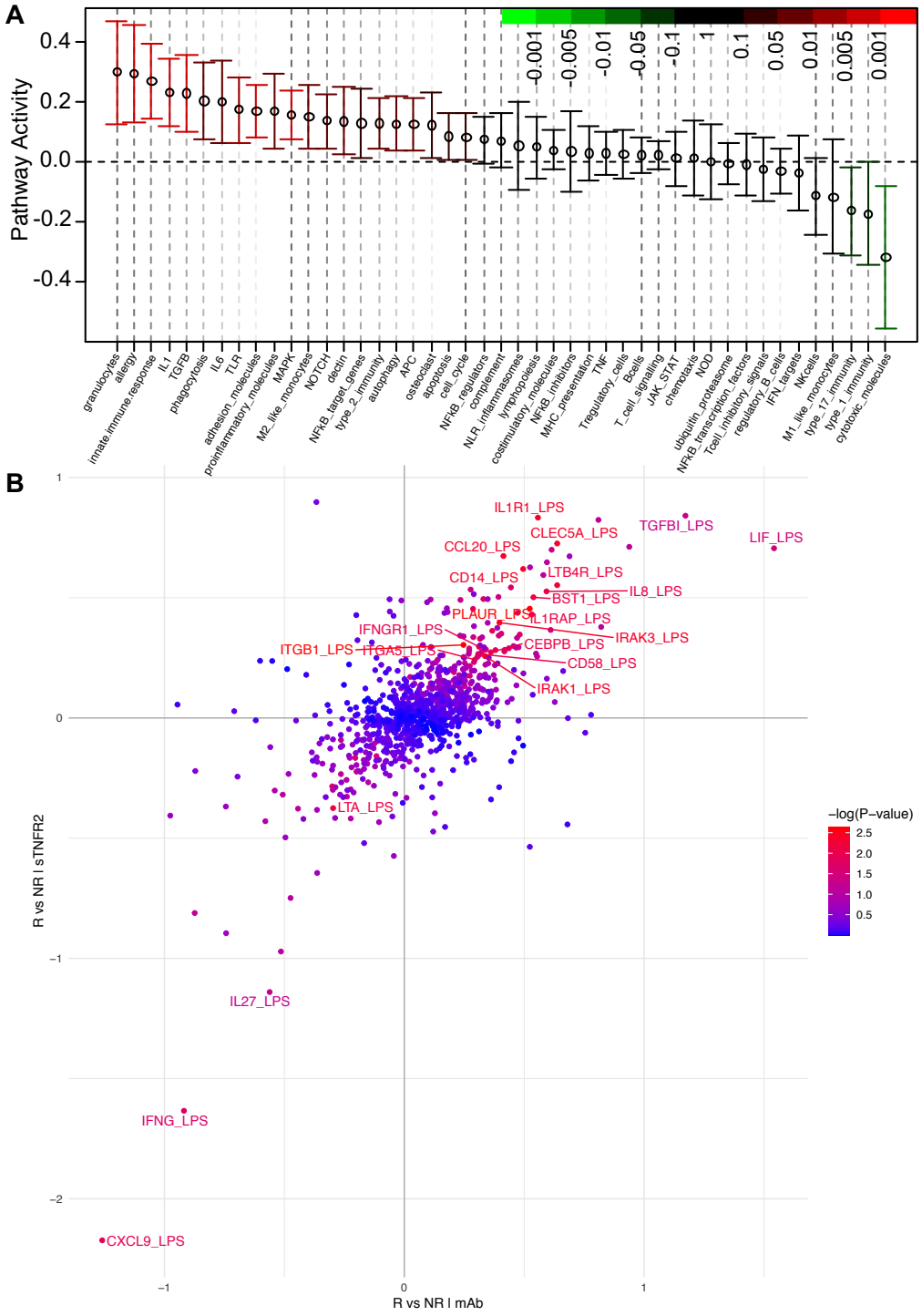


Supplementary Figure 9. TNF blockers break a TNF- and IL-1-dependent feed-forward loop of NF-κB activation in monocytes isolated from SpA patients

(**A**) Monocytes were isolated from 5 SpA patients and pre-incubated with or without TNFi (etanercept) for 10 minutes, prior to stimulation with LPS (20 ng/mL) for the indicated times. Gene expression was analyzed with the nCounter Human Immunology v2 panel. (**B, C**) Heatmaps show the top differentially expressed genes in monocytes in response to *in vitro* TNFi treatment after stimulation with LPS for 120 minutes (**B**) or 240 minutes (**C**). Orange and green rectangles distinguish samples pre-treated or not with TNFi, respectively. Gene expression analysis at the individual time points was performed using the Limma package with an adjusted *P*-value threshold of 0.1. (**D**) Expression kinetics of NF-kB target genes in LPS-stimulated monocytes cultured for the indicated times (minutes, horizontal axis). Monocytes were incubated with LPS (20 ng/mL) alone (red solid line), or pre-treated with TNFi for 10 minutes, followed by addition of LPS (blue dashed line). Shown are mean and standard deviation of 5 independent experiments. (**E**) Model for the intracellular mechanism of action of TNF-blockers.

Gene expression profiles of monocytes treated or not with Eta were strikingly different after 2 and 4 hours of LPS stimulation. A large proportion of the genes downregulated by TNFi at these time points were direct NF-kB target genes, such *NFKBIA*, *TNFAIP3*, *TNFAIP6*, or *IL1A*. The expression of NF-kB target genes in monocytes pre-treated with TNFi overlapped with untreated cultures during the first hour of stimulation, but diverged after 2 and 4 hours, compatibly with a positive feed-forward mechanism mediated by LPS-stimulated TNF production, which induces sustained activation of NF-kB and expression of its target genes, such as *IL1A* and *IL1B*, amplifying the inflammatory response.[13] Our data suggest that TNFi act by breaking the TNF- and the IL-1-dependent autocrine loops, dampening the activity of the NF-kB transcriptional cascade. Very similar results were obtained with monocytes isolated from 4 healthy donors, indicating that the action of TNFi on the NF-kB pathway is not dependent on the disease process (data not shown).

Supplementary Figure 10



Supplementary Figure 10. A. Modular transcriptional repertoire analysis reveals differential activity of signaling pathways in responders versus non-responders before treatment initiation (D0).
Differential activity of 45 gene modules (**online supplementary table 5**) generated from 456 immune-related genes (80 patients). Whole-blood cultures were stimulated with LPS. For each gene module, the mean activity fold-change and 95% confidence interval are plotted and color-coded according to their FDR-corrected P-values (means compared to fold-change zero). Confidence intervals overlapping the horizontal dotted line indicate statistically significant increased or decreased module activity comparing responders and non-responders. **B.** Patients were grouped based on the type of treatment (etanercept (sTNFR2) versus monoclonal antibodies (mAb), see **online supplementary table 1**) and differential gene expression between responders and non-responders was calculated for each group at D0 (adjusted p-value <0.05, **table 2**), and fold-changes of the differentially expressed genes were plotted for both groups. The labels identify the genes with differential expression at adj. p-value < 0.05.

Supplementary Table 1. Demographic and clinical characteristics and response to anti-TNF treatment of the 80 axSpA patients included in the study

Patient ID	Gender	Age	CRP M0	ASDAS M0	CRP M3	ASDAS M3	Reponse ASDAS	Smoke	B27	Psoriasis	Uveitis	IBD	Anti-TNF
1	M	31	5.40	1.99	2.00	0.98	NR	1	1	0	0	0	Eta
2	F	37	2.10	3.71	2.00	3.39	NR	1	1	0	1	0	Ada
3	M	19	2.00	2.28	2.00	2.08	NR	0	0	0	0	0	Eta
4	M	37	2.00	1.96	2.00	0.97	NR	1	1	1	0	0	Eta
5	M	24	47.00	4.79	2.00	1.27	R	1	1	0	0	0	Eta
6	M	53	7.00	2.64	2.00	1.26	PR	1	1	0	1	0	Eta
7	M	54	2.00	1.13	5.00	1.71	NR	0	1	0	0	0	Eta
8	M	58	5.00	2.50	2.00	1.23	PR	0	1	0	1	0	Eta
9	M	34	17.00	4.39	2.00	1.58	R	0	1	0	0	0	Eta
10	M	42	9.00	3.03	2.00	0.94	R	1	0	1	0	0	Eta
11	M	23	51.00	4.46	2.00	0.87	R	1	1	0	0	0	Eta
12	F	42	2.00	2.16	2.00	0.83	PR	1	1	0	0	0	Eta
13	M	26	0.09	3.87	1.20	2.61	PR	0	1	0	0	0	Eta
14	M	26	2.48	1.28	0.00	0.87	NR	0	0	0	0	0	Eta
15	F	40	10.73	3.35	0.00	1.10	R	0	1	0	1	0	Eta
16	M	24	1.72	2.35	0.00	1.09	PR	0	1	0	0	0	Eta
17	F	27	11.35	2.58	2.00	0.64	PR	0	1	0	0	0	Eta
18	M	47	5.41	3.49	4.10	1.69	PR	0	1	1	1	0	Eta
19	M	30	1.11	3.53	0.50	2.58	NR	1	0	0	0	0	Eta
20	F	39	1.23	2.96	0.00	2.09	NR	1	1	1	0	0	Eta
21	F	21	20.15	2.35	2.80	1.08	PR	0	1	0	0	0	Eta
23	M	20	0.53	2.02	4.00	1.38	NR	0	1	0	1	0	Eta
24	F	58	37.50	4.75	7.00	2.97	PR	1	0	0	0	0	Eta
25	M	36	0.28	1.72	0.30	0.64	PR	0	1	0	0	0	Eta
26	M	48	7.46	1.61	0.00	0.77	NR	0	0	0	0	0	Eta
27	M	33	1.24	2.30	1.30	3.29	NR	0	1	1	1	0	Ada
28	F	40	21.44	4.68	0.00	0.87	R	1	1	1	0	0	Eta
29	M	50	27.45	4.39	3.10	2.29	R	1	1	0	1	0	Gol
30	M	57	2.63	3.07	1.20	2.67	NR	1	1	0	1	0	Gol
31	M	51	33.06	4.73	0.00	3.01	PR	1	1	0	1	0	Gol
32	M	58	16.97	3.25	0.00	1.24	R	0	0	1	0	0	Eta
33	F	24	5.72	2.78	4.00	1.16	PR	0	0	0	0	0	Eta
34	M	56	3.27	3.56	0.00	2.93	NR	0	1	0	1	0	Eta
35	F	38	39.38	4.43	2.00	0.75	R	0	1	0	1	1	Ada
36	F	47	0.68	2.50	0.50	1.48	NR	0	0	0	0	0	Eta
37	M	37	14.27	3.74	0.40	2.15	PR	0	1	0	1	0	Gol
38	M	43	4.09	3.10	2.00	1.71	PR	1	1	1	0	0	Eta
39	F	34	8.31	2.30	6.40	2.27	NR	0	1	0	0	0	Eta
40	M	43	1.11	2.32	3.00	0.92	PR	0	1	0	0	0	Eta
41	M	41	6.39	2.43	2.00	1.13	PR	0	1	1	0	0	Eta
42	M	55	15.24	2.88	21.10	2.63	NR	0	1	0	1	0	Eta
43	M	43	27.20	3.63	1.10	1.91	PR	0	1	0	1	0	Ada
44	M	47	0.18	2.14	0.50	1.63	NR	1	1	0	0	0	Eta
45	M	24	0.50	2.62	0.00	2.76	NR	1	1	0	0	0	Gol
46	M	27	0.35	3.50	0.00	1.33	R	1	1	1	0	0	Gol
47	M	44	4.18	3.21	2.70	2.08	PR	1	0	0	0	0	Eta
48	M	27	0.82	2.55	0.90	1.70	NR	1	1	0	0	0	Eta
49	F	52	15.20	4.07	9.00	1.87	R	0	0	0	1	1	Ada
50	F	27	3.58	3.09	4.00	2.82	NR	0	1	0	0	0	Gol
51	F	32	1.87	3.69	0.00	3.34	NR	1	0	0	0	1	Ada
52	M	27	2.00	2.56	1.00	0.96	PR	0	1	0	1	0	Eta
53	M	42	0.82	2.12	0.00	1.56	NR	1	1	0	0	0	Eta
54	M	45	3.39	3.16	1.00	1.78	PR	1	0	0	0	0	Gol

Patient ID	Gender	Age	CRP M0	ASDAS M0	CRP M3	ASDAS M3	Reponse ASDAS	Smoke	B27	Psoriasis	Uveitis	IBD	Anti-TNF
55	M	58	16.65	3.45	7.00	2.47	NR	1	0	0	0	0	Eta
56	M	41	10.74	2.80	0.00	0.71	R	0	1	0	1	0	Gol
57	M	39	3.35	3.29	1.90	1.89	PR	1	1	0	0	0	Gol
58	M	46	21.24	3.96	1.00	1.06	R	1	0	1	1	0	Ada
59	M	27	12.21	2.17	3.40	0.86	PR	1	1	0	0	0	Gol
60	F	29	16.38	2.56	1.00	0.94	PR	1	1	0	0	0	Gol
61	M	20	17.48	3.99	0.80	0.84	R	0	1	0	0	0	Eta
62	F	23	2.55	1.88	0.00	1.15	NR	1	1	0	0	0	Eta
63	M	32	0.48	1.92	2.00	0.64	PR	0	1	1	1	0	Eta
64	M	43	10.64	4.12	2.70	1.92	R	1	1	0	0	0	Eta
65	M	39	17.70	3.13	0.30	1.27	PR	0	1	0	1	0	Ada
66	F	64	15.23	3.88	6.00	1.72	R	0	1	0	1	0	Eta
67	M	57	12.53	4.19	2.10	3.45	NR	1	1	1	0	0	Eta
68	M	22	26.76	3.87	1.90	1.13	R	1	1	0	0	0	Eta
69	M	36	9.70	3.55	1.30	1.59	R	0	1	0	1	0	Eta
70	F	31	1.40	2.70	1.00	2.21	NR	1	1	0	0	0	Eta
71	F	21	62.00	4.61	51.80	3.12	R	1	1	0	0	0	Eta
72	F	55	1.00	2.75	1.00	0.94	R	1	0	0	0	0	Eta
73	M	57	28.60	3.90	2.00	0.94	R	0	1	0	1	0	Ada
74	F	48	1.00	2.69	1.00	1.62	NR	1	1	1	1	0	Ada
75	F	33	1.00	2.61	1.00	2.73	NR	0	0	0	0	0	Ada
76	F	53	7.80	2.64	1.40	1.41	R	0	1	1	1	0	Ada
77	F	25	6.40	2.10	5.00	1.06	NR	1	1	0	0	0	Eta
78	M	30	1.90	1.54	0.60	0.64	NR	0	1	0	0	0	Eta
79	M	31	19.20	3.60	2.00	2.40	R	1	1	0	0	0	Inf
80	M	23	20.00	3.30	2.00	0.90	R	0	1	0	0	0	Gol
81	M	26	8.20	3.10	2.00	1.10	R	1	1	1	0	0	Ada

Supplementary Table 2. Drug dosage and anti-drug antibodies

Patient ID	anti-TNF	response ASDAS	Drug dosage (µg/ml)	Dosage ADAb (ng/ml)
P01	Etanercept	NR	2.8	<10
P02	Adalimumab	NR	2	<10
P03	Etanercept	NR	1.6	<10
P04	Etanercept	NR	1.2	<10
P05	Etanercept	R	<0.2	<10
P06	Etanercept	R	1.2	<10
P07	Etanercept	NR	3.2	<10
P08	Etanercept	R	1.6	<10
P09	Etanercept	R	2.8	<10
P10	Etanercept	R	2.8	<10
P11	Etanercept	R	<0.2	<10
P12	Etanercept	R	1.6	<10
P13	Etanercept	R	2.3	<10
P14	Etanercept	NR	0.7	<10
P15	Etanercept	R	2.6	<10
P16	Etanercept	R	3.1	<10
P18	Etanercept	R	0.3	<10
P19	Etanercept	R	3.7	<10
P21	Etanercept	NR	0.7	<10
P23	Etanercept	R	2.2	<10
P24	Etanercept	R	1.3	<10
P25	Etanercept	R	1.9	<10
P26	Etanercept	R	2.5	<10
P27	Adalimumab	NR	8.7	<10
P28	Etanercept	R	2.3	<10
P29	Golimumab	R	4.2	<2,5
P30	Golimumab	NR	3.6	<2,5
P31	Golimumab	NR	2.7	<2,5
P32	Etanercept	R	0.9	<10
P33	Etanercept	R	2.6	<10
P34	Etanercept	NR	4.2	<10
P35	Adalimumab	R	>20	<10
P36	Etanercept	R	3.2	<10
P37	Golimumab	NR	1	<2,5
P39	Etanercept	NR	3.8	<10
P40	Etanercept	R	1.7	<10
P41	Etanercept	R	3.3	<10
P42	Etanercept	NR	2.8	<10
P44	Etanercept	NR	1.6	<10
P45	Golimumab	NR	2	<2,5
P46	Golimumab	R	1.8	<2,5
P47	Etanercept	NR	>5	<10

Patient ID	anti-TNF	response ASDAS	Drug dosage (µg/ml)	Dosage ADAb (ng/ml)
P48	Etanercept	NR	2.3	<10
P49	Adalimumab	R	10.3	<10
P51	Golimumab	NR	2	<2,5
P53	Adalimumab	NR	10.2	<10
P54	Etanercept	R	>5	<10
P55	Etanercept	NR	2.1	<10
P56	Golimumab	R	1.9	<2,5
P57	Etanercept	NR	1.1	<10
P58	Golimumab	R	1.8	<2,5
P59	Golimumab	R	0.9	<2,5
P60	Adalimumab	R	13.8	<10
P61	Golimumab	R	<0,1	<2,5
P62	Golimumab	R	4.8	<2,5
P63	Etanercept	R	>5	<10
P64	Etanercept	R	1.8	<10
P65	Etanercept	R	>5	<10
P66	Etanercept	R	2.3	<10
P67	Adalimumab	NR	8.4	<10
P68	Etanercept	R	3.3	<10
P69	Etanercept	NR	<0,2	<10
P70	Etanercept	R	>5	<10

Supplementary Table 3. Innate and Adaptive Immune Stimuli included in TruCulture Assays

Stimulus	Concentration	Supplier	Sensor or Receptor
Null		NA	
C12-iE-DAP	4 µg / ml	Invivogen	NOD1
α-CD3 + α-CD28	0.4µg/ml + 0.33 µg/ml		TCR
CPPD	100 µg/ml	Invivogen	NLRP3 & TLR2
Gardiquimod	3 µM	Invivogen	TLR7
HK *C. albicans*	10^7 bacteria	Invivogen	complex
HK *E.coli 0111:B4*	10^7 bacteria	Invivogen	complex
HK *H. pylori*	10^7 bacteria	Invivogen	complex
HK *S. aureus*	10^7 bacteria	Invivogen	complex
IFNγ (Imukin)	1000 IU/mL	Boehringer Ingelheim	IFNγR
IL-1β	25 ng/ml	Peprotec	IL1R
IL-1β + TNFα	25 ng/ml + 10 ng/ml		IL1R + TNFR
IL-23	50 ng/ml	Miltenyi Biotech	IL23R
Influenza (live)	1:700	Charles Rivers	Complex
LPS-EB (hi)	10 ng/ml		TLR4
BCG (Immucyst)	$3 * 10^5$ bacteria	Sanofi Pasteur	complex
poly I:C	20 µg/ml	Invivogen	TLR3
R848	1 µM	Invivogen	TLR7 & TLR8
Enterotoxin SEB	0.4 µg/ml	Bernhard Nocht Institute	TCR
TNFα	10 ng/ml	Miltenyi Biotech	TNFR
WGP	40 µg/ml	Invivogen	Dectin-1
Zymosan	300 µg/mL	Sigma-Aldrich	TLR2

Abbreviations are as follows: HK, heat killed; IU, international units. The stimulation conditions used for the preparation of TruCulture tubes are listed, with the indicated dose and commercial supplier.

Supplementary Table 4. Analytes measured in the supernatants of TruCulture Assays with Luminex xMAP technology

Analytes	Abbreviation	Units	LDD	LLOQ
Brain-Derived Neurotrophic Factor	BDNF	pg/mL	18.0	56.0
Eotaxin-1	Eotaxin-1	pg/mL	99.0	117.0
Factor VII	Factor VII	pg/mL	3000.0	2400.0
Granulocyte-Macrophage Colony-Stimulating Factor	GM-CSF	pg/mL	15.0	26.0
Intercellular Adhesion Molecule 1	ICAM-1	pg/mL	4200.0	6200.0
Interferon gamma	IFN-gamma	pg/mL	6.3	6.8
Interleukin-1 alpha	IL-1 alpha	pg/mL	0.8	1.1
Interleukin-1 beta	IL-1 beta	pg/mL	2.8	8.5
Interleukin-1 receptor antagonist	IL-1ra	pg/mL	38.0	59.0
Interleukin-2	IL-2	pg/mL	49.0	55.0
Interleukin-3	IL-3	pg/mL	8.3	8.6
Interleukin-4	IL-4	pg/mL	29.0	43.0
Interleukin-5	IL-5	pg/mL	3.5	6.0
Interleukin-6	IL-6	pg/mL	5.4	6.8
Interleukin-7	IL-7	pg/mL	30.0	41.0
Interleukin-8	IL-8	pg/mL	3.9	6.1
Interleukin-10	IL-10	pg/mL	4.9	8.1
Interleukin-12 Subunit p40	IL-12p40	pg/mL	220.0	450.0
Interleukin-12 Subunit p70	IL-12p70	pg/mL	25.0	37.0
Interleukin-15	IL-15	pg/mL	670.0	1200.0
Interleukin-17	IL-17	pg/mL	2.9	8.9
Interleukin-18	IL-18	pg/mL	31.0	42.0
Interleukin-23	IL-23	pg/mL	1300.0	3200.0
Macrophage Inflammatory Protein-1 alpha	MIP-1 alpha	pg/mL	43.0	48.0
Macrophage Inflammatory Protein-1 beta	MIP-1 beta	pg/mL	56.0	59.0
Matrix Metalloproteinase-3	MMP-3	pg/mL	55.0	70.0
Matrix Metalloproteinase-9	MMP-9	pg/mL	41000.0	33000.0
Monocyte Chemotactic Protein 1	MCP-1	pg/mL	107.0	83.0
Stem Cell Factor	SCF	pg/mL	97.0	222.0
Tumor Necrosis Factor alpha	TNF-alpha	pg/mL	16.0	24.0
Tumor Necrosis Factor beta	TNF-beta	pg/mL	39.0	58.0
Vascular Endothelial Growth Factor	VEGF	pg/mL	16.0	42.0

* The least detectable dose (LDD) was determined as the mean + 3 standard deviations of 200 blank readings. Results below the LDD are more variable than results above the LDD.

† The LLOQ (Lower Limit of Quantitation) is the lowest concentration of an analyte in a sample that can be reliably detected and at which the total error meets CLIA requirements for laboratory accuracy. As the LLOQ and the LDD values are independent from each other, on occasion the LLOQ is lower than the LDD.

Supplementary Table 5. Gene modules used in QuSAGE analysis

Module	Genes
Adhesion molecules	*APP, CD164, CD2, CD36, CD44, CD58, CD6, CD9, CD97, CD99, CEACAM1, CTNNB1, CX3CR1, DPP4, FN1, ICAM1, ICAM2, ICAM3, ICAM4, ICAM5, ITGA4, ITGA5, ITGA6, ITGAE, ITGAL, ITGAM, ITGAX, ITGB1, ITGB2, LGALS3, PECAM1, PLAU, PLAUR, PTK2, S100A9, SELE, SELL, SELPLG, SPP1, SRC, TGFBI, TNFAIP6*
Allergy	*CCL18, CCL5, FCER1A, IL13RA1, LTB4R, LTB4R2*
APC (Antigen Presenting Cells)	*BATF3, CCR7, CD14, CD163, CD1D, CD209, CD80, CD83, CD86, CD8A, CX3CR1, CXCR4, ITGAL, ITGAM, ITGAX, PDCD1LG2*
Apoptosis	*APP, BAX, BCAP31, BCL10, BCL2, BCL2L11, BID, CASP1, CASP10, CASP2, CASP3, CASP8, CD2, CD27, CD44, CDKN1A, CLEC5A, CRADD, CSF2RB, CTSC, CTSS, FAS, GZMB, LEF1, LGALS3, LTBR, MCL1, PDCD2, PRF1, PTK2, RAF1, TNFRSF10C, TNFRSF8, TNFSF10, TNFSF12, TNFSF15, TP53*
Autophagy	*ABL1, ATG10, ATG12, ATG16L1, ATG5, ATG7, IFI16, PTPN22, S100A8, S100A9, TOLLIP, XBP1*
B-cells	*BCL6, BLNK, BST1, BST2, BTK, CD19, CD22, CD24, CD27, CD79A, CD79B, CD80, CD81, CD99, CR2, CXCL13, ENTPD1, IFITM1, IL4R, IL6R, IRF8, ITGA5, LEF1, LILRB3, MS4A1, PAX5, PRDM1, PRKCD, PTPN6, SYK, TNFRSF13C, TNFRSF8, TNFSF13B, TNFSF8, ZAP70*
Cell cycle	*ABL1, AHR, BAX, BCL2, BID, CCND3, CDKN1A, IKZF1, MAPK1, PML, PRKCD, PTK2, RARRES3, S100A8, S100A9, SRC*
Chemotaxis	*CCL13, CCL18, CCL19, CCL2, CCL20, CCL22, CCL23, CCL24, CCL3, CCL4, CCL5, CCL7, CCL8, CCR1, CCR2, CCR5, CCR6, CCR7, CCRL2, CD99, CX3CR1, CXCL1, CXCL10, CXCL11, CXCL13, CXCL9, CXCR1, CXCR2, CXCR3, CXCR4, CXCR6, IL16, IL8, LGALS3, PPBP*
Complement	*C1QB, C1QBP, C2, C3, CASP1, CASP10, CASP3, CCL5, CD36, CD40LG, CD46, CD59, CEBPB, CFB, CFD, CFP, CR1, CR2, CTSC, CXCL1, FCER1G, FYN, ITGAM, ITGAX, ITGB2, LTF, PLAUR, PRKCD, PSMB9, RAF1, SERPING1, SRC, TNFAIP3*
Costimulatory molecules	*ADA, CD27, CD28, CD40, CD40LG, CD48, CD6, CD79B, CD80, CD82, CD86, CLEC5A, DPP4, ICOS, ICOSLG, MBP, PDCD1LG2, TAGAP, TNFRSF4, TNFRSF8, TNFRSF9, TNFSF12, TNFSF15, TNFSF4, TNFSF8, TRAF1*
Cytotoxic molecules	*GNLY, GZMA, GZMB, GZMK, IFNG, KLRD1, KLRF1, PRF1*
Dectin	*BCL10, CARD9, CD209, CLEC4A, CLEC4E, CLEC6A, CLEC7A, MALT1, SRC, SYK, ZAP70*
Granulocytes	*CCRL2, CD164, CD24, CD44, CLEC5A, CSF2, CSF3R, CXCL1, CXCR1, CXCR2, FCGR1A.B, FCGR3A.B, IL3, IL8, ITGAL, ITGAM, ITGAX, ITGB2, LTB4R, LTB4R2, LTF, MME, NCF4, SELL*
IFN targets	*BST2, CXCL10, IFI35, IFIH1, IFIT2, IFITM1, IFNA1.13, IFNAR1, IFNAR2, IRF1, IRF3, IRF4, IRF5, IRF7, IRF8, JAK1, MX1, PSMB8, TMEM173, TYK2*
IL1	*EGR1, IL18, IL18R1, IL18RAP, IL1A, IL1B, IL1R1, IL1R2, IL1RAP, IL1RL1, IL1RN, IRAK1, IRAK2, IRAK3, IRAK4, MYD88, SIGIRR, TOLLIP, TRAF6*
IL6	*IL6, IL6R, IL6ST*
Innate immune response	*ABL1, APP, BCL10, C1QBP, CD14, CLEC5A, CLEC7A, FCER1G, IKBKG, IL1RAP, IRAK1, IRAK4, LY96, NLRP3, S100A8, S100A9, TLR2, TLR4, TOLLIP*
JAK_STAT	*CISH, JAK1, JAK2, JAK3, PTPN2, PTPN6, PTPRC_all, SOCS1, SOCS3, STAT1, STAT2, STAT3, STAT4, STAT5A, STAT5B, STAT6, TYK2*
Lymphopoiesis	*CXCR4, IKZF1, IKZF2, IKZF3, NT5E, PAX5, RUNX1*
M1-like monocytes	*CCL19, CCL20, CCL5, CCL8, CCR7, CD80, CD86, CXCL10, CXCL11, CXCL9, IDO1, IFNGR1, IL12B, IL1R1, IL23A, IL2RA, MARCO, PTGS2, SOCS3*

Module	Genes
M2-like monocytes	*CCL13, CCL18, CCL2, CCL22, CCL24, CD163, CD209, CD36, CLEC7A, EGR2, FCER1A, FN1, IL10, IL1R2, IL1RAP, IL1RN, IL21R, IL4R, IRF4, MRC1, MSR1*
MAPK	*CD83, DUSP4, MAP4K1, MAP4K2, MAP4K4, MAPK1, MAPK14, MAPKAPK2, RAF1*
MHC presentation	*B2M, BCAP31, CD74, CTSS, HLA.A, HLA.B, HLA.C, HLA.DMA, HLA.DMB, HLA.DOB, HLA.DPA1, HLA.DPB1, KLRC1, KLRC2, KLRC3, KLRC4, KLRD1, KLRF1, KLRG1, KLRK1, LAMP3, LILRA1, LILRA2, LILRA3, LILRA6, LILRB1, LILRB2, LILRB4, MR1, MS4A1, NCF4, TAP1, TAP2, TAPBP, TNFSF4, XBP1*
NFkB inhibitors	*NFKBIA, NFKBIZ, TNFAIP3*
NFkB regulators	*BCL10, BTK, CHUK, IKBKAP, IKBKB, IKBKE, IKBKG, MALT1, MAP4K4, TBK1, TRAF4*
NFkB target genes	*BCL2, BCL3, CCL13, CCL19, CCL4, CXCL2, CYBB, ICAM1, IL1B, IL8, NFKBIA, PLAU, PTGS2, TNF, TNFAIP3, TNFSF13B, TRAF1, TRAF2*
NFkB transcription factors	*NFKB1, NFKB2, RELA, RELB*
NK-cells	*CD244, FCGR3A.B, GZMA, GZMB, IFNG, IL21R, IL2RA, IL2RB, IL2RG, ITGAL, ITGB2, KLRB1, KLRC1, KLRC2, KLRD1, KLRF1, KLRG1, KLRK1, NCAM1, NCR1, PRF1, SH2D1A, SIGIRR, SLAMF6, SLAMF7*
NLR_inflammasomes	*BCL2, CASP1, GBP5, NLRP3, PYCARD*
NOD	*CARD9, NOD1, NOD2, TRAF4, TRAF6*
NOTCH	*APP, IL2RA, NCR1, NFIL3, NOTCH1, NOTCH2, TGFB1, TGFBR2*
Osteoclast	*CEBPB, CSF1, CSF1R, CTNNB1, GPR183, LILRA1, LILRA2, LILRA3, LILRA5, LILRA6, MAPK14, NFATC1, SYK, TFRC, TRAF6*
Phagocytosis	*CYBB, ETS1, FCER1A, FCER1G, FCGR1A.B, FCGR2A, FCGR2A.C, FCGR2B, FCGR3A.B, FCGRT, ICAM3, ICAM5, IRF8, ITGAL, ITGAM, ITGAX, ITGB2, MARCO, PECAM1, SLAMF1*
Proinflammatory molecules	*CCL13, CCL18, CCL19, CCL2, CCL20, CCL22, CCL23, CCL24, CCL3, CCL4, CCL5, CCL7, CCL8, CCR1, CCR2, CCR5, CD163, CMKLR1, CSF1, CSF1R, CSF2, CXCL1, CXCL2, CXCR1, CXCR2, CXCR4, IL1B, IL32, IL6, IL6R, IL6ST, IL8, LILRA5, LITAF, MIF, PTAFR, PTGER4, PTGS2, S100A8, S100A9, TNF*
Regulatory B-cells	*CD19, CD1D, CD24, CD27, CD40, CD5, CD80, CD86, ICOSLG, IL10, PAX5, TFRC, TGFB1, TNFRSF13C*
T-cell signaling	*CD247, CD28, CD3D, CD3E, CD4, CD45R0, CD45RA, CD45RB, CD7, CD8A, CD8B, FYN, IL2RA, IL2RB, IL2RG, LCK, LCP2, NFATC1, NFATC2, NFATC3, PTPN22, PTPRC_all, ZAP70*
T-cell inhibitory signals	*BTLA, CAMP, CD244, CD274, CD276, CD5, CD96, CTLA4_all, CTLA4.TM, HAVCR2, IDO1, LAG3, PDCD1LG2, sCTLA4, TIGIT, TNFRSF14*
TGFB	*MAPK1, SKI, SMAD3, SMAD5, TGFB1, TGFBI, TGFBR1, TGFBR2*
TLR	*BCL10, CD14, IRAK1, IRAK2, IRAK4, LY96, MALT1, MYD88, TBK1, TICAM1, TIRAP, TLR1, TLR2, TLR3, TLR4, TLR7, TLR8, TOLLIP*
TNF	*LTA, LTBR, TNF, TNFRSF1B, TNFSF12, TRAF1, TRAF2, TRAF3, TRAF5, TRAF6*
T-regulatory cells	*CTLA4_all, CTLA4.TM, EGR2, ENTPD1, FOXP3, IL10, IL2, IL2RA, IL2RB, IL2RG, LAG3, LGALS3, NT5E, RUNX1, sCTLA4, STAT5A, STAT5B, TGFB1*
Type 1 immunity	*BATF3, CSF2, CXCR3, EBI3, GZMB, IFNG, IFNGR1, IL12B, IL12RB1, IL27, PRF1, STAT1, STAT4, TBX21, TNF*
Type 17 immunity	*AHR, BATF, CCR6, IL12B, IL17A, IL17F, IL21, IL22, IL23A, IRF4, KLRB1, MAF, STAT3, ZBTB16*
Type 2 immunity	*CCL18, CEBPB, CXCR4, CXCR6, IL13, IL1RL1, IL4R, STAT6*
Ubiquitin / proteasome	*CUL9, PSMB10, PSMB5, PSMB7, PSMB8, PSMB9, PSMC2, PSMD7, UBE2L3*

Supplementary Table 6. Differential gene expression between D0 (before treatment initiation) and D90 after treatment initiation*

	ALL patients (n=32)		RESPONDERS n=(19)		NON RESPONDERS (n=13)	
Gene ID	**logFC**	**adj.P.Val**	**logFC**	**adj.P.Val**	**logFC**	**adj.P.Val**
ABL1	0.253176758	0.001442703	0.30586301	0.006636767	0.176173775	0.222400494
ARHGDIB	0.21558608	3.61186E-06	0.201240695	0.001991642	0.236552413	0.00577187
ATG7	-0.806553689	6.3366E-11	-0.76352244	4.30165E-06	-0.869445515	1.23085E-05
B2M	-0.186180517	0.003718453	-0.11403	0.232445	-0.291628825	0.004017264
BATF	-0.418270605	3.22532E-05	-0.25939167	0.086164035	-0.650478282	1.23085E-05
BATF3	-0.824332688	4.53682E-08	-0.76012543	0.000211741	-0.918174066	0.000377461
BCL10	-0.260307478	4.88768E-05	-0.23718	0.010894	-0.294103926	0.003440946
BCL2L11	0.310465201	1.08621E-05	0.361489029	0.001221817	0.235891913	0.016977221
BID	-0.386253923	0.003427715	-0.335769	0.07856512	-0.460039575	0.029254642
BLNK	0.642740956	0.000113894	0.83513124	0.000559507	0.361555156	0.088286622
BTK	0.407795145	1.10241E-06	0.485467354	4.54742E-05	0.294274224	0.025728486
C1QB	1.298878127	0.002256578	1.920997082	0.000728495	0.389627346	0.67250275
C1QBP	0.175416058	0.009720333	0.243298384	0.009080724	0.076203	0.597687
C3	-1.236195878	1.73288E-12	-1.10986335	4.86561E-06	-1.420835733	2.24648E-07
CARD9	-1.366533982	3.75512E-05	-1.64858428	0.000384184	-0.954306617	0.094026687
CASP1	-0.256446002	0.006838877	-0.13	0.354265	-0.441256469	0.002979257
CASP2	0.187295816	1.20777E-05	0.21786434	0.00065451	0.142619	0.070906
CASP8	0.153024118	0.00628384	0.14024682	0.09774861	0.171698632	0.060545374
CCL18	-0.945888265	0.0003865	-0.79553	0.010676	-1.16564	0.030486
CCL20	-1.287913287	4.62521E-08	-1.5420472	4.76266E-07	-0.916486796	0.03208227
CCL22	0.64122106	0.000282579	0.80783953	0.001002966	0.397701757	0.20070615
CCL23	-1.479951808	1.03024E-07	-1.12084465	4.69475E-05	-2.004800729	0.000650271
CCL3	-1.21153499	2.8261E-11	-1.17490667	3.20446E-06	-1.265068689	1.88176E-05
CCL4	-1.381514235	1.53306E-12	-1.28750358	1.79524E-06	-1.518914417	1.86125E-06
CCL7	1.397525731	0.006838877	1.357208422	0.067879816	0.303937794	0.005981057
CCL8	0.91478911	0.006054825	1.146072573	0.013102692	0.576759433	0.342077344
CCND3	0.30486017	5.86801E-07	0.305491269	0.000384184	0.303938	0.005981
CCR2	0.542001344	0.006267441	0.706916022	0.003047309	0.300972	0.496146
CCR6	0.373827331	0.00052979	0.399839161	0.009836448	0.33581	0.065575
CCRL2	-0.710506217	1.34062E-07	-0.61429394	0.001127077	-0.851124167	4.95602E-05
CD19	0.408952327	6.90133E-05	0.486926825	0.000446304	0.294989599	0.098863245
CD1D	1.58606564	1.25323E-07	2.027168669	3.20446E-06	0.941377	0.050676
CD209	1.066351628	0.00053409	1.221933753	0.002604643	0.838962368	0.145168481
CD22	-0.689190454	4.52499E-06	-0.70825783	0.00127915	-0.661322747	0.003015926
CD274	-0.169834846	0.229203138	0.027022	0.915602	-0.457548233	0.008113729
CD3E	0.142228304	0.008427814	0.127954284	0.127351976	0.16309	0.05572
CD4	0.175907644	0.007214366	0.207115	0.020234	0.130296392	0.315832453
CD44	-0.430761404	8.68282E-08	-0.36339167	0.000947654	-0.529224861	5.38549E-05
CD48	-0.15982423	0.006749758	-0.11564771	0.196980747	-0.22439	0.015094
CD53	-0.213420521	0.005641292	-0.18267	0.112217	-0.258361946	0.026996209
CD58	-0.510167608	4.77139E-08	-0.51696989	4.9506E-05	-0.500225805	0.000528138
CD74	0.255761012	0.000311337	0.288231669	0.013884888	0.208304	0.01278
CD79A	0.502219684	3.89983E-08	0.562565512	8.20946E-06	0.414021936	0.006403097
CD79B	0.470882264	1.88402E-05	0.518391159	0.000578627	0.401446188	0.041019776
CD82	-0.660170925	1.48723E-06	-0.61122447	0.002614556	-0.731708058	5.67222E-05
CD83	-0.436272449	9.93638E-05	-0.36562	0.024918	-0.53953624	0.00283801

	ALL patients (n=32)		RESPONDERS n=(19)		NON RESPONDERS (n=13)	
Gene ID	**logFC**	**adj.P.Val**	**logFC**	**adj.P.Val**	**logFC**	**adj.P.Val**
CD86	0.386163476	0.001303576	0.321470565	0.053924554	0.480714654	0.011091724
CDKN1A	-0.46189628	1.87844E-05	-0.44215002	0.004284074	-0.490756204	0.003520628
CFB	-0.50162	0.016864	-0.2442	0.489735	-0.877853933	0.001255433
CFP	-0.395356865	0.00853923	-0.29629492	0.226126483	-0.540139709	0.004705888
CLEC4E	-0.661224	1.60969E-08	-0.55272992	0.000372592	-0.819792266	2.88963E-05
CLEC5A	-0.974540425	8.55814E-06	-1.11103278	0.000132995	-0.775051593	0.021319142
CLEC6A	-0.88133631	0.001315286	-0.58318589	0.059775076	-1.317094621	0.018794245
CMKLR1	0.691054017	0.004847332	1.041467919	0.002100337	0.178911	0.712636
CR1	0.700674902	8.68282E-08	0.682214063	5.50216E-05	0.727656129	0.001976613
CSF1R	1.072382648	0.007543374	0.873350027	0.089829004	1.36327648	0.0376294
CSF3R	0.884218586	9.30584E-08	0.862728133	3.61882E-05	0.91562771	0.003520628
CTLA4.TM	0.30033701	0.002287422	0.322429	0.013816	0.268049	0.14634
CTSS	-0.291748424	0.00187452	-0.17162511	0.232253804	-0.467313265	0.000650271
CXCL13	-0.747634709	0.009915684	-1.02821	0.024732	-0.33756904	0.395990086
CXCL2	-0.512839732	0.012003273	-0.71951007	0.006462195	-0.21078	0.63822
CXCR1	1.393945583	1.35617E-09	1.289375477	5.13172E-05	1.546778815	3.05185E-05
CXCR2	1.624132901	4.78263E-10	1.498542502	4.9506E-05	1.8076881	1.01218E-05
CXCR3	0.455116599	0.003934521	0.432833349	0.062474661	0.487684	0.057755
EGR2	-1.193637301	0.005169237	-1.27371	0.051147	-1.07661219	0.078384248
FCGR3A.B	0.50565039	0.000473349	0.470280905	0.024887141	-0.12161	0.741475
GBP1	-0.30919	0.061949	-0.05915	0.856059	-0.674629045	0.00594838
GFI1	-0.331428841	0.000364426	-0.29799237	0.016222066	-0.380297523	0.01880011
GZMA	0.335012971	0.019117139	0.51511183	0.009836448	0.071792	0.797502
HLA.DMA	0.398497551	1.97889E-06	0.447338798	0.000804489	0.327114189	0.003135577
HLA.DMB	0.400221403	3.35841E-07	0.524745205	9.40411E-06	0.218225078	0.062579352
HLA.DPA1	0.399941187	2.11781E-05	0.531855391	0.0002424	0.207144	0.108904
HLA.DPB1	0.324440692	3.22532E-05	0.426154747	0.000442848	0.175781689	0.106433258
ICAM1	-0.585892693	3.82126E-09	-0.49916988	0.000283245	-0.712641417	1.23085E-05
ICAM5	-1.868511919	3.36776E-06	-2.35700327	2.63903E-05	-1.15456	0.097668
ICOSLG	-0.350217647	0.001818817	-0.28376	0.090759	-0.447349566	0.005241155
IFITM1	0.340797741	0.000117897	0.43726364	0.001416312	0.199809119	0.10777694
IFNA1.13	-0.482674487	0.008038195	-0.65052326	0.005755428	-0.23736	0.545243
IKBKAP	0.275415395	0.009915684	0.316681853	0.06165645	0.215102878	0.095971671
IKBKB	-0.586494261	4.05119E-10	-0.54132876	4.23135E-05	-0.652505384	2.07119E-05
IKBKE	-0.5041171	6.78676E-08	-0.4404755	0.000846567	-0.59713175	8.36705E-05
IKZF3	0.178708824	0.007872859	0.206137	0.020362	0.138621	0.260855
IL12B	-0.54674	0.08931	0.010747	0.984681	-1.361529376	0.004705888
IL12RB1	0.468047093	0.003956999	0.65134258	0.007385616	0.200153689	0.440823919
IL1A	-1.588912275	2.15941E-11	-1.69809729	2.14274E-07	-1.429334177	0.000314071
IL1B	-1.002162478	9.87292E-10	-1.0953298	1.79524E-06	-0.865994858	0.001269259
IL1RN	-0.651814707	2.97169E-05	-0.59560873	0.008921978	-0.733961901	0.002979257
IL21R	0.403332947	0.000582278	0.513279894	0.002100337	0.242641	0.227161
IL2RA	-0.541302821	8.09206E-05	-0.53361	0.013885	-0.552547674	0.001989851
IL2RG	-0.361122445	6.54302E-08	-0.33909897	0.00065451	-0.393310596	6.46097E-05
IL8	-0.945016376	1.10241E-06	-1.14106032	7.03263E-06	-0.658490605	0.06012345
IRAK2	-1.190504303	1.57408E-13	-1.13387517	7.19758E-07	-1.27326996	2.24648E-07
IRAK3	-0.62522298	2.44206E-06	-0.64086859	0.000155825	-0.602356317	0.005979485
IRF3	0.800892843	0.000443594	0.742095693	0.009345021	0.886827	0.05073

	ALL patients (n=32)		RESPONDERS n=(19)		NON RESPONDERS (n=13)	
Gene ID	**logFC**	**adj.P.Val**	**logFC**	**Gene ID**	**logFC**	**adj.P.Val**
IRF5	-0.365246966	0.001919254	-0.21617117	0.214886147	-0.583126974	0.001194602
IRF8	0.302224424	0.000790595	0.427388287	0.000278178	0.119292624	0.543391483
ITGA4	0.197001693	0.002680241	0.237105427	0.009836448	0.138389	0.273789
ITGA6	0.479899036	1.01785E-05	0.512806788	0.001481542	0.431803091	0.011464671
KCNJ2	-0.716060168	7.33496E-06	-0.42315	0.054942	-1.144152681	1.23085E-05
KLRB1	0.383235294	0.000150528	0.426704773	0.001982921	0.319703	0.078384
LCK	0.155076478	0.002051829	0.115774681	0.077399912	0.212517566	0.043546384
LGALS3	-0.346184251	0.000426199	-0.26579	0.069423	-0.463680833	0.002189262
LIF	-1.188262593	0.010734849	-1.87595349	0.002606424	-0.18318	0.851711
LILRA5	0.764340852	3.09327E-06	0.854163228	0.000207665	0.633061994	0.02163132
LITAF	-0.282569407	0.003504842	-0.1912827	0.195135199	-0.415988437	0.002785657
LY96	0.394479046	1.01785E-05	0.422931162	0.000680084	0.352895	0.019343
MAF	0.803491056	0.002838368	0.697089022	0.057124202	0.959001721	0.047282386
MAP4K1	0.323772798	7.34957E-07	0.317921328	0.000139679	0.332324946	0.008199687
MAP4K4	-0.731686012	2.88379E-10	-0.71535371	6.31502E-06	-0.755556305	1.23085E-05
MAPKAPK2	-0.23123666	0.000507941	-0.21108877	0.030291873	-0.260683578	0.017436201
MR1	-0.278917486	0.001462725	-0.23801	0.097427	-0.338701136	0.000505089
MS4A1	0.49613521	4.08127E-07	0.588380908	1.54772E-06	0.361314574	0.060322602
MSR1	2.090184322	5.50291E-05	2.096634806	0.002799193	2.080756691	0.021319142
NCF4	-0.502544188	1.7403E-06	-0.4912643	0.000637919	-0.519030178	0.002249831
NFATC3	0.105235214	0.004604782	0.136827971	0.009403209	0.059061	0.496146
NFIL3	-0.372637632	0.000643062	-0.35581634	0.026157748	-0.397222603	0.008199687
NFKB1	-0.79545124	3.85948E-17	-0.71464798	4.52343E-08	-0.913548312	1.15864E-08
NFKB2	-1.035590339	1.0278E-17	-0.93058387	4.52343E-08	-1.189061328	1.18783E-10
NFKBIA	-0.929375472	1.81968E-13	-0.8143065	4.30165E-06	-1.097553208	1.15864E-08
NFKBIZ	-0.660373378	1.64257E-08	-0.53605548	0.00127915	-0.842068761	3.84915E-06
NLRP3	-0.468266914	0.002203625	-0.58871229	0.002656799	-0.292231363	0.285936157
NOD2	-0.361212226	0.006838877	-0.34381	0.065951	-0.38664	0.070625
PAX5	0.684106563	3.00373E-06	0.810584093	4.23135E-05	0.499254788	0.04754084
PDCD2	0.396623252	0.000309816	0.430426776	0.002606424	0.347218	0.098863
PECAM1	0.553546251	6.90133E-05	0.471065216	0.009836448	0.674095456	0.004017264
PLAU	-1.316292651	1.97984E-08	-1.26929666	0.000140084	-1.384979097	4.02053E-05
PLAUR	-0.504355418	0.000365784	-0.57978451	0.001999694	-0.394112905	0.064915461
POU2F2	-0.482414564	9.17239E-07	-0.483269	0.000878999	-0.481165777	0.000912509
PSMB8	-0.17211	0.030245	-0.04890618	0.722042935	-0.352168114	0.004084143
PTAFR	-0.342918864	0.00060921	-0.27654	0.062596	-0.439939362	0.002767448
PTGER4	-0.597416412	2.13124E-07	-0.53083817	0.001517981	-0.694723068	4.26947E-05
PTGS2	-1.386926469	3.90116E-08	-1.59610243	3.20446E-06	-1.08121	0.015094
PTPN6	-0.325569865	0.000282579	-0.25426042	0.043644988	-0.429791368	0.004280161
PYCARD	1.130288149	3.07935E-06	1.238654912	0.00127915	0.971905957	0.000825361
RARRES3	0.191450911	0.017127959	0.306867286	0.008921978	0.02276544	0.883038304
RELA	-0.581464106	9.94774E-09	-0.46431158	0.001127077	-0.752687024	3.84915E-06
RELB	-0.432904533	1.93687E-08	-0.3725432	0.000756653	-0.521124937	1.23085E-05
S1PR1	0.392213947	7.07808E-07	0.372395484	0.000947654	0.421179392	0.001991995
sCTLA4	-0.478750817	0.000730678	-0.44104	0.053153	-0.533867162	0.001850054
SELL	0.523442023	1.35938E-10	0.537322904	4.30165E-06	0.503154582	0.000153988
SELPLG	0.757833519	8.83505E-10	0.797775589	8.20946E-06	0.699456648	0.000314071
SERPING1	0.51739366	0.00164636	0.703850838	0.004768582	0.244879	0.314625

	ALL patients (n=32)		RESPONDERS n=(19)		NON RESPONDERS (n=13)	
Gene ID	**logFC**	**adj.P.Val**	**logFC**	**Gene ID**	**logFC**	**adj.P.Val**
SIGIRR	0.407921841	1.87018E-08	0.388157515	5.04747E-05	0.436808164	0.001414008
SLAMF6	0.416915322	4.31456E-06	0.42569929	0.00214681	0.404077215	0.002437258
SLAMF7	-0.854984503	1.03024E-07	-0.59398939	0.00765582	-1.236438906	7.75456E-07
SMAD3	-0.457325657	8.20745E-07	-0.40614065	0.002213422	-0.532134521	0.000459663
SPP1	-0.794205003	0.006838877	-1.13375815	0.009734778	-0.297935018	0.525169427
SRC	-0.670388058	7.48334E-10	-0.63305952	4.9506E-05	-0.724945148	1.44397E-05
STAT3	-0.185796453	0.002739462	-0.15686506	0.084910947	-0.22808	0.030486
STAT4	-0.187332013	0.009835226	-0.15119	0.181257	-0.240157118	0.033513267
STAT5A	-0.511725699	2.11201E-08	-0.3733039	0.004269041	-0.714034475	4.03351E-07
TBK1	-0.548140736	3.17416E-09	-0.48200561	0.00013077	-0.644799762	1.23085E-05
TCF4	0.240389232	0.000182434	0.253729701	0.005921155	0.220892	0.027591
TGFBI	1.257893485	0.000934785	-0.03817041	0.799722111	1.707269505	0.008873181
TGFBR2	0.155682235	0.003427715	0.147221	0.045992	0.168048137	0.105730884
TICAM1	-0.627297527	9.94774E-09	-0.55943284	0.000251976	-0.726484375	5.38549E-05
TLR1	0.301487889	0.003795702	0.336966816	0.012181055	0.249634071	0.194842484
TLR2	-0.354026519	0.001753236	-0.38141463	0.010009451	-0.313997736	0.098863245
TLR7	1.502003515	4.85702E-07	1.630807897	4.23135E-05	1.313750958	0.008113729
TLR8	-0.24383	0.046783	-0.02036	0.923926	-0.570439977	0.000108151
TMEM173	0.212835816	0.001818817	0.249259051	0.010675997	0.159602	0.189862
TNF	-0.737837176	1.63006E-07	-0.62379687	0.001815452	-0.904511471	5.67222E-05
TNFAIP3	-0.743909873	3.73123E-12	-0.67338653	1.74548E-05	-0.846982456	1.5869E-08
TNFAIP6	-0.77937602	2.75691E-09	-0.63472702	0.000578627	-0.990786101	4.89881E-07
TNFRSF10C	1.662540627	2.75691E-09	1.539185685	5.50216E-05	1.842828618	4.26947E-05
TNFRSF13C	0.303073544	0.000626729	0.311564959	0.004009388	0.290663015	0.098863245
TNFRSF14	-0.23832668	0.000187472	-0.1966	0.039269	-0.299317481	0.001991995
TNFRSF1B	-0.331628502	0.003985162	-0.37665385	0.017275321	-0.26582	0.118605
TNFRSF8	-0.537510779	0.000157277	-0.59340591	0.004262181	-0.455817896	0.021041647
TNFRSF9	-0.737572652	7.60717E-06	-0.61598	0.012135	-0.915282656	0.00012323
TNFSF10	0.489942445	0.001425955	0.627319215	0.006396474	0.289161	0.208458
TNFSF12	0.288440067	4.37051E-06	0.244761545	0.009345021	0.352277908	0.000452173
TNFSF15	-1.050889668	3.49818E-05	-1.16887786	0.001218513	-0.87844539	0.034916538
TNFSF8	-0.564143873	6.42165E-08	-0.57483268	5.4258E-05	-0.548521777	0.002437258
TRAF1	-0.861854702	2.3542E-17	-0.81646177	4.52343E-08	-0.928198212	1.15864E-08
TRAF3	-0.499738901	6.61848E-10	-0.45991522	8.02238E-05	-0.557942744	3.84915E-06
XBP1	-0.566411149	3.14357E-11	-0.50349512	1.72885E-05	-0.65836535	8.47457E-06

*Gene expression data was analyzed in Truculture LPS stimulated samples from 32 patients. Limma analysis was performed to compare gene expression at D0 versus D90 in all 32 patients (column 2 and 3), or selectively in patients classified as Responders (column 4 and 5) or Non-responders, according to ASDAS criteria. Shown are the log fold change and adjusted p-values for the genes that resulted differentially expressed (adjusted p value equal or <0.01) in at least one of the three analyses. The grey shading indicates the comparisons that do not reach statistical significance at the adjusted p-value level of 0.05.

Supplementary Table 7. Gene Module Scoring Table when comparing D0 *vs* D7 for SEB and LPS stimulation

Gene module, SEB stimulation	log fold change	p Value	FDR
NFkB_inhibitors	-1.0811	2.02E-12	9.08E-11
NFkB_transcription_factors	-0.8426	2.51E-11	5.64E-10
NFkB_target_genes	-0.7197	4.46E-11	6.70E-10
TNF	-0.3519	1.23E-08	1.38E-07
NOD	-0.3290	0.0002	0.0009
dectin	-0.3283	4.29E-05	0.0003
IL1	-0.2952	6.18E-05	0.0003
TLR	-0.2914	0.0008	0.0031
NFkB_regulators	-0.2880	1.23E-06	1.11E-05
costimulatory_molecules	-0.2342	3.73E-06	2.8E-05
MAPK	-0.2158	0.0001	0.0007
M1_like_monocytes	-0.1913	0.0359	0.0734
NOTCH	-0.1599	0.0041	0.0142
complement	-0.1530	0.0387	0.0756
type_2_immunity	-0.1468	0.0190	0.0475
adhesion_molecules	-0.1445	0.0118	0.0312
JAK_STAT	-0.1374	0.0299	0.0641
IL6	-0.1241	0.1699	0.2548
autophagy	-0.1139	0.1506	0.2420
innate.immune.response	-0.1131	0.0951	0.1646
allergy	-0.1127	0.3181	0.4469
IFN_targets	-0.1111	0.0813	0.1510
NLR_inflammasomes	-0.1012	0.3928	0.5199
Tregulatory_cells	-0.0950	0.0839	0.1510
cell_cycle	-0.0905	0.1592	0.2470
apoptosis	-0.0806	0.1103	0.1839
phagocytosis	-0.0521	0.4715	0.5894
Tcell_inhibitory_signals	-0.0307	0.6199	0.7153
T_cell_signalling	-0.0052	0.9062	0.9483
type_17_immunity	-0.0037	0.9633	0.9852
TGFB	0.0006	0.9974	0.9974
proinflammatory_molecules	0.0143	0.8390	0.8989
type_1_immunity	0.0235	0.7952	0.8728
Bcells	0.0293	0.6046	0.7153
ubiquitin_proteasome	0.0304	0.5214	0.6341
chemotaxis	0.0324	0.6567	0.7387
lymphopoiesis	0.0458	0.4136	0.5318
MHC_presentation	0.0661	0.3278	0.4469
osteoclast	0.0872	0.3111	0.4469
regulatory_B_cells	0.1668	0.0084	0.0235
NKcells	0.2034	0.0268	0.0604
APC	0.2368	0.0011	0.0041
granulocytes	0.2622	0.0076	0.0229
cytotoxic_molecules	0.2823	0.0268	0.0604
M2_like_monocytes	0.3084	0.0073	0.0229

Gene module, LPS stimulation	log fold change	p Value	FDR
NFkB_inhibitors	-1.2400	3.44E-12	3.87E-11
NFkB_transcription_factors	-1.1598	1.78E-15	7.99E-14
NFkB_target_genes	-0.9257	4E-15	8.99E-14
NOD	-0.6677	0.0002	0.0007
IL1	-0.6262	1.47E-08	9.46E-08
TNF	-0.5435	4.41E-13	6.62E-12
M1_like_monocytes	-0.5073	0.0002	0.0006
dectin	-0.4728	0.0003	0.0008
NFkB_regulators	-0.4424	7.69E-10	6.92E-09
costimulatory_molecules	-0.4246	3.5E-08	1.97E-07
MAPK	-0.3875	6.25E-09	4.69E-08
type_1_immunity	-0.3813	0.0374	0.0732
type_2_immunity	-0.3728	0.0006	0.0017
Tregulatory_cells	-0.3393	0.0001	0.0004
type_17_immunity	-0.3152	0.0021	0.0052
TLR	-0.2818	0.0061	0.0144
IL6	-0.2740	0.0002	0.0007
NOTCH	-0.2562	0.0001	0.0004
JAK_STAT	-0.2377	0.0004	0.0011
cell_cycle	-0.2348	0.0001	0.0003
adhesion_molecules	-0.2189	0.0069	0.0156
autophagy	-0.2077	0.0177	0.0379
innate.immune.response	-0.2010	0.0362	0.0732
proinflammatory_molecules	-0.1644	0.1176	0.1864
Tcell_inhibitory_signals	-0.1397	0.0760	0.1368
apoptosis	-0.1290	0.0527	0.0988
complement	-0.1258	0.1201	0.1864
chemotaxis	-0.1028	0.3238	0.4180
ubiquitin_proteasome	-0.0873	0.0977	0.1691
phagocytosis	-0.0541	0.6510	0.7146
IFN_targets	-0.0537	0.5600	0.6300
allergy	-0.0535	0.7086	0.7416
MHC_presentation	-0.0474	0.3977	0.4774
T_cell_signalling	-0.0467	0.1170	0.1864
Bcells	-0.0173	0.7776	0.7952
NLR_inflammasomes	0.0318	0.8381	0.8381
granulocytes	0.0492	0.6799	0.7285
lymphopoiesis	0.0552	0.3251	0.4180
NKcells	0.0682	0.4773	0.5507
osteoclast	0.0878	0.3801	0.4751
M2_like_monocytes	0.1029	0.4031	0.4774
regulatory_B_cells	0.1091	0.1945	0.2917
APC	0.1140	0.2805	0.3944
TGFB	0.1404	0.2162	0.3139
cytotoxic_molecules	0.1906	0.3019	0.4117