Modeling autoimmune associated genetics in primary human T cells using CRISPR/Cas9 gene editing

Warren Anderson

A dissertation submitted in partial fulfillment
of the requirements for the degree of:

Doctor of Philosophy

University of Washington

2019

Reading Committee:

David Rawlings, chair

Richard James

Jane Buckner

Program Authorized to Offer Degree

Pathology

University of Washington

Abstract

Modeling autoimmune associated genetics in primary human T cells using CRISPR/Cas9 gene editing

Warren Anderson

Chair of Supervisory Committee:
Adjunct Professor David J. Rawlings
Department of Immunology

Abstract:

Genome wide association studies have identified genetic risk variants associated with multiple autoimmune diseases, thereby impacting large numbers of patients. Prominent examples are found within the phosphatase encoding genes *PTPN22* and *PTPN2*. Studies have shown that risk variants in these genes impact a variety of cell types, but their expression or ablation in lymphocytes alone can be sufficient to drive autoimmunity in certain mouse models, due in part, to hyper-active lymphocyte signaling. Notably, cross-sectional studies of human carriers of the *PTPN22* or *PTPN2* risk variants have shown these donors to possess T cells that are hypo-responsive to TCR or cytokine stimuli. To investigate this discrepancy and understand the functional impact of these variants, new methods of studying these genes are required. The purpose of this study is to establish methods of using gene editing approaches in primary human T cells to modify expression the genes *PTPN22* and *PTPN2* and assess the impact of altered gene expression. Using this approach we found that ablation of PTPN22 and PTPN2 in primary human $CD4^+$ T cells results in T cell responses which mirror current mouse models, with increased responses to TCR stimulation upon disruption of either gene, and enhanced responsiveness to IFNγ and IL-2 in *PTPN2* disrupted cells. Interestingly, in the case of *PTPN2* disruption, we found responses to IL-2 to be dynamic, eventually resulting in loss of responsiveness to IL-2, mirroring

current human data from carriers of a reduced expression variant. Furthermore, we have explored several efficient methods to alter the coding sequence of *PTPN22* to reflect its risk variant in non-risk donor T cells. Collectively our data shows that gene editing is a powerful tool for investigating how gene variants can contribute to disease, and that the effects of genetic risk variants may impart contextual and dynamic phenotypes on human lymphocytes. Finally, the methods we have established in this study are applicable to many other gene variants, and potentially could be utilized in multiple primary cell types.

Table of Contents

List of Figures

Acknowledgements

Regarding gene disruption studies I would like to thank collaborators Jerill Thorpe, Alice Long, and my mentor David Rawlings. For work involving cloning I would like to thank Tania Habib and Megan Tatum from the Buckner Lab at the Benaroya Research Institute. For all work presented using AAV I would like to thank Iram Khan and the Rawlings lab viral core. A special thanks to Rawlings lab members Yuhci Honaker and Karen Sommer for their help and advice in assay development and experimental protocols. I would also like to thank Jane Buckner (Benaroya Research Institute, Seattle, WA) and Rich James (Seattle Children's Research Institute, Seattle, WA) for their helpful comments and insight.

Dedication

To my wonderful wife Jennifer.

My successes are made possible by your sacrifices.

I love you,

Hubs.

Chapter 1

Introduction

Autoimmunity

It was once considered unthinkable that the body's immune system would turn against its host. The term "horror autotoxicus" referred to a class of antigens that the adaptive immune system would be prevented from reacting to, on the grounds that doing so would cause severe health consequences[1]. Autoimmunity, or the loss of immune tolerance to self-antigen, we now know occurs with startling frequency, resulting in degenerative health conditions termed autoimmune diseases. Autoimmunity is most prevalent in western countries affecting between 4 to 10% of certain populations[2], accounting for a large portion of health care spending[3] and is increasing in prevalence[4]. To date there are more than 80 distinct autoimmune diseases[5], with unique pathologies targeting singular organs such as Type 1 Diabetes (T1D), or multiple tissues systemically such as Systemic Lupus Erythematous (SLE). Autoimmune diseases can occur at different stages throughout life and present with remarkable heterogeneity[6,7]. Reflecting this heterogeneity, therapeutic responsiveness is highly variable, with frequent disease progression despite treatment, severely impacting quality of life and longevity.

Risk factors for autoimmunity

Reflecting the heterogeneity of autoimmune diseases, potential causes of these diseases are similarly diverse. One of the most striking risks regarding autoimmunity is its disproportionate burden on women. For some diseases, such as SLE and primary Sjogren's syndrome (pSS), women are between 7 and 10 times more likely to be affected by these disorders than men[8]. Reasons for disproportionate burden of some autoimmune diseases have been postulated to be the effects of estrogen and other gender specific hormones on immune cell subsets in certain organ systems[9-11]. Also, the possibility of incomplete X chromosome inactivation in cell subsets could impact immunological processes by altering

expression levels of autoimmune related genes on the X chromosome, such as *TLR7*[7,8]. But interestingly, gender discrepancy is not a ubiquitous feature of autoimmunity, with other diseases such as T1D presenting more equally among genders[12,13].

Environmental factors have been associated with several autoimmune diseases. Smoking is a potent example of environmental risks for autoimmunity, as it is a risk factor for rheumatoid arthritis (RA), wherein smoking is believed to induce elevated concentrations of autoimmune associated citrullinated proteins in the lung[14,15]. Diet and microbiota have also been implicated in increased proclivity to autoimmunity[16]. The microbiota develops in a progressive fashion after birth, and factors such as vaginal birth or the presence and duration of breast-feeding can influence microbiome flora in drastic ways that may associate with the development of T1D[17]. Composition of the microbiome has also been shown to influence immunological processes in diverse ways such as altering cytokine production and support of tolerance maintaining regulatory T cells (Tregs)[18,19]. Though the potential environmental contributors to autoimmunity are vast, many are highly context and disease specific, obscuring a mechanistic understanding of how these factors relate to human disease.

The advent of immunotherapy as a frontline cancer fighting approach has yielded a new variety of patients presenting with autoimmunity[20]. Generally used in cancer settings, these therapies, such as checkpoint inhibitors, modify natural immunological processes and empower the patient's immune system to overcome immunosuppressive barriers erected by their malignancies[21]. Interestingly, use of these therapies correlates with increased incidence of several autoimmune manifestations, despite no obvious link to known autoimmune risk factors[22,23]. Such outcomes argue that the immune system is constantly walking a fine balance between promoting aggressive responses to unwanted invasion and cancer and restraining such responses sufficiently to prevent autoimmunity.

Genetic variants that promote heightened immune responses to pathogens have been a major driving force of human evolution[24], therefore it is no surprise that genetics is a major contributor to autoimmunity. Collectively, genetic predisposition constitutes one of best-researched and well-established risk factors for driving autoimmunity, with many specific genetic risk factors[25] shared among multiple

diseases and these shared genetic risks may highlight common mechanisms underlying loss of tolerance[26]. As is true of other risk factors previously discussed, genetic components driving autoimmunity can be both simple or heterogeneous, depending on the overall heritability of disease[6], and the mechanism of action as to how such genetic variables contributing to loss of tolerance are still controversial.

Genetics of autoimmunity

The relationship between genetics and autoimmunity has been well-established due to the propensity for autoimmune diseases to cluster in specific families. Overall, genetic heritability is variable among autoimmune diseases as demonstrated by concordance between genetically identical twins[27]. For heterogeneous diseases such as RA, if one twin is diagnosed the other will be affected in 12 to 30 percent of cases, while some diseases like T1D shows more heritability with concordances between 30 to 50 percent[28], however, limitations of twin studies makes these figures difficult to refine and some argue that total heritability may be significantly higher[29]. In some rare instances, single genetic variants or de-novo mutations are sufficient to cause disease outright, such as mutations to the genes *FOXP3* or *AIRE* driving the systemic and potentially lethal autoimmune diseases: immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) and autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED) respectively[30,31]. However, for commonly diagnosed autoimmune diseases most genetic associations comprise a relatively modest increase or decrease in risk to specific forms of autoimmunity[26].

The strongest genetic contributor to autoimmunity is human leukocyte antigen (HLA), also referred to as the major histocompatibility complex (MHC)[32]. This class of molecules comprises a family of proteins that are used to present antigen to T cells, licensing them to participate in and orchestrate an array of immune responses. The genes that encode these molecules are highly polymorphic and many haplotypes are highly associated with specific autoimmune diseases, such as HLA-B27 or HLA-DR3 which are associated with increased incident of ankylosing spondylitis (AS) and T1D respectively[33,34].

Furthermore, relative risk may increase up to 5-fold depending on a person's specific combination of HLA genes[35]. The reason for this connection between specific autoimmune diseases and certain HLA haplotypes is unclear, but leading hypotheses revolve around altered binding of self-peptides in a manner that promotes self-reactivity, molecular mimicry, and possible alterations in central tolerance [35].

The role of HLA in autoimmunity has been well appreciated for decades due to ongoing investigation of HLA in other aspects of medicine[33], while recent technological advancements in DNA sequencing has made whole genome sequencing more accessible allowing in-depth studies of genetic associations to autoimmune diseases[26]. Genome-Wide Association Studies (GWAS) are case control studies that assemble the DNA sequences of large cohorts of patients diagnosed with a specific disorder, which are compared to non-diseased controls. Computational analysis quantifies variants in genetic sequences, determining whether they are enriched or diminished in disease cohorts relative to controls. Over the past fifteen years, GWAS studies have confirmed or revealed hundreds of gene variants and single nucleotide polymorphisms (SNPs) that are significantly associated with various autoimmune diseases[36,37].

While the volume of data that has emerged from GWAS is impressive, these studies do not describe how these variants contribute to disease. Also, most variants identified offer relatively modest increases in risk for any specific disease when compared to HLA haplotypes, usually bestowing an odds ratio between 1.2 to 2 fold higher than the general population[36]. Furthermore, it is not always possible to fully identify the causal variant in all cases, as many potential variants occur in linkage disequilibrium with each other[38], allowing the assessment of a risk variant to only be located within a specific haplotype block, or a defined region of a gene or the genome. Finally, even with the data provided by numerous large GWAS studies, total heritability of all autoimmune diseases are not believed to be accounted for[25].

Though the genetic mechanisms that predispose individuals to autoimmunity have not been fully elucidated by GWAS, these studies have provided strong clues as to the nature of how genetics can push an individual's immune system towards loss of tolerance. GWAS autoimmune associated variants cluster in genes that are preferentially expressed in immune cell subsets[39]. Though some gene variants identified

have alterations to the protein-coding region, such as in the genes *PTPN22* or *TYK2*, it is estimated that 90% of GWAS identified risk variants are non-coding[40]. Most of these non-coding sites are hypothesized to be estimated quantitative trait loci (eQTLs), or sites believed to control the expression levels of the genes they occupy through acting as transcription factor binding sites or sites for epigenetic modification. Follow-up studies have demonstrated that while a majority of variants localize near eQTLs, a small percent (~25%) seem to impact gene expression level at baseline in some immune cell subsets[41]. However, other studies have shown that when eQTLs are examined in relevant tissues their impact is highly tissue specific[42] and many only show expression modifying effects when those specific cell types are stimulated in a functionally relevant context[43]. Also, the effects of eQTLs may be trans-acting, or impacting the expression of genes far removed from their genomic locus[44], which makes their contribution to autoimmunity harder to predict. Therefore, genetic association studies have shown that most (though not all), non-HLA genetic risk factors for autoimmunity likely paint a complex network, affecting genomic regions that may impact expression levels of multiple genes in a highly context, and tissue specific manner[45].

Perhaps most importantly, several GWAS studies have implicated numerous risk variants found in genes such as *PTPN22*, *CTLA4*, *IL23R*, *PTPN2*, and *TYK2*[46,47], to significantly associate with many autoimmune diseases affecting multiple different organs and tissues. The association of genes in multiple diseases may highlight common mechanisms that are exploited among those diseases to eventually result in a loss of tolerance[26]. However, elucidating a mechanism of action for many variants has been complicated by the fact that these immune related genes are expressed at variable levels in different immune cell subsets[48]. For example, the autoimmune risk gene *PTPN22* is expressed to varying degrees in multiple hematopoietic subsets[49], and has been shown to function in numerous immunological pathways[50]; therefore the gene has been argued to modulate risk to diseases through a number of mechanisms including: alteration of response to and secretion of type 1 interferons[51-53] in multiple cell lineages, increased release of potential auto-antigens by neutrophils[54], and altered lymphocyte antigen receptor responsiveness[38,55,56]. While none of these hypothesizes need be mutually exclusive, it does beg the question: which mechanism, or combination of mechanisms, contributes to a given disease?

While mechanisms of action are currently still up for debate for many of the genetic variants in autoimmunity, genetic risk associations have already begun to show promise as a clinical tool, such as using HLA haplotype to avoid allergic reactions to specific drugs[57]. Recent work has shown that genetic risk scoring systems, combining data on genetic variants to calculate relative risk of disease[58], may aid clinical monitoring and treatment. Similar work has been done in the context of T1D, demonstrating that combinations of specific genetic risk factors may yield risk ratios for disease higher than a simple additive model of their individual assessed risk[59]. This shows that combined genetic risk assessment is a superior predictor of T1D relative to HLA alone, and potentially, the combined effect of a few minor risk variants may result in a substantial risk of autoimmunity. Such approaches could allow for earlier intervention in disease progression and improved disease maintenance**.** Several different treatment options for early intervention in T1D are in trial with some promising outcomes that may offer improved maintenance of β-cell function and minimal impact on overall health[60-62]. Given the similar genetic etiology of T1D to other autoimmune diseases, as implied by shared GWAS risk loci[46,47], it is conceivable that such approaches could be utilized to control other autoimmune diseases, re-enforcing the need for more information on how specific GWAS risk variants contribute to autoimmunity.

Genetic variant impacts on immune function, a potential source of controversy

As stated earlier, most GWAS identified risk variants are non-coding and highly cell type specific. Frequently the cell-type impacts of a given variant are identified through cell and mouse models, allowing hypotheses generation regarding how these variants can contribute to autoimmunity; however when these hypotheses are tested in primary human tissues from carriers of the variant, opposing phenotypes may emerge from that of mouse models. Such paradoxical findings between human studies and mouse models have been reported regarding autoimmune risk SNPs in the genes *PTPN22* (rs2476601) and *PTPN2* (rs1893217), both of which are phosphatases that regulate numerous and overlapping cell activation/ signaling pathways in multiple immune cell subsets. Interestingly, while rs1893217 in *PTPN2* is a non-coding variant that has been shown to impact expression levels in T cells specifically[63] rs2476601

in *PTPN22* is a coding variant, altering the sequence of PTPN22 in all hematopoietic subsets[64], which complicates how this SNP may contribute to autoimmunity. For example *PTPN22* rs2476601 increases the release of citrullinated peptides as a part neutrophil netosis[54]. As citrullination is a post-translational modification of peptides that is specific to subsets of cells (such as neutrophils) actively expressing key genes[65], it is believed that exposure of lymphocytes to citrullinated peptides during negative selection is minimal[66], allowing the escape of lymphocytes that are reactive to citrullinated peptides (a common feature of of RA patients)[66]. Given that *PTPN22* rs2476601 has also been shown to impact lymphocyte signaling[55,67], it is entirely possible that rs2476601 increases the abundance of pathogenic auto-antigens, and also allows for the selection of ACPA reactive lymphocytes.

Regardless of how these SNPs in *PTPN22* and *PTPN2* may impact different cell subsets, they, as with most autoimmune associated GWAS variants, have been shown to impact B and/or T cell function[45]. Recent mouse model work has shown that for autoimmune diseases which feature auto-antibody formation, such as T1D or pSS, that disruption of T cell help to B cells via prophylactic administration of CD40 or CD40L blocking antibodies is sufficient to halt the onset of disease[68]. Furthermore, CD40L blocking drugs have shown clinical efficacy in human trials of patients with active RA as well[69]. Furthermore T cell autoimmune cytolytic damage or inflammation can be driven in the absence of autoantibodies, making T cell activity essential for nearly all autoimmune disease pathology[7]. This makes understanding the impact of the rs2476601 and rs1893217 variants on T cell function critical for understanding how they are involved in autoimmune risk; unfortunately consensus on this impact has been elusive, as discussed below.

PTPN22

Fifteen years ago, it was first reported that a SNP (rs2476601) in the gene Protein Tyrosine Phosphatase Non-receptor 22 (*PTPN22)* was significantly enriched in T1D patients compared to non-diabetic controls[64]. Subsequent studies reported that the variant was also enriched in RA and SLE patients[49,70]. Over time the rs2476601 risk SNP has been shown to be associated with nearly a dozen

different autoimmune diseases, most of which strongly feature auto-antibody formation[71]. Since that time, PTPN22 has garnered much attention for two good reasons: 1) the *PTPN22* rs2476601 variant is one of the strongest non-HLA genetic factors associated with many prevalent autoimmune diseases (with an odds ratio between 1.5 to 2 for T1D, RA, and SLE) [72]; and 2) that *PTPN22* rs2476601, unlike most GWAS identified risk variants, is a coding variant, producing an arginine to tryptophan substitution at amino acid residue 620 (R620W)[64]. Due to PTPN22's known involvement in regulating antigen receptor signaling[73,74] it was reasoned that this SNP would alter protein function in a manner that is conducive to autoimmunity.

In brief, when the T cell receptor (TCR) is presented antigen on an MHC molecule by an antigen-presenting cell (APC), the immunoreceptor tyrosine-based activation motifs (ITAMs) on the TCR zeta chain are phosphorylated by Src family kinases, allowing for recruitment and activation of the protein ZAP70[75]. Binding of ZAP70 allows for the activation and disbursement of many effector proteins eventually promoting T cell adhesion, mobility, coordinated expression of transcription factors for growth, differentiation, and cytoskeletal reorganization, all hallmarks of T cell activation[76]. Concurrently, upon TCR engagement, the protein PAG releases the regulatory kinase Csk which associates with its co-inhibitory phosphatase PTPN22[77,78]. Csk and PTPN22 work in tandem to halt phosphorylation of TCR ITAMs, with PTPN22 de-phosphorylating the activating tyrosine's of Src family kinases and other key signaling proteins, such as ZAP70[73], while Csk phosphorylates inactivating residues on these proteins, altering their conformation and ceasing their activity[78,79]. This co-inhibition by PTPN22 and Csk, is a major component of restraining undesired TCR activation[80,81], and ultimately ceasing TCR signaling after antigen presentation[82]. PTPN22 consists of three domains, a catalytic domain, an interdomain, and a proline rich domain, consisting of multiple proline rich repeats that allow association with Csk and other potential binding partners[79,83]. The R620W risk variant is located within the proline rich domain required for binding to Csk and has been shown to disrupt the association of these proteins[64].

Early investigation into PTPN22 had shown that siRNA knock down in Jurkat T cell lines produces heightened TCR signaling[49] and deletion in mouse models likewise showed enhanced TCR signaling in memory T cells[84], and though they do not develop autoimmunity they do develop splenomegaly and lymphadenopathy due to significant expansion of memory/effector T cells. This

expansion of effector/memory T cells is also found in human carriers of the *PTPN22* risk variant[56]. Mouse models of the *Ptpn22* risk variant knock-in (KI) murine model have demonstrated several key autoimmune disease features including an expanded effector T cell compartment, enhanced calcium flux in response to TCR engagement, altered B cell signaling, and spontaneous autoimmunity, findings all consistent with loss of antigen receptor regulation[55,67].

While mouse models largely agree that the *Ptpn22* risk variant produces a hyper-responsive phenotype in lymphocytes, similar to what is seen in memory T cells of knockout mice; studies of human T cells have produced opposing results. Mature T and B cells derived from the peripheral blood of healthy human subjects (either homozygous or heterozygous for the *PTPN22* risk allele) exhibit reduced calcium flux, IL-2 secretion, and CD25 expression in response to antigen receptor stimulation[56,85,86]. Based on these observations, several groups have suggested that the risk variant reduces response to self-peptide, leading to the escape of self-reactive T and B cells from negative selection[56,87]. Consistent with this idea, carriers of the risk variant also exhibit an increased proportion of Th1 T cells[56,88]. This data generated from human risk SNP carriers supports the hypothesis that the *PTPN22* risk variant acts as a gain-of-function variant, driving decreased, not increased, antigen-receptor signaling in lymphocytes. This may be possible as human and mouse genes coding for PTPN22 are the most divergent examples of phosphatase orthologs between human and murine species (89% and 61% homology in their catalytic and non-catalytic domains respectively)[74]. However, recent studies demonstrated that transfection of primary human cells with siRNA targeting *PTPN22* leads to an increase in IL-2 release upon antigen stimulation[89], and likewise, use of CRISPR/Cas9 in human Jurkat T cells has shown loss of PTPN22 resulted in increased CD69 expression and IL-2 secretion in response to TCR stimulation[90]. These increased responses to TCR engagement in the absence of PTPN22, in human derived cells, shows that PTPN22 similarly regulates TCR signaling in these models as it does in mouse knockout models[84], but detailed analysis of PTPN22 loss-of-function in primary human T cells is lacking.

PTPN2

Similar to *PTPN22,* the gene Protein Tyrosine Phosphatase Non-receptor 2 (*PTPN2)* was connected to multiple autoimmune diseases nearly 15 years ago in some of the first large scale GWAS studies of multiple autoimmune cohorts[91,92]. A non-coding SNP within *PTPN2*'s Intron 7 (rs1893217), was discovered to predispose carriers to increased risk of T1D and RA and has since been connected to other autoimmune diseases such as juvenile idiopathic arthritis[93] and it may be associated with Crohn's disease[94]. This risk variant has been shown to reduce *PTPN2* transcript levels in the T cells of human carriers in a dose-dependent manner[63].

Before its connection to autoimmunity, PTPN2 (also called TCPTP) had been established to participate in TCR signaling[95]. Global *Ptpn2* mouse knockout models produce a scurfy-like phenotype, featuring stunted growth, juvenile mortality, and various immune cell subsets that are hyper-responsive to *both* antigen and cytokine stimulation[96,97]. For IL-2 family cytokine receptors, upon binding their ligand, a chain of signaling events, led by the phosphorylation of certain Janus kinases (JAK proteins), ultimately results in the phosphorylation of specific signal transducer and activator of transcription proteins (STATs) among other key signaling molecules[98]. PTPN2 had been shown to regulate JAK/STAT signaling responses to the cytokines IL-2 and IFNγ (among others), by dephosphorylating JAK and STAT proteins[96,97,99]. Interestingly, though T cells from *PTPN2* rs1893217 carriers have been shown to have reduced *PTPN2* transcript levels, memory T cells from carriers of this variant show blunted pSTAT5 response to IL-2 and IL-15[63,100,101].

Mouse studies in T cell specific *Ptpn2* knockouts have shown that loss of Ptpn2 results in hyper-active TCR signaling, enhanced pSTAT5 signaling in response to IL-2, and autoimmunity that is transferrable through CD8+ T cells alone, even when bred onto a non-autoimmune prone background[102,103]. siRNA PTPN2 knock down in cell lines have shown lower expression of *PTPN2* to yield enhanced pSTAT1 and pSTAT5 signaling in response to INFγ and IL-2 respectively[104]. Restriction of *Ptpn2* deletion to the T cell lineage on the non-obese diabetic (NOD) background, significantly accelerates diabetes onset/progression compared to *Ptpn2* competent mice. This model demonstrated

increased T cell activation, cytokine secretion, responsiveness to IL-12 and INFγ, and autoantibody formation, despite significantly expanded Treg compartments that have unaltered suppressive capability[105]. This suggests that disruption of *Ptpn2* drives autoimmunity via enhanced T cell signaling, and increased T cell help in B cell activation and affinity maturation. This NOD-*Ptpn2* disruption model is highly relevant to the human SNP, as heterozygous knockout mice offer an intermediate phenotype, with increased diabetes compared to wild-type NOD mice and increased serum autoantibodies[105]. Collectively, murine data demonstrates that dose dependent loss of Ptpn2 in T cells drives autoimmunity through, at least in part, enhanced T cell activation, standing in contrast to primary human data that demonstrated dose dependent loss of PTPN2 to result in decreased responses.

New methods for resolving GWAS genotype – phenotype discrepancies

Contradictory phenotypes between human cross-sectional studies and mouse models of GWAS variants are well exemplified in the body of data surrounding *PTPN22* and *PTPN2*. Data from mouse models of both variants would suggest a pre-disposition to autoimmunity through increased lymphocyte activation; human data would suggest these SNPs predispose their carriers to autoimmunity through reduced or modulated responses to various stimuli. Though it may be convenient to dismiss these contradictions as differences in functionality between human and mouse genetics, the strong connection of these genetic variants to not one, but dozens of autoimmune diseases, and therefore potentially millions of patients, demands a more thorough examination to account for these discrepancies. Furthermore, our ability to examine other GWAS variants such as these in a human context is limited by the availability of human carriers. For example, GWAS studies have identified a separate SNP (rs33996649) in the gene *PTPN22,* which is strongly related to protection from RA and SLE[106,107]. This creates an interesting counterbalance to the rs2476601 SNP, which increases risk for those diseases and may be equally informative as to a genetic pathway for the loss or re-enforcement of self-tolerance. However, due to the rarity of the rs33996649 SNP in the general population, cross sectional studies of the

rs33996649 variant's function on primary human T cells or the generation of mouse models has not been conducted.

Data surrounding the *PTPN22* and *PTPN2* SNPs currently seems at an impasse, but this study seeks to explore a novel approach to resolving the contradictions present regarding these variants. In recent years, major advancements have been made in the field of gene editing. This new tool has been proven to be useful for the quick generation of new genetically modified cell lines and animal models, as well as a viable option for the improved treatment of multiple genetic disorders. In addition to these uses, a body of research has proven the utility of gene editing in a human context through the manipulation of human $CD4^+$ T cell genes[108,109], which has been shown to be minimally impactful to cell viability and function. Groups have called for the use of DNA editing technology to be used for improved functional analysis of human GWAS variants[36,45]. This approach would use healthy control donor tissues and gene editing approaches to alter the DNA of GWAS variant loci to reflect autoimmune associated variants, thus allowing the consequential effects of a variant to be examined on an isogenic human background. Such an approach could potentially be an extremely powerful *in vitro* approach to GWAS variant analysis. The purpose of this study is to examine how to apply new opportunities created through gene editing to address old questions such as the discrepancies surrounding *PTPN22* and *PTPN2*. Manipulation of these genes in human $CD4^+$ T cells would allow a more thorough examination of the roles of these genes and impacts of their variants in a human context, while providing a method to study other rare but informative GWAS variants in human tissues. Furthermore, validation of this approach in primary human T cells would be informative for the incorporation of human SNP editing studies in other cell subsets that may be impacted by autoimmune associated GWAS variants.

Gene editing

The desire to perform targeted manipulation to the genome has been discussed for decades. Methods to perform homologous recombination (HR),the insertion of a DNA template based upon homology to a genomic site, have long been available, but generally work at low efficiency[110]. However, it

has been known for some time that the introduction of a double stranded break (DSB) near the site of genetic manipulation will increase HR rates by orders of magnitude[111,112]. Over the last decade, an increasing range of designer nuclease "platforms" have become available that allow investigators to achieve efficient site-specific DNA cleavage[113,114]. The increasing availably of these nuclease platforms have made gene editing a practical reality in many models ranging from cell lines to embryonic cells allowing for a new range of animal models. Over recent years some groups have shown efficient gene editing is possible within primary human cells, including T cells[109,115].

Designer nucleases

Nucleases (enzymes that cleave DNA at specific sequences) have been used for decades in cloning and other biological processes. However, for useful gene editing applications nuclease activity is required which is both: 1) site specific (one singular site within the genome) and 2) programmable, so as to be used for different applications in different genes. Previous work has been performed using zinc finger nucleases, transcription activator-like effector nucleases (TALENS), and homing endonucleases[114]. However, platforms like zinc-finger nucleases and TALENs required extensive cloning and expertise to make two separated DNA binding motifs which bind to specific regions of the genome and allow for the dimerization of a FokI nuclease. However, the gene editing field has exploded with the introduction of the programmable nuclease platform: clustered regularly interspaced short palindromic repeats (CRISPR/Cas9).

What makes CRISPR/Cas9 so useful is that it has few components and forms a single complex that localizes to a programmable 20 base-pair sequence of DNA and produces a DSB[116]. The CRISPR complex is a type of innate immune function in some prokaryotes, which allows for the removal of invasive DNA in its genome[113], and related variations can be found in many species of bacteria. CRISPR/Cas9 utilizes three components: a single-stranded CRISPR RNA molecule (crRNA) which aligns to a 20bp DNA sequence target (termed the protospacer), a transactivating RNA (tracrRNA) which complexes with the crRNA (becoming a guide RNA or gRNA) and acts as a scaffold that allows

association with a Cas9 nuclease molecule[113]. When all three components are assembled the complex will bind to the protospacer sequence if it is present in the DNA *and* is next to a protospacer adjacent motif (PAM site). The complexed Cas9 will then produce a DSB 6 base pairs upstream (5') of the PAM site. The most versatile part of this platform is that to change nuclease targeting sites requires only modifications to the protospacer sequence and that the new protospacer be located next to a PAM site. The most commonly used form of CRISPR comes from the bacteria Streptococcus *pyogenes*, which recognizes a PAM site of NGG, meaning that a 20bp protospacer sequence can be designed to target any site in the genome as long as the next 3 base pairs are NGG. This opens up far more of the genome to genetic manipulation than was previously available with other nuclease platforms.

DNA double stranded break induction and repair

After a DNA double-stranded break is created, the cell must repair the break to continue normal growth and function. DSB repair occurs through two main pathways: non-homologous end joining (NHEJ) or homology directed repair (HDR), and a number of sub-pathways[117]. Cellular choice between NHEJ and HDR is highly relevant to gene editing as it determines what genetic manipulation is permissible. This decision between NHEJ and HDR is dependent on several factors, primary of which is the phase of growth the cell is currently in[118], and the presence or absence of a homologous DNA template.

At any non-mitotic stage of cell growth, when a DSB occurs the hetero-dimeric proteins Ku70/80 will attempt to bind the free ends of DNA, protecting them from further damage[119]. The DNA bound Ku proteins, after phosphorylation by DNA-dependent protein kinase, act as a platform for the binding of several DNA end-processing factors, including the nuclease Artemis, and the DNA Ligase IV. Artemis works to trim the free ends of DNA and make them acceptable for ligation, then the ligase process joins the broken ends together[119]. Due to the error involved in Artemis activity, among other steps[120], this process frequently results in insertions or deletions of base-pairs of DNA (termed Indels), which if created in a coding region of a gene can lead to gene knockout through frame shift mutations or altering protein

folding. If the desired use of gene editing is the disruption of gene function, then Indel formation in the coding region of a gene can be desirable.

If a cell is in S or G2 phase and a DNA break occurs, then HDR is available to the cell as an alternative to NHEJ. HDR is a much less error prone pathway and offers repair of DSBs that occur regularly during genome replication with minimal occurrence of Indels and other mutations. In brief, during these cell phases altered regulation of numerous kinase and ubiquitination pathways allow for the increased exclusion of Ku proteins from the DSB locus and the binding of nuclease complexes to the free DNA[118]. Nuclease complexes initiate a series of events that lead single stranded resection of DNA[121] and then the search of these single stranded arms of DNA for homologous DNA to use as a template for extension and repair[118]. Ideally when this occurs the cell would intend to find homologous DNA from the cell's sister chromatid, but if a designer DNA template with homologous ends (frequently called homology arms) is transmitted into the nucleus at the time the DSB is generated, the cell may use this new template for HDR with the resected DNA strands[113]. If the desired use of gene editing is to produce a specific alteration to a genetic locus, then an HDR event is desirable.

Routes of delivery

Various methods have been used for delivery of nucleases, including CRISPR[122,123], into primary human T cells. However, the introduction of ribonucleoproteins (RNPs), a pre-complexed form of the CRISPR system delivered as a Cas9 protein loaded with a guide RNA sequence, has been shown to superiorly edit primary cells[109,124]. Our laboratory and others have shown efficient modification of the genome of primary human lymphocytes is achievable via delivery of RNP by electroporation and donor DNA template delivery via Adeno-Associated virus (AAV)[125,126]. Furthermore, potent HDR outcomes with minimal toxicity are possible if DNA donors are simultaneously introduced using ssDNA in the form of single-stranded oligo-deoxynucleotides (ssODNs) upon electroporation[109,127].

AAVs are a non-pathogenic form of parvovirus harboring a non-enveloped ssDNA genome[128]. AAVs of several serotypes can be constructed artificially by use of specific helper cell lines, producing

viruses that package DNA constructs, devoid of viral components, which can be used for gene editing applications[129]. When added to cell culture AAVs bind to a surface receptor protein and then make use of endosomal pathways to traffic to the nucleus where the viral capsid is disassembled releasing the DNA genome[130]. Use of AAV in gene editing has one large advantage over other methods of donor DNA template delivery, and that is packaging size[110,129]. Designer AAVs can deliver donor DNA sequences over 4 kB, and when used for the introduction of a reporter gene such as GFP, this approach allows for both high rates of gene modification within primary human T cells, and rapid purification of modified cells through fluorescence activated cell sorting (FACs).

Concerns in editing primary human cells

Gene editing has been straightforward for some applications. Use of any of the tools described, in most cell line models is permissible as many are resistant to nucleotide mediated toxicity, grow indefinitely, and extended purification protocols are acceptable to produce populations of homogenous, gene edited cultures. But if the desired outcome is the manipulation of somatic human cells, such as T cells, then several concerns must be adequately addressed. First is that primary T cells are resistant to the introduction of foreign nucleic acids through several different anti-viral mechanisms[131]. As such, the direct introduction of CRISPR RNA components or donor DNA is toxic to T cells resulting in death or inflammatory responses from nucleic acid sensing pathways such as STING or RIG-I[132,133]. To avoid this, the use of RNPs has proven to be relatively inert on cell viability as the majority of the RNA components are shielded from detection by the Cas9 protein, but even small exposures of phosphorylated ends of ssRNA contained within Cas9 molecules has been shown sufficient to produce inflammatory responses from primary cells[133]. However, commercially available CRISPR gRNAs can be pretreated with phosphatases or chemically modified to avoid such outcomes[108].

Delivery of homologous DNA templates elicits similar concerns. When cells are electroporated with ssODNs viability is minimally affected as it is believed that DNA sensing pathways are generally able to recognize only double stranded DNA[134]. However, in some situations, the STING pathway can

recognize ssDNA[132], but these concerns only come into play when the ssDNA begins to increase in size. Also, it has been shown recently that, though electroporation of plasmid DNA is extremely toxic, the electroporation of dsDNA may be tolerated for some gene editing applications[135], though limitations to this approach do apply. AAVs are also able to deliver homologous DNA to the nucleus of cells avoiding some of these issues[136], but the addition of AAV to T cell culture does also produce toxicity and phenotypic effects of T cells as well (discussed in chapter 2).

Regardless of the method taken in the editing of primary T cell genomes, there will be drawbacks that must be accounted for and controlled. Even the act of electroporation itself causes massive depolarization of the cell, and a strong stress response[137]. Also, upon DSB induction it is appreciated a p53 response that halts cell cycle till repair of the break is completed[117,138], and in some cell types this response can drive senescent phenotypes[139]. If gene editing is the most logical way to address a scientific question, then, as in all good science, appropriate controls must be established (See Chapter 2).

Questions to address

GWAS studies have produced a large set of well-validated genetic variants that predispose their carriers to a wide variety of autoimmune manifestations[5,36]. Some of these variants have been heavily studied in hundreds of publications over fifteen years, but still lack consensus as to how they promote autoimmunity in human patients. These disagreements are driven by divergent results generated from either studying these SNPs in cross-sectional human cohorts or modeling them in cell lines or mouse models. As cross-sectional studies of human lymphocytes are largely conducted on donated tissue from consenting adults, and many autoimmune diseases tend to show onset at a relatively young age, with driving manifestations likely occurring years before the presentation of clinical symptoms[68,140]. Studies using primary human tissue make use of the most clinically relevant model, while potentially risking examining the aftermath of a genetic variant's effects rather than how that variant drives a given disease. Conversely, current use of mouse models allows for more conclusive causal data on specific genetic variants, at the risk of studying such variants in a non-human model with an inbred background. The

trade-offs of these two approaches obviates the need for improved methods of generating genotype-phenotype data in primary human specimens.

With the ability to produce a targeted DSB at nearly any desired location in the genome, and the capacity to deliver of a broad array of DNA repair templates using a variety of methods, we now have the opportunity to use gene editing as a platform for the functional study of human genetic variants within primary human cells. This approach is predicted to be faster and less expensive than assembly and testing of large genotyped cohorts of human subjects (as currently required in human genetic studies) and permits generation of both control and edited cells from an isogenic donor, providing a powerful new addition to current methods for studies of human autoimmune risk genes.

Can disruption of genes in human T cells using gene editing methods demonstrate immunological function?

Previous studies of the protein phosphatases PTPN22 and PTPN2 have shown that SNPs in either gene are strongly associated with increased odds of diagnosis with one of several autoimmune diseases, including T1D and RA[91,92]. *PTPN22* rs2476601 is a coding variant that causes an arginine to tryptophan variant, disrupting the association of PTPN22 with other co-inhibitory molecules[64]; while *PTPN2* rs1893217 is a non-coding variant that impacts expression of the gene, reducing transcript copy number[63]. Mouse models of genetic ablation of *Ptpn22,* or knock-in models of its variant, along with models of T cell specific knockout of *Ptpn2* result in T cells that are hyper-proliferative and hyper-responsive to TCR stimulation, and hyper-responsive to cytokine stimulation in the case of *Ptpn2* knockout. Human examples of genetic loss-of-function of these genes have not been identified, but studies on primary human T cells from carriers of these SNPs reveal a phenotype contrasting with mouse models, with decreased TCR initiated signaling in *PTPN22* rs2476601 donors[56,85], and decreased response to IL-2 in *PTPN2* rs1893217 donors[63,100].

In this study we establish highly efficient gene editing approaches, and explore methods to characterize alterations in T cell signaling pathways that occur upon *PTPN22* or *PTPN2* disruption in T

cells from multiple primary human donor populations. We find that in T cells where either of these genes have been disrupted, T cell response to TCR stimulation or cytokine stimulation largely mimic previous findings in mouse models, showing increased responses to both. Interestingly, in the case of *PTPN2* disruption, the power of this approach allowed us to also find culturing conditions wherein hyper-responsiveness to IL-2 proved dynamic, and over time and exposure to high cytokine, enhanced IL-2 responses eventually give way to reduced IL-2 responses, mimicking findings in primary T cells from natural carriers of the *PTPN2* rs1893217 variant. This data demonstrates that while mouse models can provide a strong model for human genetics, some GWAS SNPs work to promote disease in a context dependent manner that may be difficult to model in a non-human setting.

Can the PTPN22 risk SNP be edited into the genome of non-risk human T cells?

Previous studies have sought to examine loss-of-function of PTPN22 in limited human models[49,89,90]. While such information is useful, the *PTPN22* rs2476601 risk variant has been shown to not alter expression levels of *PTPN22* but affect its binding dynamics[64,67], likely resulting in a modulation of function or localization. Many studies have shown that PTPN22 is involved in a variety of immune cell functions including regulation of TCR signaling, integrin expression, and type 1 interferon secretion[51,141,142]. Such involvement in multiple immunological pathways complicates resigning the *PTPN22* rs2476601 SNP's contribution to autoimmunity as simple loss-of-function or gain of function and obviates the need for accurate modeling of the rs2476601 risk SNP in primary human T cells.

To address these concerns, we use gene editing approaches to create *PTPN22* rs2476601 risk variant primary human CD4 T cells from non- risk donors. To do this we employ use of CRISPR/Cas9 RNPs and multiple options for the delivery of homologous DNA cassettes, which allow for the alteration of the *PTPN22* coding sequence in a trackable and efficient manner. We found that modification of the coding sequence in *PTPN22* is indeed possible with high efficiency, but no current editing approach is without potential confounders or drawbacks. Studies continue as to the functional consequence of this SNP in primary human T cells, but preliminary data would suggest that primary human T cells edited to

express the *PTPN22* rs2476601 risk variant display some overlapping features with current mouse KI models.

Chapter 2

Efficient CRISPR/Cas9 disruption of autoimmune-associated genes reveals key signaling programs in primary human T cells

Introduction

Genome-wide-association-studies have identified a subset of genetic risk variants that are shared broadly across multiple, distinct autoimmune diseases [5,6]. The shared risk of these variants suggest they impact key signaling pathways in a manner that promotes or sustains the loss of immune tolerance [36]. Identifying how perturbation of these pathways impacts autoimmunity is critical for both understanding the loss of tolerance and the development of therapeutic interventions. Notably, previous studies of some human risk variants have produced discordant data depending on the model used [143,144]. These discrepancies are likely due to a combination of factors including species specific differences between mouse and human tissues, activation state and underlying transcriptional profile in the case of cell lines, and differences in genetic background or environmental factors in cross-sectional studies using primary human cells. The complexity demonstrated by models of human genetic risk variants likely reflects context specific impacts on lymphocyte programming that makes translation of data from some models to primary human lymphocytes difficult [39,43,145]. Thus, a major challenge in understanding autoimmunity is to develop methods that accurately discern the impact of a candidate genetic variant on primary human cell function. Here, we focus on genes encoding two phosphatases with variants that are strongly associated with increased risk of multiple autoimmune diseases. These phosphatases participate in lymphocyte antigen and cytokine signaling pathways and their risk variants have shown contrasting functional results depending on the model systems used.

The gene, Protein Tyrosine Phosphatase Nonreceptor 22 (*PTPN22),* encodes for a key negative regulator of antigen receptor signaling in lymphocytes [74,79]. A single nucleotide polymorphism (SNP), rs2476601, in *PTPN22* is associated with >10 autoimmune diseases, including type 1 diabetes (T1D), rheumatoid arthritis (RA), and systemic lupus erythematous (SLE) [49,64,70], and impacts lymphocyte fate

and function [55,67,84,146]. Despite work from multiple groups, the functional impact of the risk variant remains controversial [50,72]. Mouse models of the *Ptpn22* risk variant develop autoimmune pathologies driven, in part, by dysregulation of antigen receptor signaling, leading to enhanced activation of T lymphocytes, increased IL-2 secretion, and enhanced calcium flux [55,67]. *Ptpn22* knockout mouse models exhibit a largely overlapping phenotype [84,146] and show improved clearance of LCMV infection due, at least in part, to a lower threshold for T cell activation [142,147,148]. Consistent with mouse knockout models, a previous study using siRNA knockdown of PTPN22 in human lymphocytes showed an increase in IL-2 production upon CD3/CD28 stimulation [89]. More recently, knockout *PTPN22* in the human Jurkat T cell line using CRISPR/Cas9 resulted in increased CD69 expression and IL-2 secretion in response to TCR engagement [90]. Loss-of-function in *PTPN22* has not been described in human subjects, but in contrast to risk variant knock-in mouse models, human carriers of the rs2476601 risk variant exhibit decreased TCR dependent downstream signaling [56,85]. Thus, the mechanism(s) by which PTPN22 regulates primary human T cell function remains unclear.

Protein Tyrosine Phosphatase Nonreceptor 2 (*PTPN2)* is an additional key modulator of T cell activation that functions primarily through modulation of JAK/STAT signaling [96]. A non-coding SNP within *PTPN2* (rs1893217) is associated with multiple autoimmune diseases including T1D and RA [91,92]. Loss-of-function of *PTPN2* in mouse tumor models and human cell lines have demonstrated increased pSTAT1 and pSTAT5 signaling in response to IFNγ and IL-2, respectively [104,149]. PTPN2 has also been suggested to inhibit TCR signaling in murine T cells, as T cell-specific *Ptpn2* disruption results in hyper-active TCR signaling, development of anti-nuclear antibodies, and CD8 T cell-mediated autoimmunity [102]. The *PTPN2* risk SNP causes an allele-dose dependent reduction in *PTPN2* mRNA transcripts in human lymphocytes [63]. However, again in contrast to murine and cell line data, memory T cells derived from healthy human carriers of the risk variant exhibit blunted pSTAT5 responses to IL-2 and IL-15 [63,100].

Together, the largely contradictory observations from murine and cell line models versus primary human T cells in these genes, illustrate an urgent need to find new methods to understand the functional impact of candidate autoimmune risk alleles. Specifically, new approaches are required to better

understand the alterations in T cell signaling triggered by variants in *PTPN22, PTPN2,* and other risk alleles in primary human lymphocytes.

Advances in gene editing of primary human hematopoietic cells provide a unique opportunity to address key questions regarding the effect of altered signaling programs in human primary lymphoid populations [125,150-152]. Genetic research has been transformed by the introduction of designer nucleases. Among nuclease platforms, CRISPR/Cas9 is unique in accessing its genomic target sites by a guide RNA (gRNA) sequence. Co-delivery of gRNA and Cas9 protein as ribonucleoprotein complexes (RNPs) efficiently facilitates DNA double stranded breaks at target sites [109,113]. DNA break repair via the non-homologous end-joining (NHEJ) pathway results in insertion or deletion of nucleotides leading to gene disruption. Alternatively, in the presence of a DNA donor template, the homology directed repair (HDR) pathway can be utilized for repair and/or modification of the coding sequences surrounding the DNA break [109,113,152].

While multiple studies have now utilized gene editing to generate new animal and cell models for the study of disease, work using CRISPR to study gene function in primary human T cells has shown promise [122,153,154] but remains relatively limited. Challenges to progress likely reflect both the perception that primary human T cells are difficult to edit in a consistent fashion and the requirement to rapidly assay edited populations in a functionally relevant manner. In this study, we established a robust platform to utilize gene editing to perform rapid, reproducible, and definitive analyses of gene edited primary human T cell populations. We utilized co-delivery of RNPs and short, single-stranded, deoxy-oligonucleotides (ssODNs) to efficiently introduce a stop codon within candidate genetic loci. This approach leveraged synergy gained by combining outcomes of both the NHEJ and HDR pathways, thereby permitting us to achieve highly-efficient gene disruption in populations of minimally manipulated, primary human T cells.

Using this optimized editing platform, we assessed the functional impact(s) of loss of expression of key candidate autoimmune-associated genes in an isogenic cell setting. Specifically, we achieved highly efficient gene disruption in the *ZAP70*, *PTPN22*, and *PTPN2* loci of primary human $CD4^+$ T cells. Using short-term expansion, followed by functional assays, our combined data demonstrate that loss-of-

function predominantly mimics mouse knockout models, with key and informative exceptions, demonstrating dynamic adaptation of signaling programs driven by immune experience.

Results

Loss of ZAP70 prevents TCR activation of human CD4$^+$ T cells.

Using modifications of methods developed by our laboratory for use in engineering of human primary B and T lymphoid cells [125,150-152], we designed a work-flow for the editing, expansion, subsequent rest, and stimulation of primary human CD4$^+$ T cells. We utilized CRISPR/Cas9 nucleases delivered as RNPs to facilitate gene editing of candidate signaling effectors (Fig. 2-1A). As an initial target for proof- of-principle study we elected to disrupt expression of a critical TCR signaling effector, *ZAP70*. ZAP70 is a non-receptor tyrosine kinase that functions as an immediate transducer of TCR-driven, protein phosphorylation, required to initiate down-stream signaling and transcriptional events. Both murine models and human patients with ZAP70 loss-of-function mutations fail to initiate multiple downstream signals in response to TCR engagement, resulting in severe impairment of normal T cell development [155,156]. We used *in silico* identification software to design four candidate gRNAs targeting exons 4 or 5 of *ZAP70*, regions required for gene expression. Following transfection of primary human CD4$^+$ T cells with these RNPs, a T7 nuclease assay was performed using PCR amplified genomic DNA. NHEJ mediated gene disruption was present with all gRNAs (Fig. 2-1B) and *ZAP70* G4 displayed the highest levels of cleavage activity.

For our approach, assessing the functional role of candidate genes in primary T cell populations would require performing studies using bulk-edited, isogenic test and control cell populations with uniform genetic changes. Importantly, this approach should minimize handling and time in culture to preserve functional relevance of the cells and responsiveness to subsequent TCR engagement. A critical requirement would be to achieve rapid and near complete gene disruption (~90%) of target genes in primary T cells from multiple independent donors. We reasoned that this goal would be challenging to achieve consistently with RNP delivery only. Therefore, using *ZAP70* G4, we explored using co-delivery

of a homology directed repair (HDR) template, introduced as short, single-stranded oligo-deoxynucleotides (ssODN). Previous work demonstrated that co-delivery of non-homologous DNA can promote CRISPR/Cas9 mediated NHEJ [157] and that co-delivery of homologous ssODNs can increase gene disruption rates in primary human $CD4^+$ T and B cells [125,127]. To apply this method for targeting *ZAP70*, we designed a 200bp "stop" ssODN for co-delivery with *ZAP70* G4 RNP. The ssODN was designed to introduce a stop codon through HDR in all potential reading frames using 91bp homology arms aligning to the cleavage site [125].

Consistent with the predicted increase in disruption rates, western blot analysis demonstrated >90% loss of ZAP70 protein expression with this approach. Co-delivery of G4 RNP and the stop ssODN resulted in a greater reduction in ZAP70 protein compared to cells transfected with RNP alone (Fig. 2-S1, A). Gene disruption in expanded, edited T cell populations was consistent across multiple independent donors and independent experiments (Fig. 2-1, C-D). Consistent with our protein analysis, use of stop ssODN increased allelic disruption by over 10% on average compared to RNP alone as determined by ICE analysis [158] (data not shown) and decreased variance among individual donors. As noted below, more dramatic enhancement of gene disruption rates were observed at other target loci using RNP and ssODN co-delivery.

In parallel with this approach, we established control T cell populations required for the comparative signaling and functional analyses of gene edited human T cells. We generated mock edited cells (activated, electroporated without RNP or ssODN, and cultured identically) from each donor. Importantly, editing the genome and exposure to ssDNA may have unintended effects on cell phenotype [132,138,139,159]. To account for these potential impacts, we also generated gRNAs and "stop" ssODNs targeting the human *CCR5* locus. *CCR5* encodes a chemokine receptor that, based upon individuals homozygous for a loss-of-function allele, is dispensable for T cell immune responses including responses to TCR engagement. Based on this premise, all subsequent experiments utilized isogenic mock and *CCR5* edited T cells as negative controls.

Similar to mock and *CCR5* edited control populations, *ZAP70* disruption exhibited minimal impact on cell viability (Fig. 2-S1, B-C). Loss of ZAP70 also did not impact cell expansion (Fig. 2-S1, D). As we

utilized cytokines for cell expansion, we introduced a 24-hour, cytokine free, rest period to reduce cell activation to a baseline "rested" state prior to the assessment of candidate activation signals. 24-hour withdrawal of cytokine lead to a significant reduction in basal activation as measured by surface expression of CD40L (Fig. 2-S1, E). Upon TCR stimulation with soluble anti-CD3, *ZAP70* edited cultures failed to initiate calcium flux (Fig. 2-1, E). This defect correlated with markedly reduced expression of cell activation markers including CD69 and CD25 in response to stimulation with plate bound anti-CD3 for 24-hours (Fig. 2-1, F-I-J; Fig. 2-S1, F). Together, these findings show that human $CD4^+$ T cells lacking ZAP70 are unable to respond to TCR engagement, directly replicating data from previous murine knockout studies and human T lymphocyte cell line models [155,156]. Importantly, our data demonstrate the establishment of a robust platform to achieve efficient, rapid gene disruption in bulk, human primary $CD4^+$ T cells leading to the generation of uniform gene-targeted and control, isogenic T cell populations.

PTPN22-deficient human $CD4^+$ T cells are hyper-responsive to TCR stimulation.

We next utilized our gene editing platform to determine the impact of *PTPN22* disruption in human $CD4^+$ T cells. After screening, we identified gRNAs that mediated partial reduction of PTPN22 in $CD4^+$ T cells (Fig. 2-2, A). *PTPN22* G2, targeting exon 2, exhibited the most robust editing rates. *PTPN22* G14 was designed to target exon 14 adjacent to the rs2476601 SNP. In parallel, we generated "stop" ssODNs to facilitate gene disruption. Co-delivery gRNAs and relevant stop ssODNs significantly promoted gene disruption and reduced PTPN22 expression (Fig. 2-2, B-C). The efficacy of G14 was most dramatically increased by ssODN co-delivery.

To accurately assess the contribution that ssODN delivery made to overall editing rates, we established a droplet-digital based PCR (ddPCR) assay to quantify both HDR and NHEJ in T cells. We assessed both *PTPN22* G14 RNP and *CCR5* RNP, with or without co-delivery of the relevant stop ssODN (Fig. 2-2, C; Fig. 2-S2, A). HDR (stop codon insertion) made a substantial contribution to overall editing rates (for example, 24.7% HDR using *PTPN22* G14 RNP and ssODN co-delivery). Inclusion of ssODNs also led to an increase in NHEJ events for both genes (45.5% NHEJ using *PTPN22* G14 RNP and

ssODN delivery compared to 27.8% with RNP alone; Fig. 2-2, C; Fig. 2-S2, A). Edited and mock edited cells exhibited a minimal, but significant, impact in cell viability at Day 2 (Fig. 2-S2, B), likely due to the impact of electroporation. Viability was equivalent across all groups by Day 7 (data not shown).

After editing, expansion, and cell rest as shown in Figure 1, *PTPN22* edited cells were activated with soluble or plate-bound anti-CD3. PTPN22 deficient T cells manifested an increase in overall calcium flux that was inversely proportional to the level of *PTPN22* disruption (Fig. 2-2, D-E). Upon stimulation with plate bound anti-CD3, PTPN22 deficient cells also exhibited increased secretion of IFNγ and increased cell size (Fig. 2-2, F; Fig. 2-S2, D). After TCR stimulation, PTPN22 deficient cells also exhibited increased expression of the activation marker, CD69, and of the activation/exhaustion marker, PD-1, consistent with stronger levels of TCR engagement (Fig. 2-3, A-C; Fig. 2-S2, D-E). Expression of additional activation markers, CD71 and CD25, were also increased in PTPN22 deficient cells (Fig. 2-3, D-F), and, similar to calcium flux findings, increased expression of these markers correlated inversely with PTPN22 expression.

To corroborate our findings using an alternative editing strategy, we next established $CD4^+$ populations that were identifiable by flow-cytometry as bi-allelically edited for *PTPN22* loss-of-function. Our laboratory and others have previously demonstrated efficient HDR-based gene editing of primary human T or B cells using co-delivery of designer nucleases and recombinant AAV vectors engineered to deliver a homology donor cassette [125,126,150-152]. This approach can be used to efficiently disrupt coding sequences of a target gene following HDR via introduction of a gene cassette encoding a heterologous promoter driving expression of a fluorochrome reporter. We designed rAAV6 vectors that contained an MND promoter driving expression of either a GFP or BFP reporter and a 3' WPRE element to promote efficient mRNA export (Fig. 2-4, A). Each donor cassette contained 1kB homology arms adjacent to the cleavage sites for *PTPN22* G14 or the control *CCR5* gRNA, respectively. As in Figure 1, $CD4^+$ T cells were isolated, activated and electroporated with RNPs and, in this case, simultaneously transduced with the GFP and BFP bearing rAAV6 donors. After 7 days of expansion, edited cells exhibited distinct populations including: no HDR (BFP/GFP double negative cells), HDR with expression of 1 fluorochrome (BFP or GFP single positive cells), or bi-allelic HDR with expression of both reporters (BFP/GFP double

positive cells; Fig. 2-4, B-C). HDR-targeted AAV donor integration into the *PTPN22* locus was confirmed by PCR amplification in sorted populations (Fig. 2-S3, A). The presence of AAV mediated HDR correlated with decreased levels of PTPN22 protein expression, with the lowest PTPN22 expression observed in GFP+/BFP+ populations (Fig. 2-4, D-E). This editing approach led to a modest impact on cell viability within rAAV6 transduced cultures (Fig. 2-S3, B-E).

Again, after expansion and rest, edited populations were stimulated with plate bound anti-CD3. Consistent with our results using RNP and ssODN delivery, *PTPN22* GFP+/BFP+ cells exhibited an increase in CD25, CD71, and PD-1 expression compared to control, dual positive, *CCR5* edited cells (Fig. 2-4, F-G). *PTPN22* GFP+/BFP+ cells also exhibited a modest, but significant, increase in phospho-CD3ζ compared to control populations in response to low-dose anti-CD3 stimulation, findings consistent with reduced negative regulation of proximal TCR signaling (Fig. 2-S4, A-B). Taken together, this data demonstrates that loss of *PTPN22* expression in primary human $CD4^+$ T cells leads to hyper-active TCR signaling, consistent with a key role for the phosphatase in tuning proximal signal strength in response to TCR engagement. These findings directly correlate with studies in *Ptpn22* deficient primary murine T cells and loss-of-function studies in human T cell lines [55,67,90].

Loss of PTPN2 alters regulation of IL-2 signaling in human $CD4^+$ T cells.

Enhanced T cell signaling has been proposed to mediate autoimmune pathology in the setting of PTPN2 loss-of-function [102], yet memory $CD4^+$ T cells from human subjects with the *PTPN2* rs1893217 risk SNP display reduced, rather than enhanced, responsiveness to IL-2 stimulation [63,100]. As the PTPN2 rs1893217 variant is believed to act through reduced expression of PTPN2, we hypothesized that *PTPN2* disruption in primary human $CD4^+$ T cells would be a valid approach to investigate the functional impact of PTPN2 modulation in isogenic T cell populations and to limit the impact of previous immune experience. We designed gRNAs targeting exons 2-3 of *PTPN2* (Fig. 2-5, A), sequences contained within both functional isoforms of PTPN2. Three gRNAs mediated significant NHEJ levels in $CD4^+$ T cells (Fig. 2-S5, A). Notably, the human genome also encodes for two homologous, *PTPN2* pseudo-genes on separate

chromosomes. *In silico* predictions suggested that *PTPN2* G2 exhibited the least likelihood for off-target editing of these pseudogenes. Therefore, we designed a stop ssODN for co-delivery with *PTPN2* G2 RNP. Using a stop ssODN as opposed to RNP alone, we observed significantly increased allelic disruption as determined by ICE sequencing analysis (Fig. 2-S5, B) resulting in a significant reduction of PTPN2 protein expression based on western blot analysis (Fig. 2-S5, C). Co-delivery of *PTPN2* G2 with ssODN resulted in greater than 80% reduction in PTPN2 protein levels across multiple donors (Fig. 2-5, B-C). Viability was not significantly impacted 2 days post-editing (Fig. 2-S5, E), and slightly reduced at 7 days post-editing (Fig. 2-S5, F).

Based on previous work in murine models, we predicted that loss of PTPN2 might lead to increased sensitivity to IL-2, potentially altering the phenotype of cells after expansion in IL-2 supplemented media. To account for this potential impact, we altered our work-flow (Fig. 2-5, D), decreasing the IL-2 concentration used for initial $CD4^+$ T cell activation by 100-fold (0.5ng/ml) prior to editing. Immediately post-editing, $CD4^+$ T cells were either cultured in media without IL-2 for 2 days, or, alternatively, placed into IL-2 only supplemented media (50 ng/ml) and expanded for 7 days as in Figures 1 and 2.

Two days post-editing, in media supplemented with IL-2, surface CD25 expression was increased in *PTPN2* disrupted cells compared to control populations. In contrast, this change was not present in cells cultured in the absence of IL-2 (Fig. 2-S5, D). Assays performed using rested, edited cells were also consistent with an increased responsiveness to IL-2. Following the 2-day rest without supplemental IL-2, edited and control T cells were stimulated with IL-2 for 20 minutes and assayed for pSTAT5 by flow cytometry. *PTPN2* edited cells exhibited enhanced pSTAT5 levels by increased median fluorescence (Fig. 2-5, E-F; Fig. 2-S6, A). These combined findings indicated that *PTPN2* disruption promote enhanced responsiveness to IL-2 in human $CD4^+$ T cells.

In contrast to this assay using short-term cultured cells, a 7-day expansion in IL-2 supplemented media, with subsequent rest and stimulation with IL-2 showed *PTPN2* and CCR5 edited populations to exhibit equivalent IL-2 responsiveness (Fig. 2-S6, B-C). Together, these data suggest that *PTPN2*

disruption promotes IL-2 responsiveness and that sustained IL-2 signaling down-modulates this signaling program.

We also assessed the response of 7-day expanded *PTPN2* edited cells to additional signals including TCR engagement (using plate bound anti-CD3 for 24 hours) or inflammatory cytokines (using a 20-minute exposure to IFNγ). *PTPN2* edited cells exhibited increased expression of PD-1, CD69, CD25, and CD71 in response to anti-CD3 cross-linking (Fig. 2-6, A-F; Fig. 2-S6, D-G). *PTPN2* edited cells also showed an increased response to IFNγ as assessed by measurement of STAT1 phosphorylation (Fig. 2-6, G-I).

SOCS3 is upregulated in primary T cells lacking PTPN2 or expressing the rs1893217 risk allele

The observations in *PTPN2* edited human $CD4^+$ T cells largely replicated data from murine models of *Ptpn2* deficiency that identified a key inhibitory role for this phosphatase in multiple T cell signaling cascades. Considering these observations, it became important to better understand the potential mechanism(s) responsible for loss of enhanced pSTAT5 signaling following the expansion of edited cells in IL-2 media. We hypothesized that sustained hyper-active IL-2 signaling leads to increased activity of an alternate regulatory pathway that compensates for cytokine signals; a program presumably sensitized by loss of PTPN2 expression. To test this hypothesis, we returned to our original work-flow using a seven-day, cytokine-mediated expansion of edited T cells. For this work, we utilized supplemental IL-2 as well as supplemental IL-7 and 15, cytokines critical for maintenance of memory T cell populations (Fig. 2-7, A). Following expansion in multi-cytokine media, T cell populations were rested and stimulated with IL-2 for 20 minutes. This protocol resulted in a slight reduction in viability of *PTPN2* edited cells relative to control cells (Fig. 2-S7, A-B). Under these conditions *PTPN2* edited T cells exhibited a significant reduction in the percentage of cells responding to IL-2 stimulation measured by STAT5 phosphorylation (Fig. 2-7, B-C).

To determine a possible mechanism responsible for the reversal of IL-2 responsiveness, we isolated mRNA from edited cells and quantified expression of potential negative regulators of IL-2

signaling. As expected, transcript levels for PTPN2 were strongly reduced in *PTPN2* edited cells, but not control populations (Fig. 2-S7, C). Strikingly, we observed a marked increase in expression of SOCS3 (Fig. 2-7, D), a key suppressor of multiple cytokine signaling pathways. No differences were observed in expression of other candidate negative regulators including SOCS1, PTPN11, or in relative levels IL-2Rβ expression (Fig. 2-S7, D-F).

As noted above, previous studies of T cell subsets from healthy human subjects with the *PTPN2* rs1893217 risk allele demonstrated reduced responsiveness to IL-2 stimulation, data analogous to the reduced responses of *PTPN2* edited cells cultured in IL-2, IL-7, and IL-15. Therefore, we next directly assessed whether overexpression of SOCS3 correlated with the diminished pSTAT5 response in ex-vivo memory T cells from human carriers of the rs1893217 risk SNP. PBMCs were obtained from non-risk and heterozygous risk human subjects. After isolating $CD45RO^{+}CD25^{-}CD4^{+}$ effector T cells through negative isolation, T cells were stimulated with IL-2 for 20 minutes and assayed for pSTAT5 levels by flow cytometry. Consistent with the observations in a previous study [63], T cells from risk subjects exhibited a significantly reduced percentage of pSTAT5 positive cells in response to IL-2 stimulation (Fig. 2-7, E). Next, we isolated mRNA from unstimulated $CD4^{+}$ effector T cells from these donors and assessed SOCS3 expression by RT-qPCR analysis. Consistent with our findings in *PTPN2* edited T cells, rs1893217 risk subjects expressed significantly higher levels of SOCS3 compared to non-risk donors (Fig. 2-7, F).

Discussion

In this study, we established a robust gene editing platform to rapidly address the functional consequences of loss of expression of candidate autoimmune-associated genes in primary human $CD4^{+}$ T cells. Our editing platform capitalized upon a synergy gained by using a repair ssDNA template to harness both the HDR and NHEJ repair pathways upon DNA disruption. This synergy lead to enhanced rates of gene disruption across all target loci, with minimal toxicity and consistent findings, in primary $CD4^{+}$ T cells isolated from multiple independent healthy donors. A range of previous approaches have

been utilized to enhance or select for CRISPR/Cas9 mediated gene disruption in primary human cells. These methods have included screening optimal gRNAs or editing conditions [160], co-delivery of nuclease and recombinant AAV6 homology donor templates designed to disrupt the target locus [126,152], co-transfection of gene editing "helper" proteins [161], and introduction of viral-based selection cassettes [122]. Our approach reduced the time and labor required to produce optimal editing rates relative to previous approaches. High efficiency, single-step gene editing using commercially available RNPs and chemically modified "stop" ssODNs also eliminated a requirement for viral vectors and/or selection protocols to enrich for edited cells, expediting the analysis pipeline. Importantly, following gene editing, expansion, and short-term cell rest, edited populations remained responsive to both TCR and cytokine stimulation; allowing detailed functional assessment without additional steps, including no requirement for sub-cloning, long-term culture and/or sorting of edited cell populations. Most critically, the capacity to rapidly assess functional activity in minimally manipulated, T cell populations reduced the confounding impacts of differentiation and regulatory signals that likely become operative in the context of longer-term, less efficient editing methods. These work-flow assets allowed the discovery of a key regulatory network controlling reduced response to cytokine in PTPN2 ablated CD4 T cells, which is likely an indirect as opposed to direct consequence of gene ablation.

Use of a control gene editing locus (*CCR5*) was important to this study. CCR5 is not required for TCR or cytokine signaling. Consistent with this concept, *CCR5* has been utilized as a "safe-harbor" locus for various gene therapy approaches. Of note, introduction of CRISPR gRNA and/or ssDNA can trigger innate signaling cascades and/or exert other stimulatory effects on primary human cells [132,159]. In addition, introduction of DNA double stranded breaks triggers p53-dependent inhibitory responses in many primary cell types [138,139]. Thus, to control for these potentially confounding effects, we generated equivalently edited, *CCR5* disrupted isogenic control T cell populations. As expected, delivery gRNAs with or without ssODNs, led to modest functional and phenotypic changes compared to isogenic, non-edited T cells. By using the *CCR5* control, we accounted for these non-specific impacts in primary T cells.

In parallel with our studies using RNP and ssODN co-delivery, we performed studies using RNP and rAAV6 co-delivery. AAV homology donor constructs were designed to track gene disruption using

dual expression of cis-linked fluorescent markers. Gating on dual-edited T cell populations allowed assessment of candidate gene disruption. Functional studies of *PTPN22* using this approach lead to findings equivalent to results using co-delivery of RNP and ssODNs. Comparison of the response in *PTPN22* versus *CCR5* disrupted $CD4^+$ T cells from multiple donors demonstrated enhanced cell activation in PTPN22-disrupted cells. The observed alterations to TCR-induced cell activation following AAV HDR editing were, however, more variable than data obtained using co-delivery of RNP and ssODNs. This variability likely reflects additional cellular impacts of rAAV on cell differentiation and phenotype in dual-edited cell populations that may partially obscure the functional consequences of PTPN22 deletion.

Our data of ZAP70, PTPN22, and PTPN2 loss in primary human $CD4^+$ T cells largely reflect previous findings in mouse knock-out strains and in cell line models, with important exceptions. As predicted, ZAP70 deficient, $CD4^+$ T cells were unable to respond to TCR engagement, exhibiting no increase in surface activation markers CD69 or CD25 and absent calcium flux upon CD3 stimulation. In contrast, PTPN22 disrupted T cells were hyper-responsive to TCR engagement, exhibiting enhanced calcium flux as well as increased expression of CD69, CD25, CD71, and PD-1. These combined observations were consistent with findings in murine *Zap70* [156] and *Ptpn22* knockout models [55,67,84,146], respectively. While PTPN22 disrupted T cells also exhibited increased secretion of IFNγ, IL-2 secretion was not impacted (data not shown), in contrast to previous reports using *PTPN22* targeted human T cell models [89,90]. These differences may reflect time in culture or exposure to IL-2 during expansion prior to TCR stimulation. Overall, the gene editing platform described here faithfully reproduced data obtained using gene disruption in mouse genetic studies as well as studies performed in transformed human T cell lines.

Our findings provide the first demonstration of a key role for *PTPN2* in both the TCR and cytokine receptor signaling cascades in primary human $CD4^+$ T cells. Gene disruption led to increased responsiveness to TCR engagement as demonstrated by increased activation marker expression. PTPN2 disrupted human $CD4^+$ T cells were also hyper-responsive to both IL-2 and IFNγ as demonstrated by enhanced phosphorylation of STAT5 and STAT1, respectively. These combined findings directly mirror

previous models of *Ptpn2* deficiency that identified alterations in proximal TCR and cytokine signaling programs [102,104,149].

In contrast to the above findings, culture of PTPN2 deficient human $CD4^+$ T cells in media supplemented with IL-2-family cytokines lead to paradoxical changes in cell responsiveness. In this setting, PTPN2 deficient T cells exhibited a reduction in the response to IL-2, as compared to PTPN2 competent cells. Loss of IL-2 reactivity correlated with increased expression of the inhibitory adapter protein, SOCS3. Importantly, our findings in PTPN2 disrupted, isogenic primary human T cells are mirrored in human carriers of the PTPN2 autoimmune risk SNP, rs1893217. T cells from carriers of this SNP, that leads to reduced expression of PTPN2 [63], also expressed increased levels of SOCS3 relative to non-risk donors. Consistent with this finding and with previous work [63,100], rs1893217 SNP carrier T cells displayed reduced IL-2 responsiveness. Our findings do not directly link increased SOCS3 levels to blunting of IL-2 responsiveness in PTPN2 deficient human T cells. However, our combined data demonstrate that altered expression of this key negative regulator of cytokine signaling co-occurs with a decreased response to IL-2. Together, our data regarding PTPN2 loss in primary human T cells support a model of enhanced and reduced responses to various stimuli that is context dependent and help to resolve existing discrepancies between human and murine data regarding loss of PTPN2.

The seemingly divergent and context specific phenotype of PTPN2 deficiency may help explain how perturbations in a single gene can predispose rs1893217 risk SNP carriers to a range of autoimmune diseases that may arise through functionally distinct mechanisms. Mouse models with T cell specific *Ptpn2* deletion have shown increased TCR signaling and acquisition of autoimmune pathology [102] due to enhanced T cell activity. Therefore, enhanced responses to TCR and cytokine stimuli may alter the TCR repertoire and/or promote enhanced effector function, increasing self-reactivity. Our results show that loss of PTPN2 in human T cells can produce analogous alterations in TCR signaling as reported in mice with lineage-specific *Ptpn2* deletion. Interestingly, recent work has demonstrated that *Ptpn2* deficient T cells promote enhanced rates of arthritis in SKG mice due to the conversion of Tregs to pathogenic Th17 cells [162]. In support of this finding, PTPN2 disrupted human T cells exhibit a compensatory increase in SOCS3 expression, and overexpression of SOCS3 in human Treg cells has been shown to impair growth,

suppressive function, and maintenance of FoxP3 expression [163]. Our data would thus support a model wherein PTPN2 deficiency contextually impacts Treg function. Therefore, through hyper- and hypo-responsiveness to alternative stimuli present in distinct immune settings, disruption to PTPN2 may facilitate at least two pathways that contribute to the development of autoimmunity.

Given our findings with *PTPN2,* it is conceivable that a similar context specific impact may be operative for the *PTPN22* risk variant, rs2476601. Recent work using CRISPR/Cas9 to knockout PTPN22 in Jurkat T cells [90] resulted in an increase in TCR signaling; findings consistent with data in *Ptpn22* deficient murine T cells [55,67,84,146]. The authors concluded that PTPN22 function is likely conserved between mice and humans. Our data in PTPN22 disrupted primary human $CD4^+$ T cells, while not identical, strongly supports this conclusion. As mouse models of *Ptpn22* knock-out and *Ptpn22* risk variant knock-in each demonstrate a loss-of-function phenotype [55,67], we would hypothesize that primary human T cells engineered to express the risk variant under endogenous locus control are likely to phenocopy *PTPN22* disrupted human T cells and exhibit increased TCR signaling. Of note, however, T cells from human carriers of the rs2476601 variant reproducibly display a reduction in TCR responsiveness [56,85]. Thus, by extrapolation of our experience with *PTPN2,* alternative regulatory pathway(s) active in the absence of negative regulation by PTPN22 are predicted to mediate transition into a hypo-responsive phenotype. Future work using HDR gene editing will be required to fully address how the *PTPN22* risk variant impacts human T cell function under alternative activation conditions.

In summary, the gene editing platform described in this study provides a robust capacity for uniform and rapid analysis of cell signaling following candidate gene disruption in primary human T cells. Taken together, our observations for *ZAP70*, *PTPN22*, and *PTPN2* show that loss-of-function mimics data obtained from previously established mouse genetic models and human cell lines. Further, our analysis of *PTPN2* gene disrupted T cells demonstrates dynamic effects whereby hyper-active IL-2R signaling mediates compensatory transcriptional events that modulate subsequent signaling responses. We postulate that, over time, these distinct impacts in signaling programs promote a cascade of cell intrinsic events that promote autoimmune risk, a hypothesis that correlates with observations of human subjects with the *PTPN2* risk variant. Given the broad availability of the tools and optimized methods described

here, our approach should be rapidly translatable to assess loss-of function studies in genetic targets across a broad range of primary human cell populations.

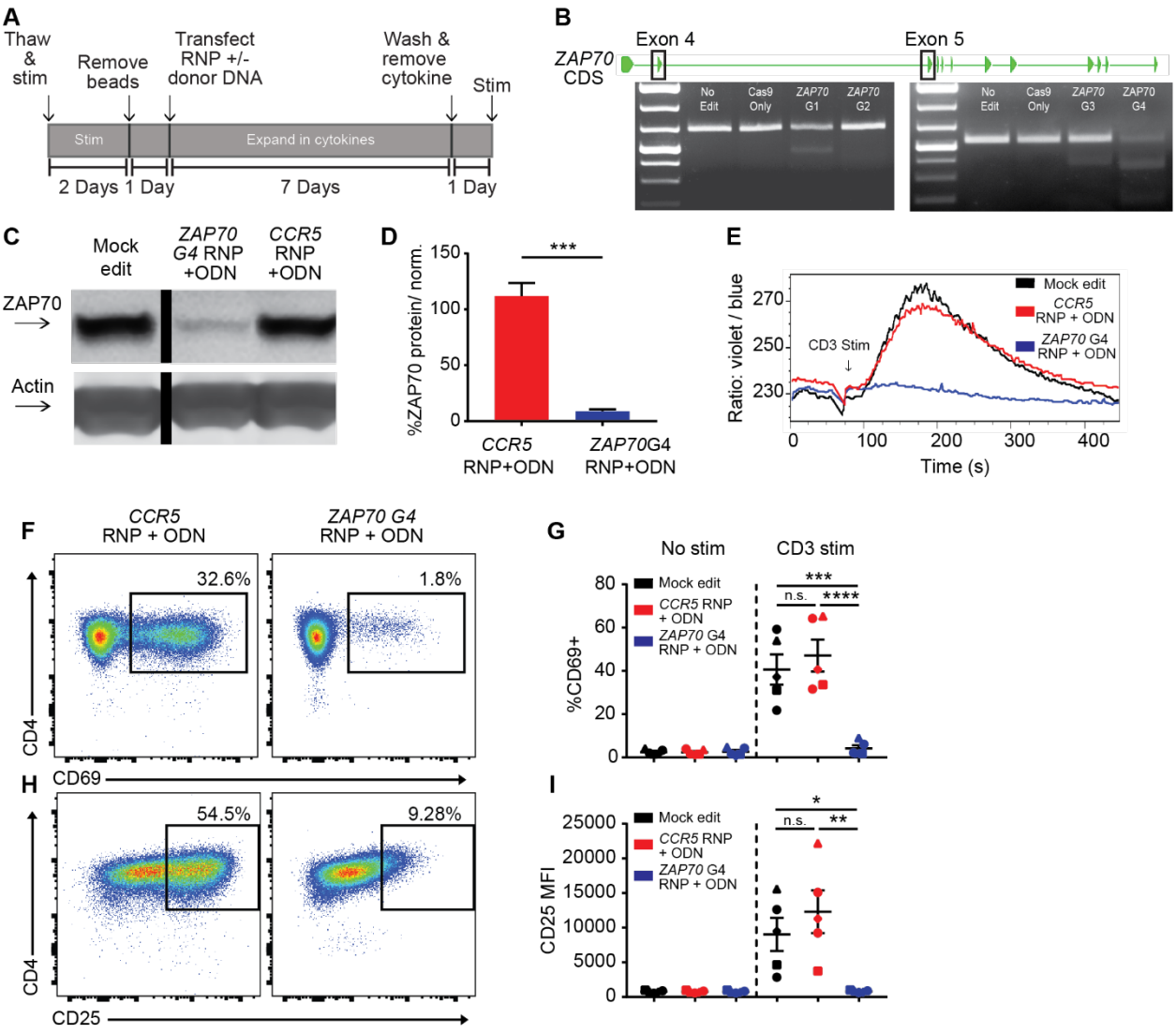

Figure 2-1. *ZAP70* disruption in primary human CD4$^+$ T cells abrogates TCR-mediated activation. (**A**) Editing protocol used to generate and assay *ZAP70* edited CD4$^+$ T cells and control T cell populations. (**B**) *ZAP70* coding exons and representative T7 assays showing RNP cleavage. *ZAP70* G1 & G2 target exon 4 and *ZAP70* G3 & G4 target exon 5. (**C**) Representative western blot of ZAP70 expression in mock, *ZAP70*, and *CCR5* edited CD4$^+$ T cells from originating from the same human donor. Lanes were run on the same gel but were noncontiguous. (**D**) Quantified ZAP70 protein expression relative to actin and normalized to mock edited values from the same T cell donor (bars represent mean +/- SEM, n=5 human samples (4 independent donors plus 1 repeat donor, repeat donors were run in separate experiments), paired t test). (**E**) Representative TCR-induced calcium flux of human CD4$^+$ T cells generated as in (**A**). Cells were stained with indo-1 AM, monitored for baseline then stimulated with anti-CD3 (arrow). (**F-I**) Human CD4$^+$ T cells edited as in (**A**) and stimmed with plate bound anti-CD3 for 24 hours. Representative flow plots of CD69 (**F**) and CD25 (**H**) in *ZAP70* and *CCR5* edited cells from the same donor. (**G&I**) Summary flow data for CD69 and CD25 expression +/- 24-hour anti-CD3 stimulation (n=5, analysis of stimmed cells only: matched one-way ANOVA with Tukey's correction). Summary graphs lines and error bars represent mean +/- SEM and shapes correspond to individual donors. RNP- ribonucleoprotein, ODN or ssODN - single stranded oligo-deoxynucleotide. All data are from 2 independent experiments. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001.

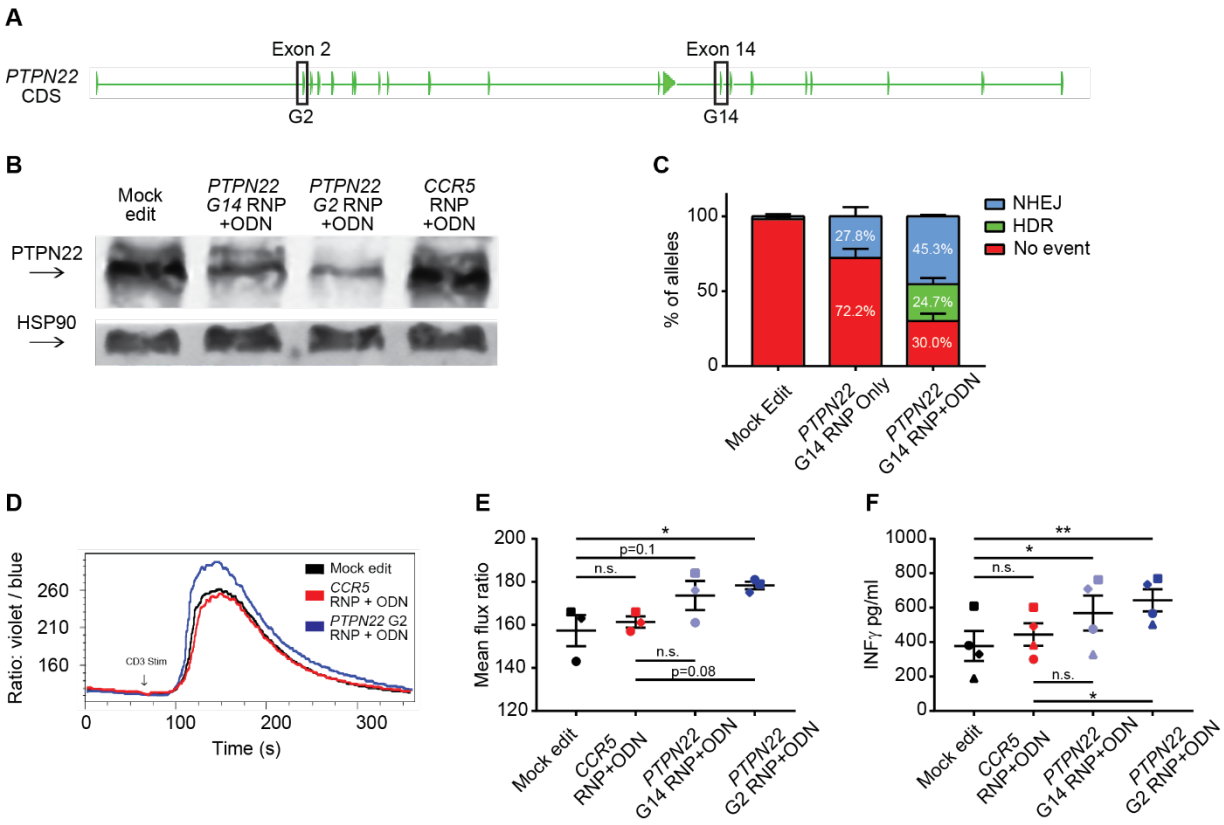

Figure 2-2. Disruption of *PTPN22* in CD4$^+$ cells results in increased TCR-triggered calcium flux. (**A**) *PTPN22* coding exons with targeted exons highlighted. (**B**) Representative western blot of PTPN22 expression in mock, *PTPN22* G14, *PTPN22* G2, and CCR5 edited CD4$^+$ T cells from the same human donor. (**C**) ddPCR analysis of editing frequencies in unedited and *PTPN22* G14 edited cells +/- stop codon containing ssODN (bars represent mean +/- SEM, n=4 independent human donors, percentages reflect summary data). (**D**) Representative TCR-induced calcium flux of human CD4$^+$ T cells generated, expanded, and rested as in Fig.1A. Cells were stained with indo-1 AM, monitored for baseline, then stimulated with anti-CD3 (arrow). (**E**) Summary of mean flux ratios for data generated as in (**D**) (graphs lines and error bars represent mean +/- SEM, n=3, matched one-way ANOVA with Tukey's correction). (**F**) IFNγ ELISA using supernatants from edited cells that had been stimulated with plate bound anti-CD3 for 48 hours. RNP – ribonucleoprotein, ODN or ssODN - single stranded oligo-deoxynucleotide, NHEJ - non-homologous end joining, HDR - homology directed repair. Shapes in summary plots correspond to individual donors. All data is from at least 2 independent experiments. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001.

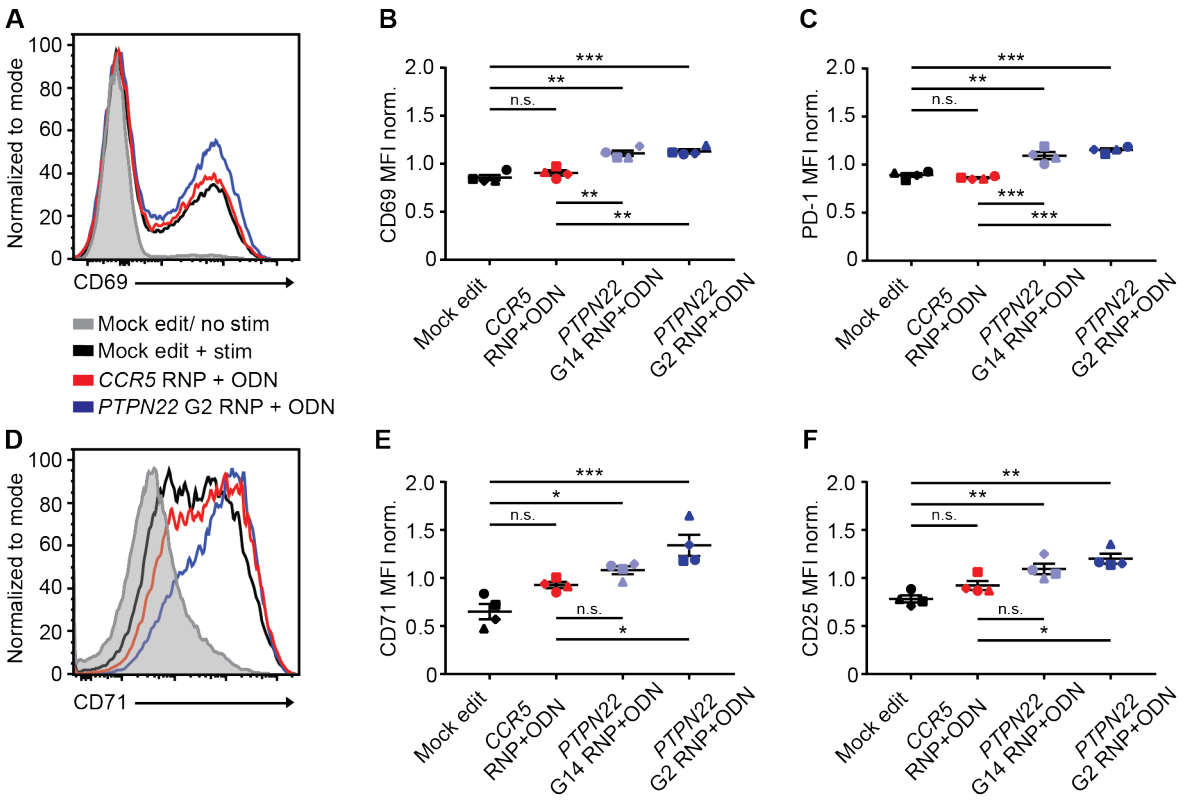

Figure 2-3. *PTPN22* disruption in CD4$^+$ cells increases cell activation in response to TCR stimulation. CD4$^+$ T cells from human donors were either mock edited, or edited with *PTPN22* G14, *PTPN22* G2, or *CCR5* RNPs and corresponding ssODNs containing stop codons. Cells were expanded and rested as in Fig.1A and subsequently stimulated with plate bound anti-CD3 for 24 hours. (**A**) Representative flow-cytometry histogram overlay of CD69 expression in different editing conditions from the same donor. (**B-C**) Summary data of median flow values for CD69 (**B**) and PD-1 (**C**) for all editing conditions following 24-hour anti-CD3 stimulation. (**D**) Representative flow overlay of CD71 expression in different editing conditions from the same donor. (**E-F**) Summary data of median flow values for CD71 (**E**) and CD25 (**F**) for all editing conditions following 24-hour anti-CD3 stimulation. Statistical analysis for all summary data (n=4) utilized a matched one-way ANOVA with Tukey's correction. Data is normalized to average MFI for all genotypes from the individual donor. Lines and error bars represent mean +/- SEM. RNP- ribonucleoprotein, ODN or ssODN - single stranded oligo-deoxynucleotide. Shapes in summary plots correspond to individual donors. All data is from 2 independent experiments. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001.

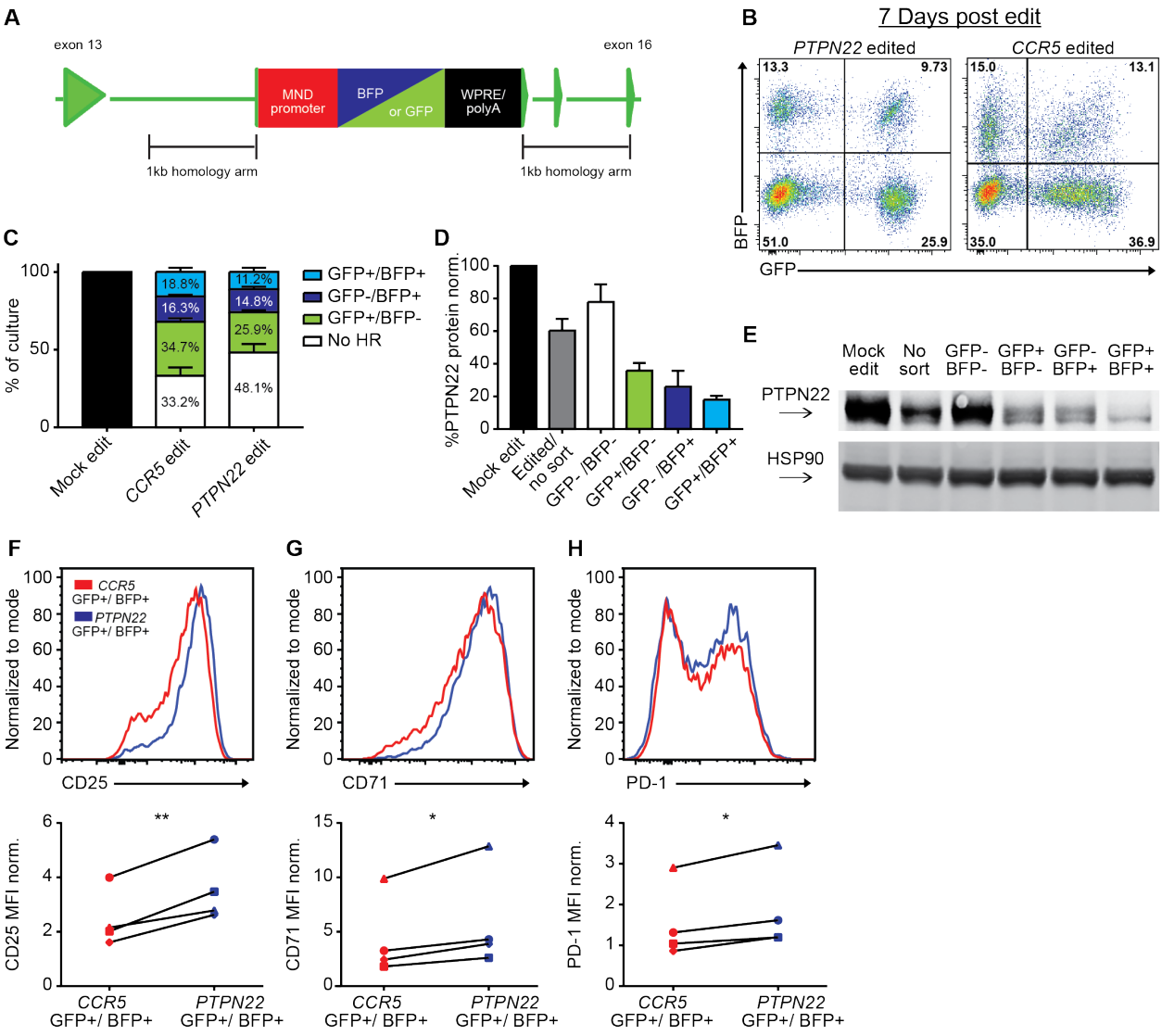

Figure 2-4. Bi-allelic disruption of *PTPN22* in CD4$^+$ cells using rAAV6 donor delivery leads to increased T cell activation in response to TCR engagement. (**A**) *PTPN22* disrupting AAVs within the *PTPN22* locus after homology directed repair. Promoter – (GFP or BFP) reporter – WPRE/polyA constructs disrupt exon 14 of *PTPN22*. Identical constructs with homology arms that align to the *CCR5* cut site used to make *CCR5* control populations. (**B**) Representative flow plots of *PTPN22* and *CCR5* AAV edited CD4$^+$ T cells from the same human donor at Day 7 post-editing. (**C**) Editing outcomes for *PTPN22* or *CCR5* AAV-edited T cells (bars represent mean +/- SEM, n=4 independent human donors, percentages reflect summary data). (**D**) Quantified PTPN22 protein expression relative to HSP90 and normalized to mock edited values from the same T cell donor (bars represent mean +/- SEM, n=4 independent human donors). (**E**) Representative western blot of PTPN22 expression in sorted *PTPN22* AAV-edited CD4$^+$ T cells from the same human donor. (**F-G**) Human CD4$^+$ T cells were edited as in (**A**) and stimulated using plate bound anti-CD3 for 48 hours. Upper panels: Representative flow overlays of CD25 (**F**) CD71 (**G**) and PD-1 (**H**) in GFP+/BFP+ *PTPN22* and *CCR5* AAV-edited cells from the same donor. Lower panels: Summary data below represent median fluorescence of GFP+/BFP+ cells normalized to the median fluorescence of stimulated, mock edited CD4$^+$ T cells from the same donor (n=4, paired t test). Lines illustrate data derived from each individual donor. RNP- ribonucleoprotein. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001.

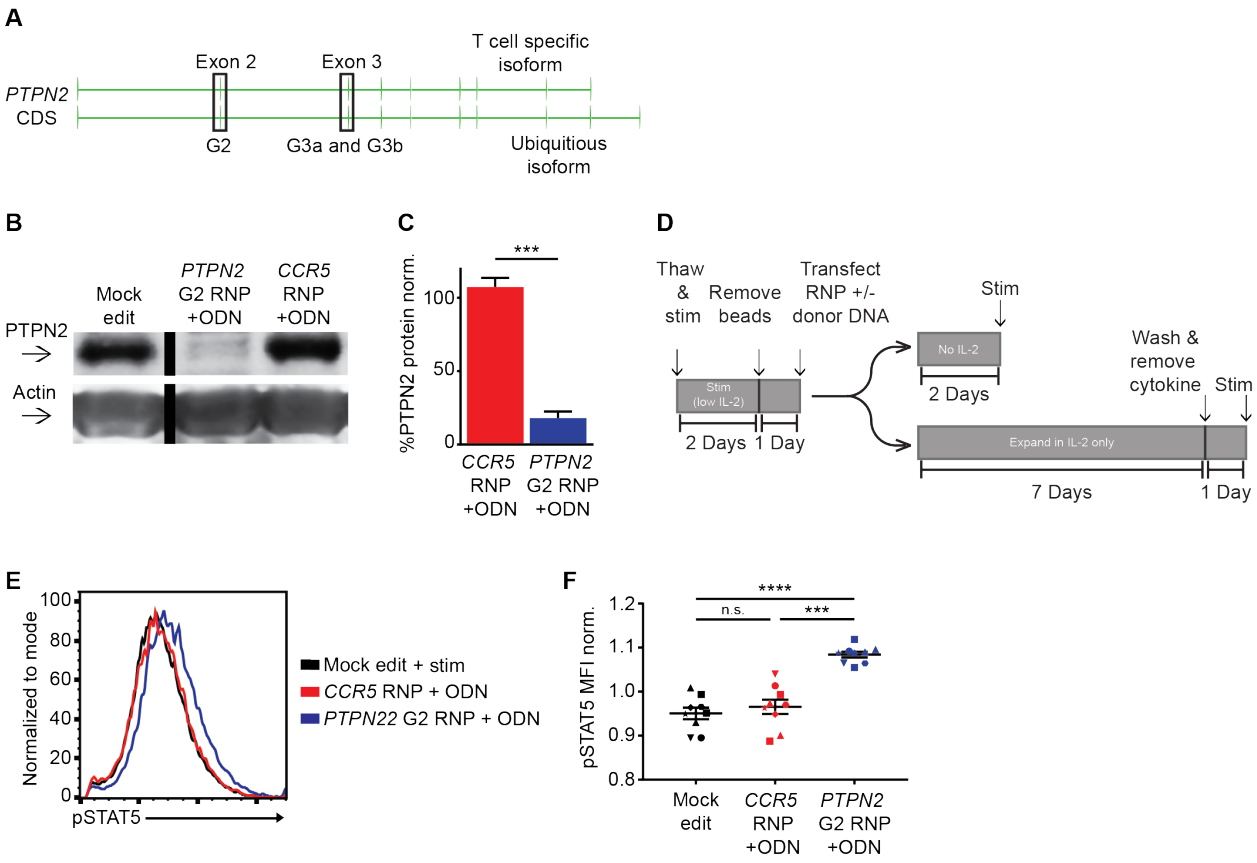

Figure 2-5. *PTPN2* disruption in CD4$^+$ cells promotes increased IL-2 signaling. (**A**) *PTPN2* gene showing structure and exons used for alternative isoform expression and location of guide RNA target sites (in highlighted exons). (**B**) Representative western blot of PTPN2 expression in mock, *PTPN2*, and *CCR5* edited CD4$^+$ T cells from the same human donor. Lanes were run on the same gel but were non-contiguous. (**C**) Quantified PTPN2 protein expression relative to actin and normalized to mock edited values from the same T cell donor (bars represent mean +/- SEM, n=6 independent human donors, paired t test). (**D**) Work-flow used to produce and assay *PTPN2* edited CD4$^+$ T cells and corresponding controls with or without IL-2 supplemented media. (**E-F**) Human CD4$^+$ T cells edited as in (**A**) and rested for 2 days in cytokine free media, were stimulated with IL-2 for 20 minutes. (**E**) Representative histogram overlay of pSTAT5 in mock, *PTPN2*, and *CCR5* edited cells from the same donor. (**F**) Summary flow data of median pSTAT5 expression post IL-2 stimulation (graph lines and error bars represent mean +/- SEM, n=9 human samples (7 independent donors plus 2 repeat donors, repeat donors were run in separate experiments), matched one-way ANOVA with Tukey's correction). RNP - ribonucleoprotein, ODN or ssODN - single stranded oligo-deoxynucleotide. Shapes in summary plots correspond to individual donors. All data is from 2 or 3 independent experiments. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001.

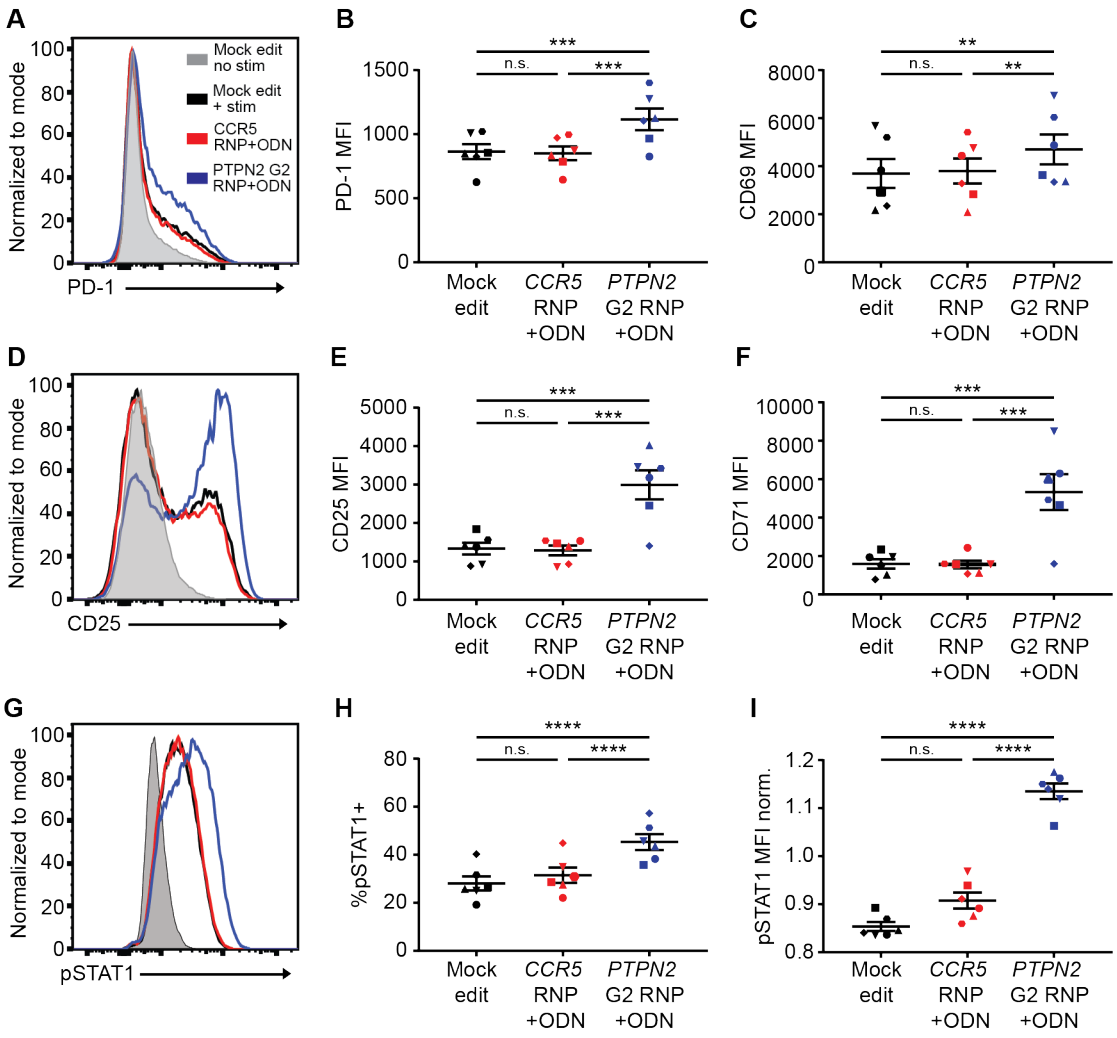

Figure 2-6. *PTPN2* disruption in CD4$^+$ cells promotes increased TCR and IFNγ signaling. CD4$^+$ T cells from human donors were either mock edited or edited with *PTPN2* G2 or *CCR5* RNPs and corresponding ssODNs as in Fig.5B; expanded for 7 days in IL-2, washed, and rested for 24 hours without cytokine. Cells were then stimulated with plate bound anti-CD3 for 24 hours (**A-F**) or IFNγ for 20 minutes (**G-I**). (**A, D, & G**) Representative flow overlay of PD-1 (**A**), CD25 (**D**), pSTAT1 (**G**) expression in edited cells from the same donor. (**B-C, E-F**) Summary data of median flow values for PD-1 (**B**), CD69 (**C**), CD25 (**E**), and CD71 (**F**) for all editing conditions in all donors after 24-hour anti-CD3 stim. (**H-I**) Summary data of percent positive and median flow values for pSTAT1 for all editing conditions in all donors after IFNγ stim. (**I**) Median pSTAT1 values normalized to average MFI of all editing conditions from the individual donor. For all summary data n=6, matched one-way ANOVA with Tukey's correction. Lines and error bars represent mean +/- SEM. RNP- ribonucleoprotein, ODN or ssODN – single stranded oligo-deoxynucleotide. Shapes in summary plots correspond to individual donors. All data is from 2 independent experiments. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001.

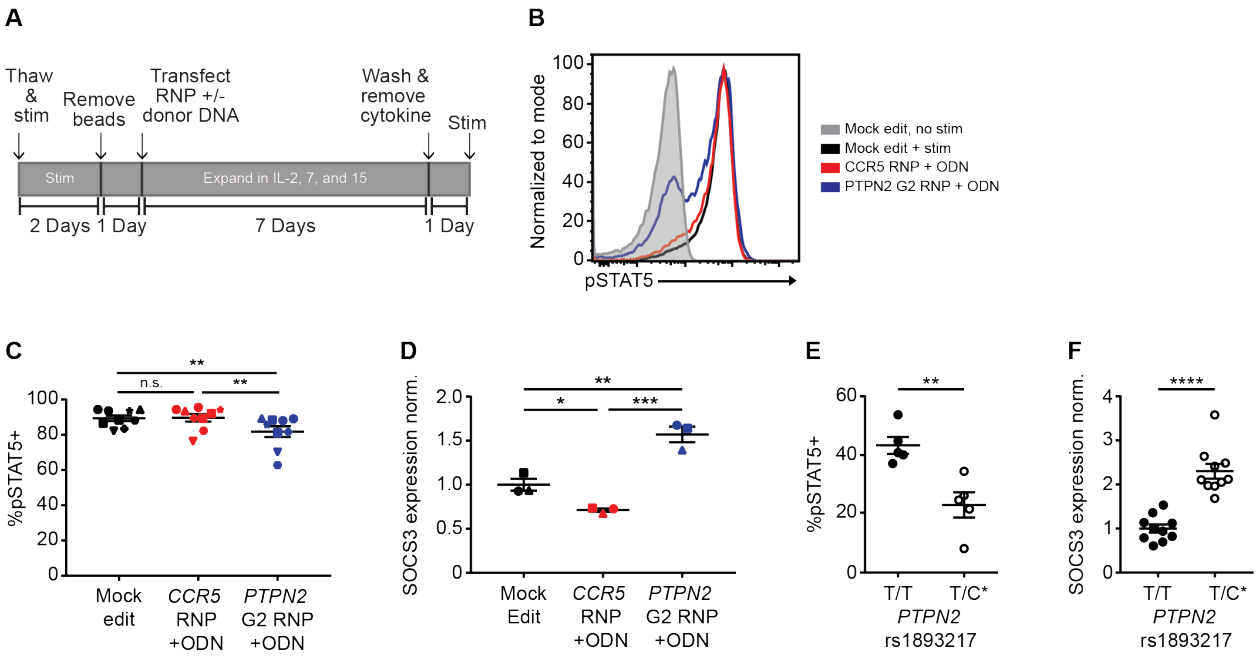

Figure 2-7. Sustained cytokine signals in *PTPN2* disrupted CD4$^+$ cells leads to loss of IL-2 response and increased SOCS3 expression. CD4$^+$ T cells from human donors were either mock edited or edited with *PTPN2* G2 or *CCR5* RNPs and corresponding ssODNs as in Fig.5B. Cells were then expanded for 7 days in IL-2, -7, and -15, washed and rested for 24 hours without cytokine, and subsequently stimulated with IL-2 for 20 minutes. (**A**) Work-flow used to produce and assay PTPN2 edited CD4$^+$ T cells and corresponding controls in cytokine media. (**B**) Flow overlay of pSTAT5 response to IL-2 stimulation in different cell populations from the same donor. (**C**) Summary data of %pSTAT5+ cells for all edited conditions after 20-minute IL-2 stim (n=9 human samples (7 independent donors plus 2 repeat donors, repeat donors were run in separate experiments), matched one-way ANOVA with Tukey's correction, representative of 3 independent experiments). (**D**) Edited CD4$^+$ T cells were expanded, and rested as in (**A**), and assessed for baseline SOCS3 expression by RT-qPCR (n=3, matched one-way ANOVA with Tukey's correction, Cq values normalized to housekeeping gene B2M, then normalized to the average adjusted Cq value of mock edited cells). (**E**) Memory CD45RO$^+$CD25$^-$CD4$^+$ T cells were isolated by negative selection and stratified by *PTPN2* rs1893217 risk (asterisk denotes risk variant). Cells were stimulated with IL-2 for 20 minutes. Response to IL-2 as measured by percent positive for pSTAT5 is shown (n=5, Mann-Whitney). (**F**) Memory CD4$^+$ T cells, obtained as in (**E**) were measured for baseline SOCS3 expression by RT-qPCR. SOCS3 Cq values were normalized to the mean of 2 housekeeping genes (B2M, RPL36AL) then normalized to the average adjusted Cq value of T/T donors (n=10, Mann-Whitney). All lines and error bars represent mean +/- SEM. RNP- ribonucleoprotein, ODN or ssODN - single stranded oligo-deoxynucleotide. (**C&D**) Shapes correspond to individual donors. (**E&F**) All dots represent individual donors. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001.

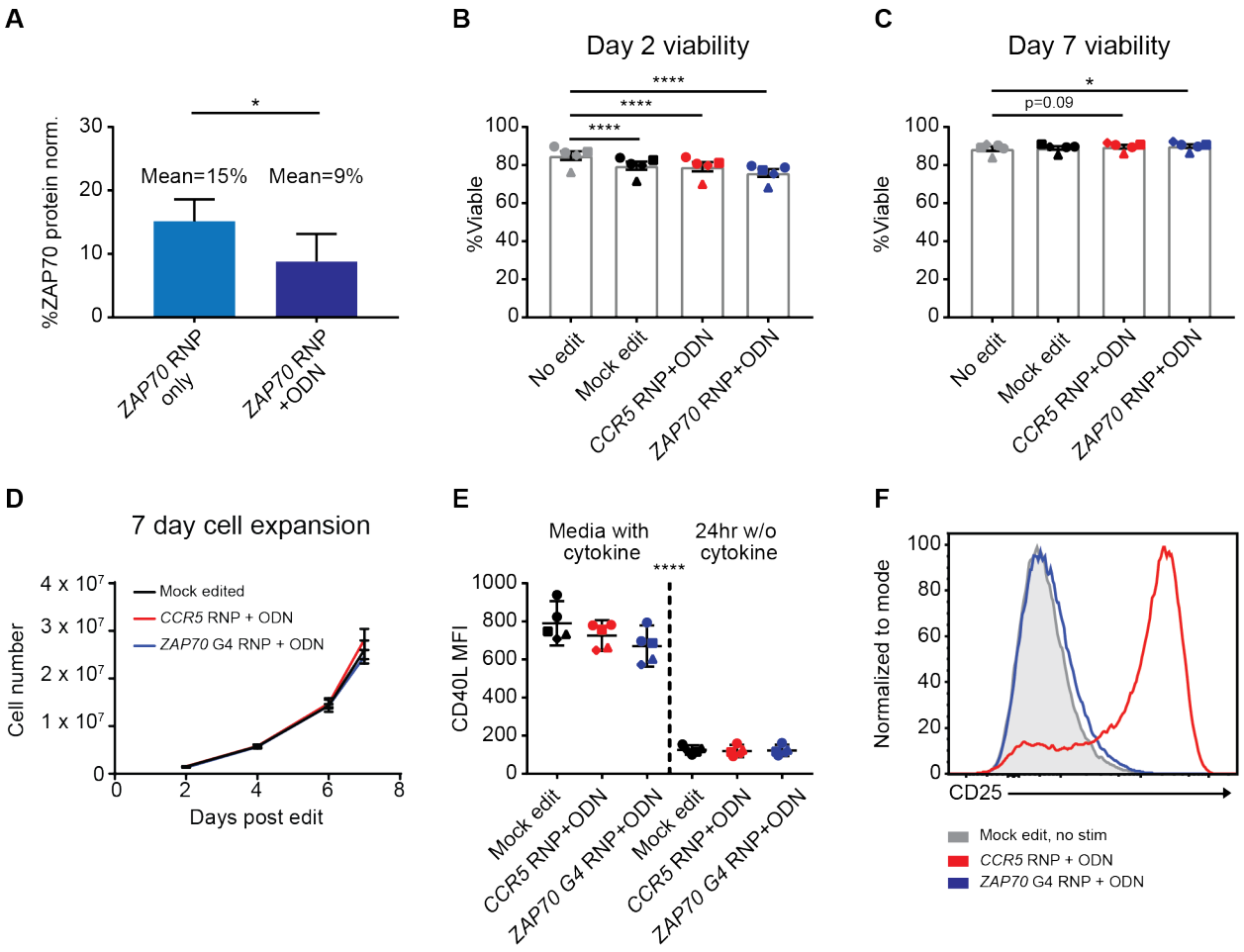

Figure 2-S1. *ZAP70* disruption in CD4$^+$ T cells is highly efficient and results in minimal impact on cell viability and expansion post-editing. (**A**) Quantified ZAP70 protein expression comparing western blots of CD4$^+$ T cells from the same donors that had been transfected with *ZAP70* G4 RNP alone vs. *ZAP70* G4 RNP plus an ssODN containing a stop codon. Protein expression values are relative to actin and normalized to mock edited values from the same T cell donor (n=5 human samples (4 independent donors plus 1 repeat donor, repeat donors were run in separate experiments), paired t test). (**B-C**) After editing, cells were expanded in high cytokine culture for 7 days. At 2 days (**B**) and 7 days (**C**) post-editing, cells were stained with viability dye and assessed via flow cytometry. Percent viable reflects the percent of events collected from each culture that were identified as single, live, lymphocytes (n=5, matched one-way ANOVA with a Dunnet post hoc test referenced to unedited cells). (**D**) Cell expansion post-editing stratified by editing condition (n=5 human samples (4 independent donors plus 1 repeat donor, repeat donors were run in separate experiments), lines represent mean, brackets represent SEM). (**E**) Summary flow data for CD40L expression in cells from the same edited cultures +/- 24-hour rest in cytokine free media (lines and error bars represent mean +/- SEM, n=5 human samples (4 independent donors plus 1 repeat sample, repeat samples were run in separate experiments), matched one-way ANOVA). (**F**) Representative overlay of flow cytometry histograms for CD25 expression in indicated editing conditions from the same donor at 24 hours post anti-CD3 stimulation. (**A-C**) Bars represent mean +/- SEM. RNP - ribonucleoprotein, ODN or ssODN - single stranded oligo-deoxynucleotide. Shapes in summary plots correspond to individual donors. All data is from 2 independent experiments. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001.

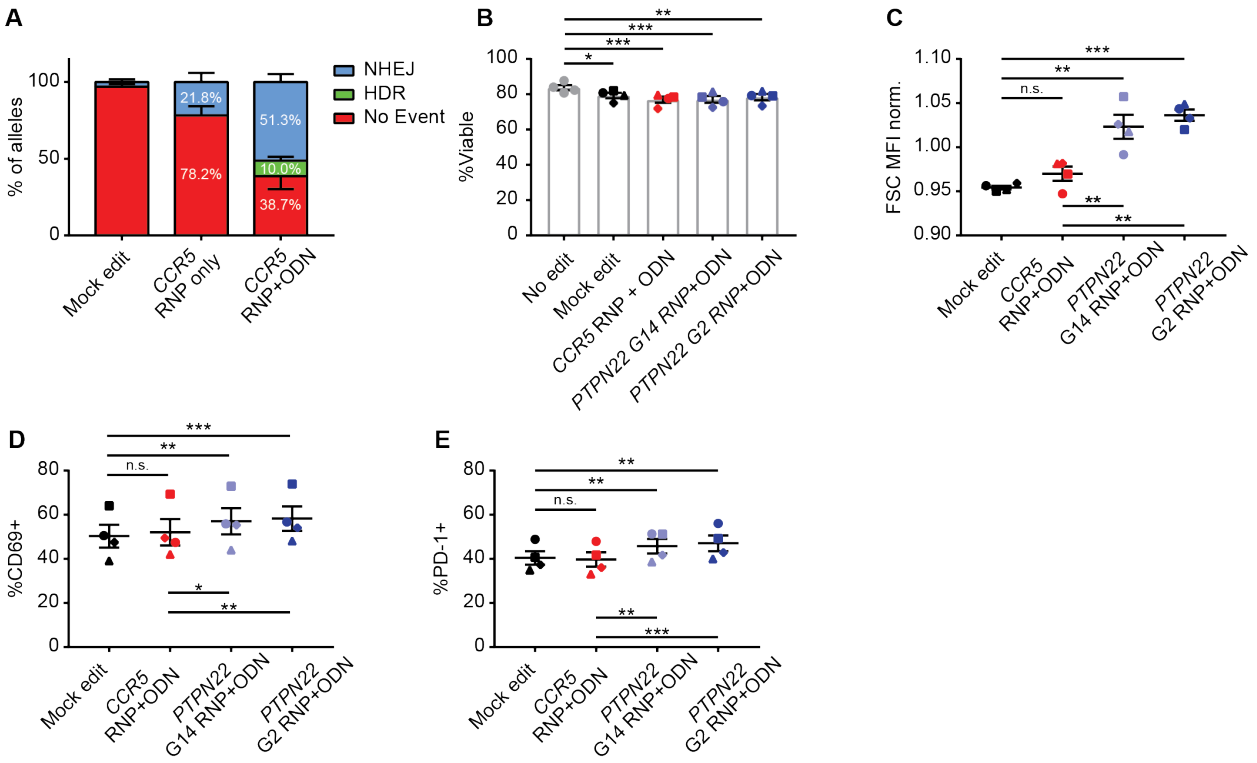

Figure 2-S2. *PTPN22* disruption in CD4$^+$ T cells promotes increased cell activation in response to TCR stimulation. (**A**) ddPCR analysis of editing frequencies in unedited and *CCR5* RNP edited cells with and without co-delivery of ssODN containing a stop codon (bars represent mean +/- SEM, n=4 independent human donors, percentages reflect summary data). (**B**) Two days post-editing cells were stained with viability dye and assessed via flow cytometry. Percent viable reflects the percent of events collected from each culture that were identified as single, live, lymphocytes. (**C-E**) Summary data of median flow values for forward scatter (**C**), percent CD69 positive (**D**), and percent PD-1 positive (**E**) for all editing conditions in all donors at 24-hour post-anti-CD3 stimulation. (**B-E**) Lines represent mean +/- SEM, n=4, matched one-way ANOVA with Dunnet post hoc test (**B**) or Tukey's correction (**C-E**), representative of 2 independent experiments. RNP – ribonucleoprotein, ODN or ssODN - single stranded oligo-deoxynucleotide. Shapes in summary plots correspond to individual donors. All data is from 2 independent experiments. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001.

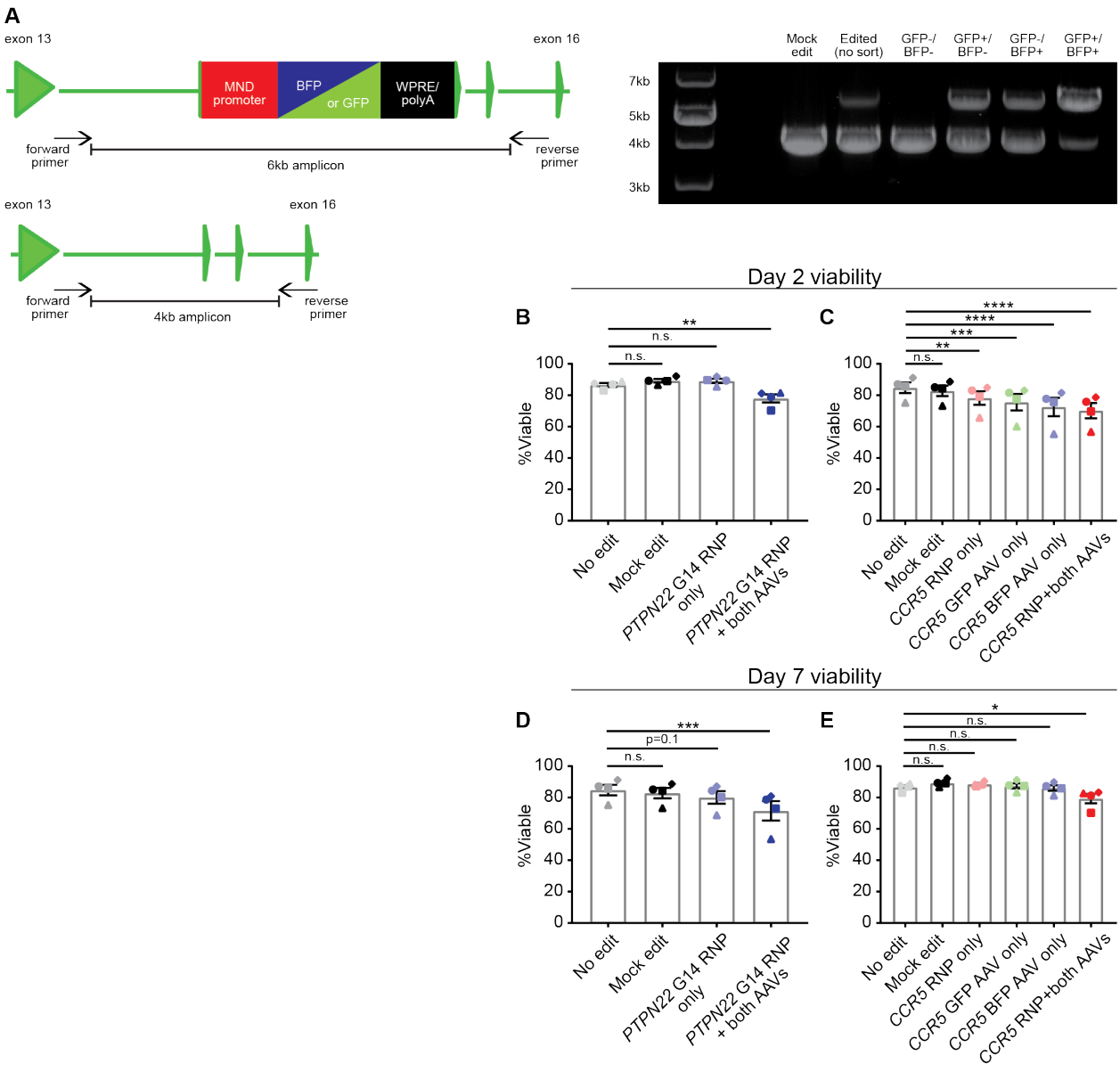

Figure 2-S3. Bi-allelic knockout of *PTPN22* in CD4$^+$ cells using rAAV6 delivered homology donor cassettes and impact on cell viability. (**A**) Left: Schematic of PCR based detection of AAV mediated HDR gene editing in gDNA. Right: Representative agarose gel of amplified gDNA from *PTPN22* bi-allelically edited CD4$^+$ T cells that FACS-sorted based on editing outcome. Large band reflects alleles with AAV homology cassette within the amplicon site. (**B-E**) Following editing, CD4$^+$ T cells were expanded in culture for 7 days. At 2 days (**B-C**) and 7 days (**D-E**) post-editing cells were stained with viability dye and assessed via flow cytometry. Cells were either unedited, mock edited, RNP edited only (*PTPN22* (**B&D**), *CCR5* (**C&E**)) or edited with a combination of RNP and AAVs containing homology donor cassettes. Percent viable reflects the percent of events collected from each culture that were identified as a single, live, lymphocytes (n=4, matched one-way ANOVA with a Dunnet post hoc test referenced to unedited cells). Bars represent mean +/- SEM. RNP - ribonucleoprotein. Shapes in summary plots correspond to individual donors. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001.

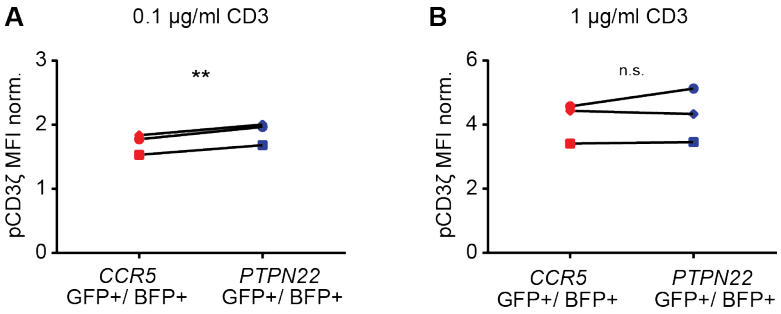

Figure 2-S4. Bi-allelic knockout of *PTPN22* in CD4$^+$ T cells using rAAV6 homology donors enhances CD3ζ phosphorylation in response to low-dose anti-CD3 crosslinking. (**A-B**) CD4$^+$ T cells bi-allelically edited (GFP+/BFP+) at either the *CCR5* or *PTPN22* loci from the same donor were stimulated using mouse anti-human CD3 (at final concentrations listed) and goat anti-mouse crosslinker (at half the concentration of anti-CD3 used). Cells were fixed at several timepoints post stimulation and analyzed for CD3ζ phosphorylation by flow cytometry. Median pCD3ζ values are shown relative to the same sample's unstimulated value. Shapes correspond to individual donors (n=3, paired t test, ** $p<0.01$).

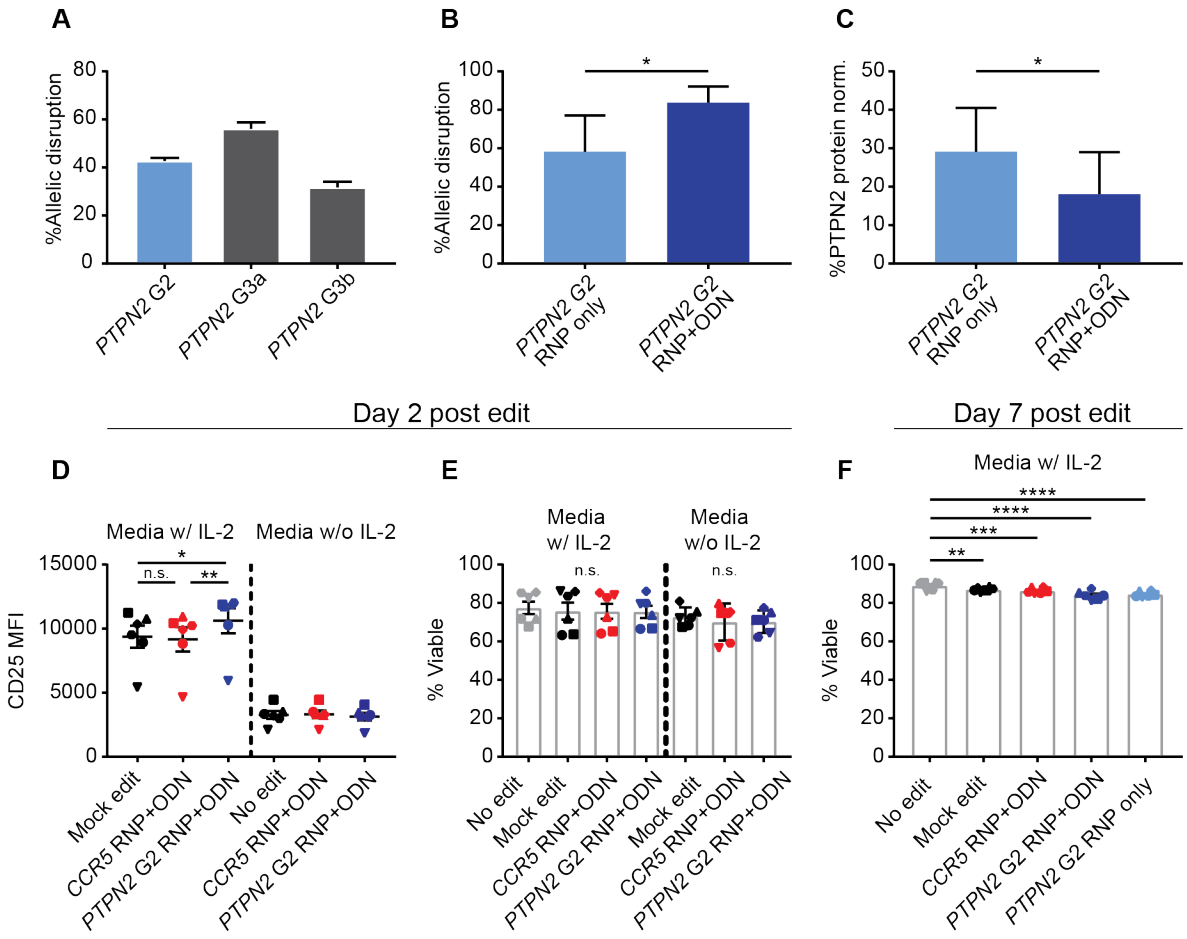

Figure 2-S5. *PTPN2* disruption in CD4$^+$ T cells is most efficient with RNP and ssODN co-delivery and impacts the cell surface phenotype in media containing IL-2. (**A**) ICE analysis of indel frequency in primary human CD4$^+$ T cells using 3 alternative gRNAs delivered as RNPs. Percent allelic disruption represents estimated indel frequency in edited cell compared to unedited controls (n=3 independent human donors). (**B**) ICE analysis of indel frequency in primary human CD4$^+$ T cells using *PTPN2* G2 RNP +/- stop codon containing ssODNs. Percent allelic disruption represents estimated indel frequency in edited cell compared to unedited controls (n=5 independent human donors, paired t test). (**C**) Quantified PTPN2 protein expression comparing western blots of CD4$^+$ T cells from the same donors edited with *PTPN2* G2 RNP +/- stop codon containing ssODN. Protein expression values are relative to actin and normalized to unedited values from the same T cell donor (n=6 independent human donors, paired t test). (**D**) Median CD25 expression in edited cells cultured for 2 days with or without IL-2 as in Fig. 2-5B (n=6, matched two-way ANOVA with Tukey's correction stratified by presence of IL-2 in media, representative of 2 independent experiments). (**E-F**) Two and seven days, post-editing cells were stained with viability dye and assessed via flow cytometry. Percent viable reflects the percent of events collected from each culture that were identified as single, live, lymphocytes (n=6, (**E**) matched two-way ANOVA with Tukey's correction, stratified by presence of IL-2 in media, (**F**) matched one-way ANOVA with a Dunnet post-hoc test referenced to unedited cells). (**A-E**) Bars and lines represent mean +/- SEM. RNP - ribonucleoprotein, ODN or ssODN - single stranded oligo-deoxynucleotide. Shapes in summary plots correspond to individual donors. All data is from 2 independent experiments. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001.

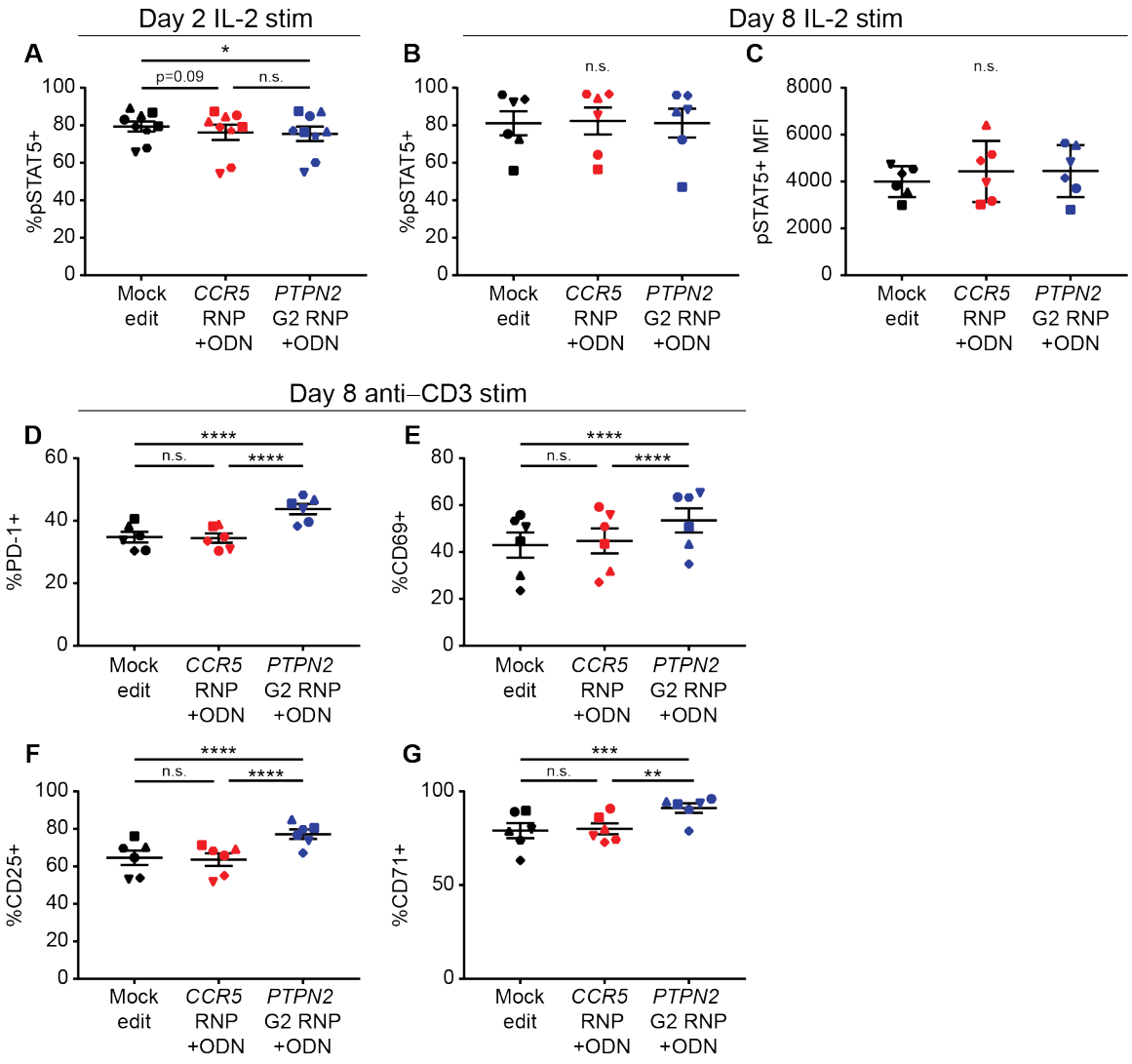

Figure 2-S6. *PTPN2* disruption in CD4$^+$ T cells alters the response to IL-2 and TCR stimulation. (**A**) Response to IL-2 stimulation for experiments described in Fig.5E&F as measured by %pSTAT5 positive over background levels (n=9). (**B-C**) Percent response (**B**) and median pSTAT5 fluorescence (**C**) following a 20-minute IL-2 stimulation. Prior to simulation, edited cells were expanded for 7 days in IL-2 supplemented media and rested for 24-hours in cytokine free media as described in Fig.5B (n=6). (**D-G**) Summary data for all donors showing percentage of cells staining positive for PD-1 (**D**), CD69 (**E**), CD25 (**F**), or CD71 (**G**) for control and edited T cell populations in response to 24-hour anti-CD3 stimulation. All data analyzed with matched one-way ANOVA with Tukey's correction. Lines and error bars represent mean +/- SEM. RNP - ribonucleoprotein, ODN or ssODN - single stranded oligo-deoxynucleotide. Shapes in summary plots correspond to individual donors. All data is from 2 independent experiments. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001.

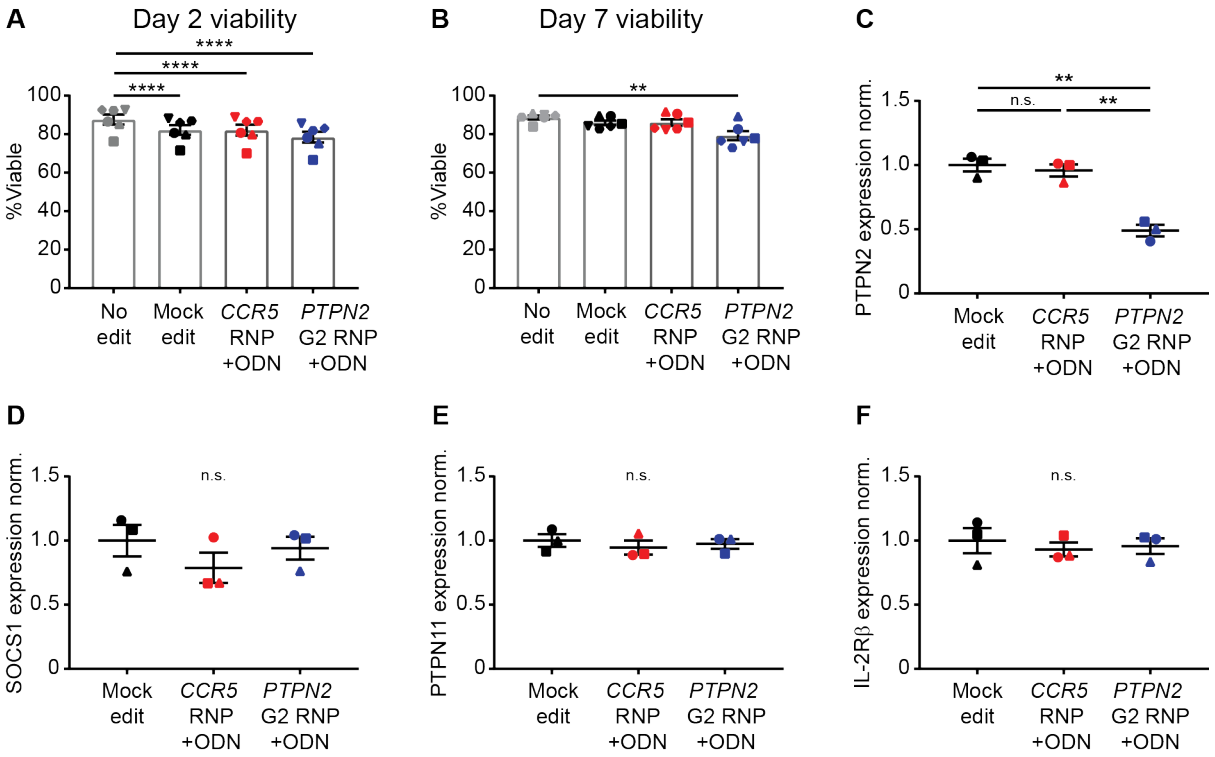

Figure 2-S7. *PTPN2* disruption in CD4$^+$ T cells reduces long-term viability in high cytokine media. (**A-B**) Viability assessment of edited and control populations at Day 2 (**A**) and Day 7 (**B**) post-editing using cytokine media (IL-2, IL-7, IL-15) as in Fig. 2-7A. Cells were stained with viability dye and assessed via flow cytometry. Percent viable reflects the percent of events collected from each culture that were identified as single, live, lymphocytes (bars represent mean +/- SEM, n=6, matched one-way ANOVA with a Dunnet post-hoc test referenced to unedited cells). (**C-F**) CD4$^+$ T cells that were edited, expanded and rested as in Fig.7A, were assessed for baseline mRNA expression levels for PTPN2 (**C**), SOCS1 (**D**), PTPN11 (**E**), and IL-2Rβ (**F**) by RT-qPCR (Lines and error bars represent mean +/- SEM, n=3, matched one-way ANOVA with Tukey's correction, Cq values normalized to housekeeping gene B2M, then normalized to the average adjusted Cq value of mock edited cells). RNP - ribonucleoprotein, ODN or ssODN - single stranded oligo-deoxynucleotide. Shapes in summary plots correspond to individual donors. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001.

gRNA Name	crRNA Sequence
ZAP70 G1	UUGCUACGACGGCCCACGAG
ZAP70 G2	CCCAGAGUAAAGUUUGCGCU
ZAP70 G3	GCACCAAGUUUGACACGCUC
ZAP70 G4	GGCAAGUACUGCAUUCCCGA
CCR5	CUCACUAUGCUGCCGCCCAG
PTPN22 G2	AAGGCAAUCUACCAAGUACA
PTPN22 G14	GACACCUGAAUCAUUUAUUG
PTPN2 G2	CCACUCUAUGAGGAUAGUCA
PTPN2 G3a	AAGGAGUUACAUCUUAACAC
PTPN2 G3b	CAGUUUAGUUGACAUAGAAG

Table 2-ST1. Guide RNA (gRNA) sequences for all genomic targets discussed. All gRNAs were complexed with IDT tracer RNA at a 1:1 molar ratio then complexed with recombinant Cas9 at a 1.2:1 molar ratio prior to electroporation with $CD4^+$ T cells.

	Assay	gRNA Assessed	Forward Primer	Reverse Primer
Primers	ICE Analysis or T7	*ZAP70* G1-G2	CCTCGAAAACCTGCCTAACA	GCACTGTGTCTTTTTGGCTC
		ZAP70 G3-G4	CAGCAGATCTGGAGGTGATG	CCTTCAGGCAGTAGATGAGC
		CCR5	TTGCATTCATGGAGGGCAAC	TGAGAGCTGCAGGTGTAATGA
		PTPN2 G2	TTGGCATCTGAAGGACCTTG	CCAAGCCCTCCTTTTCACTA
		PTPN2 G3a-G3b	AGTCACAATGGCTAATGTGCT	AAGCATAAGCAGCACTCTGT
	AAV Integration	*PTPN22* G14	ATGGAGGCAATCAAGATCTGGTAC	CCTGTCTTCACCTTGCTCTTTACA
		CCR5	CTTCTGCTAAGGAGAACTAAACCC	AAGGCAGCTTATTTCCAAATGAAT
	ddPCR	*PTPN22* G14	ATAAGCAAAAACCTCCTGGG	GCCTCAATGAACTCCTCAAA
		CCR5	TTGCATTCATGGAGGGCAAC	TGAGAGCTGCAGGTGTAATGA
	Assay	**gRNA Assessed**	**HDR Probe**	
Probes	ddPCR	*PTPN22* G14	CCTCAACCACAACTAGTCACTTATAAATG	
		CCR5	CCGCCTAAGTGACTAGCAGTGG	
			WT Probe	
	ddPCR	*PTPN22* G14	CCAGCTTCCTCAACCACAATAAATG	
		CCR5	ATGCTGCCGCCCAGTGG	
			Reference Probe	
	ddPCR	*PTPN22* G14	TCATCATCTATCCTTGGAGCAGTTG	
		CCR5	CCTCCGCTCTACTCACTGGTGT	

Table 2-ST2. Primer and probe sequences for all editing assessment assays discussed. All ddPCR probes were ordered with a 3' BHQ1 quencher; HDR ddPCR probes were ordered with a 5' FAM reporter, while all WT and Reference probes were ordered with a 5' HEX reporter.

Gene analyzed	Forward Primer	Reverse Primer
PTPN2	CGGGAGTTCGAAGAGTTGGATA	CGACTGTGATCATATGGGCTTA
SOCS3	CCCAGAAGAGCCTATTACATCTAC	CAGCTGGGTGACTTTCTCATAG
SOCS1	CTTCTGTAGGATGGTAGCACAC	GAACGGAATGTGCGGAAGT
PTPN11	GTTATGATTCGCTGTCAGGAAC	CTGCTTGAGTTGTAGTACTGTACC
IL-2Rβ	CCAGATTCTCAGAAACTGACCA	TTATGTTGCATCTGTGGGTCTC
B2M	GAGGCTATCCAGCGTACTCCA	CGGCAGGCATACTCATCTTTT

Table 2-ST3. Primer sequences for all RT-qPCR targets discussed. All targets were assessed using SYBR Green detection.

Chapter 3

CRISPR/Cas9 "SNP editing" of the autoimmune-associated gene *PTPN22* in human T cells

Introduction

Genetic variation can account for altered risk of nearly all forms of human disease. Genome wide association studies (GWAS) have identified thousands of genetic risk variants that associate with incidence of tuberculosis, heart disease, Parkinson's disease, and many other disorders[6,26]. Previous work has demonstrated that genetic associations with certain diseases can highlight potential biological mechanisms responsible or required for disease progression and obviate relevant targets for therapeutic intervention[5,6,36]. However, if a genetic variant of interest is rare or has low homology in murine settings, then options for functional study are limited.

The GWAS identified rs2476601 single nucleotide polymorphism (SNP) in the gene *PTPN22* is associated with increased risk for >10 autoimmune disorders[71]. Such involvement in multiple forms of autoimmunity indicates a role in a common biological mechanism, which is frequently exploited by autoimmune disorders to result in loss of tolerance to self-antigen. PTPN22 is a negative regulator of antigen receptor signaling in lymphocytes[73,74,79] but has also been shown to regulate several other immunological pathways, including type 1 interferon signaling[51], integrin expression[141], and neutrophil netosis[54]. The risk SNP results in a C1858T coding sequence change, converting an arginine to a tryptophan (R620W) in a proline rich domain that is used for PTPN22's association with co-regulatory molecules[64]. Murine models of this risk variant and knockout, in part, result in decreased regulation of antigen receptors, resulting in hyper-active TCR and BCR signaling driving enhanced proliferation, cytokine secretion, germinal center formation, and auto-antibody development[55,67,84].

Though the rs2476601 SNP has been studied extensively, consensus is lacking on the variant's functional impact due to strongly contrasting data from cross-sectional studies of lymphocytes derived from human carriers of the variant relative to murine models. Lymphocytes from human carriers of the rs2476601 variant show decreased antigen receptor signaling in T and B cells and diminished cytokine

production in response to *in vitro* stimulation[56,85]. This lack of agreement between human and murine data has entrenched the question of whether the R620W variant represents a loss- or gain-of-function?

Gene ablation studies in human memory T cells produces a phenotype like that of murine knockout memory T cells (see Chapter 2), arguing that the protein plays a similar role between species in T cell activation. However, studies have shown the R620W variant not to impact expression levels[67], but rather affect localization and protein interactions[64,67]. Also, as PTPN22 is involved in diverse immune cell pathways, mis-localization would likely impact various pathway uniquely, potentially producing a different phenotype than simple gene knockout would bestow. In agreement with this, NOD mice with the *Ptpn22* risk variant or *Ptpn22* knockout show different progression to disease, with the risk variant being required for auto-antibody formation[164]. This argues that the effects of the R620W variant cannot be reduced to a simple gain-of-function/ lack-of-function dichotomy and must be appreciated in its own right.

The contrasting nature of data surrounding *PTPN22* exemplifies the limits of our current approaches to studying the genetics of human disease, and to address this, some groups have called for the use of gene editing approaches for improved modeling of human genetic risk variants[45]. Gene editing generally uses site specific nucleases to create a double stranded break (DSB) in the DNA locus of relevance, while co-delivery of a homologous donor DNA template can allow for the cell to utilize homology directed repair (HDR) to mend the break in an error free manner[118]. If the donor DNA template contains minor edits to the coding sequence near the site of the DSB, then slight edits to the genome can be made[113]. Conversely if the cell is in the wrong phase of growth or a donor DNA template cannot be found, the cell will utilize non-homologous end joining (NHEJ), an error prone repair pathway to quickly fix the DSB, which frequently results in gene disruption[119]. This presents a unique challenge in the context of *PTPN22* rs2476601 as this variant represent altered coding of the *PTPN22* gene, and therefore any modification of T cells to impart the R620W edit would require highly efficient HDR while minimizing NHEJ within the coding region.

Our lab has shown that specific targeting and modification of the *PTPN22* locus is possible at high efficiency in primary human T cells using CRISPR/Cas9 ribonucleoprotein (RNP) delivery (Chapter

2). Also, we and other groups have also shown that through designer nuclease transduction and virally delivered selection constructs that functional interrogation of genetically altered primary human lymphocytes is possible[122,125,150,152]. However, previous methods for purification of edited primary human cells have been lengthy and required FACs sorting or drug-resistance selection, all factors which potentially affect immune cell function *in vitro*. This chapter looks at methods for the precise editing of the *PTPN22* R620W variant into non-risk donors, at high efficiency in bulk isolated $CD4^+$ T cell populations. We identify several approaches for the use of gene editing in this manner (Fig. 3-1, A-C), however appreciable challenges remain with each method identified.

Results

Trackable editing of the PTPN22 R620W SNP using AAV mediated reporter insertion.

Rates of HDR within a population of cells occurs at variable rates dependent on several factors such as gRNA sequence, DNA repair template homology, and donor response to stimulation[165]. For these reasons we sought to develop methods to alter the coding region of *PTPN22* exon 14 to reflect the risk variant of the R620W SNP using AAV delivered DNA repair templates. These templates would simultaneously alter the SNP site and insert a GFP cassette to allow flow-based identification of edited cells. As *PTPN22* has several splice variants and expression levels of different isoforms of PTPN22 has been linked with the presence of autoimmune disease[166,167], we felt it important to preserve as many features of endogenous expression as possible.

For editing the R620W SNP we utilized an RNP targeting near the 3' end of Intron 13 in *PTPN22* (G13), we simultaneously transduced the cells with a rAAV6 vector containing MND promoter driving expression of a GFP reporter and a 3' WPRE flanked by 1kB homology arms adjacent to the cleavage site. We termed the vector intronic repair AAV, or "irAAV" (Fig. 3-2, A). The 3' homology arm contained a C -> T transition at *PTPN22* nucleotide 1858 to create the rs2476601 SNP. As a successful HDR would encode the GFP cassette into the intron while inserting the SNP into the exon and mRNA transcripts should possess the SNP variant, but the GFP cassette should be spliced out. The GFP cassette was

encoded in reverse sequence orientation to the expression of PTPN22 and therefore expressed off the opposite strand of DNA from *PTPN22*, and as the MND promoter is unidirectional[152], it should have minimal impact on PTPN22 expression levels.

Using previously established methods (chapter 2) primary human $CD4^+$ T cells were transduced with G13 and irAAV and expanded in culture for 7 days. At 2 days and 7 days post-editing, flow cytometry was conducted to monitor GFP expression (Fig. 3-2, B). As the GFP construct contains its own promoter and tail element, this allows for the episomal expression of the GFP as well as expression after HDR. At 2 days post-edit we saw high levels of variable GFP expression, indicating viral transduction. After 7 days post-edit, GFP expression was reduced, leaving only GFP high-expressing cells, indicating consistent GFP expression and therefore likely HDR events. gDNA was taken from 7 day expanded cell cultures, and HDR edited alleles were amplified by using PCR to amplify from the GFP construct to 100bp 3' of the homology arm. Sanger sequencing of the PCR products revealed that 70-80% of GFP+ alleles were simultaneously positive for the R620W SNP (Fig. 3-2, C).

To explore if alternate AAV mediated editing approaches may be more efficient than use of G13 and the irAAV, we designed a cDNA construct to be delivered via AAV and integrate at exon 14, coding for the remainder of the PTPN22 protein with a T2A-GFP reporter. This approach was termed "cAAV". As the use of a cDNA construct from Exon 14 to 21 would code for approximately only a quarter of the PTPN22 coding region and allow the formation of disease associated splice isoforms, we felt this would result in minimal alterations of transcription. However, to avoid use of an artificial polyA tail or WPRE element which may affect transcript abundance[168], we identified a second gRNA sequence to use in conjunction with G13, targeting intron 20 of the PTPN22 gene (G20), and designed homology arms flanking the cDNA-T2A-GFP cassette to match 1kB 5' of the G13 RNP site, and 1kB 3' of the G20 RNP site (Fig. 3-3, A). With this approach successful HDR would excise approximately 20kB of the *PTPN22* gene and link the free ends of DNA with a cDNA-T2A-GFP construct, that codes for the rs2476601 risk variant and uses the endogenous *PTPN22* stop codon.

As before, primary human $CD4^+$ T cells were transduced with G13, G20, and cAAV and expanded in culture. After 4 days RNA was extracted from control and edited populations to generate

cDNA, and PTPN22 transcripts were amplified using PCR. Amplicons were created by priming from exon 2 to immediately 3' of the PTPN22 stop codon (Fig. 3-3, B). In unedited samples one band was visible, reflecting a single dominant gene product, while in cultures edited with both RNPs and cAAV two extra bands were visible, one enlarged by the expected size of the GFP product, and one reduced by the expected size of exons 14-20. Sanger sequencing primed from the GFP construct revealed that all GFP+ transcripts also contained the risk SNP. Flow cytometry was performed on days 2, 4, 7 and 10 post-editing. Cells transduced with cAAV alone showed no GFP expression, while those transduced with either G13 or G20 and cAAV showed slight emergence of GFP expressing cells. However, cells transduced with both guides showed more robust GFP+ populations (20-30% GFP+, Fig. 3-3, C). Interestingly while GFP expression was highest at day 7, for reasons that are not immediately clear, expression declined at day 10 in both donor populations.

Editing of the PTPN22 R620W SNP using alternative ssODNs.

Though we found it possible to edit the coding sequence of *PTPN22* to reflect the R620W SNP in a robust and trackable manner using AAV delivered DNA constructs, previous work had shown that AAV based editing can produce phenotypic consequences on primary T cell responses to stimulation (Fig. 2-4). Therefore, we wanted to determine the efficiency possible using ssODN mediated HDR, which we found in chapter 2 to be less impactful on cell viability and produced *CCR5* edited controls that were largely identical to un-edited controls in response to various stimulations (Fig. 2-3, 5, 6, & 7). It has been reported that ssODN mediated HDR efficiency increases upon proximity to the DNA DSB[169], therefore we designed a gRNA to bind directly over the *PTPN22* R620W locus, producing a DSB immediately 5' of the SNP site (G14). Also, we designed 3 ssODNs producing 3 potential sequence edits when used in tandem with G14 (Fig. 3-4, A). Risk or control editing ssODNs made the risk associated C1858T edit or a silent C1858A edit respectively, while both also encoded a silent A1857C edit to assist with preventing RNP cleavage after successful HDR. Finally, a knockout ssODN inserted an 11bp "stop" cassette to produce a stop codon in all 3 reading frames. All ssODNs were 200bp with equal size homology arms on 5' and 3' sides.

Primary human CD4+ cells were activated and transduced with G14 and various ssODNs using previously described methods (Chapter 2). Flow cytometry was conducted 2- and 7-days post-editing, which showed no significant reduction in viability at day 2 (Fig. 3-4, B) and equal viability at day 7 (data not shown). At 4 days post-editing DNA was collected from all conditions to quantify HDR and NHEJ rates between conditions using droplet digital PCR quantification (Fig. 3-4, C). While there was a slight trend toward increased HDR in the knockout edited populations, HDR and NHEJ rates were similar between risk and control edited groups (Fig. 3-4, C-E).

After a 7-day expansion in culture, cells were washed and rested overnight without cytokine. Cultures were then activated with plate bound anti-CD3 for 24 hours and analyzed for altered expression of surface bound activation markers. As previously reported we found that PTPN22 KO cells showed significantly increased expression of CD69, CD25, PD-1, and CD40L over un-edited controls (Fig. 3-5, A-D). Interestingly, *PTPN22* risk edited populations produced trend toward an intermediate phenotype between control edited and knockout-edited populations. As HDR/NHEJ rates between these groups were similar, this possibly suggests that edited cultures containing the risk variant possessed less regulation of the TCR than control edited cultures. Also, cells from two separate donors were tested for calcium flux after editing, expansion, and rest. PTPN22 deficient T cells were increased in overall calcium flux relative to unedited controls, as measured by area under the curve (AUC), but data from other edits was inconsistent (Fig. 3-5, E). Similar data was found for other calcium flux metrics such as peak and mean flux (data not shown).

As human carriers of the PTPN22 rs2476601 risk SNP are reported to have increased populations of Th1 memory T cells, we decided to see if altered coding for PTPN22 resulted in modulated Th1 skewing. After rest, edited cell populations were subjected to cytokine skewing conditions consisting of exogenous IL-2, IL-12, anti IL-4, and CD3/CD28 stimulation to promote the generation of a Th1 phenotype. Interestingly, *PTPN22* knockout and risk edited populations showed a trend toward reduced IFNγ expression relative to unedited and control edited populations (Fig. 3-5, F).

Cloning of PTPN22 R620W SNP ssODNs edited cells.

Though SNP editing of the *PTPN22* R620W SNP using G14 and ssODN templates is possible at relatively high efficiency, a drawback is that upon editing, the lack of a trackable marker results in a mixed population of unedited cells, bi-allelically edited, heterozygous, knockout, and monozygous cells. This mixed population makes determining overall effect size of the variant difficult. To account for this, we explored using single cell sorting to generate clone populations.

To test the feasibility of finding edited clones from this approach we first tested this with Jurkat T cells. Work flow for cloning consisted of identical editing approaches established in Figure 3A, and after 7 days of expansion cells were single cell sorted into 96 well u-bottom plates and expanded for 3 weeks. Cells were then genotyped using a nested PCR on heat disrupted cell lysate. Nested PCR products were sanger sequenced and identified as biallelically edited or monozygous for the intended edit using ICE analysis[158] (Fig. 3-6, A). After Jurkat clones were positively identified by genotype and expanded, several were assayed by anti-CD3 induced calcium flux (Fig. 3-6, B). After testing clones that were unedited, biallelic *PTPN22* risk edited, or knockout, we found no obvious in differences among genotypes for mean and peak flux, or area under the curve (Fig. 3-6, C-E), though there was a trend toward increased cells size by FSC in knockout clones relative to unedited.

We have since applied this editing/ cloning approach to primary $CD4^+$ T cells. We found this approach to work equivalently, however, upon sorting, primary cells are stimulated again using PHA and irradiated feeder cells, and before loss of expansion from the first PHA stim cells are genotyped, and positively selected for further rounds of expansion. Functional assays of edited primary human T cell clone have yet to provide conclusive data.

Discussion

In this chapter we describe multiple methods that have been tested to create functional “SNP edits” to the *PTPN22* locus of primary human $CD4^+$ T cells. We have shown that high efficiency or

trackable modifications to the coding sequence are possible in a manner that allows phenotypic study of edited populations. However, there is no approach that we have described or pursued without potential drawbacks. Work actively continues to pursue editing outcomes that allow robust investigation of GWAS SNPs on an isogenic human background. Currently one must weigh pros and cons of any approach taken herein when used to modify primary T cells.

While many studies using gene editing with nuclease and AAV delivered DNA donor templates frequently report HDR rates as readout for editing success[152,170], we found that this is insufficient for our purposes. When editing cells using G13 and irAAV we found that if our DNA donor template contains several edits that are not sequential then HDR events may not utilize the entirety of the donor template for repair of the DSB, resulting in a subset of cells that are GFP+ and *PTPN22* non-risk. As we now appreciate that successful HDR requires high proximity to the DSB[169], this creates an efficiency problem for editing coding regions of cells in this manner.

We also found distinct technical challenges when utilizing our dual RNP cAAV approach. In this set-up we did find expected outcomes of excising a significant portion of the gene and replacing it with a DNA repair template containing a cDNA risk variant of PTPN22 with aT2A linked GFP reporter. The use of dual RNPs proved a requirement to generate significant editing, however this reporter was of low value as its expression was based off the endogenous *PTPN22* promoter, which showed low transcription in our setting. Also, we saw waning expression over time in culture, likely reflecting decreasing levels of activation in expanding T cell cultures, and thus decreased expression of PTPN22. Furthermore, this approach was entirely dependent on high quality AAV production and proved difficult to replicate. Subsequent lots of AAV (even made with the same stock of plasmid DNA) produced much less dramatic results.

Our work in chapter 2 showed that use of AAV transduction in the phenotypic study of gene edited T cells produces potentially less robust data than editing using ssODN for the study of immune related gene function (chapter 2, Figure 4). Therefore, it was logical to explore SNP editing utilizing RNP and ssODN repair templates. This approach was strengthened by highly controlled factors between conditions, as all edited conditions from the same donor received the same RNP and ssODN delivery

varied by only a few nucleotides, however this approach was limited by a loss of clarity in the genetic composition of these populations. Though droplet digital PCR data showed that editing rates were similar between populations, next generation sequencing will be required to better understand the total editing outcomes of these cells.

Using ssODN mediated gene editing, we saw some preliminary results indicating that the *PTPN22* rs2476601 risk variant may bestow a loss of regulation of the TCR in T cell populations relevant to control edited cells, with trends of increased expression of some surface activation markers, though more work is clearly needed as the number of subjects tested and the effects were small. It is interesting that *PTPN22* knockout and risk variant edited T cells showed decreased INFγ expressing cells after culturing in Th1 skewing conditions (Fig. 3-4, F). Human carriers of the PTPN22 risk variant show an increased abundance of Th1 T cells, and previous work has suggested that strength of TCR signaling is a determining factor in T helper skewing outcomes[171]. We thought that increased TCR signaling may play a role in promoting Th1 phenotypes, but the outcome was opposite of what was expected. The lower concentration of INFγ expressing cells in our PTPN22 KO and risk edited T cells may reflect increased activation induced cell death due to increased TCR signaling rather than altered skewing potential, or possibly a combination of altered cytokine responses in tandem with altered TCR signaling. Further experiments will be needed to examine apoptosis rates in PTPN22 edited cells, and response to exogenous cytokines. These experiments are currently ongoing.

In this chapter we have shown that alterations to the coding sequence of *PTPN22* to produce R620W risk variant cells from non-risk donors is possible. Depending on the editing approach used, these cells can produce phenotypic data that may be informative as to how the *PTPN22* risk variant behaves in primary human T cells. These approaches also can be applied in tandem with single cell sorting and the generation of monoclonal populations in both cell lines and primary human T cells. Furthermore these approaches can be broadly applicable to nearly any genomic loci and may potentially provide functional insight on numerous other GWAS identified gene variants.

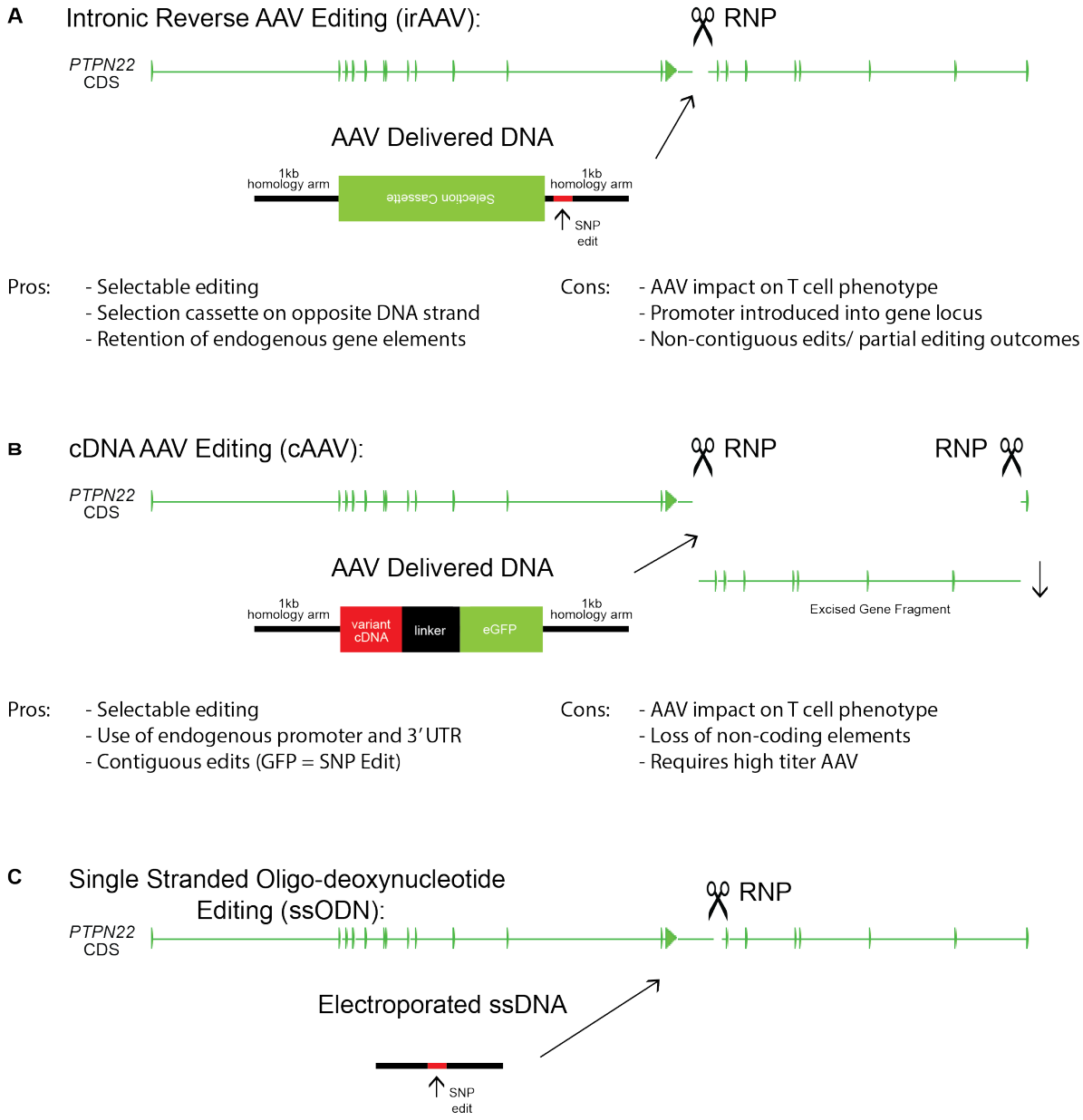

Figure 3-1. Overview of editing approaches used in modification of *PTPN22* R620W locus. (**A**) Editing using intronic RNP targeting 5' of the R620W SNP site, and an AAV designed to insert a SNP alteration as well as a GFP cassette into the cut site on the opposite strand of DNA (reversed orientation from PTPN22 gene). (**B**) Editing using two intronic RNPs targeting 5' of the R620W SNP site and 5' of the final coding exon, and an AAV designed to insert a GFP linked cDNA construct into the site, allowing for expression of a cDNA version of the PTPN22 risk variant. (**C**) Editing using an RNP targeting the R620W SNP locus and a ssDNA repair template. RNP- ribonucleoprotein.

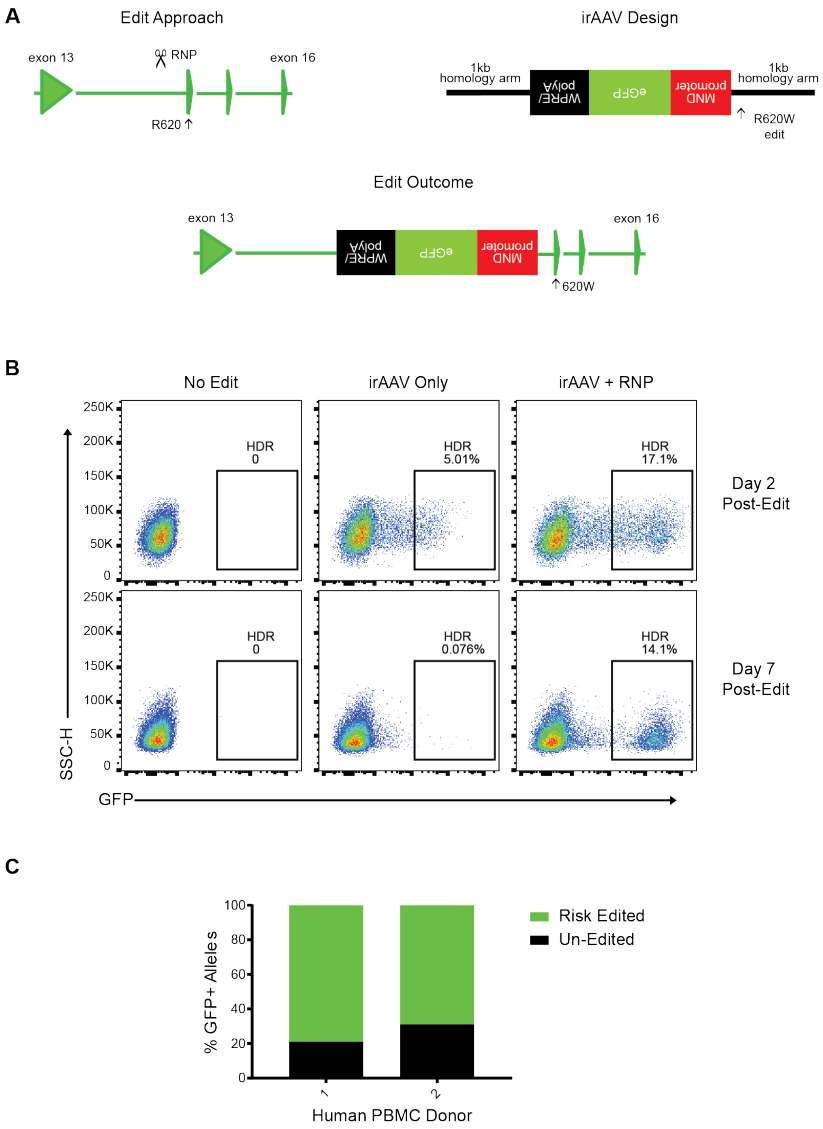

Figure 3-2. Editing of *PTPN22* R620W locus using irAAV. (**A**) Editing strategy for modifying PTPN22 R620W SNP using irAAV mediated HDR upon RNP cleavage. In brief, RNP cleavage of Intron 13 of PTPN22 allows AAV insertion of GFP cassette into the cut site. Successful HDR would allow for GFP to be expressed off opposite strand of DNA relative to PTPN22 using a unidirectional promoter, while a simultaneous coding change would be made to the R620W locus which was encoded into the 3' homology arm of the AAV. (**B**) Representative flow plots of controls and cells edited as in (**A**) at 2- and 7-days post-edit. (**C**) Bar graph of rates of conversion of C -> T (risk allele) in GFP+ alleles. RNP-ribonucleoprotein.

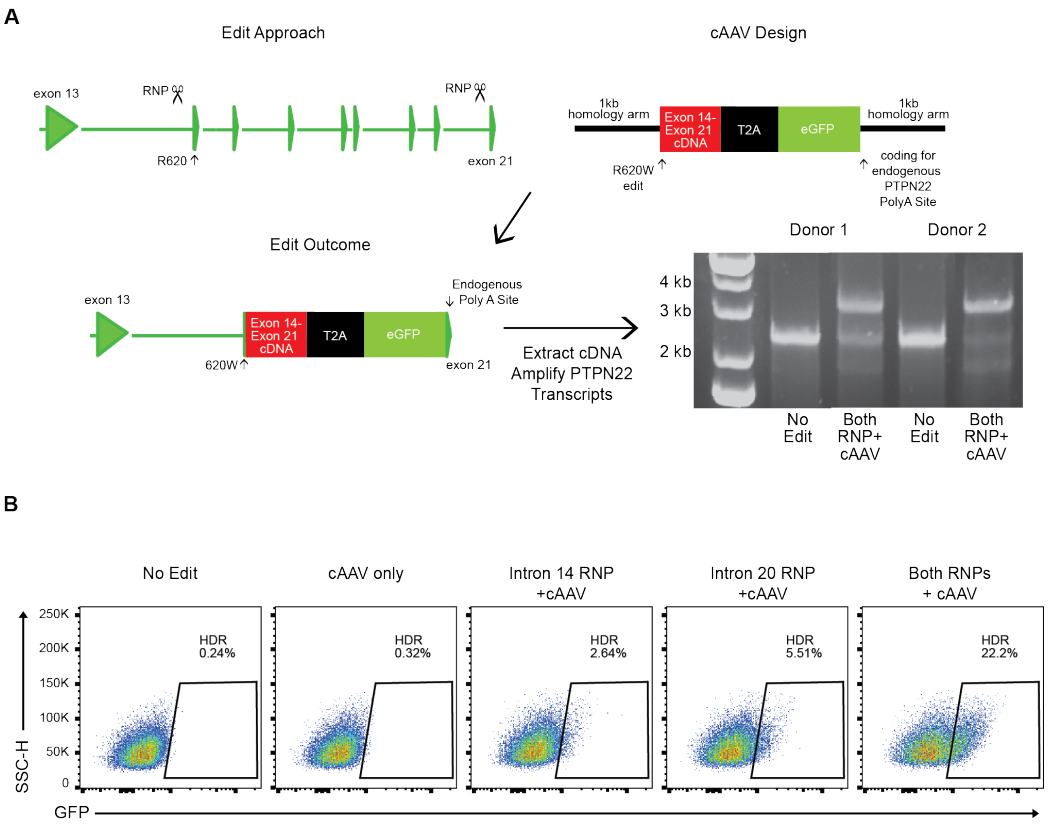

Figure 3-3. Editing of *PTPN22* R620W locus with cAAV. (**A**) Editing strategy for modifying PTPN22 R620W SNP using cAAV mediated HDR upon RNP cleavage. In brief, RNP cleavage of Intron 13 and Intron 20 of PTPN22 allows AAV insertion of cDNA-T2A-eGFP cassette into the cut site. Successful HDR would excise 20kb of the PTPN22 gene and replace with the AAV construct, while re-connecting with the genome at the PTPN22 endogenous stop codon/ PolyA site and use the gene's endogenous 3' UTR. (**B**) PCR product of cDNA extracted from unedited or edited cells from two donors. PCR product amplifies from Exon 2 of PTPN22 to 3' of the stop codon. Enlarged band represents GFP inclusion in mRNA transcripts as determined by DNA sequencing. (**C**) Representative flow plots of controls and cells edited as in (**A**) at 7-days post-edit.

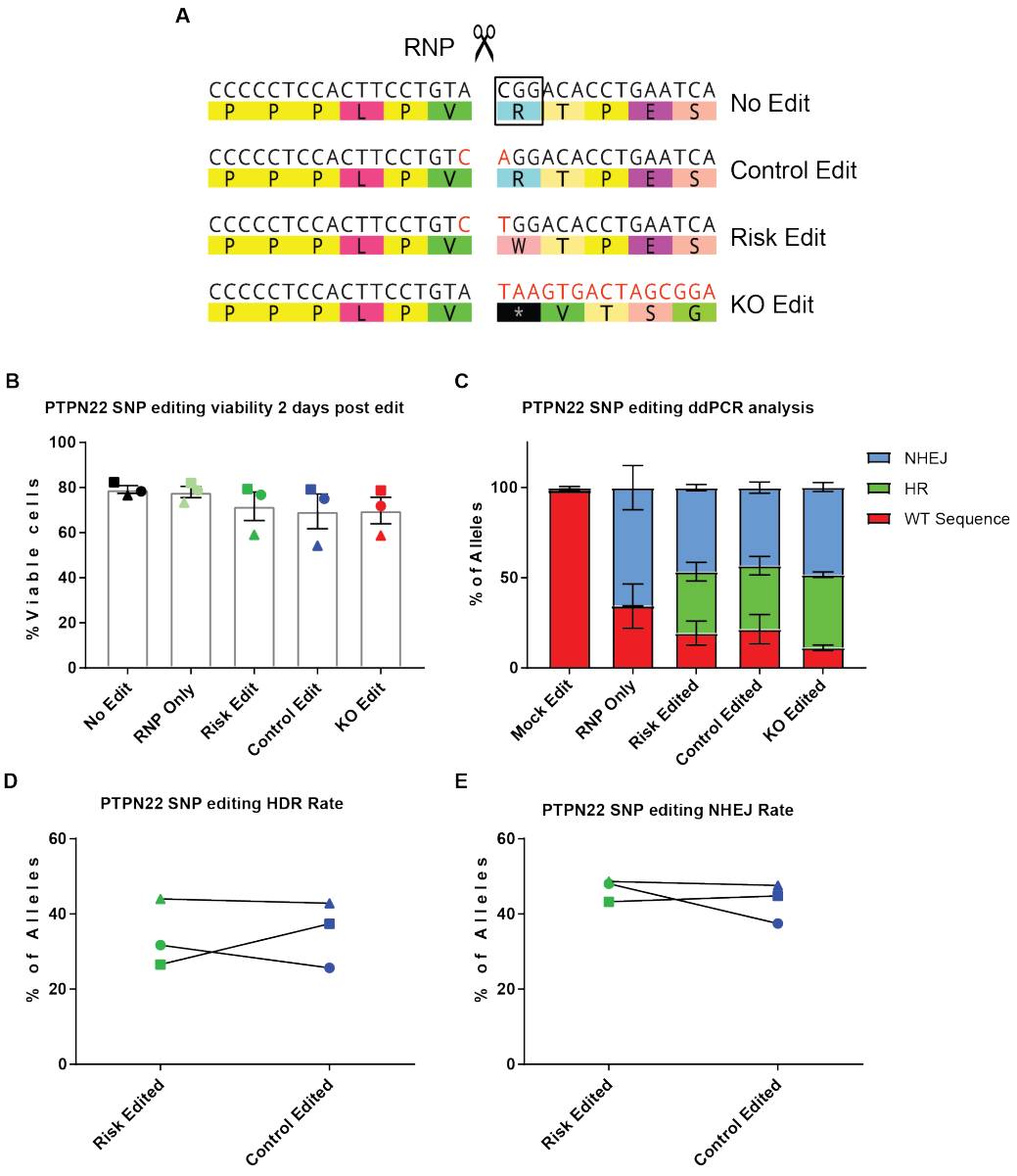

Figure 3-4. Editing of *PTPN22* R620W locus with RNP and ssODN. (**A**) Editing strategy for modifying *PTPN22* R620W SNP using RNP cleavage and ssODN repair. In brief, upon cleavage directly 5' of the *PTPN22* R620W risk SNP different populations of cells derived from the same human donor were simultaneously transfected with ssODNs coding for edits to the gene highlighted in red. Boxed codon represents residue 620 in primary PTPN22 sequence. (**B**) Editing rates of different coding outcomes as determined by ddPCR (n=3 human PBMC donors). (**C**) Two-days post-editing cells were stained with viability dye and assessed via flow cytometry. Percent viable reflects the percent of events collected from each culture that were identified as single, live, lymphocytes (n=3). (**D-E**) Rates of HDR (**D**) and NHEJ (**E**) among risk and control edited populations, connecting lines indicate identical PBMC donors. RNP- ribonucleoprotein, ODN or ssODN - single stranded oligo-deoxynucleotide. HDR- homology directed repair. NHEJ- non-homologous end joining. Shapes in summary plots correspond to individual donors.

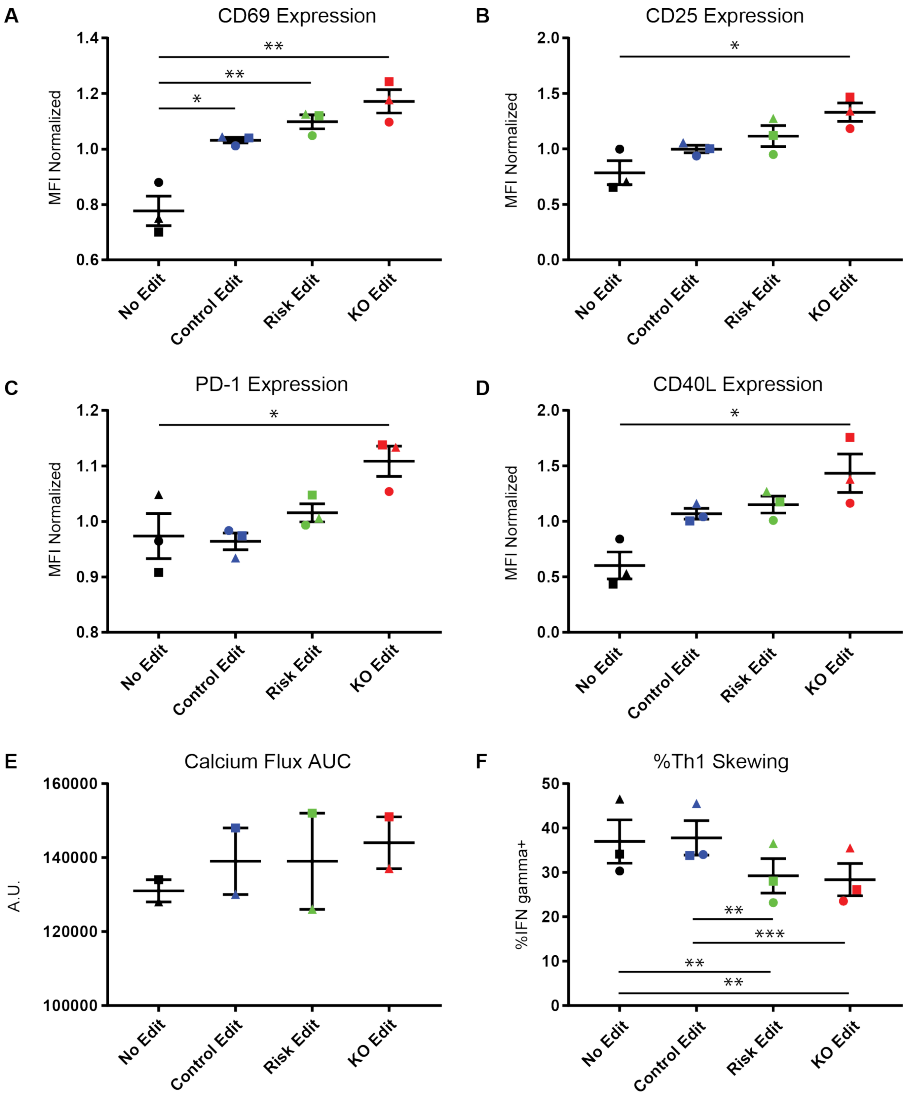

Figure 3-5. *PTPN22* risk edited T cells are potentially more responsive to TCR stim than controls. (**A-F**) CD4$^+$ T cells from human donors were either mock edited or edited with *PTPN22* RNP and corresponding ssODNs as in Fig.3-3A; expanded for 7 days in IL-2, 7, and 15 then washed and rested for 24 hours without cytokine. Cells were then stimulated with plate bound anti-CD3 for 24 hours (A-D, F) or soluble anti-CD3 (E). (A-D) Summary data of median flow values for CD69 (**A**), CD25 (**B**), PD-1 (**C**), and CD40L (**D**) for all editing conditions in all donors after 24-hour anti-CD3 stim. (**A-D**) Values normalized to average MFI of all editing conditions from the individual donor. (**E**) Summary of area under the curve for calcium flux of edited T cells. CD4$^+$ T cells generated, expanded, and rested as in Fig.3-3A were stained with indo-1 AM, monitored for baseline, then stimulated with anti-CD3. (**F**) Edited, expanded and rested cells were subjected to Th1 Skewing conditions (IL-2, IL-12, anti IL-4, And CD3/CD28 stim) for 3 days post cytokine rest. Cells were then stimmed with PMA and Ionomycin in the presence of golgi stop for 5 hours, fix/permed and stained for IFN gamma expression. All data analyzed with matched one-way ANOVA with Tukey's correction. A.U.- arbitrary units. Shapes in summary plots correspond to individual donors.

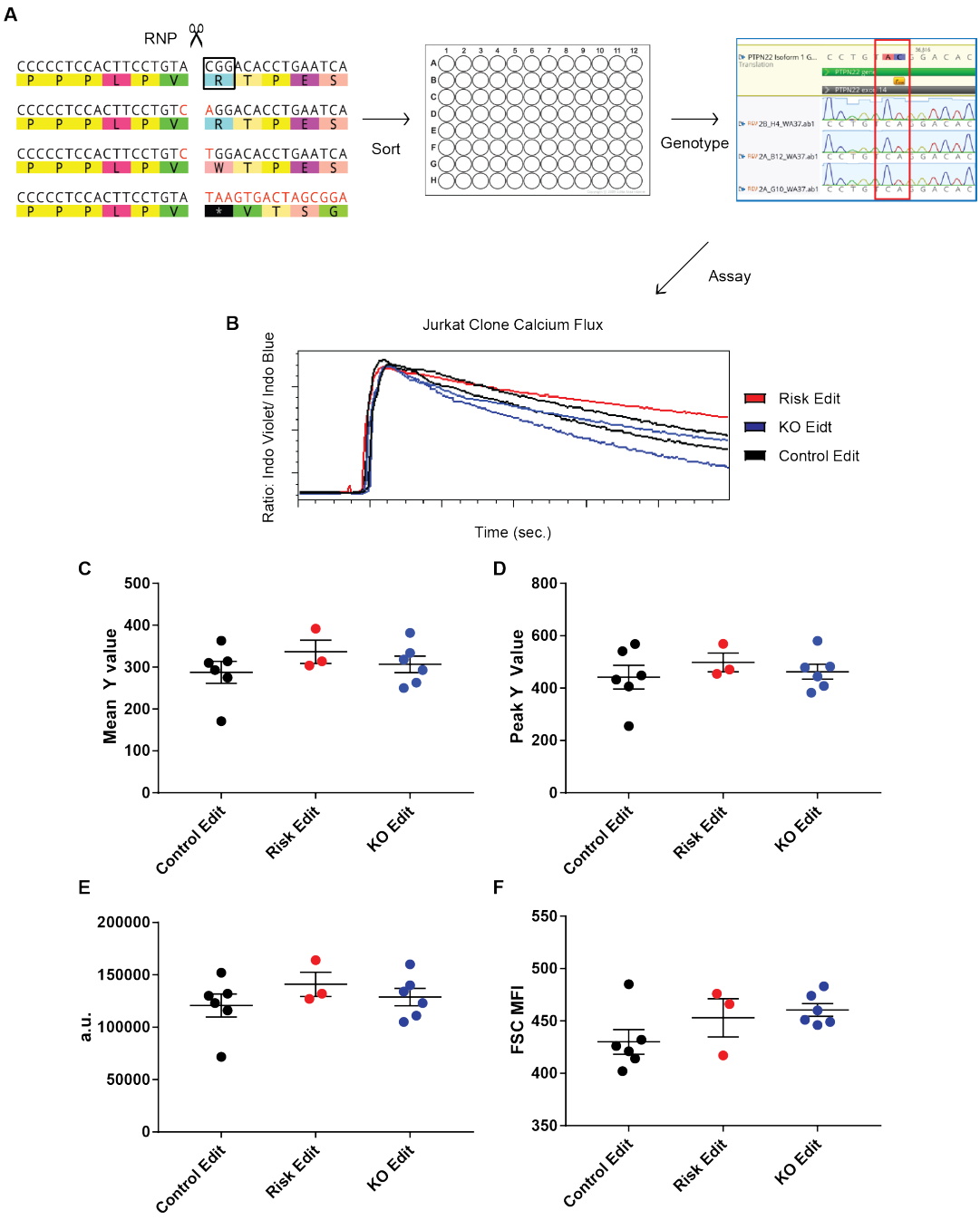

Figure 3-6. Generation of *PTPN22* edited Jurkat T cell clones using RNP and ssODN. (**A**) Editing strategy from Fig.3-3A adapted for cloning. In brief, after modifying *PTPN22*, cultures were expanded 7 days, then single cell FAC sorted. After expansion for 3 weeks, cells were genotyped using a nested PCR, then frozen. (**B**) Representative calcium flux of differently edited Jurkat T cell clones. Cells were stained with indo-1 AM, monitored for baseline, then stimulated with anti-CD3. (**C-F**) Summary data of calcium flux assays using *PTPN22* edited Jurkat clones, representing mean Y values (**C**), peak Y values (**D**), area under the curve (**E**) and mean FSC (**F**), n = 3 or 6 per group. RNP – ribonucleoprotein, ODN or ssODN - single stranded oligo-deoxynucleotide. All data is from 3 independent experiments.

Chapter 4

Concluding Remarks

Autoimmune diseases are correlated with several potential contributors, perhaps most prominent among them is genetics. Though genetics play only a supporting role in the generation of autoimmunity, overlapping genetic features shared among many diseases provide strong evidence of common biological mechanisms that produce a general loss of self-tolerance. Therefore, understanding how specific genetic variants incline their carrier toward a self-reactive phenotype may yield improved understanding of not only one specific disease but a whole range disorders, and may reveal potential targets for intervention and improved risk scoring systems for better preventative care.

Genome wide association studies have produced hundreds of genetic variants, commonly found throughout the population, which predispose their carriers to increased odds of a specific disease. While these variants most commonly make a minimal singular impact in regard to the carrier's overall risk of disease (odds ratio <2), in the case of autoimmunity they may account for a substantial proportion of total heritability of a given disease[5,36]. For example, in the case of T1D, the strongest indicator of genetic risk is most definitely HLA haplotype (50% of heritable risk[29]), but as it is estimated that 80% of the heritability of this disease is identified[172], this would mean that half of total genetic risk for T1D (and more for other diseases) would come from a patch-work collection of minor genetic risk variants and as yet missing heritability factors[25,26]. There are competing theories on what comprises the "missing" heritability residing in autoimmune diseases, be it rare but significantly impactful gene variants, weakly associated variants that are currently unidentifiable with modern study designs, or errors in estimation of total heritability of these diseases and the contributions made by individual variants[25,26]. Regardless of what is not yet known, it is likely that the combined effect of identified GWAS variants constitute a significant contribution to autoimmune genetic risk. Unfortunately, many of these variants have gone unstudied due to the sheer volume of variants identified, and/or the rarity at which they are found among the population. Furthermore, investigation of some variants have produced contrasting data depending on how that data is generated and what model is used, such as in the cases of *PTPN22* and *PTPN2*.

In this study, we have shown that gene editing approaches can be utilized in human $CD4^+$ T cells to elucidate gene function in primary human tissue through efficient gene disruption, and also allow for the modification of genetic sequences to reflect autoimmunity associated variants identified in GWAS studies. This approach has provided useful data confirming that the genes *ZAP70*, *PTPN22*, and *PTPN2* serve similar roles in human T cells as they have been shown to do in murine T cells, helping to resolve some unanswered questions regarding PTPN22 and PTPN2 function. Gene editing of primary human cells in this approach focused on T cell biology, but theoretically could be applied to other human cell culture and other genes, however, the effort involved in optimizing editing approaches for each application, which can only provide *in vitro* data regarding a single cell type at a time, is a limiting consideration for using gene editing in this fashion. In the analysis of GWAS variants of unknown or controversial significance, gene editing of primary human cells would need to be secondary to other research platforms such as mouse modeling or clinical data, which are able to offer holistic views of the effects of candidate variants on immune function. However, as we have shown, gene editing of primary human cells confirm data generated in other models translates to human cell biology, and currently is best suited for studying variants that have been shown to produce a loss of gene expression.

It is worth noting that many of the findings presented in this study showed manipulation of candidate GWAS identified genes produced modest effects on T cell activation and signaling. This would be expected as autoimmune associated GWAS variants must expand or persist within the population by offering some selective benefit that outweighs any negative effect. Therefore we would not expect most GWAS variants to severely impact critical components of the immune system and likely see variants afflict genes that have redundant or dispensable functions. When PTPN22 is ablated in a non-autoimmune prone mouse strain, they do not show disease and have subtle alterations to lymphocyte activation. Similarly, our subjects for these studies were healthy human donors with diverse genetic backgrounds and no demonstrated propensity toward autoimmunity, so it is unsurprising that ablation of PTPN22 in their T cells would also show subtle alterations to lymphocyte activation. As this technology expands, and improved homology directed repair allows more efficient SNP edits as opposed to simple gene ablation, subtle findings will likely be commonplace in the interpretation of GWAS identified variants,

especially those occupying non-coding regions of the genome. For example, CRISPR screens have recently been used in combination with mouse modeling to demonstrate that a non-coding GWAS SNP in the CD25 gene locus contributes to autoimmunity by slowing the expression of CD25 in response to TCR stimulation in a subtle and context specific manner[173]. Given our findings here it is entirely conceivable that such non-coding GWAS variants could be similarly modeled in primary human T cells.

Studies into the *PTPN22* rs2476601 risk variant have produced a deluge of data, which paints a complex portrait of ways in which PTPN22 may contribute to autoimmunity. For this study we are specifically interested in examining the role that PTPN22 plays in regulating T cell receptor signaling. Multiple murine models of *Ptpn22* knockout or the risk variant KI produce a phenotype that results in hyper-active TCR signaling, while human carriers of the risk variant display reduced TCR signaling responses to antigen. We used our gene editing platform to demonstrate for the first time, the effects of PTPN22 ablation in primary $CD4^+$ human T cells and found that in this setting, human models of *PTPN22* disruption largely align with mouse models, showing hyper-active TCR signaling through increased calcium flux, expression of surface activation markers, and cytokine secretion. These findings implicate the role of PTPN22 to be similar between mice and human in regulation of the TCR. Therefore, it remains confounding that mouse models of the rs2476601 risk variant display such opposing T cell phenotypes from human carriers of the variant.

There is data to suggest that mouse models of knockout and the risk variant, though in some ways are similar, are not identical in their effect on cellular function. One study has used CRISPR/Cas9 gene editing to generate both *Ptpn22* knockout and risk variant mice on the NOD background, and examined T1D rates between the two[164]. While the risk variant indeed increased T1D rates in female mice and increases autoantibodies, the *Ptpn22* knockout NOD mice show a milder increase in T1D progression and no autoantibodies. This would suggest that risk variant mice are perhaps providing help to B cells in a manner that *Ptpn22* knockout mice are not. Also, previous work in conventionally generated *Ptpn22* risk and knockout models on a mixed background have shown through phospho-proteomic analysis that several key signaling proteins are more strongly phosphorylated in risk variant T cells than knockout T cells after TCR stimulation[67]. This makes logical sense, as the risk variant is not shown to

alter expression levels, but rather affect binding affinity with Csk and therefor alter protein localization[64]. All of this indicates that a better understanding of how the *PTPN22* risk variant contributes to autoimmunity in human T cells will require further use gene editing to modify the coding region of *PTPN22* to reflect the risk variant in previously non-risk donors. However, as described in chapter 3 there are technical challenges that come with any approach to the editing of a gene locus while trying to preserve the remainder of the natural genetics which may play a crucial role in the pathogenesis of the SNP. Preliminary data from using the ssODN editing strategy described in chapter 3 (See Figure 3-3), suggests that this SNP may reflect an intermediate loss-of-function phenotype compared to knockout and control edited isogenic populations, displaying trends toward increased CD25 and other activation markers upon TCR engagement. More research is required, and as this approach only edits a subset of non-risk alleles, producing mixed populations of non-risk, risk, and knockout alleles, it may ultimately be difficult to parse the effects of the risk variant from the knockout alleles. Other approaches to SNP editing discussed in chapter 3 are still under investigation and may indeed provide useful insights as to how the *PTPN22* rs2476601 variant impacts human T cell activation and regulation.

Similar to *PTPN22*, studies of *PTPN2* have shown contrasting data, but on a much more limited scale. The *PTPN2* rs1893217 risk SNP has been shown to predispose carriers to increased odds of several autoimmune diseases, while also reducing expression levels of PTPN2. Murine studies of *Ptpn2* T cell specific knockouts on even non-autoimmune prone backgrounds promote enhancements to signaling in effector T cells which is strong enough to drive autoimmunity while leaving Treg function unaffected[102,105]. Interestingly, *PTPN2* rs1893217 risk carriers have no published alterations to TCR signaling, and multiple studies show them to have a reduced response to IL-2[63,100,101]. Our *PTPN2* disrupted T cells supports a holistic view of this discrepancy, showing that *PTPN2* disruption in human T cells results in a dynamic phenotype. PTPN2 deficient gene edited cells have increased response to IL-2, IFNγ, and TCR stimulation which would predict increased T cell effector function and increased T cell help to B cells, and given its implication in diseases that feature autoantibody development, blockade of CD40-CD40L interaction may prove more efficacious in carriers of the *PTPN2* rs1893217 risk SNP[69]. However, over time and exposure to chronic cytokine signaling in our model, IL-2 responses normalize and

eventually decrease relative to *PTPN2* competent cells. IL-2 signaling is critical for Tregs and promising clinical trials have been reported demonstrating that low dose IL-2 therapy may increase Treg function in some autoimmune settings, while leaving T effector function unaltered[61], reducing disease severity. Our data would also suggest that *PTPN2* rs1893217 carriers may also benefit from low dose IL-2 treatment more than non-carriers in certain settings.

One limitation in our model of *PTPN2* loss, which potentially impacts its application to rs1893217 risk carriers, is that this SNP codes for a reduction in transcript expression, not an ablation of expression as our model creates. It is conceivable that these two situations produce different phenotypes, and thus would contribute differently to disease. However, recent work with T cell specific *Ptpn2* knockout on a NOD background shows that bi-allelic loss of *Ptpn2* drives severely accelerated diabetes through enhanced T cell signaling and autoantibody generation, while heterozygous disruption of *Ptpn2* in this model is still sufficient to cause an intermediate disease progression, with enhanced diabetes and autoantibody formation relative to *Ptpn2* wild type mice[105]. This would suggest that our gene editing models reflects an extreme example of phenotypes generated by the *PTPN2* 1893217 variant, but likely still highlights the pathways through which it contributes to disease nonetheless.

As alluded to in chapter 1, investigations into autoimmunity and cancer are becoming increasingly intertwined. In fact, some GWAS variants have been shown to protect against autoimmunity at the expense of increased rates of cancer, as with a variant in the gene Tyk2 for example[174,175]. Studies have identified PTPN2 as a potential target for immunotherapy, with CRISPR screens in cancer models demonstrating PTPN2 loss to increase cell responsiveness to IFNγ, and respond with increased expression of MHC molecules, resulting in improved antigen presentation and thus improved T cell responses[149]. Also, CRISPR knockout of *Ptpn2* in mouse hematopoietic stem cells results in improved clearance of implanted tumor cells and LCMV infection[176]. In human *PTPN2* disrupted T cells, we also saw increased responsiveness to IFNγ as measured by pSTAT1 response, in addition to increased TCR activation and IL-2 mediated pSTAT5 signaling. The increased responsiveness to both IFNγ and TCR signaling suggests that targeting PTPN2 may be beneficial in human immunotherapy applications. However, in our model chronic exposure to IL-2 family cytokines eventually resulted in the accumulation

of cells that were less responsive to IL-2. Even though we saw enhanced IFNγ responses in PTPN2 disrupted T cells concomitant with depressed IL-2 responses, further research will be necessary to investigate if prolonged exposure to IFNγ, in the absence of PTPN2, will similarly erode IFNγR signaling in cancer specific lymphocytes, potentially reducing their effectiveness through cell intrinsic mechanisms.

While the loss of response to IL-2 seen in *PTPN2* disrupted T cells correlated with an increase in SOCS3 expression, it is likely that expression of other regulatory proteins are also modulated in the absence of PTPN2 in a cell intrinsic effort to restrain cytokine signaling. For this study, such compensatory responses in primary human T cells informs us of a potential mechanism that unifies data generated in mouse and human based studies, which until now seemed contradictory. Also, this data highlights a second pathway in which perturbation to *PTPN2* could contribute to autoimmunity; one being increased response of effector T cells, and the other being dynamic alterations to IL-2 signaling, which would impact Treg function. This fits with data showing that peripheral blood from *PTPN2* rs1893217 risk carriers have reduced peripheral Treg numbers, which display reduced regulatory capacity[101]. However, our study shows in principle that extreme caution must be taken when modulating primary human T cell signaling. As we seek optimal targets for a new generation of immunologic therapies, inherent redundancies within key pathways may arise to produce outcomes opposing the desired effect.

In this study we have produced a high efficiency platform for the ablation of immune related genes in primary human T cells, as well as investigated a number of ways to alter the DNA coding sequence of autoimmune associated genes. Every approach for modeling SNPs in human T cells through gene editing has inherent strengths and weaknesses, but this type of approach to genotype-phenotype assessment of GWAS variants is still in its infancy. Currently, work is ongoing to utilize multiple CRISPR gRNAs to multiplex gene disruption and investigate additive effects of loss-of-function in key regulatory genes. Also, the introduction of flag tagged versions of *PTPN22* in primary T cells is in progress to potentially examine altered binding activity of PTPN22 when coding for risk or non-risk variants. While gene editing can potentially be a useful approach for the study of human disease related genetics, in the end, there will always be limitations to the use of gene editing of primary human cells. For example, *Ptpn22* has been shown in mouse models to differentially regulate the TCR depending upon antigen

experience of the cells[77], and the risk variant has been shown to impact lymphocyte development and antigen receptor repertoire[177]. These are factors which indicate that there may be developmental and context specific effects of the *PTPN22* risk variant which culminate in the phenotype seen in human carriers but cannot be captured in the current gene editing model. However, the overlapping phenotypes between murine T cells and gene edited human T cells we have shown possible in multiple genes would argue that study of GWAS related genetics though gene editing in primary human T cells should continue to be researched and can be a powerful new approach.

Materials and Methods

Human samples and primary T cell editing.

PBMCs were collected from whole blood of consenting donors and cryopreserved at the Fred Hutchinson Cancer Research Center. Upon thaw, total $CD4^+$ T cells were isolated by negative selection (EasySep $CD4^+$, Stemcell Tech.) and cultured in Roswell Park Memorial Institute (RPMI) 1640 media supplemented with 20% FBS, 1x Glutamax (Gibco), and 1mM HEPES (Gibco). Unless otherwise noted, Cells were cultured in 50ng/ml recombinant IL-2, 5ng/ml IL-7, and 5ng/ml IL-15 (Peprotech). After thaw cells were counted and cultured at 1 million/ml in flat bottom culture plates.

CRISPR/Cas9 and ssODN reagents.

CRISPR RNAs (crRNA) targeting *ZAP70*, *PTPN22*, *PTPN2*, and *CCR5* (sequences in Table ST1) were identified using the CCTop design tool [178] and the COSMID CRISPR design tool [179], and commercially synthesized by Integrated DNA Technologies (IDT). ssODNs were commercially synthesized by (IDT; Ultramer DNA Oligonucleotides) with phosphorothioate linkages between the first and final 3 base pair sequences. crRNA and trans-activating RNA (tracrRNA; IDT) were complexed at a 1:1 ratio, as per manufacturer's instructions. crRNA:tracrRNA complexes were mixed with Cas9 nuclease (IDT) at a 1.2:1 ratio and delivered with or without ssODNs to cells by Neon electroporation (Thermo Fisher Scientific).

AAV Vectors

All AAV donor templates designed for HDR experiments were cloned into AAV plasmid backbones as previously described [125,152]. AAV templates were modified to possess 800 bp homology arm sequences homologous to the PTPN22 G14 or CCR5 RNP cut site. AAV stocks were produced as previously described [110,125,152]. All AAVs used were of serotype 6.

Gene Editing

After thaw cells were activated with CD3/CD28 Activator Beads (Gibco). After 2 days beads were magnetically removed and cells re-plated without changing media or adjusting cell number. 24 hours later cells were electroporated with 2.5µg complexed RNP +/- ssODN.

Prior to electroporation, cells were washed with PBS and resuspended in Neon Buffer T. 2.5µg of complexed RNP and (if used) 20 pmol ssODN per 3×10^5 cells was added to the resuspension so that the final cell density was 3×10^7 cells/ml. Cells were electroporated (1400 V, 10ms, 3 pulses) in 10µl Neon tips, and then transferred into pre-warmed cell culture medium with IL-2, IL-7, and IL15 (unless otherwise noted). For samples transduced with AAV, virus was added to the culture immediately after electroporation at MOIs ranging from 5,000 to 20,000 and comprising no more than 20% of the total well volume.

After editing, cells were maintained in media identical to pre-editing conditions (unless otherwise noted). Cells were counted at least every two-days using Count Bright absolute counting beads (Thermo Fisher Scientific) and split to maintain cell density of 1 to 2 million/ml. Following expansion cells were counted, washed 3 times with phosphate buffered saline, and cultured at 1 million/ml for 24 hours in cytokine free media consisting of RPMI 1640 media with 10% FBS, 1x Glutamax (Thermo Fisher Scientific), and 1mM HEPES. Cells were re-counted prior to stimulation.

T7 assays and ICE sequencing analysis.

Gene disruption was analyzed using both the T7 endonuclease 1 assay and Inference of CRISPR Edits (ICE) analysis (Synthego). Total genomic DNA was isolated from 0.5 – $1x10^6$ cells using a DNeasy Blood & Tissue Kit (Qiagen). gRNA target genomic regions were first amplified PrimeSTAR GXL DNA Polymerase (Takara Bio) with primers creating a 400 to 700bp amplicon containing the gRNA target site (Table ST2). PCR amplicons were purified with Gene-jet PCR purification kit (Thermo Fisher Scientific).

For T7 assays 300ng of purified PCR product was denatured and re-annealed in 1x NEB Buffer 2 (New England Biolabs) in 19µl total volume, after which 10 U of T7 endonuclease I (New England Biolabs) was

added to the solution for 15 minutes at 37°C then stopped with 1ul of 0.5M EDTA. The reactions were then run on a 2.5% agarose gel for 1 hour and imaged. For ICE analysis [158], 25ng of purified PCR products were sanger sequenced using BigDye v.3.1 (Life Tech.). ab1 files were uploaded to https://ice.synthego.com/#/ for ICE analysis.

ddPCR

Quantification of HDR and NHEJ rates in edited human $CD4^+$ T cells was obtained using a droplet digital PCR, dual-probe competition assay. Primers and probes are listed in Table ST2. All probes were ordered from Sigma Aldrich with a 3' Black Hole 1 Quencher. Probes specific to sequences generated by HDR insertion of stop codons were labeled with a 5' FAM reporter and used in tandem with 5' HEX labeled probes specific to wild type (WT) sequences. Editing was measured after generating droplets with 50ng of genomic DNA (gDNA), both HDR-FAM and WT-HEX probes, and primers to the editing locus producing amplicons of <500bp (1× assay, 900-nM primers and 250-nM probe) using ddPCR supermix for probes (no deoxyuridine triphosphate [dUTP]) (Bio-Rad). Reference reactions were simultaneously performed using a 5' HEX labeled control probe targeting a sequence at least 40bp 5' of the RNP cut site and the same primers as the dual probe reaction. Droplets were generated with the QX200 Droplet Generator (Bio-Rad) and amplified. All samples were run in triplicate and averaged. Fluorescence was measured using the QX200 Droplet reader (Bio-Rad) and analyzed using Quantasoft software. Editing rates were calculated as the relative frequency (%) of FAM+ corresponding to %HDR, HEX+ corresponding to %No Event, and reference – (FAM+HEX) corresponding to %NHEJ.

Western blotting

Edited primary human $CD4^+$ T cells were lysed in 1x RIPA lysis buffer on ice for 10 minutes then clarified by centrifugation. Concentration of clarified lysate was determined by BCA assay (Pierce), diluted, and suspended in 1x LDS Sample Buffer (Invitrogen). 10µg of lysate was run on 4-12% Bis-Tris NuPAGE gels in 1x MOPS buffer (Invitrogen). Protein was transferred to nitrocellulose in 1x Transfer Buffer (Invitrogen)

and 10% methanol. Non-specific binding was minimized with a 1-hour RT incubation in Odessey LI-COR Blocking Buffer. Primary antibodies were stained at 1:1000 for at least 12 hours at 4°C, excluding PTPN22 which was stained at 1:3000 for at least 12 hours, and Actin, which was stained at 1:1000 at RT for 40 minutes. Primary antibodies used were from Cell Signaling Technology: ZAP70 (99F2), HSP90 (rabbit polyclonal, Cat.# 4874), and Actin (8H10D10); from R and D Systems: PTPN22 (goat polyclonal, Cat.# AF3428); and from Sigma-Aldrich: PTPN2 (rabbit polyclonal, Cat.# SAB4200249). After primary stain membranes were washed with 1x TBST and incubated with secondary antibodies at 1:10,000 for 30 minutes at RT. Stained blots were washed and imaged on an Odyssey Infrared Imaging System (LI-COR Biotech.). Western blot quantifications were performed with ImageJ software.

Plate bound anti-CD3 stimulation, ELISA, and Th1 skewing

Stimulation plates were made in 96-well flat bottom culture plate. 100ul of PBS supplemented with LEAF purified anti-CD3 (OKT3, Biolegend) at 0.25ug/ml was added to each well and incubated at least 12 hours at 4°C. The plate was then emptied, and wells were given 100ul of cytokine free T cell media. After cells were edited, expanded for 7 days, and rested 24 hours in cytokine free media, 100ul of cells at 2 million/ml were added to each well. Plates were incubated at 37°C for 24 to 48 hours.

Two days after stimulating edited cells with plate bound anti-CD3 as described, culture supernatants were collected, and cytokine secretion levels were determined by ELISA for IL-2 (Life Technologies, Cat.# 88-7025-86) and IFNγ (Life Technologies, Cat.# 88-7316-86). Supernatants were diluted 1:40, and all experiments followed manufacturer's protocols.

For Th1 skewing assays, cells were expanded and rested as normal, then stimulated with anti-CD3 plates as performed above. Media was also supplemented with 0.25ug/ml soluble anti-CD28 (CD28.2, Biolegend), 1ug/ml anti-IL-4 (MP4-25D2, Thermo Fisher), 5ng/ml IL-2 (Peprotech), and 10ng/ml IL-12 (Peprotech). Plates were incubated at 37°C for 72 hours. Before fixation cells were stimulated in the presence of 1ug/ml Golgi Stop (Biolegend) with 80nM PMA, and 1ug/ml Ionomycin at 37°C for 5 hours. Cells were then fixed with BD Cytofix/Cytoperm as per manufacture's protocols, and stained for INFγ.

Flow cytometry and gating strategies

Flow cytometric analysis was performed on an LSR II flow cytometer (BD Biosciences) and data was analyzed using FlowJo software (Tree Star). Cells were stained with LIVE/DEAD Fixable Near-IR Dead Cell Stain Kit, as per the manufacturer's instructions and cells were stained with fluorescence labeled antibodies for 30 minutes at 4°C. Antibodies used in this study include those from Biolegend: CD3 (SK7), CD4 (RPA-T4), CD69 (FN50), PD-1 (EH12.2H7), CD71 (CY1G4), CD40L/CD154 (24-31), INFγ (4S.B3); those from BD Biosciences: CD25 (2A3), pSTAT5 (pY694, 47), pCD3ζ/pCD247 (pY142, K25-407.69); and those from Miltenyi: pSTAT1 (pY701, REA345). All antibodies were used at a dilution of 1:100, except for those staining phospho-sites which were used at a dilution of 1:10. All antibody stains were 30 minutes on ice. Gating order proceeded: lymphocytes -> singlets -> live cells. For viability, %events that were live, single cells were reported. Surface stains of other markers were subsequently gated on $CD3^+/CD4^+$ cells, then the marker of interest.

Calcium flux was measured in edited, 7 day expanded, and 24-hour rested $CD4^+$ T cells that were incubated with indo-1 AM (Life Technologies) for 45 minutes at 37°C. Cells were then washed and resuspended in HBSS media with calcium and Magnesium stimulated with 5µg/ml (final concentration) OKT3 anti-CD3. Induction of Ca2+ mobilization was determined by flow cytometry.

For pSTAT1 or pSTAT5 staining, cells that were edited and rested 48 hours without cytokine, or expanded 7 days then rested 24 hours without cytokine. All cells were serum starved for 2 hours before receiving a 20-minute stimulation with 1.25ng/ml* of recombinant human IFNγ (Peprotech) or 0.5ng/ml* of recombinant human IL-2 (Peprotech) respectively (* - final concentration). Reactions were stopped by fixing cells with a final concentration of 2% PFA for 12 minutes at 37°C. Cells were then washed and permeabilized with BD Perm Buffer III for at least 30 minutes at -20°C. Cells were then washed wand stained as described above. For pCD3z/pCD247 staining edited/rested cells were serum starved for 1 hour, before being stained with either 0.1 or 1ug/ml of mouse anti-CD3 (Biolegend) for thirty minutes on ice. Cells were then washed and cross-linked with 0.2 or 2ug/ml goat anti-mouse Ig respectively

(Southern Bio.) for 0, 2, and 5 minutes. Reactions were stopped by fixing cells with a final concentration of 2% PFA for 12 minutes at 37°C. Cells were then washed and permeabilized with BD Perm Buffer I. Unoccupied GAM was blocked with mouse Ig for 15 minutes at RT, and cells were then stained for pCD3ζ for 30 minutes at RT.

FACS sorting of AAV edited cells

AAV edited cells were expanded for seven days in culture then rested 24 hours without cytokine. Cells were then bulk sorted based on editing outcome, using a FACS Aria I. Cells were gated on size and singlets, then sorted on BFP/GFP positivity. After sorting cells were expanded with CD3/CD28 Activator Beads (Gibco) at 1 bead to 50 cells (ratios of beads to cells higher than 1:25 caused severe activation induced cell death).

Quantitative RT-PCR

RNA was extracted from $1x10^6$ cells per sample with the RNeasy Kit (Qiagen) as per the manufacturer's protocol. cDNA was generated from RNA with Maxima First Strand cDNA Synthesis Kit (Thermo Fisher Scientific). Real-Time PCR was performed on the cDNA using iTaq Universal Syber Green Supermix (Bio-Rad) and a BioRad C1000 Thermal Cycler. A list of primers used can be found in Table ST3.

Statistics

Statistical analyses were performed using GraphPad Prism 7 (GraphPad). For all testing of gene edited cells, due to the low variability in culturing conditions and lack of obvious skewing, data was assumed to maintain a normal distribution. p-values in multiple comparisons were calculated using one-way ANOVA with the Tukey or Dunnet correction; p-values in comparisons between two groups were calculated using a paired two-tailed t test. For testing done with un-edited, genotyped human T cells, p-values were

calculated with Mann-Whitney tests. Values from combined independent experiments are shown as mean ± SEM.

Study Approval (Human Subjects)

For gene editing experiments, human donor leukopaks were purchased from the Fred Hutchinson Cancer Research Center, which were obtained from consenting donors under an IRB-approved protocol and cryopreserved. For experiments without gene editing, cryopreserved PBMCs were obtained from the Benaroya Research Institute (BRI) biorepository, collected under the BRI Immune Mediated Diseases IRB. Subjects were selected from the biorepository based on *PTPN2* genotype, the absence of autoimmune disease, and lack of autoimmunity in first-degree relatives. All PBMC donors provided written informed consent for the use of their tissues in research studies. After collection, all samples were de-identified for the protection of human PBMC donors.

References:

1 Silverstein, A. M. Horror autotoxicus, Autoimmunity, and Immunoregulation: The Early History. *Transfusion Medicine and Hemotherapy* **32**, 296-302, doi:10.1159/000089116 (2005).

2 Cooper, G. S., Bynum, M. L. K. & Somers, E. C. Recent Insights in the Epidemiology of Autoimmune Diseases: Improved Prevalence Estimates and Understanding of Clustering of Diseases. *Journal of autoimmunity* **33**, 197-207, doi:10.1016/j.jaut.2009.09.008 (2009).

3 Association, A. A. R. D. The cost burden of autoimmune disease: the latest front in the war on healthcare spending. *American Autoimmune Related Diseases Association, Eastpointe, Michigan, USA* (2011).

4 Patterson, C. C. *et al.* Trends in childhood type 1 diabetes incidence in Europe during 1989–2008: evidence of non-uniformity over time in rates of increase. *Diabetologia* **55**, 2142-2147, doi:10.1007/s00125-012-2571-8 (2012).

5 Ramos, P. S., Shedlock, A. M. & Langefeld, C. D. Genetics of autoimmune diseases: insights from population genetics. *Journal of Human Genetics* **60**, 657-664, doi:10.1038/jhg.2015.94 (2015).

6 Cho, J. H. & Feldman, M. Heterogeneity of autoimmune diseases: pathophysiologic insights from genetics and implications for new therapies. *Nature Medicine* **21**, 730-738, doi:10.1038/nm.3897 (2015).

7 Theofilopoulos, A. N., Kono, D. H. & Baccala, R. The multiple pathways to autoimmunity. *Nature Immunology* **18**, 716-724, doi:10.1038/ni.3731 (2017).

8 Rubtsova, K., Marrack, P. & Rubtsov, A. V. Sexual dimorphism in autoimmunity. *The Journal of Clinical Investigation* **125**, 2187-2193, doi:10.1172/JCI78082 (2015).

9 Grimaldi, C. M., Jeganathan, V. & Diamond, B. Hormonal regulation of B cell development: 17 beta-estradiol impairs negative selection of high-affinity DNA-reactive B cells at more than one developmental checkpoint. *Journal of Immunology (Baltimore, Md.: 1950)* **176**, 2703-2710, doi:10.4049/jimmunol.176.5.2703 (2006).

10 Markle, J. G. & Fish, E. N. SeXX matters in immunity. *Trends in Immunology* **35**, 97-104, doi:10.1016/j.it.2013.10.006 (2014).

11 Ter Horst, R. *et al.* Host and Environmental Factors Influencing Individual Human Cytokine Responses. *Cell* **167**, 1111-1124.e1113, doi:10.1016/j.cell.2016.10.018 (2016).

12 Maahs, D. M., West, N. A., Lawrence, J. M. & Mayer-Davis, E. J. Chapter 1: Epidemiology of Type 1 Diabetes. *Endocrinology and metabolism clinics of North America* **39**, 481-497, doi:10.1016/j.ecl.2010.05.011 (2010).

13 National Diabetes Statistics Report | Data & Statistics | Diabetes | CDC. (2019).

14 Morgan, A. W. *et al.* Reevaluation of the interaction between HLA-DRB1 shared epitope alleles, PTPN22, and smoking in determining susceptibility to autoantibody-positive and autoantibody-negative rheumatoid arthritis in a large UK Caucasian population. *Arthritis and Rheumatism* **60**, 2565-2576, doi:10.1002/art.24752 (2009).

15 Samara, K. D. *et al.* Upregulation of citrullination pathway: From Autoimmune to Idiopathic Lung Fibrosis. *Respiratory Research* **18**, doi:10.1186/s12931-017-0692-9 (2017).

16 Scher, J. U., Littman, D. R. & Abramson, S. B. Review: Microbiome in Inflammatory Arthritis and Human Rheumatic Diseases. *Arthritis & Rheumatology* **68**, 35-45, doi:10.1002/art.39259 (2016).

17 Stewart, C. J. *et al.* Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature* **562**, 583-588, doi:10.1038/s41586-018-0617-x (2018).

18 Schirmer, M. *et al.* Linking the Human Gut Microbiome to Inflammatory Cytokine Production Capacity. *Cell* **167**, 1125-1136.e1128, doi:10.1016/j.cell.2016.10.020 (2016).

19 Wen, L. *et al.* Innate immunity and intestinal microbiota in the development of Type 1 diabetes. *Nature* **455**, 1109-1113, doi:10.1038/nature07336 (2008).

20 Pauken, K. E., Dougan, M., Rose, N. R., Lichtman, A. H. & Sharpe, A. H. Adverse Events Following Cancer Immunotherapy: Obstacles and Opportunities. *Trends in Immunology* **40**, 511-523, doi:10.1016/j.it.2019.04.002 (2019).

21 Postow, M. A., Sidlow, R. & Hellmann, M. D. Immune-Related Adverse Events Associated with Immune Checkpoint Blockade. *New England Journal of Medicine* **378**, 158-168, doi:10.1056/NEJMra1703481 (2018).

22 Immune-Related Adverse Events Associated with Immune Checkpoint Blockade | NEJM. *New England Journal of Medicine*.

23 Wolchok, J. D. *et al.* Ipilimumab efficacy and safety in patients with advanced melanoma: a retrospective analysis of HLA subtype from four trials. *Cancer Immunity Archive* **10**, 9 (2010).

24 Quintana-Murci, L. Human Immunology through the Lens of Evolutionary Genetics. *Cell* **177**, 184-199, doi:10.1016/j.cell.2019.02.033 (2019).

25 Goris, A. & Liston, A. The Immunogenetic Architecture of Autoimmune Disease. *Cold Spring Harbor Perspectives in Biology* **4**, a007260, doi:10.1101/cshperspect.a007260 (2012).

26 Stranger, B. E., Stahl, E. A. & Raj, T. Progress and Promise of Genome-Wide Association Studies for Human Complex Trait Genetics. *Genetics* **187**, 367-383, doi:10.1534/genetics.110.120907 (2011).

27 Aslani, S., Rezaei, R., Jamshidi, A. & Mahmoudi, M. Genetic and epigenetic etiology of autoimmune diseases: lessons from twin studies. *Rheumatology Research* **3**, 45-57, doi:10.22631/rr.2018.69997.1041 (2018).

28 Bogdanos, D. P. *et al.* Twin studies in autoimmune disease: Genetics, gender and environment. *Journal of Autoimmunity* **38**, J156-J169, doi:10.1016/j.jaut.2011.11.003 (2012).

29 Redondo, M. J., Jeffrey, J., Fain, P. R., Eisenbarth, G. S. & Orban, T. Concordance for Islet Autoimmunity among Monozygotic Twins. *New England Journal of Medicine* **359**, 2849-2850, doi:10.1056/NEJMc0805398 (2008).

30 Bennett, C. L. *et al.* The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. *Nature Genetics* **27**, 20, doi:10.1038/83713 (2001).

31 Nagamine, K. *et al.* Positional cloning of the APECED gene. *Nature Genetics* **17**, 393, doi:10.1038/ng1297-393 (1997).

32 Cruz-Tapias, P., Castiblanco, J. & Anaya, J.-M. *HLA Association with Autoimmune Diseases*. (El Rosario University Press, 2013).

33 Bodis, G., Toth, V. & Schwarting, A. Role of Human Leukocyte Antigens (HLA) in Autoimmune Diseases. *Rheumatology and Therapy* **5**, 5-20, doi:10.1007/s40744-018-0100-z (2018).

34 Larsen, C. E. & Alper, C. A. The genetics of HLA-associated disease. *Current Opinion in Immunology* **16**, 660-667, doi:10.1016/j.coi.2004.07.014 (2004).

35 Gough, S. C. L. & Simmonds, M. J. The HLA Region and Autoimmune Disease: Associations and Mechanisms of Action. *Current Genomics* **8**, 453-465, doi:10.2174/138920207783591690 (2007).

36 Gutierrez-Arcelus, M., Rich, S. S. & Raychaudhuri, S. Autoimmune diseases — connecting risk alleles with molecular traits of the immune system. *Nature reviews. Genetics* **17**, 160-174, doi:10.1038/nrg.2015.33 (2016).

37 Sharma, A. *et al.* Identification of non-HLA genes associated with development of islet autoimmunity and type 1 diabetes in the prospective TEDDY cohort. *Journal of Autoimmunity* **89**, 90-100, doi:10.1016/j.jaut.2017.12.008 (2018).

38 Clarke, F. *et al.* Protein tyrosine phosphatase PTPN22 is dispensable for dendritic cell antigen processing and promotion of T-cell activation by dendritic cells. *PLOS ONE* **12**, e0186625, doi:10.1371/journal.pone.0186625 (2017).

39 Marson, A., Housley, W. J. & Hafler, D. A. Genetic basis of autoimmunity. *The Journal of Clinical Investigation* **125**, 2234-2241, doi:10.1172/JCI78086 (2015).

40 Farh, K. K.-H. *et al.* Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* **518**, 337-343, doi:10.1038/nature13835 (2015).

41 Chun, S. *et al.* Limited statistical evidence for shared genetic effects of eQTLs and autoimmune-disease-associated loci in three major immune-cell types. *Nature Genetics* **49**, 600, doi:10.1038/ng.3795 (2017).

42 Gamazon, E. R. *et al.* Using an atlas of gene regulation across 44 human tissues to inform complex disease- and trait-associated variation. *Nature Genetics* **50**, 956, doi:10.1038/s41588-018-0154-4 (2018).

43 Kim-Hellmuth, S. *et al.* Genetic regulatory effects modified by immune activation contribute to autoimmune disease associations. *Nature Communications* **8**, 266, doi:10.1038/s41467-017-00366-1 (2017).

44 Javierre, B. M. *et al.* Lineage-Specific Genome Architecture Links Enhancers and Non-coding Disease Variants to Target Gene Promoters. *Cell* **167**, 1369-1384.e1319, doi:10.1016/j.cell.2016.09.037 (2016).

45 Gallagher, M. D. & Chen-Plotkin, A. S. The Post-GWAS Era: From Association to Function. *The American Journal of Human Genetics* **102**, 717-730, doi:10.1016/j.ajhg.2018.04.002 (2018).

46 Cotsapas, C. *et al.* Pervasive Sharing of Genetic Effects in Autoimmune Disease. *PLOS Genetics* **7**, e1002254, doi:10.1371/journal.pgen.1002254 (2011).

47 Cotsapas, C. & Hafler, D. A. Immune-mediated disease genetics: the shared basis of pathogenesis. *Trends in Immunology* **34**, 22-26, doi:10.1016/j.it.2012.09.001 (2013).

48 Reyes, M. *et al.* Multiplexed enrichment and genomic profiling of peripheral blood cells reveal subset-specific immune signatures. *Science Advances* **5**, eaau9223, doi:10.1126/sciadv.aau9223 (2019).

49 Begovich, A. B. *et al.* A Missense Single-Nucleotide Polymorphism in a Gene Encoding a Protein Tyrosine Phosphatase (PTPN22) Is Associated with Rheumatoid Arthritis. *American Journal of Human Genetics* **75**, 330-337 (2004).

50 Rawlings, D. J., Dai, X. & Buckner, J. H. The Role of PTPN22 Risk Variant in the Development of Autoimmunity: Finding Common Ground between Mouse and Human. *The Journal of Immunology* **194**, 2977-2984, doi:10.4049/jimmunol.1403034 (2015).

51 Holmes, D. A. *et al.* Autoimmunity-associated protein tyrosine phosphatase PEP negatively regulates IFN-α receptor signaling. *Journal of Experimental Medicine* **212**, 1081-1093, doi:10.1084/jem.20142130 (2015).

52 Vlachogiannis, N. I. *et al.* Increased frequency of the PTPN22W* variant in primary Sjogren's Syndrome: Association with low type I IFN scores. *Clinical Immunology* **173**, 157-160, doi:10.1016/j.clim.2016.10.015 (2016).

53 Wang, Y. *et al.* PTPN22 Variant R620W Is Associated With Reduced Toll-like Receptor 7–Induced Type I Interferon in Systemic Lupus Erythematosus. *Arthritis & Rheumatology* **67**, 2403-2414, doi:10.1002/art.39211 (2015).

54 Chang, H.-H., Dwivedi, N., Nicholas, A. P. & Ho, I. C. The W620 Polymorphism in PTPN22 Disrupts Its Interaction With Peptidylarginine Deiminase Type 4 and Enhances Citrullination and NETosis. *Arthritis & Rheumatology* **67**, 2323-2334, doi:10.1002/art.39215 (2015).

55 Zhang, J. *et al.* The autoimmune disease-associated PTPN22 variant promotes calpain-mediated Lyp/Pep degradation associated with lymphocyte and dendritic cell hyperresponsiveness. *Nature Genetics* **43**, 902-907, doi:10.1038/ng.904 (2011).

56 Rieck, M. *et al.* Genetic Variation in PTPN22 Corresponds to Altered Function of T and B Lymphocytes. *The Journal of Immunology* **179**, 4704-4710, doi:10.4049/jimmunol.179.7.4704 (2007).

57 Giacomini, K. M. *et al.* Genome-wide association studies of drug response and toxicity: an opportunity for genome medicine. *Nature Reviews Drug Discovery* **16**, 70, doi:10.1038/nrd.2016.234 (2017).

58 Khera, A. V. *et al.* Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nature Genetics* **50**, 1219, doi:10.1038/s41588-018-0183-z (2018).

59 Bonifacio, E. *et al.* Genetic scores to stratify risk of developing multiple islet autoantibodies and type 1 diabetes: A prospective study in children. *PLOS Medicine* **15**, e1002548, doi:10.1371/journal.pmed.1002548 (2018).

60 Ali, M. A. *et al.* Metabolic and immune effects of immunotherapy with proinsulin peptide in human new-onset type 1 diabetes. *Science Translational Medicine* **9**, eaaf7779, doi:10.1126/scitranslmed.aaf7779 (2017).

61 Seelig, E. *et al.* The DILfrequency study is an adaptive trial to identify optimal IL-2 dosing in patients with type 1 diabetes. *JCI Insight* **3**, doi:10.1172/jci.insight.99306 (2018).

62 Herold, K. C. *et al.* An Anti-CD3 Antibody, Teplizumab, in Relatives at Risk for Type 1 Diabetes. *New England Journal of Medicine* **381**, 603-613, doi:10.1056/NEJMoa1902226 (2019).

63 Long, S. A. *et al.* An autoimmune-associated variant in PTPN2 reveals an impairment of IL-2R signaling in CD4+ T cells. *Genes & Immunity* **12**, 116-125, doi:10.1038/gene.2010.54 (2011).

64 Bottini, N. *et al.* A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes. *Nature Genetics* **36**, 337-338, doi:10.1038/ng1323 (2004).

65 Nauseef, W. M. & Borregaard, N. Neutrophils at work. *Nature Immunology* **15**, 602-611, doi:10.1038/ni.2921 (2014).

66 Chemin, K., Klareskog, L. & Malmström, V. Is rheumatoid arthritis an autoimmune disease? *Current Opinion in Rheumatology* **28**, 181-188, doi:10.1097/BOR.0000000000000253 (2016).

67 Dai, X. *et al.* A disease-associated PTPN22 variant promotes systemic autoimmunity in murine models. *The Journal of Clinical Investigation* **123**, 2024-2036, doi:10.1172/JCI66963 (2013).

68 Mahmoud, T. I. *et al.* Autoimmune manifestations in aged mice arise from early-life immune dysregulation. *Science Translational Medicine* **8**, 361ra137-361ra137, doi:10.1126/scitranslmed.aag0367 (2016).

69 Karnell, J. L. *et al.* A CD40L-targeting protein reduces autoantibodies and improves disease activity in patients with autoimmunity. *Science Translational Medicine* **11**, eaar6584, doi:10.1126/scitranslmed.aar6584 (2019).

70 Kyogoku, C. *et al.* Genetic Association of the R620W Polymorphism of Protein Tyrosine Phosphatase PTPN22 with Human SLE. *American Journal of Human Genetics* **75**, 504-507 (2004).

71 Zheng, J., Ibrahim, S., Petersen, F. & Yu, X. Meta-analysis reveals an association of PTPN22 C1858T with autoimmune diseases, which depends on the localization of the affected tissue. *Genes and Immunity* **13**, 641-652, doi:10.1038/gene.2012.46 (2012).

72 Stanford, S. M. & Bottini, N. PTPN22: the archetypal non-HLA autoimmunity gene. *Nature Reviews Rheumatology* **10**, 602-611, doi:10.1038/nrrheum.2014.109 (2014).

73 Cloutier, J. F. & Veillette, A. Association of inhibitory tyrosine protein kinase p50csk with protein tyrosine phosphatase PEP in T cells and other hemopoietic cells. *The EMBO Journal* **15**, 4909-4918 (1996).

74 Cohen, S., Dadi, H., Shaoul, E., Sharfe, N. & Roifman, C. M. Cloning and Characterization of a Lymphoid-Specific, Inducible Human Protein Tyrosine Phosphatase, Lyp. *Blood* **93**, 2013-2024 (1999).

75 Rhee, I. & Veillette, A. Protein tyrosine phosphatases in lymphocyte activation and autoimmunity. *Nature Immunology* **13**, 439-447, doi:10.1038/ni.2246 (2012).

76 Chakraborty, A. K. & Weiss, A. Insights into the initiation of TCR signaling. *Nature Immunology* **15**, 798-807, doi:10.1038/ni.2940 (2014).

77 Davidson, D. *et al.* The Csk-Associated Adaptor PAG Inhibits Effector T Cell Activation in Cooperation with Phosphatase PTPN22 and Dok Adaptors. *Cell Reports* **17**, 2776-2788, doi:10.1016/j.celrep.2016.11.035 (2016).

78 Brownlie, R. J. & Zamoyska, R. T cell receptor signalling networks: branched, diversified and bounded. *Nature Reviews Immunology* **13**, 257-269, doi:10.1038/nri3403 (2013).

79 Cloutier, J.-F. & Veillette, A. Cooperative Inhibition of T-Cell Antigen Receptor Signaling by a Complex between a Kinase and a Phosphatase. *Journal of Experimental Medicine* **189**, 111-121, doi:10.1084/jem.189.1.111 (1999).

80 Manz, B. N. *et al.* Small molecule inhibition of Csk alters affinity recognition by T cells. *eLife* **4**, e08088, doi:10.7554/eLife.08088 (2015).

81 Tan, Y. X. *et al.* Inhibition of the kinase Csk in thymocytes reveals a requirement for actin remodeling in the initiation of full TCR signaling. *Nature Immunology* **15**, 186-194, doi:10.1038/ni.2772 (2014).

82 Gaud, G., Lesourne, R. & Love, P. E. Regulatory mechanisms in T cell receptor signalling. *Nature Reviews Immunology* **18**, 485, doi:10.1038/s41577-018-0020-8 (2018).

83 Gregorieff, A., Cloutier, J. F. & Veillette, A. Sequence requirements for association of protein-tyrosine phosphatase PEP with the Src homology 3 domain of inhibitory tyrosine protein kinase p50(csk). *The Journal of Biological Chemistry* **273**, 13217-13222, doi:10.1074/jbc.273.21.13217 (1998).

84 Hasegawa, K. PEST Domain-Enriched Tyrosine Phosphatase (PEP) Regulation of Effector/Memory T Cells. *Science* **303**, 685-689, doi:10.1126/science.1092138 (2004).

85 Vang, T. *et al.* Autoimmune-associated lymphoid tyrosine phosphatase is a gain-of-function variant. *Nature Genetics* **37**, 1317-1319, doi:10.1038/ng1673 (2005).

86 Arechiga, A. F. *et al.* Cutting Edge: The PTPN22 Allelic Variant Associated with Autoimmunity Impairs B Cell Signaling. *Journal of immunology (Baltimore, Md. : 1950)* **182**, 3343-3347, doi:10.4049/jimmunol.0713370 (2009).

87 Menard, L. *et al.* The PTPN22 allele encoding an R620W variant interferes with the removal of developing autoreactive B cells in humans. *The Journal of Clinical Investigation* **121**, 3635-3644, doi:10.1172/JCI45790 (2011).

88 Vang, T. *et al.* The autoimmune-predisposing variant of lymphoid tyrosine phosphatase favors T helper 1 responses. *Human Immunology* **74**, 574-585, doi:10.1016/j.humimm.2012.12.017 (2013).

89 Perri, V. *et al.* Use of short interfering RNA delivered by cationic liposomes to enable efficient down-regulation of PTPN22 gene in human T lymphocytes. *PLOS ONE* **12**, e0175784, doi:10.1371/journal.pone.0175784 (2017).

90 Bray, C. *et al.* Crispr/Cas Mediated Deletion of PTPN22 in Jurkat T Cells Enhances TCR Signaling and Production of IL-2. *Frontiers in Immunology* **9**, doi:10.3389/fimmu.2018.02595 (2018).

91 Todd, J. A. *et al.* Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. *Nature Genetics* **39**, 857-864, doi:10.1038/ng2068 (2007).

92 The Wellcome Trust Case Control, C. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661-678, doi:10.1038/nature05911 (2007).

93 Thompson, S. D. *et al.* The susceptibility loci juvenile idiopathic arthritis shares with other autoimmune diseases extend to PTPN2, COG6, and ANGPT1. *Arthritis & Rheumatism* **62**, 3265-3276, doi:10.1002/art.27688 (2010).

94 Marcil, V. *et al.* Association Between the PTPN2 Gene and Crohn's DiseaseDissection of Potential Causal Variants. *Inflammatory Bowel Diseases* **19**, 1149-1155, doi:10.1097/MIB.0b013e318280b181 (2013).

95 You-Ten, K. E. *et al.* Impaired Bone Marrow Microenvironment and Immune Function in T Cell Protein Tyrosine Phosphatase–deficient Mice. *The Journal of Experimental Medicine* **186**, 683-693, doi:10.1084/jem.186.5.683 (1997).

96 Simoncic, P. D., Lee-Loy, A., Barber, D. L., Tremblay, M. L. & McGlade, C. J. The T Cell Protein Tyrosine Phosphatase Is a Negative Regulator of Janus Family Kinases 1 and 3. *Current Biology* **12**, 446-453, doi:10.1016/S0960-9822(02)00697-8 (2002).

97 Heinonen, K. M. T-cell protein tyrosine phosphatase deletion results in progressive systemic inflammatory disease. *Blood* **103**, 3457-3464, doi:10.1182/blood-2003-09-3153 (2004).

98 Villarino, A. V., Kanno, Y. & O'Shea, J. J. Mechanisms and consequences of Jak–STAT signaling in the immune system. *Nature Immunology* **18**, 374-384, doi:10.1038/ni.3691 (2017).

99 Zhu, W., Mustelin, T. & David, M. Arginine Methylation of STAT1 Regulates Its Dephosphorylation by T Cell Protein Tyrosine Phosphatase. *Journal of Biological Chemistry* **277**, 35787-35790, doi:10.1074/jbc.C200346200 (2002).

100 Cerosaletti, K. *et al.* Multiple Autoimmune-Associated Variants Confer Decreased IL-2R Signaling in CD4+CD25hi T Cells of Type 1 Diabetic and Multiple Sclerosis Patients. *PLOS ONE* **8**, e83811, doi:10.1371/journal.pone.0083811 (2013).

101 Yang, J. H. M. *et al.* Natural Variation in Interleukin-2 Sensitivity Influences Regulatory T-Cell Frequency and Function in Individuals With Long-standing Type 1 Diabetes. *Diabetes* **64**, 3891-3902, doi:10.2337/db15-0516 (2015).

102 Wiede, F. *et al.* T cell protein tyrosine phosphatase attenuates T cell signaling to maintain tolerance in mice. *Journal of Clinical Investigation* **121**, 4758-4774, doi:10.1172/JCI59492 (2011).

103 Wiede, F., La Gruta, N. L. & Tiganis, T. PTPN2 attenuates T-cell lymphopenia-induced proliferation. *Nature Communications* **5**, 3073, doi:10.1038/ncomms4073 (2014).

104 Kleppe, M. *et al.* Deletion of the protein tyrosine phosphatase gene PTPN2 in T-cell acute lymphoblastic leukemia. *Nature Genetics* **42**, 530-535, doi:10.1038/ng.587 (2010).

105 Wiede, F. *et al.* T Cell-Specific PTPN2-Deficiency in NOD Mice Accelerates the Development of Type 1 Diabetes and Autoimmune Co-Morbidities. *Diabetes*, db181362, doi:10.2337/db18-1362 (2019).

106 Orrú, V. *et al.* A loss-of-function variant of PTPN22 is associated with reduced risk of systemic lupus erythematosus. *Human Molecular Genetics* **18**, 569-579, doi:10.1093/hmg/ddn363 (2009).

107 Rodríguez-Rodríguez, L. *et al.* The PTPN22 R263Q polymorphism is a risk factor for rheumatoid arthritis in Caucasian case–control samples. *Arthritis & Rheumatism* **63**, 365-372, doi:10.1002/art.30145 (2011).

108 Hendel, A. *et al.* Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nature Biotechnology* **33**, 985-989, doi:10.1038/nbt.3290 (2015).

109 Schumann, K. *et al.* Generation of knock-in primary human T cells using Cas9 ribonucleoproteins. *Proceedings of the National Academy of Sciences* **112**, 10437-10442, doi:10.1073/pnas.1512503112 (2015).

110 Khan, I. F., Hirata, R. K. & Russell, D. W. AAV-mediated gene targeting methods for human cells. *Nature Protocols* **6**, 482-501, doi:10.1038/nprot.2011.301 (2011).

111 Bibikova, M. *et al.* Stimulation of Homologous Recombination through Targeted Cleavage by Chimeric Nucleases. *Molecular and Cellular Biology* **21**, 289-297, doi:10.1128/MCB.21.1.289-297.2001 (2001).

112 Porteus, M. H. & Baltimore, D. Chimeric Nucleases Stimulate Gene Targeting in Human Cells. *Science* **300**, 763-763, doi:10.1126/science.1078395 (2003).

113 Sander, J. D. & Joung, J. K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature Biotechnology* **32**, 347-355, doi:10.1038/nbt.2842 (2014).

114 Gaj, T., Gersbach, C. A. & Barbas Iii, C. F. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in Biotechnology* **31**, 397-405, doi:10.1016/j.tibtech.2013.04.004 (2013).

115 Su, S. *et al.* CRISPR-Cas9 mediated efficient PD-1 disruption on human primary T cells from cancer patients. *Scientific Reports* **6**, doi:10.1038/srep20070 (2016).

116 Bak, R. O., Gomez-Ospina, N. & Porteus, M. H. Gene Editing on Center Stage. *Trends in Genetics* **0**, doi:10.1016/j.tig.2018.05.004

10.1016/j.stemcr.2017.08.026</p> (2018).

117 Bednarski, J. J. & Sleckman, B. P. At the intersection of DNA damage and immune responses. *Nature Reviews Immunology* **19**, 231, doi:10.1038/s41577-019-0135-6 (2019).

118 Schwertman, P., Bekker-Jensen, S. & Mailand, N. Regulation of DNA double-strand break repair by ubiquitin and ubiquitin-like modifiers. *Nature Reviews Molecular Cell Biology* **17**, 379-394, doi:10.1038/nrm.2016.58 (2016).

119 Chang, H. H. Y., Pannunzio, N. R., Adachi, N. & Lieber, M. R. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nature Reviews Molecular Cell Biology* **18**, 495-506, doi:10.1038/nrm.2017.48 (2017).

120 Biehs, R. *et al.* DNA Double-Strand Break Resection Occurs during Non-homologous End Joining in G1 but Is Distinct from Resection during Homologous Recombination. *Molecular Cell* **65**, 671-684.e675, doi:10.1016/j.molcel.2016.12.016 (2017).

121 Anand, R., Ranjha, L., Cannavo, E. & Cejka, P. Phosphorylated CtIP Functions as a Co-factor of the MRE11-RAD50-NBS1 Endonuclease in DNA End Resection. *Molecular Cell* **64**, 940-950, doi:10.1016/j.molcel.2016.10.017 (2016).

122 Kozhaya, L. *et al.* Functional Interrogation of Primary Human T Cells via CRISPR Genetic Editing. *The Journal of Immunology*, ji1701616, doi:10.4049/jimmunol.1701616 (2018).

123 Ehrke-Schulz, E. *et al.* CRISPR/Cas9 delivery with one single adenoviral vector devoid of all viral genes. *Scientific Reports* **7**, 17113, doi:10.1038/s41598-017-17180-w (2017).

124 Hultquist, J. F. *et al.* A Cas9 Ribonucleoprotein Platform for Functional Genetic Studies of HIV-Host Interactions in Primary Human T Cells. *Cell Reports* **17**, 1438-1452, doi:10.1016/j.celrep.2016.09.080 (2016).

125 Hung, K. L. *et al.* Engineering Protein-Secreting Plasma Cells by Homology-Directed Repair in Primary Human B Cells. *Molecular Therapy* **26**, 456-467, doi:10.1016/j.ymthe.2017.11.012 (2018).

126 Gaj, T. *et al.* Targeted gene knock-in by homology-directed genome editing using Cas9 ribonucleoprotein and AAV donor delivery. *Nucleic Acids Research* **45**, e98-e98, doi:10.1093/nar/gkx154 (2017).

127 Xu, X. *et al.* Efficient homology-directed gene editing by CRISPR/Cas9 in human stem and primary cells using tube electroporation. *Scientific Reports* **8**, 11649, doi:10.1038/s41598-018-30227-w (2018).

128 Kotterman, M. A. & Schaffer, D. V. Engineering adeno-associated viruses for clinical gene therapy. *Nature Reviews Genetics* **15**, 445-451, doi:10.1038/nrg3742 (2014).

129 Thomas, C. E., Ehrhardt, A. & Kay, M. A. Progress and problems with the use of viral vectors for gene therapy. *Nature Reviews Genetics* **4**, 346-358, doi:10.1038/nrg1066 (2003).

130 Pillay, S. *et al.* An essential receptor for adeno-associated virus infection. *Nature* **530**, 108-112, doi:10.1038/nature16465 (2016).

131 Chiang, C. & Gack, M. U. Post-translational Control of Intracellular Pathogen Sensing Pathways. *Trends in Immunology* **38**, 39-52, doi:10.1016/j.it.2016.10.008 (2017).

132 Herzner, A.-M. *et al.* Sequence-specific activation of the DNA sensor cGAS by Y-form DNA structures as found in primary HIV-1 cDNA. *Nature Immunology* **16**, 1025-1033, doi:10.1038/ni.3267 (2015).

133 Dowdy, S. F. Overcoming cellular barriers for RNA therapeutics. *Nature Biotechnology* **35**, 222-229, doi:10.1038/nbt.3802 (2017).

134 Burdette, D. L. & Vance, R. E. STING and the innate immune response to nucleic acids in the cytosol. *Nature Immunology* **14**, 19-26, doi:10.1038/ni.2491 (2013).

135 Roth, T. L. *et al.* Reprogramming human T cell function and specificity with non-viral genome targeting. *Nature* **559**, 405, doi:10.1038/s41586-018-0326-5 (2018).

136 Yin, H., Kauffman, K. J. & Anderson, D. G. Delivery technologies for genome editing. *Nature Reviews Drug Discovery* **16**, 387-399, doi:10.1038/nrd.2016.280 (2017).

137 Forjanič, T. *et al.* Electroporation induced stress response and its effect on gene electrotransfer efficacy: in vivo imaging and numerical modelling. *IEEE Transactions on Biomedical Engineering*, 1-1, doi:10.1109/TBME.2019.2894659 (2019).

138 Haapaniemi, E., Botla, S., Persson, J., Schmierer, B. & Taipale, J. CRISPR–Cas9 genome editing induces a p53-mediated DNA damage response. *Nature Medicine*, 1, doi:10.1038/s41591-018-0049-z (2018).

139 Ihry, R. J. *et al.* p53 inhibits CRISPR–Cas9 engineering in human pluripotent stem cells. *Nature Medicine*, 1, doi:10.1038/s41591-018-0050-6 (2018).

140 Heninger, A.-K. *et al.* A divergent population of autoantigen-responsive CD4+ T cells in infants prior to β cell autoimmunity. *Science Translational Medicine* **9**, eaaf8848, doi:10.1126/scitranslmed.aaf8848 (2017).

141 Sanchez-Blanco, C. *et al.* Protein tyrosine phosphatase PTPN22 regulates LFA-1 dependent Th1 responses. *Journal of Autoimmunity* **94**, 45-55, doi:10.1016/j.jaut.2018.07.008 (2018).

142 Salmond, R. J., Brownlie, R. J., Morrison, V. L. & Zamoyska, R. The tyrosine phosphatase PTPN22 discriminates weak self peptides from strong agonist TCR signals. *Nature Immunology* **advance online publication**, doi:10.1038/ni.2958 (2014).

143 Zikherman, J. & Weiss, A. Unraveling the functional implications of GWAS: how T cell protein tyrosine phosphatase drives autoimmune disease. *The Journal of Clinical Investigation* **121**, 4618-4621, doi:10.1172/JCI60001 (2011).

144 Cerosaletti, K. & Buckner, J. H. Protein tyrosine phosphatases and type 1 diabetes: genetic and functional implications of PTPN2 and PTPN22. *The review of diabetic studies : RDS* **9**, 188-200, doi:10.1900/rds.2012.9.188 (2012).

145 Jonkers, I. H. & Wijmenga, C. Context-specific effects of genetic variants associated with autoimmune disease. *Human Molecular Genetics* **26**, R185-R192, doi:10.1093/hmg/ddx254 (2017).

146 Zikherman, J. *et al.* PTPN22 deficiency cooperates with the CD45 E613R allele to break tolerance on a non-autoimmune background. *Journal of immunology (Baltimore, Md. : 1950)* **182**, 4093-4106, doi:10.4049/jimmunol.0803317 (2009).

147 Maine, C. J., Teijaro, J. R., Marquardt, K. & Sherman, L. A. PTPN22 contributes to exhaustion of T lymphocytes during chronic viral infection. *Proceedings of the National Academy of Sciences of the United States of America* **113**, E7231-E7239, doi:10.1073/pnas.1603738113 (2016).

148 Jofra, T. *et al.* Extrinsic Protein Tyrosine Phosphatase Non-Receptor 22 Signals Contribute to CD8 T Cell Exhaustion and Promote Persistence of Chronic Lymphocytic Choriomeningitis Virus Infection. *Frontiers in Immunology* **8**, doi:10.3389/fimmu.2017.00811 (2017).

149 Manguso, R. T. *et al.* In vivo CRISPR screening identifies Ptpn2 as a cancer immunotherapy target. *Nature* **547**, 413-418, doi:10.1038/nature23270 (2017).

150 Hale, M. *et al.* Engineering HIV-Resistant, Anti-HIV Chimeric Antigen Receptor T Cells. *Molecular Therapy: The Journal of the American Society of Gene Therapy* **25**, 570-579, doi:10.1016/j.ymthe.2016.12.023 (2017).

151 Hubbard, N. *et al.* Targeted gene editing restores regulated CD40L expression and function in X-HIGM T cells. *Blood*, blood-2015-2011-683235, doi:10.1182/blood-2015-11-683235 (2016).

152 Sather, B. D. *et al.* Efficient modification of CCR5 in primary human hematopoietic cells using a megaTAL nuclease and AAV donor template. *Science Translational Medicine* **7**, 307ra156-307ra156, doi:10.1126/scitranslmed.aac5530 (2015).

153 Noel, S., Lee, S. A., Sadasivam, M., Hamad, A. R. A. & Rabb, H. KEAP1 Editing Using CRISPR/Cas9 for Therapeutic NRF2 Activation in Primary Human T Lymphocytes. *The Journal of Immunology*, ji1700812, doi:10.4049/jimmunol.1700812 (2018).

154 Shifrut, E. *et al.* Genome-wide CRISPR Screens in Primary Human T Cells Reveal Key Regulators of Immune Function. *Cell* **175**, 1958-1971.e1915, doi:10.1016/j.cell.2018.10.024 (2018).

155 Elder, M. E. *et al.* Human severe combined immunodeficiency due to a defect in ZAP-70, a T cell tyrosine kinase. *Science* **264**, 1596-1599, doi:10.1126/science.8202712 (1994).

156 Kadlecek, T. A. *et al.* Differential Requirements for ZAP-70 in TCR Signaling and T Cell Development. *The Journal of Immunology* **161**, 4688-4694 (1998).

157 Richardson, C. D., Ray, G. J., Bray, N. L. & Corn, J. E. Non-homologous DNA increases gene disruption efficiency by altering DNA repair outcomes. *Nature Communications* **7**, 12463, doi:10.1038/ncomms12463 (2016).

158 Hsiau, T. *et al.* Inference of CRISPR Edits from Sanger Trace Data. *bioRxiv*, 251082, doi:10.1101/251082 (2018).

159 Wienert, B., Shin, J., Zelin, E., Pestal, K. & Corn, J. E. In vitro–transcribed guide RNAs trigger an innate immune response via the RIG-I pathway. *PLOS Biology* **16**, e2005840, doi:10.1371/journal.pbio.2005840 (2018).

160 Seki, A. & Rutz, S. Optimized RNP transfection for highly efficient CRISPR/Cas9-mediated gene knockout in primary T cells. *Journal of Experimental Medicine* **215**, 985-997, doi:10.1084/jem.20171626 (2018).

161 Gwiazda, K. S. *et al.* High Efficiency CRISPR/Cas9-mediated Gene Editing in Primary Human T-cells Using Mutant Adenoviral E4orf6/E1b55k “Helper” Proteins. *Molecular Therapy* **24**, 1570-1580, doi:10.1038/mt.2016.105 (2016).

162 Svensson, M. N. D. *et al.* Reduced expression of phosphatase PTPN2 promotes pathogenic conversion of Tregs in autoimmunity. *The Journal of Clinical Investigation* **129**, 1193-1210, doi:10.1172/JCI123267 (2019).

163 Pillemer, B. B. L., Xu, H., Oriss, T. B., Qi, Z. & Ray, A. Deficient SOCS3 expression in CD4+CD25+FoxP3+ regulatory T cells and SOCS3-mediated suppression of Treg function. *European Journal of Immunology* **37**, 2082-2089, doi:10.1002/eji.200737193 (2007).

164 Lin, X. *et al.* CRISPR-Cas9–Mediated Modification of the NOD Mouse Genome With Ptpn22R619W Mutation Increases Autoimmune Diabetes. *Diabetes* **65**, 2134-2138, doi:10.2337/db16-0061 (2016).

165 Kan, Y., Ruis, B., Takasugi, T. & Hendrickson, E. A. Mechanisms of precise genome editing using oligonucleotide donors. *Genome Research* **27**, 1099-1111, doi:10.1101/gr.214775.116 (2017).

166 Diaz-Gallo, L. M. *et al.* Differential association of two PTPN22 coding variants with Crohn's disease and ulcerative colitis. *Journal of Translational Medicine* **8**, P2, doi:10.1186/1479-5876-8-S1-P2 (2010).

167 Diaz-Gallo, L.-M. & Martin, J. PTPN22 splice forms: a new role in rheumatoid arthritis. *Genome Medicine* **4**, 13, doi:10.1186/gm312 (2012).

168 Zufferey, R., Donello, J. E., Trono, D. & Hope, T. J. Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element Enhances Expression of Transgenes Delivered by Retroviral Vectors. *Journal of Virology* **73**, 2886-2892 (1999).

169 Harmsen, T., Klaasen, S., van de Vrugt, H. & te Riele, H. DNA mismatch repair and oligonucleotide end-protection promote base-pair substitution distal from a CRISPR/Cas9-induced DNA break. *Nucleic Acids Research* **46**, 2945-2955, doi:10.1093/nar/gky076 (2018).

170 Bak, R. O. *et al.* Multiplexed genetic engineering of human hematopoietic stem and progenitor cells using CRISPR/Cas9 and AAV6. *eLife* **6**, doi:10.7554/eLife.27873 (2017).

171 Tubo, N. J. & Jenkins, M. K. TCR signal quantity and quality in CD4+ T cell differentiation. *Trends in Immunology* **35**, 591-596, doi:10.1016/j.it.2014.09.008 (2014).

172 Pociot, F. & Lernmark, Å. Genetic risk factors for type 1 diabetes. *Lancet (London, England)* **387**, 2331-2339, doi:10.1016/S0140-6736(16)30582-7 (2016).

173 Simeonov, D. R. *et al.* Discovery of stimulation-responsive immune enhancers with CRISPR activation. *Nature* **549**, 111-115, doi:10.1038/nature23875 (2017).

174 Diogo, D. *et al.* TYK2 Protein-Coding Variants Protect against Rheumatoid Arthritis and Autoimmunity, with No Evidence of Major Pleiotropic Effects on Non-Autoimmune Complex Traits. *PLoS ONE* **10**, e0122271, doi:10.1371/journal.pone.0122271 (2015).

175 Strobl, B., Stoiber, D., Sexl, V. & Mueller, M. Tyrosine kinase 2 (TYK2) in cytokine signalling and host immunity. *Frontiers in Bioscience (Landmark Edition)* **16**, 3214-3232 (2011).

176 LaFleur, M. W. *et al.* A CRISPR-Cas9 delivery system for in vivo screening of genes in the immune system. *Nature Communications* **10**, 1668, doi:10.1038/s41467-019-09656-2 (2019).

177 Metzler, M. E. & Rawlings, D. J. 1 online resource (79 pages) (University of Washington Libraries ,, Seattle, 2017).

178 Stemmer, M., Thumberger, T., Keyer, M. d. S., Wittbrodt, J. & Mateo, J. L. CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. *PLOS ONE* **10**, e0124633, doi:10.1371/journal.pone.0124633 (2015).

179 Cradick, T. J., Qiu, P., Lee, C. M., Fine, E. J. & Bao, G. COSMID: A Web-based Tool for Identifying and Validating CRISPR/Cas Off-target Sites. *Molecular Therapy. Nucleic Acids* **3**, e214, doi:10.1038/mtna.2014.64 (2014).