nature portfolio

https://doi.org/10.1038/s41588-022-01106-y

Supplementary information

Systematic discovery and perturbation of regulatory genes in human T cells reveals the architecture of immune networks

In the format provided by the authors and unedited

Supplementary note

From subsection: **Discovery of upstream regulators of IL2RA, IL-2, and CTLA4**

To confirm functional effects of SLICE editing, we initially compared the abundance of sgRNAs in the starting plasmid library to their abundance in the transduced GFP+ T cells. sgRNAs targeting either GFP or known essential genes were highly depleted in the GFP+ sorted population (Extended Data Fig. 1b and 1d). Although we were focused on discovering regulators of IL2RA, IL-2, and CTLA4, we also identified a number of genes important for T cell proliferation or survival even in the short six-day screen window by comparing the sgRNA abundance between the GFP+ sorted population and the starting plasmid library (Extended Data Fig. 1e). sgRNAs targeting the tumor suppressors PTEN and P53[1] were overrepresented, while sgRNAs targeting components of the IL-2 signaling pathway (JAK3, IL2RA, IL2RB, STAT5B), which is important for T cell fitness[2–4], were depleted (Extended Data Fig. 1e). As an additional quality control, we confirmed that the screen hits are expressed in CD4+ T cells (Extended Data Fig. 1c). Taken together, these results confirmed that we were able to uncover important regulators of cell state with the pooled perturbations.

From subsection: **Arrayed KOs validate and characterize screen results**

We selected 50 hits from the three screens, 3 positive controls, and 4 non-targeting controls for the arrayed KOs. Thirty-nine of these genes regulate at least two of the three target genes, while the remaining genes were selected because they had large effects in an individual screen. For each gene, we used two guide RNAs per target and three donors per guide. We genotyped each KO sample using amplicon sequencing[5]. In many cases, over 90% of the amplicons contained insertions or deletions demonstrating highly efficient gene KO which enabled us to assess altered cell phenotypes (Fig. 3c and Extended Data Fig. 2c-d). Five days after RNP delivery we stained the cells and performed high-throughput flow cytometry to measure how each KO affected the protein levels of IL2RA, IL-2, and CTLA4.

Our pooled screens measured relative shifts in sgRNA frequency in the tails of the distribution of protein expression of IL2RA, IL-2, and CTLA4; the arrayed flow cytometry extended these measurements to show how each KO affected the full distribution of IL2RA, IL-2, and CTLA4 protein levels. Representative flow plots show that the KO of CBFP and STAT5B increased and decreased the protein levels of IL2RA respectively (Fig. 3b). To summarize the results, we normalized median fluorescence intensity of each KO to the non-targeting controls (Fig. 3c and Extended Data Fig. 2c-d).

Additionally, since T cells are cultured with IL-2 and IL-2 regulates IL2RA expression, we tested whether any of the top regulators of IL2RA act downstream of IL-2 signaling. We knocked out top regulators of IL2RA and blocked IL-2 signaling with an anti-IL-2 antibody. While blocking IL-2 signaling reduced IL2RA levels overall, IL2RA levels in the CBFB and KLF2 KOs were still higher than their donor matched control KOs suggesting that these genes normally negatively regulate IL2RA expression at least partially independent of IL-2 signaling (Extended Data Fig.

2e-f). Additionally, blocking IL-2 signaling in the IRF4 KOs reduced IL2RA levels in 2 out of 3 donors compared to their donor matched control KOs, suggesting that normally IRF4 positively regulates IL2RA at least partially independent of IL-2 signaling. However, IL2RA levels were reduced to similar levels in JAK3 and STAT5B KOs in either exogenous IL-2 or blocking IL-2 conditions, consistent with the fact that these genes are downstream of IL-2 signaling[4].

From subsection: **Mapping downstream target genes and CREs of IL2RA regulators**

We first confirmed that the RNA-Seq data revealed meaningful changes in the CRISPR KOs. Consistent with the genotyping results that showed high rates of insertion/deletion mutations at the CRISPR target sites (Fig. 3c), RNA-Seq demonstrated that the expression of most targeted regulators was decreased in their respective KO RNA-Seq samples (Extended Data Fig. 3a). The few examples where this was not the case could be due to a lack of nonsense-mediated decay and/or feedback mechanisms on transcription. Although our screen measured IL2RA protein levels, most of the regulators that we identified also affect *IL2RA* transcript levels. Flow cytometry of IL2RA protein levels in the same cell populations collected for RNA-Seq revealed that changes in *IL2RA* mRNA levels and protein levels were highly correlated (Extended Data Fig. 3b). These results confirmed that the regulatory effects of KOs can be ascertained from the RNA-Seq data.

CRISPR KO also affected the global chromatin landscape as assessed by ATAC-Seq. Sites of chromatin accessibility that were altered upon ablation of a specific transcription factor tended to be enriched for the corresponding binding sequence motif (see methods). We analyzed how many differential ATAC-Seq peaks contained a motif associated with the ablated transcription factor. Across these knockouts, the percent of significantly changed ATAC-Seq peaks that contained a matching motif ranged from 10-80% (median 42%), suggesting that we captured many primary regulatory effects (Extended Data Fig. 3c). These data implicate specific transcription factors required for maintaining chromatin accessibility in human CD4+ T cells at sites throughout the genome.

From subsection: **Methods**

<u>Pooled sgRNA library construction</u>
We selected transcription factors (TFs) with known or inferred motifs from Lambert et al.[6], non-target controls from the Brunello sgRNA library[7], hits from a previous screen[8] and immune genes of interest from the lab. All sgRNA sequences were taken from the Brunello sgRNA library[7]. In total we included 1349 genes with an average of 4 guides per gene, 13 guides against GFP as a positive control for editing, and 593 non-targeting controls. We ordered a pooled oligo library from Twist Biosciences with flanking sequences for cloning into the plasmid backbone. Following the custom sgRNA library cloning protocol as described by Joung et al.[9], we integrated our sgRNA library into the LRG2.1 backbone featuring an improved guide RNA scaffold (Addgene, plasmid# 108098) from Grevet et al.[10] using NEBuilder HiFi DNA Assembly master mix (NEB, Cat #E2621X) according to the manufacturer's protocol. We used Endura

ElectroCompetent Cells to amplify the library per the manufacturer's protocol (Endura, Cat #60242-1). Finally, we performed maxipreps using the ZymoPure Plasmid Maxiprep kit (Zymo, Cat # D4202).

Lentiviral production

HEK 293T cells were seeded at 14 million cells in a 15 cm tissue culture treated culture dish (Corning, Cat #430599) in Opti-MEM (UCSF CCF, Cat #CCFAC008) 24 hours prior to transfection. Using Lipofectamine 3000 (Lifetech, Cat #L3000075) according to the manufacturer's protocol, cells were transfected with the sgRNA transfer plasmid, and two lentiviral packaging plasmids, pMD2.G (Addgene, Cat #12259) and psPAX2 (Addgene, Cat #12260). Cells were incubated for 5 hours at 37°C, after which time the transfection media was replaced with fresh Opti-MEM containing ViralBoost at 1x (Alstem, Cat #VB100). Cells were incubated for 24 hours and then the viral supernatant was collected and spun down at 300*g* for 5 minutes to remove cellular debris. The supernatant was then passed through a 0.45-µm filter, and subsequently mixed with one volume of cold Lentivirus Precipitation Solution (Alstem, Cat #VC125) at 4° C to every 4 volumes of lentivirus-containing supernatant. Samples were mixed well and placed at 4°C overnight. The virus was then concentrated by centrifugation at 1500*g* for 30 minutes at 4°C, after which the supernatant was discarded, and the residual sample underwent additional centrifugation at 1500*g* for 5 minutes to remove any residual supernatant. The viral pellet was then resuspended at a ratio of 1:100 of the original volume using PBS (Fisher Scientific, Cat #10010049) at 4°C. Virus was then stored until use at -80°C.

Culture media

Cells were grown in RPMI (Sigma, Cat # R0883) with 10% FCS (Sigma, Cat # F0926), with 100U/mL Pen-Strep (Gibco, Cat # 15140-122), 2mM L-Glutamine (Sigma, Cat # G7513), 10mM HEPES (Sigma, Cat # H0887), 1X MEM Non-essential Amino Acids (Gibco, Cat # 11140-050), 1mM Sodium Pyruvate (Gibco, Cat # 11360-070), and 50 U/mL IL-2 (Amerisource Bergen, Cat #10101641) at a concentration of 1E6 cells/mL.

Genomic DNA extraction and preparation for next generation sequencing

After sorting, cells were washed with PBS, counted, pelleted, and resuspended at up to 5E6 cells per 400 µl of lysis buffer (1% SDS, 50 mM Tris, pH 8, 10 mM EDTA). The remaining protocol reflects additives/procedures performed per each 400 µl of sample. 16µl of NaCl (5M) was added, and the sample was incubated on a heat block overnight at 66°C. The next morning, 8µl of RNAse A (10mg/ml, resuspended in ddH-$_2$O) (Zymo, Cat #E1008) was added, and the sample was vortexed briefly, and incubated at 37°C for 1 hour. Next, 8µl of Proteinase K (20mg/ml) (Zymo, Cat #D3001) was added, the sample was vortexed briefly, and incubated at 55°C for 1 hour. A phase lock tube (Quantabio, Cat #2302820) was prepared for each sample by spinning down the gel to the bottom of the tube at 20,000g for 1 minute and then 400µl of Phenol:Chloroform:Isoamyl Alcohol (25:24:1) was added to each tube. 400µl of the sample was then added to the phase lock tube and the tube was shaken vigorously. The sample was centrifuged at maximum speed at room temperature for 5 minutes. The aqueous phase was transferred to a low-binding eppendorf tube (Eppendorf, Cat #022431021) and then 40µl of Sodium Acetate (3M), 1µl GlycoBlue (Invitrogen, Cat # AM9515), and 600µl of room

temperature isopropanol was added. The sample was then vortexed and stored at -80°C for 30 minutes or until the sample had frozen solid. Next the sample was centrifuged at maximum speed at 4°C for 30 minutes, the pellet was washed with fresh 70% room temperature Ethanol, and allowed to air dry for 15 minutes. Pellets were then resuspended in Zymo DNA elution buffer (Zymo, Cat No: D3004-4-10), and placed on the heat block at 65°C for 1 hour to completely dissolve the genomic DNA.

sgRNA was amplified and barcoded from the genomic DNA as initially described by Joung et al.[9]. Up to 2.5 µg of genomic DNA were added to each 50 µL reaction, which included 25 µL of NEBNext Ultra II Q5 master mix (NEB, Cat #M0544L), 1.25 µL of the 10 µM forward primer and 1.25 µL of the 10 µM reverse primer, and H2O to 50 uL. The following PCR cycling conditions were used: 98°C for 3 minutes, followed by 23 cycles at 98°C for 10 seconds, 63°C for 10 seconds, and 72°C for 25 seconds, and ending with 2 minutes at 72°C. Samples were then cleaned and concentrated in Zymo Spin-V columns (Zymo, Cat #C1016-50) and eluted in 150 uL of Zymo DNA elution buffer. Up to 2 µg of each library were loaded on a 2% agarose gel, and the band at ~250 base pairs was extracted using the Zymoclean Gel DNA recovery kit (Zymo, Cat #D4008). The concentration of each sample was then measured using the Qubit dsDNA high sensitivity assay kit (Thermo Fisher Scientific, Cat #Q32854). Samples were then sequenced on an Illumina HiSeq 4000 using 10-30% PhiX (Illumina, Cat #15017872), and a custom sequencing primer. Primer sequences are listed in Supplementary Table 8.

Arrayed validation isolation, culture, and electroporation

We selected 50 candidate regulators from the screens, 3 positive controls (IL2RA, IL-2, and CTLA4), and 4 non-targeting controls for validation. Thirty-nine of these genes regulate at least two out of three target genes, while the remaining genes were selected because they had large effects in individual screens. For each gene, we selected the top two performing guides from the screen data. Primary human T cells were obtained from whole blood donors (independent of the donors used for the screens) through a protocol approved by the UCSF Committee on Human Research (CHR#13-11950), isolated and stimulated as described above. Custom crRNA plates were ordered from Dharmacon, and were assembled as RNP-ssODN complexes as described above. 48 hours after stimulation, cells were counted, pelleted, and resuspended in room temperature Lonza P3 buffer (Lonza, Cat #V4XP-3032) at 1E6 cells per 20 µL. Cells were then mixed with 100 pmol of RNP, transferred to a 96 well electroporation cuvette plate (Lonza, Cat #VVPA-1002), and nucleofected using the pulse code EH-115. After electroporation, 90 µL of pre-warmed media was immediately added to each well and plates were incubated at 37°C for 15 minutes. Wells were then split to a target culture population of 1E6 cells/mL filling all edge wells in the 96-well plate with PBS in order to avoid edge-effects and incubated at 37°C.

Arrayed validation phenotyping using flow cytometry and genotyping.

Arrayed validation plates were phenotyped at 5 days after electroporation using the sample protocol and materials as outlined in the screen in a 96-well plate format. Cells were stained for IL2RA (CD25) (Tonbo, Cat #20-0259-T100), IL-2 (Biolegend, Cat #500310), or CTLA-4 (Biolegend, Cat #349908) and analyzed with an Attune NxT Flow Cytometer with a 96-well plate-reader. All antibodies were used at a 1:25 dilution for staining.

On day 5 (sgRNA #1 Donor 1-3) or day 7 (sgRNA #2 Donor 1, 3) post-electroporation, genomic DNA was isolated from each sample using DNA QuickExtract (Lucigen, Cat #QE09050) according to the manufacturer's protocol. Primers were designed to flank each sgRNA genomic target site. Amplicons containing CRISPR edit sites were generated by adding 1.25 µL each of forward and reverse primer at 10uM to 5 µL of sample in QuickExtract, 12.5 µL of NEBNext Ultra II Q5 master mix (NEB, Cat #M0544L), and H2O to a total 25 µL reaction volume. The following PCR cycling conditions were used: 98°C for 3 minutes, 15 cycles of 94°C for 20 seconds followed by 65°C-57.5°C for 20 seconds (0.5°C incremental decreases per cycle), and 72°C for 1 minute, and a subsequent 20 cycles at 94°C for 20 seconds, 58°C for 20 seconds and 72°C for 1 minute, and a final 10 minute extension at 72°C. For some samples with non-specific amplification, the above PCR was repeated with a higher temperature touchdown annealing step: 70˚C-64˚C (0.5°C incremental decreases per cycle). Samples were then diluted 1:200 and subsequently Illumina sequencing adapters and indices were added in a second PCR reaction. Indexing reactions included 1 µL of the diluted PCR1 sample, 2.5 µL of each the forward and reverse indexing primers at 10 µM each, 12.5 µL of NEB Q5 master mix, and H2O to a total 25 µL reaction volume. The following PCR cycling conditions were used: 98°C for 30 seconds, followed by 98°C for 10 seconds, 60°C for 30 seconds, and 72°C for 30 seconds for 12 cycles, and a final extension period at 72°C for 2 minutes. Samples were quantified in a 96-well plate reader using the Quant-IT DNA high sensitivity assay kit (Invitrogen, Cat #Q33232) according to the manufacturer's protocol and pooled into 1 tube. Post pooling, samples were then SPRI purified, and quantified using an Agilent 4200 TapeStation. Samples were then sequenced on an Illumina MiniSeq or NextSeq with PE 300 reads. Primer sequences listed in Supplementary Tables 7 and 8.

Combinatorial KO phenotyping

Double KOs were performed similar to arrayed validation phenotyping above using the same guide RNAs as used for RNA/ATAC-Seq. However, for each gene, only half of the RNP volume was used relative to that which was used for single gene KOs.

Blocking IL-2 for arrayed validations

Two days after electroporation, combinatorial KOs of gene + AAVS1 control were split. Half of the cells continued to be grown in IL-2 media (50 U/mL). The other half of the cells were washed in PBS and then anti-IL-2 antibody (Invitrogen Cat #16-7027-85) was added to the media at 40 ug/mL. Fresh antibody was added at 40 ug/mL everyday for the next 3 days until IL2RA staining and analysis.

CD37 enhancer editing

Cells were isolated and electroporated using the arrayed validation protocol. The following guide RNA sequences were used to cut next to SNP rs1465697 (AGCGGAAGCTGAGCATCAGG) or in the AAVS1 locus as a control (TAAGCAAACCTTAGAGGTTC). For each homology directed repair template
(A*G*T*GAAAGTGCTACAGGCTCGAGGTGGGAAGGTCCCTGGGAGTGGGGAGGGCGCTG
GGCCGTGCCAACACCCCCAGCGGAAGCTGAGCATCAGGCAGTACTTCCTGCCTCTGCAGT

GGGGCCCG*A*G*G,
A*G*T*GAAAGTGCTACAGGCTCGAGGTGGGAAGGTCCCTGGGAGTGGGGAGGGCGCTGG
GCCGTGCCAACACCCCCAGCGGAAGCTGAGCATCAGGCCGTACTTCCTGCCTCTGCAGTG
GGGCCCG*A*G*G,
A*G*T*GAAAGTGCTACAGGCTCGAGGTGGGAAGGTCCCTGGGAGTGGGGAGGGCGCTGG
GCCGTGCCAACACCCCCAGCGGAAGCTGAGCATCAGGCTGTACTTCCTGCCTCTGCAGTG
GGGCCCG*A*G*G), 1 uL of 100 uM template was added to individual electroporation reactions. After electroporation, cells with the three homology directed repair templates were pooled together. Five days after electroporation, cells were stained for CD37 (Miltenyi Biotec Cat# 130-123-268) at a 1:50 dilution. The cells were sorted into the 15% lowest and 15% highest expressing cells. Genomic DNA was isolated as outlined for the screen. Amplicons around the SNP were generated by adding 2.5 µL each of forward (cgacgctcttccgatctCAGAGGCCACTGGGTGTG) and reverse primer (cgtgtgctcttccgatctCAGCCTGTTCTGGACCAGG) at 10 uM to 20 µL of sample, 25 µL of NEB Ultra II Q5 master mix (NEB, Cat #M0544L). The following PCR cycling conditions were used: 98°C for 30 seconds, followed by 98°C for 10 seconds, 67°C for 30 seconds, and 72°C for 30 seconds for 35 cycles, and a final extension period at 72°C for 2 minutes. Samples were then diluted 1:200 and subsequently Illumina sequencing adapters and indices were added in a second PCR reaction. Indexing reactions included 1 µL of the diluted PCR1 sample, 2.5 µL of the forward and reverse indexing primers 10 µM each, 12.5 µL of NEB Ultra II master mix, and H2O to a total 25 µL reaction volume. The following PCR cycling conditions were used: 98°C for 30 seconds, followed by 98°C for 10 seconds, 65°C for 1:15 for 15 cycles, and a final extension period at 65°C for 5 minutes. Samples were cleaned up with a 1X SPRI reaction and sequenced on a NovaSeq with paired end 150bp reads.

Collection for RNA-Seq and ATAC-Seq

The top 24 hits with the largest effects in the IL2RA validation were selected for genomic profiling via RNA-Seq and ATAC-Seq. T cells were isolated from 3 donors that were independent of the donors used for the screens or arrayed validation, using consented Leukopaks (STEMCELL) as described above. For each gene we used the sgRNA sequence that had the largest effect in the validation and ordered custom crRNAs from Dharmacon. We also ordered 8 custom crRNAs targeting the safe harbor AAVS1 locus (AAVS1 #1-8). These crRNAs were assembled as RNP-ssODN complexes and electroporated into cells as described above. Five days after electroporation, 60,000 cells were collected and used to generate ATAC-Seq libraries. A fraction of each sample was used for IL2RA flow cytometry as described above to obtain matched protein and RNA expression changes. The remaining cells were lysed with Zymo QuickRNA lysis buffer and isolated with the Zymo QuickRNA micro kit (Zymo, Cat #R1050) using the manufacturer's protocol. RNA samples were treated with 1.5 ul of Turbo DNase (Invitrogen, Cat # AM2238) and then cleaned up using the Zymo RNA-5 Clean and Concentrator (Zymo, Cat #R1016). Isolated RNA integrity and concentration was checked using Agilent RNA Screen Tapes (Agilent, Cat #5067-5576)

Statistics and analysis

Analysis of arrayed validation
Cells were gated on Lymphocytes and singlets in FlowJo Version 10.1 and the median fluorescence intensity (MFI) for APC-height for each stain was exported to csv files. Fluorescence data was imported into R. The MFI across 4 non-targeting controls was averaged per donor per plate and the MFI of each well on the plate was then normalized to the average MFI of the control wells. Individual wells with less than 30% lymphocytes were excluded from analysis due to toxicity. Knockout of HINFP appeared to be toxic and caused most of the cells to die so was excluded from analysis.

Motif enrichment in ATAC-Seq peaks and at disease-associated SNPs
Known transcription factor binding motifs were downloaded from CIS-BP[11], JASPAR2020[12] and HOCOMOCO V11[13]. Motifs in ATAC-Seq peaks or at disease-associated SNPs were identified with the motifmatchr package version 1.10.0 (DOI: 10.18129/B9.bioc.motifmatchr) using the matchMotifs function with default settings. If a given transcription had multiple binding motifs listed across these three databases, any associated motif that fell within an ATAC-Seq peak or overlapped a disease-associated SNP was considered a match.

Analysis of IL2RA CRISPRa screen data in Jurkat cells
Preprocessed bedgraphs showing coverage of CRISRPa guide RNAs enriched in IL2RA high, mid, or low bins compared to IL2RA negative samples were downloaded from Simeonov et al.[14]. (GEO: GSE98178)

Analysis of arrayed genotyping data
Adapter sequences were trimmed from fastq files using cutadapt version 2.8[15] using default settings keeping a minimum read length of 50 bp. Insertions and deletions at each CRISPR target site were then calculated using Crispresso2 version 2.0.42[16] with the following options "--quantification_window_size 3" and "--ignore_substitutions". For each guide RNA, the calculated insertion/deletion frequencies were averaged between donors.

Multiple sclerosis SNPs analysis
Multiple sclerosis (MS) SNPs from a recent multiple sclerosis GWAS meta-analysis[17] with pre-calculated Probabilistic Identification of Causal (PICS) scores were downloaded from pics2.ucsf.edu[18]. SNPs were filtered to be genome wide significant (p-value < 5*10^-8) and have a PICS probability > 0.5. SNPs that fall within 350 bp ATAC-Seq peaks were identified using the findOverlaps function in the GenomicRanges package version 1.40.0[19].

Analysis of CRISPR edits at the *CD37* enhancer
Adapter sequences were trimmed from fastq files using cutadapt version 2.8[15] using default settings keeping a minimum read length of 150 bp. Insertions and deletions at the CD37 enhancer CRISPR target site were then calculated using Crispresso2 version 2.2.6[16] with the following options "-q 30 " (Minimum average phred33 quality score to keep a read) and "--min_bp_quality_or_N 20 " (bases with a phred33 quality score less than 20 will be set to N). Reads were assigned to the reference SNP or alternative bases if they perfectly matched those sequences within a 40bp window centered around the SNP. Insertions/deletions (indels) were

quantified if they overlapped the predicted guide RNA cut site. Non indels substitutions or "N" bases were ignored.

CTLA4, IL-2, and IL2RA disease associations

Clinical phenotypes associated with coding variants in CTLA4 and IL2RA were based on the International Union of Immunological Societies' phenotype classifications[20]. Associations between CTLA4, IL-2, and IL2RA and common autoimmune diseases were based on connections between autoimmune GWAS SNPs and these genes on the Open Targets Platform[21].

Supplementary references

1. Wang, L.-H., Wu, C.-F., Rajasekaran, N. & Shin, Y. K. Loss of Tumor Suppressor Gene Function in Human Cancer: An Overview. *Cell. Physiol. Biochem.* **51**, 2647–2693 (2018).
2. Ross, S. H. & Cantrell, D. A. Signaling and Function of Interleukin-2 in T Lymphocytes. *Annual Review of Immunology* vol. 36 411–433 (2018).
3. Liao, W., Lin, J.-X. & Leonard, W. J. Interleukin-2 at the crossroads of effector responses, tolerance, and immunotherapy. *Immunity* **38**, 13–25 (2013).
4. Spolski, R., Li, P. & Leonard, W. J. Biology and regulation of IL-2: from molecular mechanisms to human therapy. *Nat. Rev. Immunol.* **18**, 648–659 (2018).
5. Leenay, R. T. *et al.* Large dataset enables prediction of repair after CRISPR-Cas9 editing in primary T cells. *Nat. Biotechnol.* **37**, 1034–1037 (2019).
6. Lambert, S. A. *et al.* The Human Transcription Factors. *Cell* **175**, 598–599 (2018).
7. Doench, J. G. *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34**, 184–191 (2016).
8. Cortez, J. T. *et al.* CRISPR screen in regulatory T cells reveals modulators of Foxp3. *Nature* **582**, 416–420 (2020).
9. Joung, J. *et al.* Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nat. Protoc.* **12**, 828–863 (2017).
10. Grevet, J. D. *et al.* Domain-focused CRISPR screen identifies HRI as a fetal hemoglobin

regulator in human erythroid cells. *Science* **361**, 285–290 (2018).

11. Weirauch, M. T. *et al.* Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* **158**, 1431–1443 (2014).
12. Fornes, O. *et al.* JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **48**, D87–D92 (2020).
13. Kulakovskiy, I. V. *et al.* HOCOMOCO: towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. *Nucleic Acids Res.* **46**, D252–D259 (2018).
14. Simeonov, D. R. *et al.* Discovery of stimulation-responsive immune enhancers with CRISPR activation. *Nature* **549**, 111–115 (2017).
15. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).
16. Clement, K. *et al.* CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat. Biotechnol.* **37**, 224–226 (2019).
17. International Multiple Sclerosis Genetics Consortium. Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. *Science* **365**, (2019).
18. Taylor, K. E., Ansel, K. M., Marson, A., Criswell, L. A. & Farh, K. K.-H. PICS2: Next-generation fine mapping via probabilistic identification of causal SNPs. *Bioinformatics* (2021) doi:10.1093/bioinformatics/btab122.
19. Lawrence, M. *et al.* Software for computing and annotating genomic ranges. *PLoS Comput. Biol.* **9**, e1003118 (2013).
20. Bousfiha, A. *et al.* Human Inborn Errors of Immunity: 2019 update of the IUIS phenotypical classification. *J. Clin. Immunol.* **40**, 66–81 (2020).
21. Ochoa, D. *et al.* Open Targets Platform: supporting systematic drug–target identification and prioritisation. *Nucleic Acids Research* vol. 49 D1302–D1310 (2021).