nature research

Corresponding author(s): John Todd (jatodd@well.ox.ac.uk)

Last updated by author(s): May 2, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Illumina ImmuoChip genotype clusters were generated using the Illumina GeneTrain2 algorithm
Data analysis	Complete documentation of all code and tools used for all analyses can be found at https://github.com/ccrobertson/t1d-immunochip-2020 Below are software tools used for key steps in the analyses: Relationship inference and genotype QC - PLINK1.9 and KING (version 2.1.3) Genotype imputation – Michigan Imputation Server: phasing with Eagle (version 2.4) and imputation with Minimac4 Variant annotation – ANNOVAR (version released on Mon, 16 Apr 2018) Defining fine-mapping regions – R package "humarray" Case-control association analysis - SNPTEST (version 2.5.4) Family-based association analysis - PLINK1.9 Meta-analysis of association results – METAL (version released on 2011-03-25) Fine-mapping of associated regions - GUESSFM and PAINTOR Haplotype analyses of fine-mapped regions – GUESSFM, R package "mice", and custom code available in github repository ATAC-seq data processing – PEPATAC, minimap2 (version 2.17), bowtie2 (version 2.3.5), Picard tools (version 2.20.2), deeptools (version 3.3.0), macs2 (version 2.1.2), featureCounts (version 1.6.4), Skewer (version 0.2.2), bedops (version 2.4.35), R package "bigWig" Chromatin accessibility enrichment analyses - custom code available in github repository, GoShifter, DESeq2 Chromatin accessibility QTL analysis – QTLtools, R packages 'limma', 'edgeR' , 'MatrixEQTL', and 'meta' QTL colocalisation - R packages 'coloc' and 'locuscomparer' Drug target prioritisation - R package 'Pi' Data visualization - R packages 'ggplot2','cowplot', 'ggbio', 'GenomicRanges', 'gridExtra', 'RColorBrewer', and 'rtracklayer'

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All univariable summary statistics for genotype association with T1D (including imputed variants) are available through the NHGRI-EBI GWAS Catalog (GCST90013445 and GCST90013446). Chromatin accessibility QTL summary statistics are available through the Type 1 Diabetes Knowledge Portal.
Publicly available ATAC-seq - Raw FASTQ files were obtained from Gene Expression Omnibus (GEO) accession number GSE118189. These data included four individuals and 25 immune cell types under resting conditions and after stimulation with anti-human CD3/CD28 dynabeads and human IL-2 (for 24 hours, T lymphocytes), F(ab)'2 anti-human IgG/IgM 35 and human IL-4 (for 24 hours, B lymphocytes), human IL-2 (for 48 hours, NK cells), or LPS (for 6 hours, monocytes)34. ATAC-seq data from pancreatic islets of five donors without glucose intolerance and five EndoCβH1 cell line replicates, under resting conditions and after stimulation with IFN-γ and IL-1β for 48 hours were downloaded from GEO, accession number GSE123404 32.
ATAC-seq data from cardiac fibroblasts (two fetal and three adult) were downloaded from the European Nucleotide Archive (https://www.ebi.ac.uk/ena/data/view/SRX2843570 and https://www.ebi.ac.uk/ena/data/view/SRX2843571), as a control cell type that we did not expect to be involved in the aetiology of T1D 37.
Epigenome annotation tracks - chromHMM 76 tracks from diverse primary human cells were obtained from the NIH Epigenome Roadmap, http://dcc.blueprint-epigenome.eu/#/md/secondary_analysis/Segmentation_of_ChIP-Seq_data_20140811
and additional immune-specific human primary and cell lines from the Blueprint consortium, https://egg2.wustl.edu/roadmap/data/byFileType/chromhmmSegmentations/ChmmModels/coreMarks/jointModel/final.
Whole blood eQTL summary statistics - Summary statistics from whole blood cis eQTL analysis from 31,683 individuals39 were downloaded from https://eqtlgen.org.
Additional databases used in the Priority Index (Pi) drug target prioritization analysis were obtained through the relational database provided in the R package 'Pi' (http://pi.well.ox.ac.uk:3010/download).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample-size calculations were performed. Samples were collected according to resource availability. This represents the largest study of T1D to date. Genetic studies of this samples size have been shown to be well-powered to detect disease-associated genetic regions.
Data exclusions	Data were excluded according to the following quality control metrics (described in the Online Methods), which are in accordance with standards in the field for quality filtering of genetic data: Genetic data, variant filtering: (1) re-annotated ImmunoChip variant positions by aligning probe sequences to hg37 and removed any variants with less than a 100% match or multiple matches at different positions in the genome; (2) removed variants with call rates less than 98%; (3) removed variants with any discordance between duplicate or monozygotic twin samples, as confirmed by genotype-inferred relationships; (4) removed variants with Mendelian inconsistencies in more than 1% of informative trios or parent-offspring pairs, based on genotypeinferred relationships. Genetic data, sample filtering: We used X chromosome heterozygosity and Y chromosome missingness to identify and exclude participants with apparent sex chromosome anomalies or resolve inconsistencies with reported sex. Additionally, pedigree-defined and genotype-inferred sample relationships were compared and samples were excluded when inconsistencies could not be resolved. Relationships between families, within and across cohorts, were also checked. For each pair of related families observed, we randomly selected one to remove from association analysis. After resolving sex and relationship issues, samples with genotype call rate less than 98% were removed and variants with genotype frequencies deviating from Hardy-Weinberg Equilibrium ($p < 5\times10\text{-}5$) in unrelated European controls were excluded. ATAC-seq sample filtering: Libraries with transcription start site (TSS) enrichment scores less than 6 or fewer than 10 million aligned reads were excluded from analyses.
Replication	This study does not include a dedicated replication cohort. We made efforts to confirm the validity of our findings by comparing genetic associations with previously reported associations (with T1D and other diseases), as well as, by confirming concordance between results from subsets of the data, including multiple case-control cohorts of different ancestries and an affected trio cohort.

Randomization

There was no randomization involved in this study. Data were observational in nature, thus there was no opportunity for randomizing a treatment. Specifically, the treatment of interest is genetic variation and the outcome of interest is type 1 diabetes status. Neither can be randomized.

Blinding

There was no blinding involved in this study. Data were observational in nature, thus blinding is irrelevant to this study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

ETS-1 Rabbit mAb
supplier: Cell Signaling Technology
catalog number: #14069
clone number: D8O8A
lot number: #2

Stat1 Rabbit mAb
supplier: Cell Signaling Technology
catalog number: #14994
clone number: D1K9Y
lot number: #5

Normal Rabbit IgG
supplier: Cell Signaling Technology
catalog number: #2729
clone number: polyclonal
lot number: #9

goat anti-human IgM/IgG/IgA antibody
supplier name: Jackson Immunoresearch
catalog number: 109-006-064
clone name: Polyclonal - RRID: AB_2337548
Final concentration: 10 ug/ml

MEGACD40L protein (recombinant human)
supplier name: ENZO lifesciences
catalog number: ALX-522-110-C010
clone name: NA
Final concentration: 0.15 ug/ml

recombinant human IL-21
supplier name: Peprotech
catalog number: 200-21
clone name: NA
Final concentration: 20 ng/ml

recombinant human IL-4
supplier name: Peprotech
catalog number: 200-04
clone name: NA
Final concentration: 20ng/ml

Validation

ETS-1 Rabbit mAb: Quality tested by immunoprecipitation and western blotting according to vendor's website (including images). The

Validation

vendor has also indicated that the antibody has been also validated by using SimpleChIP® Enzymatic Chromatin IP Kits.
STAT1 Rabbit mAb: Quality tested by immunoprecipitation and western blotting according to vendor's website. The vendor has also indicated that the antibody has been also validated by using SimpleChIP® Enzymatic Chromatin IP Kits.

Normal Rabbit IgG: Quality tested by immunoprecipitation and western blotting according to vendor's website (including images).

MMEGACD40L effect on immune cell activation was validated using the following protocol, according to the manufacturer (quantitative results provided): PBMCs (peripheral blood mononuclear cells) were incubated at 37°C, 5% CO2 for 48 hours in 48 well plates (1 x 106 cells per well) containing 200µl serum free test media with serially diluted MEGACD40L® (Prod. No. ALX-522-110), CD40L (Prod. No. ALX-522-015), or CD40L + 2µg/ml Enhancer (Prod. No. ALX-804-034). After treatment the cells were washed 3X with PBS + 1% BSA. The cells were dual stained for 2 hours at 4°C with mouse anti-human CD19 (PE conjugate) and mouse anti-human CD86 (APC conjugate). The cells were washed 2X with PBS + 1% BSA, then re-suspended In PBS. Samples were analyzed on a BD Facs Calibur flow cytometer. The data is presented as the percent of CD86 positive cells per CD19 positive B cells at each concentration.

Biological activity of the recombinant human IL-21 was determined by its ability to stimulate the proliferation of human ANBL-6 cells. According to manufacturer, the expected ED50 is ≤ 0.5 ng/ml, corresponding to a specific activity of ≥ 2 x106 units/mg.

Biological activity of the recombinant human IL-4 was determined by the dose-dependent stimulation of human TF-1 cells and the ED50 is ≤ 0.2 ng/ml, corresponding to a specific activity of ≥ 5 x 106 units/mg, according to manufacturer.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Jurkat T cells were purchased from ATCC: Clone E6-1 (ATCC TIB-152)
Authentication	ATCC authenticates each distribution lot of cell lines. In addition, morphology of the cell line was checked by microscope, cells grew at a stable proliferation ratio, T cell receptor activation protocols confirmed IL-2 secretion
Mycoplasma contamination	Jurkat cells purchased from ATCC were not tested for mycoplasma.
Commonly misidentified lines (See ICLAC register)	The cell line used in this study is not listed in the ICLAC database of commonly misidentified lines.