## 

## Interaction of rs479777 and AP003774.1 Within a Highly Active Loop Structure in Several Immune Cells May Regulate T-cell Immunoreactivity

­Team: Ay Lab

Contact1: Ferhat Ay Ph.D - [ferhatay@lji.org](mailto:ferhatay@lji.org)

Contact 2: Joaquin Reyna - [jreyna@lji.org](mailto:jreyna@lji.org)

## 

## Section A: Hypothesis Proposal

### Abstract

Type 1 diabetes (T1D) is a disease characterized by the death of βcell populations in the pancreas. Recent papers have proposed further subtyping T1D to better capture and expression the possible mechanisms that drive pathogenesis (Imagawa et al., 2000). One of the most important subtypes involves the T-cell infiltration of the pancreas. As part of our investigation we found AP003774.1, a long non-coding RNA, which is shown to be under the influence of regulatory elements, in particular transcription factor binding sites with the variant rs479777. We hypothesize that the AP003774.1 lncRNA is playing an important role in T-cell populations which goes on to impact T1D related pathways.

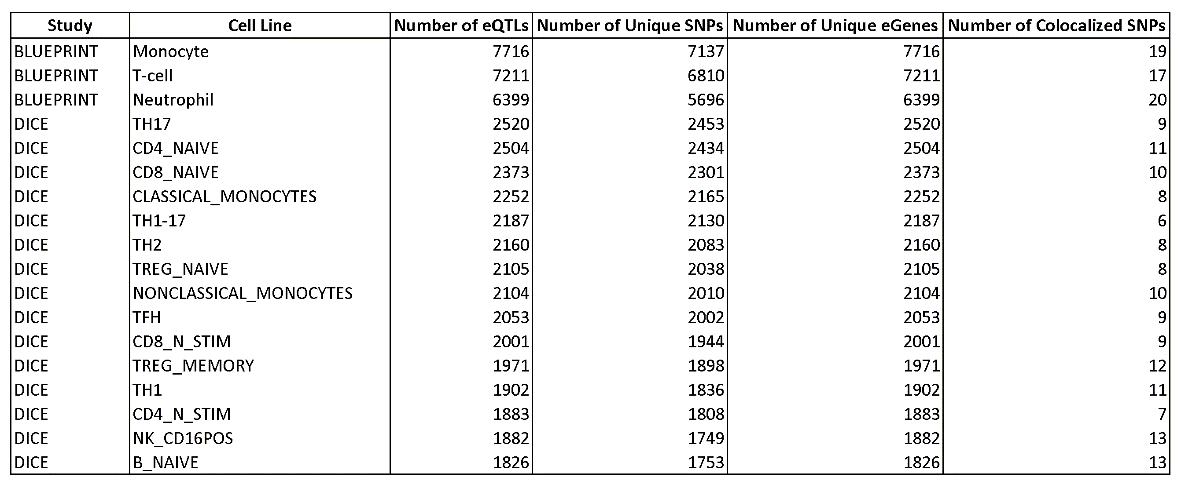
### Brief Statement

AP003774.1 is a long non-coding RNA with an eQTL involving rs479777 in T-cell immune related cell lines. This particular variant sits on a transcription factor binding site which influences expression of AP003774.1. As part of our T1D submission we hypothesize that AP003774.1 is regulating T-cell activity in T1D patients and further research is required to assess clinical interventions.

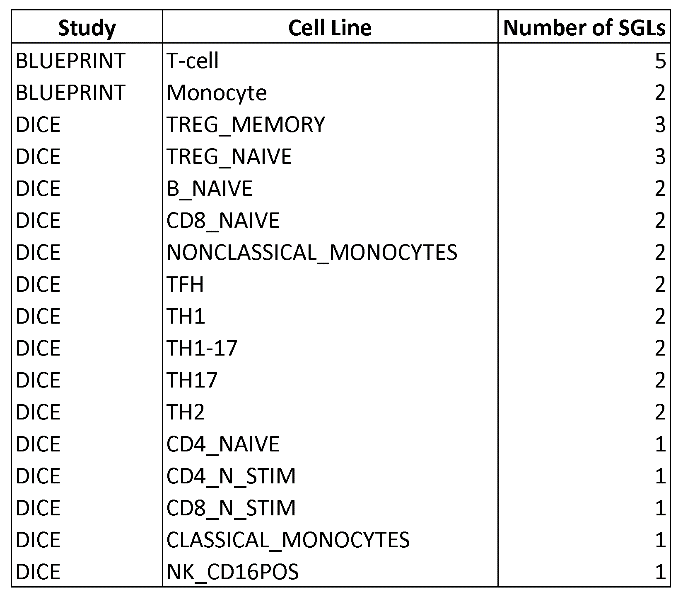
### Hypothesis

We hypothesize that dysregulation of long noncoding RNAs in T-cells influence T1D related pathways, in particular, we identified lncRNA AP003774.1 as a potential target for further T1D research efforts. We integrated several datasets including a T1D GWAS, eQTL studies in several immune cells, HiChIP in the same cell lines and several literature-based sources to prioritize this final T1D candidate gene. As part of this hypothesis we will explain each dataset, our pipeline and deep dive into the AP003774.1 locus.

Table 1: Statistics of eQTL and colocalization analyses broken down by study and cell types. The 3rd-5th columns describe eQTL based information and the last column contains colocalized SNP counts.

Type 1 diabetes has been shown to include a strong genetic component. In (Chiou et al., 2021) they performed a T1D-GWAS study using 520,580 individuals where, after fine-mapping, they prioritized 136 SNPs. In addition, they performed single-cell ATAC-seq on whole blood followed by clustering which provided open chromatin information on a cell type basis. Taking these 136 SNPs they found high enrichment within open chromatin regions. The large sample size and results from this paper assured us of the importance of these SNPs for further analysis. As part of our analyses we took the larger set of SNPs prior to fine-mapping which includes 45,994 SNPs with genome-wide significance (p-value ≤ 1e-8).

As part of our effort we integrated eQTL results between SNPs and gene expression in several immune cells. More specially, we included eQTLs from BLUEPRINT which includes monocytes, neutrophils and T-cells as well as eQTLs from DICE which include gene expression from naive B cells, classical monocytes, non-classical monocytes, T follicular helper cells, T helper (including the derivatives 1, 2, 17, 1/17), Natural killer cells, naive CD4 T cells, and naive CD8 T cells (Mu et al., 2021; Chandra et al., 2020; Schmiedel et al., 2018). BLUEPRINT cell lines have the most eGenes with a median of 7211 eGenes versus 2104 eGenes for DICE. This can be attributed to cell subtyping within the DICE dataset. This subtyping allows us to find cell type specific patterns such as the difference between monocytes and non-classical monocytes which differ (numerically) by 148 eGenes. We can also observe the great diversity of eGenes calls within the T-cell subtypes which range from 1,902 eGenes in T-helper cells to 2,520 eGenes in T-helper 17 cells (Table 1).

Several papers have coined the term post-GWAS to signify the need to solve the how, where and when of GWAS-SNP function (Pierce et al., 2020). In alignment with this need colocalization analyses have followed which calculates the posterior probability that a genetic locus contains two separate signals, in our case we analyzed the colocalization of a GWAS and eQTL signal (Wallace et al., 2021; Wallace et al., 2020; Giambartolomei et al., 2014). We performed this analysis between the (Chiou et al., 2021) GWAS summary statistics and eQTL summary statistics for each cell line resulting in a median of 19 versus 9 colocalized SNPs for BLUEPRINT and DICE, respectively. As in the eQTL analysis we saw a diverse range of colocalized SNPs for T-cells ranging from 12 for Treg memory cells to 6 in T helper 1-17 cells which may suggest that SNPs pathogenic for T1D could be biasedly hitting some T-cell subsets over others.

Colocalization made us more confident that a given SNP was somehow physiologically connected to T1D via some gene expression mechanism. Given our expertise in DNA-DNA interactions we hypothesized that colocalized SNP-gene pairs could be connected via 3D interactions, most importantly, we integrated HiChIP data which extracted loops containing the activating histone mark H3K27ac. For easier reference we denote SNP-gene pairs with loops as SGLs. We found an average of 2 SGLs per cell line and combined all such SNP-gene pairs across all cell lines to establish a set of candidate genes (Table 2). In total we would have 14 genes including BACH2, UBASH3A, PTPN22, SIRPG which have been previously associated with type 1 diabetes. Given their previous significance we focused on genes which are less known in literature, this includes: AP003774.1, AP4B1, CTSH, M6PR, RGS1, RMI2, RP11-75L1.1, RPS26, SULT1A2, and TMPRSS3.

Table 2: Number of SGLs per study and cell type.

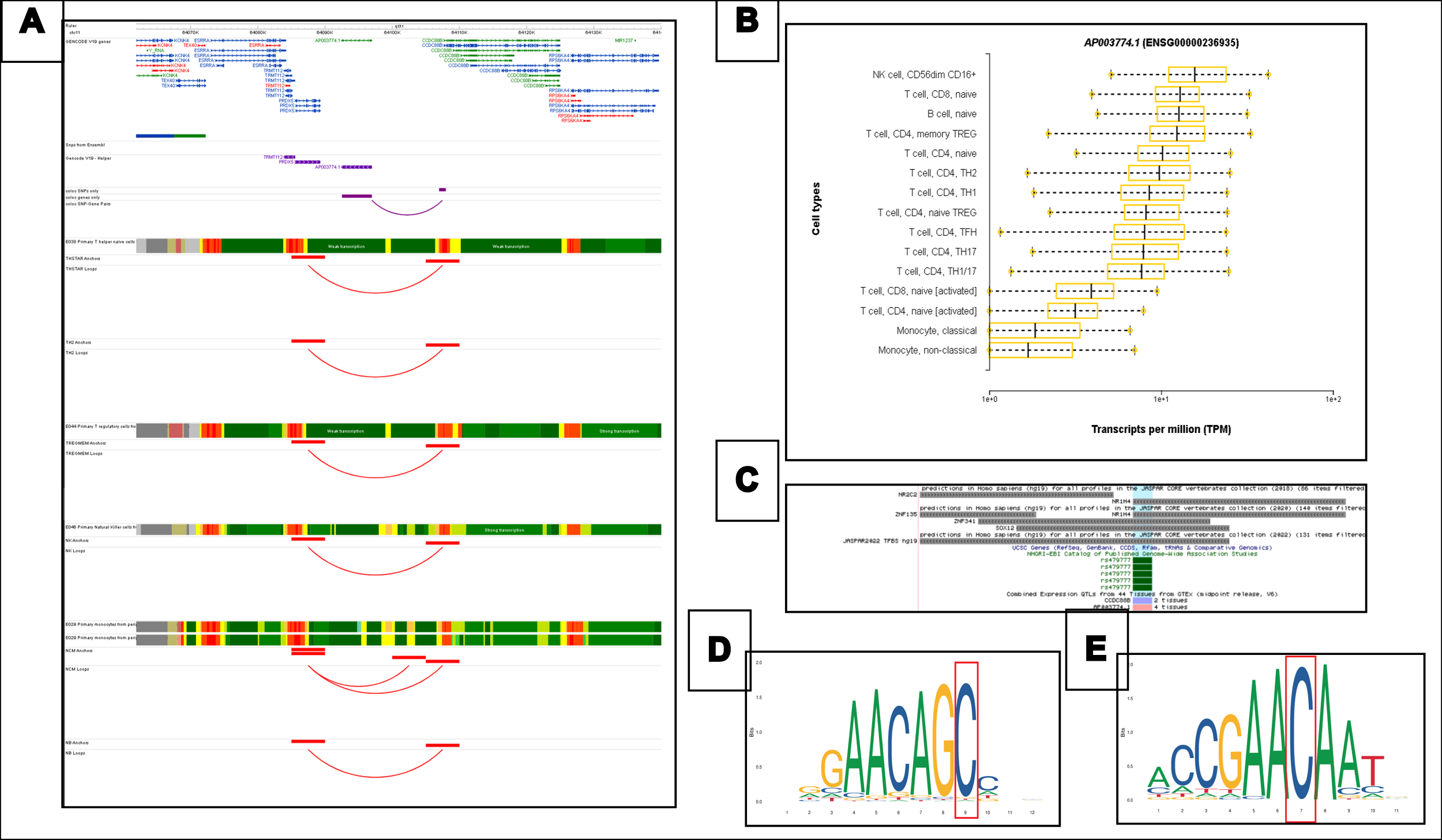
At this point we had systematically analyzed the genome and for the next half of our analyses we relied on several genome annotations to study interesting additional parameters such as overlap of the SNPs with transcription factor binding sites and ChromHMM states which indicate some type of gene regulation. We first observed that several cell lines (TH2, TH1-17, Treg memory, natural killer, non-classical monocytes and naive B cells) contained the rs479777 SNP within one anchor and the AP003774.1 gene within the loop body (figure panel A). No other such loop construct came close in terms of absolute cell lines. By adding the cell type specific chromHMM annotations we found that these loops connected two nearby transcription start sites (figure panel A). Then, we verified the expression of this gene in the DICE database which shows that AP003774.1 is indeed active in immune cell types. We further investigated this loci by adding JASPAR motif tracks which located 2 transcription factors overlapping rs479777, SOX12 and ZNF341 (figure panel C). By looking at the JASPAR website we found that rs479777 actually swaps a highly conserved C for a T nucleotide (figure panel D and E) at both TF binding sites. Given that JASPAR relies completely on in-silico analyses we also queried the ADASTRA database for ChIP-seq derived transcription factor binding sites and found that HTF4 binds in monocytes. As more datasets are included in ADASTRA we do expect that ChIP-seq in immune cells will further corroborate our findings. Lastly, we went back and added scATAC-seq from (Chiou et al., 2021) where we found an ATAC-seq peak overlapping the rs479777 SNP, most importantly in activated CD4 T, adaptive NK, classical monocytes, conventional dendritic, cytotoxic CD8 T, cytotoxic NK, megakaryocyte, memory B, memory CD8 T, naive B, naiveT, non-classical monocytes, pancreatic CD8 T, pancreatic macrophage, plasmacytoid dendritic, and regulatory T cells.

Figure 1: Investigating the rs479444/AP003774.1 Locus. A) Visualization of the AP003774.1 locus with tracks for gene annotations (Gencode V19), dbSNP, simplified gene track (purple), colocalized SNP track (purple), colocalized gene track (purple), link showing SNP-gene pairs with at least on physical loop and several ChromHMM and loop track for the following cell lines: THSTAR, TH2, TREGMEM, NK, CM, and naïve B. B) Gene expression in the DICE database for AP003774.1, C) JASPAR tracks for transcription factor binding sites, D) SOX12 motif and E) ZNF341 motif.

In addition to T1D we also examined the association of rs479777 to other autoimmune diseases. We found associations to sarcoidosis (inflammation and granuloma formation in

affected organs) (Fischer et al., 2012), Crohn’s disease (inflammatory bowel disease) (Crohn's disease., 2020) and alopecia areata (reversible hair loss) (Alopecia areata., 2017). Interestingly, we found that rs479777 was significant in a GWAS for thyroid preparations meaning individuals with this variant often require some form of medication to make up for an improper thyroid function (Sakaue et al., 2021). Having evaluated the importance of rs479777 we focused the remainder of this hypothesis on studying AP003774.1.

AP003774.1 is a long noncoding RNA with little functional annotation which continues to mask its priority as a research topic. Initially we found that the locus containing AP003774.1 is very close to the CCDC88B gene and rs479777 is very close to the 5-prime end of it. Research has shown that lncRNAs can regulate genes in cis (Gil & Ulitsky, 2019) which could mean that AP003774.1 could somehow regulate CCDC88B but possibly other genes as well. (Teimuri et al., 2018) have shown that lncRNAs in Th17 could be potential therapeutic drug targets in autoimmune diseases. (Almo et al., 2018) also show that noncoding RNAs play an important role in human T CD3+ cells and a much longer list of papers describe the role the lncRNAs can have on T-cell development, differentiation, stability of pathogenic mRNAs and several other autoimmune diseases (Dieter et al., 2021; Plasek & Valadkhan 2020; Gonzalez-Moro et al., 2020; Roy & Awasthi, 2019; Gao et al., 2018; Guo et al., 2019). Given this background, we consider AP003774.1 an important target for future T1D studies and further reveal the role of understudied RNA molecules in disease etiology.

### Potential Study Design

Our lab has generated Jurkat cell lines, a mixture of different T-cell subtypes, that express the KRAB-dCas9 construct. Using this CRISPR inhibition (CRISPRi) system we can target the SNP we found associated with AP003774.1 expression by designing guide RNAs that will inhibit the enhancer (open chromatin region) overlapping this SNP. We will quantify the expression level of AP003774.1 as well as other nearby genes in cells with the target gRNA vs controls to test our hypothesis that the specific enhancer region we found is controlling AP003774.1 gene expression in multiple cell types where the SNP showed colocalization.

We have previously used this model in (Chandra et al., 2021) to show that the FHIT gene is regulated by rs11130745 via a looping mechanism. Like in the AP003774.1 context, the FHIT gene is located within the loop body, in addition, this SNP overlaps an enhancer within one anchor and a promoter on the second anchor. A tiling approach was then used to silence nearby regions and resulted in lower FHIT gene expression.

We used a second recombination based CRISPR approach to change the rs1130745 site from a G/G to an A/A and again saw lower expression. By using this system we can better understand the role AP003774.1 plays in T-cells development and T1D pathology.

### Author list, affiliations, funding acknowledgements

Joaquin Reyna1,2

Sourya Bhattacharyya1

Nikhil Rao1,4

Abhijit Chakraborty1

Ferhat Ay1,2,3,4,5

1La Jolla Institute for Immunology, La Jolla, CA, USA.

2Bioinformatics and Systems Biology Graduate Program, University of California, San Diego, La Jolla, CA 92093-0419

3Department of Medicine, University of California San Diego, La Jolla, CA, USA.

4Department of Bioengineering, University of California San Diego, La Jolla, CA, USA.

5Division of Biological Sciences, University of California San Diego, La Jolla, CA, USA.

## B: Protocol from Protocols.io (PDF)

<https://www.protocols.io/private/B58E0187427B11ECAA1D0A58A9FEAC02>

## Section C: Table of resources used

|  |  |  |  |
| --- | --- | --- | --- |
| Name | RRID | URL | Usage |
| GWAS: Catalog of Published Genome-Wide Association Studies | SCR\_012745 | <http://www.ebi.ac.uk/gwas/> | * Downloaded T1D-GWAS summary statistics from (Chiou et al., 2021) * Downloaded scATAC-seq to investigate open chromatin within the AP003774.1 locus |
| Blueprint Epigenome | SCR\_003844 | <http://www.blueprint-epigenome.eu/> | * Downloaded eQTL summary statistics |
| JASPAR | SCR\_003030 | <http://jaspar.genereg.net/> | * Investigated transcription factor binding motifs for several SNPs |
| Roadmap Epigenomics Project | SCR\_008924 | <http://roadmapepigenomics.org/> | * Annotated the WashU Epigenome Browser with ChromHMM annotations |
| WashU Epigenome Browser | SCR\_006208 | <http://epigenomegateway.wustl.edu/> | * Visualized loops, SNPs, genes, ChromHMM and several other tracks |
| UCSC Genome Browser | SCR\_005780 | <http://genome.ucsc.edu/> | * Visualized additional tracks which were not easily accessible via the WashU Epigenome Browser |

## Section D: References

* Alopecia areata. Nat Rev Dis Primers 3, 17012 (2017). <https://doi.org/10.1038/nrdp.2017.12>
* Almo MM, Sousa IG, Maranhão AQ, Brigido MM. J Immunol Sci. (2018); 2(1): 32-36
* Chiou J, Geusz RJ, Okino ML, Han JY, Miller M, Melton R, Beebe E, Benaglio P, Huang S, Korgaonkar K, Heller S, Kleger A, Preissl S, Gorkin DU, Sander M, Gaulton
* Crohn's disease. *Nat Rev Dis Primers* 6, 23 (2020). <https://doi.org/10.1038/s41572-020-0166-0>
* Chandra V, Bhattacharyya S, Schmiedel BJ, Madrigal A, Gonzalez-Colin C, Fotsing S, Crinklaw A, Seumois G, Mohammadi P, Kronenberg M, Peters B, Ay F, Vijayanand P. Promoter-interacting expression quantitative trait loci are enriched for functional genetic variants. Nat Genet. 2021 Jan;53(1):110-119. doi: 10.1038/s41588-020-00745-3. Epub 2020 Dec 21. PMID: 33349701; PMCID: PMC8053422.
* Crohn's disease. *Nat Rev Dis Primers* 6, 23 (2020). <https://doi.org/10.1038/s41572-020-0166-0>
* Dieter C, Lemos NE, Corrêa NRF, Assmann TS, Crispim D. The Impact of lncRNAs in Diabetes Mellitus: A Systematic Review and In Silico Analyses. Front Endocrinol (Lausanne). 2021 Mar 19;12:602597. doi: 10.3389/fendo.2021.602597. PMID: 33815273; PMCID: PMC8018579.
* Fischer A, Schmid B, Ellinghaus D, Nothnagel M, Gaede KI, Schürmann M, Lipinski S, Rosenstiel P, Zissel G, Höhne K, Petrek M, Kolek V, Pabst S, Grohé C, Grunewald J, Ronninger M, Eklund A, Padyukov L, Gieger C, Wichmann HE, Nebel A, Franke A, Müller-Quernheim J, Hofmann S, Schreiber S. A novel sarcoidosis risk locus for Europeans on chromosome 11q13.1. Am J Respir Crit Care Med. 2012 Nov 1;186(9):877-85. doi: 10.1164/rccm.201204-0708OC. Epub 2012 Jul 26. PMID: 22837380.
* Gao Y, Li S, Zhang Z, Yu X, Zheng J. The Role of Long Non-coding RNAs in the Pathogenesis of RA, SLE, and SS. Front Med (Lausanne). 2018 Jul 3;5:193. doi: 10.3389/fmed.2018.00193. PMID: 30018955; PMCID: PMC6038710.
* Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, Plagnol V. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet. 2014 May 15;10(5):e1004383. doi: 10.1371/journal.pgen.1004383. PMID: 24830394; PMCID: PMC4022491.
* Gil N, Ulitsky I. Regulation of gene expression by cis-acting long non-coding RNAs. Nat Rev Genet. 2020 Feb;21(2):102-117. doi: 10.1038/s41576-019-0184-5. Epub 2019 Nov 15. PMID: 31729473.
* Gonzalez-Moro I, Olazagoitia-Garmendia A, Colli ML, Cobo-Vuilleumier N, Postler TS, Marselli L, Marchetti P, Ghosh S, Gauthier BR, Eizirik DL, Castellanos-Rubio A, Santin I. The T1D-associated lncRNA *Lnc13* modulates human pancreatic β cell inflammation by allele-specific stabilization of *STAT1* mRNA. Proc Natl Acad Sci U S A. 2020 Apr 21;117(16):9022-9031. doi: 10.1073/pnas.1914353117. Epub 2020 Apr 13. PMID: 32284404; PMCID: PMC7183221.
* Guo J, Liu Z, Gong R. Long noncoding RNA: an emerging player in diabetes and diabetic kidney disease. Clin Sci (Lond). 2019 Jun 20;133(12):1321-1339. doi: 10.1042/CS20190372. PMID: 31221822.
* Katsarou, A., Gudbjörnsdottir, S., Rawshani, A. *et al.* Type 1 diabetes mellitus. *Nat Rev Dis Primers* 3, 17016 (2017). <https://doi.org/10.1038/nrdp.2017.16>
* KJ. Interpreting type 1 diabetes risk with genetics and single-cell epigenomics. Nature. 2021 Jun;594(7863):398-402. doi: 10.1038/s41586-021-03552-w. Epub 2021 May 19. PMID: 34012112.
* Imagawa A, Hanafusa T, Miyagawa J, Matsuzawa Y. A proposal of three distinct subtypes of type 1 diabetes mellitus based on clinical and pathological evidence. Ann Med. 2000 Nov;32(8):539-43. doi: 10.3109/07853890008998833. PMID: 11127931.

## Mountjoy, E., Schmidt, E.M., Carmona, M. *et al.* An open approach to systematically prioritize causal variants and genes at all published human GWAS trait-associated loci. *Nat Genet* 53, 1527–1533 (2021). <https://doi.org/10.1038/s41588-021-00945-5>

* Mu, Z., Wei, W., Fair, B. *et al.* The impact of cell type and context-dependent regulatory variants on human immune traits. *Genome Biol* 22, 122 (2021). <https://doi.org/10.1186/s13059-021-02334-x>
* Pierce SE, Booms A, Prahl J, van der Schans EJC, Tyson T, Coetzee GA. Post-GWAS knowledge gap: the how, where, and when. NPJ Parkinsons Dis. 2020 Sep 9;6(1):23. doi: 10.1038/s41531-020-00125-y. PMID: 34497275.
* Plasek LM, Valadkhan S. lncRNAs in T lymphocytes: RNA regulation at the heart of the immune response. Am J Physiol Cell Physiol. 2021 Mar 1;320(3):C415-C427. doi: 10.1152/ajpcell.00069.2020. Epub 2020 Dec 9. PMID: 33296288; PMCID: PMC8294623.
* Roy S, Awasthi A. Emerging roles of noncoding RNAs in T cell differentiation and functions in autoimmune diseases. Int Rev Immunol. 2019;38(5):232-245. doi: 10.1080/08830185.2019.1648454. Epub 2019 Aug 14. PMID: 31411520.
* Sakaue S, Kanai M, Tanigawa Y, Karjalainen J, Kurki M, Koshiba S, Narita A, Konuma T, Yamamoto K, Akiyama M, Ishigaki K, Suzuki A, Suzuki K, Obara W, Yamaji K, Takahashi K, Asai S, Takahashi Y, Suzuki T, Shinozaki N, Yamaguchi H, Minami S, Murayama S, Yoshimori K, Nagayama S, Obata D, Higashiyama M, Masumoto A, Koretsune Y; FinnGen, Ito K, Terao C, Yamauchi T, Komuro I, Kadowaki T, Tamiya G, Yamamoto M, Nakamura Y, Kubo M, Murakami Y, Yamamoto K, Kamatani Y, Palotie A, Rivas MA, Daly MJ, Matsuda K, Okada Y. A cross-population atlas of genetic associations for 220 human phenotypes. Nat Genet. 2021 Oct;53(10):1415-1424. doi: 10.1038/s41588-021-00931-x. Epub 2021 Sep 30. PMID: 34594039.
* Schmiedel BJ, Singh D, Madrigal A, Valdovino-Gonzalez AG, White BM, Zapardiel-Gonzalo J, Ha B, Altay G, Greenbaum JA, McVicker G, Seumois G, Rao A, Kronenberg M, Peters B, Vijayanand P. Impact of Genetic Polymorphisms on Human Immune Cell Gene Expression. Cell. 2018 Nov 29;175(6):1701-1715.e16. doi: 10.1016/j.cell.2018.10.022. Epub 2018 Nov 15. PMID: 30449622; PMCID: PMC6289654.
* Skyler JS. Characterizing subgroups of type 1 diabetes. Diabetes. 2014 Nov;63(11):3578-80. doi: 10.2337/db14-1160. PMID: 25342728.
* Teimuri S, Hosseini A, Rezaenasab A, Ghaedi K, Ghoveud E, Etemadifar M, Nasr-Esfahani MH, Megraw TL. Integrative Analysis of lncRNAs in Th17 Cell Lineage to Discover New Potential Biomarkers and Therapeutic Targets in Autoimmune Diseases. Mol Ther Nucleic Acids. 2018 Sep 7;12:393-404. doi: 10.1016/j.omtn.2018.05.022. Epub 2018 Jul 11. PMID: 30195777; PMCID: PMC6128809.
* Wallace C. A more accurate method for colocalisation analysis allowing for multiple causal variants. PLoS Genet. 2021 Sep 29;17(9):e1009440. doi: 10.1371/journal.pgen.1009440. PMID: 34587156; PMCID: PMC8504726.
* Wallace C. Eliciting priors and relaxing the single causal variant assumption in colocalisation analyses. PLoS Genet. 2020 Apr 20;16(4):e1008720. doi: 10.1371/journal.pgen.1008720. PMID: 32310995; PMCID: PMC7192519.

## Section E: Conflict of Interest

No conflict to declare.

## Other - Thorough Description of Datasets

For our purposes, we used eQTLs from several immune cell types. Here are the datasets that we used and the cell types that they include:

DICE:

* Description: The DICE dataset consists of population RNA-seq data for 13 unstimulated immune cell types including various naïve and effector/memory T cell subtypes, classical and non-classical monocytes, B cells, and NK cells. The DICE dataset also includes RNA-seq data from CD4+ and CD8+ T cells that have been activated in vitro by engaging T cell receptor (TCR) complex using CD3/CD28 antibodies. Although the sample size in the DICE dataset is the smallest (n=91) among the four datasets, the large number of sorted cell types makes the DICE dataset ideal to identify cell type-specific genetic effects.
* Source types: B cells (naive), CD4+ T Cells (naive), CD4+ T Cells (stimulated), CD8+ T Cells (naive), CD8+ T Cells (stimulated), Classical Monocytes, Natural killer cell (CD16+), Non-classical Monocytes, T follicular helper cells, T helper cells subset which expression IL1/IL17, T helper cells subset which expression IL1 only, T helper cells subset which expression IL17 only, T helper cells subset which expression IL17 only, T helper 1 cells, T helper 2 cells, T reg memory cells, T reg naive cells
* Number of samples: 91 per cell type

BLUEPRINT:

* Description: The BLUEPRINT dataset consists of RNA-seq data from three cell types (classical monocytes, naïve CD4+ T cells and neutrophils) in ∼197 individuals.
* Source types: Classical monocytes, Naive CD4+, Neutrophils
* Number of samples: 197 per cell type

DGN:

* Description: The DGN consortium collected whole blood samples from 922 individuals
* Source types: Whole blood
* Number of samples per cell types: 922

GEUVADIS:

* Description: GEUVADIS collected RNA-seq data from 462 lymphoblastoid cell lines (LCL)
* Source types: Lymphoblastoid
* Number of samples per cell types: 462

## Other - Thorough Description of CRISPR

**CRISPRi targeting of enhancers using KRAB-dCas9.**

At 3 d before CRISPRi assay, KRAB-dCas9-expressing cells (mCherry positive) were sorted again to ensure that all cells expressed KRAB-dCas9. Then, 44 μM crRNA and tracrRNA (from IDT) duplex specific for each target was prepared by mixing the two RNA oligos in equimolar concentrations in a sterile microcentrifuge and heating at 95 °C followed by cooling at room temperature. Cells were transfected with 3.6 μM crRNA and tracrRNA duplex specific for the target enhancer or for the nontargeting region (from IDT) using the Neon Transfection System (Thermo Fisher Scientific) according to the manufacturer’s protocol (settings: 1,600 V, 10 ms, 3 pulses); see Supplementary Table 1e for crRNA sequences. Fresh medium (as described above) was then added and cells were maintained for 48 h. After 48 h cells were collected and knockdown efficiency for the target gene was analyzed by real-time PCR for transcript levels.

**Inhibition of targeted region using dCas9.**

At 3 d before CRISPRi assay, dCas9-expressing cells (EGFP positive) were sorted to ensure that all cells expressed dCas9. Then, 44 μM crRNA and tracrRNA (IDT) duplex specific for each target was prepared by mixing the two RNA oligonucleotides in equimolar concentrations in a sterile microcentrifuge, and heating at 95 °C followed by cooling at room temperature. RNP complexes were prepared by incubating dCas9 and crRNA–tracrRNA duplex specific for target regions for 20 min at room temperature. Cells were transfected with RNP complex using the Neon Transfection System (Thermo Fisher Scientific) according to the manufacturer’s protocol (settings: 1,600 V, 10 ms, 3 pulses); see Supplementary Table 1e for crRNA sequences. Fresh medium (as described above) was then added and cells were maintained for 48 h. After 48 h cells were collected and knockdown efficiency was analyzed by real-time PCR for transcript levels of target genes.

**Genome editing in activated CD4+ T cells.**

Naïve CD4+ T cells purified using a

magnetic-activated cell-sorting column were resuspended at a concentration of 2.5 × 105

per ml in 1 ml of prewarmed IMDM medium, supplemented with 5%

(vol/vol) heat-inactivated FBS and 2% (vol/vol) human AB serum (CellGro) and activated ex vivo with Dynabeads Human T-Activator CD3/CD28 (Thermo Fisher Scientific) at a bead-to-cell ratio of 2:1 for the indicated duration at 37 °C. Three different activation conditions were used to determine the activation condition that leads to efficient HDR efficiency as well as reproducibility of effect. After activation, Dynabeads were removed and cells were cultured in fresh medium with IL-2 for the indicated number of days. RNP complexes were prepared by incubating dCas9 with either single guide RNA (sgRNA) alone or crRNA– tracrRNA duplex specific for the target region for 20 min at room temperature. Then, 2.0 × 105

activated CD4+ T cells were washed two times with PBS before

resuspension in 8 μl of buffer T, and 80 pmol of HDR template was added to the cell suspension along with RNP complex. Cells were transfected using the Neon Transfection System (Thermo Fisher Scientific) according to the manufacturer’s protocol (settings: 1,600 V, 10 ms, 3 pulses); see Supplementary Table 1e for sgRNA and crRNA sequences. Fresh medium (as described above) was then added and cells were maintained for the indicated duration. Cells were collected, and DNA and RNA were isolated for downstream analysis. Genome editing was verified by Sanger sequencing and effects on gene expression assessed by real-time PCR for transcript levels.

Other: Full Table

|  |  |  |  |
| --- | --- | --- | --- |
| Name | RRID | URL | Usage |
| GWAS: Catalog of Published Genome-Wide Association Studies | SCR\_012745 | <http://www.ebi.ac.uk/gwas/> | * Downloaded T1D-GWAS summary statistics from (Chiou et al., 2021) |
| Blueprint Epigenome | SCR\_003844 | <http://www.blueprint-epigenome.eu/> |  |
| DICE | nan | <https://dice-database.org/> |  |
| JASPAR | SCR\_003030 | <http://jaspar.genereg.net/> |  |
| Roadmap Epigenomics Project | SCR\_008924 | <http://roadmapepigenomics.org/> |  |
| WashU Epigenome Browser | SCR\_006208 | <http://epigenomegateway.wustl.edu/> |  |
| UCSC Genome Browser | SCR\_005780 | <http://genome.ucsc.edu/> |  |
| FIVEx | nan | <https://fivex.sph.umich.edu/> |  |
| ADASTRA | nan | <https://adastra.autosome.ru/zanthar> |  |

1) the GWAS dataset, 2) eQTL datasets, 3) colocalization analyses the GWAS and eQTLs, 4) HiChIP loop data, 5) investigation of a candidate list, and 6)