A Non-Coding Variant in *SLC15A4* Modulates Enhancer Activity and Lysosomal Deacidification Linked to Lupus Susceptibility

Manish Kumar Singh[1†], Guru Prashad Maiti[1†], HariKrishna Reddy-Rallabandi[1], Mehdi Fazel-Najafabadi[1], Loren L. Looger[2], Swapan K. Nath[1*]

[1]Arthritis and Clinical Immunology Research Program, Oklahoma Medical Research Foundation, Oklahoma City OK, USA

[2]Howard Hughes Medical Institute, Department of Neurosciences, University of California San Diego, La Jolla CA, USA

† Contributed equally

*** Correspondence:**
Swapan K. Nath, Ph.D.
Swapan-Nath@omrf.org

Keywords: SLC15A4, Luciferase, ChIP-qPCR, CRISPR, Lupus, Lysosomal acidification

Running Title: SLC15A4 variant in lupus susceptibility

Abstract

Systemic lupus erythematosus (SLE) is a complex autoimmune disease with a strong genetic basis. Despite the identification of several single nucleotide polymorphisms (SNPs) near the *SLC15A4* gene that are significantly associated with SLE across multiple populations, specific causal SNP(s) and molecular mechanisms responsible for disease susceptibility are unknown. To address this gap, we employed bioinformatics, expression quantitative trait loci (eQTLs), and 3D chromatin interaction analysis to nominate a likely functional variant, rs35907548, in an active intronic enhancer of *SLC15A4*. Through luciferase reporter assays followed by chromatin immunoprecipitation (ChIP)-qPCR, we observed significant allele-specific enhancer effects of rs35907548 in diverse cell lines. The rs35907548 risk allele T is associated with increased regulatory activity and target gene expression, as shown by eQTLs and chromosome conformation capture (3C)-qPCR. The latter revealed long-range chromatin interactions between the rs35907548 enhancer and the promoters of *SLC15A4*, *GLTLD1*, and an uncharacterized lncRNA. The enhancer-promoter interactions and expression effects were validated by CRISPR/Cas9 knock-out (KO) of the locus in HL60 promyeloblast cells. KO cells also displayed dramatically dysregulated endolysosomal pH regulation. Together, our data show that the rs35907548 risk allele affects multiple aspects of cellular physiology and may directly contribute to SLE.

1 Introduction

Systemic lupus erythematosus (SLE) is a chronic autoimmune disease characterized by immune attack on the body's tissues and organs, resulting in profound dyshomeostasis and damage to skin, joints, kidneys, cardiovascular system, and nervous system, among others. The disease is very heterogeneous, with each patient presenting uniquely, thus making diagnosis, treatment, and even basic understanding challenging [1]. Gender and ethnicity significantly influence the incidence and severity of SLE, with females and individuals of African, Hispanic, and Asian ancestry being both more prone to SLE and to severe manifestations like kidney disease and frequent hospitalization. Conversely, European-ancestry SLE patients tend to exhibit more skin manifestations but less renal involvement. Therefore, managing SLE and its comorbidities requires a comprehensive internal medicine approach and an understanding of its diverse underlying pathogenic mechanisms [2, 3, 4, 5].

SLE has a significant genetic component, evidenced by familial aggregation, sibling risk ratio, and twin studies [6]. SLE appears to be a highly polygenic disease, with >100 risk loci implicated to date [7, 8]. These SNPs and associated genes are involved in biological pathways related to tolerance, cell signaling, apoptosis, and other critical immune functions [9, 10]. Molecular pathway analysis has revealed different underpinnings of immune system homeostasis and SLE risk in diverse ethnic populations [9, 11, 12], emphasizing the importance of studying the genetics of complex disease in multiple ethnicities – both to elucidate fundamental biochemical pathways and to develop personalized diagnostics and treatments.

SLC15A4 was discovered as an SLE risk locus in 2009 through GWAS on Chinese individuals. Since then, several SNPs have been associated with risk across Asian, European, Hispanic, and African ancestries [9]. However, despite strong genetic association, the actual functional SNP(s) and underlying biological mechanism(s) contributing to SLE pathogenesis are not understood [13, 14].

SLC15A4, also known as peptide/histidine transporter 1 (PHT1), is a member of the solute carrier family 15 of proton-coupled oligopeptide transporters [13, 14, 15, 16]. It is primarily located on the endolysosomal membrane of immune cells, where it transports histidine and bacterially-derived dipeptides such as the NOD2 ligand muramyl dipeptide (MDP). SLC15A4 is crucial for lysosomal acidification, as it generates a proton gradient through transport of the proton acceptor histidine. Moreover, SLC15A4 can recruit the adapter molecule “TLR adaptor interacting with SLC15A4 on the lysosome” (TASL), which regulates Toll-like receptor (TLR) function and promotes downstream signaling through type I interferon and Interferon Response Factor 5 (IRF5) [17]. Therefore, SLC15A4 plays a crucial role in regulating lysosomal function and innate immunity. SLC15A4 deficiency may promote lysosomal dysfunction and impaired autophagy, both associated with various autoimmune diseases, including SLE [18]. Consequently, these findings suggest that pharmacological intervention to restore or supplement the function of SLC15A4 may be a promising therapeutic approach for treating lupus and other endosomal TLR-dependent diseases [18, 19, 20, 21].

Here, we employ systematic bioinformatics to investigate SNPs in the *SLC15A4* locus, identifying rs35907548 as a likely regulatory variant. We show that this variant indeed underlies activity of a potent enhancer, chromatin interactions, expression of *SLC15A4* and nearby genes,

and ultimately plays a decisive role in maintaining endolysosomal acidification, critical for proper function of immune cells. Our results finely localize SLE risk of the highly associated *SLC15A4* locus and give mechanistic insight into its function.

2 Materials and Methods

2.1 Bioinformatics

We utilized RegulomeDB [22] and Ldlink [23] to prioritize SNP regulatory potential. We added information on *cis*-regulatory elements, including chromatin accessibility (ATAC-seq and DNase-I hypersensitivity), canonical histone modifications (H3K27ac, H3K4me1, H3K4me3), and RNA polymerase II (Pol II) binding, from ENCODE [24] data from immune cell lines. Transcription factor binding sites were taken from ENCODE as well.

To evaluate cell type-specific expression quantitative trait loci (eQTLs) and target gene expression, we used data from ImmuNexUT (Immune Cell Gene Expression Atlas from the University of Tokyo) [25], encompassing 28 immune cell types from both healthy donors and individuals with 10 immune diseases. To broaden our analysis, we added other eQTL databases, including eQTLgen [26], which is tailored for European SNP-based eQTLs.

2.2 Cell lines

Dual-luciferase reporter assays were performed in HEK293, lymphoblastoid cell lines (LCLs; B-cell), Jurkat (T-cell), and HL-60 (promyeloblast). ChIP-qPCR assays were performed in LCLs (GM18624, GM18603). Chromatin-conformation capture (3C) experiments were in the same LCLs. CRISPR/Cas9 gene editing was in HL-60. Endolysosomal pH measurements were in wild-type and gene-edited HL-60. All cell lines were purchased from ATCC (American Type Culture Collection). All cells were tested for mycoplasma by PCR and used between passages 4-7.

2.3 Dual-luciferase assay

To investigate whether the rs35907548 region has enhancer activity, we utilized the well-established Dual-Luciferase Reporter Assay System. A detailed description of the method is provided elsewhere [27, 28]. Briefly, we cloned the 300 bp region, rs35907548 at the middle (chr12:129,282,013-129,282,312, hg 19) locus into pGL4.26 (Promega) and co-transfected with pGL4.74 (internal control) in HEK293, LCL, Jurkat, and HL-60 cells. After 24 hours, enhancer activity was measured using a Synergy H1 spectrophotometer (BioTek). Three experimental replicates were performed per cell type. Statistical significance was assessed by Student's t-test using GraphPad PRISM; p-value < 0.05 was considered significant. We used two non-coding SNPs, rs12831705 and rs34616325, as negative controls for allele-specific luciferase activity.

2.4 Chromatin immunoprecipitation and quantitative PCR (ChIP-qPCR)

To investigate whether the rs35907548 region shows allele-specific binding to specific histone marks (H3K27ac, H3K4me1, and H3K4me3), we conducted ChIP-qPCR assays using the Magnify ChIP assay (Cat No. 492024, Thermo-Fisher, Waltham, MA, USA), following manufacturer's guidelines. A detailed description of the method is provided elsewhere [27, 28].

Briefly, 1.5-2 × 10^6 homozygous rs35907548 risk-"TT" (GM18603) and non-risk-"CC" (GM18624) genotype B-lymphoblastoid LCLs (Coriell) were cross-linked with 1% paraformaldehyde, washed, and sonicated. Immunoprecipitation was performed overnight at 4°C with antibodies against individual histone marks or other DNA-binding proteins pre-incubated with Dynamag magnetic A+G beads. After reverse crosslinking and elution, real-time qPCR analysis was performed with SYBR Green and primers flanking the rs35907548 region using an Applied Biosystems 7900HT qPCR machine. Statistical significance was assessed by Student's t-test using GraphPad PRISM software, and a p-value < 0.05 was considered significant. Detailed methods are described in Supplementary Materials and Methods.

2.5 Chromatin conformation capture (3C) with quantitative PCR (3C-qPCR)

To investigate the chromatin interactions between the promoters of the target genes and the rs35907548 region in an *ex vivo* context, we utilized 3C-qPCR in LCL and Jurkat cells. Detailed protocols are provided elsewhere [27, 28]. Briefly, cells were suspended in complete media with 10% FBS (1 × 10^6/mL media) at 70-80% confluency and cross-linked with 1% paraformaldehyde at room temperature for 10 minutes. Following quenching with 0.2 M glycine, cross-linked cells were lysed in buffer containing protease inhibitors. Cross-linked nuclei were purified, suspended in 0.5% SDS, and incubated at 62°C for 10 minutes followed by quenching with Triton X-100. Perforated nuclei were digested with 400 U *HindIII* and *DpnII* at 37°C overnight and in-nucleus ligated with T4 DNA ligase at 20°C for 4 hours. A small volume of the digested mixture was reserved to evaluate digestion efficiency. DNA was purified from ligated chromatin with proteinase K digestion, phenol-chloroform extraction, and alcohol precipitation. Purified DNA was quantified and diluted for 3C-PCR. Primers were designed to amplify several promoter regions based on restriction maps. Primers within the rs35907548 enhancer were used as common primers for other fragments. Primer sequences are in **Supplementary Table 1**. Cross-linking frequencies were calculated from PCR band intensity. Data were plotted as relative interaction frequency versus genomic distance from rs35907548.

2.6 CRISPR/Cas9 gene editing

To evaluate the functional consequences of the rs35907548 region, we utilized CRISPR/Cas9 gene editing. Detailed experimental protocols are available elsewhere [27, 28]. Short-guide RNA (sgRNA)/Cas9 RNP complexes were introduced into HL-60 cells by Neon Electroporation System. Genomic DNA was extracted three days post-transfection and Sanger-sequenced to verify deletion. Indel efficiency was determined using TIDE and ICE. Subsequently, pooled edited cells were cultured and harvested for gene expression measurements. sgRNA sequences are in **Supplementary Table 1**. Three experimental replicates were performed per cell type. Statistical significance was assessed by Student's t-test using GraphPad PRISM; p-value of < 0.05 was considered significant.

2.7 Quantitative real-time PCR (qRT-PCR)

Pooled CRISPR/Cas9-edited cells were used for RNA purification using the RNA Mini-prep kit (Zymo Research). Purity and concentration were measured using a NanoDrop spectrophotometer. Approximately 700 ng total purified RNA was used to generate cDNA with the iScript cDNA Synthesis kit (Bio-Rad). cDNA was used for PCR to quantify gene expression for *SLC15A4, GLTD1,* and lncRNA AC069262.1 using qRT-PCR LightCycler 480 Instrument II

(Roche) using specific primers and iTaq Universal SYBR Green Supermix (Bio-Rad). To normalize gene expression data, 18S rRNA was used as an internal control. Primer sequences are in **Supplementary Table 1**.

2.8 Endolysosomal pH

To assess differences in endolysosomal pH between WT and KO cells, we used the pHrodo® Red AM Intracellular pH Indicator (Thermo-Fisher) following manufacturer's protocol. Briefly, WT and KO cells (0.2×10^6 cells) were stained with 5 μM pHrodo® Red AM at 37°C for 30 minutes in 96-well plates. Cells were washed with Live Cell Imaging Media, and standard buffers containing 10 μM nigericin and 10 μM valinomycin were added to specific wells for 5 minutes to clamp intracellular pH values at 4.5, 5.5, 6.5, and 7.5. The average cellular fluorescence was measured in triplicate samples using a spectrophotometer (Synergy H1, BioTek). A standard curve was generated for WT and KO samples, showing a linear relationship between intracellular pH and relative fluorescence units.

3 Results

3.1 Systematic bioinformatics prioritizes rs35907548 as a likely regulatory SNP

We used diverse bioinformatics tools to assess the regulatory potential of all 77 high-linkage disequilibrium (LD; r^2>80%) SNPs in and around the *SLC15A4* locus (**Supplementary Table 2**). These 77 SNPs were generated from 5 significantly associated ($p<5 \times 10^{-8}$) index SNPs (rs10847697/rs1385374 [29], rs12370194 [30], rs10593112 [31], and rs11059928 [32]). RegulomeDB2 ranked all SNPs for regulatory potential, prioritizing intronic rs35907548 (**Supplementary Table 2**). We used ENCODE [24]-annotated histone marks, chromatin accessibility, and RNA Pol II occupancy to identify active chromatin, within which rs35907548 lies (**Figure 1**). ENCODE annotated rs35907548 as a "distal enhancer-like" candidate *cis*-regulatory element (cCRE). We used PCHi-C chromatin-conformation data to identify several regions topologically associated with rs35907548 in multiple immune cells (**Figure 1**). To investigate the impact of rs35907548 on target gene expression, we conducted a thorough search in publicly available expression quantitative trait locus (eQTL) databases, specifically focusing on patient-derived primary immune cells [25]. Our findings indicate that rs35907548 acts as an eQTL for *SLC15A4* in various immune cell types (**Supplementary Figure 1**). Notably, plasmacytoid dendritic cells (pDCs) exhibit the most significant eQTLs on *SLC15A4*. Finally, rs35907548 is in the middle of multiple transcription factor binding sites (**Supplementary Figure 2**), where the risk allele T is universally conserved among motifs. These transcription factors include Ikaros family zinc finger protein 1 (IKZF1), IKZF3, E74-like factor 1 (ELF1), Friend leukemia integration 1 transcription factor (FLI1), and ETS1, all of which are SLE risk genes themselves [33, 34, 35, 36, 37].

3.2 rs35907548 risk allele increases enhancer activity in diverse cell lines

We assessed allele-specific enhancer activity of rs35907548 using dual-luciferase reporter enhancer assays in both non-immune HEK293 cells and immune cells including Jurkat, HL-60, and LCL B-cells. In all cell types, the rs35907548 risk allele (TT) exhibited notably higher enhancer activity compared to the non-risk allele (CC) (p-values: < 0.0001, < 0.001, 0.016,

0.0016; as shown in **Figure 2a**). As negative controls, we selected two non-coding SNPs, rs12831705 and rs34616325. Both SNPs, at distances of 12.7 kb and 25.5 kb from the target SNP, respectively, are scored "1f" (i.e., predicted to have a similar regulatory potential as rs35907548; **Supplementary Table 2**) by RegulomeDB and are in strong linkage disequilibrium (LD) with rs35907548. Our results show insignificant regulatory activity for both negative control SNPs in HEK293 and LCL (**Supplementary Figure 3**), underscoring the effects

Using ChIP-grade antibodies, we analyzed the specific binding patterns of three regulatory histone marks—namely, H3K27ac, H3K4me1, and H3K4me3—at the SNP site within its biological context. The quantification of binding was conducted through ChIP-qPCR (**Figure 2b**). Across all three marks, we observed a distinct and heightened level of binding in GM18603 (the risk TT genotype) when compared to GM18624 (the non-risk CC genotype) (H3K27ac: $p = 0.002$, ~4-fold increase; H3K4me1: $p < 0.001$, ~20-fold increase; H3K4me3: $p < 0.001$, ~13-fold increase). This substantiates the prevailing notion that the SNP is situated within an active enhancer region, where intricate interactions involving RNA polymerase, histone marks, chromatin modulators, and allele-specific components collectively oversee the orchestration of transcriptional control at these specific loci.

3.3 The rs35907548 enhancer establishes long-range chromatin interactions with target gene promoters

Given that rs35907548 demonstrated allele-specific enhancer activity across all examined cell lines, we set out to explore its potential involvement in establishing promoter-enhancer connections through chromatin interactions. To explore these interactions, we performed 3C-qPCR experiments between the SNP and gene promoters (**Figure 2b**). Our findings aligned with previous eQTL and pCHiC data, unveiling interactions between the enhancer-SNP region and the promoters of *SLC15A4, GLT1D1*, and LncRNA *AC069262.1* (as illustrated in **Figure 1**). These collective results imply that the enhancer region encompassing rs35907548 could play an active regulatory role in governing the expression of *SLC15A4, GLT1D1*, and LncRNA *AC069262.1*.

3.4 Validating transcriptional effects with CRISPR-based genome editing

To validate the transcriptional regulatory effects of the rs35907548 locus, we used three short-guide RNAs (sgRNAs) to delete ~140 bases around rs35907548 in HL-60 cells (**Figure 4a** and **Supplementary Figure 4**). The deletion was confirmed with Sanger sequencing, and ICE analysis demonstrated high indel efficiency (63%) of pooled cells (data not shown). Subsequently, we determined expression levels of *SLC15A4, GLT1D1*, and lncRNA *AC069262*.1 in KO and WT cells. Expression levels were lower in KO cells than WT (~45%, $p < 0.0001$; ~25%, $p < 0.01$; and ~35%, $p < 0.001$, respectively; **Figure 4b**). These findings provide further evidence that rs35907548 exerts regulatory effects to influence the expression of *SLC15A4* and other target genes through its enhancer activity.

3.5 Impact of the rs35907548 enhancer on endolysosomal acidification

SLC15A4 plays a crucial role in maintaining the acidic environment of the endolysosomal compartment [38]. We found that KO cells expressed less SLC15A4; consistent with this, KO

cells showed a dramatic increase over WT in endolysosomal pH (5.3 *vs*. 4.5, $p < 0.01$; **Figure 5**). These data underscore the crucial link between SLC15A4, maintenance of the acidic environment in endolysosomal compartments, and subsequent effects on cellular pH regulation.

4 Discussion

GWAS analyses invariably produce peaks of association; these peaks are often quite large, covering many genes and 1000s of SNPs. Sifting through these genes and SNPs to identify the true causal variants and their mechanisms of action from broad GWAS peaks is challenging [39], but an essential prerequisite for understanding disease risk and opportunities for diagnosis and treatment. Several studies have linked *SLC15A4* SNPs to SLE risk in Asians and Europeans [18, 19, 40], albeit with no experimental testing.

Here, we localize *SLC15A4* SLE risk to the intronic SNP rs35907548, at the center of an enhancer modulating expression of – and physically interacting with – *SLC15A4* and nearby genes. Dual-luciferase reporter assays revealed strong enhancer activity, particularly of the risk "TT" allele, in diverse cell types. The risk allele consistently showed substantially higher binding to active histone marks. We confirmed the activity of the rs35907548 enhancer in cells by creating CRISPR/Cas9 knock-out cells, which showed significantly lower levels of *SLC15A4*, *GLT1D1*, and the uncharacterized lncRNA *AC069262.1*. SLC15A4 is a histidine transporter located primarily on the endolysosomal membrane, where it plays several pivotal roles including maintenance of endolysosomal acidification [19, 38], thus regulating protein degradation, inflammation, endocytosis, and autophagy, among other processes [41, 42, 43]. Accordingly, KO cells failed to properly acidify the endolysosome, demonstrating the profound effects of the rs35907548 enhancer on cell physiology.

Several studies have suggested SLC15A4 as a potential therapeutic target for systemic lupus erythematosus (SLE) and other autoimmune diseases [21, 38, 44]. In mouse models, the absence of *SLC15A4* has been shown to confer resistance to the development of multiple autoimmune diseases, including dextran sulfate sodium (DSS)-induced colitis (wherein *Slc15a4*$^{-/-}$ pDCs fail to produce IFNα upon TLR7 agonist R848 stimulation) and the Faslpr model of SLE [18]. This observation is significant as TLR7/8 inhibitors are currently in clinical trials for SLE [19]. These findings collectively underscore the dysregulation of lysosomal pH, particularly deacidification, in autoimmune diseases such as SLE. This dysregulation is closely linked to the activity of lysosomal hydrolases.

Interestingly, two widely used anti-malarial drugs, hydroxychloroquine and chloroquine, are commonly used to treat SLE. These drugs have many cellular targets and complicated mechanisms of action. The weak base drugs accumulate in acidic compartments like the lysosome, where they increase pH by becoming protonated. The drugs also directly inhibit lysosomal hydrolases, increasing pH by a further 2 units. Finally, they block stimulation of Toll-like receptor 9 (TLR9) family members, decreasing innate immune reactivity. Together, these effects decrease lysosome activity and inflammasome activation and mitigate inflammation. The precise mechanisms through which these drugs benefit SLE patients are still being investigated, and other cellular targets may be discovered. Despite this incomplete understanding, these drugs are critical tools in managing SLE and related inflammatory diseases [45, 46, 47].

Crucially, our study identifies *SLC15A4* as a genetic locus linked to substantially elevated SLE risk. Given its central role in governing endolysosomal pH and activity, this genetic variation likely contributes to SLE susceptibility, aligning well with the fact that endolysosomal modulator drugs are essentially the only SLE drugs of note. Our findings, as depicted in Figures 2a and 4, point to heightened SLC15A4 expression in SLE patients. Importantly, a recent study [48] lends additional support to this assertion. Despite its reliance on a limited sample size, the study reports significantly elevated *SLC15A4* mRNA in PBMCs from SLE patients compared to healthy controls. This observation bolsters the association between *SLC15A4* expression and autoimmune diseases such as SLE [49].

rs35907548 lies in the middle of multiple conserved transcription factor binding sites, including those for SLE risk genes and white blood cell factors IKZF1, IKZF3, ELF1, FLI1, and ETS1; the ancestral risk allele T is universally present in the binding motifs. It is likely that proper binding of these transcription factors is critical to enhancer function, and indeed the risk allele T shows much higher levels of active chromatin marks and enhancer activity, implying a heightened regulatory role.

The convergence of various elements – including the existence of evolutionarily conserved binding sites, the participation of transcription factors associated with immune processes and SLE, and the distinct presence of the risk allele within these motifs – collectively implies a complex interplay that likely underlies the regulatory function of the enhancer. These findings emphasize the intricate and multifaceted nature of genetic aspects of disease susceptibility and cellular function.

The adjacent locus Glycosyltransferase 1 Domain Containing 1 (*GLT1D1*) has also been flagged as an SLE risk locus [50]. GLT1D1 targets programmed cell death-ligand 1 (PD-L1) for glycosylation [51]; glycosylated PD-L1 is strongly immunosuppressive and correlates with B-cell non-Hodgkin's lymphoma progression. In addition to the links to endolysosomal acidification, TLR signaling, and autophagy through SLC15A4, future experiments will investigate the contribution of rs35907548 and other SNPs to SLE risk through *GLT1D1*. The role of lncRNA *AC069262.1* in biological processes and diseases, including SLE, remains largely unexplored. Further investigations are required to uncover its potential significance and contribution to SLE.

In the present study, our primary objective was to define susceptibility regions likely to contain functional SNP(s) linked to increased lupus risk. Leveraging bioinformatics and experiment, we identified rs35907548 within an active enhancer with the capacity to modulate expression of target genes with critical immune functions. Nevertheless, we acknowledge a limitation of our study, namely that we have not explicitly demonstrated effects of the single base-pair change on target gene expression, endolysosomal pH, or other cellular properties. To address this limitation, advanced techniques like CRISPR base-editing or prime-editing will be required. We have demonstrated numerous specific effects of the single base-pair change (transcription factor binding, enhancer activity, chromatin contacts) from both in vitro experiments and from data acquired from *in vivo* samples, which are strongly consistent with the cellular effects that we observe from CRISPR deletion of the rs35907548 locus. Future studies will validate the cellular effects of the single base-pair change, allowing stronger conclusions to be made about the effect of the rs35907548 SNP on SLE development and progression.

In summary, our study both identifies a SNP causally underlying SLE risk association at the *SLC15A4* locus and establishes an analytical and experimental framework for studying risk SNPs at *SLC15A4* and other loci. The results and framework may contribute to future investigation of therapeutic interventions and diagnostics.

5 Data availability statement

All supporting data is available upon request from the corresponding author.

6 Conflict of interest

Authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

7 Author Contributions

SKN conceived and designed the study. MKS, GPM, HRR conducted and validated the experiments. MFN, LLL, SKN organized the database and conducted bioinformatics analyses. SKN, GPM, MKS, LLL wrote the manuscript. All authors contributed to manuscript revision, and read and approved the submitted version.

8 Funding

Research reported in this publication was supported by National Institutes of Health grants R01AI172255, R21 AI168943, and R21 AI144829.

9 Acknowledgments

The authors wish to sincerely express their appreciation to Dr. Valerie Harris for her invaluable advice on setting up the pHrodo experiments. Furthermore, they extend their thanks to Ms. Louise Williamson for her technical assistance during the preparation of the manuscript.

10 Figure Legends

Figure 1. Bioinformatics. Cartoon showing the genomic location of rs35907548 (black vertical line), PCHi-C interactions from rs35907548 to the promoters (rectangle boxes) of the neighboring genes, regulatory histone marks, and Pol 2.

Figure 2. Allelic-specific regulatory effects of the rs35907548enhancer. (A) Luciferase reporter assays on 4 cell types. **(B)** ChIP-qPCR with risk (R) and non-risk (NR) LCLs.

Figure 3. 3C interaction analysis. (A) Schematic of the rs35907548 region and neighboring genes. SacI restriction enzyme sites were used to design primers. Small arrows represent primer locations and orientations. The common primer at the rs35907548 region is common for all other primers at *SLC15A4*, *GLT1D1* and *TMEM132C* promoter regions. Big arrowheads represent transcriptional start sites (TSSs). **(B-E)** Graphs show relative interactions of rs35907548 regions with different genomic regions in HEK (**B**), Jurkat **(C)** and EBV-transformed LCL B-cells **(D,**

E). The relative interaction frequency for each set of primers represents the intensity of PCR. X-axis shows genomic distance (in kb) in forward and reverse directions from rs35907548 (0 kb).

Figure 4. CRISPR/Cas9-based enhancer deletion. (A) CRISPR-based deletion, **(B)** qPCR-based target gene expression between WT and KO.

Figure 5. Comparing the endolysosomal acidification between WT and KO cells. (A) pHrodo™ Red AM-stained WT and KO cells under a fluorescent microscope. **(B)** Standard curve for WT and KO cells using pHrodo™ Red AM with Intracellular pH Calibration Buffer Kit for the translation of fluorescence ratios into pH. **(C)** pH determined for WT and KO using standard curve and manufacturer's protocols. RFU = Relative Fluorescence Units.

Supplementary Figure 1. Immune cell type-specific eQTLs for rs35907548. The genotypes and their corresponding sample sizes are shown below each box plot. (T/T) is the risk allele for rs35907548. USM = unstimulated, SM = stimulated, DN = double negative.

Supplementary Figure 2. Transcription factor binding sites overlapping rs35907548. The ancestral risk allele T (red oval) produces more favorable motifs for all TFs. Green highlights bases where the consensus motif is present.

Supplementary Figure 3. Negative control for assessing allelic-specific enhancer activity. Luciferase reporter assays were conducted on sequences surrounding the predicted inactive SNPs rs12831705 and rs34616325 in HEK-293 and LCL cells, revealing non-significant allele-specific regulatory effects.

Supplementary Figure 4. CRISPR/Cas9-based enhancer deletion of 139 bp region containing rs35907548.

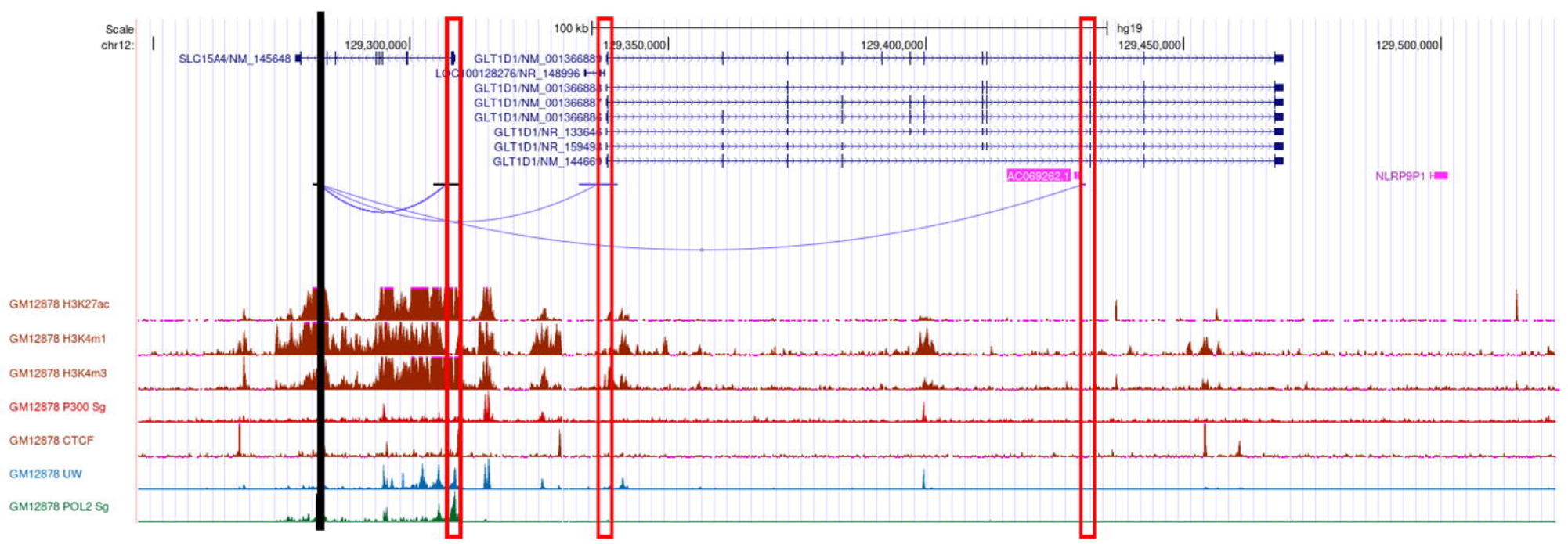
References

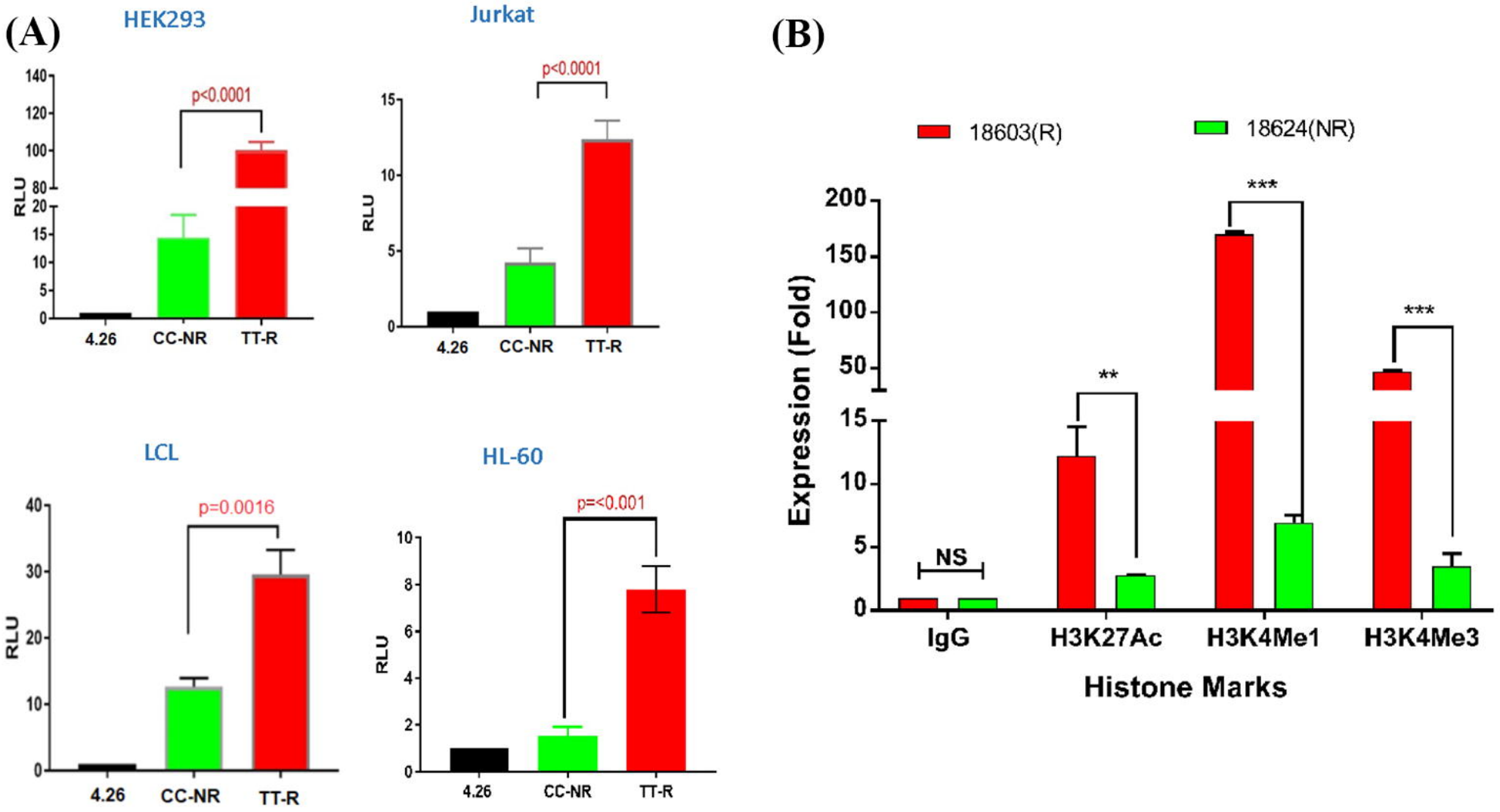
1. Crow MK. Pathogenesis of systemic lupus erythematosus: risks, mechanisms and therapeutic targets. Ann Rheum Dis. 2023.
2. Aguirre A, Izadi Z, Trupin L, Barbour KE, Greenlund KJ, Katz P, et al. Race, Ethnicity, and Disparities in the Risk of End-Organ Lupus Manifestations Following a Systemic Lupus Erythematosus Diagnosis in a Multiethnic Cohort. Arthritis Care Res (Hoboken). 2023;75(1):34-43.
3. Kabani N, Ginzler EM. Is ethnicity linked to the severity of SLE manifestations? Nat Rev Rheumatol. 2019;15(9):515-6.
4. Lewis MJ, Jawad AS. The effect of ethnicity and genetic ancestry on the epidemiology, clinical features and outcome of systemic lupus erythematosus. Rheumatology (Oxford). 2017;56(suppl_1):i67-i77.
5. Maningding E, Dall'Era M, Trupin L, Murphy LB, Yazdany J. Racial and Ethnic Differences in the Prevalence and Time to Onset of Manifestations of Systemic Lupus Erythematosus: The California Lupus Surveillance Project. Arthritis Care Res (Hoboken). 2020;72(5):622-9.
6. Kuo CF, Grainge MJ, Valdes AM, See LC, Luo SF, Yu KH, et al. Familial Aggregation of Systemic Lupus Erythematosus and Coaggregation of Autoimmune Diseases in Affected Families. JAMA Intern Med. 2015;175(9):1518-26.
7. Harley ITW, Sawalha AH. Systemic lupus erythematosus as a genetic disease. Clin Immunol. 2022;236:108953.
8. Molineros JE, Yang W, Zhou XJ, Sun C, Okada Y, Zhang H, et al. Confirmation of five novel susceptibility loci for systemic lupus erythematosus (SLE) and integrated network analysis of 82 SLE susceptibility loci. Hum Mol Genet. 2017;26(6):1205-16.
9. Owen KA, Bell KA, Price A, Bachali P, Ainsworth H, Marion MC, et al. Molecular pathways identified from single nucleotide polymorphisms demonstrate mechanistic differences in systemic lupus erythematosus patients of Asian and European ancestry. Sci Rep. 2023;13(1):5339.
10. Theofilopoulos AN, Kono DH, Baccala R. The multiple pathways to autoimmunity. Nat Immunol. 2017;18(7):716-24.
11. Owen KA, Price A, Ainsworth H, Aidukaitis BN, Bachali P, Catalina MD, et al. Analysis of Trans-Ancestral SLE Risk Loci Identifies Unique Biologic Networks and Drug Targets in African and European Ancestries. Am J Hum Genet. 2020;107(5):864-81.
12. Wang YF, Wei W, Tangtanatakul P, Zheng L, Lei Y, Lin Z, et al. Identification of Shared and Asian-Specific Loci for Systemic Lupus Erythematosus and Evidence for Roles of Type III Interferon Signaling and Lysosomal Function in the Disease: A Multi-Ancestral Genome-Wide Association Study. Arthritis Rheumatol. 2022;74(5):840-8.
13. Toyama-Sorimachi N, Kobayashi T. Lysosomal amino acid transporters as key players in inflammatory diseases. Int Immunol. 2021;33(12):853-8.
14. Yamashita T, Shimada S, Guo W, Sato K, Kohmura E, Hayakawa T, et al. Cloning and functional expression of a brain peptide/histidine transporter. J Biol Chem. 1997;272(15):10205-11.
15. Daniel H, Kottra G. The proton oligopeptide cotransporter family SLC15 in physiology and pharmacology. Pflugers Arch. 2004;447(5):610-8.

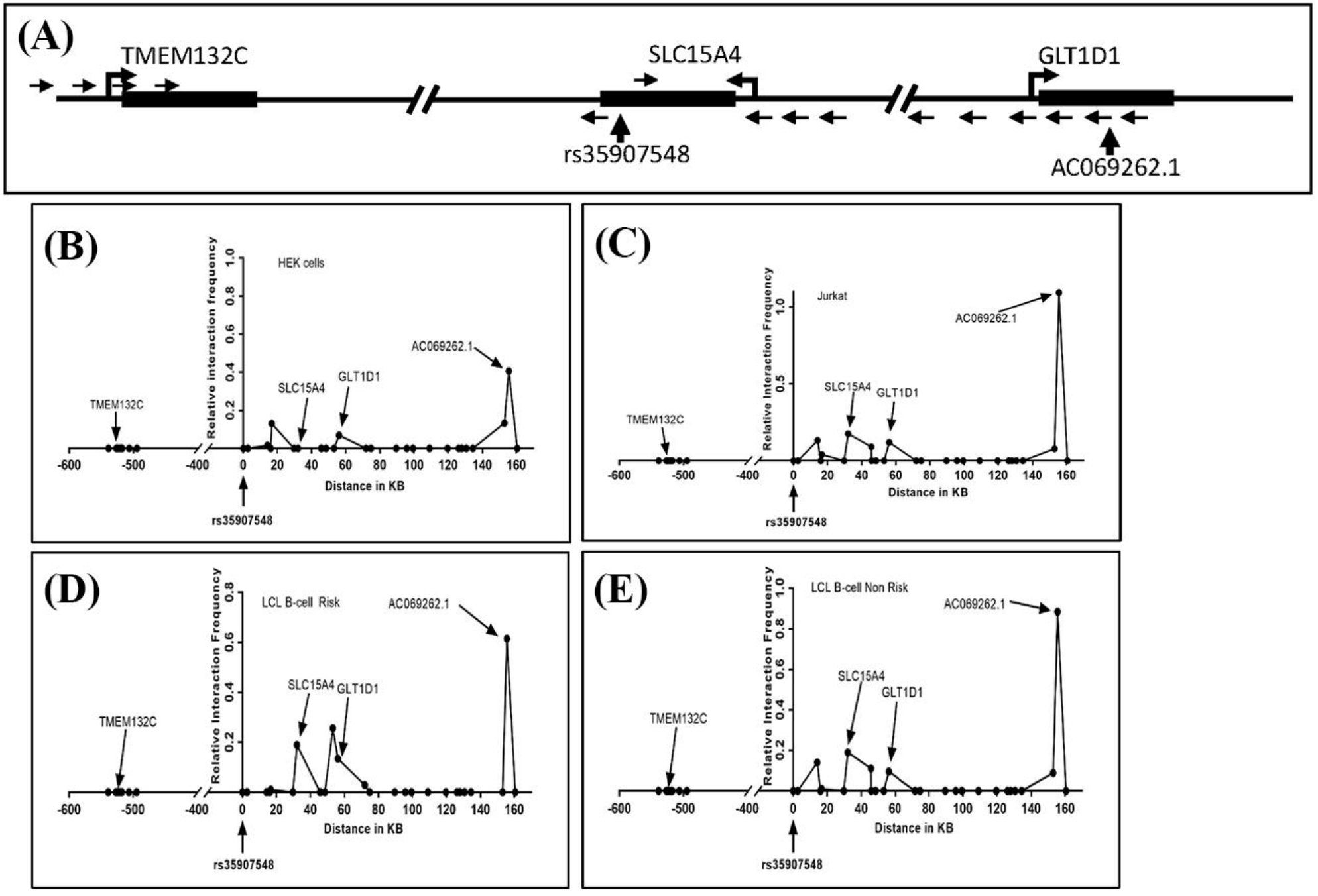
16. Smith DE, Clémençon B, Hediger MA. Proton-coupled oligopeptide transporter family SLC15: physiological, pharmacological and pathological implications. Mol Aspects Med. 2013;34(2-3):323-36.
17. Heinz LX, Lee J, Kapoor U, Kartnig F, Sedlyarov V, Papakostas K, et al. TASL is the SLC15A4-associated adaptor for IRF5 activation by TLR7-9. Nature. 2020;581(7808):316-22.
18. Baccala R, Gonzalez-Quintial R, Blasius AL, Rimann I, Ozato K, Kono DH, et al. Essential requirement for IRF8 and SLC15A4 implicates plasmacytoid dendritic cells in the pathogenesis of lupus. Proc Natl Acad Sci U S A. 2013;110(8):2940-5.
19. Katewa A, Suto E, Hui J, Heredia J, Liang J, Hackney J, et al. The peptide symporter SLC15a4 is essential for the development of systemic lupus erythematosus in murine models. PLoS One. 2021;16(1):e0244439.
20. Kobayashi T, Tsutsui H, Shimabukuro-Demoto S, Yoshida-Sugitani R, Karyu H, Furuyama-Tanaka K, et al. Lysosome biogenesis regulated by the amino-acid transporter SLC15A4 is critical for functional integrity of mast cells. Int Immunol. 2017;29(12):551-66.
21. Rimann I, Gonzalez-Quintial R, Baccala R, Kiosses WB, Teijaro JR, Parker CG, et al. The solute carrier SLC15A4 is required for optimal trafficking of nucleic acid-sensing TLRs and ligands to endolysosomes. Proc Natl Acad Sci U S A. 2022;119(14):e2200544119.
22. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. 2012;22(9):1790-7.
23. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics. 2015;31(21):3555-7.
24. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489(7414):57-74.
25. Ota M, Nagafuchi Y, Hatano H, Ishigaki K, Terao C, Takeshima Y, et al. Dynamic landscape of immune cell-specific gene regulation in immune-mediated diseases. Cell. 2021;184(11):3006-21.e17.
26. Võsa U, Claringbould A, Westra HJ, Bonder MJ, Deelen P, Zeng B, et al. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. Nat Genet. 2021;53(9):1300-10.
27. Fazel-Najafabadi M, Rallabandi HR, Singh MK, Maiti GP, Morris J, Looger LL, et al. Discovery and Functional Characterization of Two Regulatory Variants Underlying Lupus Susceptibility at 2p13.1. Genes (Basel). 2022;13(6).
28. Singh B, Maiti GP, Zhou X, Fazel-Najafabadi M, Bae SC, Sun C, et al. Lupus Susceptibility Region Containing CDKN1B rs34330 Mechanistically Influences Expression and Function of Multiple Target Genes, Also Linked to Proliferation and Apoptosis. Arthritis Rheumatol. 2021;73(12):2303-13.
29. Han JW, Zheng HF, Cui Y, Sun LD, Ye DQ, Hu Z, et al. Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. Nat Genet. 2009;41(11):1234-7.
30. Sun C, Molineros JE, Looger LL, Zhou XJ, Kim K, Okada Y, et al. High-density genotyping of immune-related loci identifies new SLE risk variants in individuals with Asian ancestry. Nat Genet. 2016.

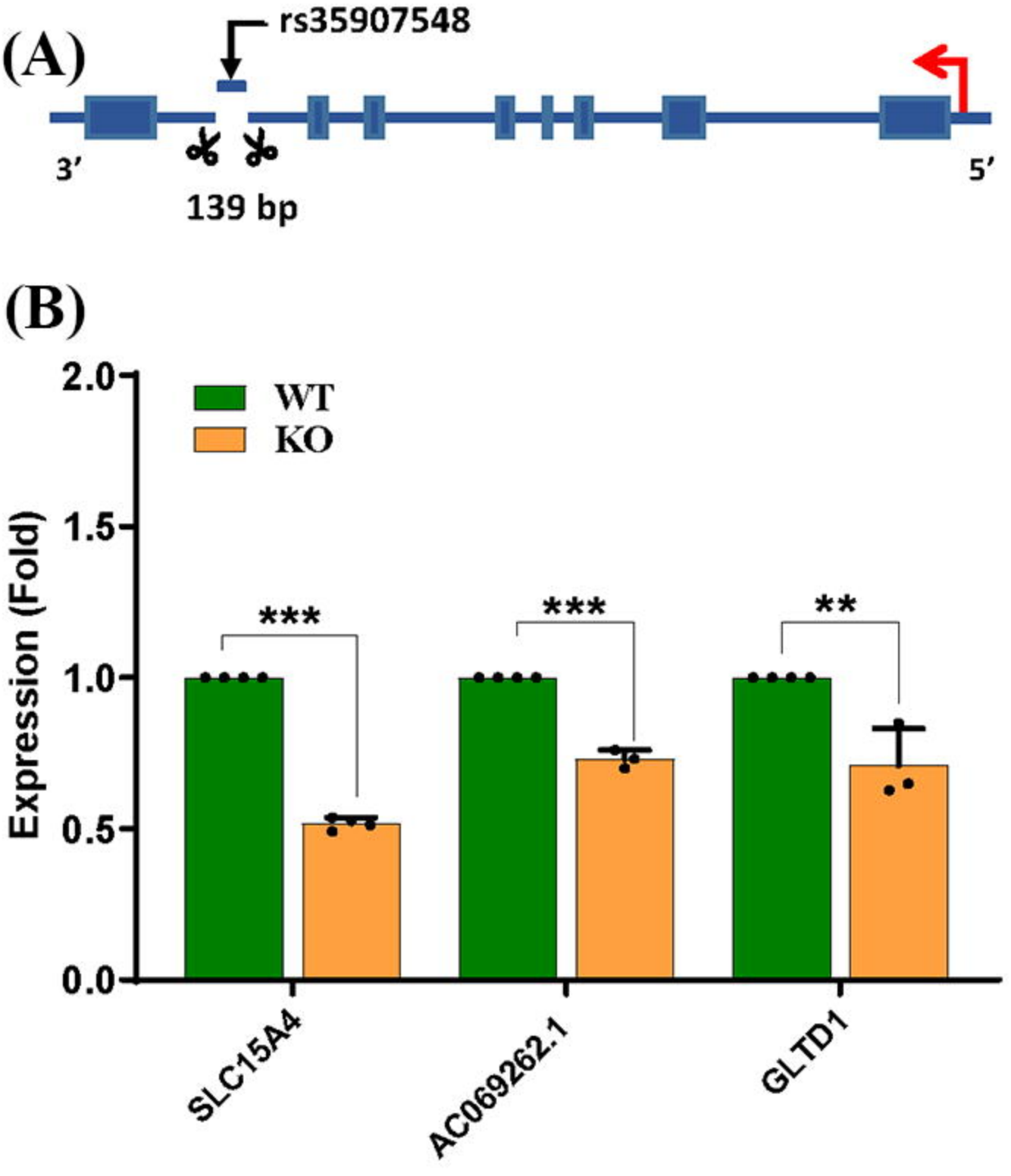
31. Langefeld CD, Ainsworth HC, Cunninghame Graham DS, Kelly JA, Comeau ME, Marion MC, et al. Transancestral mapping and genetic load in systemic lupus erythematosus. Nat Commun. 2017;8:16021.
32. Yin X, Kim K, Suetsugu H, Bang SY, Wen L, Koido M, et al. Meta-analysis of 208370 East Asians identifies 113 susceptibility loci for systemic lupus erythematosus. Ann Rheum Dis. 2020.
33. Cai X, Qiao Y, Diao C, Xu X, Chen Y, Du S, et al. Association between polymorphisms of the IKZF3 gene and systemic lupus erythematosus in a Chinese Han population. PLoS One. 2014;9(10):e108661.
34. Garrett-Sinha LA, Kearly A, Satterthwaite AB. The Role of the Transcription Factor Ets1 in Lupus and Other Autoimmune Diseases. Crit Rev Immunol. 2016;36(6):485-510.
35. Hu SJ, Wen LL, Hu X, Yin XY, Cui Y, Yang S, et al. IKZF1: a critical role in the pathogenesis of systemic lupus erythematosus? Mod Rheumatol. 2013;23(2):205-9.
36. Sato S, Lennard Richard M, Brandon D, Jones Buie JN, Oates JC, Gilkeson GS, et al. A critical role of the transcription factor fli-1 in murine lupus development by regulation of interleukin-6 expression. Arthritis Rheumatol. 2014;66(12):3436-44.
37. Yang J, Yang W, Hirankarn N, Ye DQ, Zhang Y, Pan HF, et al. ELF1 is associated with systemic lupus erythematosus in Asian populations. Hum Mol Genet. 2011;20(3):601-7.
38. Kobayashi T, Shimabukuro-Demoto S, Yoshida-Sugitani R, Furuyama-Tanaka K, Karyu H, Sugiura Y, et al. The histidine transporter SLC15A4 coordinates mTOR-dependent inflammatory responses and pathogenic antibody production. Immunity. 2014;41(3):375-88.
39. Claussnitzer M, Cho JH, Collins R, Cox NJ, Dermitzakis ET, Hurles ME, et al. A brief history of human disease genetics. Nature. 2020;577(7789):179-89.
40. He CF, Liu YS, Cheng YL, Gao JP, Pan TM, Han JW, et al. TNIP1, SLC15A4, ETS1, RasGRP3 and IKZF1 are associated with clinical features of systemic lupus erythematosus in a Chinese Han population. Lupus. 2010;19(10):1181-6.
41. Bonam SR, Wang F, Muller S. Lysosomes as a therapeutic target. Nat Rev Drug Discov. 2019;18(12):923-48.
42. Cao M, Luo X, Wu K, He X. Targeting lysosomes in human disease: from basic research to clinical applications. Signal Transduct Target Ther. 2021;6(1):379.
43. Noguchi M, Hirata N, Tanaka T, Suizu F, Nakajima H, Chiorini JA. Autophagy as a modulator of cell death machinery. Cell Death Dis. 2020;11(7):517.
44. Lazar DC, Wang WW, Chiu T-Y, Li W, Jadhav AM, Wozniak JM, et al. Chemoproteomics-guided development of SLC15A4 inhibitors with anti-inflammatory activity. bioRxiv [Preprint]. 2022. doi: 10.1101/2022.10.07.511216.
45. Ponticelli C, Moroni G. Hydroxychloroquine in systemic lupus erythematosus (SLE). Expert Opin Drug Saf. 2017;16(3):411-9.
46. Dima A, Jurcut C, Chasset F, Felten R, Arnaud L. Hydroxychloroquine in systemic lupus erythematosus: overview of current knowledge. Ther Adv Musculoskelet Dis. 2022;14:1759720x211073001.
47. Dörner T. Therapy: Hydroxychloroquine in SLE: old drug, new perspectives. Nat Rev Rheumatol. 2010;6(1):10-1.
48. Li S, Wu Q, Jiang Z, Wu Y, Li Y, Ni B, et al. miR-31-5p Regulates Type I Interferon by Targeting SLC15A4 in Plasmacytoid Dendritic Cells of Systemic Lupus Erythematosus. J Inflamm Res. 2022;15:6607-16.

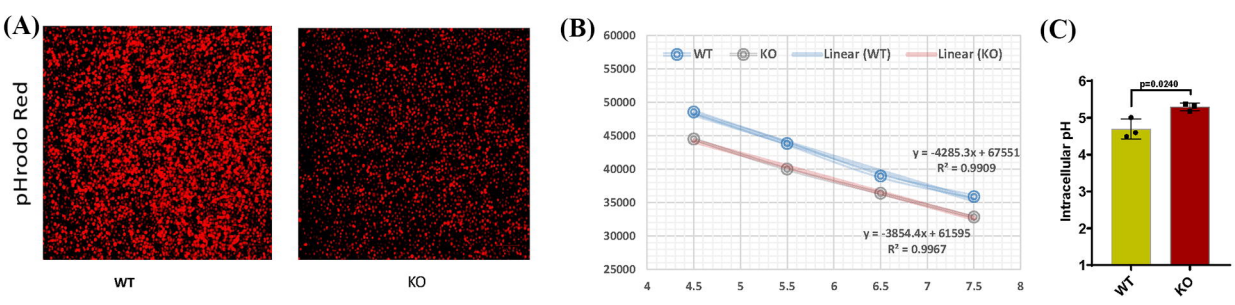
49. Fujio K. Functional Genome Analysis for Immune Cells Provides Clues for Stratification of Systemic Lupus Erythematosus. Biomolecules. 2023;13(4).
50. Joo YB, Lim J, Tsao BP, Nath SK, Kim K, Bae SC. Genetic variants in systemic lupus erythematosus susceptibility loci, XKR6 and GLT1D1 are associated with childhood-onset SLE in a Korean cohort. Sci Rep. 2018;8(1):9962.
51. Liu X, Zhang Y, Han Y, Lu W, Yang J, Tian J, et al. Overexpression of GLT1D1 induces immunosuppression through glycosylation of PD-L1 and predicts poor prognosis in B-cell lymphoma. Mol Oncol. 2020;14(5):1028-44.

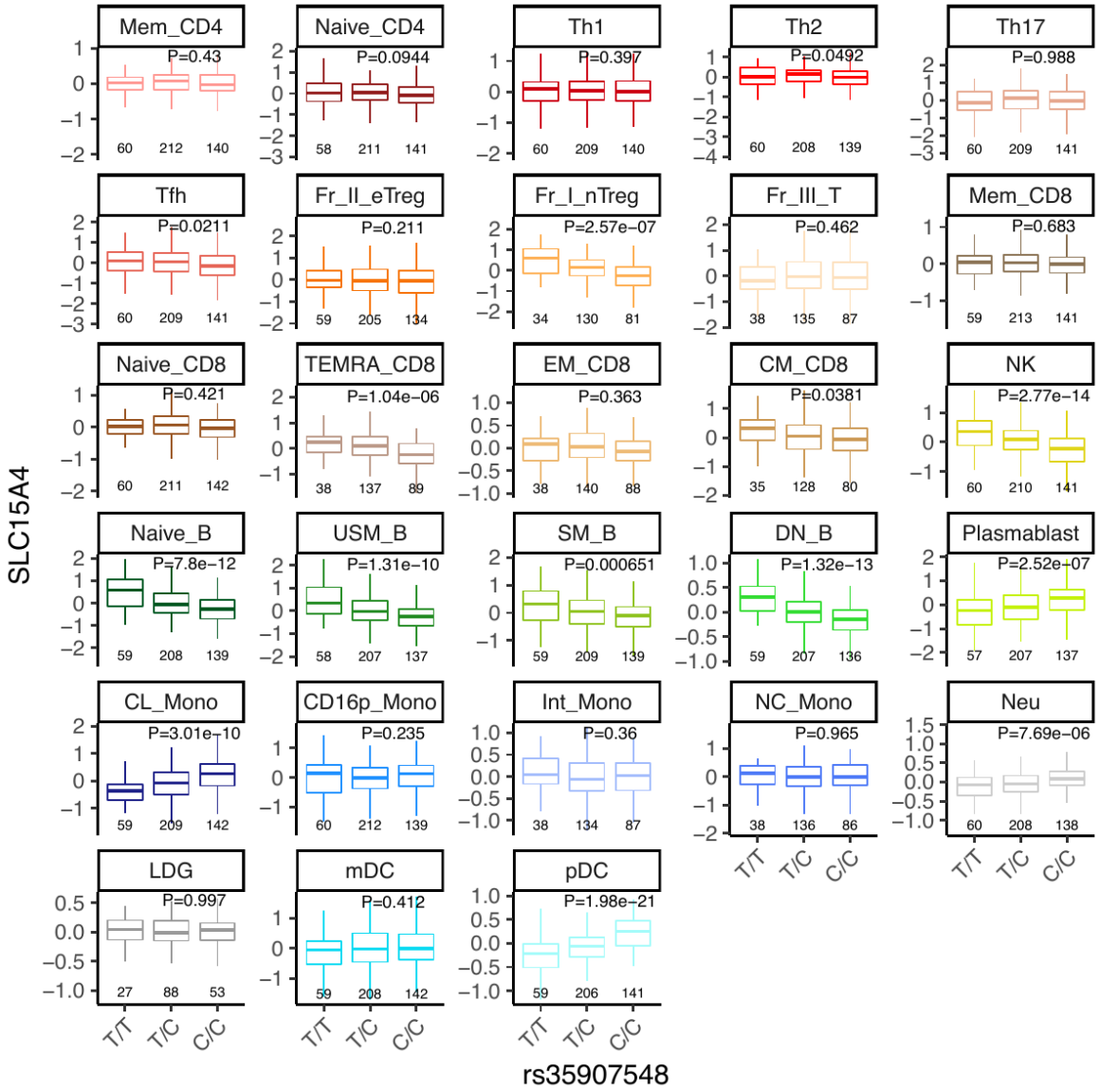

Scale
chr12:
100 kb
hg19
129,300,000
129,350,000
129,400,000
129,450,000
129,500,000
SLC15A4/NM_145648
GLT1D1/NM_001366889
LOC100128276/NR_148996
GLT1D1/NM_001366884
GLT1D1/NM_001366887
GLT1D1/NM_001366886
GLT1D1/NR_133645
GLT1D1/NR_159493
GLT1D1/NM_144669
AC069262.1
NLRP9P1
GM12878 H3K27ac
GM12878 H3K4m1
GM12878 H3K4m3
GM12878 P300 Sg
GM12878 CTCF
GM12878 UW
GM12878 POL2 Sg


(A)
HEK293
p<0.0001
RLU
140
120
100
80
20
15
10
5
0
4.26
CC-NR
TT-R
Jurkat
p<0.0001
RLU
15
10
5
0
4.26
CC-NR
TT-R
LCL
p=0.0016
RLU
40
30
20
10
0
4.26
CC-NR
TT-R
HL-60
p=<0.001
RLU
10
8
6
4
2
0
4.26
CC-NR
TT-R
(B)
18603(R)
18624(NR)
Expression (Fold)
200
150
100
50
15
10
5
0
NS
**
***
***
IgG
H3K27Ac
H3K4Me1
H3K4Me3
Histone Marks

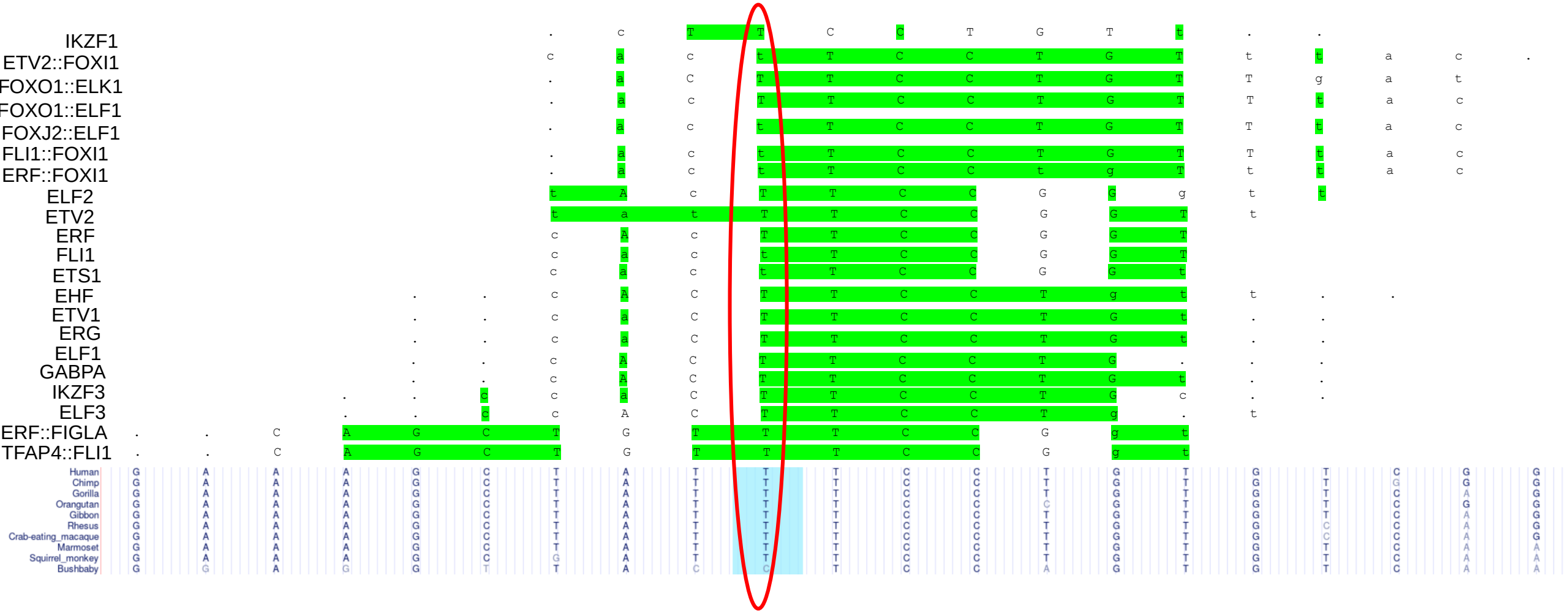

(A)
TMEM132C
SLC15A4
GLT1D1
rs35907548
AC069262.1
(B)
HEK cells
Relative interaction frequency
TMEM132C
SLC15A4
GLT1D1
AC069262.1
Distance in KB
rs35907548
(C)
Jurkat
Relative Interaction Frequency
TMEM132C
SLC15A4
GLT1D1
AC069262.1
Distance in KB
rs35907548
(D)
LCL B-cell Risk
Relative Interaction Frequency
TMEM132C
SLC15A4
GLT1D1
AC069262.1
Distance in KB
rs35907548
(E)
LCL B-cell Non Risk
Relative Interaction Frequency
TMEM132C
SLC15A4
GLT1D1
AC069262.1
Distance in KB
rs35907548


(A)
rs35907548
3'
5'
139 bp
(B)
WT
KO
Expression (Fold)
2.0
1.5
1.0
0.5
0.0
***
***
**
SLC15A4
AC069262.1
GLTD1

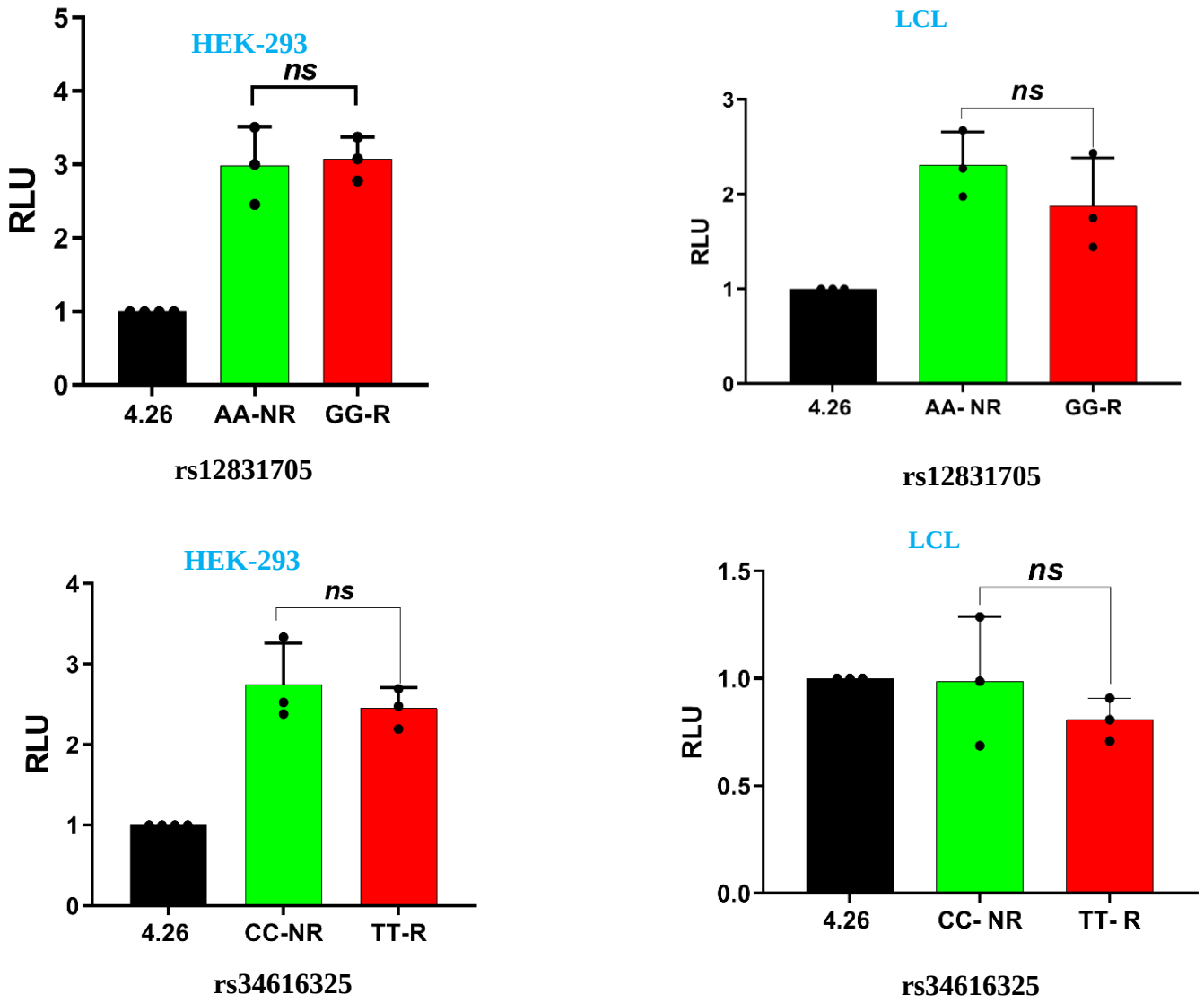

(A)
pHrodo Red
WT
KO
(B)
WT
KO
Linear (WT)
Linear (KO)
y = -4285.3x + 67551
R² = 0.9909
y = -3854.4x + 61595
R² = 0.9967
60000
55000
50000
45000
40000
35000
30000
25000
4
4.5
5
5.5
6
6.5
7
7.5
8
(C)
p=0.0240
Intracellular pH
6
5
4
3
2
1
WT
KO



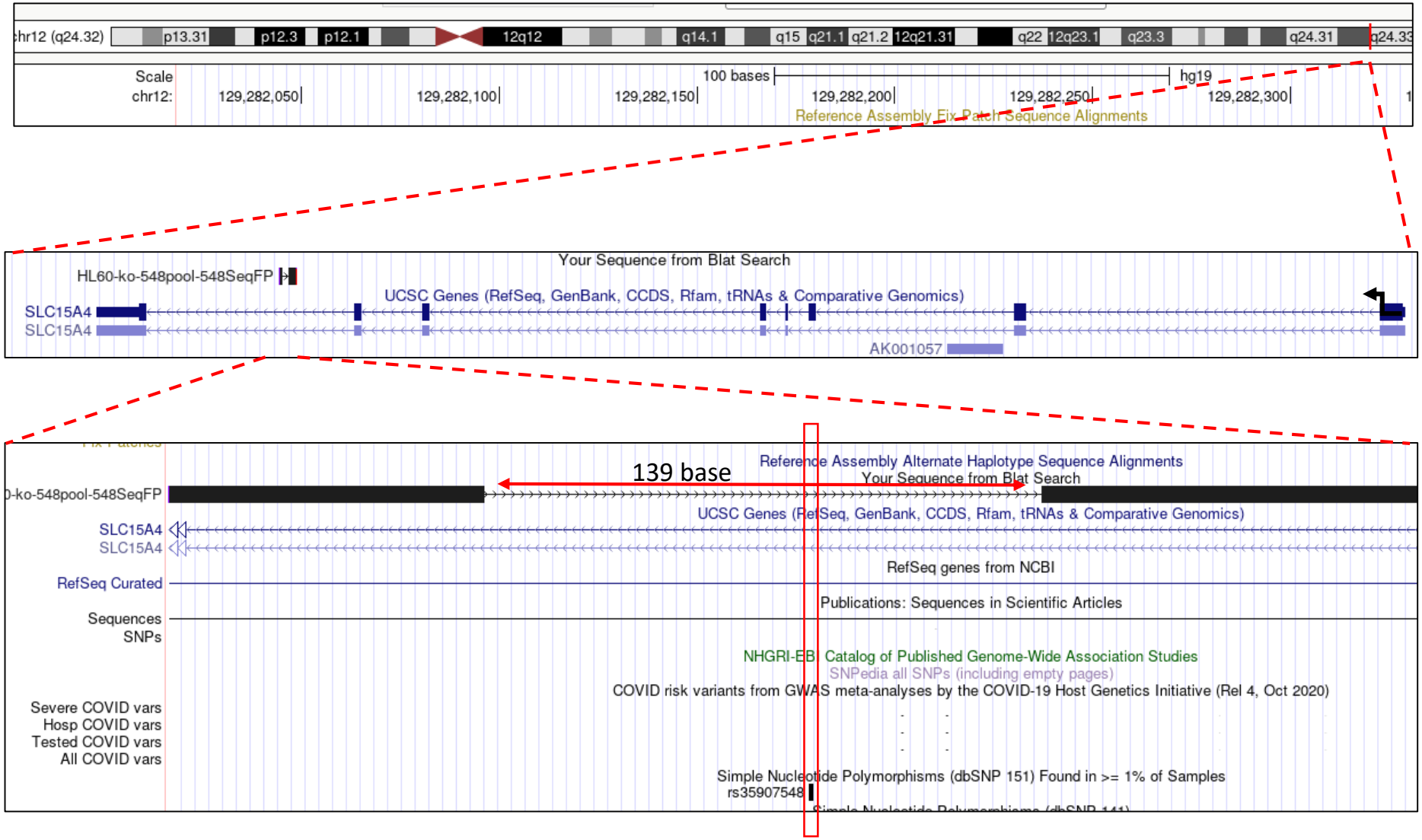
Supple Fig 1. Immune cell type-specific eQTLs for rs35907548. The genotypes and their corresponding sample sizes are shown at the bottom of each box plot. (C/C) is the risk for rs35907548. USM = unstimulated, SM = stimulated, DN = double negative.



Supp Fig. 2. Transcription factor binding sites overlapping rs35907548. The ancestral risk allele T (red oval) produces more favorable motifs for all TFs. Green highlights bases where the consensus motif is present.



Supple Fig 3: Supple Fig 3: Negative control for assessing allelic-specific enhancer activity. Luciferase reporter assays were conducted on sequences containing rs12831705 and rs34616325 across HEK293 and LCL cells, revealing non-significant allele-specific regulatory effects.



Supple Fig 4: CRISPR/Cas9-based enhancer deletion of 139 bp region containing rs35907548