TCL1A, a novel transcription factor and a co-regulator of NF-κB p65: SNP and estrogen-dependence

Ming-Fen Ho, Ph.D., Edroaldo Lummertz da Rocha, Ph.D., Cheng Zhang, Ph.D., James N. Ingle, M.D., Paul E. Goss, M.D., Ph.D., Lois E. Shepherd, M.D., Michiaki Kubo, M.D., Ph.D., Liewei Wang, M.D., Ph.D., Hu Li, Ph.D., and Richard M. Weinshilboum, M.D.

Division of Clinical Pharmacology, Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA (MH, ELR, CZ, LW, HL and RW).

Division of Hematology/Oncology, Department of Medicine, Massachusetts General Hospital Cancer Center, 55 Fruit Street, Lawrence House, LRH-302, Boston, Massachusetts 02114, USA (PG).

Canadian Cancer Trials Group, 10 Stuart Street, Kingston, Ontario K7L 3N6 Canada (LS).

Division of Medical Oncology, Department of Oncology, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA (JI).

RIKEN Center for Integrative Medical Science, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, 230-0045 Japan (MK).

Running title: *TCL1A*, a transcription factor and co-regulator of NF-κB p65

Corresponding author: Richard Weinshilboum, M.D., Mayo Clinic 200 First Street SW, Rochester, MN 55905, Phone: 507-284-2790, FAX: 507-284-4455, Email: weinshilboum.richard@mayo.edu

Number of text pages: 23

Number of tables: 0 (+4 Supplementary Tables)

Number of figures: 8 (+5 Supplementary Figures)

References: 27

Abstract, 196/250 words

Introduction: 457/750 words

Discussion: 1265/1500 words

List of abbreviations

4OH-TAM	4-hydroxytamoxifen
ChIP	Chromatin immunoprecipitation
ChIP-seq	Chromatin immunoprecipitation sequencing
Co-IP	Co-Immunoprecipitation
DAVID	The Database for Annotation, Visualization and

	Integrated Discovery
E2	Estradiol
EMSA	Electrophoretic mobility shift assay
EREs	Estrogen response elements
ERα	Estrogen receptor alpha
GO	Gene ontology
GTEx	The Genotype-Tissue Expression Project database
GWAS	Genome-wide association study
LCLs	Lymphoblastoid cell lines
NGS	Next generation sequencing
RNA-seq	RNA sequencing
SNPs	Single nucleotide polymorphisms
TCL1A	T cell leukemia 1A
TLRs	Toll like receptors
TSS	Transcription start site

Recommended section: Cellular and Molecular

Abstract

TCL1A single nucleotide polymorphisms (SNPs) have been associated with aromatase inhibitor-induced musculoskeletal adverse events. We previously demonstrated that TCL1A is estradiol (E2) inducible and plays a critical role in the regulation of cytokines, chemokines and toll like receptors in a *TCL1A* SNP genotype and estrogen-dependent fashion. Furthermore, *TCLIA* SNP-dependent expression phenotypes can be "reversed" by exposure to selective estrogen receptor modulators, e.g, 4-hydroxytamoxifen (4OH-TAM). The present study was designed to comprehensively characterize the role of TCL1A in transcriptional regulation across the genome by performing RNA-seq and ChIP-seq assays with lymphoblastoid cell lines. RNA-seq identified 357 genes that were regulated in a *TCL1A* SNP and E2-dependent fashion with expression patterns that were 4OH-TAM reversible. ChIP-seq for the same cells identified 57 TCL1A binding sites that could be regulated by E2 in a SNP-dependent fashion. Even more striking, NF-κB p65 bound to those same DNA regions. In summary, TCL1A is a novel transcription factor with expression that is regulated in a SNP and E2-dependent fashion—a pattern of expression that can be "reversed" by 4OH-TAM. Integrated RNA-seq and ChIP-seq results suggest that TCL1A also acts as a transcriptional co-regulator with NF-κB p65, an important immune system transcription factor.

Introduction

We previously performed a genome-wide association study (GWAS) that identified three SNPs located 3' of the T cell leukemia 1A(*TCL1A)* gene that were associated with aromatase inhibitor (AI)-induced musculoskeletal adverse events (Ingle et al., 2010). The top hit SNP (rs11849538) from that GWAS was in tight linkage disequilibrium (LD) with two other SNPs, rs7160302 and rs7359033 (R^2=0.92 and 0.98 respectively) (Ho et al., 2016a). Subsequently, we performed functional genomic studies using a "Human Variation Panel" of lymphoblastoid cell lines (LCLs) and found that TCL1A expression was induced by estradiol (E2), but only in cell lines homozygous for variant genotypes for the *TCL1A* SNPs (Ho et al., 2016a). These three SNPs appeared to act in concert to influence estrogen-dependent TCL1A induction (Ho et al., 2016a). However, this expression pattern could be "reversed" after estrogen receptor alpha (ERα) blockade with the active tamoxifen metabolite, 4-hydroxytamoxifen (4OH-TAM) We have also reported that mechanisms underlying this drug-induced reversal of TCL1A expression are due, at least in part, to altered ER binding to estrogen response elements (EREs), two of which are at a distance from the SNPs (Ho et al., 2016a). Even more striking, a series of downstream immune mediator genes including those encoding cytokines (Liu et al., 2012), chemokines (Ho et al., 2016a) and toll like receptors (TLRs) (Ho et al., 2017) responded in parallel with *TCL1A* SNP and estrogen-dependent transcription. These observations have potential implications for the treatment of inflammatory diseases such as rheumatoid arthritis, because these immune mediators are drug targets for the treatment of those diseases. If the *TCL1A* SNP genotypes were known, these observations raise the possibility that immune mediator gene expression could be manipulated by the use of drugs, e.g., 4OH-TAM.

Molecular mechanisms underlying the effect of the *TCL1A* SNPs that result in changes in TCL1A expression in an estrogen and 4OH-TAM-dependent fashion are not clear, and molecular mechanisms by which TCL1A expression might be associated with AI-induced musculoskeletal adverse events also remain unclear. The present study was designed to explore those molecular mechanisms.

Specifically, we set out to determine whether TCL1A might function as a transcription factor that acts broadly—genome wide—in a SNP and estrogen-dependent fashion, and to study underlying mechanisms responsible for this genomic phenomenon using a "Human Variation Panel" LCL model system. We should emphasize that this LCL panel represents a genomic data-rich cell line model system that has repeatedly demonstrated its utility in the generation and testing of pharmacogenomic hypotheses (Ingle et al., 2010; Liu et al., 2012; Ingle et al., 2013; Ho et al., 2016a; Ingle et al., 2016; Ho et al., 2017). Specifically, it made it possible for us to select LCLs for study with any common genotype or combinations of genotypes—as demonstrated by the experiments described subsequently.

Materials and methods

"Human Variation Panel" lymphoblastoid cell lines

The "Human Variation Panel" lymphoblastoid cell line (LCL) model system consists of 300 LCLs from healthy subjects of three ethnicities (100 European-American, 100 African-American

and 100 Han Chinese-American). This LCL cell line model system was utilized to perform many of the experiments described in this report. This cell line model system provides genome-wide mRNA expression as determined by Affymetrix U133 2.0 Plus GeneChip expression arrays. SNP data were generated by Illumina 550K and 510S SNP BeadChip SNP array (Illumina, San Diego, CA, USA), and genotyping data were then used to impute approximately seven million SNPs per cell line (Liu et al., 2012).

RNA Sequencing and data analysis

RNA was isolated from LCLs with known *TCL1A* SNP genotypes (two LCLs with homozygous wildtype and three LCLs with homozygous variant genotypes for the *TCL1A* SNPs). RNA-seq experiments were conducted by the Mayo Clinic Center for Individualized Medicine Medical Genomics Facility. RNA-seq libraries were prepared with the Ovation RNA-seq system v2 kit (NuGEN) according to the manufacturer's instruction, and were sequenced using an Illumina HiSeq 2000 with 6 samples in each lane using 100bp paired end index reads. Fastq files containing paired RNASeq reads were aligned with Tophat 2.0.12 (Kim et al., 2013) against the UCSC human reference genome (hg19) using Bowtie 2.2.3 with default settings (Langmead and Salzberg, 2012). Gene level counts from uniquely mapped, non-discordant read pairs were obtained using the subRead featureCounts program (v1.4.6) (Liao et al., 2013) and gene models from the UCSC hg19 Illumina iGenomes annotation package. Size factors were computed using the estimateSizeFactors function of DESeq2 and reference genes (GAPDH, ACTB) as the controlGenes parameter for estimateSizeFactors by the DESeq2 package (Love et al., 2014).

Differential expression analysis was also performed using the DESeq2 package with default parameters.

Identification of genes with TCL1A expression patterns

To identify genes with expression patterns similar to that of TCL1A after drug or hormone treatment, we mined our transcriptome data to identify genes that were up-regulated with log2FC >1 only in LCLs homozygous for wildtype sequences for the *TCL1A* SNPs in response to E2 treatment, and LCLs homozygous for variant genotypes for the *TCL1A* SNPs with non-significant changes or down-regulated (log2FC <0.5) expression in response to E2 treatment. We also searched for genes that had "reversed" gene expression patterns in response to E2+4OH-TAM treatment. Specifically, these were genes that were up-regulated with log2FC >1 only in LCLs with homozygous variant sequences for the *TCL1A* SNPs in response to E2+4OH-TAM treatment, and LCLs with homozygous wildtype genotypes for the *TCL1A* SNPs that displayed non-significant changes or down-regulation (log2FC <0.5) in response to E2+4OH-TAM treatment. Gene ontology (GO) terms for each dataset were generated using The Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.8 (https://david.ncifcrf.gov/) (Huang et al., 2008).

Chromatin immuneprecipitation sequencing (ChIP-seq) and data analysis

ChIP-seq experiments were performed in duplicate using two LCLs with known *TCL1A* genotypes. Specifically, we selected one LCL from each genotype group from the five LCLs we

used to perform RNA seq. ChIP-seq libraries were prepared using the Ovation ultralow DR Multiplex kit (NuGEN, San Carlos, CA) , and were subsequently sequenced to 51bp paired ends using an Illumina HiSeq 2000 at the Mayo Clinic Center for Individualized Medicine Medical Genomics Facility. Fastq files were aligned against the UCSC human reference genome (hg19) with Bowtie 1.1.0 using the following Bowtie parameters: --sam --chunkmbs 512 -p 4 -k 1 -m 1 -e 70 -l 51 –best. We then used the HOMER function findPeaks with default parameters to identify significantly enriched peaks in the experimental samples (Heinz et al., 2010). We used deepTools to visualize binding profiles for TCL1A around transcription start sites (-2kb,+2kb) (Ramírez et al., 2016). Briefly, sorted bam files were used as input for the deepTools bamCoverage function using the parameter --normalizeTo1x 2451960000 for normalization and generating BigWig files. We then used the function computeMatrix with binSize=20 to compute scores for genome regions and used its output to plot a heatmap of TCL1A binding profiles across vehicle and drug treatment conditions, using the deepTools function plotHeatmap. To identify regions with distinct binding intensities, we clustered binding profiles into four clusters using k-means. Next, we used the ChipSeeker package to annotate peaks identified by HOMER and study the distribution of read densities around annotated genomic regions. Motif analyses were performed using the HOMER function findMotifsGenome with the parameter: size 50 -mask, to identify transcription factor binding motifs. The accession number for the p65 ChIP-seq data is GSE55105 (Zhao et al., 2014). Overlap analysis for genome-wide TCL1A and p65 binding sites was performed using a window of +/-500bp from the center of each peak of TCL1A binding, and the binding signals of TCL1A and p65 were plotted side by side for each window, as shown in Figure 5B. In Figure 5C, we summed the ChIP-seq signals for TCL1A and p65 bindings, respectively, using a window +/-1000bp from the TSS from each of the 357 TCL1A

target genes that we studied. Correlations were then calculated between TCL1A and p65 binding for each treatment condition.

CRISPR/ Cas9 plasmids and knockout TCL1A gene

TCL1A Double Nickase Plasmid sc-402028-NIC (Santa Cruz Biotechnology, Dallas, Texas, USA) were transfected into LCLs using the program X-001 of Amaxa Nucleofector, and recovered for 24 hours in RPMI1640 medium supplemented with 15% FBS. Transfected cells were cultured in medium containing 0.2μg/ml puromycin for 10 days for selection. Cells were seeded on 96 well culture plate using RPMI1640 medium supplemented with 15% FBS for 12 weeks. TCL1A expression was quantified and normalized to the β-actin signal by Western blot analysis. Cells exhibiting <10% normalized TCL1A signal were considered to have TCL1A knocked out.

Drug treatment

LCLs with known genotypes were cultured in RPMI 1640 media (Cellgro, Manassas, VA, USA) supplemented with 15% FBS FBS (Atlanta Biologicals, Flowery Branch, GA, USA). Before estrogen treatment, cells were cultured in RPMI media containing 5% (V/V) charcoal stripped FBS for 24 hours, and were subsequently cultured in FBS free RPMI media for another 24 hours. Cells were then treated with 0.1nM of estradiol (E2) for 24 hours. In some experiments, cells were treated with 4OH-TAM (10^{-7}μM).

Real time PCR

Four cell lines with each *TCL1A* genotype were selected for functional validation. The PCR mixture contained 100ng of total RNA, 5µl of 2X VeriQuest SYBR green qPCR master mix (Affymetrix, Santa Clara, CA USA), 0.1µl of DNA polymerase, 1µl of gene specific primer and distilled water up to 10µl per reaction. GAPDH and ACTB were utilized as internal controls. Real time PCR reactions were performed in duplicate using the Applied Biosystems ViiA 7™ Real-Time PCR System (Life Technologies, Carlsbad, CA, USA). The $2^{-\Delta\Delta Ct}$ method was employed for statistical data analysis.

RNA interference and transfection

Smart pooled TCL1A, p65 siRNA and negative control were purchased from Dharmacon (Chicago, IL, USA). siRNA was transfected into the LCLs by electroporation using the nucleofector kit (Lonza, Köln, Germany). Briefly, the electroporation reaction contained $2.5x10^6$ cells, 100µl of nucleofector solution and 300nM of siRNA. After electroporation, cells were transferred into 12-well plates containing pre-equilibrated RPMI medium for use in the experiments described in this report.

Chromatin immunoprecipitation (ChIP) assay and ChIP-re-ChIP assay

ChIP assays were performed using LCLs with known *TCL1A* SNP genotypes (n=4 for each genotype) using the EpiTect ChIP OneDay Kit (Qiagen, Valencia, CA, USA). DNA-TCL1A complexes were immunoprecipitated using antibodies against TCL1A, RelA or with normal

mouse IgG (Cell Signaling Technology, Danvers, MA, USA) as a control. ChIP-re-ChIP assays were performed using the Re-ChIP-IT® Kit (Active Motif, Carlsbad, CA, USA). Real time PCR was used to quantify TCL1A binding or p65 binding. Primer sets for the ChIP assays are listed in **Supplementary Table 4**.

Co-Immunoprecipitation (Co-IP) of TCL1A and NF-κB

LCLs ($1x10^7$) were resuspended in 650µl IP lysis buffer containing 2.2µl protease inhibitor cocktail (Qiagen, Valencia, CA, USA) and were incubated on ice for 30 min. Cells were then centrifuged at 12,000 g at 4°C for 15 min. Supernatants were collected. Protein A agarose (ThermoScientific, Madison, WI, USA) was prepared and washed 3 times with IP lysis buffer. A pre-cleaning step was performed in order to clean the background. Cell lysates containing protein A agarose beads were rotated at 4°C for an hour. After centrifugation, supernatant was collected. At this point, input (50µl) was collected and stored at -80°C. Anti-TCL1A (1:50) (Cell Signaling Technology, Danvers, MA, USA) or Anti-NF-κB p65 (1:50) antibodies (Abcam, MA, USA) were used to perform IP. IgG (Cell Signaling Technology, Danvers, MA, USA), was used as negative control. Specifically, IP Samples containing protein A agarose beads were rotated at 4°C overnight. Immunoprecipitates were washed three times with ice cold lysis buffer, and proteins were eluted with 50µl 1X Laemmli loading buffer. Proteins were separated on 4-12% SDS-PAGE gels and transferred onto PDVF membranes. After blocking, membranes were incubated with primary antibodies against TCL1A or NF-κB p65 at 4°C overnight. The washed membrane was then incubated with secondary antibody (1:15000 dilution) for an hour at room

temperature. The membrane was visualized using super signal ECL substrate (Thermo Scientific, Madison, WI, USA).

Electrophoretic mobility shift assay (EMSA)

Recombinant Human TCL1A Protein (NBP1-30239) was purchased from Novus Biologicals, Littleton, CO, USA. Synthetic oligonucleotides (sense and anti-sense) were derived from TCL1A response element sequences. Double-stranded DNA probes were 3′end-labeled with biotin (IDT, Coralville, Iowa, USA). A 100-fold excess of unlabeled oligonucleotide competitors was used for the competition experiments. The reaction mixture was then loaded onto a 5% TBE native polyacrylamide gel and electrophoresis was run for 1 hour at 100 V in 0.5 × TBE buffer. The protein-DNA complexes were transferred to Biodyne B Nylon Membranes (Thermo scientific, Madison, WI, USA) and UV cross-linked. EMSA was performed using LightShift® Chemiluminescent EMSA Kit (Thermo scientific, Madison, WI, USA).

Immunofluorescence staining and confocal imaging analysis

LCLs were grown on glass coverslips and treated with E2 0.1nM for 24 hours. Cells were then fixed in 4% paraformaldehyde at room temperature for 10 min. Cells were washed in cold PBS and permeablized with 0.2% Triton X-100. After blocking for one hour with 3% BSA, cells were incubated with mouse anti-TCL1A antibody (Cell Signaling Technology, Danvers, MA, USA) overnight at 4°C. The secondary antibody (red) was ab150119 Alexa Fluor® 647 goat anti-mouse IgG (H+L) used at 1:1000 dilution for an hour. DAPI was used to stain the cell nuclei

(blue) at a concentration of 1.43μM. Slides were visualized using fluorescence microscopy (Olympus, FV1200).

Results

RNA-seq genome-wide identification of *TCL1A* SNP and estrogen-dependent transcription regulation

As a first step, RNA-seq was performed using LCLs selected for either homozygous wildtype or homozygous variant genotypes for the *TCL1A* SNPs shown graphically in **Figure 1A.** The cells were treated with vehicle, E2, or E2 plus 4OH-TAM. The drug concentrations used had been optimized as shown in **Supplementary Figure 1** (Liu et al., 2012; Ho et al., 2016a; Ho et al., 2017). We then randomly selected 39 genes for use in the validation of the RNA-seq data and found very high reproducibility based on the use of qPCR (**Supplementary Figure 2**). For example, in the presence of E2, TCL1A expression was induced in LCLs that were homozygous variant but not in those with homozygous wildtype genotypes for the *TCL1A* SNPs (**Figure 1B**). As expected, this expression pattern could be reversed in the presence of 4OH-TAM. Specifically, TCL1A expression was significantly up-regulated in cells that were homozygous wildtype for the *TCL1A* SNPs but not in those with homozygous variant genotypes (**Figure 1B**). Based on the RNA-seq data, 357 genes displayed this *TCL1A* SNP and estrogen-dependent gene expression pattern—a pattern that could be reversed by 4OH-TAM (**Figure 1C**), in parallel with the pattern of expression of TCL1A itself in response to treatment with either E2 or E2+4OH-TAM (**Figure 1B**). Specifically, a total of 357 genes could be significantly induced by E2 with

log2 fold change ≥1 **(Supplementary Table 1),** but only in LCLs homozygous for *TCL1A* SNP variant genotypes (**Figure 1C**). However, the expression of those same 357 genes was significantly induced in the presence of 4OH-TAM together with E2 in LCLs that were homozygous for wildtype genotypes for the *TCL1A* SNPs (**Figure 1C**). We also performed gene ontology (GO) analysis and found that these 357 genes clustered within pathways involved in transcriptional regulation and T cell activation (**Supplementary Table 2)** by using The Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.8 (https://david.ncifcrf.gov/) (Huang et al., 2008). These results provided additional information with regard to the role of *TCL1A* in the transcriptional regulation of immune mediators (Liu et al., 2012; Ho et al., 2016a; Ho et al., 2016b). In the next series of experiments, we set out to directly test by ChIP-seq the possibility that *TCL1A* was a novel transcription factor that acts in a SNP and estrogen-dependent fashion.

ChIP-seq genome-wide identification of TCL1A binding sites

We next examined TCL1A's binding profile on a genome-wide scale by performing ChIP-seq using the same set of LCLs that had been used in the RNA-seq studies. We clustered Chip-seq peak signals within ±2kb from the transcription start site (TSS) into four clusters using K-means. We observed that genes in cluster 1 were more strongly bound by TCL1A than genes in clusters 2 and 3, with higher enrichments in variant genotypes after E2 treatment (**Figure 2A)**. Therefore, TCL1A was found to bind preferentially to promoter regions. In addition, TCL1A also bound significantly in introns (~30% of sites) and intergenic regions (~23%) (**Figure 2B**). We also performed ChIP-qPCR experiments for 25 randomly selected sites using independent samples to

confirm the specificity of these observations (**Supplementary Figure 3**). The most significantly associated gene ontology terms for these genes involved transcription regulation and immune function related pathways (**Supplementary Table 3**). This result confirmed our previous findings with regard to the effect of TCL1A on the expression of cytokines, chemokines and their receptors (**Supplementary Figure 1**), all of which play an important role in the regulation of immune responses. Our ChIP-seq results showed that TCL1A is capable of binding to specific DNA sequences across the human genome, resulting in the regulation of the transcription of target genes.

TCL1A response elements

To identify in an unbiased fashion sequence motifs at sites occupied by TCL1A, we examined DNA sequences (50bp ± from the peak) using HOMER (http://homer.salk.edu/homer/ngs/) (Heinz et al., 2010) and MEME (Bailey et al., 2009). *De novo* motif search in each set of TCL1A binding sites from these samples revealed enrichment of the consensus sequence 5'-CCATATATGG-3', which we designated as a TCL1A consensus DNA binding motif (**Figure 2C**). We then performed EMSA using biotin labeled TCL1A binding motif sequences and recombinant TCL1A protein to confirm that TCL1A proteins could recognize and bind to this DNA motif sequence (**Figure 3A**). The next series of studies was designed to characterize TCL1A as a novel transcription factor, with effects that occur with and without 4OH-TAM exposure.

Estrogen significantly alters TCL1A occupancy of binding sites in or near target genes in a SNP-dependent fashion

In the absence of E2, TCL1A is mainly expressed in the cytoplasm and, to a lesser degree, in the cell membrane as determined by immunofluorescence staining of LCLs (**Figure 3B)**. However, TCL1A migrates and accumulates in the nucleus in the presence of E2 (**Figure 3B**). This observation is important because it supports the ability of TCL1A to translocate into the nucleus where it can function as a transcription factor, with effects that can be altered by estrogens and by the SNPs located 3' of *TCL1A*. For example, ChIP-seq showed that TCL1A was capable of binding to sites in the *CCR6* and *IL17RA* genes, both of which can be regulated by TCL1A in a SNP and estrogen-dependent fashion (Ho et al., 2016a) (**Figure 3C**). Even more striking, the binding density depended on *TCL1A* SNP genotype and E2 treatment (**Figure 3C**). Specifically, TCL1A binding density in peaks in both *CCR6* and *IL17RA* was significantly higher in cells with homozygous variant genotypes for the *TCL1A* SNPs as compared to results for cells with homozygous wildtype genotypes. These results parallel those found during the RNA-seq studies (**Figure 3D**). It should be emphasized that the SNPs were located 3'of the *TCL1A* gene, but not in *CCR6* or *IL17RA*. These observations led us to test whether the SNP and estrogen-dependent expression of TCL1A might be associated with the transcription of TCL1A target genes. Therefore, the next series of studies was designed to integrate the ChIP-seq and RNA-seq data and to explore the possible role of TCL1A in transcriptional regulation in a SNP and estrogen-dependent manner.

Integrated ChIP-seq and RNA-seq analyses reveal functional TCL1A binding sites associated with transcriptional regulation in a SNP-and estrogen-dependent fashion

To explore the consequences of SNP-dependent TCL1A binding in or near target genes in response to E2 treatment, we merged RNA-seq and ChIP-seq data by overlapping the results for genes that displayed *TCL1A* SNP-dependent gene expression and SNP-dependent TCL1A binding in response to E2 treatment. A total of 357 genes could be induced by E2 in a *TCL1A* SNP-dependent fashion, and 57 of those genes also displayed binding of TCL1A with significantly increased binding density in response to E2 treatment but only in the cells with homozygous variant genotypes for the *TCL1A* SNP genotypes (**Figure 4A**). We confirmed this observation by performing ChIP-qPCR with primers specifically designed to amplify genomic regions for TCL1A binding sites in these genes. Specifically, for all 57 TCL1A binding regions, greater TCL1A binding was observed in the presence of E2 in cells with homozygous variant genotypes for the *TCL1A* SNPs as compared to those with wildtype genotypes (**Figure 4B**). Furthermore, in the presence of 4OH-TAM, this TCL1A binding pattern was reversed (**Figure 4C**). These results were compatible with the gene expression patterns for all 57 of these genes in response to E2 treatment (**Figure 4D**), all of which also displayed a "reversal" of expression in the presence of 4OH-TAM (**Figure 4E**) using the same cell lines from which the data shown in Figures 4B and 4C were obtained. It should be pointed out once again that the SNPs were located 3' of *TCL1A* and not in the 57 genes themselves. In summary, this series of experiments demonstrated clearly that TCL1A served as a transcription factor for these target genes in a SNP and estrogen-dependent fashion.

TCL1A and NF-κB p65 co-occupy binding sites on TCL1A-responsive target genes

We also interrogated the sequences of the TCL1A binding sites (peak+/-50bp) for overrepresentation of known DNA binding motifs using HOMER. An NF-kB p65 binding motif was highly enriched in these genomic regions (with p values =1E-12 and 1E-38 for cells with homozygous wildtype or homozygous variant genotypes for the *TCL1A* SNPs, respectively). These observations led us to determine whether TCL1A might physically interact with NF-κB p65 and, as a result, might serve as a transcriptional co-regulator with NF-κB p65. Possible protein-protein interaction between TCL1A and NF-κB p65 was determined by co-immunoprecipitation (IP) using LCLs with endogenous expression of both TCL1A and NF-κB p65 (**Figure 5A**). Reverse IP verified the physical association between NF-κB p65 and TCL1A (**Figure 5A**). We next tested whether both TCL1A and NF-κB p65 might target the same genes and/or regulatory elements. To do that, we compared our TCL1A ChIP-seq dataset and publicly available p65 ChIP-seq dataset generated for LCLs (Zhao et al., 2014). There was a high degree of overlap between the locations of peaks for the two transcription factors (**Figure 5B**). It should be pointed out that the heat map for p65 binding sites shown in **Figure 5B** used a window +/-500bp centered on TCL1A peaks genome wide, which indicated that p65 bound to genome regions similar to those to which TCL1A bound. This overlap suggested a possible functional relationship between TCL1A and p65 in the regulation of gene expression, including genes involved in immune response. This overlap was particularly striking for the 357 genes shown in **Figure 1C**, genes that could be regulated by *TCL1A* in a SNPs and estrogen-dependent manner (**Figure 5C**). Therefore, we next attempted to determine whether TCL1A and the NF-κB p65 subunit might interact within the same transcriptional complex by performing ChIP-re-ChIP assay

Those experiments showed that TCL1A and NF-κB p65 co-occupied the same sites on TCL1A target genes, as shown in **Figure 6A-6B** based on the results of ChIP-re-ChIP assays performed using antibody against TCL1A, followed by antibody against NF-κB p65. Even more striking, the binding pattern for NF-κB p65 displayed a *TCL1A* SNP and estrogen-dependent pattern. Specifically, a significant increase in NF-κB p65 binding was observed in cells with homozygous variant genotypes for the *TCL1A* SNPs in response to E2 treatment (**Figure 6A**). However, this binding pattern was reversed when the cells were exposed to 4OH-TAM (**Figure 6B**), in parallel with the binding pattern for TCL1A that we had observed (**Figure 4B-4C**). It should be emphasized once again that the SNPs being studied were located 3' of *TCL1A* not in the NF-κB p65 gene or in the genes listed in **Figures 6A-6B**.

TCL1A and p65 expression displayed a positive correlation in LCLs (**Figure 7A).** Specifically, TCL1A knockout in LCLs using CRISPR-Cas9 resulted in the down-regulation of p65. NF-κB p65 knockout resulted in cell death, therefore, we use siRNA transient knockdown of this gene in our experiments. Similarly, knockdown of p65 in LCLs using siRNA caused down-regulation of TCL1A (**Figure 7A**). Furthermore, *TCL1A* SNP and estrogen-dependent binding to DNA often appeared to involve both TCL1A and NF-κB p65. For example, the SNP-dependent binding pattern for TCL1A target genes was lost after the knockdown of p65 in LCLs with known *TCL1A* SNP genotypes (**Figure 7B**). In parallel, the SNP and estrogen-dependent p65 binding for TCL1A target genes was also abolished after TCL1A was knocked out (**Figure 7C).** Furthermore, The *TCL1A* SNP and estrogen-dependent gene expression pattern was also lost after p65 was knocked down (**Figure 7D)** or TCL1A was knocked out (**Figure 7E).**

Taken together, these data suggest that a TCL1A and NF-κB p65 protein complex might contribute to *TCL1A* SNP and estrogen-dependent transcription regulation—at least for some

genes. We also observed that *TCL1A* SNP and estrogen-dependent p65 binding was lost after TCL1A knockout (data not shown). These data suggested that TCL1A and NF-κB p65 either co-occupy the same sites or that their binding sites are located very close together on TCL1A target genes, resulting in the SNP and estrogen-dependent variation in gene expression that we observed. Finally, TCL1A also appeared to interact with all NF-kB subunits except for p52 (**Figure 8**). However, relB, c-rel and p50 were not able to bind the TCL1A target genes in areas of TCL1A binding (**Supplementary Figure 4**). Taken as a whole, these results supported the conclusion that NF-κB p65 may play an important role in *TCL1A* SNP and estrogen-dependent transcriptional regulation.

Discussion

We have studied mechanisms responsible for *TCL1A* SNP and estrogen-dependent regulation of gene expression using a cell-line model system. TCL1A is expressed in immune cells including activated T lymphocytes, B lymphocytes and thymocytes (Kang et al., 2005). Not only is the LCL model system well suited for these experiments because of the dense genomic data that we have generated for these cells, but also the Genotype-Tissue Expression (GTEx) Project database https://www.gtexportal.org/home/ reports that, of all human tissues and cells included in that database, the highest expression of TCL1A is observed in LCLs (see **Supplementary Figure 5**). In the present study, we have demonstrated clearly that TCL1A is a novel transcription factor that acts in a SNP and estrogen-dependent fashion. The three SNPs involved are located 3' of the *TCL1A* gene (**Figure 1A**) and have been associated with AI- induced musculoskeletal adverse

events (Ingle et al., 2010). The present study also demonstrated that TCL1A appears to serve as a transcriptional co-regulator with the NF-kB p65 subunit. Therefore, the present study represents an important step in the process of providing functional and mechanistic explanations for the association of *TCL1A* SNPs with inflammation and the immune response.

SNPs that map to the coding regions of genes or within DNA transcription factor binding motifs are well documented to have functional effects on the regulation of gene expression (Barrett et al., 2012). However, the present series of studies began with a GWAS that identified SNPs located 3' of *TCL1A* on chromosome 14 that were associated with AI-induced musculoskeletal pain (Ingle et al., 2010). The minor allele frequency for these SNPs is approximately 19% in all major human populations (European, African, Asian and Caucasian-Americans) based on 1000 Genomes Project data (The 1000 Genomes Project Consortium, 2015). We subsequently performed a series of functional genomic studies which showed that TCL1A expression was up-regulated by E2 only in cells carrying variant sequences for these SNPs. This SNP-dependent induction depended on functional EREs located—in part—at a distance from the SNPs (**Figure 1A**) (Ho et al., 2016a). Furthermore, TCL1A induction was associated with variation in the expression of a series of immune mediators including cytokines, chemokines and TLRs (Liu et al., 2012; Ho et al., 2016a; Ho et al., 2017). Finally, this SNP-estrogen-dependent pattern of gene expression could be "**reversed**" by 4-OH-TAM (Liu et al., 2012; Ho et al., 2016a; Ho et al., 2017). The present study in which we have performed RNA-seq and ChIP-seq using LCLs with known *TCL1A* SNP genotypes has greatly extended these previous observations (**Figure 1C**). It should be emphasized that the SNPs involved in all of these studies are those 3' of *TCL1A*. These SNPs were **not** in or near genes that were regulated by TCL1A. It should also be emphasized that

basal expression levels were not significantly different for these inflammatory mediators from our previous reports or, as shown in **Figure 1C,** between the two different genotype patterns for the *TCL1A* SNPs (homozygous wildtype vs homozygous variant). These results serve to highlight the fact that genetically polymorphic variation in the expression of a transcription factor, in this case TCL1A, can have a profound influence on the transcriptional regulation of downstream genes, in this case in a SNP and estrogen dependent fashion (**Figure 1C**).

The present study also identified a consensus sequence motif for a "TCL1A response element" that appeared to be a ten base-pair palindromic inverted sequence (**Figure 2C**). This TCL1A response element consensus sequence could be found near the center of a majority of the strong TCL1A bindings peaks that we observed. However, these regions might not be the only determinant of TCL1A binding to DNA. We also observed that high TCL1A occupancy at many of its binding sites was associated with the occurrence of NF-κB p65 binding (**Figure 5B**). NF-κB plays a pivotal role in inflammation (Makarov, 2001; Lawrence, 2009). Inflammatory mediators include cytokines, chemokines and TLRs, all of which are thought to be crucial for immune function, and they can also have a profound effect on inflammation as a result of NF-κB activation (Tak and Firestein, 2001). As outlined above, TCL1A can modulate the expression of these immune mediators in a SNP and estrogen-dependent manner (Liu et al., 2012; Ho et al., 2016a; Ho et al., 2017). As a result, the present study has not only greatly expanded our previous observations, but also offers a novel mechanism by which NF-κB p65 might act as a co-regulator and contribute to *TCL1A* SNP and estrogen-dependent transcription regulation, all of which provide possible pharmacogenomic explanations for the association of *TCL1A* SNPs with musculoskeletal adverse events induced by AI therapy. Therefore, future studies will be needed

to test the possibility that, by silencing the *TCL1A* gene, TCL1A could be a drug target for the management of musculoskeletal symptoms induced by AI therapy.

Furthermore, observations made in the series of TCL1A functional genomic studies also provide insight into the expression of immune mediators that could potentially be pharmacologically manipulated in a SNP-dependent fashion, and might have potential implications for the treatment of rheumatologic disease. For example, rheumatoid arthritis (RA) displays a strong sex bias toward women and reveals profound variation in incidence and severity during periods of change in estrogen levels. The highest risk for RA is observed during the menopausal years (Goemaere et al., 1990), and the incidence of RA and the risk of "flares" are increased during the postpartum period, a time of rapidly decreased plasma estrogen levels (Peschken et al., 2012). These observations display interesting parallels to the clinical impact of the pharmaceutical lowering of estrogen levels by drugs such as the AIs. *TCL1A* SNP-dependent transcriptional regulation of immune mediators also provides insight into pharmacogenomic aspects of the regulation of the expression of inflammatory mediators, i.e, IL17A and IL17RA, both of which are deeply involved in RA pathophysiology and are therapeutic targets for RA. If these observations can ultimately be translated into the clinic, they might represent a novel mechanism by which drugs could be used to regulate the estrogen-dependent induction of inflammatory mediator expression.

The present study utilized a genomic data rich cell-line model system to uncover a relationship between TCL1A and p65 DNA binding and variation in the regulation of gene expression. Obviously, transcription factor binding is regulated in a cell type-specific manner. However, the

present study represents a critical step in the process of providing a functional and mechanistic explanation for *TCL1A* SNP and estrogen-dependent gene expression regulation. This mechanism might contribute broadly to individual variation in transcriptional regulation, in immune response and in variation in drug response—the phenotype that resulted in the discovery of this potentially important process in transcriptional regulation.

In conclusion, in the this study, we have greatly extended previous observations with regard to *TCL1A* SNP and estrogen-dependent gene expression genome-wide by performing RNA-seq and ChIP-seq using LCLs with known *TCL1A* SNP genotypes. We have demonstrated clearly that TCL1A is a novel transcription factor that acts in a SNP and estrogen-dependent fashion. The integrated RNA-seq and ChIP-seq data reported here also suggest that TCL1A may be a co-regulator of expression acting together with the NF-κB p65 subunit, a transcription factor that plays a key role in inflammation. As a result, the present study has greatly extended our understanding of the role of TCL1A in transcriptional regulation and has highlighted a novel mechanism for *TCL1A* SNP and estrogen regulation of transcription—as a co-regulator with NF-kB p65. Obviously, future studies will be required to pursue the role of TCL1A in individual variation in immunity and inflammation as well as its possible role in the pathophysiology of human disease.

Authorship Contributions

Participated in research design: Ho, Lummertz da Rocha, Zheng, Li, Ingle, Goss, Shepherd, Kubo, Wang, and Weinshilboum.

Conducted experiments: Ho

Performed data analysis and interpretation: Ho, Lummertz da Rocha, Zheng, Li, Ingle, Wang and Weinshilboum.

Wrote or contributed to the writing of the manuscript: Ho, Lummertz da Rocha, Zhang, Li, Ingle, Goss, Shepherd, Kubo, Wang, and Weinshilboum.

All authors have given final approval of the version to be published.

References

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW and Noble WS (2009) MEME Suite: tools for motif discovery and searching. *Nucleic Acids Research* **37**:W202-W208.

Barrett LW, Fletcher S and Wilton SD (2012) Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements. *Cellular and Molecular Life Sciences* **69**:3613-3634.

Goemaere S, Ackerman C, Goethals K, De Keyser F, Van der Straeten C, Verbruggen G, Mielants H and Veys EM (1990) Onset of symptoms of rheumatoid arthritis in relation to age, sex and menopausal transition. *J Rheumatol* **17**:1620-1622.

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H and Glass CK (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular cell* **38**:576-589.

Ho M-F, Bongartz T, Liu M, Kalari KR, Goss PE, Shepherd LE, Goetz MP, Kubo M, Ingle JN, Wang L and Weinshilboum RM (2016a) Estrogen, SNP-Dependent Chemokine Expression and Selective Estrogen Receptor Modulator Regulation. *Molecular Endocrinology* **30**:382-398.

Ho M-F, Ingle J, Goss P, Mushiroda T, Kubo M, Shepherd L, Wang L and Weinshilboum R (2016b) Aromatase Inhibitor-induced Musculoskeletal Symptoms and TCL1A SNP-Mediated, TLR-MyD88-Dependent NF-κB Activation: Molecular Mechanisms Involving a Crucial Adaptor Protein MyD88 *Clinical Pharmacology & Therapeutics* **99 (suppl 1)**:S5.

Ho M-F, Ingle JN, Bongartz T, R.Kalari K, E.Goss P, E.Shepherd L, T.Mushiroda, M.Kubo, Liewei.Wang and M.Weinshilboum R (2017) TCL1A SNPs and estrogen-mediated toll-like receptor-MYD88-dependent NF-kappaB activation: SNP and SERM-dependent modification of inflammation and immune response. *Molecular Pharmacology* **92**:175-184.

Huang DW, Sherman BT and Lempicki RA (2008) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protocols* **4**:44-57.

Ingle JN, Liu M, Wickerham DL, Schaid DJ, Wang L, Mushiroda T, Kubo M, Costantino JP, Vogel VG, Paik S, Goetz MP, Ames MM, Jenkins GD, Batzler A, Carlson EE, Flockhart DA, Wolmark N, Nakamura Y and Weinshilboum RM (2013) Selective Estrogen Receptor Modulators and Pharmacogenomic Variation in ZNF423 Regulation of BRCA1 Expression: Individualized Breast Cancer Prevention. *Cancer Discovery* **3**:812-825.

Ingle JN, Schaid DJ, Goss PE, Liu M, Mushiroda T, Chapman J-AW, Kubo M, Jenkins GD, Batzler A, Shepherd L, Pater J, Wang L, Ellis MJ, Stearns V, Rohrer DC, Goetz MP, Pritchard KI, Flockhart DA, Nakamura Y and Weinshilboum RM (2010) Genome-Wide Associations and Functional Genomic Studies of Musculoskeletal Adverse Events in Women Receiving Aromatase Inhibitors. *Journal of Clinical Oncology* **28**:4674-4682.

Ingle JN, Xie F, Ellis MJ, Goss PE, Shepherd LE, Chapman J-AW, Chen BE, Kubo M, Furukawa Y, Momozawa Y, Stearns V, Pritchard KI, Barman P, Carlson EE, Goetz MP, Weinshilboum RM, Kalari KR and Wang L (2016) Genetic polymorphisms in the long noncoding RNA MIR2052HG offer a pharmacogenomic basis for the response of breast cancer patients to aromatase inhibitor therapy. *Cancer Research* **76**:7012–7023.

Kang S-M, Narducci MG, Lazzeri C, Mongiovì AM, Caprini E, Bresin A, Martelli F, Rothstein J, Croce CM, Cooper MD and Russo G (2005) Impaired T- and B-cell development in Tcl1-deficient mice. *Blood* **105**:1288-1294.

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R and Salzberg SL (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology* **14**:R36.

Langmead B and Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Meth* **9**:357-359.

Lawrence T (2009) The Nuclear Factor NF-κB Pathway in Inflammation. *Cold Spring Harbor Perspectives in Biology* **1**:a001651.

Liao Y, Smyth GK and Shi W (2013) The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Research* **41**:e108-e108.

Liu M, Wang L, Bongartz T, Hawse J, Markovic S, Schaid D, Mushiroda T, Kubo M, Nakamura Y, Kamatani N, Goss P, Ingle JN and Weinshilboum RM (2012) Aromatase inhibitors, estrogens and musculoskeletal pain: estrogen-dependent T-cell leukemia 1A (TCL1A) gene-mediated regulation of cytokine expression. *Breast Cancer Research* **14**:R41.

Love MI, Huber W and Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**:550.

Makarov S (2001) NF-kappaB in rheumatoid arthritis: a pivotal regulator of inflammation, hyperplasia, and tissue destruction. *Arthritis Res* **3**:200-206.

Peschken CA, Robinson DB, Hitchon CA, Irene S, Hart D, Bernstein CN and EL-Gabalawy HS (2012) Pregnancy and the Risk of Rheumatoid Arthritis in a Highly Predisposed North American Native Population. *The Journal of Rheumatology* **39**:2253-2260.

Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F and Manke T (2016) deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Research* **44**:W160-W165.

Tak PP and Firestein GS (2001) NF-kappaB: a key role in inflammatory diseases. *J Clin Invest* **107**:7-11.

The 1000 Genomes Project Consortium (2015) A global reference for human genetic variation. *Nature* **526**:68-74.

Zhao B, Barrera Luis A, Ersing I, Willox B, Schmidt Stefanie CS, Greenfeld H, Zhou H, Mollo Sarah B, Shi Tommy T, Takasaki K, Jiang S, Cahir-McFarland E, Kellis M, Bulyk Martha L, Kieff E and Gewurz

Benjamin E (2014) The NF-κB Genomic Landscape in Lymphoblastoid B Cells. *Cell Reports* **8**:1595-1606.

Footnotes

This work was supported in part by National Institute of General Medical Sciences [U19 GM61388], National Institutes of Health [P50 CA11620, RO1 CA196648, RO1 GM28157, and RO1 CA138461], the Avon Breast Cancer Crusade and the Breast Cancer Research Funcdation, as well as by the Mayo Clinic Center for Individualized Medicine. Drs. Weinshilboum and Wang are co-founders and stockholders in OneOme.

Figure Legends

Figure 1. (A) Schematic diagram of three *TCL1A* SNPs: rs11849538, the "top hit" signal from the MA.27 musculoskeletal adverse event GWAS, rs7359033 and rs7160302. All three of these SNPs map near the 3'-terminus of *TCL1A* and are in tight LD. Locations of estrogen response elements (EREs) are shown as boxes. (B) SNP and estrogen-related variation of TCL1A mRNA expression in lymphoblastoid cell lines with known *TCL1A* SNP genotypes after exposure to E2 with or without 4-OH-TAM. A Student's *t* test was performed to compare gene expression in LCLs with differing *TCL1A* SNP genotypes for each treatment condition, ***= $p<0.001$. (C) Heat map showing expression profiles for 357 genes regulated by *TCL1A* SNPs in an estrogen-dependent fashion, all of which could be reversed by 4OH-TAM treatment, as determined by RNA-seq using LCLs with either homozygous wildtype (W/W) (n=2) or homozygous variant (V/V) (n=3) genotype for the *TCL1A* SNPs.

Figure 2. (A) Genome-wide TCL1A occupancy profiles. Regions (Tss +/-2kb) were clustered based on their profiles for TCL1A ChIP enrichment over input for all four samples using k-means clustering. Clustering was used to identify regions with distinct binding intensities. The gradient blue-to-red color indicates high-to-low counts in the corresponding genome region (B) Percent distribution of TCL1A ChIP-seq peaks for all four samples. (C) Identification of TCL1A consensus binding sequence motifs using HOMER and MEME.

Figure 3. (A) EMSA results showing that one TCL1A DNA binding motif sequence, CCATATAGG, is sufficient for DNA–TCL1A protein interaction. (B) Immunofluorescence

staining showing TCL1A nuclear translocation in response to E2 (0.1nM) treatment in LCLs. (C) Representive examples of TCL1A occupancy peaks are depicted for the *CCR6* and *IL17RA* loci in LCLs with known *TCL1A* SNP genotypes in the absence of E2 or in the presence of E2 (0.1nM). Red boxes indicate that TCL1A binding in both *CCR6* and *IL17RA* is significantly increased in response to E2 treatment, but that only occurs in cells with homozygous variant genotypes for *TCL1A* SNPs. (D) Changes in TCL1A occupancy are highly correlated with changes in mRNA expression levels for CCR6 and IL17RA. A Student's *t* test was performed to compare gene expression in LCLs with differing *TCL1A* SNPs (homozygous wildtype *.vs.* homozygous variant) for each treatment condition, ***= $p<0.001$.

Figure 4. (A) Venn diagram showing 357 gene that displayed *TCL1A* SNP and estrogen-dependent gene expression patterns as determined by RNA-seq, and 57 of those 357 genes that displayed *TCL1A* SNP and estrogen-dependent TCL1A occupancy. (B-C) TCL1A ChIP assays were performed to confirm results obtained from TCL1A ChIP-seq for all 57 genes, all of which showed greater TCL1A binding in the presence of E2, but only in cells homozygous variant for the *TCL1A* SNP genotypes. In contrast, in the presence of 4OH-TAM, this binding pattern was reversed for all 57 genes (n=4 for each genotype group). (D-E) Changes in TCL1A binding were correlated with changes in mRNA expression for all 57 genes using the same cell lines from which the data shown in Figure 4B and 4C were obtained. . Specifically, in the presence of E2, all of those genes showed significant induction only for the variant genotype. However, the expression pattern could be reversed by 4OH-TAM, thus confirming that changes in TCL1A occupancy were highly correlated with transcription.

Figure 5. (A) Co-Immunoprecipitation (IP) was used to determine whether TCL1A protein could interact with p65 in LCLs. Whole cell lysates from $1x10^7$ LCLs were immunoprecipitated with (left panel) anti-TCL1A (1:50) antibodies or anti-IgG antibodies and protein samples were immunoblotted and probed with antibodies against TCL1A. Reversed IP was preformed to confirm that p65 and TCL1A interacted (right panel). (B) Heatmap plots showing the association between TCL1A binding and p65 binding in LCLs (Zhao et al., 2014). The signals for TCL1A ChIP-seq peaks and p65 ChIP-seq peaks are shown as heatmaps using red (the strongest signal) and white (the weakest signal) color schemes. Each row shows +/-500bp centered on the TCL1A ChIP-seq peak summits. (C) TCL1A occupancy vs p65 occupancy in LCLs. Scatter plots depicting the Pearson correlations (r) between TCL1A ChIP tag density (Y axis) and p65 ChIP tag density (X axis) for the 357 TCL1A target genes are shown in **Figure 1C**.

Figure 6. (A-B) p65 ChIP-re-ChIP assays were performed to confirm the co-occupancy of TCL1A and p65 on 57 selected binding regions as shown in Figure 4B-C (n=4 for each genotype group).

Figure 7. (A) Western blot analysis after TCL1A knockout or p65 knockdown in LCLs. (B) Knockdown of p65 resulted in the abolition of SNP and estrogen-dependent TCL1A occupancy. (C) Knockdown of TCL1A abolished the *TCL1A* SNP and estrogen-dependent p65 binding. (D-E), *TCL1A* SNP and estrogen-dependent gene expression was lost after the knockdown of p65 or the knockout of TCL1A.

Figure 8. Co-Immunoprecipitation was used to determine whether TCL1A protein could interact with NF-kB subunits in LCLs.

Figure 1

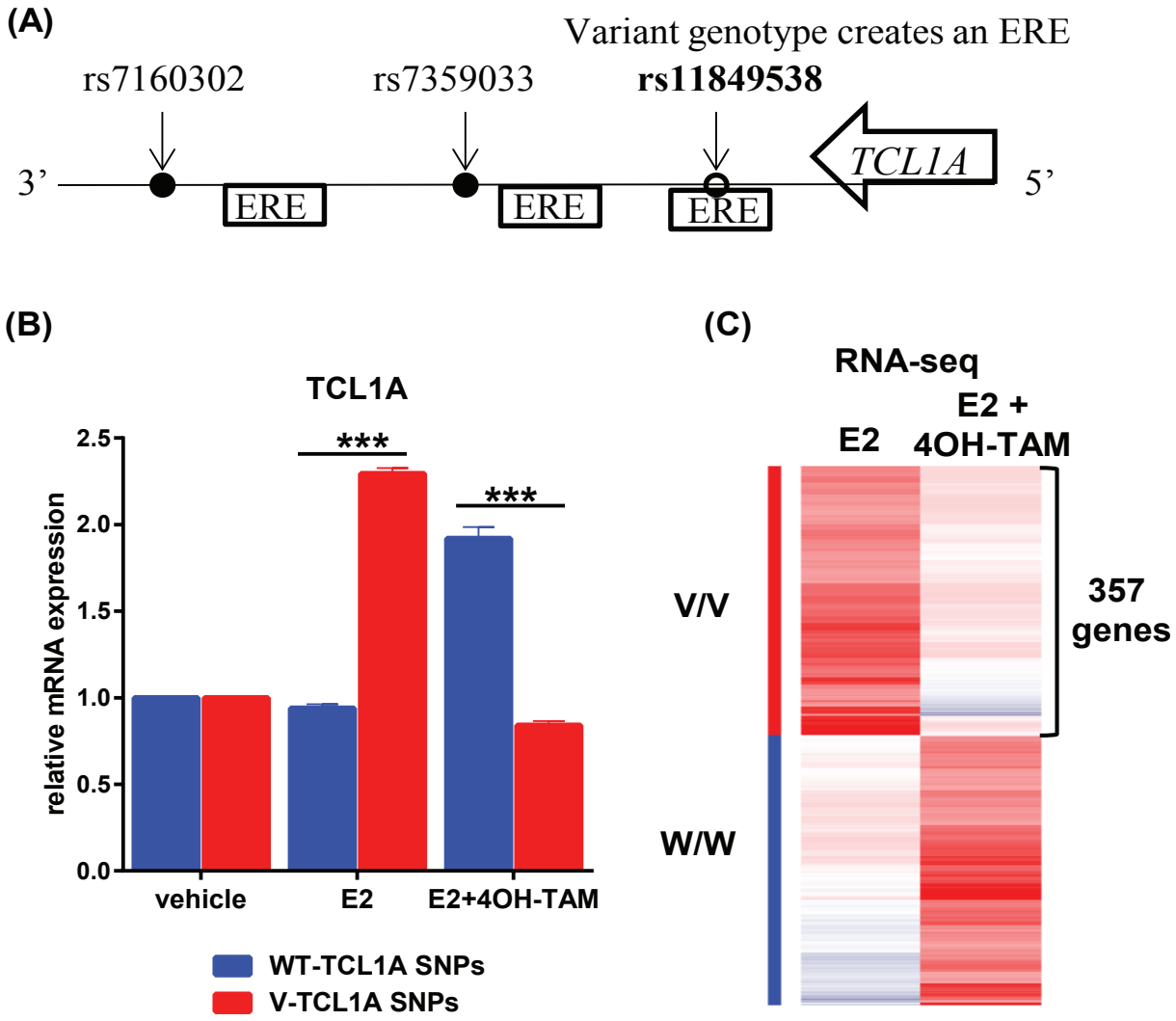

(A)
Variant genotype creates an ERE
rs7160302
rs7359033
rs11849538
3'
ERE
ERE
ERE
TCL1A
5'
(B)
TCL1A
***
***
relative mRNA expression
2.5
2.0
1.5
1.0
0.5
0.0
vehicle
E2
E2+4OH-TAM
WT-TCL1A SNPs
V-TCL1A SNPs
(C)
RNA-seq
E2
E2 +
4OH-TAM
V/V
W/W
357
genes

Figure 2

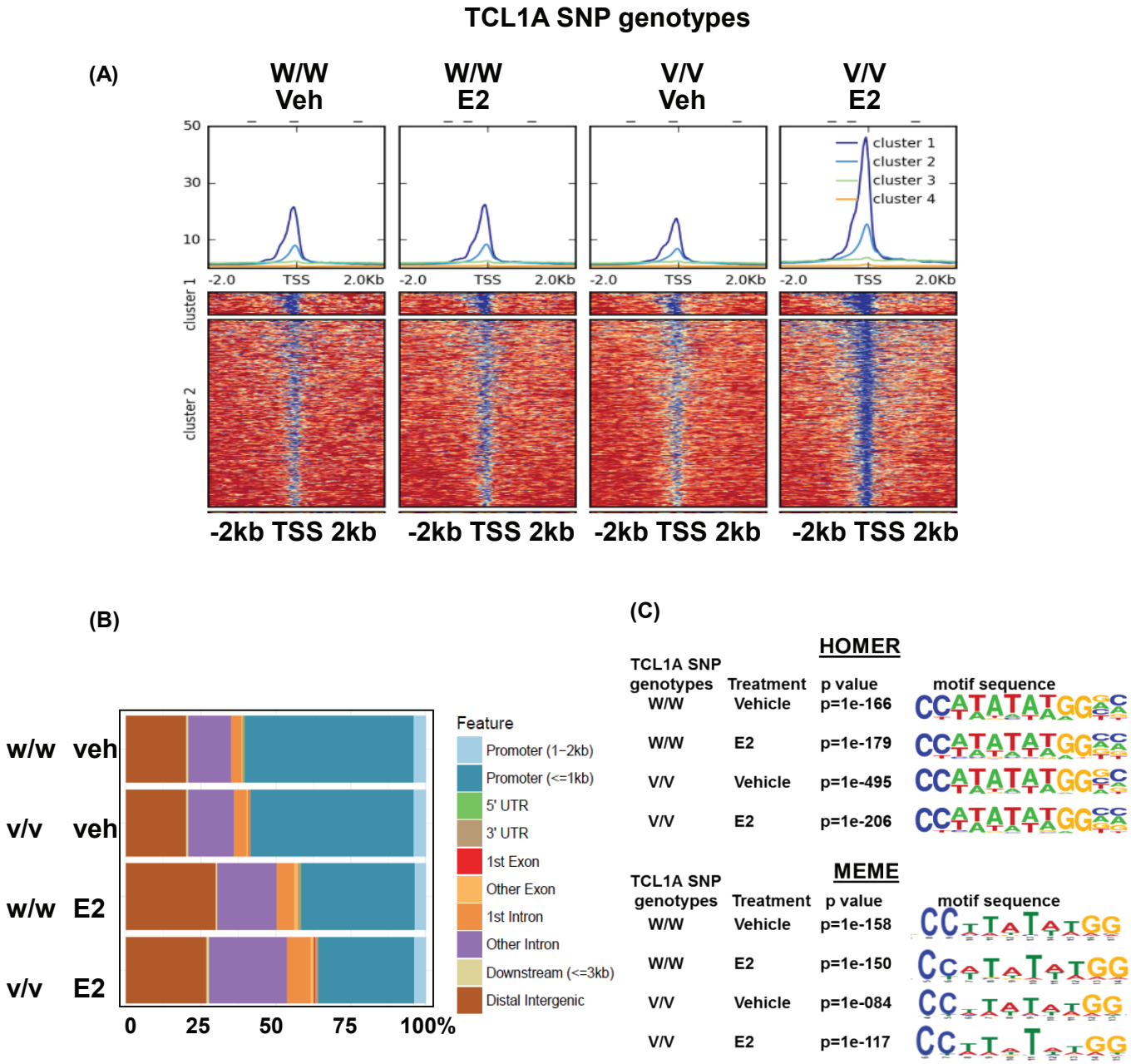

TCL1A SNP genotypes
(A)
W/W Veh
W/W E2
V/V Veh
V/V E2
cluster 1
cluster 2
cluster 3
cluster 4
-2.0
TSS
2.0Kb
-2kb TSS 2kb
(B)
w/w veh
v/v veh
w/w E2
v/v E2
0
25
50
75
100%
Feature
Promoter (1−2kb)
Promoter (<=1kb)
5' UTR
3' UTR
1st Exon
Other Exon
1st Intron
Other Intron
Downstream (<=3kb)
Distal Intergenic
(C)
HOMER
TCL1A SNP genotypes
Treatment
p value
motif sequence
W/W
Vehicle
p=1e-166
W/W
E2
p=1e-179
V/V
Vehicle
p=1e-495
V/V
E2
p=1e-206
MEME
TCL1A SNP genotypes
Treatment
p value
motif sequence
W/W
Vehicle
p=1e-158
W/W
E2
p=1e-150
V/V
Vehicle
p=1e-084
V/V
E2
p=1e-117

Figure 3

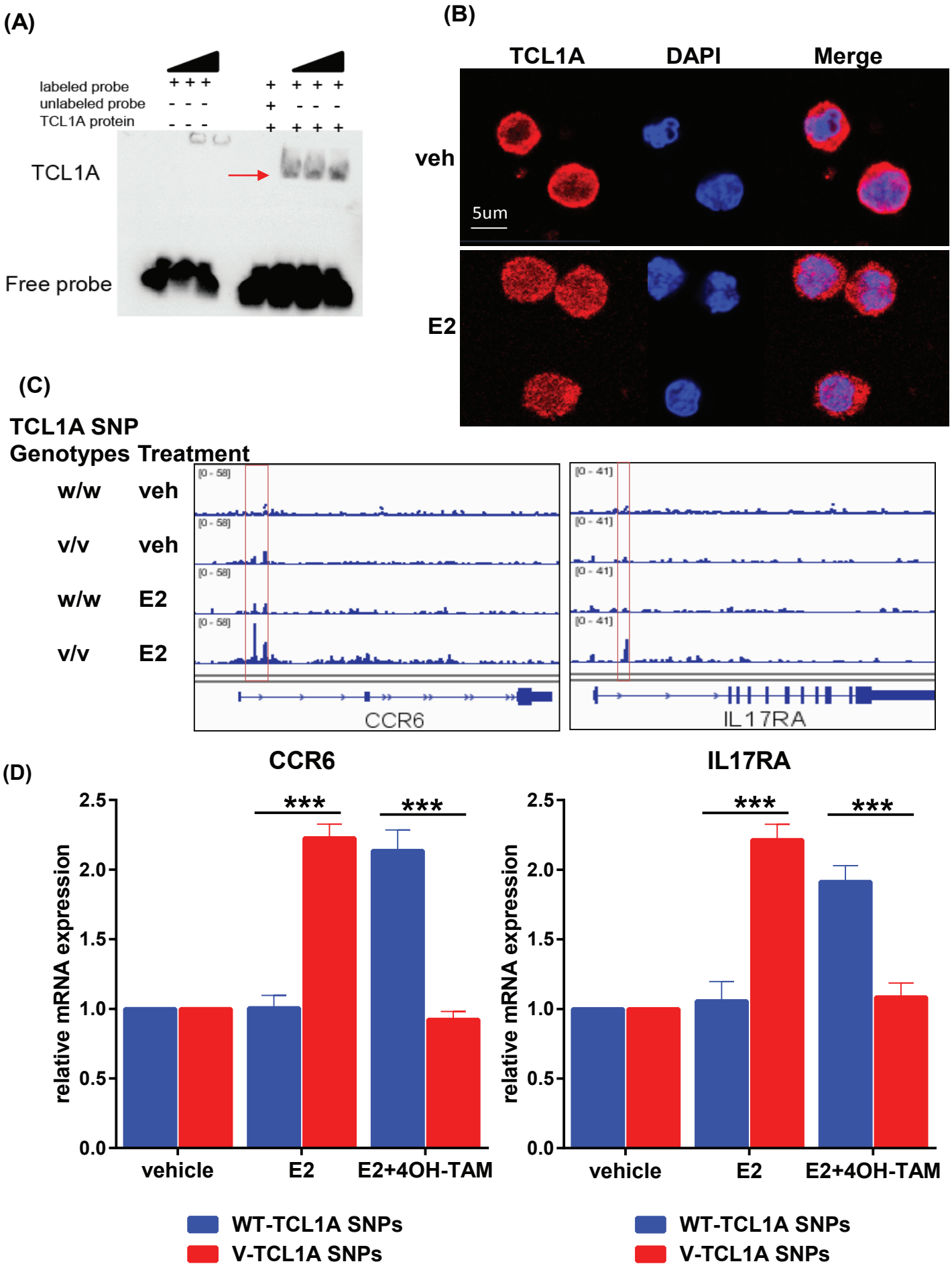

(A)
labeled probe
unlabeled probe
TCL1A protein
TCL1A
Free probe
(B)
TCL1A
DAPI
Merge
veh
5um
E2
(C)
TCL1A SNP Genotypes
Treatment
w/w veh
v/v veh
w/w E2
v/v E2
CCR6
IL17RA
(D)
CCR6
IL17RA
relative mRNA expression
***
vehicle
E2
E2+4OH-TAM
WT-TCL1A SNPs
V-TCL1A SNPs

Figure 4

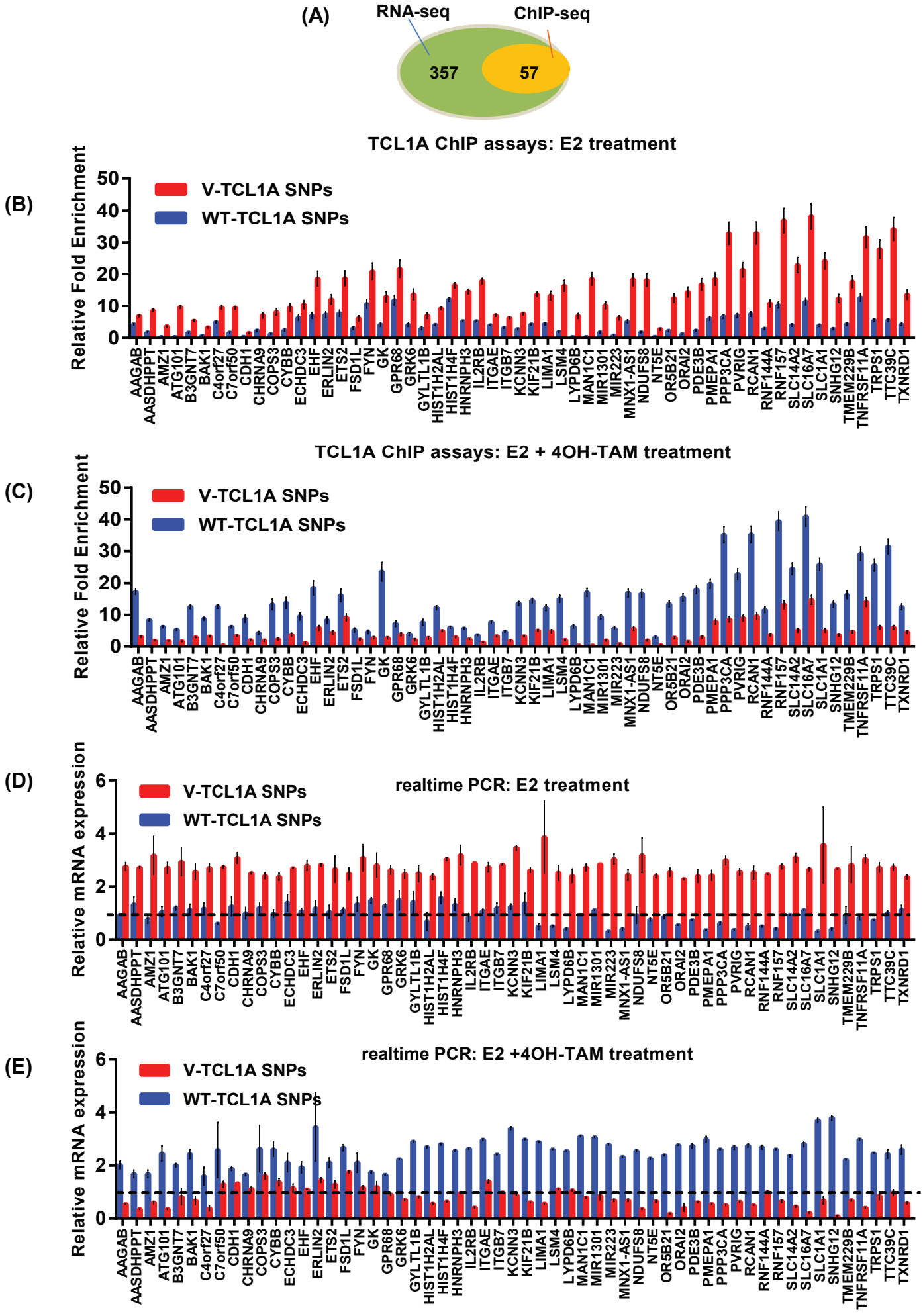

(A)
RNA-seq
ChIP-seq
357
57
(B)
TCL1A ChIP assays: E2 treatment
Relative Fold Enrichment
V-TCL1A SNPs
WT-TCL1A SNPs
(C)
TCL1A ChIP assays: E2 + 4OH-TAM treatment
Relative Fold Enrichment
V-TCL1A SNPs
WT-TCL1A SNPs
(D)
realtime PCR: E2 treatment
Relative mRNA expression
V-TCL1A SNPs
WT-TCL1A SNPs
(E)
realtime PCR: E2 +4OH-TAM treatment
Relative mRNA expression
V-TCL1A SNPs
WT-TCL1A SNPs
AAGAB AASDHPPT AMZ1 ATG101 B3GNT7 BAK1 C4orf27 C7orf50 CDH1 CHRNA9 COPS3 CYBB ECHDC3 EHF ERLIN2 ETS2 FSD1L FYN GK GPR68 GRK6 GYLTL1B HIST1H2AL HIST1H4F HNRNPH3 IL2RB ITGAE ITGB7 KCNN3 KIF21B LIMA1 LSM4 LYPD6B MAN1C1 MIR1301 MIR223 MNX1-AS1 NDUFS8 NT5E OR5B21 ORAI2 PDE3B PMEPA1 PPP3CA PVRIG RCAN1 RNF144A RNF157 SLC14A2 SLC16A7 SLC1A1 SNHG12 TMEM229B TNFRSF11A TRPS1 TTC39C TXNRD1

Figure 5

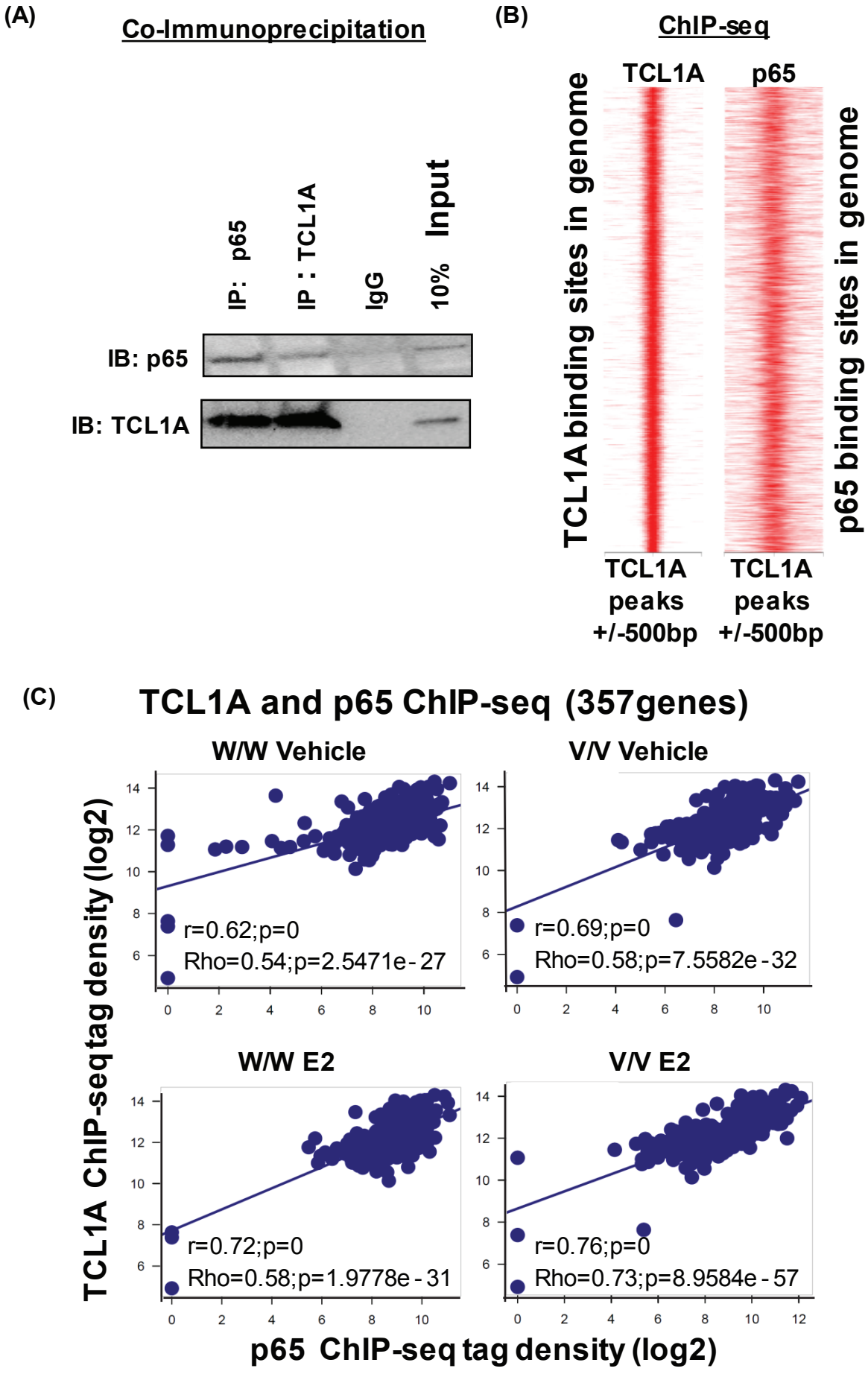

(A)
Co-Immunoprecipitation
IP: p65
IP : TCL1A
IgG
10% Input
IB: p65
IB: TCL1A
(B)
ChIP-seq
TCL1A
p65
TCL1A binding sites in genome
p65 binding sites in genome
TCL1A peaks +/-500bp
TCL1A peaks +/-500bp
(C)
TCL1A and p65 ChIP-seq (357genes)
W/W Vehicle
r=0.62;p=0
Rho=0.54;p=2.5471e-27
V/V Vehicle
r=0.69;p=0
Rho=0.58;p=7.5582e-32
W/W E2
r=0.72;p=0
Rho=0.58;p=1.9778e-31
V/V E2
r=0.76;p=0
Rho=0.73;p=8.9584e-57
TCL1A ChIP-seq tag density (log2)
p65 ChIP-seq tag density (log2)

Figure 6

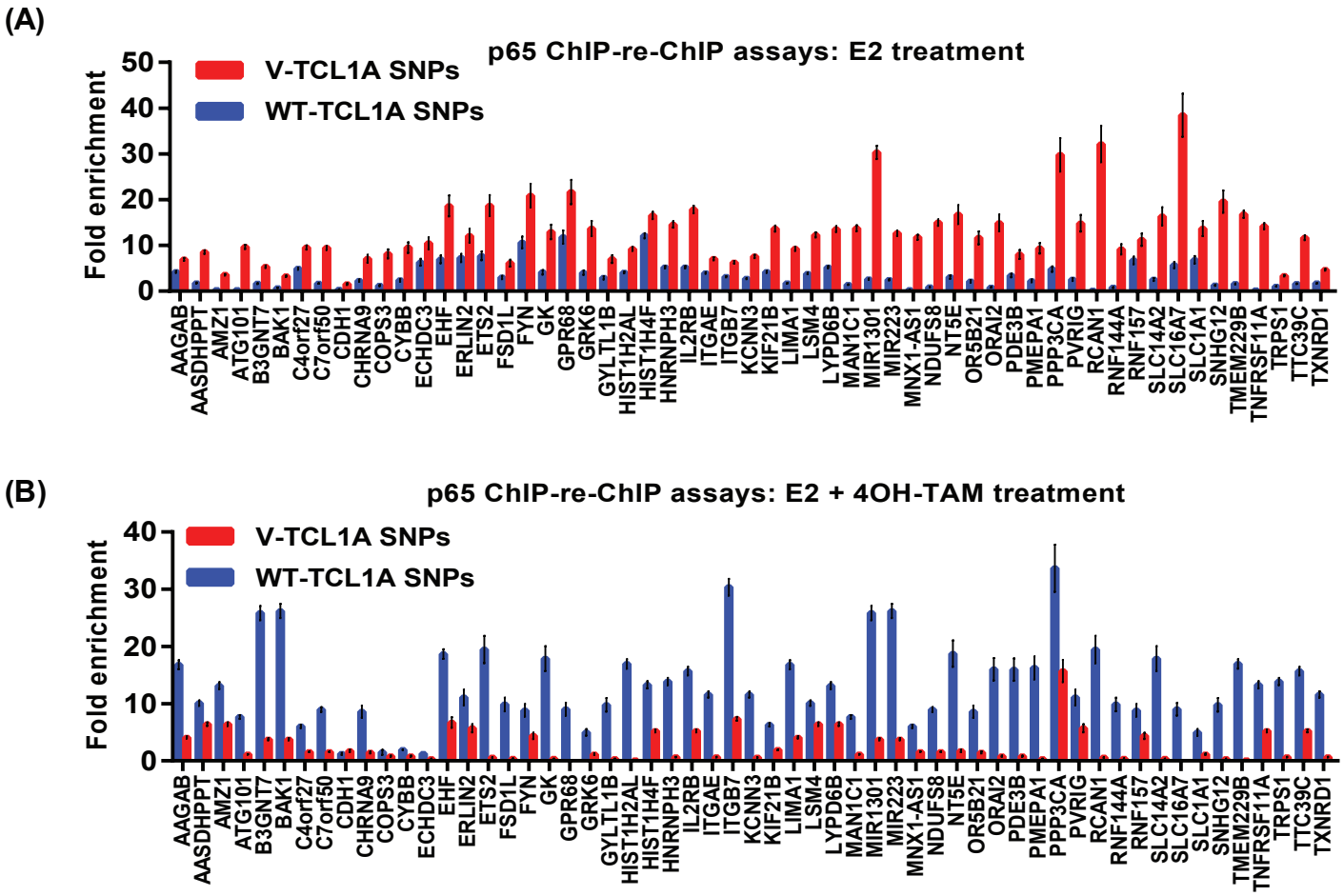

(A)
p65 ChIP-re-ChIP assays: E2 treatment
V-TCL1A SNPs
WT-TCL1A SNPs
Fold enrichment
0
10
20
30
40
50
AAGAB
AASDHPPT
AMZ1
ATG101
B3GNT7
BAK1
C4orf27
C7orf50
CDH1
CHRNA9
COPS3
CYBB
ECHDC3
EHF
ERLIN2
ETS2
FSD1L
FYN
GK
GPR68
GRK6
GYLTL1B
HIST1H2AL
HIST1H4F
HNRNPH3
IL2RB
ITGAE
ITGB7
KCNN3
KIF21B
LIMA1
LSM4
LYPD6B
MAN1C1
MIR1301
MIR223
MNX1-AS1
NDUFS8
NT5E
OR5B21
ORAI2
PDE3B
PMEPA1
PPP3CA
PVRIG
RCAN1
RNF144A
RNF157
SLC14A2
SLC16A7
SLC1A1
SNHG12
TMEM229B
TNFRSF11A
TRPS1
TTC39C
TXNRD1
(B)
p65 ChIP-re-ChIP assays: E2 + 4OH-TAM treatment
V-TCL1A SNPs
WT-TCL1A SNPs
Fold enrichment
0
10
20
30
40
AAGAB
AASDHPPT
AMZ1
ATG101
B3GNT7
BAK1
C4orf27
C7orf50
CDH1
CHRNA9
COPS3
CYBB
ECHDC3
EHF
ERLIN2
ETS2
FSD1L
FYN
GK
GPR68
GRK6
GYLTL1B
HIST1H2AL
HIST1H4F
HNRNPH3
IL2RB
ITGAE
ITGB7
KCNN3
KIF21B
LIMA1
LSM4
LYPD6B
MAN1C1
MIR1301
MIR223
MNX1-AS1
NDUFS8
NT5E
OR5B21
ORAI2
PDE3B
PMEPA1
PPP3CA
PVRIG
RCAN1
RNF144A
RNF157
SLC14A2
SLC16A7
SLC1A1
SNHG12
TMEM229B
TNFRSF11A
TRPS1
TTC39C
TXNRD1

Figure 7

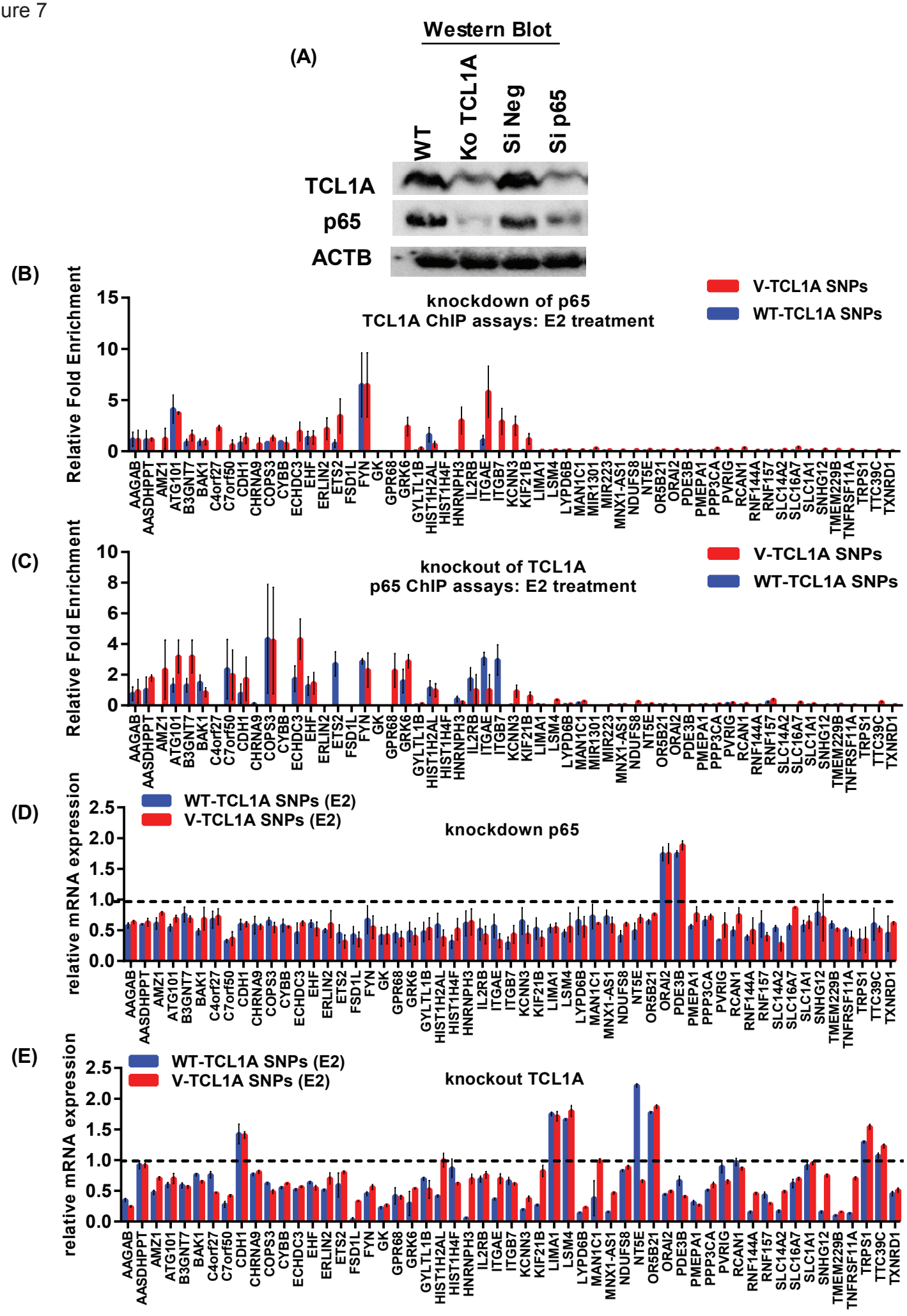
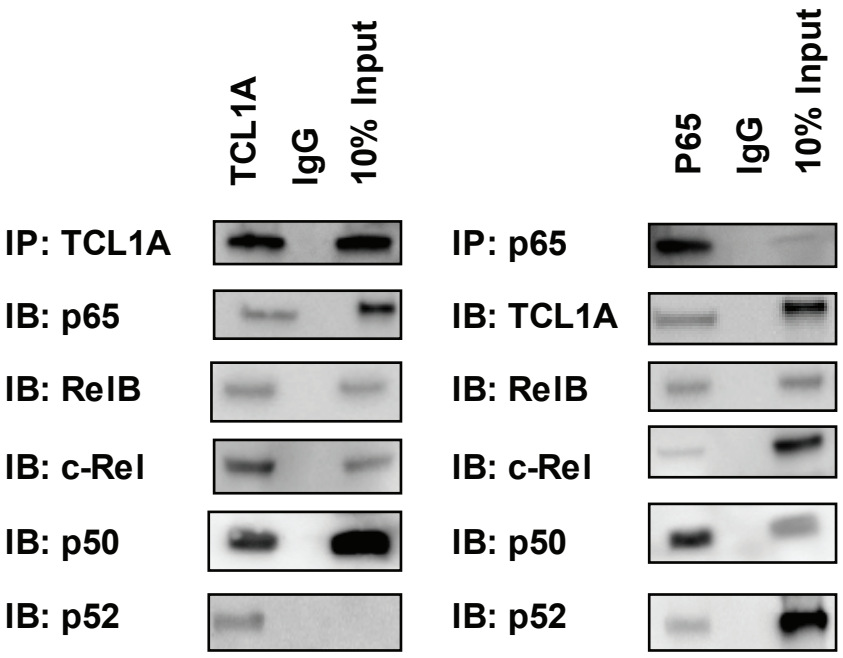

(A)
Western Blot
WT
Ko TCL1A
Si Neg
Si p65
TCL1A
p65
ACTB
(B)
Relative Fold Enrichment
knockdown of p65
TCL1A ChIP assays: E2 treatment
V-TCL1A SNPs
WT-TCL1A SNPs
(C)
Relative Fold Enrichment
knockout of TCL1A
p65 ChIP assays: E2 treatment
V-TCL1A SNPs
WT-TCL1A SNPs
(D)
relative mRNA expression
WT-TCL1A SNPs (E2)
V-TCL1A SNPs (E2)
knockdown p65
(E)
relative mRNA expression
WT-TCL1A SNPs (E2)
V-TCL1A SNPs (E2)
knockout TCL1A

Figure 8


Co-immunoprecipitation
TCL1A
IgG
10% Input
P65
IgG
10% Input
IP: TCL1A
IB: p65
IB: RelB
IB: c-Rel
IB: p50
IB: p52
IP: p65
IB: TCL1A
IB: RelB
IB: c-Rel
IB: p50
IB: p52