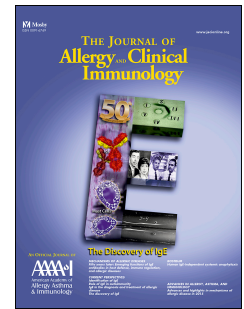


Accepted Manuscript

Genetic Variants with Gene Regulatory Effects are Associated with Diisocyanate Asthma

David I. Bernstein, MD FAAAAI, Zana L. Lummus, PhD, Banu Kesavalu, MS, Jianbo Yao, PhD, Leah Kottyan, PhD, Daniel Miller, BS, André Cartier, MD FAAAAI, Maria-Jesús Cruz, MD, Catherine Lemiere, MD, Xavier Muñoz, MD, Santiago Quirce, MD PhD, Susan Tarlo, MD FAAAAI, Joaquin Sastre, MD PhD FAAAAI, Louis Philippe Boulet, Matthew T. Weirauch, PhD, Kenneth Kaufman, PhD



PII: S0091-6749(18)30932-1

DOI: [10.1016/j.jaci.2018.06.022](https://doi.org/10.1016/j.jaci.2018.06.022)

Reference: YMAI 13503

To appear in: *Journal of Allergy and Clinical Immunology*

Received Date: 12 January 2018

Revised Date: 1 June 2018

Accepted Date: 8 June 2018

Please cite this article as: Bernstein DI, Lummus ZL, Kesavalu B, Yao J, Kottyan L, Miller D, Cartier A, Cruz M-J, Lemiere C, Muñoz X, Quirce S, Tarlo S, Sastre J, Boulet LP, Weirauch MT, Kaufman K, Genetic Variants with Gene Regulatory Effects are Associated with Diisocyanate Asthma, *Journal of Allergy and Clinical Immunology* (2018), doi: 10.1016/j.jaci.2018.06.022.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Genetic Variants with Gene Regulatory Effects are Associated with Diisocyanate Asthma

David I. Bernstein, MD FAAAAI^{1*}, Zana L. Lummus, PhD¹, Banu Kesavalu, MS¹, Jianbo Yao, PhD², Leah Kottyan, PhD^{3,4}, Daniel Miller, BS³, André Cartier, MD FAAAAI⁵, Maria-Jesús Cruz, MD⁶, Catherine Lemiere, MD⁵, Xavier Muñoz, MD⁶, Santiago Quirce, MD PhD⁷, Susan Tarlo, MD FAAAAI⁸, Joaquin Sastre, MD PhD FAAAAI⁹, Louis Philippe Boulet¹⁰, Matthew T Weirauch, PhD^{3,4,11}, Kenneth Kaufman, PhD^{3,4,12}

¹ Division of Immunology, Allergy and Rheumatology, University of Cincinnati College of Medicine, Cincinnati, Ohio (E) bernstd@ucmail.uc.edu; zana.lummus@uc.edu; kesavabr@ucmail.uc.edu

² West Virginia University, Morgantown, West Virginia (E) Jianbo.Yao@mail.wvu.edu

³ Center for Autoimmune Genomics and Etiology, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio (E) Kenneth.Kaufman@cchmc.org; Matthew.Weirauch@cchmc.org; Leah.Kottyan@cchmc.org; Daniel.Miller@cchmc.org

⁴ Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, Ohio (E) Kenneth.Kaufman@cchmc.org; Matthew.Weirauch@cchmc.org; Leah.Kottyan@cchmc.org

⁵ Université de Montréal, Hôpital du Sacré-Coeur de Montréal, Montreal, Quebec, Canada (E) andre.cartier@umontreal.ca; Catherine.Lemiere@umontreal.ca

⁶ Hospitals Vall D'Hebron, Barcelona and CIBER de Enfermedades Respiratorias CIBERES, Madrid, Spain (E) mjcruz@ir.vhebron.net; xmunoz@vhebron.net

⁷ Department of Allergy, Hospital La Paz-IdiPAZ and CIBER de Enfermedades Respiratorias CIBERES, Madrid, Spain (E) squirce@gmail.com

⁸ Department of Allergy, Fundación Jiménez Díaz and CIBER de Enfermedades Respiratorias CIBERES, Madrid, Spain (E) jsastre@fjd.es

⁹ University of Toronto, Toronto, Ontario, Canada (E) susan.tarlo@utoronto.ca

¹⁰ Université Laval, Hôpital Laval, Sainte-Foy, Québec, Canada (E) LPBoulet@med.ulaval.ca

¹¹ Divisions of Biomedical Informatics and Developmental Biology, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio (E) Matthew.Weirauch@cchmc.org

¹² Cincinnati VA Medical Center, Cincinnati, Ohio (E) Kenneth.Kaufman@cchmc.org

* To whom correspondence should be addressed: David I. Bernstein, MD, Division of Immunology, Allergy and Rheumatology, University of Cincinnati Department of Internal

Medicine, 231 Albert Sabin Way, Cincinnati, Ohio 45267-0563; fax: 513-558-3799; E-mail bernstdd@ucmail.uc.edu

This work was supported by the following grants: NIOSH/CDC Grant OH008795; DI Bernstein-PI; NIH shared instrumentation grant (S10 RR027015-01; KD Greis-PI); NIH Grant P30 AR070549 for Cincinnati Children's Hospital Medical Center Functional Genomics Core. Cincinnati Children's Hospital Medical Center CpG Award 53553 (MT Weirauch-PI); Lupus Research Institute "Novel Approaches" Award (MT Weirauch-PI); and NIH R21 HG008186 (MT Weirauch-PI).

Role of the funding sources: The funding sources had no role in the study design, collection, analysis and interpretation of data, writing of the report, or the decision to submit the article for publication.

Disclosure of potential conflict of interest: Dr. Bernstein has received research support from NIOSH-CDC and consultancy fees from the International Isocyanate Institute outside the submitted work. Dr. Cartier received personal fees and other from AZ pharma-Canada, TEVA pharma-Canada, Sanofi Genzyme-Canada, Boehringer-Ingelheim pharma-Canada, GSK pharma-Canada, IC-EBM Canada, outside the submitted work. Dr. Lemiere received grants from Institut de Recherche Robert Sauvé en santé et sécurité au Travail, grants from Canadian Institutes of Health Research, outside the submitted work. Dr. Tarlo has received reimbursement from the NIOSH-CDC grant funding this study, and support from grants during 2003-18: Workplace Safety & Insurance Board, Network of Centres of Research Expertise "Centre of Research Expertise in Occupational Disease (CRE-OD)", Verma Grant, outside the submitted work, and has a patent and Book editor Royalties with royalties paid. Dr. Sastre received payment from an NIH grant, and personal fees from Sanofi, Novartis, AstraZeneca, ALK, LETI, outside the submitted work. Dr. Boulet has received payments for participation in multicenter research studies from AIM Therapeutics, Amgen, Asmacure, AstraZeneca, Axikin, GlaxoSmithKline, Hoffman La Roche, Novartis, Ono Pharma, Sanofi, Takeda, research support from AstraZeneca, Boehringer-Ingelheim, GlaxoSmithKline, Merck, Takeda, consultancy fees from Astra Zeneca, Novartis, Methapharm, royalties for co-authorship of "Up-To-Date" (occupational asthma), grants for production of educational materials from AstraZeneca, Boehringer-Ingelheim, GlaxoSmithKline, Merck Frosst, Novartis, conference fees from AstraZeneca, GlaxoSmithKline, Merck, Novartis and Takeda. The rest of the authors declare that they have no relevant conflicts of interest.

ABSTRACT

Background: Isocyanates are major causes of occupational asthma but susceptibility and mechanisms of Diisocyanate Asthma (DA) remain uncertain.

Objective: The aim of this study is to identify DA-associated functional genetic variants via next generation sequencing (NGS), bioinformatics, and functional assays.

Methods. NGS was performed in 91 workers with DA. Fourteen loci with known DA associated single nucleotide polymorphisms (SNPs) were sequenced and compared to data from 238 unexposed subjects. Ranking of DA associated SNPs based on their likelihood to affect gene regulatory mechanisms in the lung yielded 21 prioritized SNPs. Risk (R) and non-risk (NR) oligonucleotides were tested for binding of nuclear extracts from A549, BEAS 2B, and IMR-90 lung cell lines by electrophoretic mobility shift assays (EMSA). DNA constructs were cloned into a pGL3-promoter vector for luciferase gene reporter assays.

Results. NGS detected 130 risk variants associated with DA ($3.1 \times 10^{-6} - 6.21 \times 10^{-4}$), of which 129 were located in non-coding regions. The 21 SNPs prioritized by functional genomic data sets were in or proximal to five genes: *CDH17* (n=10), *ATF3* (n=7), *FAM71A* (n=2), *TACR1* (n=1), and *ZBTB16* (n=1). EMSA detected allele-dependent nuclear protein binding in A549 cells for 8 of 21 variants. In the luciferase assay, 4 of the 21 SNPs exhibited allele-dependent changes in gene expression. DNA affinity precipitation and mass spectroscopy of rs147978008 revealed allele-dependent binding of H1 histones, which was confirmed by Western blot.

Conclusions. We identified five DA associated potential regulatory SNPs. Four variants exhibited effects on gene regulation (*ATF* rs11571537, *CDH17* rs2446824, rs2513789, and *TACR1* rs2287231). A fifth variant (*FAM71A* rs147978008) showed non-risk allele preferential binding to H1 histones. These results demonstrate that many DA-associated genetic variants likely act by modulating gene regulation.

Key Messages

- Diisocyanate-induced occupational asthma (DA) is associated with multiple disease associated genetic variants.
- Five potential regulatory single nucleotide polymorphisms (SNPs) associated with DA were identified by next generation sequencing (NGS) and potential functional roles confirmed by electrophoretic mobility shift assays (EMSA) and luciferase reporter assays.

Capsule Summary

This study advances knowledge of genetic mechanisms of occupational asthma by identification of a role for regulatory genetic variants associated with diisocyanate-induced asthma.

Key Words

Asthma

Bioinformatics

Diisocyanate

Electrophoretic mobility shift assay

Functional Genomics

Gene regulation

Genetic

Histone

Isocyanate

Luciferase assay

Next generation sequencing

Occupational asthma

Oligonucleotide

Single nucleotide polymorphism

Abbreviations

A549, A549 human lung carcinoma epithelial cell line

BEAS 2B, BEAS 2B epithelial virus transformed normal human lung/bronchus cell line

BWA, Burrows-Wheeler transform

DA, diisocyanate-induced asthma

DAPA, DNA affinity purification assay

EMSA, electrophoretic mobility shift assay

HDI, hexamethylene diisocyanate

IMR-90, IMR-90 normal human fetal lung fibroblast cell line

GATK, Genome Analysis Toolkit (tools for NGS using MapReduce programming)

GWAS, genome-wide association study

MAF, minor allele frequency

MDI, methylene diphenyl diisocyanate

NanoLC-MS/MS, Nano liquid chromatography, followed by tandem mass spectrometry analysis

NGS, next generation sequencing

NR, non-risk allelic oligonucleotide probe

OA, occupational asthma

R, risk allele oligonucleotide probe

rSNPs, non-coding regulatory SNPs

SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis

SIC, specific inhalation challenge

SNP, single nucleotide polymorphism

SNV, single nucleotide variant

TDI, toluene diisocyanate

INTRODUCTION

Occupational asthma (OA) is a common occupational lung disorder⁽¹⁾. Diisocyanate exposure accounts for approximately 25% of approved workers compensation claims for OA⁽²⁾. Causative diisocyanate chemicals include hexamethylene diisocyanate (HDI), methylene diphenyl diisocyanate (MDI), and toluene diisocyanate (TDI) that are used for manufacturing polyurethane foam products, metal surface paints, and other applications^(3, 4). These chemicals remain important respiratory sensitizers and among the ten most common agents associated with work-related asthma in the United States^(5, 6).

Despite clinical resemblance to allergic asthma, allergic mechanisms for DA have not been identified, including biomarkers of sensitization (e.g., specific IgE)^(7, 8). Thus, an unmet need exists for markers that facilitate the identification of workers with high DA susceptibility and early recognition of DA^(9, 10), thereby preventing disability and persistent asthma^(11, 12). Genetic association studies of diisocyanate-exposed workers including two GWAS studies have identified DA associated genetic variants from multiple loci, some reaching genome wide significance⁽¹³⁻²¹⁾. A potential study limitation is the small number of workers with DA able to be recruited (a relatively rare disorder). In the latter publications^(19, 20), a very well defined disease phenotype of occupational asthma confirmed by controlled SIC testing with the work-relevant isocyanate chemical could explain the large number of reported disease associated SNPs. DA-associated SNPs have been found in genes encoding antioxidant enzymes^(13, 21), HLA Class I and II molecules⁽¹⁸⁾, Th1/Th2 cytokines and *CD14*^(14, 16) and epithelial junctional proteins⁽²²⁾. We previously replicated DA associations with two *CTNNA3* (alpha-T catenin) SNPs previously reported in a genome wide association study (GWAS) of Korean workers^(17, 19). We also conducted a GWAS in 74 Caucasian workers with confirmed DA, leading to

identification of 11 SNPs in 6 genetic loci that increase risk for DA ($p < 10^{-7}$) and 38 SNPs in 7 genetic loci with suggestive significance ($p < 10^{-6}$)⁽²⁰⁾.

The aim of this study was to further characterize loci of interest from our GWAS study⁽²⁰⁾ and to identify functional risk alleles associated with confirmed DA. We used a targeted custom NGS approach to sequence the 14 genetic loci that had been identified in our previous studies^(17, 20). Confirmed genetic risk variants were then prioritized based on their potential for altering gene regulatory mechanisms in relevant cell types. Prioritized DA associated oligonucleotides were tested for binding of nuclear extract proteins by EMSA⁽²³⁾ and genotype-dependent DNA regulatory activity was investigated in gene reporter assays⁽²⁴⁾. Finally, using DNA affinity purification and mass spectroscopy, we identified specific proteins bound to DA associated genetic risk variants. Results from this study provide insight into the specific molecular mechanisms impacted by DA associated variants.

METHODS

Study participants

Ninety-one workers exposed to either hexamethylene (HDI), diphenyl methane (MDI), or toluene (TDI) diisocyanate, with DA confirmed by positive specific inhalation challenge (SIC) were studied. As a comparator group, we used 238 subjects with available DNA sequence data from the 1,000 genomic (1KG) control data set⁽²⁵⁾. All subjects were of European descent. Hypothetically, healthy control subjects could develop OA if exposed to diisocyanates, but likely, no more than the known incidence of this disorder (2-5%) vs 100% of subjects with disease. Second, absence of OA could not be assured even among a small number of “asymptomatic” workers without challenge testing, a procedure not considered ethical unless clinically indicated⁽²⁶⁾. For this reason, asymptomatic, unexposed healthy controls have been used for this and another published GWAS study of diisocyanate asthma⁽¹⁹⁾.

Of the 91 cases, 87 underwent controlled specific inhalation challenge testing to a diisocyanate to prove occupational asthma, conducted at four participating sites (2 sites in Spain and 2 sites in Quebec, Canada) in specialized laboratory facilities according to published protocols⁽²⁶⁻²⁸⁾. SIC testing was not available in 4 workers at one center (Toronto, Canada) and serial monitoring of peak expiratory flow rates (PEFR) was used as an alternative procedure to confirm DA⁽²⁹⁾. These patients had the diagnosis rigorously confirmed by serial measurement of peak flow rates recorded by workers every 4-6 hours and at least four times a day during at least 2 weeks while working and 1-2 weeks off work. In addition, methacholine challenge tests were performed at the end of a working week and at the end of at one week off work, showing a significant (at least 3-fold) improvement in methacholine PC₂₀ away from work.

The study protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki and written informed consent was obtained for all subjects. The study protocol was reviewed and renewed annually by the ethics committees of participating institutions.

Study Design

The methodologies used to identify DA-associated variants with regulatory functions are summarized in Fig 1, and described in detail below.

Next Generation Sequencing (NGS)

In order to thoroughly investigate DA-associated variants by NGS in an unbiased fashion, genomic DNA was collected from 91 subjects with DA using the QIAamp blood kit (QIAGEN Inc., Chatsworth, CA, USA). Targeted NGS was performed by the DNA Sequencing and Genotyping Core at Cincinnati Children's Hospital Medical Center, using Illumina protocols for library construction and HiSeq 2500 Rapid Sequencing. Roche Nimblegen design software (www.design.nimblegen.com) and proprietary sequence capture technology was used for target enrichment, consisting of 14 DA associated loci and flanking regions identified by GWAS^(17, 20) ranging from 24,512 to 190,905 bp, using *Homo sapiens* genome build hg19/GRCh37. A target region was defined by identifying all variants in linkage disequilibrium ($r^2 > 0.2$) with a DA associated SNP.

DNA sequence from OA samples was aligned to build 37 of the human reference genome using BWA⁽³⁰⁾. Variants were called with GATK⁽³¹⁾ and analyzed with Goldenhelix SVS software (Bozeman, MT, USA), with 66,393 variants detected in total. Variants with a call rate below 90%, genotype quality score <20 or read depth below 15 were removed. Whole genome variant calls for 238 non-Hispanic European samples were obtained from the 1000 genomes project (<http://www.internationalgenome.org>). The same QC filters were applied to the 1000

genomes samples. Principal component analysis was used to identify genetic outliers (two samples from the 1000 genomes project were dropped). An allelic chi-square test was performed, and variants with departure from Hardy-Weinberg ($p < 0.001$ in control subjects) were removed from the analysis. In total, 31,299 variants passed all quality control measurements and were present in both data sets. To prioritize the variants showing associations we initially used a p-value cutoff of <0.0001 . To correct for multiple testing, we used a false discovery rate threshold of 0.05 for a primary set of variants and 0.15 for suggestive associations. 130 variants passed these requirements and are included in the study (see Supplementary Table E1 in this article's Online Repository at www.jacionline.org). The statistical genetic associations identified genetic variants in which the minor allele frequencies differ in the DA cases and controls. Risk alleles were defined for each variant as the allele that was more common in the subjects with DA than the controls.

Computational identification of variants likely to alter gene regulatory mechanisms

To identify non-coding SNPs associated with DA that might impact transcription factor (TF) binding events in cell types relevant to DA, we compiled ChIP-seq datasets from multiple sources, including the ENCODE Consortium⁽³²⁾ the UCSC Genome Browser⁽³³⁾, Cistrome⁽³⁴⁾, and ReMap⁽³⁵⁾. Custom software was used to restrict datasets to cell types relevant to asthma (cell lines and cells derived from lung tissues, as well as immune cells) and identify DA-associated SNPs whose genomic positions intersect these datasets. SNPs were ranked based on the total number of intersecting asthma-relevant datasets (see Supplemental Table E4 in this article's Online Repository at www.jacionline.org). Additionally, we placed higher priority on genetic loci encoding genes meeting one or more of the following three criteria: 1) genes previously shown to contain SNPs achieving GWAS significance⁽²⁰⁾ (*CDH17*, *PITPNC1*,

ZBTB16, *TACR1*); 2) genes located in a genomic region flanking an intergenic SNP with significant DA association (*LOC101929565*, *ATF3*, *FAM71A*, and *CDH17*); or 3) genes known to have transcriptional regulatory activities in asthma (*ATF3*^(36, 37), *TACR1*^(38, 39)).

Cell Lines

To prepare nuclear extracts for EMSA studies, the following three lung-derived cell lines were cultured: A549 epithelial cells, IMR90 fibroblasts and BEAS 2B epithelial cells. Two of these cell lines were obtained from the American Type Culture Collection (Manassas, VA, USA). A549 human lung carcinoma epithelial cells (ATCC® CCL185™) were cultured in ATCC-formulated F12K Medium, 10% fetal bovine serum (FBS). IMR-90 normal human fetal lung fibroblasts (ATCC® CCL186™) were cultured in ATCC-formulated Eagle's Minimum Essential Medium (EMEM), 10% FBS. BEAS 2B epithelial virus transformed normal human lung/bronchus cells were a gift from Dr. Jeffrey Whitsett and were cultured in RPMI, supplemented with Penicillin-Streptomycin (10,000 U/ml) and 10% heat inactivated fetal bovine serum (Thermo Fisher Scientific, Waltham, MA, USA). Cells were cultured at 37°C with 5% CO₂. Nuclear extracts were prepared from cell lines using the NE-PER Nuclear and Cytoplasmic Extraction Kit from Thermo Fisher Scientific (Waltham, MA, USA).

Electrophoretic Mobility Shift Assays (EMSA)

To screen for nuclear protein-DNA interactions of DA-associated variants, EMSA^(40, 41) was performed using the Odyssey® Infrared Imaging Kit and protocol for EMSA analysis (LI-COR® Biosciences, Lincoln, Nebraska, USA)⁽²³⁾. Twenty-one SNPs were used to prepare 5'IMR700dye-labeled 35 bp duplexed oligonucleotide probes (Integrated DNA Technologies, Inc., Coralville, Iowa, USA). SNP oligonucleotide sequences were obtained from the NCBI dbSNP database (<https://www.ncbi.nlm.nih.gov/snp>), using human genome build 37. Each

oligonucleotide consisted of the SNP allele and 17 bases of 5' and 3' flanking sequence. Both risk and non-risk alleles (n=42) were tested (probe sequences shown in Table I). EMSA binding reactions contained 6 µg nuclear extract, 100 fmol labeled probe, 2 µl 10× binding buffer, 1 µg Poly(dI-dC), 1 µl 40% glycerol, 2 µl 25 mM dithiothreitol (DTT)/2.5% Tween 20, 1 µl 1% NP-40, 1 µl 100 mM MgCl₂, and ultrapure water (LI-COR® Biosciences, Lincoln, Nebraska, USA) in a total reaction volume of 20 µl. The products were electrophoresed on native 5% TBE (Tris-Borate-EDTA) pre-cast polyacrylamide gels (Bio-Rad Life Sciences, Hercules, CA). Gels were pre-run at 70 V/cm for 30 min in 0.5X non-denaturing TBE buffer, samples loaded, run for 60 min, then imaged using an Odyssey® CLx imager (LI-COR® Biosciences, Lincoln, Nebraska, USA).

Dual Luciferase Reporter Assays

These studies were performed to identify allele dependent effects of prioritized SNPs on gene transcription. Luciferase reporter constructs were cloned using the pGL3-promoter vector (Promega, Madison, WI, USA) as a backbone. To create fragments for insertion into this backbone, two complementary oligonucleotides were designed so that annealing of a sense and antisense oligonucleotide formed a 31bp DNA fragment with a KpnI site at the 5' end and a XhoI site at the 3' end (Supplemental Table EIII in this article's Online Repository at www.jacionline.org). DNA fragments were cloned into the vector upstream of the SV40 promoter. Both risk and non-risk allele constructs were produced for each SNP. Larger luciferase reporter constructs for SNP rs147978008 and SNP rs11571537 were made by PCR amplification of DNA fragments containing the SNPs (203 bp for rs147978008 and 205 bp for rs11571537) using a genomic DNA sample heterozygous for both SNPs and primers containing KpnI and XhoI sites. All constructs were verified by DNA sequencing.

Luciferase reporter assays were performed in 96-well cell culture plates. A549 cells (2×10^4 /well) were transfected with a pGL3 reporter construct and the pRL-CMV vector using the Lipofectamine LTX transfection reagent (Invitrogen™, Thermo Fisher Scientific, Waltham, MA, USA). The ratio of reporter construct and pRL-CMV plasmid was 19:1 (100 ng total plasmid per well). Cells were collected 48 h after transfection. Both firefly and renilla luciferase activities were measured in the same sample using the Dual-Glo Luciferase Assay System (Promega, Madison, WI, USA). Normalized firefly luciferase activity was obtained by dividing the firefly luciferase activity by the corresponding renilla luciferase activity in the same sample. Three independent experiments were performed and eight replicate samples were assayed in each experiment. Relative luciferase activity was calculated by normalizing the luciferase activity of the empty pGL3-Promoter vector to 1. Student's t-test was performed to determine differences between the risk and non-risk alleles of each SNP.

DNA affinity purification assay (DAPA) and Mass Spectrometry (MS)

DAPA was used to characterize specific proteins in the nuclear extracts from A549 and BEAS 2B cells with binding to specific oligonucleotides of interest. Duplexed 5'-biotinylated oligonucleotides (35 bp) containing SNPs were obtained from IDT (Integrated DNA Technologies, Inc., Coralville, Iowa, USA) for use as probes. Binding reactions were performed with biotinylated probes, nuclear extracts, binding buffer, binding enhancer, protease inhibitor, phosphatase inhibitor and 0.1 µg poly (dI-dC) according to protocols supplied with the µMACS Factor Finder Kit (Miltenyi Biotec, San Diego, CA, USA). DAPA eluates were collected and analyzed for protein identification by MS. Proteins were separated by silver-stained SDS-PAGE and gel bands were submitted to the University of Cincinnati Proteomics Laboratory for

identification by trypsin digestion and nano liquid chromatography coupled tandem mass spectrometry (nanoLC-MS/MS).

NanoLC-MS/MS analyses were performed on a TripleTOF 5600+ (Sciex, Toronto, Canada) attached to an Eksigent (Dublin, CA) nanoLC.ultra nanoflow system as described previously⁽⁴²⁾ but with the following changes: 1) The LC mobile phases included 0.1% formic acid in water as solvent A and 0.1% formic acid in acetonitrile as solvent B; 2) The samples were trapped and desalted on an Eksigent NanoLC Trap (ChromXP C18-CL-3 μm 120Å, 350 μm x 0.5 mm) at 1 $\mu\text{L}/\text{min}$ in 0/1% formic acid for 15 min; 3) All data were recorded using Analyst®-TF (v 1.7, build 96) software.

MS/MS fragmentation spectra were searched against an in-house FASTA database, consisting of 553,831 protein entries downloaded from the UniProt database (www.uniprot.org) on November 16, 2016 along with additional contaminating microbial reference sequences using SCIEX (Concord, Ontario, Canada. Unless otherwise noted, protein identification was reported at a 1% false discovery rate as measured against an inverse database of all 553,831 protein entries with a minimum of two peptides with fragmentation spectra reaching at least a 99% confidence level.

Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Pre-cast polyacrylamide gels (Invitrogen™ Novex™ NuPAGE™ 4-12% Bis-Tris Protein Gels, 1.0mm x 12 well) with Nu Page MES-SDS running buffer, and a Novex Mini Gel Tank (Fisher Scientific, Waltham, MA, USA) were used. Samples, containing 14 μl DAPA eluate, 1 μl of 1M DTT and 5 μl of NuPAGE LDS loading buffer 4X, were heated at 70°C, 10 min. Gels were loaded with 18 μl sample, 10 $\mu\text{l}/\text{well}$ Precision Plus Protein™ All Blue Prestained Standards (10 kDa to 250 kDa (Bio-Rad Life Technologies Corporation, Carlsbad, CA, USA).

After electrophoresis (200v, 20 min) gels were stained with Bio-Rad Silver Stain Plus kit (Bio-Rad Life Sciences, Hercules, CA, USA).

Western Blotting

Western blotting was performed to confirm histone binding previously identified in the aforementioned DAPA experiment. by Li-COR Near-infrared Western Blot Detection (<https://licor.app.boxenterprise.net/s/fxc6evxvxbub4srkqy6i9yg46l7i0xz5>), using rabbit polyclonal affinity purified anti-human histone H1.2, Abcam product number ab17677, (Abcam.com, Cambridge, MA, USA) diluted 1:500, followed by incubation with IRDye® 800CW donkey polyclonal affinity-purified anti-rabbit IgG (H+L), diluted 1:10,000, and imaged using an Odyssey® CLx imager and Image Studio Software (LI-COR® Biosciences, Lincoln, Nebraska, USA).

RESULTS

The 91 DA workers were predominantly male, of European ancestry, with a mean age of 40.2 years (Table II). Mean duration of diisocyanate exposure was 143 months. The comparator group consisted of 238 males of non-Hispanic European ancestry.

Fourteen loci were selected for NGS, based on prior DA association^(17, 20) (Table III). We identified 66,393 genetic variants, and 130 variants were significantly associated with DA (range of p values 3.13×10^{-6} – 6.21×10^{-4} see Supplemental Table E1 in this article's Online Repository at www.jacionline.org). Of these, 21 variants were previously identified in our GWAS study⁽²⁰⁾ and 110 were new DA variants located in or proximal to the following specific genes: *TACR1* (33), *CDH17* (27), *ZBTB16* (20), *ATF3* (14), *FAM71A* (7), *HDAC9* (6), *SLC6A12* (7), *CTNNA3* (4), *GALNT14* (4), *PITPNC1* (3), *NPAS3* (3), *IQSEC3* (1), *PSMD12* (1) and *KCND3* (1). Only one of the 130 SNPs was found in a coding region (synonymous SNP rs2513797).

To prioritize the NGS derived genetic variants, we performed a computational analysis of the 130 DA associated SNPs and these were ranked based on the number of overlapping functional genomics datasets in disease-relevant lung cell types (see Methods). The 21 top ranked SNPs localized to six DA risk loci within putative gene regulatory regions (Table IV). Four of these variants (rs2446823, rs2446824, rs2513791, rs1672692) were previously identified as DA risk variants in our GWAS⁽²⁰⁾. Many of the variants were located within ChIP-seq peaks for particular transcription factors in lung cell-types (Supplemental Table E2 in this article's Online Repository at www.jacionline.org), suggesting possible disease-relevant functional roles. For example, an intergenic locus 3' of the *CDH17* gene contains eight SNPs located within

ChIP-seq peaks for SPI1, CEBPB, RELA, EP300, FOSL2, and EP300 in A549 and/or IMR-90 cells).

To screen for allele-dependent genotype binding of nuclear transcription factors, we next performed EMSA with 21 SNP oligonucleotides and nuclear extracts of A549 cells. These experiments revealed that 10 of the 21 tested SNPs bound to nuclear proteins, with 8 displaying genotype-dependent binding (Fig 2 and Supplemental Fig E2 in this article's Online Repository at www.jacionline.org). Six genetic variants, rs1001304, rs11571537, rs147978008, rs2287231, rs2513789, and rs2513791 (not shown) demonstrated increased binding to the non-risk allele, while two genetic variants (rs2513788 and rs2513790) showed stronger binding to the risk allele. The rs147978008 non-risk variant probe, containing the ACA allele, showed the greatest specific binding with A549 nuclear extract proteins, while the corresponding risk variant probe, containing a deletion, showed no detectable binding.

Oligonucleotides rs147978008 (*FAM71A* locus) and rs11571537 (*ATF3* locus) showed a similar pattern of allele-dependent binding, with a preference for the non-risk allele binding to proteins in the nuclear extracts from 3 different cell lines (A549, BEAS 2B, and IMR-90) (Fig 3 and Supplemental Fig E2 in this article's Online Repository at www.jacionline.org). Inhibition with unlabeled oligonucleotides established the specificity of binding for these two variants (Supplemental Fig. E1 in this article's Online Repository at www.jacionline.org). Both genetic variants are located on chromosome 1 and fall within ChIP-seq peaks for SPI1 in IMR-90 cells and POLR2A in A549 and IMR-90 cells (Supplemental Table E2 in this article's Online Repository at www.jacionline.org).

DAPA was subsequently performed to determine specific nuclear proteins bound to rs147978008 using nuclear extract proteins from epithelial cells (A549 and BEAS 2B). Binding

activity for the rs147978008 non-risk oligonucleotide probe was assessed using SDS-PAGE analysis of the eluates obtained (Fig 4, A). Strong bands for both extracts were observed for proteins having a molecular size of 29 kDa. These two bands were analyzed by nano liquid chromatography coupled tandem mass spectrometry (nanoLC-MS/MS). In the BEAS 2B band, histones H1.4 (11 peptides match), H1.5 (4 peptides match) and H1.2 (3 peptides match) were clearly identified. For the A549 band, no protein identifications reached the 2 peptide standard, but two proteins were detected, histones H1.2 and H1.3, each with a single 99% confidence peptide.

We confirmed genotype-dependent binding of H1.2 by Western Blot for the risk and non-risk DAPA purified eluates of A549 and BEAS-2B cells (Fig 4, B). Regardless of the cellular source of nuclear lysate, the non-risk allele shows stronger binding of H1.2 than the risk allele, consistent with EMSA. These results indicate that rs147978008 in the promoter of *FAM71A* is bound by H1.2 in a genotype-dependent manner.

To investigate the 21 prioritized genetic variants for allele dependent effects on gene regulation, a luciferase reporter assay was employed. For each variant, two luciferase reporter constructs, representing the non-risk and risk alleles, were generated and transfected into A549 cells. Variants were assessed as a 31 bp region in the context of the SV40 promoter. Luciferase activity in the transfected cell lines was measured to determine the transcriptional activity for each variant (Fig 5 and Table V).

Three genetic variants showed significant genotype-dependent differences in luciferase reporter activity ($P < .05$), with the rs2446824 risk allele and the rs2513789 and rs2287231 non-risk alleles driving enhanced expression (Fig 5. A, B, C). Since rs11571537 and rs147978008 both demonstrated differential allelic binding by EMSA, we retested expanded oligonucleotide

sequences surrounding these variants (203 bp for rs147978008 and 205 bp for rs11571537). The rs147978008-203bp also failed to show an allelic difference in gene transcription activity, whereas the rs11571537-205bp non-risk allele showed enhanced luciferase activity (see Fig 5, E). Of the 10 genetic variants with EMSA binding, significant genotype-dependent luciferase reporter activity was identified in four (rs11571537, rs2446824, rs2287231, rs2513789). For the fifth SNP (rs147978008) showing allele-dependent binding with histone H1.2, an allele-dependent difference in transcriptional activity was not found.

DISCUSSION

By targeted sequencing of DA-associated loci found in a previous GWAS⁽²⁰⁾, we identified 130 non-coding variants significantly associated with DA. A limitation of GWAS is that because it is designed to capture common variations in the genome by relying on tagging common haplotype blocks, it may be unable to capture variants of low minor allele frequency ($MAF < 5\%$) and rare variants ($MAF < 0.5\%$). Therefore, precise high-throughput next generation DNA sequencing (NGS) is preferred for detecting associations with rare genetic variants and for identifying functional variants^(43, 44). Our use of NGS did not reveal any low MAF or rare variants, but was highly successful in identifying additional non-coding variants associated with DA in genomic regulatory regions, some of which may have functional activities that contribute to development of DA. This is the first study to employ NGS in the investigation of OA and DA. Although our sample size of DA workers is modest, our results are strengthened by the well-defined DA phenotype, confirmed in all but 4 subjects by controlled specific inhalation testing in a specialized laboratory with a work-relevant diisocyanate chemical.

Our results suggest that functional DA associated variants may exist in non-coding regions of at least four genetic loci [i.e., *FAM71A* (rs147978008), *ATF3* (rs1157537), *TACR1* (rs2287231), and *CDH17* (rs2446824, rs2513789)]. To identify potentially functional disease-associated non-coding variants, we used a computational prioritization strategy employing publicly deposited functional genomic datasets. Using this approach, we identified 21 top-ranked variants, of which 10 bound nuclear proteins, with eight displaying genotype-dependent binding. One of the risk variants (rs147978008) affected binding of the nuclear histone H1.2. In luciferase reporter assays, four of the 21 variants showed allele-dependent transcriptional activity. It is possible for certain transcription factors to be expressed and activated as a function

of cellular context⁽⁴⁵⁾, which may explain why more of the prioritized variants did not demonstrate genotype-dependent functional activity.

The non-risk allele for rs147978008 exhibited strong nuclear binding with EMSA but did not exhibit allele-dependent reporter activity. Nevertheless, our data demonstrate binding of the linker histone, H1.2, to the rs147978008 reference allele after DAPA purification and MS. Western blotting confirmed allele-dependent binding of histone H1.2. This SNP is located in the promoter region of *FAM71A*, the coding gene for a Golgi-resident Rab2B specific binding protein GARI-L4 (also called the FAM71A protein), with an important role in the maintenance of the compacted Golgi morphology of mammalian cells⁽⁴⁶⁾. No other function has been reported for this protein, and this is the first report of allele-dependent binding of H1 histones to the *FAM71A* promoter. Based on previous experimental evidence, H1 histones may enhance or inhibit gene transcription^(47, 48), suggesting that modulation of H1 binding by rs147978008 might result in altered *FAM71A* expression levels in a cellular context.

Three of the 21 prioritized DA genetic variants are located near the *ATF3* gene. The activating transcription factor 3 (*ATF3*) is a member of the ATF/cAMP responsive element binding (CREB) protein family of transcription factors and is believed to play a central role in the cellular adaptive response to stress signals⁽⁴⁹⁾. It is noteworthy that mutations in *ATF3* have been shown to mitigate inflammatory cell recruitment to the lung^(36, 37, 50-52), maintenance of cell-cell adhesion and barrier function⁽⁵³⁾, and maintenance and repair of genomic stability, in response to a wide range of cellular stresses including oxidative stress and DNA damage^(7, 54-56).

We found 33 DA-associated *TACR1* variants, but only one (rs2287231) was included in the final 21 top-ranked SNPs (see Table IV), and this SNP showed enhanced allele-dependent luciferase activity. The *TACR1* gene codes for the neurokinin-1 receptor (NK1R), for which the

tachykinin, substance P (SP) is the preferred ligand⁽⁵⁷⁾. Both SP and NKA, which binds preferentially to the neurokinin-2 receptor (NK2R) have been shown to play a role in airway hyperresponsiveness and airway inflammation in animal rodent models of isocyanate induced airway inflammation^(39, 58, 59). In human studies, there is indirect evidence that tachykinins are involved in asthma⁽³⁸⁾. However, the exact role of tachykinins and neuropeptides in the pathogenesis of DA is unknown.

Genetic variants mapping to the *CDH17* (cadherin 17) gene were previously identified in our GWAS⁽²⁰⁾. One *CDH17* intronic variant showed genome-wide significance ($p=1.22 \times 10^{-8}$) and 5 other non-coding variants had suggestive genome-wide significance ($p < 1 \times 10^{-6}$). *CDH17* is a member of the cadherin superfamily, which consists of genes encoding calcium-dependent, membrane-associated glycoproteins. The encoded protein is cadherin-like, consisting of an extracellular region, containing seven cadherin domains, and a transmembrane region lacking the conserved cytoplasmic domain. Cadherins are cell adhesion proteins that preferentially interact with themselves in a homophilic manner in connecting cells. Our sequencing results identified five non-coding variants at a locus 3' of the *CDH17* gene, located within ChIP-seq peaks for CEBPB (CCAAT/Enhancer Binding Protein B) in A549 and IMR90 cells. Two *CDH17* variants within this locus showed expression in the luciferase reporter assay, with rs2513789 showing preferential expression of the non-risk allele, and rs2446824 showing preferential expression by the risk allele.

In summary, this NGS study of previously recognized DA associated gene loci identified 130 non-coding disease associated SNPs. After screening for those intersecting with asthma-relevant datasets, 21 DA associated SNPs were prioritized, of which 10 exhibited nuclear binding by EMSA. Four of these SNPs exhibited allelic effects on gene transcriptional (i.e.,

luciferase) activity, suggesting that these oligonucleotides modify transcriptional events that could affect the clinical expression of DA. A possible limitation of these findings is that our results have not been replicated which will be challenging due to the rarity of Caucasian workers with confirmed DA and stored DNA. Finally, we discovered a novel SNP with differential allelic nuclear binding activity, in part attributable to binding of the histone H1.2 protein. Future studies are planned to study differential allelic binding to H1.2 and how these interactions could potentially modify gene transcription.

Table I. List of Oligonucleotide Probes Tested in EMSA

	SNP		Non-risk allele	Risk allele
1	rs2513788	5' - 3'	TA GCC CAT CCT CCT TGT [T] TT ATG GGC CCA GAT GGC	[G]
2	rs1001304	5' - 3'	TC CAT CTG GTT AAC CAA [A] GA GGG CTT CCA GGA GGA	[G]
3	rs149630836	5' - 3'	TA GAA TTG CCT ACT ATC [TCA GTA G] TC AGT AGC CCA CCA TCT C	[DEL]
4	rs11571537	5' - 3'	TG ACC ACA GAT CCC CGG [T] TG AGA GGA ATG CCC AGA	[C]
5	rs75465959	5' - 3'	CT TGA TGT GTA TGG TAA [G] TG TAA ATG AAG GTC TAT	[A]
6	rs11571563	5' - 3'	CT TTG ACC TCA TGT ATG [T] TT CCT TTA AAT GAT CAC	[G]
7	rs1672692	5' - 3'	TT CTT GTC ATA TGT AAA [G] AG GGA ATC ATT GTG GTA	[A]
8	rs17019510	5' - 3'	TT ACA CCA CAT GCC CAG [A] TA GAG ATG TAC CTG TTG	[G]
9	rs2251996	5' - 3'	TG AAG GAG CAA GTG GTG [G] TC TAC GTG AAC AGG TCA	[C]
10	rs2446823	5' - 3'	TA CAG AAG CAA AAG TTC [A] TC TGA CTG AGG GAA AGC	[C]
11	rs2446824	5' - 3'	AT CCA TTT TTC CAT TGC [G] AC ATT TGG GCA GAT GTG	[A]
12	rs2513789	5' - 3'	AT TGT TTT TTC ATT GTC [G] AT GTT ATC TAA GTC ATC	[A]
13	rs2446821	5' - 3'	AG GGT CCT CCT TCA GAA [C] AA CAT CAG CTT CCC AAT	[A]
14	rs2513790	5' - 3'	CT TTC CAA GCA CAT ACT [T] AA AAG AAT CAC TTG ACA	[C]
15	rs2513791	5' - 3'	TC AGT CAG CTC AGG CCA [C] CA TAA CAA AAT ACC ATA	[A]
16	rs74138575	5' - 3'	TC AGC ACT TCT TTC AGG [G] GA TAA TTT ATG TTC CTT	[A]
17	rs117579120	5' - 3'	GT GAG AAA CTC TTG GGA [C] AC ATC CAG AGT GAA GTT	[A]
18	rs72756369	5' - 3'	CC AGT GTG GAG CCT CTT [T] CT GTC TTG TTC ATT CTT	[A]
19	rs11571559	5' - 3'	TC TGT TTC ATG TCT AAA [C] GT GTG TCT CAT TGT TGC	[T]
20	rs2287231	5' - 3'	GC CAC TTG GCT GAT GAA [G] GA CAA CAT TGG CAG GGA	[A]
21	rs147978008	5' - 3'	TA TCC CAT TTA ATC CTC [ACA] ACA ACT GCT ACC CCC ATT	[DEL]

Table II. Subject Demographics and Exposure Information

SUBJECT GROUP	OA positive N=91	1KG CONTROLS* N=238
GENDER (M/F)	77/14	238/0
AGE Mean $\pm$ SD (Range)	40.21 $\pm$ 11.61 (22.00 to 64.81)	unknown
Race (white/other)	91/0	238/0
ETHNICITY (Fr-Canadian/Spanish/Other)	64/18/9	N/A
EXPOSURE CHEMICAL (HDI/MDI/TDI)	43/21/27	N/A
EXPOSURE DURATION Mean months $\pm$ SD	142.88 $\pm$ 93.79	N/A
SMOKING (Current/Never/Ex)	15/40/36	N/A
# PY Mean $\pm$ SD	10.02 $\pm$ 13.05	N/A

*1KG controls were selected using a composite reference panel of integrated haplotypes from the 1000 Genome Project sequence data freezes from March, 2012. Imputed genotypes were required to meet or exceed a probability threshold of 0.9, an information measure of >0.4, and the same quality control criteria used for our sequence study.

Table III. Targeted Regions of Genes Sequenced in Workers and Genomic Controls

Genes Selected for Sequencing			Area of Genome Sequenced (build 37)			
Gene	best p value	gene size (# bases)	Chr	Start	Stop	Total
ATF3	6.53E-09	551,449	1	212706176	212826619	120,443
CDH17	1.22E-08	90,690	8	95106894	95262031	155,137
CTNNA3	2.66E-02	1,777,118	10	67639776	69488449	1,848,673
GANTL4	1.33E-10	673,887	2	31100831	31385006	284,175
HDAC9	1.33E-13	910,422	7	18094072	19071635	977,563
HERC2	6.94E-14	210,995	15	28323683	28599813	276,130
KCND3	7.18E-08	213,523	1	112285944	112564647	278,703
NPAS3	2.42E-06	748,366	14	33371615	34305882	934,267
ODZ3	8.59E-09	480,463	4	183212637	183756677	544,040
PITPNC1	6.33E-07	313,431	17	65340897	65725879	384,982
PRKCA	2.55E-05	506,184	17	64266260	64839362	573,102
SLC6A12	1.47E-06	24,512	12	266743	356240	89,497
TACR1	4.08E-05	153,972	2	75241090	75509145	268,055
ZBTB16	1.68E-07	190,905	11	113897789	114159202	261,413
Total		4,214,561				6,996,180

Sequencing was performed by targeted Illumina NGS. Target regions consisted of 14 DA associated loci and flanking regions previously identified by GWAS^(17, 20). A target region was defined by identifying all variants in linkage disequilibrium ($r^2 > 0.2$) with a DA associated SNP based on the 1000 genomes project data set.

DNA sequence data from subjects with confirmed DA were aligned to *Homo sapiens* genome build hg19/GRCh37 (build 37) using BWA⁽³⁰⁾. Variants were called with GATK⁽³¹⁾ and analyzed with Goldenhelix SVS software (Bozeman, MT, USA).

Table IV. Genetic variants considered most likely to function by altering gene regulatory mechanisms in disease-relevant cells

Type of Marker*	Variant	Chr: position	Gene	genomic region	Flanking Genes (Left: :Right)	Risk Allele	Non-Risk Allele	$\chi^2 P^{**}$	Number of Data-sets***
SNV	rs1001304	1:212732686	none	intergenic	LOC10192956: :ATF3	C	T	3.15×10^{-4}	50
SNV	rs72756369	1: 212770271	ATF3	intronic		A	T	3.42×10^{-4}	27
SNV	rs11571537	1:212787025	ATF3	intronic		C	T	4.78×10^{-4}	12
SNV	rs11571559	1:212794550	ATF3	promoter		T	C	8.77×10^{-5}	4
SNV	rs11571563	1:212794873	none	intergenic	ATF3: :FAM71A	G	T	2.80×10^{-4}	5
SNV	rs74138575	1:212794997	none	intergenic	ATF3: :FAM71A	A	G	8.77×10^{-5}	6
SNV	rs75465959	1:212795196	none	intergenic	ATF3: :FAM71A	A	G	8.77×10^{-5}	6
Deletion	rs147978008	1:212796008	FAM71A	promoter		-	ACA	8.77×10^{-5}	3
SNV	rs17019510	1:212816729	none	intergenic	FAM71A: :Loc100129948	G	A	4.6×10^{-6}	7
SNV	rs2287231	2:75449129	none	intergenic	TACR1: :GAPDHP59	A	G	4.30×10^{-5}	17
SNV	rs2446824	8:95127574	none	intergenic	RPL34P18: CDH17	T	C	3.00×10^{-5}	32
SNV	rs2446823	8:95127612	none	intergenic	RPL34P18: CDH17	G	T	3.00×10^{-5}	32
Deletion	rs149630836	8:95127836	none	intergenic	RPL34P18: CDH17	-	TCAGTAG	4.82×10^{-5}	35
SNV	rs2513788	8:95127896	none	intergenic	RPL34P18: CDH17	G	T	3.00×10^{-5}	32
SNV	rs2513789	8:95128113	none	intergenic	RPL34P18: CDH17	A	G	3.00×10^{-5}	16
SNV	rs2513790	8:95128147	none	intergenic	RPL34P18: CDH17	C	T	3.00×10^{-5}	23
SNV	rs2446821	8:95128479	none	intergenic	RPL34P18: CDH17	T	G	3.00×10^{-5}	27
SNV	rs2513791	8:95128529	none	intergenic	RPL34P18: CDH17	A	C	3.00×10^{-5}	22
SNV	rs117579120	8:95159359	CDH17	intronic		A	C	3.78×10^{-5}	21
SNV	rs2251996	8:95160288	CDH17	intronic		C	G	1.61×10^{-5}	22
SNV	rs1672692	11:113945609	ZBTB16	intronic		A	G	4.26×10^{-5}	8

*SNV = Single Nucleotide Variant

** χ^2 P values for the DA association results for the DA subjects compared to the 1000 genomes database controls*** The total number of functional genomics datasets each SNP overlaps (see Supplemental Table E4 in this article's Online Repository at www.jacionline.org).

Table V. SNP Expression in luciferase assays

SNP	Comparison	P value	allele
rs11571559	C vs. pGL3	0.0014	non-risk
	T vs. pGL3	0.0325	risk
	C vs. T	0.2585	
rs11571563	G vs. pGL3	0.0038	risk
	T vs. pGL3	0.0026	non-risk
	G vs. T	0.9009	
rs75465959	A vs. pGL3	0.0453	risk
	G vs. pGL3	0.2518	non-risk
	A vs. G	0.3802	
rs2287231*	A vs. pGL3	0.9009	risk
	G vs. pGL3	0.0190	non-risk
	A vs. G	0.0138	
rs2446824*	A vs. pGL3	<0.0001	risk
	G vs. pGL3	0.0055	non-risk
	A vs. G	0.0103	
rs2446823	A vs. pGL3	<0.0001	non-risk
	C vs. pGL3	<0.0001	
	A vs. C	0.6319	
rs149630836	Del vs. pGL3	<0.0001	risk
	Ins vs. pGL3	0.0048	non-risk
	Del vs. Ins	0.1783	
rs2513789*	A vs. pGL3	0.3274	risk
	G vs. pGL3	0.0009	non-risk
	A vs. G	0.0055	
rs117579120	A vs. pGL3	<0.0001	risk
	C vs. pGL3	0.0003	non-risk
	A vs. C	0.7259	
rs2251996	C vs. pGL3	0.0003	risk
	G vs. pGL3	0.0065	non-risk
	C vs. G	0.3209	
rs1672692	A vs. pGL3	0.0002	risk
	G vs. pGL3	<0.0001	non-risk
	A vs. G	0.5028	
rs62084077	A vs. pGL3	0.0345	risk
	G vs. pGL3	0.0004	non-risk
	A vs. G	0.1134	
rs147978008-35bp	Del vs. pGL3	0.2358	risk
	Ins vs. pGL3	0.5707	Non-risk
	Del vs. Ins	0.4213	
rs11571537-35bp	C vs. pGL3	0.3074	risk
	T vs. pGL3	0.1703	Non-risk
	C vs. T	0.7650	
rs147978008-203bp	Del vs. pGL3	0.0005	risk
	Ins vs. pGL3	0.0016	non-risk
	Del vs. Ins	0.7091	
rs11571537-205bp*	C vs. pGL3	<0.0001	risk
	T vs. pGL3	0.0002	non-risk

	C vs. T	0.0257	
*genetic variants showing allele-dependent expression			

Figure Legends

Fig 1. Study design used to identify DA-associated regulatory variants

Fig 2. Allele-dependent protein-DNA interactions in A549 cells.

A. EMSA binding reactions of A549 nuclear extract proteins by oligonucleotide probes containing non-risk (N) or risk (R) alleles of genetic variants: (1) rs75465959, (2) rs1001304, (3) rs2446824, (4) rs11571537, (5) rs2513790, (6) rs2287231, (7) rs14978008, (8) rs2513788, (9) rs2513789. Variants 5 and 8 show stronger binding to the risk allele, variants 2, 4, 6, 7, and 9 show stronger binding to the non-risk allele, and genetic variants 1 and 3 show comparable risk and non-risk binding. The arrow points to a non-specific band that was seen in all the EMSA images (see supplemental Figure EI). **B.** Oligonucleotide controls for gels shown in Panel A.

Fig 3. Allele-dependent protein-DNA interactions in multiple cell lines.

EMSAs corresponding to the binding reactions seen for the non-risk (NR) and risk (R) alleles of SNP oligonucleotide probes rs147978008 (Lanes 1-6) and rs11571537 (Lanes 7-10) with BEAS 2B (Lanes 1,2), A549 (Lanes 3,4,7,8) or IMR-90 (Lanes 5,6,9,10) nuclear extracts (NE). The rs147978008 SNP shows non-risk allele-dependent binding for proteins in the nuclear extracts of all 3 cell lines. The rs11571537 also shows non-risk allele-dependent binding for proteins in A549 and IMR-90 nuclear extracts. B-BEAS 2B NE; A-A549 NE; I-IMR90 NE.

Fig 4. Histone H1.2 binds more strongly to the rs147978008 non-risk allele probe.

A. SDS-PAGE of DAPA purified rs147978008 binding proteins. Analysis of the DAPA eluates of A549 (Lanes 2 & 3) and BEAS 2B (Lanes 4 & 5) cells produced by binding to either the non-risk or risk allele of rs147978008. L1-molecular weight standards, L2-A549 extract/non-risk; L3-A549 extract/risk; L4-BEAS 2B/non-risk; L5-BEAS /risk. **B.** Western Blot of SDS-PAGE gel immunoblotted with rabbit polyclonal anti-H1.2. The lanes correspond to those shown in Panel A. The bands detected have an approximate molecular size of 29 kDa, consistent with the size of H1.2.

Fig 5. Luciferase reporter assays for genetic variants showing risk (2nd bar) and non-risk (3rd bar) gene expression for variants: **A;** rs2446824-35bp, **B;** rs2287231-35bp, **C;** 2513789-35bp and **E;** rs11571537-205bp, which show allele-dependent gene expression. **D;** rs 147978008-35bp. **F;** rs147978008-203bp, which shows increased driven by the deletion mutant (Del). A549 cells were transfected with luciferase reporter constructs containing luciferase reporter constructs for either the risk or the non-risk allele of each variant. Three independent experiments were performed and eight replicate samples were assayed in each experiment. Relative luciferase activity was calculated by taking the normalized luciferase activity of the vector (pGL3-Promoter) to be 1. Data represent the mean $\pm$ S.E. p values were determined by Student's t-test.

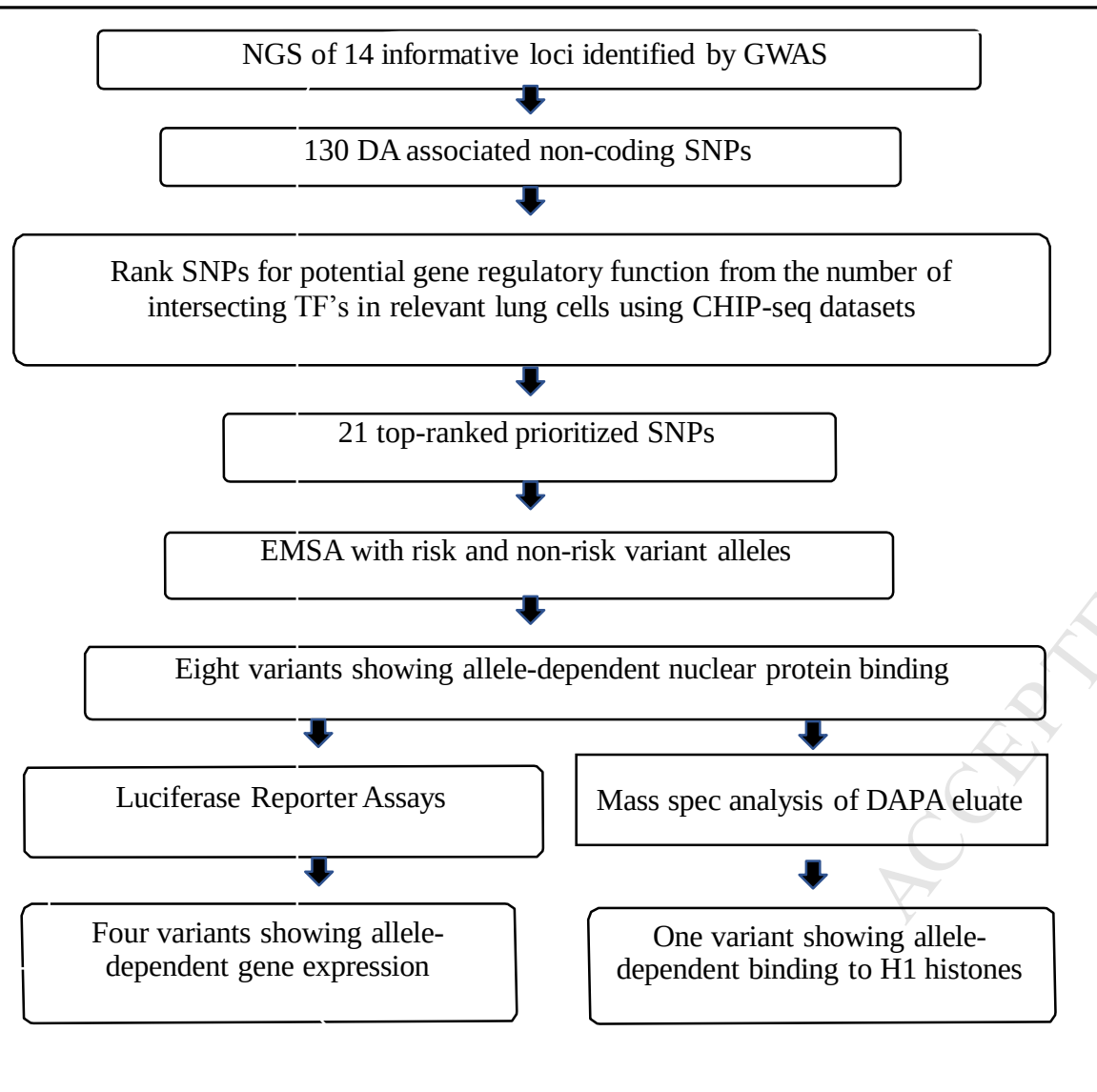
References

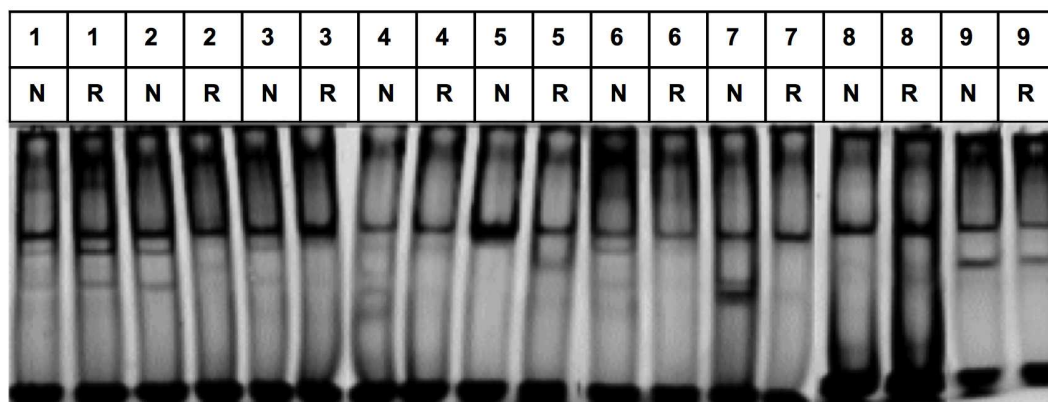
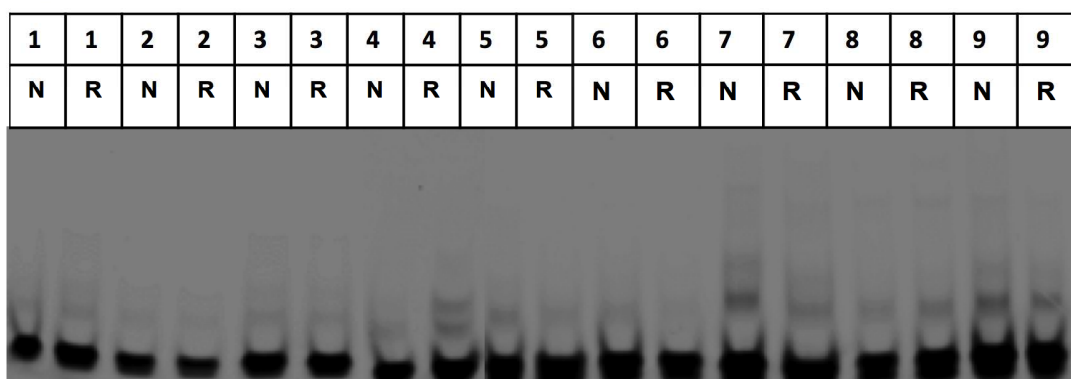
1. Meyer JD, Holt DL, Chen Y, Cherry NM, McDonald JC. SWORD '99: surveillance of work-related and occupational respiratory disease in the UK. *Occup Med (Lond)*. 2001;51(3):204-8.
2. Ribeiro M, Tarlo SM, Czyrka A, Vernich L, Luce CE, Liss GM. Diisocyanate and non-diisocyanate sensitizer-induced occupational asthma frequency during 2003 to 2007 in Ontario, Canada. *J Occup Environ Med*. 2014;56(9):1001-7.
3. Booth K, Cummings B, Karoly WJ, Mullins S, Robert WP, Spence M, et al. Measurements of airborne methylene diphenyl diisocyanate (MDI) concentration in the U.S. workplace. *J Occup Environ Hyg*. 2009;6(4):228-38.
4. Wisnewski AV, Stowe MH, Nerlinger A, Opare-Addo P, Decamp D, Kleinsmith CR, et al. Biomonitoring Hexamethylene diisocyanate (HDI) exposure based on serum levels of HDI-specific IgG. *Ann Occup Hyg*. 2012;56(8):901-10.
5. Jajosky RA, Harrison R, Reinisch F, Flattery J, Chan J, Tumpowsky C, et al. Surveillance of work-related asthma in selected U.S. states using surveillance guidelines for state health departments--California, Massachusetts, Michigan, and New Jersey, 1993-1995. *MMWR CDC Surveill Summ*. 1999;48(3):1-20.
6. Paggiaro PL, Innocenti A, Bacci E, Rossi O, Talini D. Specific bronchial reactivity to toluene diisocyanate: relationship with baseline clinical findings. *Thorax*. 1986;41(4):279-82.
7. Lummus ZL, Wisnewski AV, Bernstein DI. Pathogenesis and disease mechanisms of occupational asthma. *Immunol Allergy Clin North Am*. 2011;31(4):699-716, vi.
8. Sastre J, Vandenplas O, Park HS. Pathogenesis of occupational asthma. *Eur Respir J*. 2003;22(2):364-73.
9. Bernstein DI, Jolly A. Current diagnostic methods for diisocyanate induced occupational asthma. *Am J Ind Med*. 1999;36(4):459-68.
10. Campo P, Lummus ZL, Bernstein DI. Advances in methods used in evaluation of occupational asthma. *Curr Opin Pulm Med*. 2004;10(2):142-6.
11. Carino M, Aliani M, Licitra C, Sarno N, Ioli F. Death due to asthma at workplace in a diphenylmethane diisocyanate-sensitized subject. *Respiration*. 1997;64(1):111-3.
12. Piirila PL, Keskinen HM, Luukkonen R, Salo SP, Tuppurainen M, Nordman H. Work, unemployment and life satisfaction among patients with diisocyanate induced asthma--a prospective study. *J Occup Health*. 2005;47(2):112-8.
13. Yucesoy B, Johnson VJ, Lummus ZL, Kissling GE, Fluharty K, Gautrin D, et al. Genetic variants in antioxidant genes are associated with diisocyanate-induced asthma. *Toxicol Sci*. 2012;129(1):166-73.
14. Yucesoy B, Kashon ML, Johnson VJ, Lummus ZL, Fluharty K, Gautrin D, et al. Genetic variants in TNFalpha, TGFB1, PTGS1 and PTGS2 genes are associated with diisocyanate-induced asthma. *Journal of immunotoxicology*. 2015:1-8.
15. Mapp CE. The role of genetic factors in occupational asthma. *Eur Respir J*. 2003;22(1):173-8.
16. Bernstein DI, Kissling GE, Khurana Hershey G, Yucesoy B, Johnson VJ, Cartier A, et al. Hexamethylene diisocyanate asthma is associated with genetic polymorphisms of CD14, IL-13, and IL-4 receptor alpha. *J Allergy Clin Immunol*. 2011;128(2):418-20.
17. Bernstein DI, Kashon M, Lummus ZL, Johnson VJ, Fluharty K, Gautrin D, et al. CTNNA3 (alpha-catenin) gene variants are associated with diisocyanate asthma: a replication study in a Caucasian worker population. *Toxicol Sci*. 2013;131(1):242-6.
18. Yucesoy B, Johnson VJ, Lummus ZL, Kashon ML, Rao M, Bannerman-Thompson H, et al. Genetic variants in the major histocompatibility complex class I and class II genes are associated with diisocyanate-induced Asthma. *J Occup Environ Med*. 2014;56(4):382-7.

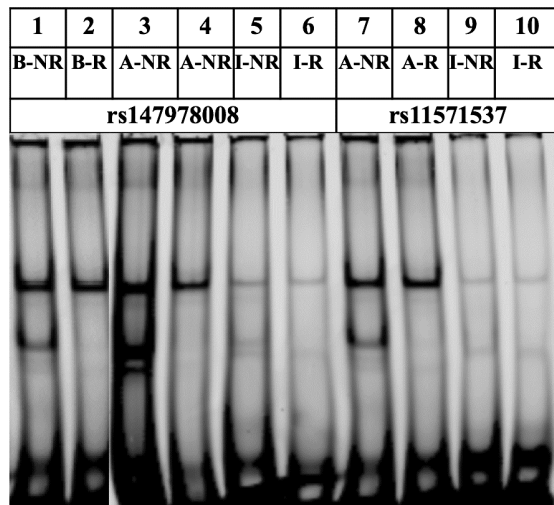
19. Kim SH, Cho BY, Park CS, Shin ES, Cho EY, Yang EM, et al. Alpha-T-catenin (CTNNA3) gene was identified as a risk variant for toluene diisocyanate-induced asthma by genome-wide association analysis. *Clin Exp Allergy*. 2009;39(2):203-12.
20. Yucesoy B, Kaufman KM, Lummus ZL, Weirauch MT, Zhang G, Cartier A, et al. Genome-Wide Association Study Identifies Novel Loci Associated With Diisocyanate-Induced Occupational Asthma. *Toxicol Sci*. 2015;146(1):192-201.
21. Yucesoy B, Kissling GE, Johnson VJ, Lummus ZL, Gautrin D, Cartier A, et al. N-Acetyltransferase 2 Genotypes Are Associated With Diisocyanate-Induced Asthma. *J Occup Environ Med*. 2015;57(12):1331-6.
22. McGeachie MJ, Wu AC, TSE SM, Clemmer GL, al. e. CTNNA3 and SEMA3D: promising loci for asthma exacerbation identified through multiple GWAS. *J Allergy Clin Immunol*. 2015;136(6):1503-10.
23. Miller DE, Patel ZH, Lu X, Lynch AT, Weirauch MT, Kottyan LC. Screening for Functional Non-coding Genetic Variants Using Electrophoretic Mobility Shift Assay (EMSA) and DNA-affinity Precipitation Assay (DAPA). *Journal of visualized experiments : JoVE*. 2016(114).
24. Chorley BN, Wang X, Campbell MR, Pittman GS, Noureddine MA, Bell DA. Discovery and verification of functional single nucleotide polymorphisms in regulatory genomic regions: current and developing technologies. *Mutat Res*. 2008;659(1-2):147-57.
25. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. *Nature*. 2015;526(7571):68-74.
26. Sastre J, Fernandez-Nieto M, Novalbos A, De Las Heras M, Cuesta J, Quirce S. Need for monitoring nonspecific bronchial hyperresponsiveness before and after isocyanate inhalation challenge. *Chest*. 2003;123(4):1276-9.
27. Cartier A, Malo J-L. Occupational challenge tests. In: Bernstein IL, Chan-Yeung M, Malo J-L, Bernstein DI, editors. *Asthma in the workplace*. Second ed. New York: Marcel Dekker, Inc.; 1999. p. 211-33.
28. Vandenplas O, Burge PS, Moscato G, Malo J-L. Functional Assessment. In: Malo J-L, Chan-Yeung M, Bernstein DI, editors. *Asthma in the Workplace*. 4 ed. Baco Raton: CRC Press; 2013. p. 133-7.
29. Tarlo SM, Balmes J, Balkissoon R, Beach J, Beckett W, Bernstein D, et al. Diagnosis and management of work-related asthma: American College Of Chest Physicians Consensus Statement. *Chest*. 2008;134(3 Suppl):1S-41S.
30. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754-60.
31. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297-303.
32. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012;489(7414):57-74.
33. Dreszer TR, Karolchik D, Zweig AS, Hinrichs AS, Raney BJ, Kuhn RM, et al. The UCSC Genome Browser database: extensions and updates 2011. *Nucleic Acids Res*. 2012;40(Database issue):D918-23.
34. Mei S, Qin Q, Wu Q, Sun H, Zheng R, Zang C, et al. Cistrome Data Browser: a data portal for ChIP-Seq and chromatin accessibility data in human and mouse. *Nucleic Acids Res*. 2017;45(D1):D658-D62.
35. Griffon A, Barbier Q, Dalino J, van Helden J, Spicuglia S, Ballester B. Integrative analysis of public ChIP-seq experiments reveals a complex multi-cell regulatory landscape. *Nucleic Acids Res*. 2015;43(4):e27.
36. Gilchrist M, Henderson WR, Jr., Clark AE, Simmons RM, Ye X, Smith KD, et al. Activating transcription factor 3 is a negative regulator of allergic pulmonary inflammation. *J Exp Med*. 2008;205(10):2349-57.

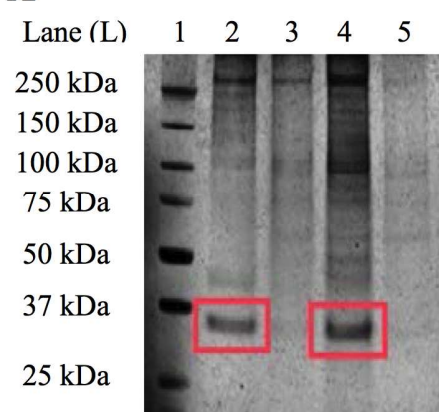
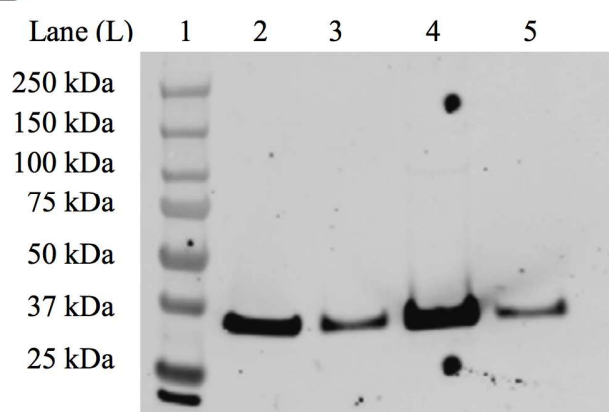
37. Roussel L, Robins S, Schachter A, Berube J, Hamid Q, Rousseau S. Steroids and extracellular signal-regulated kinase 1/2 activity suppress activating transcription factor 3 expression in patients with severe asthma. *J Allergy Clin Immunol.* 2011;127(6):1632-4.
38. Joos GF, Germonpre PR, Pauwels RA. Role of tachykinins in asthma. *Allergy.* 2000;55(4):321-37.
39. Mapp CE, Lucchini RE, Miotto D, Chitano P, Jovine L, Saetta M, et al. Immunization and challenge with toluene diisocyanate decrease tachykinin and calcitonin gene-related peptide immunoreactivity in guinea pig central airways. *Am J Respir Crit Care Med.* 1998;158(1):263-9.
40. Holden NS, Tacon CE. Principles and problems of the electrophoretic mobility shift assay. *J Pharmacol Toxicol Methods.* 2011;63(1):7-14.
41. Hellman LM, Fried MG. Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nat Protoc.* 2007;2(8):1849-61.
42. Wijeratne AB, Manning JR, Schultz Jel J, Greis KD. Quantitative phosphoproteomics using acetone-based peptide labeling: method evaluation and application to a cardiac ischemia/reperfusion model. *J Proteome Res.* 2013;12(10):4268-79.
43. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. *Nature.* 2009;461(7265):747-53.
44. Ober C, Yao TC. The genetics of asthma and allergic disease: a 21st century perspective. *Immunol Rev.* 2011;242(1):10-30.
45. Handstad T, Rye M, Drablos F, P. S. Cell-type specificity of ChIP-predicted transcription factor binding sites. *BioMed Central Genomics.* 2012;13.
46. Aizawa M, Fukuda M. Small GTPase Rab2B and Its Specific Binding Protein Golgi-associated Rab2B Interactor-like 4 (GARI-L4) Regulate Golgi Morphology. *J Biol Chem.* 2015;290(36):22250-61.
47. Hergeth SP, Schneider R. The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. *EMBO Rep.* 2015;16(11):1439-53.
48. Bednar J, Hamiche A, Dimitrov S. H1-nucleosome interactions and their functional implications. *Biochim Biophys Acta.* 2016;1859(3):436-43.
49. Hai T, Wolford CC, Chang YS. ATF3, a hub of the cellular adaptive-response network, in the pathogenesis of diseases: is modulation of inflammation a unifying component? *Gene Expr.* 2010;15(1):1-11.
50. Kwon J-W, Kwon H-K, Shin H-J, Choi Y-M, Anwaee M, Choi S. Activating transcription factor 3 represses inflammatory responses by binding to the p65 subunit of NF- κ B. *Sci Rep.* 2015.
51. Hertz AL, Bender AT, Smith KC, Gilchrist M, Amieux PS, Aderem A, et al. Elevated cyclic AMP and PDE4 inhibition induce chemokine expression in human monocyte-derived macrophages. *Proc Natl Acad Sci U S A.* 2009;106(51):21978-83.
52. De Nardo D. Toll-like receptors: Activation, signalling and transcriptional modulation. *Cytokine.* 2015;74(2):181-9.
53. Shan Y, Akram A, Amatullah H, Zhou DY, Gali PL, Maron-Gutierrez T, et al. ATF3 protects pulmonary resident cells from acute and ventilator-induced lung injury by preventing Nrf2 degradation. *Antioxid Redox Signal.* 2015;22(8):651-68.
54. Fan F, Jin S, Amundson SA, Tong T, Fan W, Zhao H, et al. ATF3 induction following DNA damage is regulated by distinct signaling pathways and over-expression of ATF3 protein suppresses cells growth. *Oncogene.* 2002;21(49):7488-96.
55. Cui H, Guo M, Xu D, Ding ZC, Zhou G, Ding HF, et al. The stress-responsive gene ATF3 regulates the histone acetyltransferase Tip60. *Nat Commun.* 2015;6:6752.
56. Zhao J, Li X, Guo M, Yu J, Yan C. The common stress responsive transcription factor ATF3 binds genomic sites enriched with p300 and H3K27ac for transcriptional regulation. *BMC Genomics.* 2016;17:335.

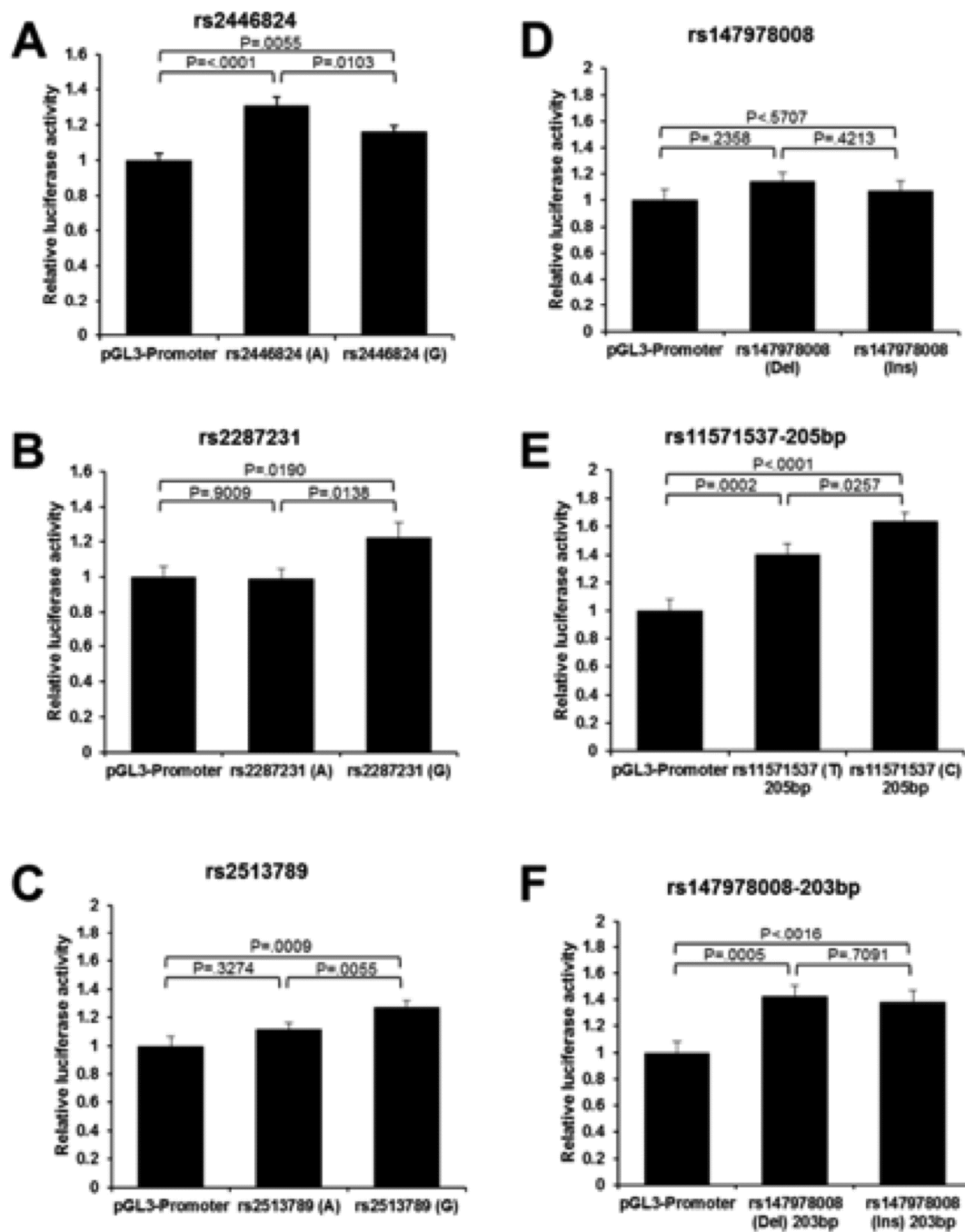
57. Douglas SD, Leeman SE. Neurokinin-1 receptor: functional significance in the immune system in reference to selected infections and inflammation. *Ann N Y Acad Sci.* 2011;1217:83-95.
58. Scheerens H, Buckley TL, Davidse EM, Garssen J, Nijkamp FP, Van Loveren H. Toluene diisocyanate-induced in vitro tracheal hyperreactivity in the mouse. *Am J Respir Crit Care Med.* 1996;154(4 Pt 1):858-65.
59. Scheerens H, Buckley TL, Muis T, Van Loveren H, Nijkamp FP. The involvement of sensory neuropeptides in toluene diisocyanate-induced tracheal hyperreactivity in the mouse airways. *Br J Pharmacol.* 1996;119(8):1665-71.



A**B**



A**B**



Supplemental Table EIV. Association Data for Significant SNPs ($p < 1 \times 10^{-3}$)

Chr	Position	Identifier	Gene	Strand	Left gene	Right gene	Reference	GWAS Chi- Squared P	Chi- Squared P	Chi- Squared FDR	Odds Ratio (Minor Allele)	OR Lower Confidenc e Bound (Minor)	OR Upper Confidence Bound (Minor)	Call Rate (Cases)	Call Rate (Controls)	Minor Allele	Minor Allele Freq. (Cases)	Minor Allele Freq. (Controls)
1	112529254	rs34115852	KCND3	-			del	NA	5.84E-04	0.142	3.11	1.58	6.13	1.000000	1.000000	A	0.0944444	0.0324232
1	212732686	rs1001304	none	+	Loc101929565**	ATF3	A*	NA	3.15E-04	0.105	0.45	0.29	0.70	1.000000	1.000000	A*	0.1500000	0.2832765
1	212745765	rs138203798	ATF3	+			AG	NA	5.81E-04	0.143	1.89	1.31	2.73	0.844444	0.972696	del	0.6184211	0.4614035
1	212764930	rs17019411	ATF3	+			G	NA	3.42E-04	0.106	3.29	1.66	6.53	1.000000	1.000000	A	0.0944444	0.0307167
1	212765432	rs77911820	ATF3	+			C	NA	1.93E-04	0.074	3.49	1.74	6.99	1.000000	1.000000	T	0.0944444	0.0290102
1	212765681	rs116604459	ATF3	+			T	NA	3.42E-04	0.106	3.29	1.66	6.53	1.000000	1.000000	C	0.0944444	0.0307167
1	212770271	rs72756369	ATF3	+			T	NA	3.42E-04	0.106	3.29	1.66	6.53	1.000000	1.000000	A	0.0944444	0.0307167
1	212787025	rs11571537	ATF3	+			T	NA	4.78E-04	0.131	3.46	1.66	7.23	1.000000	1.000000	C	0.0833333	0.0255973
1	212791460	rs8192659	ATF3	+			A	NA	3.78E-04	0.110	6.77	2.01	22.75	1.000000	1.000000	G	0.0444444	0.0068259
1	212793050	rs1803173	ATF3	+			A	NA	3.78E-04	0.110	6.77	2.01	22.75	1.000000	1.000000	G	0.0444444	0.0068259
1	212794550	rs11571559	ATF3	+			C	NA	8.77E-05	0.039	7.66	2.33	25.17	1.000000	1.000000	T	0.0500000	0.0068259
1	212794873	rs11571563	none	+	ATF3	FAM71A	T	NA	2.80E-04	0.103	6.12	2.02	18.49	1.000000	1.000000	G	0.0500000	0.0085324
1	212794997	rs74138575	none	+	ATF3	FAM71A	G	NA	8.77E-05	0.039	7.66	2.33	25.17	1.000000	1.000000	A	0.0500000	0.0068259
1	212795196	rs75465959	none	+	ATF3	FAM71A	G	NA	8.77E-05	0.039	7.66	2.33	25.17	1.000000	1.000000	A	0.0500000	0.0068259
1	212795692	rs11119989	none	+	ATF3	FAM71A	C	NA	8.77E-05	0.039	7.66	2.33	25.17	1.000000	1.000000	T	0.0500000	0.0068259
1	212796008	rs147978008	FAM71A	+			ACA	NA	8.77E-05	0.039	7.66	2.33	25.17	1.000000	1.000000	del	0.0500000	0.0068259
1	212796487	rs74138578	FAM71A	+			G	NA	3.78E-04	0.110	6.77	2.01	22.75	1.000000	1.000000	A	0.0444444	0.0068259
1	212799039	rs74138590	FAM71A	+			G	NA	3.78E-04	0.110	6.77	2.01	22.75	1.000000	1.000000	A	0.0444444	0.0068259
1	212799163	rs74138591	FAM71A	+			C	NA	3.41E-04	0.109	6.85	2.04	23.02	0.988889	1.000000	A	0.0449438	0.0068259
1	212804008	rs141048429	none	+	FAM71A	Loc 100129948	A	NA	1.05E-04	0.046	9.04	2.37	34.44	1.000000	1.000000	T	0.0444444	0.0051195
1	212808004	rs56412520	none	+	FAM71A	Loc 100129948	T	NA	2.22E-05	0.032	10.23	2.74	38.20	1.000000	1.000000	G	0.0500000	0.0051195
1	212816729	rs17019510	none	+	FAM71A	Loc 100129948	A	NA	4.60E-06	0.048	11.43	3.11	42.01	1.000000	1.000000	G	0.0555556	0.0051195
2	31121117	rs55971416	none	-	CAPN13	GALNT14	C	NA	4.06E-04	0.114	2.77	1.54	4.98	1.000000	1.000000	G	0.1222222	0.0477816
2	31213834	rs77918726	GALNT14	-			C	NA	5.20E-04	0.134	16.71	1.94	144.02	1.000000	1.000000	T	0.0277778	0.0017065
2	31214987	rs112012072	GALNT14	-			C	NA	2.97E-04	0.100	?	?	?	1.000000	1.000000	G	0.0222222	0.0000000
2	31215035	rs116366946	GALNT14	-			T	NA	2.97E-04	0.100	?	?	?	1.000000	1.000000	A	0.0222222	0.0000000
2	75448760	rs2287230	none	-	TACR1	GAPDHP59	G	NA	6.30E-06	0.039	2.18	1.55	3.08	1.000000	1.000000	A	0.6222222	0.4300341
2	75448766	rs1468899	none	-	TACR1	GAPDHP59	G	NA	4.30E-05	0.028	0.48	0.33	0.68	1.000000	1.000000	G	0.3000000	0.4726962
2	75449129	rs2287231	none	-	TACR1	GAPDHP59	G	NA	4.30E-05	0.028	0.48	0.33	0.68	1.000000	1.000000	G	0.3000000	0.4726962
2	75449986	rs4853124	none	-	TACR1	GAPDHP59	G	2.14E-06	6.30E-06	0.039	2.18	1.55	3.08	1.000000	1.000000	A	0.6222222	0.4300341
2	75451047	rs6717216	none	-	TACR1	GAPDHP59	A	3.05E-04	4.30E-05	0.028	0.48	0.33	0.68	1.000000	1.000000	A	0.3000000	0.4726962
2	75451823	rs7598278	none	-	TACR1	GAPDHP59	G	NA	1.02E-04	0.045	0.51	0.36	0.72	1.000000	1.000000	G	0.3500000	0.5153584
2	75452466	rs6546958	none	-	TACR1	GAPDHP59	T	NA	4.30E-05	0.028	0.48	0.33	0.68	1.000000	1.000000	T	0.3000000	0.4726962
2	75453100	rs2193406	none	-	TACR1	GAPDHP59	G	NA	1.59E-04	0.064	0.50	0.35	0.72	0.977778	1.000000	G	0.3068182	0.4675768
2	75453589	rs7605539	none	-	TACR1	GAPDHP59	G	NA	4.30E-05	0.028	0.48	0.33	0.68	1.000000	1.000000	G	0.3000000	0.4726962
2	75457783	rs4852364	none	-	TACR1	GAPDHP59	T	5.92E-07	7.13E-06	0.022	0.45	0.31	0.64	1.000000	1.000000	T	0.3111111	0.5017065
2	75458427	rs1108105	none	-	TACR1	GAPDHP59	T	NA	7.13E-06	0.022	0.45	0.31	0.64	1.000000	1.000000	T	0.3111111	0.5017065
2	75460555	rs12328696	none	-	TACR1	GAPDHP59	T	5.93E-07	7.13E-06	0.022	0.45	0.31	0.64	1.000000	1.000000	T	0.3111111	0.5017065
2	75460880	rs13012482	none	-	TACR1	GAPDHP59	G	NA	7.14E-05	0.034	1.97	1.41	2.77	1.000000	1.000000	C	0.5111111	0.3464164
2	75461150	rs13017325	none	-	TACR1	GAPDHP59	A	NA	7.14E-05	0.034	1.97	1.41	2.77	1.000000	1.000000	G	0.5111111	0.3464164
2	75461177	rs13017967	none	-	TACR1	GAPDHP59	G	3.46E-07	7.14E-05	0.034	1.97	1.41	2.77	1.000000	1.000000	A	0.5111111	0.3464164
2	75461645	rs11680861	none	-	TACR1	GAPDHP59	C	NA	7.14E-05	0.034	1.97	1.41	2.77	1.000000	1.000000	T	0.5111111	0.3464164
2	75461679	rs11680895	none	-	TACR1	GAPDHP59	C	NA	7.14E-05	0.034	1.97	1.41	2.77	1.000000	1.000000	T	0.5111111	0.3464164
2	75462004	rs13018989	none	-	TACR1	GAPDHP59	A	NA	7.14E-05	0.034	1.97	1.41	2.77	1.000000	1.000000	G	0.5111111	0.3464164

Supplemental Table EIV. Association Data for Significant SNPs ($p < 1 \times 10^{-3}$)

Chr	Position	Identifier	Gene	Strand	Left gene	Right gene	Reference	GWAS Chi- Squared P	Chi- Squared P	Chi- Squared FDR	Odds Ratio (Minor Allele)	OR Lower Confidenc e Bound (Minor)	OR Upper Confidence Bound (Minor)	Call Rate (Cases)	Call Rate (Controls)	Minor Allele	Minor Allele Freq. (Cases)	Minor Allele Freq. (Controls)
2	75473916	rs7588010	none	–	TACR1	GAPDHP59	C	4.08E-09	6.75E-06	0.030	2.16	1.54	3.03	1.000000	1.000000	A	0.5611111	0.3720137
2	75474706	rs7585802	none	–	TACR1	GAPDHP59	T	NA	6.67E-06	0.035	2.16	1.54	3.03	1.000000	1.000000	G	0.5666667	0.3771331
2	75476771	rs6729324	none	–	TACR1	GAPDHP59	C	NA	1.22E-04	0.052	1.99	1.40	2.84	1.000000	0.993174	A	0.3888889	0.2422680
2	75476848	rs6757865	none	–	TACR1	GAPDHP59	A	NA	4.43E-04	0.123	1.88	1.32	2.68	1.000000	1.000000	G	0.3777778	0.2440273
2	75477143	rs2160376	none	–	TACR1	GAPDHP59	G	NA	1.92E-04	0.074	1.96	1.37	2.80	0.988889	0.989761	C	0.3820225	0.2396552
2	75477345	rs2110940	none	–	TACR1	GAPDHP59	T	NA	4.86E-04	0.129	1.88	1.31	2.69	0.988889	0.996587	C	0.3764045	0.2431507
2	75478007	rs7583283	none	–	TACR1	GAPDHP59	C	NA	6.21E-04	0.148	1.85	1.30	2.65	1.000000	1.000000	T	0.3722222	0.2423208
2	75478858	rs142821201	none	–	TACR1	GAPDHP59	del	NA	5.82E-04	0.142	1.84	1.30	2.62	0.977778	0.982935	ACAACCTTT	0.4147727	0.2777778
2	75478874	rs11679944	none	–	TACR1	GAPDHP59	A	NA	3.16E-04	0.103	1.88	1.33	2.66	1.000000	1.000000	T	0.4222222	0.2798635
2	75478876	rs11679945	none	–	TACR1	GAPDHP59	A	NA	3.16E-04	0.103	1.88	1.33	2.66	1.000000	1.000000	G	0.4222222	0.2798635
2	75478974	rs11691207	none	–	TACR1	GAPDHP59	G	NA	5.68E-04	0.141	1.86	1.30	2.66	1.000000	0.996587	A	0.3722222	0.2414384
2	75479741	rs11681360	none	–	TACR1	GAPDHP59	A	NA	3.79E-04	0.109	1.86	1.32	2.64	1.000000	1.000000	G	0.4222222	0.2815700
2	75484310	rs2110934	none	–	TACR1	GAPDHP59	A	NA	3.13E-06	0.098	2.23	1.59	3.14	1.000000	1.000000	G	0.4833333	0.2952218
2	75485335	rs35879047	none	–	TACR1	GAPDHP59	del	NA	2.20E-05	0.038	2.07	1.48	2.91	1.000000	1.000000	A	0.4888889	0.3156997
2	75487574	rs2216103	none	–	TACR1	GAPDHP59	A	NA	2.04E-04	0.077	1.90	1.35	2.67	1.000000	1.000000	G	0.4611111	0.3105802
7	18431378	rs79527202	HDAC9	+			C	NA	6.14E-04	0.148	2.99	1.56	5.75	1.000000	1.000000	A	0.1000000	0.0358362
7	18970676	rs6651065	HDAC9	+			A	NA	2.97E-04	0.100	?	?	?	1.000000	1.000000	T	0.0222222	0.0000000
7	18971455	rs73322158	HDAC9	+			A	NA	5.20E-04	0.134	16.71	1.94	144.02	1.000000	1.000000	G	0.0277778	0.0017065
7	18972725	rs2892905	HDAC9	+			C	NA	2.97E-04	0.100	?	?	?	1.000000	1.000000	A	0.0222222	0.0000000
7	19036230	rs35409100	HDAC9	+			T	NA	2.97E-04	0.100	?	?	?	1.000000	1.000000	C	0.0222222	0.0000000
7	19063083	rs34917720	none	+	HDAC9	TWIST1	G	NA	5.53E-04	0.138	10.07	2.01	50.34	1.000000	1.000000	del	0.0333333	0.0034130
8	95126377	rs2446825	none	–	RPL34P18	CDH17	G	NA	1.42E-04	0.058	2.12	1.43	3.15	1.000000	1.000000	T	0.2833333	0.1569966
8	95126478	rs2513784	none	–	RPL34P18	CDH17	T	7.24E-06	1.82E-04	0.071	2.10	1.42	3.10	1.000000	1.000000	C	0.2833333	0.1587031
8	95126888	rs2513786	none	–	RPL34P18	CDH17	G	NA	3.00E-05	0.025	2.38	1.57	3.59	1.000000	1.000000	A	0.2555556	0.1262799
8	95127080	rs2513787	none	–	RPL34P18	CDH17	C	NA	3.00E-05	0.025	2.38	1.57	3.59	1.000000	1.000000	A	0.2555556	0.1262799
8	95127574	rs2446824	none	–	RPL34P18	CDH17	G*	4.64E-07	3.00E-05	0.025	2.38	1.57	3.59	1.000000	1.000000	A*	0.2555556	0.1262799
8	95127612	rs2446823	none	–	RPL34P18	CDH17	A*	2.23E-04	3.00E-05	0.025	2.38	1.57	3.59	1.000000	1.000000	C*	0.2555556	0.1262799
8	95127836	rs149630836	none	–	RPL34P18	CDH17	TCAGTAG	NA	4.82E-05	0.031	2.32	1.53	3.51	1.000000	0.993174	del	0.2555556	0.1288660
8	95127896	rs2513788	none	–	RPL34P18	CDH17	T	NA	3.00E-05	0.025	2.38	1.57	3.59	1.000000	1.000000	G	0.2555556	0.1262799
8	95128113	rs2513789	none	–	RPL34P18	CDH17	G	NA	3.00E-05	0.025	2.38	1.57	3.59	1.000000	1.000000	A	0.2555556	0.1262799
8	95128147	rs2513790	none	–	RPL34P18	CDH17	T	NA	3.00E-05	0.025	2.38	1.57	3.59	1.000000	1.000000	C	0.2555556	0.1262799
8	95128479	rs2446821	none	–	RPL34P18	CDH17	C*	NA	3.00E-05	0.025	2.38	1.57	3.59	1.000000	1.000000	A*	0.2555556	0.1262799
8	95128529	rs2513791	none	–	RPL34P18	CDH17	C	4.64E-07	3.00E-05	0.025	2.38	1.57	3.59	1.000000	1.000000	A	0.2555556	0.1262799
8	95129174	rs2513792	none	–	RPL34P18	CDH17	G	NA	3.00E-05	0.025	2.38	1.57	3.59	1.000000	1.000000	A	0.2555556	0.1262799
8	95129259	rs2446819	none	–	RPL34P18	CDH17	A	NA	3.00E-05	0.025	2.38	1.57	3.59	1.000000	1.000000	G	0.2555556	0.1262799
8	95129261	rs2513793	none	–	RPL34P18	CDH17	A	NA	3.00E-05	0.025	2.38	1.57	3.59	1.000000	1.000000	C	0.2555556	0.1262799
8	95129402	rs2446818	none	–	RPL34P18	CDH17	G	NA	3.00E-05	0.025	2.38	1.57	3.59	1.000000	1.000000	T	0.2555556	0.1262799
8	95133647	rs2513795	none	–	RPL34P18	CDH17	C	NA	2.30E-04	0.086	2.22	1.44	3.42	0.933333	0.996587	T	0.2380952	0.1232877
8	95143138	rs2513797***	CDH17	–			A	3.39E-07	2.20E-05	0.033	2.41	1.59	3.65	1.000000	1.000000	G	0.2555556	0.1245734
8	95143344	rs146033050	CDH17	–			A	NA	3.78E-04	0.110	6.77	2.01	22.75	1.000000	1.000000	G	0.0444444	0.0068259
8	95152576	rs2513804	CDH17	–			T	NA	2.81E-04	0.102	1.87	1.33	2.62	1.000000	0.996587	C	0.4833333	0.3339041
8	95156137	rs2514795	CDH17	–			G	NA	2.20E-05	0.033	2.41	1.59	3.65	1.000000	1.000000	C	0.2555556	0.1245734
8	95156145	rs2514796	CDH17	–			A	NA	2.20E-05	0.033	2.41	1.59	3.65	1.000000	1.000000	G	0.2555556	0.1245734
8	95157476	rs2340023	CDH17	–			C	2.47E-07	1.61E-05	0.034	2.45	1.62	3.71	1.000000	1.000000	T	0.2555556	0.1228669
8	95159359	rs117579120	CDH17	–			C	NA	3.78E-04	0.110	6.77	2.01	22.75	1.000000	1.000000	A	0.0444444	0.0068259

Supplemental Table EIV. Association Data for Significant SNPs ($p < 1 \times 10^{-3}$)

Chr	Position	Identifier	Gene	Strand	Left gene	Right gene	Reference	GWAS Chi- Squared P	Chi- Squared P	Chi- Squared FDR	Odds Ratio (Minor Allele)	OR Lower Confidenc e Bound (Minor)	OR Upper Confidence Bound (Minor)	Call Rate (Cases)	Call Rate (Controls)	Minor Allele	Minor Allele Freq. (Cases)	Minor Allele Freq. (Controls)
8	95160288	rs2251996	CDH17	-			G	NA	1.61E-05	0.034	2.45	1.62	3.71	1.000000	1.000000	C	0.2555556	0.1228669
8	95167247	rs2514805	CDH17	-			A	1.22E-08	3.44E-06	0.054	2.60	1.72	3.92	1.000000	1.000000	G	0.2666667	0.1228669
8	95170648	rs2514806	CDH17	-			G	NA	1.80E-05	0.035	2.46	1.61	3.74	1.000000	1.000000	A	0.2500000	0.1194539
10	67829697	rs189879649	CTNNA3	-			T	NA	5.53E-04	0.138	10.07	2.01	50.34	1.000000	1.000000	A	0.0333333	0.0034130
10	68325826	rs60424434	CTNNA3	-			T	NA	4.06E-04	0.114	2.77	1.54	4.98	1.000000	1.000000	C	0.1222222	0.0477816
10	68733954	rs574762398	CTNNA3	-			T	NA	5.17E-05	0.029	?	?	?	1.000000	1.000000	C	0.0277778	0.0000000
10	68790758	rs112756470	CTNNA3	-			G	NA	4.81E-04	0.129	7.86	2.01	30.73	1.000000	1.000000	A	0.0388889	0.0051195
10	68885274	rs550635526	ZBTB16	+			A	NA	5.17E-05	0.028	?	?	?	1.000000	1.000000	T	0.0277778	0.0000000
10	69056053	rs143425155	ZBTB16	+			T	NA	5.17E-05	0.028	?	?	?	1.000000	1.000000	C	0.0277778	0.0000000
10	69332802	rs575404339	ZBTB16	+			G	NA	3.84E-05	0.030	?	?	?	0.966667	1.000000	A	0.0287356	0.0000000
10	69450634	rs77090539	ZBTB16	+			G	NA	4.85E-05	0.030	?	?	?	0.755556	0.952218	C	0.0294118	0.0000000
11	113939539	rs1784684	ZBTB16	+			A	6.20E-07	1.45E-05	0.035	2.28	1.56	3.34	1.000000	1.000000	G	0.3222222	0.1723549
11	113939671	rs1791804	ZBTB16	+			C	NA	1.45E-05	0.035	2.28	1.56	3.34	1.000000	1.000000	T	0.3222222	0.1723549
11	113939783	rs1784683	ZBTB16	+			A	NA	1.45E-05	0.035	2.28	1.56	3.34	1.000000	1.000000	G	0.3222222	0.1723549
11	113940469	rs3018331	ZBTB16	+			C	2.83E-07	2.19E-05	0.040	2.25	1.54	3.30	1.000000	1.000000	A	0.3166667	0.1706485
11	113941756	rs2846629	ZBTB16	+			G	5.72E-07	2.48E-05	0.034	2.18	1.51	3.15	1.000000	1.000000	G	0.3500000	0.1979522
11	113942151	rs1672691	ZBTB16	+			C	1.68E-07	6.92E-05	0.037	2.10	1.45	3.04	1.000000	1.000000	T	0.3388889	0.1962457
11	113942191	rs2735199	ZBTB16	+			C	7.05E-07	3.17E-05	0.025	2.16	1.50	3.12	1.000000	1.000000	C	0.3500000	0.1996587
11	113942767	rs2852793	ZBTB16	+			A	NA	3.17E-05	0.025	2.16	1.50	3.12	1.000000	1.000000	A	0.3500000	0.1996587
11	113945609	rs1672692	ZBTB16	+			G	1.42E-07	4.26E-05	0.031	2.19	1.50	3.22	1.000000	1.000000	A	0.3111111	0.1706485
11	113945804	rs2846625	ZBTB16	+			T	NA	2.85E-05	0.037	2.23	1.52	3.25	1.000000	1.000000	T	0.3166667	0.1723549
11	113946484	rs1614059	ZBTB16	+			C	NA	4.26E-05	0.031	2.19	1.50	3.22	1.000000	1.000000	T	0.3111111	0.1706485
11	113949232	rs1784692	ZBTB16	+			T	1.17E-06	4.40E-04	0.123	1.95	1.34	2.85	1.000000	1.000000	C	0.3111111	0.1877133
11	113949740	rs1791800	ZBTB16	+			G	NA	3.26E-04	0.105	2.01	1.37	2.97	0.955556	0.989761	A	0.3081395	0.1810345
11	113954925	rs2852800	ZBTB16	+			C	NA	4.17E-05	0.032	2.04	1.45	2.88	1.000000	1.000000	C	0.4444444	0.2815700
11	113966184	rs139196790	ZBTB16	+			G	NA	1.65E-04	0.065	?	?	?	0.922222	1.000000	A	0.0240964	0.0000000
11	114026639	rs17116541	ZBTB16	+			C	NA	1.64E-04	0.066	3.71	1.80	7.67	1.000000	1.000000	A	0.0888889	0.0255973
12	277708	rs75748509	IQSEC3	+			T	NA	4.81E-04	0.129	7.86	2.01	30.73	1.000000	1.000000	A	0.0388889	0.0051195
12	293366	rs77318956	none	-	IQSEC3	SLC6A12	A	NA	2.97E-04	0.100	?	?	?	1.000000	1.000000	G	0.0222222	0.0000000
12	294021	rs113856433	none	-	IQSEC3	SLC6A12	C	NA	1.10E-04	0.047	?	?	?	0.866667	0.989761	T	0.0256410	0.0000000
12	295008	rs3890993	none	-	SLC6A12	SLC6A12	T	NA	2.97E-04	0.100	?	?	?	1.000000	1.000000	G	0.0222222	0.0000000
12	295367	rs647001	none	-	SLC6A12	SLC6A12	C	NA	5.17E-05	0.028	?	?	?	1.000000	1.000000	T	0.0277778	0.0000000
12	295370	rs529198	none	-	SLC6A12	SLC6A12	G	NA	5.17E-05	0.028	?	?	?	1.000000	1.000000	T	0.0277778	0.0000000
12	295525	rs478520	none	-	SLC6A12	SLC6A12	C	NA	5.17E-05	0.028	?	?	?	1.000000	1.000000	T	0.0277778	0.0000000
12	327925	rs200539477	none	-	SLC6A12	SLC6A12	TCCT	NA	6.01E-05	0.032	8.20	2.43	27.63	0.811111	0.972696	del	0.0547945	0.0070175
14	33659784	rs142880613	NPAS3	+			C	NA	5.53E-04	0.138	10.07	2.01	50.34	1.000000	1.000000	T	0.0333333	0.0034130
14	33660446	rs56110420	NPAS3	+			T	NA	5.53E-04	0.138	10.07	2.01	50.34	1.000000	1.000000	C	0.0333333	0.0034130
14	33861512	rs117306110	NPAS3	+			G	NA	3.78E-04	0.110	6.77	2.01	22.75	1.000000	1.000000	A	0.0444444	0.0068259
17	65367094	rs8073599	none	+	PSMD12	LOC101928074	T	6.06E-06	1.36E-04	0.057	0.46	0.31	0.69	1.000000	1.000000	T	0.1944444	0.3447099
17	65397518	rs187061404	PITPNC1	+			A	NA	5.20E-04	0.134	16.71	1.94	144.02	1.000000	1.000000	G	0.0277778	0.0017065
17	65399074	rs191465321	PITPNC1	+			C	NA	4.79E-04	0.130	16.91	1.96	145.70	0.988889	1.000000	T	0.0280899	0.0017065

* Bolded alleles are named from the reverse sequence to those shown in the NCBI dbSNP database

** Uncharacterized RNA gene

*** Synonymous coding SNP that doesn't change the amino acid sequence

Supplemental Table E2. Overlap Between Non-Coding Variants and Transcription Factor Chip-Seq Datasets in disease-relevant cell types.

SNP	Gene	Gene Strand	Transcription Factor: Cell line(s)				
			1	2	3	4-5	6-7
rs1001304	OC101929565	+	P130: IMR90	RBL2: IMR90	RUNX1: SKNO-1	POLR2A: A549; IMR90	
rs72756369	ATF3	+	P130: IMR90				
rs11571537	ATF3	+	SPI1: IMR90	POLR2A: A549; IMR90			
rs11571559	downstream ATF3	+	SPI1: IMR90	RUNX1: CD34+	POLR2A: A549; IMR90		
rs11571563	downstream ATF3	+	RUNX1: CD34+ cells	POLR2A: A549; IMR90			
rs74138575	downstream ATF3	+	RUNX1: CD34+ cells; SKNO-1	POLR2A: A549; IMR90			
rs75465959	downstream ATF3	+	RUNX1: CD34+ cells; SKNO-1	POLR2A: A549; IMR90	MAFK: IMR90		
rs147978008	downstream FAM71A	+	SPI1: IMR90	POLR2A: A549; IMR90			
rs17019510	downstream FAM71A	+	USF1: A549				
rs2287231	TACR1	-	SPI1: IMR90				
rs2446824	downstream CDH17	-	SPI1: IMR90	CEBPB: A549; IMR90			

Supplemental Table E2, continued

rs2446823	downstream CDH17	-	SPI1: IMR90	CEBPB: A549; IMR90			
rs149630836	downstream CDH17	-	RELA: A549	SPI1:IMR90	EP300: IMR90	CEBPB: A549; IMR90	FOSL2: A549
rs2513788	downstream CDH17	-	RELA: A549	SPI1:IMR90	EP300: IMR90	CEBPB: A549, IMR90	FOSL2: A549
rs2513789	downstream CDH17	—	SPI1:IMR90	CEBPB: A549; IMR90	FOSL2: A549_		
rs2513790	downstream CDH17	-	SPI1:IMR90	P130: IMR90			
rs2446821	downstream CDH17	-	P130:IMR90				
rs2513791	downstream CDH17	-	P130:IMR90				
rs117579120	inside CDH17	—	SPI1:IMR90	EP300: IMR90			
rs2251996	inside CDH17	-	RUNX1: CD34+ cells				
rs1672692	inside ZBTB16	-	RUNX1: CD34+ cells				
rs62084077	inside PITPNC1	+	SPI1: IMR90	RUNX1: CD34+ cells; SKNO-1	CTCF: A549	NR3C1: A549 FOSL2: A549,	POLR2A: IMR90, RUNX3: GM12878

Supplemental Table E3. Oligonucleotides used for preparation of luciferase reporter constructs

SNP	Gene	Allele	Forward sequence (5'-3')	Reverse sequence (5'-3')
rs1001304	LOC101929565	A	CCATCTGGTTAACCAGAGGGCTTCCAGGAGC	TCGAGCTCCTGGAAGCCCTCTTTGGTTAACAGATGGGTAC
		G	CCATCTGGTTAACCAGAGGGCTTCCAGGAGC	TCGAGCTCCTGGAAGCCCTCTTTGGTTAACAGATGGGTAC
rs72756369	ATF3	A	CAGTGTGGAGCCTCTTACTGTCTTGTTCATTCC	TCGAGGAATGAACAAGACAGTAAGAGGCTCCACACTGGTAC
		T	CAGTGTGGAGCCTCTTCTGTCTTGTTCATTCC	TCGAGGAATGAACAAGACAGTAAGAGGCTCCACACTGGTAC
rs11571537	ATF3	C	CACCACAGATCCCCGGCTGAGAGGAATGCCAC	TCGAGTGGGCATTCTCTCAGCCGGGATCTGTGGTGGTAC
		T	CACCACAGATCCCCGGTGTGAGAGGAATGCCAC	TCGAGTGGGCATTCTCTCAGCCGGGATCTGTGGTGGTAC
rs11571559	ATF3	C	CTGTTTCATGTCTAAACGTGTGTCTCATTGTTT	TCGAGAACAATGAGACACACGTTTAGACATGAAACAGGTAC
		T	CTGTTTCATGTCTAAATGTGTGTCTCATTGTTT	TCGAGAACAATGAGACACACATTTAGACATGAAACAGGTAC
rs11571563	ATF3	G	CTTGACCTCATGTATGTTCTTAAATGATCC	TCGAGGATCATTTAAAGGAACCATACATGAGGTCAAGGTAC
		T	CTTGACCTCATGTATGTTCTTAAATGATCC	TCGAGGATCATTTAAAGGAACCATACATGAGGTCAAGGTAC
rs74138575	ATF3	A	CAGCACTTCTTTCAGGAGATAATTTATGTTCCC	TCGAGGGAACATAAATTATCTCCTGAAAGAGGTGCTGGTAC
		G	CAGCACTTCTTTCAGGGGATAATTTATGTTCCC	TCGAGGGAACATAAATTATCCTGAAAGAGGTGCTGGTAC
rs75465959	ATF3	A	CTGATGTGTATGGTAAATGTAATGAAGGTCTC	TCGAGAGACCTTCATTACATTTACCATACACATCAGGTAC
		G	CTGATGTGTATGGTAAATGTAATGAAGGTCTC	TCGAGAGACCTTCATTACACTTACCATACACATCAGGTAC
rs147978008	FAM71A	-	CTCCCATTTAATCCTCACAACACTGCTACCCCC	TCGAGGGGGGTAGCAGTTGTGTGAGGATTAATGGGAGGTAC
	ACA		CTCCCATTTAATCCTCACAACACTGCTACCCCC	TCGAGGGGGGTAGCAGTTGTGTGAGGATTAATGGGAGGTAC
rs17019510	FAM71A	A	CACACCACATGCCAGATAGAGATGTACCTGTC	TCGAGACAGGTACATCTCTATCTGGGCATGTGGTGTGGTAC
		G	CACACCACATGCCAGGTAGAGATGTACCTGTC	TCGAGACAGGTACATCTCTACCTGGGCATGTGGTGTGGTAC
rs2287231	TACR1	A	CCACTTGGCTGATGAAAGACAACATTGGCAGGC	TCGAGCCTGCCAATGTTGTCTTTCATCAGCCAAGTGGGTAC
		G	CCACTTGGCTGATGAAAGACAACATTGGCAGGC	TCGAGCCTGCCAATGTTGTCTTTCATCAGCCAAGTGGGTAC
rs2446824	CDH17	A	CCATTTTTTTCATTGCAACATTGGGCAGATGC	TCGAGCATCTGCCAAATGTTGCAATGGAAGGATTAATGGGAGGTAC
		G	CCATTTTTTTCATTGCGACATTGGGCAGATGC	TCGAGCATCTGCCAAATGTCGCAATGGAAGGATTAATGGGAGGTAC
rs2446823	CDH17	A	CCAGAAGCAAAAGTTCATCTGACTGAGGGAAAC	TCGAGTTCCCTCAGTCAGATGAACCTTTGCTTCTGGGTAC
		C	CCAGAAGCAAAAGTTCCTCTGACTGAGGGAAAC	TCGAGTTCCCTCAGTCAGAGGAACCTTTGCTTCTGGGTAC
rs149630836	CDH17	-	CGAATTGCCTACTACTCTCAGTAGCCACCATC	TCGAGATGGTGGGCTACTGAGATAGTAGGCAATTCGGTAC
	TCAGTAG		CGAATTGCCTACTACTCTCAGTAGTCAGTAGCCACCATC	TCGAGATGGTGGGCTACTGACTACTGAGATAGTAGGCAATTCGGTAC
rs2513788	CDH17	G	CGCCCATCCTCTGTGTTATGGGCCAGATGC	TCGAGCATCTGGGCCATAAACAAGGAGGATGGGCGGTAC
		T	CGCCCATCCTCTGTGTTATGGGCCAGATGC	TCGAGCATCTGGGCCATAAACAAGGAGGATGGGCGGTAC
rs2513789	CDH17	A	CTGTTTTTTCATTGTCAATGTTATCTAAGTCAC	TCGAGTGACTTAGATAAATGACAAATGAAAAACAGGTAC
		G	CTGTTTTTTCATTGTGCGATGTTATCTAAGTCAC	TCGAGTGACTTAGATAAATGACAAATGAAAAACAGGTAC
rs2513790	CDH17	C	CTTCAAGCACATACTCAAAGAATCACTTGAC	TCGAGTCAAGTGATTCTTTGAGTATGTGCTTGAAGGTAC
		T	CTTCAAGCACATACTTAAAGAATCACTTGAC	TCGAGTCAAGTGATTCTTTAAGTATGTGCTTGAAGGTAC
rs2446821	CDH17	A	CGGTCTCTCTCAGAAAAACATCAGCTTCCAC	TCGAGTGGGAAGCTGATGTTTTCTGAAGGAGGACCGGTAC
		C	CGGTCTCTCTCAGAACACATCAGCTTCCAC	TCGAGTGGGAAGCTGATGTTGTTCTGAAGGAGGACCGGTAC
rs2513791	CDH17	A	CAGTCAGCTCAGGCCAACATAACAATAACCAC	TCGAGTGGTATTTTGTATGTTGGCCTGAGCTGACTGGTAC
		C	CAGTCAGCTCAGGCCAACATAACAATAACCAC	TCGAGTGGTATTTTGTATGTTGGCCTGAGCTGACTGGTAC
rs117579120	CDH17	A	CGAGAACTCTTGGGAAACATCCAGAGTGAAGC	TCGAGCTTCACTCTGGATGTTCCCAAGAGTTTCTCGGTAC
		C	CGAGAACTCTTGGGACACATCCAGAGTGAAGC	TCGAGCTTCACTCTGGATGTGTCCTCAAGAGTTTCTCGGTAC
rs2251996	CDH17	C	CAAGGAGCAAGTGGTCTACGTGAACAGGTG	TCGAGACCTGTTACGTAGAGCACCCTGCTCTTGGTAC
		G	CAAGGAGCAAGTGGTCTACGTGAACAGGTG	TCGAGACCTGTTACGTAGACCACCTGCTCTTGGTAC
rs1672692	ZBTB16	A	CCTGTGCATATGTAAGAGGGAATCATTGTGGC	TCGAGCCACAATGATTCCTTTTACATATGACAAGGGTAC
		G	CCTGTGCATATGTAAGAGGGAATCATTGTGGC	TCGAGCCACAATGATTCCTCTTACATATGACAAGGGTAC
rs11571537*	ATF3	C/T	GGCCGGTACCACTGGGTCTCGTGAGGAT (PCR primer)	GGCCCTCGAGGTTGGTCCCAAGGGACAAT (PCR primer)
rs147978008*	FAM71A	-/ACA	GGCCGGTACCACTGCTTTCAGGGAGGAA (PCR primer)	GGCCCTCGAGACTGTGAGACCTCGAGCAA (PCR primer)

Luciferase reporter constructs were cloned using the pGL3-promoter vector as a backbone. Fragments for insertion into this backbone consisted of two complementary oligonucleotides, designed so that annealing of a sense and antisense oligonucleotide formed a 31 bp with a KpnI site at the 5' end and a XhoI site at the 3' end. The DNA fragments were then cloned into the pGL3-promoter vector upstream of the SV40 promoter. Two constructs were produced for each SNP (for the R and NR alleles).

*Larger luciferase reporter constructs for SNP rs147978008 and SNP rs11571537 were made by PCR amplification of DNA fragments containing the SNPs (203 bp for rs147978008 and 205 bp for rs11571537). All constructs were verified by DNA sequencing.

Table EIV. Overlapping functional genomics datasets: Sources used

This spreadsheet contains the results of the prioritization of genetic variants based on their overlap with functional genomics datasets. Datasets relevant to gene regulation were gathered from multiple sources (see below), and restricted to those performed in lung-derived cell lines. The "Summary" worksheet lists the total number of datasets each variant intersects, broken down by data source. The "All dataset intersections" worksheet lists all datasets intersected by each variant.

Dataset key	Data type	PMID
Cistrome_ChIP	ChIP-seq	22495751
DNaseClusters	DNase-seq	22955619
DNaseI_Duke	DNase-seq	23193274
FANTOM5_enhancer_organ	Enhancer expression	24670763
Misc_ChIP	ChIP-seq	23193258
Pazar_ChIP	ChIP-seq	18971253
ReMap_ChIP	ChIP-seq	25477382
Roadmapepigenomics_ActiveChromatin	HMM (model)	25693563
Roadmapepigenomics_Histone_narrow	ChIP-seq	25693563
TxnFactorChIPV3	ChIP-seq	23193274

ll lines and tissues

ACCEPTED MANUSCRIPT

Table EIV. Overlapping functional genomics datasets: Summary

ID	Total score	Cistrome_ChIP	DNaseClusters	DNaseI_Duke	FANTOM5_enhancer_organ	Misc_ChIP	Pazar_ChIP	ReMap_ChIP	Roadmapepigenomics_ActiveChromatin	Roadmapepigenomics_Histone_narrow	TxnFactorChIPV3
rs1001304	50	0	0	1	0	0	4	1	4	37	3
rs149630836	35	3	0	2	0	0	0	0	5	22	3
rs2446823	32	1	0	0	0	0	0	0	5	24	2
rs2446824	32	1	0	0	0	0	0	0	5	24	2
rs2513788	32	3	1	2	0	0	0	0	5	18	3
rs13022860	29	0	0	0	0	0	0	0	2	27	0
rs2446821	27	0	0	0	0	0	3	0	3	21	0
rs72756369	27	0	0	0	0	0	2	0	3	22	0
rs2513790	23	1	0	0	0	0	1	0	4	17	0
rs2251996	22	0	0	0	0	0	0	2	1	19	0
rs2513791	22	0	0	0	0	0	3	0	3	16	0
rs117579120	21	2	0	0	0	0	0	0	3	16	0
rs2287231	17	1	0	1	0	0	0	0	2	13	0
rs4853124	17	0	0	0	0	0	0	0	2	15	0
rs2513789	16	1	0	0	0	0	0	0	4	8	3
rs7588010	15	1	0	0	0	0	0	0	1	13	0
rs11571537	12	1	0	0	0	0	0	0	2	4	5
rs1468899	10	0	0	0	0	0	0	0	0	10	0
rs2287230	10	0	0	0	0	0	0	0	0	10	0
rs1672692	8	0	0	0	1	1	0	0	1	5	0
rs58895081	8	0	0	0	0	0	0	0	3	5	0
rs59038579	8	0	0	0	0	0	0	0	3	5	0
rs17019510	7	0	0	0	0	0	0	0	0	4	3
rs1784692	7	0	0	0	0	0	0	0	1	6	0
rs2446819	7	0	0	0	0	0	0	0	0	7	0
rs2513784	7	0	0	0	0	0	0	0	0	7	0
rs2513792	7	0	0	0	0	0	0	0	0	7	0
rs2513793	7	0	0	0	0	0	0	0	0	7	0
rs56412520	7	0	0	0	0	0	0	0	1	6	0
rs13018989	6	0	0	0	0	0	0	0	1	5	0
rs74138575	6	0	1	0	0	1	0	2	0	0	2
rs75465959	6	0	0	0	0	1	0	2	0	0	3
rs11571563	5	0	1	0	0	1	0	1	0	0	2
rs17116541	5	0	0	0	0	0	0	0	1	4	0
rs1784684	5	0	0	0	0	0	0	0	1	4	0
rs2513786	5	0	0	0	0	0	0	0	0	5	0
rs2846625	5	0	0	0	1	1	0	0	1	2	0
rs11571559	4	1	0	0	0	1	0	0	0	0	2
rs1791800	4	0	0	0	0	0	0	0	1	3	0
rs1803173	4	0	0	0	0	0	0	0	0	4	0
rs11119989	3	0	0	0	0	0	0	0	0	1	2
rs141048429	3	0	0	0	0	0	0	2	0	1	0
rs147978008	3	1	0	0	0	0	0	0	0	0	2
rs17019411	3	0	0	0	0	0	0	0	0	3	0
rs191465321	3	0	0	0	0	0	0	0	1	2	0
rs2513787	3	0	0	0	0	0	0	0	1	2	0
rs13017325	2	0	0	0	0	0	0	0	0	2	0
rs13017967	2	0	0	0	0	0	0	0	0	2	0

rs139196790	2	0	0	0	0	0	0	0	1	1	0
rs1784683	2	0	0	0	0	0	0	0	0	2	0
rs1791804	2	0	0	0	0	0	0	0	0	2	0
rs2160376	2	0	0	0	0	0	0	0	0	2	0
rs2216103	2	0	0	0	0	0	0	0	0	2	0
rs2446818	2	0	0	0	0	0	0	0	0	2	0
rs2852800	2	0	0	0	0	0	0	0	1	1	0
rs34917720	2	0	0	0	0	0	0	0	0	2	0
rs35879047	2	0	0	0	0	0	0	0	0	2	0
rs529198	2	0	0	0	0	0	0	0	0	2	0
rs647001	2	0	0	0	0	0	0	0	0	2	0
rs6729324	2	0	0	0	0	0	0	0	0	2	0
rs74138590	2	0	0	0	0	0	0	0	0	2	0
rs74138591	2	0	0	0	0	0	0	0	0	2	0
rs11679944	1	0	0	0	0	0	0	0	0	1	0
rs11679945	1	0	0	0	0	0	0	0	0	1	0
rs11691207	1	0	0	0	0	0	0	0	0	1	0
rs142821201	1	0	0	0	0	0	0	0	0	1	0
rs142880613	1	0	0	0	0	0	0	0	0	1	0
rs1614059	1	0	0	0	0	0	0	0	1	0	0
rs187061404	1	0	0	0	0	0	0	0	0	1	0
rs2110940	1	0	0	0	0	0	0	0	0	1	0
rs2513795	1	0	0	0	0	0	0	0	0	1	0
rs2514806	1	0	0	0	0	0	0	0	0	1	0
rs2852793	1	0	0	0	0	0	0	0	0	1	0
rs34115852	1	0	0	0	0	0	0	0	0	1	0
rs35409100	1	0	0	0	0	0	0	0	0	1	0
rs55971416	1	0	0	0	0	0	0	0	0	1	0
rs575404339	1	0	0	0	0	0	0	0	0	1	0
rs6546958	1	0	0	0	0	0	0	0	0	1	0
rs6717216	1	0	0	0	0	0	0	0	0	1	0
rs74423934	1	0	0	0	0	0	0	0	0	1	0
rs75748509	1	0	0	0	0	0	0	0	0	1	0
rs7583283	1	0	0	0	0	0	0	0	0	1	0
rs77911820	1	0	0	0	0	0	0	0	0	1	0
rs8192659	1	0	0	0	0	0	0	0	0	1	0

ACCEPTED MANUSCRIPT									
Table EIV. Overlapping functional genomics datasets: All dataset interactions									
snp_chr	snp_st	snp_ed	snp_rsid	cell	dataset	peak_chr	peak_st	peak_ed	peak_element
chr1	212732686	212732686	rs1001304	A549	DNaseI_Duke	chr1	212731677	212732775	.
chr1	212732686	212732686	rs1001304	IMR90_Quiescent	Pazar_Chip	chr1	212732403	212732747	P130
chr1	212732686	212732686	rs1001304	IMR90_shRbQuiescent	Pazar_Chip	chr1	212732455	212732844	P130
chr1	212732686	212732686	rs1001304	IMR90_shRbSenescent	Pazar_Chip	chr1	212732417	212732863	P130
chr1	212732686	212732686	rs1001304	IMR90_Senescent	Pazar_Chip	chr1	212732403	212732747	RBL2
chr1	212732686	212732686	rs1001304	skno1_leukemic	ReMap_ChIP	chr1	212731941	212733027	RUNX1
chr1	212732686	212732686	rs1001304	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_ActiveChromatin	chr1	212732600	212733000	2_TssAFlnk
chr1	212732686	212732686	rs1001304	Fetal_Lung	Roadmapepigenomics_ActiveChromatin	chr1	212732400	212732800	2_TssAFlnk
chr1	212732686	212732686	rs1001304	IMR90	Roadmapepigenomics_ActiveChromatin	chr1	212732400	212732800	2_TssAFlnk
chr1	212732686	212732686	rs1001304	NHLF_Lung_Fibroblasts	Roadmapepigenomics_ActiveChromatin	chr1	212732600	212732800	2_TssAFlnk
chr1	212732686	212732686	rs1001304	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr1	212731769	212733008	H2A.Z
chr1	212732686	212732686	rs1001304	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212732470	212732761	H2A.Z
chr1	212732686	212732686	rs1001304	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr1	212732446	212732713	H2A.Z
chr1	212732686	212732686	rs1001304	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212732505	212732821	H2AK5ac
chr1	212732686	212732686	rs1001304	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212732485	212733060	H2BK120ac
chr1	212732686	212732686	rs1001304	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212732492	212733345	H2BK12ac
chr1	212732686	212732686	rs1001304	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212732496	212732806	H2BK15ac
chr1	212732686	212732686	rs1001304	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212732498	212733410	H2BK20ac
chr1	212732686	212732686	rs1001304	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212732497	212732761	H2BK5ac
chr1	212732686	212732686	rs1001304	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212732503	212732786	H3K14ac
chr1	212732686	212732686	rs1001304	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212732466	212733010	H3K18ac
chr1	212732686	212732686	rs1001304	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212732526	212732760	H3K23ac
chr1	212732686	212732686	rs1001304	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr1	212731853	212732880	H3K27ac
chr1	212732686	212732686	rs1001304	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212732467	212732758	H3K27ac
chr1	212732686	212732686	rs1001304	Lung	Roadmapepigenomics_Histone_narrow	chr1	212732420	212732939	H3K27ac
chr1	212732686	212732686	rs1001304	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr1	212731464	212733333	H3K27ac
chr1	212732686	212732686	rs1001304	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212732484	212733107	H3K4ac
chr1	212732686	212732686	rs1001304	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr1	212732491	212733097	H3K4me1
chr1	212732686	212732686	rs1001304	Fetal_Lung	Roadmapepigenomics_Histone_narrow	chr1	212732487	212732871	H3K4me1
chr1	212732686	212732686	rs1001304	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212732502	212732788	H3K4me1
chr1	212732686	212732686	rs1001304	Lung	Roadmapepigenomics_Histone_narrow	chr1	212732602	212732814	H3K4me1
chr1	212732686	212732686	rs1001304	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr1	212732583	212732743	H3K4me1
chr1	212732686	212732686	rs1001304	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr1	212731576	212732994	H3K4me2
chr1	212732686	212732686	rs1001304	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212732465	212732770	H3K4me2
chr1	212732686	212732686	rs1001304	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr1	212732468	212732687	H3K4me2
chr1	212732686	212732686	rs1001304	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr1	212731415	212732950	H3K4me3
chr1	212732686	212732686	rs1001304	Fetal_Lung	Roadmapepigenomics_Histone_narrow	chr1	212731403	212732800	H3K4me3
chr1	212732686	212732686	rs1001304	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212731422	212732706	H3K4me3
chr1	212732686	212732686	rs1001304	Lung	Roadmapepigenomics_Histone_narrow	chr1	212732550	212732690	H3K4me3
chr1	212732686	212732686	rs1001304	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212732532	212732726	H3K56ac
chr1	212732686	212732686	rs1001304	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr1	212731756	212732944	H3K9ac
chr1	212732686	212732686	rs1001304	Fetal_Lung	Roadmapepigenomics_Histone_narrow	chr1	212732403	212732931	H3K9ac
chr1	212732686	212732686	rs1001304	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212732491	212732781	H3K9ac
chr1	212732686	212732686	rs1001304	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr1	212731485	212732692	H3K9ac
chr1	212732686	212732686	rs1001304	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212732462	212733092	H4K5ac
chr1	212732686	212732686	rs1001304	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212732473	212732829	H4K8ac
chr1	212732686	212732686	rs1001304	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212732493	212733413	H4K91ac
chr1	212732686	212732686	rs1001304	A549+DEX_100nM	TxnFactorChIPV3	chr1	212731321	212732859	POLR2A
chr1	212732686	212732686	rs1001304	A549+EtOH_0.02pct	TxnFactorChIPV3	chr1	212731321	212732859	POLR2A
chr1	212732686	212732686	rs1001304	IMR90	TxnFactorChIPV3	chr1	212731321	212732859	POLR2A
chr1	212795692	212795692	rs11119989	Lung	Roadmapepigenomics_Histone_narrow	chr1	212795685	212795866	H3K27ac
chr1	212795692	212795692	rs11119989	A549	TxnFactorChIPV3	chr1	212794247	212796320	POLR2A
chr1	212795692	212795692	rs11119989	IMR90	TxnFactorChIPV3	chr1	212794247	212796320	POLR2A
chr1	212787025	212787025	rs11571537	IMR90	Cistrome_ChIP	chr1	212786962	212787139	SPI1
chr1	212787025	212787025	rs11571537	Lung	Roadmapepigenomics_ActiveChromatin	chr1	212787000	212787600	7_Enh
chr1	212787025	212787025	rs11571537	NHLF_Lung_Fibroblasts	Roadmapepigenomics_ActiveChromatin	chr1	212786200	212787200	7_Enh
chr1	212787025	212787025	rs11571537	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr1	212786961	212787239	H3K27ac
chr1	212787025	212787025	rs11571537	Lung	Roadmapepigenomics_Histone_narrow	chr1	212786919	212787134	H3K4me1
chr1	212787025	212787025	rs11571537	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212787003	212787190	H3K79me2
chr1	212787025	212787025	rs11571537	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr1	212786983	212787821	H3K79me2
chr1	212787025	212787025	rs11571537	A549	TxnFactorChIPV3	chr1	212780052	212787859	POLR2A
chr1	212787025	212787025	rs11571537	A549	TxnFactorChIPV3	chr1	212780052	212787859	POLR2A
chr1	212787025	212787025	rs11571537	A549+DEX_100nM	TxnFactorChIPV3	chr1	212780052	212787859	POLR2A
chr1	212787025	212787025	rs11571537	A549+EtOH_0.02pct	TxnFactorChIPV3	chr1	212780052	212787859	POLR2A
chr1	212787025	212787025	rs11571537	IMR90	TxnFactorChIPV3	chr1	212780052	212787859	POLR2A
chr1	212794550	212794550	rs11571559	IMR90	Cistrome_ChIP	chr1	212794493	212794639	SPI1
chr1	212794550	212794550	rs11571559	CD34+	Misc_ChIP	chr1	212794103	212795547	RUNX1
chr1	212794550	212794550	rs11571559	A549	TxnFactorChIPV3	chr1	212794247	212796320	POLR2A
chr1	212794550	212794550	rs11571559	IMR90	TxnFactorChIPV3	chr1	212794247	212796320	POLR2A
chr1	212794873	212794873	rs11571563	A549	DNaseClusters	chr1	212794860	212795010	.
chr1	212794873	212794873	rs11571563	CD34+	Misc_ChIP	chr1	212794103	212795547	RUNX1
chr1	212794873	212794873	rs11571563	cd34_nont8_21	ReMap_ChIP	chr1	212794861	212795311	RUNX1
chr1	212794873	212794873	rs11571563	A549	TxnFactorChIPV3	chr1	212794247	212796320	POLR2A
chr1	212794873	212794873	rs11571563	IMR90	TxnFactorChIPV3	chr1	212794247	212796320	POLR2A
chr2	75478874	75478874	rs11679944	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr2	75478746	75479010	H3K9me3
chr2	75478876	75478876	rs11679945	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr2	75478746	75479010	H3K9me3
chr2	75478974	75478974	rs11691207	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr2	75478746	75479010	H3K9me3
chr8	95159359	95159359	rs117579120	IMR90	Cistrome_ChIP	chr8	95159307	95159632	EP300
chr8	95159359	95159359	rs117579120	IMR90	Cistrome_ChIP	chr8	95158856	95159656	SPI1
chr8	95159359	95159359	rs117579120	Fetal_Lung	Roadmapepigenomics_ActiveChromatin	chr8	95158800	95160000	7_Enh
chr8	95159359	95159359	rs117579120	IMR90	Roadmapepigenomics_ActiveChromatin	chr8	95158200	95160600	7_Enh
chr8	95159359	95159359	rs117579120	NHLF_Lung_Fibroblasts	Roadmapepigenomics_ActiveChromatin	chr8	95159000	95159800	7_Enh
chr8	95159359	95159359	rs117579120	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr8	95159257	95159680	H2A.Z
chr8	95159359	95159359	rs117579120	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95158901	95159756	H2BK120ac
chr8	95159359	95159359	rs117579120	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95159331	95159472	H2BK15ac
chr8	95159359	95159359	rs117579120	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95158995	95159770	H2BK20ac

ACCEPTED MANUSCRIPT											
chr8	95159359	95159359	rs117579120	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95158851	95159949	H2BK5ac		
chr8	95159359	95159359	rs117579120	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95159313	95159500	H3K14ac		
chr8	95159359	95159359	rs117579120	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95158613	95159932	H3K18ac		
chr8	95159359	95159359	rs117579120	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95158095	95160597	H3K27ac		
chr8	95159359	95159359	rs117579120	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr8	95159225	95159575	H3K27ac		
chr8	95159359	95159359	rs117579120	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95159285	95159474	H3K4ac		
chr8	95159359	95159359	rs117579120	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95159292	95159754	H3K4me1		
chr8	95159359	95159359	rs117579120	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95158848	95159940	H3K4me2		
chr8	95159359	95159359	rs117579120	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr8	95159299	95159591	H3K4me2		
chr8	95159359	95159359	rs117579120	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95159313	95159745	H4K5ac		
chr8	95159359	95159359	rs117579120	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95159263	95159497	H4K8ac		
chr8	95159359	95159359	rs117579120	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95159272	95159496	H4K91ac		
chr2	75461150	75461150	rs13017325	Fetal_Lung	Roadmapepigenomics_Histone_narrow	chr2	75461028	75461238	H3K4me1		
chr2	75461150	75461150	rs13017325	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75461092	75461402	H4K91ac		
chr2	75461177	75461177	rs13017967	Fetal_Lung	Roadmapepigenomics_Histone_narrow	chr2	75461028	75461238	H3K4me1		
chr2	75461177	75461177	rs13017967	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75461092	75461402	H4K91ac		
chr2	75462004	75462004	rs13018989	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_ActiveChromatin	chr2	75461800	75462600	7_Enh		
chr2	75462004	75462004	rs13018989	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75461941	75462108	H2BK120ac		
chr2	75462004	75462004	rs13018989	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75461846	75462091	H2BK15ac		
chr2	75462004	75462004	rs13018989	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75461942	75462091	H2BK20ac		
chr2	75462004	75462004	rs13018989	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75461956	75462072	H3K4ac		
chr2	75462004	75462004	rs13018989	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr2	75461739	75462028	H3K4me1		
chr2	75467382	75467382	rs13022860	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_ActiveChromatin	chr2	75466000	75467400	1_TssA		
chr2	75467382	75467382	rs13022860	IMR90	Roadmapepigenomics_ActiveChromatin	chr2	75466200	75467400	1_TssA		
chr2	75467382	75467382	rs13022860	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75467319	75467448	H2AK5ac		
chr2	75467382	75467382	rs13022860	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75467310	75467455	H2AK9ac		
chr2	75467382	75467382	rs13022860	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75467323	75468950	H2BK120ac		
chr2	75467382	75467382	rs13022860	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75467272	75467926	H2BK12ac		
chr2	75467382	75467382	rs13022860	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75467328	75467586	H2BK15ac		
chr2	75467382	75467382	rs13022860	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75467306	75468275	H2BK20ac		
chr2	75467382	75467382	rs13022860	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75467312	75468123	H2BK5ac		
chr2	75467382	75467382	rs13022860	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75467301	75467509	H3K14ac		
chr2	75467382	75467382	rs13022860	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75467233	75468250	H3K18ac		
chr2	75467382	75467382	rs13022860	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75467329	75467503	H3K23ac		
chr2	75467382	75467382	rs13022860	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr2	75464313	75469071	H3K27ac		
chr2	75467382	75467382	rs13022860	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75467221	75467920	H3K27ac		
chr2	75467382	75467382	rs13022860	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr2	75465527	75467384	H3K27ac		
chr2	75467382	75467382	rs13022860	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75467266	75468307	H3K4ac		
chr2	75467382	75467382	rs13022860	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr2	75467264	75468646	H3K4me1		
chr2	75467382	75467382	rs13022860	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75467322	75468200	H2BK15ac		
chr2	75467382	75467382	rs13022860	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr2	75465582	75467552	H3K4me2		
chr2	75467382	75467382	rs13022860	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75467199	75467526	H3K4me2		
chr2	75467382	75467382	rs13022860	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr2	75465624	75468883	H3K4me3		
chr2	75467382	75467382	rs13022860	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75467322	75467507	H3K4me3		
chr2	75467382	75467382	rs13022860	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75467298	75467424	H3K79me1		
chr2	75467382	75467382	rs13022860	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr2	75467131	75468869	H3K79me2		
chr2	75467382	75467382	rs13022860	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75467296	75468213	H3K79me2		
chr2	75467382	75467382	rs13022860	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr2	75464440	75468204	H3K9ac		
chr2	75467382	75467382	rs13022860	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75467302	75468909	H4K5ac		
chr2	75467382	75467382	rs13022860	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75467298	75468241	H4K8ac		
chr2	75467382	75467382	rs13022860	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75467299	75468274	H4K91ac		
chr11	113966184	113966184	rs139196790	Lung	Roadmapepigenomics_ActiveChromatin	chr11	113942800	113966400	6_EnhG		
chr11	113966184	113966184	rs139196790	Lung	Roadmapepigenomics_Histone_narrow	chr11	113966167	113966326	H3K4me1		
chr1	212804008	212804008	rs141048429	cd34_nont8_21	ReMap_ChIP	chr1	212803300	212804115	RUNX1		
chr1	212804008	212804008	rs141048429	sem	ReMap_ChIP	chr1	212803401	212804071	RUNX1		
chr1	212804008	212804008	rs141048429	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr1	212803712	212804051	H3K27me3		
chr2	75478857	75478857	rs142821201	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr2	75478746	75479010	H3K9me3		
chr14	33659784	33659784	rs142880613	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr14	33659584	33660168	H3K27me3		
chr2	75448766	75448766	rs1468899	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448497	75448798	H2AK5ac		
chr2	75448766	75448766	rs1468899	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448462	75448843	H2BK120ac		
chr2	75448766	75448766	rs1468899	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448450	75448812	H2BK12ac		
chr2	75448766	75448766	rs1468899	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448471	75448780	H2BK15ac		
chr2	75448766	75448766	rs1468899	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448492	75448806	H2BK20ac		
chr2	75448766	75448766	rs1468899	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448666	75448786	H3K18ac		
chr2	75448766	75448766	rs1468899	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448644	75448791	H3K4ac		
chr2	75448766	75448766	rs1468899	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448634	75448775	H4K5ac		
chr2	75448766	75448766	rs1468899	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448659	75448806	H4K8ac		
chr2	75448766	75448766	rs1468899	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448608	75448820	H4K91ac		
chr1	212796008	212796010	rs147978008	IMR90	Cistrome_ChIP	chr1	212795777	212796102	SP1		
chr1	212796008	212796010	rs147978008	A549	TxnFactorChIPV3	chr1	212794247	212796320	POLR2A		
chr1	212796008	212796010	rs147978008	IMR90	TxnFactorChIPV3	chr1	212794247	212796320	POLR2A		
chr8	95127836	95127842	rs149630836	IMR90	Cistrome_ChIP	chr8	95127764	95128000	EP300		
chr8	95127836	95127842	rs149630836	A549	Cistrome_ChIP	chr8	95127830	95128012	RELA		
chr8	95127836	95127842	rs149630836	IMR90	Cistrome_ChIP	chr8	95127737	95128151	SP1		
chr8	95127836	95127842	rs149630836	A549	DNaseI_Duke	chr8	95127800	95128025	.		
chr8	95127836	95127842	rs149630836	IMR90	DNaseI_Duke	chr8	95127772	95128096	.		
chr8	95127836	95127842	rs149630836	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_ActiveChromatin	chr8	95127400	95128600	7_Enh		
chr8	95127836	95127842	rs149630836	Fetal_Lung	Roadmapepigenomics_ActiveChromatin	chr8	95127200	95128000	7_Enh		
chr8	95127836	95127842	rs149630836	IMR90	Roadmapepigenomics_ActiveChromatin	chr8	95127000	95128600	7_Enh		
chr8	95127836	95127842	rs149630836	Lung	Roadmapepigenomics_ActiveChromatin	chr8	95127200	95128400	7_Enh		
chr8	95127836	95127842	rs149630836	NHLF_Lung_Fibroblasts	Roadmapepigenomics_ActiveChromatin	chr8	95127200	95128600	7_Enh		
chr8	95127836	95127842	rs149630836	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127767	95127897	H2A.Z		
chr8	95127836	95127842	rs149630836	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr8	95127752	95127895	H2A.Z		
chr8	95127836	95127842	rs149630836	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127478	95127910	H2AK5ac		
chr8	95127836	95127842	rs149630836	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95126858	95127940	H2BK120ac		
chr8	95127836	95127842	rs149630836	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127146	95127939	H2BK12ac		
chr8	95127836	95127842	rs149630836	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127505	95127846	H2BK15ac		
chr8	95127836	95127842	rs149630836	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95126870	95127963	H2BK20ac		

ACCEPTED MANUSCRIPT											
chr8	95127836	95127842	rs149630836	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127478	95127901	H2BK5ac		
chr8	95127836	95127842	rs149630836	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127489	95127893	H3K14ac		
chr8	95127836	95127842	rs149630836	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127129	95127929	H3K18ac		
chr8	95127836	95127842	rs149630836	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127506	95127842	H3K23ac		
chr8	95127836	95127842	rs149630836	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127487	95127943	H3K27ac		
chr8	95127836	95127842	rs149630836	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr8	95127433	95128005	H3K27ac		
chr8	95127836	95127842	rs149630836	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127127	95127953	H3K4me1		
chr8	95127836	95127842	rs149630836	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr8	95127787	95128025	H3K4me1		
chr8	95127836	95127842	rs149630836	Fetal_Lung	Roadmapepigenomics_Histone_narrow	chr8	95127505	95127954	H3K4me1		
chr8	95127836	95127842	rs149630836	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127134	95127974	H3K4me1		
chr8	95127836	95127842	rs149630836	Lung	Roadmapepigenomics_Histone_narrow	chr8	95127630	95127922	H3K4me1		
chr8	95127836	95127842	rs149630836	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr8	95127364	95127900	H3K4me1		
chr8	95127836	95127842	rs149630836	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127458	95127935	H4K5ac		
chr8	95127836	95127842	rs149630836	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127422	95127914	H4K8ac		
chr8	95127836	95127842	rs149630836	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127183	95127930	H4K91ac		
chr8	95127836	95127842	rs149630836	A549	TxnFactorChIPV3	chr8	95127817	95128141	CEBPB		
chr8	95127836	95127842	rs149630836	IMR90	TxnFactorChIPV3	chr8	95127817	95128141	CEBPB		
chr8	95127836	95127842	rs149630836	A549+EtOH_0.02pct	TxnFactorChIPV3	chr8	95127724	95128134	FOSL2		
chr11	113946484	113946484	rs1614059	Lung	Roadmapepigenomics_ActiveChromatin	chr11	113942800	113966400	6_EnhG		
chr11	113945609	113945609	rs1672692	lung_expressed_enhancers	FANTOM5_enhancer_organ	chr11	113945404	113945849			
chr11	113945609	113945609	rs1672692	CD34+	Misc_ChIP	chr11	113943838	113946241	RUNX1		
chr11	113945609	113945609	rs1672692	Lung	Roadmapepigenomics_ActiveChromatin	chr11	113942800	113966400	6_EnhG		
chr11	113945609	113945609	rs1672692	Lung	Roadmapepigenomics_Histone_narrow	chr11	113944563	113945791	H3K27ac		
chr11	113945609	113945609	rs1672692	Lung	Roadmapepigenomics_Histone_narrow	chr11	113945356	113945815	H3K36me3		
chr11	113945609	113945609	rs1672692	Fetal_Lung	Roadmapepigenomics_Histone_narrow	chr11	113945575	113945758	H3K4me1		
chr11	113945609	113945609	rs1672692	Lung	Roadmapepigenomics_Histone_narrow	chr11	113944556	113946133	H3K4me1		
chr11	113945609	113945609	rs1672692	Lung	Roadmapepigenomics_Histone_narrow	chr11	113945601	113945725	H3K9me3		
chr1	212764930	212764930	rs17019411	Fetal_Lung	Roadmapepigenomics_Histone_narrow	chr1	212764902	212765156	H3K27me3		
chr1	212764930	212764930	rs17019411	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212764561	212764932	H3K27me3		
chr1	212764930	212764930	rs17019411	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr1	212764660	212764988	H3K27me3		
chr1	212816729	212816729	rs17019510	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr1	212816592	212816960	H2A.Z		
chr1	212816729	212816729	rs17019510	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212816586	212816735	H2BK12ac		
chr1	212816729	212816729	rs17019510	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212816623	212816738	H3K4ac		
chr1	212816729	212816729	rs17019510	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212816626	212816736	H3K9me1		
chr1	212816729	212816729	rs17019510	A549+DEX_100nM	TxnFactorChIPV3	chr1	212816462	212816846	USF1		
chr1	212816729	212816729	rs17019510	A549+EtOH_0.02pct	TxnFactorChIPV3	chr1	212816462	212816846	USF1		
chr1	212816729	212816729	rs17019510	A549+EtOH_0.02pct	TxnFactorChIPV3	chr1	212816462	212816846	USF1		
chr11	114026639	114026639	rs17116541	Lung	Roadmapepigenomics_ActiveChromatin	chr11	114025200	114030000	6_EnhG		
chr11	114026639	114026639	rs17116541	Lung	Roadmapepigenomics_Histone_narrow	chr11	114025419	114027162	H3K27ac		
chr11	114026639	114026639	rs17116541	IMR90	Roadmapepigenomics_Histone_narrow	chr11	114026621	114026841	H3K27me3		
chr11	114026639	114026639	rs17116541	Lung	Roadmapepigenomics_Histone_narrow	chr11	114026466	114027452	H3K36me3		
chr11	114026639	114026639	rs17116541	Lung	Roadmapepigenomics_Histone_narrow	chr11	114026589	114027431	H3K4me1		
chr11	113939783	113939783	rs1784683	Lung	Roadmapepigenomics_Histone_narrow	chr11	113939513	113939833	H3K36me3		
chr11	113939783	113939783	rs1784683	Lung	Roadmapepigenomics_Histone_narrow	chr11	113939649	113939882	H3K4me1		
chr11	113939539	113939539	rs1784684	Lung	Roadmapepigenomics_ActiveChromatin	chr11	113937400	113939600	6_EnhG		
chr11	113939539	113939539	rs1784684	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr11	113939285	113939563	H3K27me3		
chr11	113939539	113939539	rs1784684	IMR90	Roadmapepigenomics_Histone_narrow	chr11	113939408	113939631	H3K27me3		
chr11	113939539	113939539	rs1784684	Lung	Roadmapepigenomics_Histone_narrow	chr11	113939513	113939833	H3K36me3		
chr11	113939539	113939539	rs1784684	Lung	Roadmapepigenomics_Histone_narrow	chr11	113939384	113939557	H3K4me1		
chr11	113949232	113949232	rs1784692	Lung	Roadmapepigenomics_ActiveChromatin	chr11	113942800	113966400	6_EnhG		
chr11	113949232	113949232	rs1784692	Lung	Roadmapepigenomics_Histone_narrow	chr11	113949069	113949376	H3K27ac		
chr11	113949232	113949232	rs1784692	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr11	113947756	113949424	H3K27me3		
chr11	113949232	113949232	rs1784692	Fetal_Lung	Roadmapepigenomics_Histone_narrow	chr11	113948470	113949424	H3K27me3		
chr11	113949232	113949232	rs1784692	IMR90	Roadmapepigenomics_Histone_narrow	chr11	113949224	113949364	H3K27me3		
chr11	113949232	113949232	rs1784692	Lung	Roadmapepigenomics_Histone_narrow	chr11	113949191	113949390	H3K36me3		
chr11	113949232	113949232	rs1784692	Fetal_Lung	Roadmapepigenomics_Histone_narrow	chr11	113948826	113949276	H3K4me1		
chr11	113949740	113949740	rs1791800	Lung	Roadmapepigenomics_ActiveChromatin	chr11	113942800	113966400	6_EnhG		
chr11	113949740	113949740	rs1791800	Lung	Roadmapepigenomics_Histone_narrow	chr11	113949679	113949896	H3K27ac		
chr11	113949740	113949740	rs1791800	Lung	Roadmapepigenomics_Histone_narrow	chr11	113949629	113949741	H3K36me3		
chr11	113949740	113949740	rs1791800	Lung	Roadmapepigenomics_Histone_narrow	chr11	113949377	113949852	H3K4me1		
chr11	113939671	113939671	rs1791804	Lung	Roadmapepigenomics_Histone_narrow	chr11	113939513	113939833	H3K36me3		
chr11	113939671	113939671	rs1791804	Lung	Roadmapepigenomics_Histone_narrow	chr11	113939649	113939882	H3K4me1		
chr1	212793050	212793050	rs1803173	Fetal_Lung	Roadmapepigenomics_Histone_narrow	chr1	212792751	212793078	H3K36me3		
chr1	212793050	212793050	rs1803173	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212792367	212793999	H3K36me3		
chr1	212793050	212793050	rs1803173	Lung	Roadmapepigenomics_Histone_narrow	chr1	212792365	212793331	H3K36me3		
chr1	212793050	212793050	rs1803173	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr1	212792795	212793247	H3K36me3		
chr17	65397518	65397518	rs187061404	IMR90	Roadmapepigenomics_Histone_narrow	chr17	65397425	65397539	H3K79me2		
chr17	65399074	65399074	rs191465321	Lung	Roadmapepigenomics_ActiveChromatin	chr17	65397800	65399200	7_Enh		
chr17	65399074	65399074	rs191465321	Lung	Roadmapepigenomics_Histone_narrow	chr17	65399017	65399196	H3K4me1		
chr17	65399074	65399074	rs191465321	IMR90	Roadmapepigenomics_Histone_narrow	chr17	65399004	65399108	H3K79me2		
chr2	75477345	75477345	rs2110940	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75477250	75477564	H3K9me3		
chr2	75477143	75477143	rs2160376	Fetal_Lung	Roadmapepigenomics_Histone_narrow	chr2	75476957	75477208	H3K9me3		
chr2	75477143	75477143	rs2160376	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75477053	75477189	H3K9me3		
chr2	75487574	75487574	rs2216103	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr2	75487369	75487652	H3K36me3		
chr2	75487574	75487574	rs2216103	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr2	75487026	75487632	H3K79me2		
chr8	95160288	95160288	rs2251996	cd34_nont8_21	ReMap_ChIP	chr8	95160241	95160615	RUNX1		
chr8	95160288	95160288	rs2251996	sem	ReMap_ChIP	chr8	95160091	95160631	RUNX1		
chr8	95160288	95160288	rs2251996	IMR90	Roadmapepigenomics_ActiveChromatin	chr8	95158200	95160600	7_Enh		
chr8	95160288	95160288	rs2251996	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95160169	95160416	H2AK5ac		
chr8	95160288	95160288	rs2251996	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95160091	95160539	H2BK120ac		
chr8	95160288	95160288	rs2251996	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95160051	95160559	H2BK12ac		
chr8	95160288	95160288	rs2251996	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95160060	95160467	H2BK15ac		
chr8	95160288	95160288	rs2251996	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95160091	95160573	H2BK20ac		
chr8	95160288	95160288	rs2251996	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95160000	95160568	H2BK5ac		
chr8	95160288	95160288	rs2251996	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95160093	95160405	H3K14ac		
chr8	95160288	95160288	rs2251996	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95159992	95160579	H3K18ac		
chr8	95160288	95160288	rs2251996	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95160110	95160391	H3K23ac		
chr8	95160288	95160288	rs2251996	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95158095	95160597	H3K27ac		

ACCEPTED MANUSCRIPT											
chr8	95160288	95160288	rs2251996	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95160093	95160535	H3K4ac		
chr8	95160288	95160288	rs2251996	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95160014	95160593	H3K4me1		
chr8	95160288	95160288	rs2251996	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95160025	95160462	H3K4me2		
chr8	95160288	95160288	rs2251996	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95160101	95160386	H3K56ac		
chr8	95160288	95160288	rs2251996	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95160067	95160438	H3K9ac		
chr8	95160288	95160288	rs2251996	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95160247	95160400	H3K9me1		
chr8	95160288	95160288	rs2251996	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95160034	95160588	H4K5ac		
chr8	95160288	95160288	rs2251996	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95160087	95160420	H4K8ac		
chr8	95160288	95160288	rs2251996	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95160060	95160518	H4K91ac		
chr2	75448760	75448760	rs2287230	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448497	75448798	H2AK5ac		
chr2	75448760	75448760	rs2287230	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448462	75448843	H2BK120ac		
chr2	75448760	75448760	rs2287230	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448450	75448812	H2BK12ac		
chr2	75448760	75448760	rs2287230	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448471	75448780	H2BK15ac		
chr2	75448760	75448760	rs2287230	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448492	75448806	H2BK20ac		
chr2	75448760	75448760	rs2287230	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448666	75448786	H3K18ac		
chr2	75448760	75448760	rs2287230	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448644	75448791	H3K4ac		
chr2	75448760	75448760	rs2287230	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448634	75448775	H4K5ac		
chr2	75448760	75448760	rs2287230	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448659	75448806	H4K8ac		
chr2	75448760	75448760	rs2287230	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448608	75448820	H4K91ac		
chr2	75449129	75449129	rs2287231	IMR90	Cistrome_ChIP	chr2	75448994	75449483	SP1		
chr2	75449129	75449129	rs2287231	IMR90	DNaseI_Duke	chr2	75449080	75449330	.		
chr2	75449129	75449129	rs2287231	IMR90	Roadmapepigenomics_ActiveChromatin	chr2	75449000	75450200	7_Enh		
chr2	75449129	75449129	rs2287231	NHLF_Lung_Fibroblasts	Roadmapepigenomics_ActiveChromatin	chr2	75449000	75450200	7_Enh		
chr2	75449129	75449129	rs2287231	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr2	75448997	75449423	H2A.Z		
chr2	75449129	75449129	rs2287231	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr2	75448934	75449246	H2A.Z		
chr2	75449129	75449129	rs2287231	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448890	75449161	H2BK120ac		
chr2	75449129	75449129	rs2287231	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448874	75449138	H2BK12ac		
chr2	75449129	75449129	rs2287231	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448895	75449149	H2BK20ac		
chr2	75449129	75449129	rs2287231	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448921	75449145	H3K27ac		
chr2	75449129	75449129	rs2287231	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr2	75448867	75449198	H3K27ac		
chr2	75449129	75449129	rs2287231	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448926	75449153	H3K4ac		
chr2	75449129	75449129	rs2287231	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448961	75449209	H3K4me1		
chr2	75449129	75449129	rs2287231	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr2	75448914	75449145	H3K4me1		
chr2	75449129	75449129	rs2287231	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448950	75449137	H4K5ac		
chr2	75449129	75449129	rs2287231	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448938	75449153	H4K8ac		
chr2	75449129	75449129	rs2287231	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75448909	75449153	H4K91ac		
chr8	95129402	95129402	rs2446818	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95129384	95129544	H2BK120ac		
chr8	95129402	95129402	rs2446818	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95129374	95129508	H2BK20ac		
chr8	95129259	95129259	rs2446819	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95129158	95129268	H2AK5ac		
chr8	95129259	95129259	rs2446819	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95128107	95129338	H2BK120ac		
chr8	95129259	95129259	rs2446819	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95128125	95129273	H2BK12ac		
chr8	95129259	95129259	rs2446819	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95129122	95129390	H2BK15ac		
chr8	95129259	95129259	rs2446819	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95128133	95129305	H2BK20ac		
chr8	95129259	95129259	rs2446819	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95128892	95129299	H3K4ac		
chr8	95129259	95129259	rs2446819	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95128817	95129351	H4K91ac		
chr8	95128479	95128479	rs2446821	IMR90_Senescent	Pazar_Chip	chr8	95128370	95128583	P130		
chr8	95128479	95128479	rs2446821	IMR90_shRbQuiescent	Pazar_Chip	chr8	95128130	95129000	P130		
chr8	95128479	95128479	rs2446821	IMR90_shRbSenescent	Pazar_Chip	chr8	95128244	95128636	P130		
chr8	95128479	95128479	rs2446821	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_ActiveChromatin	chr8	95127400	95128600	7_Enh		
chr8	95128479	95128479	rs2446821	IMR90	Roadmapepigenomics_ActiveChromatin	chr8	95127000	95128600	7_Enh		
chr8	95128479	95128479	rs2446821	NHLF_Lung_Fibroblasts	Roadmapepigenomics_ActiveChromatin	chr8	95127200	95128600	7_Enh		
chr8	95128479	95128479	rs2446821	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr8	95128169	95128484	H2A.Z		
chr8	95128479	95128479	rs2446821	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95128180	95128511	H2A.Z		
chr8	95128479	95128479	rs2446821	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr8	95128165	95128500	H2A.Z		
chr8	95128479	95128479	rs2446821	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95128177	95128711	H2AK5ac		
chr8	95128479	95128479	rs2446821	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95128107	95129338	H2BK120ac		
chr8	95128479	95128479	rs2446821	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95128125	95129273	H2BK12ac		
chr8	95128479	95128479	rs2446821	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95128147	95128731	H2BK15ac		
chr8	95128479	95128479	rs2446821	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95128133	95129305	H2BK20ac		
chr8	95128479	95128479	rs2446821	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95128106	95128513	H2BK5ac		
chr8	95128479	95128479	rs2446821	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95128123	95128705	H3K14ac		
chr8	95128479	95128479	rs2446821	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95128101	95128805	H3K18ac		
chr8	95128479	95128479	rs2446821	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95128154	95128526	H3K23ac		
chr8	95128479	95128479	rs2446821	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95128125	95128777	H3K27ac		
chr8	95128479	95128479	rs2446821	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr8	95128089	95128776	H3K27ac		
chr8	95128479	95128479	rs2446821	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95128111	95128845	H3K4ac		
chr8	95128479	95128479	rs2446821	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr8	95128117	95128554	H3K4me1		
chr8	95128479	95128479	rs2446821	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95128127	95128662	H3K4me1		
chr8	95128479	95128479	rs2446821	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95128106	95128663	H3K56ac		
chr8	95128479	95128479	rs2446821	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95128125	95128847	H4K5ac		
chr8	95128479	95128479	rs2446821	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95128113	95128774	H4K8ac		
chr8	95128479	95128479	rs2446821	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95128108	95128754	H4K91ac		
chr8	95127612	95127612	rs2446823	IMR90	Cistrome_ChIP	chr8	95127447	95127694	SP1		
chr8	95127612	95127612	rs2446823	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_ActiveChromatin	chr8	95127400	95128600	7_Enh		
chr8	95127612	95127612	rs2446823	Fetal_Lung	Roadmapepigenomics_ActiveChromatin	chr8	95127200	95128000	7_Enh		
chr8	95127612	95127612	rs2446823	IMR90	Roadmapepigenomics_ActiveChromatin	chr8	95127000	95128600	7_Enh		
chr8	95127612	95127612	rs2446823	Lung	Roadmapepigenomics_ActiveChromatin	chr8	95127200	95128400	7_Enh		
chr8	95127612	95127612	rs2446823	NHLF_Lung_Fibroblasts	Roadmapepigenomics_ActiveChromatin	chr8	95127200	95128600	7_Enh		
chr8	95127612	95127612	rs2446823	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr8	95127558	95127692	H2A.Z		
chr8	95127612	95127612	rs2446823	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127478	95127910	H2AK5ac		
chr8	95127612	95127612	rs2446823	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127522	95127685	H2AK9ac		
chr8	95127612	95127612	rs2446823	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95126858	95127940	H2BK120ac		
chr8	95127612	95127612	rs2446823	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127146	95127939	H2BK12ac		
chr8	95127612	95127612	rs2446823	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127505	95127846	H2BK15ac		
chr8	95127612	95127612	rs2446823	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95126870	95127963	H2BK20ac		
chr8	95127612	95127612	rs2446823	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127478	95127901	H2BK5ac		
chr8	95127612	95127612	rs2446823	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127489	95127893	H3K14ac		
chr8	95127612	95127612	rs2446823	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127129	95127929	H3K18ac		

chr8	95127612	95127612	rs2446823	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127506	95127842	H3K23ac
chr8	95127612	95127612	rs2446823	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127487	95127943	H3K27ac
chr8	95127612	95127612	rs2446823	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr8	95127433	95128005	H3K27ac
chr8	95127612	95127612	rs2446823	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127127	95127953	H3K4ac
chr8	95127612	95127612	rs2446823	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr8	95127452	95127720	H3K4me1
chr8	95127612	95127612	rs2446823	Fetal_Lung	Roadmapepigenomics_Histone_narrow	chr8	95127505	95127954	H3K4me1
chr8	95127612	95127612	rs2446823	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127134	95127974	H3K4me1
chr8	95127612	95127612	rs2446823	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr8	95127364	95127900	H3K4me1
chr8	95127612	95127612	rs2446823	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr8	95127552	95127776	H3K4me2
chr8	95127612	95127612	rs2446823	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127562	95127814	H3K56ac
chr8	95127612	95127612	rs2446823	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127548	95127693	H3K9me1
chr8	95127612	95127612	rs2446823	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127458	95127935	H4K5ac
chr8	95127612	95127612	rs2446823	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127422	95127914	H4K8ac
chr8	95127612	95127612	rs2446823	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127183	95127930	H4K91ac
chr8	95127612	95127612	rs2446823	A549	TxnFactorChIPV3	chr8	95127455	95127717	CEBPB
chr8	95127612	95127612	rs2446823	IMR90	TxnFactorChIPV3	chr8	95127455	95127717	CEBPB
chr8	95127574	95127574	rs2446824	IMR90	Cistrome_ChIP	chr8	95127447	95127694	SPI1
chr8	95127574	95127574	rs2446824	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_ActiveChromatin	chr8	95127400	95128600	7_Enh
chr8	95127574	95127574	rs2446824	Fetal_Lung	Roadmapepigenomics_ActiveChromatin	chr8	95127200	95128000	7_Enh
chr8	95127574	95127574	rs2446824	IMR90	Roadmapepigenomics_ActiveChromatin	chr8	95127000	95128600	7_Enh
chr8	95127574	95127574	rs2446824	Lung	Roadmapepigenomics_ActiveChromatin	chr8	95127200	95128400	7_Enh
chr8	95127574	95127574	rs2446824	NHLF_Lung_Fibroblasts	Roadmapepigenomics_ActiveChromatin	chr8	95127200	95128600	7_Enh
chr8	95127574	95127574	rs2446824	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr8	95127558	95127692	H2A.Z
chr8	95127574	95127574	rs2446824	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127478	95127910	H2AK5ac
chr8	95127574	95127574	rs2446824	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127522	95127685	H2AK9ac
chr8	95127574	95127574	rs2446824	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95126858	95127940	H2BK120ac
chr8	95127574	95127574	rs2446824	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127146	95127939	H2BK12ac
chr8	95127574	95127574	rs2446824	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127505	95127846	H2BK15ac
chr8	95127574	95127574	rs2446824	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95126870	95127963	H2BK20ac
chr8	95127574	95127574	rs2446824	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127478	95127901	H2BK5ac
chr8	95127574	95127574	rs2446824	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127489	95127893	H3K14ac
chr8	95127574	95127574	rs2446824	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127129	95127929	H3K18ac
chr8	95127574	95127574	rs2446824	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127506	95127842	H3K23ac
chr8	95127574	95127574	rs2446824	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127487	95127943	H3K27ac
chr8	95127574	95127574	rs2446824	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr8	95127433	95128005	H3K27ac
chr8	95127574	95127574	rs2446824	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127127	95127953	H3K4ac
chr8	95127574	95127574	rs2446824	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr8	95127452	95127720	H3K4me1
chr8	95127574	95127574	rs2446824	Fetal_Lung	Roadmapepigenomics_Histone_narrow	chr8	95127505	95127954	H3K4me1
chr8	95127574	95127574	rs2446824	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127134	95127974	H3K4me1
chr8	95127574	95127574	rs2446824	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr8	95127364	95127900	H3K4me1
chr8	95127574	95127574	rs2446824	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr8	95127552	95127776	H3K4me2
chr8	95127574	95127574	rs2446824	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127562	95127814	H3K56ac
chr8	95127574	95127574	rs2446824	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127548	95127693	H3K9me1
chr8	95127574	95127574	rs2446824	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127458	95127935	H4K5ac
chr8	95127574	95127574	rs2446824	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127422	95127914	H4K8ac
chr8	95127574	95127574	rs2446824	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127183	95127930	H4K91ac
chr8	95127574	95127574	rs2446824	A549	TxnFactorChIPV3	chr8	95127455	95127717	CEBPB
chr8	95127574	95127574	rs2446824	IMR90	TxnFactorChIPV3	chr8	95127455	95127717	CEBPB
chr8	95126478	95126478	rs2513784	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95126436	95126719	H2BK120ac
chr8	95126478	95126478	rs2513784	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95126466	95126722	H2BK12ac
chr8	95126478	95126478	rs2513784	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95126447	95126730	H2BK15ac
chr8	95126478	95126478	rs2513784	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95126419	95126743	H2BK20ac
chr8	95126478	95126478	rs2513784	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95126452	95126714	H3K14ac
chr8	95126478	95126478	rs2513784	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95126473	95126708	H3K4ac
chr8	95126478	95126478	rs2513784	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95126441	95126704	H4K91ac
chr8	95126888	95126888	rs2513786	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95126858	95127940	H2BK120ac
chr8	95126888	95126888	rs2513786	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95126858	95127058	H2BK12ac
chr8	95126888	95126888	rs2513786	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95126868	95127071	H2BK15ac
chr8	95126888	95126888	rs2513786	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95126870	95127963	H2BK20ac
chr8	95126888	95126888	rs2513786	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95126856	95127017	H4K91ac
chr8	95127080	95127080	rs2513787	IMR90	Roadmapepigenomics_ActiveChromatin	chr8	95127000	95128600	7_Enh
chr8	95127080	95127080	rs2513787	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95126858	95127940	H2BK120ac
chr8	95127080	95127080	rs2513787	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95126870	95127963	H2BK20ac
chr8	95127896	95127896	rs2513788	IMR90	Cistrome_ChIP	chr8	95127764	95128000	EP300
chr8	95127896	95127896	rs2513788	A549	Cistrome_ChIP	chr8	95127830	95128012	RELA
chr8	95127896	95127896	rs2513788	IMR90	Cistrome_ChIP	chr8	95127737	95128151	SPI1
chr8	95127896	95127896	rs2513788	A549	DNaseClusters	chr8	95127880	95128030	.
chr8	95127896	95127896	rs2513788	A549	DNaseI_Duke	chr8	95127800	95128025	.
chr8	95127896	95127896	rs2513788	IMR90	DNaseI_Duke	chr8	95127772	95128096	.
chr8	95127896	95127896	rs2513788	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_ActiveChromatin	chr8	95127400	95128600	7_Enh
chr8	95127896	95127896	rs2513788	Fetal_Lung	Roadmapepigenomics_ActiveChromatin	chr8	95127200	95128000	7_Enh
chr8	95127896	95127896	rs2513788	IMR90	Roadmapepigenomics_ActiveChromatin	chr8	95127000	95128600	7_Enh
chr8	95127896	95127896	rs2513788	Lung	Roadmapepigenomics_ActiveChromatin	chr8	95127200	95128400	7_Enh
chr8	95127896	95127896	rs2513788	NHLF_Lung_Fibroblasts	Roadmapepigenomics_ActiveChromatin	chr8	95127200	95128600	7_Enh
chr8	95127896	95127896	rs2513788	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127767	95127897	H2A.Z
chr8	95127896	95127896	rs2513788	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127478	95127910	H2AK5ac
chr8	95127896	95127896	rs2513788	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95126858	95127940	H2BK120ac
chr8	95127896	95127896	rs2513788	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127146	95127939	H2BK12ac
chr8	95127896	95127896	rs2513788	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95126870	95127963	H2BK20ac
chr8	95127896	95127896	rs2513788	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127478	95127901	H2BK5ac
chr8	95127896	95127896	rs2513788	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127129	95127929	H3K18ac
chr8	95127896	95127896	rs2513788	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127487	95127943	H3K27ac
chr8	95127896	95127896	rs2513788	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr8	95127433	95128005	H3K27ac
chr8	95127896	95127896	rs2513788	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127127	95127953	H3K4ac
chr8	95127896	95127896	rs2513788	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr8	95127787	95128025	H3K4me1
chr8	95127896	95127896	rs2513788	Fetal_Lung	Roadmapepigenomics_Histone_narrow	chr8	95127505	95127954	H3K4me1
chr8	95127896	95127896	rs2513788	IMR90	Roadmapepigenomics_Histone_narrow	chr8	95127134	95127974	H3K4me1
chr8	95127896	95127896	rs2513788	Lung	Roadmapepigenomics_Histone_narrow	chr8	95127630	95127922	H3K4me1

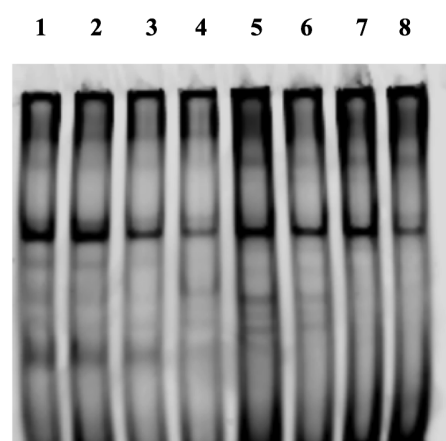
ACCEPTED MANUSCRIPT											
chr11	113945804	113945804	rs2846625	Lung	Roadmapepigenomics_Histone_narrow	chr11	113944556	113946133	H3K4me1		
chr11	113942767	113942767	rs2852793	Fetal_Lung	Roadmapepigenomics_Histone_narrow	chr11	113942599	113943303	H3K27me3		
chr11	113954925	113954925	rs2852800	Lung	Roadmapepigenomics_ActiveChromatin	chr11	113942800	113966400	6_EnhG		
chr11	113954925	113954925	rs2852800	Lung	Roadmapepigenomics_Histone_narrow	chr11	113954921	113955119	H3K4me1		
chr1	112529253	112529253	rs34115852	Fetal_Lung	Roadmapepigenomics_Histone_narrow	chr1	112528659	112529599	H3K27me3		
chr7	19063083	19063083	rs34917720	IMR90	Roadmapepigenomics_Histone_narrow	chr7	19063012	19063720	H2BK120ac		
chr7	19063083	19063083	rs34917720	IMR90	Roadmapepigenomics_Histone_narrow	chr7	19063056	19063191	H2BK20ac		
chr7	19036230	19036230	rs35409100	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr7	19035885	19036356	H2A.Z		
chr2	75485334	75485334	rs35879047	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr2	75485060	75485543	H3K9me3		
chr2	75485334	75485334	rs35879047	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75485181	75485345	H3K9me3		
chr2	75449986	75449986	rs4853124	IMR90	Roadmapepigenomics_ActiveChromatin	chr2	75449000	75450200	7_Enh		
chr2	75449986	75449986	rs4853124	NHLF_Lung_Fibroblasts	Roadmapepigenomics_ActiveChromatin	chr2	75449000	75450200	7_Enh		
chr2	75449986	75449986	rs4853124	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr2	75449293	75450041	H2A.Z		
chr2	75449986	75449986	rs4853124	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75449365	75450218	H2BK120ac		
chr2	75449986	75449986	rs4853124	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75449396	75450207	H2BK12ac		
chr2	75449986	75449986	rs4853124	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75449394	75450222	H2BK15ac		
chr2	75449986	75449986	rs4853124	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75449534	75450230	H2BK20ac		
chr2	75449986	75449986	rs4853124	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75449393	75450191	H3K14ac		
chr2	75449986	75449986	rs4853124	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75449379	75450208	H3K18ac		
chr2	75449986	75449986	rs4853124	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75449546	75450227	H3K4ac		
chr2	75449986	75449986	rs4853124	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75449551	75450162	H3K4me1		
chr2	75449986	75449986	rs4853124	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr2	75449507	75450051	H3K4me1		
chr2	75449986	75449986	rs4853124	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75449506	75450031	H3K56ac		
chr2	75449986	75449986	rs4853124	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75449720	75450185	H3K9me1		
chr2	75449986	75449986	rs4853124	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75449595	75450220	H4K5ac		
chr2	75449986	75449986	rs4853124	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75449550	75450204	H4K8ac		
chr2	75449986	75449986	rs4853124	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75449530	75450230	H4K91ac		
chr12	295370	295370	rs529198	IMR90	Roadmapepigenomics_Histone_narrow	chr12	295140	295482	H2A.Z		
chr12	295370	295370	rs529198	Lung	Roadmapepigenomics_Histone_narrow	chr12	295334	295458	H3K9me3		
chr2	31121117	31121117	rs55971416	Lung	Roadmapepigenomics_Histone_narrow	chr2	31120940	31121273	H3K9me3		
chr1	212808004	212808004	rs56412520	Fetal_Lung	Roadmapepigenomics_ActiveChromatin	chr1	212808000	212809600	7_Enh		
chr1	212808004	212808004	rs56412520	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212807917	212808036	H2BK15ac		
chr1	212808004	212808004	rs56412520	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212807882	212808035	H3K18ac		
chr1	212808004	212808004	rs56412520	Fetal_Lung	Roadmapepigenomics_Histone_narrow	chr1	212807912	212808046	H3K4me3		
chr1	212808004	212808004	rs56412520	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr1	212807721	212808142	H3K9me3		
chr1	212808004	212808004	rs56412520	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212807710	212808073	H3K9me3		
chr1	212808004	212808004	rs56412520	Lung	Roadmapepigenomics_Histone_narrow	chr1	212807913	212808037	H3K9me3		
chr10	69332802	69332802	rs575404339	Lung	Roadmapepigenomics_Histone_narrow	chr10	69332660	69332810	H3K9me3		
chr11	114025754	114025754	rs58895081	IMR90	Roadmapepigenomics_ActiveChromatin	chr11	114025600	114025800	12_EnhBiv		
chr11	114025754	114025754	rs58895081	Lung	Roadmapepigenomics_ActiveChromatin	chr11	114025200	114030000	6_EnhG		
chr11	114025754	114025754	rs58895081	Fetal_Lung	Roadmapepigenomics_ActiveChromatin	chr11	114025000	114026600	7_Enh		
chr11	114025754	114025754	rs58895081	Lung	Roadmapepigenomics_Histone_narrow	chr11	114025419	114027162	H3K27ac		
chr11	114025754	114025754	rs58895081	IMR90	Roadmapepigenomics_Histone_narrow	chr11	114025386	114025758	H3K27me3		
chr11	114025754	114025754	rs58895081	Fetal_Lung	Roadmapepigenomics_Histone_narrow	chr11	114025067	114026567	H3K4me1		
chr11	114025754	114025754	rs58895081	Lung	Roadmapepigenomics_Histone_narrow	chr11	114025526	114026266	H3K4me1		
chr11	114025754	114025754	rs58895081	Fetal_Lung	Roadmapepigenomics_Histone_narrow	chr11	114025575	114025846	H3K9ac		
chr11	114025757	114025759	rs59038579	IMR90	Roadmapepigenomics_ActiveChromatin	chr11	114025600	114025800	12_EnhBiv		
chr11	114025757	114025759	rs59038579	Lung	Roadmapepigenomics_ActiveChromatin	chr11	114025200	114030000	6_EnhG		
chr11	114025757	114025759	rs59038579	Fetal_Lung	Roadmapepigenomics_ActiveChromatin	chr11	114025000	114026600	7_Enh		
chr11	114025757	114025759	rs59038579	Lung	Roadmapepigenomics_Histone_narrow	chr11	114025419	114027162	H3K27ac		
chr11	114025757	114025759	rs59038579	IMR90	Roadmapepigenomics_Histone_narrow	chr11	114025386	114025758	H3K27me3		
chr11	114025757	114025759	rs59038579	Fetal_Lung	Roadmapepigenomics_Histone_narrow	chr11	114025067	114026567	H3K4me1		
chr11	114025757	114025759	rs59038579	Lung	Roadmapepigenomics_Histone_narrow	chr11	114025526	114026266	H3K4me1		
chr11	114025757	114025759	rs59038579	Fetal_Lung	Roadmapepigenomics_Histone_narrow	chr11	114025575	114025846	H3K9ac		
chr12	295367	295367	rs647001	IMR90	Roadmapepigenomics_Histone_narrow	chr12	295140	295482	H2A.Z		
chr12	295367	295367	rs647001	Lung	Roadmapepigenomics_Histone_narrow	chr12	295334	295458	H3K9me3		
chr2	75452466	75452466	rs6546958	Lung	Roadmapepigenomics_Histone_narrow	chr2	75452398	75452528	H3K27ac		
chr2	75451047	75451047	rs6717216	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75450992	75451149	H2BK120ac		
chr2	75476771	75476771	rs6729324	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr2	75476374	75476794	H3K9me3		
chr2	75476771	75476771	rs6729324	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75476546	75476790	H3K9me3		
chr1	212770271	212770271	rs72756369	IMR90_shRbQuiescent	Pazar_Chip	chr1	212769501	212770837	P130		
chr1	212770271	212770271	rs72756369	IMR90_shRbSenescent	Pazar_Chip	chr1	212769880	212770388	P130		
chr1	212770271	212770271	rs72756369	NHLF_Lung_Fibroblasts	Roadmapepigenomics_ActiveChromatin	chr1	212770000	212770400	2_TssAFlnk		
chr1	212770271	212770271	rs72756369	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_ActiveChromatin	chr1	212770200	212771000	7_Enh		
chr1	212770271	212770271	rs72756369	IMR90	Roadmapepigenomics_ActiveChromatin	chr1	212770200	212771200	7_Enh		
chr1	212770271	212770271	rs72756369	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212770250	212770522	H2AK5ac		
chr1	212770271	212770271	rs72756369	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212769792	212770609	H2BK120ac		
chr1	212770271	212770271	rs72756369	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212769826	212770664	H2BK12ac		
chr1	212770271	212770271	rs72756369	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212769865	212770565	H2BK15ac		
chr1	212770271	212770271	rs72756369	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212769856	212770624	H2BK20ac		
chr1	212770271	212770271	rs72756369	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212769739	212770638	H2BK5ac		
chr1	212770271	212770271	rs72756369	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212769888	212770607	H3K14ac		
chr1	212770271	212770271	rs72756369	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212769778	212770612	H3K18ac		
chr1	212770271	212770271	rs72756369	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212769880	212770518	H3K23ac		
chr1	212770271	212770271	rs72756369	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr1	212769714	212770585	H3K27ac		
chr1	212770271	212770271	rs72756369	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212769822	212770562	H3K27ac		
chr1	212770271	212770271	rs72756369	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr1	212768628	212771201	H3K27ac		
chr1	212770271	212770271	rs72756369	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212769825	212770594	H3K4ac		
chr1	212770271	212770271	rs72756369	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212770238	212770572	H3K4me1		
chr1	212770271	212770271	rs72756369	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr1	212770233	212770714	H3K4me1		
chr1	212770271	212770271	rs72756369	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212769760	212770372	H3K4me2		
chr1	212770271	212770271	rs72756369	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr1	212769903	212770590	H3K4me2		
chr1	212770271	212770271	rs72756369	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212769838	212770451	H3K56ac		
chr1	212770271	212770271	rs72756369	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212769842	212770330	H3K9ac		
chr1	212770271	212770271	rs72756369	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212769823	212770613	H4K5ac		
chr1	212770271	212770271	rs72756369	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212769855	212770562	H4K8ac		
chr1	212770271	212770271	rs72756369	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212769873	212770610	H4K91ac		
chr1	212794997	212794997	rs74138575	A549	DNaseClusters	chr1	212794860	212795010	.		

chr1	212794997	212794997	rs74138575	CD34+	Misc_ChIP	chr1	212794103	212795547	RUNX1
chr1	212794997	212794997	rs74138575	cd34_nont8_21	ReMap_ChIP	chr1	212794861	212795311	RUNX1
chr1	212794997	212794997	rs74138575	skno1_leukemic	ReMap_ChIP	chr1	212794971	212795282	RUNX1
chr1	212794997	212794997	rs74138575	A549	TxnFactorChIPV3	chr1	212794247	212796320	POLR2A
chr1	212794997	212794997	rs74138575	IMR90	TxnFactorChIPV3	chr1	212794247	212796320	POLR2A
chr1	212799039	212799039	rs74138590	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212798910	212800527	H3K27me3
chr1	212799039	212799039	rs74138590	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr1	212798696	212799466	H3K27me3
chr1	212799163	212799163	rs74138591	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212798910	212800527	H3K27me3
chr1	212799163	212799163	rs74138591	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr1	212798696	212799466	H3K27me3
chr2	31118932	31118934	rs74423934	Fetal_Lung	Roadmapepigenomics_Histone_narrow	chr2	31118048	31119373	H3K27me3
chr1	212795196	212795196	rs75465959	CD34+	Misc_ChIP	chr1	212794103	212795547	RUNX1
chr1	212795196	212795196	rs75465959	cd34_nont8_21	ReMap_ChIP	chr1	212794861	212795311	RUNX1
chr1	212795196	212795196	rs75465959	skno1_leukemic	ReMap_ChIP	chr1	212794971	212795282	RUNX1
chr1	212795196	212795196	rs75465959	IMR90	TxnFactorChIPV3	chr1	212795093	212795377	MAFK
chr1	212795196	212795196	rs75465959	A549	TxnFactorChIPV3	chr1	212794247	212796320	POLR2A
chr1	212795196	212795196	rs75465959	IMR90	TxnFactorChIPV3	chr1	212794247	212796320	POLR2A
chr12	277708	277708	rs75748509	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr12	277522	277763	H3K27me3
chr2	75478007	75478007	rs7583283	Lung	Roadmapepigenomics_Histone_narrow	chr2	75477998	75478163	H3K9me3
chr2	75473916	75473916	rs7588010	IMR90	Cistrome_ChIP	chr2	75473670	75474232	SPI1
chr2	75473916	75473916	rs7588010	NHLF_Lung_Fibroblasts	Roadmapepigenomics_ActiveChromatin	chr2	75473200	75474400	7_Enh
chr2	75473916	75473916	rs7588010	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr2	75473871	75474426	H2A.Z
chr2	75473916	75473916	rs7588010	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75473527	75473955	H2BK120ac
chr2	75473916	75473916	rs7588010	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75473536	75473936	H2BK12ac
chr2	75473916	75473916	rs7588010	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75473537	75473956	H2BK20ac
chr2	75473916	75473916	rs7588010	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75473525	75474691	H2BK5ac
chr2	75473916	75473916	rs7588010	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75473533	75473956	H3K18ac
chr2	75473916	75473916	rs7588010	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75473526	75473936	H3K27ac
chr2	75473916	75473916	rs7588010	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr2	75473838	75474052	H3K36me3
chr2	75473916	75473916	rs7588010	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75473533	75473948	H3K4ac
chr2	75473916	75473916	rs7588010	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75473540	75473931	H3K4me2
chr2	75473916	75473916	rs7588010	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75473604	75473951	H3K56ac
chr2	75473916	75473916	rs7588010	A549_EtOH_0.02pct_Lung_Carcinoma	Roadmapepigenomics_Histone_narrow	chr2	75473747	75474168	H3K79me2
chr2	75473916	75473916	rs7588010	IMR90	Roadmapepigenomics_Histone_narrow	chr2	75473542	75473942	H4K8ac
chr1	212765432	212765432	rs77911820	NHLF_Lung_Fibroblasts	Roadmapepigenomics_Histone_narrow	chr1	212765277	212765526	H3K27me3
chr1	212791460	212791460	rs8192659	IMR90	Roadmapepigenomics_Histone_narrow	chr1	212791254	212792323	H3K36me3

Supplemental Figure Legends

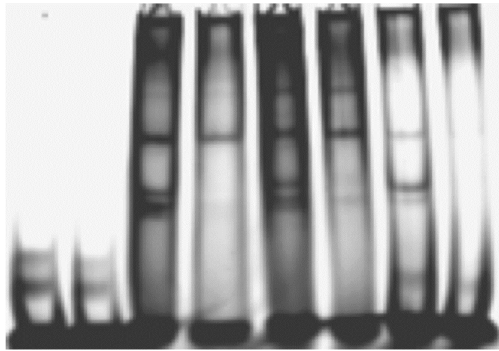
Fig E1. Competitive inhibition experiment showing specificity of non-risk oligonucleotide probes for rs147978008 and rs11571537 for EMSA band with A549 nuclear extract. The labeled and unlabeled oligonucleotides were co-incubated with the nuclear extract prior to running EMSA. The arrow points to a non-specific band that was observed in all the EMSA gels. This band does not show inhibition by unlabeled oligonucleotides. **Legend:** Lane 1 - A549 nuclear extract +rs11571537 - 50 Femtomole labeled, Lane 2 - A549 nuclear extract +rs rs11571537 - 50 fm unlabeled+50 fm labeled (Equimolar), Lane 3 - A549 nuclear extract +rs rs11571537 - 5000 fm unlabeled+50 fm labeled (100 Molar excess unlabeled) , Lane 4 - A549 nuclear extract +rs11571537 - 10000 fm unlabeled+50 fm labeled (200 Molar excess unlabeled), Lane 5 - A549 nuclear extract +rs147978008 - 50 Femtomole labeled, Lane 6 - A549 nuclear extract +rs147978008 - 50 fm unlabeled+50 fm labeled (Equimolar), Lane 7 - A549 nuclear extract +rs147978008 - 5000 fm unlabeled+50 fm labeled (100 Molar excess unlabeled), Lane 8 - A549 nuclear extract +rs147978008 - 10000 fm unlabeled+50 fm labeled (200 Molar excess unlabeled)

Fig E2. EMSA gels showing Allele-specific binding of nuclear extract (NE) proteins in 3 different cell lines for non-risk (NR) and risk (R) alleles of oligonucleotide probes. Lane 1-NR oligonucleotide negative control, Lane 2-Risk oligonucleotide negative control, Lane 3-NR oligonucleotide with A549 NE, Lane 4-R oligonucleotide with A549 NE, Lane 5- NR oligonucleotide with BEAS 2B NE, Lane 6-R oligonucleotide with BEAS 2B NE, Lane 7-NR oligonucleotide with IMR-90 NE, Lane 8-R oligonucleotide with IMR-90 NE.

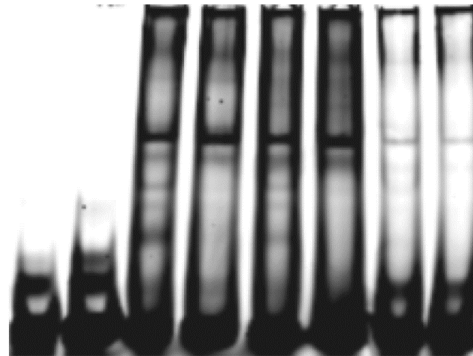


A. rs147978008

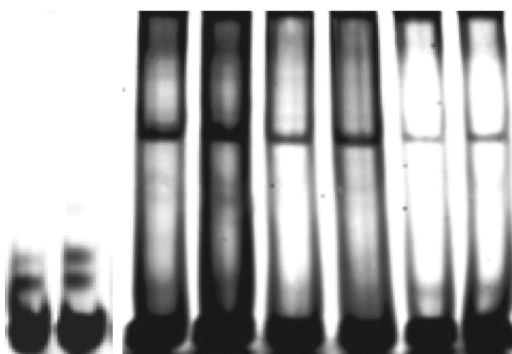
1 2 3 4 5 6 7 8

**B. rs11571537**

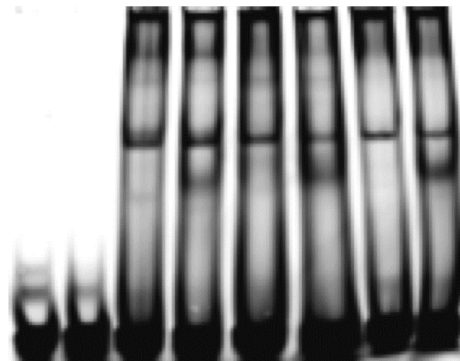
1 2 3 4 5 6 7 8

**C. rs2513788**

1 2 3 4 5 6 7 8

**D. rs2513790**

1 2 3 4 5 6 7 8

**E. rs2287231**

1 2 3 4 5 6 7 8

