Integrated Genomics Approaches Identify Mediators of Transcriptional and Epigenetic Responses to Afghan Desert Particulate Matter in Human Small Airway Epithelial Cells

Arnav Gupta, Sarah K. Sasse, Reena Berman, Margaret A. Gruca, Robin D. Dowell, Hong Wei Chu, Gregory P. Downey, Anthony N. Gerber

Abstract

New respiratory diseases in personnel deployed to Southwest Asia after September 11th raise major concerns about the impacts of airborne particulate matter. Although regulations exist, the knowledge about how particulates influence disease states is limited, precluding the appropriate recognition, prevention and treatment of deployment-related lung diseases. We applied two genomics assays, Precision Run-on sequencing (PRO-seq) and the assay for transposase accessible chromatin with sequencing (ATAC-seq), to characterize the small airway epithelial cell response to Afghan desert particulate matter (APM). In addition to cellular pathways that regulate susceptibility to further insults, such as cytochrome P450 or P53, we identified novel pathways associated with APM exposure, such as GRHL2 loss or TEAD3 activation. We further demonstrated that TEAD3 activation was unique to APM exposure despite similar inflammatory responses when compared with wood smoke particle (WSP) exposure. Our results establish the utility of an integrated genomics approach in characterizing responses to exposures and identify cellular and genomic targets for advanced investigation of the pathogenesis of deployment-related respiratory disease.

Introduction

The development of respiratory disease following post-9/11 deployment to Southwest Asia (SW Asia) and exposure to associated airborne hazards, collectively termed deployment-related lung diseases (DLD), is an important clinical problem that has had a major adverse impact on the long-term health of veterans. DLD encompasses several respiratory conditions, such as asthma, constrictive bronchiolitis and chronic rhinosinusitis (1), which develop following deployment and exposure in individuals in the absence of known prior respiratory illness or predisposition. Epidemiologic studies have previously associated exposure with emergent respiratory disease, both in the context of SW Asia deployment (1, 2) and domestic air pollution (3). Additional studies of airborne exposures extend their impact beyond inflammatory and remodeling diseases of the respiratory tract and implicate these exposures in the development of lung cancer, cardiovascular disease and neuropsychiatric disease (4–7). This raises

additional concerns about long-term sequelae of deployment exposures. However, our mechanistic understanding of the impacts of deployment exposures in relationship to disease pathogenesis and long-term risk remains fragmentary.

Previous studies have used cell and animal-based models of particulate matter (PM) exposure to identify potentially relevant cellular, physiologic and pathophysiologic responses. Although such studies fail to capture the complexity of true inhalational particulate exposure, they have been useful in defining PM-regulated pathways relevant to human exposures, such as Type I inflammation. For example, wood smoke particle (WSP) exposure resulted in increased inflammatory markers in the bronchoalveolar lavage (BAL) fluid of healthy human subjects (8) and induced pro-inflammatory gene expression in submerged airway epithelial cell culture (9). Building on these physiologic consistencies, other studies identified mediators that may contribute to airways diseases, such as mucus abnormalities (10) or cilia dysfunction (11), as well as long term sequelae, such as cellular dysplasia and malignancy (12). Additional studies connected inflammatory responses to particulate exposure with development of fibrosis (13) or allergic sensitization (14), justifying the potential for particulate exposures to contribute to new-onset disease. More recently, particulate matter models specific to SW Asia deployment (Afghan Desert Particulate matter or APM) have demonstrated physiologic responses relevant to DLD, such as inflammation or airway hyperresponsiveness (15), and induction of cytokines such as CXCL8 and IL-33 (16). However, these studies did not identify the mechanistic basis for these responses, nor did they examine factors that many distinguish the effects of APM in comparison to other PMs that are associated with a different spectrum of disease risk.

We previously applied a combination of nascent transcript sequencing using Precision Run-on sequencing (PRO-seq) and chromatin profiling using the assay for transposase accessible chromatin with sequencing (ATAC-seq) to define the mechanistic basis for the immediate early effects of WSP on airway epithelial cells. Our data indicated that acute exposure to particulate matter under highly controlled experimental conditions results in broad transcriptional and epigenetic responses (17), including induction of NFkB activity. Furthermore, PRO-seq provides an indirect measure of the activity of transcriptional regulatory elements (TREs) based on associated enhancer RNA levels. Using this metric, we identified thousands of TREs that respond directly to WSP and linked these changes in TRE activity to epigenetic regulation. We hypothesized that APM also induces changes in epigenetics and TRE activity that define pathways unique to APM exposure. Here, we address this hypothesis through combinatorial

application of PRO-seq and ATAC-seq to APM-exposed primary human small airway epithelial cells and comparison to the effects of WSP exposure.

Methods

Cell culture

Cell culture of primary small airway epithelial cells (SAECs) was performed as described previously (17). Briefly, de-identified primary human small airway epithelial cells (< 2 mm diameter) were obtained from the National Jewish Health Biobank from donors with no history of deployment or chronic lung disease. Cells were first co-cultured with irradiated NIH/3T3 (ATCC) fibroblasts in F-medium containing 1 uM Y-27632 (APEX Bio) and then passaged onto tissue culture dishes double-coated with Type I bovine collagen solution (Advanced Biomatrix) and grown to confluence in BronchiaLife Epithelial Medium with the complete LifeFactors Supplement Kit (Lifeline Cell Technology). All exposures were performed in this medium. All cells were maintained in 5% CO_2 at 37°C.

Afghan Desert Particulate Matter – Preparation and Exposure

Afghan Desert Particulate Matter (APM) was a generous gift from Drs. Greg Downey and Hong Wei Chu. The composition and cell culture exposure conditions of this material have been characterized previously (16). Briefly, topsoil from Bagram Air Force Base in the Parwan Province of Afghanistan was collected in August 2009 and shipped to the United States, where it was autoclaved and irradiated in accordance with Department of Agriculture standards. Scanning electron microscopy determined that 89.5% of the particles were < 2.5 um in diameter, 10.4% between 2.5 and 10 um and 0.1% greater than 10 um. Dry APM was resuspended in 1X Phosphate buffered saline (PBS, Dulbecco) and vortexed prior to addition to cell culture media. After addition, media was gently agitated to ensure even distribution of the particles.

Wood Smoke Particles – Preparation and Exposure

Wood Smoke Particles (WSP) were generously gifted by Dr. Andrew Ghio at the Environmental Protection Agency Laboratory in North Carolina and generated as described previously (8). WSP composition, preparation and exposure were described previously by our laboratory (17). The mean diameter of the freshly generated wood smoke particles was 0.14 um. After receipt of dry WSP, particles were suspended in 1X PBS and disaggregated by sonication and sterilized under UV light for 20 minutes. WSP particle suspension was then added directly to cell culture

media at concentrations of 0.001 mg/ml to 1 mg/ml as indicated in the text and gently agitated to evenly disperse particles.

Interleukin 8 Quantification by Enzyme linked Immunosorbent Assay (ELISA)

Quantification of IL-8 was performed by a commercial ELISA kit (R&D systems). Briefly, plates were prepared with anti-IL-8 capture antibody prior to addition of sample. Cell supernatant was harvested following exposure and applied directly to prepared plates. Anti-IL-8 detection antibody was then added, and IL-8 levels were quantified from streptavidin-Horseradish peroxidase associated with the detection antibody using absorbance read by an Infinite M1000 plate reader (Tecan).

Precision Run-on Sequencing (PRO-seq)

Cells required treatment and harvest on two separate days to generate enough nuclei to perform PRO-seq on two replicates per condition (untreated control (Ctrl) vs APM). For each set of replicates, Sm1 cells (small airway epithelial cells from donor 1) were seeded on 5 x 15 cm double collagen-coated plates per treatment and grown to confluence (3-4 days), then left untreated (Ctrl) or treated with APM (50 µg/cm^2) for 20 hours. Cells were harvested and nuclei prepared as reported previously (18). PRO-seq libraries for all replicates were then prepared simultaneously by subjecting one aliquot of 1e7 nuclei/sample to 3-minute nuclear run-on reactions in the presence of Biotin-11-CTP (PerkinElmer) following our previously detailed protocol (19). Uniquely indexed libraries were pooled and sequenced on an Illumina NextSeq instrument using 75 bp single-end reads by the BioFrontiers Sequencing Facility at the University of Colorado Boulder.

PRO-seq Computational Analysis

PRO-seq data were processed using a standardized Nextflow pipeline (https://github.com/Dowell-Lab/Nascent-Flow). A complete pipeline report detailing all software programs and versions utilized and a detailed quality control report including trimming, mapping, coverage, and complexity metrics are included in Supplemental Files SR1-SR3. TDF coverage files output by the pipeline, normalized by reads per million mapped, were visualized using the Integrative Genomics Viewer (IGV) genome browser (v. 2.8.0). FStitch (v. 1.0) and Tfit (v. 1.0) were used to identify regions with bidirectional transcriptional activity as described (17). Reference Sequences for hg38 were downloaded from the UCSC genome browser (May 18, 2018). DESeq2 (v. 1.20.0, Bioconductor release v. 3.7) was used to determine differentially

transcribed genes between control and APM exposure. Functional annotation of differentially transcribed genes was performed using DAVID v. 6.8 (20). For bidirectional comparisons, all predicted bidirectional Tfit calls were aggregated using mergeBed (argument -d 60) from BEDTools (v. 2.28.0) to generate an annotation file. Counts were then calculated for each sample using multicov (BEDTools v.2.28.0), and DESeq2 was used to identify differentially transcribed bidirectionals between treatments.

Assay for Transposase Accessible Chromatin using Sequencing (ATAC-seq)

Sm1 cells were grown as described above for PRO-seq and aliquots of cells were taken from confluent 15 cm tissue culture dishes after vehicle (PBS) or APM treatment. Sm2 and Sm3 cells were grown to confluence in 6-well tissue culture dishes and treated with vehicle (PBS), APM or WSP for 20 hours. Cells were rinsed and scraped in ice-cold PBS, then ~50K cells from each treatment were pelleted and processed in duplicate for Omni-ATAC-seq using the protocol developed by Corces et al (21). Uniquely indexed libraries were pooled and sequenced on an Illumina NextSeq using 37 bp single- or paired-end reads by the BioFrontiers Sequencing Facility at the University of Colorado Boulder or on an Illumina NovaSeq using 150 bp paired-end reads at the University of Colorado Anschutz Medical Campus Shared Genomic Resource.

ATAC-seq Computational Analysis

Computational analysis of ATAC-seq data was performed as described previously (17). Briefly, ATAC-seq reads were trimmed for adapters, minimum length, and minimum quality using the bbduk tool from the BBMap Suite (v. 38.73) with arguments ‘ref=adapters.fa ktrim=r qtrim=10 k=23 mink=11 hdist=1 maq=10 minlen=20’. Samples sequenced using 150 bp reads were further trimmed to 37 bp in length for consistency using the additional argument 'ftr=36'. Trimmed reads were mapped to the human genome (hg38; downloaded from the UCSC genome browser on September 16, 2019, with corresponding hisat2 index files) using hisat2 (v. 2.1.0). Duplicate reads were removed using picard tools (v.2.26.6; http://broadinstitute.github.io/picard/) MarkDuplicates and deduplicated and sorted BAM files were converted to bedGraph coverage format using genomeCoverageBed from the BEDTools suite (v. 2.29.2). Read coverage was normalized to reads per million mapped using a custom python script and files converted to TDF format using igvtools (v. 2.5.3) for visualization in IGV. MACS2 (v. 2.1.4) callpeak with ‘--SPMR’ argument was applied to sorted BAM files for each replicate pair of samples with parameters set to call peaks with $\log_2$ fold change > 1 above baseline with q < 1e-5 as significant.

Differential Transcription Factor Motif Enrichment

Transcription factor enrichment analysis for both PRO- and ATAC-seq datasets was performed using TFEA (v.1.0; https://github.com/Dowell-Lab/TFEA; (22)). Briefly, sets of consensus regions of interest (ROIs) are first defined for all Tfit-called bidirectionals (for PRO-seq) or all MACS2-called peaks (for ATAC-seq) using the muMerge function of TFEA. TFEA then calculates read coverage for each replicate within each ROI and applies DESeq2 to determine differentially represented ROIs by treatment condition and assign rank as a function of p-value. For each ranked ROI, the FIMO tool (23) from the MEME suite scans the 3 kb sequence surrounding the ROI center (bidirectional origin or ATAC-seq peak summit) for positions of transcription factor consensus motif matches (with p-value cutoff of 1e-5), represented by a curated set of position weight matrices (24). Enrichment scores (E-scores) for each motif are then calculated and corrected for sequence content to reduce known biases associated with local GC enrichment and p-values are determined using Z-scores. Motif displacement distributions are displayed as dot-plots where the location of all motif instances within $\pm$1500 bp are indicated relative to the ROI center on the vertical axis, and the rank of each ROI based on signal loss or gain is indicated on the horizontal axis.

Western Blotting

Western blotting and protein detection were performed using standard protocols (25) in SAECs cultured to confluence in 6-well plates and treated with vehicle control or APM. Primary antibodies purchased included anti-GRHL2 (Sigma HPA004820), anti-pan-TEAD (Cell Signaling Technologies 13295S) and anti-GAPDH (Santa Cruz sc-25778). Secondary antibody was ECL Donkey anti-Rabbit IgG, HRP-linked $F(ab')_2$ fragment (NA9340) from GE Healthcare/Amersham.

RNA purification and quantitative reverse transcription-PCR (qRT-PCR)

RNA purification and qRT-PCR was performed as described previously (17). Briefly, SAECs were grown to confluence in 12-well tissue culture dishes and treated with vehicle (PBS) or APM for 20 hours. Cells were harvested in TRIzol (Life Technologies) and RNA purified by PureLink RNA Mini Kit (Life Technologies) prior to qRT-PCR, performed with normalization to RPL19 as previously detailed (18).Sequences of primers used for qRT-PCR analysis are provided in Supplemental Table ST1.

Statistics

Statistical comparisons for IL-8 ELISA and qRT-PCR assays were made by 2-tailed t-test using the Bonferroni correction when appropriate. These analyses were conducted using statistical software embedded in Open Office (Apache OpenOffice v 4.1.7 http://www.openoffice.org/welcome/credits.html).

All High Throughput Sequencing Data is publicly available through the Gene Expression Omnibus (GEO) with accession GSE201152.

Results

Primary small airway epithelial cells (SAECs) secrete IL-8 in response to APM and WSP exposure in submerged culture

We exposed SAECs to a previously validated concentration of APM (50 µg/cm^2 culture surface area) (16) for 20 hours, based on a previous time course experiment, and demonstrated an increase in interleukin 8 (IL-8) secretion in cells from three healthy non-smoking donors, consistent with previous studies (16) (Figure 1a). We next sought to compare APM exposure with an additional particulate exposure model that resulted in similar inflammatory responses but had different clinical associations. We selected wood smoke particles (WSP), a highly relevant model for wildfire smoke exposure, and measured IL-8 secretion in response to varied doses of WSP exposure (Supplemental Figure SF1). Primary SAECs from the same three donors exhibited a similar increase in IL-8 secretion in response to APM exposure at 50 µg/cm^2 and WSP exposure at 1 mg/ml (Figure 1b). Therefore, we selected these doses of APM (50 µg/cm^2) and WSP (1 mg/ml) for subsequent genome-wide analysis and comparisons.

Nascent transcript sequencing identified sustained transcriptional pathways in response to APM exposure

PRO-seq identifies nascent transcripts with high temporal and spatial resolution, and thus can define both gene transcription and the transcription of short-lived non-coding RNAs, such as enhancer RNAs, which are associated with active transcriptional regulatory elements (TREs). We first employed PRO-seq to characterize gene transcriptional responses to APM in SAECs. Our previous study demonstrated marked temporal variation in transcriptional responses within the first two hours of particulate exposure (26). Therefore, we anticipated nascent transcription after 20 hours of exposure would represent a steady-state response to APM. We performed

PRO-seq on a primary culture of human SAECs, Sm1, after 20 hours of APM exposure and used DESeq2 to identify gene transcription changes induced by APM. We identified 946 genes with increased transcription and 531 genes with decreased transcription in response to APM (Figure 2a). Representative PRO-seq genomic tracks visualized in the IGV Genome Browser for *GAS1* (downregulated) and *CXCL8* (upregulated) are demonstrated in Figure 2b. DAVID functional annotation (20) identified common pathways among upregulated genes that included cell-cell adhesion ($p = 0.0064$) and keratin-associated signaling ($p = 0.00013$). Pathways shared among downregulated genes included cytochrome P450 ($p = 2e\text{-}6$) and glycolysis ($p = 0.00064$). The top 10 functional annotation pathways for downregulated and upregulated genes are depicted in Figure 2c. Thus, aligned with prior reports of particulate exposures, APM exposure causes significant changes in gene expression associated with specific biologic processes.

Nascent enhancer RNA transcripts identify transcriptional regulatory elements and transcription factors that mediate APM exposure

To determine regulatory mechanisms underpinning changes in gene expression caused by APM, we used Transcription-fit (Tfit) (27) to interrogate PRO-seq transcripts for enhancer RNA (eRNA) signatures, which demarcate transcriptional regulatory elements (TREs). TRE activity is estimated by quantification of eRNA transcripts at the specific locus. Differential expression analysis of 18,214 TREs identified 207 TREs with significantly increased eRNA transcription ($p_{adj} < 0.05$) and 310 TREs with decreased eRNA transcription (Figure 3a). Representative TREs neighboring *AKR1C1* (downregulated) and *MICAL2* (upregulated) are demonstrated in Figure 3b.

We next used Transcription Factor Enrichment Analysis (TFEA) to identify transcription factor binding motifs associated with gain or loss in TRE activity after APM exposure. We identified 19 transcription factor binding motifs associated with increased TRE activity (enriched) and 8 motifs associated with decreased TRE activity (depleted) (Figure 3c). Dot plots depict co-clustering of EGR1 motifs with enriched features (increased TRE activity) and P53 motifs with depleted features (decreased TRE activity) (Figure 3d). Changes in APM-regulated TRE activity are therefore associated with both enrichment and depletion of motifs for specific transcription factors.

ATAC-seq (Assay for Transposase Accessible Chromatin with sequencing) identified chromatin accessibility changes induced by APM exposure

Dynamic changes in TRE activity are frequently associated with chromatin structure alterations. To examine how APM modifies chromatin structure in small airway epithelial cells, we used ATAC-seq to profile chromatin accessibility in vehicle and APM treated Sm1 cells. We identified 203,838 discrete genomic loci of chromatin accessibility in Sm1 cells with and without APM exposure. To determine whether APM exposure alters chromatin accessibility on a genome-wide basis, we compared the pre- and post- exposure ATAC-seq datasets using TFEA. With this analysis, we identified 128 motifs that were significantly enriched after APM in comparison to vehicle, whereas 101 motifs were depleted ($p_{adj} < 1e-5$). The top 10 differential motifs based on Enrichment scores are depicted in Figure 4c. In Figure 4d, these data are visualized for two specific motifs using dot plots, which depict the co-clustering of the TEAD3 motif with enriched chromatin accessibility features and GRHL2 with depleted features, implying gain of TEAD3 and loss of GRHL2 activity with respect to epigenetic remodeling.

ATAC-seq profiling distinguishes exposure- and donor-specific transcriptional responses

To determine whether epigenetic responses to particulate matter in SAECs are stereotyped or instead may be influenced by donor or particulate characteristics, we performed ATAC-seq on SAECs from two additional donors after 20 hours of exposure to APM or WSP. Differential accessibility analysis demonstrated generally stereotyped responses to exposures (Figure 5a). We again employed TFEA to identify transcription factor binding motifs associated with differential accessibility (Figure 5b). Complete TFEA results for all transcription factor binding motifs are listed in supplemental tables ST2 – ST5. Select transcription factor binding motif patterns are demonstrated as dot plots in Figure 5c. P53 motif depletion was shared between both exposures and both SAECs. TEAD3 enrichment was demonstrated in both Sm2 and Sm3 in response to APM exposure but not WSP. GRHL2 was depleted in only Sm3 cells exposed to APM. Thus, the SAEC response to particulate exposure differs based on both the donor and the underlying particulate composition, and TEAD3 activity is implicated as a distinguishing feature in epigenetic responses to APM.

Integrating genomics approaches identifies TRE targets of APM activity

To further investigate potential regulatory roles for transcription factors identified by TFEA, such as TEAD3, GRHL2 and P53, we first performed Western Blot on whole cell lysates from Sm2 and Sm3 cells exposed to APM. TEAD3 and GRHL2 were expressed at the protein level,

supporting roles in epigenetic and transcriptional responses to APM in these cell types (Figure 6a). Changes in expression levels in response to exposure did not directly correlate with effects on motif Enrichment scores, suggesting the exposures modulate factor activity post-transcriptionally. Similarly, although P53 was expressed at the mRNA level, changes in expression in response to APM did not correlate with Enrichment scores (Figure 6b). However, gene expression of three canonical P53 target genes, *TP53INP1*, *SESN1* and *DDIT4*, decreased following APM exposure, correlating with the depletion of the P53 motif in APM treated cells, and gene expression of four canonical TEAD targets, *CCN1*, *CCN2*, *MYC* and *FGFBP1* increased, correlating with TEAD3 motif enrichment (Figure 6b). In aggregate, these data suggest post-transcriptional regulation of TEAD and P53 activity by APM, a feature of both TEAD and P53 factors that has been reported previously (28, 29), and suggest an important role for post-transcriptional control of epigenetic responses to exposures.

To further investigate roles of P53 and TEAD3 in APM responses, we interrogated PRO-seq identified TREs neighboring P53 and TEAD3 target genes using the JASPAR library of transcription factor binding motifs (30). We identified canonical P53 binding sequences in TREs which, based on congruent activity between the TRE and neighboring genes, are implicated in directly regulating transcription of *TP53INP1* and *SESN1* (Figure 6c). We also identified TEAD3 sequences that colocalize with TREs linked to control of *CCN1* and *CCN2* transcription (Figure 6d). These data further support the roles of P53 and TEAD3 in mediating APM responses and demonstrate the application of integrated transcriptional genomics to identify transcription factors and specific genomic regions that mediate responses to APM exposure.

Discussion

The mechanistic basis for the development of deployment- and exposure-associated respiratory disease is poorly understood. Here, we approached this problem by exposing primary SAECs to APM, a highly relevant deployment-specific exposure, and characterizing the genomic responses. We identified several pathways that may be involved in the development of DLD and potential neoplastic implications of deployment exposures, such as loss of P53, GRHL2 or cytochrome P450 signaling. Furthermore, we identified transcription factor pathways that mediate either transcriptional responses, such as the EGR family, or chromatin accessibility change, such as the TEAD family. We identified pathways specific to APM exposure by comparing effects of APM and WSP exposure, with TEAD activation, a pathway that had not previously associated with particulate exposures, as specific to APM. In aggregate, our data

indicate that APM causes epigenetic changes in airway epithelial cells as a consequence of changes in transcription factor activity, and defines specific effects of APM in comparison to WSP.

We identified several response pathways in SAECs exposed to APM that may affect disease susceptibility to additional insults, supporting a multi-hit hypothesis for the development of DLD. APM exposure resulted in downregulation of cytochrome P450 signaling, which is responsible for catabolism of polyaromatic compounds (31). Even transient loss of this program can result in increased susceptibility to additional environmental exposures, as increased polyaromatic hydrocarbon composition in particulate exposures is associated with increased cytotoxicity (32) and several polyaromatic compounds are known carcinogens (33). Thus, a loss of cytochrome activity may predispose to aberrant responses and additional injury in response to repetitive exposures, which are typically a feature of deployment-related exposures.

Abnormal p53 signaling responses, as evidenced by our data indicating a loss of accessible chromatin at p63/p53 binding motifs after APM exposure, may also promote impaired cellular responses to further insults. Reduced p53 activity has previously been implicated in distal bronchiolar pathology and a failure of apoptosis in response to RSV infection (34), a common cause of acute bronchiolitis. Moreover, p63 expression is reduced in constrictive bronchiolitis arising from mustard gas exposure (35). Although additional studies are needed to establish a causal relationship, our work suggests that reduced p63/p53 activity may similarly predispose to bronchiolitis in the setting of APM inhalation. Our data also implicate a loss of GRHL2 in association with APM, and GRHL2 has been demonstrated to work with p63/p53 in determining epithelial cell fate at the epigenetic level need citation (36). Although additional studies are needed to determine a causal role for loss of p53 signaling and GRHL2 in relation to DLD, our data suggest that increasing p53 activity via agents such as idasanutlin, which is currently in clinical development for myeloid cancers, may have therapeutic utility in DLD.

Our data also indicate APM-mediated activation of the TEAD family of transcription factors. This novel finding may represent a unique feature of APM exposure or composition, as WSP exposure did not result in TEAD enrichment in chromatin accessibility features. TEAD transcription factors are classically identified as downstream effectors of the Hippo-YAP-TAZ signal transduction pathway that promote development and cell proliferation (37). Several small molecule compounds exist that can activate or block Hippo pathway signaling (38, 39), and the

TEAD family may represent another potential therapeutic target for prevention or treatment of DLD. Ultimately, further investigation is necessary to characterize the effects of each of these pathways in development of cellular pathophysiology or disease. However, our novel application of genomic methods has identified several interesting candidates, including pathways specific to deployment exposures.

The complexity of exposure biology is a major limitation in the study of exposure-associated health impacts, including the mechanisms underlying DLD. Several factors contribute to this, including the complexity of the exposure interface with the respiratory system and the breadth of responses in exposed tissue. Our integrated genomics-based method using a simple model exposure system establishes a basic discovery system to identify pathways that are potentially relevant to DLD-related disease. Our comprehensive approach has several advantages in the study of complex airborne hazards. ATAC-seq and PRO-seq reads map to all regions of the genome, including intergenic regions that contain a majority of regulatory elements. Our approach also quantifies the activity and epigenetic changes within these elements, enabling unbiased bioinformatics-based identification of driving transcriptional regulators that mediate these changes. Supporting the translational potential of our methods, a number of the pathways (e.g. P53) we defined had previously been linked to similar lung pathologies in human samples.

Although our study was not powered for detecting individual-level variations in exposure responses, we demonstrated variation in genomic response patterns between SAECs from different donors. For example, GRHL2 loss was present in Sm1 and Sm3 but not Sm2. This variation may depend on several factors, such as age, sex, systemic disease or exposure history. Future studies that incorporate donor-level data can investigate the effect of systemic factors or genetic variants on responses to APM exposure. Finally our approach also detected differences between the genomic responses to APM and WSP exposures. It is tempting to speculate that these differences we defined in our reductionist model portend the well-known clinical differences in response to deployment-associated particulates and other more commons sources of PM. Despite these advantages, our model system had several limitations. SAECs were cultured and exposed in submerged cell culture without differentiation into a cuboidal ciliated epithelium, which is characteristic of small airways (40). The dose and composition of exposures may also diverge between our model and true inhalational exposures experienced by military personnel deployed to Southwest Asia. However, despite these limitations, our study identified several plausible and novel features of APM exposure that could only be identified

through the use of sophisticated genomics tools in our model system. Extending our findings to *in vivo* models and even more sophisticated *in vitro* systems, such as precision cut lung slices, will be crucial for validating the novel APM-associated pathways we have reported here.

References

1. **Krefft SD**, **Meehan R**, **Rose CS**. Emerging spectrum of deployment-related respiratory diseases. *Curr. Opin. Pulm. Med.* 21 Lippincott Williams and Wilkins: 185–192, 2015.
2. **Krefft SD**, **Rose CS**, **Nawaz S**, **Miller YE**. Deployment-Related Lung Disorders. [Online]. *Fed Pract* 32: 32–38, 2015. http://www.ncbi.nlm.nih.gov/pubmed/30766070 [24 Nov. 2020].
3. **Kim KH**, **Kabir E**, **Kabir S**. A review on the human health impact of airborne particulate matter. *Environ Int* 74: 136–143, 2015. doi: 10.1016/J.ENVINT.2014.10.005.
4. **Li R**, **Zhou R**, **Zhang J**. Function of PM2.5 in the pathogenesis of lung cancer and chronic airway inflammatory diseases. *Oncol. Lett.* 15 Spandidos Publications: 7506–7514, 2018.
5. **Hamanaka RB**, **Mutlu GM**. Particulate Matter Air Pollution: Effects on the Cardiovascular System. *Front Endocrinol (Lausanne)* 9: 680, 2018. doi: 10.3389/FENDO.2018.00680.
6. **Lee BJ**, **Kim B**, **Lee K**. Air Pollution Exposure and Cardiovascular Disease. *Toxicol Res* 30: 71, 2014. doi: 10.5487/TR.2014.30.2.071.
7. **Calderón-Garcidueñas L**, **Stommel EW**, **Rajkumar RP**, **Mukherjee PS**, **Ayala A**. Particulate Air Pollution and Risk of Neuropsychiatric Outcomes. What We Breathe, Swallow, and Put on Our Skin Matters. *Int J Environ Res Public Health* 18: 11568, 2021. doi: 10.3390/IJERPH182111568.
8. **Ghio AJ**, **Soukup JM**, **Case M**, **Dailey LA**, **Richards J**, **Berntsen J**, **Devlin RB**, **Stone S**, **Rappold A**. Exposure to wood smoke particles produces inflammation in healthy volunteers. *Occup Environ Med* 69: 170–175, 2012. doi: 10.1136/oem.2011.065276.
9. **Ghio AJ**, **Soukup JM**, **Dailey LA**, **Tong H**, **Kesic MJ**, **Budinger GRS**, **Mutlu GM**. Wood Smoke Particle Sequesters Cell Iron to Impact a Biological Effect. *Chem Res Toxicol* 28: 2104–11, 2015. doi: 10.1021/acs.chemrestox.5b00270.
10. **Memon TA**, **Nguyen ND**, **Burrell KL**, **Scott AF**, **Almestica-Roberts M**, **Rapp E**, **Deering-Rice CE**, **Reilly CA**. Wood Smoke Particles Stimulate MUC5AC Overproduction by Human Bronchial Epithelial Cells Through TRPA1 and EGFR Signaling. *Toxicol Sci* 174: 278–290, 2020. doi: 10.1093/toxsci/kfaa006.
11. **Cao Y**, **Chen M**, **Dong D**, **Xie S**, **Liu M**. Environmental pollutants damage airway

epithelial cell cilia: Implications for the prevention of obstructive lung diseases. *Thorac Cancer* 11: 505, 2020. doi: 10.1111/1759-7714.13323.

12. **Kyung SY**, **Jeong SH**. Particulate-matter related respiratory diseases. *Tuberc. Respir. Dis. (Seoul).* 83 Korean National Tuberculosis Association: 116–121, 2020.
13. **Chen YC**, **Chuang TY**, **Liu CW**, **Liu CW**, **Lee TL**, **Lai TC**, **Chen YL**. Particulate matters increase epithelial-mesenchymal transition and lung fibrosis through the ETS-1/NF-κB-dependent pathway in lung epithelial cells. *Part Fibre Toxicol* 17: 1–14, 2020. doi: 10.1186/S12989-020-00373-Z/FIGURES/8.
14. **Maes T**, **Provoost S**, **Lanckacker EA**, **Cataldo DD**, **Vanoirbeek JAJ**, **Nemery B**, **Tournoy KG**, **Joos GF**. Mouse models to unravel the role of inhaled pollutants on allergic sensitization and airway inflammation. *Respir Res 2010 111* 11: 1–16, 2010. doi: 10.1186/1465-9921-11-7.
15. **Berman R**, **Kopf KW**, **Min E**, **Huang J**, **Downey GP**, **Alam R**, **Chu HW**, **Day BJ**. IL-33/ST2 signaling modulates Afghanistan particulate matter induced airway hyperresponsiveness in mice. *Toxicol Appl Pharmacol* 404: 115186, 2020. doi: 10.1016/j.taap.2020.115186.
16. **Berman R**, **Downey GP**, **Dakhama A**, **Day BJ**, **Chu HW**. Afghanistan Particulate Matter Enhances Pro-Inflammatory Responses in IL-13-Exposed Human Airway Epithelium via TLR2 Signaling. *Toxicol Sci* 166: 345–353, 2018. doi: 10.1093/TOXSCI/KFY217.
17. **Gupta A**, **Sasse SK**, **Gruca MA**, **Sanford L**, **Dowell RD**, **Gerber AN**. Deconvolution of multiplexed transcriptional responses to wood smoke particles defines rapid aryl hydrocarbon receptor signaling dynamics. *J Biol Chem* 0: 101147, 2021. doi: 10.1016/J.JBC.2021.101147.
18. **Sasse SK**, **Gruca M**, **Allen MA**, **Kadiyala V**, **Song T**, **Gally F**, **Gupta A**, **Pufall MA**, **Dowell RD**, **Gerber AN**. Nascent transcript analysis of glucocorticoid crosstalk with TNF defines primary and cooperative inflammatory repression. *Genome Res* 29: 1753–1765, 2019. doi: 10.1101/gr.248187.119.
19. **Gally F**, **Sasse SK**, **Kurche JS**, **Gruca MA**, **Cardwell JH**, **Okamoto T**, **Chu HW**, **Hou X**, **Poirion OB**, **Buchanan J**, **Preissl S**, **Ren B**, **Colgan SP**, **Dowell RD**, **Yang I V.**, **Schwartz DA**, **Gerber AN**. The MUC5B-associated variant rs35705950 resides within an enhancer subject to lineage- and disease-dependent epigenetic remodeling. *JCI Insight* 6, 2021. doi: 10.1172/JCI.INSIGHT.144294.
20. **Huang DW**, **Sherman BT**, **Lempicki RA**. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4: 44–57, 2009. doi:

10.1038/nprot.2008.211.

21. **Corces MR**, **Trevino AE**, **Hamilton EG**, **Greenside PG**, **Sinnott-Armstrong NA**, **Vesuna S**, **Satpathy AT**, **Rubin AJ**, **Montine KS**, **Wu B**, **Kathiria A**, **Cho SW**, **Mumbach MR**, **Carter AC**, **Kasowski M**, **Orloff LA**, **Risca VI**, **Kundaje A**, **Khavari PA**, **Montine TJ**, **Greenleaf WJ**, **Chang HY**. An improved ATAC-seq protocol reduces background and enablesinterrogation of frozen tissues. *Nat Methods* 14: 959, 2017. doi: 10.1038/NMETH.4396.
22. **Rubin JD**, **Stanley JT**, **Sigauke RF**, **Levandowski CB**, **Maas ZL**, **Westfall J**, **Taatjes DJ**, **Dowell RD**. Transcription factor enrichment analysis (TFEA) quantifies the activity of multiple transcription factors from a single experiment. *Commun Biol 2021 41* 4: 1–15, 2021. doi: 10.1038/s42003-021-02153-7.
23. **Grant CE**, **Bailey TL**, **Noble WS**. FIMO: Scanning for occurrences of a given motif. *Bioinformatics* 27: 1017–1018, 2011. doi: 10.1093/bioinformatics/btr064.
24. **Lambert SA**, **Jolma A**, **Campitelli LF**, **Das PK**, **Yin Y**, **Albu M**, **Chen X**, **Taipale J**, **Hughes TR**, **Weirauch MT**. The Human Transcription Factors. *Cell* 172 Cell Press: 650–665, 2018.
25. **Altonsy MO**, **Sasse SK**, **Phang TL**, **Gerber AN**. Context-dependent cooperation between nuclear factor κB (NF-κB) and the glucocorticoid receptor at a TNFAIP3 intronic enhancer: A mechanism to maintain negative feedback control of inflammation. *J Biol Chem* 289: 8231–8239, 2014. doi: 10.1074/jbc.M113.545178.
26. **Gupta A**, **Sasse SK**, **Gruca MA**, **Sanford L**, **Dowell RD**, **Gerber AN**. Deconvolution of multiplexed transcriptional responses to wood smoke particles defines rapid aryl hydrocarbon receptor signaling dynamics. *J Biol Chem* 297: 101147, 2021. doi: 10.1016/J.JBC.2021.101147.
27. **Azofeifa JG**, **Dowell RD**. A generative model for the behavior of RNA polymerase. *Bioinformatics* 33: 227–234, 2017. doi: 10.1093/bioinformatics/btw599.
28. **Zhang J**, **Chen X**. Posttranscriptional Regulation of p53 and Its Targets by RNA-Binding Proteins. *Curr Mol Med* 8: 845, 2008. doi: 10.2174/1566524087867337 48.
29. **Huh HD**, **Kim DH**, **Jeong H-S**, **Park HW**. Regulation of TEAD Transcription Factors in Cancer Biology. *Cells* 8: 600, 2019. doi: 10.3390/CELLS8060600.
30. **Castro-Mondragon JA**, **Riudavets-Puig R**, **Rauluseviciute I**, **Berhanu Lemma R**, **Turchi L**, **Blanc-Mathieu R**, **Lucas J**, **Boddie P**, **Khan A**, **Perez NM**, **Fornes O**, **Leung TY**, **Aguirre A**, **Hammal F**, **Schmelter D**, **Baranasic D**, **Ballester B**, **Sandelin A**, **Lenhard B**, **Vandepoele K**, **Wasserman WW**, **Parcy F**, **Mathelier A**. JASPAR 2022:

The 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* 50: D165–D173, 2022. doi: 10.1093/NAR/GKAB1113.

31. **Shimada T**, **Fujii-Kuriyama Y**. Metabolic activation of polycyclic aromatic hydrocarbons to carcinogens by cytochromes P450 1A1 and 1B1. *Cancer Sci* 95: 1–6, 2004. doi: 10.1111/J.1349-7006.2004.TB03162.X.
32. **Wang B**, **Li N**, **Deng F**, **Buglak N**, **Park G**, **Su S**, **Ren A**, **Shen G**, **Tao S**, **Guo X**. Human bronchial epithelial cell injuries induced by fine particulate matter from sandstorm and non-sandstorm periods: Association with particle constituents. *J Environ Sci (China)* 47: 201–210, 2016. doi: 10.1016/J.JES.2015.12.015.
33. **Hecht SS**. Potential Carcinogenic Effects of Polynuclear Aromatic Hydrocarbons and Nitroaromatics in Mobile Source Emissions - Air Pollution, the Automobile, and Public Health - NCBI Bookshelf [Online]. [date unknown]. https://www.ncbi.nlm.nih.gov/books/NBK218134/ [11 Apr. 2022].
34. **Eckardt-Michel J**, **Lorek M**, **Baxmann D**, **Grunwald T**, **Keil GM**, **Zimmer G**. The Fusion Protein of Respiratory Syncytial Virus Triggers p53-Dependent Apoptosis. *J Virol* 82: 3236, 2008. doi: 10.1128/JVI.01887-07.
35. **Ghanei M**, **Chilosi M**, **Akbari HMH**, **Motiei-Langroudi R**, **Harandi AA**, **Shamsaei H**, **Bahadori M**, **Tazelaar HD**. Use of Immunohistochemistry Techniques in Patients Exposed to Sulphur Mustard Gas. *Patholog Res Int* 2011: 1–7, 2011. doi: 10.4061/2011/659603.
36. **Latil M**, **Nassar D**, **Beck B**, **Boumahdi S**, **Wang L**, **Brisebarre A**, **Dubois C**, **Nkusi E**, **Lenglez S**, **Checinska A**, **Vercauteren Drubbel A**, **Devos M**, **Declercq W**, **Yi R**, **Blanpain C**. Cell-Type-Specific Chromatin States Differentially Prime Squamous Cell Carcinoma Tumor-Initiating Cells for Epithelial to Mesenchymal Transition. *Cell Stem Cell* 20: 191-204.e5, 2017. doi: 10.1016/J.STEM.2016.10.018.
37. **Lin KC**, **Park HW**, **Guan KL**. Regulation of the Hippo pathway transcription factor TEAD. *Trends Biochem Sci* 42: 862, 2017. doi: 10.1016/J.TIBS.2017.09.003.
38. **Aihara A**, **Iwawaki T**, **Abe-Fukasawa N**, **Otsuka K**, **Saruhashi K**, **Mikashima T**, **Nishino T**. Small molecule LATS kinase inhibitors block the Hippo signaling pathway and promote cell growth under 3D culture conditions. *J Biol Chem* 298: 101779, 2022. doi: 10.1016/J.JBC.2022.101779/ATTACHMENT/46711C20-8AFE-46B9-AC3A-A429C72D720A/MMC1.PDF.
39. **Crawford JJ**, **Bronner SM**, **Zbieg JR**. Hippo pathway inhibition by blocking the YAP/TAZ-TEAD interface: a patent review. *Expert Opin Ther Pat* 28: 867–873, 2018. doi:

10.1080/13543776.2018.1549226.

40. **van den Bosch WB**, **James AL**, **Tiddens HAWM**. Structure and function of small airways in asthma patients revisited. *Eur Respir Rev* 30: 1–18, 2021. doi: 10.1183/16000617.0186-2020.

40**.** **Bushnell B.** BBMap: A Fast, Accurate, Splice-Aware Aligner [Online]. *United States N. p., 2014*: [date unknown]. https://www.osti.gov/biblio/1241166-bbmap-fast-accurate-splice-aware-aligner [9 Aug. 2021].

41. **Andrews S.** FastQC: A Quality Control Tool for High Throughput Sequence Data [Online] [Online]. 2010. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ [9 Aug. 2021].

42. **Azofeifa J, Allen MA, Lladser M, Dowell R.** FStitch: A fast and simple algorithm for detecting nascent RNA transcripts. In: *ACM BCB 2014 - 5th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics*. Association for Computing Machinery, Inc, p. 174–183.

43. **Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP.** Integrative genomics viewer. *Nat. Biotechnol.* 29 NIH Public Access: 24–26, 2011.

44. **Quinlan AR, Hall IM.** BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 26: 841–842, 2010. doi: 10.1093/bioinformatics/btq033.

45. **Kim D, Paggi JM, Park C, Bennett C, Salzberg SL.** Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* 37: 907–915, 2019. doi: 10.1038/s41587-019-0201-4.

46. **Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R.** The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25: 2078–2079, 2009. doi: 10.1093/bioinformatics/btp352.

47. **Liao Y, Smyth GK, Shi W.** FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30: 923–930, 2014. doi: 10.1093/bioinformatics/btt656.

48. **Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS.** Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9: R137, 2008. doi: 10.1186/gb-2008-9-9-r137.

49. **Love MI**, **Huber W**, **Anders S**. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15, 2014. doi: 10.1186/S13059-014-0550-8.

Figure Legends

Figure 1 – Particulate exposure induces IL-8 secretion in SAECs derived from three independent healthy donors. ELISA analysis of IL-8 concentration in the supernatant of Sm1, Sm2 and Sm3 SAECs exposed to (A) APM (50 ug/cm^2 of well surface area) or (B) WSP (1 mg/ml of cell culture media) for 20 hours in submerged cell culture. Bars indicate mean IL-8 concentration ($\pm$SD); n $\geq$ 3/group, *p < 1e-3 for indicated comparisons.

Figure 2 – PRO-seq reveals sustained transcriptional programs associated with APM exposure. (A) Volcano plot illustrates differentially regulated nascent gene transcripts in Sm1 cells treated with APM for 20 hours. Each point represents one gene, with red indicating p_{adj} < 0.05, light orange representing $\log_2$ fold change > 1, green indicating both p_{adj} < 0.05 and $\log_2$ fold change > 1 and black indicating neither condition was satisfied. The number of genes meeting both criteria (green dots) that were down- or upregulated is shown. (B) Representative examples of downregulated (left) and upregulated (right) genes shown as PRO-seq tracks visualized in the Integrative Genomics Viewer (IGV) genome browser based on counts per million mapped reads (vertical scales). Positive (blue) indicates reads annotated to the sense strand while negative (red) data reflect reads annotated to the antisense strand. The TSS and direction of transcription are indicated by arrows at the top of each panel. (C) Bar graphs display top 10 most significantly enriched functional annotation terms output by DAVID Functional Annotation Clustering applied to downregulated (left) and upregulated (right) transcripts.

Figure 3 – TREs defined by PRO-seq bidirectional signatures identify transcription factors that respond to APM exposure. (A) Volcano plot illustrates differentially regulated TRE transcripts in Sm1 cells treated with APM for 20 hours. Each point represents one TRE, with green signifying both p_{adj} < 0.05 and $\log_2$ fold change > 1, and the number of TREs meeting both criteria (green dots) that were down- or upregulated indicated. (B) IGV-visualized PRO-seq tracks of Tfit-called bidirectionals (boxed in black and magnified beside each panel) and nearest genes exhibiting a similar pattern of regulation as representative examples of downregulated and upregulated TREs. (C) Transcription factor binding motifs identified by TFEA. Depleted motifs are encountered more frequently in TREs with loss of transcriptional signal while enriched motifs are encountered more frequently in TREs with gain in transcriptional signal. Motifs listed meet the threshold of p_{adj} < 1e-5 and are ranked by Enrichment (E) score, a relative measure of motif enrichment within TREs. (D) Representative dot-plots for one enriched (EGR1) and one depleted (TP53) motif depict co-clustering of enriched motifs with TRE signal gain or

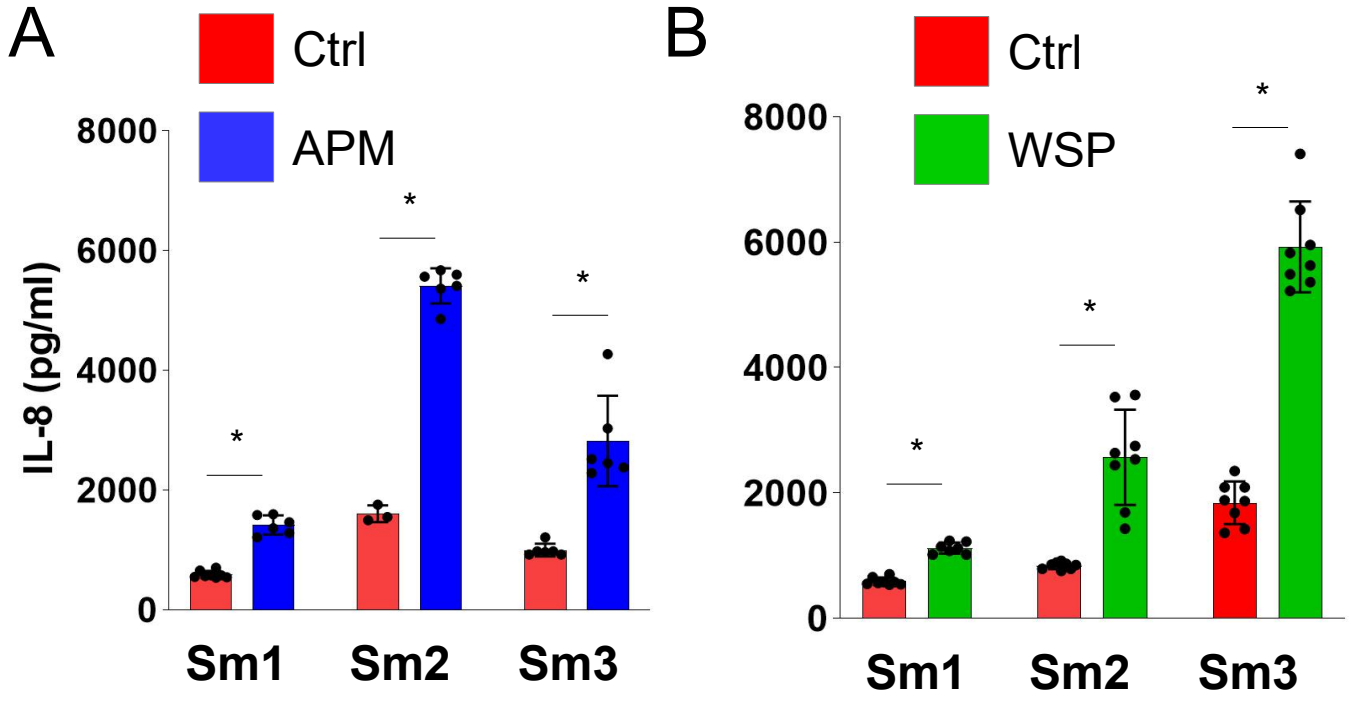
depleted motifs with TRE signal loss. Each dot represents a motif instance found ±1500 bp of a TRE. The horizontal axis represents each TRE ranked from greatest signal loss (left) to greatest signal gain (right) and the vertical axis represents the distance between the motif instance and the center of the TRE. E score and p_{adj} value for comparison between control and APM exposure for each motif is listed.

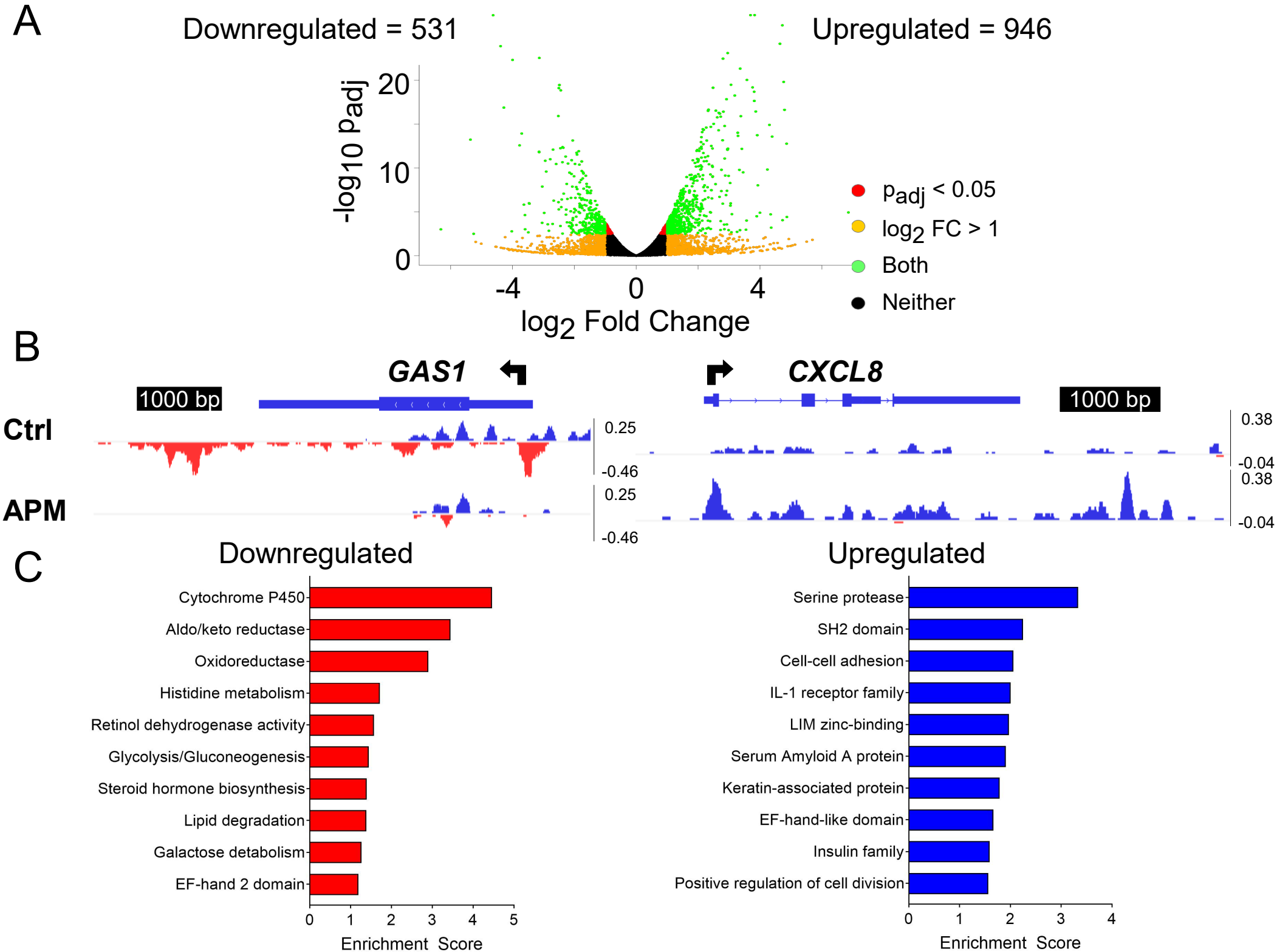
Figure 4 – ATAC-seq identifies transcription factors associated with chromatin accessibility changes in response to APM exposure. (A) Pyramid plot illustrates genomic regions of chromatin accessibility ranked by accessibility change in response to APM exposure. Each point represents one discrete region of chromatin accessibility, with a heatmap of colors depicting the degree of accessibility change. The most significant accessibility gain is red, no accessibility change is yellow and most significant accessibility loss is green. (B) TFEA plot demonstrates transcription factor binding motif enrichment in chromatin accessibility features. Each motif within the reference set is an individual point. Events represent number of instances of each motif in the entire dataset of chromatin accessibility features on a log scale. Here, E score is a quantitative assessment of motif enrichment within accessibility regions, with a positive enrichment score representing motif enrichment and negative E score representing motif depletion. Motifs achieving statistical significance ($p_{adj} < 1e-5$) are circled in red. (C) Bar graphs display the top 10 depleted or enriched transcription factor binding motifs identified by TFEA and ranked by E score. (D) Representative dot-plots for one enriched (TEAD3) and one depleted (GRHL2) motif identified by TFEA performed on ATAC-seq-identified chromatin accessibility features. Each dot represents a motif encountered within ±1500 bp of an accessibility region, ranked by accessibility gain on the horizontal axis and distance to accessibility region center on the vertical axis. E score and p_{adj} value for comparison between control and APM exposure for each motif is listed.

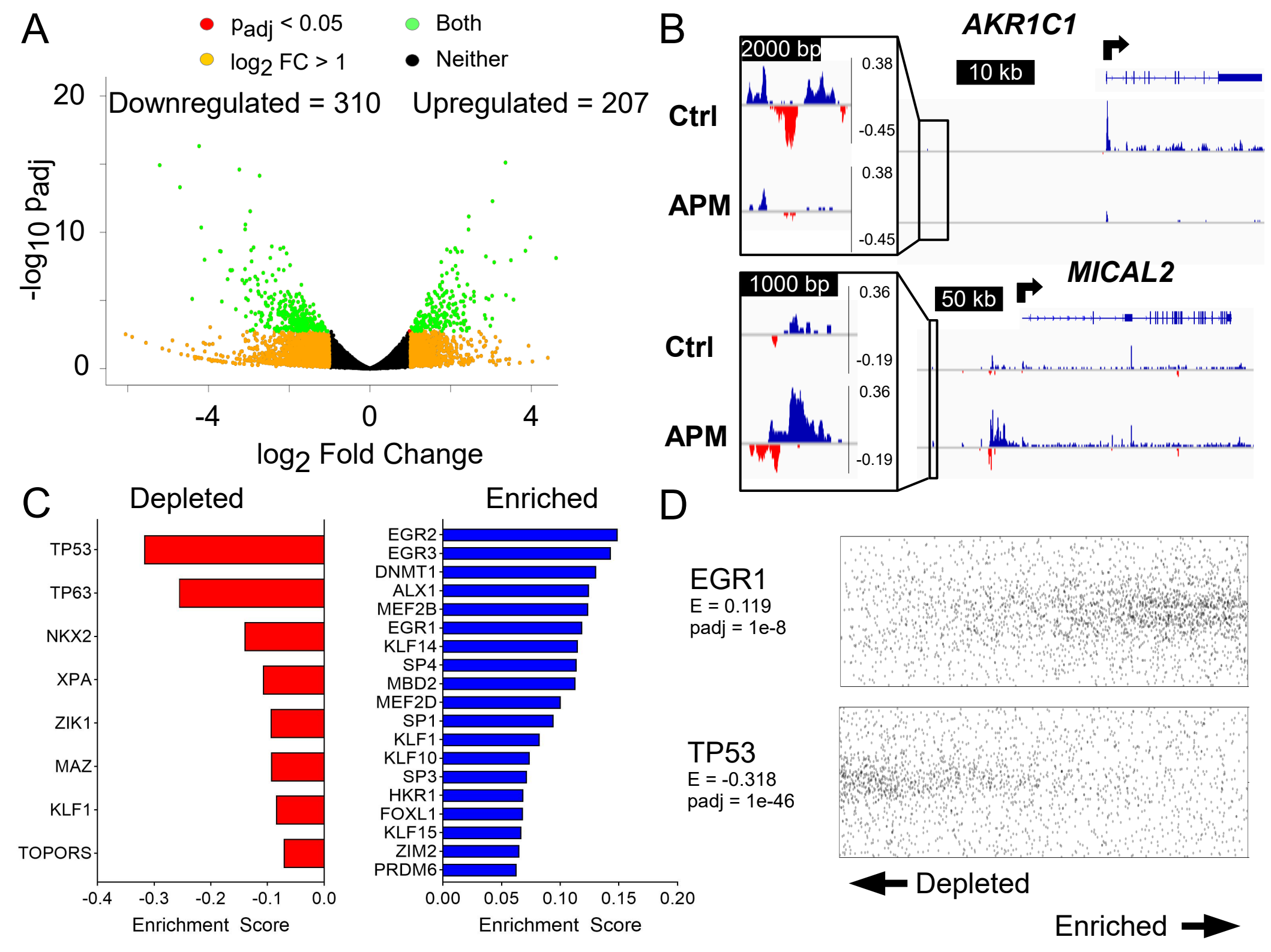
Figure 5 – Differential ATAC-seq profiles following APM and WSP exposure in two additional cell lines identify exposure- and individual-specific responses. (A) Pyramid plots illustrate genomic regions of chromatin accessibility change in Sm2 and Sm3 cells exposed to APM and WSP. Each point is a genomic region of chromatin accessibility that is colored on a red, yellow and green heatmap to demonstrate rank of chromatin accessibility change in response to exposure. (B) TFEA plots demonstrate transcription factor binding motifs associated with accessibility changes. Each point represents an individual motif within the reference set with significantly enriched or depleted motifs circled in red, based on $p_{adj} < 1e-5$.

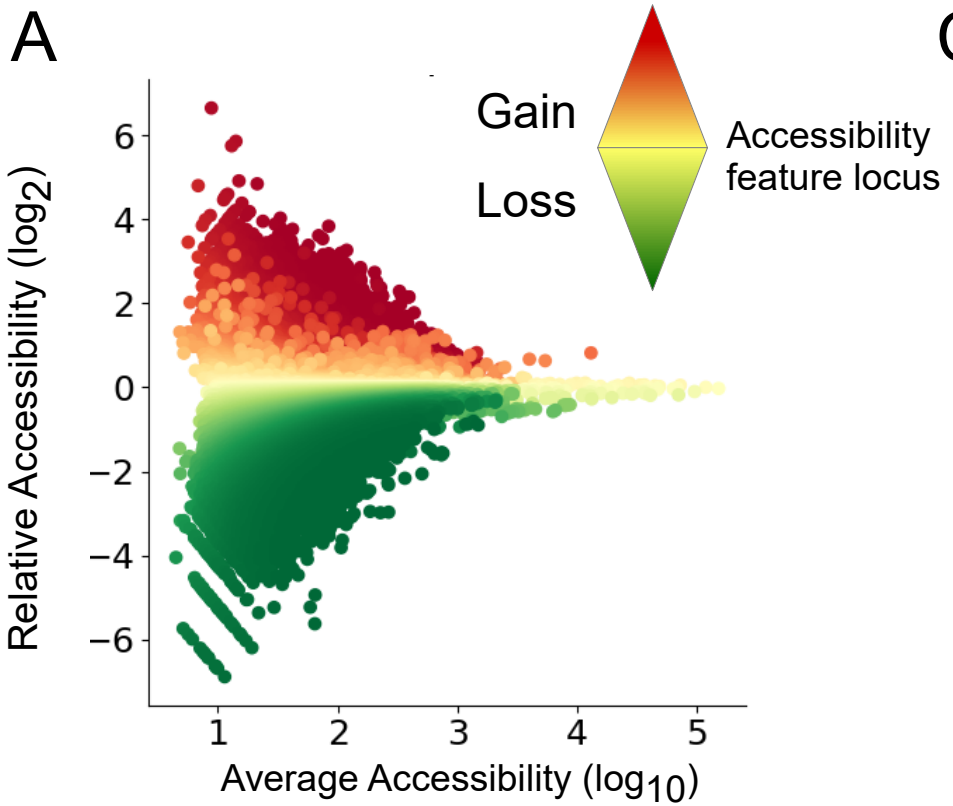
Number of significantly enriched or depleted motifs are listed. (C) Dot plots for selected transcription factor binding motifs. Each dot represents a motif encountered within ±1500 bp of an accessibility region, ranked by accessibility gain on the horizontal axis and distance to accessibility region center on the vertical axis. Depleted motifs are highlighted in red and enriched motifs are highlighted in blue. E score and p_{adj} value for comparison between control and exposure for each motif is listed.

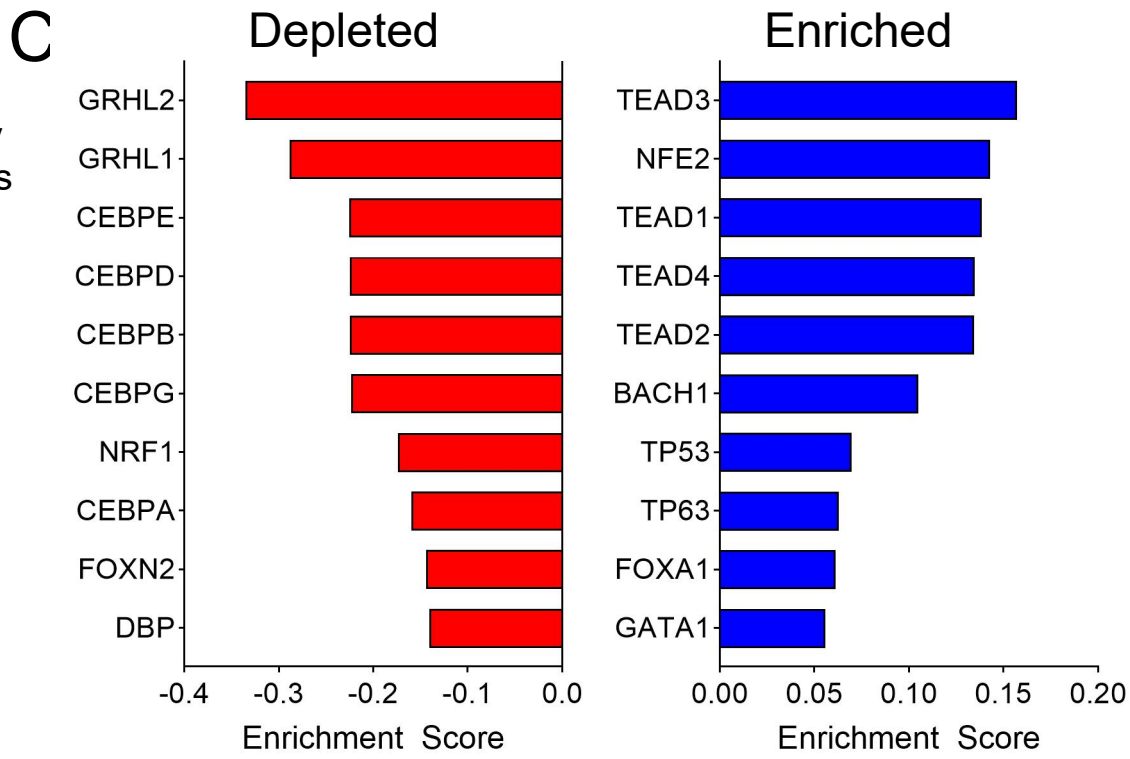
Figure 6 – Transcriptional targets of TFEA-identified transcription factors are regulated by APM-induced TREs. (A) Western blot demonstrates expression of TEAD3 and GRHL2 proteins in control and APM-exposed Sm2 and Sm3 cells; GAPDH served as a loading control. (B) qRT-PCR analysis of indicated gene expression in Sm2 or Sm3 cells following APM exposure, including canonical P53 targets *TP53INP1, SESN1* and *DDIT4,* and canonical TEAD targets *CCN1*, *CCN2*, *MYC* and *FGFBP1*. Bars represent mean normalized C_T values on a $\log_2$ scale (±SD) relative to vehicle-treated controls; n = 4/group, *$p < 0.05$ vs. vehicle. (C) IGV-visualized PRO-seq tracks of Tfit-called TREs neighboring P53 target genes (boxed in black and magnified beside each panel) exhibit a similar pattern of regulation and harbor canonical P53 binding sequences. (D) IGV-visualized PRO-seq tracks of Tfit-called TREs neighboring TEAD3 target genes (boxed in black and magnified beside each panel) exhibit a similar pattern of regulation and harbor canonical TEAD3 binding sequences.

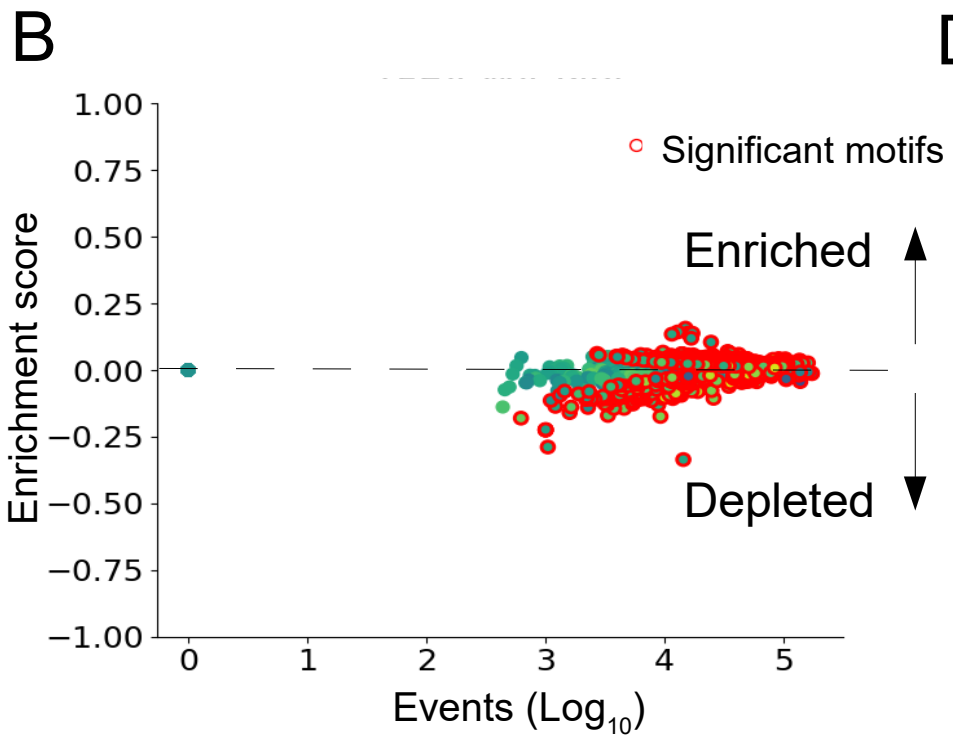

A
Ctrl
APM
IL-8 (pg/ml)
8000
6000
4000
2000
0
*
*
*
Sm1
Sm2
Sm3
B
Ctrl
WSP
8000
6000
4000
2000
0
*
*
*
Sm1
Sm2
Sm3

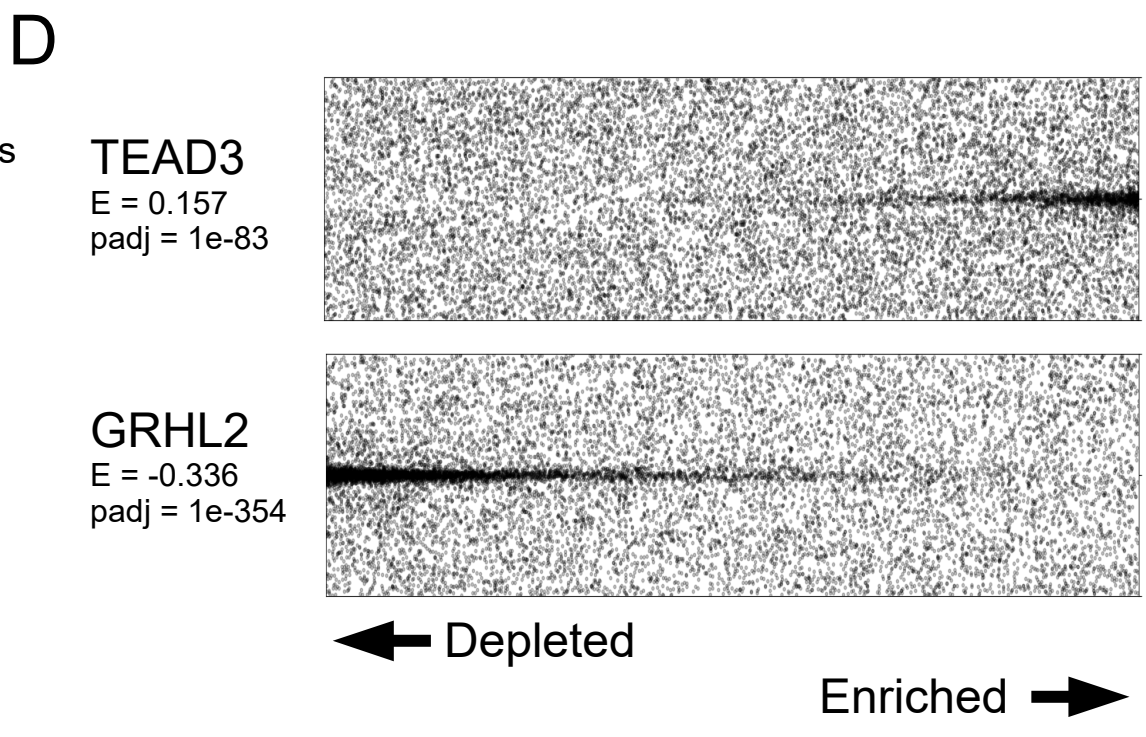

A
Downregulated = 531
Upregulated = 946
-log10 padj
20
10
0
-4
0
4
log2 Fold Change
padj < 0.05
log2 FC > 1
Both
Neither
B
GAS1
CXCL8
1000 bp
1000 bp
Ctrl
APM
0.25
-0.46
0.25
-0.46
0.38
-0.04
0.38
-0.04
C
Downregulated
Cytochrome P450
Aldo/keto reductase
Oxidoreductase
Histidine metabolism
Retinol dehydrogenase activity
Glycolysis/Gluconeogenesis
Steroid hormone biosynthesis
Lipid degradation
Galactose detabolism
EF-hand 2 domain
0 1 2 3 4 5
Enrichment Score
Upregulated
Serine protease
SH2 domain
Cell-cell adhesion
IL-1 receptor family
LIM zinc-binding
Serum Amyloid A protein
Keratin-associated protein
EF-hand-like domain
Insulin family
Positive regulation of cell division
0 1 2 3 4
Enrichment Score

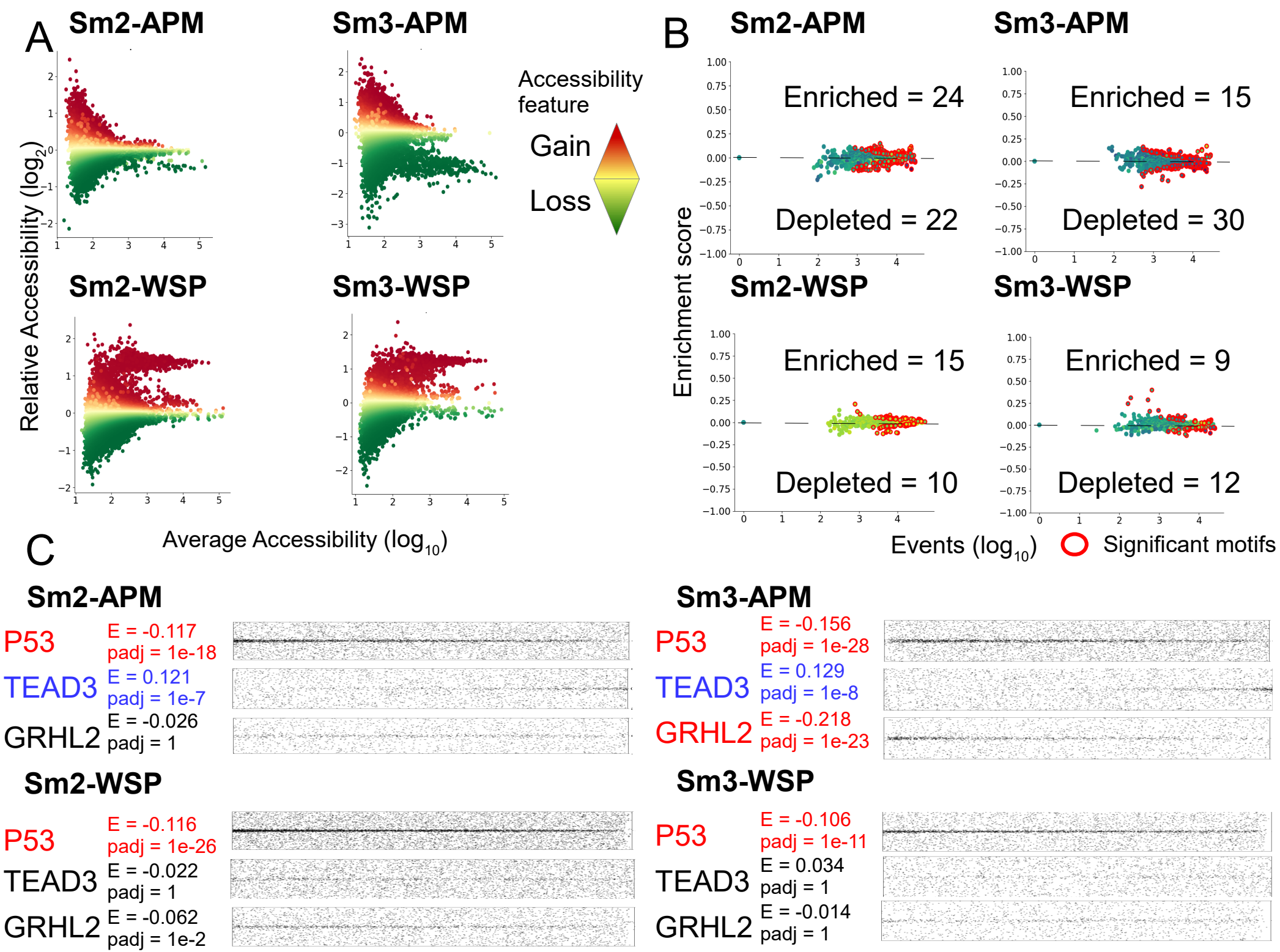

A
p_{adj} < 0.05
Both
log_2 FC > 1
Neither
Downregulated = 310
Upregulated = 207
-log10 padj
20
10
0
-4
0
4
log2 Fold Change
B
AKR1C1
2000 bp
10 kb
Ctrl
APM
0.38
-0.45
0.38
-0.45
MICAL2
1000 bp
50 kb
Ctrl
APM
0.36
-0.19
0.36
-0.19
C
Depleted
Enriched
TP53
TP63
NKX2
XPA
ZIK1
MAZ
KLF1
TOPORS
-0.4
-0.3
-0.2
-0.1
0.0
Enrichment Score
EGR2
EGR3
DNMT1
ALX1
MEF2B
EGR1
KLF14
SP4
MBD2
MEF2D
SP1
KLF1
KLF10
SP3
HKR1
FOXL1
KLF15
ZIM2
PRDM6
0.00
0.05
0.10
0.15
0.20
Enrichment Score
D
EGR1
E = 0.119
padj = 1e-8
TP53
E = -0.318
padj = 1e-46
Depleted
Enriched

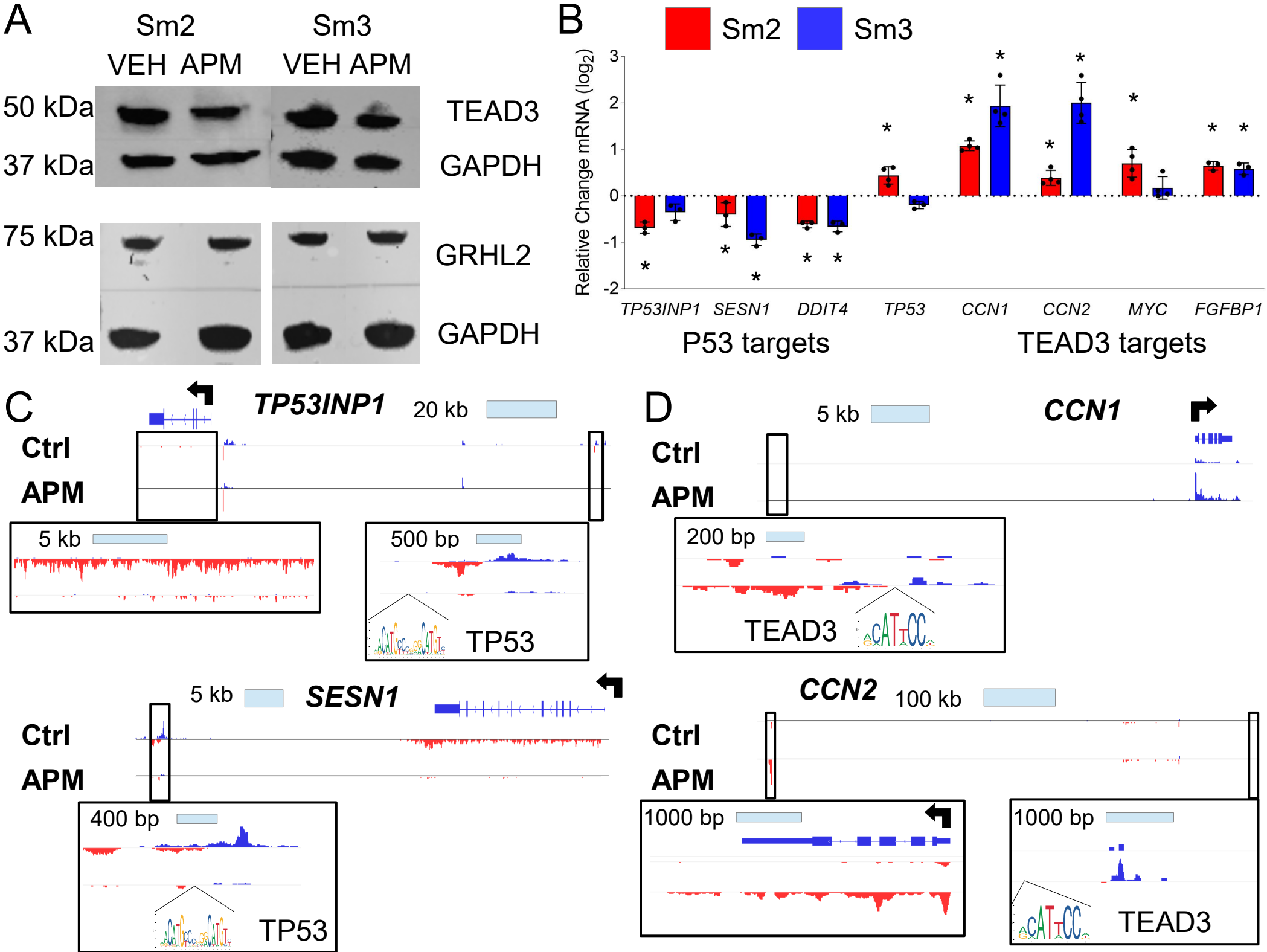

A
Gain
Loss
Accessibility feature locus
Relative Accessibility (log_2)
Average Accessibility (log_{10})
6
4
2
0
−2
−4
−6
1
2
3
4
5


C
Depleted
Enriched
GRHL2
GRHL1
CEBPE
CEBPD
CEBPB
CEBPG
NRF1
CEBPA
FOXN2
DBP
TEAD3
NFE2
TEAD1
TEAD4
TEAD2
BACH1
TP53
TP63
FOXA1
GATA1
-0.4
-0.3
-0.2
-0.1
0.0
0.00
0.05
0.10
0.15
0.20
Enrichment Score
Enrichment Score


B
Significant motifs
Enriched
Depleted
Enrichment score
Events (Log_{10})
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
0
1
2
3
4
5


D
TEAD3
E = 0.157
padj = 1e-83
GRHL2
E = -0.336
padj = 1e-354
Depleted
Enriched


A
Sm2-APM
Sm3-APM
Sm2-WSP
Sm3-WSP
Relative Accessibility (log2)
Average Accessibility (log10)
Accessibility feature
Gain
Loss
B
Sm2-APM
Enriched = 24
Depleted = 22
Sm3-APM
Enriched = 15
Depleted = 30
Sm2-WSP
Enriched = 15
Depleted = 10
Sm3-WSP
Enriched = 9
Depleted = 12
Enrichment score
Events (log10)
Significant motifs
C
Sm2-APM
P53 E = -0.117 padj = 1e-18
TEAD3 E = 0.121 padj = 1e-7
GRHL2 E = -0.026 padj = 1
Sm3-APM
P53 E = -0.156 padj = 1e-28
TEAD3 E = 0.129 padj = 1e-8
GRHL2 E = -0.218 padj = 1e-23
Sm2-WSP
P53 E = -0.116 padj = 1e-26
TEAD3 E = -0.022 padj = 1
GRHL2 E = -0.062 padj = 1e-2
Sm3-WSP
P53 E = -0.106 padj = 1e-11
TEAD3 E = 0.034 padj = 1
GRHL2 E = -0.014 padj = 1


A
Sm2
Sm3
VEH APM
VEH APM
50 kDa
37 kDa
TEAD3
GAPDH
75 kDa
37 kDa
GRHL2
GAPDH
B
Sm2
Sm3
Relative Change mRNA (log2)
3
2
1
0
-1
-2
TP53INP1
SESN1
DDIT4
TP53
CCN1
CCN2
MYC
FGFBP1
P53 targets
TEAD3 targets
C
TP53INP1
20 kb
Ctrl
APM
5 kb
500 bp
TP53
SESN1
5 kb
Ctrl
APM
400 bp
TP53
D
5 kb
CCN1
Ctrl
APM
200 bp
TEAD3
CCN2
100 kb
Ctrl
APM
1000 bp
1000 bp
TEAD3