TITLE:

The cistrome response to hypoxia in human umbilical vein endothelial cells

AUTHORS

Ayush Singh (1), Viktor Pastukh (2), Justin Roberts (2), Zachary M. Turpin (1), Zehta S. Glover (1), Grant T. Daly (2), Jane M. Benoit (1), Mark N. Gillespie (2), and Hank W. Bass (1)

Author Affiliations:

(1) Department of Biological Science, Florida State University, Tallahassee, USA

(2) Department of Pharmacology, University of South Alabama College of Medicine, Mobile, Alabama, USA

Author Emails and ORCIDs: [HWB: *please add or update you email & ORCID below*]
AS, ayushsingh5906@gmail.com, 0000-0003-0184-4486
VP, vpastukh@southalabama.edu,
JR, jtroberts@southalabama.edu,
ZMT, zachturpin92@gmail.com,
ZSG, zehtashea@gmail.com,
GTD, daly@southalabama.edu, 0000-0002-4109-0546,
JB, jbenoit@bio.fsu.edu, 0000-0001-9166-6219
MNG, mgillesp@southalabama.edu,
HWB, bass@bio.fsu.edu, 0000-0003-0522-0881

Corresponding Author: HWB

ABSTRACT:

Transcriptional signaling during hypoxic stress is primarily governed by hypoxia-inducible transcription factors (HIFs), which bind to hypoxia response elements (HREs) in gene regulatory regions. However, only a small proportion (~1%) of known HREs are occupied by HIFs during hypoxia, suggesting the involvement of additional pathways. To address this gap, we utilized MOA-seq, an MNase-based assay that enables genome-wide, high-resolution (<30 bp) identification of transcription factor (TF) occupancy footprints within open chromatin. We applied this cistrome profiling approach to nuclei from cultured endothelial cells exposed to normoxia or hypoxia for 1, 3, or 24 hours, revealing thousands of hypoxia-modulated genomic sites with dynamic TF footprints. The affected genes were enriched in canonical hypoxia-induced pathways, such as angiogenesis. Motif analysis identified over 100 candidate TFs potentially mediating this complex multifaceted genomic response. We followed up these observations with a quantitative footprint clustering analysis to gain mechanistic insight into how concurrent TF regulatory networks might choreograph these responses. We found 10 distinct MOA-defined kinetic clusters, half of which contained HIF1A motifs. HIF1A-proximal binding sites revealed candidate co-activators, whereas motifs in non-HIF1A clusters implicated HIF1A-independent pathways. These findings highlight the utility of cistrome analysis in unraveling the genomic response to hypoxia.

KEY POINTS:

1. MOA-seq mapped TF occupancy at 21,765 sites in normoxia, including 7,444 beyond the known ENCODE cCREs.
2. Hypoxia for 1, 3, and 24h changes the cistrome occupancy at thousands of genes.
3. Clustering analysis distinguished candidate HIF1A-associated cofactors and HIF1A-independent TFs.

Introduction

The regulation of gene expression through transcriptional control is a fundamental process that drives phenotypic diversity. Transcription factors (TFs) are crucial in this process, binding to specific DNA sequences either directly or as components of larger protein complexes, at promoters or enhancer regions. Mapping these binding sites across the genome is essential for understanding active transcriptional programs and identifying cis-regulatory elements and motifs. Traditional methods such as chromatin immunoprecipitation followed by sequencing (ChIP-seq) or assay for transposase-accessible chromatin with sequencing (ATAC-seq) have been widely used for this purpose. However, ChIP-seq is constrained by the requirement for high-quality antibodies and its focus on individual TFs, limiting its scalability for comprehensive cistrome analysis. ATAC-seq does not require antibodies, but dermacates relatively large nucleosome-free accessible chromatin regions with peak sizes averaging 200-300 bp (1).

In contrast to these traditional methods, the recently developed MOA-seq (MNase-defined cistrome-Occupancy Analysis) offers a robust alternative, capable of globally identifying protein-occupied loci within accessible chromatin regions with high resolution peaks of ~30 bp (2). The term cistrome is used here to refer to the set of cis-acting targets occupied by trans-acting factors on a genome-wide scale, a genome-wide location of occupied transcription factor binding-sites. MOA-seq was initially developed and validated in maize (*Zea mays* L.), a model genetic organism with a diploid genome size comparable to humans, combining the strengths of chromatin accessibility profiling with the precision of MNase digestion. High resolution MNAse footprinting in plants has proven useful for dissecting the complex regulatory landscapes in various tissues and environmental contexts (3–6).

Here we used MOA-seq to investigate changes in chromatin structure induced in human umbilical vein endothelial cells (HUVECs) by hypoxia, a condition characterized by reduced oxygen availability, which is critical in various physiological and pathological processes. HUVECs were employed as a model system because of the availability of

83 previous reports on larger scale chromatin remodeling induced by hypoxia in this cell
84 type (7).

85 The application of MOA-seq to this experimental system offers a high-resolution view of
86 transcriptional regulation under hypoxic conditions, enabling precise identification of TF-
87 binding sites and cis-regulatory elements. Time-resolved profiling at 1, 3, and 24 hours
88 of hypoxic exposure identified over 5,000 responsive TF occupancy sites across the
89 genome. Unsupervised clustering of these dynamic regions revealed distinct regulatory
90 trajectories, encompassing both canonical HIF1A-associated elements and alternative,
91 HIF1A-independent programs. Together, this work demonstrates the power of MOA-seq
92 to resolve chromatin-level responses to hypoxic stress and uncovers regulatory
93 complexity that may be overlooked by conventional chromatin profiling methods.

94 Materials and Methods

95 MOA-seq library preparation and sequencing

96 HUVECs were cultured to 90–95% confluency and subjected to hypoxia (2% O_2) for 1,
97 3, or 24 hours. Following hypoxia treatment, cells were formaldehyde-fixed and stored
98 at -80°C until processing. To isolate nuclei, fixed cells from a 10 cm dish were thawed
99 and resuspended in 500 µL Nucleus Isolation Buffer (NIB: 0.3 M sucrose, 2 mM
100 MgOAc, 2.1% Triton X-100, 10 mM HEPES pH 7.4, 1 mM $CaCl_2$) and incubated at room
101 temperature for 5 minutes. Cells were pelleted (5 minutes, 1000g), washed in fresh NIB,
102 and resuspended in 500 µL NIB. A small aliquot (10 µL) was retained for microscopy
103 checks of nuclear integrity. The remaining nuclei suspension was aliquoted into eight
104 1.5 mL tubes (60 µL per tube).

105 Micrococcal nuclease (MNase) digestion was performed using a serial dilution series of
106 MNase (starting at 1280 U/mL and decreasing two-fold to 20 U/mL). Each enzyme
107 dilution (6.7 µL) was added to corresponding nuclei aliquots, and digestion proceeded at
108 37°C for 15 minutes. Reactions were stopped by adding 5 µL of 0.5 M EGTA-NaOH (pH

109 8.0). Decrosslinking was performed overnight at 65°C with 360 µL ddH_2O, 50 µL 10%
110 SDS, 50 µL 1.5 M NaCl, and 5 µL 20 mg/mL Proteinase K. DNA was extracted through
111 sequential phase separation with Tris-saturated phenol, followed by phenol:chloroform
112 (1:1), and chloroform alone. DNA was precipitated by adding 0.05 volumes of 3M
113 sodium acetate (pH 5.2), 2 µL 25 mg/mL linear polyacrylamide, and 2 volumes of -20°C
114 ethanol. The mixture was incubated at -20°C for at least 30 minutes and then pelleted
115 by centrifugation (20,000g, 15 minutes, 4°C). The pellet was washed in 70% ethanol,
116 air-dried, and resuspended in 100 µL TE buffer containing 40 µg/mL RNase A. RNA
117 was removed by incubation at room temperature for 1 hour, and DNA was re-
118 precipitated as previously described. Final DNA pellets were resuspended in 30 µL TE
119 buffer, quantified by spectrophotometry, and analyzed on a 1% agarose gel. DNA
120 quality was further assessed using an Agilent 2100 Bioanalyzer, Q-PCR, and a Qubit
121 2.0 Fluorometer.

122 Libraries were prepared using NEBNext® Ultra™ II DNA Library Prep Kits (Cat.
123 E7103S). The indexed libraries were size-selected to enrich for sub-nucleosomal (<150
124 bp) genomic insert sequences and submitted for paired-end 150 bp reads on the
125 Illumina HiSeq (FSU College of Medicine, Translational Science Lab). Two biological
126 replicates were analyzed per condition.

127 For control MOA libraries, DNA was isolated from A549 human cells (ATCC, CCL-185),
128 cultured in F12K media supplemented with fetal bovine serum, 1% L-glutamine, and 1%
129 penicillin-streptomycin. DNA extraction was performed using TRIzol reagent (cat no.
130 15596026) following the manufacturer's protocol. DNA quality was assessed via
131 agarose gel electrophoresis, and concentrations were determined using a NanoDrop
132 UV spectrophotometer. DNA yields were sufficient for multiple MNase digestions.
133 MNase digestion conditions were optimized based on MOA-seq chromatin digestion,
134 scaled to 1/64th concentration. Digestion was performed to achieve a partial digest
135 pattern similar to MOA-seq. Digest volumes were scaled to 60 µL to yield ~1 µg of DNA
136 per digest for Illumina library preparation. Digestion patterns were assessed via gel
137 electrophoresis, and selected samples were pooled to increase DNA input to 1 µg per
138 sample. Library preparation for control MOA libraries followed the NEBNext Ultra II DNA

139 Library Prep Kit for Illumina (NEB #E7103). After adapters were added, samples were
140 flash-frozen, and library-specific index sequences were incorporated using the dual
141 indices method by the NGS Library Facility at FSU. DNA libraries were sequenced with
142 a target read depth of 80 million reads per sample and stored at -80°C until sequencing.

143 MOA-seq data processing and fragment center coverage

144 The sequence data reads (in FASTQ format) were trimmed for adapters using Cutadapt
145 (8) with default parameters. Because the sequence read chemistry generates
146 overlapping reads, the trimmed paired-end reads for each biological replicate were
147 subsequently merged using FLASh (9) with the parameters "-t 40 --cap-mismatch-quals
148 -M 100 -d -q". The resulting merged read files represented 91-97% of each library, and
149 they were aligned to the human reference genome (hg38) using the bwa-mem2 mem
150 function (10) with the parameters "-M -t 40."

151 Alignment files (SAM) were converted to BAM format using SAMtools (11), with
152 additional filtering applied to retain only reads with a mapping quality score (MAPQ) of
153 20 or greater. PCR duplicates were also removed. The average fragment length was
154 calculated using the stat function in SAMtools (11), while the effective genome size was
155 determined using unique-kmers.py (12) with the average fragment length and the hg38
156 genome FASTA as inputs. BAM files were converted to BED files using the bamToBed
157 function in Bedtools (13). The alignment and pre-processing metrics are provided in
158 Supplementary Table S1.

159 To achieve greater resolution, fragment center tracks (also referred to as 'frenters' as
160 described by Savadel et al., 2021) were created. In this process, each mapped read
161 was shortened to its central 20 bp using a custom awk command. For reads with an
162 uneven number of bases, the central base was taken and extended by 10 bp. For reads
163 with an even number of base pairs, the central base was randomly selected from one of
164 the two central bases and then extended by 10 bp. Bedgraphs were generated using
165 the genomeCoverageBed function in Bedtools (13) with a counts-per-million (CPM)
166 scaling factor to normalize for sequencing depth. Finally, BigWig files for UCSC

167 Genome Browser visualization were produced using the bedGraphToBigWig utility (14),
168 a UCSC binary tool.

169 Within sample and differential peak calling

170 We determined that MACS3 (15) was the most suitable tool for MOA peak calling due to
171 its flexibility across various scenarios and input parameters (16). Fragment center BED
172 files for each biological replicate pair per treatment were used as input to the callpeak
173 function of MACS3 with the following parameters "-p .001 --nomodel --extsize 20 --min-
174 length 20 --max-gap 20 --keep-dup all." The resulting peak sets were analyzed using
175 the Irreproducibility Discovery Rate (IDR) framework (17) to assess data consistency,
176 identifying reproducible peaks between biological replicates within each sample,
177 referred to as within-sample peaks. In accordance with ENCODE ChIP-seq guidelines
178 (18), we further evaluated data quality by assessing pooled pseudo-replicate
179 consistency and self-consistency for each individual replicate (Supplementary Table
180 S2). To minimize artifacts, within-sample peaks were filtered using the intersect function
181 in Bedtools to exclude blacklist regions (19) and MOA control regions, defined by
182 chromatin-free total DNA light digests with MNase.

183 To identify hypoxia-induced binding sites (gained, lost, or shared), we used the bdgdiff
184 function in MACS3, comparing each hypoxia time point (1h, 3h, 24h) to the normoxic
185 sample. Combined replicate BED files for each sample were used as input for bdgdiff
186 analysis, with the minimum peak length and maximum peak gap parameters set to 20.
187 The optimal likelihood ratio (LHR) cutoff was determined using the elbow method.
188 Specifically, we plotted the number of peaks against the total peak length as a function
189 of LHR and identified the optimal LHR as the inflection point, occurring just before a
190 sharp increase in slope (Supplementary Table S2). This threshold was chosen to
191 maximize sensitivity while minimizing noise in peak detection.

192 ## MOA-seq enrichment analyses

193 We gathered various ENCODE-published datasets for MOA-seq enrichment analysis.
194 Bulk RNA-seq gene quantifications (20) (ENCFF892OHT) were obtained for endothelial
195 cells and used to demonstrate the correlation between MOA normoxic coverage and
196 gene expression at the transcription start site (TSS) and the transcription end site
197 (TES). Candidate cis-regulatory elements (cCREs) and DNase data from multiple cell
198 types (21, 22) were collected for their genomic coordinates and served as key features
199 for plotting local average normoxic MOA coverage and analyzing normoxic peak base
200 pair intersections.

201 To evaluate the base pair intersections of normoxic peaks with ENCODE cCREs and
202 DNase data, a probability distribution was generated by intersecting shuffled normoxic
203 peak coordinates with the published dataset genomic coordinates using Bedtools 100
204 times. This approach determined the expected number of base pairs that would
205 intersect by chance. The expected base pair observations were plotted to create a
206 probability distribution, which was then used to calculate a z-score (Supplementary
207 Table S3). The z-score quantified how many standard deviations the observed number
208 of intersected base pairs deviated from the expected number.

209 ChIP-seq peaks of various transcription factors for the HUVEC cell type were obtained
210 from the ReMap 2022 ChIP-seq database (23) (Experiments and peak counts listed in
211 Supplementary Table S4). Their genomic coordinates were used as central features for
212 plotting the local average normoxic MOA coverage. Coverage heatmaps were
213 generated using Deeptools plotHeatmap (24), focusing only on ChIP-seq peaks with
214 non-zero MOA coverage within the 400-bp window. These peaks represented
215 approximately 33% of the HUVEC cell type ReMap dataset, which is unsurprising given
216 that many ReMap peaks were derived from experiments conducted under different
217 conditions.

218 Hypoxia-induced binding sites or diff-MOA were visualized using heatmaps generated
219 with Deeptools plotHeatmap, overlaying hypoxia and normoxia coverage. Gain sites

220 were sorted by hypoxia coverage values, while loss sites were sorted by normoxia
221 coverage values. The diff-MOA from all time points were intersected with the gene
222 bodies of protein-coding genes from GENCODE version 45 (25) using Bedtools
223 intersect, with gene start and end coordinates extended by 200 base pairs to capture
224 potential regulatory elements downstream of the TSS or TES. These genes were then
225 grouped based on whether the intersecting diff-MOA was located within a cCRE,
226 outside a cCRE, or if the gene contained both (Supplementary Table S5). Any identified
227 gene sets were submitted to ENRICHR (26, 27) for transcription, pathway, and gene
228 ontology enrichment analysis. To assess the statistical significance of the transcriptional
229 pathways for the diff-MOA gene lists, we generated randomized size-matched gene lists
230 for ENRICHR input to simulate a random genomic background for comparison
231 (Supplementary Table S5).

232 Gain and loss diff-MOA peaks by hour were submitted separately and in combination to
233 ENRICHR for pathway analysis (Supplementary Table S6). We then examined the
234 relationship between diff-MOA and differentially expressed genes (DEGs) by identifying
235 DEGs that also had a gain or loss diff-MOA peak. These subsets of genes were
236 submitted to ENRICHR for pathway analysis using the MSigDB Hallmark and
237 WikiPathways Human libraries. To assess whether the diff-MOA-identified subset of
238 DEGs exhibited stronger statistical enrichment, we created three size-matched gene
239 subsets, randomly sampled from the list of DEGs for each time point. Each subset was
240 submitted to ENRICHR, and the average statistical significance from the three inputs
241 was calculated to determine whether a random, size-matched subset of DEGs
242 performed comparably (Supplementary Table S7).

243 Hierarchical clustering analysis

244 All diff-MOA peaks, including gains and losses at 1h, 3h, or 24h of hypoxia, were
245 merged and windowed at 30 bp centered on the merged peaks. Normalized read-depth
246 coverage counts for each time point were calculated for each window using the
247 coverage function in Bedtools (13). These normalized counts were standardized using

248 z-score normalization for each window or peak across the three hypoxia time points and
249 used as input for clustering analysis. A hierarchical dendrogram was generated using
250 the hclust function in R. The optimal number of clusters (k = 10) was determined using
251 the elbow method, where the within-cluster sum of squares (WCSS) stabilized and
252 plateaued (Supplementary Table S8). The resulting 10 clusters (hypoxia MOA clusters
253 or HMCs) were visualized using ggplot2 (28), with each line in the cluster graphs
254 representing a single differential peak and its MOA coverage across the four sample
255 time points. The unstandardized MOA coverage values for each hour and the genomic
256 coordinates for each merged diff-MOA peak, organized by cluster, are provided in
257 Supplementary Table S9.

258 Motif discovery and analysis

259 Motif discovery, enrichment, and subsequent analyses were performed using
260 XSTREME from the MEME suite (29). Default settings were applied to the diff-MOA
261 peaks, with the Markov order set to M = 0 (Supplementary Zip File 1). Genomic
262 sequences from the diff-MOA peaks were extracted using the getfasta function in
263 Bedtools and used as input for XSTREME. Motif enrichment analysis was conducted
264 using the HOCOMOCOv11 transcription factor motif database (30). Additionally, the
265 genomic sequences from each of our HMC clusters were used as input for XSTREME
266 to identify cluster-specific TF motifs (Supplementary Table S10, Supplementary Zip File
267 2).
268
269 The motif enrichment tool SEA from the MEME suite was used to identify HIF1A motifs
270 in the hypoxia-induced footprint gain sites for 1h, 3h, and 24h (Supplementary Zip File
271 3). For each time point, we plotted the local average hypoxic and normoxic MOA
272 coverage, centered on the genomic coordinates of the HIF1A motifs corresponding to
273 that time point. We then examined the distribution of these HIF1A motifs across the
274 previously defined HMC clusters using Bedtools intersect, which allowed us to group the
275 HMC clusters based on the prevalence of HIF1A motif occurrence, resulting in our
276 HIF1A (hmc2, 3, 4, 5, 9) and NON-HIF1A (hmc1, 6, 7, 8, 10) cluster groups.

277 Additionally, we visualized the proportion of HIF1A motifs within each HMC cluster
278 attributable to the 1h, 3h, and 24h time points (Supplementary Table S11). To identify
279 TF motifs enriched in the HIF1A and NON-HIF1A cluster groups, we applied a threshold
280 requiring each TF motif identified from the previous XSTREME motif analysis of our
281 HMC clusters to be present in at least two clusters for the NON-HIF1A group (any two
282 of hmc1, 6, 7, 8, 10) and in at least three clusters for the HIF1A group (any three of
283 hmc2, 3, 4, 5, 9) (Supplementary Table S12).

284 To identify motifs co-occurring near the HIF1A motifs, we extracted the central base pair
285 genomic coordinates of the peaks containing HIF1A motifs in our HIF1A cluster group
286 (hmc2, 3, 4, 5, 9) and extended the window by 50 bp on either side, resulting in a 101-
287 bp window. These genomic coordinates were then analyzed using the motif enrichment
288 tool SEA from the MEME suite (Supplementary Zip File 4). From the TF motifs enriched
289 in our HIF1A and NON-HIF1A cluster groups identified in our previous analysis, we
290 marked with an asterisk those that were also found to co-occur with HIF1A motifs in the
291 immediate region (Supplementary Table S12).

292 Gene Set Enrichment Analysis

293 Differential gene expression analysis was performed using DESeq2 (31) to process raw
294 read counts and identify differentially expressed genes (DEGs). Genes with insufficient
295 read counts across replicates were filtered out. DESeq2 was used to normalize the data
296 and calculate log2 fold changes (Log2FC) for each condition (1h, 3h, 24h) relative to the
297 baseline (0h). Significant DEGs were identified for each time point using an adjusted p-
298 value threshold of ≤ 0.05. These DEGs were then combined into a single list, retaining
299 the corresponding Log2FC values for each time point.

300 From this combined list, only DEGs associated with a defined cluster were included in
301 subsequent analyses, as not all DEGs corresponded to a cluster. The selected DEGs
302 were organized into their respective clusters, and gene set enrichment analysis (GSEA)
303 (32, 33) was performed for each time point using the DEGs within a given cluster and
304 their corresponding Log2FC values. The gseGO function from the clusterProfiler R
305 package was used to identify enriched biological processes (BP ontology). Enrichment
306 parameters included a minimum gene set size of 3, a maximum gene set size of 800,
307 and a p-value cutoff of 0.05. For each cluster and time point, normalized enrichment
308 scores (NES), p-values, and gene ratios were calculated, where gene ratios
309 represented the proportion of genes contributing to enrichment within each gene set.

310 The outputs of the enrichment analysis, including cluster-specific enrichment results for
311 each time point, were exported as CSV files for further interpretation (Supplementary
312 Zip File 5, 6, 7). This approach enabled the identification of hypoxia-induced biological
313 processes and their temporal dynamics across gene expression by cluster.

314 Results

315 Transcription Factor Occupancy Profiling with MNase "MOA-seq" 316 in normoxic and hypoxic HUVECs

317 To investigate hypoxia-induced changes in transcription factor occupancy, we applied
318 MOA-seq to HUVECs under normoxic and hypoxic conditions. MOA-seq enables
319 genome-wide identification of protein-bound cis-regulatory elements through
320 micrococcal nuclease (MNase) digestion of fixed chromatin, followed by high-resolution
321 footprint detection (2). For this study, we adapted the assay to human cells using
322 chromatin fixation and MNase digestion protocols previously optimized for mammalian
323 systems (34). This allowed us to generate reproducible, high-resolution cistrome maps
324 across three time points of hypoxic exposure (1, 3, and 24 hours), enabling quantitative
325 analysis of TF occupancy dynamics during the hypoxic response.

326 The MOA-seq experimental scheme is summarized in Figure 1, showing the samples
327 (Fig. 1A) used in this study. These four samples are designated as 0h, normoxia, as the
328 pre-treatment control, and 1h, 3h, or 24h to indicate the time of exposure to hypoxia, 2%
329 oxygen in this case. Cells were fixed with formaldehyde to capture the endogenous
330 chromatin structure at the time of fixation. The DNA-protein interactions are revealed as
331 nuclease digestion patterns (Fig. 1B) based on measures of fragment abundances
332 resulting from the light/partial chromatin digests. By making libraries from light digests
333 (Supplementary Table S1), we enriched for recovery of open, accessible, and nuclease
334 hypersensitive chromatin, similar to other methods that use nucleases to probe
335 chromatin structure, such as ATAC-seq, and DNAse I hypersensitivity profiling, ideal for
336 nucleosome occupancy and vacancy mapping (35–37). However, unlike these gold-
337 standard assays, MOA-seq was specifically developed to map small particles in open
338 chromatin to generate high-resolution coverage peak maps based on fragment center
339 (frenter) profiling (bottom of Fig. 1B). In maize, MOA-seq footprints were shown to not
340 only overlap with the majority of ATAC peaks, but they also appeared in regions not
341 previously identified as open or regulatory chromatin, expanding the cistrome
342 occupancy catalog. Specifically, this approach uncovered hundreds of motif families
343 collectively found in < 0.1% of the genome, unprecedented in scope and resolution (2).
344 For these reasons, we used MOA profiling to map the hypoxia-responsive changes in
345 chromatin structure in human cells.

346 MOA-seq coverage and footprints align with ENCODE regulatory 347 sites and other genomic regions

348 Given that MOA-seq primarily maps sites of TF occupancy (2), we compared the human
349 MOA-seq data to well-curated regulatory regions. We found that, as in maize, MOA-seq
350 coverage colocalizes with regulatory regions, as summarized in Figure 2. MOA
351 coverage profiles exhibited characteristic peaks of open chromatin at the transcription
352 start (TSS) and end (TES) sites of genes. From the average coverage plots centered on
353 the TSS, we observed local peaks just upstream of the TSS and a local depletion within
354 the gene bodies (Fig. 2A). At the TES, we observed a local double peak flanking

355 transcript model end, which is near the polyA addition signal (Fig. 2B). At both ends of
356 the genes, the average MOA coverage scales with transcript abundance quartiles (Fig.
357 2A,B).

358 We next examined MOA-seq coverage and peaks in relation to published regulatory
359 elements. For this and all subsequent analyses, we plotted and called peaks on the
360 MOA-seq read coverage using only the central 21 bp of each paired-end read-aligned
361 fragment. These 21 bp fragment centers, called frenters", increase the resolution of
362 MNAse-based footprints (2). We next next examined the Encyclopedia of DNA
363 Elements (ENCODE) catalog of candidate cis-Regulatory Elements (cCREs), which are
364 comprised of nearly one million elements classified in humans according to chromatin
365 signatures such as DNaseI hypersensitivity and histone modifications (38). The five
366 subclasses of cCREs are shown using the UCSC/ENCODE labels and color-coded as
367 promoter-like signatures (PLS/prom, red), proximal enhancer-like signatures
368 (pELS/enhP, orange), distal enhancer-like signatures (dELS/enhD, yellow), DNase-
369 H3K4me3 (DNase-H3K4me3/K4m3, pink), and CTCF-only (CTCF-only/CTCF, blue).
370 We observed striking and conspicuous co-occurrence of MOA-seq read alignment
371 averages at all five cCRE subtypes (Fig. 2C).

372 We called 21,765 high-confidence peaks in normoxic HUVECs (0h sample)
373 (Supplementary Table S2) using the IDR framework (17) and found that their bp-wise
374 intersections were 10s to 100s of standard deviations away from their corresponding
375 randomized distributions, with CTCF and distal enhancer-like signatures (dELS) being
376 the highest with Z-scores exceeding 200 (Fig. 2D; Supplementary Table S3**)**. Z-score
377 analysis with DNaseI clusters (DHS sites) showed the highest correspondence with
378 HUVECs compared to five other cell types, indicating that MOA-seq captures open
379 chromatin with cell-type specificity (Fig. 2E; Supplementary Table S3).

380 Interestingly, we noticed that a large proportion of MOA peaks were not predicted by
381 cCRE or DNaseI clusters. We quantified MOA peak abundance relative to DHS vs.
382 non-DHS sites and cCRE vs non-cCRE sites (Fig 2F), with genomic regions illustrating
383 MOA peaks at known or unclassified regulatory regions (Fig. 2G). Remarkably, 40% of

384 the ~21.8K MOA peaks were not under DHS clusters whereas 34% of the MOA peaks
385 were not under cCRE regions.

386 Enrichment of normoxic MOA at ReMap ChIP summits for 13 transcription factors from HUVEC-specific datasets

388 Given the MOA-seq assay maps of cistrome occupancy, we expected that read
389 coverage would show peaks concordant with the location of known and mapped
390 transcription factor (TF) binding. To test for this, we limited the analysis to HUVEC
391 datasets for ReMAP ChIP summits (23) for TFs mapped in HUVECs. The human
392 ReMap consolidates regulatory regions from integrated analyses of multiple DNA-
393 binding sequencing experiments and can be filtered by organism, cell type, and species.
394 Here we combined ReMap data from all the experiments from HUVECs for 13 TFs that
395 were highly expressed in endothelial cells (39) and with ChIP-seq data available
396 summarized in Figure 3. Of these, MEF2C, ERG, FLI1, PROX1, and ETS1 were also
397 noted as characteristic TFs for HUVECs (39).

398 The average normoxia MOA coverage peaked at ChIP-seq summits (Fig. 3A, 3B) , as
399 expected based on the strong cCRE signals. For each individual TF analyzed, we
400 observed a local enrichment of MOA coverage (Fig. 3C). The RAD21 and MEF2C
401 showed the highest local MOA coverage enrichments. Local MOA coverages for each
402 TF (Fig. 3D) are plotted and vary in plot smoothness primarily according to the number
403 of peaks in the datasets, with RAD21 having over 60,000 peaks, whereas EPAS1 has
404 only 109 peaks (Supplementary Table S4)**.**

405 Hypoxia causes temporally-specific changes in MOA-defined chromatin structure.

407 Having demonstrated that MOA-seq is highly informative for mapping TF occupancy
408 within human cell cultures, we used it to define genomic regions with differential MOA
409 coverage following hypoxia. These regions, termed “diff-MOA” sites as summarized in

410 Figure 4, should reflect early chromatin responses to the hypoxic stress signaling. We
411 generated a set of diff-MOA peaks with the bdgdiff function of MACS3 using the elbow
412 method to optimize for signal-to-noise as described in the methods (Supplementary
413 Table S2). For each hypoxia time sample, we identified three categories of peaks
414 relative to the 0h normoxia: gained, lost, or shared. We observed a total diff-MOA (gain
415 + loss) peak count of 1,978 after 1 hour of hypoxia, 1,932 after 3 hours of hypoxia, and
416 2,586 after 24 hours of hypoxia (Fig. 4A). Specific examples of these regions, which
417 reflect hypoxia-induced changes in chromatin structure, are shown as browser
418 screenshots (Fig. 4B) or global trends as heatmaps sorted on hypoxia MOA coverage
419 for "Gain" peaks or sorted on normoxia coverage for "Loss" peaks (Fig. 4C).

420 For all hypoxia-induced diff-MOA peaks at any time point, we found that a surprisingly
421 large number occurred in regions not classified as cCREs. These non-cCRE regions
422 could represent biologically non-informative sites peculiar to MNase, or conversely, they
423 could represent biologically informative sites that are not well captured by current cCRE
424 annotations. To investigate this, we performed a TF enrichment analysis using gene
425 lists from all time points and their associated diff-MOA peaks, comparing gene lists from
426 diff-MOA plus cCRE regions with those from diff-MOA at non-cCRE sites (Fig. 4D;
427 Supplementary Table S5). Remarkably, both groups, cCRE-intersecting and non-cCRE,
428 showed comparable enrichment for ENCODE and CHEA consensus TFs from ChipX
429 (Fig. 4E) using the gene set enrichment analysis tools of ENRICHR (26, 27, 40). These
430 findings clearly established that the extra sites substantially expanded the functional
431 cistrome by nearly two fold beyond the gold-standard ENCODE annotations. We also
432 checked randomized gene lists of comparable sizes and none of these exceeded the
433 adjusted P-value of 0.05 (Supplementary Table S5). We therefore concluded that MOA-
434 seq enables precise and sensitive delineation of candidate genomic regulatory regions
435 in HUVECs exposed to hypoxia.

436 ## Comparison of differential MOA with differential gene expression

437 We next addressed the question of whether time-dependent changes in the hypoxic
438 transcriptome were biologically related to the MOA-defined cistrome dynamics. For this,
439 we compared DEGs with differentially bound MOA sites (diff-MOA). For the three time
440 points, the number of diff-MOA genes ranged from 1,092-1,401. The DEGs displayed
441 somewhat more time-dependent variation, with 853 at 1h, 1,830 at 3h, and 2,521 at
442 24h. A comparison of hypoxia-induced diff-MOA genes to the DEGs is shown in Figure
443 5. Remarkably, the gene sets enumerated by diff-MOA alone identified significant
444 enrichments from the ENRICHR libraries, Human MSigDB Hallmark and WikiPathway
445 Human gene sets (Fig. 5A), including the well-known hypoxia-responsive “VEGFA
446 VEGFR2 signaling" pathway with adjusted p-values of 0.034 at 3h, and 0.015 at 24h
447 (Supplementary Table S6). The ability to detect these pathway enrichments solely from
448 small particle mapping supports the idea that MOA-seq maps functional elements for
449 biological pathways, even in the absence of expression data.

450 We next explored the relationship between diff-MOA and DEGs by binary classification
451 of DEGs as up or down-regulated and diff-MOA peaks as gained or lost footprints. Only
452 a moderate proportion of genes were on both differential lists. For these doubly
453 differential genes, they showed a slight tendency to have more diff-MOA loss than diff-
454 MOA gain (Fig. 5B, Supplementary Table S7). We next wanted to ask if combining
455 these doubly differential genes had stronger enrichment than a number-matched
456 sample of DEGs alone using the same Human MSigDB Hallmark and WikiPathway
457 Human gene sets. Indeed the MOA-DEG combination gene sets displayed stronger
458 enrichment values than a comparably sized random subset of genes from the DEGs
459 alone (Fig. 5C, Supplementary Table S7). These findings establish that diff-MOA gene
460 sets alone identify functionally relevant pathways and that the MOA-DEG combination
461 shows higher enrichments than comparably sized DEG gene sets. This illustrates the
462 value of combining orthogonal datasets to study stress-induced responses.

Motif discovery and analysis at diff-MOA peaks identify candidate hypoxia-responsive gene regulatory network cis-elements

Because the diff-MOA peaks are small, averaging 24-32 bp overall (Fig. 4A), and likely primarily reflect TF occupancy (Fig. 3), the corresponding genomic sequences are ideal for motif discovery and analysis. We divided the diff-MOA peaks into two sets—gain and loss—and submitted their corresponding genomic sequences to the MEME motif analysis suite, with the top results shown in Figure 6 (Supplementary Zip File 1). This resulted in six lists of motifs, grouped as diff-MOA GAIN (Fig. 6A) or diff-MOA LOSS (Fig. 6B) for each of the three hypoxia exposure times. For each list of motifs, the top de novo motif is shown first, followed by motifs matching the abundant TF families (CTCF, FOS, JUN), and then motifs for the top 10 individual TFs, sorted alphabetically according to E-value. The percentage of peaks with a match is indicated, ranging from 2% for ZN502 and ZN582 to 67% for SMAD3. Notably, some TFs have matching or similar consensus sequences. Based on this analysis, and supported by the GO enrichment findings, these motifs help nominate regulatory TF candidates as factors that govern the myriad response pathways to hypoxia in these cell culture models. Given the vast nature of this global assay, which captured on a genome-wide scale over 100,000 peaks from within-sample data, and over 6,000 diff-MOA peak regions, it is not surprising that our pooled motifs list far exceeded 100 motif families from the diff-MOA alone.

Clustering MOA coverage kinetics delineates TF motifs and gene ontologies.

To better address the complexity and scale of what we found, we followed up with a quantitative footprint clustering analysis to gain mechanistic insight into how concurrent TF regulatory networks might modulate these myriad responses. For this, we used a MOA peak region coverage clustering approach that captures all the relationships for every genomic region mapped as gain or loss diff-MOA for any time point. Such

490 unsupervised machine learning clustering has proven valuable for extracting constituent
491 patterns in gene expression data. Instead of transcript levels, however, our approach
492 uses normalized MOA read coverage as the quantitative variable.

493 We defined 5,383 hypoxia-induced regions as 30-bp windows centered on diff-MOA
494 peaks. The k-means clustering with the elbow method (Supplementary Table S8)
495 resulted in 10 clusters, named hmc1 to hmc10 to indicate hypoxia MOA clusters 1-10,
496 as shown in Figure 7. Rather than reflecting a single regulatory pathway, these clusters
497 revealed multiple concurrent but mechanistically distinct transcriptional networks
498 governing the hypoxic response. Some clusters were enriched for early TF binding
499 events, potentially marking chromatin priming events, whereas others showed later-
500 stage occupancy patterns, corresponding to sustained transcriptional activity or
501 secondary responses to hypoxia.

502 The number of diff-MOA peak regions in the clusters ranged from 64 for hmc10 to 967
503 for hmc8, with most clusters having over 400 regions (Supplementary Table S9). Each
504 cluster was individually used as input for MEME motif analysis, and the top motifs for
505 each cluster are listed under their respective panels (Fig. 7A, Supplementary Zip File 2).
506 The clustering was remarkably successful at subdividing the motif families as indicated
507 by the unique lists of top TFs per cluster. For instance, the ETV4 motif was a top motif
508 in hmc1, hmc3, and hmc8, whereas CTCFL was primarily in hmc10, and hmc2. Nearly
509 all the motifs from the genome-wide analysis were retained but partitioned into one or
510 more clusters (Supplementary Table S10).

511 Given that the motifs were found in different clusters, we wondered if the cluster-specific
512 genes would reveal distinct biological processes. Of the entire set of 5,383 diff-MOA
513 regions, nearly 60% were assigned to genes. This assignment was based on overlap
514 with the gene model regions, with a 200 bp window on either side to capture proximal
515 promoters and downstream flanks. These cluster-specific gene sets, along with their
516 respective fold-change values from our DEGs, were used for GO analysis, which
517 specifically utilizes transcript inputs relative to time 0h. Gene Set Enrichment Analysis
518 (GSEA) is a computational method that determines whether a priori-defined sets of

519 genes show statistically significant, concordant differences between two biological
520 states—in our case, hypoxia. In this way, we could discern enrichments for genes
521 whose transcripts were activated versus suppressed by hypoxia (blue or red,
522 respectively, Fig. 7B,C), a more informative analysis than simple over-representation
523 analysis of enrichment.

524 We first looked at GSEA GO enrichments for cases where the GO term had a significant
525 score for all three time points, 1h, 3h, and 24h (Fig 7B). We found a total of seven GO
526 categories. The predominant pattern was for gene activation, seen for 17 of the 21 data
527 points. The only GO category that showed more suppression than activation was the
528 *carbohydrate derivative biosynthetic process* (GO:1901137) from hmc7. Interestingly,
529 *circulatory system development* (GO:0072359) from hmc7 and *blood vessel*
530 *development* (GO:0001568) from hmc3 were repressed at 1h, but then activated at 3h
531 and 24h. All the others, including *cellular response to stress* (GO:0033554) from hmc9
532 and *cellular response to hypoxia* (GO:0071456) from hmc9 were activated at all three
533 timepoints.

534 We also looked at the most significant GO categories that were detected in only two of
535 the three time points, displaying the top results for each pair group (Fig 7C) or listing the
536 full set (Supplementary Zip File 5, 6, 7). Among these, we find notable the detection of
537 biologically and hypoxia-relevant GO terms such as response to oxidative stress, DNA
538 repair, regulation of angiogenesis, and ncRNA processing. On the other hand, some
539 categories have less obvious but potentially important metabolic roles such as
540 organonitrogen compound biosynthesis, nucleoside triphosphate metabolism, and
541 protein localization to organelle. Similarly, several other processes appear to reflect
542 cellular or organism developmental processes such as neuron or positive regulation of
543 apoptosis, cell proliferation, and multicellular organism development.

544 These findings indicate that responses may be variably modulated in the very earliest
545 time points, 1 and 3 hours after hypoxia. For instance, three GO categories switch from
546 repressed at 1h to activated at 3h. These are *blood vessel development* (hmc3),
547 *circulatory system development* (hmc7), are suppressed at 1h, activated at 3h, and they

548 remain activated at 24h. Interestingly, *negative regulation of response to stimulus*
549 (hmc3) also shows this switching pattern of suppressed 1h then activated at 3h. This
550 pathway may temporarily allow for some acute hypoxia signaling, possibly via
551 metabolite flux, which could relate to the activation of *carbohydrate derivative*
552 *biosynthetic process* pathways (hmc7) at 1h followed by suppression at 3h and 24h.
553 These observations open a range of new questions that could be further addressed with
554 additional temporal resolution during the first few hours of stress exposure, when the
555 cistrome may be responding in ways not yet detectable through transcript abundance
556 changes.

557 Clustering partitions the genomic response into HIF1A-associated 558 TF cofactors and HIF1A-independent factors.

559 To further investigate the role of HIF1A and the potential impact of cofactors on
560 hypoxia's effect on the chromatin landscape, as summarized in Figure 8, we examined
561 these regions for co-occurring or proximal non-HIF motifs. We utilized a consensus
562 motif, vdACGTGh, from the HIF1A ChIP-seq dataset model,
563 HIF1A_HUMAN.H11MO.0.C, within the HOCOMOCOv11 database (30). This motif is
564 also known to share significant similarity with other hypoxia-responsive elements
565 (HREs) (41).

566 Using the MEME suite SEA tool, we filtered the hypoxia-induced gain of footprint peaks,
567 retaining only those that contained an entire HIF1A motif (Supplementary Zip File 3).
568 The gain of footprints is conspicuous when MOA coverage is centered on these HIF1A
569 motifs (Figure 8A). We observed that of the 2,640 total hypoxia-induced footprint gain
570 peaks from all time points, 775 of them (~30%) had at least one HIF1A motif, the
571 majority of which has only one HIF1A motif per peak. With increasing duration of
572 exposure to hypoxia, we observed an increasing number of HIF1A-containing footprints;
573 45 motifs at 6% of the 1h peaks, 213 motifs at 29% of 3h peaks, and 720 motifs at 43%
574 of peaks of the 24h peaks (Figure 8A).

575 Next, we examined the distribution of HIF1A motifs across the previously defined HMC
576 clusters (Supplementary Table S11). The HMC clusters were divided into two main
577 groups based on HIF1A motif prevalence: the HIF1A Group (HMCs 2, 3, 4, 5, and 9)
578 and the NON-HIF1A Group with none or very few HIF1A motifs (HMCs 1, 6, 7, 8, and
579 10) (Figure 8B). Interestingly, at each time point, different clusters exhibited the highest
580 count of HIF1A motifs, indicating a dynamic shift in motif occupancy over time. Notably,
581 at 1 hour, HMC2 showed the highest count with 43 HIF1A motifs; at 3 hours, HMC4 had
582 147 motifs; and by 24 hours, HMC9 showed the largest count with 390 motifs (Figure
583 8C).

584 To identify transcription factor motifs enriched in our two cluster groups (HIF1A vs.
585 NON-HIF1A), we used the motifs discovered from our previous analysis (Figure 7A).
586 We required each TF to be present in at least two clusters for the NON-HIF1A group
587 and in at least three clusters for the HIF1A group (Supplementary Table S12). Of the 69
588 TF motifs identified, 14 motifs, including ETV1 and SMAD2, were unique to the HIF1A
589 group. A total of 21 motifs, including MYF6 and MYOD1, were unique to the NON-
590 HIF1A group, and 34 motifs, including well-known factors like FOS, JUND, and ELK1,
591 were shared between the two groups (Figure 8D).

592 This prompted us to examine which of these TF motifs were in immediate proximity (+/-
593 50 bp) to the HIF1A motif as a way to detect potential HIF1A-interacting TFs
594 (Supplementary Zip File 4). The HIF1A-proximal motifs defined in this way (asterisks,
595 Figure 8D) were found in 64% of motifs in the uniquely HIF1A group, 47% of motifs in
596 the shared group, but only 9.5% of the uniquely non-HIF1A group. From this, we
597 conclude that the co-occurrence of TF motifs in the HIF1A group may identify candidate
598 interacting factors. In contrast, TF motifs enriched in the NON-HIF1A group may identify
599 a separate set of factors that function independently of HIF1A, while still participating in
600 hypoxia-induced regulatory networks or pathways. This differential motif distribution
601 highlights the complexity of transcriptional regulation under hypoxia, and it will be
602 important to further investigate how these distinct sets of TFs might work together or in
603 parallel to shape the broader cistrome response.

604 Together, these findings suggest that HIF1A binding does not occur in isolation but
605 within broader regulatory networks, including potential co-occupancy with ETS, AP-1,
606 and oxidative stress-responsive TFs. The identification of non-HIF1A clusters enriched
607 for distinct TF motifs implies the existence of parallel hypoxia-responsive regulatory
608 programs that may function independently of classical HIF signaling.

609 In summary, we demonstrate the use of MOA-seq with MNase footprinting to study
610 small particle dynamics in HUVECs following hypoxia, generating a time-resolved atlas
611 of over 5,000 genomic regions with significant cis-regulatory changes. This approach
612 revealed dynamic remodeling of TF occupancy across the genome, identifying genes
613 with biologically relevant functional ontologies. By clustering these dynamic footprints,
614 we distinguished between HIF1A-associated and HIF1A-independent regulatory
615 programs, both of which likely contribute to the transcriptional control of the hypoxic
616 response.

617 Discussion

618 In this study, we successfully applied the MOA-seq (MNase-defined cistrome-
619 Occupancy Analysis) assay to human umbilical vein endothelial cells (HUVECs) to map
620 transcription factor (TF) occupancy and chromatin dynamics in response to hypoxia at
621 multiple time points. Our results show that this strategy not only recapitulates known
622 regulatory regions annotated by ENCODE (38) but also uncovers a substantial number
623 of additional candidate *cis*-regulatory elements. A strength of this approach lies in its
624 simplicity as a single assay (Figure 1) that, although straightforward to execute,
625 provides a breadth of regulatory data for integration, elucidating pathways, and
626 identifying candidate CREs.

627 These experiments generated high-resolution maps of TF occupancy that closely align
628 with established regulatory features. The MOA-seq coverage profiles of open and
629 accessible chromatin exhibited characteristic peaks at transcription start sites (TSS) and
630 transcription end sites (TES), correlating with gene expression levels (Figure 2A,B).
631 Importantly, a significant portion of MOA-seq peaks (~34%) did not overlap with

632 ENCODE candidate CREs or DHS sites, demonstrating the ability of MOA-seq to
633 capture additional regulatory regions missed by conventional assays (Figure 2F). This
634 finding is consistent with previous studies in maize, where MOA-seq identified hundreds
635 of motif families in less than 0.1% of the genome, vastly expanding the known cistrome
636 (2).

637 Our analysis showed that the number of diff-MOA peaks increased over time relative to
638 the 0-hour normoxia samples, reflecting a growing list of regions that respond to
639 hypoxia (Figure 4A). Collectively, these diff-MOA peaks, both gains and losses, were
640 distributed among both cCRE and non-cCRE regions (Figure 4D), with an average of
641 two diff-MOA peaks per gene (Supplementary Table S5). The fact that both cCRE-
642 overlapping and non-cCRE diff-MOA regions showed comparable enrichment for TF
643 motifs suggests that these additional sites are biologically significant (Figure 4E). This
644 comprehensive analysis of chromatin fine-structural changes in hypoxia should provide
645 a more nuanced but also holistic view of the complex cellular adaptations to low oxygen
646 conditions.

647 The combination of diff-MOA genes and differentially expressed genes (DEGs) revealed
648 enhanced detection of functionally relevant pathways (Figure 5C). Although only a
649 moderate overlap was observed between diff-MOA genes and DEGs, the combined
650 gene sets showed stronger enrichment for hypoxia-related pathways, such as VEGFA-
651 VEGFR2, Hedgehog, and mTORC1 signaling (Figure 5B,C). This suggests that
652 changes in TF occupancy and chromatin accessibility, as detected by MOA-seq,
653 precede or complement transcriptional changes captured by RNA-seq. In this context, it
654 may be important to compare MOA-seq particle dynamics with transcription rates
655 measured more directly via nascent RNA or run-on RNA-seq experiments. The
656 identification of functional pathways from diff-MOA peaks alone highlights the utility of
657 MOA-seq in capturing early regulatory events (Figure 5A).

658 The use of fragment center (frenters) in MOA-seq enabled high-resolution peak calling,
659 which strengthened the identification of TF motifs and candidate TF binding sites
660 (Figure 6). With over 100,000 peaks overall, and more than 6,000 hypoxia-induced diff-

661 MOA peaks, we identified a large set of regions and motifs. To more effectively handle
662 the scale and complexity of our results, we implemented a peak clustering approach
663 based on the assumption that shared variation in chromatin dynamics reflects shared
664 functional or genomic response pathways. This method not only made the data more
665 manageable but also captured sample relationships beyond the pairwise comparison of
666 the three hypoxia time points to normoxia. We identified ten clusters of hypoxia-induced
667 regions, each exhibiting unique patterns of chromatin accessibility changes over time
668 (Fig. 7A). Motif analysis on a cluster-by-cluster basis demonstrated that certain TF
669 motifs are preferentially associated with specific kinetic patterns. For instance, CTCF,
670 FOS, MAFB, and ETV4 were broadly enriched in the diff-MOA datasets (Figure 6), but
671 after clustering, each of these motifs was limited to a smaller subset of the ten clusters
672 (Figure 7A).

673 Integrating these cluster-specific regions with differential gene expression data and
674 conducting GSEA uncovered that each cluster is linked to distinct biological processes.
675 Notably, some clusters were associated with the activation of stress response and cell
676 growth processes (cluster 9), angiogenesis and metabolic processes (cluster 7), while
677 others were associated primarily with DNA repair (cluster 4). All of these processes can
678 be reasonably implicated in the adaptive response to hypoxia, including DNA repair
679 pathways, which have been linked to hypoxia-regulated transcription (42). Interestingly,
680 certain gene ontologies exhibited temporal shifts from repression to activation (Figure
681 7B,C; red vs. blue circles, respectively), highlighting dynamic regulatory mechanisms in
682 response to hypoxia. For example, pathways related to blood vessel development
683 shifted from being repressed at 1 hour to being activated at 3 and 24 hours post-
684 hypoxia (Figure 7B). These findings suggest that clustering MOA-seq coverage kinetics
685 is an effective strategy to partition and investigate the diverse transcriptional regulatory
686 networks that modulate the temporal orchestration of gene regulation at the chromatin
687 level during hypoxic stress.

688 Among these different networks, we observed that some were affiliated with HIF1A,
689 while others were not. Given that HIF1A is a well-known hypoxia-activated transcription
690 factor, we focused on the clusters enriched for HIF1A and the other TFs present in

691 those clusters (Fig. 8). This cohort of transcription factors (Fig. 8D, TFs with *) whose
692 motifs are frequently located near HIF1A motifs can be considered potential co-binding
693 or cooperative regulatory TFs. Among the 27 TFs co-enriched with HIF1A, we observed
694 three distinct groups: the ETS family, the FOS/JUN family, and the NFE2L2 and MAF
695 family.

696 For the ETS family, TFs such as ELK1 are known to interact directly with HIF1A to
697 upregulate genes, including the proto-oncogene CD133, a cancer stem cell marker
698 gene involved in angiogenesis and adaptation to hypoxic conditions (43). The ETS
699 family gene ETV4 plays a broad role in hypoxia-induced signaling. ETV4 forms
700 complexes with both HIF1A and HIF2A to enhance the transcription of hypoxia-
701 responsive genes, promoting cellular survival, angiogenesis, and metabolic adaptation,
702 particularly in cancer (44). Given that nearly 8% of hypoxically induced transcripts
703 require ETV4 for efficient induction (44), it is not surprising that we uncovered it using
704 our chromatin TF particle dynamics profiling. For the FOS/JUN family, components of
705 the AP-1 complex with broad gene activation functions, we also expected to find these.
706 For instance, the AP-1 complex interacts with HIF1A to regulate angiogenesis, and
707 more specifically, FOSL2 is known to bind to HIF1A target genes (45, 46). These AP-1
708 complex factors were clearly enriched as HIF1A co-occurrence motifs in our MOA
709 clusters, adding to the growing evidence of their role in the regulatory response to
710 hypoxia.

711 For the NFE2L2/NRF2 and MAF family, the co-enrichment with HIF1A motifs is
712 particularly intriguing, given their role in binding oxidative stress response elements
713 (47). NRF2 upregulates shared target genes with HIF1A to cooperatively promote tumor
714 growth and progression and directly influences HIF1A expression by binding to a
715 conserved antioxidant response element (ARE) upstream of the HIF1A locus (48). This
716 binding event drives ROS-responsive gene expression and enhances HIF1A
717 expression, particularly in cancers with oncogenic NRF2 mutations (49). In cancer
718 models, NRF2 activation has been shown to stabilize HIF1A protein by directly
719 interacting with its oxygen-dependent degradation domain, enhancing the transcription
720 of angiogenic and survival genes (50). Our detection of NRF2 and MAF motifs in

721 proximity to HIF1A-binding sites supports a model in which oxidative stress and hypoxia
722 signaling converge at shared cis-regulatory elements. This co-occupancy may reflect a
723 coordinated adaptation to redox and oxygen stress, particularly in endothelial contexts
724 where both pathways are active. Together, these findings highlight NRF2's dual function
725 as a transcriptional regulator and post-translational stabilizer of HIF1A, suggesting an
726 integrated regulatory axis that could contribute to context-specific hypoxia responses
727 and may be of interest in disease settings such as cancer and ischemia (49, 50).

728 We recognize that many of the interactions described here are already known and,
729 based on single-gene experiments or other more laborious approaches, may have been
730 easy to predict. But, it is important to note that we uncovered these relationships in an
731 orthogonal assay independent of transcriptomic or genetic data. In this regard, finding
732 these well-known transcriptional regulators validates the biological relevance of
733 observations from MOA-seq. Furthermore, it strengthens our confidence in, and interest
734 in, the other, less obvious discoveries of TFs, which we refer to as HIF1A-independent
735 pathways.

736 One of the significant challenges in understanding the role of HIF1A is the fact that,
737 despite numerous studies, there is remarkably little agreement about HIF1A target
738 genes and true HIF1A-binding sites (51). Increasingly, it is recognized that HIF1A may
739 operate in concert with co-factors or other transcription factors at sites where HIF1A
740 ChIP-seq peaks or HIF1A DNA motifs do not occur. This suggests that a multitude of
741 other TFs and chromatin-modifying activities may modulate genomic responses to
742 hypoxia. In support of this idea, we document TF-like footprint changes that respond to
743 hypoxia, but at locations not enriched for HIF1A motifs, which we refer to as our non-
744 HIF1A TFs from clusters 1, 6-8, and 10 (Fig. 8).

745 Among the 19 transcription factors enriched in diff-MOA clusters lacking strong HIF1A
746 motif co-occurrence, some have previously reported interactions with HIF1A, while
747 others may represent novel candidates. For example, MYOD1 has been shown to
748 influence hypoxia signaling indirectly through MyoD-mediated activation of the
749 noncanonical Wnt/β-catenin pathway, which in turn modulates HIF1A activity in the

750 context of myogenic differentiation (52). Similarly, OLIG2 and HIF1A have been
751 reported to cooperatively regulate noncanonical target genes in oligodendrocyte
752 progenitor cells (OPCs), impairing differentiation under hypoxic conditions (53). Notably,
753 this OLIG2-HIF1A interaction appears to be cell-type specific, raising the possibility that
754 OLIG2 motif enrichment in our dataset reflects broader hypoxia co-responsiveness,
755 even in the absence of direct HIF1A co-occurrence. Together, these observations
756 suggest that TFs in non-HIF1A clusters may operate through parallel or context-specific
757 pathways, and highlight the potential of MOA-seq to uncover diverse transcriptional
758 regulators beyond the canonical HIF axis.

759 Several transcription factors enriched in non-HIF1A diff-MOA clusters have limited or no
760 previously established relationship to HIF1A but emerged from our analysis as
761 potentially important regulators of the hypoxia cistrome. One such factor is ZEB1, which
762 is induced by hypoxia and may act through parallel, HIF1A-independent pathways. Both
763 HIF1A and ZEB1 have been implicated in promoting epithelial-to-mesenchymal
764 transition (EMT), and a HIF1A/ZEB1 axis has been proposed as a driver of malignant
765 transformation in certain cancers (54). Another notable factor is TF7L1 (also known as
766 TCF7L1), which is enriched exclusively in diff-MOA cluster 6. This cluster was
767 associated with genes exhibiting a positive response to oxidative stress at 3 hours post-
768 hypoxia (Supplementary Zip File 6). TF7L1 is upregulated in multiple aggressive tumor
769 types and has been shown to promote antioxidant responses via the KEAP1/NRF2
770 pathway (55), suggesting a potential role in early transcriptional priming during hypoxic
771 stress. These observations point to previously underappreciated transcriptional
772 regulators that may play temporally distinct or cell-type–specific roles in the broader
773 hypoxia response.

774 The distinction between HIF1A-associated and HIF1A-independent transcription factors
775 underscores the complexity of the hypoxic transcriptional landscape and raises several
776 mechanistic possibilities. HIF1A-associated TFs identified through our motif co-
777 occurrence analysis may work in concert with HIF1A, enhancing the expression of
778 genes critical for immediate adaptation to low oxygen conditions, such as those involved
779 in angiogenesis, metabolism, and survival pathways. These cooperative interactions

780 may refine the hypoxic response and contribute to its specificity across diverse cellular
781 contexts. In contrast, HIF1A-independent TFs may regulate alternative or
782 complementary pathways, potentially driving cell-type–specific functions, developmental
783 programs, or immune responses. This broader regulatory input could support long-term
784 or context-dependent adaptations beyond the direct influence of HIF1A. By resolving
785 distinct temporal patterns in TF occupancy, our clustering analysis offers a mechanistic
786 framework for understanding how concurrent transcriptional networks coordinate the
787 multifaceted response to hypoxia.

788 Future studies should integrate MOA-seq with additional epigenomic and transcriptomic
789 datasets to build a comprehensive model of gene regulation under hypoxia. Expanding
790 the analysis to additional cell types, time points, or environmental perturbations will
791 further clarify how chromatin remodeling and TF network dynamics vary across
792 biological contexts. Functional validation of novel regulatory elements identified here will
793 be essential to define their roles in gene regulation and disease progression.

794 In summary, cistrome occupancy profiling provides a detailed and expanded view of
795 how the HUVEC genome changes under hypoxic conditions. The assay's ability to
796 capture both known and novel regulatory elements, along with temporal changes in TF
797 occupancy, offers valuable insights into the complex regulatory networks orchestrating
798 the hypoxic response. Our findings highlight the importance of comprehensive
799 chromatin mapping for uncovering the full spectrum of hypoxia-responsive regulatory
800 elements and lay the groundwork for future efforts to dissect transcriptional control
801 mechanisms in health and disease.

802

References

1. Yan,F., Powell,D.R., Curtis,D.J. and Wong,N.C. (2020) From reads to insight: a hitchhiker's guide to ATAC-seq data analysis. *Genome Biol*, **21**, 22.

2. Savadel,S.D., Hartwig,T., Turpin,Z.M., Vera,D.L., Lung,P.-Y., Sui,X., Blank,M., Frommer,W.B., Dennis,J.H., Zhang,J., *et al.* (2021) The native cistrome and sequence motif families of the maize ear. *PLoS Genet.*, **17**, e1009689.

3. Parvathaneni,R.K., Bertolini,E., Shamimuzzaman,M., Vera,D.L., Lung,P.-Y., Rice,B.R., Zhang,J., Brown,P.J., Lipka,A.E., Bass,H.W., *et al.* (2020) The regulatory landscape of early maize inflorescence development. *Genome Biol.*, **21**, 165.

4. Hu,G., Grover,C.E., Vera,D.L., Lung,P.-Y., Girimurugan,S.B., Miller,E.R., Conover,J.L., Ou,S., Xiong,X., Zhu,D., *et al.* (2024) Evolutionary dynamics of chromatin structure and duplicate gene expression in diploid and allopolyploid cotton. *Mol. Biol. Evol.*, **41**.

5. Engelhorn,J., Snodgrass,S.J., Kok,A., Seetharam,A.S., Schneider,M., Kiwit,T., Singh,A., Banf,M., Khaipho-Burch,M., Runcie,D.E., *et al.* (2023) Genetic variation at transcription factor binding sites largely explains phenotypic heritability in maize. *bioRxiv*, 10.1101/2023.08.08.551183.

6. Rodgers-Melnick,E., Vera,D.L., Bass,H.W. and Buckler,E.S. (2016) Open chromatin reveals the functional maize genome. *Proc. Natl. Acad. Sci. U. S. A.*, **113**, E3177–84.

7. Niskanen,H., Tuszynska,I., Zaborowski,R., Heinäniemi,M., Ylä-Herttuala,S., Wilczynski,B. and Kaikkonen,M.U. (2018) Endothelial cell differentiation is encompassed by changes in long range interactions between inactive chromatin regions. *Nucleic Acids Res.*, **46**, 1724–1740.

8. Martin,M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, **17**, 10–12.

9. Magoč,T. and Salzberg,S.L. (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*, **27**, 2957–2963.

10. Vasimuddin,M., Misra,S., Li,H. and Aluru,S. (2019) Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems. In *2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*. IEEE, pp. 314–324.

11. Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G., Durbin,R. and 1000 Genome Project Data Processing Subgroup (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.

12. Crusoe,M.R., Alameldin,H.F., Awad,S., Boucher,E., Caldwell,A., Cartwright,R., Charbonneau,A., Constantinides,B., Edvenson,G., Fay,S., *et al.* (2015) The khmer software package: enabling efficient nucleotide sequence analysis. *F1000Res*, **4**, 900.

13. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.

840 14. Kent,W.J., Zweig,A.S., Barber,G., Hinrichs,A.S. and Karolchik,D. (2010) BigWig and
841 BigBed: enabling browsing of large distributed datasets. *Bioinformatics*, **26**, 2204–2207.

842 15. Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nusbaum,C.,
843 Myers,R.M., Brown,M., Li,W., *et al.* (2008) Model-based analysis of ChIP-Seq (MACS).
844 *Genome Biol.*, **9**, R137.

845 16. Eder,T. and Grebien,F. (2022) Comprehensive assessment of differential ChIP-seq tools
846 guides optimal algorithm selection. *Genome Biol.*, **23**, 119.

847 17. Li,Q., Brown,J.B., Huang,H. and Bickel,P.J. (2011) Measuring reproducibility of high-
848 throughput experiments. *aoas*, **5**, 1752–1779.

849 18. Landt,S.G., Marinov,G.K., Kundaje,A., Kheradpour,P., Pauli,F., Batzoglou,S.,
850 Bernstein,B.E., Bickel,P., Brown,J.B., Cayting,P., *et al.* (2012) ChIP-seq guidelines and
851 practices of the ENCODE and modENCODE consortia. *Genome Res*, **22**, 1813–1831.

852 19. Amemiya,H.M., Kundaje,A. and Boyle,A.P. (2019) The ENCODE Blacklist: Identification of
853 Problematic Regions of the Genome. *Sci. Rep.*, **9**, 9354.

854 20. An integrated encyclopedia of DNA elements in the human genome (2012) *Nature*, **489**, 57–
855 74.

856 21. Sabo,P.J., Hawrylycz,M., Wallace,J.C., Humbert,R., Yu,M., Shafer,A., Kawamoto,J., Hall,R.,
857 Mack,J., Dorschner,M.O., *et al.* (2004) Discovery of functional noncoding elements by
858 digital analysis of chromatin structure. *Proc Natl Acad Sci U S A*, **101**, 16837–16842.

859 22. Sabo,P.J., Kuehn,M.S., Thurman,R., Johnson,B.E., Johnson,E.M., Cao,H., Yu,M.,
860 Rosenzweig,E., Goldy,J., Haydock,A., *et al.* (2006) Genome-scale mapping of DNase I
861 sensitivity in vivo using tiling DNA microarrays. *Nat Methods*, **3**, 511–518.

862 23. Hammal,F., de Langen,P., Bergon,A., Lopez,F. and Ballester,B. (2022) ReMap 2022: a
863 database of Human, Mouse, Drosophila and Arabidopsis regulatory regions from an
864 integrative analysis of DNA-binding sequencing experiments. *Nucleic Acids Res.*, **50**,
865 D316–D325.

866 24. Ramírez,F., Dündar,F., Diehl,S., Grüning,B.A. and Manke,T. (2014) deepTools: a flexible
867 platform for exploring deep-sequencing data. *Nucleic Acids Res.*, **42**, W187–91.

868 25. Frankish,A., Carbonell-Sala,S., Diekhans,M., Jungreis,I., Loveland,J.E., Mudge,J.M.,
869 Sisu,C., Wright,J.C., Arnan,C., Barnes,I., *et al.* (2023) GENCODE: reference annotation for
870 the human and mouse genomes in 2023. *Nucleic Acids Res*, **51**, D942–D949.

871 26. Chen,E.Y., Tan,C.M., Kou,Y., Duan,Q., Wang,Z., Meirelles,G.V., Clark,N.R. and Ma'ayan,A.
872 (2013) Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool.
873 *BMC Bioinformatics*, **14**, 128.

874 27. Xie,Z., Bailey,A., Kuleshov,M.V., Clarke,D.J.B., Evangelista,J.E., Jenkins,S.L.,
875 Lachmann,A., Wojciechowicz,M.L., Kropiwnicki,E., Jagodnik,K.M., *et al.* (2021) Gene Set
876 Knowledge Discovery with Enrichr. *Curr Protoc*, **1**, e90.

877 28. Wickham,H. ggplot2 Springer New York.

878 29. Timothy L. Bailey, James Johnson, Charles E. Grant, William S. Noble (2015) The MEME
879 Suite. *Nucleic Acids Research*, **43**.

880 30. Kulakovskiy,I.V., Vorontsov,I.E., Yevshin,I.S., Sharipov,R.N., Fedorova,A.D.,
881 Rumynskiy,E.I., Medvedeva,Y.A., Magana-Mora,A., Bajic,V.B., Papatsenko,D.A., *et al.*
882 (2018) HOCOMOCO: towards a complete collection of transcription factor binding models
883 for human and mouse via large-scale ChIP-Seq analysis. *Nucleic Acids Res*, **46**, D252–
884 D259.

885 31. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and
886 dispersion for RNA-seq data with DESeq2. *Genome Biol*, **15**, 550.

887 32. Subramanian,A., Tamayo,P., Mootha,V.K., Mukherjee,S., Ebert,B.L., Gillette,M.A.,
888 Paulovich,A., Pomeroy,S.L., Golub,T.R., Lander,E.S., *et al.* (2005) Gene set enrichment
889 analysis: a knowledge-based approach for interpreting genome-wide expression profiles.
890 *Proc Natl Acad Sci U S A*, **102**, 15545–15550.

891 33. Mootha,V.K., Lindgren,C.M., Eriksson,K.-F., Subramanian,A., Sihag,S., Lehar,J.,
892 Puigserver,P., Carlsson,E., Ridderstråle,M., Laurila,E., *et al.* (2003) PGC-1α-responsive
893 genes involved in oxidative phosphorylation are coordinately downregulated in human
894 diabetes. *Nature Genetics*, **34**, 267–273.

895 34. Benoit,J., Sheikhbahaei,M.K. and Dennis,J. (2021) Chromatin dynamics: Nucleosome
896 occupancy and sensitivity as determinants of gene expression and cell fate. *J Cancer Biol*,
897 **2**, 51–55.

898 35. Minnoye,L., Marinov,G.K., Krausgruber,T., Pan,L., Marand,A.P., Secchia,S.,
899 Greenleaf,W.J., Furlong,E.E.M., Zhao,K., Schmitz,R.J., *et al.* (2021) Chromatin accessibility
900 profiling methods. *Nat Rev Methods Primers*, **1**.

901 36. Chawla,A., Nagy,C. and Turecki,G. (2021) Chromatin Profiling Techniques: Exploring the
902 Chromatin Environment and Its Contributions to Complex Traits. *Int. J. Mol. Sci.*, **22**.

903 37. Klein,D.C. and Hainer,S.J. (2020) Genomic methods in profiling DNA accessibility and factor
904 localization. *Chromosome Res.*, **28**, 69–85.

905 38. ENCODE Project Consortium, Moore,J.E., Purcaro,M.J., Pratt,H.E., Epstein,C.B.,
906 Shoresh,N., Adrian,J., Kawli,T., Davis,C.A., Dobin,A., *et al.* (2020) Expanded
907 encyclopaedias of DNA elements in the human and mouse genomes. *Nature*, **583**, 699–
908 710.

909 39. Payne,S., Neal,A. and De Val,S. (2024) Transcription factors regulating vasculogenesis and
910 angiogenesis. *Dev. Dyn.*, **253**, 28–58.

911 40. Kuleshov,M.V., Jones,M.R., Rouillard,A.D., Fernandez,N.F., Duan,Q., Wang,Z., Koplev,S.,
912 Jenkins,S.L., Jagodnik,K.M., Lachmann,A., *et al.* (2016) Enrichr: a comprehensive gene set
913 enrichment analysis web server 2016 update. *Nucleic Acids Res.*, **44**, W90–7.

914 41. Cabaj A, Moszyńska A, Charzyńska A, Bartoszewski R, Dąbrowski M (2022) Functional and
915 HRE motifs count analysis of induction of selected hypoxia-responsive genes by HIF-1 and
916 HIF-2 in human umbilical endothelial cells. *Cellular Signalling*, **90**, 110209.

917 42. Cowman,S., Pizer,B. and Sée,V. (2021) Downregulation of both mismatch repair and non-
918 homologous end-joining pathways in hypoxic brain tumour cell lines. *PeerJ*, **9**, e11275.

919 43. Ohnishi,S., Maehara,O., Nakagawa,K., Kameya,A., Otaki,K., Fujita,H., Higashi,R.,
920 Takagi,K., Asaka,M., Sakamoto,N., *et al.* (2013) hypoxia-inducible factors activate CD133
921 promoter through ETS family transcription factors. *PLoS One*, **8**, e66255.

922 44. Wollenick,K., Hu,J., Kristiansen,G., Schraml,P., Rehrauer,H., Berchner-Pfannschmidt,U.,
923 Fandrey,J., Wenger,R.H. and Stiehl,D.P. (2012) Synthetic transactivation screening reveals
924 ETV4 as broad coactivator of hypoxia-inducible factor signaling. *Nucleic Acids Res*, **40**,
925 1928–1943.

926 45. Zhang,Y., Wang,S., Hu,H. and Li,X. (2022) A systematic study of HIF1A cofactors in hypoxic
927 cancer cells. *Sci Rep*, **12**, 18962.

928 46. Leiherer,A., Geiger,K., Muendlein,A. and Drexel,H. (2014) Hypoxia induces a HIF-1α
929 dependent signaling cascade to make a complex metabolic switch in SGBS-adipocytes.
930 *Mol Cell Endocrinol*, **383**, 21–31.

931 47. Fuse,Y. and Kobayashi,M. (2017) Conservation of the Keap1-Nrf2 System: An Evolutionary
932 Journey through Stressful Space and Time. *Molecules*, **22**.

933 48. Bae,T., Hallis,S.P. and Kwak,M.-K. (2024) Hypoxia, oxidative stress, and the interplay of
934 HIFs and NRF2 signaling in cancer. *Exp Mol Med*, **56**, 501–514.

935 49. Lacher,S.E., Levings,D.C., Freeman,S. and Slattery,M. (2018) Identification of a functional
936 antioxidant response element at the HIF1A locus. *Redox Biol*, **19**, 401–411.

937 50. Zheng,J., Kim,S.-J., Saeidi,S., Kim,S.H., Fang,X., Lee,Y.-H., Guillen-Quispe,Y.N.,
938 Ngo,H.K.C., Kim,D.-H., Kim,D., *et al.* (2023) Overactivated NRF2 induces pseudohypoxia in
939 hepatocellular carcinoma by stabilizing HIF-1α. *Free Radic Biol Med*, **194**, 347–356.

940 51. Yfantis,A., Mylonis,I., Chachami,G., Nikolaidis,M., Amoutzias,G.D., Paraskeva,E. and
941 Simos,G. (2023) Transcriptional Response to Hypoxia: The Role of HIF-1-Associated Co-
942 Regulators. *Cells*, **12**.

943 52. Cirillo,F., Resmini,G., Ghiroldi,A., Piccoli,M., Bergante,S., Tettamanti,G. and Anastasia,L.
944 (2017) Activation of the hypoxia-inducible factor 1α promotes myogenesis through the
945 noncanonical Wnt pathway, leading to hypertrophic myotubes. *FASEB J*, **31**, 2146–2156.

946 53. Allan,K.C., Hu,L.R., Scavuzzo,M.A., Morton,A.R., Gevorgyan,A.S., Cohn,E.F.,
947 Clayton,B.L.L., Bederman,I.R., Hung,S., Bartels,C.F., *et al.* (2021) Non-canonical Targets
948 of HIF1a Impair Oligodendrocyte Progenitor Cell Function. *Cell Stem Cell*, **28**, 257–
949 272.e11.

950 54. Yoshimoto,S., Tanaka,F., Morita,H., Hiraki,A. and Hashimoto,S. (2019) Hypoxia-induced
951 HIF-1α and ZEB1 are critical for the malignant transformation of ameloblastoma via TGF-β-
952 dependent EMT. *Cancer Med*, **8**, 7822–7832.

953 55. Zhang,B., Wu,J., Cai,Y., Luo,M., Wang,B. and Gu,Y. (2019) TCF7L1 indicates prognosis
954 and promotes proliferation through activation of Keap1/NRF2 in gastric cancer. *Acta
955 Biochim Biophys Sin (Shanghai)*, **51**, 375–385.

Data Availability

All MOA-seq data discussed in this publication have been deposited at NCBI SRA under accession number PRJNA1217882. MOA bigwig and peak files are available on Gene Expression Omnibus (ID: GSE289863). MOA processing and custom analysis scripts have been deposited in the GitHub repository (https://github.com/as5906/HypoxiaMOA-HUVEC).

Acknowledgements

This work was supported in part by an NSF Plant Genome Research Program grant to HWB (NSF IOS awards 1444532, 2025811); a Phi Beta Kappa MJ Hay Award and Biological Science Faculty Award for Undergraduate Scholarship to AS; and an American Heart Association grant to GTD (#830166).

Figure Legends

Figure 1. Transcription Factor Occupancy Profiling with MNase "MOA-seq" in HUVECs following hypoxia for 1, 3, and 24h. MOA-seq, MNase-defined cistrome-Occupancy Analysis, (Savadel et al. 2021), was developed to globally identify protein-bound cis-regulatory sequences in a single robust assay using modified nuclease profiling of fixed chromatin and customized bioinformatic analysis pipelines. MOA-seq was applied to HUVEC cells stressed with hypoxia to map short (1h and 3h) or long (24h) exposure-induced changes in TF occupancies. **(A)** The treatment scheme for the four samples used in this study: 0h = normoxia; 1h, 3h, and 24h refer to hypoxia (2% oxygen) exposure times. **(B)** Diagrammatic summary of the MOA-seq, starting with formaldehyde-fixed cells to preserve chromatin structure including bound transcription factors (TFs) at their cognate binding sites (cis-elements). Light/partial digest with MNase preferentially cuts DNA within nucleosome-free/accessible chromatin regions (NRF/ACR), but not where protein occupancy occurs. Recovery of de-crosslinked, small DNA fragments from ideal partial digests (yellow asterisks, inset gel) is used for NGS library production. MOA-seq coverage profiles of aligned read pair fragments (orange) and their fragment centers (frenters, dark orange) reveal footprints for analysis. The frenter peaks reflect protein-protected footprints within the NFR/ACRs (e.g. cis-element 1 and 2) but do not mark unoccupied cis-elements (e.g. cis-element 3). Peak regions are used for comparative genomics and motif discovery and analysis.

986
987 **Figure 2. MOA-seq maps gene regulatory features at ENCODE cCREs and additional sites**
988 **with HUVEC cell-type specificity in normoxic cells.** MOA-seq coverage profiles were
989 compared to ENCODE and other genomic annotations related to gene expression and
990 regulation. **(A,B)** Profile trend plots of MOA coverage showing peaks at the transcription start
991 (TSS) and end (TES) sites of genes, with peak heights positively correlating with transcript
992 abundance quartiles. **(C)** Profile trend plots of MOA frenter coverage profiles colocalize with
993 cCRE subtypes (K4m3, DNase-H3K4me3; dELS, distal enhancer-like signature; pELS, proximal
994 enhancer-like signature; PLS, promoter-like signature). **(D)** Intersection significance is shown by
995 Z-scores calculated from peak bp intersections with cCRE subtypes. **(E)** Intersection
996 significance is shown by Z-scores calculated from peak bp intersections with DNaseI
997 hypersensitive (DHS) peaks from multiple different cell types. **(F)** The abundance of MOA-peaks
998 relative to DHS regions or known cCRE regions. **(G)** Genome browser session showing an
999 example of a MOA peak within a cCRE and DHS region (top) and another example in a non-
1000 cCRE and non-DHS region (bottom), with tracks labeled on the left.
1001
1002 **Figure 3. Enrichment of MOA at ReMap ChIP summits for 13 characteristically HUVEC**
1003 **transcription factors.** MOA-seq enrichment at previously characterized TF binding sites as
1004 defined by ChIP-seq, using ReMAP (Hammal et al., 2021) ChIP summits for 13 canonical
1005 HUVEC TFs (Payne et al., 2023). Profile trend **(A)** and heatmap **(B)** plots of MOA 0hr frenters
1006 coverage centered on ReMap HUVEC ChIP peak summits. Data plotted were from regions with
1007 non-zero MOA coverage of ~250k summits from a total of ~750k spanning various studies and
1008 treatments. **(C)** Bar chart showing total unfiltered enrichment scores (MOA cov summit/flanks)
1009 for a select subset of TF ChIPs from experiments conducted in HUVEC only, with nearby (< 10
1010 bp) summits for MOA and ChIP. **(D)** Profile trend plots of local MOA frenter coverage for the
1011 individual TFs shown in panel C, grouped primarily by signal height.
1012
1013 **Figure 4. Summary of hypoxia-induced differential MOA (diff-MOA) peaks and genes,**
1014 **along with the enrichment analysis for regulation by specific transcription factors.** To
1015 identify genomic regions that changed their TF occupancy footprints in response to hypoxia,
1016 bdgdiff (MACS3) of MOA frenter coverage was used with the elbow method. **(A)** Table of diff-
1017 MOA peak numbers and sizes for diff-MOA peaks after hypoxia relative to 0h. Total diff-MOA
1018 peaks (gain + loss) were 1,978, 1,932, and 2,586 after 1h, 3h, and 24h hypoxia, respectively.
1019 **(B)** Genome browser session showing two examples each for diff-MOA gain, diff-MOA loss, or
1020 shared for 24h hypoxia treatment. **(C)** MOA-seq frenter coverage heatmaps of the diff-MOA
1021 (GAIN vs LOSS of footprints relative to normoxia) for each time point. Heatmaps were sorted by
1022 total MOA signal averaged over the 1 kb window centered on the peaks. In addition, sorting was
1023 done for diff-MOA gain sites using the hypoxic time point as the primary sort, and for diff-MOA
1024 loss sites, using the normoxic time point as the primary sort. **(D)** Venn diagram of genes with
1025 diff-MOA footprints for cCRE versus non-cCRE regions. Genes in the intersection had multiple
1026 diff-MOA footprints, with at least one of each cCRE type. **(E)** Enrichments (adjusted p-values)
1027 for gene lists derived from diff-MOA peaks in cCRE versus non-cCRE regions. Results from
1028 ENRICHR (https://maayanlab.cloud/Enrichr/) using the libraries for ENCODE and ChEA

1029 Consensus TFs from ChIP-X. The plot was sorted by P-values for the total Diff-MOA peaks
1030 (grey bars) and all significant TFs (dashed lines) are shown.
1031
1032 **Figure 5. Comparison of differential MOA with differential gene expression.**
1033 Comparison of hypoxia-induced diff-MOA genes to hypoxia-induced differentially expressed
1034 genes (DEGs) with regard to gene enrichment analyses, changes in RNA levels, and gain or
1035 loss of diff-MOA footprints. (**A**) Gene lists defined by diff-MOA (gain or loss) for each hypoxia
1036 time point were used as input for the ENRICHR libraries, and top hits are shown for two
1037 libraries, Human MSigDB Hallmark and WikiPathway Human gene sets. The significance (-
1038 log10(FDR)) is proportional to the circle size; gene list sizes are indicated at the bottom. (**B**)
1039 Direction relationships of diff-MOA (gain or loss) versus DEG (transcript up or down) is plotted
1040 for each time point. (**C**) Analysis of genes showing both diff-MOA and DEGs. The top shows
1041 Venn diagrams for each time point. The bottom shows ENRICHR results using the same two
1042 libraries as in panel A. To compare the intersected diff-MOA and DEG gene sets, we sampled
1043 the larger DEG gene sets using the intersection size-matched gene number. For each such
1044 comparison, the average (n=3) of the -log10(FDR) was less significant or insignificant.
1045
1046 **Figure 6. Motif discovery and analysis at diff-MOA peaks identify candidate hypoxia-**
1047 **responsive gene regulatory network cis-elements.**
1048 The diff-MOA peaks were split into six sets based on gain (**A**) or loss (**B**) of footprint at each of
1049 the three hypoxia treatment timepoints. The genomic sequences under the peaks were used as
1050 input for MEME motif analysis. For each list of motifs, three results are shown; the top *de novo*
1051 motif and its sequence LOGO, the TF families (italics) and their members (bold in parentheses),
1052 and the top 10 individual TFs sorted alphabetically. For this last list each motif shows the
1053 background-adjusted expected number of motifs is listed as MEME's Eval and the percentage
1054 of peaks with the indicated motif is listed (%), and the number of additional motifs with an Eval
1055 less than 5e-2 is indicated (red text).
1056
1057 **Figure 7. Cluster analysis and GSEA GO enrichment results.**
1058 We defined 5,383 hypoxia-induced peak regions as 30-bp windows centered on all diff-MOA
1059 peaks. The R-package k-means clustering of the MOA coverage values with the elbow method
1060 resulted in 10 clusters. (**A**) The hypoxia MOA clusters were named "hmc1" through "hmc10",
1061 plotted for each peak region, along with the total number of peaks and DEGs associated with
1062 each cluster. The peak regions for each cluster were separately used as input for MEME motif
1063 analysis, and the top motifs for each are listed under their respective panels. (**B, C**) The cluster-
1064 specific DEGs and their fold change transcript abundance for each hypoxia treatment hour were
1065 used as input for the gene set enrichment program, gseGO in clusterProfiler.
1066 Resulting enrichments are shown and plotted across the time points for significant pathways, for
1067 suppressed (pink) or activated (blue) genes. Plots list the GO term, the hmc cluster number,
1068 and the GO ID for each pathway for those found to be shared among all three time points (left)
1069 or shared between any two of the three time points (right).
1070
1071 **Figure 8. Analysis of potential TFs and candidate cofactors at hypoxia-induced footprint**
1072 **gain sites containing HIF1A motifs.** The hypoxia-induced gain of footprint peaks was filtered
1073 to retain only those that included a HIF1A motif using the MEME suite SEA tool. (**A**) MOA-seq
1074 average profile plots for coverage centered on these HIF1A motifs at 1h, 3h, and 24h hypoxia.
1075 Each is coplotted with the 0h normoxia control (black line). (**B**) Table of HIF1A motif numbers
1076 found in the diff-MOA peaks at each hypoxia time point. The 10 clusters from Fig. 7 are listed

separately and sorted into two sets, the HIF1A group with a total of 972 HIF1A motifs (HMCs 2,3,4,5,9), and the NON-HIF1A group with a total of 5 HIF1A motifs (HMC 1,6,7,8,10). (**C**) Plot of total HIF1A motif percentages by cluster across the three time points. (**D**) Three lists of enriched TF-enriched cluster-associated motifs: 55 Non-HIF1A-group TFs defined as those found in at least two of the associated clusters; 48 HIF1A-group TFs defined as those found in at least three of the associated clusters; and 34 TFs shared between the two lists as summarized by the Venn diagram. Motifs for TFs that are implicated as possible HIF1A co-activators are marked with an asterisk to indicate co-occurrence enrichment relative to promoter cCRE regions with HIF1A motifs that were used as a control.

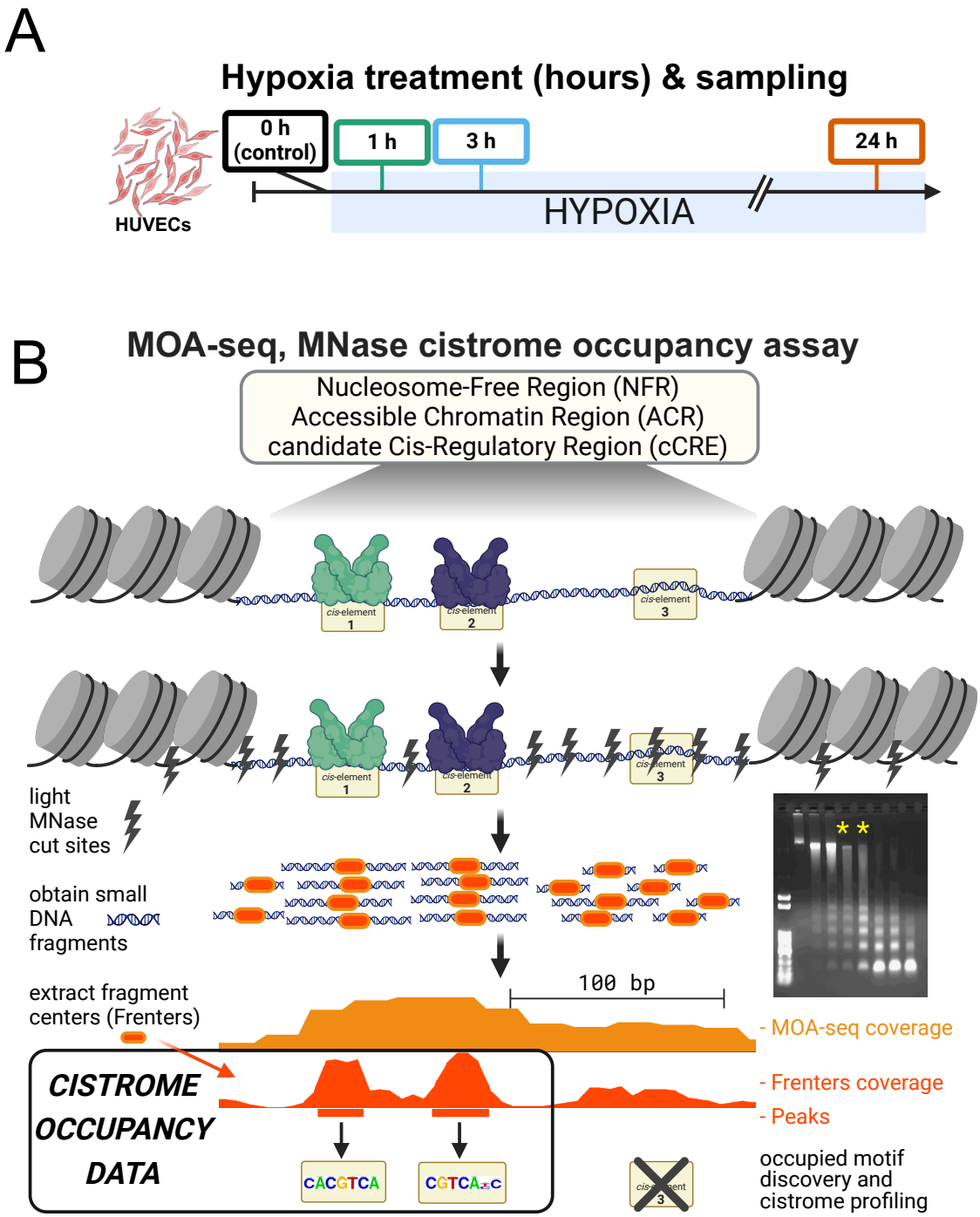


Figure 1. Transcription Factor Occupancy Profiling with MNase "MOA-seq" in HUVECs following hypoxia for 1, 3, and 24h. MOA-seq, MNase-defined cistrome-Occupancy Analysis, (Savadel et al. 2021), was developed to globally identify protein-bound cis-regulatory sequences in a single robust assay using modified nuclease profiling of fixed chromatin and customized bioinformatic analysis pipelines. MOA-seq was applied to HUVEC cells stressed with hypoxia to map short (1h and 3h) or long (24h) exposure-induced changes in TF occupancies. **(A)** The treatment scheme for the four samples used in this study: 0h = normoxia; 1h, 3h, and 24h refer to hypoxia (2% oxygen) exposure times. **(B)** Diagrammatic summary of the MOA-seq, starting with formaldehyde-fixed cells to preserve chromatin structure including bound transcription factors (TFs) at their cognate binding sites (cis-elements). Light/partial digest with MNase preferentially cuts DNA within nucleosome-free/accessible chromatin regions (NRF/ACR), but not where protein occupancy occurs. Recovery of de-crosslinked, small DNA fragments from ideal partial digests (yellow asterisks, inset gel) is used for NGS library production. MOA-seq coverage profiles of aligned read pair fragments (orange) and their fragment centers (frenters, dark orange) reveal footprints for analysis. The frenter peaks reflect protein-protected footprints within the NFR/ACRs (e.g. cis-element 1 and 2) but do not mark unoccupied cis-elements (e.g. cis-element 3). Peak regions are used for comparative genomics and motif discovery and analysis.

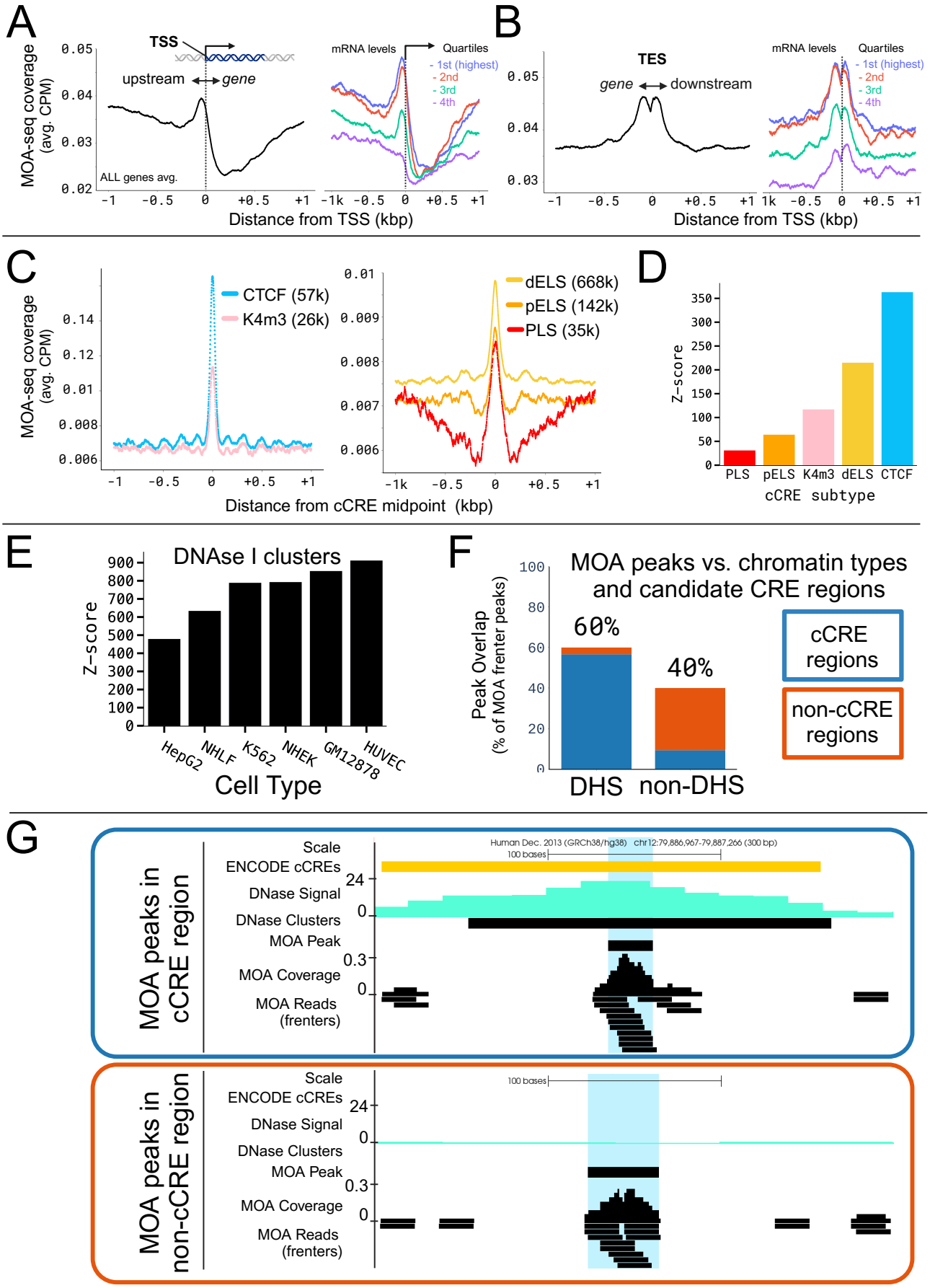

A
TSS
upstream
gene
MOA-seq coverage (avg. CPM)
ALL genes avg.
Distance from TSS (kbp)
mRNA levels
Quartiles
- 1st (highest)
- 2nd
- 3rd
- 4th
B
TES
gene
downstream
Distance from TSS (kbp)
C
CTCF (57k)
K4m3 (26k)
dELS (668k)
pELS (142k)
PLS (35k)
Distance from cCRE midpoint (kbp)
D
Z-score
PLS
pELS
K4m3
dELS
CTCF
cCRE subtype
E
DNAse I clusters
Z-score
HepG2
NHLF
K562
NHEK
GM12878
HUVEC
Cell Type
F
MOA peaks vs. chromatin types and candidate CRE regions
Peak Overlap (% of MOA frenter peaks)
60%
40%
DHS
non-DHS
cCRE regions
non-cCRE regions
G
MOA peaks in cCRE region
Scale
ENCODE cCREs
DNase Signal
DNase Clusters
MOA Peak
MOA Coverage
MOA Reads (frenters)
Human Dec. 2013 (GRCh38/hg38) chr12:79,886,967-79,887,266 (300 bp)
100 bases
MOA peaks in non-cCRE region

Figure 2 Continued.

Figure 2. MOA-seq maps gene regulatory features at ENCODE cCREs and additional sites with HUVEC cell-type specificity in normoxic cells. MOA-seq coverage profiles were compared to ENCODE and other genomic annotations related to gene expression and regulation. **(A,B)** Profile trend plots of MOA coverage showing peaks at the transcription start (TSS) and end (TES) sites of genes, with peak heights positively correlating with transcript abundance quartiles. **(C)** Profile trend plots of MOA frenter coverage profiles colocalize with cCRE subtypes (K4m3, DNase-H3K4me3; dELS, distal enhancer-like signature; pELS, proximal enhancer-like signature; PLS, promoter-like signature). **(D)** Intersection significance is shown by Z-scores calculated from peak bp intersections with cCRE subtypes. **(E)** Intersection significance is shown by Z-scores calculated from peak bp intersections with DNaseI hypersensitive (DHS) peaks from multiple different cell types. **(F)** The abundance of MOA-peaks relative to DHS regions or known cCRE regions. **(G)** Genome browser session showing an example of a MOA peak within a cCRE and DHS region (top) and another example in a non-cCRE and non-DHS region (bottom), with tracks labeled on the left.

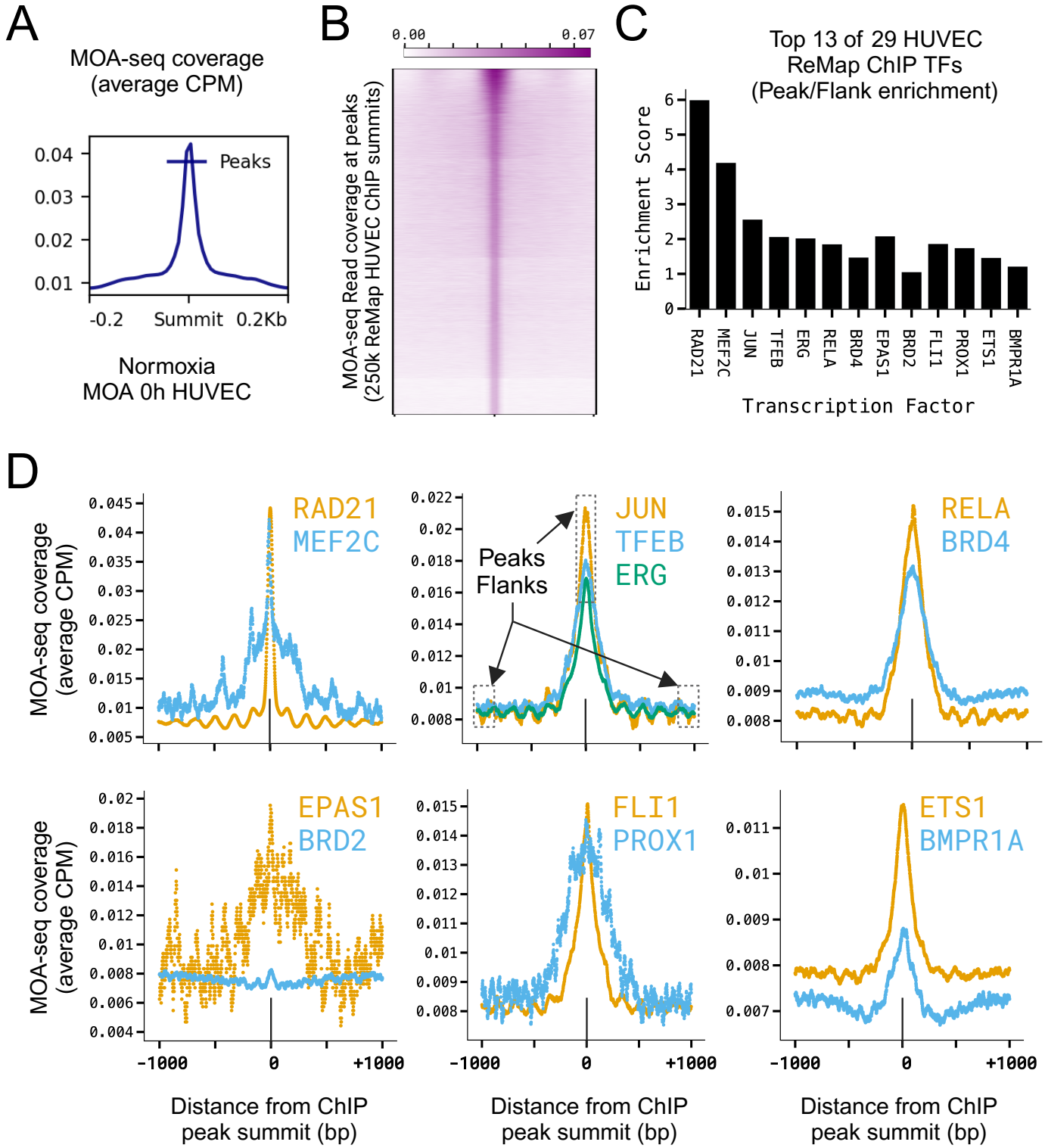
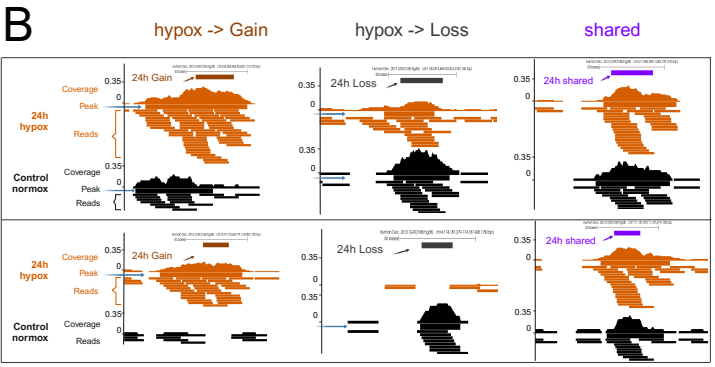


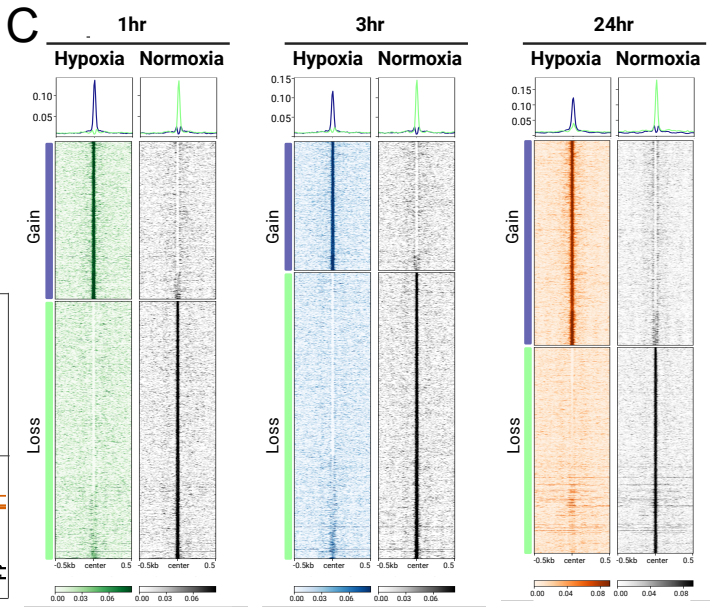
Figure 3. Enrichment of MOA at ReMap ChIP summits for 13 characteristically HUVEC transcription factors. MOA-seq enrichment at previously characterized TF binding sites as defined by ChIP-seq, using ReMAP (Hammal et al., 2021) ChIP summits for 13 canonical HUVEC TFs (Payne et al., 2023). Profile trend **(A)** and heatmap **(B)** plots of MOA 0hr frenters coverage centered on ReMap HUVEC ChIP peak summits. Data plotted were from regions with non-zero MOA coverage of ~250k summits from a total of ~750k spanning various studies and treatments. **(C)** Bar chart showing total unfiltered enrichment scores (MOA cov summit/flanks) for a select subset of TF ChIPs from experiments conducted in HUVEC only, with nearby (< 10 bp) summits for MOA and ChIP. **(D)** Profile trend plots of local MOA frenter coverage for the individual TFs shown in panel C, grouped primarily by signal height.

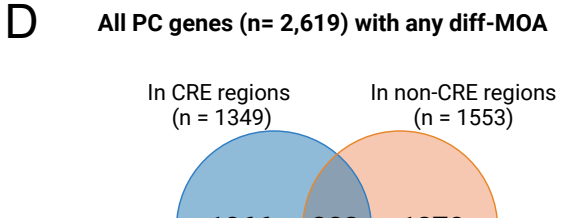
A

Differential MOA (diff-MOA) peaks

Type	GAIN			LOSS			SHARED		
Hypoxia	1h	3h	24h	1h	3h	24h	1h	3h	24h
Number	751	600	1,289	1,227	1,332	1,297	9,056	7,555	10,712
Avg bp	24	24	26	24	24	24	31	30	32
Total kb	18	15	33	29	32	31	283	228	347







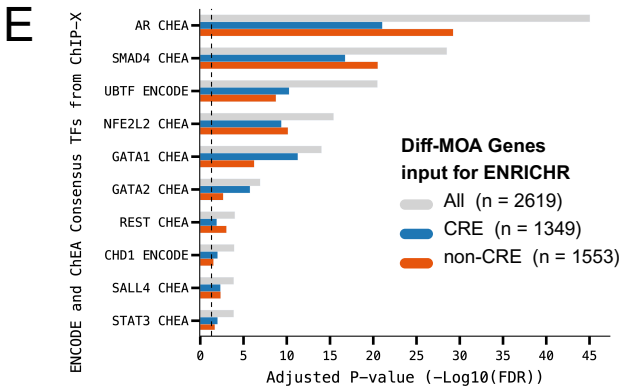


Figure 4. Summary of hypoxia-induced differential MOA (diff-MOA) peaks and genes, along with the enrichment analysis for regulation by specific transcription factors. To identify genomic regions that changed their TF occupancy footprints in response to hypoxia, bdgdiff (MACS3) of MOA frenter coverage was used with the elbow method. **(A)** Table of diff-MOA peak numbers and sizes for diff-MOA peaks after hypoxia relative to 0h. Total diff-MOA peaks (gain + loss) were 1,978, 1,932, and 2,586 after 1h, 3h, and 24h hypoxia, respectively. **(B)** Genome browser session showing two examples each for diff-MOA gain, diff-MOA loss, or shared for 24h hypoxia treatment. **(C)** MOA-seq frenter coverage heatmaps of the diff-MOA (GAIN vs LOSS of footprints relative to normoxia) for each time point. Heatmaps were sorted by total MOA signal averaged over the 1 kb window centered on the peaks. In addition, sorting was done for diff-MOA gain sites using the hypoxic time point as the primary sort, and for diff-MOA loss sites, using the normoxic time point as the primary sort. **(D)** Venn diagram of genes with diff-MOA footprints for cCRE versus non-cCRE regions. Genes in the intersection had multiple diff-MOA footprints, with at least one of each cCRE type. **(E)** Enrichments (adjusted p-values) for gene lists derived from diff-MOA peaks in cCRE versus non-cCRE regions. Results from ENRICHR (https://maayanlab.cloud/Enrichr/) using the libraries for ENCODE and ChEA Consensus TFs from ChIP-X. The plot was sorted by P-values for the total Diff-MOA peaks (grey bars) and all significant TFs (dashed lines) are shown.

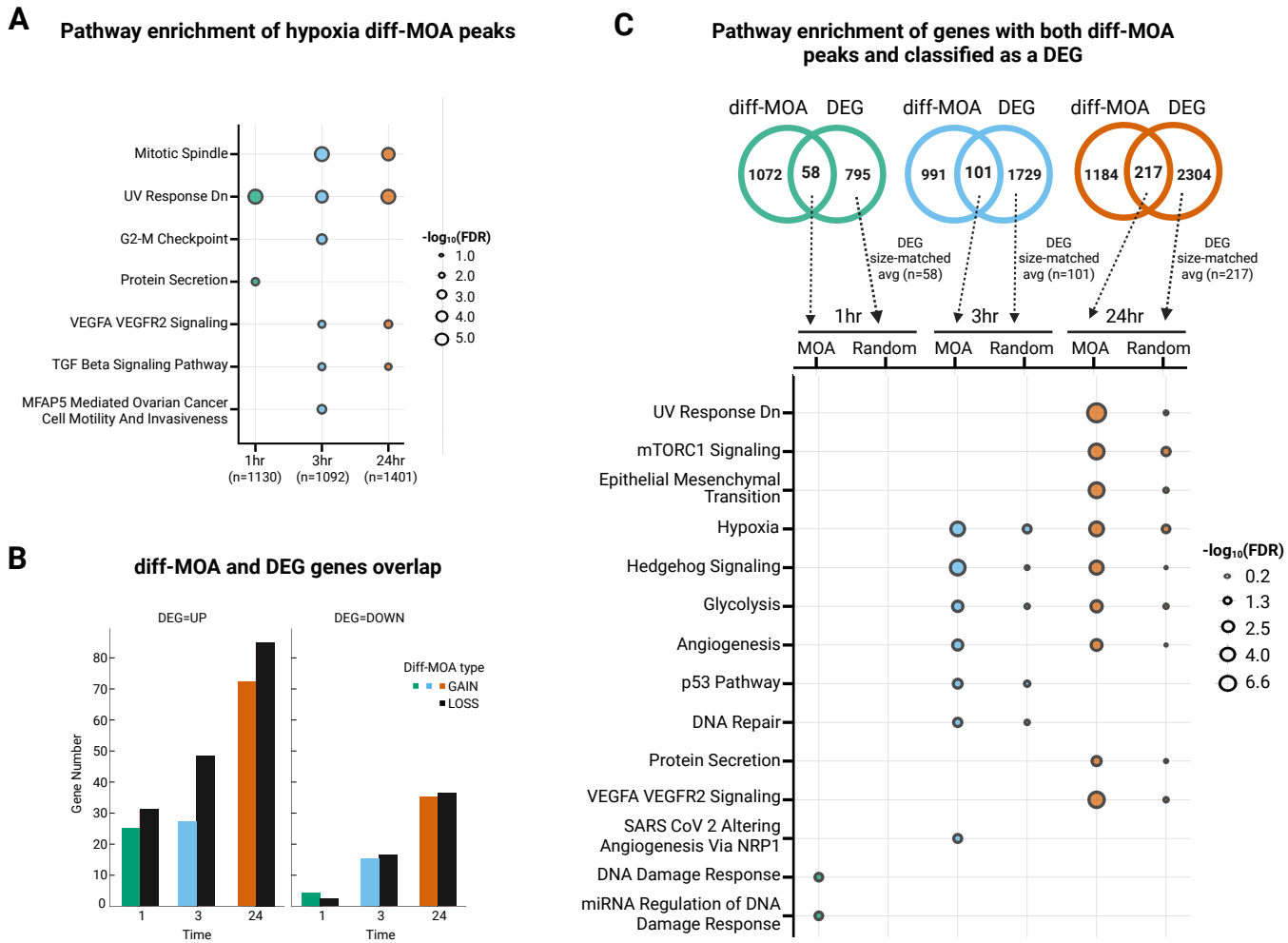


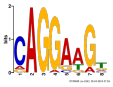
Figure 5. Comparison of differential MOA with differential gene expression. Comparison of hypoxia-induced diff-MOA genes to hypoxia-induced differentially expressed genes (DEGs) with regard to gene enrichment analyses, changes in RNA levels, and gain or loss of diff-MOA footprints. (**A**) Gene lists defined by diff-MOA (gain or loss) for each hypoxia time point were used as input for the ENRICHR libraries, and top hits are shown for two libraries, Human MSigDB Hallmark and WikiPathway Human gene sets. The significance (-log10(FDR)) is proportional to the circle size; gene list sizes are indicated at the bottom. (**B**) Direction relationships of diff-MOA (gain or loss) versus DEG (transcript up or down) is plotted for each time point. (**C**) Analysis of genes showing both diff-MOA and DEGs. The top shows Venn diagrams for each time point. The bottom shows ENRICHR results using the same two libraries as in panel A. To compare the intersected diff-MOA and DEG gene sets, we sampled the larger DEG gene sets using the intersection size-matched gene number. For each such comparison, the average (n=3) of the -log10(FDR) was less significant or insignificant.

A

difMOA GAIN

1h Hypox (n=751)

De Novo (top of 2)



1-CAGGAAGT

TF *Families*

CTCF (**CTCFL**)
FOS (**FOSB**, **FOSL1/2**)
JUN (**JUNB**, **JUND**)

TFs (top 10)

TF	Eval	%
AP2A	3.9e-3	12%
BCL6	5.1e-3	4%
BHA15	5.7e-3	16%
ETV4	7.8e-4	9%
ITF2	1.1e-5	16%
MAFB	1.7e-5	54%
NDF2	8.1e-3	22%
SMAD3	3.6e-6	55%
STAT6	6.5e-3	26%
TBX3	3.7e-5	42%

19 more < 5e-2

3h Hypox (n=600)

De Novo (top of 1)

1-CAGGAAR

TF *Families*

(no CTCF, FOS, JUN)

TFs (top 10)

TF	Eval	%
ATOH1	2.4e-4	12%
COE1	8.1e-4	10%
ELF3	4.4e-3	18%
ETV4	1.9e-6	19%
ETV5	6.5e-4	37%
MAFB	1.3e-6	54%
SALL4	7.7e-4	22%
SMAD2	1.9e-3	28%
SMAD3	4.5e-4	67%
ZIC3	1.9e-3	14%

12 more < 5e-2

24h Hypox (n=1289)

De Novo (top of 2)

1-YCTSCTG

TF *Families*

CTCF (**CTCFL**)
FOS (**FOSB**,**FOSL1/2**)
JUN (**JUNB**,**JUND**)

TFs (top 10)

TF	Eval	%
ATOH1	8.0e-14	21%
BACH2	8.7e-13	16%
ETV4	3.7e-12	21%
HTF4	3.9e-13	16%
ITF2	8.1e-15	13%
LYL1	9.8e-13	29%
NDF2	1.6e-16	17%
NF2L2	2.2e-15	10%
NFE2	2.0e-12	9%
SMAD3	1.0e-14	25%

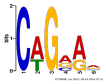
82 more < 5e-2

B

difMOA LOSS

1h Hypox (n=1227)

De Novo (top of 1)



1-CWGRAR

TF *Families*

(no CTCF, FOS, JUN)

TFs (top 10)

TF	Eval	%
BACH2	9.6e-4	8%
ETV4	6.1e-3	10%
ITF2	8.2e-3	38%
MAFB	1.6e-5	27%
MYOD1	2.7e-2	6%
NFAC4	2.6e-3	11%
PTF1A	2.2e-2	7%
SMAD3	1.8e-4	29%
SOX9	4.6e-3	11%
ZBT18	1.1e-2	8%

3 more < 5e-2

3h Hypox (n=1332)

De Novo (top of 2)

1-CTGAAAAAN

TF *Families*

CTCF (**CTCFL**)

TFs (top 10)

TF	Eval	%
ASCL1	2.6e-4	13%
ATOH1	1.8e-6	10%
HTF4	9.5e-5	38%
ISL1	3.0e-5	44%
NDF1	4.9e-8	12%
NDF2	3.5e-7	21%
OLIG2	2.6e-4	21%
TF7L1	7.1e-4	19%
TGIF1	7.5e-4	13%
ZBT18	1.9e-5	19%

22 more < 5e-2

24h Hypox (n=1297)

De Novo (top of 1)

1-RVAGRAAAWR

TF *Families*

(no CTCF, FOS, JUN)

TFs (top 10)

TF	Eval	%
ELK4	2.9e-3	11%
MAFB	6.5e-3	42%
PO2F1	3.8e-9	3%
SMAD3	3.5e-3	19%
SMCA1	2.0e-3	27%
TEAD1	1.5e-3	16%
TF7L1	4.0e-3	43%
ZN502	2.1e-6	2%
ZN582	4.0e-3	2%
ZNF85	3.0e-3	3%

8 more < 5e-2

Figure 6 Continued.

Figure 6. Motif discovery and analysis at diff-MOA peaks identify candidate hypoxia-responsive gene regulatory network cis-elements. The diff-MOA peaks were split into six sets based on gain (**A**) or loss (**B**) of footprint at each of the three hypoxia treatment timepoints. The genomic sequences under the peaks were used as input for MEME motif analysis. For each list of motifs, three results are shown; the top *de novo* motif and its sequence LOGO, the TF families (italics) and their members (bold in parentheses), and the top 10 individual TFs sorted alphabetically. For this last list each motif shows the background-adjusted expected number of motifs is listed as MEME's Eval and the percentage of peaks with the indicated motif is listed (%), and the number of additional motifs with an Eval less than 5e-2 is indicated (red text).

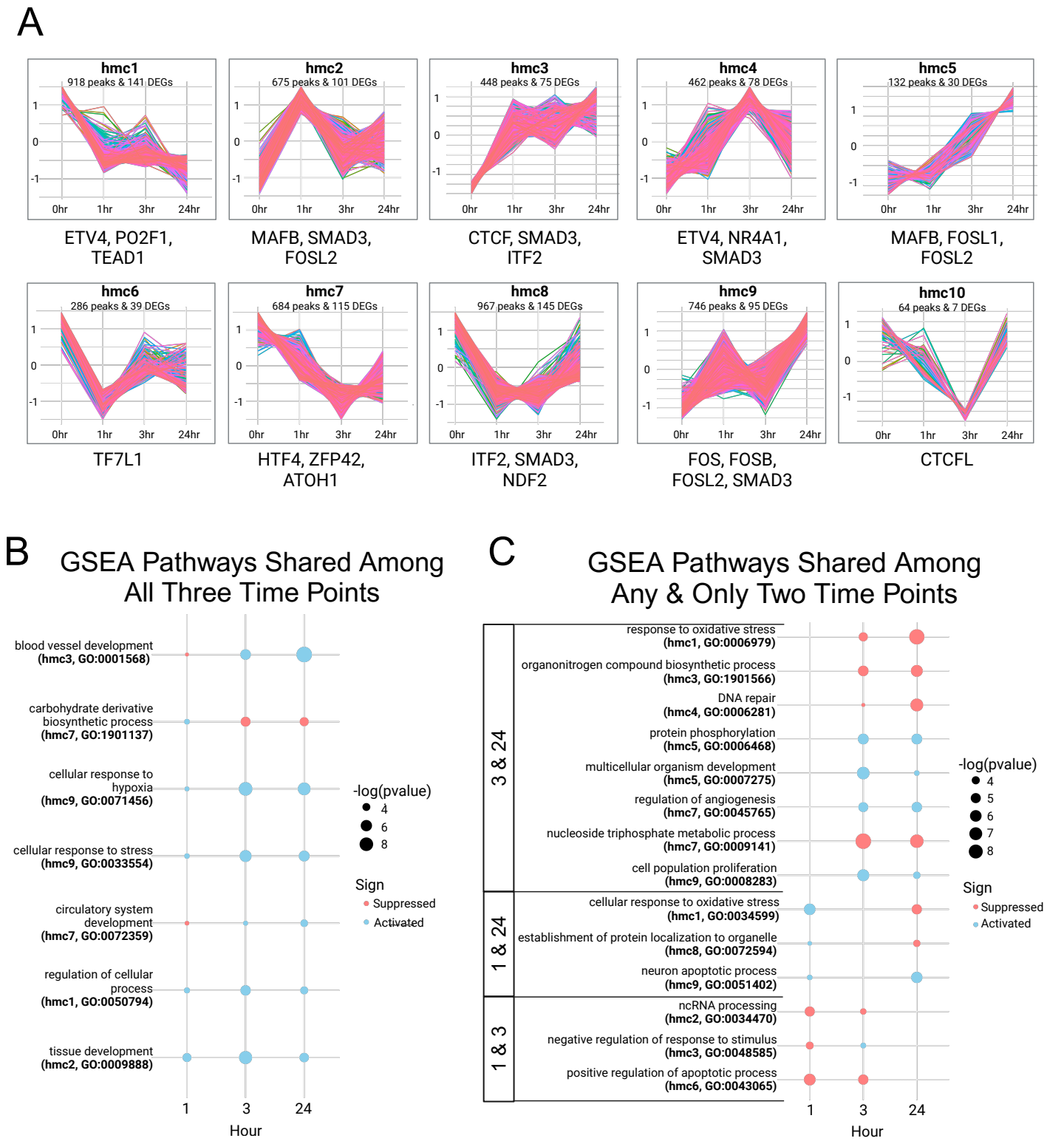
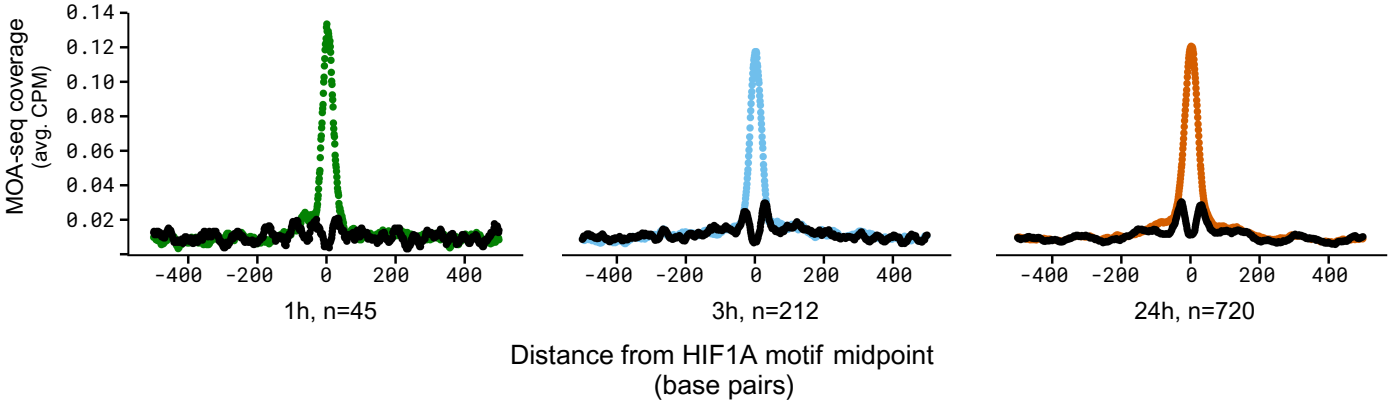


Figure 7. Cluster analysis and GSEA GO enrichment results.
We defined 5,383 hypoxia-induced peak regions as 30-bp windows centered on all diff-MOA peaks. The R-package k-means clustering of the MOA coverage values with the elbow method resulted in 10 clusters. (**A**) The hypoxia MOA clusters were named "hmc1" through "hmc10", plotted for each peak region, along with the total number of peaks and DEGs associated with each cluster. The peak regions for each cluster were separately used as input for MEME motif analysis, and the top motifs for each are listed under their respective panels. (**B, C**) The cluster-specific DEGs and their fold change transcript abundance for each hypoxia treatment hour were used as input for the gene set enrichment program, gseGO in clusterProfiler. Resulting enrichments are shown and plotted across the time points for significant pathways, for suppressed (pink) or activated (blue) genes. Plots list the GO term, the hmc cluster number, and the GO ID for each pathway for those found to be shared among all three time points (left) or shared between any two of the three time points (right).

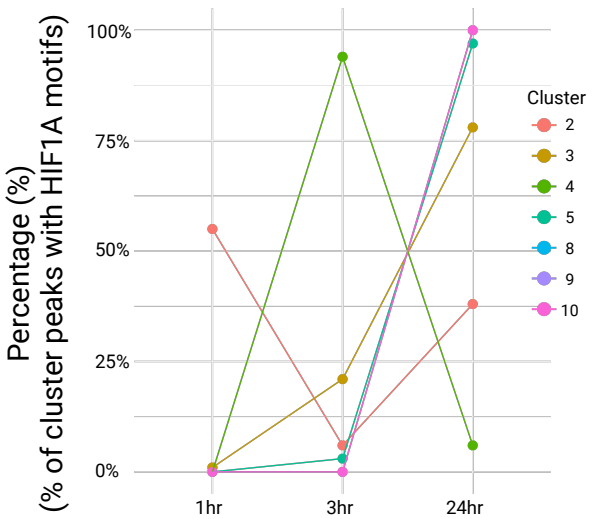
A



B

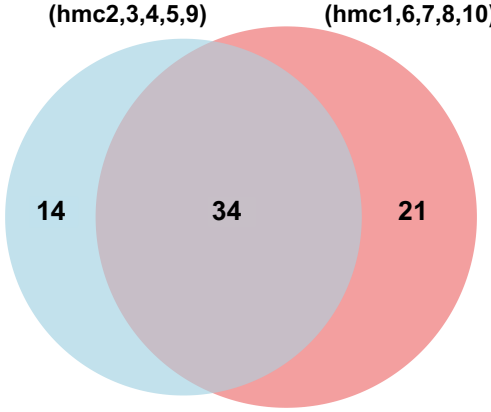
		HIF1A motifs			
Group	Cluster (peak no.)	1h	3h	24h	Total
HIF1A (n=972)	2 (675)	43	5	30	78
	3 (448)	2	57	210	269
	4 (462)	0	147	9	156
	5 (132)	0	2	76	78
	9 (746)	0	1	390	391
NON-HIF1A (n=5)	1 (918)	0	0	0	0
	6 (286)	0	0	0	0
	7 (684)	0	0	0	0
	8 (967)	0	0	3	3
	10 (64)	0	0	2	2
	Hour Total	45	212	720	

C



D

** = co-occurrence with HIF1A motifs in immediate region*



*EHF FOXO1
*ELK1 *JUND
*ERG *NFE2L2
*ETV1 NFATC2
*ETV2 *NFE2
FEZF1 SMAD2
*FOSL2 TCF3

MYF6 REL *ZFP42
MYOD1 RORA *ZIC3
NKX21 SALL4 ZN257
NR1D1 SMAD4 ZN350
OLIG2 TF7L1 ZN418
OSR2 TF7L2 ZN667
PRDM1 ZBT18 ZNF85

ATOH1 *BCL11A BCL6 *BHLHA15 *CTCF *CTCFL *ELK4 *ETS2 *ETV4 *ETV5
*FEV *FOS *FOSB *FOSL1 HTF4 ITF2 *JUN LEF1 LYL1 *MAF *MAFB NEUROD2
NFATC1 *NR4A1 SMAD3 SMARCA1 STAT4 STAT6 TBX3 TEAD1 TEAD4 TGIF1 ZEB1 ZN563

Figure 8 Continued.

Figure 8. Analysis of potential TFs and candidate cofactors at hypoxia-induced footprint gain sites containing HIF1A motifs. The hypoxia-induced gain of footprint peaks was filtered to retain only those that included a HIF1A motif using the MEME suite SEA tool. (**A**) MOA-seq average profile plots for coverage centered on these HIF1A motifs at 1h, 3h, and 24h hypoxia. Each is coplotted with the 0h normoxia control (black line). (**B**) Table of HIF1A motif numbers found in the diff-MOA peaks at each hypoxia time point. The 10 clusters from Fig. 7 are listed separately and sorted into two sets, the HIF1A group with a total of 972 HIF1A motifs (HMCs 2,3,4,5,9), and the NON-HIF1A group with a total of 5 HIF1A motifs (HMC 1,6,7,8,10). (**C**) Plot of total HIF1A motif percentages by cluster across the three time points. (**D**) Three lists of enriched TF-enriched cluster-associated motifs: 55 Non-HIF1A-group TFs defined as those found in at least two of the associated clusters; 48 HIF1A-group TFs defined as those found in at least three of the associated clusters; and 34 TFs shared between the two lists as summarized by the Venn diagram. Motifs for TFs that are implicated as possible HIF1A co-activators are marked with an asterisk to indicate co-occurrence enrichment relative to promoter cCRE regions with HIF1A motifs that were used as a control.