Supplemental Methods S1. Materials and Methods

Cell Culture

Cells were cultured in standard conditions at 37˚C with 5% CO_2 as follows: UM-RC-2 Dulbecco's Modified Eagle Medium (DMEM) with low glucose (Sigma-Aldrich) with 10% fetal bovine serum (FBS); 786-0 Roswell Park Memorial Institute (RPMI) 1640 with 10% FBS; MCF-7 Eagle's Minimal Essential Medium (EMEM) with 10% FBS, 1% non-essential amino acids, 10 µg/mL insulin; ZR-75-1 RPMI 1640 with 10% FBS; SUM149 Human Mammary Epithelial Cell (HuMEC) media with 1% penicillin-streptomycin, 2% FBS; SUM229 Ham's F12 media with 5% FBS, 5 µg/mL insulin, 1 µg/mL hydrocortisone, 10 mM HEPES, pH 7.5, 1% penicillin-streptomycin; MDA-MB-231 DMEM with 10% FBS; A673, EWS894, EWS502, and SK-NM-C RPMI 1640 with 15% FBS, 1% L-Glutamine; HUVEC (Lonza) EGM-Plus Endothelial Cell Growth Media (Lonza CC-5035) plus supplements (Lonza CC-4542).

FAIRE assay for fixed cells

Five million cells (EWS894, UM-RC-2, MCF-7, HUVEC) were crosslinked in 1% formaldehyde in biological triplicate and nuclei were prepared using the Active Motif Chromatin Preparation kit (53046) according to manufacturer's instructions. 0.25 million formaldehyde crosslinked *Drosophila* S2 cells were mixed with the human tumor cells just prior to nuclei preparation. Nuclei were resuspended at a concentration of 1 million nuclei per 90 µL of Lysis Buffer A (Simon et al. 2012) and 10 µL of nanodroplets (Triangle Biotechnology, Inc., FF101-5000 and (Kasoji et al. 2015)) was added to each sonication tube (Triangle Biotechnology, Inc., PL101-BT100) prior to sonication. Sonication was performed in a Covaris E110 instrument using conditions described in (Chiarella et al. 2018) with nanodroplets (EWS894, 5 seconds; UM-RC-2, 30 seconds; MCF-7, 1.5 minutes; HUVEC, 2 minutes) or without nanodroplets (EWS894 and MCF-7, 16 minutes; UM-RC-2, 10 minutes, HUVEC sonicated with a probe sonicator as described in (Pattenden et al. 2016)). Following sonication, insoluble debris was removed by centrifugation and samples were digested with 200 µg RNase (Qiagen) at 37˚C for 30 minutes. Ten µL of sample was removed for input then 90 µL Lysis Buffer A (Simon et al. 2012) was added to bring total volume to 100 µL. Inputs were digested with 40 mg Proteinase K (Worthington) for 1 hour at 55˚C, followed by 2 hours at 80˚C to reverse the crosslinks, then purified using a silica matrix column (Zymo Research, ChIP DNA Clean and Concentrate kit, D5201). FAIRE was performed on the remaining 90 µL of sonicated chromatin by

purification on a silica matrix column as described in (Pattenden et al. 2016). DNA was quantified by fluorometry (Qubit dsDNA High Sensitivity Assay Kit, Invitrogen). Average peak fragment size of input and FAIRE DNA was confirmed using Agilent TapeStation.

FAIRE-qPCR

FAIRE and input DNA samples were diluted to between 4 ng/μL and 0.03 ng/μL. Two μL of each DNA sample was transferred to a 384-well optically clear qPCR plate in duplicate. Eight μL of Sso Advanced Universal SYBR Green Supermix with ROX (Bio-Rad) was added per well and qPCR was performed on the ViiA 7 Real-Time PCR system. Percent FAIRE signal compared to input signal was determined using the ΔCt method (Livak and Schmittgen 2001). Primers used in this study are listed in Supplemental Table 2.

FAIRE-seq and ATAC-seq alignment

We aligned the reads post filtering to the hg19 genome using Star v.2.5.2b (--outFilterScoreMin 1 --outFilterMultimapNmax 1 --alignMatesGapMax 1000 --outFilterMismatchNmax 2 --chimJunctionOverhangMin 15 --chimSegmentMin 15) (Dobin et al. 2013). (Reads from *Drosophila* S2 cells were spiked in for normalization, but not used for analysis). Reads that mapped to more than one genomic location were removed (SAMtools view -f 256), as well as any genes that aligned to blacklisted areas of the genome. BAM files were then converted to BED files (BEDtools v.2.26 BAMtoBED) and the reads shift extended to 150 bp in length. These BEDfiles were then converted to BEDgraphs using genomecov, scaling by the total number of reads aligned, and converted to bigWigs using BEDGraphTobigWig. Murine reads were processed and filtered by a parallel pipeline, aligning to the mm9 construct. BigWig files produced post filtering and alignment of FAIRE-seq sequencing reads were uploaded (UCSC Genome Browser) for visualization.

Peak calling and overlap

Peaks were called using MACS2 callpeak (MACS2 v.2.1.2); significant peaks were defined by default parameters of adjusted q-value <0.05 (Zhang et al. 2008). To identify overlapping peaks between FAIRE data sets we called peaks on the pooled replicate BED files and calculated the percent of overlapping peaks using BEDtools v.2.29 intersect. When comparing older block peaks to younger block peaks, we randomly down sampled the aligned reads from the older blocks to match that of the younger blocks (Older blocks had greater sequencing depth

associated with our experimental design). Our EWS894 FAIRE data was comprised of two older blocks (with three technical replicates per block) and three younger blocks. We defined our older blocks peak set as those peaks present in either of the 2 older blocks (peaks had to be present across all 3 technical replicates). We defined our younger peak set as peaks present in at least 2 of the 3 blocks.

Defining FFPE FAIRE tumor-type specific regions

Using the peaks called within each replicate of each cell line FAIRE sample from FFPE tissue xenografts, cell line specific peaks were defined as the shared peak intervals across all replicates within each cell line using BEDtools v.2.29 intersect. Cell line specific peaks were then compiled into one union set of peaks and categorized them by their tumor-type classification (ccRCC, Ewing sarcoma, breast ER-positive, breast ER-negative). Outlier replicates were excluded from this analysis as defined by the PCA and Spearman's correlation coefficients. Peaks of significantly differential signal were found by performing each tumor-type DESeq2 comparison (ccRCC vs not, Ewing sarcoma vs. not, etc.). The TSRs for each tumor-type were categorized as the upregulated regions with an adjusted p-value less than 0.05 and a $\log_2$ fold change greater than 0.5.

Principle component analyses

Genome wide signal of normalized FAIRE-seq signal per sample in bigWig format was condensed into the mean signal per 300 bp bin, with 100 bp shifts using the multibigWigSummary bins tool (deepTools v.3.2.0 (Ramirez et al. 2016)). Mean signal in a set of peaks was calculated per sample using the multibigWigSummary BED-file tool, again using the FAIRE samples in bigWig format as input. We conducted principal component analyses on various combinations of our data sets using the plotPCA function of deepTools, using the $\log_2$ of the entire matrix as output from multiBigWigSummary bins (--ntop 0, --$\log_2$ --transpose).

Heatmaps of FAIRE signal

Normalized FAIRE signal in each base pair around the center of peaks in peak sets were enumerated using the computeMatrix reference-point tool of the deepTools v.3.2.0 package (-bs 1) from FAIRE bigWigs as input.

Heatmap of this signal around peak sets was depicted using the plotHeatmap tool, sorting samples as needed by particular FAIRE sample (--sortUsingSamples).

Correlation analyses

Spearman's correlation coefficients between normalized FAIRE signal from bigWigs was computed among FAIRE samples and displayed as a heatmap using the plotCorrelation tool (deepTools v.3.2.0). Any regions with zero signal across all FAIRE samples per analysis were skipped (--skipZeros –plotNumbers -c Spearman's -p heatmap). Distances between the samples were calculated by complete linkage with clusters determined by the nearest point algorithm.

Lineplot of average signal around TSSs

Mean normalized signal around TSSs was computed by enumerating the normalized signal per base, 3 Kb around the center of each TSS from FAIRE sample bigWig files using the computeMatrix reference-point tool (deepTools v.3.2.0). BEDfile of TSS locations was downloaded in RefSeq format using the UCSC Genome Browser Table Browser on 07/11/2018. This signal was then plotted using the plotProfile tool (--averageType mean).

Genomic region enrichment analysis

Percentages of peaks in peak sets at genomic features were determined using the Cis-regulatory Element Annotation System (CEAS) for hg19 alignments. Percent of promoter regions <=1000 bp, 1000-2000 bp, and 2000-3000 bp were combined to represent entire promoter region.

Motif analysis

Enrichment of known transcription factor motifs was identified using the findMotifsGenome.pl of the HOMER v.4.10 package (hg19, -size given, -mask, default genome background) (Heinz et al. 2010). For comparison of specific transcription factors, the presence of relevant motifs was searched within the known motif output. Failure to detect a motif resulted in a p-value E of 0.

Deconvolution of murine stromal cell chromatin accessibility signal from FAIRE xenograft samples

To identify murine reads, FASTQ files were aligned to the mm9 genome (add Star v.2.5.2b). After filtering and processing as previously described, regions of enrichment were identified by MACS2 callpeak (MACS2 v.2.1.2) with significance defined as default parameters with adjusted q-value <0.05. To deconvolute murine chromatin signal into subpopulations, scATAC-seq from endothelial, dendritic, macrophages, B lymphocytes, and hepatocytes was downloaded and the peak matrix of indicating the presence of a specific peak identified in cells across the 5 tissue types (Mouse sci-ATAC-seq Atlas (Cusanovich et al. 2018)). Cell-type specific peaks were those present in at least 75% of cells of a specific type. Any peaks that were shared across cell types were filtered out, as well as any peaks at transcript start sites (as defined by RefSeq hg19 construct). The mean normalized FAIRE signal in each of xenograft replicate was calculated for each of the three xenografted cell lines at the scATAC cell-type specific peaks, 2 Kb around the center using the computeMatrix reference-point tool of the deepTools v.3.2.0 package (--referencePoint center -a 1000 -b 1000 -bs 1). The mean signal per replicate for each sample was plotted using the plotProfile tool of this same package (--yMin 0 --yMax .3 –perGroup).

References

Chiarella AM, Quimby AL, Mehrab-Mohseni M, Velasco B, Kasoji SK, Davis IJ, Dayton PA, Hathaway NA, Pattenden SG. 2018. Cavitation Enhancement Increases the Efficiency and Consistency of Chromatin Fragmentation from Fixed Cells for Downstream Quantitative Applications. *Biochemistry* **57**: 2756-2761.

Cusanovich DA, Hill AJ, Aghamirzaie D, Daza RM, Pliner HA, Berletch JB, Filippova GN, Huang X, Christiansen L, DeWitt WS et al. 2018. A Single-Cell Atlas of In Vivo Mammalian Chromatin Accessibility. *Cell* **174**: 1309-1324 e1318.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15-21.

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

Kasoji SK, Pattenden SG, Malc EP, Jayakody CN, Tsuruta JK, Mieczkowski PA, Janzen WP, Dayton PA. 2015. Cavitation Enhancing Nanodroplets Mediate Efficient DNA Fragmentation in a Bench Top Ultrasonic Water Bath. *PLoS One* **10**: e0133014.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**: 402-408.

Pattenden SG, Simon JM, Wali A, Jayakody CN, Troutman J, McFadden AW, Wooten J, Wood CC, Frye SV, Janzen WP et al. 2016. High-throughput small molecule screen identifies inhibitors of aberrant chromatin accessibility. *Proc Natl Acad Sci U S A* **113**: 3018-3023.

Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F, Manke T. 2016. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**: W160-165.

Simon JM, Giresi PG, Davis IJ, Lieb JD. 2012. Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA. *Nat Protoc* **7**: 256-267.

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