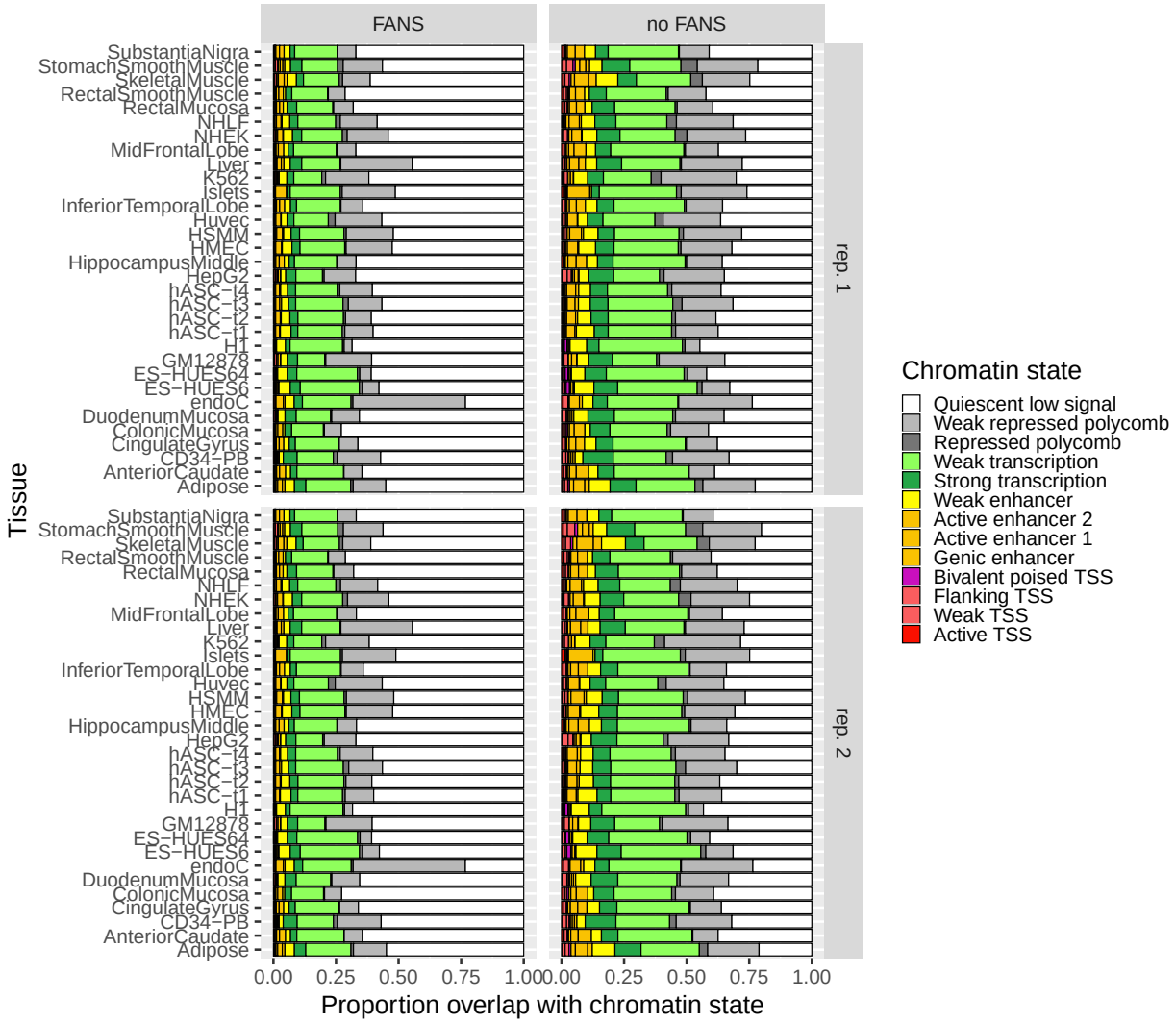

Figure S1. Chromatin state overlap for TSS-distal (> 5kb from TSS) ATAC-seq peaks from the FANS and non-FANS snATAC-seq libraries.

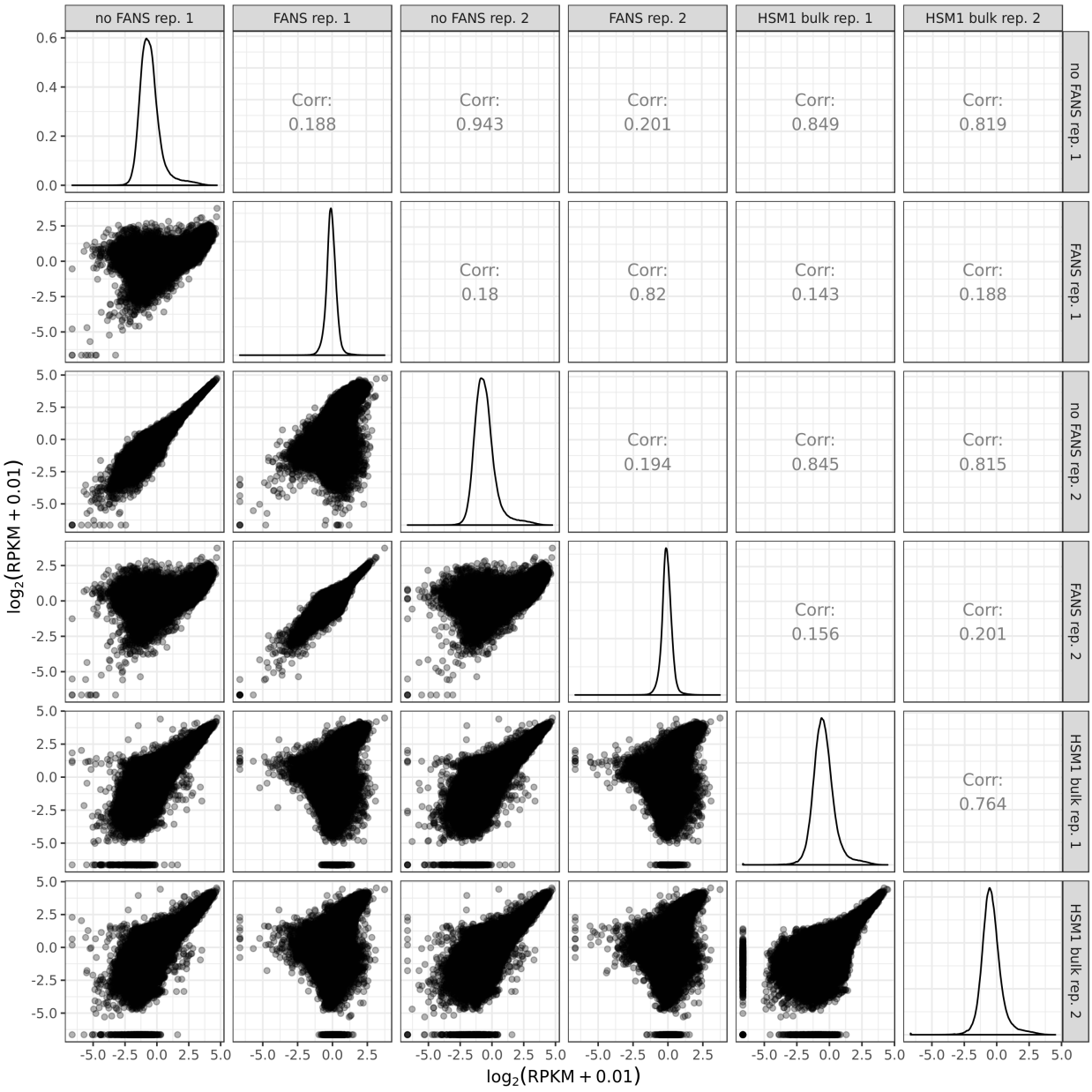

Figure S2. Correlation between FANS snATAC-seq, non-FANS snATAC-seq, and standard bulk ATAC-seq libraries. Each point represents one peak.

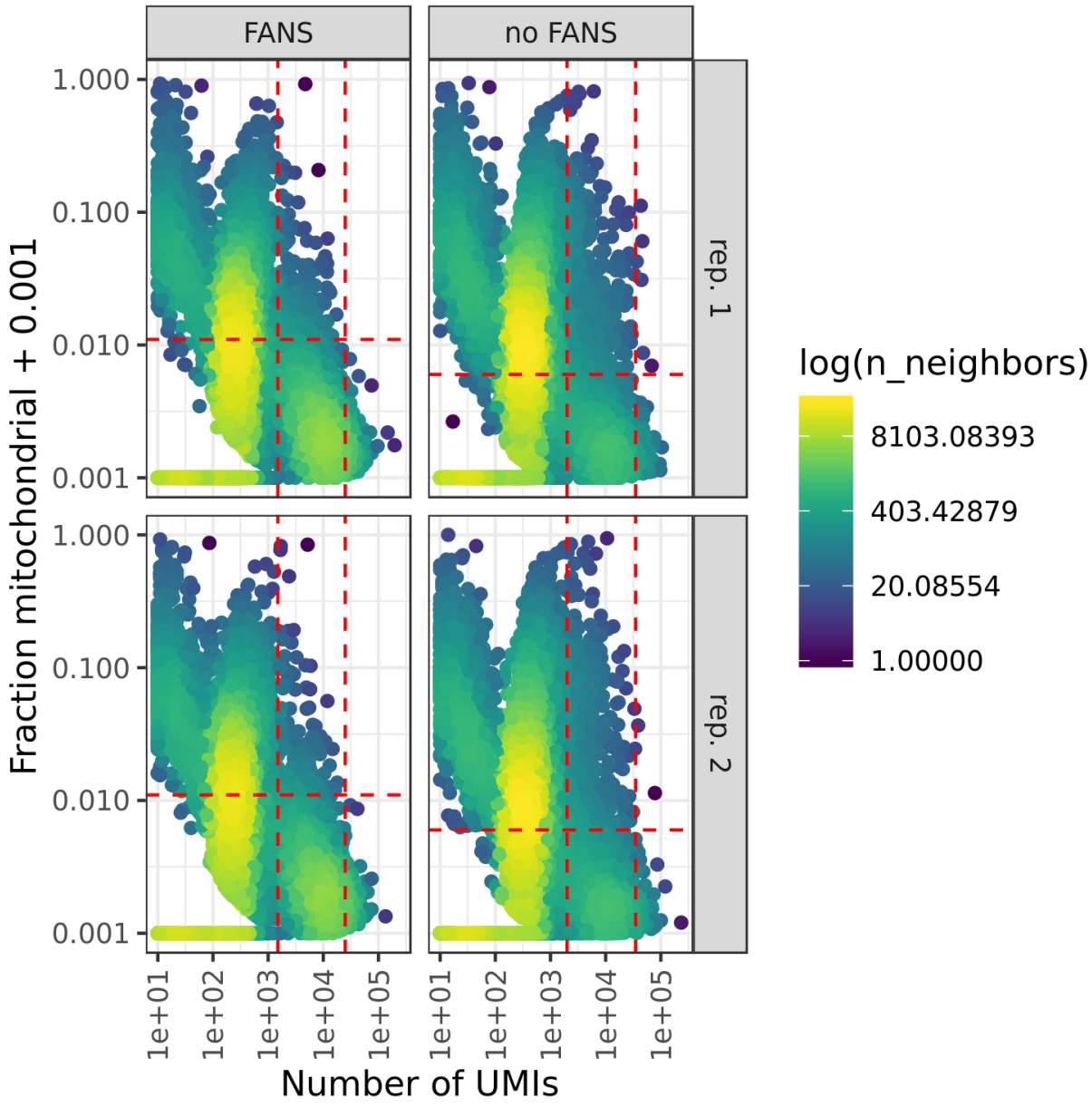

Figure S3. QC thresholds for FANS and non-FANS snRNA-seq libraries. Dashed lines represent thresholds for minimum number of UMIs, maximum number of UMIs, and maximum fraction of mitochondrial UMIs.

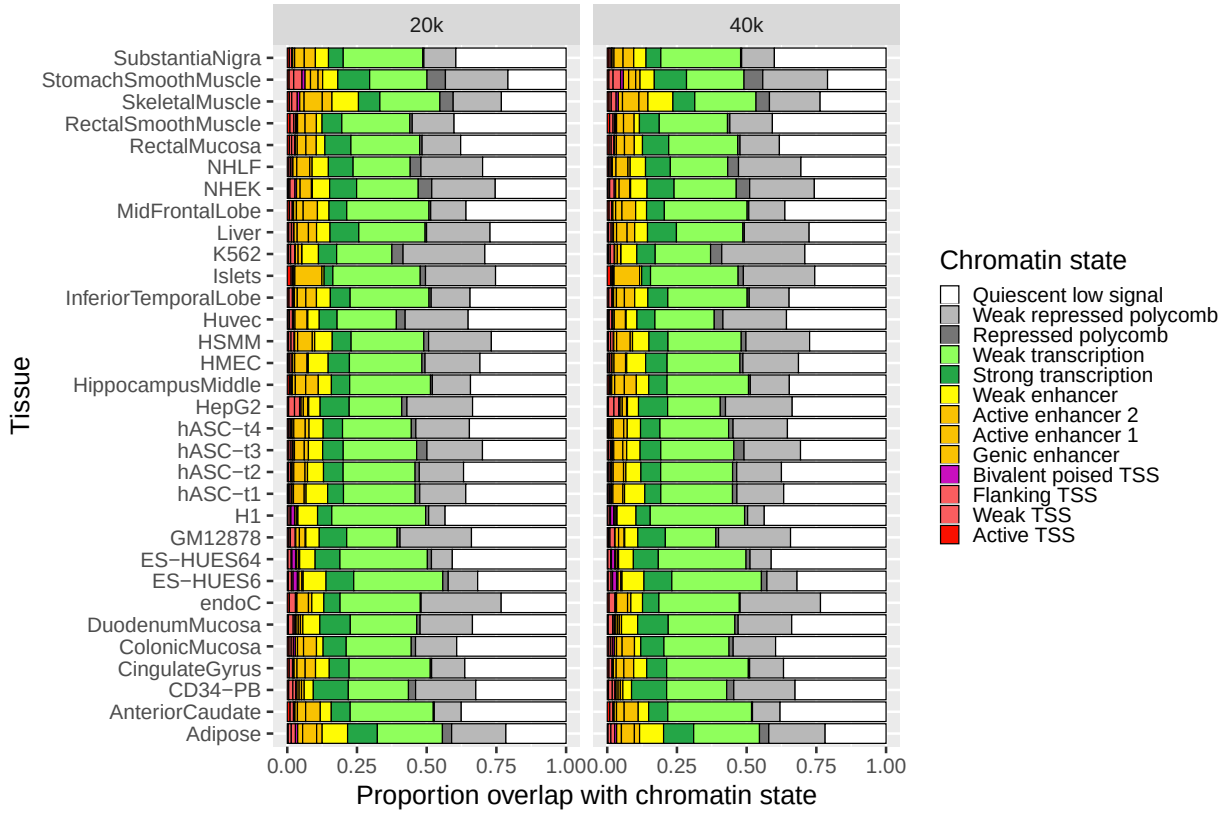

Figure S4. Chromatin state overlap for TSS-distal (>5 kb from TSS) ATAC-seq peaks from the 20k and 40k nucleus FANS snATAC-seq libraries.

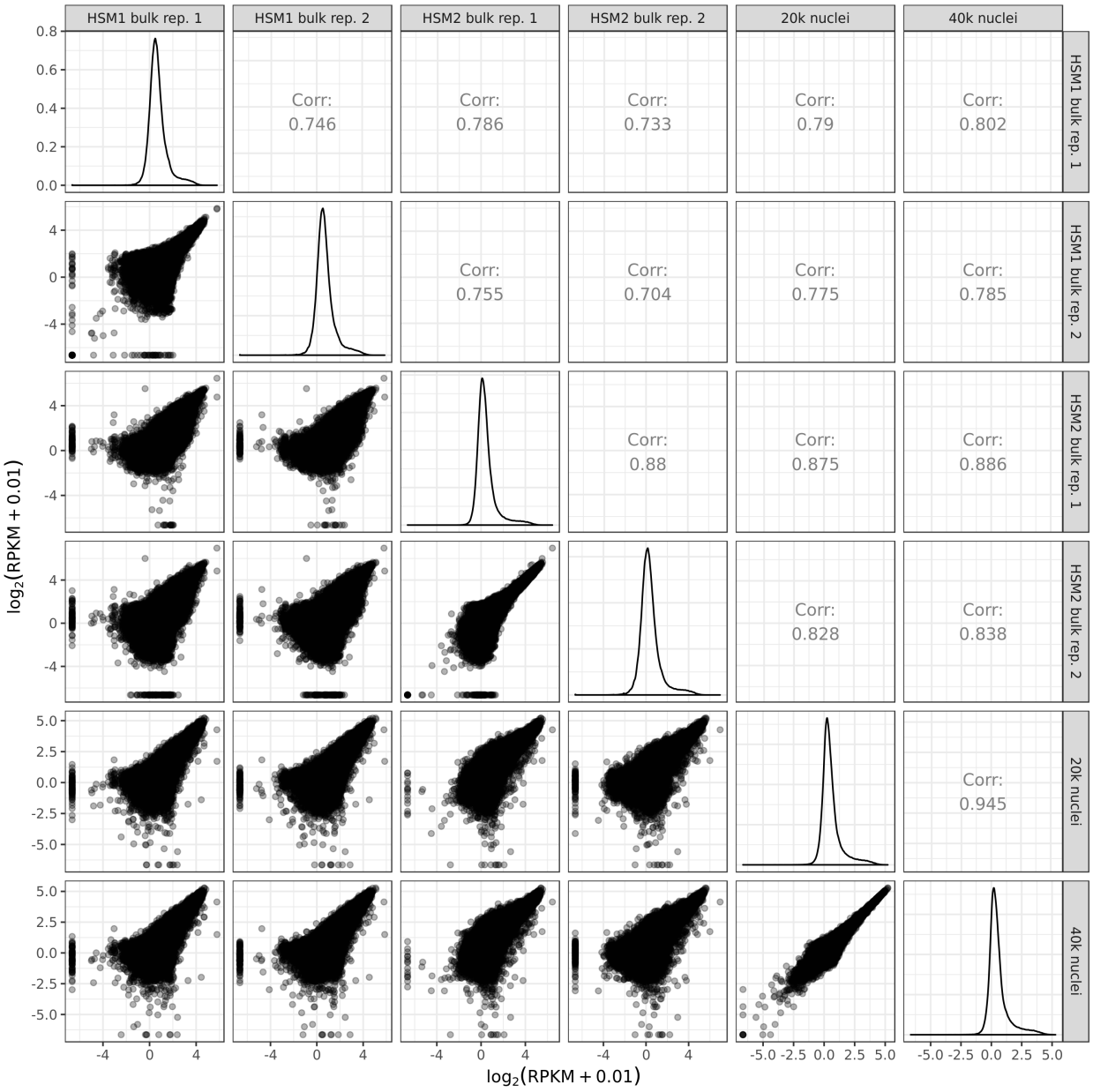

Figure S5. Correlation between 20k and 40k nucleus snATAC-seq libraries and standard bulk ATAC-seq libraries. Each point represents one peak.

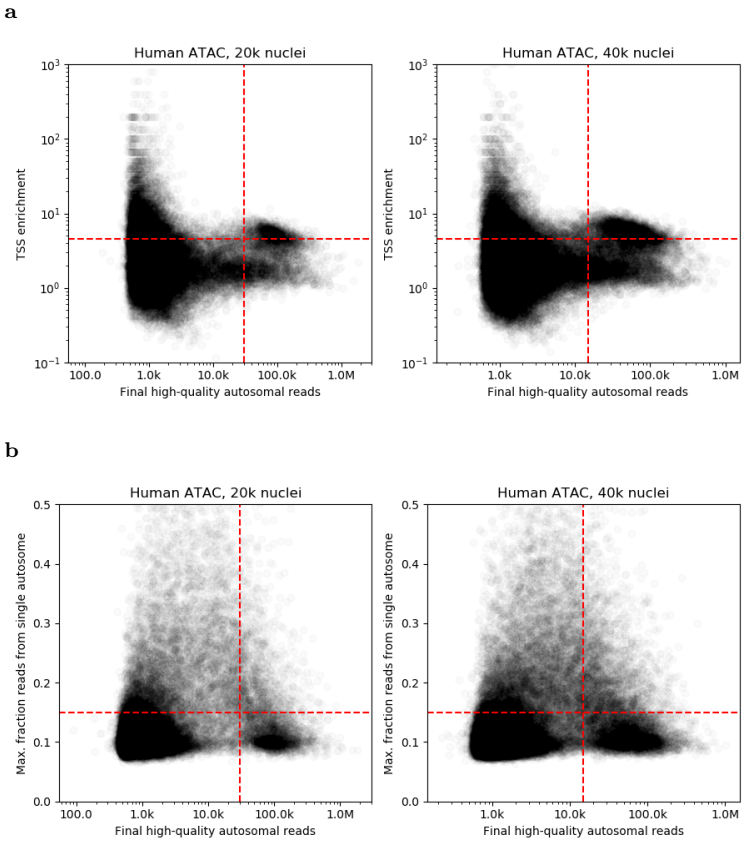

Figure S6. QC thresholding for the 20k and 40k nuclei input snATAC-seq libraries. (a) Dashed lines represent thresholds for minimum number of reads, maximum number of reads, and minimum TSS enrichment. (b) Dashed lines represent thresholds for minimum number of reads, maximum number of reads, and the maximum fraction of reads derived from a single autosome (imposed to filter out nuclei showing abberant per-chromosome coverage).

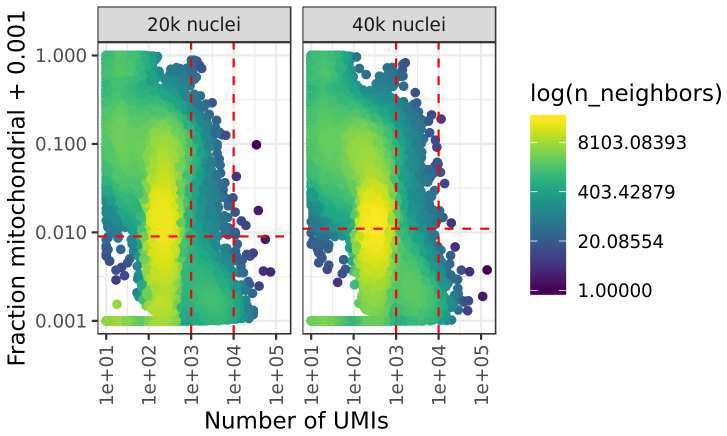

Figure S7. QC thresholds for the 20k and 40k nuclei input snRNA-seq libraries. Dashed lines represent thresholds for minimum number of UMIs, maximum number of UMIs, and maximum fraction of mitochondrial UMIs.

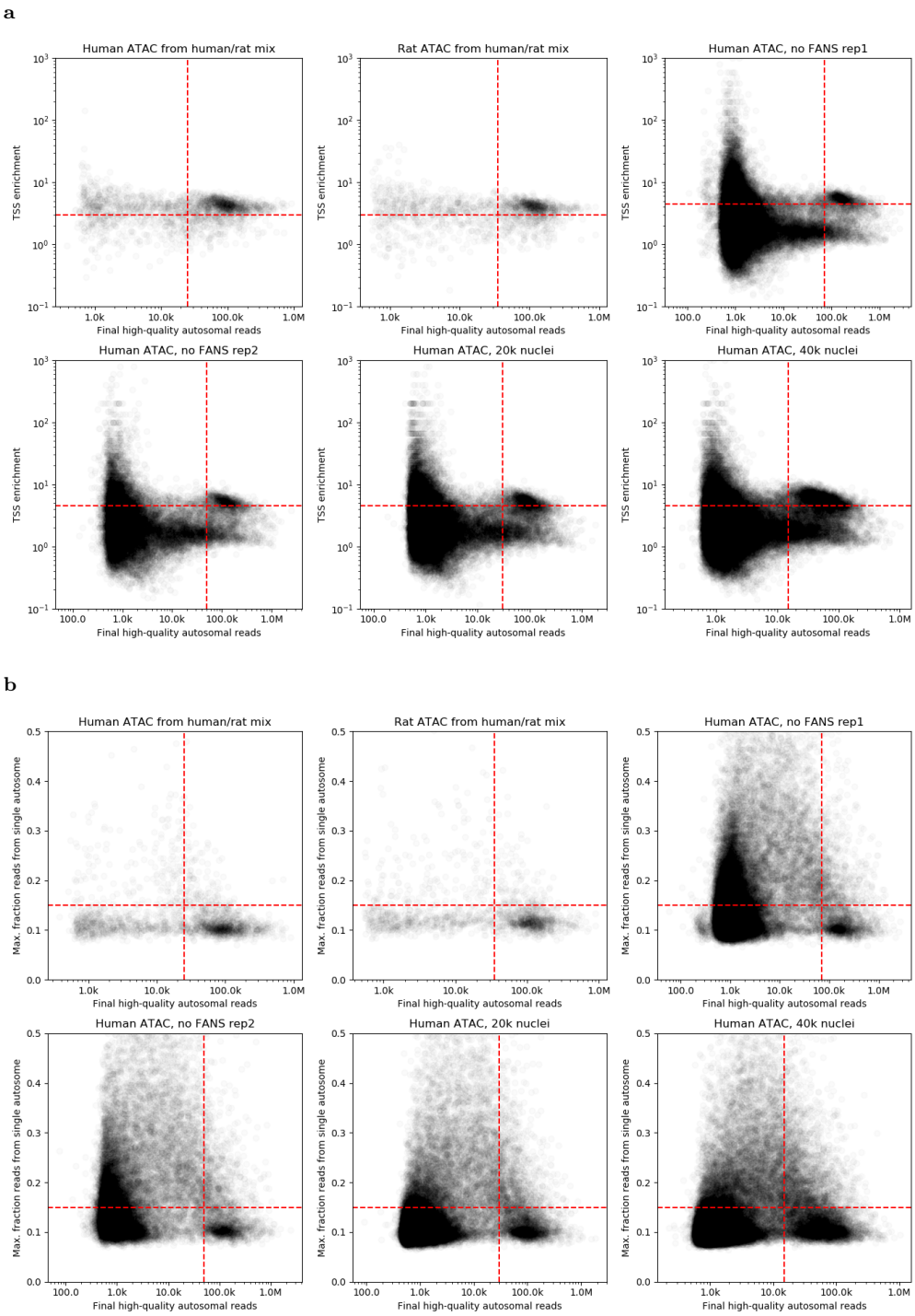

Figure S8. QC thresholds for all snATAC-seq libraries used in cell type clustering and downstream analyses. (a) Dashed lines represent thresholds for minimum number of reads and minimum TSS enrichment. (b) Dashed lines represent thresholds for minimum number of reads and the maximum fraction of reads derived from a single autosome (imposed to filter out nuclei showing abberant per-chromosome coverage).

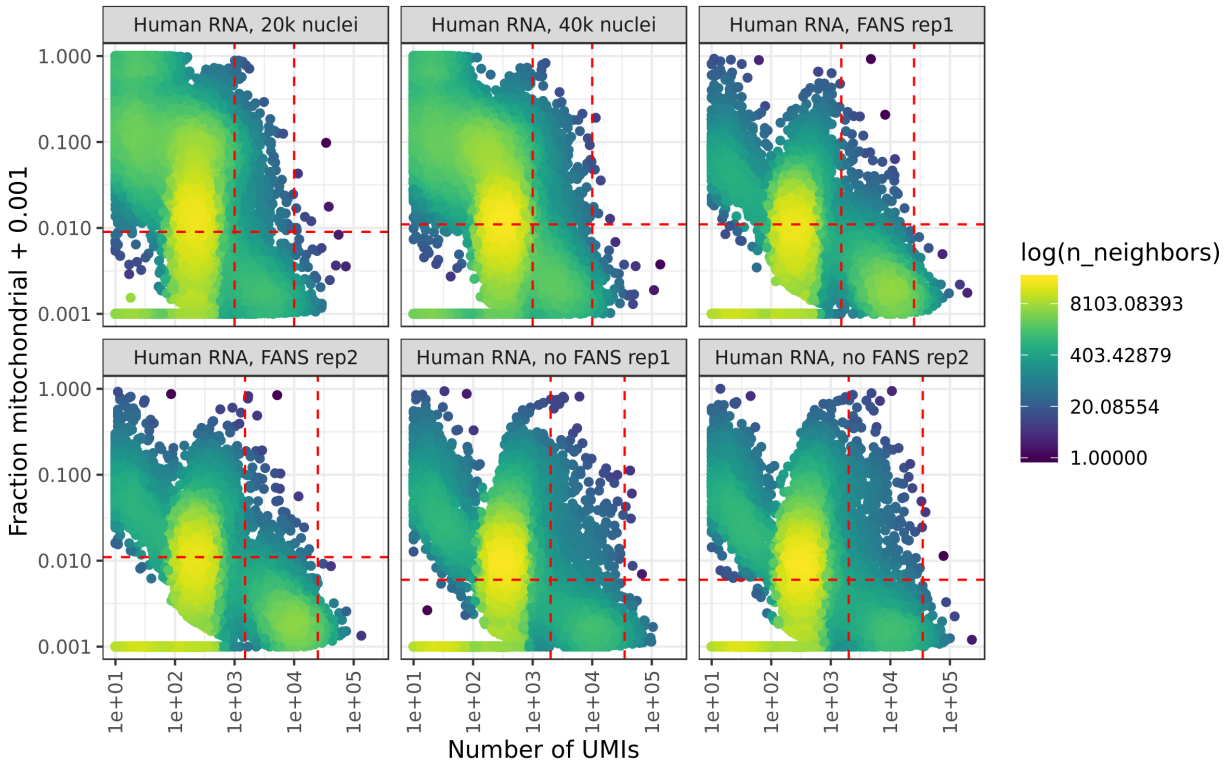

Figure S9. QC thresholds for all snRNA-seq libraries used in cell type clustering and downstream analyses. Dashed lines represent thresholds for minimum number of UMIs, maximum number of UMIs, and maximum fraction of mitochondrial UMIs.

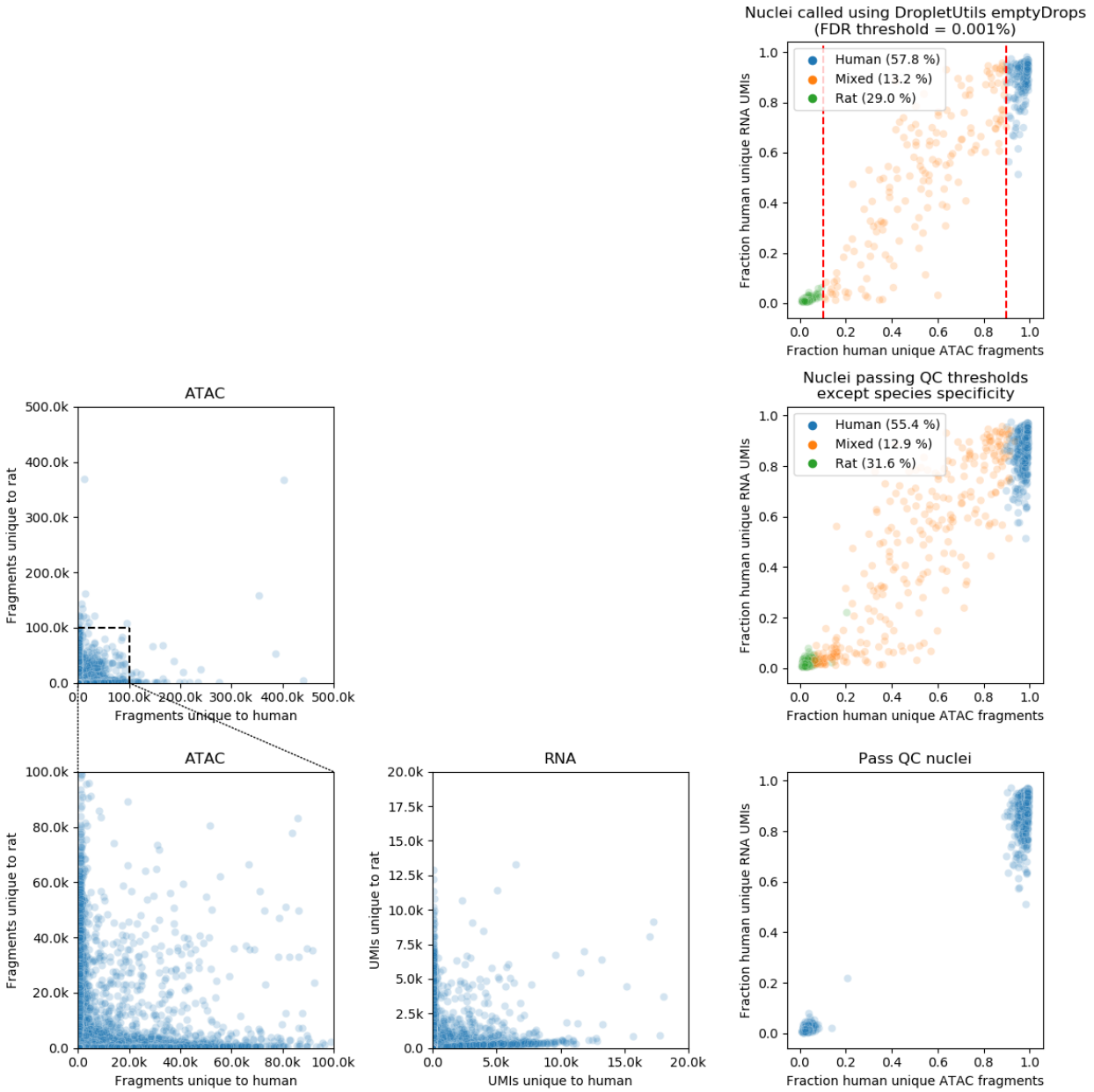

Figure S10. Number of ATAC fragments per 10x nuclear barcode mapping only to human or rat (left), number of UMIs per 10x nuclear barcode mapping only to human or rat (center), and fraction of species-unique ATAC fragments from human vs fraction of species-unique RNA UMIs from human in either nuclei called by DropletUtils' emptyDrops() function (Lun et al. 2019; top right) or nuclei passing imposed QC thresholds (middle and bottom right; middle shows all nuclei passing QC thresholds except the species specificity requirement; bottom shows all nuclei passing QC thresholds including successful assignment to a species). The greater deviation away from 0 or 1 and towards 0.5 along the RNA axis (y-axis) relative to the ATAC axis (x-axis) in the right-most panels is also observed in similar joint chromatin accessibility - gene expression data from another platform (Ma et al. 2020) and might be explained by the source of ambient RNA contamination being different than the source of ambient DNA contamination. The mixing rate noted in the upper-right panel is likely an upper bound estimate, as some nuclei marked as mixed may simply have higher ambient contamination; consistent with this hypothesis, many of the nuclei labeled here as 'mixed' still lean heavily towards one species (only 7% of nuclei lie between 0.2 and 0.8 on the ATAC axis).

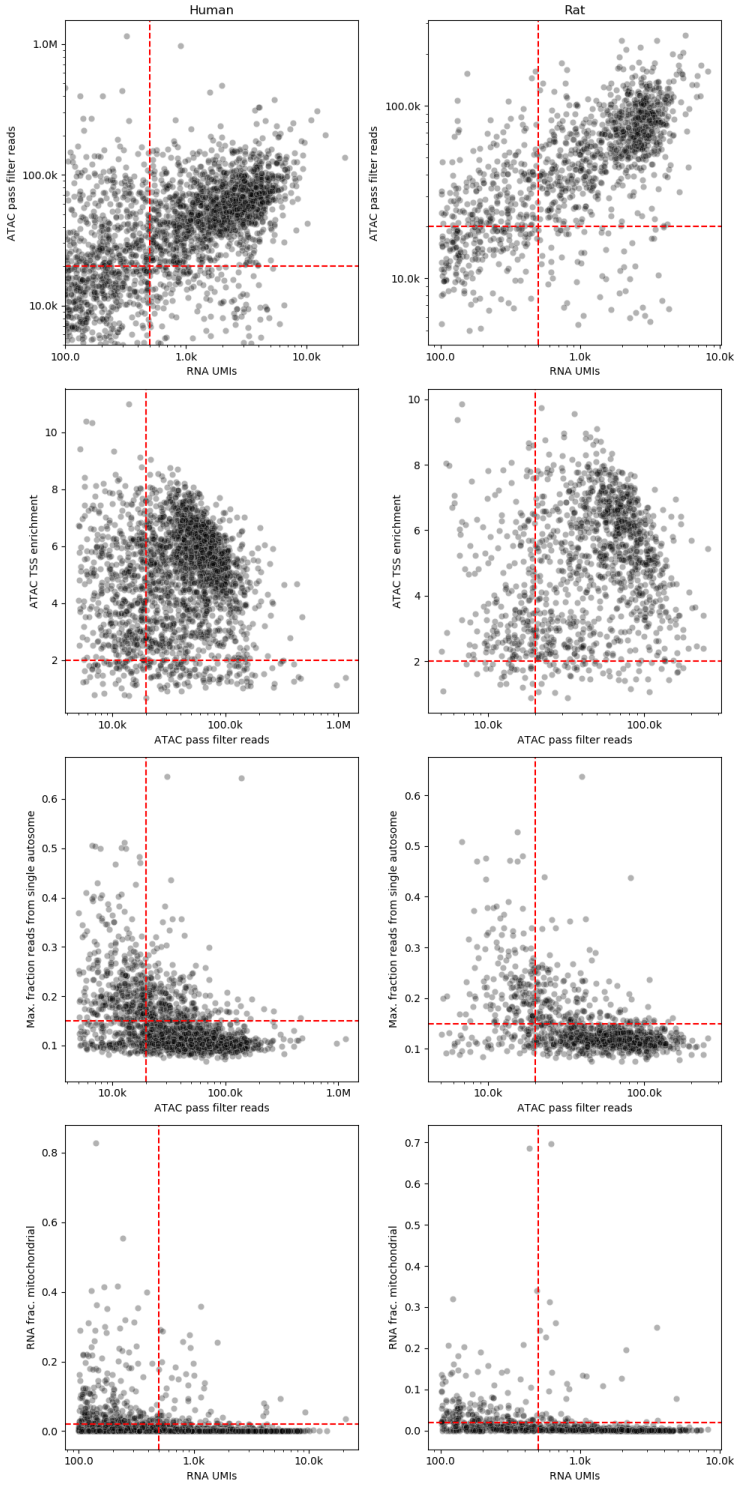

Figure S11. QC thresholds for multi-ome library

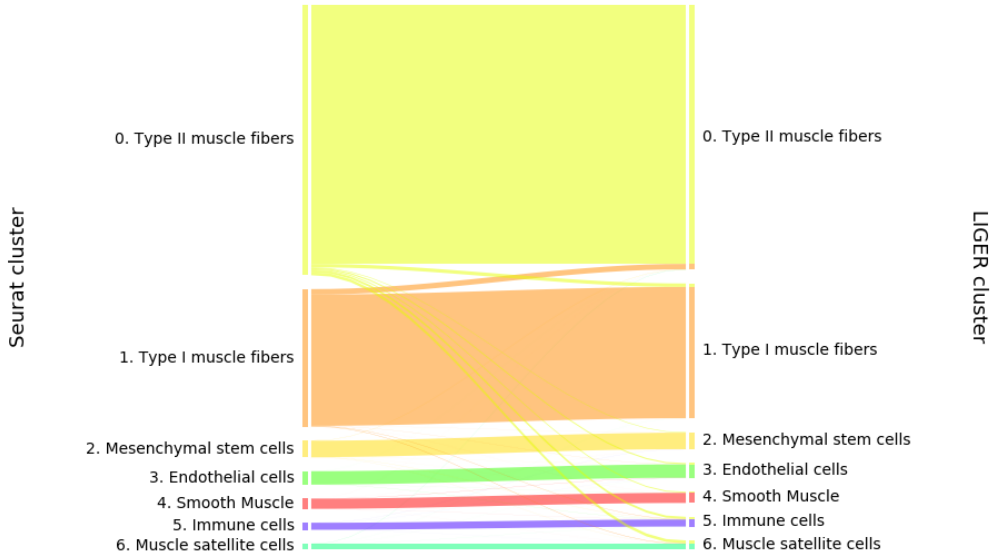

Figure S12. Comparison between clustering with Seurat and clustering with LIGER. Cluster assignments are largely concordant (95.9% of nuclei are assigned to the same cluster).

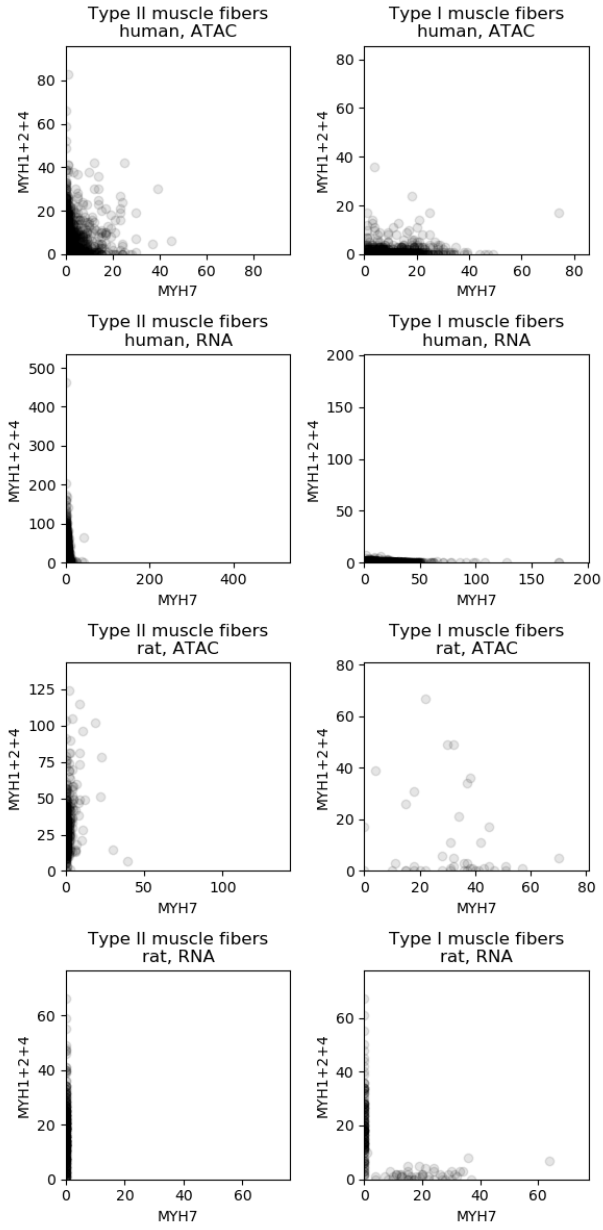

Figure S13. snATAC-seq fragment counts (gene promoter + gene body) and snRNA-seq read counts derived from the Type II muscle fiber myosin heavy chain genes (MYH1, MYH2, MYH4) or the Type I muscle fiber myosin heavy chain gene (MYH7) for human and rat nuclei. Each point represents a single nucleus. Type I muscle fibers/Type II muscle fibers headers represent the cluster to which each nucleus was assigned.

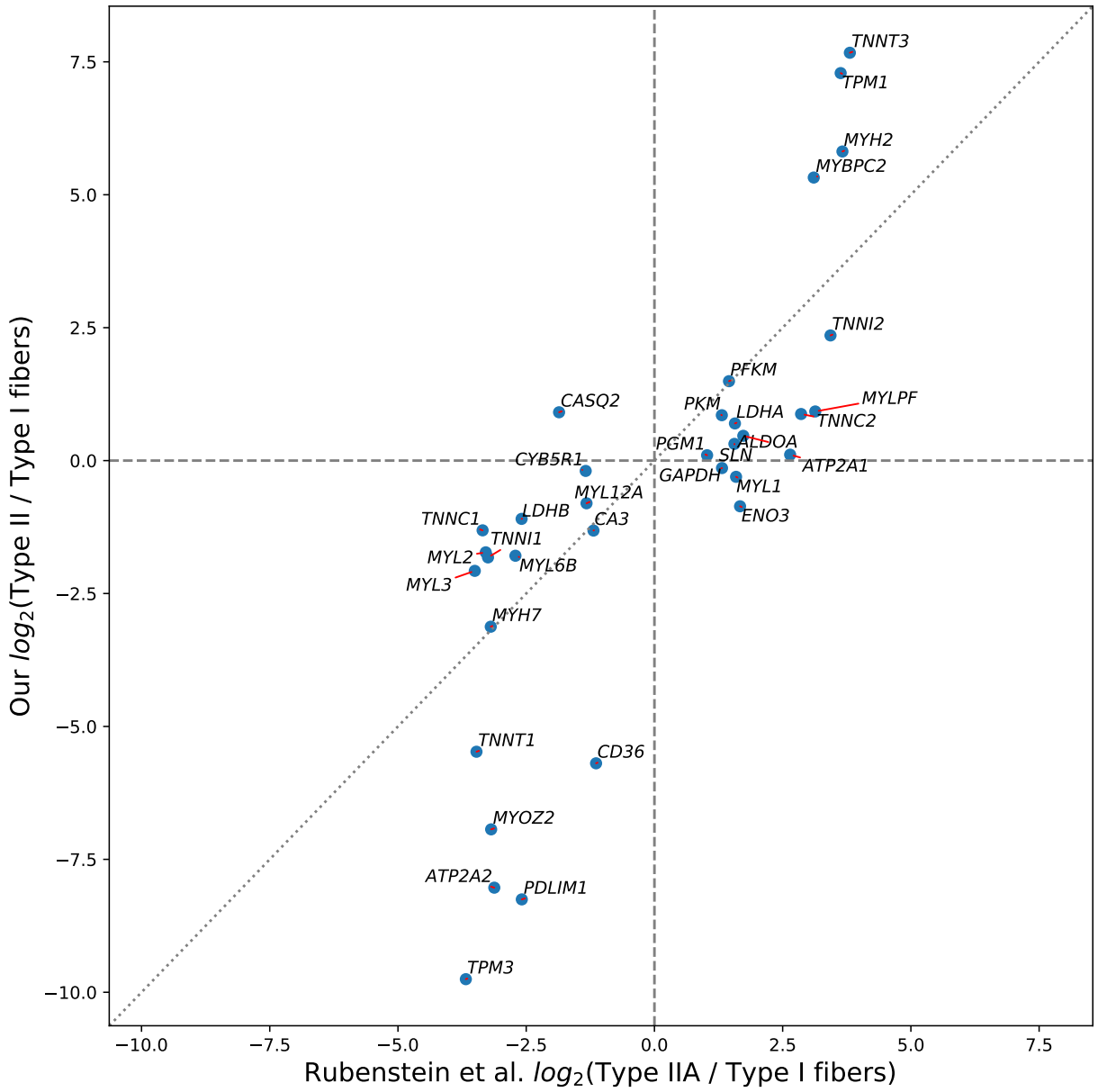

Figure S14. Log_2(fold change) for Type II vs Type I muscle fiber gene expression, showing the genes with the largest fold changes between fiber types based on data from Rubenstein et al. (Rubenstein et al. Table S4). Rubenstein et al. performed RNA-seq on pooled type I and pooled type II muscle fibers, and determined the 20 genes with the largest fold change in type II relative to type I fibers, and the 20 genes with the largest fold change in the other direction, along with p-values for differential expression. The 34 genes (of those 40 genes) that were differentially expressed are shown here.

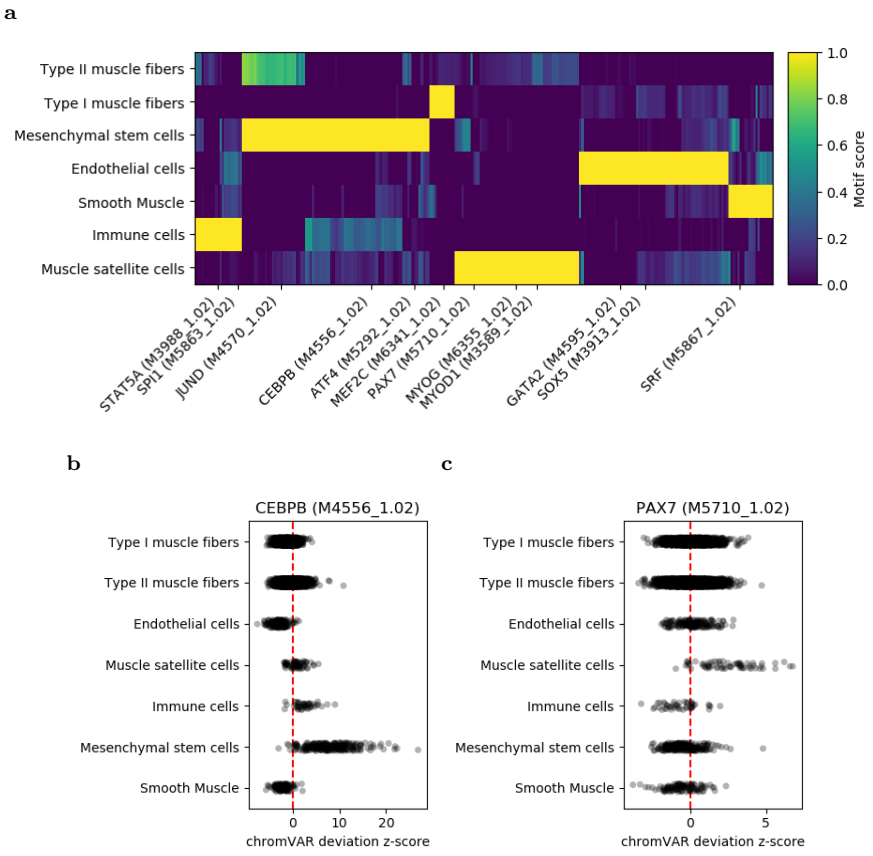

Figure S15. Motif accessibility in muscle cell type clusters. (A) Mean motif chromVAR deviation z-score, normalized between 0 and 1 for each motif. Motifs relatively specific to a single cell type (see Methods) and with mean deviation z-score > 2 in a cluster are shown; notable examples are labelled. An interactive, HTML version of this heatmap is available at https://github.com/ParkerLab/2021-03-sn-muscle/blob/master/FigS15-interactive.html. (B) and (C) Per nucleus deviation z-scores for two of the labelled motifs.

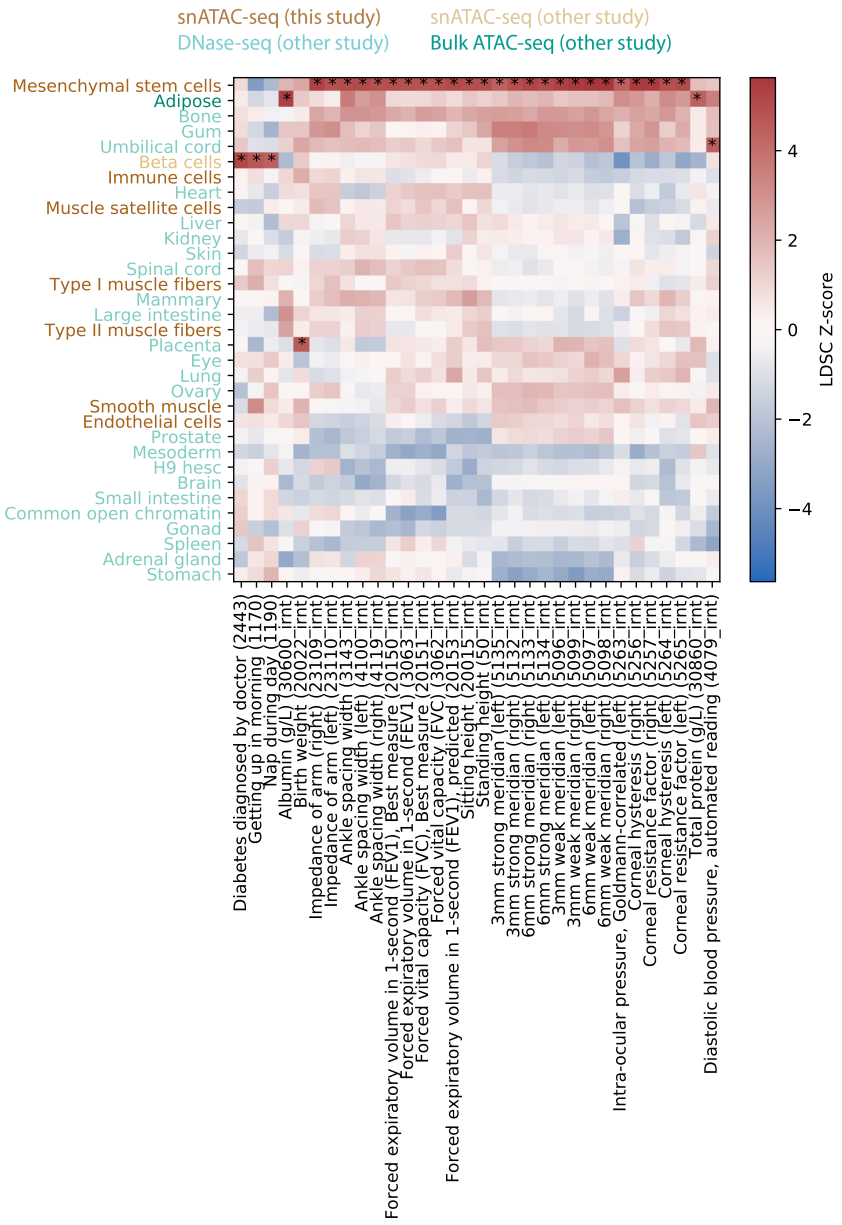

Figure S16. UK Biobank LDSC partitioned heritability results for traits for which at least one cell type showed significant heritability enrichment after Benjamini-Yekutieli correction (LDSC coefficient $p < 0.05$) across all cell types and traits (rat peaks projected into human coordinates). Asterisks denote the significant cell type - trait combinations. One trait name has been shortened to preserve space: 'Diseases of veins, lymphatic vessels and lymph nodes, not elsewhere classified' has been shortened to 'Diseases of veins, lymphatic vessels and lymph nodes'

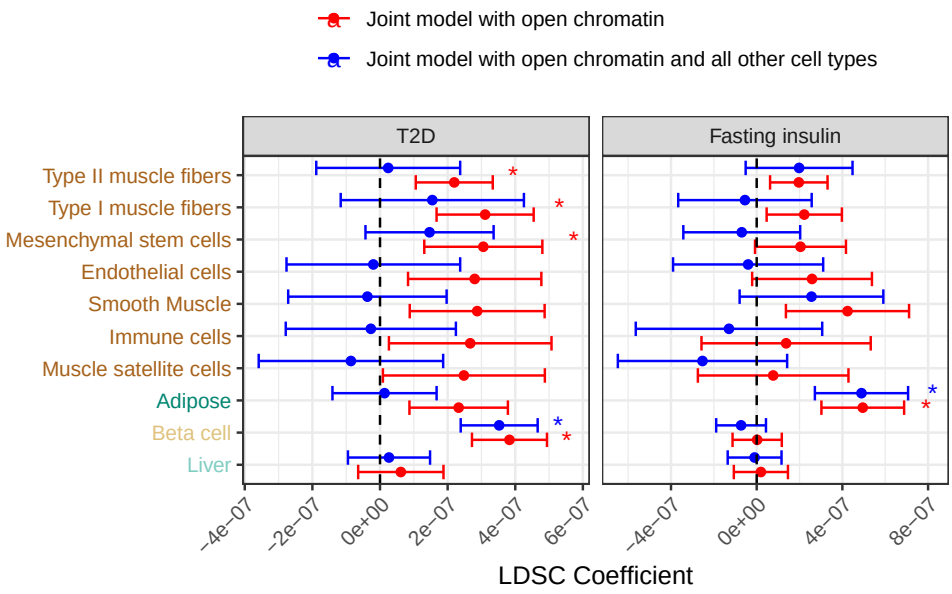

Figure S17. LDSC partitioned heritability results for T2D (BMI-unadjusted) and Fasting insulin GWAS (BMI-adjusted), using rat peak calls projected into human coordinates for the muscle cell types. Results are shown for pancreatic beta cell, adipose, and liver open chromatin regions as well. First, for each of the ten cell types, one model was run adjusting for cell type-agnostic annotations from the LDSC baseline model and common open chromatin regions (this is the joint model with open chromatin). Then, a single model containing those same annotations and all ten cell types was run (this is the joint model with open chromatin and all other cell types). Asterisk represents Bonferroni significance ($p < 0.05$ after adjusting for two traits, ten cell types, and two models per cell type = 40 tests).

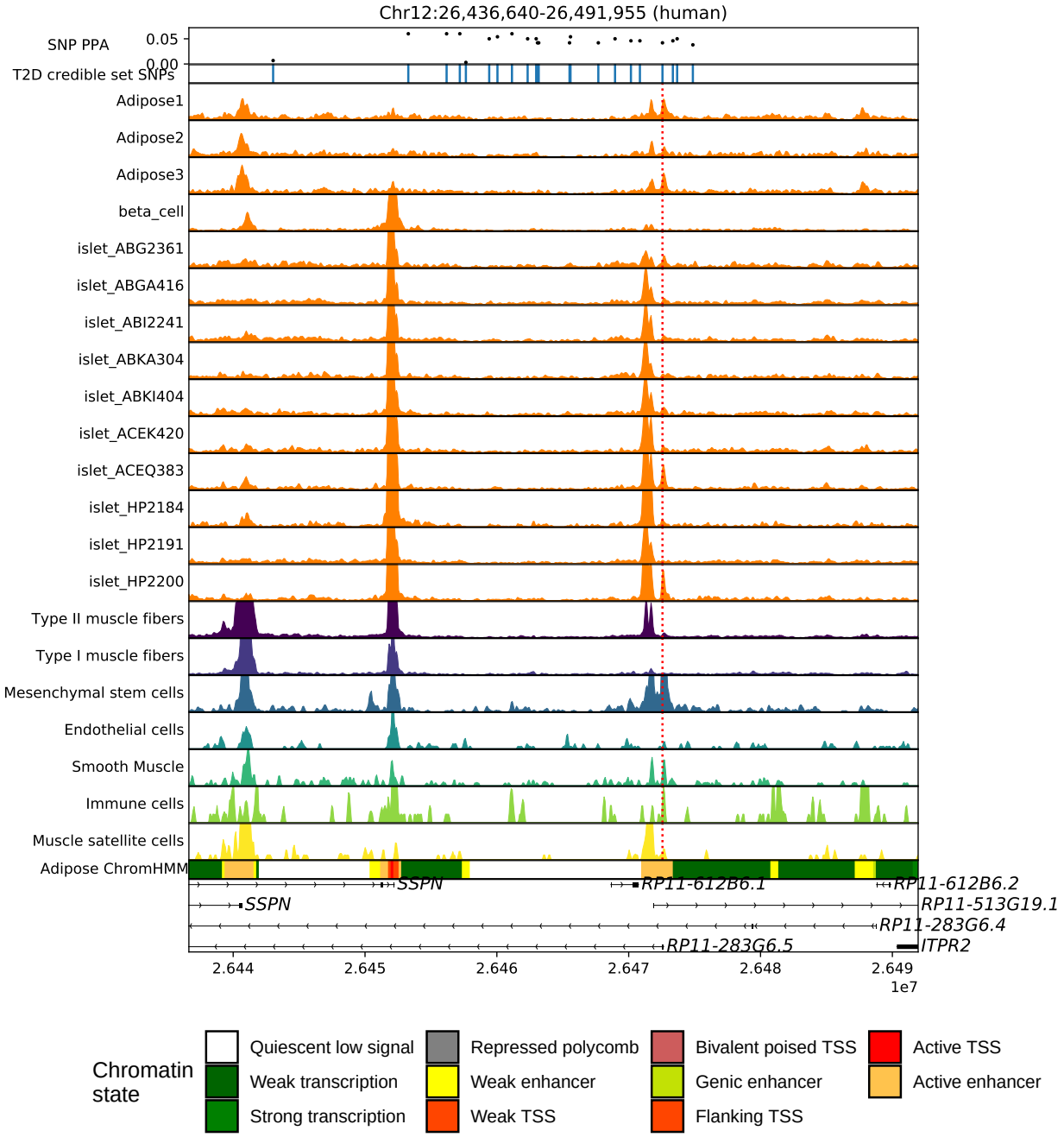

Figure S18. ATAC-seq signal in bulk adipose, bulk islet, single-nucleus pancreatic beta cell, or our muscle cell types at the *ITPR2* locus. Position of SNP rs7132434 is indicated by the long vertical red line. All ATAC-seq tracks are normalized to 1M reads and have the same y-axis range corresponding to 0 reads per million to 1 reads per million.

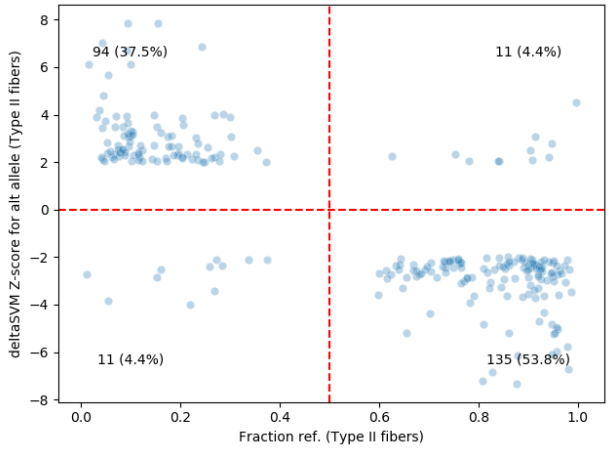

Figure S19. Comparison between deltaSVM z-score and ATAC-seq allelic bias directionality, in type II muscle fiber nuclei. Positive deltaSVM scores indicate greater predicted accessibility for the alt. allele. SNPs shown are those that have absolute deltaSVM z-score of at least 2, and show ATAC-seq allelic bias (FDR < 1%, using binomial test w/ expected fraction ref. = 0.5; only SNPs with min. coverage = 15 and at least one ref and one alt allele observed were tested)

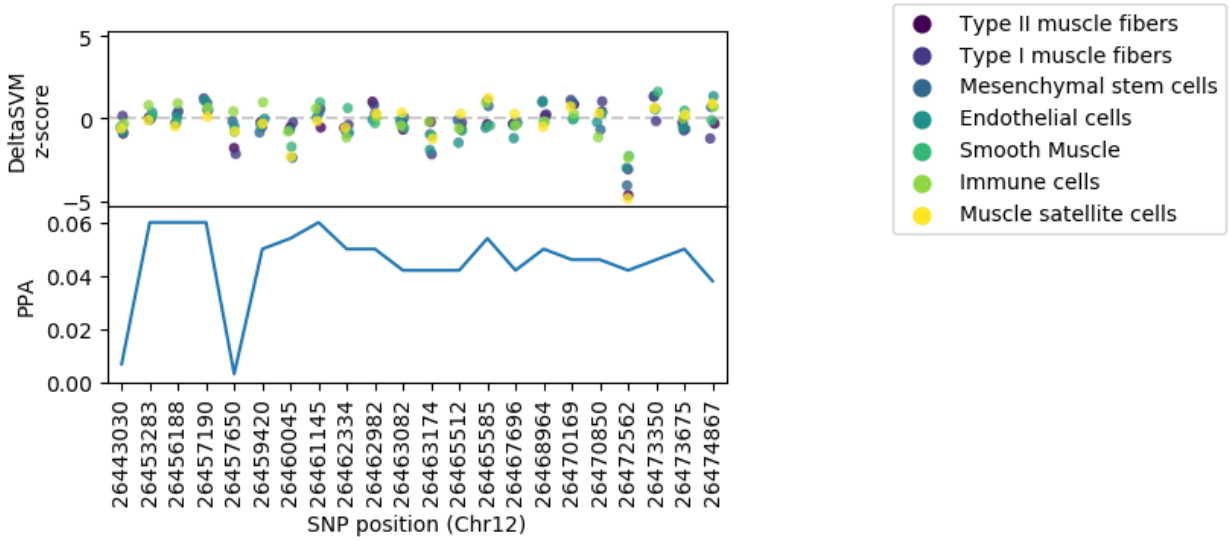

Figure S20. DeltaSVM z-score for each cell type and credible SNP at the *ITPR2* locus.

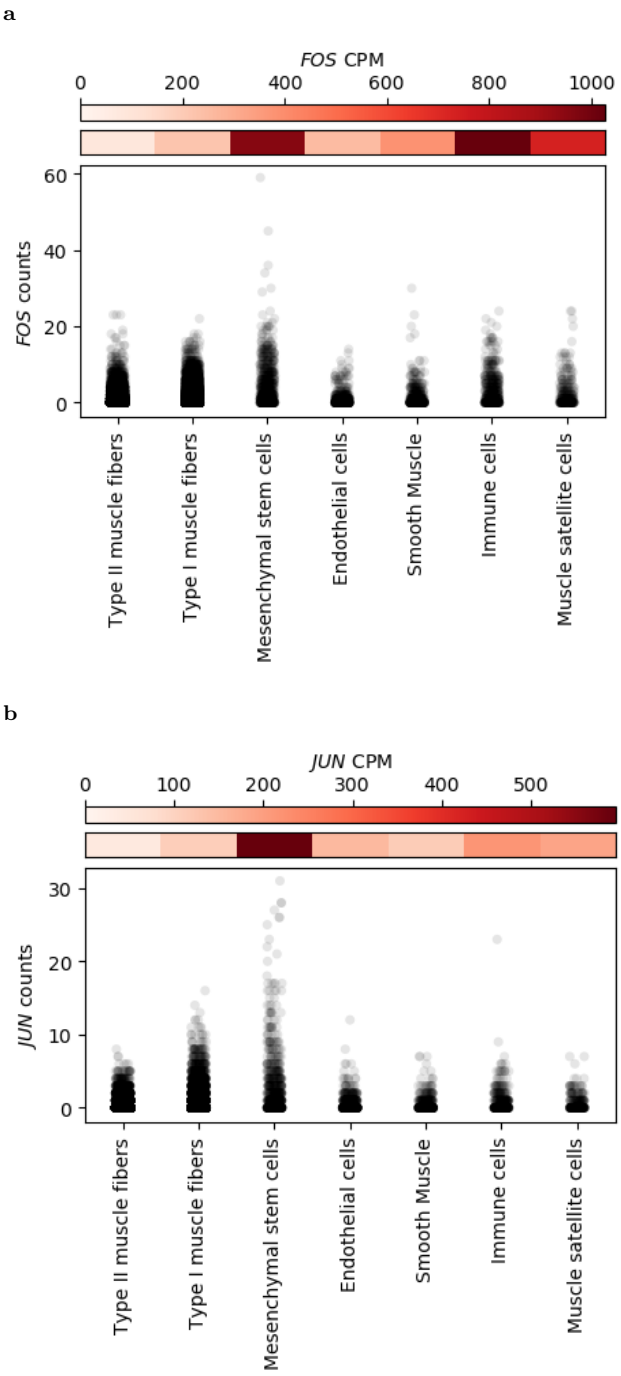

Figure S21. Epression of two AP-1 subunits in each cluster (human).

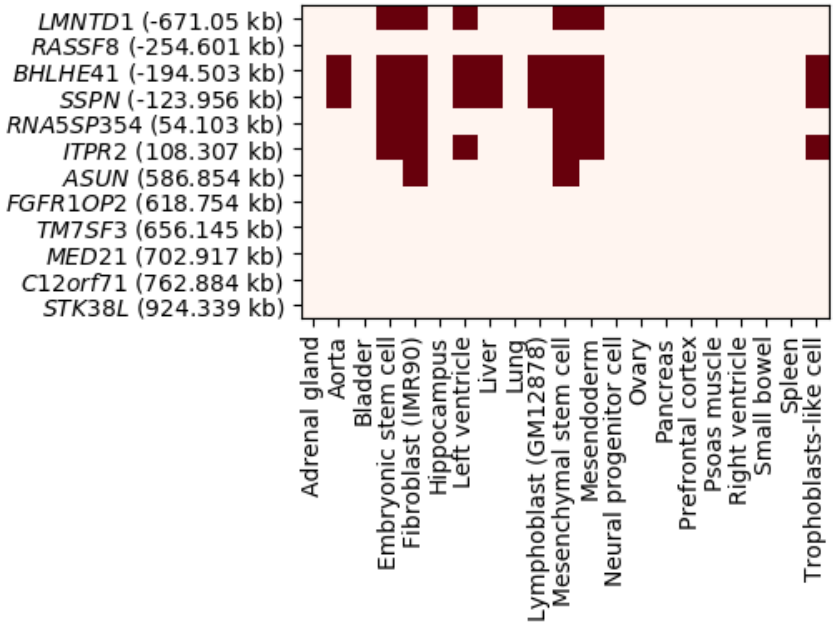

Figure S22. Presence (dark red) or absence (light red) of Hi-C connection between ATAC-seq peak containing SNP rs7132434 and gene TSS within 1 Mb of the SNP, across 21 tissues and cell types profiled in from Schmitt et al. 2016.

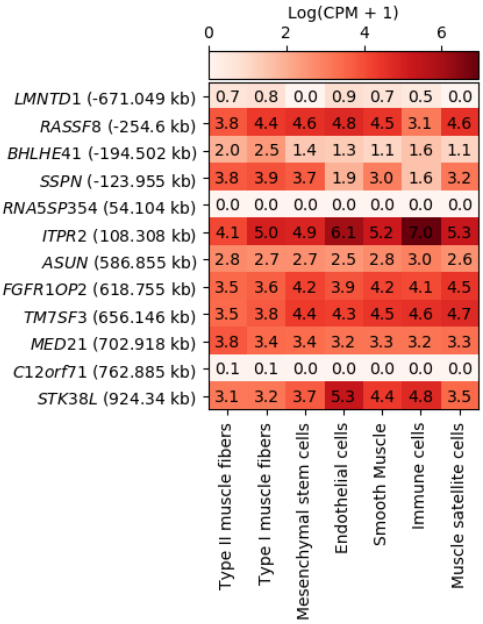

Figure S23. Expression (log(CPM + 1)) of genes with TSS within 1 Mb of SNP rs7132434

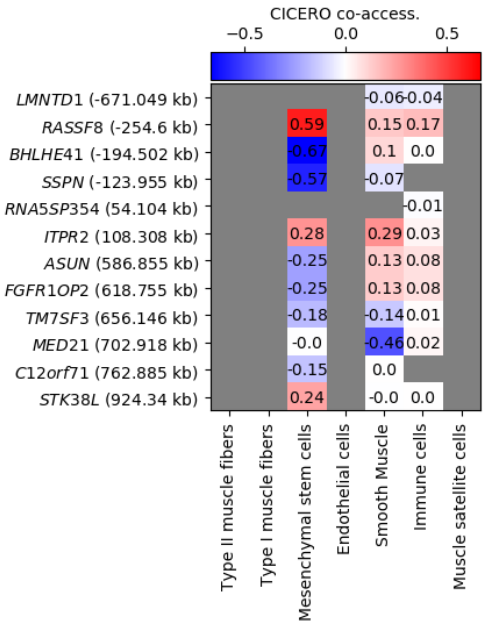

Figure S24. CICERO co-accessibility scores between ATAC-seq peak containing SNP rs7132434 and gene TSS within 1 Mb of the SNP. Grey = missing data (e.g., no gene TSS peak or CICERO coaccessibility of 'NA').

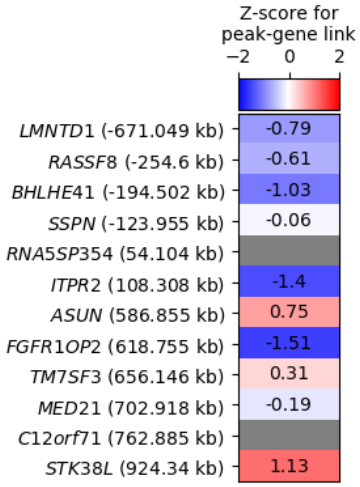

Figure S25. Z-score for peak - gene expression correlation (calculated using Signac software package LinkPeaks function), using signal of ATAC-seq peak containing SNP rs7132434 and genes within 1 Mb of the SNP. Z-score represents Spearman correlation of the peak - gene combination relative to correlation between the gene and control peaks. Grey = missing data (gene - peak pair not tested, e.g. because fewer than 5 nuclei showed a read count for both the gene and the peak).

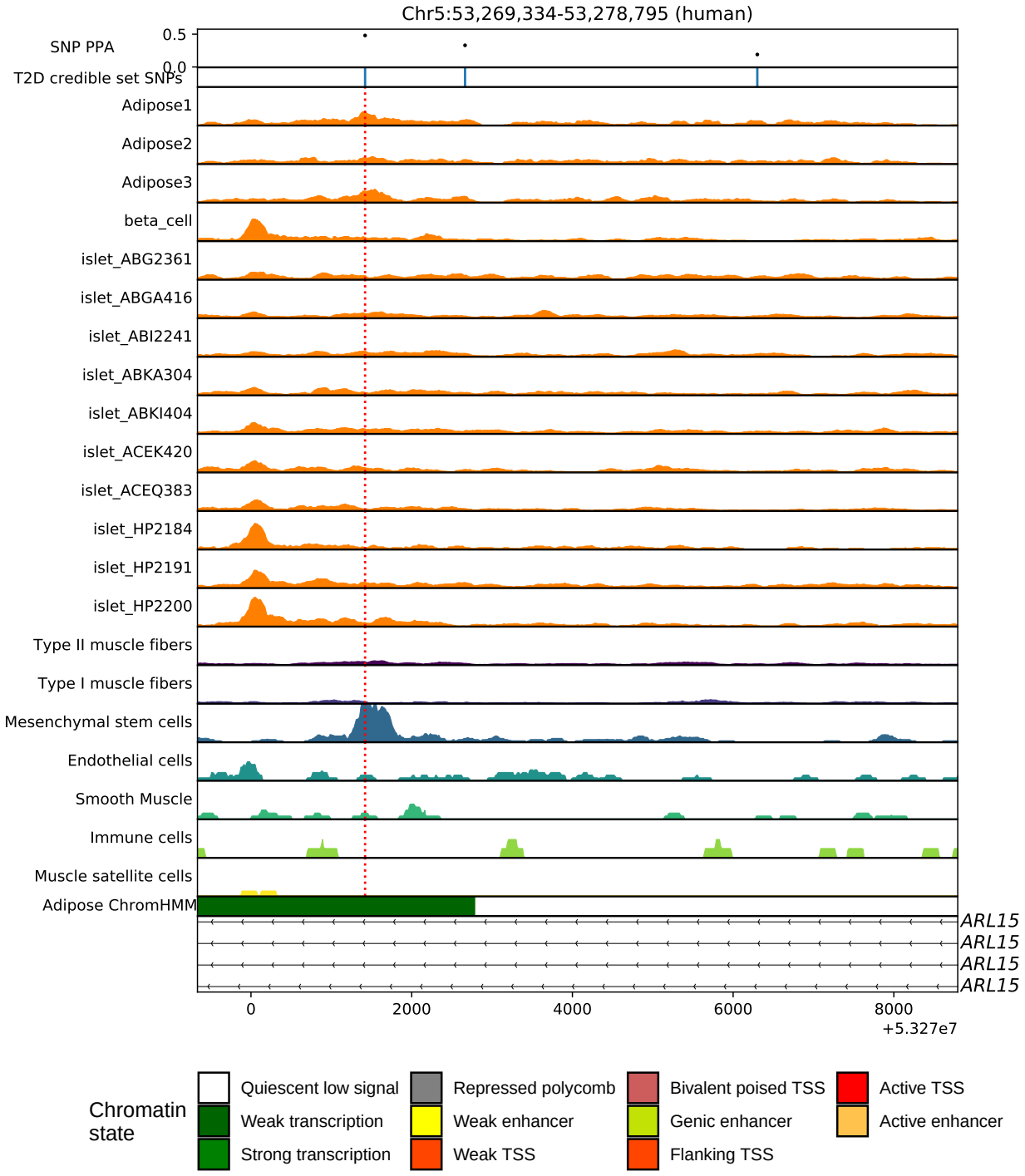

Figure S26. ATAC-seq signal in bulk adipose, bulk islet, single-nucleus pancreatic beta cell, or our muscle cell types at the *ARL15* locus. Position of SNP rs702634 is indicated by the long vertical red line. All ATAC-seq tracks are normalized to 1M reads and have the same y-axis range corresponding to 0 reads per million to 1 reads per million.

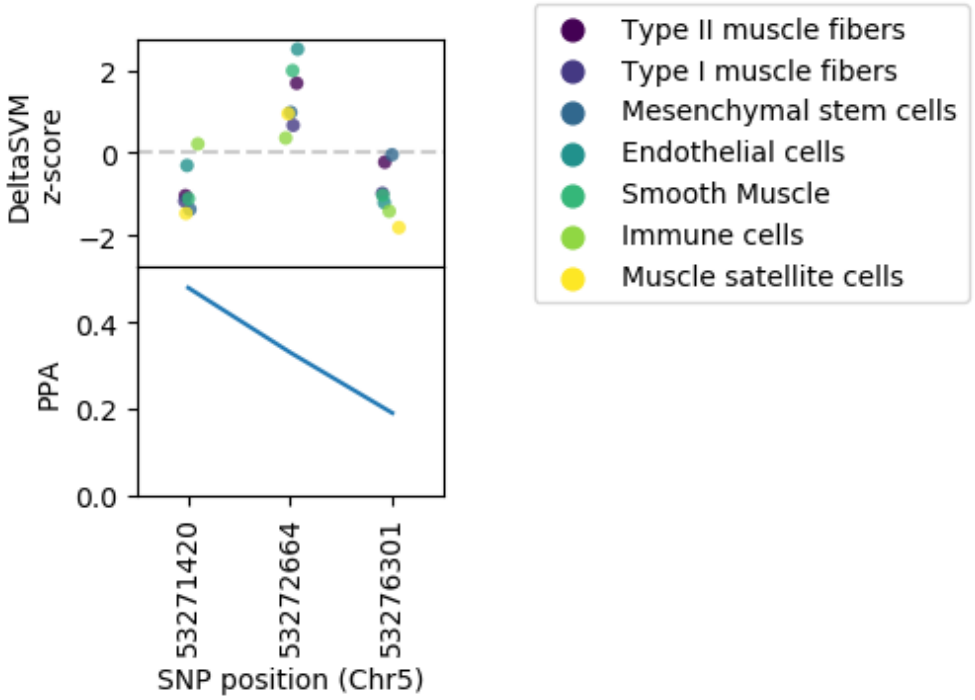

Figure S27. DeltaSVM z-score for each cell type and credible SNP at the *ARL15* locus.

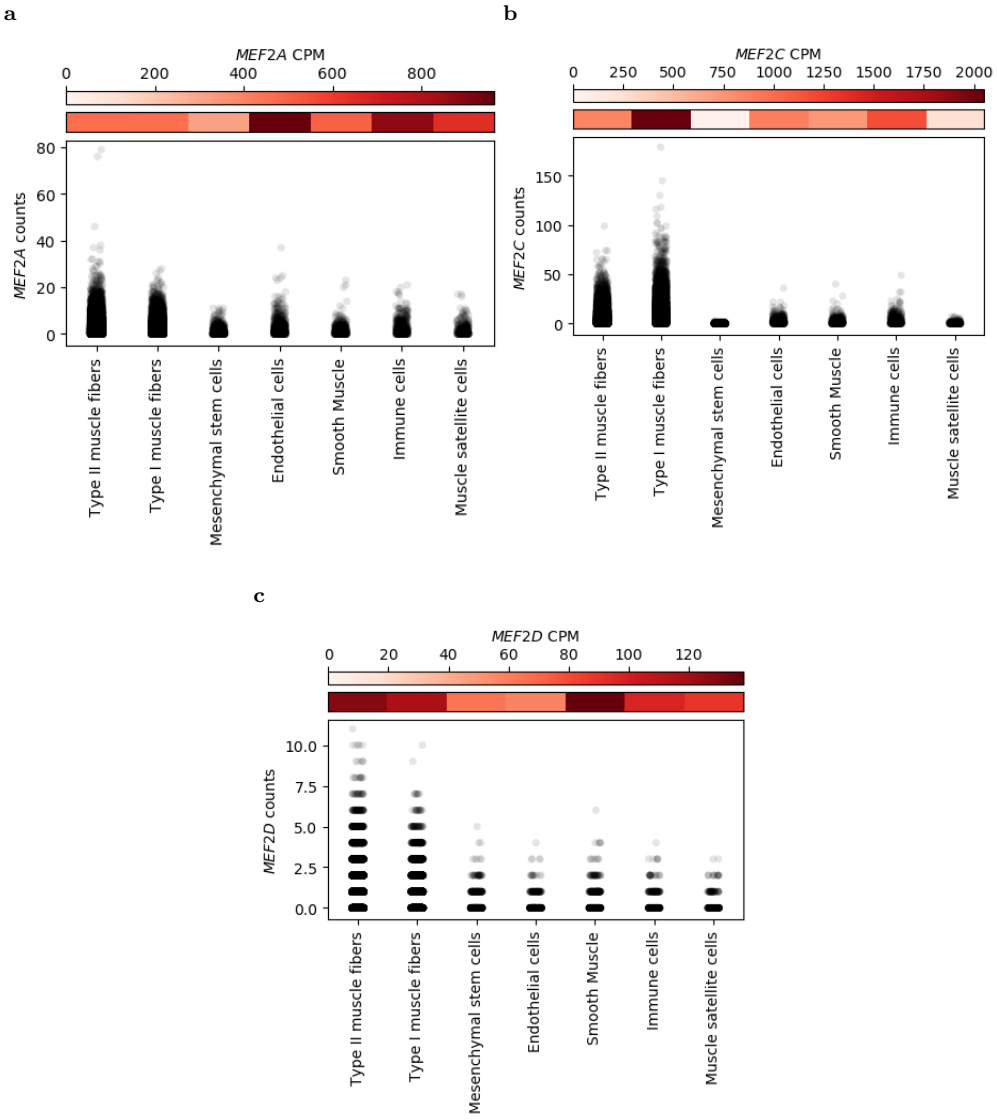

Figure S28. MEF2A/C/D expression in each cluster (human). Due to an overlapping transcript annotation, MEF2B expression is not quantified.

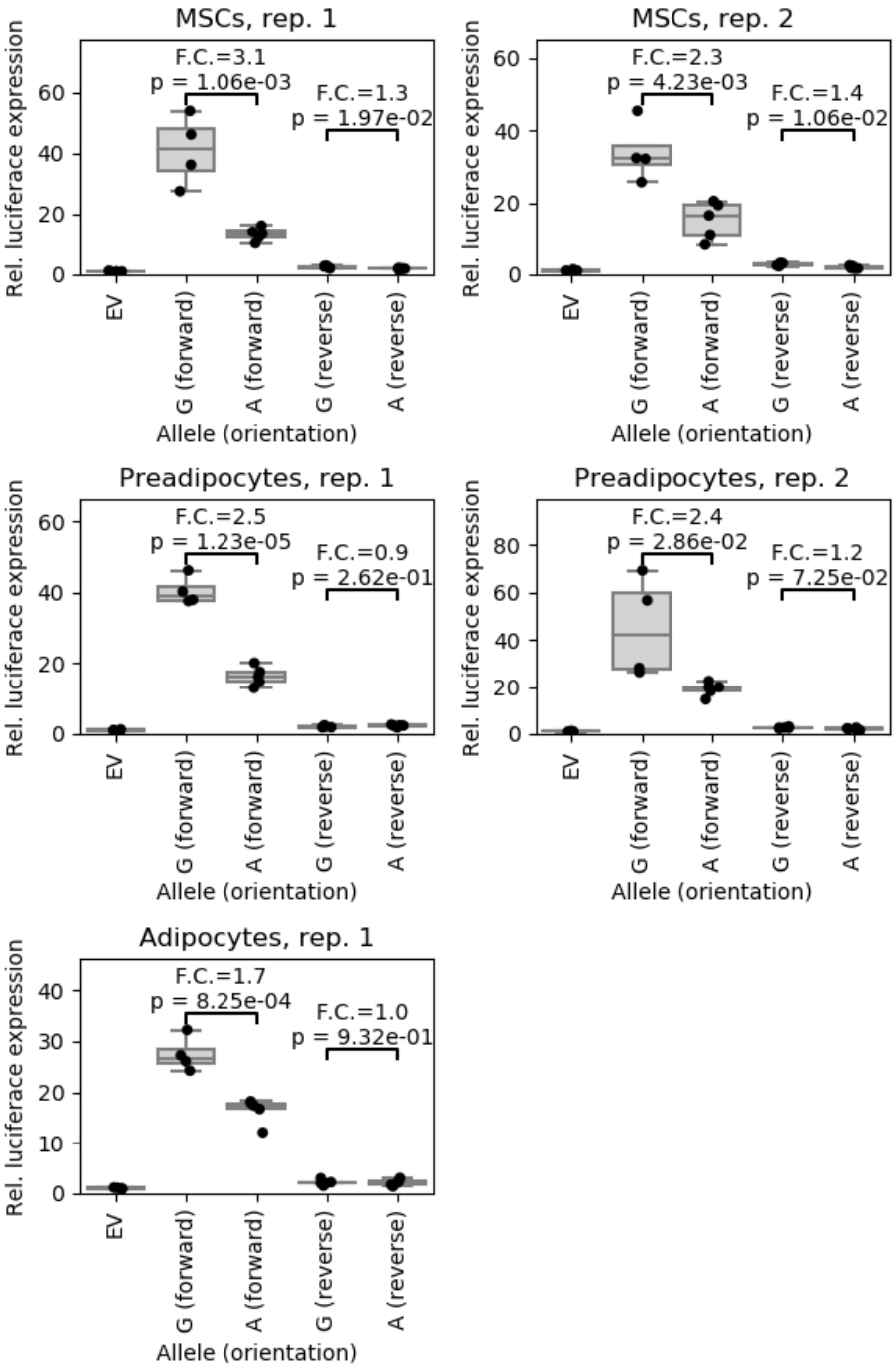

Figure S29. Luciferase assay using construct containing either allele of SNP rs702634, in human adipose-derived mesenchymal stem cells, pre-adipocytes (24 hours of mesenchymal stem cell adipogenic differentiation), and adipocytes (11 days of adipogenic mesenchymal stem cell differentiation). Each point represents one clone; p-values computed using a two-sided unpaired t-test.

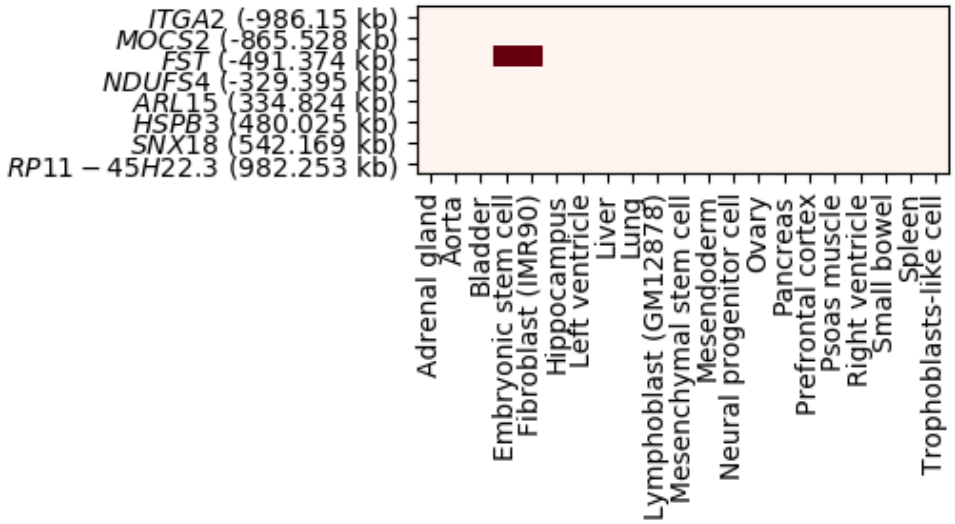

Figure S30. Presence (dark red) or absence (light red) of Hi-C connection between ATAC-seq peak containing SNP rs702634 and gene TSS within 1 Mb of the SNP, across 21 tissues and cell types profiled in from Schmitt et al. 2016.

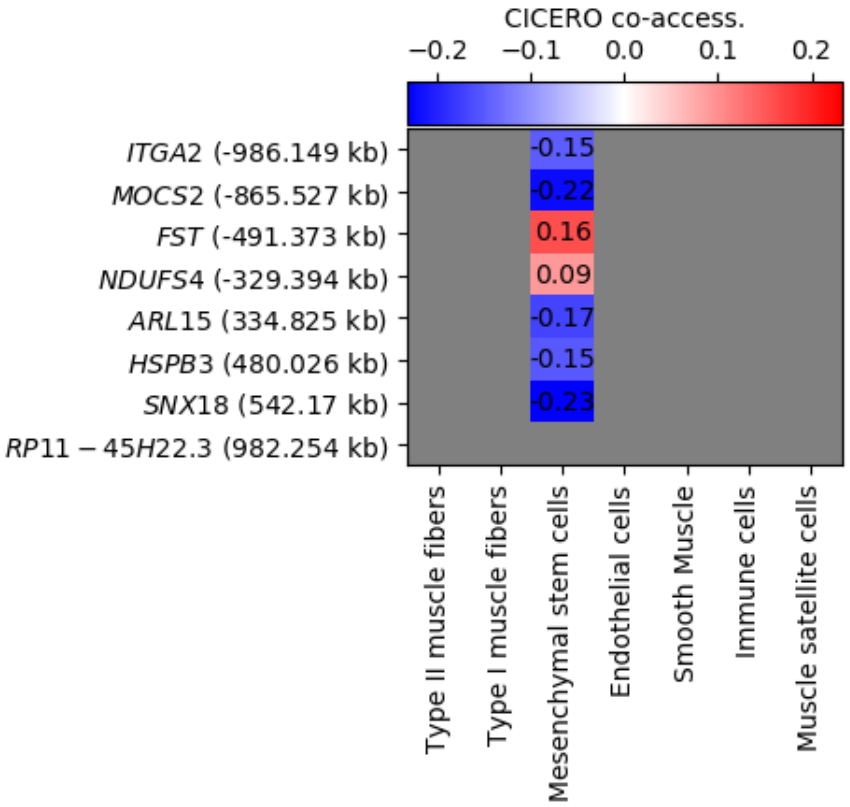

Figure S31. CICERO co-accessibility scores between ATAC-seq peak containing SNP rs702634 and gene TSS within 1 Mb of the SNP. Grey = missing data (e.g., no gene TSS peak or CICERO coaccessibility of 'NA').

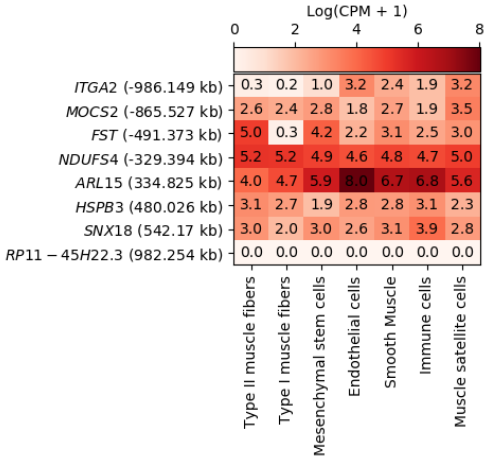

Figure S32. Expression (log(CPM + 1)) of genes with TSS within 1 Mb of SNP rs702634

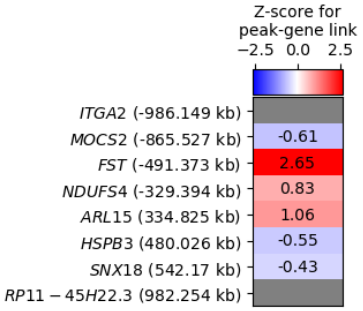

Figure S33. Z-score for peak - gene expression correlation (calculated using Signac software package LinkPeaks function), using signal of ATAC-seq peak containing SNP rs702634 and genes within 1 Mb of the SNP. Z-score represents Spearman correlation of the peak - gene combination relative to correlation between the gene and control peaks. Grey = missing data (gene - peak pair not tested, e.g. because fewer than 5 nuclei showed a read count for both the gene and the peak).

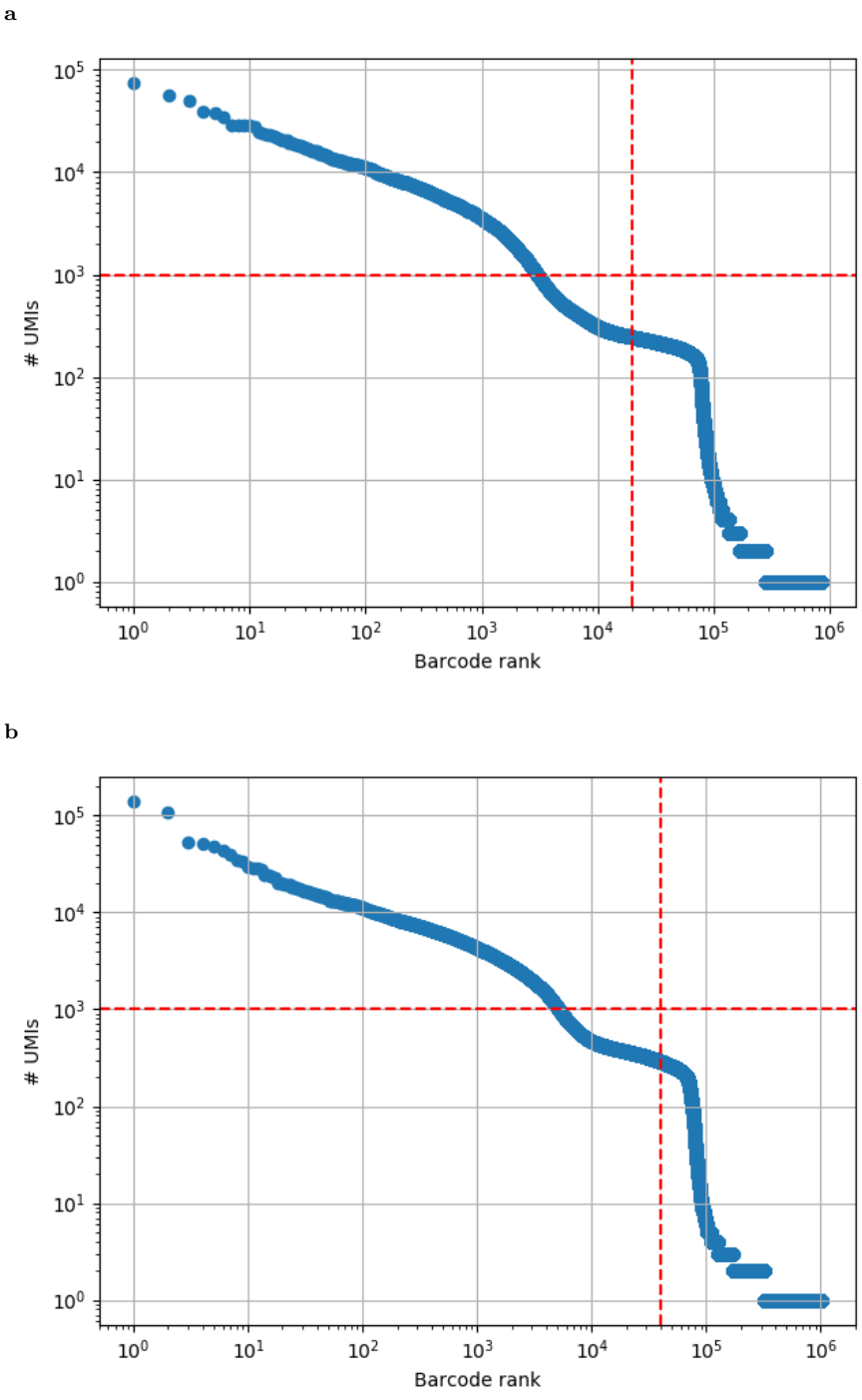

Figure S34. Barcode rank plots for snRNA-seq libraries with (A) 20k or (B) 40k nuclei loaded. Horizontal red line represents 1000 UMIs (min UMI threshold for a nucleus to be called). Vertical red line represents the rank corresponding to the number of nuclei ostensibly loaded