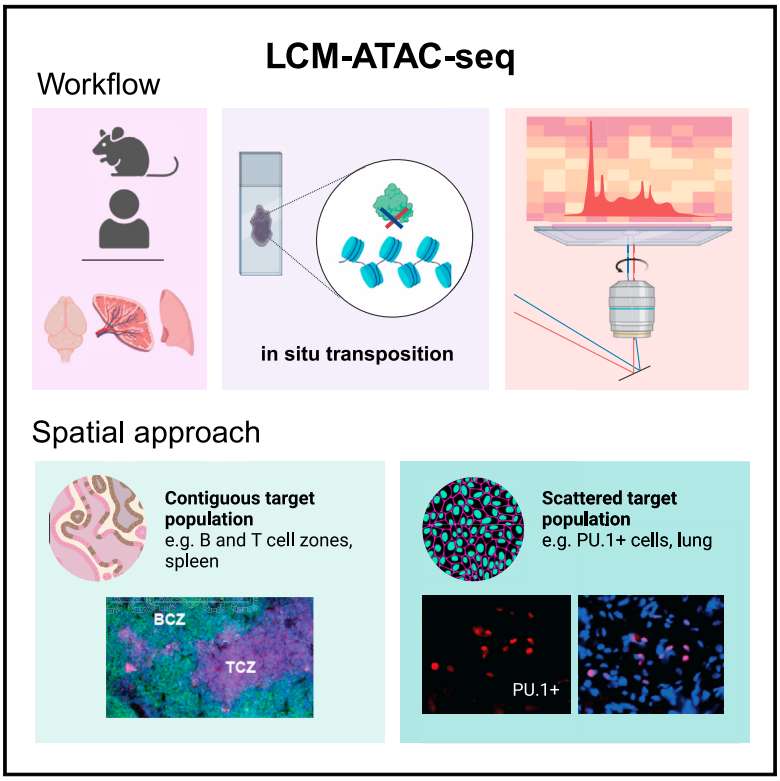
Cell Reports
Methods

Report

Chromatin accessibility profiling of targeted cell populations with laser capture microdissection coupled to ATAC-seq

Graphical abstract



Authors

Caterina Carraro, Lorenzo Bonaguro, Rachana Srinivasa, ..., Marc Beyer, Adem Saglam, Joachim L. Schultze

Correspondence

caterina.carraro@dzne.de (C.C.), a.saglam1980@gmail.com (A.S.)

In brief

Carraro et al. establish a complete workflow for the spatial characterization of chromatin accessibility in tissues of interest based on laser capture microdissection coupled to ATAC-seq. The technology can be easily combined with a morphological staining, allowing the targeting of both contiguous and scattered cell populations with a mini-bulk resolution.

Highlights

- Spatially resolved chromatin accessibility profiling with LCM-ATAC-seq
- Laser capture microdissection and *in situ* tagmentation enables targeted tissue profiling
- LCM-ATAC-seq can be combined with cellular staining to investigate target populations
- Not only contiguous but also scattered cell populations can be analyzed by LCM-ATAC-seq

Carraro et al., 2023, Cell Reports Methods *3*, 100598
October 23, 2023
https://doi.org/10.1016/j.crmeth.2023.100598

CellPress

Cell Reports Methods

CellPress
OPEN ACCESS

Report

Chromatin accessibility profiling of targeted cell populations with laser capture microdissection coupled to ATAC-seq

Caterina Carraro,[1,2,9,*] Lorenzo Bonaguro,[1,2,3] Rachana Srinivasa,[1,2] Martina van Uelft,[1,2] Victoria Isakzai,[1,2] Jonas Schulte-Schrepping,[1,2,3] Prerna Gambhir,[1,2] Tarek Elmzzahi,[1,4] Jessica V. Montgomery,[1] Hannah Hayer,[1,2] Yuanfang Li,[1,4] Heidi Theis,[3] Michael Kraut,[3] Krishnaa T. Mahbubani,[5] Anna C. Aschenbrenner,[1] Ireen König,[6] Eugenio Fava,[7] Hans-Ulrich Fried,[6] Elena De Domenico,[1,2,3] Marc Beyer,[1,3,4,8] Adem Saglam,[1,3,8,*] and Joachim L. Schultze[1,2,3,8]

[1]Systems Medicine, Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE) e.V., Bonn, Germany
[2]Genomics and Immunoregulation, Life & Medical Sciences (LIMES) Institute, University of Bonn, Bonn, Germany
[3]PRECISE Platform for Genomics and Epigenomics, Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE) e.V. and University of Bonn, Bonn, Germany
[4]Immunogenomics & Neurodegeneration, Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE) e.V., Bonn, Germany
[5]Department of Surgery, University of Cambridge, and Cambridge NIHR Biomedical Research Centre, Cambridge, UK
[6]Core Research Facilities and Services, Light Microscope Facility (LMF), Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE) e.V., Bonn, Germany
[7]Core Research Facilities and Services, Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE) e.V., Bonn, Germany
[8]These authors contributed equally
[9]Lead contact
*Correspondence: caterina.carraro@dzne.de (C.C.), a.saglam1980@gmail.com (A.S.)
https://doi.org/10.1016/j.crmeth.2023.100598

MOTIVATION Spatial epigenomics is still underexplored compared to the advanced evolution of spatial transcriptomics technologies. Different methods have been proposed recently, with distinct advantages and drawbacks concerning the achievable resolution, the possibility to overlay a morphological staining, and the overall applicability. In this study, we report the targeted use of LCM for the investigation of chromatin accessibility. This spatial methodology enables the analysis of scattered tissue cell populations at a mini-bulk resolution, with the possibility to integrate cellular or morphological stainings on the same processed section.

SUMMARY

Spatially resolved omics technologies reveal context-dependent cellular regulatory networks in tissues of interest. Beyond transcriptome analysis, information on epigenetic traits and chromatin accessibility can provide further insights on gene regulation in health and disease. Nevertheless, compared to the enormous advancements in spatial transcriptomics technologies, the field of spatial epigenomics is much younger and still underexplored. In this study, we report laser capture microdissection coupled to ATAC-seq (LCM-ATAC-seq) applied to fresh frozen samples for the spatial characterization of chromatin accessibility. We first demonstrate the efficient use of LCM coupled to *in situ* tagmentation and evaluate its performance as a function of cell number, microdissected areas, and tissue type. Further, we demonstrate its use for the targeted chromatin accessibility analysis of discrete contiguous or scattered cell populations in tissues via single-nuclei capture based on immunostaining for specific cellular markers.

INTRODUCTION

Omics technologies aim for the comprehensive detection of biomolecules and traits in target tissues and cell populations, i.e., genomic variation, epigenetic modifications, and transcriptional output. Being powerful tools to disentangle diverse layers of biological complexity of tissues and organs, omics allowed the construction of granular atlases describing the heterogeneity of transcriptional programs across cell types and states.[1,2] Technological advances finally enable the investigation of omics layers up to the single-cell level with spatial resolution by combined analysis of cellular localization and transcriptional output,

Table 1. Comparison between spatial-ATAC-seq, sciMAP-ATAC, solid-phase-capture spatial ATAC-seq, and LCM-ATAC-seq

	Spatial ATAC-seq	sciMAP-ATAC	Solid-phase-capture spatial ATAC-seq	LCM-ATAC-seq
Founding technology	*in situ* transposition + microfluidic device	microbiopsy punches	*in situ* transposition + solid-phase capture	*in situ* transposition + LCM
Tagmentation mode	*in situ*	post-punched	*in situ*	*in situ*
Spatial resolution	d: 20 μm spots	h: 100–300 μm d: 250–500 μm microbiopsies	d: 55 μm spots 100 μm interspot distance	single captured nuclei
ATAC-seq resolution	d: 20 μm spots	single nucleus (by dissociation)	d: 55 μm spots 100 μm interspot distance	50 single captured nuclei
Morphological marker-based selection	not possible	possible within the specified spatial resolution	possible within the specified spatial resolution	possible within the specified ATAC-seq resolution
Suggested approach	agnostic	hypothesis-driven/agnostic	agnostic	hypothesis-driven

The previously reported spatial ATAC-seq, sciMAP-ATAC, and solid-phase-capture spatial ATAC-seq technologies are here compared with LCM-ATAC-seq. LCM, laser capture microdissection; h, height; d, diameter.

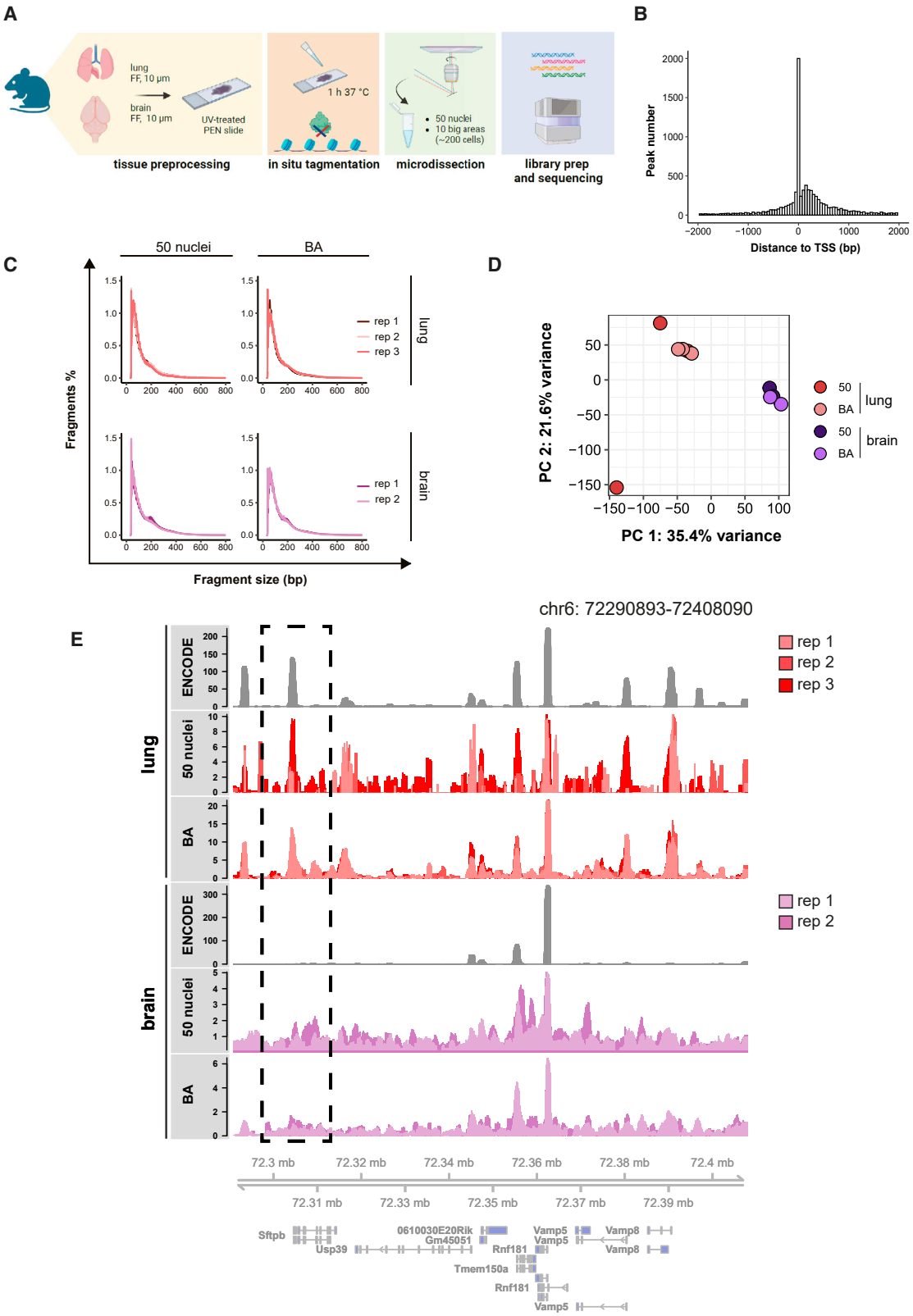
thereby linking phenotypes to spatial context.[3–5] Recent years have seen the development of a multitude of sequencing and microscopy-based spatial transcriptomic technologies, i.e., based on highly multiplexed fluorescence *in situ* hybridization, *in situ* sequencing, or *in situ* barcoding, and numerous additional technologies are in the pipeline.[6]

Mechanisms of epigenetic regulation are critical for the control of transcriptional dynamics.[7] Indeed, modulation of chromatin accessibility is at the basis of cell differentiation, reprogramming, and maintenance of cell identity.[8] Different technologies were developed to investigate chromatin accessibility, i.e., DNase-seq, MNase-seq, and ATAC-seq,[9] all capable of addressing epigenetic heterogeneity across cell types and tissues up to the single-cell level.[10,11] Retaining the localization of profiled cells or nuclei within a tissue would allow us to connect epigenetic to spatial information. To address this point, Thornton et al.[12] recently developed sciMAP-ATAC (single-cell combinatorial indexing on microbiopsies assigned to positions for the assay for transposase accessible chromatin) for spatial single-cell ATAC-seq based on microbiopsy punching and combinatorial indexing, capable of assessing patterns of chromatin accessibility with a quality comparable to non-spatial sci-ATAC-seq protocols. Nevertheless, the spatial resolution of this approach is limited by the diameter size of collected microbiopsies (minimum 100 × 250 μm), and a morphological characterization of the tissue can only be acquired on a separate section and post-registered after punching. In a subsequent study, Deng et al. reported spatial ATAC-seq,[13] which relies on *in situ* transposition[14] and microfluidic deterministic barcoding. This second example can notably reach a resolution of 20 μm spots but does not allow us to further implement a morphological marker-based selection of nuclei/areas, and thus it is indicated mostly for hypothesis-free analyses. Interestingly, the method has been recently reported for the spatial co-profiling of the epigenome and transcriptome.[15] Recently, Llorens-Bobadilla et al. reported the use of solid-phase-capture spatial ATAC-seq,[16] an epigenomic analog of the spatial transcriptomics (ST) technology,[17] also based on *in situ* transposition and spatial barcoding, reaching a resolution of 55 μm diameter spots. Beyond chromatin accessibility, technologies able to investigate other epigenetic traits are also under development, as recently reported, e.g., for epigenomic MERFISH (multiplexed error-robust fluorescence *in situ* hybridization) capturing specific histone modifications with single-molecule resolution.[18]

In this work, we report the use of ultraviolet (UV)-based laser capture microdissection (LCM) coupled with ATAC-seq for the investigation of chromatin accessibility in fresh frozen tissues (LCM-ATAC-seq; Table 1). LCM allows us to isolate subpopulations of cells from tissue slices via direct microscopic visualization and targeted laser microdissection, enabling collected samples to be processed for different downstream analyses including genotyping and RNA and protein profiling.[19] LCM exploits a UV or infrared (IR) laser for the microscopic dissection of tissue areas, which are harvested through integrated collector systems for further processing. So far, LCM has been connected with omics technologies to assess the genome, transcriptome, or proteome from cells within tissues[20] but not for the analysis of the chromatin landscape, e.g., via ATAC-seq. Notably, Nichterwitz et al. first described the use of LCM with single-cell resolution, which was coupled to Smart-seq2 for the analysis of mRNA in mouse and human brain tissue samples[21] and was further developed into the Geo-seq protocol by Chen et al.[22] In addition, methods have been reported coupling LCM to mass spectrometry for proteomics analyses, both on fresh frozen and formalin-fixed paraffin-embedded (FFPE) tissues, as well as lipidomics and metabolomics studies.[23–26] Addressing epigenetic regulation, LCM has been coupled with reduced representation bisulfite sequencing (LCM-RRBS) to gain insights on DNA methylation changes, e.g., associated with gonadectomy-induced adrenocortical neoplasia in the mouse.[27] These studies illustrate that LCM can be utilized to obtain spatial single-cell-level information for important and yet unaddressed omics layers.

Here, we investigated if LCM can be extended to the assessment of nucleosome-free chromatin areas. To this end, we established a workflow for LCM-ATAC-seq analysis of mouse brain, lung, and spleen tissues, evaluating different resolution and microdissection strategies. We then demonstrated its use for the targeted analysis of distinct intra-tissue populations and finally optimized it for the characterization of scattered cell populations in healthy human lung samples.

CellPress
OPEN ACCESS



(legend on next page)

RESULTS

LCM-ATAC-seq allows characterizing tissue-specific chromatin accessibility

The applicability of the LCM-ATAC-seq protocol was first assessed in an untargeted approach on mouse fresh frozen tissues as a function of cell number, microdissected areas, and tissue type (Figure 1A). We isolated lung lower right lobe and cerebrum samples from 6-month-old C57BL/6J mice to evaluate the LCM-ATAC-seq workflow on tissues with distinct histological properties and processing standards. *In situ* tagmentation was performed by incubating 10 μm lung and brain sections with the Tn5 transposase for 1 h at 37°C and further proceeded with the UV laser microdissection (for details, see the STAR Methods). In order to investigate the resolution of our approach, we collected both mini-bulks composed of either 50 single nuclei captured randomly across the tissue based on DAPI staining or 10 larger dissected areas (big areas [BAs]), each one of roughly 20 cells (for a total of ∼200 nuclei), at least in duplicate. The possibility to downscale to single-nuclei collection was excluded due to the insufficient quality of isolated material after LCM (data not shown). Cross-linking of the tagmented DNA was then reversed, libraries were generated, and paired-end sequencing was performed. The resulting reads were then further processed for downstream analysis, including alignment using bowtie2 and peak calling using MACS2, as specified in the STAR Methods.

We first evaluated the quality of the obtained dataset using established descriptors for ATAC-seq data.[28] The alignment quality control (QC) attested to the purity of the obtained libraries, with almost all samples reaching rates over 98% (Figure S1A). As shown in Figure 1C, profiles of fragment sizes showed a nucleosome pattern distribution (fragment size distribution [FSD]) in line with published spatial ATAC-seq technologies,[12,13] with the appearance of a mono- and a di- (sometimes also tri-)nucleosome fingerprint demonstrating the successful completion of the transposition process. Further, reads generated by Tn5 transposition are expected to accumulate in promoter regions at transcription start site (TSS) sites of open chromatin. To this end, TSS enrichment scores (TSS-ESs; Figure S1B) showed values always above 3 across samples (3–10), reaching very high scores in the case of lung BA samples (>10). These results demonstrated a high enrichment of TSSs in the data with only limited background enrichment in the obtained libraries. In order to define a joint model for the dataset, we called peaks after merging the sequencing data of all samples (see STAR Methods) and identified 35,423 peaks mostly distributed in promoter regions (Figure S1C). As expected, reads mapping to peaks were mostly found around TSSs (Figure 1B). The fraction of total reads falling in peaks (FRiP) was always above 20% across samples, which is lower than the 30% reported for bulk ATAC-seq data[29] but in line with other spatial ATAC-seq examples,[13] further attesting to a good signal-to-noise ratio. Additionally, a slight difference in the main QC metrics across conditions suggested that the performance of the technology might vary depending on the tissue type (e.g., higher mitochondrial reads in brain compared with lung samples; Figure S1B) and quality of the starting material.

To determine whether LCM-ATAC-seq could retain original tissue features, we performed principal-component analysis (PCA) (Figure 1D), demonstrating distinct clustering of lung and brain samples. To further illustrate tissue specificity, we visualized in Figure 1E the sample-wise non-aggregated accessibility profile of the *Sftpb* locus, a well-known lung-specific gene,[30] showing promoter peaks exclusively present in lung samples and absent in brain tissue. Vice versa, the *Rundc3a* brain-specific locus[31] showed promoter peaks only enriched in brain samples (Figure S1E). Peaks obtained from brain and lung samples showed high correlation with their respective ENCODE bulk ATAC-seq references (Figure S1D), further demonstrating the robustness of the obtained profiles of accessibility. We next performed a differential accessibility analysis between lung and brain by comparing samples of 50 cells and 10 BAs separately (adjusted p value [padj] < 0.05, $\log_2$(fold change) threshold = 1). As shown in Figures 2A and 2B, a total of 2,423 peaks were identified as differentially accessible (differentially accessible regions [DARs]) either up- (922 peaks: 738 in BA comparison, 435 in 50 cells comparison) or downregulated (1,501 peaks: 722 in BA comparison, 1,192 in 50 cells comparison) in the lung compared with brain samples. Of note, roughly half of the up- or downregulated DARs were identified both in the 50 cells and the 10 BA comparisons, attesting that LCM-ATAC-seq could capture tissue specificities up to a 50 nuclei resolution. Though cells and BAs were picked up randomly, we cannot exclude slightly different cell-type compositions across samples from the same tissue, which could be responsible for the incomplete overlap between DARs between BA and 50 nuclei samples. To further investigate the regulation of these regions across samples beyond their actual statistical significance, we assessed the accessibility of detected union of DARs mapping in promoters, elements that are strictly involved in the epigenetic regulation of gene expression. As expected, the accessibility of these regions was consistent across samples belonging to the same tissue (Figure 2C), and two distinct clusters of DARs upregulated either in brain (cluster 1) or lung samples (cluster 2) were identified. Further, gene set enrichment analysis (GSEA) based on the genes linked to the identified differentially accessible promoter regions on the two DAR clusters (Figure 2D) confirmed the

Figure 1. LCM-ATAC-seq allows characterizing of tissue-specific chromatin accessibility

(A) Schematic representation of the mouse lung/brain LCM-ATAC-seq workflow.

(B) Overall transcription start site (TSS) enrichment profile of analyzed mouse lung/brain LCM-ATAC-seq samples.

(C) Fragment size distribution (FSD) across analyzed mouse lung/brain samples.

(D) Principal-component analysis (PCA) of the analyzed LCM-ATAC-seq mouse lung/brain dataset.

(E) *Sftpb* locus accessibility across analyzed mouse lung/brain groups (normalized bigwig tracks overlayed condition-wise). Reference bulk ATAC-seq ENCODE bigwig tracks are also reported (mouse lung bulk: ENCSR102NGD dataset, ENCFF435PTI.bigwig; mouse brain bulk: ENCSR310MLB dataset, ENCFF561KNB.bigwig). FF, fresh frozen; BA, big areas (10 dissected areas, 20 cells each, for a total of ∼200 cells).

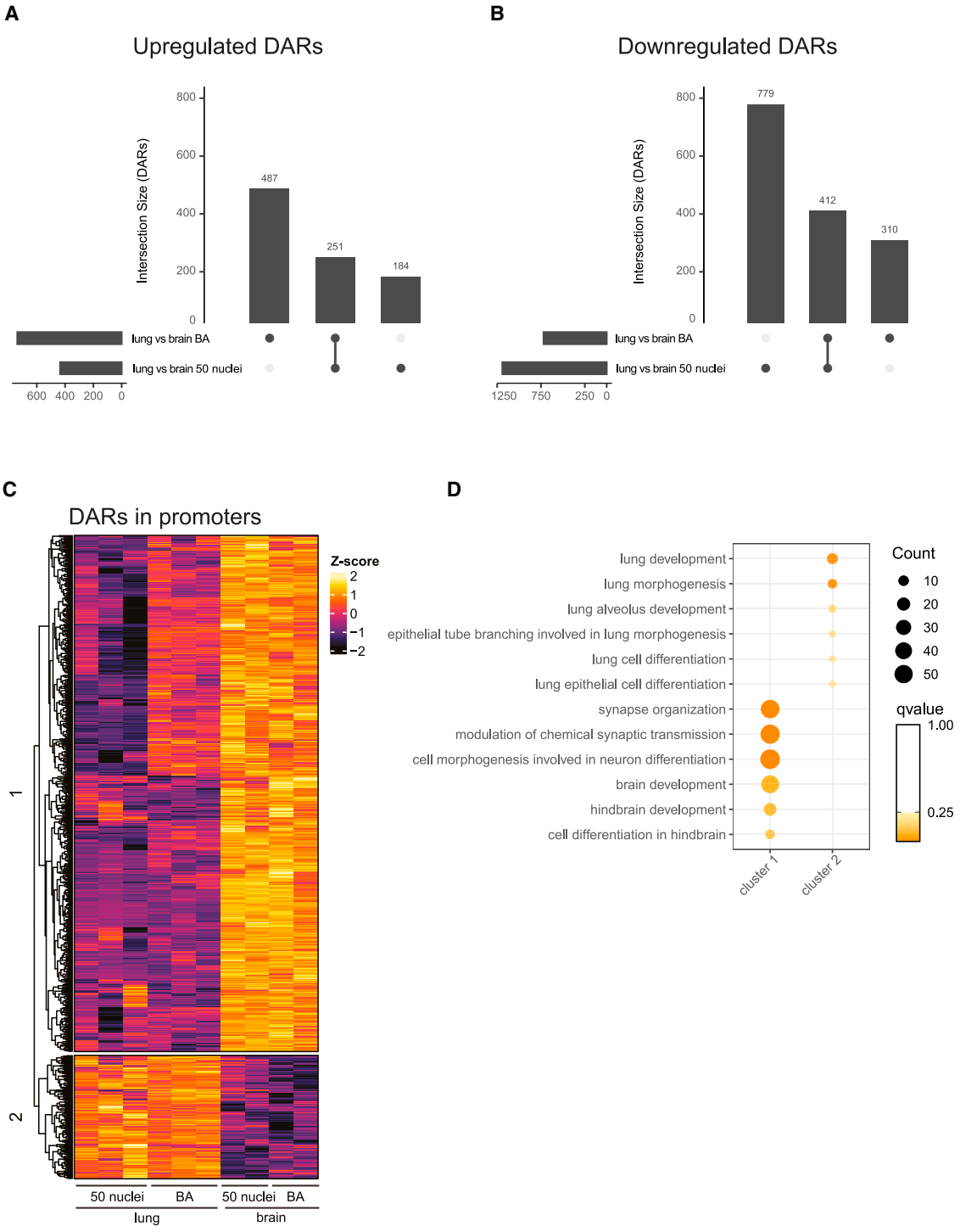


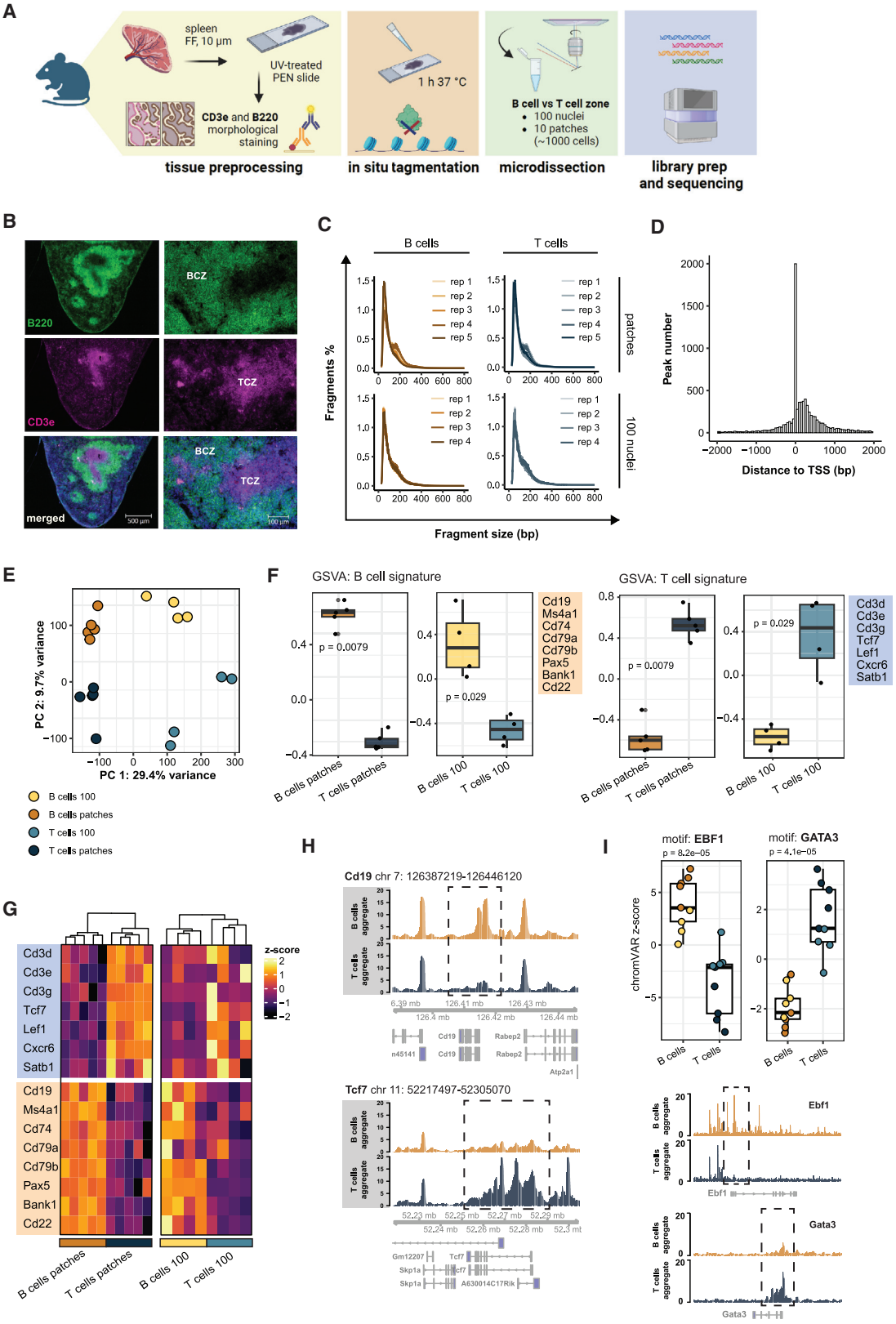
Figure 2. LCM-ATAC-seq profiles describe tissue-specific chromatin landscapes

(A) Number of upregulated differentially accessible regions (DARs) in mouse lung compared with brain samples (padj < 0.05, $\log_2$(fold change) threshold = 1). Separate tissue-wise comparisons were performed for 50 nuclei versus 10 big area (BA) samples.

(B) Number of downregulated DARs in mouse lung compared with brain samples (padj < 0.05, $\log_2$(fold change) threshold = 1). Separate tissue-wise comparisons were performed for 50 nuclei versus 10 BA samples.

(C) Hierarchically clustered heatmap reporting the relative accessibility of the union of all DARs mapping in promoter regions across samples.

(D) Gene set enrichment analysis (GSEA) on GO database (biological process, q < 0.25) enrichment of gene-associated DARs in clusters 1 and 2.



(legend on next page)

enrichment of neuronal biological processes in cluster 1 and, respectively, of lung-related terms in cluster 2.

Overall, here, we demonstrated that LCM-ATAC-seq can be used to retrieve tissue-specific chromatin landscapes with spatial resolution from as few as 50 nuclei.

Investigating the chromatin accessibility of targeted tissue cell populations

We next explored how LCM-ATAC-seq could be combined with cell surface marker staining for the targeted characterization of distinct populations within a tissue (Figure 3A). To this aim, we immunostained fresh frozen mouse spleen sections for B220 (B cell marker) and CD3e (T cell marker) to detect B and T cell zones (BCZs and TCZs) in the spleen (Figure 3B). Paraformaldehyde (PFA) fixation followed by antibody staining was employed as described in STAR Methods. Then, we performed LCM-ATAC-seq from collected B and T cell populations derived from stained regions as mini-bulks of either 10 patches (each of ~100 cells, in total ~1,000 collected cells; Figure S2A) or 100 single captured nuclei, analyzed at least in quadruplicates.

We first checked the QC metrics, which demonstrated again high alignment rates (Figure S2B) and the classical ATAC-seq FSD (Figure 3C). TSS-ESs and FRiP scores (Figure S2C) were in line with the previous dataset (TSS-ESs between 3 and 9; FRiP scores between 10% and 20%) but were generally lower for the 100 nuclei mini-bulks, although they were still in the range of previously published spatial ATAC-seq technologies.[13,16] Beyond pointing toward tissue- and size-specific differences, the higher noise observed in the 100 nuclei samples may also derive here from the introduction of a PFA fixation step (necessary for the immunostaining), which was previously reported to affect chromatin quality.[32] As a next step, we called peaks after merging sequencing data of all samples and identified 27,388 peaks mostly located in promoter regions (Figure S2D). Reads mapping in peak regions mostly mapped again around TSSs, as expected (Figure 3D).

We further inspected the ability of LCM-ATAC-seq to capture cell-type-specific chromatin accessibility profiles in B versus T cell populations. Promisingly, PCA showed that samples clustered according to their zone of isolation (BCZ, TCZ, separating in PC2) and cutting strategy/size (patches, 100 nuclei, separating in PC1) as shown in Figure 3E. We then computed per-gene chromatin accessibility and ran a gene set variation analysis (GSVA), checking for B cell and T cell signature enrichment across samples. Canonical markers were selected to build up such signatures (B cell signature: *Cd19*, *Ms4a1*, *Cd74*, *Cd79a*, *Cd79b*, *Pax5*, *Bank1*, *Cd22*; T cell signature: *Cd3d*, *Cd3e*, *Cd3g*, *Tcf7*, *Lef1*, *Cxcr6*, *Satb1*[33]). Notably, T and B cell signatures were significantly enriched in TCZ- and BCZ-isolated samples, respectively, both at the level of patches and 100 nuclei (Figure 3F). Per-gene accessibility of key signature markers also unveiled the expected phenotypes, with B cell marker accessibility higher in BCZ samples compared with in TCZ samples, and vice versa (Figure 3G). Further endorsing our hypothesis, accessibility at key marker loci showed exclusive peaks in the *Cd19* and *Ms4a1* gene bodies only represented in the B cell aggregates, as well as peaks in the *Tcf7* and *Cd3e* loci only observed in our T cell samples (Figures 3H and S2E). As a last step to confirm B and T cell sample identities, we ran a chromVAR transcription factor motif deviations analysis.[34] As shown in Figure S3A, the top 25 motifs enriched in B and T cells showed similar *Z* score deviations across BCZ versus TCZ samples, with the two groups clustering separately in the heatmap. Among these top 25 motifs, we identified transcription factors known to be exclusively active in B cells, such as EBF1,[35] and T cells, i.e., GATA3.[36] As expected, BCZ samples showed significantly higher EBF1 motif *Z* scores and a more accessible *Ebf1* locus (Figure 3I). Conversely, TCZ samples resulted in significantly higher GATA3 motif *Z* scores and a more accessible *Gata3* locus (Figure 3I).

Overall, this second analysis demonstrated the feasibility to couple LCM-ATAC-seq with a cell surface marker staining for the targeted chromatin profiling of discrete cell populations down to a 100 nuclei mini-bulk resolution.

Assessing the chromatin accessibility of scattered tissue cell populations

After assessing the ability of the technology to characterize targeted but contiguous cell populations within the same tissue, we further investigated the possibility to analyze the chromatin profile of scattered or rare tissue cell populations, now also inspecting the applicability of LCM-ATAC-seq on human material. As a proof of principle, we selected PU.1 as a nuclear immune transcription factor particularly expressed in myeloid cells to target them in the human lung parenchyma.[37] To validate the technology and determine its reproducibility across distinct parenchyma sampling regions and donors, we collected mini-bulks of 100 PU.1-positive (PU.1+) single captured nuclei from two left lobe (LL) locations (L1: lower LL; L2: upper LL) with at least two

Figure 3. Investigating the chromatin accessibility of targeted tissue cell populations

(A) Schematic representation of the mouse spleen LCM-ATAC-seq workflow.

(B) Representative immunofluorescence B220 and CD3e staining of a mouse spleen section (5× magnification left column, 20× magnification right column).

(C) FSD across analyzed samples.

(D) Overall TSS enrichment profile of analyzed mouse spleen LCM-ATAC-seq samples.

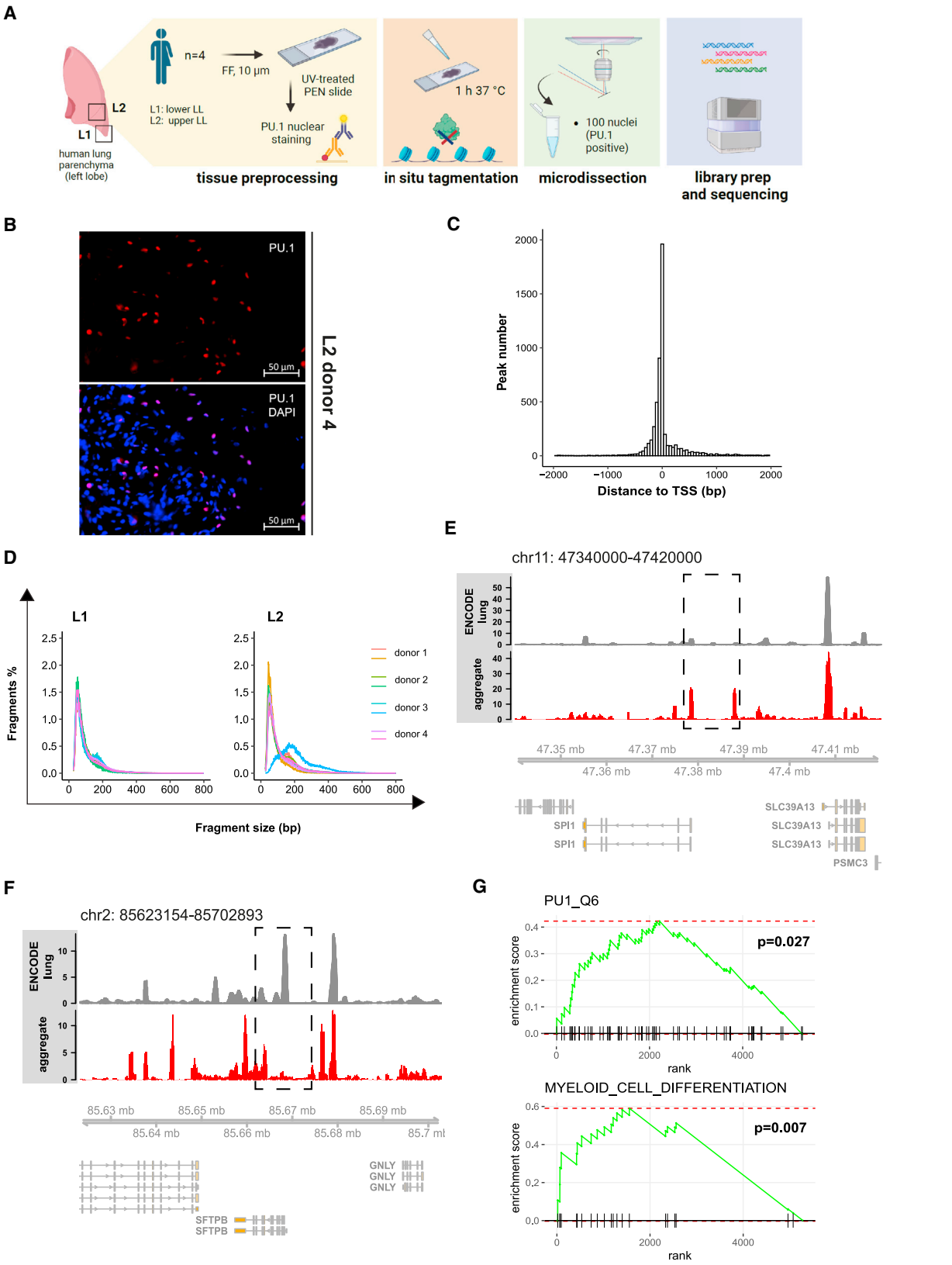
(E) PCA of the analyzed mouse spleen LCM-ATAC-seq dataset.

(F) GSVA enrichment for B (left) and T cell (right) signatures across mouse spleen samples, plotted across groups. Wilcoxon test was applied to test significance, and p values are reported.

(G) Hierarchically clustered (by column) heatmap reporting the per-gene accessibility of analyzed mouse spleen samples for key B and T cell marker genes.

(H) *Cd19* and *Tcf7* loci accessibility in normalized TCZ versus BCZ samples aggregate bigwig tracks.

(I) ChromVAR *Z* scores for B and T cell samples obtained for the motifs EBF1 and GATA3. Wilcoxon test was applied to test significance, and p values are reported. The bottom panel reports *Ebf1* and *Gata3* loci accessibility in normalized TCZ versus BCZ samples aggregate bigwig tracks. FF, fresh frozen; BCZ, B cell zone; TCZ, T cell zone; B cell/T cells 100, 100 nuclei mini-bulk from BCZ or TCZ; B cell/T cells patch, 10 patches mini-bulk from BCZ or TCZ.



(legend on next page)

technical replicates from 4 donors (Figure 3A). A fast PFA fixation followed by immunostaining protocol was employed for PU.1 staining (Figures 3B and S4A; see STAR Methods for details).

In line with previous observations, the QC showed high alignment rates (Figure S4B) and the typical ATAC-seq FSD (Figure 4D) in our data. As observed in the spleen dataset, TSS-ES and FRiP scored generally lower compared with the first unfixed mouse dataset, with values varying across samples (Figure S4C; TSS-ESs between 0.4 and 3; FRiP scores between 5% and 10%), further indicating that the PFA fixation together with the lower quality of the starting postmortem material likely introduce unwanted noise in the data. Nevertheless, the obtained scores are in line with other examples from previously published spatial ATAC-seq technologies.[13,16] The higher noise resulted in a lower number of called peaks (7,305 peaks from merged samples), but a high enrichment in promoter regions was obtained (Figure S4D). Although lower in number, reads in peaks mostly mapped around TSSs (Figure 4C), confirming the robustness of called peak regions.

To determine if PU.1+ nuclei isolation via LCM coupled to ATAC-seq would enrich myeloid cells, we first determined chromatin landscapes at loci of typical myeloid cell markers. Indeed, when comparing it with a reference human lung bulk ATAC-seq profile, we observed higher accessibility in our PU.1+ nuclei aggregate profile at the *SPI1* (PU.1) gene promoter (Figure 4E) and at the promoter of the macrophage marker *CD68*, strongly indicating successful targeting of myeloid cells (Figure S4E). In contrast, when assessing lung-related genes outside the immune compartment, for example *SFTPB*, encoding for the surfactant protein B mainly expressed by alveolar type II (AT2) cells,[38] we observed a loss of specific peaks detected in reference whole-lung tissue (Figure 4F), providing additional evidence for the enrichment of myeloid cells. Further confirming the enrichment for mainly PU.1+ myeloid cells, we report the significant enrichment of the obtained accessibility profiles for signatures including both PU.1 target genes and myeloid cell differentiation genes (Figure 4G), performed via one-tailed GSEA (for details, see STAR Methods).

Collectively, these data demonstrate that LCM-ATAC-seq enables the determination of spatial chromatin landscapes of scattered cell populations based on specific marker genes within a tissue of interest.

DISCUSSION

Spatial omics technologies offer novel tools to tackle tissue biology in homeostasis and disease.[3] The organization of chromatin accessibility represents an important layer of cellular regulation and has been described to foreshadow cellular differentiation and response capacities.[9] However, the analysis of chromatin accessibility patterns in spatial contexts is still underexplored. To expand the selection of tools mapping chromatin states within tissue architectures, we report the use of LCM coupled to ATAC-seq.

We set up an LCM-ATAC-seq protocol and first validated it in mouse tissues exploring the effect of mini-bulk size, microdissection strategy, and tissue type. QC profiles of the analyzed brain and lung samples confirmed the applicability of our approach for the 2D investigation of chromatin states. Notably, LCM-ATAC-seq was demonstrated to efficiently retain tissue-specific chromatin landscapes with a resolution of 50 cells both in brain and lung samples, with only partial loss in the overall quality of the samples.

Further on, we showed how to combine LCM-ATAC-seq with cell surface staining for the targeted characterization of distinct contiguous populations within the same tissue. To this aim, we immunostained fresh frozen mouse spleen sections for B220 (B cell marker) and CD3e (T cell marker) to detect B and T cell zones in the spleen. Results demonstrated the ability of LCM-ATAC-seq to efficiently isolate these discrete subpopulations by retaining their cell-type-specific epigenetic landscape. Nevertheless, QC metrics evidenced the lower quality of PFA-fixed immunostained samples, especially when downsizing to 100 nuclei samples, as already described.[32]

As a step forward, we demonstrated the ability of our protocol to analyze scattered single-nuclei captured populations in human lung parenchyma samples (n = 4 donors, upper and lower LL each). PU.1 was selected as a myeloid cell marker to test this hypothesis: tissues were first immunostained for PU.1, and only marker-positive nuclei were collected and processed via LCM-ATAC-seq. Of note, results demonstrated the ability of our approach to isolate highly enriched PU.1^{+} populations and to profile their chromatin accessibility. Nevertheless, PFA fixation was shown here as well to have an impact on the overall quality of the analyzed chromatin. This suggests that alternative methods for sample fixation and subsequent immunostaining might be worth exploring in future research.

Overall, this study reported a complete workflow for the efficient use of LCM combined with ATAC-seq, describing the potential of such a technology applied to fresh frozen samples, of particular interest for the targeted analysis of chromatin accessibility of relatively rare or scattered cell populations within tissues. As we showcased the use of LCM-ATAC-seq on

Figure 4. Assessing the chromatin accessibility of scattered tissue cell populations

(A) Schematic representation of the LCM-ATAC-seq workflow in human left lobe (LL) lung parenchyma samples. L1: lower LL; L2: upper LL.

(B) Representative immunofluorescence PU.1 staining of an L2 section from donor 4 (40× magnification).

(C) Overall TSS enrichment profile of analyzed human LCM-ATAC-seq samples.

(D) FSD across analyzed human samples.

(E) *SPI1* locus accessibility in normalized PU.1 aggregate track versus lung bulk ENCODE reference (human lung bulk: ENCSR647AOY dataset, ENCFF210HIS.bigwig).

(F) *SFTPB* locus accessibility in normalized PU.1 aggregate bigwig track versus lung bulk ENCODE reference (human lung bulk: ENCSR647AOY dataset, ENCFF210HIS.bigwig).

(G) One-tailed fast gene set enrichment analysis for signatures PU1_Q6 (PU.1 transcription factor targets) and GO_BP_MYELOID_CELL_DIFFERENTIATION on PU.1 targeted cells rank of chromatin accessibility. FF, fresh frozen.

different tissues and inspecting different targeting approaches, we further envision its versatile application also on different settings (e.g., tumor biology, neurobiology). Furthermore, the technique may allow labs with access to the NanoString GeoMx technology to easily complement RNA sequencing (RNA-seq) studies with chromatin accessibility in adjacent regions for the multi-omics characterization of tissues of interest.

Limitations of the study

The QC metrics computed across the different datasets suggest that LCM-ATAC-seq could not fulfill the gold-standard criteria of bulk ATAC-seq analysis but that it is in line with examples reported for other spatial ATAC-seq technologies.[13,16] Reasons for this could be (1) the generally lower number of cells analyzed per sample compared with in a bulk experiment, which is why we refer here to mini-bulks; (2) the overall protocol length, limited throughput, and discreet laboriousness compared with in a classical bulk experiment, especially when combined with cell surface staining requiring PFA fixation, as reported[32]; and (3) the variable performance of LCM-ATAC-seq based on intrinsic tissue properties and preservation. Taken together, these points set some limitations in the applicability of our technology and may be reasons why (1) single-nucleus resolution could not be reached and (2) a generally lower number of peaks were called per dataset due to background noise. As a general remark, sparsity is also common in single-cell (sc)ATAC-seq data, as reported by Chen et al.,[39] where peak detection rate per cell varies between 5% and 10% on average across different technologies: this is probably a more reasonable reference to our 50 and 100 nuclei mini-bulks, as already inspected in other spatial ATAC-seq studies.[13,16] In future work, it might be worth performing comparative studies. Nevertheless, the distribution of reads in peaks around TSSs in all datasets as well as the downstream biological evaluation suggest the robustness of the obtained peak sets in the different inspected scenarios.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANTS DETAILS
- METHOD DETAILS
  - Tissues preprocessing and *in situ* tagmentation
  - B220 and CD3e immunofluorescent staining of mouse spleen tissues
  - PU.1 immunofluorescent staining of human lung tissues
  - Laser-capture microdissection and sample collection
  - Library preparation and sequencing
  - Bioinformatic analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.crmeth.2023.100598.

ACKNOWLEDGMENTS

We thank Charlotte Kröger, Kevin Baβler, and Theodore S. Kapellos for the helpful discussion during the development of the technology. We also thank Frank Bradke for kindly providing us with the secondary antibodies used in the LCM-ATAC-seq mouse spleen experiment. Figures 1A, 3A, and 4A and the graphical abstract were created with BioRender.com. The work was supported by the European Union's H2020 Research and Innovation Program under grant agreement no. 874656 (discovAIR); the ImmunoSensation2 Bonn Cluster of Excellence; Bundesministerium für Bildung und Forschung (BMBF-funded excellence project Diet-Body-Brain [DietBB, 01EA1809A]) to J.L.S.; the Horizon 2020 Framework Programme (EU project SYSCID under grant number 733100) to J.L.S.; Deutsche Forschungsgemeinschaft (German Research Foundation [DFG] CRC SFB 1454 Metaflammation - project number 432325352) to J.L.S. and A.C.A.; BMBF-funded grant iTREAT (01ZX1902B) to J.L.S.; Deutsche Forschungsgemeinschaft (German Research Foundation [DFG] ImmuDiet BO 6228/2-1 - project number 513977171) to L.B.; the German Research Foundation (IRTG2168 272482170, SFB1454-432325352, and EXC2151-390873048) to M.B.; RIA HORIZON Research and Innovation Actions HORIZON-HLTH-2021-DISEASE-04-07 grant no. 101057775 (NEUROCOV) to M.B., J.L.S., and A.C.A.; and the Else Kroner-Fresenius Foundation (2018_A158) to M.B.

AUTHOR CONTRIBUTIONS

Conceptualization, C.C., A.S., J.L.S., M.B., and E.D.D.; methodology, C.C., A.S., I.K., H.-U.F., and E.F.; validation, C.C., A.S., R.S., V.I., and P.G.; formal analysis, C.C., L.B., A.S., R.S., V.I., M.v.U., J.S.-S., T.E., J.V.M., H.H., Y.L., H.T., and M.K.; resources, K.T.M.; original draft preparation, C.C.; review and editing, all authors; supervision, J.L.S., A.S., M.B., and E.D.D.; project administration, M.v.U.; funding acquisition, J.L.S., A.C.A., L.B., and M.B. All authors have read and agreed to the published version of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: March 27, 2023
Revised: August 4, 2023
Accepted: September 5, 2023
Published: September 29, 2023

REFERENCES

1. Rood, J.E., Maartens, A., Hupalowska, A., Teichmann, S.A., and Regev, A. (2022). Impact of the Human Cell Atlas on medicine. Nat. Med. *28*, 2486–2496. https://doi.org/10.1038/s41591-022-02104-7.
2. Deprez, M., Zaragosi, L.-E., Truchi, M., Becavin, C., Ruiz García, S., Arguel, M.-J., Plaisant, M., Magnone, V., Lebrigand, K., Abelanet, S., et al. (2020). A Single-Cell Atlas of the Human Healthy Airways. Am. J. Respir. Crit. Care Med. *202*, 1636–1645. https://doi.org/10.1164/rccm.201911-2199OC.
3. Moses, L., and Pachter, L. (2022). Museum of spatial transcriptomics. Nat. Methods *19*, 534–546. https://doi.org/10.1038/s41592-022-01409-2.

4. Rao, A., Barkley, D., França, G.S., and Yanai, I. (2021). Exploring tissue architecture using spatial transcriptomics. Nature *596*, 211–220. https://doi.org/10.1038/s41586-021-03634-9.

5. Madissoon, E., Oliver, A.J., Kleshchevnikov, V., Wilbrey-Clark, A., Polanski, K., Richoz, N., Ribeiro Orsi, A., Mamanova, L., Bolt, L., Elmentaite, R., et al. (2023). A spatially resolved atlas of the human lung characterizes a gland-associated immune niche. Nat. Genet. *55*, 66–77. https://doi.org/10.1038/s41588-022-01243-4.

6. Tian, L., Chen, F., and Macosko, E.Z. (2023). The expanding vistas of spatial transcriptomics. Nat. Biotechnol. *41*, 773–782. https://doi.org/10.1038/s41587-022-01448-2.

7. Gibney, E.R., and Nolan, C.M. (2010). Epigenetics and gene expression. Heredity *105*, 4–13. https://doi.org/10.1038/hdy.2010.54.

8. Atlasi, Y., and Stunnenberg, H.G. (2017). The interplay of epigenetic marks during stem cell differentiation and development. Nat. Rev. Genet. *18*, 643–658. https://doi.org/10.1038/nrg.2017.57.

9. Klemm, S.L., Shipony, Z., and Greenleaf, W.J. (2019). Chromatin accessibility and the regulatory epigenome. Nat. Rev. Genet. *20*, 207–220. https://doi.org/10.1038/s41576-018-0089-8.

10. Minnoye, L., Marinov, G.K., Krausgruber, T., Pan, L., Marand, A.P., Secchia, S., Greenleaf, W.J., Furlong, E.E.M., Zhao, K., Schmitz, R.J., et al. (2021). Chromatin accessibility profiling methods. Nat. Rev. Methods Primers *1*, 10. https://doi.org/10.1038/s43586-020-00008-9.

11. Buenrostro, J.D., Wu, B., Litzenburger, U.M., Ruff, D., Gonzales, M.L., Snyder, M.P., Chang, H.Y., and Greenleaf, W.J. (2015). Single-cell chromatin accessibility reveals principles of regulatory variation. Nature *523*, 486–490. https://doi.org/10.1038/nature14590.

12. Thornton, C.A., Mulqueen, R.M., Torkenczy, K.A., Nishida, A., Lowenstein, E.G., Fields, A.J., Steemers, F.J., Zhang, W., McConnell, H.L., Woltjer, R.L., et al. (2021). Spatially mapped single-cell chromatin accessibility. Nat. Commun. *12*, 1274. https://doi.org/10.1038/s41467-021-21515-7.

13. Deng, Y., Bartosovic, M., Ma, S., Zhang, D., Kukanja, P., Xiao, Y., Su, G., Liu, Y., Qin, X., Rosoklija, G.B., et al. (2022). Spatial profiling of chromatin accessibility in mouse and human tissues. Nature *609*, 375–383. https://doi.org/10.1038/s41586-022-05094-1.

14. Chen, X., Shen, Y., Draper, W., Buenrostro, J.D., Litzenburger, U., Cho, S.W., Satpathy, A.T., Carter, A.C., Ghosh, R.P., East-Seletsky, A., et al. (2016). ATAC-see reveals the accessible genome by transposase-mediated imaging and sequencing. Nat. Methods *13*, 1013–1020. https://doi.org/10.1038/nmeth.4031.

15. Zhang, D., Deng, Y., Kukanja, P., Agirre, E., Bartosovic, M., Dong, M., Ma, C., Ma, S., Su, G., Bao, S., et al. (2023). Spatial epigenome-transcriptome co-profiling of mammalian tissues. Nature *616*, 113–122. https://doi.org/10.1038/s41586-023-05795-1.

16. Llorens-Bobadilla, E., Zamboni, M., Marklund, M., Bhalla, N., Chen, X., Hartman, J., Frisén, J., and Ståhl, P.L. (2023). Solid-phase capture and profiling of open chromatin by spatial ATAC. Nat. Biotechnol. *41*, 1085–1088. https://doi.org/10.1038/s41587-022-01603-9.

17. Ståhl, P.L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J.F., Magnusson, J., Giacomello, S., Asp, M., Westholm, J.O., Huss, M., et al. (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science *353*, 78–82. https://doi.org/10.1126/science.aaf2403.

18. Lu, T., Ang, C.E., and Zhuang, X. (2022). Spatially resolved epigenomic profiling of single cells in complex tissues. Cell *185*, 4448–4464.e17. https://doi.org/10.1016/j.cell.2022.09.035.

19. Espina, V., Wulfkuhle, J.D., Calvert, V.S., VanMeter, A., Zhou, W., Coukos, G., Geho, D.H., Petricoin, E.F., and Liotta, L.A. (2006). Laser-capture microdissection. Nat. Protoc. *1*, 586–603. https://doi.org/10.1038/nprot.2006.85.

20. Datta, S., Malhotra, L., Dickerson, R., Chaffee, S., Sen, C.K., and Roy, S. (2015). Laser capture microdissection: Big data from small samples. Histol. Histopathol. *30*, 1255–1269. https://doi.org/10.14670/HH-11-622.

21. Nichterwitz, S., Chen, G., Aguila Benitez, J., Yilmaz, M., Storvall, H., Cao, M., Sandberg, R., Deng, Q., and Hedlund, E. (2016). Laser capture microscopy coupled with Smart-seq2 for precise spatial transcriptomic profiling. Nat. Commun. *7*, 12139. https://doi.org/10.1038/ncomms12139.

22. Chen, J., Suo, S., Tam, P.P., Han, J.-D.J., Peng, G., and Jing, N. (2017). Spatial transcriptomic analysis of cryosectioned tissue samples with Geo-seq. Nat. Protoc. *12*, 566–580. https://doi.org/10.1038/nprot.2017.003.

23. Herrera, J.A., Mallikarjun, V., Rosini, S., Montero, M.A., Lawless, C., Warwood, S., O'Cualain, R., Knight, D., Schwartz, M.A., and Swift, J. (2020). Laser capture microdissection coupled mass spectrometry (LCM-MS) for spatially resolved analysis of formalin-fixed and stained human lung tissues. Clin. Proteonomics *17*, 24. https://doi.org/10.1186/s12014-020-09287-6.

24. Moulédous, L., Hunt, S., Harcourt, R., Harry, J.L., Williams, K.L., and Gutstein, H.B. (2003). Proteomic analysis of immunostained, laser-capture microdissected brain samples. Electrophoresis *24*, 296–302. https://doi.org/10.1002/elps.200390026.

25. Shen, S., Li, J., Huo, S., Ma, M., Zhu, X., Rasam, S., Duan, X., Qu, M., Titus, M.A., and Qu, J. (2021). Parallel, High-Quality Proteomic and Targeted Metabolomic Quantification Using Laser Capture Microdissected Tissues. Anal. Chem. *93*, 8711–8718. https://doi.org/10.1021/acs.analchem.1c01026.

26. Knittelfelder, O., Traikov, S., Vvedenskaya, O., Schuhmann, A., Segeletz, S., Shevchenko, A., and Shevchenko, A. (2018). Shotgun Lipidomics Combined with Laser Capture Microdissection: A Tool To Analyze Histological Zones in Cryosections of Tissues. Anal. Chem. *90*, 9868–9878. https://doi.org/10.1021/acs.analchem.8b02004.

27. Schillebeeckx, M., Schrade, A., Löbs, A.K., Pihlajoki, M., Wilson, D.B., and Mitra, R.D. (2013). Laser capture microdissection-reduced representation bisulfite sequencing (LCM-RRBS) maps changes in DNA methylation associated with gonadectomy-induced adrenocortical neoplasia in the mouse. Nucleic Acids Res. *41*, e116. https://doi.org/10.1093/nar/gkt230.

28. 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. https://doi.org/10.1186/s13059-020-1929-3.

29. Hitz, B.C., Lee, J.-W., Jolanki, O., Kagda, M.S., Graham, K., Sud, P., Gabdank, I., Seth Strattan, J., Sloan, C.A., Dreszer, T., et al. (2023). The ENCODE uniform analysis pipelines. Preprint at bioRxiv. https://doi.org/10.1101/2023.04.04.535623.

30. Hawgood, S., and Poulain, F.R. (1995). Functions of the surfactant proteins: a perspective. Pediatr. Pulmonol. *19*, 99–104. https://doi.org/10.1002/ppul.1950190205.

31. Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., et al. (2015). Tissue-based map of the human proteome. Science *347*, 1260419. https://doi.org/10.1126/science.1260419.

32. Zhang, H., Rice, M.E., Alvin, J.W., Farrera-Gaffney, D., Galligan, J.J., Johnson, M.D.L., and Cusanovich, D.A. (2022). Extensive evaluation of ATAC-seq protocols for native or formaldehyde-fixed nuclei. BMC Genom. *23*, 214. https://doi.org/10.1186/s12864-021-08266-x.

33. Beyer, M., Thabet, Y., Müller, R.U., Sadlon, T., Classen, S., Lahl, K., Basu, S., Zhou, X., Bailey-Bucktrout, S.L., Krebs, W., et al. (2011). Repression of the genome organizer SATB1 in regulatory T cells is required for suppressive function and inhibition of effector differentiation. Nat. Immunol. *12*, 898–907. https://doi.org/10.1038/ni.2084.

34. Schep, A.N., Wu, B., Buenrostro, J.D., and Greenleaf, W.J. (2017). chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. Nat. Methods *14*, 975–978. https://doi.org/10.1038/nmeth.4401.

35. Nechanitzky, R., Akbas, D., Scherer, S., Györy, I., Hoyler, T., Ramamoorthy, S., Diefenbach, A., and Grosschedl, R. (2013). Transcription factor EBF1 is essential for the maintenance of B cell identity and prevention of alternative fates in committed cells. Nat. Immunol. *14*, 867–875. https://doi.org/10.1038/ni.2641.

36. Ho, I.-C., Tai, T.-S., and Pai, S.-Y. (2009). GATA3 and the T-cell lineage: essential functions before and after T-helper-2-cell differentiation. Nat. Rev. Immunol. *9*, 125–135. https://doi.org/10.1038/nri2476.

37. Heng, T.S.P., and Painter, M.W.; Immunological Genome Project Consortium (2008). The Immunological Genome Project: networks of gene expression in immune cells. Nat. Immunol. *9*, 1091–1094. https://doi.org/10.1038/ni1008-1091.

38. Wang, Y., Tang, Z., Huang, H., Li, J., Wang, Z., Yu, Y., Zhang, C., Li, J., Dai, H., Wang, F., et al. (2018). Pulmonary alveolar type I cell population consists of two distinct subtypes that differ in cell fate. Proc. Natl. Acad. Sci. USA *115*, 2407–2412. https://doi.org/10.1073/pnas.1719474115.

39. Chen, H., Lareau, C., Andreani, T., Vinyard, M.E., Garcia, S.P., Clement, K., Andrade-Navarro, M.A., Buenrostro, J.D., and Pinello, L. (2019). Assessment of computational methods for the analysis of single-cell ATAC-seq data. Genome Biol. *20*, 241. https://doi.org/10.1186/s13059-019-1854-5.

40. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359. https://doi.org/10.1038/nmeth.1923.

41. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

42. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

43. Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res. *42*, W187–W191. https://doi.org/10.1093/nar/gku365.

44. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137. https://doi.org/10.1186/gb-2008-9-9-r137.

45. Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T., and Carey, V.J. (2013). Software for computing and annotating genomic ranges. PLoS Comput. Biol. *9*, e1003118. https://doi.org/10.1371/journal.pcbi.1003118.

46. Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics *32*, 3047–3048. https://doi.org/10.1093/bioinformatics/btw354.

47. DeBerardine, M. (2023). BRGenomics for analyzing high-resolution genomics data in R. Bioinformatics *39*, btad331. https://doi.org/10.1093/bioinformatics/btad331.

48. Conway, J.R., Lex, A., and Gehlenborg, N. (2017). UpSetR: an R package for the visualization of intersecting sets and their properties. Bioinformatics *33*, 2938–2940. https://doi.org/10.1093/bioinformatics/btx364.

49. Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics *32*, 2847–2849. https://doi.org/10.1093/bioinformatics/btw313.

50. Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinf. *14*, 7. https://doi.org/10.1186/1471-2105-14-7.

51. Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M.N., and Sergushichev, A. (2016). Fast gene set enrichment analysis. Preprint at bioRxiv. https://doi.org/10.1101/060012.

52. Ou, J., Liu, H., Yu, J., Kelliher, M.A., Castilla, L.H., Lawson, N.D., and Zhu, L.J. (2018). ATACseqQC: a Bioconductor package for post-alignment quality assessment of ATAC-seq data. BMC Genom. *19*, 169. https://doi.org/10.1186/s12864-018-4559-3.

53. Wickham, H. (2016). ggplot2 - Elegant Graphics for Data Analysis (Springer-Verlag New York). https://doi.org/10.1007/978-0-387-98141-3.

54. Luecken, M.D., Zaragosi, L.-E., Madissoon, E., Sikkema, L., Firsova, A.B., De Domenico, E., Kümmerle, L., Saglam, A., Berg, M., Gay, A.C.A., et al. (2022). The discovAIR project: a roadmap towards the Human Lung Cell Atlas. Eur. Respir. J. *60*, 2102057. https://doi.org/10.1183/13993003.02057-2021.

55. Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., et al. (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. *5*, R80. https://doi.org/10.1186/gb-2004-5-10-r80.

56. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.

CellPress
OPEN ACCESS

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Hamster monoclonal anti-mouse CD3e	BD Biosciences	Cat#553058; RRID: AB_394591
Rat monoclonal anti-mouse B220	BD Biosciences	Cat#550286; RRID: AB_393581
Goat anti-Rat IgG (H + L) Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 488	ThermoFischer	Cat#A-11006; RRID: AB_2534074
Goat anti-Syrian Hamster IgG (H + L) Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 647	ThermoFischer	Cat#A-21451; RRID: AB_2535868
Goat anti-Syrian Hamster IgG (H + L) Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 488	ThermoFischer	Cat#A-21110; RRID: AB_2535759
Goat anti-Rat IgG (H + L) Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 568	ThermoFischer	Cat#A-11077; RRID: AB_2534121
Rabbit monoclonal anti-human PU.1	ThermoFischer	Cat#MA5-15064; RRID: AB_10986949
Goat anti-Rat IgG (H + L) Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 488	ThermoFischer	Cat#A-11008; RRID: AB_143165
Biological samples		
Healthy adult lung parenchyma tissue	Cambridge Biorepository for Translational Medicine (CBTM)	Cat#15/EE/0152
Chemicals, peptides, and recombinant proteins		
Recombinant Tn5	In-house	
Digitonin	ProMega	G944A
Normal goat serum	MP Biomedicals	08642921
NEBNext HiFi PCR Master Mix	NewEngland Biosciences	M0541S
AMPure XP Bead-Based Reagent	Beckman Coulter	A63881
DAPI solution	Invitrogen	D1306
Critical commercial assays		
Qubit HS dsDNA assay	Invitrogen	Q32851
NovaSeq 6000 S1 Reagent Kit v1.5 (100 cycles)	Illumina	20028319
NovaSeq 6000 SP Reagent Kit v1.5 (100 cycles)	Illumina	20028401
Deposited data		
Mouse dataset	This paper	GEOseries: GSE237825
Human dataset	This paper	EGA: EGAS00001006167
Mouse lung bulk ATAC-seq reference	ENCODE	ENCSR102NGD dataset, ENCFF435PTI.bigwig
Mouse brain bulk ATAC-seq reference	ENCODE	ENCSR310MLB dataset, ENCFF561KNB.bigwig
Human lung bulk ATAC-seq reference	ENCODE	ENCSR647AOY dataset, ENCFF210HIS.bigwig
Mouse reference genome NCBI build 38, GRCm38	Genome Reference Consortium	https://www.ncbi.nlm.nih.gov/grc/mouse/
Human reference genome NCBI build 38, GRCh38	Genome Reference Consortium	http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/human/
Experimental models: Organisms/strains		
Mouse: C57BL/6J	The Jackson Laboratory	JAX: 000664
Oligonucleotides		
IDT DNA/RNA unique dual (UD) indexes, set A	IDT for Illumina	20026121
Software and algorithms		
R		https://www.r-project.org/
bcl2fastq2	Illumina	
Bowtie2	Langmead et al.[40]	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
samtools	Li et al.[41]	http://samtools.sourceforge.net/
Trimmomatic	Bolger et al.[42]	
Picard		http://broadinstitute.github.io/picard
deeptools	Ramírez et al.[43]	N/A
samtools	Ernst et al.[41]	N/A
MACS2	Zhang et al.[44]	N/A
GenomicRanges	Lawrence et al.[45]	N/A
MultiQC	Ewels et al.[46]	N/A
ChIPseeker	Bioconductor	N/A
BRGenomics	DeBerardine[47]	N/A
Gviz	Bioconductor	N/A
DESeq2	Bioconductor	N/A
UpSetR<	Conway et al.[48]	N/A
ComplexHeatmap	Gu et al.,[49] Bioconductor	N/A
GSVA	Hänzelmann et al.[50]	N/A
Fgsea	Korotkevich et al.[51]	N/A
ATACseqQC	Ou et al.[52]	N/A
Ggpubr	Kassambara	https://rpkgs.datanovia.com/ggpubr/
ChromVAR	Schep et al.[34]	N/A
ggplot2	Wickham[53]	N/A
Other		
R scripts to reproduce the analysis	This paper	Zenodo: https://doi.org/10.5281/zenodo.8269143
Mouse raw data	This paper	GEO: SuperSeries GSE237825 (brain and lung: GSE218216, spleen: GSE237824)
Human raw data	This paper	EGA: EGAS00001006167

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Caterina Carraro (caterina.carraro@dzne.de).

Materials availability

This study did not generate unique reagents.

Data and code availability

- For the mouse datasets, raw data have been deposited at gene expression omnibus (GEO) and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. Human LCM-ATAC-seq raw data have been deposited on the European Genome-phenome Archive (EGA) and are available upon reasonable request to the Data Access Committee (DAC). Accession numbers are listed in the key resources table. Only samples derived from left lobe parenchyma L1 and L6 (referred to as L2 in this paper) locations are reported in this study. ENCODE bulk ATAC-seq reference bigwig files were downloaded from https://www.encodeproject.org/(mouse lung bulk ENCODE: ENCSR102NGD, ENCFF435PTI.bigwig; mouse brain bulk ENCODE: ENCSR310MLB, ENCFF561KNB.bigwig; human lung bulk ENCODE: ENCSR647AOY, ENCFF210HIS.-bigwig).
- All the original code to reproduce the analyses reported in this manuscript has been deposited on GitHub at the following link https://github.com/ccarraro/LCM-ATAC-seq and is publicly available as the date of publication. An archival DOI to access scripts and R environments to reproduce the analyses is listed in the key resources table.
- Any other information required to reproduce the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANTS DETAILS

All mice were housed under standard SPF conditions with a 12-h dark/light cycle and food and water *ad libitum*. All experiments were carried out in agreement with the responsible Institutional Animal Care and Use Committee (North Rhine Westphalia). Mouse lung and brain tissues were collected from 6-month-old C57BL/6J male mice. Mouse spleen was collected from a 9-month-old C57BL/6J male animal. Human lung tissues were collected within the HCA DiscovAIR initiative from deceased transplant organ donors (donors 1, 2, 4: males; donor 3 female) by the Cambridge Biorepository for Translational Medicine (CBTM) with informed consent from the donor families and approval from the NRES Committee of East of England – Cambridge South (15/EE/0152).[5,54]

METHOD DETAILS

Tissues preprocessing and *in situ* tagmentation

Mice were first anesthetized with isoflurane. After the depth of anesthesia had been confirmed by testing for the toe pinch withdrawal reflex, the mice were sacrificed by cervical dislocation. The lung lower right lobe, cerebrum and spleen were dissected and briefly washed in ice-cold PBS. The tissues were then transferred into 10 mm × 10 mm cryomolds filled with tissue cryo-preservation medium (VWR International) at room temperature and quickly placed in pre-cooled isopentane on dry ice until the cryo-preservation medium turned completely white (ca. 2 min). The fresh frozen tissues were stored at −80°C until use and analyzed at least in technical duplicates. As concerns human samples, we analyzed fresh frozen human lung parenchyma specimens from n = 4 donors and two left lobe locations (L1: lower LL; L2: upper LL) at least in technical duplicates. Fresh frozen OCT-embedded tissue blocks were long-term stored at −80°C.

For all samples, 10 μm sections were collected on Polyethylene Naphthalate (PEN) membrane slides (Carl Zeiss, 415190-9041-000). The PEN membrane slides were previously treated in a UV crosslinker (UVP, CL-1000) with 254 nm wavelength light for 30 min with 200 mJ/cm^2 energy exposure (UV-activated PEN slides) in order to increase the adherence of the tissue section to the PEN membrane. The UV-activated PEN slides were immediately processed or stored for no longer than one month at −80°C in sealed containers.

Frozen slides were warmed up to room temperature for 1 min before proceeding. Whenever additional antibody staining was necessary, tissues were fixed in PFA 2% (B220, CD3e) or 1% (PU.1) (Sigma-Merck) for 10 min and washed for 5 min with PBS. The excess PBS was removed with KimTech tissues and a hydrophobic contour was drawn around the tissue sections using a PAP pen (Vector Labs). *In situ* tagmentation was achieved by applying 50 μL per section of in-house Tn5 diluted in transposition buffer and incubating for 1 h at 37°C in a humidity chamber. The transposition buffer consisted of THS-seq buffer (33 mM Tris-Acetate pH 7.8, 66 mM Potassium-Acetate, 10 mM Magnesium-Acetate, DMF 16%) and digitonin 0.01% (Promega). The Tn5 enzyme was stored at −20°C and kept on ice while in use. Sections were washed three times with PBS and either first stained (human dataset) or directly LCM processed (mouse datasets).

B220 and CD3e immunofluorescent staining of mouse spleen tissues

Mouse spleen tissue sections were first incubated with 100 μL blocking buffer consisting of 5% normal goat serum (08642921, MP biomedicals) diluted 1:2 in PBST (PBS with 0.4% Triton X-100, Sigma-Merck) for 30 min at room temperature. Samples were then incubated with a 100 μL mix of primary antibodies anti-CD3e (145-2C11 monoclonal hamster IgG anti-mouse CD3e, BD Biosciences) and anti-B220 (RA3-6B2 monoclonal rat IgG anti-mouse CD45R (B220), BD Biosciences) diluted 1:25 in PBST with 5% goat serum 2.5% BSA (ThermoFischer) for 45 min at room temperature. Afterward, sections were washed three times with PBST. Following, 100 μL of secondary antibodies mix diluted 1:25 (CD3e: Goat anti-Syrian Hamster IgG (H + L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 488, Invitrogen; B220: Goat anti-Rat IgG (H + L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 568) in PBS with 2.5% BSA were added and slides were incubated for 30 min at room temperature. To reduce the incubation time, in line with previously reported protocols for short staining,[21] we increased the antibody concentrations compared to the one suggested by the vendor (1:200 to 1:25). Sections were then washed three times with PBST in the dark, stained 5 min with 100 μL 4′,6-diamidino-2-phenylindole (DAPI) 1 μM solution (Invitrogen) at room temperature in the dark and washed three times with PBS. The last PBS layer was kept on the slide before moving to LCM. Instead, representative staining images in Figure 3B were acquired using an epifluorescence microscope employing a different secondary antibodies mix diluted 1:50 (CD3e: Goat anti-Syrian Hamster IgG (H + L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 647, Invitrogen; B220: Goat anti-Rat IgG (H + L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 488, Invitrogen) at 5X or 20X magnification as reported in the legend (Zeiss).

PU.1 immunofluorescent staining of human lung tissues

Human lung tissue sections were first incubated with 100 μL blocking buffer consisting of 2% normal goat serum (08642921, MP biomedicals) diluted 1:5 in PBST for 45 min at room temperature. Samples were then incubated with 100 μL of primary antibody anti-PU.1 (E.388.3 monoclonal rabbit IgG anti-human PU.1, ThermoFischer) diluted 1:100 in PBS with 1% BSA (ThermoFischer) for 15 min at room temperature. To reduce the incubation time, in line with previously reported protocols for short staining,[21] we 10-fold increased the antibody concentration compared to the one suggested by the vendor. Afterward, sections were washed three times with PBST. Following, 100 μL of secondary antibody 1:250 (Goat anti-Rabbit IgG (H + L) Cross-Adsorbed Secondary Antibody,

Alexa Fluor 488, Invitrogen) in PBS with 1% BSA were added and slides were incubated for 15 min at room temperature. Sections were then washed three times with 0.4% PBST in the dark, stained 5 min with 100 μL 4′,6-diamidino-2-phenylindole (DAPI) 1 μM solution (Invitrogen) at room temperature in the dark and washed three times with PBS. The last PBS layer was kept on the slide before moving to LCM. Instead, representative staining images in Figure 2B were obtained employing a different secondary antibody diluted 1:250 (Goat anti-Rabbit IgG (H + L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 555, Invitrogen) to minimize the tissue autofluorescence signal and acquired using an epifluorescence microscope at 40X magnification (Zeiss).

Laser-capture microdissection and sample collection

Slides were further imaged and processed using a UV-based laser microdissection microscope (PALM MicroBeam, Zeiss). Regions or nuclei of interest were manually selected at 5X or 40X magnification using the PALMRobo software (Zeiss, v 4.6) based on either DAPI (mouse data) or either B220, CD3e, PU.1 signal (human lung data). Upon selection, PBS was removed from the top of each section and replaced with 80% ethanol which was left to dry for no more than 30 min. Afterward, selected areas were cut and collected in PCR tube caps filled with 4.5 μL of collection buffer using a SingleCap Collector 200 (RoboMover PALM equipment, Zeiss). Collection buffer was composed by 2.5 μL of reverse crosslinking buffer (50 mM Tris-HCl pH 8, 50 mM NaCl, 20 μg/mL Proteinase K (Sigma-Merck), 0.2% SDS) and 2 μL of 1:5 unique dual indexes mix (IDT for Illumina). Cutting parameters (i.e., cutting and catapulting energy) were finely tuned before each session to exclude tissue damage but ensure efficient collection. After collection in tube caps, samples were spinned down at 1000g for 30s and placed on a thermal cycler at 65°C for 1h, for unfixed mouse tissue sections, or overnight, for fixed human tissue sections, with lid temperature 100°C to release the Tn5 enzyme. In case of mouse brain and lung samples, either 50 single-capture random nuclei mini-bulks or 10 areas of 20 cells each on average (∼200 cells mini-bulks) were collected at least in duplicates. For mouse spleen samples, both mini-bulks of 100 single-capture $B220^+$ or $CD3e^+$ cells and mini-bulks including 10 regions (each of 100 cells on average, resulting in roughly 1000 cells mini-bulks) were collected at least in quadruplicates. For human lung samples, mini-bulks of 100 single-captured PU.1 positive nuclei were collected in duplicates for each donor and lung position.

Library preparation and sequencing

To generate ATAC-seq libraries, 5μL of 10% Tween 20 (ThermoFischer) was first added to reverse crosslinked samples, followed by 10 μL NEBNext HiFi PCR Master Mix (New England Biolabs) for further amplification. We checked library size distribution via tapestation using D5000 assay on a Tapestation 4200 instrument (both from Agilent) and quantified the libraries via Qubit HS dsDNA assay (Invitrogen). We clustered the libraries at 250p.m. final clustering concentration on a NovaSeq6000 instrument using SP or S1 v1.5 chemistry (Illumina) and sequenced paired-end 2 × 50 cycles before demultiplexing using bcl2fastq2 v2.20.

Bioinformatic analysis

Level 1: preprocessing and QC. Adapter trimming was performed using Trimmomatic (v 0.36)[42] and sequencing reads were aligned using bowtie2 (v 2.4.5) against the mm10 mouse reference or hg38 human reference genome.[40] Following, duplicated reads were removed using the Picard *dedup* function, and the transposase-induced offset was corrected using the deeptools (v 3.1.3) *alignmentSieve* function.[43] The sorting and indexing of bam files was performed using samtools (v 1.14),[41] while MACS2 (v 2.2.7.1) was employed for the subsequent peak calling[44] on a merged bam file of all samples. Alignment statistics were generated via the MultiQC tool.[46] Sequencing reads in these unified peak regions were quantified in R (v 4.1.0) using the *summarizeOverlaps* function (GenomicAlignments package, v 1.28.0).[45] Raw counts were prefiltered to exclude low-count peaks (<10 reads, mouse brain/lung: 35,423 peaks, mouse spleen: 27,388 peaks, human: 7,305 peaks). FSD, TSS enrichment (BRGenomics v 1.4.0, ATACseqQC v 1.16.0),[47,52] log10(unique reads), FRiP score (defined as reads in peaks/(reads in gaps+reads in peaks)) and percentage of mitochondrial reads were utilized as parameters to describe the quality of ATAC-seq data.

Level 2: Further description of the dataset. For further dataset analyses, we normalized and rlog-transformed counts following the DESeq2 (Bioconductor, v 1.32.0) pipeline using default parameters.[55,56] Peak regions were annotated using *ChIPseeker* (Bioconductor, v 1.28.3), which defines the categories as being unique following a specific priority (http://bioconductor.org/packages/devel/bioc/vignettes/ChIPseeker/inst/doc/ChIPseeker.html#peak-annotation). The *annotatePeak* function employs a default window of 3,000 bp upstream and downstream of the TSS to annotate promoters (see *tssRegion* parameter for adjustment). All present peaks were used as input for the PCA. Pairwise correlations between the ENCODE bigwig references (mouse lung bulk: ENCSR102NGD dataset, ENCFF435PTI.bigwig; mouse brain bulk: ENCSR310MLB dataset, ENCFF561KNB.bigwig) and the merged.bigwig from lung versus brain mouse samples were computed using the functions *multiBigwigSummary* (*bins* mode) and *plotCorrelation* (deeptools v 3.1.3), Spearman correlation coefficients were also reported. Gene region tracks were plotted using the Gviz package (Bioconductor, mouse v 1.36.2, human v 1.38.4) *DataTrack* function and RPKM-normalized bigwig files, using also reference bulk ATAC-seq ENCODE bigwig files (mouse lung bulk: ENCSR102NGD dataset, ENCFF435PTI.bigwig; mouse brain bulk: ENCSR310MLB dataset, ENCFF561KNB.bigwig; human lung bulk: ENCSR647AOY dataset, ENCFF210HIS.bigwig).

Level 3: Main biological analysis. The call for DARs was performed for mouse lung vs. brain comparisons (separate nuclei number, 50 cells versus 10 BA) considering a padj<0.05 threshold (IHW multiple testing) and $\log_2$(fold-change) threshold = 1. Upset plots were produced using the UpSetR package (version 1.4.0).[48] Only DARs mapping in promoters were included in the heatmap (Figure 2C). Reference promoter peaks in case of multiple mapping to the same gene were chosen based on higher variance. Normalized and

rlog-transformed row-scaled counts were used as input for the heatmap (ComplexHeatmap version 2.8.0).[49] Hierarchical clustering was applied to identify blocks of DARs with similar regulations across conditions as reported in the presented heatmap. GSEA based on the gene ontology (GO) biological process database was employed for functional enrichments on DAR-associated genes in the two clusters. All enrichment dotplots report the count and q-value associated with each term, when $q < 0.25$. In the mouse spleen dataset, gene set variation analysis (GSVA v 1.40.1)[50] was applied to investigate differences in B and T cells signature enrichments in the isolated mouse spleen B and T cell populations at the chromatin level and coupled to a Wilcoxon test to underline statistical significance (ggpubr v 0.6.0). In the mouse spleen dataset, *per gene* chromatin accessibility score was defined as the sum of normalized rlog-transformed counts in peaks mapping on the specific gene. *Per gene* chromatin accessibility was employed as an input for the heatmap in Figure 3G. ChromVAR[34] (v 1.14.0) deviations and z-scores were calculated based on the ATACseqTFEA (Bioconductor, v 1.3.0) motif built, which consists of a curated collection of jasper2018, jolma2013 and cisbp_1.02 from package motifDB (v 1.28.0) and merged by distance smaller than 1e-9 calculated by MotIVmotifDistances function (v 1.42.0). The merged motifs were exported by motifStack (v 1.30.0). Top 25 motifs reported in Figure S3A were selected based on the median *Z* score across B and T cells samples. In the human dataset, one-tailed GSEA (fgsea package v 1.18.0, scoreType = "pos")[51] applied on the chromatin accessibility rank was performed based on the whole GO sets (c5.all.v2022.1.Hs.symbols.gmt) and transcription factor target sets (c3.tft.v2022.1.Hs.symbols.gmt). In detail, enrichment plots for the GO_BP_MYELOID_CELL_DIFFERENTIATION and PU1_Q6 signatures were reported (Figure 4G). As input for such analysis, the median across chromatin accessibility scores was calculated across samples, with *per gene* chromatin accessibility score defined as the sum of normalized rlog-transformed counts in peaks mapping on the specific gene.

QUANTIFICATION AND STATISTICAL ANALYSIS

The analysis was performed on R (v 4.1.0): the specific packages used for the analysis, their version and relevant parameters used are explained in the Method details sections. The adopted statistical tests, the considered significance levels, and the number of biological replicates are also reported in figure legends. Boxplots are in the style of Tukey, where the center of the box represents the median of values, hinges represent the 25th and 75th percentile, and the whiskers are extended not further than the 1.5×IQR (inter quartile range).