SINGLE-CELL GENOMICS

Joint profiling of chromatin accessibility and gene expression in thousands of single cells

Junyue Cao[1,2], Darren A. Cusanovich[1]*†, Vijay Ramani[1]*, Delasa Aghamirzaie[1], Hannah A. Pliner[1], Andrew J. Hill[1], Riza M. Daza[1], Jose L. McFaline-Figueroa[1], Jonathan S. Packer[1], Lena Christiansen[3], Frank J. Steemers[3], Andrew C. Adey[4,5], Cole Trapnell[1,6,7]‡, Jay Shendure[1,6,7,8]‡

Although we can increasingly measure transcription, chromatin, methylation, and other aspects of molecular biology at single-cell resolution, most assays survey only one aspect of cellular biology. Here we describe sci-CAR, a combinatorial indexing–based coassay that jointly profiles chromatin accessibility and mRNA (CAR) in each of thousands of single cells. As a proof of concept, we apply sci-CAR to 4825 cells, including a time series of dexamethasone treatment, as well as to 11,296 cells from the adult mouse kidney. With the resulting data, we compare the pseudotemporal dynamics of chromatin accessibility and gene expression, reconstruct the chromatin accessibility profiles of cell types defined by RNA profiles, and link cis-regulatory sites to their target genes on the basis of the covariance of chromatin accessibility and transcription across large numbers of single cells.

The concurrent profiling of multiple classes of molecules—for example, RNA and DNA—within single cells has the potential to reveal causal regulatory relationships and to enrich the utility of organism-scale single-cell atlases. However, to date, nucleic acid "coassays" rely on physically isolating each cell, limiting their throughput to a few cells per study (fig. S1A and table S1) (1–6).

Single-cell combinatorial indexing (sci) methods use split-pool barcoding to uniquely label the nucleic acid contents of single cells or nuclei (7–13). Here we describe sci-CAR, which jointly profiles single-cell chromatin accessibility and mRNA (CAR) in a scalable fashion. sci-CAR effectively combines sci–ATAC sequencing (sci-ATAC-seq) and sci-RNA-seq into a single protocol (Fig. 1) by the following steps: (i) Nuclei are extracted, with or without fixation, and distributed to wells. (ii) A first RNA-seq "index" is introduced by in situ reverse transcription (RT) with a polythymidine [poly(T)] primer that bears a well-specific barcode and a unique molecular identifier (UMI). (iii) A first ATAC-seq index is

introduced by in situ tagmentation with Tn5 transposase that bears a well-specific barcode. (iv) All nuclei are pooled and redistributed by fluorescence-activated cell sorting to multiple plates. (v) After second-strand synthesis of cDNA, nuclei in each well are lysed, and the lysate is split into RNA- and ATAC-dedicated portions. (vi) To provide a second priming site for amplification of 3′ cDNA tags, the RNA-dedicated lysate is subjected to transposition with unindexed Tn5 transposase. 3′ cDNA tags are amplified with primers corresponding to the Tn5 adaptor and RT primer. These primers also bear a well-specific barcode that is the second RNA-seq index. (vii) The ATAC-seq–dedicated lysate is amplified with primers specific to the barcoded Tn5 adaptors from (iii). These primers also bear a well-specific barcode that is the second ATAC-seq index. (viii) Amplicons from RNA-seq– and ATAC-seq–dedicated lysates are respectively pooled and sequenced. Each sequence read is associated with two barcodes that correspond to each round of indexing. As with other sci protocols, most nuclei pass through a unique combination of wells, thereby receiving a unique combination of barcodes that can be used to group reads derived from the same cell. Because the barcodes introduced to RNA-seq and ATAC-seq libraries correspond to specific wells, we can link the mRNA and chromatin accessibility profiles of individual cells.

We applied sci-CAR to a cell-culture model of cortisol response, wherein dexamethasone (DEX), a synthetic mimic of cortisol, activates glucocorticoid receptor (GR), which binds to thousands of locations across the genome, altering the expression of hundreds of genes (14–17). We collected human lung adenocarcinoma–derived

A549 cells after 0, 1, or 3 hours of 100 nM DEX treatment and performed a 96 well (first round indexing)–by–576 well (second round indexing) sci-CAR experiment. The three time points were each represented by 24 wells during the first round of indexing, whereas the remaining 24 wells contained a mixture of human embryonic kidney (HEK) 293T and NIH/3T3 (mouse) cells (fig. S1B).

We obtained sci-RNA-seq profiles for 6093 cells (median 3809 UMIs) and sci-ATAC-seq profiles for 6085 cells (median 1456 unique reads) (fig. S1, C to E). For both data types, reads assigned to the same cell overwhelmingly mapped to one species (fig. S1, F and G). We obtained roughly equivalent UMIs per cell from "RNA-only" plates processed in parallel, albeit at a lower sequencing depth per cell. Aggregated transcriptomes of coassayed versus RNA-only plates were well correlated [Pearson correlation coefficient (r) = 0.97 to 0.98; fig. S2]. By contrast, although coassayed versus "ATAC-only" plates were similar in data quality and well correlated in aggregate (fig. S3), ATAC-only plates had ~10-fold higher molecular complexity. The lower efficiency of the coassay for ATAC is likely explained by factors including buffer modifications and our use of only half the lysate.

There were 4825 cells (70% of either set) for which we recovered both transcriptome and chromatin accessibility data. To confirm that paired profiles truly derived from the same cells, we asked whether cells from mixed human-mouse wells were consistently assigned as human or mouse. Indeed, 1423/1425 (99%) of coassayed cells from those wells were assigned the same species label from both sci-RNA-seq and sci-ATAC-seq profiles (Fig. 2A).

We next examined the time course of GR activation. DEX treatment of A549 cells increased both transcription and promoter accessibility of markers of GR activation, including genes NFKBIA, SCNN1A, CKB, PER1, and CDH16 (14, 16) (fig. S4, A and B). Unsupervised clustering or t-distributed stochastic neighbor embedding (t-SNE) visualization of either sci-RNA-seq or sci-ATAC-seq profiles readily separated clusters corresponding to untreated and DEX-treated cells (Fig. 2, B and C). Reassuringly, cells from coassay plates and single-assay plates of either type were intermixed (fig. S4C).

Of coassayed cells in clusters 1 and 2 of sci-ATAC-seq data, 88 and 93% were found in corresponding sci-RNA-seq clusters (fig. S4, D and E). Cells with concordant versus discordant assignments did not significantly differ in read depth ($P > 0.1$, Welch two-sample t test) but notably fell on the border between clusters 1 and 2 in either t-SNE (Fig. 2D and fig. S4F). Whereas most discordant cells (70%) were from 0 hours, the remainder tended to derive from 1 hour rather than 3 hours (5% of 1-hour cells versus 1% of 3-hour cells, $P = 2.2 \times 10^{-16}$, Fisher's exact test). Although we cannot rule out that this is due to imperfect clustering, these discordantly assigned cells potentially reflect transitional states in GR activation.

Differential expression (DE) analysis of sci-RNA-seq data revealed significant changes in

[1]Department of Genome Sciences, University of Washington, Seattle, WA, USA. [2]Molecular and Cellular Biology Program, University of Washington, Seattle, WA, USA. [3]Illumina Inc., San Diego, CA, USA. [4]Department of Molecular and Medical Genetics, Oregon Health & Science University, Portland, OR, USA. [5]Knight Cardiovascular Institute, Portland, OR, USA. [6]Allen Discovery Center for Cell Lineage Tracing, Seattle, WA, USA. [7]Brotman Baty Institute for Precision Medicine, Seattle, WA, USA. [8]Howard Hughes Medical Institute, University of Washington, Seattle, WA, USA.
*These authors contributed equally to this work.
†Present address: Asthma & Airway Disease Research Center and Department of Cellular and Molecular Medicine, The University of Arizona, Tucson, AZ, USA.
‡Corresponding author. Email: coletrap@uw.edu (C.T.); shendure@uw.edu (J.S.)

2613 genes [5% false discovery rate (FDR)] (table S2). For comparison, a similar analysis with bulk RNA-seq data of DEX treatment in A549 cells at 0 versus 3 hours (18) identified 870 DE genes, 536 of which were also DE here. Log$_2$ fold changes were well correlated between the datasets for DE genes ($r = 0.86$, fig. S4G).

Differential accessibility (DA) analysis of sci-ATAC-seq profiles identified significant changes at 4763 sites (5% FDR) (table S3). For comparison, a similar analysis of bulk deoxyribonuclease (DNase)–seq data from DEX-treated A549 cells at 0 versus 3 hours (18) identified 672 DA sites, 544 of which were also DA here. Log$_2$ fold changes were well correlated between the datasets for DA sites [Spearman's rank correlation coefficient (rho) = 0.68, fig. S4H].

Of our DA sites, 701 (15%) were promoters, of which 175 overlapped with DE transcripts. Transcripts for genes with DA promoters that were not DE were detected in significantly fewer cells than genes with DA promoters that were DE (median 10 versus 25%, $P < 5 \times 10^{-5}$, unpaired two-sample permutation test based on 20,000 simulations), suggesting that we may be insufficiently powered to detect DE at many genes with DA promoters. For the 175 genes that are both DA and DE, the log$_2$ fold changes were modestly correlated (rho = 0.63, fig. S4I), and 130/175 (74%) exhibited directional concordance (exact two-sided binomial test, $P = 9 \times 10^{-11}$).

We ordered cells along a pseudotime trajectory with Monocle (19) based on the top 1000 DE genes (fig. S5A). Cells were ordered consistently with the time course (Fig. 2E). Of note, the aforementioned cells from 1 hour whose cluster assignments were discordant (Fig. 2D and fig. S4F) occurred significantly earlier in pseudotime than cells with concordant assignments ($P = 3 \times 10^{-5}$, Wilcoxon rank sum test, fig. S5B). Of the 2613 DE genes, 979 (37%) increased and 1111 (43%) decreased in expression along pseudotime, whereas 523 (20%) exhibited transient changes (fig. S5, C and D, and tables S2 and S4). We exploited the coassay to examine the dynamics of chromatin accessibility across RNA-defined pseudotime, identifying opening (47%), closing (32%), and transient (21%) DA sites (fig. S5E and tables S3 and S5). There were 11 genes that showed significant changes in both gene expression and promoter accessibility along pseudotime (5% FDR for both), with well-correlated dynamics (Fig. 2F and fig. S5, F to H).

We converted the ATAC-seq (cell-by-site) matrix to a [cell-by–transcription factor (TF) motif] matrix, simply by counting occurrences of each motif in all accessible sites for each cell (20). The motifs of 91/399 (23%) of expressed TFs were DA across the treatment conditions (5% FDR) (tables S6 and S7). Where chromatin immunoprecipitation (ChIP) sequencing data was available for the same time course (18), we observed consistent dynamics of increasing motif-associated accessibility (fig. S6A) and TF binding to accessible sites (fig. S6B). Motif-accessibility dynamics across expression-defined pseudotime are summarized in fig. S6C. The motif of the canonical GR NR3C1 was the most

activated, even though its expression decreased (Fig. 2G), consistent with its activation by recruitment from the cytosol rather than by increased expression. By contrast, KLF9 is a direct target of GR activation via a feed-forward loop (21). Consistent with this, we observed that both its expression and its motif accessibility increase along pseudotime (Fig. 2G and fig. S6, D and E).

Single-cell RNA-seq studies have recently characterized the transcriptomes of diverse cell types represented in the mammalian kidney (22–24). However, little is known about the epigenetic landscapes that underlie these cell type–specific gene expression programs. To investigate this, we isolated and fixed nuclei from whole kidneys of two 8-week-old male mice (fig. S7A). From one sci-CAR experiment, we obtained sci-RNA-seq profiles for 13,893 nuclei (median 1011 UMIs; fig. S7B) and sci-ATAC-seq profiles for 13,395 nuclei (median 7987 unique reads; fig. S7C). There were 11,296 cells for which we recovered both transcriptome and chromatin accessibility profiles.

We compared sci-CAR transcriptomes with a recently published single-cell RNA-seq dataset of the same tissue generated by Drop-seq (24). After correcting for gene-length biases (Drop-seq is biased toward shorter transcripts, and sci-RNA-seq toward longer transcripts), aggregated transcriptomes were reasonably well correlated ($r = 0.73$, fig. S7D). Semisupervised clustering of 10,727 sci-CAR transcriptomes (>500 UMIs) identified 14 groups, ranging in size from 74 (0.7%) to 2358 (22.0%) cells (Fig. 3A and fig. S7, E and F). Es-

tablished markers identified nearly all expected cell types (fig. S8, A and B). The expression profiles of proximal tubule cells separate them into three subtypes, including S1/S2 cells (Slc5a12+, Gatm+, Alpl+, and Slc34a1+), S3 type 1 cells (Slc34a1+ and Atp11a+), and S3 type 2 cells (Atp11a+ and Rnf24+) (fig. S8C) (25, 26). The smallest cluster is positive for cell-cycle progression markers (Mki67 and Cenpp) and may represent an actively proliferating subpopulation (fig. S8D) (25, 26). Cell-type proportions were well correlated between replicate kidneys, with the exception of paranephric body adipocytes (1.2 versus 0.4%), likely owing to technical variation in kidney dissection because these reside superficial to the renal fascia (fig. S7E).

We identified 8774 genes that were DE across the 14 cell types (5% FDR), including 1771 with more than twofold greater expression in the highest versus second-highest cell type (fig. S9, A and B, and tables S8 and S9). New marker genes were identified, such as Daam2 for renal pericytes and Calcr for collecting duct intercalated cell B (fig. S9, C and D) (25, 26). We examined the expression of solute carrier transporters, because these are essential to a principal function of the kidney. Of these, 208/345 (60%) were DE in subsets of renal tubule cell types, many corresponding to known and potentially as yet uncharacterized reabsorption specificities (Fig. 3B, fig. S9E, and table S10).

We compared aggregated sci-CAR chromatin accessibility profiles with published bulk ATAC-seq data on adult mouse kidney (18) and found them to be reasonably well correlated ($r = 0.75$;
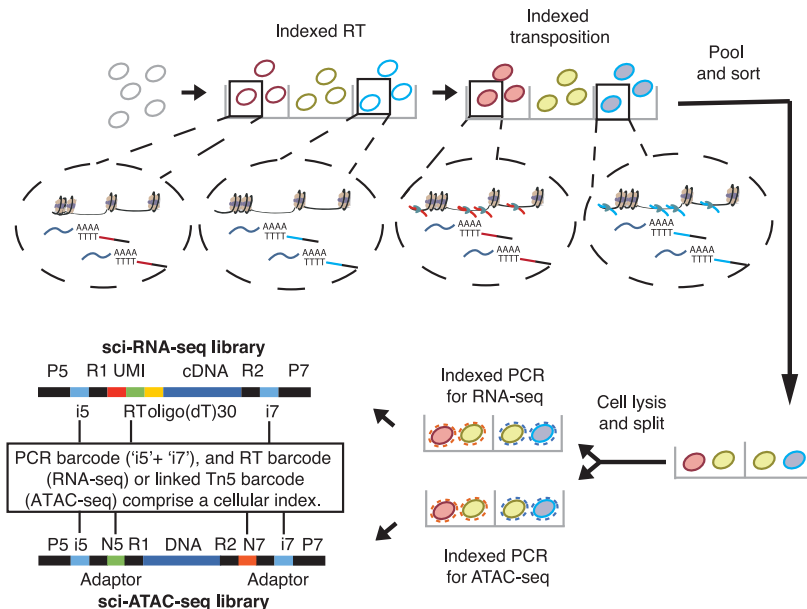
Fig. 1. sci-CAR workflow. For a more detailed explanation of key steps, see the text. From extracted nuclei, a first RNA-seq index is introduced by RT and a first ATAC-seq index by transposition. The nuclei are pooled and sorted, cDNA is synthesized, and nuclei are lysed. The lysate is split into portions for RNA-seq and ATAC-seq. For RNA-seq, index2 and read1 cover the i5 index, UMI, and RT barcode, and index1 and read2 cover the i7 index and cDNA fragment. For ATAC-seq, read1 and read2 cover the genomic DNA sequence, and index 1 and index 2 cover the Tn5 and polymerase chain reaction (PCR) barcodes. P, Illumina P5 or P7 adaptor sequence; R, annealing sites for Illumina sequencing primers; i, Illumina sequencing index; N, Tn5 transposase index.

fig. S10, A and B). Across all genes, aggregate promoter accessibility correlated with aggregate gene expression (rho = 0.26; fig. S10C). Nonetheless, a considerable challenge for single-cell ATAC-seq data, relative to single-cell RNA-seq data, is the sparsity of the resulting matrices (8). Thus, our initial efforts to cluster coassayed cells solely on the basis of their ATAC-seq profiles failed to discover the expected diversity of cell types. We therefore sought to leverage the coassay aspect of these data to recover the chromatin landscapes of individual cell types.

As a first approach, we simply annotated cell types from transcriptional profiles for ~96% of the 11,296 cells that were successfully coassayed. We then aggregated ATAC-seq signals for each cell type separately, followed by peak calling (27). As a second approach, we also developed an algorithm to combine the ATAC-seq profiles of cells with highly similar RNA-seq profiles before clustering (fig. S7A). For cells from each RNA-seq–defined cell type, we identified subsets of cells with highly similar expression profiles (a mean of 50 cells assigned to each of 222

"pseudocells"). We then aggregated the ATAC-seq profiles of each pseudocell and performed t-SNE on these. In contrast with single-cell ATAC-seq data, pseudocell chromatin accessibility profiles corresponding to the same cell types clustered together (Fig. 3C). Overall, these analyses illustrate how coassay data can be leveraged to overcome the relative sparsity of single-cell ATAC-seq data and define chromatin accessibility profiles even for closely related cell types.

We identified 22,026 DA sites across the 14 mouse kidney cell types, including 2096 promoters
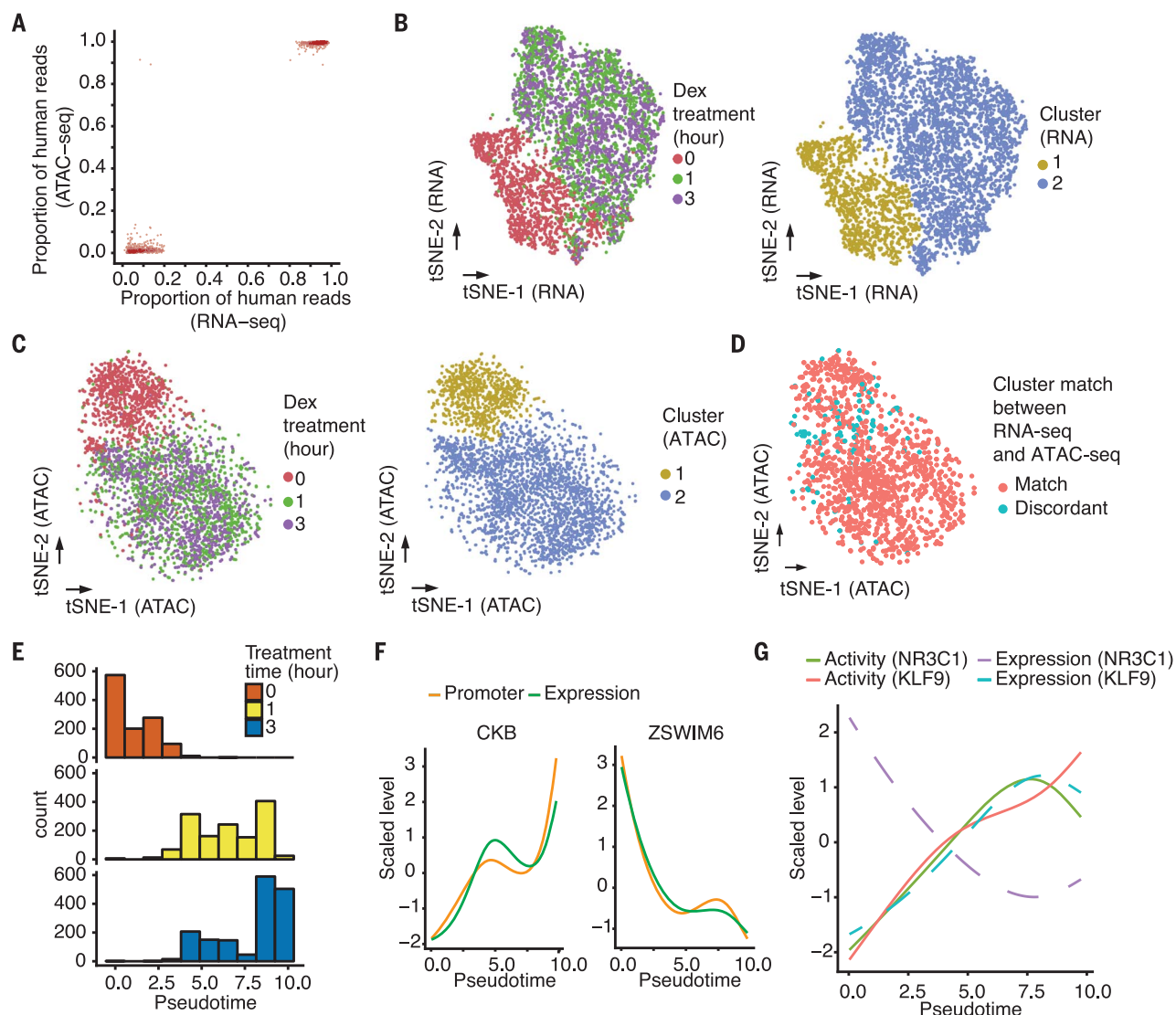

Fig. 2. Joint profiling of chromatin accessibility and transcription in DEX-treated A549 cells. (**A**) Scatter plot showing the proportion of human reads, out of all reads that map uniquely to the human or mouse reference genomes, for cells in which both RNA-seq profiles and ATAC-seq profiles were obtained. Only HEK293T (human) and NIH/3T3 (mouse) cells are plotted. (**B**) t-SNE visualization of A549 cells (RNA-seq) including cells from both sci-CAR and sci-RNA-seq–only plates, colored by DEX treatment time (left) or unsupervised clustering ID (right). (**C**) t-SNE visualization of A549 cells (ATAC-seq) including cells from both sci-CAR and sci-ATAC-seq–only plates, colored by DEX treatment time (left) or unsupervised clustering ID (right). (**D**) t-SNE visualization of A549 cells (ATAC-seq) with linked RNA-seq profiles. If the cell is in cluster 1 (or cluster 2) in both RNA-seq and ATAC-seq, then it is labeled as "Match," otherwise it is labeled "Discordant." (**E**) Distribution of cells from different DEX-treatment time points in gene expression pseudotime inferred by trajectory analysis. Pseudotime units are arbitrary. (**F**) Smoothed line plot showing scaled (with the R function scale) gene expression and promoter accessibility of *CKB* and *ZSWIM6* across pseudotime. The unscaled, unsmoothed data are shown in fig. S5, F and G. (**G**) Smoothed line plot showing the scaled mRNA level and activity change of transcription factors NR3C1 and KLF9 across pseudotime. The unscaled, unsmoothed data are shown in fig. S6, D and E.

and 19,930 distal sites (5% FDR; Fig. 3D; fig. S10, D and E; and tables S11 and S12). In some cases, DA at a gene's promoter was concordant with DE (fig. S11, A and B), but this was the exception rather than the rule. Out of 2096 genes with a DA promoter in at least one cell type, 132 genes were also DE (1% FDR) with a greater than twofold difference between the first- and second-ranked cell type. Although promoter accessibility and expression of these genes across cell types are positively correlated (median rho = 0.17), most (112/132 or 85%) exhibited maximal promoter accessibility and gene expression in different cell types (fig. S11C). The relatively weaker correlation, compared with what we observed in the A549 DEX time series (rho = 0.63; fig. S4I), is potentially a consequence of the fact that, in the A549 cells, we were comparing changes in promoter accessibility versus expression, whereas here we are comparing absolute enrichment of accessibility at promoters versus expression.

We sought to link distal cis-regulatory elements to their target genes on the basis of the covariance of chromatin accessibility and gene expression across large numbers of coassayed cells. As the sparsity of our single-cell profiles makes this challenging, we worked with the previously described 222 pseudocells (fig. S12A). For each gene, we computed correlations between its expression and the adjusted accessibility of all sites within 100 kb of its transcriptional start site (TSS) using LASSO (least absolute shrinkage and selection operator).

Within the top 2000 DE genes (ranked by q value), we linked 1260 distal sites to 321 genes (median three sites per gene, out of a median 19 sites within 100 kb of the TSS tested; fig. S12, B and C, and table S13). Of the sites, 44% were linked to the nearest TSS and 21% to the second-nearest TSS (fig. S12D). Distal site-gene linkages were significantly closer than all possible pairs tested (mean 41 kb for links versus 48 kb for all pairs tested; $P < 5 \times 10^{-5}$, unpaired permutation test based on 20,000 simulations; fig. S12E).

To evaluate the possibility that the links were artifacts of regularized regression, we permuted the sample IDs of the chromatin accessibility matrix and performed the same analysis. After this permutation, only four links were identified (fig. S12B). To control for correlations between closely located accessible sites in the genome, we separately permuted the peak IDs. This yielded 216 links, or just 17% as many links as without permutation (fig. S12B).
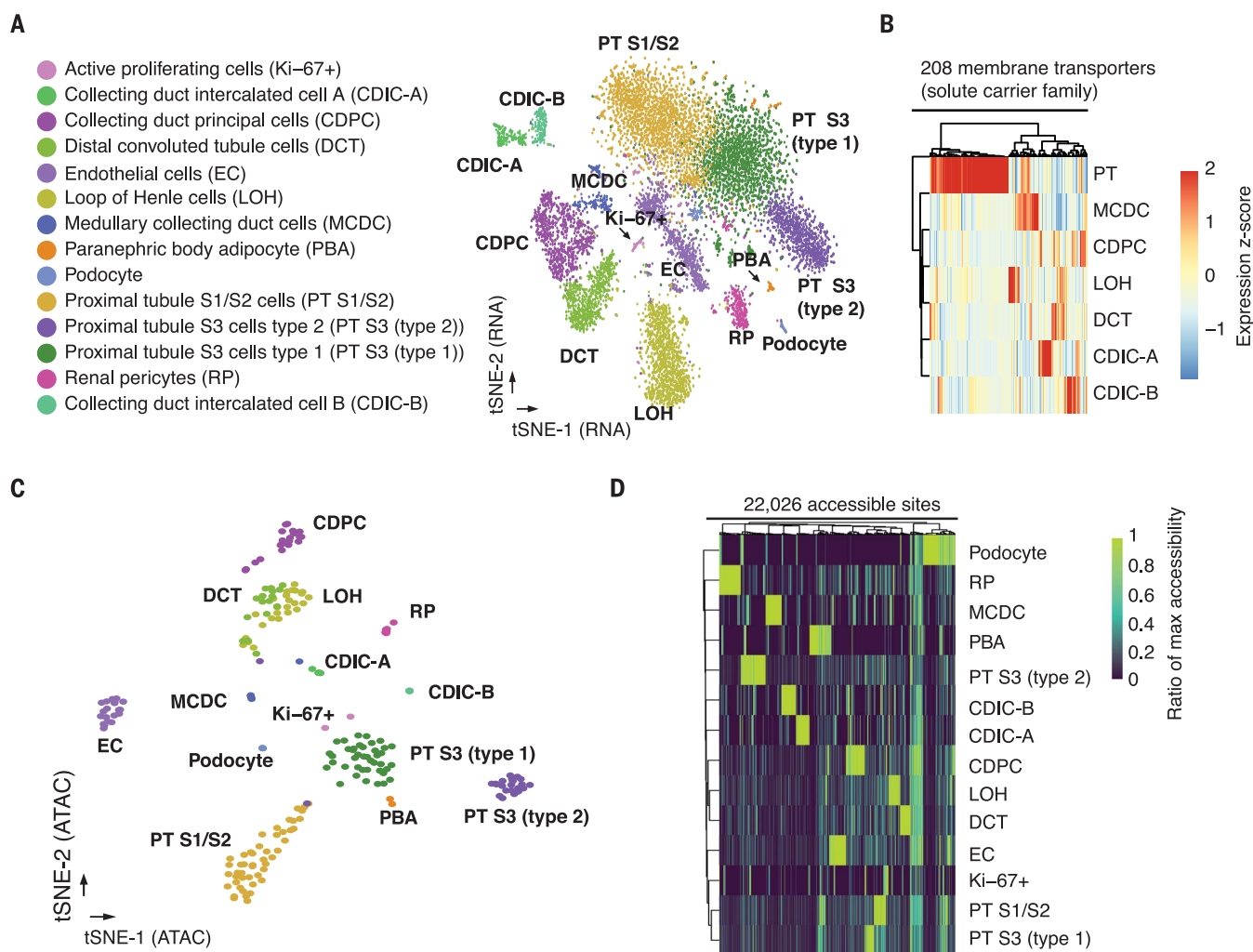

Fig. 3. sci-CAR enables joint profiling of chromatin accessibility and transcription in mouse kidney. (**A**) t-SNE visualization of mouse kidney nuclei (RNA-seq). Cell types are assigned on the basis of established marker genes. (**B**) Heatmap showing the relative expression of genes from the solute carrier group of membrane transport proteins in consensus transcriptomes of each cell type estimated by RNA-seq data from the coassay. The raw expression data (UMI count matrix) was log-transformed, column-centered, and scaled (by using the R function scale), and the resulting values clamped to (−2, 2). (**C**) t-SNE visualization of mouse kidney nuclei (ATAC-seq) after aggregating cells with highly similar transcriptomes (pseudocells), colored by cell types identified from RNA-seq. (**D**) Heatmap showing the relative chromatin accessibility of cell type–specific sites for each cell type estimated by ATAC-seq data from the coassay. The raw aggregated ATAC-seq data (read count matrix) was normalized first by the total number of reads for each cell type and then by the maximum accessibility score across all cell types.
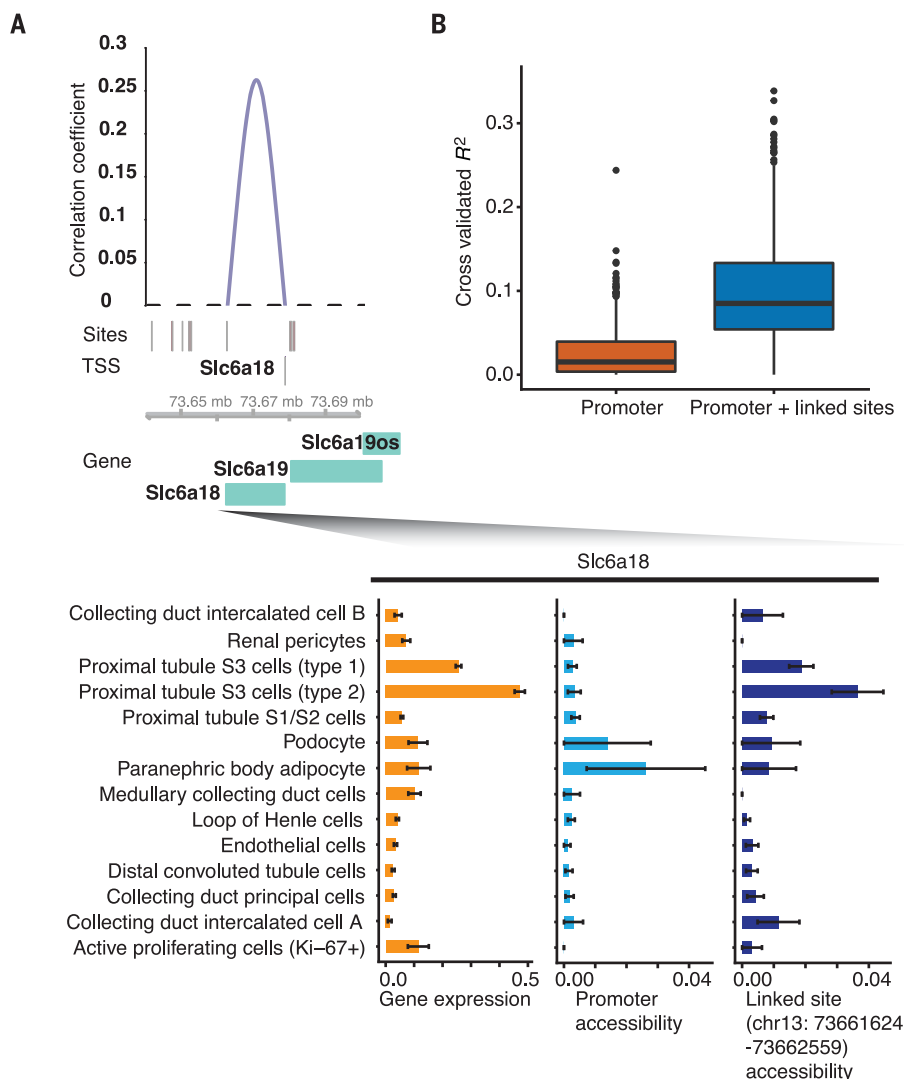
Fig. 4. Linking cis-regulatory elements to regulated genes on the basis of covariance in single-cell coassay data. (**A**) Top: Genome browser plot showing links between accessible distal regulatory sites and the gene *Slc6a18*. The height corresponds to the correlation coefficient. Bottom: Bar plots showing the average expression, promoter accessibility, and linked site accessibility for cell type–specific marker gene *Slc6a18* across different cell types. Gene expression values for each cell were calculated by dividing the raw UMI count by cell-specific size factors. Site accessibilities for each cell were calculated by dividing the raw read count by cell-specific size factors. Error bars represent SEM. chr13, chromosome 13. (**B**) Two linear regression models were built to predict gene expression differences between cell types. The first model predicts changes on the basis of promoter accessibility alone. The second model predicts changes on the basis of the chromatin accessibility of the promoter and distal sites that are linked to it. The boxplot shows the cross-validated coefficient of determination (R^2) calculated for each gene from the two models.

The 321 genes with linked distal sites were specifically expressed in a variety of cell types (fig. S12F). For example, the link with the highest correlation is between distal convoluted tubule cell marker gene *Slc12a3* and a site 36-kb downstream of its TSS that overlaps its last exon (fig. S13). The accessibility of this linked site was modestly more specific to distal convoluted tubule cells than the *Slc12a3* promoter. By contrast, the accessible site closest to the *Slc12a3* promoter (only 216 base pairs away) was not linked to the *Slc12a3* promoter by our approach nor is its accessibility specific to distal convoluted tubule cells. Similarly, a marker gene for loop of Henle cells, *Slc12a1*, is linked to two distal sites (fig. S14), both of which exhibit accessibility specific to loop of Henle cells. By contrast, the nearest accessible site (9 kb from the TSS), which was not linked, does not exhibit this specificity.

Links between distal cis-regulatory elements and their target genes can be useful for explaining differential expression across cell types. For example, the cell type–specific expression of *Slc6a18*, a marker gene for type 2 proximal tubule S3 cells, is not mirrored by cell type–specific pro-

moter accessibility (fig. S11C). However, from our covariance approach, its TSS is linked to a site 16 kb away whose accessibility is correlated with *Slc6a18* expression (Fig. 4A). To quantify the utility of the links between distal cis-regulatory elements and their target genes identified from sci-CAR data, we constructed a linear regression model to predict gene expression differences based on chromatin accessibility at promoters only versus promoters together with linked distal sites. Including linked distal sites improved predictions by fourfold ($P < 5 \times 10^{-5}$, paired permutation test based on 20,000 simulations; Fig. 4B).

Our analyses illustrate the advantages of a single-cell coassay over assays that solely profile transcription or chromatin accessibility. sci-CAR is compatible with fresh or fixed nuclei and, like other sci-seq techniques, can encode multiple samples per experiment. Its throughput can potentially be increased by additional rounds of split-pool indexing (*13*). With 384-well–by–384-well–by–384-well sci-CAR, one could potentially coassay millions of single cells per experiment. A limitation of sci-CAR is the sparsity of the resulting

data, particularly with respect to chromatin accessibility. This can potentially be overcome in the future through protocol optimizations, particularly of cross-linking conditions. A second limitation is that, although we were able to link distal elements and target genes on the basis of covariance of accessibility and expression, these data remain correlative and involve a minority of DE genes and DA elements.

Notwithstanding these limitations, sci-CAR expands the potential of combinatorial indexing for scalably profiling single-cell molecular phenotypes and may be particularly useful in the context of organism-scale single-cell atlases. With further development, we anticipate that additional DNA and RNA coassays may be realized by simply integrating other sci-seq protocols together with sci-RNA-seq (e.g., methylation plus transcripts, chromosome conformation plus transcripts, or DNA sequence plus transcripts) (*8–13*). A longer-term goal is to adapt single-cell combinatorial indexing to span the central dogma, such that aspects of DNA, RNA, and protein species can be concurrently assayed from each of many single cells.

REFERENCES AND NOTES

1. S. J. Clark et al., Nat. Commun. **9**, 781 (2018).
2. C. Angermueller et al., Nat. Methods **13**, 229–232 (2016).
3. Y. Hou et al., Cell Res. **26**, 304–319 (2016).
4. Y. Hu et al., Genome Biol. **17**, 88 (2016).
5. S. Pott, eLife **6**, e23203 (2017).
6. F. Guo et al., Cell Res. **27**, 967–988 (2017).
7. S. Amini et al., Nat. Genet. **46**, 1343–1349 (2014).
8. D. A. Cusanovich et al., Science **348**, 910–914 (2015).
9. V. Ramani et al., Nat. Methods **14**, 263–266 (2017).
10. Y. Yin et al., bioRxiv 338053 [Preprint]. 12 June 2018. https://doi.org/10.1101/338053.
11. S. A. Vitak et al., Nat. Methods **14**, 302–308 (2017).
12. R. M. Mulqueen et al., Nat. Biotechnol. **36**, 428–431 (2018).
13. J. Cao et al., Science **357**, 661–667 (2017).
14. T. E. Reddy et al., Genome Res. **19**, 2163–2171 (2009).
15. S. John et al., Nat. Genet. **43**, 264–268 (2011).
16. T. E. Reddy, J. Gertz, G. E. Crawford, M. J. Garabedian, R. M. Myers, Mol. Cell. Biol. **32**, 3756–3767 (2012).
17. C. M. Vockley et al., Cell **166**, 1269–1281.e19 (2016).
18. ENCODE Project Consortium, Nature **489**, 57–74 (2012).
19. X. Qiu et al., Nat. Methods **14**, 979–982 (2017).
20. J. D. Buenrostro et al., Nature **523**, 486–490 (2015).
21. Y. Chinenov, M. Coppo, R. Gupte, M. A. Sacta, I. Rogatsky, BMC Genomics **15**, 656 (2014).
22. L. Chen et al., Proc. Natl. Acad. Sci. U.S.A. **114**, E9989–E9998 (2017).
23. X. Han et al., Cell **172**, 1091–1107.e17 (2018).
24. J. Park et al., Science **360**, 758–763 (2018).
25. The Human Protein Atlas; www.proteinatlas.org.
26. M. Uhlen et al., Science **357**, eaan2507 (2017).
27. Y. Zhang et al., Genome Biol. **9**, R137 (2008).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6409/1380/suppl/DC1
Materials and Methods
Figs. S1 to S14
Tables S1 to S13
References (28–53)

Science

Joint profiling of chromatin accessibility and gene expression in thousands of single cells

Junyue Cao, Darren A. Cusanovich, Vijay Ramani, Delasa Aghamirzaie, Hannah A. Pliner, Andrew J. Hill, Riza M. Daza, Jose L. McFaline-Figueroa, Jonathan S. Packer, Lena Christiansen, Frank J. Steemers, Andrew C. Adey, Cole Trapnell and Jay Shendure

Single-cell chromatin and RNA analysis

Single-cell analyses have begun to provide insight into the differences among and within the individual cells that make up a tissue or organism. However, technological barriers owing to the small amount of material present in each single cell have prevented parallel analyses. Cao *et al.* present sci-CAR, a pooled barcode method that jointly analyzes both the RNA transcripts and chromatin profiles of single cells. By applying sci-CAR to lung adenocarcinoma cells and mouse kidney tissue, the authors demonstrate precision in assessing expression and genome accessibility at a genome-wide scale. The approach provides an improvement over bulk analysis, which can be confounded by differing cellular subgroups.
Science, this issue p. 1380

ARTICLE TOOLS	http://science.sciencemag.org/content/361/6409/1380
SUPPLEMENTARY MATERIALS	http://science.sciencemag.org/content/suppl/2018/08/29/science.aau0730.DC1
RELATED CONTENT	http://science.sciencemag.org/content/sci/361/6409/1330.full
REFERENCES	This article cites 39 articles, 8 of which you can access for free http://science.sciencemag.org/content/361/6409/1380#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions