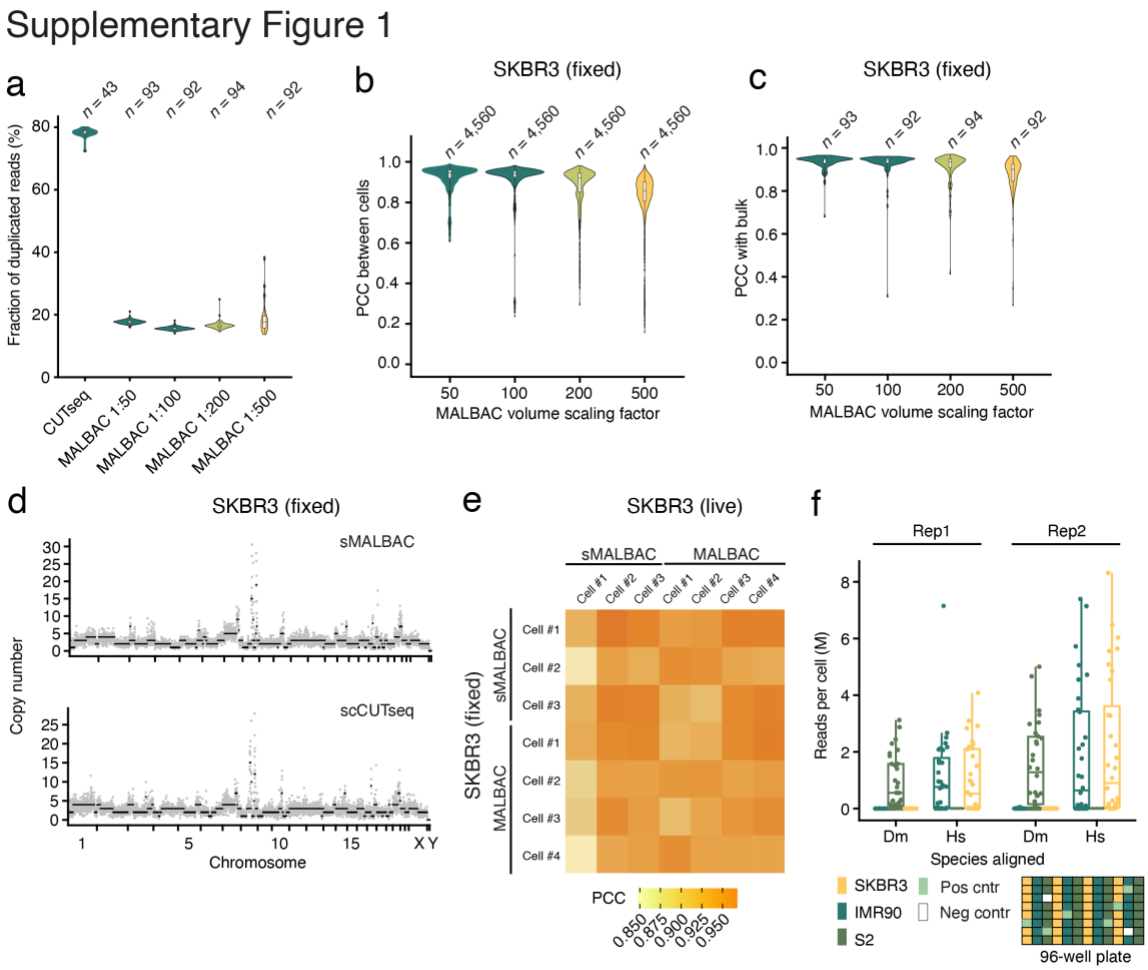
SUPPLEMENTARY INFORMATION

High clonal diversity and spatial genetic admixture in early prostate cancer and surrounding normal tissue

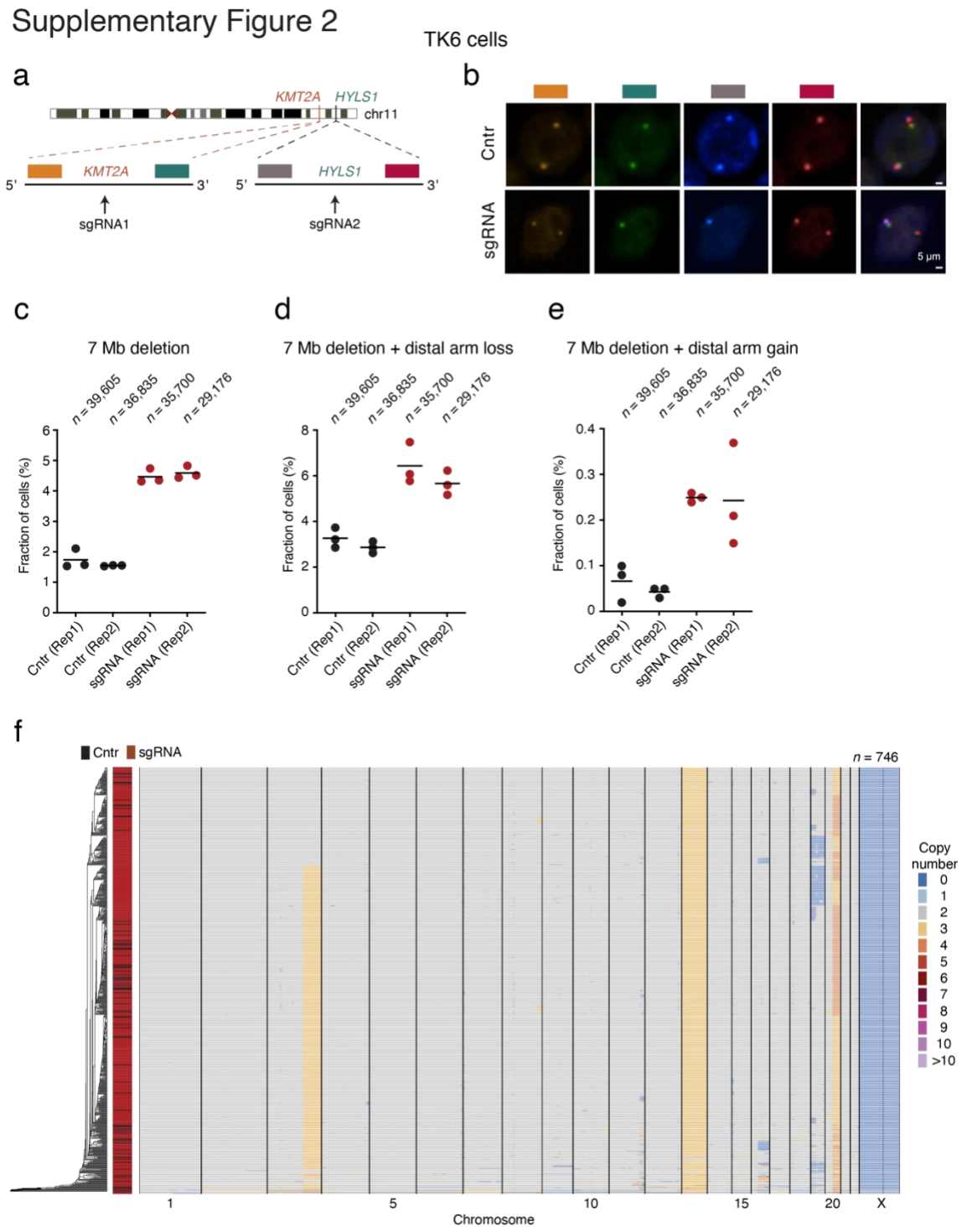
Ning Zhang, Luuk Harbers, Michele Simonetti, Gabriel M.C. Longo, Enrico Berrino, Niklas Schultz, Firas Tarish, Peng Su, Bo Han, Wanzhong Wang, Sofia Onorato, Qifeng Yang, Thomas Helleday, Vassilis Roukos, Caterina Marchiò, Magda Bienko, Nicola Crosetto

1. Supplementary Figures



Supplementary Fig. 1. Technical performance and reproducibility of scCUTseq. (**a**) Fraction of read duplicates obtained by performing standard CUTseq on individual cells or MALBAC by scaling reagent volumes 50, 100, 200 and 500 times. *n*, number of single cells sequenced. (**b**) Pearson's correlation coefficient (PCC) for all possible pair-wise comparisons between the segmented copy number profiles of individual SKBR3 cells processed by scCUTseq, using different MALBAC reagent volume scaling factors. *n*, number of pair-wise comparisons in each group. (**c**) Same as in (b) but comparing the copy number profile of each individual cell with the copy number profile of the corresponding cell line determined by bulk CUTseq. *n*, number of cells compared to bulk CUTseq. (**d**) Example of genome-wide copy number profiles of individual mildly fixed SKBR3 cells determined by performing a 1:200 scaled down version of MALBAC followed either by standard library preparation (sMALBAC) or by CUTseq (scCUTseq). Each gray dot represents a 500 kilobases (kb) genomic bin. Black dots indicate

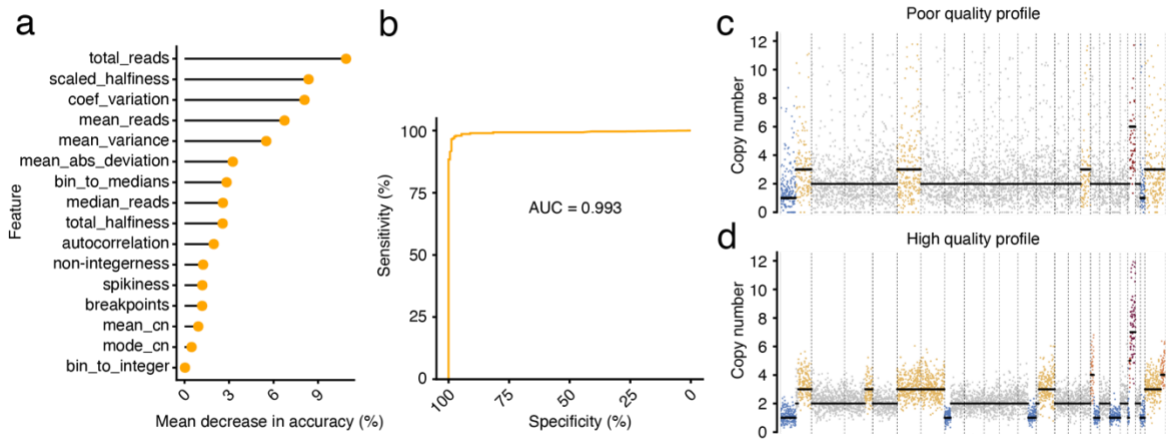
copy number levels determined by circular binary segmentation (see **Methods**). (**e**) Correlation matrix showing the degree of similarity between the segmented copy number profiles of individual fixed or non-fixed (live) SKBR3 cells obtained by performing standard MALBAC or a 1:200 scaled down version of it (sMALBAC) followed by standard library preparation. PCC, Pearson's correlation coefficient. (**f**) Distributions of scCUTseq reads per cell after alignment to the *Homo sapiens* (Hs) or *Drosophila melanogaster* (Dm) reference genomes for three different cell types (Hs IMR90 and SKBR3 cells; Dm S2 cells) sorted in different columns of a 96-well plate as shown in the bottom scheme. Pos cntr, positive control consisting of 20 pg of genomic DNA from each cell line used as input for scCUTseq. Neg cntr, negative control consisting of Nuclease-Free Water. In all the boxplots in the figure, each box spans from the 25^{th} to the 75^{th} percentile and whiskers extend from –1.5×IQR to +1.5×IQR from the closest quartile, where IQR is the inter-quartile range. Each dot corresponds to single well in the plate scheme shown below.



Supplementary Fig. 2. scCUTseq sensitivity assessment. (**a**) Scheme of fluorescence in situ hybridization (FISH) probes (colored rectangular bars) surrounding the *KMT2A* and *HYLS1* gene loci on chr11, used to detect a 7 Mb deletion induced by CRISPR-Cas9 using two different small-guide RNAs (sgRNA) (see also **Fig. 1d**). (**b**) Visualization of the FISH probes shown in (a) in TK6 cells either transfected with non-targeting sgRNA (Cntr) or with both sgRNAs

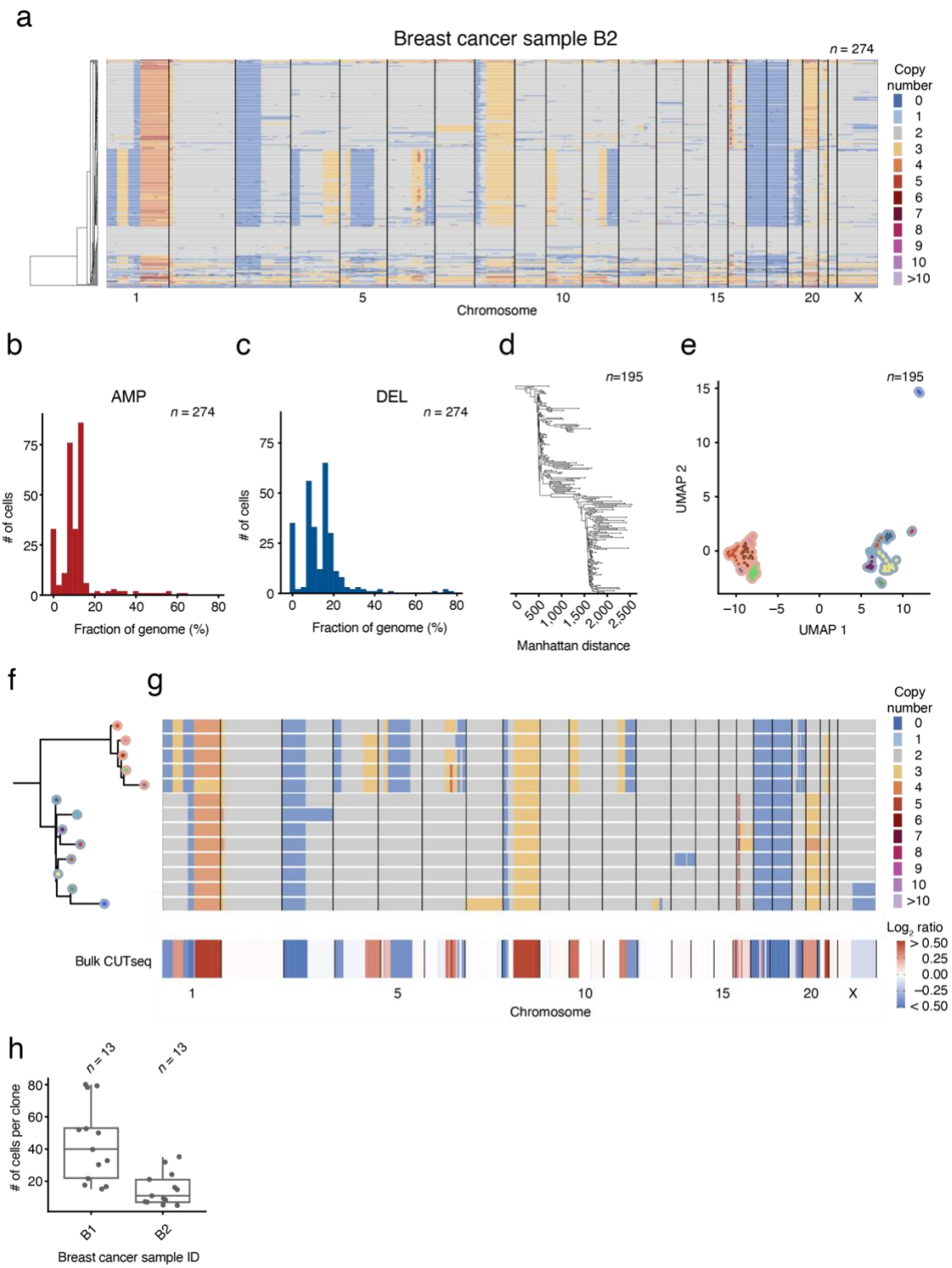
shown in (a) (sgRNA). The colored bars on the top correspond to the FISH probes shown in (a). In sgRNA treated cells in which the correct 7 Mb deletion event has occurred, the probes downstream of the *KMT2A* locus (green) and upstream of the *HYLS1* locus (blue) are detected only on one chr11 homologue, as expected. Scale bars, 5 μm. (**c**) Fraction of TK6 cells carrying the exact 7 Mb deletion on chr11. Each dot represents the fraction of cells containing the chromosomal rearrangement from a FISH technical replicate (one out of three wells of a 96-well plate). Horizontal black bars represent the mean. Rep, biological replicates. (**d**) Same as in (c), but for cells in which the 3′ portion of chr11 after the 7 Mb deletion was lost. (**e**) Same as in (c), but for cells in which the 3′ portion of chr11 after the 7 Mb deletion was amplified. (**f**) Hierarchically clustered copy number profiles of control (Cntr) and sgRNA-treated (sgRNA) TK6 cells. *n*, number of single cells sequenced.

Supplementary Figure 3



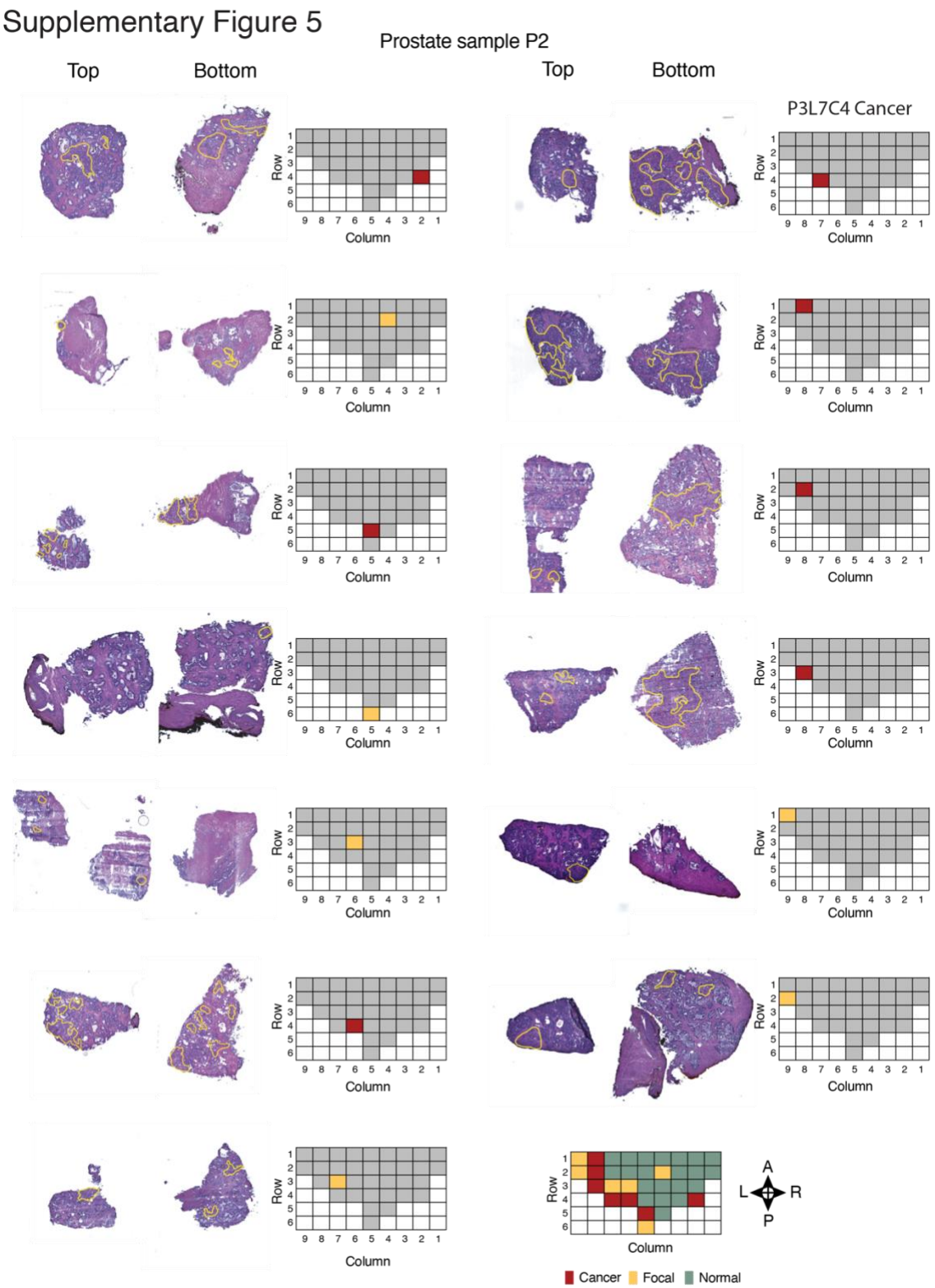
Supplementary Fig. 3. Random forest classifier of scCUTseq single-cell CNA profiles. (**a**) Features used by the random forest classifier to distinguish between high- and low-quality scCUTseq copy number profiles. For a description of each feature, see **Supplementary Table 2**. (**b**) Receiver operating characteristic curve analysis of the random forest classifier performance. AUC, area under the curve. (**c, d**) Examples of poor (c) and high (d) quality copy number profiles distinguished by the random forest classifier. Each dot represents a 500 kilobases (kb) genomic bin. Black, blue, yellow, and red dots indicate, respectively, copy number levels determined by circular binary segmentation, chromosomal deletions, gains, and amplifications.

Supplementary Figure 4



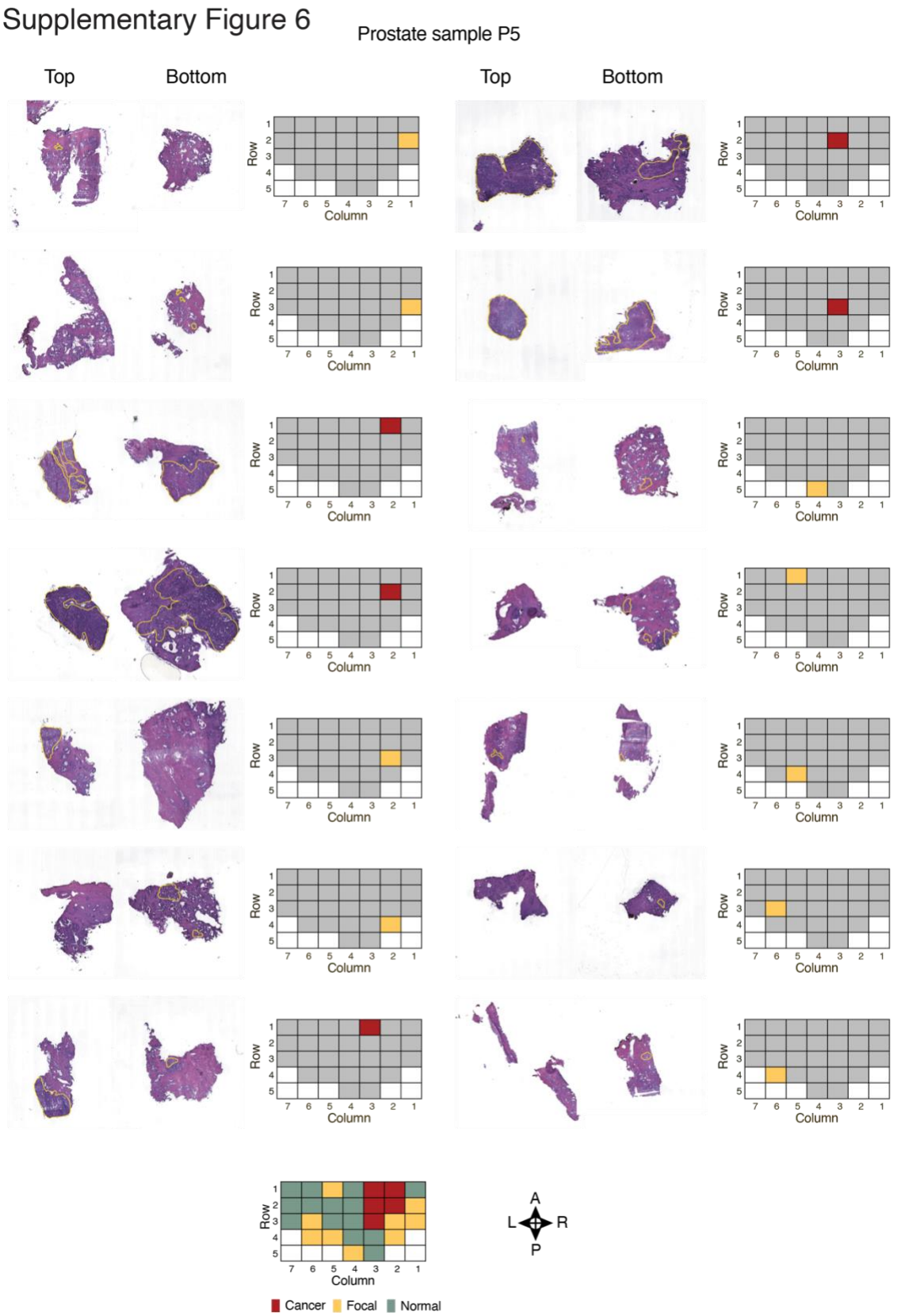
Supplementary Fig. 4. Copy number heterogeneity and clonal structure in the second breast cancer sample (B2) in which scCUTseq was tested. (**a**) Hierarchically clustered copy number profiles of individual nuclei extracted from a frozen biopsy. *n*, number of single nuclei

sequenced. (**b**) Distribution of the fraction of the genome amplified (> 2 copies) in the cells shown in (a). (**c**) Same as in (b), but for deletions. (**d**) Balanced minimal evolution tree of the non-diploid cells in (a) based on their copy number profiles. *n*, number of non-diploid cells. (**e**) Dimensionality reduction by Uniform Manifold Approximation and Projection (UMAP) of the copy number profiles of the cells clustered in (d). *n*, number of cells. Colored dot groups represent clones, whereas colored outlines indicate so-called superclones. Each dot represents a single cell. (**f**) Balanced minimal evolution tree of the clones shown in (e). (**g**) Median copy number profile of each clone shown in (f). The heatmap on the bottom shows the copy number profile obtained by applying bulk CUTseq to gDNA extracted from leftover nuclei that were not used for sorting. (**h**) Distributions of the number of cells in each clone shown in (e) and **Fig. 2e**. *n*, number of clones. In each boxplot, the box spans from the 25th to the 75th percentile and whiskers extend from the minimum to the maximum value. The horizontal line represents the median.



Supplementary Fig. 5. Pathological annotation of hematoxylin-eosin (H&E) stained tissue sections from the top and the bottom of each tissue cube in sample P2. Tumor regions are delineated by the yellow lines on top of each H&E-stained section. For each pair of sections,

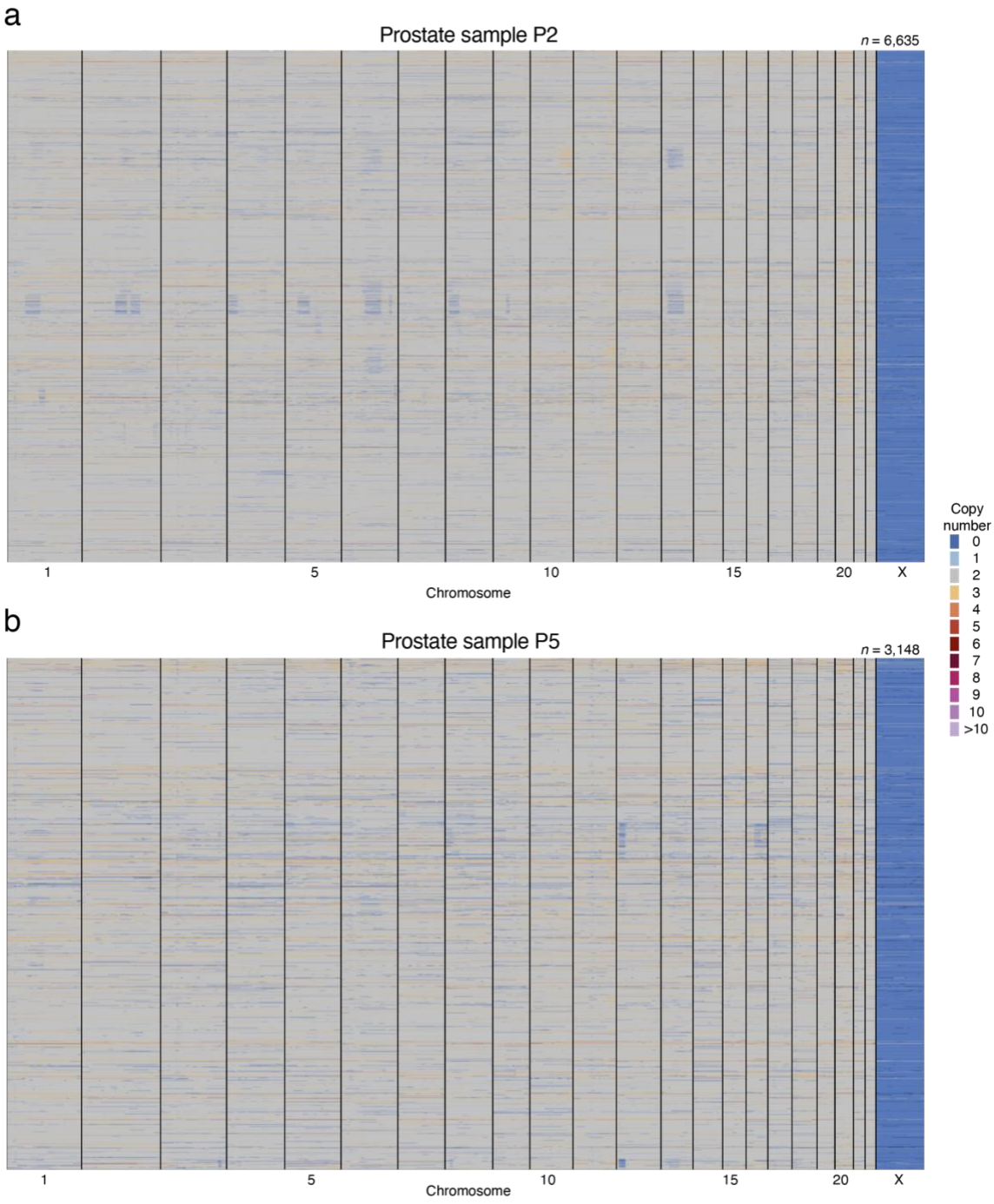
the position of the corresponding tissue cube (region) is depicted in the map on the right. Only regions containing tumor cells are shown. The map on the bottom right is the same as the one shown in **Fig. 3b**. Two board-certified pathologists examined the sections independently and classified them based on the percentage of tumor cells.



Supplementary Fig. 6. Pathological annotation of hematoxylin-eosin (H&E) stained tissue sections from the top and the bottom of each tissue cube in sample P5. Tumor regions are

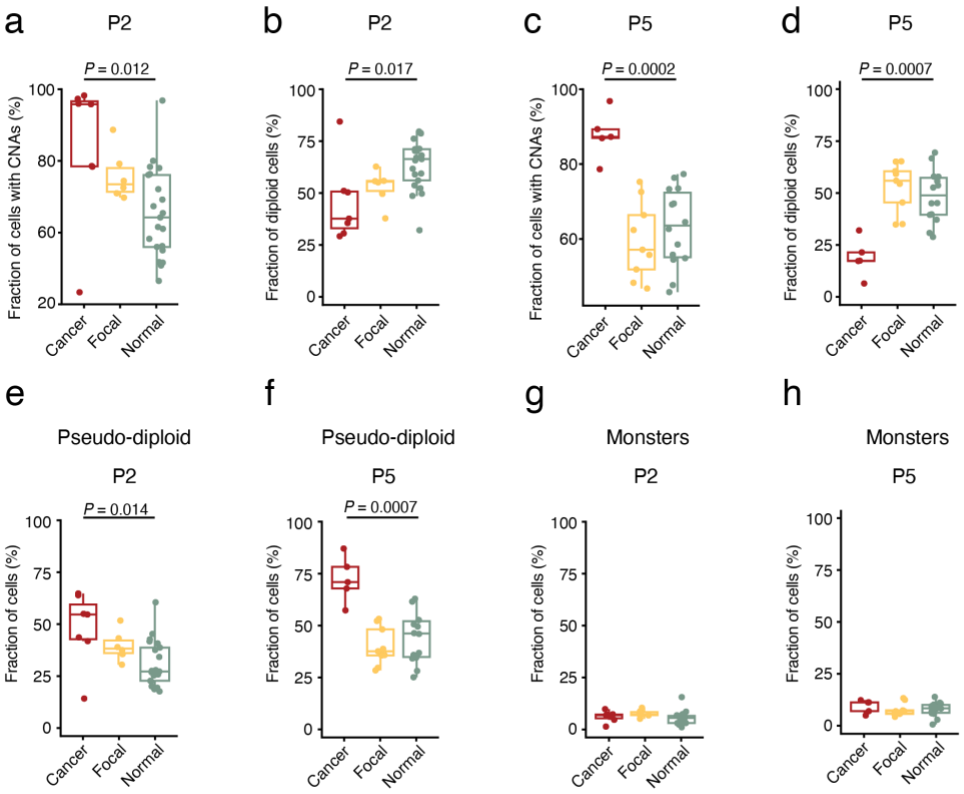
delineated by the yellow lines on top of each H&E-stained section. For each pair of sections, the position of the corresponding tissue cube (region) is depicted in the map on the right. Only regions containing tumor cells are shown. The map on the bottom is the same as the one shown in **Fig. 3e**. Two board-certified pathologists examined the sections independently and classified them based on the percentage of tumor cells.

Supplementary Figure 7



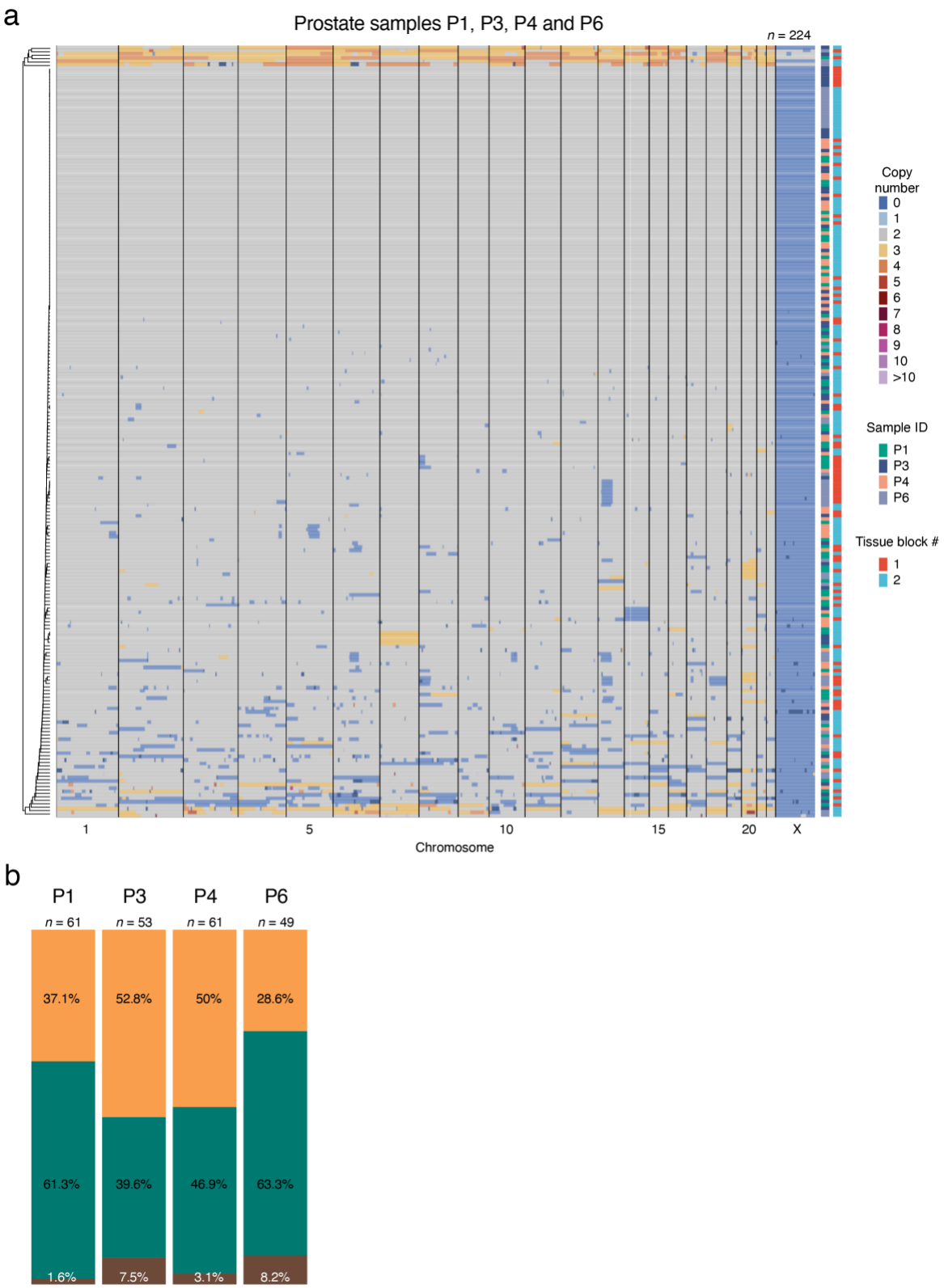
Supplementary Fig. 7. Copy number profiles in prostate tissue from prostatectomies performed in patients diagnosed with early-stage localized prostate cancer. (**a**) Copy number profiles from the first prostatectomy sample (P2), in which all tissue cubes were analyzed by scCUTseq. *n*, number of single cells. (**b**) Same as in (a), but for the second prostate sample (P5) in which we also performed spatially resolved scCUTseq.

Supplementary Figure 8



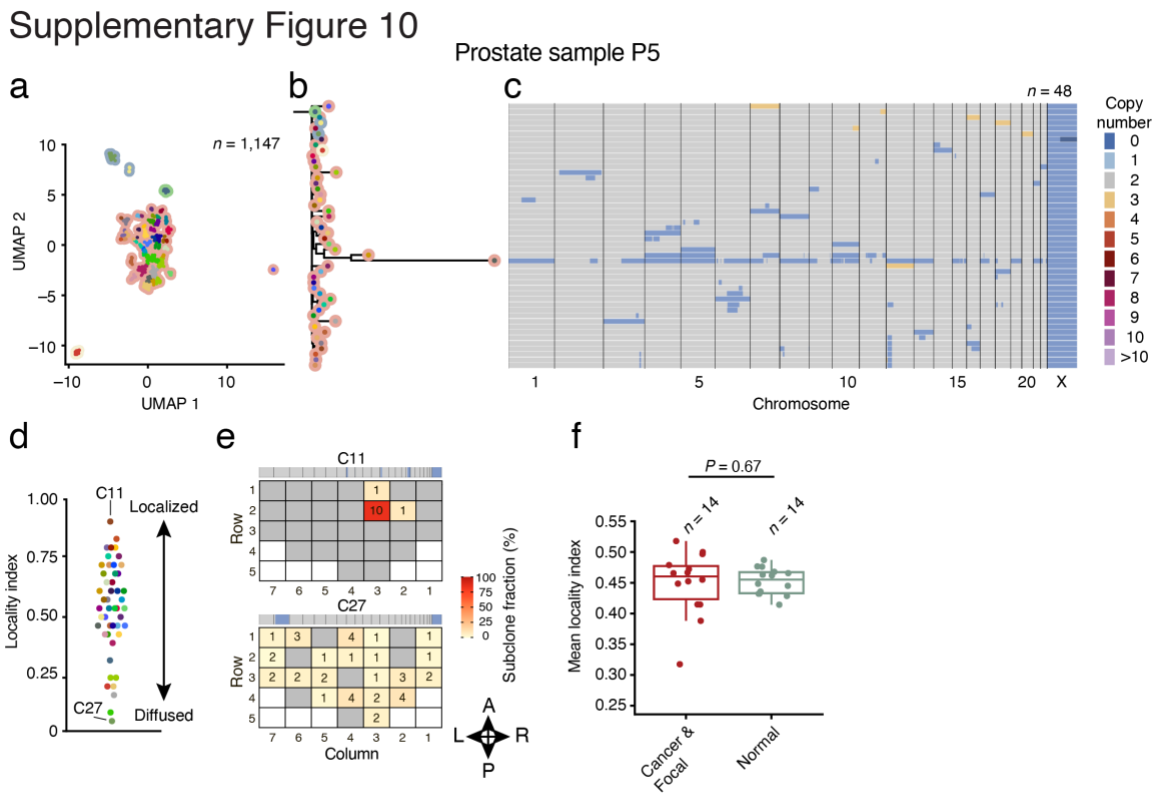
Supplementary Fig. 8. Distribution of different cell types across normal and tumor regions profiled by scCUTseq in P2 and P5 samples. (**a**) Fraction of cells with at least one CNA event in tumor (>50% of tumor cells), focal (10–50% of tumor cells) and normal regions in P2. Each dot represents one region. *P*, Wilcoxon test, two-tailed. (**b**) Same as in (a), but for diploid cells. (**c-d**) Same as in (a) and (b), respectively, but for P5. (**e**) Fraction of cells classified as pseudo-diploid in tumor, focal and normal regions in P2. (**f**) Same as in (e), but for P5. (**g-h**) Same as in (e) and (f), respectively, but for monster cells. In each boxplot in the figure, the box spans from the 25th to the 75th percentile and whiskers extend from the minimum to the maximum value. The horizontal line represents the median.

Supplementary Figure 9



Supplementary Fig. 9. Copy number profiles in the other four prostatectomy samples (P1, P3, P4, P6) that were profiled by scCUTseq. (**a**) Hierarchically clustered scCUTseq profiles from

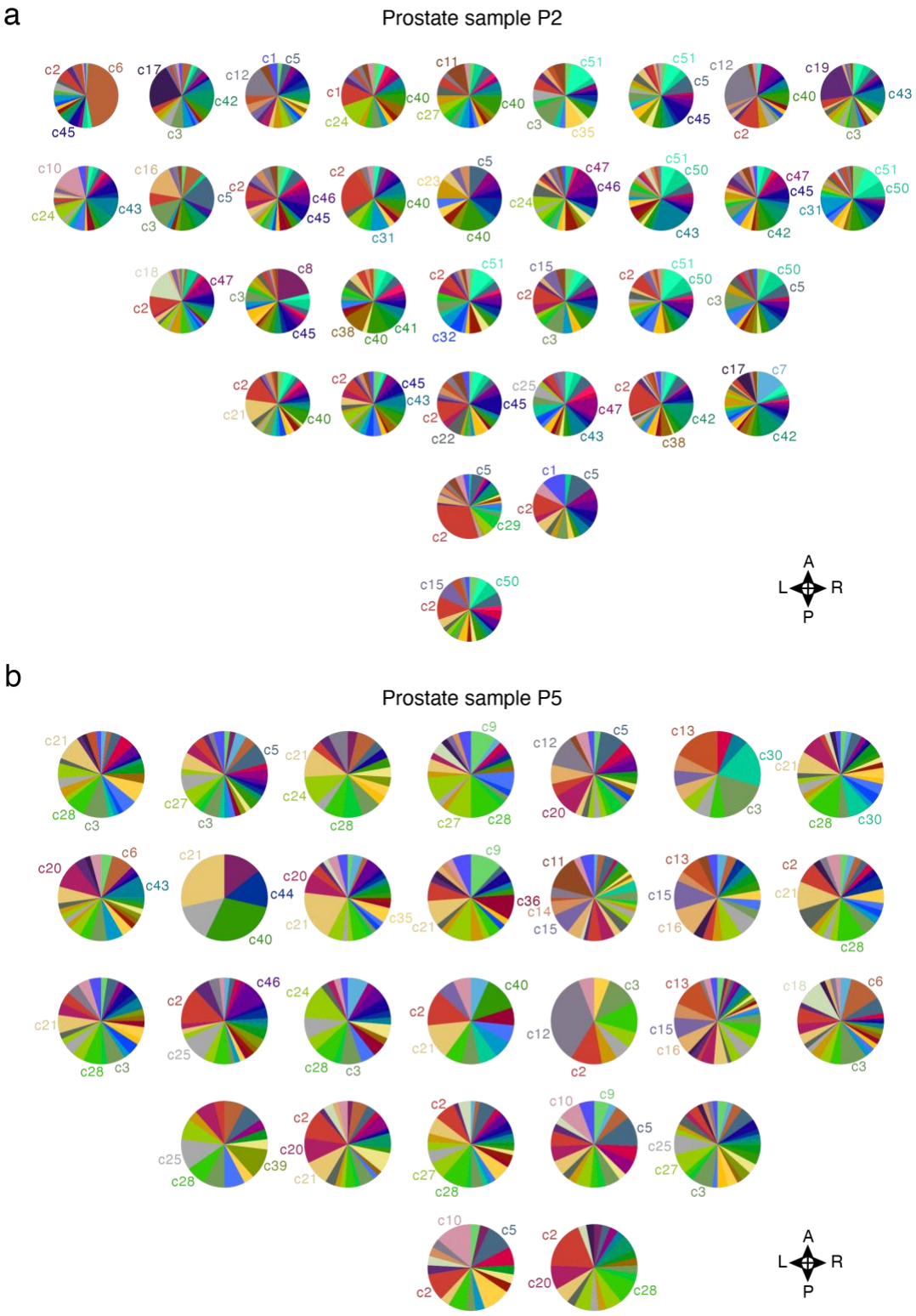
all the cells sequenced in the four samples. (**b**) Fraction of diploid, 'pseudo-diploid and 'monster' cells in the sampled shown in (a). *n*, number of single cells analyzed.



Supplementary Fig. 10. Clonal diversity and spatial heterogeneity of the pseudo-diploid cells. (**a**) UMAP dimensionality reduction of the copy number profiles of pseudo-diploid cells identified in sample P5. *n*, number of cells. Dots of the same color represent clones, whereas colored outlines indicate so-called superclones. Each dot represents a single cell. (**b**) Balanced minimal evolution tree of the clones shown in (a). (**c**) Median copy number profile of each clone shown in (b). (**d**) Distribution of the locality index for each pseudo-diploid cell clone shown in (a). The color of each clone is the same as in (a). (**e**) Schematic maps of the spatial distribution of the most focal (top) and most spread (bottom) pseudo-diploid clones in P5. Each grid represents the tissue map, with each cell in the grid representing a tissue cube from which nuclei were isolated and profiled by scCUTseq (see scheme in **Fig. 3a**). Grey cells indicate tissue regions in which the corresponding clone was not detected. White cells indicate absence of prostate tissue. The bar on top of each grid shows the genome-wide median copy number profile of the corresponding clone. Blue bars indicate deletions. Vertical black bars mark the boundaries between consecutive chromosomes. The anatomical orientation of the grids is shown by the four arrows on the bottom right. A, anterior. P, posterior. L, left. R, right. (**f**) Distributions of the locality index of pseudo-diploid clones in different P5 tissue regions. *P*, Wilcoxon's test, two-tailed. In each boxplot, the box spans from the 25th to the 75th percentile

and whiskers extend from the minimum to the maximum value. The horizontal line represents the median.

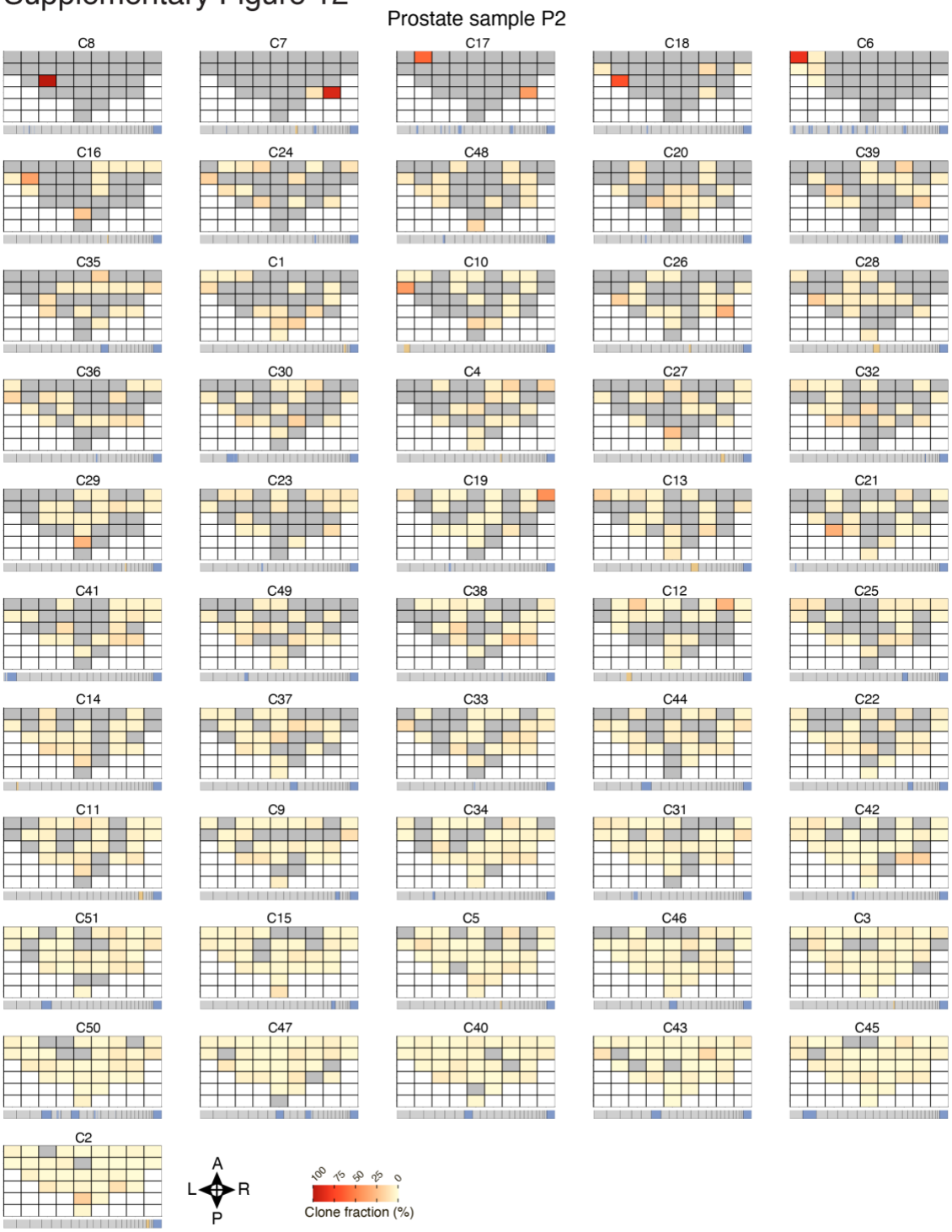
Supplementary Figure 11



Supplementary Fig. 11. Proportion of pseudo-diploid clones throughout all the tissue regions sampled by scCUTseq in samples P2 (**a**) and P5 (**b**). Each pie chart corresponds to one tissue cube (region) The ID of the three most frequent clones (c) in each tissue block is shown. The

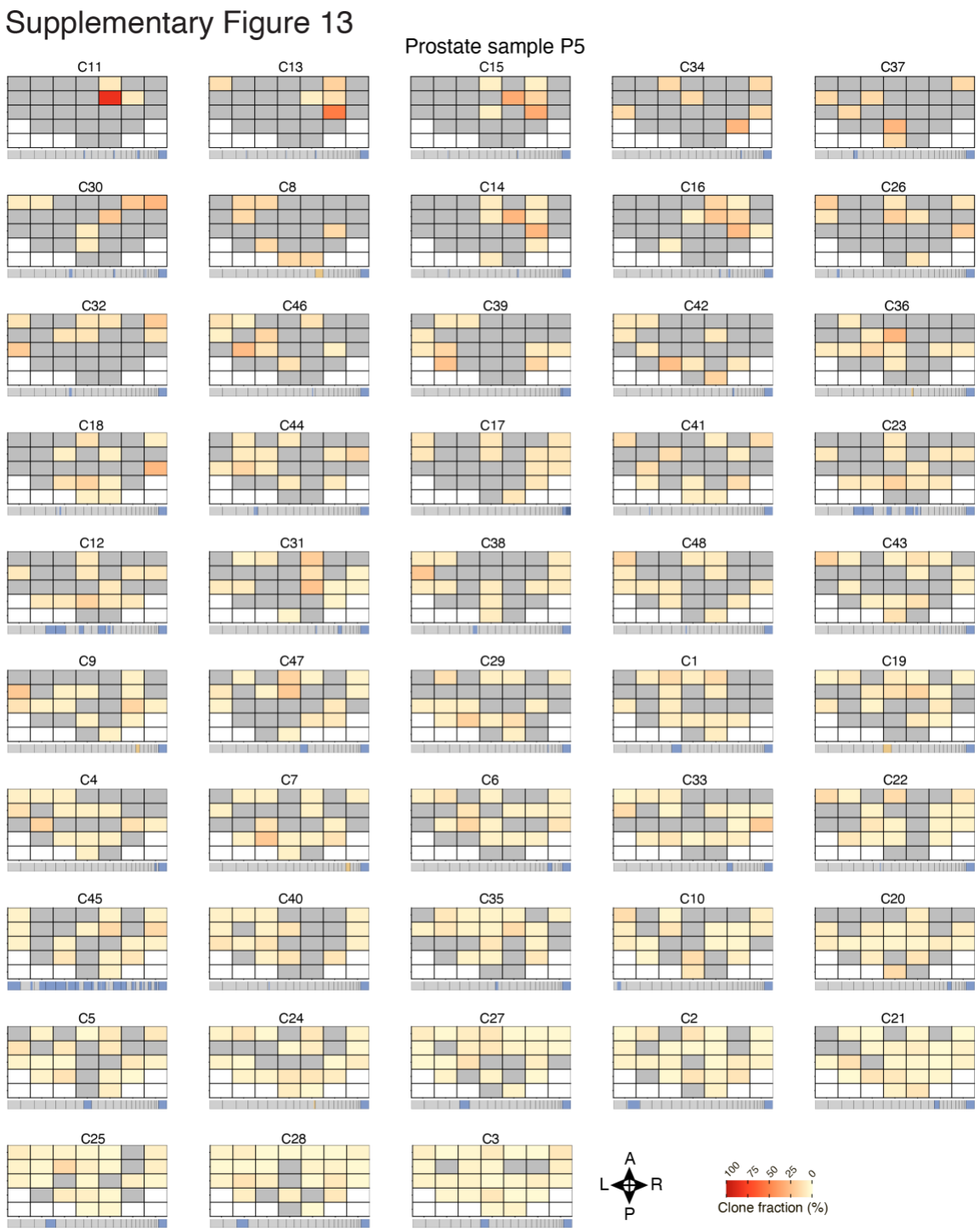
anatomical orientation of the tissues is shown by the four arrows on the right. A, anterior. P, posterior. L, left. R, right.

Supplementary Figure 12



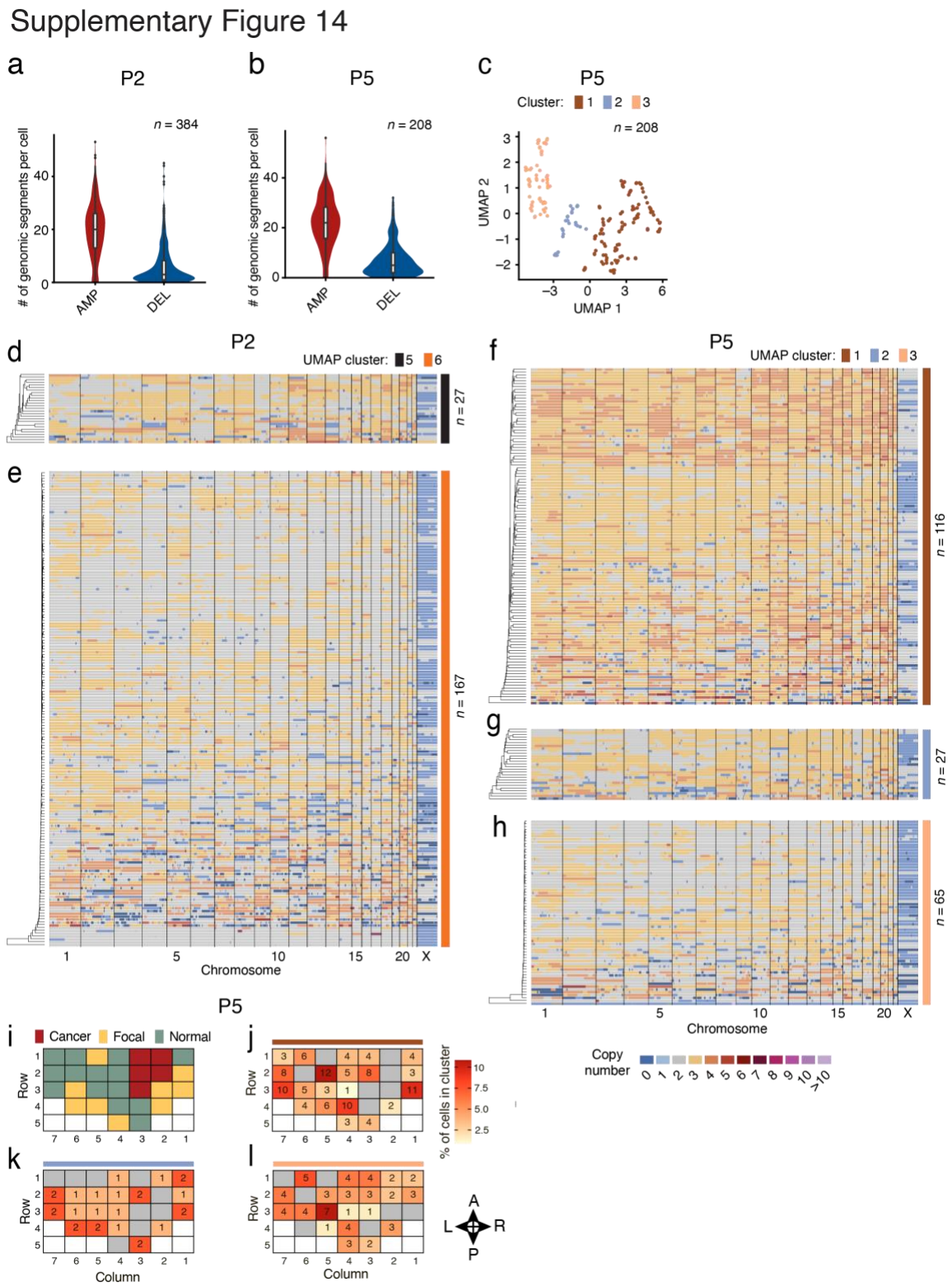
Supplementary Fig. 12. Spatial distribution of pseudo-diploid clones in sample P2. Each grid represents the tissue map of one clone (indicated on top of the grid), and the regions in which the clone was found are color-coded based on the percentage of cells belonging to that clone. Grey cells indicate tissue regions in which the corresponding clone was not detected. White cells indicate absence of prostate tissue. The bar below each grid shows the genome-wide

median copy number profile of the corresponding clone. Blue bars indicate deletions, orange bars amplifications. Vertical black bars mark the boundaries between consecutive chromosomes. The anatomical orientation of the grids is shown by the four arrows on the bottom right. A, anterior. P, posterior. L, left. R, right.



Supplementary Fig. 13. Spatial distribution of pseudo-diploid cell clones in sample P5. Each grid represents the tissue map of one clone (indicated on top of the grid), and the regions in which the clone was found are color-coded based on the percentage of cells belonging to that clone. Grey cells indicate tissue regions in which the corresponding clone was not detected. White cells indicate absence of prostate tissue. The bar below each grid shows the genome-

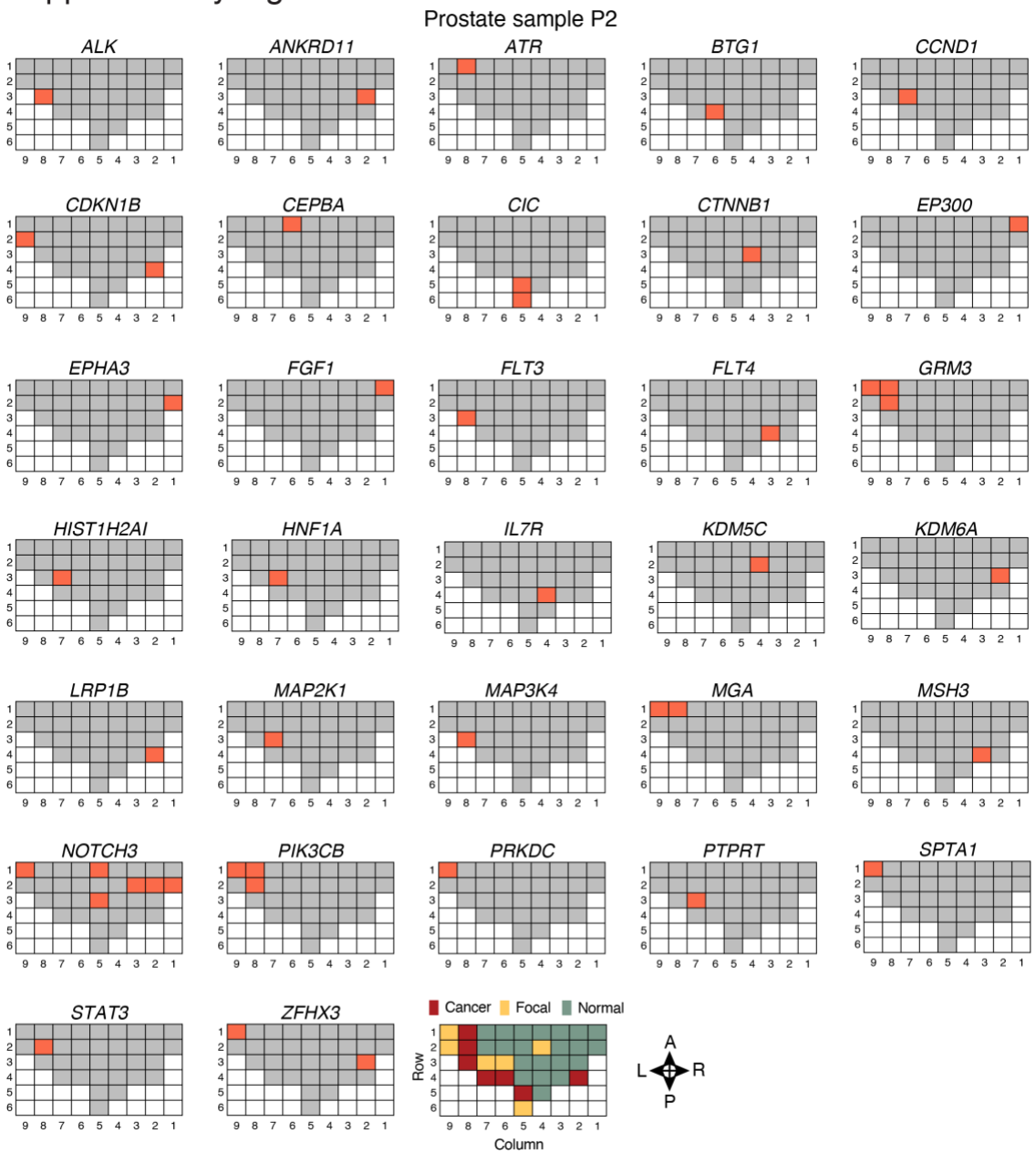
wide median copy number profile of the corresponding clone. Blue bars indicate deletions, orange bars amplifications. Vertical black bars mark the boundaries between consecutive chromosomes. The anatomical orientation of the grids is shown by the four arrows on the bottom right. A, anterior. P, posterior. L, left. R, right.



Supplementary Fig. 14. Spatial distribution of monster cells. (**a**) Distributions of the numbers of genomic segmented either amplified (AMP) or deleted (DEL) across all the monster cells identified in sample P2. In both violin plots, each box spans from the 25th to the 75th percentile and whiskers extend from –1.5×IQR to +1.5×IQR from the closest quartile, where IQR is the

inter-quartile range. Dots: outliers (data falling outside whiskers). (**b**) Same as in (a), but for P5. (**c**) UMAP dimensionality reduction of the CNA profiles of monster cells in P5. *n*, number of cells. Dots of the same color belong to the same cluster. Each dot represents a single cell. (**d, e**) Hierarchically clustered single-cell copy number profiles of monster cells identified in P2 and assigned to UMAP clusters 5 to 6 in **Fig. 4g**. (**f-h**) Hierarchically clustered single-cell copy number profiles of monster cells identified in P5 and assigned to UMAP clusters 1 to 3 in (c). (**i**) Schematic map displaying the pathological classification of each region profiled by scCUTseq in P5 (same as in **Fig. 3e**). (**j**) Map showing the relative frequency and number of monster cells assigned to UMAP cluster 1, in each region of sample P5. (**k, l**) Same as in (j), but for cells assigned to UMAP clusters 2 and 3, respectively. In all the maps in (i-l), each cell in the grid represents a tissue cube from which nuclei were isolated and profiled by scCUTseq (see scheme in **Fig. 3a**). Grey cells indicate tissue regions in which the corresponding clone was not detected. White cells indicate absence of prostate tissue. The anatomical orientation of all the maps in is shown by the four arrows on the bottom right. A, anterior. P, posterior. L, left. R, right.

Supplementary Figure 15



Supplementary Fig. 15. Spatial distribution of mutated cancer genes in the P2 sample in which we performed targeted deep sequencing. In all the maps, each cell in the grid represents a tissue cube from which nuclei were isolated and profiled by scCUTseq (see scheme in **Fig. 3a**). Grey cells indicate tissue regions in which the corresponding mutation was not detected. White cells indicate absence of prostate tissue. A schematic map displaying the pathological classification of each region profiled by scCUTseq in P2 is shown on the bottom right (same as in **Fig. 3b**). The anatomical orientation of all the maps is shown by the four arrows on the bottom right. A, anterior. P, posterior. L, left. R, right.

2. Supplementary Methods

Step-by-step scCUTseq protocol

REAGENTS

- DPBS, with $MgCl_2$ and $CaCl_2$ (Sigma, Cat. No. D8662)
- TrypLE Express Enzyme (1×), phenol red (Thermo Fisher Scientific, Cat. No. 12605036)
- PBS (10×), pH 7.4, (Thermo Fisher Scientific, Cat. No. AM9625)
- EDTA (0.5 M), pH 8.0 (Thermo Fisher Scientific, Cat. No. AM9260G)
- 16% Paraformaldehyde aqueous solution (EMS, Cat. No. 15710)
- Glycine (Fisher Scientific, Cat. No. BP381)
- Nuclease-Free Water (Thermo Fisher Scientific, Cat. No. AM9932)
- Sodium citrate tribasic dihydrate (Sigma, Cat. No. C8532)
- Spermine tetrahydrochloride (Sigma, Cat. No. S1141)
- Tris (Hydroxymethyl) aminomethane (Sigma, Cat. No. 252859)
- IGEPAL CA-630 (Sigma, Cat. No. I8896)
- HCl, molecular biology grade, (Sigma, Cat. No. H1758)
- Trypsin-EDTA (0.25%), phenol red (Thermo Fisher Scientific, Cat. No. 25200072)
- Collagenase (Worthington, Cat. No. CLS-7 LS005332)
- Dispase II (Gibco, Cat. No. 17105-041)
- Trypsin inhibitor from chicken egg white, Type II-O (Sigma, Cat. No. T9253)
- Ribonuclease A from bovine pancreas, Type I-A (Sigma, Cat. No. R4875)
- Methanol (Sigma, Cat. No. 179957)
- Acetic Acid (Sigma, Cat. No. A6283)
- Hoechst 33342 Solution (20 mM, 12.3mg/ml) (Thermo Fisher Scientific, Cat. No. 62249)
- Absolute Ethanol (VWR, Cat. No. 20816.367)
- Vapor-Lock (Qiagen, Cat. No. 981611)
- MALBAC Single Cell WGA Kit (Yikon Genomics, Cat. No. YK001A)
- SYBR Green I Nucleic Acid Gel Stain (10,000X) (Thermo Fisher Scientific, Cat. No. S7563)
- CutSmart buffer (NEB, Cat. No. B7204S)
- NlaIII-HF (NEB, Cat. No. R0125L)
- scCUTseq oligonucleotide adapters (ordered from IDT)
- Rapid DNA Ligation Kit (Thermo Fisher Scientific, Cat. No. K1423)

- T4 DNA Ligase (Thermo Fisher Scientific, Cat. No. EL0011)
- UltraPure BSA (50 mg/mL) (Thermo Fisher Scientific, Cat. No. AM2616)
- ATP Solution (100mM) (Thermo Fisher Scientific, Cat. No. R0441)
- MEGAscript T7 Transcription Kit (Thermo Fisher Scientific, Cat. No. AM1334)
- DNase I, RNase-free (Thermo Fisher Scientific, Cat. No. AM2222)
- RA3 adaptor and RTP, RP1 and RPI primers (custom-made by IDT based on the sequences in the TruSeq Small RNA Library Preparation kit, Illumina)
- RNaseOUT Recombinant Ribonuclease Inhibitor (Invitrogen, Cat. No. 10777019)
- T4 RNA ligase 2, truncated (NEB, Cat. No. M0242L)
- SuperScript IV Reverse Transcriptase (Thermo Fisher Scientific, Cat. No. 18090050)
- NEBNext UltraII Q5 PCR Mastermix (NEB, cat. no. M0544L)
- Agencourt RNAClean XP (Beckman Coulter, Cat. No. A63987)
- Agencourt AMPure XP (Beckman Coulter, Cat. No. A63881)

CONSUMABLES

- Sterile Disposable Scalpel (No. 10) (VWR, Cat. No. 233-5363)
- Eppendorf DNA LoBind tubes 5 mL (Sigma, Cat. No. EP0030108310)
- Eppendorf DNA LoBind tubes 0.5 mL (Sigma, Cat. No. EP0030108035)
- Eppendorf DNA LoBind tubes 1.5 mL (Sigma, Cat. No. EP0030108051)
- 15 mL conical tubes (Fisher Scientific, Cat. No. 430790)
- Falcon 5 mL Round Bottom Polystyrene Tube, Sterile (Fisher Scientific, Cat. No. 10100151)
- Sapphire Filter tips, low retention (Greiner Bio-One, Cat. No. 771265, 773265, 738265, 750265)
- 50 μm CellTrics cell strainer (Sysmex, Cat. No. 04-004-2327)
- 20 μm CellTrics cell strainer (Sysmex, Cat. No. 04-004-2325)
- 10 μm CellTrics cell strainer (Sysmex, Cat. No. 04-004-2324)
- microTUBE-50 AFA Fiber Screw-Cap (25) (Covaris, Cat. No. 520166)
- 8 microTUBE-50 AFA Fiber Strip V2 (12) (Covaris, Cat. No. 520174)
- 96-well plates (Thermo Fisher Scientific, Cat. No. 4316813)
- 384-well plates (Thermo Fisher Scientific, Cat. No. 4483320)
- Nunc Rectangular Dish (Thermo Fisher Scientific, Cat. No. 267060)
- Qubit Assay Tubes (Thermo Fisher Scientific, Cat. No. Q32856)
- Qubit RNA BR Assay Kit (Thermo Fisher Scientific, Cat. No. Q10211)
- Qubit dsDNA BR Assay Kit (Thermo Fisher Scientific, Cat. No. Q32850)

- Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Cat. No. Q32851)
- Bioanalyzer High Sensitivity DNA Kit (Agilent, Cat. No. 5067-4627)

EQUIPMENT

- Cell counter (Invitrogen Countess II FL Automated Cell Counter or equivalent)
- Centrifuge (Eppendorf Microcentrifuge 5424 and 5810 or equivalent)
- Fluorescent Cell Imager (BIO-RAD ZOE Fluorescent Cell Imager or equivalent)
- Sample mixer (Thermo Scientific Tube Revolver Rotator or equivalent)
- Cell sorter (BD FACSAria Fusion or equivalent)
- Non-contact liquid handler (Dispendix GmbH I.DOT)
- Incubator (Binder GmbH KB 53 or Boekel Scientific 241000 or equivalent)
- Thermoshaker (Eppendorf Thermomixer F or equivalent)
- PCR cycler for 384-well (QuantStudio 5 384-well or equivalent)
- PCR cycler for tubes or strips (Analytik Jena Biometra TRIO or equivalent)
- Sonication device (Covaris ME220 Focused-ultrasonicator or equivalent)
- SpeedVac vacuum concentrator (Savant SpeedVac DNA 130 or equivalent)
- Magnetic stand (Invitrogen DynaMag-2/-5 Magnet or equivalent)
- Fluorometers (Qubit 2.0 Fluorometer or equivalent)
- Fragment analyzer (Agilent 2100 Bioanalyzer or equivalent)

PROCEDURE

Cell preparation

1. Wash the cells with 1× DPBS containing $MgCl_2$ and $CaCl_2$ pre-warmed to 37 °C and aspirate the DPBS from the flask
2. Add TrypLE Express Enzyme and incubate at 37 °C to dissociate cells
3. Add equal volume of culture medium to neutralize the enzyme and resuspend the cells
4. Count cell number and take 1×10^6 cells

NOTE: For suspension cells, step 1–3 can be skipped

5. Pellet cells at 100–400×g for 5 min (differs per cell type)
6. Wash cells with 1× PBS/5 mM EDTA once
7. Pellet cells at 100–400×g for 5 min

NOTE: For live cells, step 8–12 can be skipped for immediate sorting

8. Resuspend cells in 3 mL 1× PBS/5 mM EDTA, incubate at room temperature for 1 min
9. Add 3 mL of 8% PFA (in 1× PBS/5 mM EDTA) and pipette up and down few times

NOTE: Using 5% BSA pre-coated tubes can reduce cell loss cause by fixation step

10. Incubate for 10 min in darkness
11. Add 315 µL of 2.5 M Glycine (the final concentration is 125 mM)
12. Pellet cells at 300×g for 5 min
13. Resuspended cells in 1 mL 1× PBS/5mM EDTA (final density ~1×10^6/mL) and transfer the cell suspension to BD FACS tubes
14. Keep the live/fixed cells on ice for immediate use, or store the fixed cells in 1× PBS/5 mM EDTA/0.05% NaN_3 at 4 °C for up to three months

Nuclei extraction from frozen tissue

1. Before each processing, incubate Spermine Solution and Stop Solution at room temperature for at least 10 min. Thaw the enzymatic stock solutions and place them on ice
2. Prepare a dry ice bucket
3. For each sample to be processed, pre-chill on dry ice a pair of sterile scalpels and one sterile Petri dish
4. Prepare fresh Tissue Lysis Solution before each use (volumes for one sample)

Trypsin-EDTA (0.25%), phenol red	24 µL
Collagenase (8 mg/mL)	25 µL
Dispase II (8 mg/mL)	25 µL
Spermine Solution pH 7.6	1.926 mL

5. Retrieve the cryovial containing the tissue fragment(s) from –80 °C freezer and always keep on dry ice
6. Transfer the tissue fragment into the pre-chilled Petri dish placed on top of the dry ice
7. From the 2 mL of tissue lysis solution, pipette 200 µL on top of the tissue fragment
8. Incubate for 2–3 minutes until the tissue lysis solution on top of the tissue fragments freezes

NOTE: The freezing step in tissue lysis solution provides a more homogeneous tissue density and increases the efficiency of tissue mincing, enzymatic disaggregation, and nuclei recovery

9. Using a chilled disposable sterile scalpel, mince the tissue fragment(s) embedded in Tissue Lysis Solution thoroughly
10. Remove the Petri dish (containing the minced tissue and Tissue Lysis Solution) from dry ice and bring to room temperature

11. Continue mincing the tissue until the Tissue Lysis Solution is completely thawed and the tissue fragments can flow through a 1 mL tip without clogging
12. Use the remaining 1800 µL of Tissue Lysis Solution to rinse and transfer all the tissue fragments from the Petri dish back to a 5 mL LoBind tube
13. Place the tube on a sample mixer and incubate at room temperature for 15 min, rotating at 20 rpm (or very low speed). Check the tissue digestion every 5 min

NOTE: The tissue that is not successfully digested in 15 min, will likely not digest with longer incubation

14. Add 2 mL of Stop Solution and mix by inverting the tube gently
15. Wash the 50 µm cell strainer with 1 mL of spermine solution
16. Filter the nuclei suspension through the 50 µm cell strainer and collect the flowthrough in a 5 mL LoBind tube
17. Centrifuge the flowthrough containing the nuclei at 300×g for 5 min at room temperature

NOTE: Increasing the centrifugation speed or time will result in increased nuclei clumping and lead to excessive amounts of cellular debris

18. Carefully discard the supernatant and resuspend the nuclei in 400 µL of Nuclei Fixation Solution, by pipetting up and down

NOTE: For live nuclei, the nuclei fixation step can be skipped for immediate sorting

19. Incubate the nuclei suspension on ice for 15 min
20. Centrifuge the nuclei suspension at 300×g for 5 min at room temperature
21. Carefully discard the supernatant and resuspend the nuclei in 1 mL 1× PBS/5 mM EDTA
22. Filter the nuclei suspension through either 10 or 20 µm cell strainer into BD FACS tubes (test which strainer is more suitable for the nuclei of each tissue)
23. Keep the nuclei suspension on ice for immediate use, or store the fixed nuclei suspension in 1× PBS/5 mM EDTA/0.05% NaN_3 at 4 °C for up to two months

Single cell / nuclei sorting

NOTE: Independently of which FACS machine is used, always set the nozzle size to 100 µm and make sure the FACS droplet volume is below 10 nL. Excessive volume will dilute the downstream reactions and excessive EDTA will inhibit enzymatic activities. 384-well plates are better suited when dispensing nanoliter volumes as the conical bottom and size of the wells allows easier visualization of the dispensed droplets inside the Vapor-Lock (oil phase).

<u>For live or fixed cells</u>

1. Dilute the Hoechst stock solution (12.3 mg/ul) 1:10 with water. Add 2 μL of the diluted Hoechst solution to each 1 mL of cell suspension (the final concentration will be 2.46 ng/ul) and mix gently and thoroughly
2. Incubate the cell suspension for at least 30–40 min at 37 °C, while shaking at 600 rpm in darkness
3. After staining, the tubes can be kept on ice in darkness until sorting

For live or fixed nuclei

1. Dilute the Hoechst stock solution (12.3 mg/ μL) 1:10 with water. Add 2 μL of the diluted Hoechst solution to each 1 mL of nuclei suspension (the final concentration will be 2.46 ng/μL) and mix gently and thoroughly
2. Incubate the nuclei suspension for 2 min on ice in darkness
3. After staining, we recommend keeping the tubes on ice or at +4 °C in darkness until sorting

Sorting

1. Aliquot 5 μL of Vapor-Lock into each well inside the targeted region of 384-well plates, using a multiple-channel pipette
2. Bring the FACS tubes, 384-well plates pre-filled with Vapor-Lock and the adhesive films to the FACS machine
3. When sorting, set 10 nL as the FACS droplet volume and 100 μm as the FACS nozzle size
4. Leave wells empty randomly to be used for positive or negative controls
5. After sorting, seal each 384-well plate immediately with adhesive film and spin down at 300×g for 5 min. Store the plates at 4 °C for intermediate use, or place them at –20 °C for long-term storage.

Preparation of positive and negative controls

1. Pre-warm the 384-well plates at 4 °C for 10 min if they were stored at –20 °C
2. Serially dilute bulk genomic DNA (gDNA) extracted from the same sample from which the single cell or nuclei suspension was prepared down to a concentration of 2 ng/μL. This serves as positive control
3. Dispense positive and negative controls using I.DOT:
   i. For positive controls, dispense 10 nL of 2 ng/μL gDNA (20 pg of gDNA per well)
   ii. For negative controls, dispense 10 nL of Nuclease-Free Water

MALBAC Cell Lysis

NOTE: The following MALBAC reagent volumes are intended for a single well of a 384-well plate. We recommend preparing a mix for at least 48 wells every time, as otherwise the volume would be too low to be handles with a manual pipette

1. Prepare the Cell Lysis Reaction Mix by combining the following components and mix well by pipette up and down before adding the mix to the source well in I.DOT

Cell Lysis Buffer	30 nL
Cell Lysis Enzyme	0.6 nL

2. Dispense 30 nL per well. Shake and centrifuge the plate

NOTE: From now on, shake the plate in a ThermoMixer at 1,000 rpm for 1 min and centrifuge at 3,220×g for 5 min after each dispensing round on I.DOT

3. Place the plate in a PCR thermocycler with the lid set at 80 °C and run the following program:
    i. 50 °C, 2 h
    ii. 80 °C, 10 min
    iii. 4 °C, hold

NOTE: Lysis at 50 °C can be reduced to 1 h, as we observed no substantial differences. When handling multiple plates together, we recommend using an incubator instead of a thermocycler. Proceed immediately to the next step or keep the plates at –20 °C for long-term storage

MALBAC Pre-Amplification

1. Prepare the Pre-Amplification Reaction Mix by combining the following components and mixing well before adding to the source well in I.DOT:

Pre-Amplification Buffer	150 nL
Pre-Amplification Enzyme Mix	5 nL

2. Dispense 150 nL of Pre-Amplification Reaction Mix per well. Shake and centrifuge the plate
3. Place the plate in a PCR thermocycler and run the following program:
    i. 94 °C, 3 min
    ii. 20 °C, 40 sec
    iii. 30 °C, 40 sec
    iv. 40 °C, 30 sec
    v. 50 °C, 30 sec
    vi. 60 °C, 30 sec

vii. 70 °C, 4 min

viii. 95 °C, 20 sec

ix. 58°C, 10 sec

Go to step ii and repeat 8 times

x. 4 °C, forever

NOTE: Proceed immediately to the next step

MALBAC Exponential Amplification

1. Prepare the Amplification Reaction Mix by combining the following components and mixing well before adding to the source well in I.DOT:

Amplification Buffer	150 nL
Amplification Enzyme Mix	4 nL
SYBR Green I (4X)	11 nL

2. Dispense 160 nL of Amplification Reaction Mix per well. Shake and centrifuge the plate
3. Place the plate in a PCR thermocycler and run the following program:

i. 94 °C , 30 sec

ii. 94 °C, 20 sec

iii. 58 °C, 30 sec

iv. 72 °C , 3 min

Go to step ii and repeat 16 times

v. 4°C, hold

NOTE: Proceed immediately to the next step or place the plates at –20°C for long-term storage. Sample amplification efficiency can be monitored using a real-time thermal cycler by adding SYBR Green I dye to the Amplification Reaction Mix at a final concentration of 0.125X. Data analysis should be performed on raw background-subtracted (not on baseline cycle normalized) fluorescence, and the instrument/software should be set to the appropriate mode. It is not necessary to monitor the MALBAC amplification efficiency for each plate once the method has been set up successfully.

DNA Digestion

1. Prepare the Digestion Mix by combining the following components and mixing well before adding the mix to the source well in I.DOT:

CutSmart Buffer (10×)	50 nL
NlaIII-HF (10,000 U/ml)	100 nL

2. Dispense 150 nL per well. Shake and centrifuge the plate
3. Place the plates in a PCR thermocycler and run the following program:
    i. 37 °C, 1 h
    ii. 65 °C, 20 min
    iii. 4 °C, hold

NOTE: When handling many plates together, we recommend using an incubator instead of a thermocycler. Proceed immediately to the next step

Ligation of scCUTseq adapters

1. Dispense 300 nL of 33 nM scCUTseq adapter per well

NOTE: Use a different CUTseq barcode per well so that the content of multiple wells can be pooled together after ligation

2. Prepare the Ligation Mix by combining the following components and mixing well before adding the mix to the source well in I.DOT:

Reagents when using Rapid Ligase:

Nuclease-Free Water	50 nL
T4 Rapid Ligase Buffer (5×)	300 nL
ATP (10 mM)	120 nL
BSA (50 mg/ml)	30 nL
T4 Rapid Ligase	200 nL

Alternatively, when using Standard Ligase:

Nuclease-Free Water	200 nL
T4 Ligase Buffer (10×)	150 nL
ATP (10 mM)	120 nL
BSA (50 mg/ml)	30 nL
T4 Standard Ligase	200 nL

3. Dispense 700 nL per well. Shake and centrifuge the plate
4. Incubate at 22 °C for 30 min for rapid ligase, and 1h for standard ligase
5. Using a multiple-channel pipette, add 5 µL of EDTA 33 mM in Nuclease-Free Water (the final concentration will be 25 mM) to each well to stop the ligation reaction and increase the volume per well

NOTE: Adding 5 µL of EDTA before pooling the content of multiple wells is critical as it will help increase the recovery and prevent mis-ligation between different adapters during the pooling step

6. Pool the contents of all the wells in the plate by placing the plate upside-down onto a rectangular plastic dish (Nunc) covered with Parafilm
7. Seal the aligned skirts with tape and centrifuge at 120×g for 1 min
8. Collect all the liquid in the dish and transfer it into one 5 mL DNA LoBind tube
9. Briefly centrifuge the tube and carefully discard the upper oil phase of the liquid (containing the Vapor-Lock)
10. Measure the volume of the lower aqueous phase with a pipette

NOTE: Proceed immediately to the next step

DNA cleanup

1. Mix the aqueous solution containing DNA with a 1.2× vol./vol. ratio of AMPure XP beads pre-warmed at room temperature
2. Mix thoroughly and incubate for 10 min at room temperature
3. Place the sample on a magnetic stand
4. Incubate for at least 10 min until the liquid becomes clear
5. Remove and discard the liquid
6. Wash the beads twice with freshly prepared 80% ethanol at room temperature (the volume of freshly prepared ethanol should be sufficient to cover the beads)
7. Air-dry the beads for 5–10 min at room temperature

NOTE: Do not over-dry the beads, since this may result in a lower DNA yield

8. Remove the sample from the magnetic stand
9. Resuspend the beads in 200 µL of Nuclease-Free Water
10. Incubate for 2 min at room temperature
11. Place the sample back on the magnetic stand
12. Incubate for at least 5 min until the liquid becomes clear
13. Transfer 190 µL of the liquid into a clean 1.5 mL DNA LoBind tube
14. Check the DNA concentration using 1 µL of DNA solution and the Qubit DNA HS kit

NOTE: Proceed immediately to the next step or place the samples at –20 °C for long term storage

Sonication

1. Transfer 250 ng of purified DNA into microTUBE-50 AFA Fiber Screw-Cap Covaris tube and add Nuclease-Free Water to it to bring the volume up to 55 µL

NOTE: We usually load one Covaris tube with no more than 500 ng of DNA, as loading more DNA would result in incomplete fragmentation

2. Sonicate the sample on a Covaris ME220 Focused-ultrasonicator

NOTE: We typically set a target peak of 200 bp. We perform sonication even though genomic DNA has already been fragmented by the restriction enzyme used in scCUTseq, since we have found that this results in libraries of higher quality and in more reproducible results

SpeedVac

1. Transfer the sonicated DNA sample from the Covaris tube into a clean 0.5 mL DNA LoBind tube
2. Concentrate the sample using a SpeedVac Vacuum Concentrator on medium heating mode until the volume of the sample reaches 8 µL

In vitro transcription (IVT)

1. Mix the following reagents on ice, and add 12 µL of the mix to the sample:

rNTP mix	8 µL
T7 Polymerase Buffer (10×)	2 µL
T7 Polymerase	1.5 µL
RNaseOUT Recombinant Ribonuclease Inhibitor	0.5 µL

NOTE: We prepare the rNTP mix by mixing equal volumes of each separate rNTP solution provided in the MEGAscript T7 Transcription Kit

2. Incubate the sample for 14 hours at 37 °C in a PCR thermocycler with the lid set at 70 °C

NOTE: 2h of incubation also work well. Proceed immediately to the next step

RNA cleanup

1. Add 1 µL of DNase I (RNase-free) to the IVT product
2. Incubate for 15 min at 37 °C
3. Bring the volume up to 50 µL by adding 29 µL of Nuclease-Free Water, then mix with 90 µL (1.8× vol./vol.) of RNAClean XP beads pre-warmed at room temperature
4. Mix thoroughly and incubate for 10 min at room temperature
5. Place the sample on a magnetic stand
6. Incubate for at least 5 min until the liquid becomes clear
7. Remove and discard the supernatant
8. Wash the beads twice with 200 µL of freshly prepared 70% ethanol at room temperature

9. Air-dry the beads at room temperature

NOTE: Do not over-dry the beads, since this may result in a lower RNA yield

10. Remove the sample from the magnetic stand
11. Resuspend the beads in 10 µL of Nuclease-Free Water
12. Incubate for 2 min at room temperature
13. Place the sample back on the magnetic stand
14. Incubate for at least 5 min until the liquid becomes clear
15. Transfer 8.8 µL of the liquid to a clean 0.5 mL DNA LoBind tube
16. Check the RNA concentration using 1 µL of RNA solution and the Qubit-BR RNA kit

NOTE: Proceed immediately to the next step

RA3 adapter ligation

1. Add 1 µL of 10 µM RA3 adapter to 7.8 µL of purified RNA
2. Incubate for 2 min at 70 °C in a PCR thermocycler, then immediately place the sample on ice
3. Add 3.2 µL of the following mix:

RNA Ligase Buffer (10×)	1.2 µL
RNaseOUT Recombinant Ribonuclease Inhibitor	1 µL
T4 RNA Ligase, truncated	1 µL

4. Incubate for 2 hours at 25 °C in a PCR thermocycler with the lid set at 30 °C

NOTE: Proceed immediately to the next step

Reverse transcription (RT)

1. Add 2 µL of 10 µM RTP primer
2. Incubate for 2 min at 70 °C in a PCR thermocycler, then immediately place the sample on ice
3. Add 11 µL of the following mix:

SSIV Buffer (5×)	5 µL
dNTPs (12.5 mM)	1 µL
DTT (0.1 M)	2 µL
RNaseOUT Recombinant Ribonuclease Inhibitor	1 µL
SSIV Reverse Transcriptase	2 µL

4. In a PCR thermocycler, perform the following program:
    i. 50 °C, 20 min

ii. 65 °C, 10 min

iii. 4 °C, hold

NOTE: Proceed immediately to the next step or place the samples at –20°C for long term storage

Library indexing and amplification

1. Add 8 µL of the desired indexed Illumina primer at 10 µM
2. Add 167 µL of the following mix:

Nuclease-Free Water	59 µL
NEBNext UltraTM II PCR Master Mix	100 µL
RP1 primer (10 µM)	8 µL

3. Split each sample into 4 PCR tubes (each tube should contain 50 µL)
4. In a PCR thermocycler perform the following program:
   i. 98 °C, 30 sec
   ii. 98 °C, 10 sec
   iii. 60 °C, 30 sec
   iv. 65 °C, 45 sec
   Go to step ii and repeat 10 times
   v. 65 °C, 5 min
   vi. 4 °C, hold

NOTE: Adjust the number of PCR cycles based on the input in the IVT step. Proceed immediately to the next step or place the samples at –20°C for long term storage

Library cleanup

1. Pool the content of the 8 PCR tubes into a clean 1.5 mL LoBind tube
2. Add 0.8× vol./vol. ratio of AMPure XP beads pre-warmed at room temperature
3. Mix thoroughly and incubate for 10 min at room temperature
4. Place the sample on a magnetic stand
5. Incubate for at least 5 min until the liquid becomes clear
6. Remove and discard the supernatant
7. Wash the beads twice with 1 mL of freshly prepared 80% ethanol at room temperature
8. Air-dry the beads at room temperature

NOTE: Do not over-dry the beads, since this may result in a lower DNA yield

9. Remove the sample from the magnetic stand

10. Resuspend the beads in 50 µL of Nuclease-Free Water
11. Incubate for 2 min at room temperature
12. Place the sample back on the magnetic stand
13. Incubate for at least 5 min until the liquid becomes clear
14. Transfer the solution into a clear 1.5 mL DNA LoBind tube
15. Check the library concentration using the Qubit dsDNA HS kit
16. Check the fragment distribution on a Bioanalyzer 2100 using a DNA HS chip

NOTE: Place the libraries at –20 °C for long term storage

Standard MALBAC

We washed the harvested cells in 1× PBS/5 mM EDTA at room temperature and resuspended them in the same buffer at a density of 10^6/mL. We then placed the cells either on ice for immediate sorting (live cells) or fixed them to store them for longer periods before sorting. For fixation, we added an equal volume of 1× PBS/5 mM EDTA/8% methanol-free paraformaldehyde (PFA, Thermo Fisher Scientific, cat. no. 28908) to the cell suspension and pipetted the solution up and down ten times. After 10 min incubation in darkness, we added 2.5 M glycine to each tube to reach a final concentration of 125 mM to quench any residual unreacted PFA. We then washed the cells in 1× PBS/5 mM EDTA at room temperature, resuspended them at a density of 10^6/mL in the same buffer, and stored them at +4 °C in 1× PBS/5 mM EDTA/0.05% NaN_3. We have successfully used fixed cells prepared in this manner, that were kept at +4 °C several weeks up to three months. Before sorting, we transferred the cells into FACS-compatible tubes, stained them with 2.46 ng/mL Hoechst 33342 (Thermo Fisher Scientific, cat. no. 62249) and incubated for 40 min at 37 °C in darkness, while rotating. We sorted fixed and live single cells in a 96-well plate pre-filled with MALBAC lysis reaction mix (5 µL of lysis buffer + 0.5 µL of lysis enzyme) using a BD FACSJazz Cell Sorter (BD Biosciences) based on forward and side scatter properties. We centrifuged the plate at 600×g for 3 min. In parallel we used one well as negative control by adding 1 µL of Nuclease-Free Water (Thermo Fisher Scientific, cat. no. 4387936) into the lysis mix and another well as positive control by adding 1 µL of gDNA extracted from fixed cells at a final concentration of 30 pg/µL into the lysis mix. We incubated the plate at 50 °C for 1 h and 80 °C for 10 min in a PCR thermocycler. After lysis, we added 31 µL of MALBAC pre- amplification reaction mix (30 µL of pre-amp buffer + 1 µL of pre-amp enzyme mix) into each well, and centrifuged the plate at 600×g for 3 min followed by incubation at 94 °C for 3 min to denature DNA and then by 8 cycles of quasilinear amplification (20 °C for 40 s, 30 °C for 40 s, 40 °C for 30 s, 50 °C

for 30 s, 60 °C for 30 s, 70 °C for 4 min, 95 °C for 20 s and 58 °C for 10 s) in a PCR thermocycler. After pre-amplification, we added 30.8 µL of MALBAC amplification mix (30 µL of amp buffer + 0.8 µL of amp enzyme mix) into each well, centrifuged the plate at 600×g for 3 min followed by incubation at 94 °C for 30 s to denature DNA and then by 14 cycles of exponential amplification (94 °C for 20 s, 58 °C for 30 s, 72 °C for 3 min) in a PCR thermocycler. We performed lysis, pre-amplification and amplification steps using reagents included in the MALBAC kit (Yikon Genomics, cat. no. Y001A). We purified the amplified material from each well separately using 1.8 vol./vol. ratio of Agencourt Ampure XP beads (Beckman Coulter, Cat. No. A63881) and measured the concentration using Qubit DNA HS kit (Thermo Fisher Scientific, Cat. No. Q32851). We spared 200 ng of DNA for the following steps. First, we sheared the samples using Covaris ME220 Focused-ultrasonicator with a target peak set at 200 base pairs (bp) and then performed library preparation using the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB, cat. no. E7645S). We checked the fragment distribution of the final libraries on a Bioanalyzer 2100 using DNA HS chip (Agilent, cat. no. 5067-4626). To reduce the MALBAC reagent volumes and therefore the cost per cell, we tested different scaled-down versions of MALBAC (sMALBAC) by reducing the reagent volumes in the lysis, pre-amplification, and amplification steps 50, 100, 200 or 500 times. To this end, we first sorted fixed or live single cells in a 384-well plate pre-filled with 5 µL of Vapor-Lock (Qiagen, cat. no. 981611) and dispensed the scaled reagent volumes using the I.DOT nanodispenser (CELLINK) in 384-well plates. We then collected the MALBAC products manually from each well by adding 5 µL of Nuclease-Free Water to each well, and then fragmented each sample, and prepared sequencing library as described above.

Cas9 expressing TK6 cell line

We cultured TK6 lymphoblastoid cells in RPMI-1640 medium with 5 % horse serum (Gibco, cat. no. 11510516), supplemented with 2 mM L-glutamine (Gibco, cat. no. A2916801), 100 U/ml penicillin/100 µg/ml streptomycin (Gibco, cat. no. 15140122), and 100 µM sodium-pyruvate (Gibco, cat. no. 11360070) at 37 °C and 5% CO2. To establish TK6-Cas9 cells stably expressing SpCas9, we transduced the parental cells with viral particles produced using a Cas9 lentiviral expression vector (pLenti-Cas9 Blast, Addgene cat. no. 52962, kind gift from Feng Zhang). Briefly, we produced lentiviral particles in HEK293T cells by transfecting them with 4 µg of the Cas9 lentiviral expression vector and 1 µg each of lentiviral packaging plasmids, pMDLg/pRRE, pRSV-REV, pMD2.G (Addgene, cat. no. 12251, 12253, and 12259, respectively, all kind gifts from Didier Trono) using the X-tremeGENE HP DNA Transfection

Reagent (Roche, cat. no. 6366244001). We pooled the virus-containing supernatant on two consecutive days from two 10 cm dishes of HEK293 transfected with the same plasmids and concentrated the supernatant 100 times using Lenti-X Concentrator (TaKaRa, cat. no. 631231). We then added the concentrated virus to 106 TK6 cells together with 8 µg/ml of Polybrene (Sigma-Aldrich, cat. no. TR-1003-G). One day after infection, we exchanged the medium and added 5 µg/ml Blasticidin to select the cells that had been successfully transfected. We prepared single-cell clones from limiting dilutions in 96-well plates and checked each clone for Cas9 expression and activity by immunoblotting and immunofluorescence microscopy using an anti-Cas9 antibody (Active Motif, cat. no. 61577), and by performing a T7 endonuclease assay following electroporation of the cells with specific gRNAs.

Generation of a 7 Mb deletion on chr11 using CRISPR/Cas9

We targeted the *KMT2A* (hg38 chr11:118488514-118488533) and *HYLS1* (hg38 chr11:125899369-125899388) loci, which are approximately 7 Mb apart on the q-arm of chr11. We purchased guide RNAs (gRNAs) (*KMT2A*: TTTGGGTTTTAGTAGTCCAC; *HYLS1*: ATGGAAGAACTTCTACCTGA) from Dharmacon and assembled them into active complexes by adding crRNA-tracrRNA components (Dharmacon, cat. no. U-002005-5) according to the manufacturer's instructions. We electroporated the sgRNA complexes to TK6-Cas9 cells established as described above, using the Neon Transfection System (Thermo Fisher Scientific, cat. no. MPK10025) with the following program: 1350 V, 10 msec, 3 pulses. We used 106 cells with a final amount of 1.6 nmol of gRNA in each electroporation. Afterwards, we kept the cells in growing medium for two days at 37 °C and 5% CO2 to allow for the 7 Mb genomic region between the *KMT2A* and *HYLS1* loci to be deleted in a fraction of the cells.

High-throughput DNA FISH

We prepared FISH probes targeting the KMT2A-5′ (Thermo Fisher, cat. no. CTD2159M9) KMT2A-3′ (BACPAC Resources CHORI, cat. no. RP11-59N1), HYLS1-5′ (BACPAC Resources CHORI, cat. no. RP11-712D22) and HYLS-3′ (BACPAC Resources CHORI, cat. no. RP11-50B3) loci using bacterial artificial chromosomes (BACs) and performing nick translation with the Nick Translation kit (Abbott Molecular, cat. no. 7J0001) and using the following fluorescently labeled dUTPS: AlexaFluor 488-5-dUTP (Thermo Fisher Scientific, cat. no. C11397); AlexaFluor 568-5-dUTP (Thermo Fisher Scientific, cat. no. C11399); AlexaFluor 647-AHA-dUTP (Thermo Fisher Scientific, cat. no. A32763); CF405S-dUTP

(Biotium, cat. no. 40004). We plated TK6-Cas9 cells onto poly-L-lysine coated 96-well glass-bottom imaging plates (PerkinElmer, cat. no. 6055308) and spun the plates at 400×g in a swing-out rotor centrifuge for 20 sec to allow for cells to get attached to the glass. Cells were fixed in 1× PBS/4% paraformaldehyde (PFA) (Thermo Fisher Scientific, cat. no. 158127) for 15 min at room temperature and washed the cells with 1× PBS three times to remove the unreacted PFA. We permeabilized the cells in 1× PBS/0.5% saponin (Sigma-Aldrich, cat. no. 47036)/0.5% Triton X-100 (Sigma-Aldrich, cat. no. T8787) for 20 min at room temperature, followed by two washes in 1× PBS and then 15 min incubation in 0.1 N HCl at room temperature. We then washed the cells in 2× SSC at room temperature and incubated in them in 2× SSC/50% Formamide (Sigma-Aldrich, cat. no. F9037) for 20 min at room temperature. We precipitated 80 ng of each labelled BAC probe by adding 3 µg of COT-1 DNA (GeneON, cat. no. 3001), 20 µg of tRNA (Invitrogen cat. no. 10702487) and 2 volumes of 100% ethanol. Probes were centrifuged at 16,000×g for 20 min at 4 °C. We resuspended the DNA pellets in 30 µL of hybridization mix containing 2× SSC/50% formamide/10% dextran sulfate (Sigma-Aldrich, cat. no. 67578)/1% Tween-20 (Sigma-Aldrich, cat. no. P9416) and added 30 µL of each probe in hybridization mix to a single well of a 96-well plate. We performed denaturation performed at 85 °C for 10 min on a slide moat (Boekel Scientific, cat. no. 10630394), followed by plate spinning at 400×g for 20 sec. We then placed the plate in humidified chamber and incubated overnight at 37 °C. The next day, we washed the wells three times with 1× SSC at 45 °C, 5 min each, followed by three consecutive washes with 0.1× SSC at 45 °C, 5 min each. Finally, we washed the wells once with 1× PBS at room temperature and stored the plate at 4 °C until imaging. We imaged the plates on the Opera Phenix high content screening confocal microscope (PerkinElmer) operated by the Harmony 4.8 software, using a 40× NA=0.8 water immersion lens (Olympus) and a 1.3 Megapixel CCD camera with pixel binning 2, corresponding to a pixel size of 299 nm. For each condition (transfection with KMT2A/HYLS1 sgRNAs or non-targeting sgRNA control), we imaged 50 fields with 11 planes in *z* per field in three technical triplicates per experiment and two biological experiments.

Calculation of the fraction of TK6-Cas9 edited cells

To calculate the percentage of TK6-Cas9 cells in which the expected 7 Mb deletion between *KMT2A* and *HYLS1* had occurred, we first segmented nuclei in the DNA FISH images based on the fluorescence background signal in maximally projected images, using a custom-made pipeline built in the Harmony software. We performed spot detection of all four colored FISH

spots in different channels using a built-in analysis block in the Harmony software. To determine 3D distances between spots, we calculated the Euclidean spot-to-spot distances from the position of spots in maximally projected images in *x,y* and the distances between the *z*-planes of their brightest pixels. We corrected the distances between the *z* planes of spots in different colors for shifts due to chromatic aberration in *z*, by determining the offset of spot detection between the channels in *z* for one genomic locus simultaneously stained with all four different colored probes. We performed all the calculations using custom made R scripts (available upon request) with text files of analyses derived from the Harmony software as input. We called cells as harboring the 7 Mb deletion when the KMT2A-3′ and HYLS1-5′ probes (see scheme in **Supplementary Fig. 2a**) were not detected, and the distance between the KMT2A-5′ and HYLS1-3′ probes was smaller than a threshold set based on their average distance in non-targeted control cells. Similarly, we called cells as harboring the deletion plus chromosome 11 arm loss, when all probes were detected only once per cell, except the KMT2A-5′, which was detected twice. Finally, we called cells as harboring the deletion plus an amplification of 3′ arm of chromosome 11, when cells had two KMT2A-5′ probes, one KMT2A-3′ probe, one HYLS1-5′ probe and three HYLS1-3′ probes.

Targeted gene sequencing

We prepared libraries for targeted deep sequencing using the Illumina TruSight Oncology 500 panel (Illumina, cat. no. 20040765). This panel is 1.94 Mb in size, encompassing the full coding sequencing of 523 cancer-related genes (coding size: 1.2 Mb). We sonicated 80 ng of genomic DNA extracted from each region in the P2 sample using a Covaris Focused-ultrasonicator (Covaris) and then prepared libraries and performed two rounds of hybridization-based target capture following the manufacturer's instructions. We sequenced all the libraries on Illumina NovaSeq 6000 aiming at reaching minimum 500X read depth. We processed raw data using the TruSight Oncology 500 v2.2 Local App (Illumina) to generate fastq files by aligning the reads to the human reference sequence GRCh37 (hg19). We used the same application to perform QC and somatic variant calling using the tumor-only pipeline.

3. Supplementary Tables

Supplementary Table 1. Summary of sequencing runs. Because of its size, this table is provided as a separate Excel file.

Supplementary Table 2. List of metrics describing manually annotated CNA profiles used to train a random forest classifier. Because of its size, this table is provided as a separate Excel file.

Supplementary Table 3. Pathological characteristics of the prostate samples used in this study. Because of its size, this table is provided as a separate Excel file.

Supplementary Table 4. Locality index of the pseudo-diploid clones identified in prostate samples P2 and P5. Because of its size, this table is provided as a separate Excel file.

Supplementary Table 5. Frequency of amplification or deletion of COSMIC genes in the five most localized pseudo-diploid cell clones in prostate samples P2 and P5. Because of its size, this table is provided as a separate Excel file.

Supplementary Table 6. List of SNVs identified by spatially resolved targeted deep sequencing in prostate sample P2. Because of its size, this table is provided as a separate Excel file.

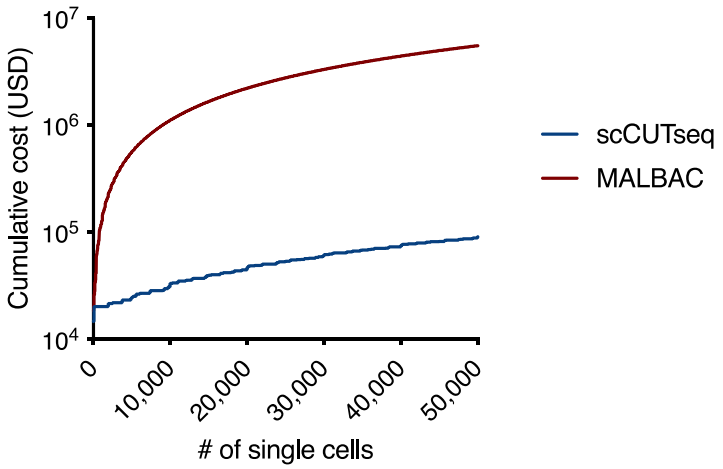
Supplementary Table 7. Sequence of the CUTseq oligonucleotide adapters used in this study. Because of its size, this table is provided as a separate Excel file.

Supplementary Table 8. List of reagents and relative costs for scCUTseq, standard MALBAC or ACT.

4. Supplementary Notes

1. Cumulative cost analysis

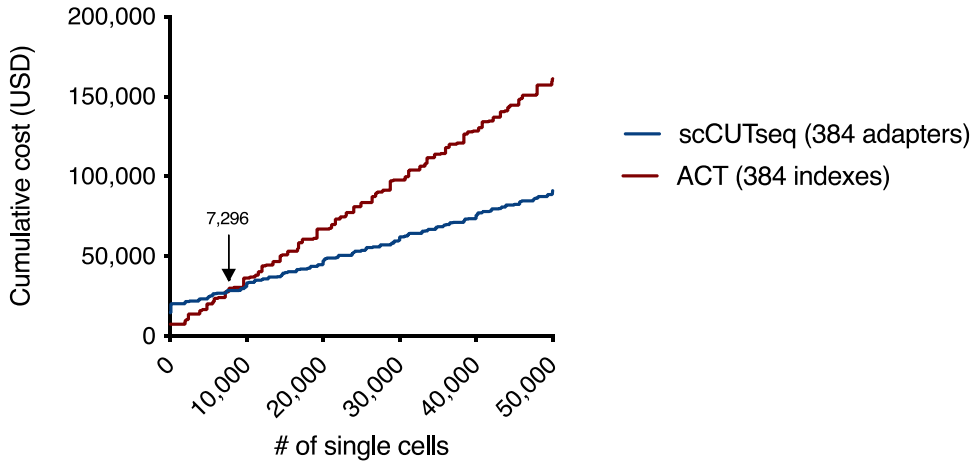
To assess the cost-effectiveness of scCUTseq, we first compared it to the standard MALBAC method[29], which in scCUTseq is used to perform whole-genome amplification prior to barcoding genomic DNA (gDNA) in individual cells (see **Fig. 1a**). In scCUTseq, the volume of MALBAC reagents used in the lysis, pre-amplification and amplification steps is 200-fold lower compared to standard MALBAC. Furthermore, while in the latter a single sequencing library needs to be generated from a whole genome amplified cell, in scCUTseq multiple cells (typically, 96 or 384, depending on the number of CUTseq adapters available) are pooled into the same library after barcoding gDNA in each individual cell. A list of reagent costs (based on prices as of June 2021) used to make the cost comparison is provided in **Supplementary Table 8**. Note that sequencing costs are excluded, since they only depend on the number of single cells sequenced and target sequencing depth. As shown in the **Supplementary Notes Figure 1** below, the cumulative cost grows at a much faster rate for standard MALBAC compared to scCUTseq and needs to be plotted on logarithmic scale for the difference between the two curves to be visible. This is expected from the fact that in scCUTseq 200-fold less MALBAC reagents are used, and 384 cells are pooled into the same library, drastically reducing the cost per cell. Assuming to sequence 10,000 cells as we did in this study, the reagent costs for making libraries would be 33.6-fold higher using standard MALBAC compared to scCUTseq (1,123,221 USD *vs.* 33,335 USD, respectively), highlighting the cost-effectiveness of scCUTseq.



Supplementary Notes Figure 1. Plot showing how the cost for preparing libraries for single-cell copy number profiling grows with the number of single cells processed using standard

MALBAC (red) *vs.* scCUTseq (blue). Note that the y-axis in on logarithmic scale to be able to distinguish between the two curves.

Next, we performed a side-by-side comparison with acoustic cell tagmentation (ACT)[17], a recently described method for scDNA-seq, which, similarly to scCUTseq, also uses a nanodispensing device (Echo 525, Beckman Coulter) to reduce the volume of reagents delivered to each single cell in 96- or 384-well plates. However, unlike scCUTseq, ACT uses a scaled-down version of the Illumina's Nextera library preparation kit (25-fold less reagent volumes per cell compared to the recommended volumes per sample in the Nextera kit) to index individual cells with the Tn5 transposase, followed by 18–16 polymerase chain reaction (PCR) cycles before pooling multiple cells together for sequencing. To compare the two methods, we computed the cumulative costs of preparing ready-to-sequence libraries for 50,000 cells using scCUTseq *vs.* ACT, assuming to use 384 CUTseq adapters (as we did in this study) and purchasing all the 384 indexes available in the Nextera kit. All the reagent costs used for this analysis are listed in **Supplementary Table 8**. We did not include sequencing costs in our analysis, as these would be the same for scCUTseq and ACT, for any given target depth per cell. As shown in the **Supplementary Notes Figure 2** below, up to 7,296 single cells processed, ACT is more cost-effective compared to scCUTseq. Above 7,296 cells, scCUTseq becomes progressively more cost-effective. For 10,000 cells (approximately the number of single cells from prostate samples that we sequenced in this study), scCUTseq is ~10% less costly than ACT, while for 50,000 cells the difference would increase to ~77%.



Supplementary Notes Figure 2. Plot showing how the cost for preparing libraries for single-cell copy number profiling grows with the number of single cells, using scCUTseq *vs.* ACT.

The higher cost for scCUTseq below 7,296 cells can be explained by the need to initially purchase more reagents compared to ACT. However, as shown in **Supplementary Table 8**, many of these reagents are then sufficient to process thousands or even millions (as in the case of CUTseq oligo adapters) of cells, thus making scCUTseq less costly than ACT above 7,296 cells.