PNAS RESEARCH ARTICLE | CELL BIOLOGY OPEN ACCESS

Structural variant and nucleosome occupancy dynamics postchemotherapy in a HER2+ breast cancer organoid model

Maja Starostecka[a,b,1], Hyobin Jeong[a,c,1], Patrick Hasenfeld[a], Eva Benito-Garagorri[a], Tania Christiansen[a,d], Catherine Stober Brasseur[a], Maise Gomes Queiroz[a], Marta Garcia Montero[e,2], Martin Jechlinger[e,f,3], and Jan O. Korbel[a,d,3]

Affiliations are included on p. 10.

Edited by Benjamin Neel, New York University Department of Medicine, New York, NY; received July 31, 2024; accepted January 8, 2025

The most common chemotherapeutics induce DNA damage to eradicate cancer cells, yet defective DNA repair can propagate mutations, instigating therapy resistance and secondary malignancies. Structural variants (SVs), arising from copy-number-imbalanced and -balanced DNA rearrangements, are a major driver of tumor evolution, yet understudied posttherapy. Here, we adapted single-cell template-strand sequencing (Strand-seq) to a HER2+ breast cancer model to investigate the formation of doxorubicin-induced de novo SVs. We coupled this approach with nucleosome occupancy (NO) measurements obtained from the same single cell to enable simultaneous SV detection and cell-type classification. Using organoids from TetO-CMYC/TetO-Neu/MMTV-rtTA mice modeling HER2+ breast cancer, we generated 459 Strand-seq libraries spanning various tumorigenesis stages, identifying a 7.4-fold increase in large chromosomal alterations post-doxorubicin. Complex DNA rearrangements, deletions, and duplications were prevalent across basal, luminal progenitor (LP), and mature luminal (ML) cells, indicating uniform susceptibility of these cell types to SV formation. Doxorubicin further elevated sister chromatid exchanges (SCEs), indicative of genomic stress persisting posttreatment. Altered nucleosome occupancy levels on distinct cancer-related genes further underscore the broad genomic impact of doxorubicin. The organoid-based system for single-cell multiomics established in this study paves the way for unraveling the most important therapy-associated SV mutational signatures, enabling systematic studies of the effect of therapy on cancer evolution.

breast cancer | organoids | single-cell multi-omics | structural variation

The induction of DNA damage, whether direct or indirect, is a fundamental mechanism through which most cancer chemotherapies and radiotherapies achieve cell death (1). The incorrect repair of this damage can result in mutations that are inherited by the progeny of damaged cells, thereby contributing to further tumor progression, secondary cancers, and long-term adverse effects (2–9). Chemotherapy frequently results in the emergence of dormant or minimal residual disease (MRD) states, whereby a subpopulation of cancer cells enters a quiescent phase, evading therapeutic interventions for extended periods (6). Mutations arising within these quiescent cells during treatment can lead to diverse survival strategies, ultimately driving lethal relapses (6). In recent years, comprehensive cancer genomic studies and the identification of mutational signatures (10) involving base substitutions (11) and small insertions and deletions (indels) (11) from platinum-based drugs (12–14), 5-fluorouracil (13, 15), and temozolomide (16, 17) have enhanced our understanding of the genomic consequences of chemotherapy. However, while SVs such as copy-number alterations and balanced DNA rearrangements represent a frequent source of driver mutations in cancer (18, 19), they have remained underexplored pertaining to their emergence during chemotherapy, as well as postchemotherapy.

This current knowledge gap is mainly due to technical difficulties in identifying somatic SVs as they appear in tissues (18). Bulk whole-genome sequencing (WGS) studies struggle to capture the full extent of intratumor heterogeneity, especially SVs present with a low cellular fraction in a tissue (18, 19), limiting our understanding of subclonal DNA rearrangements during drug exposure. And while platinum-based drugs have been shown to induce different classes of SV contributing to metastasis by bulk WGS (14), important gaps remain in our understanding of how common cancer treatments directly cause SV formation. Notably, single-cell studies, which could provide insights into chemotherapy-induced DNA rearrangement rates, have been lacking, contributing to missing knowledge about the consequences of chemotherapies on chromosomal instability.

Significance

Chemotherapeutics effectively kill cancer cells but can also induce mutations in surviving cells, leading to therapy resistance and secondary malignancies. We analyzed the DNA rearrangements caused by doxorubicin in murine HER2+ breast cancer organoids, focusing on genomic structural variants (SVs). Using single-cell sequencing, we revealed extensive therapy-induced SVs and genomic instability in a subset of posttreatment cells. Additionally, by advancing computational methods to analyze nucleosome occupancy (NO) in different cell types of the mammary lineage, we link SVs to gene expression changes that could potentially lead to breast cancer progression. This research underscores the need for strategies to minimize drug-induced genomic damage to improve treatment efficacy and patient outcomes.

Competing interest statement: The following author has previously disclosed a patent application (no. EP19169090) that is relevant to this manuscript: J.O.K. The remaining authors declare no competing interests.

This article is a PNAS Direct Submission.

[1]M.S. and H.J. contributed equally to this work.

[2]Present address: Bristol Myers Squibb, Leiden, The Netherlands.

[3]To whom correspondence may be addressed. Email: Martin.Jechlinger@molit.eu or Jan.Korbel@embl.org.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2415475122/-/DCSupplemental.

Published February 24, 2025.

Here, we employ Strand-seq (20), a single-cell technology based on bromodeoxyuridine (BrdU) incorporation during cell division, to unravel the complexity of treatment-induced somatic SVs. Unlike droplet-based single-cell technologies, Strand-seq, when coupled with a computational framework integrating information from read depth, strand state, and haplotype phase, robustly identifies a wide variety of SV classes in single cells, including de novo arising DNA rearrangements (21–23). These include chromosomal aneuploidies, deletions, duplications, and complex events. Strand-seq balances resolution and cost-effectiveness by enabling the detection of singleton SV events at fine scale (200 kb or larger), while also maintaining high throughput at low sequencing cost. Furthermore, Strand-seq facilitates the identification of functional impacts of somatic SVs by simultaneously analyzing nucleosome occupancy (NO) within the same cell (24). We adapted the Strand-seq technology to allow its use in a 3D mouse-based organoid system modeling breast cancer to investigate the effect of a common cancer therapeutic on SV formation. We focused on the detection of the SV footprint of doxorubicin (DXR), a DNA intercalator and inhibitor of type II topoisomerases. Doxorubicin has a broad spectrum of use for the treatment of childhood and adult cancers (25), including invasive and metastatic breast carcinoma. Despite disrupting critical DNA replication and transcription processes, which could lead to DNA alterations, a mutational signature of doxorubicin has not been yet specified (26). To explore the formation of SVs under chemotherapy within a tissue context, we employed Strand-seq in organoids. This approach maintains the cellular architecture of the primary tissue and allows for longitudinal monitoring throughout cancer treatment (27–29). By applying this system, we show that doxorubicin is associated with an extensive SV burden and elevated genomic heterogeneity, an increased frequency of sister chromatid exchanges (SCEs), as well as deregulated NO in all three main cell types present in murine mammary gland organoids. Our findings imply that large deletions and complex rearrangements represent a key mutational pattern associated with doxorubicin chemotherapy treatment.

Results

Doxorubicin-Induced Genomic Alterations in a HER2-Positive Breast Cancer Model. To explore therapy-induced DNA rearrangement processes in single cells, we set out to systematically analyze the types and frequency of doxorubicin-induced genomic alterations in mouse-derived organoids modeling breast cancer. We utilized the inbred TetO-CMYC/TetO-Neu/MMTV-rtTA mouse strain, which faithfully models HER2+ breast cancer in the context of CMYC activation (27, 30). Usage of this model system allows to control the overexpression of oncogenes by the addition of doxycycline (DOX) to the medium, while at the same time excluding the impact of genetic polymorphisms and allowing for tight control over the experimental environment. To assess the genotoxic effect of the drug, we treated the mammary gland organoids derived from these transgenic mice with a sublethal concentration of doxorubicin [100 nM, falling within the range of values detected in the blood of breast cancer patients undergoing chemotherapy (31), *SI Appendix*, Fig. S1] for 72 h and then incubated without a drug for 7 d. Such recovery time was necessary for cells to start dividing again for at least one cell division, allowing Strand-seq library preparation (32) from organoids according to the protocol we describe in the *Methods*, which involves incubating intact organoids with BrdU (Fig. 1*A*).

Using this protocol, we generated 459 Strand-seq libraries (post-quality filtering; mean of 354,506 uniquely mapped fragments per cell, *SI Appendix*, Fig. S2 and Dataset S1) from organoid-derived single cells at different stages of tumorigenesis before and after the drug treatment, including 130 libraries after the drug exposure (100 nM) and 84 from a vehicle (DMSO) control (33). The high rate of successfully generated libraries passing single-cell data quality control (42% on average) indicates that Strand-seq libraries are readily generated from organoids with this protocol (Fig. 1 *B*–*D* and *SI Appendix*, Fig. S3). This extends the application of Strand-seq to organoids, further broadening the application range of this powerful single-cell genomic technology.

We utilized the scTRIP (21) and PloidyAssignR (34) (*SI Appendix*, *Supplementary Methods*) computational frameworks to perform single-cell discovery of SVs and chromosomal aneuploidies in each Strand-seq library. After excluding germline SVs characteristic for the strain used in this study (Dataset S2 and *SI Appendix*, *Supplementary Methods*), we identified a very low level of de novo SV formation, and similarly a low frequency of aneuploidies, in cells undergoing doxycycline induction to initiate tumor development without the application of doxorubicin (Fig. 1*E*). However, analysis of 130 Strand-seq libraries derived from organoids treated with 100 nM doxorubicin shows that doxorubicin induces a wide range of large-scale chromosomal alterations of 200 kb in size or larger—seen in 46 cells, which represents a 7.4-fold enrichment compared to the tumor-induced control population. All of these detected events (*SI Appendix*, *Supplemental Data File 1*) are singletons (present only in a single cell) implying that these events most likely represent de novo chromosomal aberrations (*SI Appendix*, *Supplementary Note 1*). These results suggest that treatment with doxorubicin results in considerably elevated SV formation and genomic heterogeneity levels, exceeding levels arising in this HER2+ breast cancer model without chemotherapy application by far.

We subsequently performed a detailed analysis of the classes of chromosomal aberration arising in these samples. Following the drug treatment, we observe a statistically significant increase in complex SVs as well as large deletions and duplications ($P < 0.05$; Fig. 1*F*) compared to doxycycline induction without doxorubicin. We additionally observe an enrichment in chromosome gains but not in chromosome losses ($P < 0.05$) under this condition. The 46 genomically rearranged cells exhibit on average of 2 somatic alterations per genome (ranging from 1 to 7) with deletions representing the most common alteration (Fig. 2*A*). Interestingly, 13 complex rearrangements involving clustered deletions, amplifications, and/or inversions, detected in 9 cells, are seen exclusively after exposure to doxorubicin, mirroring prior observations of complex rearrangements seen in immortalized retinal epithelial cells subjected to doxorubicin (35). Deletions and duplications, both terminal and intrachromosomal, are present on all chromosomes (Fig. 2*B*) and we find no bias toward a certain chromosome having a higher number of SVs than the other chromosomes (permutation test, 10,000 permutations, $P = 0.6608$, *SI Appendix*, Fig. S4). This suggests that each chromosome is equally vulnerable to DNA damage caused by doxorubicin. All detected deletions and duplications are longer than 5 Mb with a median length of 51.5 Mb for deletions and 49.8 Mb for duplications (Fig. 2*C*), underscoring the generally large-scale nature of DNA changes induced by the drug. Taken together, these data demonstrate an increased frequency of large deletions and the presence of complex events as a consequence of doxorubicin treatment.

Strand-seq allows for the measurement of changes in the frequency of SCEs (32), which represent markers of double-strand breaks (DSBs) repaired by homologous recombination (HR) pathways during the S phase and mitosis. Along with genomic rearrangements, SCEs serve as valid indicators of genomic instability

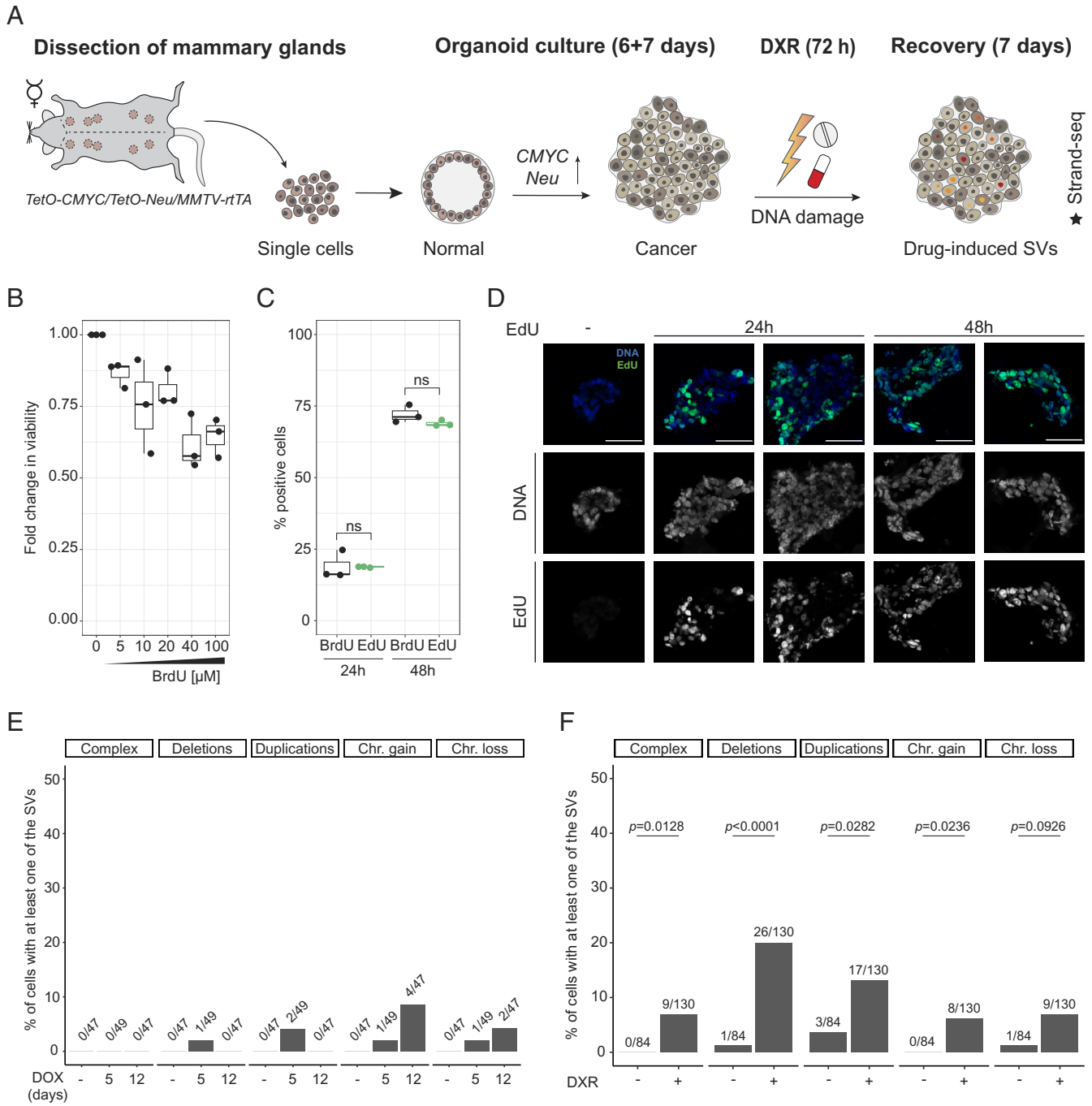

Fig. 1. Establishing a Strand-seq-based experimental workflow to study drug-induced genomic rearrangements in organoids. (*A*) Schematic overview of the experimental design. Single epithelial cells from adult virgin tritransgenic mice (TetO-CMYC/TetO-Neu/MMTV-rtTA) were cultured in 3D. Upon doxycycline-induced oncogene (CMYC and Neu) overexpression, organoids exhibited uncontrolled proliferation and lumen loss, mimicking a tumor state. After 7 d, organoids were treated with doxorubicin (or DMSO control) for 72 h and allowed to recover for up to 7 d (depending on the doxorubicin concentration). After the recovery period, the organoids were subjected to single-cell DNA sequencing to analyze the drug-induced SVs. (*B*) To optimize Strand-seq for murine mammary gland organoids, varying concentrations of BrdU were tested on newly formed organoids (two days after seeding). After 72 h of treatment, 20 μM BrdU was selected for subsequent experiments due to increased cytotoxicity at higher concentrations. Results from three independent experiments are presented. (*C*) We aimed to confirm uniform BrdU incorporation in murine mammary gland organoids induced with doxycycline using immunofluorescence. However, due to potential sample damage from DNA denaturation, we opted for Click-iT EdU labeling, which circumvents this issue. Murine mammary gland organoids induced with doxycycline were incubated with 20 μM BrdU or 20 μM EdU for 24 and 48 h. Flow cytometry analysis based on Hoechst quenching (BrdU) or Click-iT staining (EdU) showed similar incorporation efficiencies (Student's *t* test). Results represent three independent experiments with at least 10,000 single nuclei or cells recorded. (*D*) Doxycycline-induced murine mammary gland organoids were cultured with or without 20 μM EdU for 24 or 48 h. Organoids embedded in gels were sectioned into 8 μm slices, fixed, and labeled for EdU with Click-iT chemistry (green), with DNA counterstained using DAPI (blue). Representative images of single organoids demonstrate EdU incorporation in cells throughout the structures, both inside and on the rim. (Scale bar, 50 μm.) (*E*) Bar plot summarizing genomic rearrangements detected in Strand-seq libraries from cells never induced with doxycycline (47 cells), induced with 200 ng/mL doxycycline (DOX) for 5 (49 cells) or 12 d (47 cells). (*F*) Summary of different types of large scale chromosomal alterations detected in single cells of control (84 cells) and doxorubicin-treated (DXR) (130 cells) samples (two-sided Fisher's exact test).

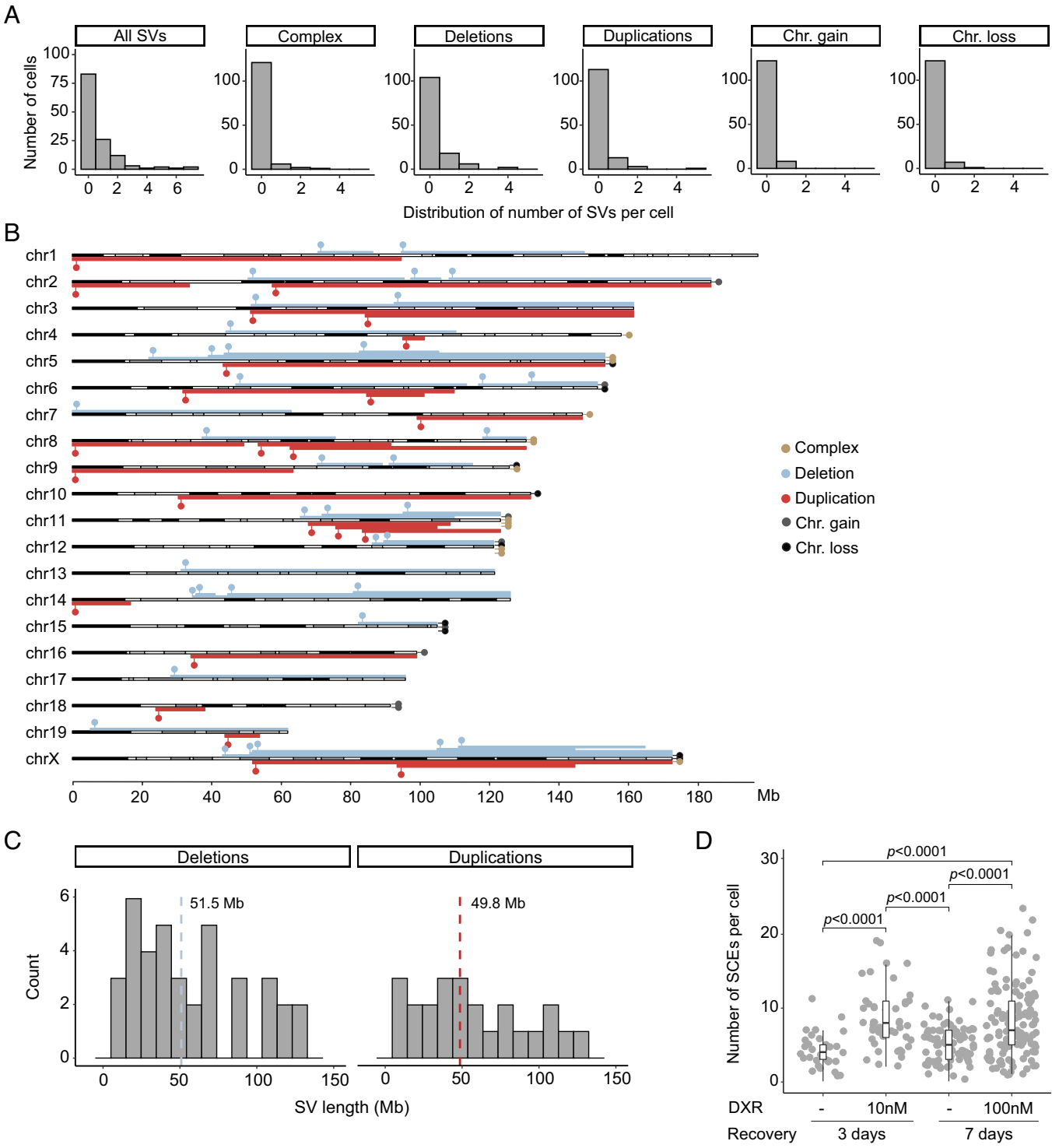

Fig. 2. Increased frequency of SVs and SCEs following doxorubicin treatment. (*A*) Number of chromosomal alterations by cell and aberration class, identified in 130 Strand-seq libraries derived from doxorubicin-treated cells. (*B*) Genome-wide karyogram representing the location and size of singleton genomic alterations post-doxorubicin treatment. Each dot represents a single event and is color-coded by class. Bars, with dots positioned above, correspond to the length of identified deletions and duplications. For complex events, chromosome gains, and losses, the dots are shown at the end of the corresponding chromosomes. (*C*) Size distribution of copy-number variants (deletions, duplications) detected in 130 single cells following doxorubicin treatment (aneuploidies not included). The median size for each SV class is indicated with a dashed line. The *P* values were determined using the Shapiro–Wilk normality test: for deletions $P = 0.0216$, for duplications $P = 0.1894$. (*D*) Number of SCEs detected in single cells of murine mammary gland organoids treated with two different concentrations of doxorubicin, 10 nM (49 cells) or 100 nM (130 cells), and their matching controls (33 and 84 cells, respectively). The statistical analysis was performed using the Kruskal–Wallis rank-sum test, and with a Wilcoxon rank-sum test for multiple-pairwise comparison between groups; only statistically significant differences are indicated.

(22, 32). We annotated SCEs in each Strand-seq library from doxorubicin-treated and control cells (*Methods*, *SI Appendix*, *Supplementary Note 2*). We included additional 49 cells treated with a lower concentration of doxorubicin (10 nM) and 33 cells of a corresponding vehicle control, both with shorter recovery time postexposure. The cells exposed to the lower drug concentration exhibit a significantly lower overall frequency of large-scale genomic alterations compared to those treated with 100 nM doxorubicin ($P = 0.0028$), and there are no significant differences in duplications and deletions under lower drug concentration when

compared to the control group (*SI Appendix*, Fig. S5). However, regardless of the doxorubicin concentration, all treated cells show a significantly higher frequency of SCEs compared to controls ($P < 0.0001$), suggesting elevated genomic stress caused by the drug, even at fairly low doxorubicin concentrations (Fig. 2*D*).

Subsequently, we explored the persistence of doxorubicin-induced SVs over an extended period, aiming to track potential SV subclonal expansion dynamics. We analyzed 20 Strand-seq-based single-cell genomes generated from our breast cancer organoid system after a longer recovery period post-doxorubicin treatment (21 d instead of 7 d). Notably, this prolonged recovery period correlated with a lower frequency of large-scale genomic alterations (*SI Appendix*, Fig. S6*A*). Specifically, 5 cells carried 6 singleton events, including chromosome gains (2 cells) and loss (1 cell) and deletions (3 cells). There was no evidence of clonal expansion or persistence of previously observed SVs. Furthermore, we observed that the frequency of SCE events had reverted to normal levels (*SI Appendix*, Fig. S6*B*), consistent with those observed in cells before doxorubicin treatment. These data collectively demonstrate that the genomic heterogeneity in breast cancer cells subject to doxorubicin treatment rises dramatically at first and subsequently declines, with cells impacted by SVs likely to experience negative selection. Irrespective of this, the wide diversity of chromosomal alterations generated by the drug could facilitate phenotypic adaptation to selective pressures during cancer therapy (36).

Exploring the Cell Lineage–Specific Sensitivity to Doxorubicin. As the structural and cellular complexity of the mammary gland is preserved in this HER2+ breast cancer model, characterizing the combined genetic and cellular heterogeneity before and after the treatment would provide a better understanding of how different cell types are affected by the same drug. Therefore, having identified the chromosomal alterations after doxorubicin treatment, we sought to determine whether there are differences in the frequency of SVs among the various cell types present in murine mammary gland organoids (*SI Appendix*, Fig. S7). The major epithelial cell types in human and murine mammary glands are luminal progenitor (LP) and mature luminal (ML) cells, as well as basal (B) cells. Luminal cells are responsible for producing and secreting milk proteins, whereas basal cells play a role in milk ejection and ductal transport (37). We harnessed the scNOVA computational framework, which integrates the discovery of somatic SVs and micrococcal nuclease-based NO measurements from Strand-seq data, allowing for single-cell multiomic analyses of chromosomal alterations (22, 24). We considered that utilizing lineage-specific genes could enable scNOVA to predict cell types in this organoid system, relying on a single-nucleus (sn)ATAC-seq-based reference epigenome dataset previously established for the mammary gland (38), which we reasoned could be suitable as a reference atlas to allow cell typing (22) in single-cell multiomics analyses (Fig. 3*A*).

We hence adapted the scNOVA framework to develop an ATAC-seq-based supervised classifier inferring the most likely cell type for each single-cell Strand-seq library (*Methods*) (39). This classifier uses the activity of transcription factor motifs as the feature set, capitalizing on the observation that chromatin accessibility measured by ATAC-seq is generally presumed to be inversely correlated with NO measurements obtained via micrococcal nuclease digestion (40). Comparing two model cell lines (mouse embryonic stem cells (mESCs) from 129 × C57BL/6 J strain and mouse embryonic fibroblasts (MEFs) isolated from FVB/NJ strain), we indeed find a pronounced inverse correlation between motif occupancy measured by Strand-seq and motif accessibility derived from ATAC-seq data (41, 42) ($R = -0.87$, $P < 2.2e^{-16}$; Fig. 3 *B* and *C*). Encouraged by this finding we trained our model using a previously generated (38) snATAC-seq dataset from adult murine mammary glands in which one of three major mammary cell types, LP, ML, or B, was labeled for each single-cell library providing us a ground truth of cell-type-specific chromatin accessibility in the motifs (*Methods*). We first converted the snATAC-seq count matrix (peak by cells) into a motif accessibility matrix and subsequently used this matrix to build a classifier using partial least squares discrimination analysis (PLSDA; *Methods*). For feature selection, we calculated variable importance in projection (VIP) values to measure the discrimination power for each motif. Motifs with significant VIP values compared to null distribution (FDR 10%) were used to finalize the model and to evaluate its performance by leave-one-out cross-validation which resulted in the AUC 0.9964, 0.9904, and 0.9983 for the classification of B, LP, and ML cell types, respectively (*SI Appendix*, Fig. S8). Our snATAC-seq-based classifier relies on 23 transcription factor motifs (stringent feature selection criteria using FDR of VIP < 10%) that are well known to regulate different mammary cell states (Fig. 3*D*). To predict cell types of single cells profiled by Strand-seq, we converted the NO count matrix (peak by cells) into a motif occupancy matrix (motifs by cells), and inverted this matrix into a motif accessibility matrix (motifs by cells). We utilized the resulting matrix as an input to the classifier which provided information about the most likely cell type from the mammary lineage, ensuring all cells were annotated.

Applying the classifier to all generated samples for this study, we successfully identified libraries representing all three distinct mammary epithelial cell types (*SI Appendix*, Fig. S9). In order to confirm the preservation of cell-type-specific chromatin accessibility in mammary gland organoids, we performed scATAC-seq on organoids kept in culture for 7 d. We find that cell type proportions closely align with those observed in Strand-seq data, following scNOVA cell-type annotation from the corresponding sample (*SI Appendix*, Fig. S10*A*). An additional motif-based analysis likewise shows consistency between the reference dataset and the cultured organoids (*SI Appendix*, Fig. S10*B*). We next leveraged scNOVA to test for statistically significant differences in SV frequency among these cell types. Notably, our analysis suggests that all types of large-scale genomic rearrangements observed in the 130 cells following doxorubicin treatment are consistently present across B, LP, and ML cells (Fig. 3*E*). We also observe a consistent frequency of SCEs between cell types within each sample (Fig. 3*F*) suggesting that all three cell types within mammary gland organoids are similarly prone to DNA damage and de novo chromosomal alterations.

Impact of Genomic Instability on NO and Pathway Dysregulation. Besides cell typing, we additionally inferred altered gene expression based on the changes in NO at gene bodies, using the scNOVA tool (24), aiming to identify deregulated pathways in doxorubicin-treated cells with and without SVs. We first established training and test datasets to generate a scNOVA model for the mouse genome and benchmarked its performance (*Methods*). In agreement with the robustness of this approach, we estimate its predictive accuracy between two distinct murine cell types, mESCs and MEFs, achieving an AUC of 0.7682 for the 10 most differentially expressed genes.

Having established the robustness of the model, we utilized scNOVA to analyze 130 single cells post-doxorubicin treatment (100 nM, 7 d of recovery), categorizing them based on the presence of chromosomal alterations. Among the cells, 84 exhibited no SVs or aneuploidy, 32 had SVs only, 3 displayed aneuploidy exclusively, and 11 harbored both SVs and aneuploidy. Subsequently, we

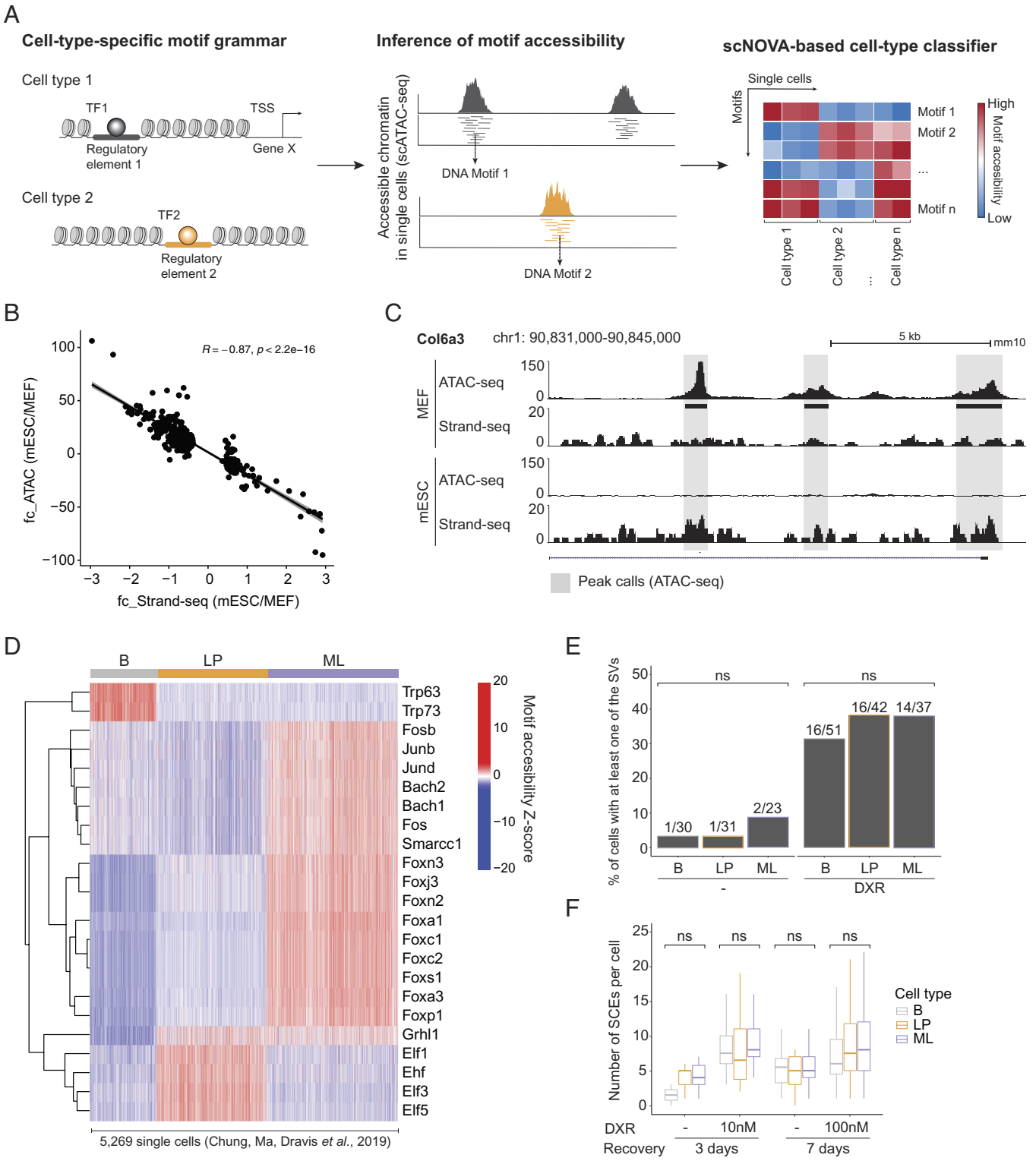

Fig. 3. Doxorubicin-associated genomic instability affects all major cell types of mammary gland organoids. (*A*) Schematic representation of the scNOVA-based cell-type classifier. Cell-type-specific gene expression relies on regulatory element interactions, defined by distinct DNA motifs and transcription factors. Using a reference scATAC-seq dataset of murine mammary gland cells, we infer cell-type-specific differential motif accessibility for model training. The resulting classifier accurately predicts the most likely cell type of a cell from which the Strand-seq library was generated. (*B*) Scatter plot of log2 fold changes (fc) of chromatin accessibility in *cis*-regulatory elements (CREs) between mESC and MEFs measured by ATAC-seq (y-axis) and the nucleosome occupancy (NO) between them measured by Strand-seq (x-axis). (*C*) Browser track snap show of representative locus nearby Col6a3. Three areas with gray background are the peak callings from ATAC-seq data showing that the gain of accessibility (ATAC-seq track) corresponds to the loss of NO (Strand-seq track). (*D*) Heatmap summarizing differential motif accessibility between three main cell types (B, LP, and ML) present in murine mammary gland. (*E*) All single-cell libraries of murine mammary gland organoids were annotated with the most likely cell type (B, LP, or ML) according to the output of the classifier. Barplot of the frequency of large-scale chromosomal alterations (y-axis) in different cell types (B, LP, and ML) treated with 100 nM doxorubicin or not. Significance tested with the Freeman–Halton extension to Fisher's exact test (2 × 3 contingency table). (*F*) Boxplot of SCE number in different cell types treated with two concentrations of doxorubicin and their corresponding controls. No significant difference in the SCE frequency between cell types was observed (Kruskal–Wallis rank-sum test with the Wilcoxon rank-sum test for multiple-pairwise comparison between groups).

applied scNOVA to our doxorubicin-treated cells, focusing on the comparison between the 84 cells not harboring alterations (abbreviated as WT, for wild type) and those 43 harboring SVs. After regressing out cell type and sample effects, we found 72 genes with increased NO suggesting decreased gene activity, and 156 with decreased NO suggesting increased gene expression (Fig. 4*A* and Dataset S3). Pursuing gene set overrepresentation analysis (43), we observe a significant decrease (FDR<0.1) in NO in members of pathways involved in breast cancer progression, including pathways related to MAPK (for example, *Rps6ka3*, *Pdgfb*), oxytocin (for example, *Prkag1*, *Prkacb*), and calcium signaling (for example, *Cacng1*, *Cacng5*). We also observe dysregulation in chemokine signaling, known to influence therapeutic outcomes (44), including genes displaying increased (for example, *Nfkbia*, *Cd40*, *Ccl26*) and decreased NO at their gene bodies (for example, *Ccr6*, *Ptk2b*) (Fig. 4*B* and Datasets S4 and S5).

We next investigated patterns of differential gene activity in cells with aneuploidies (3 + 11 cells) compared to the 84 WT cells, using scNOVA. This analysis yielded a distinct set of deregulated genes (73 genes with increased NO, 288 genes with decreased NO, Fig. 4*C* and Dataset S6), with gene set overrepresentation analysis implicating genes associated with the Rho cycle (for example, *Bcap31*, *Vav2*, *Grb7*) and Golgi functions (for example, *Arf3*, *Kif4*) in gene overactivation (Fig. 4*D* and Datasets S4 and S5). Members of the Rho GTPase family of proteins, some of them operating at the Golgi complex (45), were previously shown to be crucial for regulating processes associated with mitotic cell division, including chromosome segregation, which when aberrant, may lead to aneuploidies (46). Notably, cells with aneuploidies showed the decreased NO on genes whose products are involved in the MAPK pathway, suggesting a general connection of chromosomal abnormalities with this cancer-associated pathway.

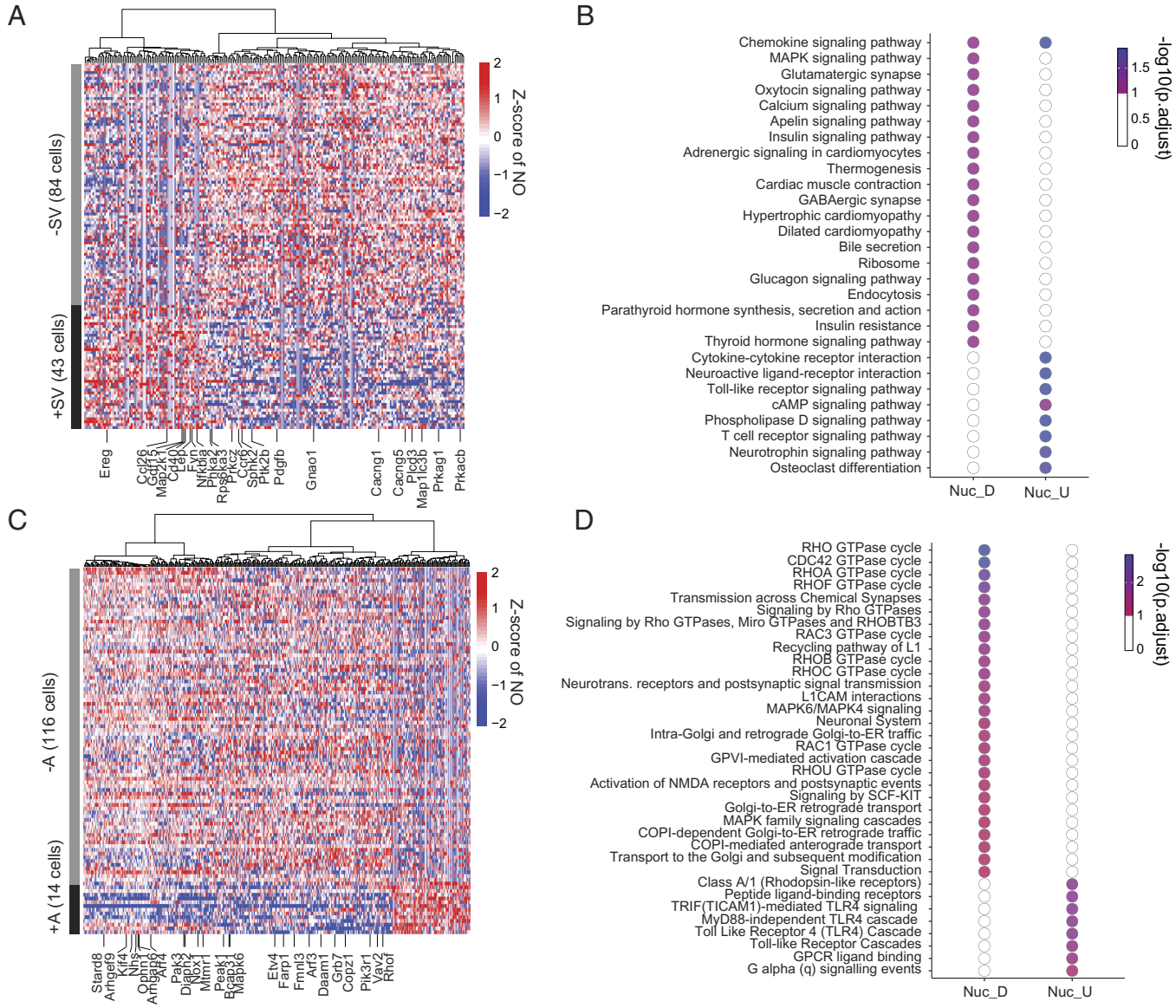

Fig. 4. Identification of deregulated pathways with scNOVA in cells with doxorubicin-induced genomic alterations. (*A*) Heatmap of 228 genes with differential nucleosome occupancy (diffNO) identified by scNOVA in cells with or without SVs after doxorubicin treatment, generated after regressing out the contribution of individual cell types and sequencing batches. The Y-axis corresponds to each single cell, and identified genes with diffNO are on the X-axis. Genes relevant for breast cancer progression are highlighted. The relative change in NO is color-coded from red (increased NO) to blue (decreased NO). (*B*) Pathways overrepresented by the genes included in the heatmap with diffNO, either showing decreased NO (Nuc_D) or increased NO (Nuc_U) (FDR < 0.1). (*C*) As in (*A*) but representing 361 genes with diffNO in cells post-doxorubicin treatment with or without aneuploidies (*A*). (*D*) As in (*B*) but presenting the data from enrichment analysis of the genes with diffNO depending on the aneuploidy status.

We additionally analyzed NO changes between cells harboring SVs and aneuploidies versus those without, for each cell type individually, uncovering a consistent number of deregulated genes across all three cell types (*SI Appendix*, Fig. S11 and Datasets S7–S12). Nonetheless, when repeating the analysis at the pathway level, we find a higher number of pathways with deregulated NO in luminal cells compared to basal cells (*SI Appendix*, Fig. S11 and Datasets S4 and S5). Overall, these data suggest that the treatment with doxorubicin induces, in addition to elevating chromosomal instability, global changes in NO at gene bodies, affecting genes involved in cancer progression and cell migration.

Discussion

By integrating single-cell multiomic approaches we analyzed the genomes of 459 single cells derived from murine mammary gland organoids, both before and after exposure to doxorubicin. Large-scale chromosome alterations such as complex SVs, large deletions and duplications, chromosome gains, and increased frequency of SCEs emerged as mutational patterns that could be associated with doxorubicin treatment in this organoid system. These findings underscore that doxorubicin commonly exerts genotoxic effects beyond single or double base substitutions, and highlight the importance of studying SVs to capture the full extent of treatment-induced DNA damage. We caution that since the resolution of Strand-seq is currently set to 200 kb (21), rearrangements smaller than this size were not the focus of this study (*SI Appendix*, *Supplementary Note 3*). Furthermore, we did not observe an increased frequency of SVs in cells without doxorubicin treatment, despite the high proliferation rate induced by overexpression of oncogenes (CMYC and Neu) used in our model system. This observation supports the hypothesis that the reported effects on chromosomal instability are doxorubicin-induced rather than caused by intrinsic mutational processes the cancer genome is subject to.

Additionally, our data suggest a dose dependency of the doxorubicin-induced SV burden. We observed that cells treated with a lower concentration of the drug exhibited fewer SVs, though the increased frequency of SCEs was consistently present. This indicates that while higher doses of doxorubicin induce more extensive structural damage, possibly through overwhelming error-prone DNA repair pathways rather than HR, even lower doses may already have an effect on genomic stability.

Understanding the implications of doxorubicin-induced genomic instability is crucial not only for immediate treatment response but also for guiding long-term outcomes and assessing the risk of secondary malignancies. Unlike studies based on bulk WGS, which concentrate on cancer cells that have undergone clonal expansion, our single-cell approach allowed us to characterize the genetic diversity of cancer cells following one round of chemotherapy. We find that the resulting cell population is highly heterogenous in terms of genomic alterations. Such genetic diversity and the presence of various SVs could pave the way for a dominant cell clone to emerge thus contributing to relapses in the clinics, particularly considering the typical regimen in clinical settings of multiple drugs received over an extended period of time.

Reinforcing these points, a recent comparative study on breast cancer patients receiving neoadjuvant treatment, including doxorubicin (the GeparSixto trial), identified groups that were at high risk of relapse despite initial pathological complete response (pCR) (47). In addition, epidemiological studies provide indirect evidence linking doxorubicin treatment to genomic instability and an increased risk of secondary malignancies, as observed in certain patient populations following treatment (48, 49). While in-depth characterization of SVs present in the distinct breast cancer subtypes posttreatment cannot be readily achieved in a clinical setting, organoid technology is well suited to recapitulate in vivo patient tumor settings (28). Here, we have methodologically opened organoid technology to in-depth single-cell multiomic characterization with Strand-seq—a single-cell technology operating at a unique spectrum of throughput and resolution (21, 22, 24, 50)—paving the way to systematic investigation of the direct and long-term molecular effects of cancer drugs. Although the murine organoid model was selected due to its experimental tractability, including the ability to tightly control BrdU incorporation in a setting of reduced intratumor heterogeneity, we envision that the approach described here could in the future be expanded to human cancer organoids as well as in vivo murine models.

Taking the analysis one step further, we also examined the susceptibility of different cell types to doxorubicin-induced DNA DSBs and resulting mutations using our murine organoid system. The scNOVA single-cell multiomic approach employed, which integrates NO–based cell typing with SV discovery, offers a streamlined method for identifying SVs while preserving cell type information. By choosing snATAC-seq data for model training, our methodological procedure eliminates the need for index sorting on mammary cells previously required by the scMNase-seq protocol (22), a laborious procedure requiring higher cell numbers. Our findings demonstrate that all three epithelial cell types within mammary gland organoids, regardless of their differentiation states and functional roles, are similarly impacted. This suggests that there are no significant differences in repair pathway preferences (51) resulting in chromosomal instability based on cell lineage within the mammary system. Besides cell typing, we additionally took advantage of the functionality of scNOVA to infer altered gene expression from Strand-seq libraries. By combining two redouts, genomic and epigenomic, from the same single cells, we show that doxorubicin treatment not only increases genomic instability through SVs and aneuploidies but also induces widespread changes in NO in cells with genomic alterations leading to changes in gene expression affecting crucial cancer-related pathways. Doing so, we could explore the global impact of SVs and aneuploidies on the epigenetic landscape.

We acknowledge a few current limitations of our study. While Strand-seq allows detecting de novo copy-number-imbalanced SVs in single cells, copy-balanced inversions and translocations need to be present in at least two cells to be confidently detected (21) and were thus not focused on in this study. As a lower coverage single-cell sequencing method, Strand-seq maps SV breakpoints to intervals rather than precise locations. Whereas our study focused on the global impact of somatic SVs, future studies may also explore local effects of SVs on the local chromatin environment (24). For gene expression inference using NO, we determined an AUC of 0.77 for the top-10 most differentially expressed genes. In the future, technologies coupling RNA-seq and DNA-seq in the same individual cell could be utilized to complement our approach. Furthermore, nanoliter-based Strand-seq platforms (50) could enable increased cell throughput in the future.

In conclusion, the findings from our study, underscore the pervasive impact of doxorubicin on de novo SV generation across various epithelial cell types, emphasizing the need for tailored therapeutic strategies that account for the widespread chromosomal instability induced by this chemotherapeutic agent. The presented approach is geared to analyze SVs and at the same time infer cell types and NO on a single-cell level and will therefore be able to evaluate different therapeutic strategies in meaningful organoid models. Future research leveraging single-cell multiomic approaches will be instrumental in uncovering the precise

mechanisms of chemotherapy-induced DNA damages with the aim to optimize cancer therapies toward lower long-term risk of tumor recurrence.

Methods

Animals and Genotyping. A mouse colony of the TetO-CMYC/TetO-Neu/MMTV-rtTA strain (in FVB background) (27, 30) was bred and maintained in the LAR (Laboratory Animal Resources) facility at EMBL Heidelberg under veterinary supervision. The animals were maintained in individually ventilated plastic cages (Tecniplast) in an air-conditioned (temperature 22 °C ± 2 °C, humidity 50% ± 10%) and light-controlled room (illuminated from 07:00 to 19:00 h). Mice were fed 1318 P autoclavable diet (Altromin, Germany) ad libitum. All animal care and procedures performed in this study conformed to the EMBL Guidelines for the Use of Animals in Experiments and were reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) (approval # MJ160070). The presence of transgenes was confirmed by diagnostic genotyping as described before (30).

3D Culture of Mouse Mammary Gland Organoids and Doxorubicin Treatment. Female virgin mice aged 8 to 10 wk were killed by cervical dislocation, and all 10 mammary glands were collected. The mammary tissue was digested overnight (max. 16 h) in 5 mL of digestion medium (DMEM/F12 (Lonza) with 25 mM HEPES, supplemented with 1% penicillin/streptomycin (Thermo Fisher Scientific), 150 U Collagenase type 3 (Worthington Biochemical Corporation) and 20 μg Liberase (Roche)) at 37 °C and 5% CO_2, in a loosely capped 50 mL polypropylene conical tube. The glands were then mechanically disrupted by pipetting with a 5 mL pipette, washed with phosphate-buffered saline (PBS), and centrifuged at 300×*g* for 5 min. The layer of fat and medium with PBS was removed, and 5 mL of 0.25% Trypsin-EDTA (Thermo Fisher Scientific) was added to the cell pellet. After incubation for 45 min at 37 °C and 5% CO_2 the enzymatic reaction was stopped by the addition of 40 mL of serum-supplemented media (DMEM/F12 (Thermo Fisher Scientific) with 25 mM HEPES, 1% penicillin/streptomycin (Thermo Fisher Scientific), and 10% fetal bovine serum (FBS) (Thermo Fisher Scientific)). The cells were centrifuged again, the pellet was resuspended in mammary epithelial cell basal medium (MEBM) (PromoCell) enriched with MEpiCGS (ScienCell Research Laboratories), and the cell suspension was transferred to collagen-coated plates (Corning). During overnight incubation at 37 °C and 5% CO_2, epithelial cells adhere to the surface of the plate while cells of other cell types and dead cells float in the medium. On the following day, the medium with nonepithelium cells was removed, and after washing in PBS and trypsinization (incubation with 0.25% Trypsin-EDTA for 5 to 7 min at 37 °C and 5% CO_2, followed by inactivation with serum-supplemented media), the epithelial cells were detached from collagen-coated plates. The cells were centrifuged, resuspended in MEBM with MEpiCGS, and counted. 10,000 cells were seeded per 1 well of a 12-well plate in 90 μL of ice-cold solution 4:1 of Matrigel (Corning): PBS. Matrigel-PBS-cell solution droplets were dispensed into the bottom of wells and incubated for 30 to 45 min at 37 °C and 5% CO_2 until the mixture solidified. Then 1.5 mL of MEBM with MEpiCGS was added to each well, and the organoids were let to grow at 37 °C and 5% CO_2. The medium was changed one day after seeding, and then every second day. If required by experiment conditions, the organoids were induced with 200 ng/mL of doxycycline (doxycycline hyclate, Sigma) 7 d after seeding. 200 ng/mL of doxycycline was then always added to the medium. Organoids were treated with either 10 nM or 100 nM doxorubicin (doxorubicin-hydrochloride, Sigma-Aldrich, D1515) [or corresponding DMSO (Sigma-Aldrich, D8418) concentration] for 72 h and then let to recover for up to 7 d in a drug-free medium before being processed for Strand-seq.

Dissociation of 3D Structures. Gels containing organoids (growing in single wells of 12-well plates) were incubated for 2 h at 37 °C with 75 U Collagenase type III (Worthington Biochemical Corporation) and 10 μg Liberase (Roche) and then disintegrated completely by mechanical disruption with pipetting up and down with a 1,000 μL pipette. The suspension from each well was transferred to its 15 mL falcon and centrifuged at 300×*g* for 5 min. After washing once in pre-warmed PBS and removing any leftovers of Matrigel that remained above the cell pellets, the pellets were resuspended in 200 μL of 0.25% Trypsin-EDTA (Thermo Fisher Scientific) and incubated for 5 min in 37 °C water bath. Trypsin was then deactivated with 5 mL medium containing DMEM/F12 (Thermo Fisher Scientific) with 25 mM HEPES, 1% penicillin/streptomycin (Thermo Fisher Scientific), and 10% FBS (Thermo Fisher Scientific). Suspensions of single cells were centrifuged at 300×*g* for 5 min, and the pellets were washed once in PBS. The suspension of single cells was further processed for Strand-seq.

Strand-Seq Library Preparation from Organoids. Organoids were incubated first with 20 μM BrdU (Sigma, B5002) for exactly one cell cycle (incubation time varied depending on the condition, from 24 to 42 h), dissociated (as described above) and the single nuclei were isolated and analyzed by flow cytometry using nuclei from a BrdU–negative control to set up gates. Then single BrdU-containing nuclei were sorted into individual wells of a 96-well plate [100 μm nozzle, BD FACSAria Fusion Flow Cytometer (BD Biosciences)] containing 5 μL of freeze buffer, centrifuged for 5 min at 4 °C at full speed, and immediately frozen at −80 °C. Strand-seq libraries were prepared according to a previously published procedure (20, 32) using a Beckman Coulter Biomek FxP liquid-handling robotic system (Beckman Coulter) and then sequenced on an Illumina NextSeq 500 sequencing platform (MID-mode, 75 base pair paired-end sequencing protocol).

Strand-Seq Data Analysis. Somatic structural variant (SV) calling and NO profiling was performed using the previously described scTRIP and scNOVA computational methods. scTRIP was originally developed to analyze Strand-seq data from the human genome (21). In this study, to analyze Strand-seq data from the mouse genome, we adapted the Mosaicatcher pipeline (21) by creating a blacklist of mouse reference genome (mm10), and the normalization scheme accordingly. In this setting, we performed SV calling of the Strand-seq using the Mosaicatcher algorithm, which uses fixed width bins (default: 100 kb) and then proceeds with joint segmentation of single cells, performing segmentation on a multivariate input (52). This algorithm uses a cost matrix, to determine the cost (summed squared error) of every possible consecutive segment. While the same direction of change was assumed in all samples in the original implementation of the segmentation approach, we adapted the algorithm to calculate this cost matrix for each cell and strand separately. Additionally, we adapted the cost matrix to penalize segments which are below 200 kb in size, as a means of avoiding oversegmentation. After the segmentation, we classified the genotype for each segment using a Bayesian model, which yielded SV calls larger than minimum segment size of 200 kb. The "strict" SV caller derived from the original scTRIP and optimized for the detection of SVs with clonal frequency ≥ 5%, was used to annotate germline SVs present in murine mammary gland organoids, while the "lenient" caller was applied to detect SVs present in a single-cell only. All chromosomes are interpretable in the dataset, even when paternal and maternal alleles share the same strand label (e.g., in WW or CC states)–in these cases, read depth was employed to make single-cell-based SV calls.

SCEs were identified in single cells as points on chromosome plots where reads mapping to both Watson and Crick strands switch to reads mapping to either the Watson or the Crick strand (without affecting the average read count). Coordinates of SCEs were extracted from the output of scTRIP, and to correct for Strand-seq resolution (200 kb), each SCE breakpoint was converted into a region ± 100 kb.

scNOVA was previously developed for the analysis of human samples, using the hg38 human reference genome. In this study, we adapted the scNOVA pipeline to be able to analyze mouse data. For this adjustment, we retrained the CNN model of scNOVA using the training set of MEFs and mESCs. For this purpose, we generated both Strand-seq and RNA-seq (*SI Appendix, Supplementary Methods*) to infer expressed genes from NO at the genebodies. We obtained an overall AUC of 0.872 for the prediction of gene expression status using leave-one-chromosome out cross-validation. The adjusted scNOVA pipeline takes single-cell Strand-seq bam files as input, and extracts NO in genebodies as well as in *cis*-regulatory elements (CREs).

scNOVA analysis for cell typing. First, we built the supervised cell-type classifier using the motif accessibility profile using previously published scATAC-seq data (38) and applying partial least square discriminant analysis (PLS-DA). This classifier enables assignment of single cells into one of the three cell types in the mammary gland: B (basal), LP (luminal progenitor), and ML (mature luminal). This classifier requires the input matrix that represents single-cell motif accessibility. To obtain this matrix from Strand-seq libraries, we extracted NO at mammary gland CREs using scNOVA (24), which provides single-cell NO at each of the CREs. Using chromVAR (53), we converted this matrix into a single-cell motif NO matrix. By multiplying this matrix by (-1), we derived a single-cell motif accessibility matrix [motif-by-single cells]. Using this matrix as input for

the cell-type classifier, we assigned each of the single-cell Strand-seq libraries into one of the three possible cell types.

scNOVA analysis to infer altered gene activity. To infer altered gene activity in the cells harboring SVs compared to WT cells (comparison1), as well as in the cells harboring aneuploidy compared to WT cells (comparison2), we performed scNOVA analysis using PLS-DA (24). FDR < 0.1 was used as a threshold to define genes with significantly altered gene activities. We further performed gene set overrepresentation analysis using ConsensusPathDB (43) to identify the enriched biological process and pathways among the genes with significantly altered activities.

Confirming the Negative Correlation Between Strand-Seq and ATAC-Seq. To evaluate the association between NO derived from Strand-seq and chromatin accessibility obtained from ATAC-seq, we first generated Strand-seq data from two model cell lines (mESCs from 129 × C57BL/6 J strain and MEFs isolated from FVB/NJ strain). In parallel, we collected bulk ATAC-seq data from the respective cell lines (mESCs and MEFs) from the previously published literature (41, 42). The ATAC-seq data were first aligned to the mm10 reference genome using bowtie2 (54). Peak calling was performed using MACS2 (55) for each of the cell lines and combined into the consensus peaks using bedtools to define cis-regulatory regions (CREs) (56). For those CREs, we compared the fold changes between mESC and MEF in terms of chromatin accessibility measured by ATAC-seq, and NO measured by Strand-seq. We then calculated the Spearman correlation coefficient between the log2 fold changes from ATAC-seq and Strand-seq.

Statistical Analysis. All statistical tests used are reported, where applied, in the main text or figure legends.

Data, Materials, and Software Availability. Computational code, Strand-seq genomic data have been deposited in https://github.com/jeongdo801/scNOVA_mouse/ (39); ENA (PRJEB76147) (33). All study data are included in the article and/or supporting information.

ACKNOWLEDGMENTS. We wish to thank the Korbel group for fruitful discussions and thank the EMBL core facilities and services for support in computing (IT), sequencing (GeneCore), animals (Laboratory Animal Resources), and cell sorting (Flow Cytometry). J.O.K acknowledges core funding from the EMBL, funding to the SATURN-3 consortium from the Federal Ministry of Education and Research (01KD2206D), as well as an ERC Consolidator Grant (MOSAIC; Grant no. 773026) supporting this study. Portions of the paper (particularly part of the *Methods*) were developed from the dissertation of M.S. at Heidelberg University (doi: 10.11588/heidok.00033263).

Author affiliations: [a]European Molecular Biology Laboratory, Genome Biology Unit, Heidelberg 69117, Germany; [b]Faculty of Biosciences, Collaboration for joint PhD degree between European Molecular Biology Laboratory and Heidelberg University, Heidelberg 69120, Germany; [c]Department of Systems Biology, College of Life Science and Biotechnology, Yonsei University, Seoul 03722, Republic of Korea; [d]Bridging Research Division on Mechanisms of Genomic Variation and Data Science, German Cancer Research Center, Heidelberg 69120, Germany; [e]European Molecular Biology Laboratory, Cell Biology and Biophysics Unit, Heidelberg 69117, Germany; and [f]Molecular and Information Technology Institute for Personalized Medicine gGmbH, Heilbronn 74076, Germany

Author contributions: M.S., H.J., M.J., and J.O.K. designed the research; M.S., H.J., P.H., E.B.-G., C.S.B., M. Gomes Queiroz, and M. Garcia Montero performed research; H.J. and T.C. contributed new reagents/analytic tools; M.S., H.J., M.J., and J.O.K. analyzed data; P.H. nuclei sorting, Strand-seq library preparation; E.B.-G. preparing cryosections, Strand-seq library preparation; T.C. wrote the PloidyAssignR code; C.S.B. and M. Gomes Queiroz Strand-seq library preparation; M. Garcia Montero mouse husbandry, training in organoid culture; and M.S., H.J., M.J., and J.O.K. wrote the paper.

1. W. M. C. van den Boogaard, D. S. J. Komninos, W. P. Vermeij, Chemotherapy side-effects: Not all DNA damage is equal. *Cancers* **14**, 627 (2022).
2. L. R. Yates *et al.*, Genomic evolution of breast cancer metastasis and relapse. *Cancer Cell* **32**, 169–184.e7 (2017).
3. K. Abe *et al.*, Doxorubicin causes ferroptosis and cardiotoxicity by intercalating into mitochondrial DNA and disrupting Alas1-dependent heme synthesis. *Sci. Signal.* **15**, eabn8017 (2022).
4. C. Demoor-Goldschmidt, F. de Vathaire, Review of risk factors of secondary cancers among cancer survivors. *Br. J. Radiol.* **92**, 20180390 (2019).
5. D. Cardinale *et al.*, Early detection of anthracycline cardiotoxicity and improvement with heart failure therapy. *Circulation* **131**, 1981–1988 (2015).
6. M. L. De Angelis, F. Francescangeli, F. La Torre, A. Zeuner, Stem cell plasticity and dormancy in the development of cancer therapy resistance. *Front. Oncol.* **9**, 626 (2019).
7. J. Youk *et al.*, Quantitative and qualitative mutational impact of ionizing radiation on normal cells. *Cell Genom* **4**, 100499 (2024).
8. S. Behjati *et al.*, Mutational signatures of ionizing radiation in second malignancies. *Nat. Commun.* **7**, 12605 (2016).
9. E. Kocakavuk *et al.*, Radiotherapy is associated with a deletion signature that contributes to poor outcomes in patients with cancer. *Nat. Genet.* **53**, 1088–1096 (2021).
10. L. B. Alexandrov, S. Nik-Zainal, D. C. Wedge, P. J. Campbell, M. R. Stratton, Deciphering signatures of mutational processes operative in human cancer. *Cell Rep.* **3**, 246–259 (2013).
11. L. B. Alexandrov *et al.*, The repertoire of mutational signatures in human cancer. *Nature* **578**, 94–101 (2020).
12. A. Boot *et al.*, In-depth characterization of the cisplatin mutational signature in human cell lines and in esophageal and liver tumors. *Genome Res.* **28**, 654–665 (2018).
13. O. Pich *et al.*, The mutational footprints of cancer therapies. *Nat. Genet.* **51**, 1732–1740 (2019).
14. S. Gonzalez, N. Lopez-Bigas, A. Gonzalez-Perez, Copy number footprints of platinum-based anticancer therapies. *PLoS Genet.* **19**, e1010634 (2023).
15. S. Christensen *et al.*, 5-Fluorouracil treatment induces characteristic T>G mutations in human cancer. *Nat. Commun.* **10**, 4571 (2019).
16. L. B. Alexandrov *et al.*, Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
17. J. Wang *et al.*, Clonal evolution of glioblastoma under therapy. *Nat. Genet.* **48**, 768–776 (2016).
18. M. R. Cosenza, B. Rodriguez-Martin, J. O. Korbel, Structural variation in cancer: Role, prevalence, and mechanisms. *Annu. Rev. Genomics Hum. Genet.* **23**, 123–152 (2022).
19. ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium, Pan-cancer analysis of whole genomes. *Nature* **578**, 82–93 (2020).
20. A. D. Sanders, E. Falconer, M. Hills, D. C. J. Spierings, P. M. Lansdorp, Single-cell template strand sequencing by Strand-seq enables the characterization of individual homologs. *Nat. Protoc.* **12**, 1151–1176 (2017).
21. A. D. Sanders *et al.*, Single-cell analysis of structural variations and complex rearrangements with tri-channel processing. *Nat. Biotechnol.* **38**, 343–354 (2020).
22. K. Grimes *et al.*, Cell-type-specific consequences of mosaic structural variants in hematopoietic stem and progenitor cells. *Nat. Genet.* **56**, 1134–1146 (2024), 10.1038/s41588-024-01754-2.
23. M. R. Cosenza *et al.*, Origins of de novo chromosome rearrangements unveiled by coupled imaging and genomics. bioRxiv [Preprint] (2024). https://doi.org/10.1101/2024.08.15.607890 (Accessed 17 November 2024).
24. H. Jeong *et al.*, Functional analysis of structural variants in single cells using Strand-seq. *Nat. Biotechnol.* **41**, 832–844 (2022), 10.1038/s41587-022-01551-4.
25. F. Yang, S. S. Teves, C. J. Kemp, S. Henikoff, Doxorubicin, DNA torsion, and chromatin dynamics. *Biochim. Biophys. Acta* **1845**, 84–89 (2014).
26. B. Szikriszt *et al.*, A comprehensive survey of the mutagenic impact of common cancer cytotoxics. *Genome Biol.* **17**, 99 (2016).
27. M. Jechlinger, K. Podsypanina, H. Varmus, Regulation of transgenes in three-dimensional cultures of primary mouse mammary cells demonstrates oncogene dependence and identifies cells that survive deinduction. *Genes Dev.* **23**, 1677–1688 (2009).
28. H. Clevers, Modeling development and disease with organoids. *Cell* **165**, 1586–1597 (2016).
29. K. M. Havas *et al.*, Metabolic shifts in residual breast cancer drive tumor recurrence. *J. Clin. Invest.* **127**, 2091–2105 (2017).
30. K. Radic Shechter *et al.*, Metabolic memory underlying minimal residual disease in breast cancer. *Mol. Syst. Biol.* **17**, e10141 (2021).
31. Y. Harahap, P. Ardiningsih, A. Corintias Winarti, D. J. Purwanto, Analysis of the doxorubicin and doxorubicinol in the plasma of breast cancer patients for monitoring the toxicity of doxorubicin. *Drug Des. Devel. Ther.* **14**, 3469–3475 (2020).
32. E. Falconer *et al.*, DNA template strand sequencing of single-cells maps genomic rearrangements at high resolution. *Nat. Methods* **9**, 1107–1112 (2012).
33. M. Starostecka *et al.*, Datasets from "Structural variant and nucleosome occupancy dynamics postchemotherapy in a HER2+ breast cancer organoid model." European Nucleotide Archive (ENA). https://www.ebi.ac.uk/ena/browser/text-search?query=PRJEB76147. Deposited 28 May 2024.
34. https://github.com/lysfyg/PloidyAssignR: PloidyAssignR v1.0.0–First Release. The PloidyAssignR computational algorithm will be described in detail elsewhere.
35. B. R. Mardin *et al.*, A cell-based model system links chromothripsis with hyperploidy. *Mol. Syst. Biol.* **11**, 828 (2015).
36. L. Sansregret, B. Vanhaesebroeck, C. Swanton, Determinants and clinical implications of chromosomal instability in cancer. *Nat. Rev. Clin. Oncol.* **15**, 139–150 (2018).
37. N. Y. Fu, E. Nolan, G. J. Lindeman, J. E. Visvader, Stem cells and the differentiation hierarchy in mammary gland development. *Physiol. Rev.* **100**, 489–523 (2020).
38. C.-Y. Chung *et al.*, Single-cell chromatin analysis of mammary gland development reveals cell-state transcriptional regulators and lineage relationships. *Cell Rep.* **29**, 495–510.e6 (2019).
39. M. Starostecka *et al.*, Computational Scripts from "Structural variant and nucleosome occupancy dynamics postchemotherapy in a HER2+ breast cancer organoid model." Github. https://github.com/jeongdo801/scNOVA_mouse. Deposited 8 June 2024.
40. L. Minnoye *et al.*, Chromatin accessibility profiling methods. *Nat. Rev. Methods Primers* **1**, 10 (2021).
41. D. Bunina *et al.*, Genomic rewiring of SOX2 Chromatin interaction network during differentiation of ESCs to postmitotic neurons. *Cell Syst* **10**, 480–494.e8 (2020).
42. M. Naval-Sanchez *et al.*, Benchmarking of ATAC sequencing data from BGI's low-cost DNBSEQ-G400 instrument for identification of open and occupied chromatin regions. *Front Mol Biosci* **9**, 900323 (2022).
43. A. Kamburov, R. Herwig, ConsensusPathDB 2022: Molecular interactions update as a resource for network biology. *Nucleic Acids Res.* **50**, D587–D595 (2022).
44. D. J. Propper, F. R. Balkwill, Harnessing cytokines and chemokines for cancer therapy. *Nat. Rev. Clin. Oncol.* **19**, 237–253 (2022).
45. M. M. Mysior, J. C. Simpson, Emerging roles for Rho GTPases operating at the Golgi complex. *Small GTPases* **12**, 311–322 (2021).
46. M. Circop, Rho GTPases as regulators of mitosis and cytokinesis in mammalian cells. *Small GTPases* **5**, e29770 (2014).
47. J. Huober *et al.*, Identifying breast cancer patients at risk of relapse despite pathological complete response after neoadjuvant therapy. *NPJ Breast Cancer* **9**, 23 (2023).

48. S. I. M. Neppelenbroek *et al.*, Doxorubicin exposure and breast cancer risk in survivors of adolescent and adult Hodgkin lymphoma. *J. Clin. Oncol.* **42**, 1903–1913 (2024).
49. M. Adamska *et al.*, Acute myeloid leukemia post cytotoxic therapy in breast cancer survivors-over 23 years of single center analysis. *J. Clin. Med.* **13**, 989 (2024).
50. V. C. T. Hanlon *et al.*, Construction of Strand-seq libraries in open nanoliter arrays. *Cell Rep Methods* **2**, 100150 (2022).
51. H. Kim *et al.*, Differential DNA damage repair and PARP inhibitor vulnerability of the mammary epithelial lineages. *Cell Rep.* **42**, 113256 (2023).
52. W. Huber, J. Toedling, L. M. Steinmetz, Transcript mapping with high-density oligonucleotide tiling arrays. *Bioinformatics* **22**, 1963–1970 (2006).
53. A. N. Schep, B. Wu, J. D. Buenrostro, W. J. Greenleaf, chromVAR: Inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat. Methods* **14**, 975–978 (2017).
54. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
55. Y. Zhang *et al.*, Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
56. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).