Ductal carcinoma in situ develops within clonal fields of mutant cells in morphologically normal ducts

Hutten, S.J.; Messal, H.A.; Lips, E.H.; Sheinman, M.; Ciwinska, M.; Braams, E.; ... ; Grand Challenge PRECISION Consorti

Journal of Pathology
J Pathol July 2024; **263:** 360–371
Published online 23 May 2024 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/path.6289

ORIGINAL ARTICLE

Ductal carcinoma *in situ* develops within clonal fields of mutant cells in morphologically normal ducts

Stefan J Hutten[1,2†], Hendrik A Messal[1,2†], Esther H Lips[1], Michael Sheinman[2,3‖], Marta Ciwinska[4], Esmee Braams[1], Carolien van der Borden[1], Petra Kristel[1], Saskia Stoffers[1], Lodewyk FA Wessels[2,3], Grand Challenge PRECISION Consortium[‡], Jos Jonkers[1,2§], Jacco van Rheenen[1,2§], Jelle Wesseling[1,5,6§] and Colinda LGJ Scheele[4*§]

[1] *Division of Molecular Pathology, The Netherlands Cancer Institute, Amsterdam, The Netherlands*
[2] *Oncode Institute, Amsterdam, The Netherlands*
[3] *Division of Molecular Carcinogenesis, The Netherlands Cancer Institute, Amsterdam, The Netherlands*
[4] *Department of Oncology, VIB Center for Cancer Biology, KU Leuven, Leuven, Belgium*
[5] *Department of Pathology, Netherlands Cancer Institute–Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands*
[6] *Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands*

**Correspondence to: CLGJ Scheele, Department of Oncology, VIB Center for Cancer Biology, KU Leuven, Leuven, 3000, Belgium.*
E-mail: colinda.scheele@kuleuven.be

†These authors contributed equally to this work.
‡The complete list of the Grand Challenge PRECISION Consortium is provided in the Acknowledgements section.
§These authors are Senior authors.
‖*Present address:* Institute for Advanced Studies, Sevastopol State University, Sevastopol

Abstract

Mutations are abundantly present in tissues of healthy individuals, including the breast epithelium. Yet it remains unknown whether mutant cells directly induce lesion formation or first spread, leading to a field of mutant cells that is predisposed towards lesion formation. To study the clonal and spatial relationships between morphologically normal breast epithelium adjacent to pre-cancerous lesions, we developed a three-dimensional (3D) imaging pipeline combined with spatially resolved genomics on archival, formalin-fixed breast tissue with the non-obligate breast cancer precursor ductal carcinoma *in situ* (DCIS). Using this 3D image-guided characterization method, we built high-resolution spatial maps of DNA copy number aberration (CNA) profiles within the DCIS lesion and the surrounding normal mammary ducts. We show that the local heterogeneity within a DCIS lesion is limited. However, by mapping the CNA profiles back onto the 3D reconstructed ductal subtree, we find that in eight out of 16 cases the healthy epithelium adjacent to the DCIS lesions has overlapping structural variations with the CNA profile of the DCIS. Together, our study indicates that pre-malignant breast transformations frequently develop within mutant clonal fields of morphologically normal-looking ducts.

Keywords: field cancerization; ductal carcinoma *in situ*; pre-malignant transformation; spatial genomics; 3D pathology; copy number aberration

Received 30 November 2023; Revised 28 February 2024; Accepted 30 March 2024

No conflicts of interest were declared.

Introduction

Around one in ten women with invasive breast cancer have local recurrences within 10 years, even though they received breast-conserving surgery with free surgical margins, combined with radiation therapy and systemic adjuvant therapy [1,2]. This observation led to the speculation that recurrence may develop from the same lobe or ductal tree that was already predisposed to develop cancer, a hypothesis that is often referred to as the 'sick lobe' hypothesis [3,4]. The sick lobe model postulates that cells that acquire genetic or epigenetic alterations, which are required for tumorigenesis but are not sufficient to induce tumors, may spread over large areas of the ductal tree, thereby predisposing the still normal-behaving lobes for cancer development [5]. Although the large-scale presence of molecular abnormalities is often referred to as field cancerization [6], we prefer to use the term mutant field clonalization, because the involved ducts are not yet tumorigenic.

In support of the sick lobe hypothesis, pioneering studies showed that, in some cases, the normal breast

tissue adjacent to invasive breast cancer (IBC) exhibits abnormalities, such as genomic aberrations [7–9], increased proliferation [10–12], or changes in gene expression [13–17]. However, estimation of the frequencies of mutant clonal fields remained difficult due to the technical limitation of the two-dimensional (2D) histological analysis used, which cannot inform whether normal ducts and nearby lesions constitute segments of the same ductal tree or if they originate from separate trees within the same breast. Moreover, these pioneering studies were restricted to invasive disease, and may not be representative for the situation at the pre-invasive stage of breast cancer. Sequencing of neutral mitochondrial DNA mutations suggested that clonal fields may also exist around pre-cancerous lesions such as ductal carcinoma *in situ* (DCIS) [18]. Unfortunately, such neutral mutations are rare and therefore cannot be used for an extensive characterization of field clonalization around these early lesions. Moreover, since only a single mutation is analyzed, it is unknown whether field clonalization is driven by a single mutagenic event or by a continuous and multi-step process of mutation accumulation.

To study whether field clonalization is driven by a continuous accumulation of aberrations and to explore if these clonal fields around early breast lesions are commonly found, we developed a method to combine 3D reconstructions of human breast epithelium containing DCIS lesions with spatially resolved DNA copy number aberration (CNA) sequencing of the same DCIS lesions and adjacent histologically normal duct epithelium. CNA sequencing performs well with minimal DNA quantities, making it particularly adept for analyzing DNA samples derived from small tissue regions. Using the CNA profiles of local tissue regions, we correlated multiple CNAs between histologically normal ducts and adjacent DCIS regions and inferred their clonal relationships. Using this 3D image-guided characterization method, we built high-resolution spatial maps of the CNA profiles of the DCIS lesions and the surrounding normal epithelium, and studied the extent of heterogeneity within a single DCIS lesion and its adjacent histologically normal ducts. Mapping of the spatial CNA profiles onto the 3D reconstructed ductal subtree showed that in half of the cases, the healthy epithelium bordering the DCIS lesions has structural variations that overlap with the CNA profile of the DCIS, indicating that a sick lobe is indeed present in the histologically healthy ducts prior to pre-malignant transformation.

Materials and methods

Sample selection

From the Core Facility Molecular Pathology and Bio-banking (CF-MPB) within The Netherlands Cancer Institute we obtained 30 formalin-fixed, paraffin-embedded (FFPE) tissue blocks from 28 women who underwent surgery after DCIS diagnosis. The local institutional review board (IRB) approved the study protocol (CFMPB 688) and all patients provided informed consent. Tissue samples were selected based on DCIS morphology, using already available H&E slides, and DCIS lesion size, as more than five FFPE blocks had to be available to allow us to use a full FFPE block. Samples containing immune infiltrate or invasive components were excluded. Slides were examined by breast pathologists at the Antoni van Leeuwenhoek Hospital for the presence of DCIS and morphologically normal ductal epithelium. All DCIS blocks were between 3 and 5 years old. More detailed information on the 28 FFPE tissue samples used in this study is provided in supplementary material, Table S1.

3D imaging of patient FFPE DCIS resections

FFPE blocks were immunolabeled and tissue-cleared using a modified FLASH protocol [19,20]. Intact FFPE-embedded DCIS resections were recovered from the paraffin block with a razor blade, followed by deparaffinization in HistoChoice® (# H2779-1L; Sigma-Aldrich, St Louis, MO, USA) at 54 °C for 2 h. Tissue pieces were washed three times in 100% MeOH (Sigma-Aldrich) at room temperature, at least 1 h per wash, followed by immersion in dichloromethane (Sigma-Aldrich) for 3 h. Dichloromethane was refreshed for a further overnight incubation. The resections were washed in 100% MeOH twice for 1 h and bleached by immersion in 15% DMSO (Sigma-Aldrich) and 15% H_2O_2 (Sigma-Aldrich) in MeOH. The bleaching solution was refreshed after 6 h and the tissues were incubated overnight. The samples were then rehydrated through incubations in 75% and 30% MeOH in distilled H_2O (dH_2O) (1 h each), followed by two washes in dH_2O for 1 h. To enable immunolabeling, the samples were incubated in the FLASH antigen retrieval solution [200 mM boric acid, 4 M urea, and 8% 3-(decyldimethylammonio)propanesulfonate inner salt (CAS 15163-36-7); all from Sigma-Aldrich] in dH_2O (pH ∼7). Samples were incubated at room temperature for 1 h, followed by overnight incubation at 37 °C, after which the solution was refreshed and the temperature was increased to 54 °C for 24 h. Next, the tissues were washed in PBT (0.2% Triton X-100 in PBS, both from Sigma-Aldrich) for a minimum of three times and 1 h per wash at room temperature. Blocking before antibody labeling was carried out for 3 h in blocking buffer (10% FBS, 1% BSA, 5% DMSO, 0.2% Triton X-100, 0.02% sodium azide in PBS; all from Sigma-Aldrich). Samples were incubated with an α-smooth muscle actin (αSMA) antibody, clone 1A4, mouse IgG2a (# A2547, Sigma-Aldrich), diluted 1:1,000 in blocking buffer; a Krt8 antibody, clone TROMA-I, rat [Developmental Studies Hybridoma Bank (DSHB), Iowa City, IA, USA], diluted 1:50 in blocking buffer; and a HER2 antibody, clone 29D8, raised in rabbit (# 2165; Cell Signaling Technology, Danvers, MA, USA), diluted 1:100 in blocking buffer, for 80 h at room temperature. The tissue pieces were washed four times in PBS, 30 min per wash,

and incubated with Alexa Fluor™ conjugated secondary antibodies 1:1,000 and 1:1,000 Hoechst 33342 (Sigma-Aldrich) in blocking buffer for 80 h. The secondary antibodies used were donkey anti-rabbit Alexa Fluor 568 (# A10042), donkey anti-mouse Alexa Fluor 568 (# A10037), donkey anti-rat Alexa Fluor 488 (# A21208), goat anti-rat Alexa Fluor 647 (# A48265), goat anti-rabbit Alexa Fluor 488 (# A11008), and goat anti-mouse Alexa Fluor 647 (# A21235) (all from Invitrogen/Thermo Fisher Scientific, Waltham, MA, USA). Immunolabeled samples were washed four times for 30 min in PBS, followed by dehydration by a series of incubations in 30%, 50%, 75%, and 100% (twice) MeOH in dH_2O, each incubation for 3 h. Next, samples were immersed in 30% and then 70% methyl salicylate (Sigma-Aldrich) for 3–6 h per incubation, followed by two times incubation in 100% methyl salicylate for 3–6 h. After 2 days, the solvent was replaced with a 2:1 mixture of benzyl benzoate and benzyl alcohol (both from Sigma-Aldrich). Samples were kept in the solvent until imaging. Imaging was carried out on an inverted multiphoton confocal microscope (Leica TCS SP8 MP; Leica Microsystems, Mannheim, Germany) with a 25× water immersion objective (FLUOTAR VISIR 25×/0.95 W). Tiled *z*-scans capturing the entire FFPE blocks were acquired in Resonant Mode (8-bit) with 512 × 512 or 256 × 256 pixel format, 8,000 Hz scan speed, 1.25 zoom, 2× line average, and 5–15 µm *z*-steps. Fluorophores were excited simultaneously with an Insight X3 (Spectra-Physics, Milpitas, CA, USA) tunable two-photon laser at 800 nm. Three HyD-RLD detectors (Leica Microsystems) were used to simultaneously acquire second harmonic generation (SHG; 390–410 nm), Hoechst emission (420–500 nm), and Alexa Fluor™ 568 emission (580–620 nm). *Z*-compensation of the detector gains was used to correct for lower detection levels in deeper tissue layers due to scattering of the emitted fluorescence. Imaris Viewer 9.7.2 (Oxford Instruments, Abingdon, UK) was used for 3D visualization of the datasets.

Sample preparation for spatially resolved CNA analyses

The 16 FFPE DCIS tissue blocks were completely cut into 10-µm slides. Slides were arranged in sets A to E, and each section contained 40 (numbered 1 to 40) continuously sectioned slides. Subsequently, even-numbered slides were stained with toluene blue and further processed for DNA isolation. The odd-numbered slides were stained with H&E and used for 3D reconstruction of the ductal subtree.

Micro-dissection

The pathologist scored the tumor percentage and indicated the DCIS regions as well as morphologically normal tissue for isolation on an H&E-stained slide. Twenty Toluene blue-stained slides per tissue block section were manually dissected by scraping the areas off under a stereomicroscope (Axio Zoom.V16; ZEISS, Jena, Germany) using a needle, with DCIS areas and normal tissue collected separately. The scraped off tissue was stored in PKD digestion buffer (# 80234; QIAGEN, Germantown, MD, USA) and stored at 4 °C for up to a week. DNA and RNA were isolated simultaneously using an Allprep DNA/RNA FFPE isolation kit (# 80234, QIAGEN) by using a QIAcube, following the manufacturer's protocol. Following isolation, DNA concentration was determined using a spectrophotometer (NanoDrop 2000, Thermo Fisher Scientific) and using a Qubit dsDNA HS Assay Kit (# Q32851, Invitrogen).

CNA sequencing

A maximum of 2 µg of double-stranded genomic DNA was fragmented by Covaris shearing (Covaris Ltd, Brighton, UK) to obtain fragment sizes of 160–200 bp. Samples were purified using 2X Agencourt AMPure XP PCR Purification Beads according to the manufacturer's instructions (# A63881; Beckman Coulter, Brea, CA, USA). The sheared DNA samples were quantified and qualified on a BioAnalyzer system using the DNA 7500 assay kit (#5067-1506; Agilent Technologies, Santa Clara, CA, USA). With an input of maximum 1 µg of sheared DNA, library preparation for Illumina sequencing was performed using the KAPA HyperPrep Kit (# KK8504; KAPA Biosystems, Wilmington, NC, USA). During library amplification, six to eight PCR cycles were used to obtain enough yield for exome capture. After library preparation, the libraries were cleaned up using 1X AMPure XP beads (Beckman Coulter). All DNA libraries were analyzed on a BioAnalyzer system using the DNA 7500 chips (Agilent Technologies) for determining the molarity. Up to 13 uniquely indexed samples were mixed together by equimolar pooling. The pools were analyzed on the Agilent Technologies 2100 Bioanalyzer. Pools were diluted to 10 nM and the concentration of DNA fragments with target adapter sequences was measured by qPCR using a KAPA Library Quantification Kit (# KK4824, KAPA Biosystems). The pool was subjected to sequencing on an Illumina HiSeq 2500 machine, each pool in one lane of a single-read 65 bp run, following the manufacturer's instructions (Illumina, San Diego, CA, USA). FASTQ files were aligned to the human reference genome GRCh38 (hg38) using BWA v.0.7.17 aligner [21] and converted to BAM files using SAMtools v.1.10 [22]. Duplicate reads were marked using Picard v.2.25.0 (http://broadinstitute.github.io/picard) and removed. Relative copy number profiles were obtained using QDNAseq R package v.1.26.0 [23] with the following protocol: Read counts from the BAM files were obtained using the binReadCounts function using the bins generated by the getBinAnnotations function with the 100 kb bin size and filtered using the applyFilters function. The correction was estimated using the estimateCorrection function and the noise was estimated using getNoise, expectedVariance. The bin counts were bias-corrected, normalized, and smoothened using

correctBins, normalizeBins, smoothOutlierBins. We also removed outliers using the winsorize function from the copynumber R package (version 1.30.0) [24]. Multi-sample segmentation of the bins was done using the segmentBins function. To get the CN calls after cellularity correction, we ran the runACE function from the ACE R package (version 1.9.3) [25] and obtained adjusted segments and calls using the getadjustedsegments and ACEcall functions. We tested the identified CNAs against copy number polymorphisms in the general population using the dbVar database of large scale genomic variants (using the UCSC genome browser) [26]. The identified copy-number polymorphisms in the general population are short gains/losses, much shorter than the regions identified in the morphologically normal regions in the patient samples. Furthermore, the losses and gains in the general population are detected with a low population frequency, whereas we identified CNAs in morphologically normal regions in 50% of the 16 analyzed patient samples, arguing against the detection of a general polymorphism in our samples.

3D reconstructions of H&E slides

The H&E slides were scanned using Aperio ScanScope with a 10× objective (Leica Microsystems). Digital images were converted to RGB.tiff files using FIJI, and subsequently image stacks were aligned, both manually and automatically using Match software (developed by Herke Jan Noordmans, University Medical Center Utrecht, available upon request). DCIS lesions and normal ducts were selected using thresholding, and subsequent background noise reduction in FIJI. Image stacks were imported into the LAS X 3D visualization module (Leica Microsystems) to recreate 3D projections of the tissues.

Immunohistochemistry (IHC)

Immunohistochemistry for ER and HER2 on FFPE tumor sections was performed using a BenchMark ULTRA autostainer (Ventana Medical Systems, Santa Clara, CA, USA). Paraffin sections were cut at 3 μm, heated at 75 °C for 28 min, and deparaffinized in the instrument using EZ Prep solution (Ventana Medical Systems). Heat-induced antigen retrieval was carried out using Cell Conditioning 1 (CC1; Ventana Medical Systems) for 36 min at 95 °C. ER was detected using a Ready-to-Use anti-ER antibody, clone SP1 (# 790-4324; Roche Diagnostics, Rotkreuz, Switzerland), incubated for 32 min at 36 °C; HER2 was detected using an anti-HER2 antibody, clone SP3, 1:100 dilution overnight at 4 °C (# MA5-16348, Thermo Fisher Scientific). Bound ER antibody was detected using an UltraView Universal DAB Detection Kit (Ventana Medical Systems); HER2 was detected using Polymer-HRP Anti-Rabbit Envision (Dako, Agilent Technologies) incubated for 30 min and visualized with DAB incubated for 3–20 min (Sigma-Aldrich). Slides were counterstained with Hematoxylin and Bluing Reagent (Ventana Medical Systems).

Results

3D tissue clearing reveals HER2 overexpression in morphologically normal tissue

DCIS lesions can stretch over several centimeters of breast tissue [27,28]. In the histopathological analysis of DCIS, the DCIS-containing ducts present often intermingled with normal ducts and terminal duct lobular units (TDLUs) [29]. This presentation is apparent in the 2D view of tissue sections. However, this 2D analysis cannot inform on the spatial relationship of the DCIS lesions and normal ducts in the same breast area. To better understand the 3D relationship between DCIS and neighboring normal ducts, we tissue-cleared FFPE-DCIS resections from 14 DCIS patients for 3D volumetric high-resolution confocal imaging (Figure 1A and supplementary material, Table S1). DCIS and normal ducts could be readily recognized in the 3D space, and their morphological presentation in optical sections mirrored their presentation in the H&E-stained tissue section of the same block. Interestingly, 3D mapping revealed direct connections between DCIS ducts and normal ducts, where DCIS lesions were connected to normal TDLUs or normal larger ducts. This shows that DCIS grows in the architecture of the pre-existing ductal trees, and in direct contact with adjacent normal ductal segments (Figure 1B). Of note, we observed that two spatially separated DCIS lesions developed within sample 2, indicating that these aberrations developed independently, originating from the same ductal subtree (Figure 1B).

To understand the implications of this direct epithelial connection on the distribution of cancerous traits, we analyzed the five HER2-positive DCIS resections in our cohort, using HER2 overexpression as an immunological indicator of tissue transformation. Interestingly, in two cases, we found that HER2 overexpression was not limited to DCIS but could also be found in rare clusters of tightly packed cells that resided in presumably morphologically normal TDLUs in the same FFPE block (Figure 1C). This indicates that cancerous traits may be shared between DCIS and adjacent morphologically normal ductal segments.

3D reconstructions of human breast epithelium combined with CNA sequencing

To estimate the frequency of aberrant clonal fields surrounding DCIS (i.e. the presence of CNAs in the surrounding healthy epithelium of DCIS lesions), we selected FFPE tissue blocks from 16 different patients containing non-invasive DCIS lesions large enough for proper DNA isolation (supplementary material, Table S1). To understand the spatial relationship between healthy and DCIS ducts, we developed a pipeline to map detected CNAs onto 3D reconstructed tissues (Figure 2A). It should be noted that two tissue blocks were obtained from the resected DCIS material for patients 9 and 11, allowing us to perform both tissue

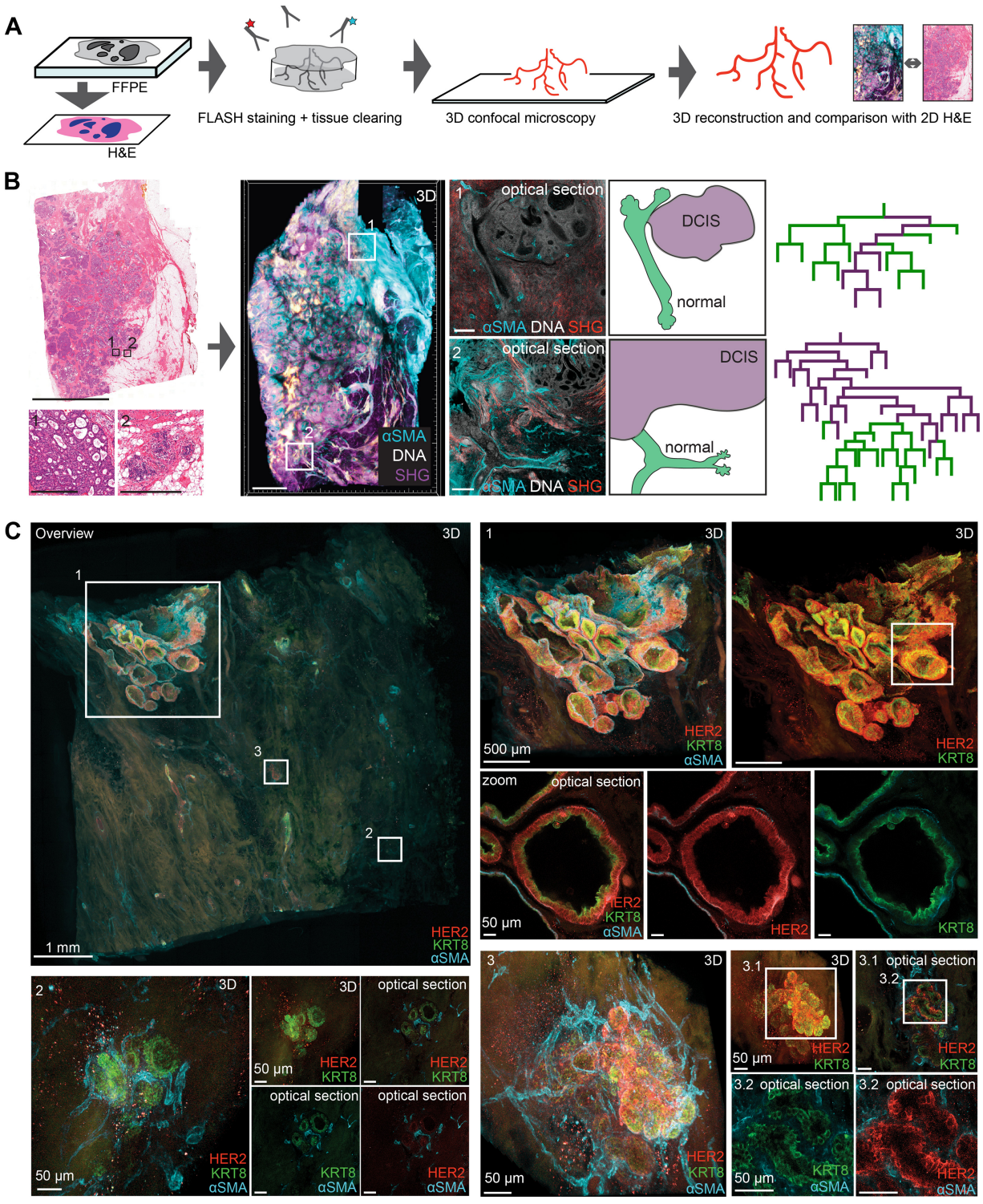

Figure 1. Tissue clearing and 3D imaging of archival DCIS-patient material reveals aberrant protein expression in adjacent normal ducts. (A) Scheme illustrating tissue clearing workflow. One tissue section of each DCIS-FFPE block was analyzed by H&E staining and histopathological evaluation. The remaining block was dewaxed, immunolabeled, and tissue-cleared, followed by 3D confocal microscopy. (B) 3D analysis and ductal mapping of a DCIS resection (sample 2). Left: 2D H&E section of the analyzed FFPE block. Scale bar: 1 cm. Insets show high magnifications of DCIS (1) and normal TDLUs (2) occupying the same block. Scale bars: 0.5 mm. Middle: 3D reconstruction of the same FFPE block after tissue clearing and immunolabeling for αSMA (cyan), nuclei (white), and fibrillary collagen [second harmonic generation (SHG), pink/red]. Scale bar: 3 mm. Insets are high magnifications of the optical sections showing DCIS connected to a large normal duct (1′) and DCIS connected to a normal TDLU (2). Scale bars: 0.2 mm. Right: drawings illustrating the connection of DCIS and normal ducts of the optical section, and dendrograms mapping the DCIS (magenta) and normal ducts (green) in the indicated areas. (C) 3D analysis of HER2-positive DCIS FFPE material (sample 1) stained for HER2 (red), KRT8 (green), and αSMA (cyan). Left: overview of the FFPE block. Scale bar: 1 mm. (1) High magnification of the indicated inset with a DCIS lesion. Scale bar: 0.5 mm. Below: optical sections of the indicated area in 1 demonstrate HER2-positive epithelium in the DCIS ducts. Scale bars: 50 µm. (2) Morphological normal TDLU at the indicated area. Optical sections show HER2-negative epithelium. Scale bars: 50 µm. (3) Morphologically normal TDLU in the indicated region. Optical sections show local clusters of HER2-positive cells (3.1 and 3.2). Scale bars: 50 µm.

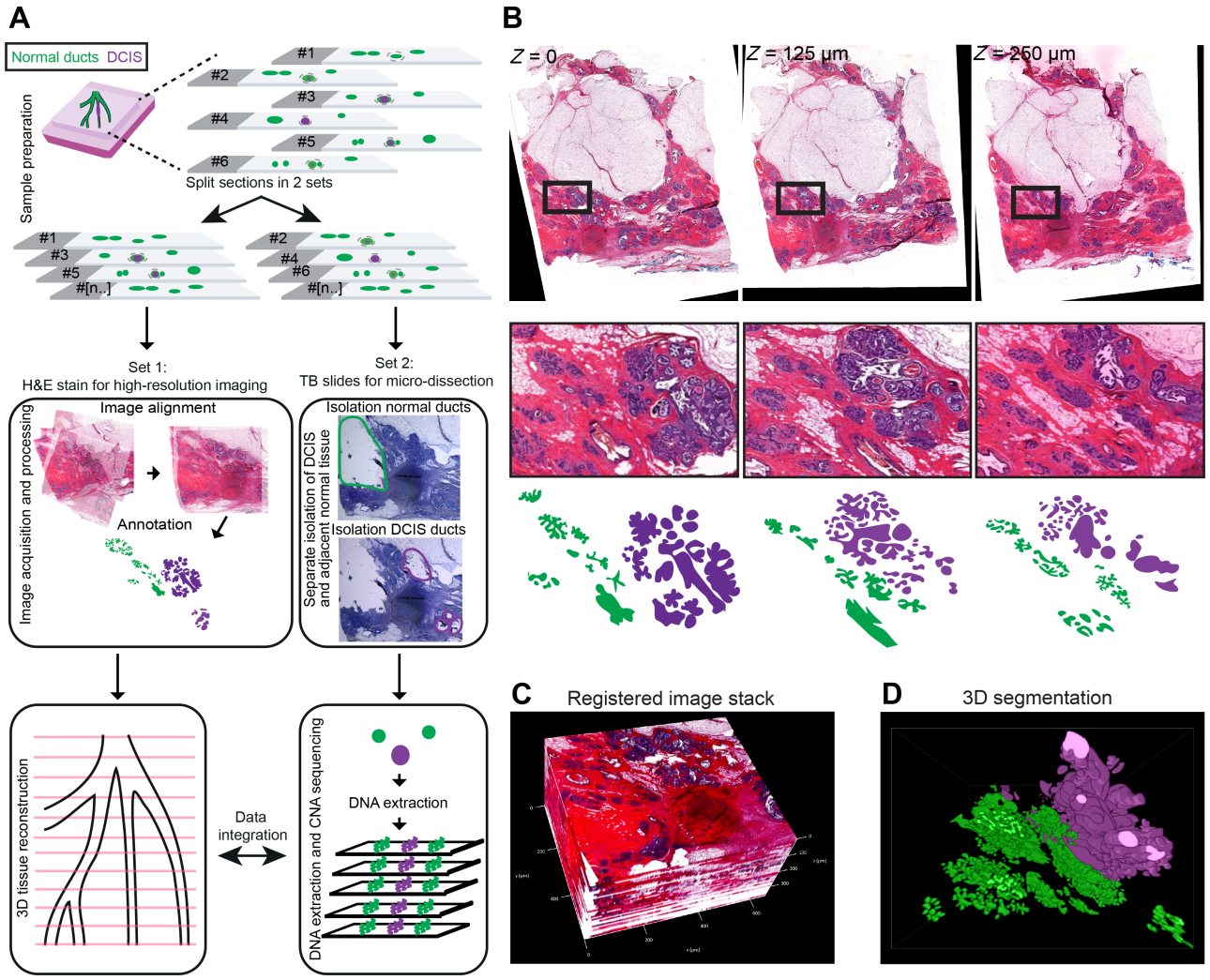

Figure 2. Identification of neighboring regions and 3D reconstruction of DCIS normal epithelial subtrees. (A) Schematic overview of the strategy used for image-guided characterization of epithelial field cancerization around the DCIS lesion. FFPE blocks are sectioned and imaged. Half of the tissue sections are used to isolate DNA from the DCIS and normal tissue; the other half are used for 3D reconstruction of the normal and DCIS containing ducts. (B–D) Workflow example for sample 11. (B) H&E sections of different *z*-sections, where DCIS lesions are indicated in purple and morphologically healthy tissue in green. (C) Image stack reconstructed for sample 11 from all the different *z*-sections. (D) 3D segmentation for sample 11 of healthy ducts (green) and DCIS lesions (purple).

clearing and CNA sequencing/3D reconstruction on the same DCIS tissue (supplementary material, Table S1). Each of the 16 tissue blocks were sliced into 10-µm sections, and the sections were split into two sets. The odd-numbered sections were stained with H&E, scanned with high resolution, and digital images were used for annotation of DCIS and healthy tissue by a pathologist (Figure 2B). The annotated images were registered, stacked, and aligned, allowing for the 3D reconstruction of DCIS and normal epithelium using 3D analysis software (Figure 2C,D). The even-numbered sections of the same tissues were stained with toluene blue. For each even-numbered section, the areas containing DCIS and the areas containing morphologically normal epithelium were manually micro-dissected and collected separately (supplementary material, Figure S1). DNA isolation and subsequent CNA sequencing (CNAseq) were performed on pooled samples derived from 20 sections (Figure 2A and supplementary material, Figure S2A).

CNAseq reveals epithelial field clonalization in a subset of samples

Using this pipeline, a spatially resolved map of CNAs was created for all 16 samples, in which the location of the detected CNAs could be mapped back onto the 3D structure of the DCIS ducts intermingled with or connected to the normal epithelial ducts. Using the 3D reconstructions of the 2D H&E sections, we manually built 2D schematic representations of the DCIS-normal mammary trees, including all the connecting ducts between the DCIS and the normal epithelium to identify the spatial relationship between the identified DCIS lesion and histologically healthy mammary ducts (Figure 2D). Combining the spatial information and the CNAseq data then allowed us to test whether epithelial field clonalization plays a role in the early events of pre-malignant lesion formation. The CNA profiles of the DCIS lesions were compared with the CNA profiles of the adjacent healthy epithelium with a *Z*-resolution

of 400 µm (supplementary material, Figure S2A). Interestingly, in addition to the previously identified overlap of HER2 overexpression between DCIS and nearby TDLUs (Figure 1C), we also identified instances of shared CNAs between the DCIS lesions and the connected normal epithelium. For instance, in sample 11, where our 3D reconstruction demonstrated that the normal epithelium and DCIS lesion are connected and part of the same ductal tree (Figure 3A,B), we identified a shared copy number gain on chromosome 6q25.1 throughout all sections of the normal epithelium (Figure 3B and supplementary material, Figure S2B). Chromosome 6q25.1 contains the *ESR1* gene, which is known to be amplified in a proportion of ER-positive breast cancer cases [30]. To validate this finding, we determined ER protein expression patterns in additional tissue sections derived from sample 11 and could confirm elevated ER expression levels in morphologically normal ducts and lobules in close proximity to the DCIS lesion (Figure 3C). This suggests that the DCIS lesion and the adjacent morphologically normal tissue originated from the same clonal field sharing the same ancestors. In addition, these data indicate that the spread of mutant cells with 6q25.1 gain may predispose the ductal tree to transformation.

Field clonalization can precede DCIS formation

To estimate whether mutant field clonalization precedes DCIS formation frequently, we examined the tissues of 16 patients. We found that in 50% of the DCIS samples (8 out of 16 patients), the micro-dissected normal ducts showed one or more CNAs (samples 9, 11, 17, 20, 22, 23, 25, and 26), whereas in the other 50% of DCIS samples, the normal epithelium appeared genomically unperturbed (Figure 4A and supplementary material,

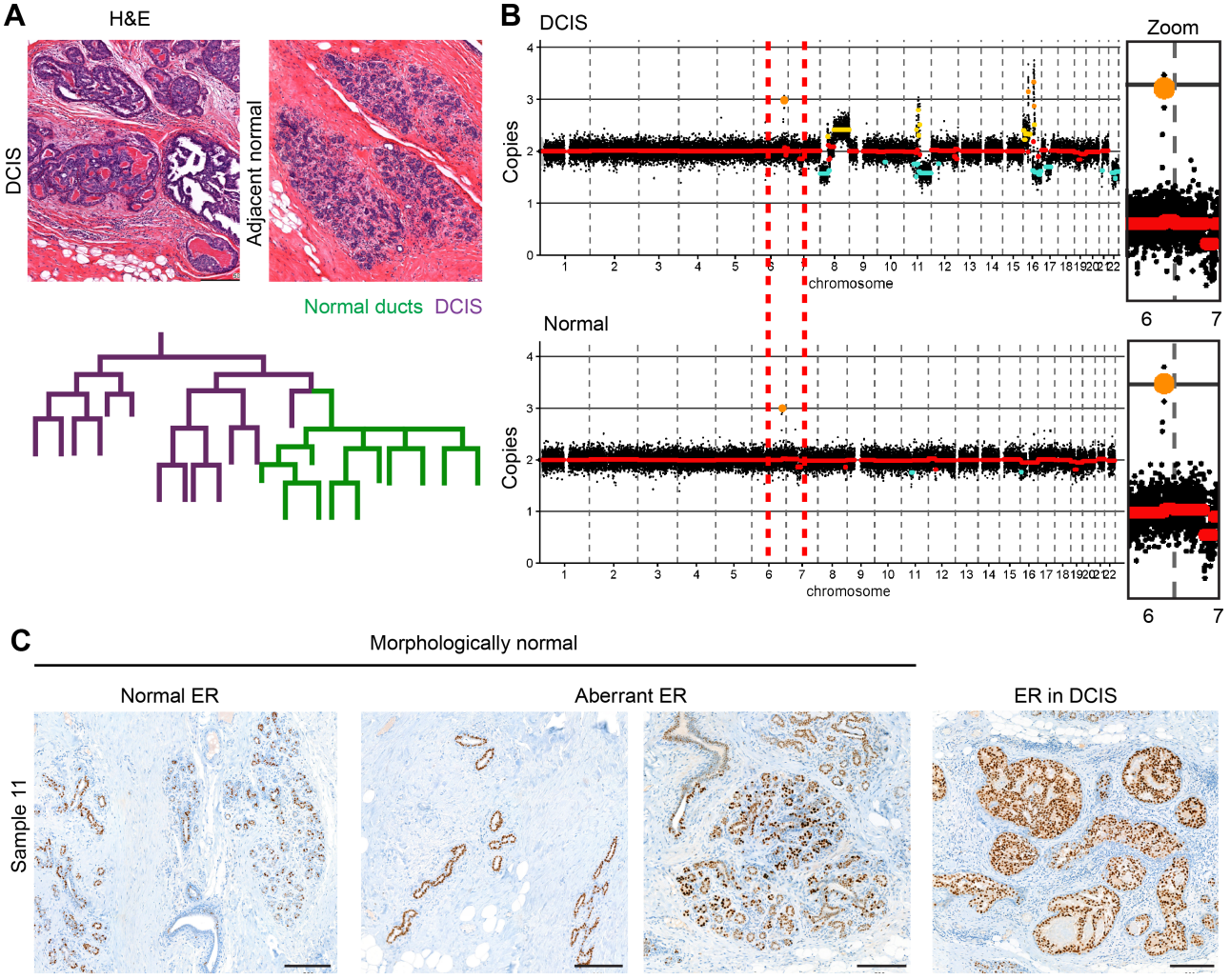

Figure 3. Shared CNA between morphologically normal epithelial tissue and adjacent DCIS regions. (A) Schematic structure of the subtree, based on the reconstruction in Figure 2D. H&E images (left) of a representative DCIS lesion and adjacent healthy ducts present in the subtree. Purple lines indicate the location of the DCIS lesion; green lines represent the histologically healthy mammary epithelium. (B) CNA plot of the DCIS lesion (top) and adjacent healthy tissue (bottom) of sample 11. The red dotted line indicates a shared copy number gain on chromosome 6, with a zoom shown on the left of the shared CNA in the DCIS lesion and healthy ducts. The *x*-axis represents the different chromosomes; the *y*-axis represents the copy number for each chromosome (log_2 ratio). (C) Immunohistochemistry labeling of ER in morphologically normal ducts and lobules (left panels) adjacent to DCIS lesions (right panel) revealing elevated ER protein levels in morphologically normal tissue and neighboring DCIS lesions, corresponding to the shared gain of a region in chromosome 6 containing the *ESR1* gene. Scale bars: 200 µm.

Figure S3). Importantly, all copy number profiles of normal mammary epithelium samples with CNAs showed a partial overlap with the CNA profile of the accompanying DCIS. For instance, in sample 11, the DCIS showed eight CNAs, of which only one was detected in the adjacent histologically healthy epithelium (Figure 3B). Aberrations found in pathologically healthy epithelium often included amplifications or deletions of genes found in the COSMIC Cancer Gene Census (CGC) [31], known to play a role in cancer formation, such as *ERBB2* and *ESR1* (Figure 4A). In addition to the amplification of *ESR1* in sample 11, we validated *ERBB2* amplification, detected by CNA analysis in sample 20, at the protein level by analyzing HER2 expression levels in morphologically normal ducts adjacent to the DCIS lesions (Figure 4B). The aberrations in histologically healthy tissue generally showed a lower $\log_2$ ratio for both amplification and deletions (supplementary material, Figures S2B and S3), indicating that the adjacent normal mammary epithelium consists of multiple epithelial clones intermingled in the same ducts, containing either healthy or mutant epithelial cells. Indeed, the HER2-amplification pattern in the morphologically normal ducts showed upregulation in only a subset of the epithelial cells lining the ducts, confirming this notion (Figure 4B). The eight samples containing mutant epithelial cells did not share any characteristics in terms of histological features, size, and number of aberrations, indicating that the presence of cells with CNAs is not linked to a specific driver aberration and can even be present in low-grade DCIS (Figure 4A and supplementary material, Figure S3 and Table S2). Moreover, we did not find evidence for correlation of fields of specific CNAs with DCIS grade, growth pattern, size, or molecular subtype (Figure 4A). In conclusion, our data show that in a subset of patients (50%), the normal epithelium connected to the DCIS lesion contains CNAs shared with the DCIS, suggesting

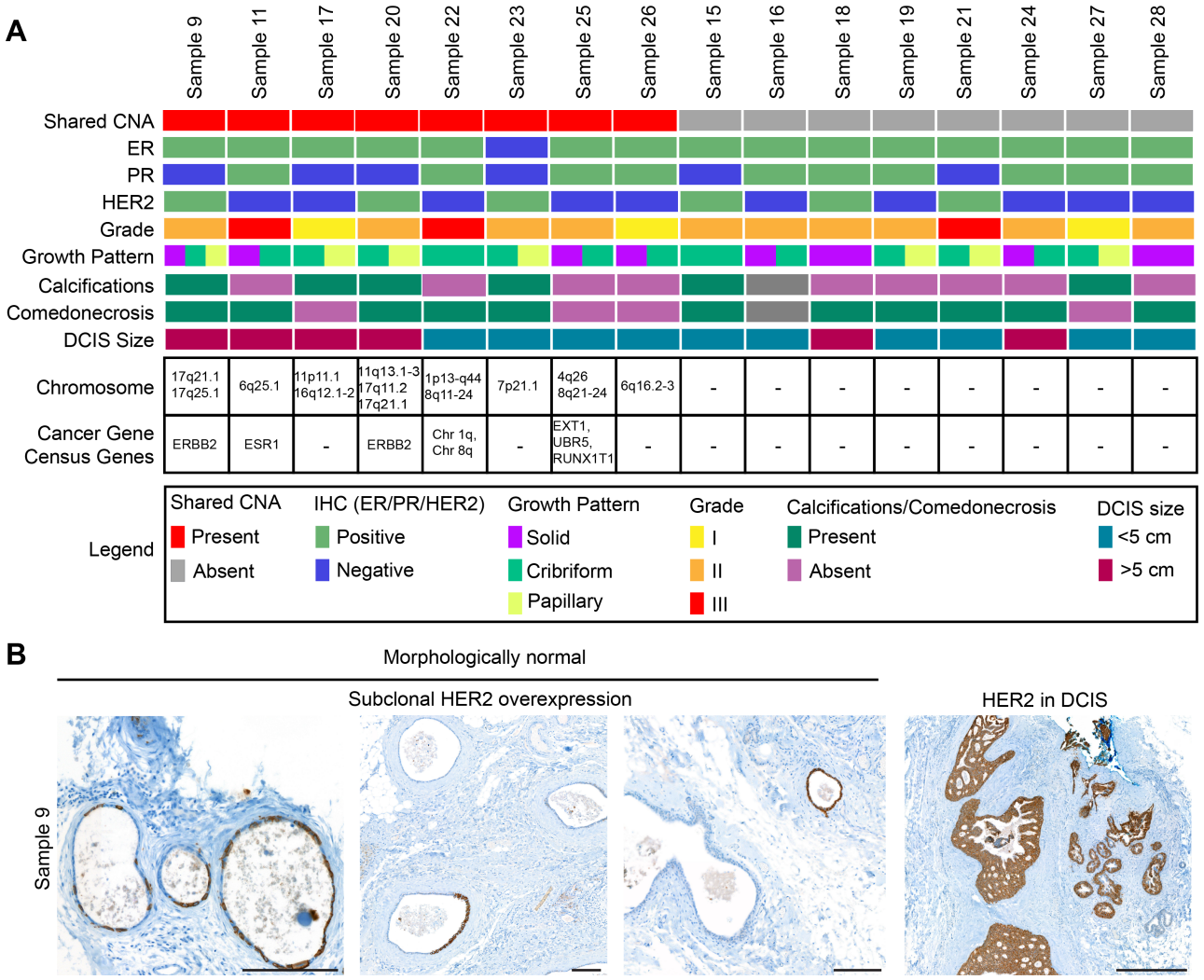

Figure 4. Field clonalization can precede DCIS formation. (A) Schematic overview of all 16 samples, the presence of shared aberrations in the pathologically healthy tissue, and the characteristics of the DCIS lesion, including ER, PR, HER2 status, grade, growth pattern, presence of calcifications or comedonecrosis, and DCIS size, as well as aberrations found in genes within the COSMIC Cancer Gene Census [31]. (B) Immunohistochemistry for HER2 in morphologically normal ducts and lobules (left panels) adjacent to DCIS lesions (right panel) revealing elevated HER2 protein expression levels in a subset of cells within the morphologically normal tissue, validating the shared gain of a region in chromosome 17 containing the *ERBB2* gene. Scale bar: 200 µm (left panels); 500 µm (right panel).

that mutant field clonalization occurs prior to DCIS formation, which provides support for the sick lobe hypothesis.

Field clonalization is a continuous process of mutation accumulation

Next, we analyzed whether field clonalization is a single event or a continuous process of accumulating genetic alternations. Interestingly, we did not observe widespread CNA intra-sample heterogeneity between the different spatially distinct sections of the same DCIS tissue samples (supplementary material, Figures S3 and S4). Instead, single DCIS lesions appeared homogeneous in terms of their CNA profile, indicating that the genetic heterogeneity at CNA level within a DCIS lesion is limited and that a single DCIS lesion is largely clonal. In contrast to the DCIS, we did observe heterogeneity in the histologically normal epithelium surrounding the DCIS lesions. For example, in sample 17, we observed that the normal epithelium was closely intermingled with the DCIS lesion (Figure 5A). Spatially resolved CNAseq in this sample identified a shared loss on chromosome 11p11 between the DCIS lesion and surrounding healthy tissue in all parts of the analyzed tissue (supplementary material, Figure S2B). In contrast, although an amplification of 16q12 was present in all parts for the DCIS lesion of sample 17, this specific amplification was only found within parts C and D of the normal tissue (Figure 5B,C and supplementary material, Figure S2B). This local CNA heterogeneity in the normal epithelium indicates that one of the cells in the clonal field lost chromosome 11p11, additionally amplified 16q12, and that this subclone spread over a larger area within the ductal tree. Importantly, one or multiple cells in the later subclone acquired additional genetic and/or microenvironmental triggers to transform into DCIS. This suggests that multiple aberrations can be fixed sequentially in morphologically normal tissue over time prior to transformation, thus predisposing parts of these clonal

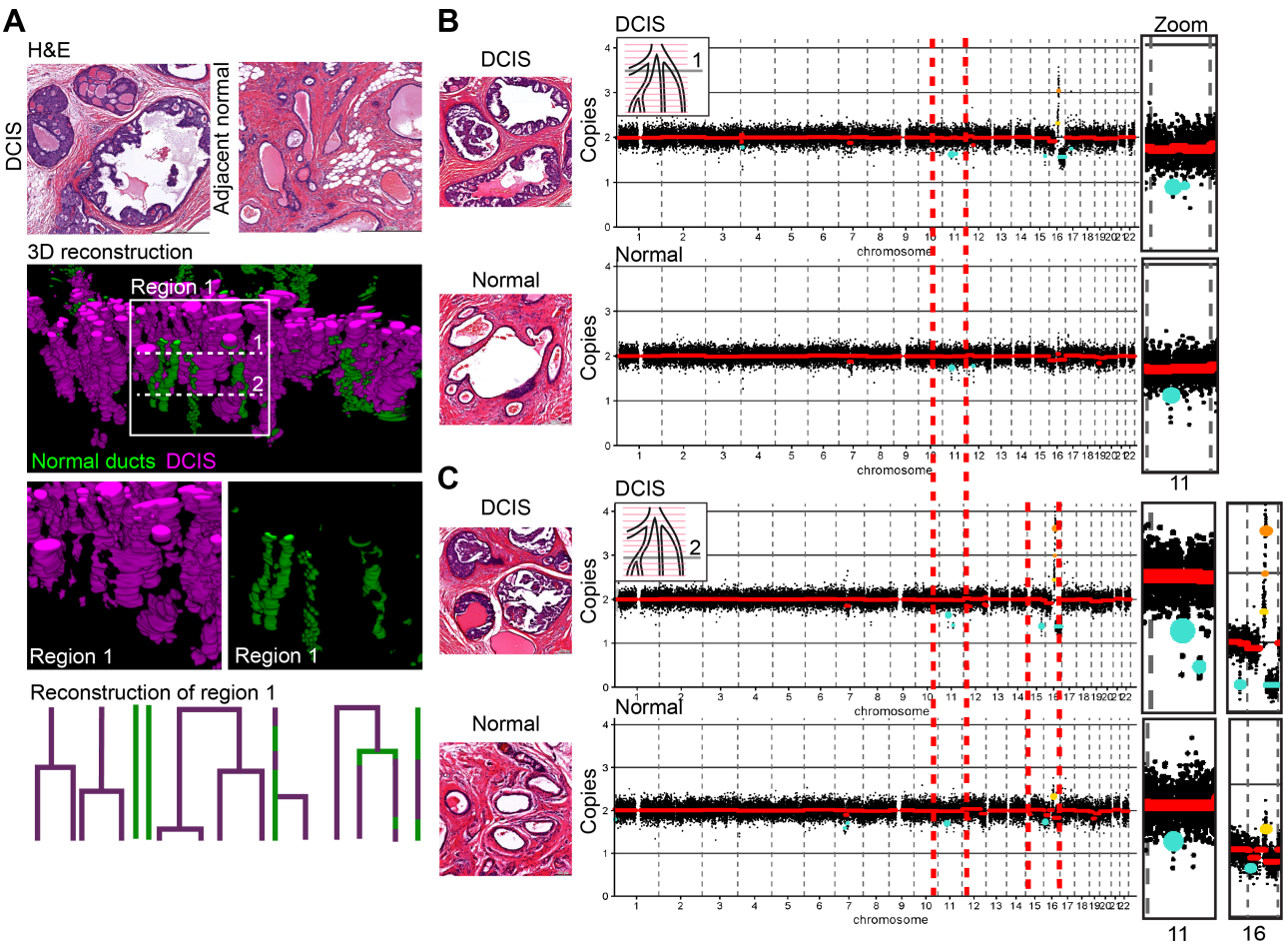

Figure 5. Morphologically normal areas present CNA heterogeneity. (A) Top: H&E images of a representative DCIS lesion and adjacent healthy ducts present in the subtree. Middle: 3D segmentation for sample 17, showing only the healthy ducts (green), DCIS lesion (purple), or combination. The white box indicates regions sequenced separately, present in the 3D segmentation. Bottom: schematic structure of the subtree for sample 17. Purple lines indicate the location of the DCIS lesion; green lines represent the histologically healthy mammary epithelium. (B) CNA plot of the DCIS lesion (top) and adjacent normal tissue (bottom) for region 1, with on the right a zoom of the shared aberration and on the left an example H&E of the DCIS and normal region in this section. The *x*-axis represents the different chromosomes; the *y*-axis represents the copy number for each chromosome ($\log_2$ ratio). (C) CNA plot of the DCIS lesion (top) and adjacent normal tissue (bottom) for region 2, with on the right a zoom of the two shared aberrations and on the left an example H&E of the DCIS and normal region in this section. The *x*-axis represents the different chromosomes; the *y*-axis represents the copy number for each chromosome ($\log_2$ ratio).

fields of healthy tissue to malignant transformation to different degrees. Importantly, these data also imply that the spread of clones over larger regions (i.e. clonalization) is a continuous process that leads to the accumulation of mutations.

Discussion

The existence of 'sick lobes' in the breasts of some DCIS patients was postulated many years ago [3,29] but to our knowledge has never been validated experimentally. To study whether mutant field clonalization can precede DCIS formation, we used 3D imaging of intact archival breast tissue samples from 14 DCIS patients, which revealed that DCIS lesions are connected to morphologically normal ducts and can arise independently in the same duct. Furthermore, using HER2-positive DCIS cases, in which we used HER2 overexpression as a marker of transformation, we identified HER2 overexpression in morphologically normal TDLUs adjacent to the HER2-overexpressing DCIS lesion in two out of five cases. This suggests that expansion of morphologically normal cells carrying *ERBB2* amplification can precede lesion formation.

To study whether other aberrations, besides *ERBB2* amplifications, can be detected in morphologically normal epithelial cells, we developed a method for image-guided characterization of the genomic profile of DCIS and nearby morphologically normal ducts and lobules in DCIS patient samples. We showed that the histologically normal mammary epithelium can already contain cells carrying genomic aberrations that are shared with the adjacent DCIS lesion. 3D reconstructions confirmed that these potentially predisposed, healthy ducts are connected to the DCIS lesion. In our sample set, we did not observe any specific DCIS characteristics related to field clonalization indicating that this phenomenon occurs in different subtypes and grades. It is important to mention that our sample set does not include triple-negative DCIS, a rare subtype of DCIS. Thus, our pipeline for image-guided and spatially resolved CNA profiling of clinical archival material confirms that mutant epithelial cells are abundantly present around pre-cancerous lesions in a subset of cases.

The CNA profiles of the histologically healthy breast epithelium show a smaller log ratio of the aberrations, suggesting that only a fraction of the healthy epithelial duct cells is composed of the mutant clone with predisposition to DCIS initiation, whereas the remainder may still consist of non-mutant epithelial clones. An alternative explanation could be that these samples were contaminated by true DCIS cells during the process of micro-dissection. However, based on the extent of the amplifications observed in morphologically normal tissue, 10–15% of isolated cells would need to derive from contamination with DCIS cells, which is a large proportion and therefore unlikely. Additionally, the verification of aberrant expression of ER and HER2 in morphologically normal tissues in samples 11 and 20, as well as 3D imaging of HER2-positive FFPE tissue, which shows field clonalization of HER2 overexpression in an independent manner, further supports that our findings do not entail contaminations.

Our 3D analyses reveal that mutant clonal fields can extend over large parts of the breast epithelial tree and may display a certain level of heterogeneity, suggesting a continuous process of aberration accumulation prior to transformation. The acquisition of additional aberrations in an existing mutant clonal field may lead to subclones with increased likelihood of transformation. Indeed, we observed independent DCIS lesions developing synchronously in the same ductal subtree (Figure 1B). This observation is in line with a recent study performing base-specific *in situ* sequencing in 2D tissue sections with DCIS, which demonstrated the presence of multiple microscopically segregated transformed subclones which were macroscopically part of the breast epithelial tree [32]. Together with our results, this study fits a model in which mutant, but non-transformed, clonal fields present the first step towards transformation, locally elevating the risk for transformation within that part of the breast epithelium. Acquisition of additional aberrations may lead to DCIS formation, and if by chance multiple subclones within the mutant field acquire the right aberrations, multi-clonal DCIS lesions may develop within the same part of the ductal tree [33,34].

Although not very frequent, ipsilateral recurrence after DCIS resection is observed in 3% of patients in the first 5 years after surgery [35,36]. Mutant field clonalization in healthy ducts surrounding the DCIS lesion could be the source of this recurrence, as the remaining morphologically untransformed ducts that surrounded the resected DCIS lesion were already predisposed to develop a subsequent DCIS/IBC with similar characteristics [37]. However, future studies will be required to confirm this finding. First, a larger cohort of patients should be studied to confirm mutant field clonalization in DCIS patients. Moreover, these studies should include a follow-up period to determine whether patients with field clonalization have an elevated risk of developing an ipsilateral recurrence. Ultimately, future studies need to elucidate whether mutant field clonalization of morphologically normal ducts is a risk factor for ipsilateral recurrence that should be taken into account in clinical practice.

Acknowledgements

We would like to acknowledge the NKI Core Facility Molecular Pathology & Biobanking (CFMPB) for providing the tissue samples and the Genomics Core Facility (GCF) for sequencing the samples. Finally, we thank the Grand Challenge PRECISION Consortium Steering Group: Jelle Wesseling, Jos Jonkers, Jacco van Rheenen, Esther H Lips, Marjanka Schmidt, Lodewyk FA Wessels, Proteeti Bhattacharjee

(Netherlands Cancer Institute, Amsterdam, The Netherlands), Alastair Thompson (Baylor College of Medicine, Houston, TX, USA), Serena Nik-Zainal, Helen Davies (University of Cambridge, Cambridge, UK), Elinor J Sawyer (King's College London, London, UK), Andrew Futreal, Nicholas Navin (MD Anderson Cancer Center, Houston, TX, USA), E Shelley Hwang (Duke University School of Medicine, Durham, NC, USA), Fariba Behbod (Kansas University Medical Center, Kansas City, KS, USA), Daniel Rea (University of Birmingham, Birmingham, UK), Hilary Stobart (Independent Cancer Patients' Voice, UK), Deborah Collyar (Patient Advocates in Research, USA), Donna Pinto (DCIS 411, USA), Ellen Verschuur, and Marja van Oirsouw (Borstkankervereniging Nederland, The Netherlands).

Author contributions statement

SJH, HAM, EHL, JW, JvR and CLGJS conceived the study and designed experiments. SJH, HAM, MC, EB, EHL and CLGJS performed experiments and analysis. Preparation for CNAseq was carried out by SJH, with the help of PK, CvdB and SS. CNAseq analysis was performed by EB and MS. HAM performed tissue clearing experiments. EHL, LFAW, JvR, JW, JJ and CLGJS supervised the study. SJH, HAM and CLGJS compiled the figures and wrote the manuscript. All authors approved the final version of the manuscript.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. Early Breast Cancer Trialists' Collaborative Group (EBCTCG), Darby S, McGale P, *et al*. Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10,801 women in 17 randomised trials. *Lancet* 2011; **378:** 1707–1716.
2. Mamounas EP, Anderson SJ, Dignam JJ, *et al*. Predictors of locoregional recurrence after neoadjuvant chemotherapy: results from combined analysis of National Surgical Adjuvant Breast and Bowel Project B-18 and B-27. *J Clin Oncol* 2012; **30:** 3960–3966.
3. Tot T. The theory of the sick breast lobe and the possible consequences. *Int J Surg Pathol* 2007; **15:** 369–375.
4. Tan MP, Tot T. The sick lobe hypothesis, field cancerisation and the new era of precision breast surgery. *Gland Surg* 2018; **7:** 611–618.
5. Gadaleta E, Thorn GJ, Ross-Adams H, *et al*. Field cancerization in breast cancer. *J Pathol* 2022; **257:** 561–574.
6. Slaughter DP, Southwick HW, Smejkal W. Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. *Cancer* 1953; **6:** 963–968.
7. Forsberg LA, Rasi C, Pekar G, *et al*. Signatures of post-zygotic structural genetic aberrations in the cells of histologically normal breast tissue that can predispose to sporadic breast cancer. *Genome Res* 2015; **25:** 1521–1535.
8. Jakubek YA, Chang K, Sivakumar S, *et al*. Large-scale analysis of acquired chromosomal alterations in non-tumor samples from patients with cancer. *Nat Biotechnol* 2020; **38:** 90–96.
9. Ellsworth DL, Ellsworth RE, Liebman MN, *et al*. Genomic instability in histologically normal breast tissues: implications for carcinogenesis. *Lancet Oncol.* 2004; **5:** 753–758.
10. Moinfar F, Man YG, Arnould L, *et al*. Concurrent and independent genetic alterations in the stromal and epithelial cells of mammary carcinoma: implications for tumorigenesis. *Cancer Res* 2000; **60:** 2562–2566.
11. Cavalli LR, Singh B, Isaacs C, *et al*. Loss of heterozygosity in normal breast epithelial tissue and benign breast lesions in *BRCA1/2* carriers with breast cancer. *Cancer Genet Cytogenet* 2004; **149:** 38–43.
12. Trujillo KA, Hines WC, Vargas KM, *et al*. Breast field cancerization: isolation and comparison of telomerase-expressing cells in tumor and tumor adjacent, histologically normal breast tissue. *Mol Cancer Res* 2011; **9:** 1209–1221.
13. Graham K, Ge X, de Las Morenas A, *et al*. Gene expression profiles of estrogen receptor-positive and estrogen receptor-negative breast cancers are detectable in histologically normal breast epithelium. *Clin Cancer Res* 2011; **17:** 236–246.
14. Aran D, Camarda R, Odegaard J, *et al*. Comprehensive analysis of normal adjacent to tumor transcriptomes. *Nat Commun* 2017; **8:** 1077.
15. Ruan X, Liu H, Boardman L, *et al*. Genome-wide analysis of loss of heterozygosity in breast infiltrating ductal carcinoma distant normal tissue highlights arm specific enrichment and expansion across tumor stages. *PLoS One* 2014; **9:** e95783.
16. Wang X, Stolla M, Ring BZ, *et al*. p53 alteration in morphologically normal/benign breast tissue in patients with triple-negative high-grade breast carcinomas: breast p53 signature? *Hum Pathol* 2016; **55:** 196–201.
17. Abdalla M, Tran-Thanh D, Moreno J, *et al*. Mapping genomic and transcriptomic alterations spatially in epithelial cells adjacent to human breast carcinoma. *Nat Commun* 2017; **8:** 1245.
18. Cereser B, Jansen M, Austin E, *et al*. Analysis of clonal expansions through the normal and premalignant human breast epithelium reveals the presence of luminal stem cells. *J Pathol* 2018; **244:** 61–70.
19. Messal HA, Almagro J, Zaw Thin M, *et al*. Antigen retrieval and clearing for whole-organ immunofluorescence by FLASH. *Nat Protoc* 2021; **16:** 239–262.
20. Hutten SJ, de Bruijn R, Lutz C, *et al*. A living biobank of patient-derived ductal carcinoma *in situ* mouse-intraductal xenografts identifies risk factors for invasive progression. *Cancer Cell* 2023; **41:** 986–1002.e9.
21. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 2009; **25:** 1754–1760.
22. Li H, Handsaker B, Wysoker A, *et al*. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009; **25:** 2078–2079.
23. Scheinin I, Sie D, Bengtsson H, *et al*. DNA copy number analysis of fresh and formalin-fixed specimens by shallow whole-genome sequencing with identification and exclusion of problematic regions in the genome assembly. *Genome Res* 2014; **24:** 2022–2032.
24. Nilsen G, Liestøl K, Van Loo P, *et al*. Copynumber: efficient algorithms for single- and multi-track copy number segmentation. *BMC Genom* 2012; **13:** 591.
25. Poell JB, Mendeville M, Sie D, *et al*. ACE: absolute copy number estimation from low-coverage whole-genome sequencing data. *Bioinformatics* 2018; **35:** 2847–2849.
26. Lappalainen I, Lopez J, Skipper L, *et al*. DbVar and DGVa: public archives for genomic structural variation. *Nucleic Acids Res* 2013; **41:** D936–D941.
27. Holland R, Hendriks JH, Vebeek AL, *et al*. Extent, distribution, and mammographic/histological correlations of breast ductal carcinoma *in situ*. *Lancet* 1990; **335:** 519–522.

28. Bellamy CO, McDonald C, Salter DM, *et al*. Noninvasive ductal carcinoma of the breast: the relevance of histologic categorization. *Hum Pathol* 1993; **24:** 16–23.
29. Tot T. DCIS, cytokeratins, and the theory of the sick lobe. *Virchows Arch* 2005; **447:** 1–8.
30. Holst F, Moelans CB, Filipits M, *et al*. On the evidence for *ESR1* amplification in breast cancer. *Nat Rev Cancer* 2012; **12:** 149.
31. Sondka Z, Bamford S, Cole CG, *et al*. The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. *Nat Rev Cancer* 2018; **18:** 696–705.
32. Lomakin A, Svedlund J, Strell C, *et al*. Spatial genomics maps the structure, nature and evolution of cancer clones. *Nature* 2022; **611:** 594–602.
33. Casasent AK, Schalck A, Gao R, *et al*. Multiclonal invasion in breast tumors identified by topographic single cell sequencing. *Cell* 2018; **172:** 205–217.e12.
34. Casasent AK, Edgerton M, Navin NE. Genome evolution in ductal carcinoma *in situ*: invasion of the clones. *J Pathol* 2017; **241:** 208–218.
35. Elshof LE, Schaapveld M, Schmidt MK, *et al*. Subsequent risk of ipsilateral and contralateral invasive breast cancer after treatment for ductal carcinoma *in situ*: incidence and the effect of radiotherapy in a population-based cohort of 10,090-women. *Breast Cancer Res Treat* 2016; **159:** 553–563.
36. Thompson AM, Clements K, Cheung S, *et al*. Management and 5-year outcomes in 9938 women with screen-detected ductal carcinoma *in situ*: the UK Sloane Project. *Eur J Cancer* 2018; **101:** 210–219.
37. Visser LL, Elshof LE, Schaapveld M, *et al*. Clinicopathological risk factors for an invasive breast cancer recurrence after ductal carcinoma *in situ* – a nested case–control study. *Clin Cancer Res* 2018; **24:** 3593–3601.

SUPPLEMENTARY MATERIAL ONLINE

Figure S1. Micro-dissection of DCIS and normal neighboring regions

Figure S2. Consecutive sectioning strategy reveals epithelial field cancerization in healthy epithelium

Figure S3. Spatially resolved CNAseq results demonstrate shared CNAs between normal and DCIS ducts in a subset of samples

Figure S4. Subset of samples without shared CNAs between normal and DCIS ducts.

Table S1. Details for the 28 FFPE tissue samples used in this study

Table S2. Contingency table for patient characteristics of FFPE tissue samples used in this study, sorted by presence or absence of CNAs in morphologically normal epithelium adjacent to DCIS