Ultrasensitive Detection and Monitoring of Circulating Tumor DNA Using Structural Variants in Early-Stage Breast Cancer

Mitchell J. Elliott[1], Karen Howarth[2], Sasha Main[3], Jesús Fuentes Antrás[4], Philippe Echelard[5], Aaron Dou[1], Eitan Amir[1], Michelle B. Nadler[1], Elizabeth Shah[6], Celeste Yu[6], Scott Bratman[7], Taylor Bird[1], June Roh[1], Elza C. de Bruin[8], Christopher Rushton[2], Yilun Chen[2], Sergii Gladchuk[2], Anthony M. George[2], Sofia Birkeälv[2], Miguel Alcaide[2], Lucia Oton[2], Girish Putcha[2], Samuel Woodhouse[2], Philippe L. Bedard[1], Lillian L. Siu[1], Hal K. Berman[9], and David W. Cescon[1]

ABSTRACT

Purpose: The detection of circulating tumor DNA (ctDNA) after curative-intent therapy in early-stage breast cancer is highly prognostic of disease recurrence. Current ctDNA assays, mainly targeting single-nucleotide variants, vary in sensitivity and specificity. Although increasing the number of single-nucleotide variants in tumor-informed assays improves sensitivity, structural variants (SV) may achieve similar or better sensitivity without compromising specificity. SVs occur across all cancers, linked to genomic instability and tumorigenesis, with unique tumor- and patient-specific breakpoints occurring throughout the genome. SVs in breast cancer are underexplored, and their potential for ctDNA detection and monitoring has not been fully evaluated.

Experimental Design: We retrospectively analyzed a tumor-informed SV-based ctDNA assay in a cohort of patients with early-stage breast cancer (n = 100, 568 timepoints) receiving neoadjuvant systemic therapy, evaluating ctDNA dynamics and lead times to clinical recurrence in the postoperative period.

Results: ctDNA was detected in 96% (91/95) of participants at baseline with a median variant allele frequency of 0.15% (range: 0.0011%–38.7%); of these, 10% (9/91) had a variant allele frequency <0.01%. ctDNA detection at cycle 2 (C2) of neoadjuvant therapy was associated with a higher likelihood of distant recurrence (log-rank $P = 0.047$) and enhanced residual cancer burden prognostication (log-rank $P = 0.041$). ctDNA was detected prior to distant recurrence in all cases (100% sensitivity) with a median lead time of 417 days (range: 4–1,931 days).

Conclusions: These results demonstrate the clinical validity of ultrasensitive ctDNA detection and monitoring using SVs. Prospective trials are required to evaluate ctDNA-guided treatment strategies.

Introduction

The detection of circulating tumor DNA (ctDNA) after curative-intent therapy in early-stage breast cancer (EBC) is strongly associated with disease recurrence (1–7). First-generation tumor-informed ctDNA assays targeting multiple tumor-specific single-nucleotide variants (SNV) by multiplex PCR or next-generation sequencing (NGS) have demonstrated analytical and clinical validity (4, 5, 8, 9). Tracking 100 to 1,000 of SNVs enables ctDNA detection at an ultrasensitive level [limit of detection at 95% certainty (LoD95) <0.001%/10 parts per million (ppm)], providing highly sensitive and specific results (10–12). However, other types of genomic alterations are also found in ctDNA, offering alternative strategies for the detection of molecular residual disease (MRD; ref. 13).

Somatic structural variants (SV) are well-established hallmarks of cancer, arising from and contributing to genomic instability and oncogenesis (14, 15). SVs, including breakpoints and rearrangements, are highly tumor specific and often reflect the underlying tumor biology (15). Despite their prevalence, most SVs lack established clinical relevance, and their broader role in breast cancer remains incompletely understood. While the detection of chromosomal rearrangements in ctDNA has been shown to detect MRD, the use of SVs for ctDNA detection and monitoring in EBC has not been comprehensively evaluated (16–18). Tumor-specific SVs can be interrogated with digital PCR (dPCR), which may permit sensitive and specific detection of low levels of ctDNA in the presence of abundant normal cell-free DNA (cfDNA). Multiplex dPCR offers technical and operational advantages over NGS for this application, eliminating the need for high-depth sequencing and reducing the risk of sequencing errors. SVs are resistant to single-bp PCR and NGS errors that may impact the specificity of SNV-based approaches.

In this study, we evaluated a personalized SV-based dPCR ctDNA assay constructed from tumor-only whole genome sequencing (WGS) in a cohort of previously untreated patients with stages I–III

[1]Division of Medical Oncology and Hematology, Department of Medicine, Princess Margaret Cancer Centre, University of Toronto, Toronto, Canada. [2]SAGA Dx, Morrisville, North Carolina. [3]Department of Medical Biophysics, University of Toronto, Toronto, Canada. [4]NEXT Oncology, Hospital Universitario Quirónsalud Madrid, Madrid, Spain. [5]Department of Pathology, Université de Sherbrooke, Sherbrooke, Canada. [6]Cancer Genomics Program, Princess Margaret Cancer Centre, Toronto, Canada. [7]Division of Radiation Oncology, Princess Margaret Cancer Centre, University of Toronto, Toronto, Canada. [8]Oncology R&D, AstraZeneca, Cambridge, United Kingdom. [9]Department of Pathology and Laboratory Medicine, University Health Network, Toronto, Canada.

Corresponding Authors: David W. Cescon, Princess Margaret Cancer Centre, 610 University Avenue, Toronto, ON M5G 2M9, Canada. E-mail: dave.cescon@uhn.ca; and Mitchell J. Elliott, mitchell.elliott@uhn.ca

Clin Cancer Res 2025;31:1520–32

doi: 10.1158/1078-0432.CCR-24-3472

Translational Relevance

Circulating tumor DNA (ctDNA) detection has an emerging role in the management of breast cancer, offering a noninvasive means to monitor disease status and detect recurrence. Ultrasensitive levels of detection, high specificity, and feasibility of repeated testing are desirable characteristics for ctDNA assays. In this study, we demonstrate that longitudinal monitoring of tumor-specific structural variants—common alterations across many tumor types, including all subtypes of breast cancer—permits the detection of ctDNA in nearly all participants within this cohort of untreated patients with early-stage breast cancer, prior to the initiation of neoadjuvant therapy. On-treatment ctDNA monitoring enabled the early identification of patients at high risk of recurrence, and ctDNA was detected prior to relapse in participants across all receptor subtypes. These findings demonstrate the clinical validity of this assay approach for ultrasensitive ctDNA monitoring.

EBC of all receptor subtypes receiving standard neoadjuvant therapy (NAT; ref. 19). We explored ctDNA dynamics during NAT and their relationship to clinical outcomes and evaluated the ability of serial ctDNA monitoring to anticipate disease recurrence in the adjuvant setting. This study aimed to characterize the performance of this new class of assay and its potential for clinical use.

Materials and Methods

Burden of pan-cancer and breast cancer SVs

Paired tumor-normal WGS data from 16,247 patients with cancer was obtained from the Genomics England 100,000 Genomes Project made possible through access to data in the National Genomic Research Library v5.1, which is managed by Genomics England Limited (a wholly owned company of the UK Department of Health and Social Care; refs. 20, 21). Samples were previously sequenced to a mean of 60× coverage and 30× coverage for the tumor biopsy and matched germline sample, respectively. After removing serial samples collected from the same patients, existing somatic SV calls were obtained from the Genomics England Research Environment release v18 (December 21, 2023), as generated by Manta (22). Raw Manta VCF files were converted to BEDPE files using Viola-SV (23). To remove artifact SVs stemming from regions with poor mappability, we aggregated additional Manta-derived germline SV calls from 2,000 unrelated patients with rare diseases in release v18 (December 21, 2023) and removed any somatic SV in which both breakpoints fell within 2 bp of a germline-derived SV call. To select for high-confidence SV, SVs were subsequently filtered to remove those that (i) failed Manta's filters, (ii) were <50 bp, and (iii) had a breakpoint homologous sequence >2 bp, or (iv) in which either breakpoint was called a nonprimary contig. Downstream statistical analysis and visualizations were generated using R4.0.2 and RStudio 2022.12.0, using the R package Tidyverse. For breast cancer–specific analysis, we limited the subset to 1,180 breast cancer cases in which the estrogen receptor (ER) and human epidermal growth factor receptor 2 (HER2) statuses were available. Student *t* tests were used to calculate the SV burden across different breast cancer subtypes.

To assign the copy number of a SV, genome-wide copy number calls generated from Canvas were obtained from Genomics England Research Environment Release v18 (December 21, 2023; ref. 24). The copy number of each breakpoint was assigned by intersecting the Canvas-derived copy number segments with the breakpoint positions using BEDTools (RRID: SCR_006646; ref. 25). If no Canvas-derived copy number state was available (commonly in centromeric or telomeric regions), we defaulted to a copy number state of 2. In case of a disagreement in the copy number state assigned to each SV breakpoint, the lower copy number state was chosen. The associated copy number state of the top 10% of SVs in each sample was used for downstream analyses.

SV-based ctDNA assay workflow

The Pathlight assay (SAGA Dx) is a tumor-informed ctDNA assay that uses WGS of tumor tissue and multiplex dPCR analysis of plasma-derived cell free DNA (cfDNA) to detect ctDNA. The assay can be divided into two main components: fingerprint generation and orthogonal validation followed by cfDNA analysis.

In fingerprint generation and orthogonal validation, DNA was extracted from patient formalin-fixed, paraffin-embedded (FFPE) tumor tissue using the Mag-Bind FFPE DNA/RNA kit (Omega Biotek). Extraction quality control was performed, including FFPE DNA quantification by fluorometry, using the Qubit 1× dsDNA Broad Range (BR) Assay (Thermo Fisher Scientific) and Qubit Flex Fluorometer (Thermo Fisher Scientific). Library preparation was performed using 22 to 100 ng of FFPE DNA. Unique device identifier barcode design included the use of full-length adapters carrying 10-bp dual-indexes (QIAseq UDI Y-Adapters, QIAGEN). Library QC was performed by automated electrophoresis using the TapeStation D1000 and High Sensitivity D1000 ScreenTape assays and the 4200 TapeStation System (Agilent Technologies). Libraries were pooled, with PhiX Control Library added at 1.5%, then sequenced using the NovaSeq X (Illumina) with 2 × 150 bp cycling. Library pool loading concentration was 120 pmol/L with a target coverage of 15×. The resulting sequencing data were demultiplexed using bcl2fastq (RRID: SCR_015058), and FASTQ files were processed with FASTP (RRID: SCR_016962) to remove read-through adapter sequences. Processed sequencing reads were aligned against the human reference genome T2T-chm13v2.0 using BWA-MEM2 (RRID: SCR_022192), and duplicate marking was performed using biobambam2 (RRID: SCR_003308). Quality control was performed using the Picard toolkit (RRID: SCR_006525) and FastQC (RRID: SCR_014583). After SV calling, candidate germline SVs were removed using gnomAD v4.0 (RRID: SCR_014964), removing any candidate SVs in which both breakpoints fell within 2 bp of an event listed in gnomAD. dPCR primers were generated for each SV with a primer sequence length of 16 to 56 bps and a melting temperature of 43 to 62°C. Up to 16 SVs were then selected for each individual to generate a tumor-specific dPCR fingerprint. Oligonucleotides were obtained from Integrated DNA Technologies.

All fingerprint SVs identified by WGS were orthogonally validated against the tumor FFPE DNA, matched buffy coat DNA, and unmatched normal genomic DNA. This process removes primer sets that contain SVs that are germline or lead to spurious DNA amplification. Matched buffy coat DNA was extracted, with 75 μL target minimum input volume, using the Mag-Bind Blood & Tissue DNA HDQ kit (Omega Bio-tek). Extraction QC was performed, including DNA quantification by fluorometry, using the Qubit 1× dsDNA BR Assay kit (Thermo Fisher Scientific) and the Qubit Flex Fluorometer instrument (Thermo Fisher Scientific). Unmatched normal genomic DNA was from human adult normal tissue (peripheral blood leukocytes) from a single donor and was

obtained from BioChain. The unmatched normal genomic DNA was fragmented by sonication to mimic the fragment size distribution of cfDNA using 8 AFA-TUBE TPX Strips (Covaris) and the ML230 Focused-ultrasonicator instrument (Covaris) by following the manufacturer's protocol for fragmentation to the 175-bp target size. Fragmentation QC was performed, including DNA quantification and fragmentation verification. Quantification was performed by fluorometry using the Qubit 1× dsDNA BR Assay kit (Thermo Fisher Scientific) and the Qubit Flex Fluorometer instrument (Thermo Fisher Scientific), and fragmentation verification was performed by automated electrophoresis using the cfDNA ScreenTape assay (Agilent Technologies) and the 4200 and 4150 TapeStation Systems (Agilent Technologies). Orthogonal validation was performed by dPCR using QIAcuity Nanoplate 8.5k 96-well (QIAGEN) plates and QIAcuity Digital PCR Systems (QIAGEN). Tumor FFPE DNA (1.28 ng input) was analyzed as the positive control and matched buffy coat DNA (16 ng input) and sheared unmatched normal genomic DNA (160 ng input) as the negative controls. About 14 μL of reaction mix was added per nanoplate reaction setup with SV primers spanning the fusion sequence breakpoint used for specific and accurate SV detection with a total of 40 cycles of dPCR as the endpoint. Hydrolysis probes (TAMRA, CY5, FAM, HEX, and ROX) were used for signal detection. An infrared reference dye was used for nanoplate partition identification. Nanoplate imaging was performed 5 minutes after thermal cycling. A total of four SVs and a positive control amplicon were analyzed per well, with a total of four nanowells per sample type. Therefore, including the controls, a total of 12 wells were analyzed per cfDNA sample. Eight samples were analyzed per 96-well nanoplate. Up to eight nanoplates can be processed simultaneously on an QIAcuity Eight system. Data analysis of the dPCR results included processing of relative fluorescence signal data with an automated thresholding algorithm.

cfDNA analysis consisted of SV preamplification followed by multitarget SV detection using four dPCR reactions per sample. A maximum of four somatic SVs were detected per reaction (resulting in a maximum of 16 somatic SVs per patient). Analysis was performed using patient cfDNA, tumor FFPE DNA, and unmatched normal genomic DNA. Patient cfDNA was extracted from 0.5 to 5.5 mL of plasma using the QIAamp Circulating Nucleic Acid Kit (QIAGEN). cfDNA extraction QC was performed, including cfDNA quantification by fluorometry, using the Qubit 1× dsDNA High Sensitivity (HS) Assay kit (Thermo Fisher Scientific) and the Qubit Flex Fluorometer instrument (Thermo Fisher Scientific). cfDNA analysis was performed by dPCR using QIAcuity Nanoplate 8.5k 96-well (QIAGEN) plates and QIAcuity Digital PCR System (QIAGEN), with 5 ng to 2 μg input. Tumor FFPE DNA (1.28 ng input) was analyzed as the positive controls and sheared unmatched normal genomic DNA (160 ng input) as the negative controls. dPCR was performed using conditions as described above. The analysis of dPCR results included processing of relative fluorescence signal data with a custom automated thresholding algorithm. A positive "ctDNA detected" result was called when a dPCR signal was detected for at least one SV.

Analytical validation of an SV-based ctDNA assay workflow

Cell line-derived DNA was obtained from the suppliers listed below for the preclinical analytical validation experiments. A fingerprint was designed for the breast cancer cell line BT474 ($HER2^+$ breast carcinoma, Cytion—Cell Lines Service GmbH, product ID: 300131GD5; RRID: CVCL_0179), and LoD95 was determined using the probit method for a standard cfDNA input amount (70 ng, RRID: CVCL_1C78, product ID: NA24385, Coriell DNA, fragmented to mimic cfDNA). An LoD95 confirmation study was performed on two additional cell lines, including SK-BR-3 (breast carcinoma, Cytion—Cell Lines Service GmbH, product ID: 300333GD5, RRID: CVCL_0033; Supplementary Fig. S2B) and FaDu (HTB-43; hypopharyngeal squamous cell carcinoma, Leibniz Institute DSMZ–German Collection of Microorganisms and Cell Cultures GmbH, product ID: ACC 784, FaDu DNA-5, RRID: CVCL_1218). All cell lines were authenticated using short tandem repeat sequencing and tested to be *Mycoplasma* free by the manufacturer [using the PlasmoTest—Mycoplasma Detection Kit (InvivoGen) and the certus QC—mycoADVANCED detection kit (Certus)].

Patient recruitment

This clinical evaluation consists of a retrospective analysis performed on a cohort of participants with available samples at the time of primary analysis. Patients with EBC of any receptor subtype (triple-negative, TNBC; HER-positive, $HER2^+$; and ER-positive/HER2-negative, or $ER^+/HER2^-$) receiving standard NAT at the Princess Margaret Cancer Centre were prospectively enrolled from October 2016 for serial blood collection and banking (Supplementary Fig. S1A). From August 2017, accrual had transitioned to the Liquid Biopsy Evaluation and Repository Development at Princess Margaret (LIBERATE) cohort (NCT03702309), with the last participant enrolled in March 2024. Clinical decisions and treatment selection were made by the treating physicians and participants in keeping with the local standard of care. This observational study was approved by the University Health Network Research Ethics Board (#17-5962 and #23-5446) in accordance with the Declaration of Helsinki; all patients provided written informed consent. There was no return of ctDNA results to treating clinicians, given the retrospective nature of the analyses.

Clinical specimens

FFPE tumor tissue from diagnostic biopsies or surgical resections (residual disease) was used for fingerprint generation. Pathology review was performed targeting ≥20% tumor cellularity (percentage of total cells present in the specimen provided for sequencing). Either 5-μm sections on slides or 1- to 3-mm^2 cores were used for genomic DNA extraction, in accordance with institutional policies, to conserve diagnostic tissue. After DNA extraction, WGS was performed by SAGA Dx. Serial blood samples (3× Streck blood collection tubes) were collected at baseline, during treatment, perioperatively, and during follow-up. After surgery, blood was collected at the time of routine standard-of-care follow-ups. Samples were double spun, aliquoted, and stored at −80°C.

Participants with tissue samples meeting the above criteria and available plasma were selected from the larger cohort for analysis. Approximately 4 mL of plasma per timepoint was used; germline DNA from baseline buffy coat was used to exclude germline variation and chromatin immunoprecipitation SV from assay design. SAGA Dx was blinded to clinical outcomes at the time of data generation.

Clinical variables

Baseline clinical stage from medical records is reported in accordance with the American Joint Committee on Cancer eighth edition (prognostic staging). Tumor size (T-stage) and nodal involvement (N-stage) were assessed via diagnostic MRI, mammogram, and/or ultrasound, along with pathologic assessment where

available (e.g., axillary fine-needle aspirate). Overall clinical stage was based on investigator assessment and local staging practices prior to NAT. Hormone receptor (ER ≥1%; progesterone receptor, PR ≥1%) and HER2 status (IHC 2+/FISH+ or IHC 3+) were assessed clinically on diagnostic biopsies using ASCO/CAP guidelines. Participant data, including the age at diagnosis, sex, systemic therapy (neoadjuvant/adjuvant chemotherapy, targeted therapy, endocrine therapy), surgeries, pathology [e.g., residual cancer burden (RCB) score], and clinical recurrence, were extracted through manual review of electronic health records. The RCB score was calculated as previously described (26, 27). Participant age, sex, and weight, as biological variables, were not considered for formal statistical analyses.

Statistical analysis

Overall survival using data from the Genomics England 100,000 Genomes Project were evaluated in all patients with breast cancer and those with HER2$^-$ breast cancer according to the definitions provided in the Genomics England release v18. The impact of SV burden above and below the median was assessed, adjusting for stage subtype (where available) using a Cox proportional hazard model.

For the clinical EBC cohort, standard clinical follow-up was delivered by treating physicians, with the date of clinical recurrence defined as the date of imaging-confirmed metastatic disease (outside the breast or local/regional lymph nodes) or local recurrence (within the ipsilateral breast and any remaining local/regional lymph nodes). Recurrence outcomes were last updated on September 30, 2024, and patients without recurrence were censored at their last follow-up. The predefined analyses in the EBC cohort included the evaluation of assay sensitivity and ctDNA levels at baseline, during, and after treatment, and examination of the relationships between ctDNA detection and clinical outcomes across the cohort, by receptor subtype and stage. The primary analysis assessed the association between ctDNA detection and distant recurrence-free interval (DRFI), defined as the time from pathologic diagnosis of invasive breast cancer to metastatic recurrence. Secondary endpoints considered invasive disease-free interval (iDFI), which included local recurrence. One participant who had an unrelated death prior to surgery contributed to baseline sensitivity only. No power calculations or definitions of error were performed.

"Baseline" samples were collected before starting NAT. On-treatment collection timepoints were designated as follows: "Pre-cycle 2 (C2)" referred to specimens collected before the second NAT cycle, whereas "mid-NAT" referred to specimens collected before the midpoint of their neoadjuvant regimen [e.g., docetaxel and cyclophosphamide (TC), precycle 3; docetaxel and carboplatin with anti-HER2 therapy (e.g., TCHP), precycle 4; or anthracycline and taxane with or without pembrolizumab or anti-HER2 therapy, precycle 5]. "Preoperative" referred to the last cycle or after finishing NAT but before surgery, and "postoperative" (Post-Op) within 60 days after completing definitive surgery. "Follow-up" timepoints were collected ≥60 days after surgery with a planned frequency of every 6 to 12 months corresponding with routine clinical follow-up. One participant treated with neoadjuvant endocrine therapy contributed to baseline assay sensitivity and recurrence lead-time analyses but was excluded from on-treatment ctDNA dynamics analyses. The Kaplan–Meier method was used to estimate DRFI and iDFI rates. Lead time was calculated as the interval (in days) between the first Post-Op or adjuvant ctDNA detection and clinical recurrence. Statistical analyses were conducted using GraphPad Prism (version 10; RRID: SCR_002798) and R (version 4.1.0). All tests were two-sided with $P < 0.05$ considered significant unless otherwise stated, and no corrections were applied for multiple significance testing.

Data availability

This research was made possible through access to data in the National Genomic Research Library, which is managed by Genomics England Limited (a wholly owned company of the UK Department of Health and Social Care). The National Genomic Research Library holds data provided by patients and collected by the NHS as part of their care and data collected as part of their participation in research. The National Genomic Research Library is funded by the National Institute for Health Research and NHS England. The Wellcome Trust, Cancer Research UK, and the Medical Research Council have also funded research infrastructure. Research on the de-identified Genomics England patient data used in this publication can be carried out in the Genomics England Research Environment subject to a collaborative agreement that adheres to patient-led governance. All interested readers will be able to access the data in the same manner that the authors accessed the data. For more information about accessing the data, interested readers may contact research-network@genomicsengland.co.uk or access the relevant information on the Genomics England website: https://www.genomicsengland.co.uk/research. Summarized clinical data, selected SVs for assay generation, and ctDNA results are provided in the Supplementary Tables. Additional de-identified participant data are available for academic purposes on request from the corresponding authors: Drs. David W. Cescon (dave.cescon@uhn.ca) or Mitchell Elliott (Mitchell.elliott@uhn.ca). The WGS data of 100/100 patients have been deposited at the European Genome-phenome Archive (EGA), which is hosted by the European Bioinformatics Institute and the Centre for Genomic Regulation, under accession number EGAS50000000799. Controlled access is required to ensure that data use is not for profit or commercial purposes. Data are available by submitting a data access request via the EGA portal (see https://ega-archive.org/access/request-data/how-to-request-data/ for detailed guidance). The remaining data are available within the article and Supplementary Materials provided.

Results

SVs are widespread across cancer types

To evaluate the suitability of SVs as a target candidate for ctDNA monitoring, we obtained somatic SV calls from 16,247 patients with cancer enrolled in the 100,000 Genomes Project (20). Existing SV calls were filtered to remove genome mapping artifacts. In total, 1,292,794 eligible SVs were detected across the 16,247 tumors evaluated (Supplementary Table S1), with SV burden varying significantly among cancer types (**Fig. 1A**). Notably, high SV burdens were observed in upper gastrointestinal tumors (median SV burden = 85, $n = 265$), sarcoma (median SV burden = 80, $n = 1819$), ovarian cancer (median SV burden = 78, $n = 634$), and breast cancer (median SV burden = 60, $n = 3{,}009$). Within breast cancer cases, a significantly higher SV burden was observed in HER2$^+$ and TNBC than in ER$^+$ (**Fig. 1B**; $P < 0.0001$). To assess amplified SVs, which can contribute multiple targetable DNA fragments per cancer cell, those within the top 10% of amplified SVs in each sample were included in the analysis (Supplementary Fig. S1B). The copy number of the top 10% of amplified SVs was

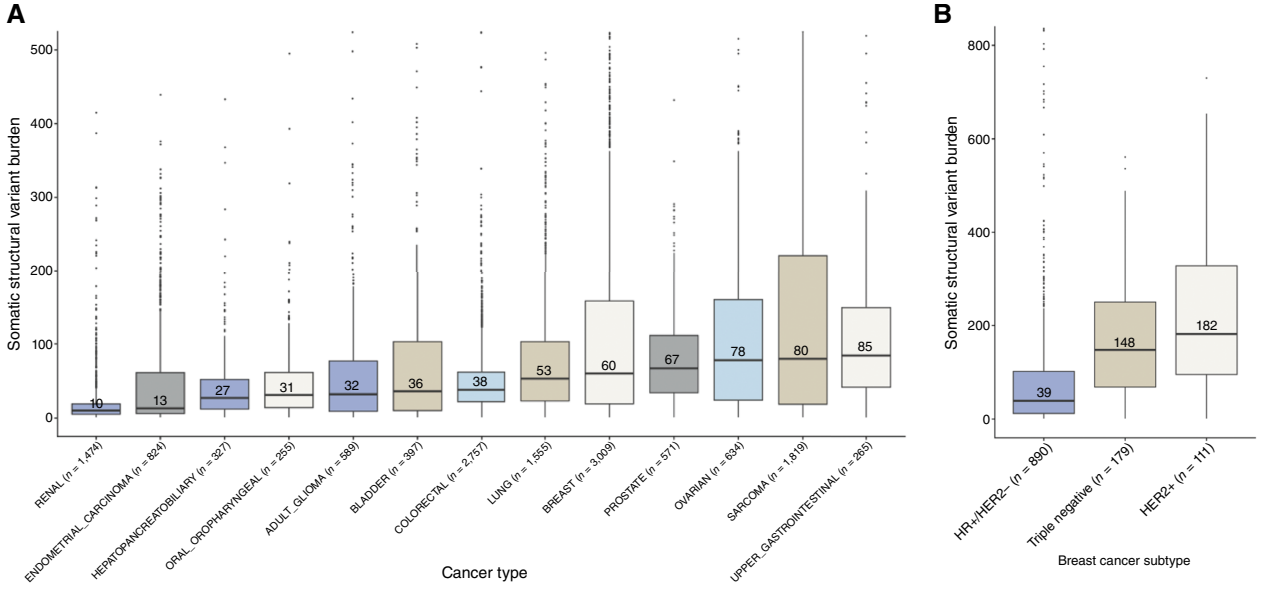

Figure 1.

SV burden and associated copy number state across solid tumors and within breast cancer subtypes. **A,** Somatic SV burden across adult solid tumor types, as obtained from the 100,000 Genomes Project. The median SV burden for each cancer type is annotated on the corresponding boxplot. Note that cancer types with <100 samples are excluded from visualization. **B,** Somatic SV burden within breast cancer clinical receptor subtypes, as derived from the 100,000 Genomes Project.

significantly higher in $HER2^+$ than in $HER2^-$ cancers (Supplementary Fig. S1C; $P < 0.0001$).

The SV number was prognostic in breast cancer (Supplementary Figures): a higher number of SVs (above the median) was associated with significantly worse overall survival [Supplementary Fig. S1D; HR = 1.62; 95% confidence interval (CI), 1.16–2.26; $P = 0.0045$], independent of the stage. When evaluated by receptor subtype, SV burden remained prognostic in $ER^+/HER2^-$ (Supplementary Fig. S1E; $P = 0.0023$) but not in $HER2^+$ (Supplementary Fig. S1F; $P = 0.76$) or in TNBC (Supplementary Fig. S1G; $P = 0.35$), controlled for stage.

Analytical performance of a dPCR-based SV ctDNA assay

The high prevalence of SVs across breast cancer subtypes supported the rationale for developing a tumor-informed dPCR-based ctDNA assay capable of targeting up to 16 SVs per tumor. To evaluate the analytical performance of this assay, validation studies were performed to characterize analytical sensitivity and specificity using the standard workflow (Supplementary Fig. S2A). For analytical sensitivity, a fingerprint was designed for the breast cancer cell line BT474 ($HER2^+$ breast carcinoma). The LoD95 was determined using the probit method for 70 ng of DNA (fragmented to mimic cfDNA) using a BT474 DNA dilution range from a tumor fraction of 0.0005% (5 ppm) to 0.00004% (0.4 ppm; **Fig. 2A**). The estimated LoD95 was 0.00052% (5 ppm, **Fig. 2A**; Supplementary Table S2A) with variants detected as low as four parts in 10 million (0.00004% or 0.4 ppm in 22% of the cases; **Fig. 2B**). Due to the nature of the SV dPCR assay, a single-molecule SV can be clearly identified with multiple positive dPCR partitions (**Fig. 2B**). An LoD95 confirmation study was performed on additional cell line–derived DNA including SK-BR-3 (breast carcinoma; Supplementary Fig. S2B) and FaDu (HTB-43, hypopharyngeal squamous cell carcinoma; Supplementary Fig. S2C). A colorectal cfDNA clinical sample was also evaluated (Supplementary Fig. S2D). An LoD95 of 0.00052% was verified in all cases. dPCR has potential advantages over NGS-based techniques for ctDNA sensitivity due to its capacity for a larger cfDNA input in the reaction, enabling detection of rare ctDNA fragments. To evaluate this, an input of 300 ng was tested, yielding an LoD95 of 0.00011% (1.1 ppm), with detection as low as one in 10 million (Supplementary Table S2B).

To characterize analytical specificity, 24 unique SV-based ctDNA assays (derived from five lung, five breast, five head and neck, five colorectal, and four ovarian cancer cases) were used to analyze 217 cancer-free donor cfDNA samples and 217 cancer-free genomic DNA buffy coat samples. None of the 434 samples yielded a positive result, demonstrating 100% analytical specificity across 5,268 SVs tested (Supplementary Fig. S2E).

EBC clinical cohort identification and SV assay design

Given the excellent laboratory analytical performance, clinical validity was evaluated in a real-world cohort of participants with EBC. Tumor specimens from 111 participants underwent assay preparation. All tumor DNA samples passed sequencing quality control. Tumor-only WGS permitted successful fingerprint generation for 95% (105/111) of participants; six samples failed assay design as fewer than four SVs were detected (Supplementary Fig. S3A; four ER^+, one $HER2^+$, and one TNBC). Four participants, for whom panels were designed, were identified through chart review as having developed metastatic disease early in their treatment course. Additionally, one participant had only adjuvant samples collected, none of which showed detectable ctDNA. These five cases were excluded from the primary analysis (Supplementary Fig. S3A; Supplementary Table S3). One hundred participants had evaluable data and were included in the primary cohort (**Table 1**).

The median WGS coverage for the 100 included participants was 21× (range: 11–46×). A total of 2,614 breakpoints were identified,

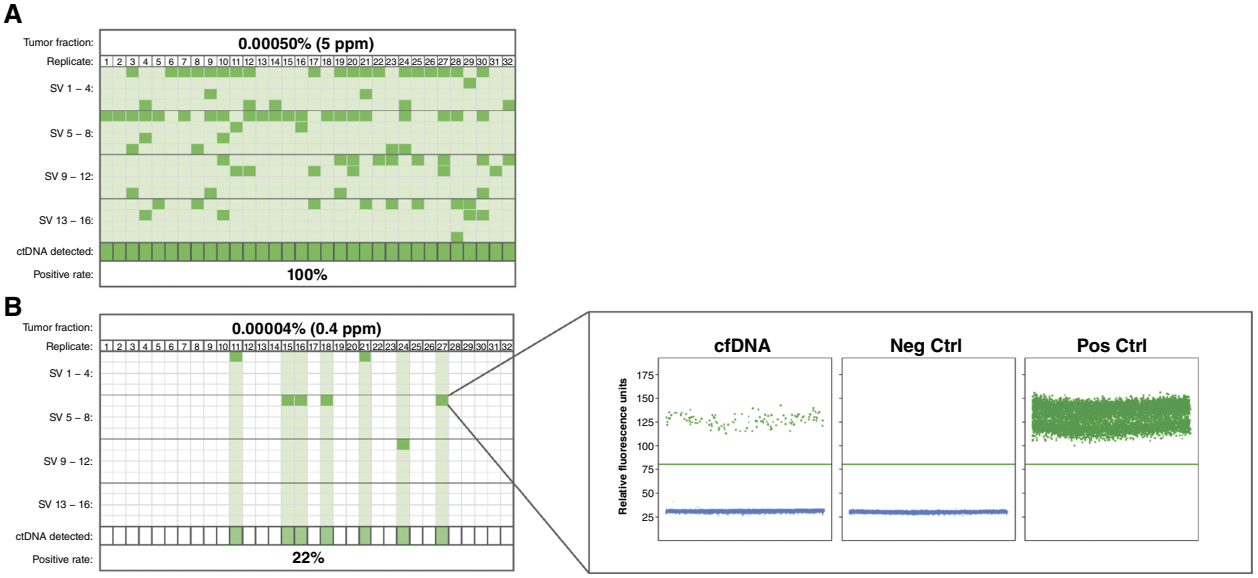

Figure 2.

Contrived cfDNA samples prepared for assessment of assay analytical validity. **A,** dPCR output of the LoD95 study at a dilution of 0.0005% (5 ppm) in BT474 demonstrating 100% ctDNA detection. Illustration of replicate number (1–32) vs. SV number (1–16). Dark green–filled cells indicate positive SV results. **B,** Example dPCR data (1D plots) for one positive SV result, including positive and negative controls, are shown. A representative threshold value is illustrated.

corresponding to 1,307 SV junctions, each of which was unique in this cohort. (**Fig. 3A**). The individual participant SV landscape is described in Supplementary Table S4 and Supplementary Material S1. Validated events involving chromosome 8 (**Fig. 3B**) and chromosome 17 (**Fig. 3C**) in all participants are illustrated as examples. Even though breakpoints on chromosomes 8 and 17 are common events in breast cancer, all SV junctions were unique to individual participants' tumors. Personalized fingerprints were designed for each participant prioritizing unique breakpoint junctions from highly amplified SVs. A median of 14 SVs were selected for fingerprint design (range: 4–16; Supplementary Table S5). These final panels were used to test for ctDNA in extracted cfDNA from each available plasma timepoint.

Participant characteristics

The 100 participants included for the final ctDNA analysis represented all receptor subtypes: TNBC ($n = 28$), ER$^+$ ($n = 33$), and HER2$^+$ ($n = 39$). All participants were females, and the median age was 50 years (range: 24–79 years), with most individuals being premenopausal ($n = 58$). The median follow-up time from baseline blood collection was 3.3 years (range: 0.5–7.7 years). The clinical characteristics of this cohort are representative of a typical NAT population, including a high proportion of node positivity (69%), Nottingham grade 3 (64%), and larger tumor size (35% T3/4). Most participants' NAT included an anthracycline and taxane–based combination ($n = 83$); one individual (in whom ctDNA was detected at baseline) received neoadjuvant endocrine therapy. At the time of data cutoff, 24 clinical recurrences (24.0%) had occurred. Full-cohort characteristics are summarized in **Table 1**.

Clinical characteristics associated with ctDNA detection

Five of the 100 participants did not have baseline plasma samples available for analysis. A median of 4 mL of plasma (range: 0.7–5.2 mL) and the resulting cfDNA (median: 38.8 ng, range: 15.7–314.9 ng) were used for baseline ctDNA assessment. ctDNA was detected in 96% (91/95) of baseline samples, with a median variant allele frequency (VAF) of 0.15% (**Fig. 3D**; range: 0.0011%–38.7%); one timepoint for one participant was detected but was unquantifiable due to dPCR saturation. ctDNA was identified in 260/568 (46.0%) of all analyzed samples (median cfDNA input: 52.6 ng, range: 14.3–1708.8 ng) with a VAF range of 0.00006% to 38.7%, including 97/260 (37.5%) with VAF <0.01% (100 ppm) and 34/260 (13.1%) with VAF <0.001% (10 ppm; **Fig. 3D**; Supplementary Table S6). There was no difference observed between cfDNA input and the presence or absence of detectable ctDNA in this study (Supplementary Fig. S3B; detected: 79.25 ng, not detected: 67.88 ng; $P = 0.13$). The rate of baseline ctDNA detection was similar across clinical receptor subtypes: TNBC 96.0% (23/24), ER$^+$ 94.0% (30/32), and HER2$^+$ 97.4% (38/39). Baseline ctDNA was detected in 66% (2/3) of participants of stage I, 96% (54/56) of stage II, and 97% (35/36) of stage III disease.

A comparison of clinical characteristics and baseline ctDNA positivity was not performed as ctDNA was detected in nearly all participants. There was a trend toward higher VAF in participants with TNBC (**Fig. 3E**; $P = 0.079$). Baseline VAF was also higher in those with node positivity at baseline (**Fig. 3F**; $P = 0.0003$). However, no significant association was observed between baseline VAF and tumor size (**Fig. 3G**; $P = 0.28$), clinical stage at diagnosis (**Fig. 3H**; $P = 0.19$), or Nottingham grade as assessed on the diagnostic biopsy (**Fig. 3I**; $P = 0.18$). The baseline VAF did not differ in those with higher versus lower number of variants included in the fingerprint (Supplementary Fig. S3C, $P = 0.63$; Pearson $r = -0.036$, $P = 0.73$). Notably, baseline VAF was higher in participants who subsequently experienced recurrence (**Fig. 3J**; $P = 0.0032$) but did not differ when evaluated by receptor subtype (Supplementary Fig. S3D–S3F). There was no significant difference in SV fingerprint copy number among those who experienced recurrence and those who did not (**Fig. 3K**; $P = 0.96$).

Table 1. Baseline participant characteristics in the EBC clinical validation cohort.

	Cohort characteristics			
	All patients (*N* = 100)	**TNBC (*n* = 28)**	**ER$^+$/HER2$^-$ (*n* = 33)**	**HER2$^+$ (*n* = 39)**
Follow-up from baseline, median (years; range)	3.3 (0.5–7.7)	2.8 (0.5–7.6)	2.9 (0.5–7.7)	3.8 (0.6–7.6)
Age, median (range)	50 (24–79)	47.5 (29–68)	52 (31–73)	48 (24–79)
Menopausal status, *n* (%)				
Post	42 (42.0%)	9 (32.0%)	15 (45.5%)	18 (46.2%)
Pre	58 (58.0%)	19 (68.0%)	18 (54.5%)	21 (53.8%)
Tumor size, *n* (%)				
T1/T2	65 (65.0%)	20 (71.0%)	17 (51.5%)	28 (71.8%)
T3/T4	35 (35.0%)	8 (29.0%)	16 (48.5%)	11 (28.2%)
Nodal status, *n* (%)				
Node-negative	31 (31.0%)	13 (46.0%)	6 (18.2%)	12 (30.8%)
Node-positive	69 (69.0%)	15 (54.0%)	27 (81.8%)	27 (69.2%)
Stage, *n* (%)				
I	4 (4.0%)	4 (14.3%)	0 (0.0%)	0 (0.0%)
II	59 (59.0%)	18 (64.3%)	13 (39.4%)	28 (71.8%)
III	37 (37.0%)	6 (21.4%)	20 (60.6%)	11 (28.2%)
Nottingham grade (diagnostic biopsy), *n* (%)				
2	26 (26.0%)	2 (7.0%)	15 (45.5%)	9 (23.1%)
3	74 (74.0%)	26 (93.0%)	18 (54.5%)	30 (76.9%)
Histology, *n* (%)				
IDC	93 (93.0%)	26 (93.0%)	30 (90.9%)	37 (94.9%)
Other	7 (7.0%)	2 (7.0%)	3 (9.1%)	2 (5.1%)
NAT, *n* (%)				
Anthracycline + taxane	83 (83.0%)	25 (89.0%)	31 (94.0%)	27 (69.2%)
Endocrine therapy	1 (1.0%)	0 (0.0%)	1 (3.0%)	0 (0.0%)
Taxane	5 (5.0%)	3 (11.0%)	1 (3.0%)	1 (2.6%)
Taxane + platinum	11 (11.0%)	0 (0.0%)	0 (0.0%)	11 (28.2%)
Pathological outcome, *n* (%)				
RCB-0	27 (27.0%)	9 (32.0%)	2 (6.1%)	16 (41.0%)
RCB-1	29 (29.0%)	7 (25.0%)	9 (27.3%)	13 (33.3%)
RCB-2	32 (32.0%)	9 (32.0%)	14 (42.4%)	9 (23.1%)
RCB-3	11 (11.0%)	2 (7.0%)	8 (24.2%)	1 (2.6%)
Missing	1 (1.0%)	1 (4.0%)	0 (0.0%)	0 (0.0%)
Recurrences, *n* (%)				
No	76 (76.0%)	19 (67.9%)	21 (63.6%)	36 (92.3%)
Yes (local)	2 (2.0%)	1 (3.6%)	0 (0.0%)	1 (2.6%)
Yes (distant)	22 (22.0%)	8 (28.5%)	12 (36.4%)	2 (5.1%)

Abbreviation: IDC, invasive ductal carcinoma.

Longitudinal ctDNA assessment

A longitudinal representation of ctDNA timepoints for each participant is shown by disease subtype: ER$^+$ (**Fig. 4A**), TNBC (**Fig. 4B**), and HER2$^+$ (**Fig. 4C**). The preoperative ctDNA detection rate was 23.7% (9/38) and was observed exclusively in participants who had residual disease. Detection of ctDNA at the preoperative timepoint was not associated with a distant recurrence event ($P = 0.17$). Individual ctDNA dynamics are represented in kinetic graphs for all participants (**Fig. 5A** and **B**; Supplementary Fig. S4A–S4C).

Treatment-related changes in ctDNA and clinical outcome

The stage (Supplementary Fig. S5A; $P = 0.021$) and response to neoadjuvant chemotherapy (RCB-status; Supplementary Fig. S5B; $P = 0.0002$) were independently prognostic in this cohort. On-treatment samples during NAT were collected for analysis of ctDNA dynamics and their association with DRFI. Clearance of ctDNA prior to the administration of C2 was associated with improved outcomes in all participants (**Fig. 5C**; HR: undefined; $P = 0.047$). The presence of ctDNA at C2, together with residual disease at the time of surgery, was associated with the highest risk of recurrence (**Fig. 5D**; $P = 0.041$). The prevalence of initial clinical stage at baseline differed in those who experienced [I: 1/12 (8.3%), II: 10/12 (83.4%), and III: 1/12 (8.3%)] or did not experience [I: 1/65 (1.5%), II: 33/65 (50.8%), and III: 31/65 (47.7%)] ctDNA clearance ($P = 0.0143$).

ctDNA clearance was also evaluated at a mid-NAT timepoint in all participants. Clearance was associated with improved outcomes (Supplementary Fig. S5C; HR: 3.13, 95% CI, 1.24–7.90; $P = 0.033$) and enhanced prediction of outcome when stratified by RCB-status (Supplementary Fig. S5D; $P = 0.027$). The prevalence of initial clinical stage at baseline differed in those who experienced [I: 2/32 (6.3%), II: 23/32 (71.9%), and III: 7/32 (21.8%)] or did not experience [I: 0/41 (0%), II: 16/41 (39.0%), and III: 25/41 (61.0%)] ctDNA clearance ($P = 0.0007$).

Postoperative and follow-up ctDNA detection

Of the 24 participants who experienced disease recurrence, 19 had evaluable postoperative and adjuvant follow-up samples. In a landmark analysis, ctDNA detected at the postoperative timepoint

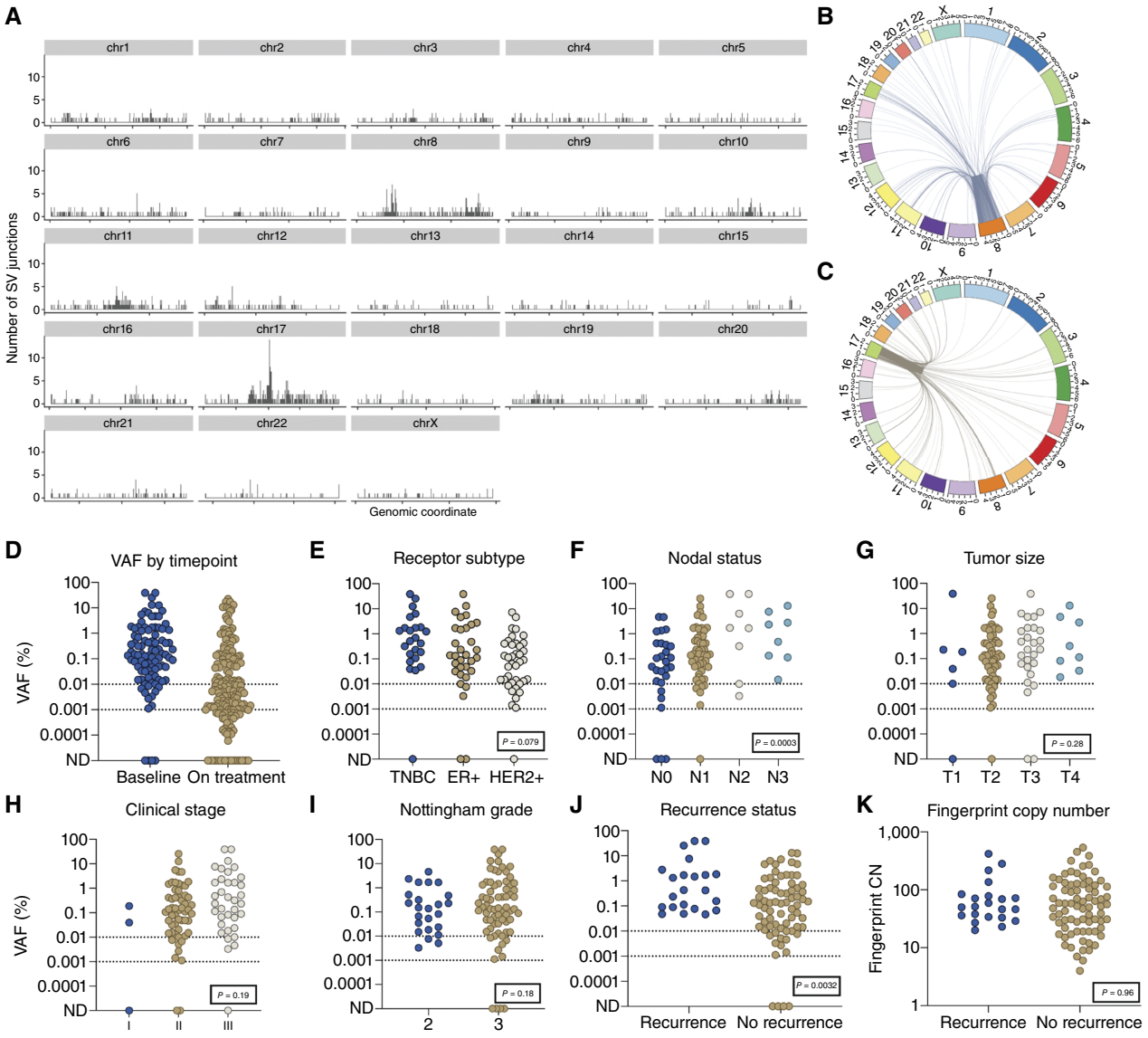

Figure 3.

Tumor-specific SV-based assay analysis. **A,** Histogram of the genomic coordinates of validated SV breakpoints genome wide in the EBC cohort ($n = 100$) using 100,000 bp bins. **B** and **C,** Example Circos plots showing the breakpoint coordinates of all validated SV junctions in which one of the breakpoints falls on chromosome 8 or chromosome 17. **D,** Representation of VAF at baseline ($n = 95$) and in all samples ($n = 568$) tested in the EBC cohort. Comparison of baseline VAF and routine clinical variables: (**E**) receptor subtype, (**F**) clinical nodal status, (**G**) tumor size, (**H**) clinical stage, and (**I**) Nottingham grade as assessed on the diagnostic biopsy and (**J**) in participants with and without recurrence (including local recurrence). **K,** Fingerprint copy number in those with and without recurrence. Chr, chromosome; CN, copy number; N, nodal status; ND, not detected; T, tumor size.

(5/34; 14.7%) was uncommon but significantly associated with recurrence (Supplementary Fig. S5E; HR: 7.37, 95% CI, 1.09–49.79; $P = 0.0002$); all participants with ctDNA detected at this timepoint experienced disease recurrence.

All participants with ctDNA detected postoperatively or during follow-up that were also detectable at baseline experienced subsequent clinical recurrence. In a binary analysis, any ctDNA detection after surgery or during follow-up was strongly associated with DRFI (**Fig. 5E**; Supplementary Fig. S5F; HR: undefined; $P < 0.0001$) and iDFI (**Fig. 5F**; HR: 43.60, 95% CI, 12.2–156.2; $P < 0.0001$).

One participant, who did not have ctDNA detected at baseline, had a positive test at their last follow-up (VAF = 0.004%) but no clinical recurrence at the time of data cutoff. Because the follow-up time from this positive sample to data cutoff (342 days) is shorter than the median lead time observed (see below), a definitive interpretation cannot be made.

Lead time assessment to clinical recurrence

The 24 recurrences observed included two local and 22 distant events. Of these 24, 19 had postoperative and/or follow-up samples enabling the evaluation of ctDNA lead time. All distant

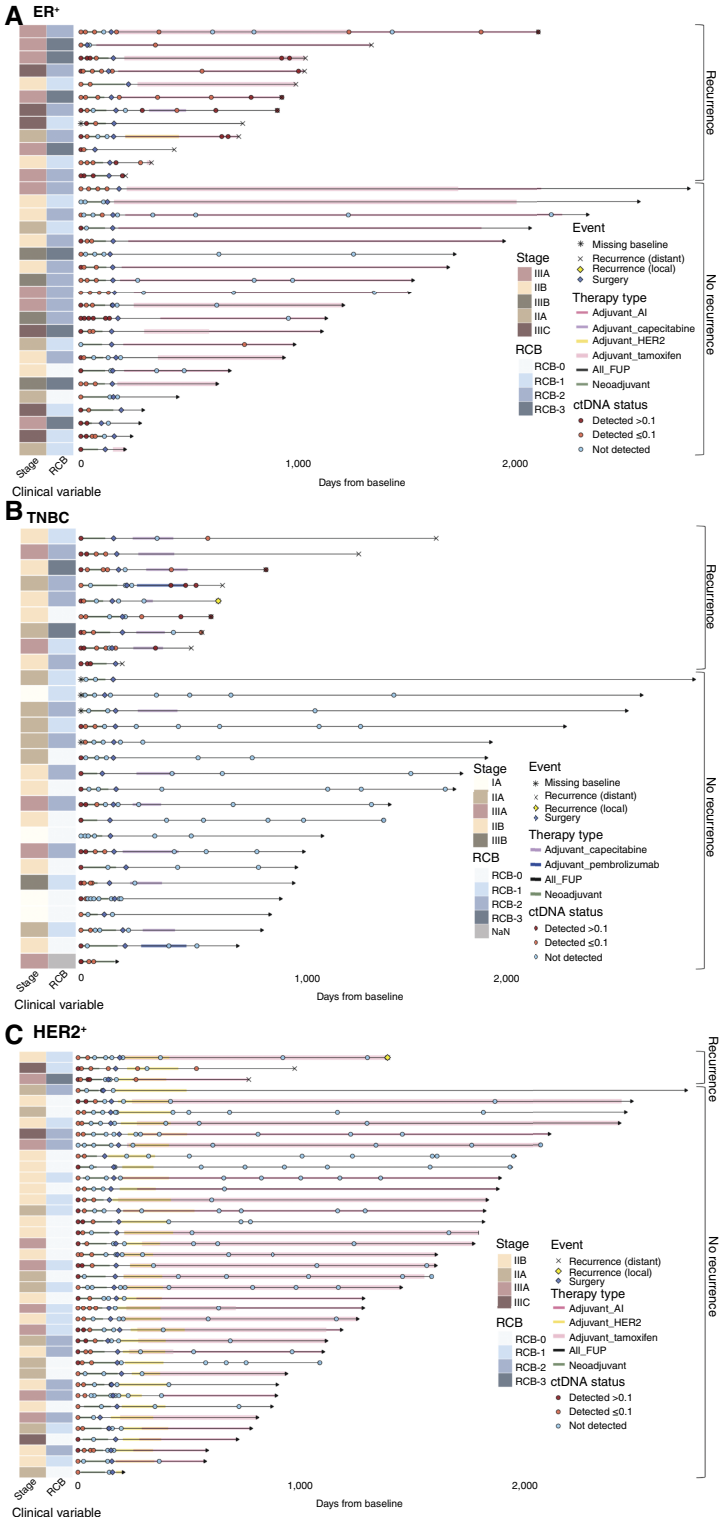

Figure 4.

Swimmer plots and clinical events. Participants' clinical timeline and timeline of plasma collection for ctDNA analysis for (**A**) ER$^+$, (**B**) TNBC, and (**C**) HER2$^+$ EBC. Plots are broken down by participants with and without clinical recurrence and individual stage (I, II, and III) as well as RCB is defined. FUP, follow-up.

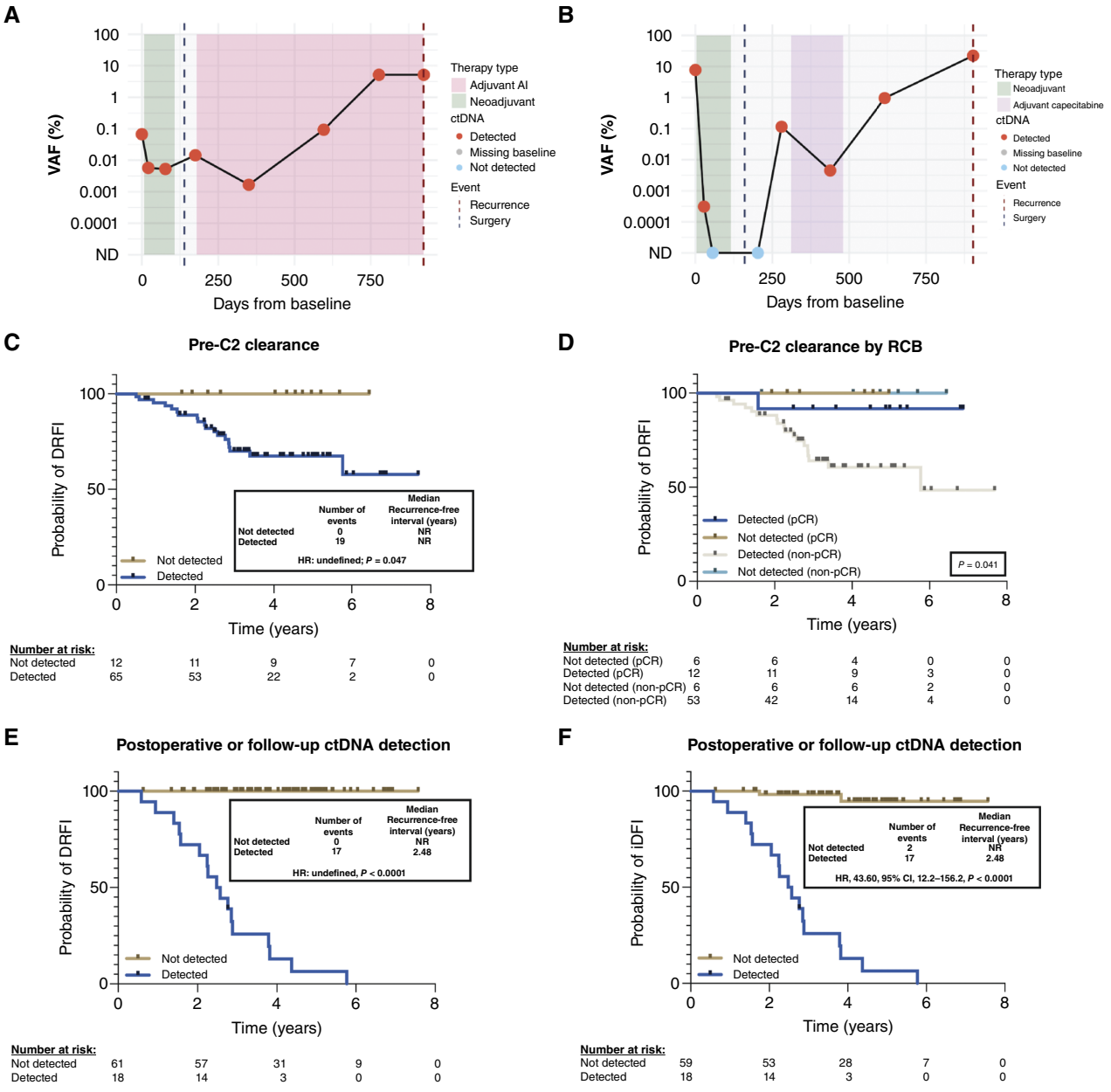

Figure 5.

On-treatment ctDNA clearance, and postoperative ctDNA detection and association with clinical outcomes. **A** and **B,** Representative longitudinal plots of ctDNA detection with reference to clinical timelines. **C,** Association between pre-C2 ctDNA clearance and association with DRFI. **D,** Association between pre-C2 ctDNA clearance and association with DRFI, stratified by RCB. **E,** Association between postoperative or follow-up ctDNA detection and association with DRFI. **F,** Association between postoperative or follow-up ctDNA detection and association with iDFI.

recurrences in evaluable patients were preceded by ctDNA detection, with a median lead time of 417 days (range: 4–1,931 days; Supplementary Table S7). Two participants had incidentally discovered asymptomatic distant recurrences on scheduled clinical imaging (Supplementary Table S7). ctDNA was not detected in the follow-up for the two participants who experienced local recurrences: one (an incidental 2-mm tumor discovered after a plastic surgery procedure) may have represented a new primary ($ER^+/HER2^+$ on original specimen, $ER^-/HER2^+$ on local recurrence), whereas the second occurred 328 days after the last available plasma collection.

Discussion

This study describes the analytical and clinical validation of a ctDNA method designed to enable ultrasensitive ctDNA monitoring using a tumor-informed SV-based dPCR assay. A key contribution of this study is the characterization of SVs as potential biomarkers

for ctDNA detection, distinguishing this approach from most tumor-informed ctDNA assays that target SNVs. SVs are well-established features of cancer genomes, reflecting the underlying biology and genomic instability of tumors. A compelling aspect of SVs for this application is their patient- and tumor-specific breakpoints, with more than 1.2 million identified across more than 10,000 cancer genomes analyzed in the 100,000 Genomes Project (20). The widespread occurrence of SVs across several tumor histologies, as evident in the analysis of the Genomics England data set, suggests that this approach could be broadly applicable for ctDNA detection and monitoring in other cancer types. Our study reinforces prior findings that SV burden varies significantly among breast cancer subtypes. Specifically, $HER2^+$ and TNBC tumors exhibit a higher SV burden than ER^+ tumors. Copy number profiles are also notably higher in $HER2^+$ cancers than in $HER2^-$ cancers. Further work is required to evaluate the relationship between the type, distribution, and clinical outcomes associated with SVs in breast cancer.

In this real-world cohort of participants with EBC, there was no overlap in SVs between individuals among the more than 2,500 identified breakpoints. This finding is particularly notable in cancers (like $HER2^+$ breast cancer) where SVs are commonly associated with driver alterations but arise from DNA breakpoints unique to each tumor. The unique nature of these breakpoints minimizes false positives that can arise from CHIP or NGS errors, which can occur with SNV-based detection methods. The sensitivity of the dPCR approach targeting these SVs is enhanced by the selection of amplified variants in the fingerprint design; as such variants have additional copies per cell and are therefore more likely to be present in a given plasma quantity when present in circulation. Altogether, this assay exhibits a low LoD95 and detects ctDNA in many samples with even lower ctDNA abundance.

The analytical validation of this assay using simulated ctDNA resulted in an LoD95 of 0.00052% (5 ppm) with industry standard inputs (70 ng) and up to 0.00011% (1.1 ppm) with an input of 300 ng. These results are consistent with ultrasensitive detection (LoD95 <0.001%/10 ppm), comparable with other second-generation MRD assays (11, 28). Although *in silico* analysis of the more than 10,000 WGS Genomics England cases supported a high likelihood of assay specificity, this was importantly demonstrated in a specificity study with 24 independent tumor fingerprints tested against 217 healthy donor–derived plasma samples. Although this assay uses shallow-depth tumor WGS for fingerprint design, this level of analytical sensitivity and specificity was achieved with the use of multiplex dPCR, which is widely available and efficiently deployable for repeated measurements.

Clinical validation in a real-world cohort demonstrated both the feasibility and performance of this assay. This cohort included participants with higher-risk features, comorbidities, or adherence issues commonly excluded from clinical trial populations using contemporary regimens, which may have contributed to the lower rate of pCR and the higher recurrence rate observed. The WGS pipeline for assay generation seemed unaffected by FFPE tissue artifacts, which are often a concern in routinely processed specimens. The assay design failure rate was notably low (~5%), with all failures resulting from fewer than the required four validated SVs needed to create a fingerprint. These findings show that tumor-informed dPCR-based ctDNA assays are feasible, using routine diagnostic biopsy tissue in standard clinical settings. cfDNA input can impact ctDNA assay sensitivity, where mutant genome equivalents present for analysis may be limiting at low tumor fractions or input ranges. Although no differences were observed in cfDNA input between samples with and without ctDNA detection, this remains an important consideration for the clinical application of this assay. Given the standard input range of 5 to 1,500 ng, larger plasma volumes could theoretically enhance sensitivity.

Although on-treatment ctDNA detection may differ on the basis of the timings of collections and therapies administered, the detection of ctDNA at baseline in similar patient cohorts may facilitate some cross-study examination of assay sensitivity. ctDNA was detected in 96% of participants at baseline, without notable differences in detection rates across breast cancer receptor subtypes, differing from trends observed in some previous studies. The approach used here seems to translate to high clinical sensitivity, particularly in ER^+ EBC that has historically been associated with a lower rate of baseline ctDNA detection (8, 9, 29). This observation in ER^+ disease is notable given that nearly half (45.5%) of the participants had grade 2 tumors and study recruitment was not restricted to those with high genomic risk, both factors potentially associated with lower ctDNA shedding (9). The high sensitivity of this SV-based assay underscores its potential for disease monitoring in EBC, potentially increasing the number of evaluable participants in future prospective clinical trials.

The dynamics of ctDNA clearance during treatment was prognostic, a finding potentially enabled by the assay's high sensitivity. Specifically, failure to clear ctDNA prior to the second cycle (C2) and mid-NAT was associated with a higher risk of distant metastatic recurrence, highlighting that persistent ctDNA detection during curative-intent treatment is a poor prognostic marker. These data are consistent with previous reports that ctDNA persistence is associated with a higher likelihood of disease recurrence in this setting using assays with lower analytical sensitivity (8, 30). On-treatment ctDNA monitoring in combination with the RCB-score also significantly enhanced prognostication, suggesting that longitudinal monitoring could complement standard-of-care practices. These results support the development of strategies to individualize therapy or prognostic estimates using molecular response.

In a landmark analysis of available postoperative timepoints, the detection of ctDNA was associated with a 100% risk of subsequent recurrence, despite the delivery of standard adjuvant therapy. While some participants had temporary clearance of ctDNA in the adjuvant setting after initiation of standard-of-care therapy, subsequent recurrence was detected prior to clinical relapse. Although adjuvant therapy is an important component of curative strategies for patients treated with NAT, the relative risk reduction of current approaches is modest, and so the absence of evident "cures" among Post-Op $ctDNA^+$ individuals in this cohort is perhaps unsurprising. The persistence or recurrence of ctDNA during or after adjuvant therapy would, however, permit the evaluation of switching or escalation strategies in such individuals at the highest risk. Repeated sampling identified additional impending recurrences in participants whose initial postoperative sample was negative (including a patient with TNBC treated with immunotherapy), illustrating the role of surveillance in enhancing clinical sensitivity.

Despite the relatively sparse follow-up sampling, the median lead time for ctDNA detection before distant recurrence was 417 days (and as long as 1,931 days). Irrespective of the timing, the detection of ctDNA after completion of definitive surgery was followed by distant recurrence in all but one case where follow-up may be considered immature. This observation is consistent with the notion that ctDNA, when measured with high sensitivity and high specificity, is an early marker for impending relapse. Although the detection of ctDNA in the absence of clinically evident metastatic disease does not necessarily imply that intervention will alter patient outcomes, the observed lead time and specificity are key attributes that could support use in interception trials. Additionally, the ability to detect and quantify low-level

results supports both ctDNA identification and surrogate response assessment. Prospective studies are required to evaluate such strategies and characterize the proportion of true MRD (as opposed to clinically occult metastatic disease) detected when accompanied by concurrent or reflex radiographic imaging. The inability to detect two local recurrences in this study may be related to technical factors, such as the frequency of sample collection, or to the development of new primary breast cancers (which would not be expected to share SVs), and biological factors like the disruption of draining lymph nodes.

Our study has several limitations. The analysis of the breakpoints in the 100,000 Genomes Project data was facilitated by the sequencing depth and use of frozen (non-FFPE) samples, eliminating artifacts from library preparation or sample preservation and enabling an idealized evaluation. Although the higher burden of SVs observed in other cancer types within the 100,000 Genomes Project cohort supports the validity of this approach in other tumor types, clinical studies are required to evaluate this hypothesis. In addition, as a tumor-informed assay, prospective testing requiring tumor WGS would take at least 2 to 3 weeks to the first result; clinical application should consider this limitation. By contrast, repeat testing with dPCR could be performed much more rapidly (hours to days). Furthermore, the relatively short follow-up period (especially for ER^+ disease) and timing of collections (which took place during the COVID-19 pandemic, when hospital visits were intentionally minimized) are limitations, though provide an assessment of this approach under "real-world" conditions. As a result, not all participants who experienced a recurrence had postsurgical specimens available for analysis and could not be considered in the surveillance and lead time analyses. In addition, real world, prospective analysis of early timepoints (baseline and pre-C2) may be challenging for immediate decision making, given the time required for personalized assay development. This cohort enrolled participants prior to the full adoption of now standard-of-care treatments, including chemo-immunotherapy and adjuvant cyclin-dependent kinase 4/6 inhibitors, which could alter adjuvant treatment kinetics and associations. In addition, the relatively small size and number of events in this cohort restricts the ability to control for confounding clinical variables and their association with on-treatment kinetics. Larger studies are required to definitively evaluate these associations. Finally, this is a retrospective analysis performed on a single-center cohort, which could contain bias related to study recruitment, standard-of-care practice for NAT selection, and follow-up frequency. Future studies would benefit from prospective multicenter cohorts.

Conclusion

In conclusion, these findings demonstrate the feasibility of a novel SV-based approach for sensitive and specific ctDNA detection in solid tumors and establish the validity for this method as a tool for ultrasensitive ctDNA detection, treatment monitoring, and prognostication in EBC. The ability to detect ctDNA in nearly all participants at baseline suggests clinical advantages over first generation ctDNA assays, especially for ER^+ disease. The clinical importance of this sensitivity is further supported by the long lead times observed before metastatic recurrence in the adjuvant and follow-up settings. These data should motivate future prospective studies to evaluate ctDNA-guided strategies and study the clinical use of ctDNA detection.

Authors' Disclosures

K. Howarth reports other support from SAGA Dx during the conduct of the study as well as other support from NeoGenomics outside the submitted work. E. Amir reports personal fees from Pfizer and Novartis, as well as other support from Novartis outside the submitted work. M.B. Nadler reports personal fees from Novartis and Exact Sciences outside the submitted work. S. Bratman reports personal fees from Adela and EMD Serono and grants from AstraZeneca outside the submitted work as well as a patent for ctDNA mutation analysis issued, licensed, and with royalties paid from Roche and a patent for ctDNA methylation analysis issued, licensed, and with royalties paid from Adela. E.C. de Bruin reports being an AstraZeneca employee and holding AstraZeneca shares. C. Rushton reports personal fees from SAGA Dx during the conduct of the study. Y. Chen reports other support from SAGA Dx during the conduct of the study; in addition, Y. Chen has a patent for 63/650,048 pending to SAGA Dx, a patent for 63/650,061 pending to SAGA Dx, a patent for 63/497,872 pending to SAGA Dx, a patent for 63/402,511 pending to SAGA Dx, and a patent for 63/402,512 pending to SAGA Dx. S. Gladchuk reports other support from SAGA Dx during the conduct of the study as well as other support from SAGA Dx outside the submitted work; in addition, S. Gladchuk reports a patent for 63/402,512 pending to SAGA Dx. A.M. George reports other support from SAGA Dx during the conduct of the study; in addition, A.M. George has a patent for 63/497,872 pending, a patent for 18/240,416 pending, a patent for 63/650,048 pending, a patent for 63/650,052 pending, a patent for 63/650,061 pending, a patent for 63/402,511 pending, a patent for 63/402,512 pending, a patent for 63/348,855 pending, and a patent for 63/348,857 pending. S. Birkeälv reports other support from SAGA Dx during the conduct of the study as well as a patent for 18/240435 pending to SAGA Dx. M. Alcaide reports other support from SAGA Dx during the conduct of the study, as well as a patent for 63/497872 pending to SAGA Dx, 18/240416 pending to SAGA Dx, 63/650048 pending to SAGA Dx, 63/650052 pending to SAGA Dx, and 63/650061 pending to SAGA Dx. L. Oton reports other support from SAGA Dx during the conduct of the study and outside the submitted work, as well as a patent for 63/497872 pending to SAGA Dx, 63/650048 pending to SAGA Dx, and 63/650061 pending to SAGA Dx. G. Putcha reports personal fees from SAGA Dx during the conduct of the study and from Natera and Optum Genomics outside the submitted work. S. Woodhouse reports other support from Saga Dx during the conduct of the study, as well as patents 63/650048, 63/650052, and 63/650061 pending to SAGA Dx. P.L. Bedard reports grants from AstraZeneca, Bicara Therapeutics, Bayer, Boehringer Ingelheim, Merck, Novartis, Roche Genentech, Lego-Chem Biosciences, Medicenna, Zymeworks, Eli Lilly, Gilead, Takeda, GlaxoSmithKline, Bristol Myers Squibb, Amgen, and Pfizer outside the submitted work, as well as being Uncompensated Advisory for Janssen, Zymeworks, Repare Therapeutics, Lilly, Seagen, and Roche Genentech. L.L. Siu reports personal fees from Merck, Pfizer, AstraZeneca, Roche, GlaxoSmithKline, Voronoi, Arvinas, Navire, Relay, Daiichi Sankyo, Amgen, Marengo, Medicenna, Tubulis, LTZ Therapeutics, Pangea, and Break Through Cancer; grants from Novartis, Bristol Myers Squibb, Pfizer, Boehringer Ingelheim, GlaxoSmithKline, Roche Genentech, AstraZeneca, Merck, Celgene, Astellas, Bayer, AbbVie, Amgen, Symphogen, Mirati, BioNTech, 23andMe, and EMD Serono; and personal fees and other support from Agios Pharmaceuticals and Treadwell Therapeutics outside the submitted work. D.W. Cescon reports financial support from AstraZeneca and other support from SAGA Dx during the conduct of the study as well as research support from Grail, Guardant Health, Inivata/NeoGenomics, Knight, and ProteinQure; personal fees and research support from AstraZeneca, GenomeRx, Gilead, GlaxoSmithKline, Merck, Pfizer, and Roche; and personal fees from Daiichi Sankyo, Lilly, Novartis, and SAGA Dx outside the submitted work. No disclosures were reported by the other authors.

Authors' Contributions

M.J. Elliott: Conceptualization, data curation, formal analysis, supervision, funding acquisition, investigation, visualization, methodology, writing–original draft, project administration, writing–review and editing. **K. Howarth:** Resources, methodology, writing–review and editing. **S. Main:** Visualization. **J. Fuentes Antrás:** Investigation. **P. Echelard:** Data curation, investigation, methodology. **A. Dou:** Data curation. **E. Amir:** Writing–review and editing. **M.B. Nadler:** Writing–review and editing. **E. Shah:** Project administration. **C. Yu:** Project administration. **S. Bratman:** Supervision. **T. Bird:** Data curation. **J. Roh:** Data curation. **E.C. de Bruin:** Conceptualization. **C. Rushton:** Data curation, formal analysis. **Y. Chen:** Data curation. **S. Gladchuk:** Data curation. **A.M. George:** Data curation. **S. Birkeälv:** Project administration. **M. Alcaide:** Software, methodology. **L. Oton:** Software, methodology. **G. Putcha:** Project administration, writing–review and editing. **S. Woodhouse:** Data curation, formal analysis. **P.L. Bedard:** Conceptualization, supervision, funding acquisition. **L.L. Siu:** Conceptualization, funding acquisition. **H.K. Berman:** Data curation. **D.W. Cescon:** Conceptualization, resources, supervision, funding acquisition, investigation, methodology, writing–original draft, writing–review and editing.

Acknowledgments

The authors would like to thank the participants and their families for their participation in this study. This study was supported by the BMO Financial Group

Chair in Precision Genomics at the Princess Margaret Cancer Centre through the Liquid Biopsy Evaluation and Repository Development at Princess Margaret (LIBERATE) program (NCT03702309); the Canadian Cancer Society Research Institute (to D.W. Cescon); the Princess Margaret Foundation (DH Gales Family Foundation); the Hold 'Em for Life and Canadian Association of Medical Oncology (CAMO) fellowships (to M.J. Elliott); a Canadian Cancer Society Research Training Award—PhD award #708002 (to S. Main); and AstraZeneca.

Note

Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

Received October 24, 2024; revised December 9, 2024; accepted January 7, 2025; published first January 9, 2025.

References

1. Siravegna G, Marsoni S, Siena S, Bardelli A. Integrating liquid biopsies into the management of cancer. Nat Rev Clin Oncol 2017;14:531–48.
2. Donaldson J, Park BH. Circulating tumor DNA: measurement and clinical utility. Annu Rev Med 2018;69:223–34.
3. Cescon DW, Bratman SV, Chan SM, Siu LL. Circulating tumor DNA and liquid biopsy in oncology. Nat Cancer 2020;1:276–90.
4. McDonald BR, Contente-Cuomo T, Sammut SJ, Odenheimer-Bergman A, Ernst B, Perdigones N, et al. Personalized circulating tumor DNA analysis to detect residual disease after neoadjuvant therapy in breast cancer. Sci Transl Med 2019;11:eaax7392.
5. Coombes RC, Page K, Salari R, Hastings RK, Armstrong A, Ahmed S, et al. Personalized detection of circulating tumor DNA antedates breast cancer metastatic recurrence. Clin Cancer Res 2019;25:4255–63.
6. Zhou Q, Gampenrieder SP, Frantal S, Rinnerthaler G, Singer CF, Egle D, et al. Persistence of ctDNA in patients with breast cancer during neoadjuvant treatment is a significant predictor of poor tumor response. Clin Cancer Res 2022;28:697–707.
7. Riva F, Bidard FC, Houy A, Saliou A, Madic J, Rampanou A, et al. Patient-specific circulating tumor DNA detection during neoadjuvant chemotherapy in triple-negative breast cancer. Clin Chem 2017;63:691–9.
8. Magbanua MJM, Swigart LB, Wu HT, Hirst GL, Yau C, Wolf DM, et al. Circulating tumor DNA in neoadjuvant-treated breast cancer reflects response and survival. Ann Oncol 2021;32:229–39.
9. Magbanua MJM, Brown Swigart L, Ahmed Z, Sayaman RW, Renner D, Kalashnikova E, et al. Clinical significance and biology of circulating tumor DNA in high-risk early-stage HER2-negative breast cancer receiving neoadjuvant chemotherapy. Cancer Cell 2023;41:1091–102.e4.
10. Parsons HA, Blewett T, Chu X, Sridhar S, Santos K, Xiong K, et al. Circulating tumor DNA association with residual cancer burden after neoadjuvant chemotherapy in triple-negative breast cancer in TBCRC 030☆. Ann Oncol 2023;34:899–906.
11. Garcia-Murillas I, Cutts R, Abbott C, Boyle SM, Pugh J, Chen R, et al. Ultra-sensitive ctDNA mutation tracking to identify molecular residual disease and predict relapse in patients with early breast cancer. J Clin Oncol 2024;42:1010.
12. Black JRM, Frankell A, Veeriah S, Colopi M, Hill M, Abbott C, et al. LBA55 an ultra-sensitive and specific ctDNA assay provides novel pre-operative disease stratification in early stage lung cancer. Ann Oncol 2023;34:S1294.
13. Leary RJ, Kinde I, Diehl F, Schmidt K, Clouser C, Duncan C, et al. Development of personalized tumor biomarkers using massively parallel sequencing. Sci Transl Med 2010;2:20ra14.
14. Collins RL, Brand H, Karczewski KJ, Zhao X, Alföldi J, Francioli LC, et al. A structural variation reference for medical and population genetics. Nature 2020;581:444–51.
15. Li Y, Roberts ND, Wala JA, Shapira O, Schumacher SE, Kumar K, et al. Patterns of somatic structural variation in human cancer genomes. Nature 2020;578:112–21.
16. Santonja A, Cooper WN, Eldridge MD, Edwards PAW, Morris JA, Edwards AR, et al. Comparison of tumor-informed and tumor-naïve sequencing assays for ctDNA detection in breast cancer. EMBO Mol Med 2023;15:e16505.
17. Olsson E, Winter C, George A, Chen Y, Howlin J, Tang MHE, et al. Serial monitoring of circulating tumor DNA in patients with primary breast cancer for detection of occult metastatic disease. EMBO Mol Med 2015;7:1034–47.
18. McDonald BR, Dennison KL, Schussman AL, McGregor SM, Pockaj BA, Murtaza M. Abstract 2416: personalized minimal residual disease detection using tumor-derived structural variants in cell-free DNA. Cancer Res 2024;84:2416–16.
19. Ma X, Shao Y, Tian L, Flasch DA, Mulder HL, Edmonson MN, et al. Analysis of error profiles in deep next-generation sequencing data. Genome Biol 2019;20:50.
20. Turnbull C, Scott RH, Thomas E, Jones L, Murugaesu N, Pretty FB, et al. The 100 000 Genomes Project: bringing whole genome sequencing to the NHS. BMJ 2018;361:k1687.
21. Caulfield M, Davies J, Dennys M, Elbahy L, Fowler T, Hill S, et al. National genomic research library, figshare. 2020 [cited 23 Apr 2020]. Available from: 10.6084/M9.FIGSHARE.4530893.V7.
22. Chen X, Schulz-Trieglaff O, Shaw R, Barnes B, Schlesinger F, Källberg M, et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. Bioinformatics 2016;32:1220–2.
23. Sugita I, Matsuyama S, Dobashi H, Komura D, Ishikawa S. Viola: a structural variant signature extractor with user-defined classifications. Bioinformatics 2022;38:540–2.
24. Roller E, Ivakhno S, Lee S, Royce T, Tanner S. Canvas: versatile and scalable detection of copy number variants. Bioinformatics 2016;32:2375–7.
25. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 2010;26:841–2.
26. Yau C, Osdoit M, van der Noordaa M, Shad S, Wei J, de Croze D, et al. Residual cancer burden after neoadjuvant chemotherapy and long-term survival outcomes in breast cancer: a multicentre pooled analysis of 5161 patients. Lancet Oncol 2022;23:149–60.
27. Symmans WF, Peintinger F, Hatzis C, Rajan R, Kuerer H, Valero V, et al. Measurement of residual breast cancer burden to predict survival after neoadjuvant chemotherapy. J Clin Oncol 2007;25:4414–22.
28. Kurtz DM, Soo J, Co Ting Keh L, Alig S, Chabon JJ, Sworder BJ, et al. Enhanced detection of minimal residual disease by targeted sequencing of phased variants in circulating tumor DNA. Nat Biotechnol 2021;39:1537–47.
29. Ignatiadis M, Brandao M, Maetens M, Ponde N, Martel S, Drisis S, et al. Neoadjuvant biomarker research study of palbociclib combined with endocrine therapy in estrogen receptor positive/HER2 negative breast cancer: the phase II NeoRHEA trial. Ann Oncol 2018;29(Suppl 8):viii57.
30. Elliott MJ, Antras JF, Echelard P, Dou A, Veitch Z, Bedard P, et al. Abstract PS06–08: longitudinal neoadjuvant and post-operative evaluation of circulating tumor DNA in early breast cancer using a tumor-informed assay: updated analysis of the TRACER cohort. Cancer Res 2024;84:PS06-08.