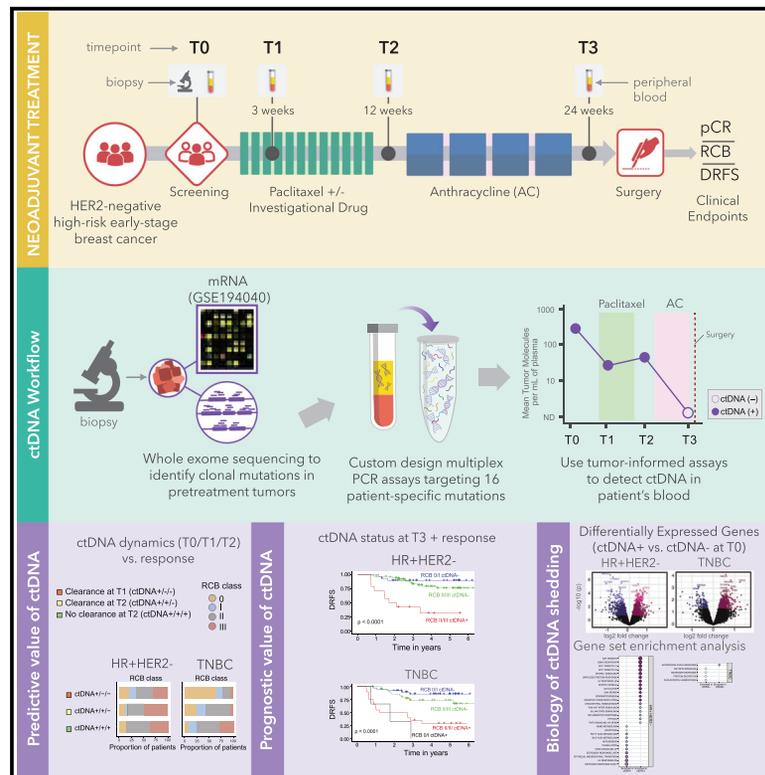


# Cancer Cell

## Clinical significance and biology of circulating tumor DNA in high-risk early-stage HER2-negative breast cancer receiving neoadjuvant chemotherapy

### Graphical abstract



### Authors

Mark Jesus M. Magbanua,  
Lamorna Brown Swigart,  
Ziad Ahmed, ..., Angela M. DeMichele,  
Hope S. Rugo, Laura J. van 't Veer

### Correspondence

mark.magbanua@ucsf.edu

### In brief

Magbanua et al. examine the dynamics of ctDNA in plasma of high-risk early-stage breast cancer patients receiving neoadjuvant chemotherapy. Understanding the predictive and prognostic value of ctDNA and biology of ctDNA shedding in different breast cancer subtypes can inform the use of ctDNA for treatment selection to improve patient outcomes.

### Highlights

- Early clearance of ctDNA in triple-negative patients associates with good response
- ctDNA dynamics during neoadjuvant chemotherapy predict clinical outcomes
- ctDNA negativity associates with improved survival despite having residual cancer
- Expression analysis reveals pathways associated with ctDNA shedding

Article

# Clinical significance and biology of circulating tumor DNA in high-risk early-stage HER2-negative breast cancer receiving neoadjuvant chemotherapy

Mark Jesus M. Magbanua,<sup>1,14,\*</sup> Lamorna Brown Swigart,<sup>1</sup> Ziad Ahmed,<sup>1</sup> Rosalyn W. Sayaman,<sup>1</sup> Derrick Renner,<sup>2</sup> Ekaterina Kalashnikova,<sup>2</sup> Gillian L. Hirst,<sup>1</sup> Christina Yau,<sup>1</sup> Denise M. Wolf,<sup>1</sup> Wen Li,<sup>1</sup> Amy L. Delson,<sup>3</sup> Smita Asare,<sup>4</sup> Minetta C. Liu,<sup>2,5,13</sup> Kathy Albain,<sup>6</sup> A. Jo Chien,<sup>1</sup> Andres Forero-Torres,<sup>7</sup> Claudine Isaacs,<sup>8</sup> Rita Nanda,<sup>9</sup> Debu Tripathy,<sup>10</sup> Angel Rodriguez,<sup>2</sup> Himanshu Sethi,<sup>2</sup> Alexey Aleshin,<sup>2</sup> Matthew Rabinowitz,<sup>2</sup> Jane Perlmutter,<sup>3</sup> W. Fraser Symmans,<sup>10</sup> Douglas Yee,<sup>11</sup> Nola M. Hylton,<sup>1</sup> Laura J. Esserman,<sup>1</sup> Angela M. DeMichele,<sup>12</sup> Hope S. Rugo,<sup>1</sup> and Laura J. van 't Veer<sup>1</sup>

<sup>1</sup>University of California, San Francisco, San Francisco, CA 94143, USA

<sup>2</sup>Natera, Inc., Austin, TX 78753, USA

<sup>3</sup>UCSF Breast Science Advocacy Core, San Francisco, CA 94143, USA

<sup>4</sup>Quantum Leap Healthcare Collaborative, San Francisco, CA 94118, USA

<sup>5</sup>Mayo Clinic, Rochester, MN 55905, USA

<sup>6</sup>Loyola University Chicago, Maywood, IL 60153, USA

<sup>7</sup>University of Alabama at Birmingham, Birmingham, AL 35233, USA

<sup>8</sup>Georgetown University, Washington, DC 20007, USA

<sup>9</sup>University of Chicago, Chicago, IL 60637, USA

<sup>10</sup>University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

<sup>11</sup>University of Minnesota, Minneapolis, MN 55455, USA

<sup>12</sup>University of Pennsylvania, Philadelphia, PA 19104, USA

<sup>13</sup>Present address: Natera, Inc., Austin, TX 78753, USA

<sup>14</sup>Lead contact

\*Correspondence: [mark.magbanua@ucsf.edu](mailto:mark.magbanua@ucsf.edu)

<https://doi.org/10.1016/j.ccell.2023.04.008>

## SUMMARY

Circulating tumor DNA (ctDNA) analysis may improve early-stage breast cancer treatment via non-invasive tumor burden assessment. To investigate subtype-specific differences in the clinical significance and biology of ctDNA shedding, we perform serial personalized ctDNA analysis in hormone receptor (HR)-positive/HER2-negative breast cancer and triple-negative breast cancer (TNBC) patients receiving neoadjuvant chemotherapy (NAC) in the I-SPY2 trial. ctDNA positivity rates before, during, and after NAC are higher in TNBC than in HR-positive/HER2-negative breast cancer patients. Early clearance of ctDNA 3 weeks after treatment initiation predicts a favorable response to NAC in TNBC only. Whereas ctDNA positivity associates with reduced distant recurrence-free survival in both subtypes. Conversely, ctDNA negativity after NAC correlates with improved outcomes, even in patients with extensive residual cancer. Pretreatment tumor mRNA profiling reveals associations between ctDNA shedding and cell cycle and immune-associated signaling. On the basis of these findings, the I-SPY2 trial will prospectively test ctDNA for utility in redirecting therapy to improve response and prognosis.

## INTRODUCTION

Circulating tumor DNA (ctDNA) analysis offers a non-invasive approach to assessing tumor burden during treatment. Monitoring ctDNA can potentially improve the management of early-stage breast cancer by replacing standard-of-care tissue biopsy for DNA-based biomarker evaluation.<sup>1,2</sup> However, breast cancer is a collection of many diseases; thus, studies that seek to demonstrate the clinical impact of ctDNA need to account for the existing heterogeneity across different breast tumor types.

Patients with early-stage, locally advanced breast cancer can benefit from neoadjuvant chemotherapy (NAC; treatment given before surgery), mainly because tumor response during and after treatment can be measured and used to inform subsequent treatment decisions.<sup>3</sup> Patients who achieve a pathologic complete response (pCR), defined as the absence of invasive cancer in the breast and regional lymph nodes, have a significantly decreased risk for metastatic recurrence and death compared with those with residual cancer after NAC.<sup>4-6</sup> For patients who do not achieve pCR, understanding the relationship between residual disease and risk for recurrence presents the opportunity

to evaluate new therapies following surgery in the post-neoadjuvant setting.<sup>7,8</sup>

However, response rates to NAC vary among receptor subtypes.<sup>9,10</sup> For example, results from I-SPY2, a neoadjuvant trial in high-risk early-stage breast cancer evaluating the efficacy of standard chemotherapy with or without investigational drugs, show that among 10 early arms of the trial, the maximum estimated pCR rate is highest in the hormone receptor (HR)-negative/HER2-positive group at 74%.<sup>9</sup> In contrast, triple-negative breast cancer (TNBC) and HR-positive/HER2-negative groups have lower maximum estimated pCR rates of 60% and 30%, respectively.<sup>9</sup> Given that the objective is for each patient to achieve a pCR, there is an unmet need, especially in HER2-negative breast cancer, to develop early predictors of response, in part to de-escalate therapy and limit toxicity for patients who achieve pCR early during treatment or to switch treatment strategy in the case of early non-response.

ctDNA monitoring shows great promise for such a biomarker. However, given the differences in biology, response, and clinical histories between HR-positive/HER2-negative breast cancer and TNBC, we hypothesize that the predictive and prognostic value of ctDNA varies between the two subtypes. Although recent studies in early-stage breast cancer have investigated the clinical significance of serial ctDNA assessment during NAC, several focused on a single breast cancer subtype (TNBC<sup>11–13</sup> or HER2-positive breast cancers),<sup>14</sup> and others had insufficient sample sizes to evaluate the predictive value of ctDNA across breast cancer subtypes<sup>15–18</sup> or did not examine the prognostic significance of ctDNA.<sup>16,19</sup>

In this study, we address these limitations by examining the predictive and prognostic value of ctDNA in a relatively large cohort representing two major breast cancer subtypes: HR-positive/HER2-negative breast cancer and TNBC. We present a follow-up study to our previous work investigating the clinical significance of ctDNA in 84 patients with high-risk early-stage breast cancer receiving NAC.<sup>15</sup> In our expanded cohort of 283 patients (145 with HR-positive/HER2-negative breast cancer and 138 with TNBC) and 1,024 serial blood samples, we evaluate the predictive and prognostic value of ctDNA. In addition to pCR as a response endpoint,<sup>4</sup> we examine the relationship of ctDNA to residual cancer burden (RCB), a measure of the extent of residual disease in the breast and axillary lymph nodes following NAC.<sup>20,21</sup>

We also hypothesize that gene expression analysis could elucidate mechanisms that govern ctDNA shedding,<sup>22</sup> as it is unclear why some breast tumors shed high amounts of ctDNA whereas others have undetectable levels. In addition, tumor size and breast cancer subtype appear to be associated with ctDNA shedding,<sup>15,23–25</sup> but these factors cannot fully explain differences in the concentrations of ctDNA in the blood of breast cancer patients. To this end, we leverage I-SPY2 trial data to examine genes and pathways associated with ctDNA shedding before treatment in these two subtypes. Findings from this study, both subtype-specific clinical associations and related to the biology of ctDNA shedding, may help maximize and fine-tune the use of ctDNA as a biomarker of response and survival in patients with high-risk early-stage breast cancer receiving NAC.

## RESULTS

### Patients and samples

The study involved high-risk early-stage breast cancer patients in the neoadjuvant trial, I-SPY2. A total of 295 HER2-negative patients with available pretreatment biopsies were included in this biomarker study (Figure S1A; Table 1; STAR Methods). Plasma samples were collected at 4 time points: pretreatment (T0); 3 weeks after the initiation of treatment (T1); at 12 weeks, between paclitaxel-based and anthracycline (AC) regimens (T2); and after NAC before surgery (T3) (Figure S1B). ctDNA was detected in plasma using a personalized and tumor-informed test (STAR Methods). Samples that did not pass quality-control measures were excluded from the analysis (Figure S1A; STAR Methods). Of the 295 patients, 283 (96%) were evaluable, with 1,024 plasma samples. ctDNA and clinical data are available in Table S1.

### Clinical correlates of ctDNA

Consistent with previous observations,<sup>15,26</sup> pretreatment (Figure S2A) and concentration (Figure S2B) (expressed as mean tumor molecules [MTM] per milliliter of plasma) were higher in TNBC than in HR-positive/HER2-negative breast cancer patients. ctDNA positivity remained significantly higher in TNBC across all time points (Figure S2A).

Furthermore, we found that higher pretreatment (T0) ctDNA positivity and concentration were associated with clinicopathologic variables that confer aggressiveness (e.g., larger tumor and node positivity in both subtypes and higher grade and MammaPrint score in the HR-positive/HER2-negative group) (Figure S3).

### ctDNA concentration is more strongly associated with pCR and RCB in TNBC compared with HR-positive/HER2-negative breast cancer patients

We examined the relationship of ctDNA concentration (MTM/mL) with response to NAC. In this cohort, the pCR rates were 14.5% (21 of 145) in HR-positive/HER2-negative breast cancer patients and 23.2% (32 of 138) in TNBC patients (Table 1).

We compared the ctDNA concentrations in patients grouped according to response to NAC. In the HR-positive/HER2-negative group, no significant differences in the distribution of ctDNA concentrations across all time points were observed between patients who had pCR vs. no pCR (Figure 1A). However, in TNBC, ctDNA concentrations were significantly higher at all time points in patients who did not achieve pCR compared with those who did (Figure 1B). Similar findings were observed when ctDNA concentrations were compared across RCB classes in both subtypes (Figures 1C and 1D).

### Early ctDNA clearance is a significant predictor of pCR and RCB in TNBC but not in HR-positive/HER2-negative breast cancer patients

We investigated whether ctDNA clearance was predictive of pCR or RCB in patients who were ctDNA positive at T0 (Figures 1E–1J). Complete ctDNA data for the first 3 time points were available from 83 (57.2%) HR-positive/HER2-negative patients and 85 (61.6%) TNBC patients. In both groups, similar proportions of patients cleared early at T1 (ctDNA+/-/-): 27.3% in

**Table 1. Clinicopathologic characteristics of patients grouped according to breast cancer subtype**

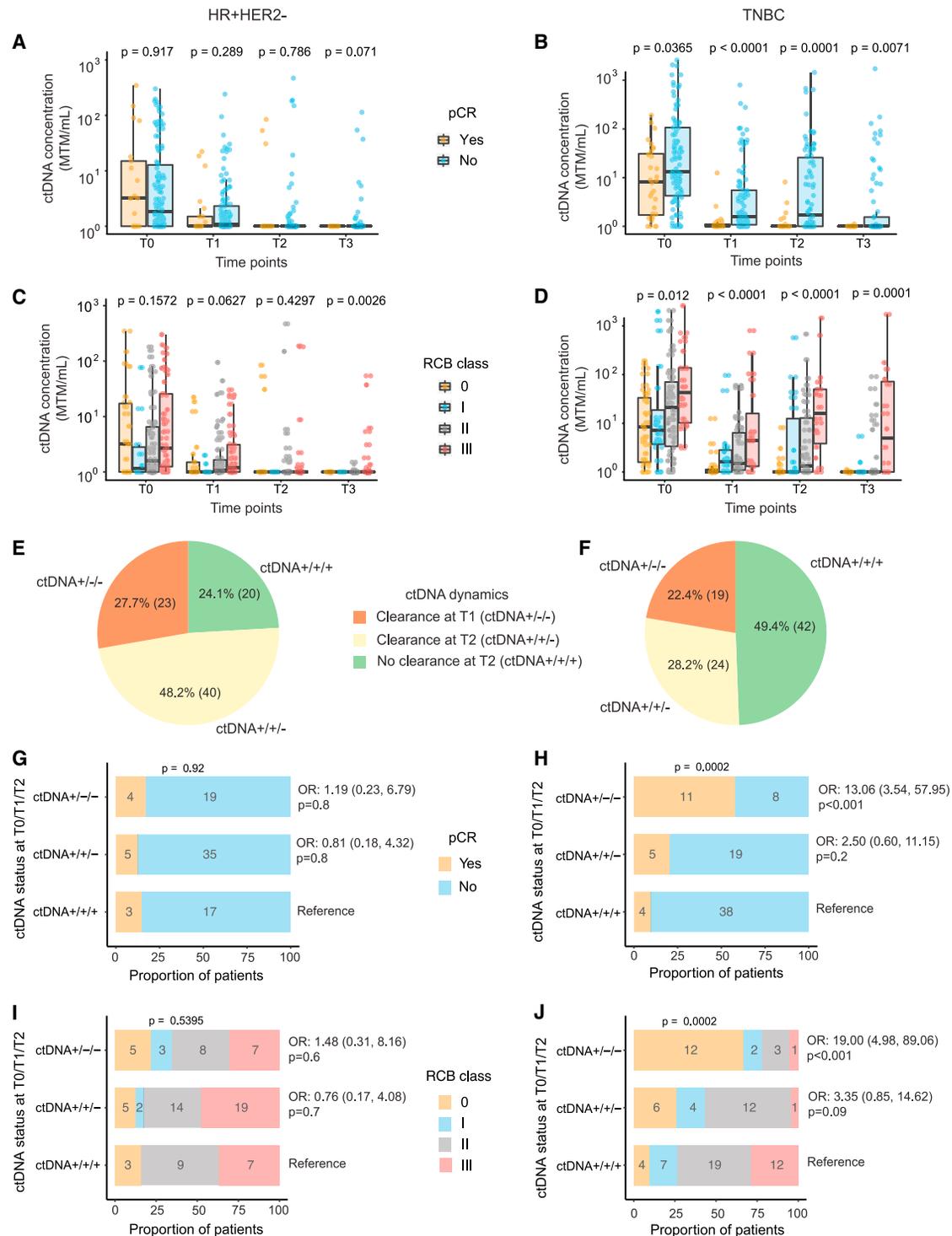
	All patients (n = 283)		HR+HER2– (n = 145 [51.2%])		TNBC (n = 138 [48.8%])		Fisher p value
	n	%	n	%	n	%	
Clinical T stage							0.6777
T1/T2	170	60.1	82	56.6	88	63.8	
T3/T4	75	26.5	39	26.9	36	26.1	
Missing	38	13.4	24	16.6	14	10.1	
Clinical N stage							0.4348
Node negative	118	41.7	55	37.9	63	45.7	
Node positive	118	41.7	62	42.8	56	40.6	
Missing	47	16.6	28	19.3	19	13.8	
Grade							<0.0001
1/2	68	24	56	38.6	12	8.7	
3	156	55.1	62	42.8	94	68.1	
Missing	59	20.8	27	18.6	32	23.2	
MammaPrint score							<0.0001
High 1	120	42.4	104	71.7	16	11.6	
High 2	163	57.6	41	28.3	122	88.4	
Treatment arm							0.4105
Paclitaxel	136	48.1	72	49.7	64	46.4	
Paclitaxel + PD-1 inhibitor	44	15.5	27	18.6	17	12.3	
Paclitaxel + MK-2206	39	13.8	17	11.7	22	15.9	
Paclitaxel + ganitumab	14	4.9	8	5.5	6	4.3	
Irinotecan + talazoparib	13	4.6	4	2.8	9	6.5	
Paclitaxel + ganetespib	12	4.2	4	2.8	8	5.8	
Paclitaxel + AMG 386	8	2.8	3	2.1	5	3.6	
Paclitaxel + PD-1 inhibitor 8-cycle	8	2.8	4	2.8	4	2.9	
Paclitaxel + ABT 888 + carboplatin	5	1.8	4	2.8	1	0.7	
Paclitaxel + neratinib	3	1.1	2	1.4	1	0.7	
SGN-LIV1A	1	0.4	0	0	1	0.7	
Pathologic complete response (pCR) <sup>a</sup>							0.0683
pCR	53	18.7	21	14.5	32	23.2	
no pCR	230	81.3	124	85.5	106	76.8	
Residual cancer burden (RCB)							0.0010
RCB-0	56	19.8	22	15.2	34	24.6	
RCB-I	31	11	10	6.9	21	15.2	
RBC-II	115	40.6	61	42.1	54	39.1	
RBC-III	75	26.5	51	35.2	24	17.4	
Missing	6	2.1	1	0.7	5	3.6	
Age (mean)							0.1440
	49.2		50.1		48.2		

<sup>a</sup>Per I-SPY2 protocol, patients who received study treatment but switched to non-protocol therapy, did not go to surgery, or withdrew consent before surgery were considered to have a non-pCR for efficacy and biomarker analysis (see [STAR Methods](#)).

the HR-positive/HER2-negative group and 22.4% in the TNBC group (Figures 1E and 1F). Among the HR-positive/HER2-negative patients, the largest group, representing about half, consisted of those who cleared their ctDNA late at T2 (ctDNA+/-/-; n = 40 [48.2%]), while in the TNBC group, the largest group consisted of those who remained positive until T2 (ctDNA+/-/+; n = 42 [49.4%]).

There was no significant association between ctDNA clearance and pCR in HR-positive/HER2-negative patients (p =

0.92) (Figure 1G). In contrast, the proportion of patients who achieved pCR was significantly higher in TNBC patients who cleared their ctDNA early by T1 (ctDNA+/-/-) compared with those who cleared at T2 (ctDNA+/-/-) or remained positive for ctDNA until T2 (ctDNA+/-/+ (p = 0.0002) (Figure 1H). Early clearance was a significant predictor of pCR in the TNBC group (odds ratio, 13.06; 95% confidence interval [CI], 3.54–57.95) but not in the HR-positive/HER2-negative group (Figures 1G and 1H). Similar results were observed when associations between



**Figure 1. Predictive value of ctDNA in high-risk early-stage HER2-negative breast cancer receiving neoadjuvant chemotherapy (NAC)**

(A–D) Distribution of ctDNA concentration expressed as mean tumor molecules (MTM) per milliliter of plasma across time points during NAC in hormone receptor (HR)-positive/HER2-negative (HR+HER2–; left panel) and triple-negative breast cancer (TNBC; right panel) patients grouped according to (A and B) pathologic complete response (pCR) and (C and D) residual cancer burden (RCB) class. ctDNA was analyzed in plasma collected at pretreatment (T0), 3 weeks after treatment initiation (T1), 12 weeks after treatment initiation between paclitaxel-based treatment and anthracycline regimens (T2), and after NAC before surgery (T3). For each boxplot, the center line represents the median value (50th percentile), the box contains the 25th to 75th percentiles of the data distribution, the whiskers represent the 5th and 95th percentiles, and the dots beyond the upper and lower bounds are outliers.

(legend continued on next page)

ctDNA clearance and RCB were examined (Figures 1I and 1J). Early clearance was a significant predictor of RCB-0/I in the TNBC group (odds ratio, 19.00; 95% CI, 4.98–89.06) (Figure 1J) but not in the HR-positive/HER2-negative group (Figure 1I). Altogether, our analysis revealed a significant predictive value of early ctDNA clearance in TNBC patients but not in HR-positive/HER2-negative breast cancer patients.

### ctDNA positivity before, during, and after NAC is a significant negative prognostic factor

To evaluate the prognostic value of ctDNA, we examined its correlation with distant recurrence-free survival (DRFS; STAR Methods).

ctDNA positivity at all time points was significantly associated with inferior DRFS in both subtypes (log rank  $p < 0.05$  for all; Figure S4). The correlation between ctDNA positivity and poor DRFS was strongest at T2 (at 12 weeks) and T3 (after NAC before surgery) (log rank  $p < 0.0001$  for all). ctDNA positivity remained a significant negative prognostic factor for DRFS in bivariate Cox proportional hazards regression models that adjusted for pCR, except for T1 in patients with TNBC (Figure S5). No DRFS events were observed in TNBC patients who tested ctDNA negative at T0, and thus no accurate estimates of hazard ratio and 95% CI were calculated.

### ctDNA status after NAC refines the prognostic value of pCR and RCB

We examined whether ctDNA status after NAC before surgery (T3) could further refine risk stratification by pCR. Patients were classified into 4 groups using ctDNA status at T3 (ctDNA positive or ctDNA negative) and whether the patient achieved pCR or not.

In HR-positive/HER2-negative patients, the stratification yielded only 3 groups because all patients who achieved pCR were ctDNA negative at T3. As expected, these patients had the most favorable DRFS (Figure 2A). For those who failed to achieve pCR, DRFS was significantly worse in ctDNA-positive patients compared with ctDNA-negative patients (hazard ratio, 5.89; 95% CI, 2.68–12.98). Similar results were observed in TNBC: ctDNA-positive patients had a significantly increased risk for relapse and death compared with ctDNA-negative patients (hazard ratio, 3.79; 95% CI, 1.87–7.68; Figure 2B). One patient with TNBC who achieved a pCR was ctDNA positive at T3; this patient experienced a recurrence one year and 9 months after study entry.

We repeated the same analyses to assess whether combining ctDNA status at T3 and dichotomized RCB class (RCB-0/I vs. RCB-II/III) could improve risk stratification. In the HR-positive/HER2-negative group, patients classified as RCB-II/III who tested ctDNA positive at T3 had a significantly higher risk for metastatic recurrence and death than RCB-II/III patients who tested negative (hazard ratio, 5.65; 95% CI, 2.45–12.99) (Figure 2C). Survival analysis in TNBC showed similar results: ctDNA positivity at T3 was associated with significantly worse out-

comes (hazard ratio, 3.84; 95% CI, 1.70–8.66; Figure 2D). In addition, all 3 TNBC patients with RCB-0 ( $n = 1$ ) and RCB-I ( $n = 2$ ) who tested positive for ctDNA at T3 experienced metastatic recurrence.

In summary, among non-responders (no pCR or RCB-II/III), a ctDNA-negative result at T3 was significantly associated with improved survival compared with those who tested positive for ctDNA. These results suggest that risk stratification using established prognostic factors, such as pCR and RCB, especially in patients with poor response to NAC, can be improved by adding information from ctDNA testing after NAC.

### ctDNA dynamics during NAC are prognostic in both subtypes

We then examined the prognostic significance of ctDNA dynamics in each subtype. Patients with available ctDNA data for all 4 time points were classified into 5 groups on the basis of ctDNA dynamics (STAR Methods). The distributions of patients on the basis of ctDNA dynamics are shown in Figures 2E and 2F. The largest group among HR-positive/HER2-negative patients was group 1 (i.e., patients who tested negative at T0 and remained negative; 31%; Figure 2E) and among TNBC patients was group 4 (i.e., patients who were ctDNA positive at T0 and cleared at T3; 31.4%; Figure 2F).

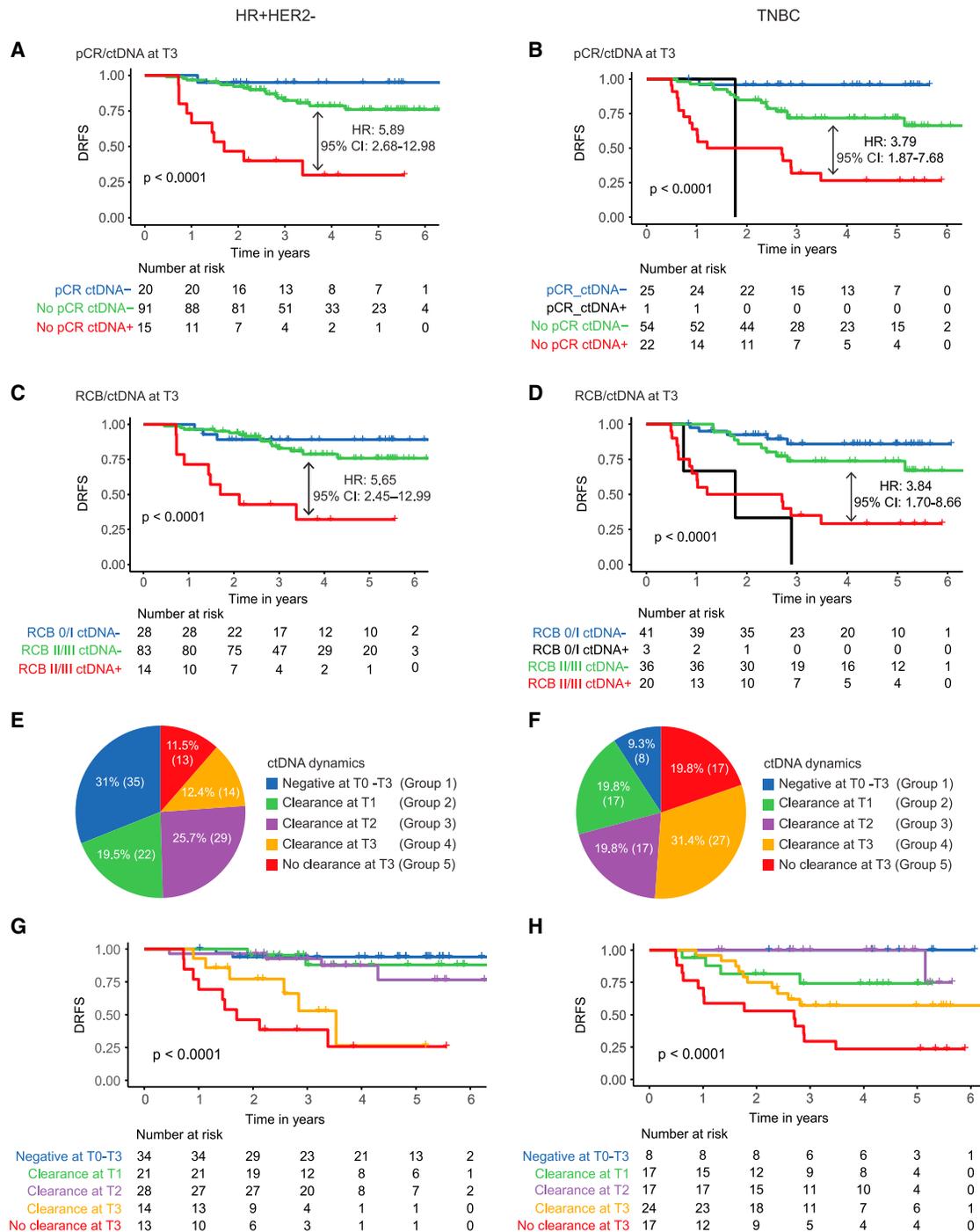
Consistent with the results of previous analysis examining the prognostic impact of ctDNA status at each time point, we observed that in both subtypes, patients who were ctDNA negative at T0 and consistently remained negative (group 1) had the most favorable survival (Figures 2G and 2H). In this group, 2 of the 35 HR-positive/HER2-negative patients experienced metastatic recurrence (Figure 2G). Of the 35, 74.3% ( $n = 26$ ) did not achieve pCR. In the same group in TNBC, none of the 8 patients relapsed or died (Figure 2H); 50% of them (4 of 8) did not achieve pCR. Consistent with previous analyses, we observed that failure to clear ctDNA (group 5) was associated with the poorest DRFS. Interestingly, in the TNBC group, patients who cleared their ctDNA after paclitaxel-based treatment (group 3) had favorable survival, with only 1 metastatic recurrence observed after a median follow-up duration of 4 years and 2 months (Figure 2H). Of the 17 patients in group 3, 13 (76.5%) did not achieve pCR.

### Genes expression correlates of ctDNA shedding

Next, we determined genes and pathways associated with ctDNA shedding by performing differential expression and gene set enrichment analysis (GSEA) of pretreatment tumor gene expression (T0) between ctDNA-positive and ctDNA-negative patients (Figure S6). Our analysis revealed 153 genes significantly differentially expressed (DE) (Benjamini-Hochberg [BH]-adjusted  $p < 0.05$ ) and >2,000 genes nominally DE ( $p < 0.05$ ) in the HR-positive/HER2-negative subtype (Figures S6A and S6B) and ~900 nominally DE genes in TNBC (Figures S6A and S6C).

(E–J) Patients who were ctDNA positive at T0 were classified into 3 groups on the basis of ctDNA dynamics from T0 to T2. Distribution of ctDNA dynamic patterns in (E) HR+HER2– and (F) TNBC patients with complete ctDNA data from T0–T2.

(G–J) Association of ctDNA dynamics with (G and H) pCR and (I and J) RCB. RCB was divided into 4 classes: RCB-0 (no residual disease equivalent to pCR), RCB-I (minimal burden), RCB-II (moderate burden), and RCB-III (extensive burden). Logistic regression was used to estimate the odds ratio (OR) and 95% confidence interval of achieving (G and H) pCR or (I and J) RCB-0/I.



**Figure 2. Prognostic value of ctDNA in high-risk early-stage HER2-negative breast cancer receiving neoadjuvant chemotherapy (NAC)**

(A–D) Survival curves of patients grouped according to combined ctDNA status after NAC before surgery (ctDNA positive or ctDNA negative at T3) and (A and B) pCR (pCR or no pCR) or (C and D) binarized RCB (RCB-0/I or RCB-II/III) in hormone receptor (HR)-positive/HER2-negative subset (HR+HER2–; left panel) and triple-negative breast cancer (TNBC; right panel). RCB was divided into 2 classes: RCB-0/I (no residual disease equivalent to pCR) and RCB-I (minimal burden) vs. RCB-II (moderate burden) and RCB-III (extensive burden). Hazard ratios (HR) and 95% confidence intervals (CI) shown were estimated from Cox proportional hazards regression models.

(E–H) Distribution of ctDNA dynamic patterns in (E) HR+HER2– and (F) TNBC patients with complete ctDNA data for all time points.

(G and H) Survival curves of (G) HR+HER2– and (H) TNBC patients grouped according to ctDNA dynamics. Distant recurrence-free survival (DRFS) was the survival endpoint. p values were calculated using the log rank test.

In the HR-positive/HER2-negative subtype, GSEA revealed enrichment of 28 Molecular Signatures Database (MSigDB) Hallmark gene sets (56%) significantly associated with ctDNA status at baseline (BH-adjusted  $p < 0.05$ ), with one other nominally enriched ( $p < 0.05$ ) (Figure 3A, top panel). Of these, 11 gene sets, including metabolic and estrogen response gene sets, had higher expression levels in ctDNA-negative patients. In ctDNA-positive patients, we observed significant enrichment of 18 gene sets—including higher expression levels of cell cycle and proliferation gene sets, inflammatory cytokines (IL-6, TNF- $\alpha$ ), and inflammatory response immune-associated gene sets. In TNBC, GSEA identified nominal enrichment of 5 gene sets ( $p < 0.05$ ) (Figure 3A, bottom panel), including immune-associated interferon- $\alpha$  response and TGF- $\beta$  signaling.

To explore the impact of proliferation, we next performed differential expression analysis in the HR-positive/HER2-negative subtype, adjusting for MammaPrint status, and found fewer DE genes detected (Figure S6C). The 3 genes that remained significantly DE (adjusted  $p < 0.05$ ) after adjusting for MammaPrint status include *N92541*, *CYBRD1*, and *RCBTB1*, all upregulated in ctDNA-negative patients. Despite adjustment for MammaPrint status, GSEA still captured gene sets associated with cell cycle processes (Figure S6D). These analyses suggest a strong association between proliferation and ctDNA release in the HR-positive/HER2-negative subtype.

Immune cells in the tumor microenvironment have been proposed to be involved in the shedding of ctDNA into blood.<sup>27</sup> We performed leading-edge analysis (LEA)<sup>28</sup> to identify genes (referred to as leading-edge genes) with the highest impact on the enrichment signal of immune response pathways associated with ctDNA shedding. In the HR-positive/HER2-negative subtype, LEA identified key genes within enriched pathways in ctDNA-positive patients, including 32 genes involved in IL-6/JAK/STAT3 signaling (e.g., *CCL7*, *BAK1*, *TNFRSF12A*); 76 genes involved in inflammatory response (e.g., *PVR*, *CCL7*, *ADRM1*); and 68 genes involved in TNF- $\alpha$  signaling via NF- $\kappa$ B (e.g., *VEGFA*, *ATF3*, *HBEGF*) (Figures 3B and S6Ei–S6Eiii). In TNBC, LEA identified key genes, including 45 genes in interferon- $\alpha$  response (e.g., *LY6E*, *TRIM26*, *TRIM14*) enriched in ctDNA-positive patients and 16 genes in TGF- $\beta$  signaling (e.g., *SMAD1*, *LTP2*, *TRIM33*) enriched in ctDNA-negative patients (Figures 3C, S6Fi, and S6Fii).

Next, we considered the 190 genes that were commonly DE ( $p < 0.05$ ) across subtypes with concordant up- or down-regulation associated with ctDNA release (Figure S6G). Protein-protein interaction (PPI) network analysis identified 99 genes that form the main PPI network with 9 gene communities (Figure 3D). PPI functional enrichment of each gene community (STRING database) identified the association of 5 communities with Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (Figure 3D). These results suggest that ctDNA shedding at baseline in both HR-positive/HER2-negative breast cancer and TNBC tumors was associated with active carbon and amino acid metabolism, high proliferation, and high levels of immune activity.

## DISCUSSION

In this study involving 283 patients, we examined the clinical significance and gene expression correlates of ctDNA shedding in

HER2-negative breast cancers (HR-positive/HER2-negative breast cancer and TNBC), representing the majority of breast cancer cases.<sup>29</sup>

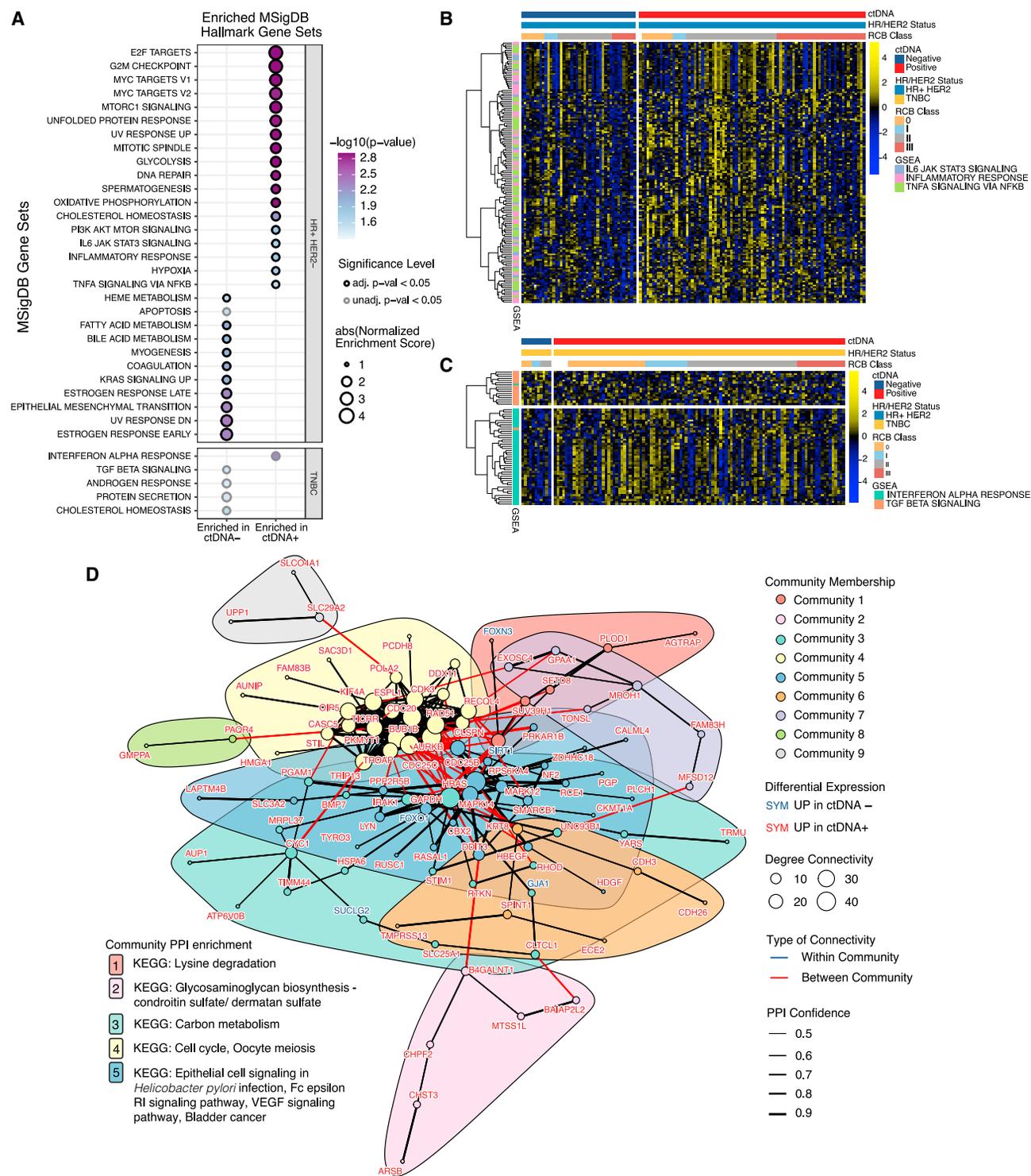
CtDNA levels at all time points were higher in TNBC than in HR-positive/HER2-negative breast cancer patients. The higher ctDNA positivity rate in TNBCs may be due to the high proliferation and cell turnover rates in this subtype,<sup>30</sup> both of which are associated with increased ctDNA shedding.<sup>24</sup>

Moreover, ctDNA positivity and dynamics were more strongly associated with response to NAC in TNBC than in HR-positive/HER2-negative breast cancer. This may be due to inherent differences in the biology, response to treatment, and disease trajectories between the two subtypes.<sup>9,10,31,32</sup> For example, we found that early clearance of ctDNA at 3 weeks after initiation of therapy was predictive of pCR and RCB in TNBC but not in the HR-positive/HER2-negative subset. These results can inform decisions for treatment redirection (escalation/de-escalation) in the neoadjuvant setting to increase the likelihood of patients achieving a pCR. Switching to more effective therapy in TNBC patients who do not clear ctDNA at 3 weeks may improve treatment response.

We examined the prognostic significance of serial ctDNA status and dynamics. In both subtypes, ctDNA positivity at pretreatment was significantly associated with an increased risk for metastatic recurrence or death, confirming previous observations.<sup>26</sup> In most cases, ctDNA positivity at different time points remained a significant negative prognostic factor for DRFS even after adjustment for the effects of pCR. These results suggest that ctDNA reflects therapeutic efficacy and could serve as an early surrogate marker of survival.

To assess whether ctDNA can refine risk stratification by pCR and RCB, we grouped patients on the basis of ctDNA status after NAC before surgery (ctDNA positive or ctDNA negative at T3) and either pCR status (pCR or no pCR) or binarized RCB class (RCB-0/I or RCB-II/III). In both subtypes, patients who were ctDNA positive after NAC and had a poor response to NAC (no pCR or RCB-II/III) showed the worst survival outcome of all the groups compared. Most important, our analysis revealed that ctDNA negativity after NAC is associated with favorable DRFS, even in those who did not achieve pCR or those with moderate or extensive RCB (RCB-II/III). Residual cancers that do not shed detectable levels of ctDNA may differ in biology from those that do and may represent a less aggressive type of cancer with lower metastatic potential. These findings could help inform patient decisions on whether to have additional treatment after primary NAC.

We next examined the prognostic significance of ctDNA dynamics in patients with ctDNA data for all time points. As expected, patients in both subtypes who did not clear their ctDNA after NAC before surgery (T3) had the poorest survival outcomes, whereas patients who tested ctDNA negative at all time points appeared to have the best survival outcomes. We speculate that patients who were ctDNA negative at all time points and did not achieve pCR may have pretreatment and residual tumors that are less aggressive and proliferative and, therefore, are less likely to recur. Thus, patients with undetectable levels of ctDNA may be eligible for treatment de-escalation in both the neoadjuvant and adjuvant settings. If our observations are confirmed in a larger cohort, consistent negative tests for ctDNA before, during,



**Figure 3. Differentially expressed genes and enriched gene sets associated with ctDNA shedding at pretreatment**

(A) GSEA of MSigDB Hallmark gene sets in hormone receptor (HR)-positive/HER2-negative subset (HR+HER2-) (top) and triple-negative breast cancer (TNBC) (bottom) subtypes. Gene sets with significant (BH-adjusted  $p < 0.05$ ) and nominally significant ( $p < 0.05$ ) enrichment in ctDNA-positive or ctDNA-negative patients are depicted with p value significance and normalized enrichment scores (NES) annotated.

(B and C) Gene expression heatmap of leading-edge genes in enriched immune-associated gene sets: (B) IL-6/JAK/STAT3 signaling, inflammatory response, and TNF-alpha signaling via NF- $\kappa$ B in HR+HER2- and (C) interferon alpha response and TGF-beta signaling in TNBC subtypes.

(D) STRING protein-protein interaction (PPI) of genes that are commonly differentially expressed ( $p < 0.05$ ) in both HR+HER2- and TNBC with concordant fold change direction across subtypes. The PPI for the 99 genes comprising the main network is shown with community membership, differential expression direction, degree and type of connectivity, and PPI confidence level annotated. Each community's STRING PPI KEGG functional enrichment is listed.

and after NAC could be an early surrogate of favorable survival in both subtypes.

Our exploratory analysis of pretreatment gene expression revealed common and unique patterns potentially associated with ctDNA shedding across and within subtypes. Although our power to detect DE genes was limited in TNBC because of an imbalance between the ctDNA-positive and ctDNA-negative groups, PPI enrichment of commonly DE genes ( $p < 0.05$ ) identified cell cycle genes associated with ctDNA shedding in both HR-positive/HER2-negative and TNBC subtypes, suggesting an important role of tumor cell proliferation in ctDNA release. Indeed, associations of cell cycle-related gene sets with ctDNA release were particularly significant in the HR-positive/HER2-negative subtype, even after adjustment for MammaPrint status. Transcriptional analysis of urothelial tumors from ctDNA-positive patients also showed upregulation of cell cycle genes.<sup>22</sup> Additionally, gene expression analysis suggested a key role of specific immune response pathways in ctDNA release, consistent with preclinical studies demonstrating the potential role of immune cells in the shedding of DNA into circulation.<sup>33,34</sup>

Recent studies in early-stage breast cancers have shown that ctDNA before<sup>17,26</sup> and after NAC<sup>11–13,15,17,18</sup> were prognostic of poor survival. Because our study analyzed ctDNA at 4 time points, we demonstrated the prognostic value, not only of pretreatment and post-NAC ctDNA but also of ctDNA detection as early as 3 and 12 weeks after initiation of treatment. Moreover, we classified patients into groups using the results from serial ctDNA tests and demonstrated the prognostic significance of ctDNA dynamics. In addition, our study showed that combining information from ctDNA status after NAC and treatment response could improve risk stratification, especially in patients with no pCR. In contrast, none of the previous studies examined the prognostic significance of ctDNA dynamics and combined ctDNA and response information. A limitation of this study is the heterogeneity in the treatment received by patients. Approximately 50% of patients received standard chemotherapy (paclitaxel plus AC), while the rest received various investigational drugs combined with paclitaxel. We speculate that ctDNA, like radiographic imaging, could serve as a measure of tumor burden regardless of treatment received. Another limitation is that the patient-specific mutations selected for the design of each personalized ctDNA assay were based on the tumor mutational profiles of tissue biopsy taken before treatment. Less dominant clones in the untreated tumor could be selected in response to treatment resulting in the emergence of new mutations that the initial assay would miss. Because of the small sample size available for analysis, we could not evaluate the predictive and prognostic significance of ctDNA in patients with HER2-positive disease, representing approximately 15% of breast cancers.<sup>29</sup> We also acknowledge limitations in the sensitivity and specificity of the ctDNA assay used in the study. For example, some patients with residual cancer tested negative for ctDNA after NAC. This represents a potential issue with the performance of the assay and also suggests that a ctDNA test alone may not provide reliable information regarding disease status and patient outcomes. Thus, improvements in assay sensitivity are needed, and combining ctDNA test results with other tests (e.g., imaging and pathology) could improve the detection of residual disease and the prediction of metastatic recurrence.

Importantly, with the potential for false negatives and false positives, caution should be exercised when using ctDNA information to de-escalate or escalate treatment.

Here, we present the results of a large study that examined the predictive and prognostic value of ctDNA as well as the biology of ctDNA shedding in early-stage HER2-negative breast cancer receiving NAC. Our findings can inform the design of future clinical trials that seek to evaluate the clinical utility of ctDNA in the neoadjuvant setting. To this end, we plan to expand our ctDNA analyses to include 1,000 patients enrolled in the I-SPY2 trial to validate the findings described in this report, expand investigations across all breast cancer subtypes (including HER2-positive patients), and compare ctDNA trajectories in patients who received immunotherapy agents vs. those who did not.

The I-SPY2 trial will prospectively test ctDNA for utility in redirecting therapy to improve patient outcomes. Results from ctDNA tests, combined with those from MRI<sup>15,25</sup> and pathology of on-treatment tumor biopsy, will guide early therapy changes (treatment escalation) in non-responding patients to increase the likelihood of achieving a pCR. Also, information from these tests will aid decisions regarding treatment de-escalation for patients predicted to have a pCR by providing an option for early surgery to limit exposure to toxicity of unnecessary treatments.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead contact
  - Materials availability
  - Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - The I-SPY2 trial
  - Patients
  - Clinical samples
- **METHOD DETAILS**
  - Quality control and analytic dataset
  - ctDNA analysis
  - Gene expression analysis
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
  - Statistical analysis of the clinical and gene expression data
- **ADDITIONAL RESOURCES**

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.ccell.2023.04.008>.

## ACKNOWLEDGMENTS

Thanks to the I-SPY2 Biomarker Working Group, patients, advocates, and investigators. The work reported in this paper is funded in part by the NIH/NCI (grant R01CA255442), NIH/NCI I-SPY2+ (grant PO1-CA210961), NIH/NCI Imaging (grant 28XS197 P-0518835), NIH/NCI Cancer Cell Map Initiative (CCMI) (grant U54CA209891), NIH/NCI Cancer Center Support Grant (CCSG) (grant P30-CA82103), NIH/NHGRI Big Data (grant U54-HG007990), Breast Cancer Research Foundation (grant BCRF-20-142), the Breast Cancer Research

Foundation (grant BCRF-20-165), Breast Cancer Research – Atwater Trust, Stand Up to Cancer, the California Breast Cancer Research Program, and Give Breast Cancer the Boot. Natera provided the ctDNA assay at no cost. Specimens and clinical data were provided by Quantum Leap Healthcare Collaborative, sponsor of the I-SPY2 trial. The authors are solely responsible for the study design, analysis of data, interpretation of results, and manuscript preparation.

#### AUTHOR CONTRIBUTIONS

M.J.M.M., L.J.v.V., L.B.S., Z.A., R.W.S., D.M.W., and C.Y. designed the study and interpreted the data. G.L.H. manages the Biomarker Working Group. S.A. manages the I-SPY trials operations. Natera employees (D.R., E.K., A.R., H.S., A.A., M.C.L., and M.R.) generated the ctDNA data. J.P. and A.L.D. are patient advocates. For I-SPY2 investigators, M.C.L., K.A., A.J.C., A.F.-T., C.I., R.N., and D.T. were drug chaperones; L.J.v.V., N.M.H., W.F.S., H.S.R., A.M.D., and D.Y. are working group leads; and L.J.E. is the principal investigator. The authors have approved the final manuscript and agreed to be accountable for all aspects of the work.

#### DECLARATION OF INTERESTS

R.W.S. owns stock in Pfizer Inc., AstraZeneca, and Moderna Inc. D.R., E.K., A.R., H.S., A.A., M.C.L., and M.R. are employees of and/or hold stock or stock options in Natera Inc. A.L.D. reports honoraria from the Department of Defense and the California Breast Cancer Research Program (CBCRP). M.C.L. reports funding from Eisai, Exact Sciences, Genentech, Genomic Health, GRAIL, Menarini Silicon Biosystems, Merck, Novartis, and Seattle Genetics; participation on advisory boards (no personal compensation) of Adela, AstraZeneca, Celgene, Roche/Genentech, Genomic Health, GRAIL, Ionis, Merck, Pfizer, Seattle Genetics, Syndax; meeting support from Agena, AstraZeneca, Celgene, Cynvenio, Genomic Health, GRAIL, Ionis, Menarini Silicon Biosystems, Merck, Pfizer. KA reports support from Merck, Seattle Genetics, Amgen, Genentech-Roche; Daiichi Sankyo, and AstraZeneca; participation on an advisory board for Genomic Health/Exact Sciences, Genentech-Roche, and a data and safety monitoring board for Seattle Genetics/Axio. A.J.C. reports funding from Novartis. A.F.-T. is an employee of Seagen. C.I. reports funding from Tesaro/GlaxoSmithKline, Seattle Genetics, Pfizer, AstraZeneca, Bristol Myers Squibb, Genentech, Novartis, PUMA, Eisai, Sanofi, ION, and Gilead. R.N. reports funding from Arvinas, AstraZeneca, Celgene, Corcept Therapeutics, Genentech/Roche, Gilead/Immunomedics, Merck, OBI Pharma Inc., OncoSec Medical, Pfizer, Relay Therapeutics, Seattle Genetics, Sun Pharmaceutical Industries Ltd., Taiho Pharmaceutica, BeyondSpring Inc., FUJIFILM Pharmaceuticals, Infinty Pharmaceuticals Inc., ITeos Therapeutics, and Seagen. J.P. reports honoraria from Methods in Clinical Research. W.F.S. reports funding from AstraZeneca and Pfizer; owns stock in IONIS Pharmaceuticals and Eiger Biopharmaceuticals; and receives royalties for patents licensed by the MD Anderson Cancer Center to Delphi Diagnostics, Inc. D.Y. reports funding from Fusion Pharmaceutical, Boehringer Ingelheim, Martell Diagnostics, and Akston Biosciences. L.J.E. reports funding from Merck & Co.; participation on an advisory board for Blue Cross Blue Shield; and personal fees from UpToDate. A.M.D. reports funding from Pfizer, Genentech, Novartis, Inivata Ltd., and Calithera Biosciences. H.S.R. reports funding from Pfizer, Merck, Novartis, Lilly, Roche, Daiichi, Seattle Genetics, MacroGenics, Sermonix, Boehringer Ingelheim, Polypor, AstraZeneca, Ayala, Astellas, Gilead, Puma, Samsung, Chugai, Blueprint, NAPO, and GE Healthcare. L.J.v.V. is a part-time employee and owns stock in Agendia. All other authors declare no competing interests.

#### INCLUSION AND DIVERSITY

We worked to ensure ethnic or other types of diversity in the recruitment of human subjects. One or more of the authors of this paper self-identifies as an underrepresented ethnic minority in their field of research or within their geographical location. One or more of the authors of this paper self-identifies as a member of the LGBTQIA+ community. One or more of the authors of this paper self-identifies as living with a disability. While citing references scientifically relevant for this work, we also actively worked to promote gender balance in our reference list.

Received: September 12, 2022

Revised: January 30, 2023

Accepted: April 12, 2023

Published: May 4, 2023

#### REFERENCES

1. Abbosh, C., and Swanton, C. (2021). ctDNA: an emerging neoadjuvant biomarker in resectable solid tumors. *PLoS Med.* 18, e1003771. <https://doi.org/10.1371/journal.pmed.1003771>.
2. Pascual, J., Attard, G., Bidard, F.C., Curigliano, G., De Mattos-Arruda, L., Diehn, M., Italiano, A., Lindberg, J., Merker, J.D., Montagut, C., et al. (2022). ESMO recommendations on the use of circulating tumour DNA assays for patients with cancer: a report from the ESMO Precision Medicine Working Group. *Ann. Oncol.* 33, 750–768. <https://doi.org/10.1016/jannonc.2022.05.520>.
3. Korde, L.A., Somerfield, M.R., Carey, L.A., Crews, J.R., Denduluri, N., Hwang, E.S., Khan, S.A., Loibl, S., Morris, E.A., Perez, A., et al. (2021). Neoadjuvant chemotherapy, endocrine therapy, and targeted therapy for breast cancer: ASCO guideline. *J. Clin. Oncol.* 39, 1485–1505. <https://doi.org/10.1200/JCO.20.03399>.
4. I-SPY2 Trial Consortium, Yee, D., DeMichele, A.M., Yau, C., Isaacs, C., Symmans, W.F., Albain, K.S., Chen, Y.Y., Krings, G., Wei, S., Harada, S., et al. (2020). Association of event-free and distant recurrence-free survival with individual-level pathologic complete response in neoadjuvant treatment of stages 2 and 3 breast cancer: three-year follow-up analysis for the I-SPY2 adaptively randomized clinical trial. *JAMA Oncol.* 6, 1355–1362. <https://doi.org/10.1001/jamaoncol.2020.2535>.
5. Spring, L.M., Fell, G., Arfe, A., Sharma, C., Greenup, R., Reynolds, K.L., Smith, B.L., Alexander, B., Moy, B., Isakoff, S.J., et al. (2020). Pathologic complete response after neoadjuvant chemotherapy and impact on breast cancer recurrence and survival: a comprehensive meta-analysis. *Clin. Cancer Res.* 26, 2838–2848. <https://doi.org/10.1158/1078-0432.CCR-19-3492>.
6. Cortazar, P., Zhang, L., Untch, M., Mehta, K., Costantino, J.P., Wolmark, N., Bonnefoi, H., Cameron, D., Gianni, L., Valagussa, P., et al. (2014). Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. *Lancet* 384, 164–172. [https://doi.org/10.1016/S0140-6736\(13\)62422-8](https://doi.org/10.1016/S0140-6736(13)62422-8).
7. von Minckwitz, G., Huang, C.S., Mano, M.S., Loibl, S., Mamounas, E.P., Untch, M., Wolmark, N., Rastogi, P., Schneeweiss, A., Redondo, A., et al. (2019). Trastuzumab emtansine for residual invasive HER2-positive breast cancer. *N. Engl. J. Med.* 380, 617–628. <https://doi.org/10.1056/NEJMoa1814017>.
8. Masuda, N., Lee, S.J., Ohtani, S., Im, Y.H., Lee, E.S., Yokota, I., Kuroi, K., Im, S.A., Park, B.W., Kim, S.B., et al. (2017). Adjuvant capecitabine for breast cancer after preoperative chemotherapy. *N. Engl. J. Med.* 376, 2147–2159. <https://doi.org/10.1056/NEJMoa1612645>.
9. Wolf, D.M., Yau, C., Wulfkuhle, J., Brown-Swigart, L., Gallagher, R.I., Lee, P.R.E., Zhu, Z., Magbanua, M.J., Sayaman, R., O'Grady, N., et al. (2022). Redefining breast cancer subtypes to guide treatment prioritization and maximize response: predictive biomarkers across 10 cancer therapies. *Cancer Cell* 40, 609–623.e6. <https://doi.org/10.1016/j.ccell.2022.05.005>.
10. Houssami, N., Macaskill, P., von Minckwitz, G., Marinovich, M.L., and Mamounas, E. (2012). Meta-analysis of the association of breast cancer subtype and pathologic complete response to neoadjuvant chemotherapy. *Eur. J. Cancer* 48, 3342–3354. <https://doi.org/10.1016/j.ejca.2012.05.023>.
11. Cavallone, L., Aguilar-Mahecha, A., Lafleur, J., Brousse, S., Aldamry, M., Roseshter, T., Lan, C., Alirezaie, N., Bareke, E., Majewski, J., et al. (2020). Prognostic and predictive value of circulating tumor DNA during neoadjuvant chemotherapy for triple negative breast cancer. *Sci. Rep.* 10, 14704. <https://doi.org/10.1038/s41598-020-71236-y>.
12. Ortolan, E., Appierto, V., Silvestri, M., Miceli, R., Veneroni, S., Folli, S., Pruneri, G., Vingiani, A., Belfiore, A., Cappelletti, V., et al. (2021). Blood-based

genomics of triple-negative breast cancer progression in patients treated with neoadjuvant chemotherapy. *ESMO Open* 6, 100086.

13. Riva, F., Bidard, F.C., Houy, A., Saliou, A., Madic, J., Rampanou, A., Hego, C., Milder, M., Cottu, P., Sablin, M.P., et al. (2017). Patient-specific circulating tumor DNA detection during neoadjuvant chemotherapy in triple-negative breast cancer. *Clin. Chem.* 63, 691–699. <https://doi.org/10.1373/clinchem.2016.262337>.
14. Rothé, F., Silva, M.J., Venet, D., Campbell, C., Bradbury, I., Rouas, G., de Azambuja, E., Maetens, M., Fumagalli, D., Rodrik-Outmezguine, V., et al. (2019). Circulating tumor DNA in HER2-amplified breast cancer: a translational research substudy of the NeoALTTO phase III trial. *Clin. Cancer Res.* 25, 3581–3588. <https://doi.org/10.1158/1078-0432.Ccr-18-2521>.
15. Magbanua, M.J.M., Swigart, L.B., Wu, H.T., Hirst, G.L., Yau, C., Wolf, D.M., Tin, A., Salari, R., Shchegrova, S., Pawar, H., et al. (2021). Circulating tumor DNA in neoadjuvant-treated breast cancer reflects response and survival. *Ann. Oncol.* 32, 229–239. <https://doi.org/10.1016/j.annonc.2020.11.007>.
16. McDonald, B.R., Contente-Cuomo, T., Sammut, S.-J., Odenheimer-Bergman, A., Ernst, B., Perdignes, N., Chin, S.-F., Farooq, M., Mejia, R., Cronin, P.A., et al. (2019). Personalized circulating tumor DNA analysis to detect residual disease after neoadjuvant therapy in breast cancer. *Sci. Transl. Med.* 11, eaax7392. <https://doi.org/10.1126/scitranslmed.aax7392>.
17. Li, S., Lai, H., Liu, J., Liu, Y., Jin, L., Li, Y., Liu, F., Gong, Y., Guan, Y., Yi, X., et al. (2020). Circulating tumor DNA predicts the response and prognosis in patients with early breast cancer receiving neoadjuvant chemotherapy. *JCO Precis. Oncol.* 4, 244–257. <https://doi.org/10.1200/PO.19.00292>.
18. Lin, P.H., Wang, M.Y., Lo, C., Tsai, L.W., Yen, T.C., Huang, T.Y., Huang, W.C., Yang, K., Chen, C.K., Fan, S.C., et al. (2021). Circulating tumor DNA as a predictive marker of recurrence for patients with stage II-III breast cancer treated with neoadjuvant therapy. *Front. Oncol.* 11, 736769. <https://doi.org/10.3389/fonc.2021.736769>.
19. Zhou, Q., Gampenrieder, S.P., Frantal, S., Rinnerthaler, G., Singer, C.F., Egle, D., Pfeiler, G., Bartsch, R., Wette, V., Pichler, A., et al. (2022). Persistence of ctDNA in patients with breast cancer during neoadjuvant treatment is a significant predictor of poor tumor response. *Clin. Cancer Res.* 28, 697–707. <https://doi.org/10.1158/1078-0432.CCR-21-3231>.
20. Yau, C., Osdoit, M., van der Noordaa, M., Shad, S., Wei, J., de Croze, D., Hamy, A.S., Laé, M., Reyat, F., Sonke, G.S., et al. (2022). Residual cancer burden after neoadjuvant chemotherapy and long-term survival outcomes in breast cancer: a multicentre pooled analysis of 5161 patients. *Lancet Oncol.* 23, 149–160. [https://doi.org/10.1016/S1470-2045\(21\)00589-1](https://doi.org/10.1016/S1470-2045(21)00589-1).
21. Symmans, W.F., Peintinger, F., Hatzis, C., Rajan, R., Kuerer, H., Valero, V., Assad, L., Poniacka, A., Hennessy, B., Green, M., et al. (2007). Measurement of residual breast cancer burden to predict survival after neoadjuvant chemotherapy. *J. Clin. Oncol.* 25, 4414–4422. <https://doi.org/10.1200/JCO.2007.10.6823>.
22. Powles, T., Assaf, Z.J., Davarpanah, N., Banchereau, R., Szabados, B.E., Yuen, K.C., Grivas, P., Hussain, M., Oudard, S., Gschwend, J.E., et al. (2021). ctDNA guiding adjuvant immunotherapy in urothelial carcinoma. *Nature* 595, 432–437. <https://doi.org/10.1038/s41586-021-03642-9>.
23. Abbosh, C., Birkbak, N.J., Wilson, G.A., Jamal-Hanjani, M., Constantin, T., Salari, R., Le Quesne, J., Moore, D.A., Veeriah, S., Rosenthal, R., et al. (2017). Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. *Nature* 545, 446–451. <https://doi.org/10.1038/nature22364>.
24. Avanzini, S., Kurtz, D.M., Chabon, J.J., Moding, E.J., Hori, S.S., Gambhir, S.S., Alizadeh, A.A., Diehn, M., and Reiter, J.G. (2020). A mathematical model of ctDNA shedding predicts tumor detection size. *Sci. Adv.* 6, eabc4308. <https://doi.org/10.1126/sciadv.abc4308>.
25. Magbanua, M.J.M., Li, W., Wolf, D.M., Yau, C., Hirst, G.L., Swigart, L.B., Newitt, D.C., Gibbs, J., Delson, A.L., Kalashnikova, E., et al. (2021). Circulating tumor DNA and magnetic resonance imaging to predict neoadjuvant chemotherapy response and recurrence risk. *npj Breast Cancer* 7, 32. <https://doi.org/10.1038/s41523-021-00239-3>.
26. Garcia-Murillas, I., Chopra, N., Comino-Méndez, I., Beaney, M., Tovey, H., Cutts, R.J., Swift, C., Kriplani, D., Afentakis, M., Hrebien, S., et al. (2019). Assessment of molecular relapse detection in early-stage breast cancer. *JAMA Oncol.* 5, 1473–1478. <https://doi.org/10.1001/jamaoncol.2019.1838>.
27. Papadopoulos, N. (2020). Pathophysiology of ctDNA release into the circulation and its characteristics: what is important for clinical applications. *Recents Results Cancer Res.* 275, 163–180.
28. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* 102, 15545–15550. <https://doi.org/10.1073/pnas.0506580102>.
29. Howlader, N., Altekruse, S.F., Li, C.I., Chen, V.W., Clarke, C.A., Ries, L.A.G., and Cronin, K.A. (2014). US incidence of breast cancer subtypes defined by joint hormone receptor and HER2 status. *J. Natl. Cancer Inst.* 106, dju055. <https://doi.org/10.1093/jnci/dju055>.
30. Healey, M.A., Hirko, K.A., Beck, A.H., Collins, L.C., Schnitt, S.J., Eliassen, A.H., Holmes, M.D., Tamimi, R.M., and Hazra, A. (2017). Assessment of Ki67 expression for breast cancer subtype classification and prognosis in the Nurses' Health Study. *Breast Cancer Res. Treat.* 166, 613–622. <https://doi.org/10.1007/s10549-017-4421-3>.
31. Dai, X., Li, T., Bai, Z., Yang, Y., Liu, X., Zhan, J., and Shi, B. (2015). Breast cancer intrinsic subtype classification, clinical use and future trends. *Am. J. Cancer Res.* 5, 2929–2943.
32. Li, X., Zhou, J., Xiao, M., Zhao, L., Zhao, Y., Wang, S., Gao, S., Zhuang, Y., Niu, Y., Li, S., et al. (2021). Uncovering the subtype-specific molecular characteristics of breast cancer by multiomics analysis of prognosis-associated genes, driver genes, signaling pathways, and immune activity. *Front. Cell Dev. Biol.* 9, 689028. <https://doi.org/10.3389/fcell.2021.689028>.
33. Choi, J.J., Reich, C.F., 3rd, and Pisetsky, D.S. (2005). The role of macrophages in the in vitro generation of extracellular DNA from apoptotic and necrotic cells. *Immunology* 115, 55–62. <https://doi.org/10.1111/j.1365-2567.2005.02130.x>.
34. Park, J., Wysocki, R.W., Amoozgar, Z., Maiorino, L., Fein, M.R., Jorns, J., Schott, A.F., Kinugasa-Katayama, Y., Lee, Y., Won, N.H., et al. (2016). Cancer cells induce metastasis-supporting neutrophil extracellular DNA traps. *Sci. Transl. Med.* 8, 361ra138. <https://doi.org/10.1126/scitranslmed.aag1711>.
35. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). Limma powers differential expression analyses for RNA-seq and microarray studies. *Nucleic Acids Res.* 43, e47. <https://doi.org/10.1093/nar/gkv007>.
36. Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M.N., and Sergushichev, A. (2021). Fast gene set enrichment analysis. Preprint at bioRxiv. <https://doi.org/10.1101/060012>.
37. Szklarczyk, D., Gable, A.L., Nastou, K.C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N.T., Legeay, M., Fang, T., Bork, P., et al. (2021). The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* 49, D605–D612. <https://doi.org/10.1093/nar/gkaa1074>.
38. R Core Team (2020). *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing).
39. Therneau, T.M., and Grambsch, P.M. (2000). *Modeling Survival Data : Extending the Cox Model* (Springer).
40. Berry, D.A. (2015). The Brave New World of clinical cancer research: adaptive biomarker-driven trials integrating clinical practice with clinical research. *Mol. Oncol.* 9, 951–959. <https://doi.org/10.1016/j.molonc.2015.02.011>.
41. Rugo, H.S., Olopade, O.I., DeMichele, A., Yau, C., van 't Veer, L.J., Buxton, M.B., Hogarth, M., Hylton, N.M., Paoloni, M., Perlmutter, J., et al. (2016). Adaptive randomization of veliparib-carboplatin treatment in breast cancer. *N. Engl. J. Med.* 375, 23–34. <https://doi.org/10.1056/NEJMoa1513749>.

42. Park, J.W., Liu, M.C., Yee, D., Yau, C., van 't Veer, L.J., Symmans, W.F., Paoloni, M., Perlmutter, J., Hylton, N.M., Hogarth, M., et al. (2016). Adaptive randomization of neratinib in early breast cancer. *N. Engl. J. Med.* 375, 11–22. <https://doi.org/10.1056/NEJMoa1513750>.
43. Pusztai, L., Yau, C., Wolf, D.M., Han, H.S., Du, L., Wallace, A.M., String-Reasor, E., Boughey, J.C., Chien, A.J., Elias, A.D., et al. (2021). Durvalumab with olaparib and paclitaxel for high-risk HER2-negative stage II/III breast cancer: results from the adaptively randomized I-SPY2 trial. *Cancer Cell* 39, 989–998.e5. <https://doi.org/10.1016/j.ccell.2021.05.009>.
44. Piccart, M., van 't Veer, L.J., Poncet, C., Lopes Cardozo, J.M.N., Delalogue, S., Pierga, J.Y., Vuylsteke, P., Brain, E., Vrijaldenhoven, S., Neijenhuis, P.A., et al. (2021). 70-gene signature as an aid for treatment decisions in early breast cancer: updated results of the phase 3 randomised MINDACT trial with an exploratory analysis by age. *Lancet Oncol.* 22, 476–488. [https://doi.org/10.1016/S1470-2045\(21\)00007-3](https://doi.org/10.1016/S1470-2045(21)00007-3).
45. Howlander, N., Cronin, K.A., Kurian, A.W., and Andridge, R. (2018). Differences in breast cancer survival by molecular subtypes in the United States. *Cancer Epidemiol. Biomarkers Prev.* 27, 619–626. <https://doi.org/10.1158/1055-9965.EPI-17-0627>.
46. Schmid, P., Cortes, J., Dent, R., Pusztai, L., McArthur, H., Kümmel, S., Bergh, J., Denkert, C., Park, Y.H., Hui, R., et al. (2022). Event-free survival with pembrolizumab in early triple-negative breast cancer. *N. Engl. J. Med.* 386, 556–567.
47. Cardoso, F., van 't Veer, L.J., Bogaerts, J., Slaets, L., Viale, G., Delalogue, S., Pierga, J.Y., Brain, E., Causeret, S., DeLorenzi, M., et al. (2016). 70-Gene signature as an aid to treatment decisions in early-stage breast cancer. *N. Engl. J. Med.* 375, 717–729. <https://doi.org/10.1056/NEJMoa1602253>.
48. Bratman, S.V., Yang, S.Y.C., Iafora, M.A.J., Liu, Z., Hansen, A.R., Bedard, P.L., Lheureux, S., Spreafico, A., Razak, A.A., Shchegrova, S., et al. (2020). Personalized circulating tumor DNA analysis as a predictive biomarker in solid tumor patients treated with pembrolizumab. *Nat. Cancer* 1, 873–881. <https://doi.org/10.1038/s43018-020-0096-5>.
49. Christensen, E., Birkenkamp-Demtröder, K., Sethi, H., Shchegrova, S., Salari, R., Nordentoft, I., Wu, H.T., Knudsen, M., Lamy, P., Linskrog, S.V., et al. (2019). Early detection of metastatic relapse and monitoring of therapeutic efficacy by ultra-deep sequencing of plasma cell-free DNA in patients with urothelial bladder carcinoma. *J. Clin. Oncol.* 37, 1547–1557. <https://doi.org/10.1200/jco.18.02052>.
50. Reinert, T., Henriksen, T.V., Christensen, E., Sharma, S., Salari, R., Sethi, H., Knudsen, M., Nordentoft, I., Wu, H.T., Tin, A.S., et al. (2019). Analysis of plasma cell-free DNA by ultradeep sequencing in patients with stages I to III colorectal cancer. *JAMA Oncol.* 5, 1124–1131. <https://doi.org/10.1001/jamaoncol.2019.0528>.
51. Coombes, R.C., Page, K., Salari, R., Hastings, R.K., Armstrong, A., Ahmed, S., Ali, S., Cleator, S., Kenny, L., Stebbing, J., et al. (2019). Personalized detection of circulating tumor DNA antedates breast cancer metastatic recurrence. *Clin. Cancer Res.* 25, 4255–4263. <https://doi.org/10.1158/1078-0432.CCR-18-3663>.
52. Loupakis, F., Sharma, S., Derouazi, M., Murgioni, S., Biason, P., Rizzato, M.D., Rasola, C., Renner, D., Shchegrova, S., Koyen Malashevich, A., et al. (2021). Detection of molecular residual disease using personalized circulating tumor DNA assay in patients with colorectal cancer undergoing resection of metastases. *JCO Precis. Oncol.* 5, 1166–1177. <https://doi.org/10.1200/po.21.00101>.
53. Dhakal, B., Sharma, S., Balcioglu, M., Shchegrova, S., Malhotra, M., Zimmermann, B., Billings, P.R., Harrington, A., Sethi, H., Aleshin, A., and Hari, P.N. (2022). Assessment of molecular residual disease using circulating tumor DNA to identify multiple myeloma patients at high risk of relapse. *Front. Oncol.* 12, 786451. <https://doi.org/10.3389/fonc.2022.786451>.
54. Csardi, G., and Nepusz, T. (2006). The igraph software package for complex network research. *Int. J. Complex Syst.* 1695, 1–9.
55. Newman, M.E.J., and Girvan, M. (2004). Finding and evaluating community structure in networks. *Phys. Rev. E* 69, 026113. <https://doi.org/10.1103/PhysRevE.69.026113>.
56. Brandes, U., Delling, D., Gaertler, M., Gorke, R., Hoefer, M., Nikoloski, Z., and Wagner, D. (2008). On modularity clustering. *IEEE* 20, 172–188. <https://doi.org/10.1109/Tkde.2007.190689>.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Biological samples</b>		
Plasma samples	I-SPY2 Trial	<a href="https://clinicaltrials.gov/ct2/show/NCT01042379">https://clinicaltrials.gov/ct2/show/NCT01042379</a>
Tumor biopsy before treatment	I-SPY2 Trial	<a href="https://clinicaltrials.gov/ct2/show/NCT01042379">https://clinicaltrials.gov/ct2/show/NCT01042379</a>
<b>Critical commercial assays</b>		
Signatera™	Natera	<a href="https://www.natera.com/oncology/signatera-advanced-cancer-detection/">https://www.natera.com/oncology/signatera-advanced-cancer-detection/</a>
<b>Deposited data</b>		
Transcriptomic	Wolf et al. <sup>9</sup>	Gene Expression Omnibus (GEO) GSE194040 (mRNA); <a href="https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE194040">https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE194040</a>
Patient-level ctDNA and clinical data	This study	Table S1 ctDNA and clinical data, related to Figures 1 and 2.
<b>Software and algorithms</b>		
limma v.3.48.3	Ritchie et al. <sup>35</sup>	<a href="https://bioconductor.org/packages/release/bioc/html/limma.html">https://bioconductor.org/packages/release/bioc/html/limma.html</a>
fgsea v.1.18.	Korotkevich et al. <sup>36</sup>	<a href="https://bioconductor.org/packages/release/bioc/html/fgsea.html">https://bioconductor.org/packages/release/bioc/html/fgsea.html</a>
STRING v11.5	Szklarczyk et al. <sup>37</sup>	<a href="https://string-db.org/">https://string-db.org/</a>
stats R package (v.3.6.3)	R Core Team <sup>38</sup>	<a href="https://stat.ethz.ch/R-manual/R-devel/library/stats/html/stats-package.html">https://stat.ethz.ch/R-manual/R-devel/library/stats/html/stats-package.html</a>
survival R package (v.3.1–12)	Therneau and Grambsch <sup>39</sup>	<a href="https://CRAN.R-project.org/package=survival">https://CRAN.R-project.org/package=survival</a>

### RESOURCE AVAILABILITY

#### Lead contact

Further information and requests for resources or data should be directed to the I-SPY Data Access and Publications Committee coordinator ([ispy2dapc@quantumleaphealth.org](mailto:ispy2dapc@quantumleaphealth.org)).

#### Materials availability

This study did not generate new unique reagents.

#### Data and code availability

Transcriptomic data used in this study are available in NCBI's Gene Expression Omnibus (GEO) GSE194040 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE194040>). ctDNA and clinical data are available in Table S1 ctDNA and clinical data, Related to Figures 1 and 2. No original code was developed for this work.

### EXPERIMENTAL MODEL AND SUBJECT DETAILS

#### The I-SPY2 trial

I-SPY2 (NCT01042379; IND 105139) is an ongoing, open-label, multicenter adaptive, randomized phase 2 platform trial of neoadjuvant therapy for high-risk early-stage breast cancer. The overview, design, patient eligibility, and oversight of the trial have been previously described in detail.<sup>40–43</sup> I-SPY2 evaluates multiple experimental treatments against a common control arm, in parallel. The study stratifies patients within 8 subtypes based on HR, HER2, and MammaPrint scores, and the combinations of these subtypes define 10 biomarker signatures. An adaptive algorithm based on Bayesian probabilities of benefit vs. control is used to randomize within a subtype.<sup>40–42</sup> The primary endpoint of the trial is pCR (ypT0/is, ypN0), defined as no residual invasive cancer in either breast or lymph nodes, evaluated at the time of surgery. Adaptive randomization in I-SPY2 preferentially assigns patients to trial arms according to continuously updated Bayesian probabilities of pCR rates within each biomarker signature. The demonstration of statistical superiority in pCR rate (vs. the control arm) for any of the 10 pre-defined biomarker signatures determines the “graduation” of an experimental arm.

I-SPY2 is an intent-to-treat trial. Per protocol, patients who received study treatment but switched to non-protocol therapy, withdrew consent before surgery, or did not receive surgery were considered to have non-pCR. Because response in those situations cannot be attributed to the treatment the patient was initially randomized to, a non-pCR was assigned for efficacy and biomarker analyses. Response to study treatment (and non-protocol therapy) was also assessed at surgery using the residual cancer burden

(RCB) method.<sup>20,21</sup> While RCB is a pathology-based approach to measure the amount of residual invasive cancer in the breast and regional lymph nodes, pCR (per I-SPY2 protocol) is a combination of pathology and trial rules/administration geared to evaluating the efficacy of a specific treatment arm. For example, when a patient assigned to an experimental arm drops out of the trial protocol, the patient is considered to have non-pCR even if, by pathology, the patient eventually achieves pCR/RCB-0 using the RCB method.

In I-SPY2, the observed pCR rates are often lower than those reported by other studies. One contributing factor is that, per the I-SPY2 protocol (discussed above), any patient who receives non-protocol therapy is considered non-pCR. Another factor is the use of the RCB method for assessing pCR, which entails a more comprehensive evaluation of the resected surgical specimens and typically results in increased identification of cases with minimal residual disease (non-pCR).

## Patients

The ctDNA study was conducted in the context of the I-SPY2 trial. Patients had MammaPrint high tumors and therefore are at high risk of metastatic recurrence within five years after diagnosis.<sup>44</sup>

This study involved patients with HER2-negative tumors. Breast cancer is generally classified into one of the 4 receptor subtypes based on hormone receptor (HR; estrogen and progesterone receptor) and HER2 expression. The HER2-negative receptor subtypes—HR-positive/HER2-negative and HR-negative/HER2-negative (also known as triple-negative breast cancer or TNBC)—represent about 85% of all breast cancers [73% HR-positive/HER2-negative and 12% TNBC].<sup>29</sup> Receptor subtypes have distinct biological characteristics that are reflected in differences in clinical outcomes<sup>45</sup> and recommended treatment modalities in the early-stage setting<sup>3</sup>; for example, HR-positive/HER2-negative tumors are treated with endocrine therapy with or without chemotherapy, while HER2-positive tumors are treated with HER2-targeted drugs in combination with chemotherapy. Early-stage TNBC (stage II and higher) is most often treated with the checkpoint inhibitor, pembrolizumab, in the neoadjuvant setting in combination with chemotherapy, followed by pembrolizumab post-surgery.<sup>46</sup>

In the I-SPY2 trial, HER2-negative patients in the control arm of I-SPY2 received paclitaxel followed by AC. Investigational regimens are given in combination with paclitaxel or as a replacement for paclitaxel. The trial limits eligibility to women >18 years with stage II or III breast cancer >2.5cm and high MammaPrint score<sup>47</sup>; these are patients at high risk of metastatic recurrence within 5 years after diagnosis. The I-SPY2 protocol was approved by Institutional Review Boards at all participating institutions and all patients signed written informed consent. Patients included in the current study were HER2-negative, with pretreatment tumor biopsy specimens available and whose plasma samples were analyzed for ctDNA. These patients were enrolled in I-SPY2 between March 2010 and July 2018. All patients provided written informed consent for subsequent use of their specimens for research purposes.

Of the 283 evaluable patients, 26.5% (75) were stage T3/T4; 41.7% (118) were node-positive; 55.1% (156) had grade 3 disease; 57.6% (163) were MammaPrint (ultra) high-risk 2; and 48.1% (136) received standard NAC, while the rest received NAC combined with an investigational drug (Table 1). Clinicopathologic characteristics and treatment assignment were balanced between subtypes except for a significantly higher proportion of patients with grade 3 (68.1% vs. 42.8%) and MammaPrint (ultra) high-risk 2 tumors (88.4% vs. 28.3%), and RCB-0 (24.6% vs. 15.2%) in the TNBC group compared with the HR-positive/HER2-negative group (all Fisher's exact test  $p < 0.05$ ).

## Clinical samples

Blood samples for ctDNA analysis were collected at pretreatment (T0), 3 weeks after treatment initiation (T1), and 12 weeks after treatment initiation between paclitaxel-based treatment and anthracycline regimens (T2), and after NAC before surgery (T3). Tumor biopsy for whole exome sequencing (WES) was collected at pretreatment (T0). Cell-free DNA (cfDNA) isolated from plasma served as the input for ctDNA analysis. Pretreatment core biopsies were subjected to WES, and germline DNA isolated from buffy coats was also sequenced and used as matched normal control. Peripheral blood was collected into EDTA-containing tubes and centrifuged at 1100–1300 g for 20 min at room temperature. The buffy coat and plasma were then aspirated and immediately frozen. Pretreatment core biopsies were collected from the primary breast tumor and mounted in Tissue-Tek O.C.T. embedding media. A section from the frozen tissue was stained with hematoxylin and eosin (H&E) for pathologic evaluation to assess the percent tumor content. For tissue samples that met the tumor content requirement (30%), 8–10 cryosections of 30  $\mu$ m thickness were further collected as ribbons using a microtome into cryovials. All samples (buffy coat, plasma, and tissue) were immediately frozen and stored at  $-80^{\circ}\text{C}$  until further processing.

Clinical samples were collected and stored at participating I SPY 2 trial sites. Samples were then shipped overnight on dry ice to the I-SPY Laboratory at the University of California San Francisco for accessioning and storage at  $-80^{\circ}\text{C}$ . Buffy coat and tissue samples were shipped to a commercial vendor for WES and plasma samples were shipped to Natera, Inc. for ctDNA analysis.

## METHOD DETAILS

### Quality control and analytic dataset

Of the 295 patients with available pretreatment biopsies, 283 (96%) tumors were successfully analyzed by WES. In total, 1,024 plasma samples from 283 HER2-negative patients (145 HR-positive/HER2-negative and 138 TNBC) with WES data comprised the final analytic dataset (Figure S1A). This dataset includes ctDNA data from 65 HER2-negative patients (223 plasma samples) from previously published work.<sup>15</sup> Of the 283 patients, 251 were part of the I-SPY2-990 mRNA/RPPA Data Resource<sup>9</sup> with pretreatment gene expression.

### ctDNA analysis

ctDNA was analyzed in plasma collected at pretreatment (T0), 3 weeks after treatment initiation (T1), 12 weeks after treatment initiation between paclitaxel-based treatment and AC regimens (T2), and before surgery (T3) using the Signatera™ test, a bespoke multiplex polymerase chain reaction (PCR) next generation sequencing-based assay. The test detects up to 16 patient-specific somatic mutations selected from WES data of pretreatment biopsies. The methods for ctDNA analysis, including cfDNA extraction, quantification, library preparation, provenance testing, WES workflow, and bioinformatics pipeline, have been previously described in detail.<sup>15</sup> Briefly, WES data derived from pretreatment core biopsies and matched normal blood samples from each patient were analyzed to select a set of 16 personalized (patient-specific), somatic, clonal, single nucleotide variants (SNVs) for multiplex PCR testing.<sup>15,22,48–51</sup> The multiplex PCR primers were designed and synthesized to track ctDNA in a patient's plasma. Plasma samples with at least 2 of the 16 variants detected were defined as ctDNA-positive. ctDNA concentration was reported as mean tumor molecules (MTM) per mL of plasma.<sup>15</sup> The MTM per mL of plasma was calculated by dividing the total number of mutant molecules by the number of targets detected. The number of mutant tumor molecules per mL of plasma was calculated as follows:

$$\text{cfDNA extracted (ng)} \times 1000 \text{ pg} \times \text{haploid genome equivalent (hGE)} \times \text{Variant Allele Fraction (VAF)} \div \\ 1 \text{ ng} \times 3.3 \text{ pg per hGE} \times \text{Plasma volume for extraction (mL)}$$

The personalized and tumor-informed ctDNA test used in this study has been clinically validated in breast cancer<sup>15,25,51</sup> and other cancer types.<sup>22,52,53</sup> Bratman and colleagues used the same ctDNA technology to show that pretreatment ctDNA concentration and ctDNA dynamics were significantly correlated with response and survival in advanced cancer patients treated with pembrolizumab.<sup>48</sup> Others have demonstrated that the same ctDNA test can guide patient selection to identify urothelial cancer patients who will benefit from adjuvant atezolizumab.<sup>22</sup>

### Gene expression analysis

Full transcriptome gene expression data from pretreatment biopsy were generated for each I-SPY2 patient on an Agilent microarray. The expression data were used to calculate the MammaPrint Score, which is required to assess patient eligibility for the trial. In a recent publication from our group, gene expression data were used to classify each patient into one of the 5 response-predictive subtypes.<sup>9</sup> Here, we used the gene expression data to discover genes and pathways associated with pretreatment ctDNA positivity.

Of the 283 patients, 251 had tumor gene expression (GSE194040)<sup>9</sup> and ctDNA data at pretreatment. Gene expression between ctDNA-positive and ctDNA-negative patients at pretreatment for each subtype: HR-positive/HER2-negative (n=131, 66% ctDNA-positive) and triple-negative breast cancer (TNBC, n=120, 91% ctDNA-positive) was evaluated independently for differential expression (limma v3.48.3),<sup>35</sup> gene set enrichment (fgsea v1.18.00),<sup>36</sup> and protein-protein interaction (STRING v11.5).<sup>37</sup>

## QUANTIFICATION AND STATISTICAL ANALYSIS

This study examined the clinicopathologic and molecular correlates of ctDNA positivity. Given the differences in the biology and clinical histories between HR-positive/HER2-negative and TNBC, we hypothesized that the predictive and prognostic value of ctDNA may vary between the two subtypes. The analyses performed in the study are outlined in [Figure S1B](#).

### Statistical analysis of the clinical and gene expression data

The response endpoints used in the study were pCR and RCB, the latter of which is divided into 4 classes: RCB-0 (no residual disease equivalent to pCR), RCB-I (minimal burden), RCB-II (moderate burden), and RCB-III (extensive burden). The survival endpoint was distant recurrence-free survival (DRFS), defined as the time interval between the date of patient consent for treatment and the date of clinical diagnosis of metastatic recurrence or death by any cause. A  $p < 0.05$  was considered significant for the statistical tests described below.

#### ctDNA vs. clinicopathologic characteristic

Association between ctDNA status and categorical variables was assessed using Fisher's exact test. For continuous clinical variables, a t-test or analysis of variance (ANOVA) was performed to compare means between groups. The ctDNA concentration (MTM/mL) was compared across clinicopathologic groupings using the Wilcoxon rank-sum test (2 groups) or the Kruskal-Wallis test (3 or more groups).

#### ctDNA vs. response

A Fisher's exact test was used to compare the proportions of patients who achieved pCR vs. those who did not, or by RCB class, stratified based on their ctDNA dynamics. To assess the association between ctDNA dynamics and response, patients who tested ctDNA-positive at T0 were grouped according to ctDNA status at T0, T1, and T2: early clearance at T1 (3 weeks after initiation of treatment, ctDNA+/-/-), late clearance at T2 (at 12 weeks, after paclitaxel-based treatment, ctDNA+/-/-) and no clearance at T2 (ctDNA +/+/+). We also used logistic regression to estimate odds ratios and 95% confidence intervals of achieving a pCR (or RCB-0/I).

#### ctDNA vs. survival

Of the 283 patients, 272 (96.1%) had DRFS data, of whom 74 (27.2%) experienced metastatic recurrence or death. The median follow-up was 3.10 years (range 0.46–7.6) for the HR-positive/HER2-negative group and 3.12 years (range 0.31–7.91) for the TNBC group. In the HR-positive/HER2-negative group, 32/142 (22.5%) experienced a DRFS event with 28 distant recurrences and 4 deaths, while in the TNBC group, 42/130 (32.3%) experienced a DRFS event with 33 distant recurrences and 9 deaths.

We investigated whether ctDNA positivity at different time points was associated with DRFS. We also examined whether ctDNA status after NAC (T3) could further refine risk stratification by pCR and RCB class (see main text).

Furthermore, we examined the prognostic significance of ctDNA dynamics in each subtype. Patients with available ctDNA data for all 4 time points [115 (79.3%) HR-positive/HER2-negative and 88 (63.8%) TNBC patients] were classified into 5 groups based on ctDNA dynamics: Group 1 includes patients who were ctDNA negative at T0 and remained negative until T3; Groups 2, 3, and 4 include patients who were ctDNA-positive at T0 and cleared their ctDNA at T1, T2, or T3, respectively; and Group 5 includes those who were positive at T0 and did not clear ctDNA at T3. We excluded 2 patients from each subtype from the analysis because they did not fall into one of the 5 groups based on ctDNA dynamics. For the HR-positive/HER2-negative subtype, one patient was ctDNA- $-/+/-/-$  and the other ctDNA- $-/-/-/+$ ; and for the TNBC subtype: 2 patients were ctDNA- $-/-/+/-$ . A total of 110 of 115 HR-positive/HER2-negative and 83 of 88 TNBC patients had follow-up data for survival analysis.

Hazard ratios and 95% confidence intervals were estimated using Cox proportional hazards regression analysis. Survival curves were plotted using Kaplan-Meier analysis, and p values were calculated using a log rank test. In multivariable analysis, we chose pCR as a covariate based on recent findings in the I-SPY2 trial, showing a strong prognostic impact of pCR on survival in neoadjuvant-treated patients.<sup>4</sup> The R package “survival” was used for Cox proportional hazards model, Kaplan-Meier survival analysis, and log rank tests.

### **Differential expression analysis**

Differential expression analysis on the global transcriptome data ( $m = 19,134$  genes, GSE194040)<sup>9</sup> between ctDNA-positive and ctDNA-negative patients at pretreatment was performed in limma (v3.48.3) with no covariates.<sup>35</sup> Additionally, to explore the impact of proliferation in the HR-positive/HER2-negative subtype, differential expression analysis was also performed with adjustment for MammaPrint status.

Linear modeling was performed on the log<sub>2</sub>-transformed expression data; the empirical Bayes moderation of computed statistics was applied with intensity-trend allowed for the prior variance. Significantly differentially expressed (DE) genes were defined as those with BH adjusted  $p < 0.05$ . For TNBC subtypes with unbalanced groups, we also report genes with nominally significant DE at  $p < 0.05$ . Genes with log<sub>2</sub> fold changes (lfc) were upregulated in ctDNA-positive patients, and those with negative lfc were upregulated in ctDNA-negative patients. Genes were considered commonly DE between HR-positive/HER2-negative and TNBC subtypes if they were at least nominally DE ( $p < 0.05$ ) in each subtype and had concordant direction of log<sub>2</sub> fold changes.

DE genes were visualized via hierarchical clustering heatmaps (pheatmap v1.0.12) using hclust Ward's clustering criterion (ward.D2) agglomerative method with Euclidean distances as the distance metric. Protein-protein interactions (PPI) and network analysis were performed using the STRING database (v.11.5, <https://string-db.org/>) with minimum interaction confidence of 0.4. PPI network was visualized in igraph (v.1.2.6),<sup>54</sup> and only the largest fully connected main network is visualized. To identify gene communities—defined as natural divisions of densely connected subgroups or community structures determined algorithmically<sup>55</sup>—community detection was performed on this main network using igraphcluster\_optimal algorithm.<sup>56</sup> Each community was analyzed in STRING for functional network enrichment (FDR  $p < 0.05$ ), and representative KEGG pathways were annotated.

### **Gene set enrichment analysis**

Gene set enrichment analysis (GSEA) of Molecular Signatures Database Hallmark (H,  $m = 50$ ) gene sets was performed in fgsea (v1.18.00) in HR-positive/HER2-negative and TNBC subtypes.<sup>36</sup> Fast GSEA was performed on ordered DE t-statistics using 1000 permutations with minimum and maximum gene set sizes set to 15 and 500, respectively. Significantly enriched gene sets were defined as those with enrichment BH adjusted at  $p < 0.05$ ; nominally enriched gene sets were defined at  $p < 0.05$ . Gene sets with positive normalized enrichment scores (NES) were enriched in ctDNA-positive patients, and those with negative NES were enriched in ctDNA-negative patients. Enrichment plots were generated for immune-associated gene sets with at least nominal enrichment ( $p < 0.05$ ) in either subtype. Leading-edge genes—core member genes that contribute to enrichment score—belonging to these immune-associated gene sets were visualized via hierarchical clustering heatmaps (pheatmap v1.0.12) using hclust Ward's clustering criterion (ward.D2) agglomerative method with Euclidean distances as the distance metric.

## **ADDITIONAL RESOURCES**

Patient-level data are available in [Table S1](#) ctDNA and clinical data, Related to [Figures 1](#) and [2](#).

**Supplemental information**

**Clinical significance and biology of circulating  
tumor DNA in high-risk early-stage HER2-negative  
breast cancer receiving neoadjuvant chemotherapy**

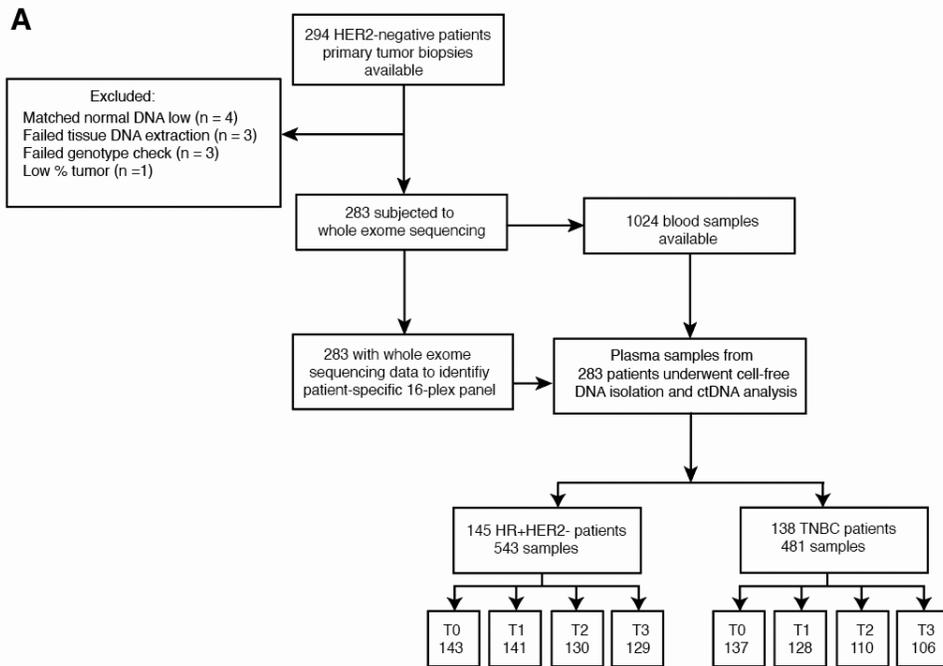
**Mark Jesus M. Magbanua, Lamorna Brown Swigart, Ziad Ahmed, Rosalyn W. Sayaman, Derrick Renner, Ekaterina Kalashnikova, Gillian L. Hirst, Christina Yau, Denise M. Wolf, Wen Li, Amy L. Delson, Smita Asare, Minetta C. Liu, Kathy Albain, A. Jo Chien, Andres Forero-Torres, Claudine Isaacs, Rita Nanda, Debu Tripathy, Angel Rodriguez, Himanshu Sethi, Alexey Aleshin, Matthew Rabinowitz, Jane Perlmutter, W. Fraser Symmans, Douglas Yee, Nola M. Hylton, Laura J. Esserman, Angela M. DeMichele, Hope S. Rugo, and Laura J. van 't Veer**

## **Supplemental Information**

### **The clinical significance and biology of circulating tumor DNA in high-risk early-stage HER2-negative breast cancer receiving neoadjuvant chemotherapy**

**Authors:** Mark Jesus M. Magbanua, Lamorna Brown-Swigart, Ziad Ahmed, Rosalyn W. Sayaman, Derrick Renner, Ekaterina Kalashnikova, Gillian L. Hirst, Christina Yau, Denise M. Wolf, Wen Li, Amy L. Delson, Smita Asare, Minetta C. Liu, Kathy Albain, A. Jo Chien, Andres Forero-Torres, Claudine Isaacs, Rita Nanda, Debu Tripathy, Angel Rodriguez, Himanshu Sethi, Alexey Aleshin, Matthew Rabinowitz, Jane Perlmutter, W. Fraser Symmans, Douglas Yee, Nola M. Hylton, Laura J. Esserman, Angela M. DeMichele, Hope S. Rugo, and Laura J. van 't Veer

## Supplementary Figures

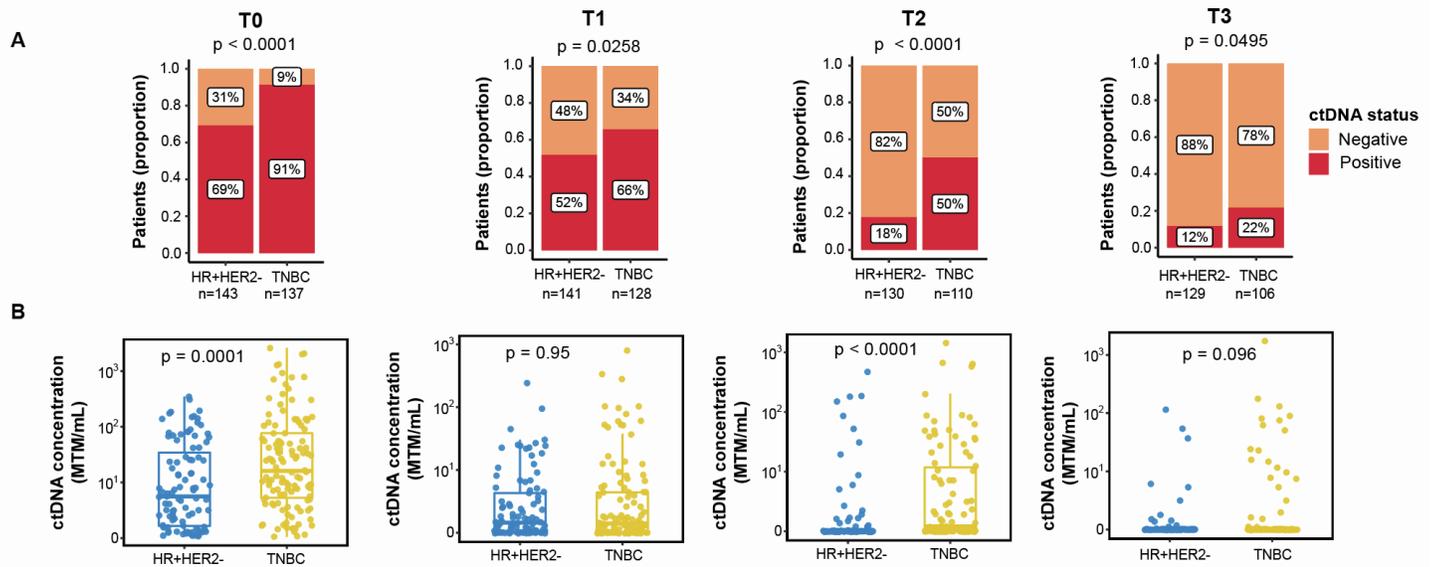


**B**

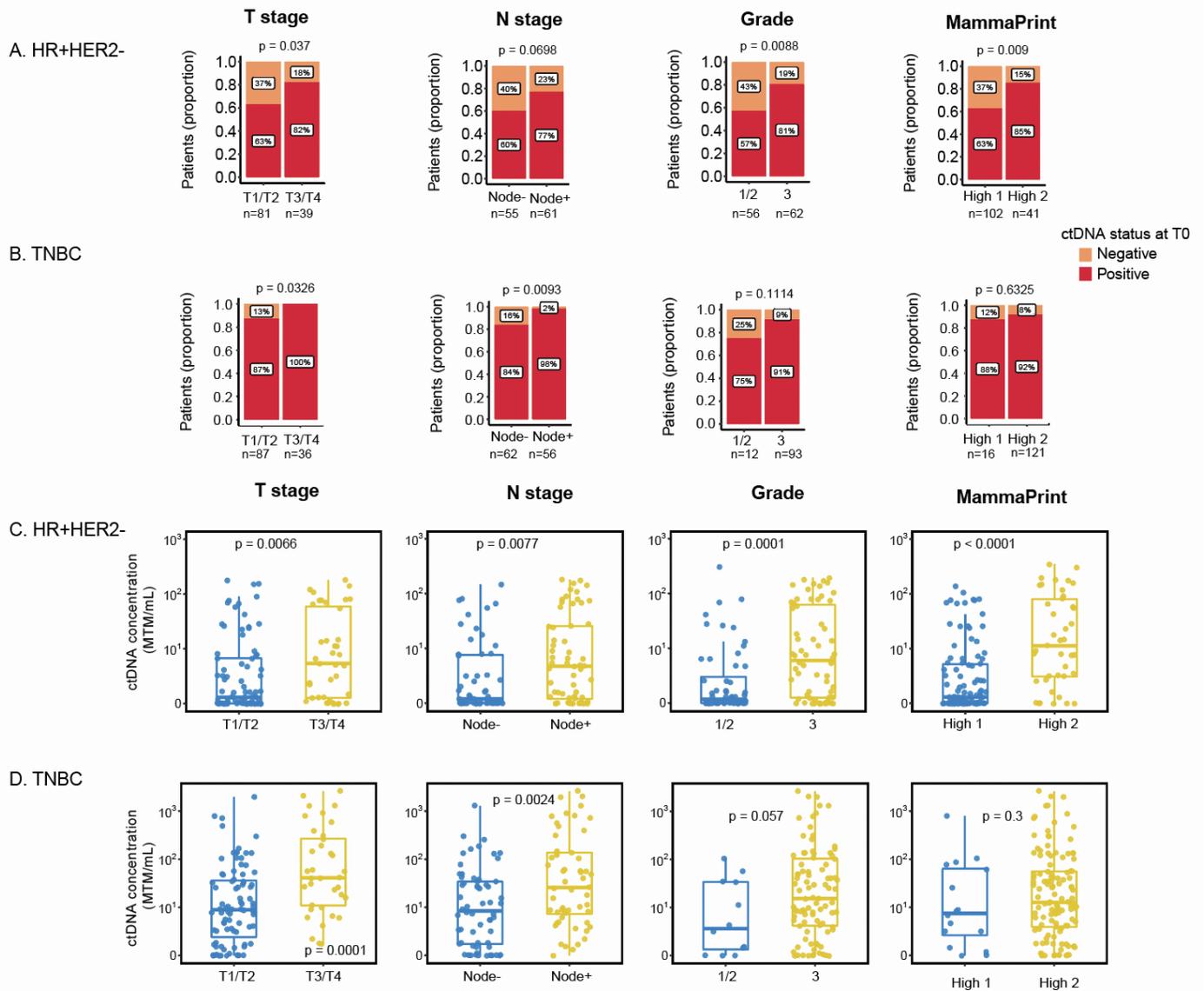
Analysis performed to evaluate association between		Sample collection (Time from treatment start)				Figure #
		T0 (0 weeks)	T1 (3 weeks)	T2 (12 weeks)	T3 (24 weeks)	
ctDNA status (+/-)	HR+HER2- vs. TNBC					Fig. S2A
	Clinicopathologic variables					Fig. S3A-B
	Gene expression					Fig. 3; Fig. S6
	Survival (DRFS)					Fig. S4
	Survival (DRFS) ~ ctDNA + pCR					Fig. S5
ctDNA status (+/-) and response (pCR, RCB)	Survival (DRFS)					Fig. 2A-D
ctDNA concentration (MTM/mL)	HR+HER2- vs. TNBC					Fig. S2B
	Clinicopathologic variables					Fig. S3C-D
	Response (pCR, RCB)					Fig. 1A-D
ctDNA dynamics	Response (pCR, RCB)	→				Fig. 1E-J
	Survival (DRFS)	→				Fig. 2E-H

- Understanding biology of ctDNA shedding
- Evaluating prognostic value of ctDNA
- Evaluating predictive value of ctDNA

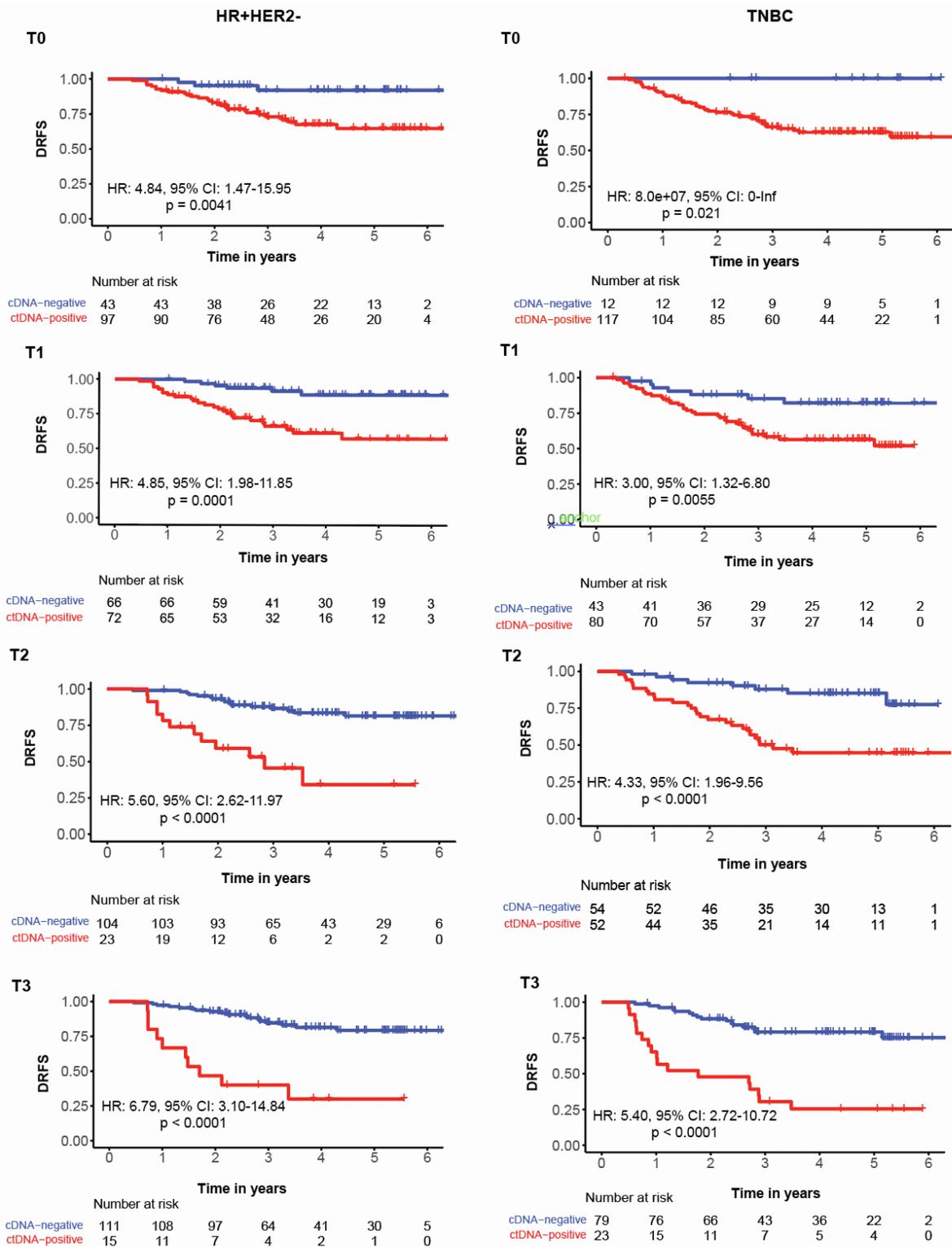
**Figure S1. Patients, samples, and analytical approach,** Related to Table 1 and Figures 1, 2, and 3. **A.** Flow chart showing the number of evaluable patients and samples used for the final analytic cohort consisting of 145 hormone receptor-positive/HER2-negative (HR+HER2-) and 138 triple-negative breast cancer (TNBC) patients; **B.** Sample collection time points and analytical approach. ctDNA was analyzed in plasma collected at pretreatment (T0), 3 weeks after treatment initiation (T1), 12 weeks after treatment initiation between paclitaxel-based treatment and anthracycline regimens (T2), and after NAC before surgery (T3). The response endpoints were pathologic complete response (pCR) and residual cancer burden (RCB). The survival endpoint was distant recurrence-free survival (DRFS).



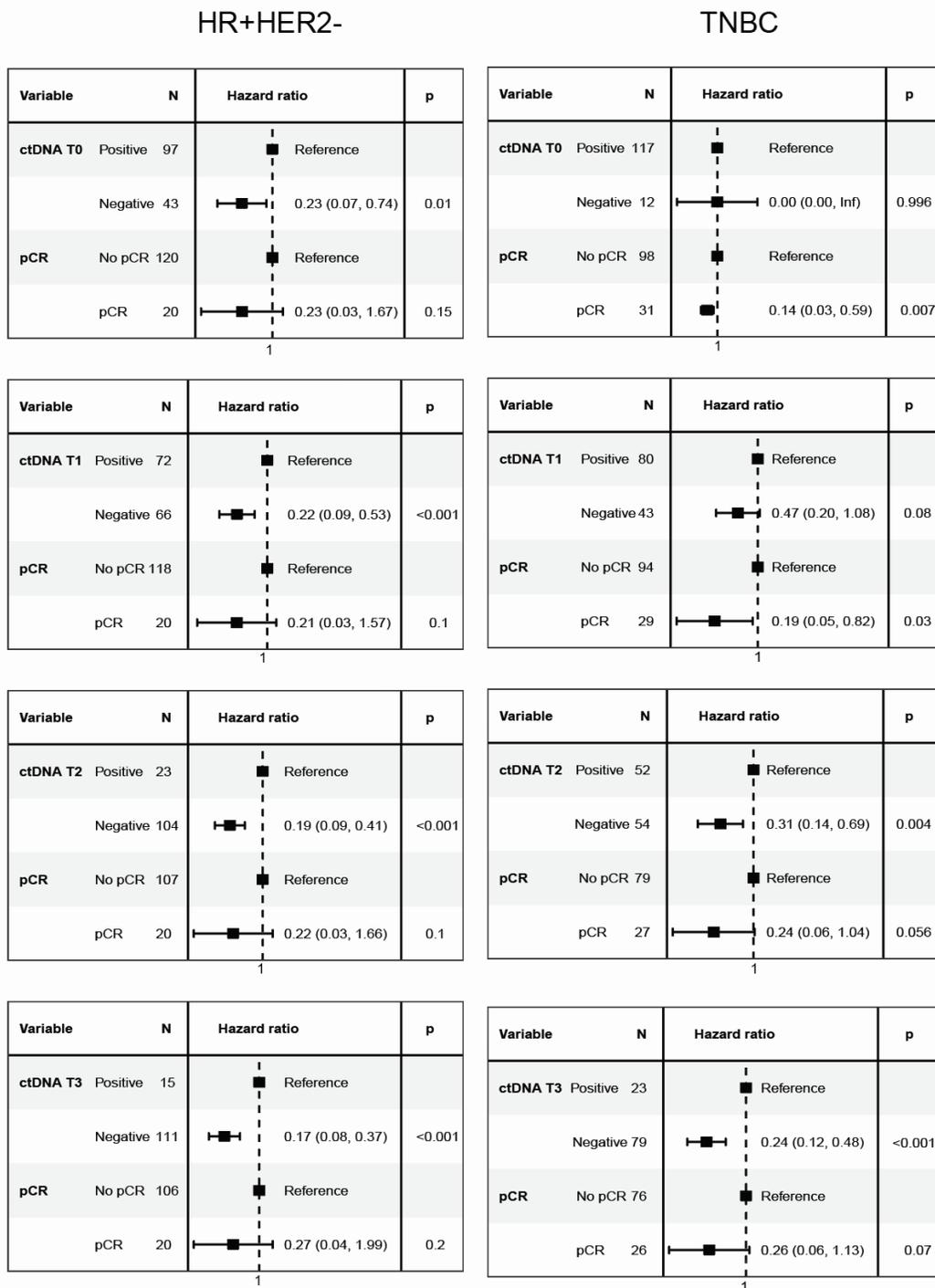
**Figure S2. Comparison of ctDNA positivity and ctDNA concentrations before, during, and after neoadjuvant chemotherapy (NAC) in plasma of HER2-negative patients, Related to Figure 1. A.** Percent ctDNA positivity and **B.** ctDNA concentration (expressed as mean tumor molecules per ml of plasma or MTM/ml) across all time points in patients with hormone receptor-positive/HER2-negative (HR+HER2-) and triple-negative breast cancer (TNBC). ctDNA was analyzed in plasma collected at pretreatment (T0), 3 weeks after treatment initiation (T1), 12 weeks after treatment initiation between paclitaxel-based treatment and anthracycline regimens (T2), and after NAC before surgery (T3).



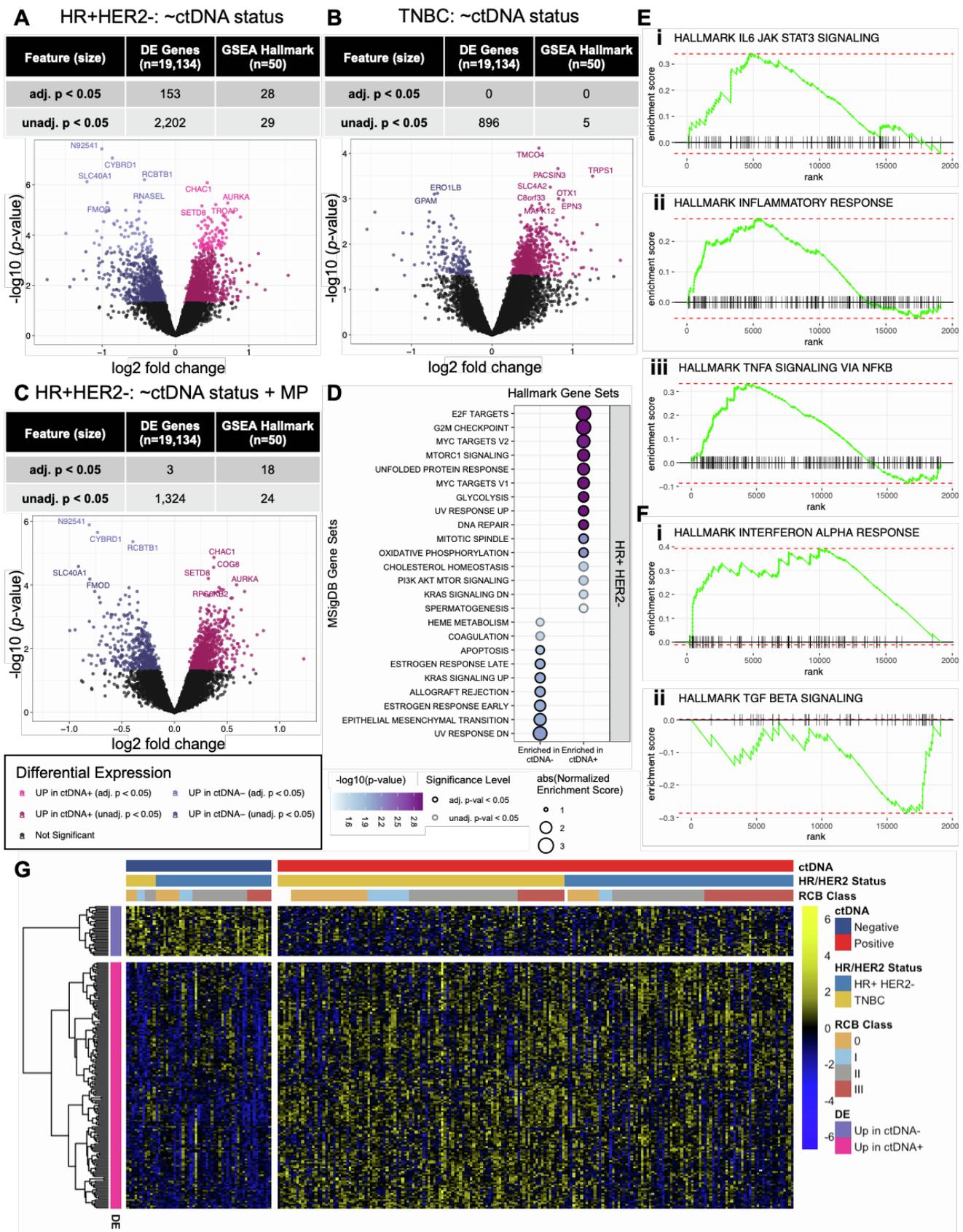
**Figure S3. Association between pretreatment ctDNA (T0) and clinicopathologic variables,** Related to Table 1. Association of ctDNA positivity in **A.** hormone receptor-positive/HER2-negative subset (HR+HER2-) and **B.** triple-negative breast cancer (TNBC) groups with clinicopathologic variables (dichotomized). P values were calculated using Fisher's exact test. The distribution of ctDNA concentration (MTM/ml) in ctDNA-positive patients with **C.** HR+HER2- and **D.** TNBC patients grouped according to dichotomized clinicopathologic variables. P values were calculated using the Wilcoxon rank-sum test.



**Figure S4. ctDNA positivity across all time points during neoadjuvant chemotherapy (NAC) is a significant negative prognostic factor for distant recurrence-free survival (DRFS), Related to Figure 2.** Hormone receptor-positive/HER2-negative (HR+HER2-, left panel) and triple-negative breast cancer (TNBC, right panel) patients were stratified according to ctDNA status (ctDNA-positive vs. ctDNA-negative) at each time point: T0 (pretreatment); T1 (3 weeks after treatment initiation); T2 (12 weeks after treatment initiation between paclitaxel-based treatment and anthracycline regimens); and T3 (after NAC before surgery). Hazard ratios (HR) and 95% confidence intervals (CI) were estimated from Cox proportional hazards regression models.



**Figure S5. Correlation of ctDNA positivity after neoadjuvant chemotherapy (NAC) and pathologic complete response (pCR) with distant recurrence-free survival, Related to Figure 2. Forest plots showing hazard ratio and 95% confidence intervals estimated from bivariate Cox proportional hazards regression models with predictors, ctDNA after NAC (T3) adjusting for the effects of pathologic complete response (pCR) in patients with hormone receptor-positive/HER2-negative (HR+HER2-, left panel) and triple-negative breast cancer (TNBC, right panel).**



**Figure S6. Differentially expressed genes and enriched gene sets associated with ctDNA shedding at pretreatment, Related to Figure 3. A-C.** Differentially expressed genes between ctDNA-positive and ctDNA-negative patients at pretreatment within each breast cancer subtype (HR-positive/HER2-negative n=131, 66%

ctDNA-positive; TNBC n=120, 91% ctDNA-positive). (**Top**) The number of significant [Benjamini-Hochberg (BH) adjusted  $p < 0.05$ ] and nominally significant ( $p < 0.05$ ) differentially expressed genes and enriched MSigDB Hallmark gene sets between in ctDNA-positive and ctDNA-negative patients and (**Bottom**) volcano plots showing differentially expressed  $\log_2$  fold changes vs.  $-\log_{10}$  p-values in **A.** HR-positive/HER2-negative (HR+HER2-), **B.** triple-negative breast cancer (TNBC), and **C.** HR+HER2- subtype adjusted for MammaPrint (MP) status. The top 10 differentially expressed genes are annotated; **D.** GSEA of MSigDB Hallmark gene sets in HR+HER2- subtype adjusting for MP status. Gene sets with significant (BH adjusted  $p < 0.05$ ) and nominally significant ( $p < 0.05$ ) enrichment in ctDNA-positive or ctDNA-negative patients are depicted with p-value significance and normalized enrichment scores (NES). GSEA enrichment plots for immune-associated gene sets ( $p < 0.05$ ): **E.** (i) IL6 JAK STAT3 signaling, (ii) inflammatory response, and (iii) TNF-alpha signaling via NF- $\kappa$ B in HR+HER2-; **F.** (i) Interferon alpha response, and (ii) TGFB signaling in TNBC. **G.** Gene expression heatmap of the 190 genes commonly differentially expressed ( $p < 0.05$ ) in both HR+HER- and TNBC with concordant fold change direction across subtypes.