nature communications

Article

https://doi.org/10.1038/s41467-025-57293-9

Determinants of response and molecular dynamics in HER2+ER+ breast cancers from the NA-PHER2 trial receiving HER2-targeted and endocrine therapies

Received: 6 March 2024

Accepted: 12 February 2025

Published online: 04 March 2025

Check for updates

Maurizio Callari[1], Matteo Dugo[2], Marco Barreca[1,3], Balázs Győrffy[4,5,6], Barbara Galbardi[2], Lucia Vigano[2], Alberta Locatelli[2], Chiara Dall'Ara[2], Marina Ferrarini[2], Giancarlo Bisagni[7], Marco Colleoni[8], Mauro Mansutti[9], Claudio Zamagni[10], Lucia Del Mastro[11,12], Stefania Zambelli[2], Antonio Frassoldati[13], Olivia Biasi[8], Lajos Pusztai[14], Pinuccia Valagussa[1], Giuseppe Viale[1,8], Luca Gianni[1,16] & Giampaolo Bianchini[2,15,16]

Improved outcomes in HER2+ female breast cancer have resulted from chemotherapy and anti-HER2 therapies. However, HER2+ER+ cancers exhibit lower response rates. The phase 2 NA-PHER2 trial (NCT02530424) investigated chemo-free preoperative HER2 blockade (trastuzumab + pertuzumab) and CDK4/6 inhibition (palbociclib) with or without endocrine therapy (fulvestrant) in HER2+ER+ breast cancer. Clinical endpoints (i.e. Ki67 dynamics and pathological complete response) were previously reported. Here we report on the biomarker analysis, secondary objective of the study. Through RNA sequencing and tumour infiltrating lymphocytes (TIL) assessment in serial biopsies, we identified biomarkers predictive of pCR or Day14 Ki67 response and unveiled treatment-induced molecular changes. High immune infiltration and low ER signalling correlated with pCR, while TP53 mutations associated with high Day14 Ki67. Stratification based on Ki67 at Day14 and at surgery defined three response groups (Ki67 HighHigh, LowHigh, LowLow), with divergent tumour and stroma expression dynamics. The HighHigh group showed dysfunctional immune infiltration and overexpression of therapeutic targets like PAK4 at baseline. The LowLow group exhibited a Luminal A phenotype by the end of treatment. This study expands our understanding of drivers and dynamics of HER2+ER+ tumour response, towards treatment tailoring.

The prognosis of Human Epidermal Growth Factor Receptor-2 positive (HER2 +) breast cancers has improved substantially with the introduction of HER2-targeted neoadjuvant/adjuvant therapies[1]. The overall most effective regimens include multi-drug chemotherapy combined with dual HER2 blockade. Among HER2+ cancers, large scale molecular and clinical differences remain between the estrogen receptor-positive (ER +) and -negative (ER-) subgroups[2]. HER2+ER+ cancers have lower pathological complete response (pCR) rates to

A full list of affiliations appears at the end of the paper. e-mail: luca.gianni@fondazionemichelangelo.org; bianchini.giampaolo@hsr.it

neoadjuvant therapy than HER2+ER- tumours[3–6]. Therefore, they represent the highest proportion of HER2+ cancers that are candidate to adjuvant T-DM1 after neoadjuvant therapy, to which, however, they respond similarly to ER- tumours[7].

ER positivity only partially captures the variability in response to treatment; other aspects of tumour biology (e.g. proliferation) and its microenvironment (e.g. immune infiltration) have a relevant role[8]. Consequently, biomarker-guided alternative approaches that could improve patients' outcome while reducing unnecessary toxicities are needed in the HER2+ER+ subgroup.

In HER2+ER+ cancers, HER2 targeted therapies are additive/synergistic with endocrine therapy, and response rates can be further improved with the addition of CDK4/6 inhibitors[9,10]. In the MonarchHER randomized Phase II trial, abemaciclib plus trastuzumab plus fulvestrant demonstrated improved progression free survival compared to chemotherapy plus trastuzumab in metastatic HER2+ER+ breast cancer[11]. In the two-cohorts NA-PHER2 Phase II neoadjuvant trial (NCT02530424) we tested the efficacy of HER2 and CDK4/6 inhibition in combination with endocrine therapy (Fulv cohort: fulvestrant + palbociclib + trastuzumab + pertuzumab)[12], or without endocrine therapy (NoFulv cohort: palbociclib + trastuzumab + pertuzumab)[13]. The primary endpoint was the change in Ki67 expression at Day14[14,15], and secondary endpoints included clinical response and pCR rates. In the Fulv cohort, geometric mean Ki67 levels decreased from 31.9% at baseline to 4.3% at week 2, and the pCR rate was 27%[12]. In the NoFulv cohort, Ki67 expression also decreased from 33.4% to 5.5%, and the pCR rate was 19.2%[13]. These pCR rates are similar to what was observed in HER2+ER+ cancers treated with dual HER2 blockade and docetaxel in the NeoSphere trial[16].

The heterogeneity in clinical responses indicate inter-tumour heterogeneity. The identification of tumour ecosystem features that drive response or resistance is required for more effective patient selection for therapy.

In this study, we use serial tumour biopsies at baseline, Day14, and after completion of all neoadjuvant therapy from the NA-PHER2 trial to discover predictive biomarkers and study the changes in the tumour and its microenvironment during treatment. We performed whole transcriptome RNA sequencing and histologic assessment of tumour infiltrating lymphocytes (TIL), on top of existing Ki67 immunohistochemistry assessment. Our goal was to identify predictive markers for pCR and for Day14 Ki67 response and describe response-dependent gene expression changes. Only few studies have examined longitudinal transcriptomic and immune marker changes during neoadjuvant treatment of breast cancer[17–21] and to our knowledge, this is the first such analysis in HER2+ER+ breast cancers receiving anti-HER2 and a CDK inhibitor.

Results

Molecular analysis of NA-PHER2 trial

As part of the secondary objectives of the study, we conducted molecular analyses on samples collected during the NA-PHER2 trial. Transcriptomic profiles were generated from 143 samples (baseline $n = 53$, Day14 $n = 49$, Surgery $n = 41$) corresponding to 55 of the 58 patients enrolled in the NA-PHER2 trial (94.8%) (Supplementary Table 1, Fig. 1A and Supplementary Fig. 1A, B). Histological assessment of stromal (sTILs) and intra-tumoral infiltrating lymphocytes (iTILs), and tumour cell Ki67 positivity were determined centrally (see Methods Section). We verified that no major unbalances were present between the two non-randomised cohorts at baseline (Supplementary Fig. 1C–F). Figure 1B shows the expression pattern of 904 most variant and highly expressed genes across the dataset, annotated with available clinico-pathological features. The main biological processes with well characterised roles in breast cancer biology are clearly distinguishable, including: (i) ER-regulated genes (e.g. *ESR1, PGR, GATA3*), (ii) HER2-amplicon associated genes, (iii) proliferation related genes (e.g. *MKI67, AURKA, TOP2A*), (iv) immune infiltration (e.g. *CD3D, GZMB, CXCL10*), (v) interferon signalling associated genes (e.g. *ISG15, OAS1*), and (vi) extracellular matrix related genes (e.g. *SPARC*, *MMP2*). Sample clustering highlighted two main groups, with the smallest mostly including Day14 and Surgery samples characterised by low expression of proliferation genes and high expression of stromal/EMC genes (Fig. 1B). There were good agreements between histologic Ki67 level and proliferation related genes expression as well as sTILs and immune gene expression, confirming the high quality of the transcriptomic data (Supplementary Fig. 1G, H).

Tumour and TME features associated with pCR

We analysed IHC and transcriptomic profiles to identify the biological processes associated with pCR in the overall population and in the Fulv and NoFulv cohorts separately.

Genes associated with pCR were identified by differential analysis. We considered either the baseline or Day14 gene expression profiles (Fig. 2A, Supplementary Fig. 2A, B, Supplementary Fig. 3). Differential analysis was complemented with a geneset enrichment analysis (Fig. 2B and Supplementary Fig. 4A) and a gene to cell type annotation. Such annotation was obtained by interrogating previously published single-cell transcriptomics data of HER2+ clinical specimens[22]. As detailed in the methods, the goal was to identify, whenever possible, the cell type prevalently expressing each gene (Fig. 2A and Supplementary Fig. 2C, D).

We found a clear positive association between high levels of immune-related genes and achievement of pCR. This association was qualitatively independent of the treatment cohort and the timepoint analysed, although stronger at baseline. Among the top up-regulated genes associated with pCR, we found markers of cytotoxic T cells (e.g. *CD8B, PRF1, GZMB*), B cells (e.g. *CD79B, CD19*), chemokines (e.g. *CXCL10, CXCL11*) and interferon signalling (e.g. *OASL, IFIH1*) (Fig. 2A). Many of the up-regulated genes were prevalently expressed by myeloid and lymphoid cells at both time points. A relevant exception was *MUCL1*, a breast cancer specific gene[23,24] prevalently expressed by tumour cells (Fig. 2A). *MUCL1* expression is promoted by HER2 through the PI3K/Akt/mTOR pathway and HER2+ breast cancer cells dependent on *MUCL1* for proliferation[24]. Mesenchymal genesets were also enriched in the pCR group, while proliferation-related genesets showed a different pattern of association in the two cohorts. However, evaluation of key proliferation marker genes highlighted modest differences between the pCR and RD groups, not reaching statistical significance, except for *AURKA* in the NoFulv cohort (Supplementary Fig. 4B). Consequently, while tumour proliferation could play some role in response to treatment, this is not a major driver in our data.

Higher expression of *ESR1* and estrogen-related genes (HALLMARK_ESTROGEN_RESPONSE_EARLY geneset) were associated with residual disease at surgery (RD) in the pooled cohort and showed the same trend in both treatment cohorts separately (Fig. 2A, B).

The simple combination of genes taken from the Immune-axis and the ER-axis could strongly discriminate cancers with pCR versus RD in the overall cohort. Indeed, the categorised expression of *ESR1* and *B2M* genes could identify an *ESR1*-low *B2M*-high group with 72.7% pCR rate and an *ESR1*-high *B2M*-low group with as low as 5.3% pCR rate (Fig. 2C). Several other gene combinations from the same biological axes could achieve similar discriminating power (Supplementary Fig. 4C).

Next, we used baseline and Day14 gene expression profiles to train regularised logistic regression models predicting pCR and selected the most parsimonious model with the least features using 10-fold cross-validation. Cross-validation procedure was repeated 100 times and model performance quantified using the area under the ROC curve (AUC) (Fig. 2D and Supplementary Fig. 5A). The best AUC values were obtained using baseline data (mean baseline AUC = 0.86, Day14 AUC = 0.61; $p < 2e\text{-}16$). In line with the results from the differential analysis, this suggested that, in this clinical context, transcriptomic

https://doi.org/10.1038/s41467-025-57293-9

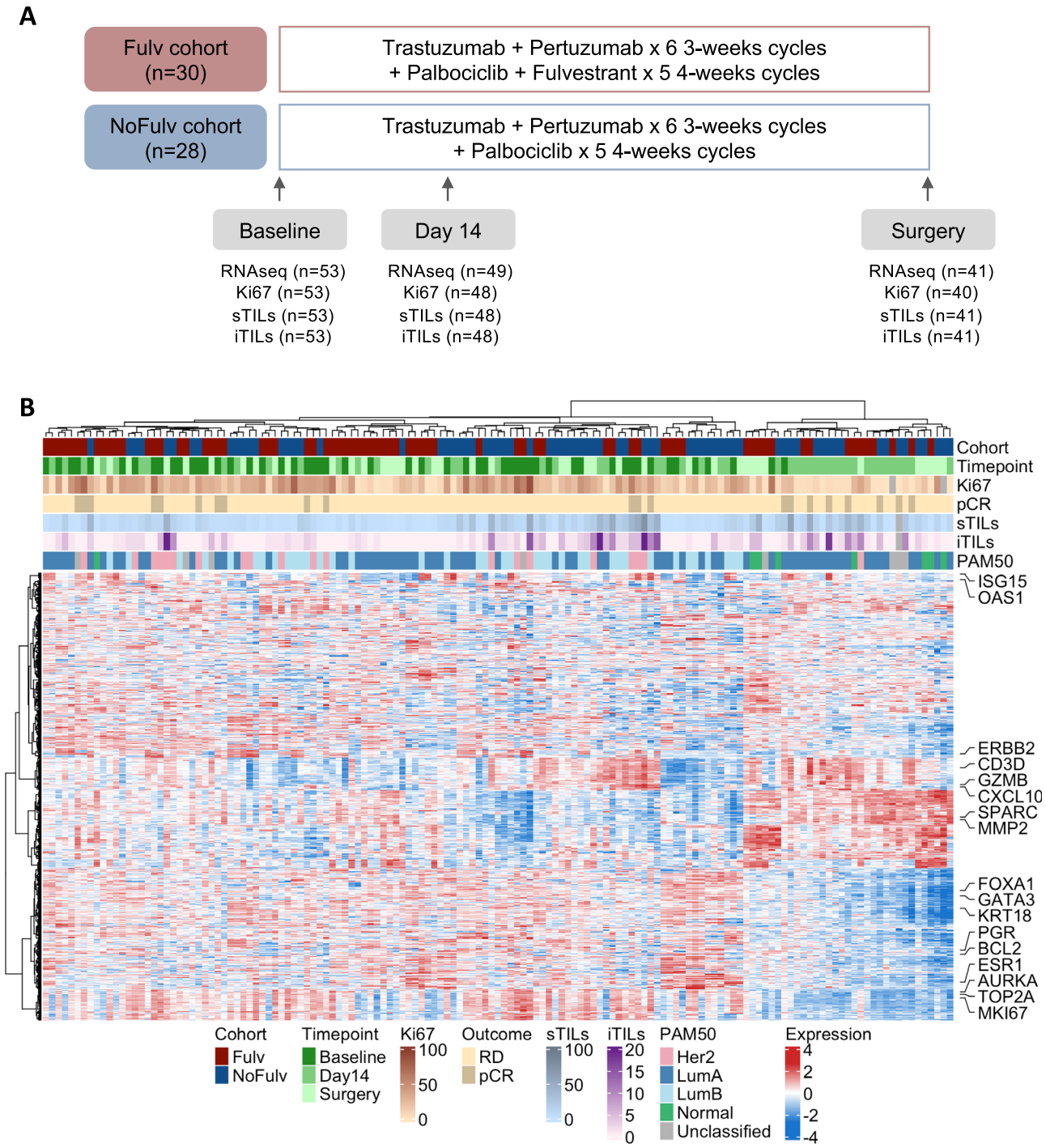


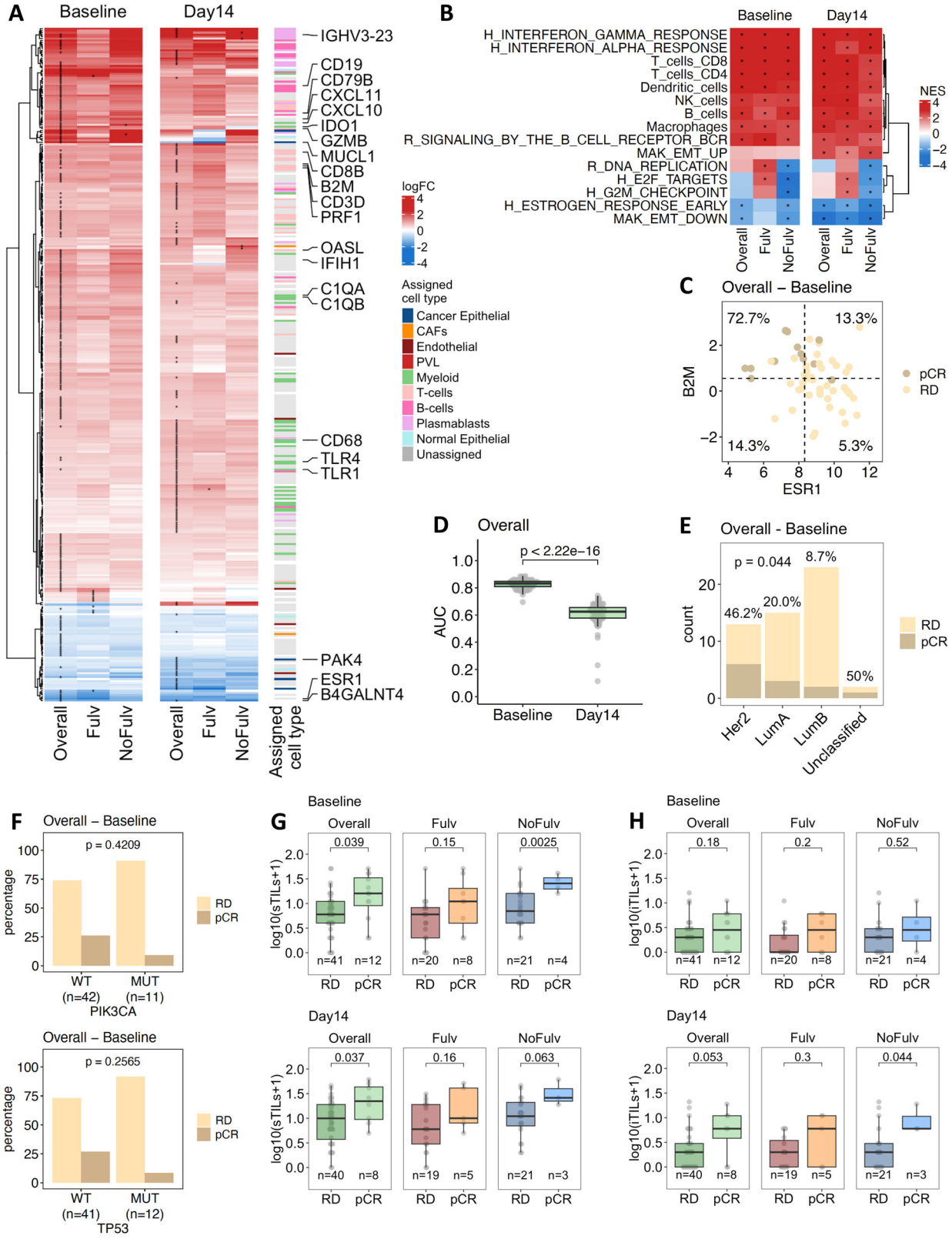
Fig. 1 | Translational research on the NA-PHER2 trial. A Schematic representation of the NA-PHER2 trial and associated serial sample collection. The number of samples for which RNA sequencing (RNA-seq), Ki67 quantification and TIL scoring were performed are indicated for each timepoint. **B** Annotated heatmap of 904 most expressed and variant genes. sTILs = stromal tumour infiltrating lymphocytes; iTILs = intra-tumoral infiltrating lymphocytes; pCR = pathologic complete response; RD = residual disease. Source data are provided as a Source Data file.

analysis of the Day14 timepoint is unlikely to be more informative than baseline biopsy in predicting pCR.

Transcriptomic data were also used to generate a PAM50 classification of all samples (Supplementary Fig. 5B). Classification of samples from a specific subgroup (i.e. HER2+ER+) and on- or post-treatment poses specific computational challenges that we overcame as detailed in the methods section. In the overall cohort, the probability of achieving pCR was 46.2% in HER2-enriched 20.0% in Luminal A and 8.7% in Luminal B cases ($p = 0.04$, Fig. 2E).

Mutations in *PIK3CA* and *TP53* genes at baseline were called from the RNA sequencing data (Supplementary Fig. 5C, Supplementary Table 2). While mutation calling from RNA-seq is not ideal, we checked the robustness of our calls (Supplementary Fig. 5D, E). Mutations in these two genes were similarly distributed across the PAM50 subtypes, except for *TP53* mutations not detected in the HER2-enriched subtype (Supplementary Fig. 5F). No significant associations with pCR were found (*PIK3CA* $p = 0.42$, *TP53* $p = 0.26$), but the response rate was numerically lower when either a *PIK3CA* or *TP53* mutation was detected (Fig. 2F).

https://doi.org/10.1038/s41467-025-57293-9



Consistent with the gene expression data, high baseline and Day14 stromal TILs were associated with pCR in the overall cohort ($p = 0.04$). The same trend was also seen when the two treatment cohorts were analysed separately (Fig. 2G). Association of intratumoral TILs with pCR reached borderline significance at Day14 ($p = 0.053$, Fig. 2H), while TIL dynamics between baseline and Day14 were not more informative (Supplementary Fig. 5G, H).

Overall, we identified both tumour and TME baseline features associated with achievement of pCR at surgery. Higher infiltration of immune cells, an HER2-enriched PAM50 phenotype and low activation of ER-mediated signalling were associated with higher probability of pCR. These associations were not qualitatively affected by the addition of fulvestrant to CDK4/6 inhibition and HER2 targeted therapies. Day14 biopsy did not add predictive value beyond what

Fig. 2 | Association of molecular features with pCR. A Heatmap of differentially expressed genes showing association with pCR (* = FDR < 10%) compared to patients with residual disease (RD) in at least one of the indicated comparisons. The association was evaluated in the overall population and separately for the two trial cohorts and using either baseline or Day14 measurements. Manually selected representative genes are highlighted. Gene to cell type annotation based on an external HER2+ scRNA-seq dataset[22] is reported. Genes are assigned to a specific cell type when their expression is significantly higher in that cell type according to the interrogated single-cell dataset. Unassigned genes are in grey, genes not measured/expressed in the single cell dataset are in white; see Methods for details. An extended version reporting all the gene symbols is presented in Supplementary Fig. 3. **B** Heatmap of representative manually selected genesets showing a positive or negative enrichment in patients achieving pCR compared to RD (* = FDR < 0.1%). A comprehensive list of all significant genesets is presented in Supplementary Fig. 4A. **C** Scatterplot of ESR1 and B2M expression at baseline in the overall population. Dashed lines indicate the first tertile and the median value for ESR1 and B2M respectively. **D** Area under the ROC curve (AUC) distribution after fitting 100 internally cross-validated regularized logistic regression models using either baseline or Day14 gene expression as candidate features to predict pCR. AUC distributions were compared using two-sided Student's t-test. **E** Association with pCR for the PAM50 subtypes defined at baseline, two-sided Fisher's test. **F** Association between presence of *PIK3CA* (top) or *TP53* (bottom) mutation and pCR, two-sided Fisher's test. **G** Association of sTILs with pCR in the overall population and per treatment cohort, quantified either at baseline or in Day14 biopsy, two-sided Student's t-test. **H** Association of iTILs with pCR in the overall population and per treatment cohort, quantified either at baseline or in Day14 biopsy, two-sided Student's t-test. All boxplots are defined as follow: centre line = median; box limits = upper and lower quartiles; whiskers = 1.5x interquartile range; points = outliers. sTILs = stromal tumour infiltrating lymphocytes; iTILs = intra-tumoral infiltrating lymphocytes; pCR = pathologic complete response; RD = residual disease; NES = Normalised Enrichment score. Source data are provided as a Source Data file.

provided by the baseline biopsy, with the possible exception of iTILs.

Tumour and TME features associated with Day14 Ki67

Proliferation suppression on Day14, indicated by Ki67 expression <10%, is an early efficacy measure of neoadjuvant endocrine therapy and is associated with good prognosis independent of pCR[14,15]. In the NA-PHER2 trial, all patients had >10% Ki67 expression at baseline (required as eligibility criteria)[12,13]. Ki67 levels on Day14 decreased similarly in both treatment cohorts, while a rebound was observed at surgery, more pronounced in the NoFulv cohort ($p = 0.017$, Fig. 3A). We found no significant association between Ki67 levels at Day14 and pCR ($p = 0.935$), suggesting that Ki67 response and pCR could capture distinct treatment response dynamics (Fig. 3B).

To identify predictive markers of Ki67 response, we compared baseline gene expression profiles of cancers that had high (i.e. >10%, $n = 16$) or low (i.e. ≤10%, $n = 31$) Day14 Ki67 in the overall populations and in the Fulv or NoFulv cohorts separately. No genes were differentially expressed after multiple testing correction (Supplementary Fig. 6A). Using geneset enrichment analysis, protein translation gene sets emerged as positively associated with high Day14 Ki67 in both treatment cohorts. In the overall and NoFulv cohorts, lower expression of mesenchymal gene signatures and high expression of proliferation-related gene sets were associated with lack of Ki67 reduction (Fig. 3C and Supplementary Fig. 6B). PAM50 subtypes, TIL counts and *PIK3CA* mutations were not associated with Day14 Ki67 (Fig. 3D, E, G, H). However, we found a significant positive association between somatic *TP53* mutation and lack of Ki67 reduction ($p = 0.029$, Fig. 3F). In light of the differential analysis results, we investigated the link between baseline proliferation (as estimated by *MKI67* gene expression) and *TP53* mutational status. *TP53* mutant tumours had significantly higher expression of *MKI67* at baseline (Supplementary Fig. 6C). Hence, we compared their ability to predict Day14 Ki67 using logistic regression. Only *TP53* mutation was significantly associated with Day14 Ki67 and the association remained significant in a bivariate model (OR = 5.87 P = 0.036, Supplementary Table 3). This can be explained noticing that all *TP53* mutant have quite high *MKI67* levels but not all tumours with high proliferation are *TP53* mutant.

All together, these results indicate that the identification of tumours that would have an early drop in proliferation during treatment is challenging and markers associated with Ki67 drop (e.g. mutation of *TP53*) are distinct from those predicting pCR.

Early changes induced by treatment

Next, we examined treatment induced transcriptomic changes by comparing Day14 profiles with baseline. Because treatment can reduce tumour cell content, we adjusted the expression levels for tumour cell purity using our previously developed Breast Cancer Purity Score (BCPS)[25]. Indeed, the BCPS dropped at Day14 and maintained similar levels in surgical samples from the Fulv cohort while a rebound was observed in the NoFulv cohort ($p = 0.046$, Supplementary Fig. 7A, B).

Comparison of gene expression profiles between Day14 and baseline, in the overall cohort and separately for the two treatment cohorts, indicated a decrease in proliferation, with MKI67 among the top DE genes (Fig. 4A, B, Supplementary Fig. 7C, Supplementary Fig. 8). This is consistent with the previously reported changes in Ki67 levels by IHC (Fig. 3A).

Among the large number of differentially expressed genes, we detected both up- and down-regulation of genes prevalently expressed by cancer associated fibroblasts (CAF), endothelial, myeloid and lymphoid cells, suggesting significant changes in the TME phenotype (Supplementary Fig. 7D, Supplementary Fig. 8). At the same time, we did not find significant changes in TIL count, only a slight upward trend between baseline and Day14 tissues (Fig. 4C, D).

Overall, the analysis of Day14 biopsies, compared to baseline showed a decrease in the expression of cell cycle related genes and a remodelling of the tissue microenvironment.

Longitudinal changes in tumours from patients with residual disease

To further study treatment induced changes in the NA-PHER2 trial, 122 longitudinal samples including baseline, Day14, and surgical samples from 43 patients with RD were analysed.

No major differences emerged in the gene expression dynamics between the two treatment cohorts (Supplementary Table 4). This justified focusing on the overall cohort, maintaining an adjustment for treatment. Time course analysis identified 655 genes with significant changes over time (Fig. 5A, Supplementary Fig. 9A, Supplementary Fig. 10). We applied unsupervised clustering (see Methods) to operationally group together genes with similar timepoint associated dynamics and 6 distinct patterns emerged. For each cluster, gene enrichment analysis was performed and significant enrichments were found for cluster 1,2 and 3 (Supplementary Fig. 9B). Additionally, gene to cell type annotation was added based on the external single-cell data (Fig. 5A, Supplementary Fig. 10).

Genes in cluster 1 showed a strong upregulation at surgery. They were linked to inflammation and interleukin signalling (e.g. *CXCL1/2, CXCR1, IL1R2, IL6 and IL33*) and activation of the leptin/complement cascade (e.g. *LEP, FCN1/2, MASP1, C6, C7*). Accordingly, they were often annotated as prevalently expressed by multiple TME cell types (Fig. 5A, Supplementary Fig. 9B and Supplementary Fig. 10). *CD36*, in this cluster, has been associated to drug resistance[26], including resistance to anti-HER2 treatment[27]. *MYC*, also in this cluster, has been associated with endocrine resistance[28] and could mediate overcoming treatment-induced cell cycle arrest. Cluster 2 included genes downregulated at surgery. Multiple CAF-expressed genes (i.e. *WNT2* and *COL11A1*),

https://doi.org/10.1038/s41467-025-57293-9

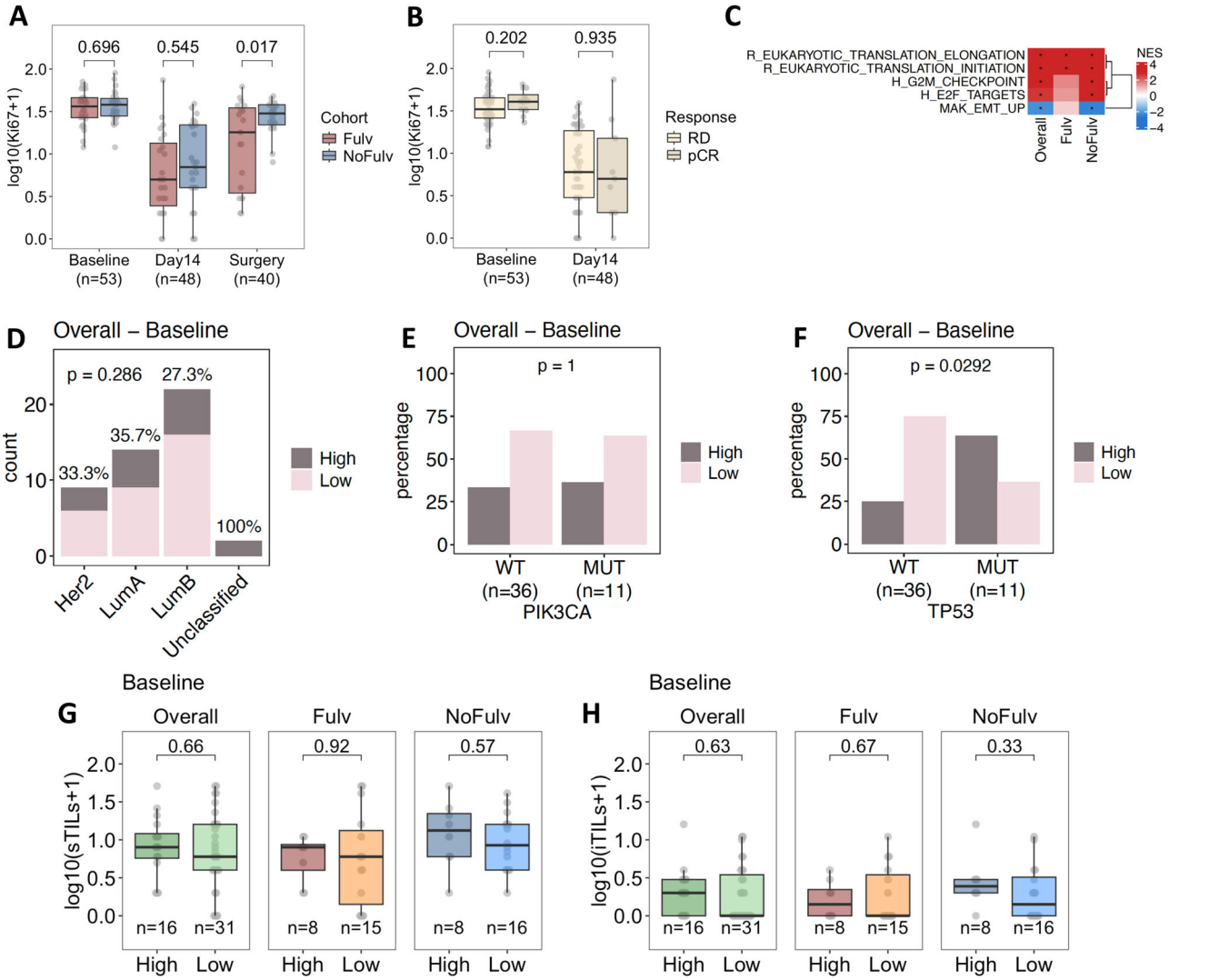


Fig. 3 | Association of baseline molecular features with Ki67 at Day14. **A** Differences in Ki67 dynamics during treatment in the Fulv and NoFulv cohorts, two-sided Student's t-test. **B** Evaluation of differences in Ki67 positivity in tumours from patients achieving or not pCR (overall cohort), two-sided Student's t-test. **C** Selected genesets with a significant enrichment in the Day14 Ki67 high group (i.e. Ki67 > 10%) compared to low (i.e. Ki67 ≤ 10%), * = gene permutation FDR < 0.1%). **D** Association of PAM50 subtypes defined at baseline with Day14 Ki67, two-sided Fisher's test. **E** Association between presence of *PIK3CA* mutation and Day14 Ki67, two-sided Fisher's test. **F** Association between presence of *TP53* mutation and Day14 Ki67, two-sided Fisher's test. **G** Association of baseline sTILs with Day14 Ki67 in the overall population and per treatment cohort, two-sided Student's t-test. **H** Association of baseline iTILs with Day14 Ki67 in the overall population and per treatment cohort, two-sided Student's t-test. All boxplots are defined as follow: centre line = median; box limits = upper and lower quartiles; whiskers = 1.5x interquartile range; points = outliers. sTILs = stromal tumour infiltrating lymphocytes; iTILs = intra-tumoral infiltrating lymphocytes; pCR = pathologic complete response; RD = residual disease; NES = Normalised Enrichment score. Source data are provided as a Source Data file.

reportedly overexpressed in CAFs compared to normal fibroblasts[29–31], were part of this cluster, suggesting a remodelling of CAFs towards a normal phenotype. Cluster 3 contained mostly genes linked to DNA replication and proliferation (e.g. *TOP2A, MKI67*) that mimic IHC Ki67 dynamics (Fig. 3A and Supplementary Fig. 9B). A few myeloid-expressed genes (e.g. *CXCL8* and *IDO1*) also belonged to this cluster. Less clear was the interpretation of clusters 4-6, containing few genes and mostly not linked to specific cell-types (Fig. 5A, Supplementary Fig. 10).

PAM50 dynamics were likely driven by the observed proliferation pattern and changes in tumour cellularity, with high prevalence of Luminal A at Day14 and an increasing number of normal-like samples over treatment (Supplementary Fig. 9C). Finally, we did not find genes indicating variation in lymphocytic infiltration, an observation confirmed by TILs quantification (Fig. 5B, C).

Overall, the longitudinal analysis of gene expression data from cancers with RD unveiled a profound remodelling of the TME, together with the overexpression of genes involved in inflammation and resistance by the end of treatment.

Molecular changes according to Ki67 dynamics

Besides Day14 Ki67, Ki67 levels at surgery have been reported as prognostic after neoadjuvant treatment[32]. Consequently, we further stratified RD patients ($n = 39$) into three groups: the HighHigh (i.e. Ki67 > 10% both at Day14 and at surgery), the LowHigh and the LowLow group. Two patients with HighLow Ki67 were included in the Low-Low group (Supplementary Fig. 11A). Time-course analysis identified 207 genes with distinct dynamics in any of the three groups (Supplementary Fig. 11B) and 9 distinct time-variant expression patterns emerged by unsupervised analysis. A gene enrichment analysis and the gene to

https://doi.org/10.1038/s41467-025-57293-9

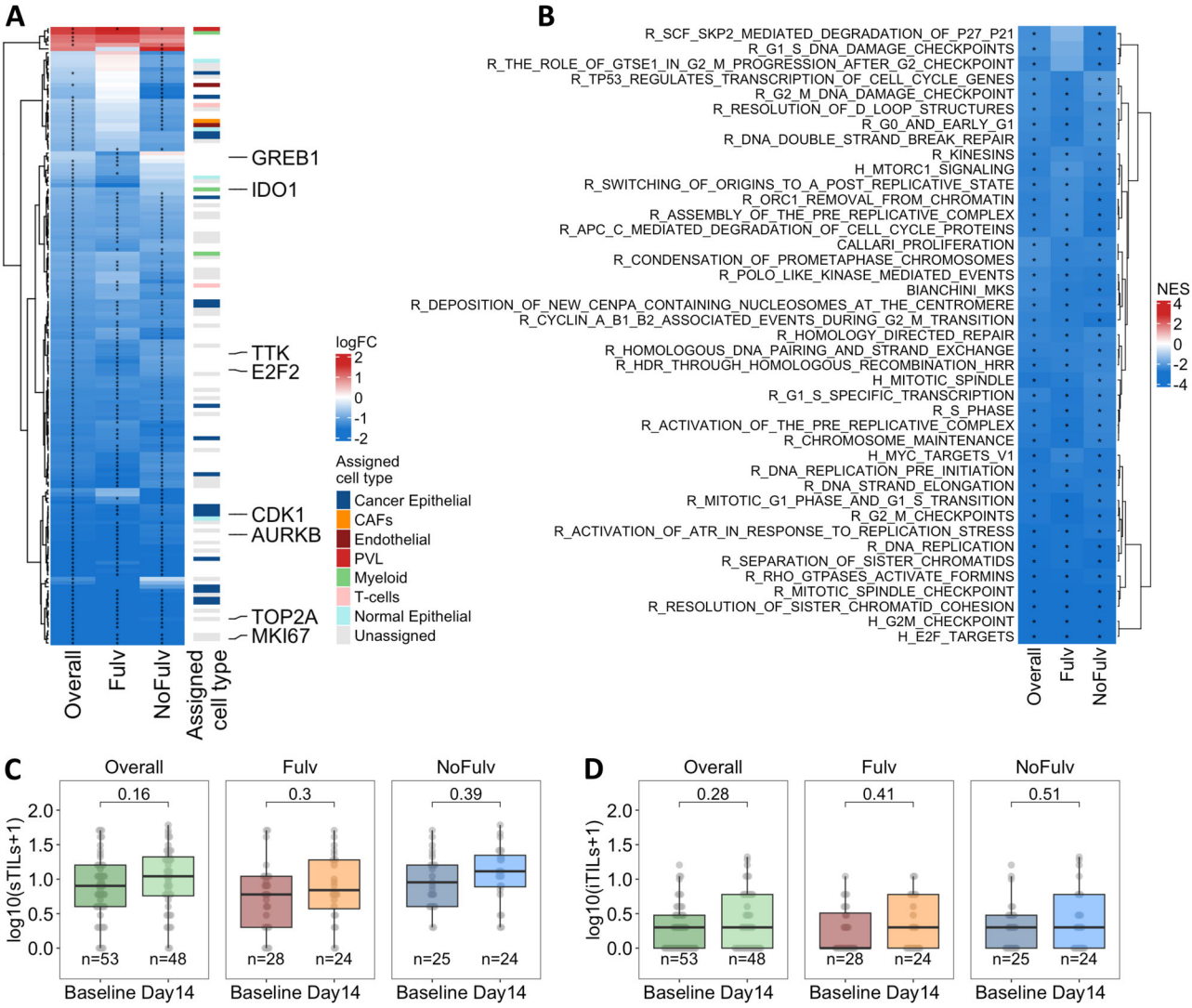


Fig. 4 | Early treatment-induced molecular changes. A Heatmap of genes differentially expressed at Day14 compared to baseline. The comparison was carried out in the overall population and separately for the two trial cohorts. Manually selected representative genes are highlighted. Gene to cell type annotation based on an external HER2+ scRNA-seq dataset[22] is reported. Only genes with |logFC|>1 are reported, * = FDR < 10%. An extended version, reporting all genes with FDR < 10% in at least one comparison, is presented in Supplementary Fig. 8. **B** Heatmap of genesets significantly enriched in either the overall population or one of the two treatment cohorts, * = gene permutation FDR < 0.1%. **C** Evaluation of changes in sTILs at Day14 compared to baseline in the overall population and separately for the Fulv and NoFulv cohorts, two-sided Student's t-test. **D** Evaluation of changes in iTILs at Day14 compared to baseline in the overall population and separately for the Fulv and NoFulv cohorts, two-sided Student's t-test. All boxplots are defined as follow: centre line = median; box limits = upper and lower quartiles; whiskers = 1.5x interquartile range; points = outliers. sTILs = stromal tumour infiltrating lymphocytes; iTILs = intra-tumoral infiltrating lymphocytes. Source data are provided as a Source Data file.

cell type annotation using external single-cell data were performed (Fig. 6A, Supplementary Fig. 11C and Supplementary Fig. 12).

Cluster 1 was enriched in proliferation-related genes, often annotated as cancer-specific (Fig. 6A, Supplementary Fig. 11C). For example, *MKI67*, *CCNB1* and *E2F1* are members of this cluster, with a pattern matching the Ki67-based group definition. Cluster 2 included genes downregulated at surgery in the LowLow group. A few of them, are involved in the mitochondrial metabolism, biosynthesis of coenzyme A and coenzyme Q10 (*NADK, PPCS* and *COQ2*). *COQ2*, in particular, was prevalently expressed by myeloid cells, suggesting a potential alteration in the functional phenotype of this cell type[33]. Cluster 3 showed a few genes up-regulated at surgery particularly in the LowLow group (e.g. *ETS2* and *ADAMTS1*) and linked to TNFα signalling by enrichment analysis. Clusters 4-6 contained genes with divergent dynamics in the LowLow group, often with a differential expression already present at baseline. Among them, the ER regulated gene *TMEM26*[34], Ligand dependent nuclear receptor *LCORL*, cell-cycle regulators *YPEL1* and *YPEL4* and neuronal genes *NRN1* and *SYT12*. In cluster 7, genes were upregulated in the LowLow and downregulated in the HighHigh at surgery. Among them was *PAPPA*, an activator of the IGF pathway[35]. Less insightful were cluster 8 and cluster 9 containing three and one gene respectively (Fig. 6A, Supplementary Fig. 12). In line with gene expression and Ki67 dynamics, 6 out of 8 LowLow tumours were classified as Luminal A at surgery (Supplementary Fig. 11D).

Finally, both baseline sTILs and iTILs tended to be higher in the HighHigh group already at baseline, and the difference became larger in the later timepoints (Fig. 6B, C).

In summary, by stratifying patients with RD based on Ki67 dynamics, we identified distinct transcriptional directories for the HighHigh, LowHigh and LowLow groups. The LowLow showed the most divergent dynamics, in both the tumour and the TME.

https://doi.org/10.1038/s41467-025-57293-9

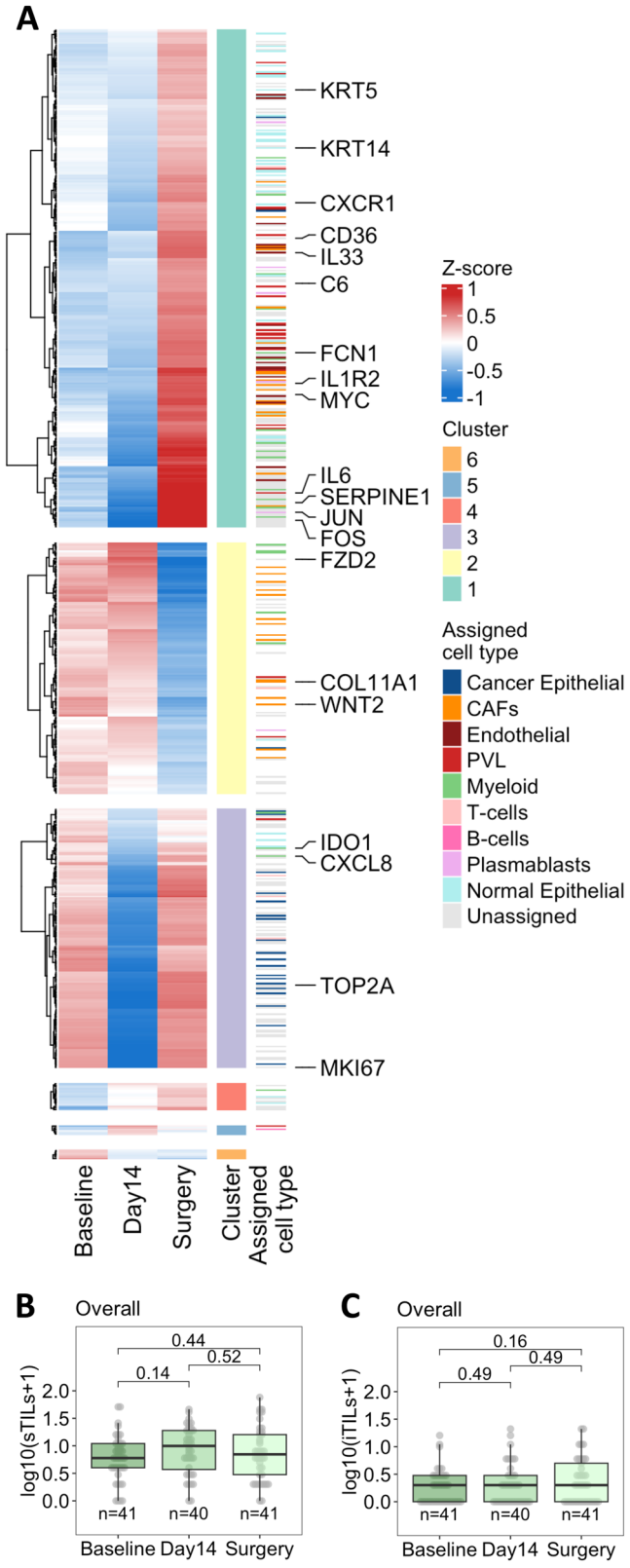


Fig. 5 | Treatment induced changes in tumours from patients with residual disease at surgery. A Heatmap of genes differentially expressed in a cohort-adjusted time course analysis. Six distinct patterns were identified by clustering analysis and manually selected representative genes are indicated. Gene to cell type annotation based on an external HER2+ scRNA-seq dataset[22] is reported. An extended version reporting all the gene symbols is presented in Supplementary Fig. 10. **B** Evaluation of changes in sTIL levels during neoadjuvant treatment in the subgroup of patients with residual disease at surgery, two-sided Student's t-test. **C** Evaluation of changes in iTIL levels during neoadjuvant treatment in the subgroup of patients with residual disease at surgery, two-sided Student's t-test. All boxplots are defined as follow: centre line = median; box limits = upper and lower quartiles; whiskers = 1.5x interquartile range; points = outliers. sTILs = stromal tumour infiltrating lymphocytes; iTILs = intra-tumoral infiltrating lymphocytes. Source data are provided as a Source Data file.

Tumours with distinct response patterns have different transcriptomic profiles at baseline

The observation that LowLow tumours already differed from the others at baseline (Fig. 6A) and similarities between the pCR and HighHigh group in terms of immune infiltration (Fig. 6B, C) prompted the direct comparison between pCR, HighHigh, LowHigh and LowLow groups in the overall cohort at baseline.

Both sTILs and iTILs were the highest in the pCR group, followed by the HighHigh group, with numerically but not statistically lower values (Fig. 7A, B). Association of the four response groups with baseline *PIK3CA* or *TP53* mutation and PAM50 classification, did not reach statistical significance (Supplementary Fig. 13A–C). However, while half of cases achieving pCR were HER2-enriched, none of the LowLow group belonged to this subtype (Supplementary Fig. 13C).

At transcriptomic level, 178 genes were differentially expressed ($p < 0.005$ and FDR < 0.3), which could be grouped in four main clusters (Fig. 7C and Supplementary Fig. 14). Cluster 2 showed a significant gene enrichment for IFN response and immune-related genesets (Supplementary Fig. 13D). In line with the TILs analysis, genes that reflect immune cell abundance (e.g. *B2M, RPSA*), drive immune cell recruitment (e.g. *AQP9*) or promote an immune suppressive/pro-cancer environment (e.g. *CCL2*) were expressed at similar levels in the pCR and HighHigh groups. However, Type I and Type II IFN response genes (e.g. *IFIH1, OASL*,) and Toll-Like Receptor signalling activator (e.g. *MYD88*) showed high expression in the pCR group but not in the HighHigh group. These innate immunity pathways promote adaptive anti-tumour immune response by stimulating lymphocytes and dendritic cell anti-tumour activity. Indeed, genes linked to an active immune response (e.g. *CCR5, CYTH1*) and consequent activation of a negative feedback (e.g. *LAG3, MARCH1*) were up in the pCR group (Fig. 7C and Supplementary Fig. 14). Overall, the balance between pro- and anti-tumour immune activity favoured the anti-tumour activity in the pCR group but not in the High/High group.

Genes upregulated in the HighHigh group populated cluster 1 (Fig. 7C and Supplementary Fig. 14). Among them, YTH N6-Methyladenosine RNA Binding Protein F3 (*YTHDF3*) is a negative regulator of type I interferon response[36], in line with pattern in cluster 2. It is worth noting *PAK4*; prevalently expressed by the tumour cells, has been associated with lower T cell and dendritic cell infiltration and resistance to immune-checkpoint inhibitors[37,38]. *PAK4* is a master gene and a potential therapeutic target for which inhibitors are being clinically tested[39]. Together with other cancer expressed genes in cluster 1 (e.g. *MAP3K9* and *EDN2*), it could activate the MEK/ERK pathway to overcome dual HER2 blockade.

Genes in the large cluster 3 were upregulated in the LowLow group (Fig. 7C and Supplementary Fig. 14). While no significant geneset enrichment was found (Supplementary Fig. 13D), of note is the high expression of *ESR1* and *IFG1R*, together with neuronal gene *BCAN* and *B4GALNT4*, negatively associated with pCR (Fig. 2A). Cluster 4 contained few TME genes up-regulated in the LowHigh group.

Finally, we specifically evaluated HER2 mRNA expression across the 4 groups. Highest expression levels were found in the pCR group followed by the HighHigh group, with lower values observed in the LowHigh and LowLow groups (Fig. 7D).

https://doi.org/10.1038/s41467-025-57293-9

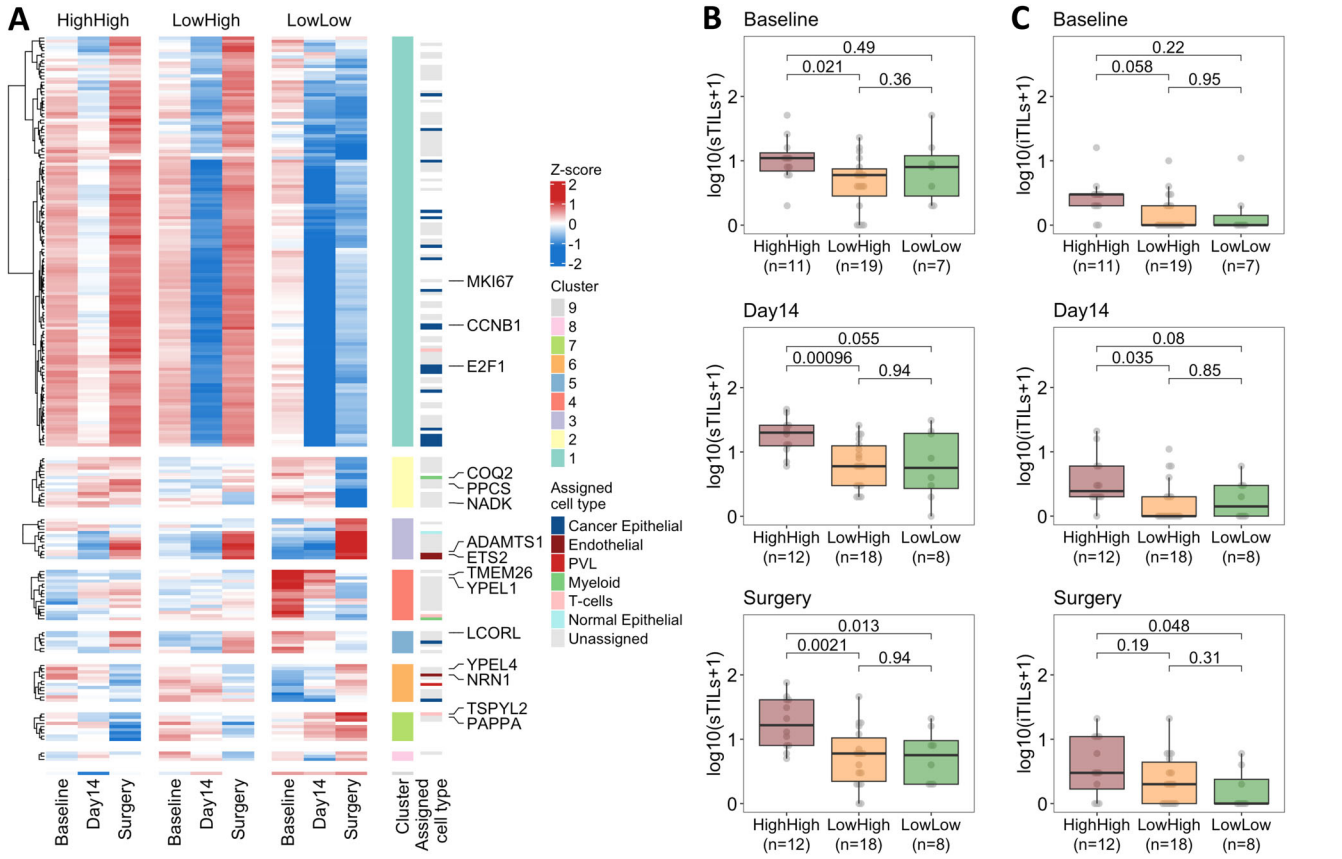


Fig. 6 | Gene expression and TIL changes during neoadjuvant treatment in patients with residual disease at surgery stratified by Ki67 dynamics. **A** Heatmap of genes differentially expressed in a time course analysis comparing Ki67 LowLow, LowHigh and HighHigh groups. Nine distinct patterns were identified and representative genes indicated. Gene to cell type annotation based on an external HER2+ scRNA-seq dataset[22] is reported. An extended version reporting all the gene symbols is presented in Supplementary Fig. 12. **B** Evaluation of sTIL levels in the three groups defined by Ki67 dynamics at baseline, Day14 and surgery, two-sided Student's t-test. **C** Evaluation of iTIL levels in the three groups defined by Ki67 dynamics at baseline, Day14 and surgery, two-sided Student's t-test.All boxplots are defined as follow: centre line = median; box limits = upper and lower quartiles; whiskers = 1.5x interquartile range; points = outliers. sTILs = stromal tumour infiltrating lymphocytes; iTILs = intra-tumoral infiltrating lymphocytes. Source data are provided as a Source Data file.

Overall, we found that transcriptional signatures at baseline could stratify tumours based on their response to the combined targeted treatments in NA-PHER2.

Discussion

In HER2+ breast cancer, response to standard treatment is significantly lower in HER2+ER+ cases compared to HER2+ER- disease[16,40]. A chemo-free strategy that might be equally effective in HER2+ER+ cancers, based on the concomitant blockade of HER2 and CDK4/6, with or without endocrine therapy was investigated in the NA-PHER2 Phase II neoadjuvant trial[12,13]. Results showed a significant drop in Ki67 positivity after 14 days of treatment and a 19.2 to 27% pCR rate depending on study arm, opening to the opportunity for treatment de-escalation. However, molecular tools for patient stratification and better understanding of the biological impact of treatment on clinical disease are necessary to implement improved treatment strategies.

We obtained pathological evaluation of TILs and the transcriptomic profile of the tumours before treatment, after two weeks on treatment and at surgery for over 90% of patients enrolled in the NA-PHER2 trial. Transcriptomic profiles were used for gene expression quantification, molecular subtyping and detection of somatic mutations in the *PIK3CA* and *TP53* genes. By integrating our bulk RNA-seq data with publicly available single-cell transcriptomic data[22], we identified, whenever possible, which cell type the gene expression signal was prevalently coming from, complementary information aiding the biological interpretation.

Achievement of pCR is an early patient-level surrogate of excellent long term recurrence free survival[41]. However, in the ER+ disease, the clinical benefit from neoadjuvant therapy extends beyond patients with pCR. Cancers with low Ki67 after 14 days of endocrine therapy also have excellent long-term survival[14,15] and this was adopted as the primary endpoint of the NA-PHER2 trial[12,13]. Therefore, we searched for molecular features associated with either pCR or with low Ki67 at Day14.

Immune infiltration and low *ESR1* mRNA expression were major determinants of the probability to achieve pCR. By simply combining the expression of *ESR1* and *B2M*, we could identify a group with 73% probability of pCR and a group with as low as a 5% chance to achieve pCR. Additional tumour intrinsic features associated with pCR included high expression of *MUCL1* and an HER2-enriched PAM50 subtype. More complex machine learning predictors also achieved high internally cross-validated performances. Overall, the baseline expression profile turned out to be more informative than data from the Day14 biopsy to predict pCR, with the notable exception of iTILs.

Lower *ESR1* expression among patients with pCR (although all enrolled patients have centrally verified ER positive tumours by immunohistochemistry) is in agreement with earlier clinical observations of higher pCR rates in the HER2+ER- disease[16]. The association of immune infiltration and HER2-enriched PAM50 subtype with pCR is

https://doi.org/10.1038/s41467-025-57293-9

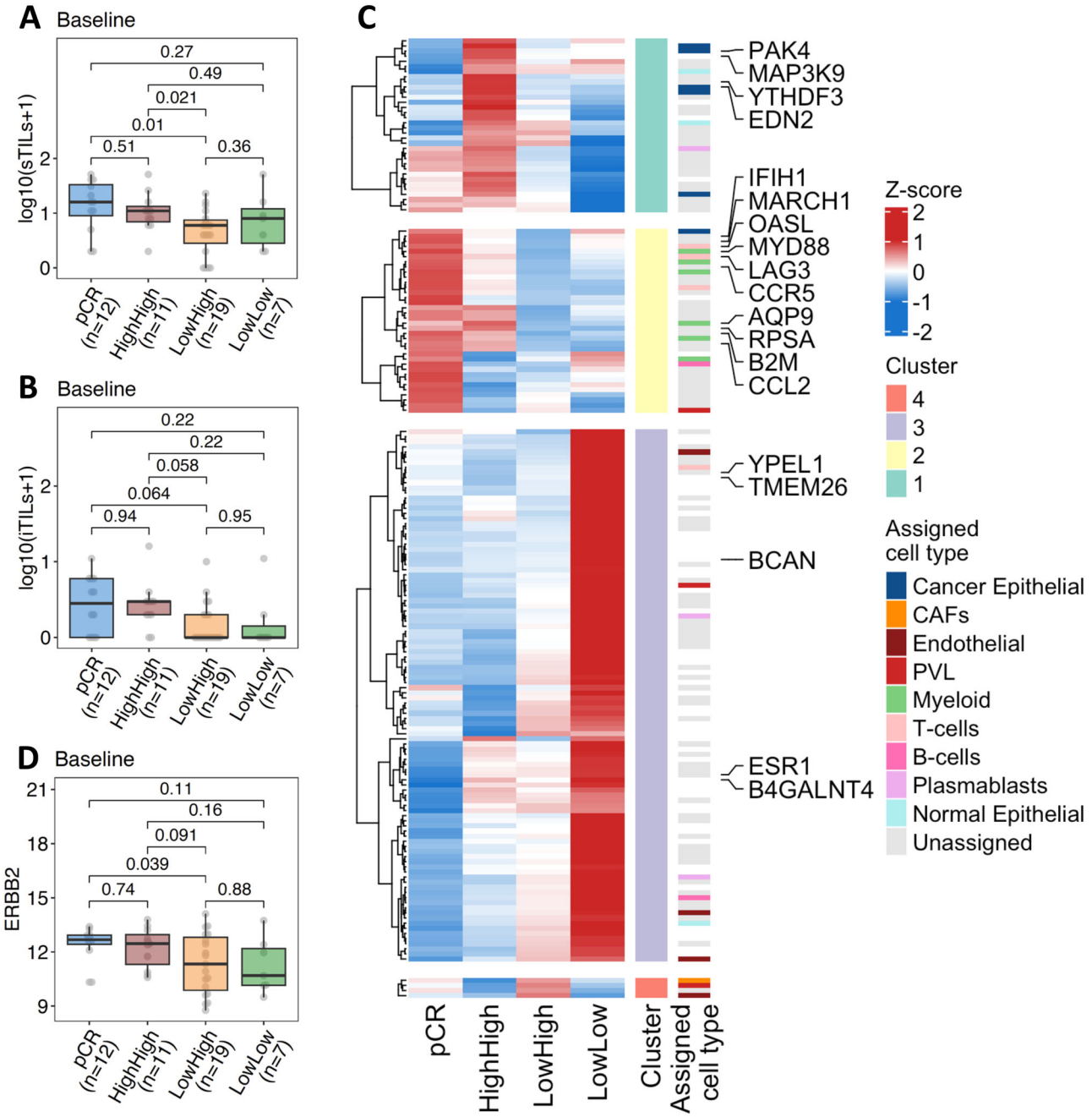


Fig. 7 | Differential analysis between four distinct response groups. Response to treatment was stratified in four distinct groups: the pCR group and three groups with residual disease at surgery but with distinct Ki67 levels defined by Day14 and surgery Ki67 (i.e. HighHigh, LowHigh, LowLow). **A** Comparison of sTIL levels between the four response groups. Differences were evaluated by two-sided Student's t-test. **B** Comparison of iTIL infiltration between the four response groups. Differences were evaluated by two-sided Student's t-test. **C** Heatmap of differentially expressed genes across the four groups. Annotation of the cell type prevalently expressing each gene is reported, when available in the HER2+ single-cell dataset[22]. Four gene clusters were identified by unsupervised analysis and manually selected genes are highlighted. An extended version of the heatmap is reported as Supplementary Fig. 14. **D** Comparison of HER2 expression levels in the four response groups. Differences were evaluated by two-sided Student's t-test. All boxplots are defined as follow: centre line = median; box limits = upper and lower quartiles; whiskers = 1.5x interquartile range; points = outliers. sTILs = stromal tumour infiltrating lymphocytes; iTILs = intra-tumoral infiltrating lymphocytes. Source data are provided as a Source Data file.

also in line with previous reports in the HER2+ disease, and were also seen in the CherLOB[42], NeoSPHERE[3], CALGB40601[43] and NOAH[44] trials. Both anti-HER2 monoclonal antibodies[8] and CDK4/6 inhibitors[45] have been shown to activate anti-tumour immune response and could be primarily responsible for the complete eradication of highly immune infiltrated HER2+ER+ tumours.

It is relevant to understand what is driving the engagement of the immune system with cancer cells at baseline. Such response was not per se sufficient to eradicate the tumour, but it became so upon treatment. We saw in the pCR group high activation of the RIG-1/MDA5 pathway at baseline, which is physiologically triggered by the detection of viral RNA, leading to innate immune response and IFN pathway activation. We hypothesise that in the context of cancer, a viral mimicry phenomenon could be triggered by aberrantly transcribed endogenous retrotransposons[46]. This mechanism could be enhanced by the CDK4/6 inhibitor Palbociclib, since CDK4/6 inhibitors have been

shown to activate expression of endogenous retroviral elements[47] and, consequently, an anti-tumour immune response.

Persistence of high Ki67 at Day14 is a sign of resistance to treatment and in our analysis associated with higher expression in translation and proliferation genes at baseline and presence of *TP53* mutations. When compared in a logistic regression model, only *TP53* mutation at baseline was significantly associated with Day14 Ki67. Mutated p53 has been linked to primary endocrine resistance in luminal cancers receiving perioperative endocrine therapy as part of the WSG-ADAPT[48] and POETIC[49] trials.

Our results clearly indicate that achievement of pCR or low Day14 Ki67 are determined by distinct biological processes. This is further supported by the observation that Day14 Ki67 levels were similar in pCR and RD patients.

After two weeks of treatment, the transcriptomic analysis confirmed the overall reduction of proliferation and unveiled transcriptional changes in the whole tumour ecosystem. When including the surgical samples for patients with RD, we observed the activation of inflammation/innate immunity-related signals at surgery that were not accompanied by quantitative changes in TILs.

The RD group was heterogeneous. By combining Ki67 levels at Day14 and at surgery, we could define three distinct groups (HighHigh, LowHigh and LowLow). Transcriptional trajectories were different in the three groups, with the LowLow diverging the most. Some transcriptomic differences between the response groups were already present at baseline, enabling the opportunity to predict not only pCR but also the proliferation dynamics induced by the NA-PHER2 treatment. While long term follow-up was not collected in our trial, it is reasonable to hypothesise that the LowLow group could have a good long-term prognosis despite not achieving pCR. Indeed, at the end of treatment it was characterised by low proliferation, lower HER2 levels, high expression of *ESR1* and prevalently a PAM50 Luminal A classification. Such features have been associated with endocrine sensitivity and the cytostatic effect observed in this group is likely to be mostly driven by Fulvestrant. Accordingly, Ki67 at surgery was significantly lower in the Fulv cohort. Intriguingly, in the LowLow group we observed several neuronal genes either up-regulated at baseline (e.g. *BCAN*, Fig. 7), or induced during treatment (e.g. *NRN1*, Fig. 6). The crosstalk between the nervous system and cancer is increasingly recognised and neural signalling can influence cancer growth and progression[50,51]. The functional role of these neuronal genes in HER2+ER+ disease warrants further investigation.

The HighHigh group arguably benefitted the least from the neoadjuvant therapy and represents the population that may be best treated with alternative therapies. This group had some unexpected commonalities with the pCR group, sharing quantitatively similar but qualitatively different immune infiltration at baseline. The lack of activation of IFN response, which hampers an effective anti-tumour immune response, could be partially explained by the overexpression of *PAK4* and *YTHDF3*[36]. *PAK4* expression is implicated in immune cell exclusion and resistance to PD-1 blockade for which clinical inhibitors are now being developed[37–39]. The substantial TIL presence in these non-responding cancers also highlights the limitations of histologic TIL quantification, which does not capture phenotypic information and immune fitness.

Possible limitations of our study lay on the relatively small sample size without long-term follow-up and the lack of a true validation, due to the absence of independent cohorts of patients receiving analogous treatment combinations. In the analysis of transcriptomic data, we acknowledge that RNA-seq is not the gold standard to identify somatic mutations but satellite analyses corroborated the validity of our calls. In our gene to cell type annotation, the number of tumours included in the single-cell dataset could limit the generalizability of the annotation, but we controlled for the false discovery rate as specified in the methods.

Our findings, however, shed light on the mechanisms and dynamics associated with distinct responses of clinical HER2+ER+ tumours to the targeting of the major driver pathways and contribute to the design of de-escalation or escalation strategies fitting the individual therapeutic needs in HER2+ breast cancer. Indeed, patients with high immune infiltration and low ER signalling could be spared chemotherapy and may be candidate to receive the targeted treatments investigated in the NA-PHER2 trial. On the contrary, the HighHigh group of tumours that is quite resistant to the administered drug combination could require treatments targeting PAK4 or with drugs reversing the immune-suppressive microenvironment.

Methods

Study design and sample collection

NA-PHER2 is an open-label, multicohort, exploratory phase 2 study run in seven hospitals and research institutions in Italy (https://clinicaltrials.gov/study/NCT02530424). The trial was approved by the Ethics Committees listed in Supplementary File 1. The criteria and methods were previously reported[12,13]. Briefly, patients were eligible if they had previously untreated, histologically confirmed, unilateral, invasive, HER2-positive (3+ on immunohistochemistry or neu-amplified) and ER-positive (>10% of cells in the tumour expressing ER) breast cancer. Other inclusion criteria were age 18 years or older, Eastern Cooperative Oncology Group (ECOG) performance status of 1 or less, and tumour classified between cT1c and cT4a–d. Key exclusion criteria were metastatic disease, bilateral breast cancer, other malignant disease, inadequate bone marrow or renal function, impaired liver function, impaired cardiac function, uncontrolled hypertension, pregnancy, and refusal to use contraception. Patients received neoadjuvant treatment with the combined regimen of trastuzumab, pertuzumab, palbociclib with (cohort A, herein referred to as Fulv) or without addition of fulvestrant (cohort B, herein referred to as NoFulv).

The study was undertaken according to Good Clinical Practice guidelines and the Declaration of Helsinki. All patients provided written informed consent for participation in the study and for provision of tumour tissue. The study protocol and informed consent form were approved by independent ethics committees at every participating institution and the relevant competent authority.

We obtained tumour samples locally from all patients at baseline, 2 weeks on treatment, and before surgery by biopsy with a 14-gauge needle (preferably with image guidance). Samples were fixed and embedded in paraffin blocks for analysis at the central referral laboratory (European Institute of Oncology, Milan, Italy). We labelled all tumour blocks to indicate the timing of the biopsy (i.e., baseline, Day14, or surgery).

Pathological assessment and immunohistochemistry

As part of the trial protocol, all samples for each timepoint (i.e. at baseline for HER2, ER and Ki67 status to confirm eligibility; at week 2 and at surgery for Ki67 and pathological response) were assessed centrally by one pathologist. Ki67 assessments was performed at the central laboratory using MIB-1 antibody (Dako, dilution 1:200); other kits used for retrospective assessment of final eligibility on the basis of HER2 and ER positivity were HercepTest kit (Dako), HER2 IQFISH pharmDx (Dako), and ER/PR PharmDx kit (Dako) used according to the manufacturer's instructions. All Dako kits were provided by Agilent Technologies Italia (Cernusco sul Naviglio, Milan, Italy).

TILs (sTILs and iTILs) were retrospectively quantified by the central laboratory. They were evaluated visually following the recommendations of the International Immuno-Oncology Biomarker Working Group[52].

RNA sequencing
Gene expression data was generated for the three timepoints using RNA sequencing on total RNA samples derived from formalin-fixed, paraffin-embedded (FFPE) tissue sections. RNA and miRNA extraction was performed with miRNeasy FFPE Kit (Qiagen Cat No./ID: 217504) and the quality of the extracted RNA was evaluated using RNA kit (DNF-489) on Fragment Anlayzer (AATI).

RNA-seq libraries were produced using NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina according to supplier recommendations (NEB). The capture was then performed on cDNA libraries with the Twist Human Core Exome Enrichment System according to supplier recommendations (Twist Bioscience). The obtained eluted-enriched DNA samples was then sequenced on an Illumina NovaSeq as paired-end 100 bp reads. Image analysis and base calling was performed using Illumina Real Time Analysis with default parameters (version 3.4.4).

Transcriptomic data processing
Base call files from each sequencing run were converted to fastq format using bcl2fastq conversion software (version 2.20). A quality control of forward and reverse fastq reads was employed using FastQC (version 0.11.4). The adapter content was removed using Trimmomatic and the resulting fastq file was aligned to the Ensembl GRCh37 *Homo sapiens* reference using HISAT2. The number of reads per gene were quantified using featureCounts and converted to transcripts per kilobase million (TPM) when needed. Correlation analysis excluded the presence of batch effects and highlighted the presence of one outlier sample removed from downstream analyses (Supplementary Fig. 1A). A total of 17314 genes with a read count>50 in at least 7 samples (i.e. 5% of all samples) were kept for downstream analyses.

PIK3CA and *TP53* Mutation calling from RNA-seq data
Variant detection from aligned reads was performed using FreeBayes. Genetic variant annotation and functional effect prediction was carried out on the output VCF files using SnpEff. Candidate mutations in the *PIK3CA* and *TP53* (excluding all synonymous variants) were annotated with the Cancer Genome Interpreter[53] for the positive selection of breast cancer driver alterations according to the provided classification coming from integration of knowledge collected across the public domain (e.g. OncoKB) with computational methods. We verified that our calls were quantitatively comparable with the one in the TCGA HER2 + /ER+ group using the two-sided Fisher's exact test statistics. Additionally, we checked *TP53* expression levels to identify possible false negatives, since *TP53* truncating mutations could cause lack of expression, making them undetectable in RNA-seq data. Association with pCR or Day14 Ki67 was evaluated using two-sided Fisher's exact test.

Differential gene expression
To compare two or multiple groups, raw counts were transformed in log2 counts per million using voom() function (mean-variance modelling at the observational level) as implemented in the limma R package (version 3.52.2). Linear models were fit and contrasts of interest extracted as implemented in the limma R package (version 3.52.2). Genes with a FDR < 10% (Benjamini-Hockberg multiple testing correction) were considered significantly differentially expressed, unless otherwise specified.

Geneset enrichment analysis
Geneset enrichment analysis was performed using the gsea() function of the phenoTest R package (version 1.44.0). All genes were ranked based on the t statistics obtained from the comparison of interest. We tested for enrichment the (i) C2 Reactome and Hallmark geneset collection (version 7.4) downloaded from the MsSigDB (https://www.gsea-msigdb.org/gsea); (ii) 18 immune and TME signatures optimised to estimate population abundance in breast cancer[54]; (iii) additional signatures previously reported by us and others, as detailed in Supplementary Table 5. Genesets containing <4 genes or > 200 genes were filtered out before testing. Gene permutation was applied to estimate the enrichment significance. Genesets with FDR < 0.1% and |NES | > 2.5 were considered significantly enriched. Results were represented as heatmaps showing NES and FDR for each geneset in each comparison.

Predictive models
Multi-gene predictors were trained using voom transformed gene expression data from either the baseline or Day14 timepoints. Forty-seven patients with both timepoints available were included in this analysis. 8116 genes with mean expression >0 and standard deviation >0.5 were used as candidate features. We applied a regularised logistic regression with elastic-net penalty ($\alpha = 0.5$) as implemented in the glmnet R package (version 4.1.4). Regularisation parameter Lambda (defining the subset of genes to include) was selected by 10-fold cross validation, out of 100 possible values tested. The whole process was then repeated 100 times. For each run, the internally cross-validated performance was evaluated by computing the area under the ROC curve (AUC) using the ROCit R Package (version 2.1.1). The features selected at each iteration were recorded allowing to identify the most recurrent features, defined as those selected >95/100 iterations.

Longitudinal analysis
In all longitudinal analyses, we took into account the treatment-associated changes in tumour purity, as quantified by our previously developed BCPS[25] (Supplementary Fig. 7A). Expression levels [i.e. (log2CPM)] were adjusted per gene by taking the residuals of the linear regression model estimating the relationship between the BCPS and the gene. In the differential analysis, |log2(fold change)|>1 in at least one timepoint was required on top of FDR < 10% threshold. Presence of multiple samples from the same patient was modelled using duplicateCorrelation() function as implemented in the limma R package (version 3.52.2).

Gene clustering and enrichment analysis
Differentially expressed genes from either the longitudinal analysis or the four response group comparison were clustered based on their average expression across timepoints and/or groups of interest using the ConsensusClusterPlus R package[55] (version 1.60.0). ConsensusClusterPlus starts by subsetting a proportion of samples (set to default 0.8). Each subsample was clustered by agglomerative hierarchical clustering and partitioned into k groups. We explored a range of k between 2 and 15. The process was repeated 100 times. Pairwise consensus values, i.e. the proportion of clustering runs in which two items are grouped together, are calculated for each k. Then, for each k, a final agglomerative hierarchical consensus clustering using distance of 1−consensus values was completed and pruned to k groups. K was primarily selected by evaluating the provided metrics, i.e. the relative change in the area under the cumulative distribution function curve, additionally checking the clusters gene content and functional enrichments described hereafter.

For the list of genes in each cluster, enrichment analysis (using the same pool of genesets used for the geneset enrichment analysis) was evaluated by hypergeometric test and corrected for multiple testing using the hypeR R package (version 1.12.0).

PAM50 classification and association with clinical variables
The PAM50 molecular subtype classifier[56] uses a centroid-based classifier requiring a gene scaling step to be suitable for the classification of transcriptomic data obtained with profiling technologies different from the original training set. If scaling is applied to a dataset with significantly different subtype proportions or where some of the subtypes are missing, a major bias can be introduced during scaling, as also previously

reported[43]. This is the case for the NA-PHER2 dataset, which included only HER2+ER+ tumours. Additionally, on- and post-treatment samples could differ from the pre-treatment samples included in the training set. To overcome such shortcomings in the PAM50 classification, we implemented a strategy based on the integration of the NA-PHER2 dataset with the TCGA dataset[57] as detailed below.

RNA-seq expression data for 1226 breast cancer samples of the TCGA were downloaded using TCGAbiolinks (version 2.24.0). 924 samples with available ER and HER2 status information were integrated with the NA-PHER2 dataset. All 50 genes used by the PAM50 classifier were present in both datasets and used for the following analyses. Batch effect between TCGA and NA-PHER2 datasets was removed by computing gene-wise median difference between NA-PHER2 baseline samples and HER2+ER+TCGA samples. Including only the NA-PHER2 baseline samples preserved any likely difference in treated samples and avoid biasing the classification of baseline samples. Such median differences were used as scaling factors to be applied to all NA-PHER2 samples. In the original PAM50 training set, 60% of samples were ER+ and 40% ER. Consequently, 466 ER + TCGA samples were discarded to obtain the same proportion of ER+ samples in the combined TCGA-NA-PHER2 dataset. Z-scores were computed per gene only considering baseline samples, to avoid introducing a treatment-related bias, as mentioned above. The scaling factors (i.e. mean and standard deviation) computed on the baseline samples were then applied to Day14 and surgical timepoints. PAM50 classification was achieved using the genefu R package (version 2.26.0). instrinsic.cluster.predict() and molecular.subtyping() functions were modified to skip the scaling step and use our custom scaled integrated dataset. After classification of all samples, probability to belong to each subtype were evaluated and if the probability difference with the second best hit was <0.025, the sample was marked as unclassified.

Gene assignment to cell types using single-cell transcriptomic data

To annotate whether the expression of each gene could be prevalently attributed to a specific cell type of the tumour ecosystem, we interrogated an external single-cell dataset. Single-cell transcriptomic profiles of GSE176078 dataset[22] were downloaded from GEO repository and a SeuratObject was created. A total of 19311 cells from 5 HER2+ patients were used. Pseudobulk samples were obtained by grouping the raw reads per patient and per cell type. Raw reads were TMM normalised using the calcNormFactors() function (edgeR R package, version 3.38.4) and voom transformed using the voom() function (limma R package, version 3.52.2). A total of 14625 genes with maximum logCPM above 3.74 were kept for downstream analyses. The threshold was data driven and a minimum in the distribution was identified using the optimize() R function. For each gene, we identified the cell type with the highest and second highest expression, respectively. The log Fold Change (logFC) was computed between the two cell types and if logFC > 1, the difference was statistically evaluated using limma R package (version 3.52.2). Computed *p*-values were adjusted using the Benjamini-Hochberg method and a gene was assigned to the cell type with highest expression if FDR < 0.1. Genes differentially expressed in the NA-PHER2 were then annotated with the assigned cell-type, when available.

Other statistical analysis

Correlation between two continuous variables was quantified by computing the Pearson correlation coefficient. Differences in the distribution of continuous variables between two groups were evaluated by two-sided Student's t-test. Differences in the distribution of categorical variables among two or more groups were evaluated by two-sided Fisher's exact test.

Association between *MKI67* expression and *TP53* mutation with pCR was evaluated with univariate or bivariate logistic regression models.

In boxplots, the centre line represents the median, box limits represent upper and lower quartiles while whiskers represent 1.5× the interquartile range.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Raw RNA sequencing data for all samples from the NA-PHER2 trial used in this study have been deposited at the European Genome-phenome Archive (https://web2.ega-archive.org/) under accession code: EGAS50000000248. Controlled access will be granted upon request, as detailed on the EGA website. Processed data and molecular annotations have been deposited at ArrayExpress under accession code: E-MTAB-14448. For data protection reasons, clinical information collected during the trial can be obtained upon request to Fondazione Michelangelo by contacting the corresponding authors or by emailing translational@fondazionemichelangelo.org. The remaining data are available within the Article, Supplementary Information or Source Data file. Source data are provided with this paper.

Code availability

No custom computational method has been developed. Analyses were carried out in R/Bioconductor and relevant packages and their version are reported in the methods section.

References

1. Loibl, S. & Gianni, L. HER2-positive breast cancer. *Lancet* **389**, 2415–2429 (2017).
2. Brandão, M. et al. What is the real impact of estrogen receptor status on the prognosis and treatment of HER2-Positive Early Breast Cancer? *Clin. Cancer Res.* **26**, 2783–2788 (2020).
3. Bianchini, G. et al. Immune modulation of pathologic complete response after neoadjuvant HER2-directed therapies in the NeoSphere trial. *Ann. Oncol.* **26**, 2429–2436 (2015).
4. Gianni, L. et al. Neoadjuvant and adjuvant trastuzumab in patients with HER2-positive locally advanced breast cancer (NOAH): Follow-up of a randomised controlled superiority trial with a parallel HER2-negative cohort. *Lancet Oncol.* **15**, 640–647 (2014).
5. Gianni, L. et al. 5-year analysis of neoadjuvant pertuzumab and trastuzumab in patients with locally advanced, inflammatory, or early-stage HER2-positive breast cancer (NeoSphere): a multicentre, open-label, phase 2 randomised trial. *Lancet Oncol.* **17**, 791–800 (2016).
6. Cortazar, P. et al. Pathological complete response and long-term clinical benefit in breast cancer: The CTNeoBC pooled analysis. *Lancet* **384**, 164–172 (2014).
7. von Minckwitz, G. et al. Trastuzumab Emtansine for Residual Invasive HER2-Positive Breast Cancer. *N. Engl. J. Med.* **380**, 617–628 (2019).
8. Bianchini, G. & Gianni, L. The immune system and response to HER2-targeted treatment in breast cancer. *Lancet Oncol.* **15**, e58–e68 (2014).
9. Viganò, L. et al. Modulation of the Estrogen/erbB2 Receptors Cross-talk by CDK4/6 Inhibition Triggers Sustained Senescence in Estrogen Receptor– and ErbB2-positive Breast Cancer. *Clin. Cancer Res* **28**, 2167–2179 (2022).
10. Agostinetto, E. et al. CDK4/6 and PI3K inhibitors: A new promise for patients with HER2-positive breast cancer. *Eur. J. Clin. Invest* **51**, e13535 (2021).
11. Tolaney, S. M. et al. Abemaciclib plus trastuzumab with or without fulvestrant versus trastuzumab plus standard-of-care chemotherapy in women with hormone receptor-positive, HER2-positive advanced breast cancer (monarcHER): a randomised, open-label, phase 2 trial. *Lancet Oncol.* **21**, 763–775 (2020).

12. Gianni, L. et al. Neoadjuvant treatment with trastuzumab and pertuzumab plus palbociclib and fulvestrant in HER2-positive, ER-positive breast cancer (NA-PHER2): an exploratory, open-label, phase 2 study. *Lancet Oncol.* **19**, 249–256 (2018).
13. Gianni, L. et al. Effects of neoadjuvant trastuzumab, pertuzumab and palbociclib on Ki67 in HER2 and ER-positive breast cancer. https://doi.org/10.1038/s41523-021-00377-8.
14. Sheri, A. & Dowsett, M. Developments in Ki67 and other biomarkers for treatment decision making in breast cancer. *Ann. Oncol.* **23**, x219–x227 (2012).
15. Smith, I. et al. Long-term outcome and prognostic value of Ki67 after perioperative endocrine therapy in postmenopausal women with hormone-sensitive early breast cancer (POETIC): an open-label, multicentre, parallel-group, randomised, phase 3 trial. *Lancet Oncol.* **21**, 1443–1454 (2020).
16. Gianni, L. et al. Efficacy and safety of neoadjuvant pertuzumab and trastuzumab in women with locally advanced, inflammatory, or early HER2-positive breast cancer (NeoSphere): A randomised multicentre, open-label, phase 2 trial. *Lancet Oncol.* **13**, 25–32 (2012).
17. Silwal-Pandit, L. et al. The longitudinal transcriptional response to neoadjuvant chemotherapy with and without bevacizumab in breast cancer. *Clin. Cancer Res.* **23**, 4662–4670 (2017).
18. Matikas, A. et al. Dynamic evaluation of the immune infiltrate and immune function genes as predictive markers for neoadjuvant chemotherapy in hormone receptor positive, HER2 negative breast cancer. *Oncoimmunology* **7**, e1466017 (2018).
19. Park, Y. H. et al. Chemotherapy induces dynamic immune responses in breast cancers that impact treatment outcome. *Nat. Commun.* **11**, 1–14 (2020).
20. Hoogstraat, M. et al. Comprehensive characterization of pre- and post-treatment samples of breast cancer reveal potential mechanisms of chemotherapy resistance. *NPJ Breast Cancer* **8**, 1–12 (2022).
21. Brasó-Maristany, F. et al. Phenotypic changes of HER2-positive breast cancer during and after dual HER2 blockade. *Nat. Commun.* **11**, 1–11 (2020).
22. Wu, S. Z. et al. A single-cell and spatially resolved atlas of human breast cancers. *Nat. Genet* **53**, 1334–1347 (2021).
23. Miksicek, R. J. et al. Identification of a novel breast- and salivary gland-specific, mucin-like gene strongly expressed in normal and tumor human mammary epithelium. *Cancer Res* **62**, 2736–2740 (2002).
24. Conley, S. J. et al. HER2 drives Mucin-like 1 to control proliferation in breast cancer cells. *Oncogene 2016* **3532**, 4225–4234 (2016).
25. Barreca, M. et al. Development and validation of a gene expression-based Breast Cancer Purity Score. *npj Precis. Oncol.* **81**, 1–11 (2024).
26. Ruan, C., Meng, Y. & Song, H. CD36: an emerging therapeutic target for cancer and its molecular mechanisms. *J. Cancer Res. Clin. Oncol.* **148**, 1551–1558 (2022).
27. Ligorio, F. et al. Targeting lipid metabolism is an emerging strategy to enhance the efficacy of anti-HER2 therapies in HER2-positive breast cancer. *Cancer Lett.* **511**, 77–87 (2021).
28. Green, A. R. et al. MYC functions are specific in biological subtypes of breast cancer and confers resistance to endocrine therapy in luminal tumours. *Br. J. Cancer* **114**, 917–928 (2016).
29. Unterleuthner, D. et al. Cancer-associated fibroblast-derived WNT2 increases tumor angiogenesis in colon cancer. *Angiogenesis* **23**, 159–177 (2020).
30. Aizawa, T. et al. Cancer-associated fibroblasts secrete Wnt2 to promote cancer progression in colorectal cancer. *Cancer Med.* **8**, 6370–6382 (2019).
31. Nallanthighal, S., Heiserman, J. P. & Cheon, D. J. Collagen Type XI Alpha 1 (COL11A1): A Novel Biomarker and a Key Player in Cancer. *Cancers (Basel)* **13**, 1–22 (2021).
32. Tanei, T. et al. Prognostic significance of Ki67 index after neoadjuvant chemotherapy in breast cancer. *Eur. J. Surg. Oncol.* **37**, 155–161 (2011).
33. Zuo, H. & Wan, Y. Metabolic Reprogramming in Mitochondria of Myeloid Cells. *Cells* **9**, 5 (2019).
34. Nass, N. et al. Expression of transmembrane protein 26 (TMEM26) in breast cancer and its association with drug response. *Oncotarget* **7**, 38408–38426 (2016).
35. Oxvig, C. & Conover, C. A. The Stanniocalcin-PAPP-A-IGFBP-IGF Axis. *J. Clin. Endocrinol. Metab.* **108**, 1624–1633 (2023).
36. Zhang, Y. et al. RNA-binding protein YTHDF3 suppresses interferon-dependent antiviral responses by promoting FOXO3 translation. *Proc. Natl. Acad. Sci. USA* **116**, 976–981 (2019).
37. Su, S. et al. PAK4 inhibition improves PD1 blockade immunotherapy in prostate cancer by increasing immune infiltration. *Cancer Lett.* **555**, 216034 (2023).
38. Abril-Rodriguez, G. et al. PAK4 inhibition improves PD-1 blockade immunotherapy. *Nat. Cancer* **1**, 46–58 (2019).
39. Yuan, Y. et al. PAK4 in cancer development: Emerging player and therapeutic opportunities. *Cancer Lett.* **545**, 215813 (2022).
40. Schneeweiss, A. et al. Pertuzumab plus trastuzumab in combination with standard neoadjuvant anthracycline-containing and anthracycline-free chemotherapy regimens in patients with HER2-positive early breast cancer: A randomized phase II cardiac safety study (TRYPHAENA). *Ann. Oncol.* **24**, 2278–2284 (2013).
41. Van MacKelenbergh, M. T. et al. Pathologic Complete Response and Individual Patient Prognosis After Neoadjuvant Chemotherapy Plus Anti-Human Epidermal Growth Factor Receptor 2 Therapy of Human Epidermal Growth Factor Receptor 2-Positive Early Breast Cancer. *J. Clin. Oncol.* **41**, 2998–3008 (2023).
42. Dieci, M. V. et al. Integrated evaluation of PAM50 subtypes and immune modulation of pCR in HER2-positive breast cancer patients treated with chemotherapy and HER2-targeted agents in the CherLOB trial. *Ann. Oncol.* **27**, 1867–1873 (2016).
43. Carey, L. A. et al. Molecular heterogeneity and response to neoadjuvant human epidermal growth factor receptor 2 targeting in CALGB 40601, a randomized phase III trial of paclitaxel plus trastuzumab with or without lapatinib. *J. Clin. Oncol.* **34**, 542–549 (2016).
44. Prat, A. et al. Research-based PAM50 subtype predictor identifies higher responses and improved survival outcomes in HER2-positive breast cancer in the NOAH study. *Clin. Cancer Res* **20**, 511–521 (2014).
45. Fassl, A., Geng, Y. & Sicinski, P. CDK4 and CDK6 kinases: From basic science to cancer therapy. *Sci. (80-.)* **375**, eabc1495 (2022).
46. Chen, R., Ishak, C. A. & De Carvalho, D. D. Endogenous retroelements and the viral mimicry response in cancer therapy and cellular homeostasis. *Cancer Discov.* **11**, 2707–2725 (2021).
47. Goel, S. et al. CDK4/6 inhibition triggers anti-tumour immunity. *Nat* **548**, 471–475 (2017).
48. Grote, I. et al. TP53 mutations are associated with primary endocrine resistance in luminal early breast cancer. *Cancer Med* **10**, 8581–8594 (2021).
49. Schuster, E. F. et al. Genomic Instability and TP53 Genomic Alterations Associate With Poor Antiproliferative Response and Intrinsic Resistance to Aromatase Inhibitor Treatment. *JCO Precis. Oncol.* **3**, 1–11 (2019).
50. Mancusi, R. & Monje, M. The neuroscience of cancer. *Nat* **618**, 467–479 (2023).
51. Magnon, C. & Hondermarck, H. The neural addiction of cancer. *Nat. Rev. Cancer* **23**, 317–334 (2023).
52. Hendry, S. et al. Assessing Tumor-infiltrating Lymphocytes in Solid Tumors: A Practical Review for Pathologists and Proposal for a Standardized Method from the International Immunooncology Biomarkers Working Group: Part 1: Assessing the Host Immune

Response, TILs in Invasive Breast Carcinoma and Ductal Carcinoma in Situ, Metastatic Tumor Deposits and Areas for Further Research. *Adv. Anat. Pathol.* **24**, 235–251 (2017).
53. Tamborero, D. et al. Cancer Genome Interpreter annotates the biological and clinical relevance of tumor alterations. *Genome Med* **10**, 1–8 (2018).
54. Cast, O. et al. Comprehensive benchmarking and integration of tumor microenvironment cell estimation methods. *Cancer Res* **79**, 6238–6246 (2019).
55. Wilkerson, M. D. & Hayes, D. N. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* **26**, 1572–1573 (2010).
56. Parker, J. S. et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.* **27**, 1160–1167 (2009).
57. Hutter, C. & Zenklusen, J. C. The Cancer Genome Atlas: creating lasting value beyond its data. *Cell* **173**, 283–285 (2018).

Acknowledgements

This work has been supported by Fondazione Michelangelo, responsible for the study design and data collection. Additional support was received by Fondazione AIRC per la ricerca sul cancro (IG2018 - ID21787, P.I. GB), the Breast Cancer Research Foundation (grant BCRF 21-181 to LG) and Pfizer. BGy was supported by National Research, Development and Innovation Office (PharmaLab, RRF-2.3.1-21-2022-00015). LP was supported by the Breast Cancer Research Foundation investigator award (BCRF-22-133) and a Susan Komen Leadership Grant (SAC220225). These founding agencies have no role in data analysis and manuscript writing.

Author contributions

L.G., G.B. and M.Ca. conceived the study and planned the computational and statistical analyses; M.Ca. performed the analyses and drafted the manuscript; M.D., M.B., B.Gy., B.Ga. contributed to the computational and statistical analyses; L.V., A.L., C.D.A., M.F. contributed to the biological interpretation of the results; G.B., M.Co., M.M., C.Z., L.D.M., S.Z., A.F. contributed to the collection of samples and clinical information; O.B. and G.V. performed the IHC evaluations; L.P. contributed to data analysis and interpretation of the results; P.V. collected and curated the clinical information. All authors contributed to subsequent drafts and agreed with submission of the manuscript for publication.

Competing interests

G.V. received grants/research supports from: Roche/Genentech, Ventana Medical Systems, Dako/Agilent Technologies; honoraria or consultation feesfrom: Ventana, Dako/Agilent, Roche, MSD Oncology, AstraZeneca, Daiichi Sankyo, Pfizer, Eli Lilly. All the authors declare no conflicts of interest. L.P. has received consulting fees and honoraria for advisory board participation from Pfizer, Astra Zeneca, Merck, Novartis, Bristol-Myers Squibb, Stemline-Menarini, GlaxoSmithKline, Genentech/Roche, Personalis, Daiichi, Natera, Exact Sciences and institutional research funding from Seagen, GlaxoSmithKline, AstraZeneca, Merck, Pfizer and Bristol Myers Squibb. M.Co. received an institutional research grant from Roche. G.B. reports personal fees for consultancy/honorarium/advisory role: Lilly, Novartis, Pfizer, Roche, AstraZeneca, Amgen, MSD, Chugai, Sanofi, Daiichi Sankyo, EISAI, Gilead, Menarini/stemline, Exact Science, Seagen, Agendia; institutional fee for research grant from Gilead. L.G. reports grants from Roche, Breast Cancer Research Foundation (BCRF), and Pfizer during the conduct of the study; personal fees from AstraZeneca, Ely Lilly and Company, Roche, Pfizer, Seattle Genentics, Artemida Pharma, Synaffix, Menarini Ricerche, and Biomedical Insights; grants and personal fees from Zymeworks and Revolution Medicine; other support from METIS Precision Medicine, Amgen, and QU Biologics outside the submitted work; in addition, L.G. has a patent for EU N. 12195182 issued and a patent for EU N.12196177.5 issued.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-025-57293-9.

Correspondence and requests for materials should be addressed to Luca Gianni or Giampaolo Bianchini.

Peer review information *Nature Communications* thanks Laurence Buisseret, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2025

[1]Fondazione Michelangelo, Milan, Italy. [2]IRCCS San Raffaele Hospital, Milan, Italy. [3]Department of Biotechnology and Biosciences, University of Milano-Bicocca, Milan, Italy. [4]Dept. of Bioinformatics, Semmelweis University, Budapest, Hungary. [5]Dept. of Biophysics, Medical School, University of Pecs, Pecs, Hungary. [6]Cancer Biomarker Research Group, Institute of Molecular Life Sciences, Research Centre for Natural Sciences, Budapest, Hungary. [7]AUSL – IRCCS, Reggio Emilia, Italy. [8]IEO, European Institute of Oncology, IRCCS, Milan, Italy. [9]Udine Academic Hospital, Udine, Italy. [10]IRCCS Azienda Ospedaliero Universitaria, Bologna, Italy. [11]Department of Internal Medicine and Medical Specialties (DiMI), School of Medicine, Università di Genova, Genoa, Italy. [12]Department of Medical Oncology, UO Clinica di Oncologia Medica, IRCCS Ospedale Policlinico San Martino, Genoa, Italy. [13]Azienda Ospedaliera Universitaria S. Anna, Ferrara, Italy. [14]Department of Internal Medicine, Section of Medical Oncology, Yale School of Medicine, New Haven, CT, USA. [15]UniSR San Raffaele University, Milan, Italy. [16]These authors jointly supervised this work: Luca Gianni, Giampaolo Bianchini. ✉e-mail: luca.gianni@fondazionemichelangelo.org; bianchini.giampaolo@hsr.it