nature communications

Article

https://doi.org/10.1038/s41467-024-54621-3

Identification of HER2-positive breast cancer molecular subtypes with potential clinical implications in the ALTTO clinical trial

Received: 26 March 2024

Accepted: 13 November 2024

Published online: 29 November 2024

Check for updates

Mattia Rediti[1,2], David Venet[1], Andrea Joaquin Garcia[1], Marion Maetens[3], Delphine Vincent[1], Samira Majjaj[1], Sarra El-Abed[4], Serena Di Cosimo[5], Takayuki Ueno[6], Miguel Izquierdo[7], Martine Piccart[8], Lajos Pusztai[9], Sherene Loi[10,11], Roberto Salgado[10,12], Giuseppe Viale[13], Françoise Rothé[1] & Christos Sotiriou[1]

In HER2-positive breast cancer, clinical outcome and sensitivity to HER2-targeted therapies are influenced by both tumor and microenvironment features. However, we are currently unable to depict the molecular heterogeneity of this disease with sufficient granularity. Here, by performing gene expression profiling in HER2-positive breast cancers from patients receiving adjuvant trastuzumab in the ALTTO clinical trial (NCT00490139), we identify and characterize five molecular subtypes associated with the risk of distant recurrence: immune-enriched, proliferative/metabolic-enriched, mesenchymal/stroma-enriched, luminal, and ERBB2-dependent. Additionally, we validate the biological profiles of the subtypes and explore their prognostic/predictive value in external cohorts, namely the NeoALTTO trial (NCT00553358), SCAN-B (NCT02306096), I-SPY2 (NCT01042379), METABRIC and TCGA. Immune-enriched tumors present better survival outcomes, in contrast to mesenchymal/stroma-enriched and proliferative/metabolic-enriched tumors, while luminal and ERBB2-dependent tumors are characterized by low and high rates of pathological complete response, respectively. Of note, these molecular subtypes provide the rationale for treatment approaches leveraging the heterogeneous biology of HER2-positive breast cancer.

Since the demonstration of the prognostic and predictive role of HER2 amplification in breast cancer[1], the clinical outcome of early-stage HER2-positive breast cancer has been profoundly reshaped. Indeed, the introduction of the monoclonal anti-HER2 antibody trastuzumab in the adjuvant setting[2] paved the way for a revolution in the treatment landscape of HER2-positive breast cancer, and was followed by the development of effective treatment strategies in both the adjuvant[3–8] and neoadjuvant/post-operative setting[9–14].

[1]Breast Cancer Translational Research Laboratory, Institut Jules Bordet, Hôpital Universitaire de Bruxelles (H.U.B), Université Libre de Bruxelles (ULB), Brussels, Belgium. [2]IFOM ETS, the AIRC Institute of Molecular Oncology, Milan, Italy. [3]Laboratory for Translational Breast Cancer Research, KU Leuven, Leuven, Belgium. [4]Breast International Group, Brussels, Belgium. [5]Department of Advanced Diagnostics, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy. [6]Breast Surgical Oncology, Breast Oncology Center, Cancer Institute Hospital, Japanese Foundation for Cancer Research, Tokyo, Japan. [7]Novartis Pharma AG, Basel, Switzerland. [8]Institut Jules Bordet, Hôpital Universitaire de Bruxelles (H.U.B), Université Libre de Bruxelles (ULB), Brussels, Belgium. [9]Yale School of Medicine, Yale Cancer Center, New Haven, CT, USA. [10]Division of Cancer Research, Peter MacCallum Cancer Centre, Melbourne, VIC, Australia. [11]The Sir Peter MacCallum Department of Medical Oncology, The University of Melbourne, Parkville, VIC, Australia. [12]Department of Pathology, ZAS Hospitals, Antwerp, Belgium. [13]Division of Pathology, IEO, European Institute of Oncology IRCCS, Milan, Italy. e-mail: christos.sotiriou@hubruxelles.be

While the mainstay of HER2-positive breast cancer treatment is represented by targeting the HER2 tyrosine kinase, the presence of higher levels of the HER2 protein alone is not sufficient to guarantee durable responses. Several resistance mechanisms may occur, and the prediction of response and relapse-risk at the patient's level remains challenging[15]. Indeed, prognosis and treatment response can be impacted by a complex network of both tumor cells-intrinsic features [e.g., HER2 downstream signaling and *ERBB2* gene amplification[16–20], luminal differentiation and estrogen receptor (ER) levels[16,19–21], proliferation[16,19]] and microenvironment characteristics [e.g., immune infiltration described by tumor-infiltrating lymphocytes (TILs) and immune signatures[16,19,22–25], stroma activation[19,26–28]]. Overall, data collected in biomarker studies point towards a substantial biological heterogeneity across HER2-positive tumors, with different processes driving treatment resistance and outcome. Nonetheless, this inter-tumor heterogeneity is currently not considered when planning treatment strategies.

The therapeutic algorithm of HER2-positive breast cancer is mostly modulated by tumor stage and, in patients receiving neoadjuvant therapies, by the assessment of response at surgery[29]. Importantly, not considering the biology of the disease could potentially lead to over- or under-treatment in a large proportion of patients. In this regard, the HER2DX assay has shown the ability to predict response [three groups based on the likelihood of achieving pathological complete response (pCR)] and outcome (two prognostic groups) in HER2-positive breast cancer patients using 4 gene modules describing immune activation, luminal phenotype, proliferation and the expression of the HER2 amplicon, as well as clinical features[16]. In addition to these advancements, a refined definition of molecular subgroups of HER2-positive breast cancer may potentially allow to better tailor treatment strategies.

In this retrospective exploratory analysis of the phase III ALTTO trial[6], we identify gene expression-based features and five molecular subtypes associated with distant recurrence and overall survival in a case-control cohort of patients receiving adjuvant trastuzumab. These subtypes, depicting both tumor-intrinsic and microenvironment features, reveal a substantial heterogeneity associated with outcome in HER2-positive breast cancer. Moreover, aiming to validate our findings, we identify the molecular subtypes in several external datasets, including the phase III neoadjuvant NeoALTTO trial as well as HER2-positive tumors from the METABRIC, TCGA, I-SPY2, and SCAN-B studies. In-depth gene expression profiling and evaluation of the prognostic/predictive value of the HER2-positive subtypes demonstrate the robustness of our findings, which provide the rationale for the evaluation of tailored treatment approaches modulated by the biology of this disease.

Results

Baseline patients' characteristics of the ALTTO case-control cohort

A case-control cohort (1:2 ratio) of patients affected by HER2-positive breast cancer was identified in the trastuzumab-only arm of the ALTTO clinical trial, based on the presence (cases) or absence (controls) of a distant recurrence event after median follow-up of 6.9 years. A total of 134 cases and 268 controls matched for clinicopathological characteristics (details in METHODS) were identified (total $N = 402$), and their characteristics are summarized in Supplementary Data 1. The median age of this cohort was 51 years (interquartile range 44-58). The case-control cohort, balanced for the clinicopathological characteristics used in the matching procedure, included a high-risk population enriched for distant relapse events (occurring in 1 patient out of 3), with higher proportions of >2 cm, node-positive and G3 tumors compared to the whole ALTTO trastuzumab arm (Supplementary Data 2). The median follow-up for this cohort was 7 years.

After quality check (Supplementary Fig. 2, details in METHODS), RNA sequencing data were analyzed for 386/402 (96%) patients, including 128 cases and 258 controls. The characteristics of this cohort did not differ compared to the whole case-control population and are reported in Supplementary Data 1. The CONSORT diagram showing the selection of the patients is depicted in Supplementary Fig. 1. All the subsequent analyses are referred to the RNA sequencing cohort.

Association of intrinsic subtypes and gene expression signatures with clinical outcome

As intrinsic subtypes showed a predictive and prognostic value in HER2-positive breast cancers[10,19,20,30], we first aimed to explore their prognostic role in this case-control cohort. Intrinsic subtypes were identified with the AIMS classifier[31]. Most of the tumors were classified as HER2-enriched (HER2-E; $N = 257/386$, 66.6%), followed by luminal B ($N = 82/386$, 21.2%) (Supplementary Fig. 3). As expected, the majority of luminal A ($N = 11/12$) and B ($N = 80/82$) tumors were hormone receptor-positive, while basal-like tumors ($N = 16/17$) were almost all hormone receptor-negative.

The HER2-E subtype was associated with worse distant relapse-free interval (DRFI) and overall survival (OS) at univariable and multivariable analyses (Supplementary Fig. 4), while luminal A/B tumors showed improved outcomes compared to the other AIMS subtypes.

We then evaluated the prognostic role of 34 gene expression signatures spanning tumor- and microenvironment-related biological processes, as well as *ERBB2* and *ESR1* expression levels (sources and genes included available in Supplementary Data 3).

At univariable and multivariable analyses (Supplementary Fig. 5), immune signatures, including B cell-related ones such as the immunoglobulin G (IgG) signature included in the HER2DX assay[16,32] and a signature derived from spatial transcriptomics data describing tertiary lymphoid structures (TLS)[33], were associated with a lower risk of metastatic recurrence. Conversely, higher levels of signatures depicting activation of HER2 downstream signaling (AKT/mTOR, RAS), as well as metabolism (glycolysis, pentose phosphate pathway), IGF1 signaling, stroma activation, and hypoxia showed negative prognostic value. Similar results were observed for OS, with the addition of the luminal HER2DX signature[16] being associated with better outcomes (Supplementary Fig. 6). Correlations between signatures (Supplementary Fig. 7, Supplementary Data 4) showed that those describing similar processes (e.g., signatures related to immune response, or stroma activation, or proliferation) were, as expected, highly correlated, whereas signatures having a similar impact on the outcome but describing different processes could be weakly or non-significantly correlated (e.g., luminal HER2DX with immune signatures, both associated to improved OS; HER2 STROMA signature with IGF1 signaling, glycolysis or AKT/mTOR signaling).

Overall, these results confirmed the positive prognostic effect of immune activation and luminal differentiation in patients treated with adjuvant trastuzumab, while several biological pathways including metabolic and stroma-related processes, as well as the HER2-E AIMS subtype, were associated with worse outcome.

Identification of HER2-positive breast cancer molecular subtypes associated with the risk of distant relapse

Our findings highlighted the association of heterogeneous biological processes to the risk of metastatic recurrence. Thus, we reasoned that adopting a supervised top-down approach could allow the identification of prognostic subgroups characterized by distinct molecular profiles (i.e., starting from the outcome information, selecting relevant features able to detect distinct subgroups). As utilizing gene expression signatures can potentially limit the findings to pre-specified processes, we decided to identify single genes associated with DRFI starting from the whole transcriptome (details in METHODS). At a multivariable analysis adjusted for

https://doi.org/10.1038/s41467-024-54621-3

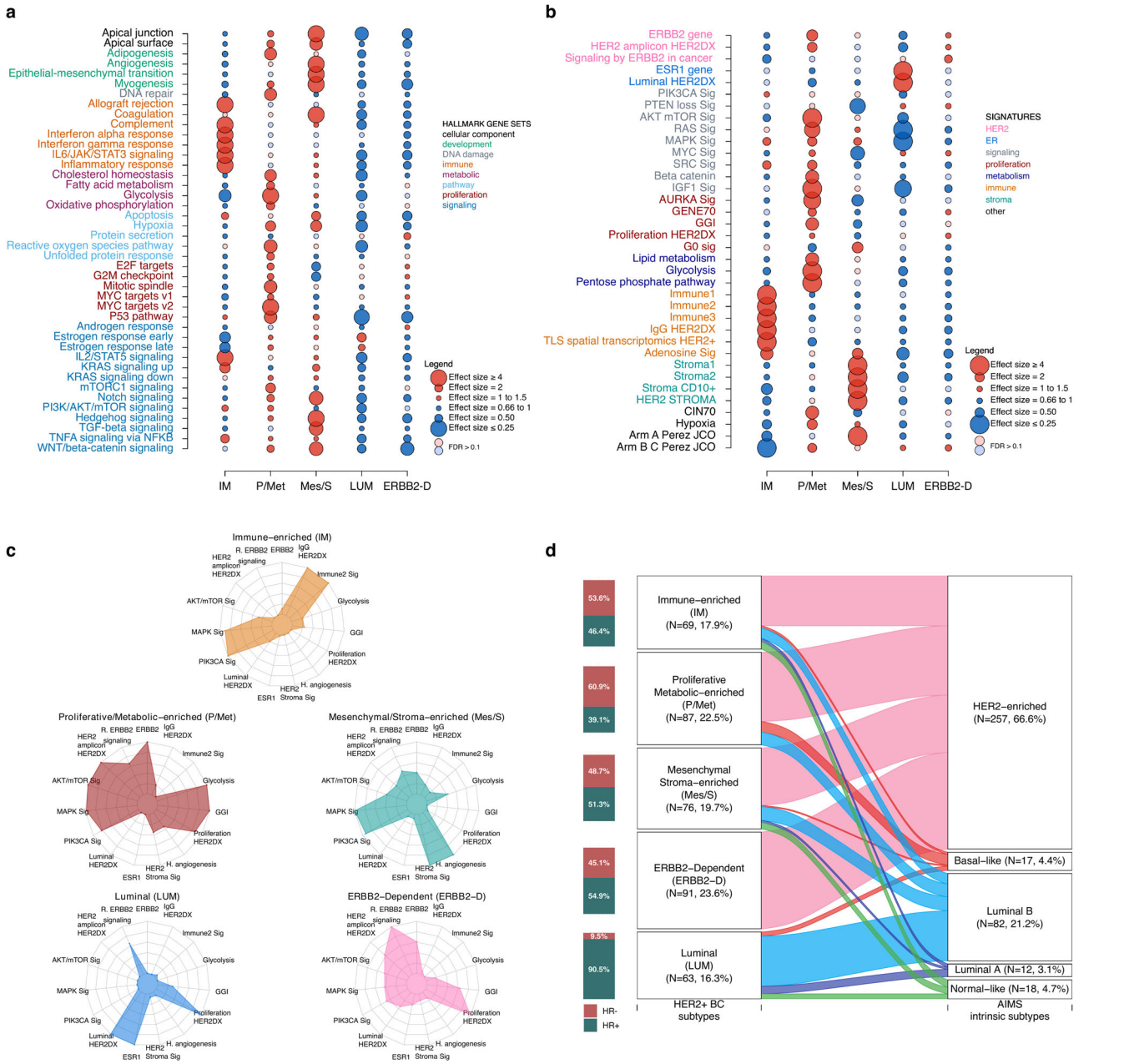

Fig. 1 | Characteristics of the HER2-positive breast cancer subtypes identified in ALTTO. a Characterization of the 5 HER2-positive breast cancer subtypes using hallmark gene sets computed performing gene set variation analysis. **b** Comparisons across HER2-positive breast cancer subtypes using single genes (*ERBB2*, *ESR1*) and gene expression signatures. For (**a**) and (**b**), Wilcoxon rank sum test was used to compare each subtype against the others (i.e., one vs. rest). The effect size represents the direction of the association (red, positive if >1, blue, negative if <1), and was calculated by applying a linear regression model. FDRs were obtained by adjusting *P* values with Benjamini & Hochberg method. The dimension of the circle varies proportionally to the effect size (see figure legend). For visualization, effect sizes ≤0.25, ≥0.66 and ≤1, >1 and ≤1.5, and ≥4 have fixed size; FDRs >0.1 are shown with lighter colors. **c** Radar plots showing the main features of the subtypes, using a selected set of signatures and single genes (*ERBB2*, *ESR1*). Median levels of the signature/gene levels for each subtype were min/max rescaled between 0 and 1 for visualization purposes. The score for H. angiogenesis was derived from GSVA, while the other signatures were computed as weighted mean as described in the methods. **d** Alluvial plot showing the distribution of AIMS intrinsic subtypes across HER2-positive breast cancer subtypes, and vice versa. The proportions of hormone receptor-positive and negative tumors in each HER2-positive breast cancer subtype are also shown. The LUM and ERBB2-D subtypes were derived from NMF and k-means-based clustering (cluster 4) after evaluating differences in survival according to AIMS subtypes (HER2-enriched vs. rest). All *P* values are two-sided. All analyses are on *N* = 386 ALTTO samples (IM, *N* = 69; P/Met, *N* = 87; Mes/S, *N* = 76; LUM, *N* = 63; ERBB2-D, *N* = 91). Source data are provided as a Source Data file. BC breast cancer, ER estrogen receptor, ERBB2-D ERBB2-dependent, FDR false discovery rate, H. Hallmark, HR hormone receptor, IM immune-enriched, LUM luminal, Mes/S mesenchymal/stroma-enriched, NMF non-negative matrix factorization, P/Met proliferative/metabolic-enriched, R. Reactome.

clinicopathological characteristics and AIMS subtypes, 1245 genes remained significantly associated with DRFI after controlling for multiple testing [false discovery rate (FDR) < 0.05; Supplementary Data 5]. Those genes were then summarized into 4 factors by applying dimensionality reduction through non-negative matrix factorization (NMF; Supplementary Fig. 8a). The number of factors was decided by measuring the mean Kullback-Leibler divergence for 1 to 8 factors across different NMF runs, and based on our a priori knowledge derived from the gene expression signature analyses (details in the METHODS). Although the occurrence of a distant

https://doi.org/10.1038/s41467-024-54621-3

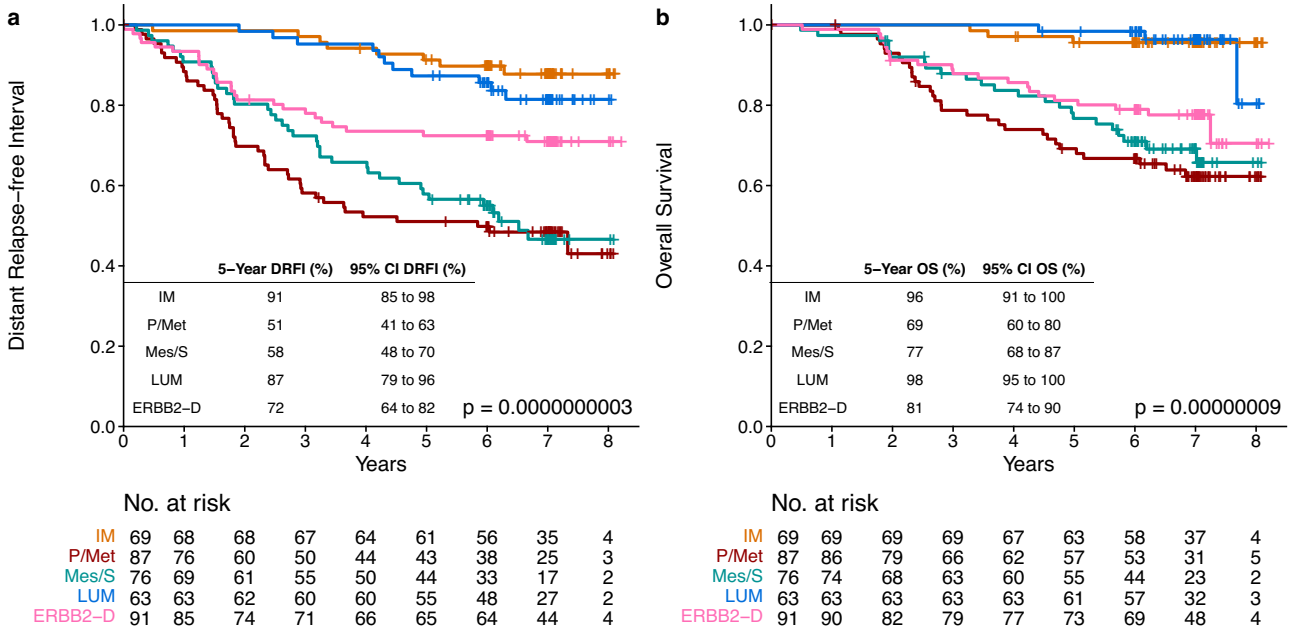

Fig. 2 | Survival analysis according to HER2-positive breast cancer subtypes identified in ALTTO. a Kaplan–Meier plot showing DRFI in ALTTO according to the 5 HER2-positive subtypes identified from the integration of NMF, k-means clustering and AIMS intrinsic subtypes. **b** Kaplan–Meier plot showing OS according to the 5 HER2-positive subtypes. *P* values (two-sided) are from log-rank test. All analyses are on $N = 386$ ALTTO samples (IM, $N = 69$; P/Met, $N = 87$; Mes/S, $N = 76$; LUM, $N = 63$; ERBB2-D, $N = 91$). CI confidence interval, DRFI distant relapse-free interval, ERBB2-D ERBB2-dependent, IM immune-enriched, LUM luminal, Mes/S mesenchymal/stroma-enriched, NMF non-negative matrix factorization, OS overall survival, P/Met proliferative/metabolic-enriched.

relapse event in the history of the disease could be explained by two groups of genes associated with either better or worse prognosis, identifying a higher number of factors through NMF (i.e., more than two) could allow the identification of different processes explaining similar outcomes, as we have shown in the previous section. We next performed k-means clustering on the factor scores obtained, to classify tumors into distinct subgroups enriched for the identified factors (Supplementary Fig. 8b). This analysis led us to the identification 4 prognostic clusters of HER2-positive breast cancer (Supplementary Fig. 9), which were further characterized performing Gene Set Variation Analysis (GSVA) on the hallmark gene sets and integrating the AIMS intrinsic subtype information.

GSVA analysis showed that clusters 1, 2, and 3 presented distinct biology described by immune signaling (cluster 1), proliferation and metabolism (cluster 2), and stroma-related processes (cluster 3), whereas cluster 4, the largest one ($N = 154$), did not show peculiar features, suggesting a more mixed nature (Supplementary Fig. 10, Supplementary Data 6). We then evaluated whether AIMS intrinsic subtypes could still provide prognostic information within each cluster. Differences in outcome, particularly for OS, were observed only in cluster 4, in which non-HER2-E tumors (the majority, $N = 54/63$, being luminal A/B) presented better prognosis (Supplementary Fig. 11). Indeed, multivariable analyses showed that, in cluster 4, HER2-E tumors presented worse outcome compared to the rest [for DRFI: hazard ratio (HR) = 2.3; 95% confidence interval (CI), 1.1–5; *P* value = 0.028; for OS: HR = 6.6; CI, 1.9–23; *P* value = 0.00059]. Following this observation and given the prognostic role of luminal signaling/subtypes described in the previous section, we divided cluster 4 into two groups according to AIMS intrinsic subtypes (HER2-E vs. others), resulting in 5 final HER2-positive subtypes.

Based on gene expression profiling (Fig. 1, Supplementary Data 7 and 8), we identified immune-enriched tumors (IM, $N = 69$, from cluster 1, presenting higher levels of immune signatures), proliferative/metabolic-enriched tumors (P/Met, $N = 87$, from cluster 2, characterized by proliferation signatures as well as glycolysis, oxidative phosphorylation, AKT/mTOR, IGF1, MAPK, MYC signaling, and high *ERBB2* expression), mesenchymal/stroma-enriched tumors [Mes/S, $N = 76$, from cluster 3, mainly characterized by stroma-related signatures, epithelial-mesenchymal transition (EMT), angiogenesis, TGF-beta/Notch/Hedgehog/WNT and MAPK signaling, as well as low proliferation]; moreover, from cluster 4 split based on AIMS, we identified the subtypes luminal (LUM, $N = 63$, with high estrogen/luminal signaling), and ERBB2-dependent (ERBB2-D, $N = 91$, characterized by *ERBB2* signaling, high *ERBB2* levels and intermediate proliferation).

Of interest, AIMS intrinsic subtypes, particularly HER2-E tumors, were redistributed across the HER2-positive subtypes (Fig. 1d, Supplementary Data 9). This is expected in consideration of the different purposes of the approaches used to identify the subtypes (focusing on HER2-positive breast cancer and outcome in our case, depicting general breast cancer molecular profiles for intrinsic subtypes[34]). Nonetheless, this highlights a certain degree of heterogeneity within AIMS intrinsic subtypes (particularly HER2-E) when considering both tumor and microenvironment features in the context of HER2-positive breast cancer. In addition, the majority of the basal-like became part of the P/Met subtype, while, also by design, most luminal A/B tumors were included in the LUM subtype. The IM and LUM tumors showed better DRFI compared to the other subtypes (Fig. 2), presenting 5-year DRFI of 91% (CI, 85–98%) and 87% (CI, 79–96%), respectively, as well as excellent OS (in IM, 5-year OS of 96%; CI, 91–100; in LUM, 5-year OS of 98%, CI, 95–100%). A worse outcome was instead observed in P/Met and Mes/S tumors.

With regards to clinicopathological characteristics, hormone receptor status differed significantly across subtypes (Supplementary Data 9). In particular, almost all (90.5%) LUM tumors were hormone receptor-positive; however, only a minority of hormone receptor-positive tumors were classified as LUM ($N = 57/212$, 26.9%).

Overall, these findings unveil how inter-tumor heterogeneity of HER2-positive breast cancer, depicted by different transcriptomic profiles capturing both tumor-intrinsic and microenvironment features, is able to provide prognostic information in patients treated with adjuvant trastuzumab.

https://doi.org/10.1038/s41467-024-54621-3

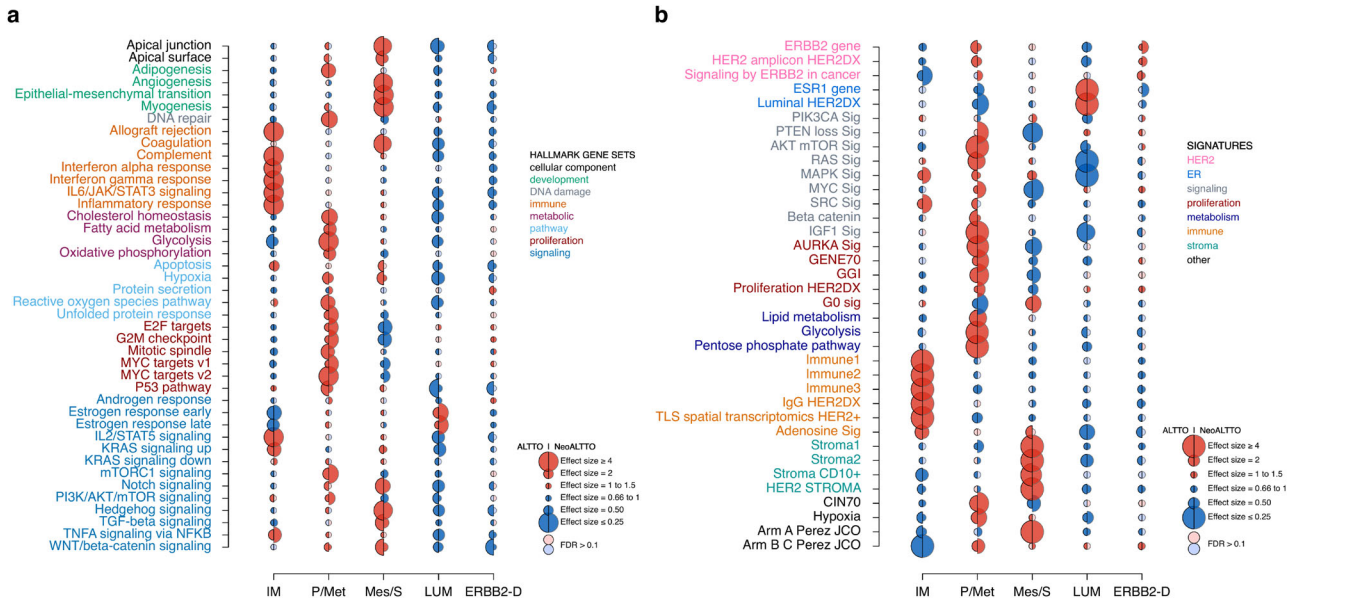

Fig. 3 | Validation of the biological characteristics of the 5 HER2-positive breast cancer subtypes using hallmark gene sets and single genes/gene signatures in NeoALTTO, and comparison with ALTTO original subtypes. a Comparisons across HER2-positive breast cancer subtypes using hallmark gene sets computed performing gene set variation analysis in ALTTO (left half of the circles) and NeoALTTO (right half of the circles). **b** Comparisons across HER2-positive breast cancer subtypes using single genes (*ERBB2*, *ESR1*) and gene expression signatures in ALTTO (left half of the circles) and NeoALTTO (right half of the circles). Wilcoxon rank-sum test was used to compare each subtype against the others (i.e., one vs. rest). The effect size represents the direction of the association (red, positive if >1, blue, negative if <1), and was calculated by applying a linear regression model. FDRs were obtained by adjusting *P* values with Benjamini & Hochberg method. The dimension of the circle varies proportionally to the effect size (see figure legend). For visualization, effect sizes ≤0.25, ≥0.66 and ≤1, >1 and ≤1.5, and ≥4 have fixed size; FDRs >0.1 are shown with lighter colors. The left halves of the circle represent the effect size for the original subtypes derived from the integration of NMF, k-means clustering and AIMS intrinsic subtypes in ALTTO, while the right halves represent results for NeoALTTO. All *P* values are two-sided. All analyses are on $N = 386$ ALTTO samples and $N = 254$ NeoALTTO samples (IM, $N = 42$; P/Met, $N = 41$; Mes/S, $N = 59$; LUM, $N = 52$; ERBB2-D, $N = 60$). Source data are provided as a Source Data file. ER estrogen receptor, ERBB2-D ERBB2-dependent, FDR false discovery rate, IM immune-enriched, LUM luminal, Mes/S mesenchymal/stroma-enriched, NMF non-negative matrix factorization, P/Met proliferative/metabolic-enriched.

Development of a gene expression-based classifier to identify the HER2-positive breast cancer subtypes in external cohorts

Following the identification of the HER2-positive molecular subtypes in ALTTO, we aimed to identify them in external early-stage breast cancer cohorts, namely NeoALTTO[19], I-SPY2[35], METABRIC[36], TCGA[34], and SCAN-B[37], while maintaining their most relevant features.

To address this objective, we developed a gene expression-based classifier in ALTTO (details regarding the development and application of the classifier reported in the METHODS section). First, we performed differential expression analysis comparing each subtype with the rest. A total of 365 differentially expressed genes (FDR < 0.01, $|\log_2(\text{fold change})| > 1.5$; Supplementary Data 10) were identified and used in a LASSO (least absolute shrinkage and selection operator) multinomial classifier, which further reduced the number of features to a final set of 87 genes. Gene coefficients for each subtype are available in Supplementary Data 11. Assessing the final classifier in the ALTTO cohort, we obtained an F1 score (i.e., the harmonic mean of precision and recall) macro average of 0.875 and an accuracy of 0.876 (cross-validation performance and details are reported in the METHODS). The largest discordance was given by 10 samples that were reclassified from ERBB2-D to LUM (Supplementary Data 12). As these two groups were defined using AIMS (i.e., splitting cluster 4) and not directly by NMF/k-means, we could expect to have more uncertainty; to explore this, we evaluated the most relevant genes in these subtypes, namely *ESR1* and *ERBB2*. Indeed, a possible explanation for this discordance may be the higher *ESR1* expression levels in samples reclassified from ERBB2-D to LUM compared to the ones concordant for ERBB2-D (Supplementary Fig. 12).

In ALTTO, the subtypes identified with this approach conserved their biological characteristics compared to the original ones (Supplementary Fig. 13, Supplementary Data 13 and 14) as well as their prognostic value (Supplementary Fig. 14). In addition to performing classification, this methodology allows to compute signature scores for each subtype as a continuous variable as well. At uni- and multivariable analyses (Supplementary Fig. 15), IM and LUM tumors showed better DRFI and OS (multivariable analyses; IM: HR = 0.35; CI, 0.19–0.63; FDR = 0.00051 for DRFI; HR = 0.26; CI, 0.11–0.61; FDR = 0.0017 for OS; LUM: HR = 0.45; CI, 0.25–0.81; FDR = 0.0062 for DRFI; HR = 0.22; CI, 0.08–0.63; FDR = 0.0018 for OS), while P/Met and Mes/S tumors presented worse outcomes. Similar results were observed for their respective signature scores.

Despite intrinsic challenges related to the applications of classifiers across platforms with technical differences and datasets with potential biases (discussed in the METHODS), the classifier was able to detect the HER2-positve subtypes when applied to the selected external cohorts, with distributions that are shown in Supplementary Fig. 16. The highest proportions of IM tumors were found in I-SPY2 and SCAN-B (24.9% and 23.6%, respectively), while LUM tumors were particularly represented in TCGA (24.4%) and SCAN-B (26.6%).

To assess the robustness of the subtypes, we evaluated whether their gene expression profiles as described by hallmark gene sets and signatures were similar to the original ones in ALTTO. Importantly, in NeoALTTO (Fig. 3, Supplementary Data 15 and 16) as well as in the other cohorts (Supplementary Figs. 17 and 18, Supplementary Data 17–24) we observed remarkable similarities, with each subtype conserving distinct profiles characterized by the most relevant biological features. In addition, in I-SPY2 we compared our subtypes according to signature scores and biomarker levels (including microarray and reverse phase protein array) obtained from Wolff et al.[35]. Immune signatures, with the exception of a mast cell signature, were

https://doi.org/10.1038/s41467-024-54621-3

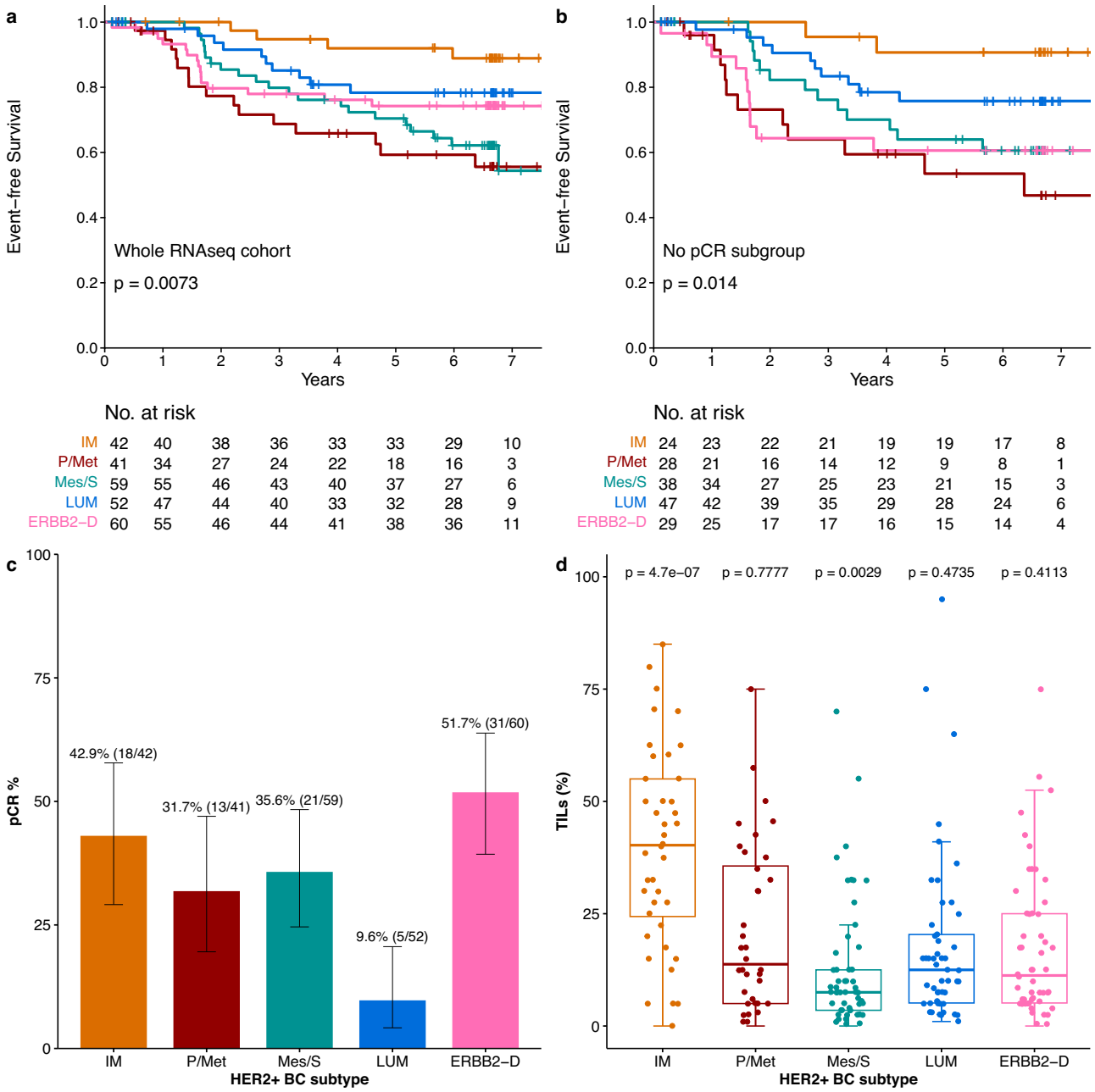

Fig. 4 | Event-free survival, pathological complete response rates, and TIL levels according to HER2-positive breast cancer subtypes in NeoALTTO. **a** Kaplan–Meier plot showing EFS in NeoALTTO according to the 5 HER2-positive subtypes identified with the gene expression-based classifier. **b** Kaplan–Meier plot showing EFS according to the 5 HER2-positive subtypes in the subset with no pCR (ypT0/is) at surgery. **c** Rates (%) of pCR (ypT0/is) according to the 5 HER2-positive subtypes in the NeoALTTO trial. **d** TIL levels according to the 5 HER2-positive subtypes in the NeoALTTO trial. For EFS, *P* values are from log-rank test. For the comparisons of TIL levels across the subtypes (i.e., one vs. the rest), *P* values are derived from Wilcoxon rank sum tests. All *P* values are two-sided. In the bar plots, the whiskers indicate the 95% confidence interval. In box plots, the boxes are defined by the upper and lower quartile; the median is shown as a bold-colored horizontal line; whiskers extend to the most extreme data point which is no more than 1.5 times the interquartile range from the box. Analyses in (**a**) and (**c**) are on $N = 254$ NeoALTTO samples (IM, $N = 42$; P/Met, $N = 41$; Mes/S, $N = 59$; LUM, $N = 52$; ERBB2-D, $N = 60$); analyses in (**b**) are on $N = 166/254$ NeoALTTO samples with no pCR (IM, $N = 24$; P/Met, $N = 28$; Mes/S, $N = 38$; LUM, $N = 47$; ERBB2-D, $N = 29$); analyses in (**d**) are on $N = 233/254$ NeoALTTO samples with TILs available (IM, $N = 40$; P/Met, $N = 36$; Mes/S, $N = 53$; LUM, $N = 50$; ERBB2-D, $N = 54$). BC breast cancer, EFS event-free survival, ERBB2-D ERBB2-dependent, IM immune-enriched, LUM luminal, Mes/S mesenchymal/stroma-enriched, pCR pathological complete response, P/Met proliferative/metabolic-enriched, RNAseq RNA sequencing, TIL tumor-infiltrating lymphocyte.

higher in IM, while the P/Met subtype presented higher levels of proliferation signatures, the Mes/S subtype was enriched for an extracellular matrix gene module, luminal signatures were higher in LUM, and the "HER2ness" signatures[35] mainly characterized the ERBB2-D subtype (Supplementary Fig. 19, Supplementary Data 25). Overall, the main biological features of the HER2-positive subtypes were maintained in different datasets, suggesting the robustness of our findings, and providing a proof-of-concept for the identification of our subtypes in other cohorts.

Validation of the clinical characteristics of the HER2-positive breast cancer subtypes in external cohorts

We next aimed to explore the impact of the HER2-positive breast cancer subtypes on survival outcomes in external cohorts, to confirm

https://doi.org/10.1038/s41467-024-54621-3

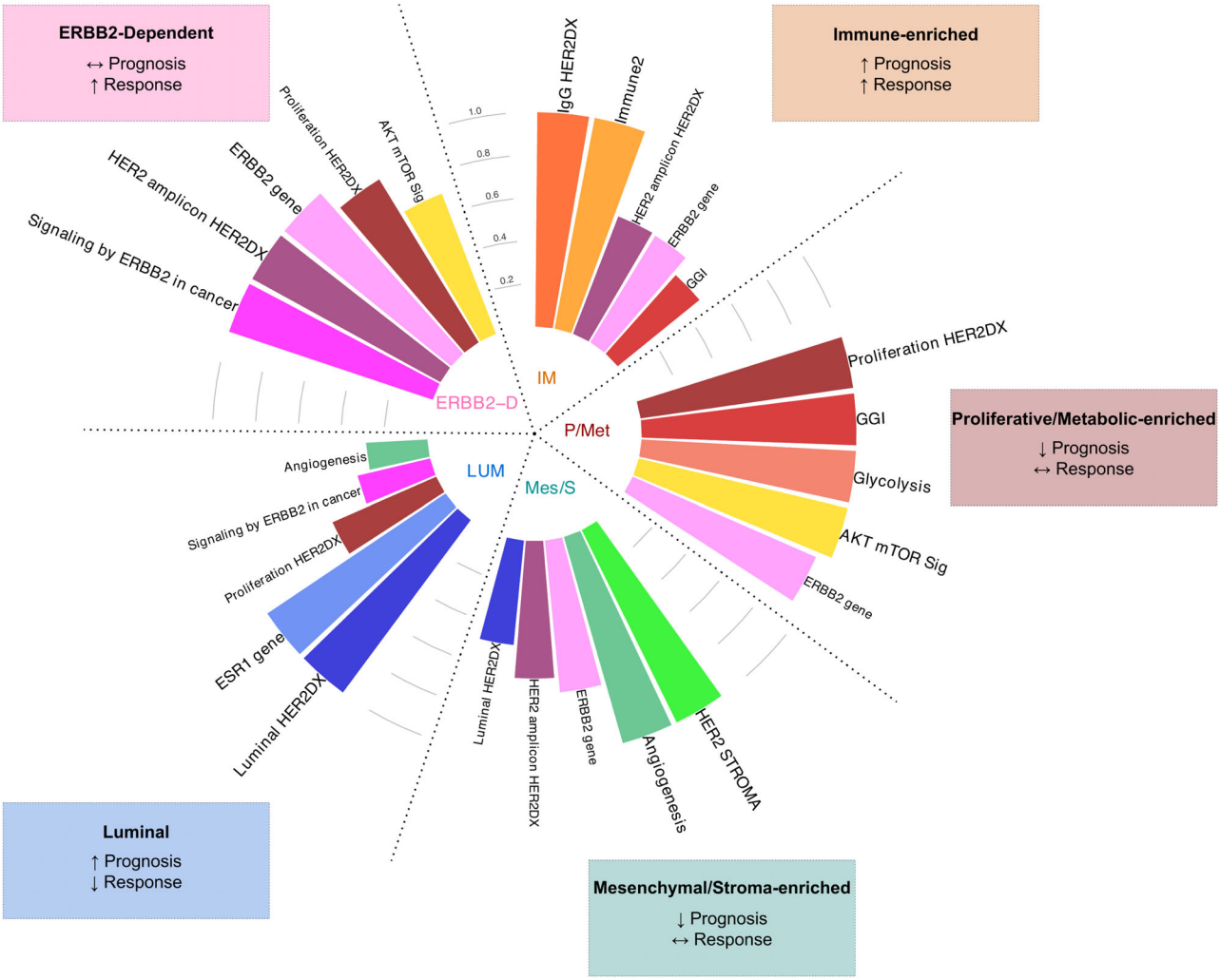

Fig. 5 | Main characteristics of the 5 HER2-positive breast cancer subtypes. Circular bar plot showing the main gene expression-related features of the HER2-positive subtypes. Scores for gene signatures, expression levels of single genes (*ERBB2* and *ESR1*), and gene set variation (angiogenesis) scores (total of 13 scores) were min/max rescaled in each dataset (ALTTO $N = 386$, NeoALTTO $N = 254$, I-SPY2 $N = 245$, TCGA $N = 131$, METABRIC $N = 236$, SCAN-B $N = 819$), and median levels computed for each subtype in a merged dataset ($N = 2071$). For visualization purposes, values were again min/max rescaled between 0 and 1, and top 5 scores in each subtype are shown. Colored panels for each subtype summarize the main clinical characteristics in terms of prognosis and response to neoadjuvant therapy. Source data are provided as a Source Data file. ERBB2-D ERBB2-dependent, IM immune-enriched, LUM luminal, Mes/S mesenchymal/stroma-enriched, P/Met proliferative/metabolic-enriched.

their prognostic value outside of ALTTO. As we initially focused on the occurrence of distant relapse and overall survival in ALTTO, we evaluated endpoints including distant relapse-free interval/survival and OS whenever available. In addition, we evaluated their ability to predict pCR in the neoadjuvant NeoALTTO and I-SPY2 trials.

In NeoALTTO, event-free survival (EFS, median follow-up of 6.7 years) was significantly different across subtypes, both in the whole RNA sequencing cohort (log-rank P value = 0.0073) and, interestingly, in the subgroup not achieving pCR (ypT0/is; log-rank P value = 0.014) (Fig. 4a, b). Significant differences in terms of OS were observed as well (Supplementary Figs. 20a, b). In particular, IM tumors presented better outcomes compared to the other subtypes, followed by LUM. Of note, IM and ERBB2-D tumors presented higher pCR rates (42.9% and 51.7%, respectively), while only 9.6% of the LUM tumors achieved pCR (Fig. 4c). Rates of pCR by arm are reported in Supplementary Fig. 20c and suggest a sensitivity to dual HER2 blockade for IM and ERBB2-D tumors, even though the small numbers demand cautious interpretation. Uni- and multi-variable analyses for EFS and pCR confirmed these results, while results for OS were not significant after adjusting for multiple tests, possibly due to the lower number of events observed for OS compared to EFS (Supplementary Fig. 21). Moreover, in the 233/254 patients with TILs data available, IM tumors, as expected, presented higher TILs as compared to other tumors, while lower TIL levels were observed in the Mes/S subtype (Fig. 4d).

Differences in sensitivity to neoadjuvant treatments were confirmed in the I-SPY2 cohort. Indeed, IM and ERBB2-D tumors presented pCR rates of 67.2% and 72.4%, respectively, while the LUM subtype achieved pCR in 20.4% of the cases (Supplementary Fig. 22a). The association of IM, ERBB2-D, and LUM subtypes with pCR was further confirmed at uni- and multi-variable analyses (Supplementary Fig. 23). Responses by subtype and treatment arm are reported in Supplementary Fig. 22b, and although these results are purely descriptive due to the small numbers in each arm, we noticed interesting signals pointing toward high activity of T-DM1 + pertuzumab in IM and ERBB2-D, as well as of paclitaxel + pertuzumab + trastuzumab in P/Met and ERBB2-D. Intriguingly, all the three patients with Mes/S tumors who received the antiangiogenic compound AMG 386 achieved pCR.

The prognostic effect of the HER2-positive subtypes was confirmed in the SCAN-B study, when evaluating in the whole population RFI (log-rank P value = 0.00049), DRFI (log-rank P value = 0.002), and

OS (log-rank P value < 0.0001) (Supplementary Fig. 24a–c), as well as when limiting the analysis to the subgroup receiving anti-HER2 treatment (Supplementary Fig. 24d-f). In particular, IM tumors presented a better prognosis compared to the others, whereas P/Met tumors had the worse outcome, as also shown in the uni- and multi-variable analyses (Supplementary Fig. 25).

Similarly, IM and LUM tumors presented better prognosis in the METABRIC dataset when evaluating OS, despite the lack of anti-HER2 therapy[36] (Supplementary Fig. 26 and 27). In TCGA, no progression-free interval (PFI) events were observed in IM tumors (Supplementary Fig. 28a), but the low total number of events (12/131 patients) precludes reliable analyses in this regard. In addition, the TIL regional fraction[38] (available for 113/131 patients) was higher in IM tumors compared to the rest of the subtypes (Supplementary Fig. 28b), confirming the ability of the IM subtype to capture immune-infiltrated tumors.

Finally, we compared the clinicopathological characteristics across the five subtypes in each study (Supplementary Data 26–30, also showing the clinicopathological characteristics of each one of the evaluated cohorts). Of note, the differences in hormone receptor status, with LUM tumors being predominantly (>90%) characterized by hormone receptor-positive tumors, were found across all cohorts. Distributions of intrinsic subtypes were also similar to those in ALTTO, with LUM tumors being mostly Luminal A/B, and ERBB2-D including almost exclusively HER2-E.

Since specific classes were identified in METABRIC, namely integrative clusters (IntClust), and I-SPY2, namely Immune+ and DNA repair deficiency (DRD)+, we also evaluated their distribution across the five HER2-positive subtypes (Supplementary Data 29 and 30).

In METABRIC, IntClust subtypes were identified in a large cohort of breast cancers including HER2-positive, hormone receptor-positive, and triple-negative tumors[36,39]. As expected, all subtypes were primarily enriched for IntClust 5 (including *ERBB2*-amplified cancers[36]), while 19.2% of LUM tumors were classified as IntClust 1 (including mainly estrogen receptor-positive tumors[36]), 23.1% of IM as IntClust 4 (characterized by immune signatures[36]), and 11.4% of P/Met as IntClust 10 (mainly basal-like tumors characterized by genomic instability[36]), showing some overlap with the HER2-positive subtypes.

In I-SPY2, the majority of IM tumors, although not all of them, were Immune+ (85.2%), suggesting a large overlap between the two classifications. Interestingly, 22.4% of P/Met tumors were DRD+, suggesting that at least some of those HER2-positive breast cancers may have a phenotype suggestive of DNA repair deficiency.

Of utmost importance, these results showed the prognostic value of the subtypes across different cohorts. In addition, the subtypes have the potential to predict response to neoadjuvant anti-HER2 treatments, with LUM tumors having lower pCR rates, while ERBB2-D and IM tumors showed higher chance of achieving pCR.

Discussion

HER2-positive breast cancer is characterized by remarkable heterogeneity in terms of response to anti-HER2 therapies, clinical outcomes, as well as molecular profiles[17,34,40]. In the last years, numerous efforts have been made to identify prognostic and predictive biomarkers in this disease, leading to an ongoing evolution of the treatment landscape[15,41]. In this regard, multiparametric tests such as HER2DX represent a novelty[16]. While such research efforts represent remarkable steps in the direction of a potential clinical application, at the current stage we lack granularity when describing molecular processes which could guide development of novel treatment approaches beyond the identification of prognostic/predictive categories.

In the present study, we highlighted different biological features associated with distant recurrence and overall survival in a case-control cohort from the trastuzumab-only arm of the ALTTO phase III adjuvant trial. In particular, we identified 5 prognostic HER2-positive subtypes depicting both tumor-intrinsic and microenvironment characteristics, namely immune-enriched (IM), luminal (LUM), proliferative/metabolic-enriched (P/Met), mesenchymal/stroma-enriched (Mes/S), and ERBB2-dependent (ERBB2-D). By developing a gene expression-based classifier to identify these subtypes, we validated our findings in external cohorts, including the NeoALTTO phase III trial, SCAN-B, I-SPY2, TCGA, and METABRIC. Of note, the HER2-positive subtypes presented consistent gene expression profiles as well as clinical characteristics across the cohorts, and retained their prognostic value. In particular, IM tumors presented the best outcome in all cohorts, followed by the LUM subtype. A predictive role beyond clinicopathological features was also shown in the neoadjuvant NeoALTTO and I-SPY2 cohorts, particularly for the LUM and ERBB2-D subtypes, associated respectively with lower and higher pCR rates, as well as for IM tumors, which presented high pCR rates. In consideration of their different biology and sensitivity to neoadjuvant therapies, the HER2-positive breast cancer subtypes (main features summarized in Fig. 5) may benefit from tailored treatment approaches. This is relevant, as similar clinical outcomes can be explained by different biological profiles associated with distinct therapeutic vulnerabilities.

The predictive and prognostic effects of immune-related features have been addressed extensively in the last years[16,19,21–25,42]. In fact, part of the activity of anti-HER2 therapies is mediated by the immune system, for instance through the activation of antibody-dependent cellular cytotoxicity[43]. IM tumors, regardless of pCR, presented excellent outcome in NeoALTTO and may avoid escalation with post-operative T-DM1 in case of residual disease at surgery[14], sparing patients from potentially unnecessary toxicities. While several immunotherapeutic approaches are under development in this disease[44,45], we may hypothesize a role for immune-checkpoint inhibitors in combination with HER2-targeting agents for IM tumors, for instance in the context of chemo-free treatment strategies.

In addition, intrinsic molecular subtype studies showed that luminal HER2-positive tumors present good prognosis[16,20,21,30,46], while, conversely, having a negative impact on pCR[16,19,20,46]. Indeed, LUM tumors, given the low probability of achieving pCR, may not benefit from further treatment escalation, particularly in the neoadjuvant phase. However, we could hypothesize that extended adjuvant endocrine treatment[47] or the use of CDK4/6 inhibitors in combination with hormonal therapy and trastuzumab[48,49] may be reasonable approaches to be tested in patients with LUM HER2-positive/hormone receptor-positive tumors. Interestingly, while hormone receptor status is an important stratification factor in HER2-positive breast cancer, a luminal phenotype is not the predominant characteristic of all hormone receptor-positive tumors. In fact, not all hormone receptor-positive tumors were classified as LUM, whereas almost the totality of LUM tumors was hormone receptor-positive.

Following the results of the APT trial, adjuvant paclitaxel in combination with trastuzumab is now the preferred regimen for small, node-negative tumors directly undergoing surgery, which present excellent prognosis[50]. We may argue that this strategy could be explored in tumors presenting favorable biology (e.g., IM or LUM tumors) beyond clinical staging, particularly in consideration of the low response to neoadjuvant treatments for the LUM subtype. De-escalation approaches driven by treatment response, such as the one evaluated in the PHERGain trial[51], could also further benefit from integrating the tumor biology into the algorithm.

A key aspect for sensitivity to anti-HER2 treatments is represented by HER2 signaling, which can be influenced by *ERBB2* expression and copy number levels, as well as by mutations in the downstream proteins[16–20]. In this regard, the ERBB2-D subtype, mostly characterized by HER2 signaling (suggesting "HER2-addiction"[17]) and intermediate proliferation, presented excellent sensitivity to neoadjuvant treatments, particularly to double HER2-blockade in NeoALTTO and I-SPY2, although these results are limited by the small number of patients in the different treatment arms.

Importantly, P/Met and Mes/S tumors may represent an unmet medical need in HER2-positive breast cancer. In line with our findings, two clusters, proliferative and EMT-enriched, both associated with poor prognosis, have been identified in a previous analysis of different breast cancer datasets[52]. P/Met tumors are characterized by metabolic reprogramming[53–55], activated by AKT/mTOR and IGF1 signaling[56,57], and proliferation, which is also sustained by the parallel RAS/MAPK transduction pathway[53,56]. Proliferation has been shown to positively impact response to neoadjuvant therapies, while, conversely, being associated with worse survival outcomes[16,19,58]. These characteristics could make P/Met tumors exquisitely sensitive to dose-dense chemotherapy in combination to HER2-targeting agents[59,60]. Furthermore, strategies targeting the complex network of metabolic processes in tumor cells are currently under development[56,61], and may represent a valuable strategy in P/Met tumors together with inhibitors of the PI3K/AKT/mTOR or MAPK pathways[15,62].

Reactive stroma and expression of collagen genes have been shown to impact sensitivity to adjuvant and neoadjuvant anti-HER2 treatments[19,26,27]. More in general, the extracellular matrix contributes to treatment resistance with several mechanisms, including the creation of a physical barrier for therapeutic agents as well as immune cells, and integrin signaling[63]. Targeting tumor stroma at different levels represents a potential approach in Mes/S tumors[63,64]. For instance, as angiogenesis is one of the main features of this subtype, treatments such as AMG 386, tested in I-SPY2[35], may be worth exploring. Moreover, inhibitors of Notch, Hedgehog, WNT, and TGF-beta could block cancer-associated fibroblast signaling characterizing these tumors[63,64]. Finally, the low proliferation of Mes/S tumors lead us to hypothesize a potential role for de-escalating chemotherapy. A growing area of interest is represented by the evaluation of biomarkers for novel antibody-drug conjugates (ADCs), which are becoming part of the treatment algorithm in HER2-positive breast cancer[15]. In this regard, investigating the efficacy of ADCs in P/Met and Mes/S tumors is of particular interest to improve their outcomes and evaluate biomarkers related to the microenvironment (e.g., stroma composition and impact on the bystander effect) and tumor cells (e.g., proliferative activity).

Of interest, further studies implementing novel technologies, e.g., single-cell RNA sequencing, spatial transcriptomics, and their integration[65], have the potential to further dissect the heterogeneity of this disease and may refine molecular subtypes depicted by bulk RNA sequencing. For instance, single-cell data could allow the identification of coexisting tumor clusters with distinct biological profiles. Moreover, different compositions in terms of immune cell subpopulations at the single-cell level may characterize subgroups of IM tumors. In addition, spatial transcriptomics, allowing spatial mapping of gene expression data, may highlight different geographic patterns of immune infiltration in IM tumors, or allow a more precise evaluation of cancer-related processes (e.g., EMT) in different tumor areas of Mes/S tumors.

Our study presents some limitations. Patients in the case-control cohort received adjuvant trastuzumab plus chemotherapy in the context of the ALTTO trial. However, being a high-risk population enriched for distant recurrence events, at least part of them, according to current recommendations[29], could be treated with (neo)adjuvant trastuzumab in combination with pertuzumab due to their clinical stage, or, with regards to patients with small and node-negative tumors, following the adjuvant APT regimen. The composition of the ALTTO case-control cohort may affect the generalizability of our findings. Nonetheless, the results obtained in the validation cohorts explored, presenting a greater variability in terms of staging, showed robust biological and clinical features associated with the HER2-positive subtypes. It remains to be evaluated the role of the identified subtypes in the metastatic setting. Moreover, trastuzumab still represents the main backbone of HER2-targeting regimens and constitutes a reliable starting point for comparisons. When developing and validating the HER2-positive breast cancer subtypes, we encountered challenges related to the different sequencing platforms and sample types (detailed in METHODS), and, at the current stage, the identification of our subtypes lacks a specific standardized assay which will allow prospective validation. Nonetheless, our approach allowed the identification of the subtypes in different cohorts while retaining their most relevant characteristics, supporting the robustness of our findings and allowing retrospective validation in additional studies. In addition, while biological profiles, as well as observed prognosis and pCR rates, allow us to hypothesize differential sensitivity to specific treatments across subtypes, this will need specific experimental and clinical validation.

The strengths of the present work consist of the use of a well-annotated case-control cohort balanced for clinicopathological characteristics in the context of the phase III ALTTO trial, and the use of distant recurrence as endpoint to identify the subtypes. This ensured the identification of biological features responsible of metastatic relapse, the most frequent cause of death in patients affected by breast cancer[66]. The link between the subtypes and survival outcomes was indeed confirmed when exploring the validation cohorts. Moreover, our findings are in line with previous studies evaluating biomarkers in HER2-positive breast cancer, supporting their reliability.

In conclusion, we identified five HER2-positive breast cancer subtypes in ALTTO, unraveling the heterogeneity of this disease in terms of molecular profiles associated with clinical outcome. Of note, their biological features and clinical behavior were validated in external datasets, demonstrating the consistency of our findings. Additional validation in cohorts implementing therapeutic interventions including the use of (neo)adjuvant pertuzumab[12,13] or (de-)escalated regimens[8,14,50] is warranted. Overall, our results provide the rationale for the exploration of novel treatment strategies and optimization approaches in HER2-positive breast cancer, leveraging the diverse biology of this disease.

Methods

Studies included in the present analysis

The ALTTO trial study design and outcomes have been previously published[6]. Briefly, between June 2007 and July 2011 the ALTTO phase III clinical trial randomized 8381 patients with primary HER2-positive breast cancer to receive adjuvant trastuzumab (T), lapatinib (L) or T + L combined with chemotherapy, into 4 arms: 52 weeks of T (Arm 1), 52 weeks of L (Arm 2), 12 weeks of trastuzumab followed by a 6-week washout and 34 weeks of L (Arm 3), and T + L for 52 weeks (Arm 4). The anti-HER2 therapy could be administered sequentially after the adjuvant chemotherapy (Design 1), concurrently (12 weeks) with a taxane and after anthracycline-based chemotherapy (Design 2), or concurrently (18 weeks) with docetaxel plus carboplatin (Design 2B). The primary endpoint was disease-free survival (DFS), defined as the time from randomization to invasive recurrence at local, regional or distant sites; contralateral invasive breast cancer; second non-breast malignancy; or death from any cause. Secondary endpoints included OS, defined as time from randomization to death from any cause and time-to-distant recurrence, and DRFI, defined as the time from randomization to recurrence at any distant sites (ignoring locoregional recurrences and second primary cancers). In 2011, due to futility to demonstrate noninferiority of L versus T, the L arm was closed, and patients free of disease were offered adjuvant T. This led to a modification of the original statistical plan. Overall, T + L did not significantly improve DFS and OS compared to T alone[6,67]. In ALTTO, clinical data were collected at Institut Jules Bordet and Frontier Science Scotland; analyses are based on the on ALTTO Study database with clinical cutoff date on 1 July 2016.

The phase III NeoALTTO trial randomized 455 HER2-positive early-stage BC patients to receive neoadjuvant T, L or T + L, and results have been previously published[9,68,69]. Patients were recruited between

January 5, 2008, and May 27, 2010. The primary endpoint was pathological complete response (pCR) defined as absence of invasive tumor cells in the breast (ypT0/is), later amended to the absence of invasive tumor cells in the breast and in the axillary lymph nodes (ypT0/is ypN0). In the present analysis, the original definition of pCR (ypT0/is) is used. Event-free survival (EFS) was a secondary endpoint, defined as the time from randomization to the first event, including breast cancer relapse after surgery, second primary malignant neoplasm, and death without recurrence for women who received surgery for breast cancer, or, for those who did not undergo surgery, non-completion of neoadjuvant therapy due to progressive disease or death during clinical follow-up. OS was defined as time from randomization to death from any cause.

Ethics committee and relevant health authorities at each participating site approved the ALTTO and NeoALTTO (Supplementary Data 31 and 32) studies and all patients provided written informed consent including future biomarker research in both trials. The current analysis has been approved by the TransALTTO scientific committee and was conducted in accordance with the Declaration of Helsinki.

For NeoALTTO, results from RNA sequencing have been previously published[19,70]. RNA sequencing data and PAM50 subtypes obtained from frozen tumor samples (RiboZero depletion of ribosomal RNA) are available for 254 patients out of the 455 patients enrolled in the trial[19], while baseline stromal TILs data are available for 233 out of the 254 patients[22]. In NeoALTTO, clinical data were collected at Institut Jules Bordet and Frontier Science Scotland; analyses are based on the clinical study database frozen on May 26, 2016.

The I-SPY2[35], METABRIC[36], TCGA[34], and SCAN-B[37] studies, including their gene expression data, have been previously published and are publicly available.

In the ongoing, adaptive, randomized phase II neoadjuvant I-SPY2 trial, pCR was defined as ypT0/is ypN0. For I-SPY2, gene-level expression microarray data from frozen samples were obtained from GEO, GSE194040, and clinical data from ref. 35. Probes including non-available values were removed. HER2-positive tumors were selected based on the annotation in the clinical data ($N = 245$). Patients affected by HER2-positive breast cancer in the I-SPY2 study were treated in 6 arms, namely paclitaxel + trastuzumab, T-DM1 + pertuzumab, paclitaxel + pertuzumab + trastuzumab, paclitaxel + neratinib, paclitaxel + MK-2206 (an AKT inhibitor) + trastuzumab, paclitaxel + AMG 386 (an anti-angiopoietin peptibody) + trastuzumab. These treatments were followed by doxorubicin + cyclophosphamide.

For METABRIC, gene expression microarray data from frozen samples were downloaded from cBioPortal (https://www.cbioportal.org/), and clinical data obtained from ref. 39. Probes including non-available values were removed. Duplicated genes were selected choosing the copy with the highest standard deviation. HER2-positive tumors were selected based on the annotation in the clinical data ($N = 236$). Patients in METABRIC did not receive anti-HER2 therapy.

TCGA raw counts (from frozen samples, poly(A) library preparation) and clinical data were obtained using the R package *TCGAbiolinks* (v2.24.3)[71]. Raw counts were normalized to transcript per million (TPM) with the R package *IOBR* (v0.99.9; function "count2tpm")[72]. Survival data for TCGA were obtained from ref. 73. Here, we focused on progression-free interval as suggested in ref. 73, although the analyses in this dataset were limited by the small number of events recorded, possibly also due to short follow-up time[73]. HER2-positive breast cancers selected as described in ref. 74 (i.e., HER2 score 3+ at immunohistochemistry or score 2+ with amplification at fluorescence in situ hybridization), excluding stage IV tumors ($N = 131$). The TIL regional fraction (available for 113 out of 131 patients) was obtained from ref. 38, while PAM50 intrinsic subtypes were obtained from ref. 75.

For SCAN-B, gene expression data from frozen tissue [poly(A) mRNA[76]] as Fragments Per Kilobase per Million (FPKM) reads and clinical data were downloaded from https://data.mendeley.com/datasets/yzxtxn4nmd/3[77]. For the present analyses, we included only the "follow-up" cohort as described in ref. 37, and HER2-positive tumors selected based on the annotation in the clinical data ($N = 819$). When specified, in SCAN-B FPKM gene expression data were converted to TPM data using the formula:

$$TPM = \frac{FPKM}{\sum_{all\ genes} FPKM} \times 10^6 \quad (1)$$

as, for instance, done in ref. 78. In SCAN-B not all patients affected by HER2-positive breast cancer received anti-HER2 therapies, as reported in the clinical data.

The ALTTO (NCT00490139), NeoALTTO (NCT00553358), I-SPY2 (NCT01042379), and SCAN-B (NCT02306096) trials are registered at https://www.clinicaltrials.gov.

Sample collection in ALTTO, selection of the case-control cohort

Patients enrolled in the ALTTO trial had centrally confirmed HER2 positive (3+ by immunohistochemistry and/or positive by fluorescence in situ hybridization) invasive breast cancer, following the 2007 American Society of Clinical Oncology/College of American Pathologists guidelines to define HER2 positivity[79]. Hormone receptor positivity was determined based on the presence of ≥1% of tumor cells expressing estrogen and/or progesterone receptors. Out of the 2097 patients enrolled in the trastuzumab-only arm of the ALTTO trial, RNA was extracted for 1357 patients from formalin-fixed paraffin-embedded (FFPE) tumor cores from surgical samples, in the area with the highest tumor cellularity as evaluated by a breast cancer specialized pathologist (G.V.). Extraction of genomic material (dual extraction for DNA and RNA using the RecoverAll™ Total Nucleic Acid Isolation Kit for FFPE (Life Technologies-Invitrogen™, Thermo Fisher) was performed at the European Institute of Oncology (Milan, Italy, responsible G.V.), one of the ALTTO central laboratory, which had access to samples from all countries enrolling patients in ALTTO except USA and China. Total RNA was quantified using the NanoQuant Plate (Tecan). A case-control approach was adopted to select cases with distant relapse and controls with no relapse, with a 1:2 ratio (Supplementary Fig. 1, showing the CONSORT diagram). In detail, we excluded patients who received neoadjuvant chemotherapy, and with less than 200 ng of total RNA available, as well as those who died without presenting a relapse event. After a median follow-up of 6.9 years (clinical cutoff date on 1 July 2016), a total of 134 patients with distant recurrence and satisfying these criteria were identified. These patients were matched with a 1:2 ratio to controls who did not develop distant or locoregional breast cancer recurrence using the R package *MatchIt* (v4.4.0)[80] according to tumor size (≤2 cm, >2 to ≤5 cm, >5 cm), nodal status (N0, N1, N2-N3) and hormone receptor status (positive vs. negative) with exact matching (the "nearest" matching option was used for study design—sequential vs. concomitant—and histological grade). A total of 402 patients were selected, with $N = 134$ cases and $N = 268$ controls (Supplementary Data 1).

RNA sequencing data processing in ALTTO

RNA sequencing was performed at the BRIGHTcore sequencing facility (Université libre de Bruxelles) using the Illumina Novaseq 6000 sequencer according to standard procedures. In detail, starting from ≥200 ng of total RNA, strand-specific cDNA libraries were constructed using the TruSeq Stranded Total RNA Library Prep Gold for Illumina paired-end sequencing (2 × 101 bp), and sequencing performed with a target read depth for rRNA-depleted total RNA (RiboZero Gold) of 30×10^6 reads. Fastq reads were trimmed using Trimmomatic v0.39[81]. Transcript (ENSEMBL v98) abundance estimates were generated by Salmon (v1.5.1)[82] based on STAR alignment to the human reference GRCh38/hg38 (using GENCODE v38 for the gene positions). Duplicated reads (often artifactual due to the low library

complexity of many samples, probably arising from the FFPE nature of the samples) were removed after alignment. The R package *tximport* (v1.24.0)[83] was used to obtain gene-level estimates and transcripts per million (TPM) expression levels. Samples with less than 1 million of non-duplicated reads mapping to the transcriptome and/or presenting a proportion of duplicated reads >80% were considered of low quality and excluded (Supplementary Fig. 2a, b). The genes *AC010970.1*, *AC017002.1*, *FP236383.4*, *FP236383.5*, *FP671120.6*, *FP671120.7*, associated with low quality samples were filtered out, and TPM recomputed. Thus, RNA sequencing data were obtained for 386 patients out of 402, with no differences in terms of clinicopathological characteristics between the two cohorts (Supplementary Data 1). Gene counts were normalized using the variance stabilizing transformation (VST) method (function "varianceStabilizingTransformation" with option blind = TRUE from the R package *DESeq2*[84], v1.36.0), after filtering genes with low expression (average of counts <10 across samples). A principal component analysis (PCA) plot on VST normalized genes and a heatmap of the top 500 most variable genes annotated by hormone receptor status are shown in Supplementary Fig. 2c, d, to depict the ability of the gene expression data to capture the biological differences between hormone receptor-positive and negative HER2-positive tumors in the ALTTO case-control cohort. In the heatmap, gene expression values are centered removing their mean and scaled dividing them by their standard deviation; rows and columns are clustered using Spearman as distance measure and Ward.D2 as clustering method.

RNA sequencing data processing in NeoALTTO

Out of the 455 patients enrolled in the NeoALTTO trial, RNA was successfully sequenced from frozen tumor samples in 254 patients as previously described[19]. Gene expression data were obtained as described in ref. 70. Briefly, starting from BAM files, read pairs were trimmed using Trimmomatic v0.39[81], and Salmon v1.5.1[82] was used for alignment to the human reference GRCh38/hg38 (patch 13), using GENCODE v38 for the gene positions. The R package *tximport* (v1.16.1)[85] was used to obtain gene-level estimates and TPM-normalized gene expression levels. Intrinsic subtypes have been previously obtained in ref. 19.

Computation of intrinsic subtypes and gene expression signatures, gene set variation analysis, identification of prognostic genes

Intrinsic subtypes were calculated on TPM gene expression data using the R package *AIMS* (Absolute Intrinsic Molecular Subtyping; v1.28.0)[31]. A pool of 36 gene expression signatures/single genes (Supplementary Data 3, reporting genes, coefficients, as well as the sources, i.e., PMID or Reactome[86]) was computed on VST normalized gene expression data in ALTTO. Gene signature scores were computed as weighted mean of the genes, using the coefficients as weights (+1/−1 based on the association with the biological process described in the original publication, or exact values derived from the original publication).

The hallmark gene sets[87] (version 7.4) were downloaded from MSigDB (https://www.gsea-msigdb.org/gsea/msigdb) with the R package *msigdbr* (v7.4.1)[88] and used to perform gene set variation analysis[89] (R package *GSVA* v1.44.2, method "GSVA") on the VST normalized gene expression data in ALTTO. The following 8 hallmark signatures were removed due to the lack of association with tumor processes or microenvironment: peroxisome, pancreas beta cells, spermatogenesis, bile acid metabolism, heme metabolism, UV response up, UV response down, xenobiotic metabolism. Gene signatures/GSVA scores were computed in the external datasets following the same procedure as for ALTTO. For uniformity, gene counts in NeoALTTO and TCGA data were also VST normalized after filtering low expressed genes. In METABRIC and I-SPY2, gene expression values as obtained from the sources were used for the calculation of gene signature/GSVA scores, while $\log_2$(FPKM + 1) data (adjusted for library protocol) were used in the SCAN-B dataset.

When comparing genes/signature scores and GSVA results across the HER2-positive breast cancer subtypes, the values were centered removing their mean and scaled dividing them by their standard deviation. Each group was compared with the others (1 group vs. rest). The effect size was obtained by applying a linear regression model, and the *P* values were obtained with a Wilcoxon rank sum test. *P* values were adjusted for multiple testing according to Benjamini & Hochberg method (i.e., FDR), and considered significant when <0.1 for these comparisons.

Prognostic genes in ALTTO were identified using a multivariable model for DRFI (see the paragraph "Statistical analyses."). Genes significant for FDR < 0.05 were considered for non-negative matrix factorization.

Non-negative matrix factorization and clustering, identification of HER2-positive breast cancer subtypes

Non-negative matrix factorization (NMF) was performed using the R package *NNLM*[90] (https://github.com/linxihui/NNLM; v0.4.4; function "nnmf") on raw counts of the prognostic genes significantly associated with DRFI, using the mean Kullback-Leibler divergence as loss function and method = "lee"[91]. The "W" and "H" components obtained were rescaled so that both the sum of the gene weights/factor and the sum of the factor scores/sample were 1. The number of factors (i.e., ranks) was decided by adapting the method suggested by the Authors of the *NNLM* package[90], guiding the choice of the number of ranks via missing value imputation. In detail, the mean Kullback-Leibler divergence was computed on 10% of randomly masked gene expression values (comparing values imputed by NMF with the real ones) across 50 runs for ranks 1 to 8, and identifying the rank in which the improvement of the mean Kullback-Leibler divergence (average across the 50 runs) started to decrease (Supplementary Fig. 29). In addition, we considered the a priori knowledge given by the results from gene expression signature analysis for DRFI, describing at least three groups of processes associated with prognosis, i.e., immune signatures, stroma-related signatures, signaling/metabolic pathways, and a group of signatures not associated with DRFI. Overall, we considered the solution with 4 factors acceptable, and NMF was run on the gene expression matrix including all the genes associated with DRFI. K-means clustering (nstart = 1000, $k = 4$) was performed on the factor scores identified by NMF, allowing the categorization of the samples. Selecting $k = 4$ led to clusters that were representative of the NMF factors. Heatmaps depicting the NMF gene weights/factor (i.e., feature matrix, "W" component in NMF) and the factor scores/sample (i.e., coefficient matrix, "H" component in NMF, ordered according to k-means clustering) are shown in Supplementary Fig. 8. The clusters were further refined into 5 subtypes using the AIMS intrinsic subtype information (HER2-enriched vs. rest), based on survival differences observed within cluster 4 (described in the RESULTS section).

Development of a gene expression-based classifier for the HER2-positive breast cancer subtypes in ALTTO

In order to validate the subtypes identified in ALTTO in other datasets, we deemed it useful to develop a gene expression-based classifier. For this objective, we followed a two-step feature selection. First, differential expression analysis comparing one subtype vs. the rest was performed using *DESeq2*[84] on gene-level abundance estimates from Salmon (functions "DESeqDataSetFromTximport" and "DESeq"), after filtering genes with very low expression (sum of the reads across all samples <10), and specifying "alpha = 0.01" in the function "results". The "lfcShrink" function was used to perform $\log_2$(fold change) shrinkage according to the "ashr" method[92]. Genes were considered significantly differentially expressed for FDR < 0.01, and selected for the next step when $|\log_2(\text{fold change})| > 1.5$ (see paragraph

"Cross-validation of the procedure to develop the gene expression-based classifier."). To ensure that genes selected in ALTTO were present in the external validation cohorts, we further filtered genes in common across all other cohorts, excluding all genes with a median TPM equal to 0 in ALTTO, NeoALTTO, TCGA, median TPM (obtained from FPKM) equal to 0 in SCAN-B (considering separately the three different library protocols, i.e., dUTP, NeoPrep, and TruSeq) and with missing values in METABRIC and I-SPY2. This choice was also motivated by the application of a methodology relying on the median values of each gene in the different cohorts to apply the classifier, as explained in the following section. A total of 365 genes were retained to develop a Least Absolute Shrinkage and Selection Operator (LASSO) multinomial classifier, using the R package *caret* (v6.0-93). This type of model was chosen as it allows to both classify samples as well as to compute gene signatures using the coefficients provided by the LASSO model. Several steps were taken to reduce overfitting.

In detail, the classifier was developed in the whole cohort starting from $\log_2(\text{TPM}+1)$ gene expression values in ALTTO, using a 10-fold × 10 times cross validation (train and test sets stratified for the subtypes). In the *caret* function "trainControl", we specified the options selectionFunction = "oneSE" (which lead to the selection of the simplest model within one standard error of the optimal one, aiming to reduce the probability of overfitting); sampling = "smote" (an hybrid sampling method to ensure balance of the classes[93] in the training sets); and classProbs = TRUE. A grid of 100 lambdas was used in the "expand.grid" function with alpha fixed at 1 (LASSO penalty), and the best lambda was chosen based on the F1 score measure (defined as the harmonic mean between precision and recall, and computed from with the R package *MLmetrics* v1.1.1, specified in the "trainControl" function) during the repeated cross validation procedure. In the *caret* "train" function we specified the options method = "glmnet", family = "multinomial", type.measure = "class", type.multinomial = "grouped" (which applies a grouped LASSO penalty to the variables), intercept = FALSE, and preprocessing included centering by the mean and scaling by the standard deviation.

The mean and standard deviation values (influenced by the smote sampling) as well as the coefficients assigned to each gene for the five subtypes were extracted from the final model, to allow the application of the classifier in ALTTO as well as external cohorts. For this aim, signature scores are computed for each subtype as weighted mean of the genes, using the coefficients derived from LASSO as weights, and the subtypes are assigned based on the highest value.

Validation of the classifier for HER2-positive breast cancer subtypes in external cohorts

As the subtypes were identified from genes associated with the risk of metastatic recurrence in ALTTO, the step concerning the external validation was necessary to validate and explore in different cohorts their prognostic/predictive values and biological profiles. Gene expression data can be influenced by several technical aspects, including type of sample (e.g., FFPE vs. fresh frozen), method of RNA selection (e.g., poly(A) enrichment vs. ribosomal RNA depletion)[94,95], and technological platform (e.g., RNA sequencing vs. microarray). Indeed, in-silico validation of classification methods can be challenging due to these technical hurdles. Moreover, identifying stable subtypes within a tumor type is not trivial, and the several, only partially overlapping, triple-negative breast cancer classifications described in the years are an example of this[96–99].

Coefficients of the LASSO model are relative to the magnitude of the gene expression values, and can therefore be sensitive to technical differences. In an attempt to minimize these differences, we implemented a methodology suggested by Bush et al.[100], i.e., the ratio-based correction on TPM gene values. This approach assumes that differences in the expression values across platforms may be related by technical factors (poly(A) vs. ribosomal RNA depletion in the case of the original publication[100]). We extended this concept to make gene expression values more comparable across the different datasets, to compute the HER2-positive subtypes. In detail, we calculated the medians of the TPM values of the genes of interest (included in the classifier) in ALTTO, as well as the median of the same genes in each one of the evaluated validation cohorts [TPM in NeoALTTO and TCGA, TPM converted from FPKM in SCAN-B (considering the different library protocols separately, which were merged after the ratio-based correction), intensities in METABRIC and I-SPY2 after reversed $\log_2$ transformation, i.e., 2^gene expression values]. In each sample, we rescaled the value of the genes of interest using the following formula:

$$gene\ A\ value_{sample} = \frac{gene\ A\ median_{ALTTO}}{gene\ A\ median_{new\ cohort}} \times gene\ A\ original\ value_{sample} \quad (2)$$

where "sample" belongs to "new cohort".

Next, a $\log_2$("new value" + 1) transformation was applied. This framework ensures that the median of each gene is comparable across studies, making the application of the LASSO coefficients, as well as the following standardization using the mean and standard deviation derived from the LASSO model, more reliable. However, despite this transformation, microarray data have a more limited dynamic range compared to RNA sequencing[101]. Therefore, for METABRIC and I-SPY2 we opted to perform standardization using the mean and standard deviation computed in each one of these two cohorts. After standardization, the signature scores for each subtype are computed as described in the previous section (i.e., as weighted mean of the genes), and subtypes assigned based on the highest value.

We acknowledge that this methodology has some limitations. Without having a ground-truth in external datasets, we could not avoid the well-known issue of test set bias related to the normalization required to reduce platform-specific artifacts[102]. In our case, this can impact the median values of each gene in each cohort, which could be influenced by the composition of the cohort; moreover, we assume that differences in gene expression levels across platforms/conditions (e.g., use of poly(A)/ribosomal depletion, RNA sequencing/microarray, FFPE/fresh-frozen) are mostly related to technical factors, while differences in biological aspects may also play a role. Despite these aspects, the robustness of our findings suggests that the process is able to detect subtypes with consistent features across different cohorts.

The classifier, including the procedure for median-ratio gene transformation and appropriate standardization based on the platform, is implemented in an R function available at https://github.com/BCTL-Bordet/HER2_subtypes, which allows the identification of our subtypes in HER2-positive breast cancer cohorts with gene expression data available.

Cross-validation of the procedure to develop the gene expression-based classifier

In order to evaluate the performance of the procedure to build the gene expression-based classifier on unseen data, we performed a tenfold cross-validation of the whole process, including the differential expression analysis between the subtypes and the development of the LASSO classifier. In detail, we split the ALTTO dataset into 10 folds (train and test sets), stratified by subtype composition. Each train set was again divided into an inner train and validation set recapitulating the procedure to develop the LASSO classifier in the whole cohort, and used to tune the best lambda in each fold, with the same criteria described for the whole cohort. In each fold, after differential expression analysis performed in the train set, genes with FDR < 0.01 as well as median TPM > 0 in both the train and corresponding test set were retained, together with the set of genes present in the external datasets as described in the paragraph "Development of a gene

expression-based classifier for the HER2-positive breast cancer subtypes in ALTTO". For differential expression, we tested different |$\log_2$(fold change)| thresholds (i.e., 0.585, 1, 1.5, and 2) and compared with a Wilcoxon rank-sum test metrics obtained in the 10 test sets (thus, in a total of 40 test sets), namely F1 score macro average, F1 score weighted average, balanced accuracy average (those three computed as average across the 5 subtypes), and accuracy. For all metrics, no differences were noted between the |$\log_2$(fold change)| thresholds 0.585, 1, and 1.5, while for a |$\log_2$(fold change)| >2 the performance was significantly lower (Supplementary Fig. 30).

For |$\log_2$(fold change)| >0.585, >1, >1.5, >2, respectively, the mean (computed across the 10 folds) F1 score macro averages (across subtypes) were 0.72 (CI, 0.68–0.77), 0.7 (CI, 0.67–0.74), 0.69 (CI, 0.64–0.74), and 0.54 (CI, 0.5–0.58); mean F1 score weighted averages were 0.74 (CI, 0.69–0.78), 0.72 (CI, 0.69–0.75), 0.71 (CI, 0.66–0.75), and 0.56 (CI, 0.53–0.6); mean balanced accuracy averages were 0.83 (CI, 0.81–0.86), 0.82 (CI, 0.8–0.84), 0.81 (CI, 0.79–0.84), and 0.72 (CI, 0.7–0.75); and mean accuracies were 0.73 (CI, 0.68–0.78), 0.71 (CI, 0.68–0.74), 0.69 (CI, 0.65–0.74), and 0.55 (CI, 0.51–0.58). The solution filtering differentially expressed genes by |$\log_2$(fold changes)| >1.5 was chosen as it provided comparable performance with respect to solutions using lower thresholds (which would start from a higher number of genes with smaller differences between groups), as we hypothesized that reducing the number of genes used to train the LASSO model while selecting the ones with large differences across groups could facilitate the classification of samples in other cohorts, potentially reducing overfitting related to the selection of genes with closer expression values between subtypes in ALTTO and improving the generalization of the model.

Statistical analysis

The Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK) criteria were followed for this study[103].

Univariable and multivariable Cox proportional hazard models were used for survival analyses. In univariable analysis, *P* values were obtained with the likelihood ratio test. For multivariable analyses, *P* values were derived by an ANOVA on nested Cox models. For survival analyses aimed at exploring the role of intrinsic subtypes, gene signatures and selecting the prognostic genes for NMF in ALTTO, we controlled for tumor size (≤2 cm vs. >2 cm), nodal status (N0 vs. N+), hormone receptor status, age, timing of chemotherapy (sequential vs. concomitant), grade (3 vs. rest), and, in analyses not evaluating AIMS intrinsic subtypes, AIMS HER2-enriched vs. others.

When evaluating the prognostic value of the HER2-positive subtypes and their signature scores in ALTTO and in the external cohorts, multivariable analyses were performed controlling for age (not available for I-SPY2), tumor stage (T1 vs. rest in ALTTO, METABRIC, TCGA, SCAN-B; T2 vs. rest in NeoALTTO; not available for I-SPY2), nodal status (N0 vs. rest; not available for I-SPY2), grade (G3 vs. rest; not available for I-SPY2 and TCGA), hormone receptor status, treatment received when the information was available (trastuzumab + lapatinib vs. single agents in NeoALTTO; anti-HER2 vs. no anti-HER2 in SCAN-B; treatment arms in I-SPY2).

For pCR analyses in NeoALTTO and I-SPY2, logistic regressions were used, and when performing multivariable analyses *P* values were derived by an ANOVA on nested logistic models.

For continuous variables, hazard ratios/odds ratios and confidence intervals in the forest plots were computed after centering the variable by removing its mean and scaling by dividing the variable by its standard deviation. Multivariable analyses were performed on patients with all variables included in the analysis available.

Wilcoxon rank sum (for comparisons between two groups) and Kruskal-Wallis (for comparisons between three or more groups) tests were used to compare continuous variables according to categorical variables. Fisher's test was performed to compare categorical variables. Correlations were assessed by calculating the Spearman's rank correlation coefficient on pairwise complete observations, and considered significant for $P < 0.05$.

Kaplan-Meier survival curves were used to represent survival outcomes according to categories (i.e., clusters, HER2-positive subtypes), and *P* values obtained with log-rank test.

All *P* values were two-sided. False discovery rates were obtained adjusting *P* values with the Benjamini & Hochberg method, whenever specified. *P* values were considered significant when <0.05, FDR significance levels were determined as specified for each analysis. All confidence intervals are 95%. In box plots, the boxes are defined by the upper and lower quartile, the median is shown as a bold horizontal line and whiskers extend to the most extreme data point which is no more than 1.5 times the interquartile range from the box. In bar plots, the error bars represent the 95% confidence interval.

All statistical analyses were performed using the R software (v4.2.1)[104]. The analyses in the present manuscript were performed at the Institut Jules Bordet/Université Libre de Bruxelles.

Reporting summary

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

Data availability

For reproducibility purposes, the RNA sequencing data (fastq files) from ALTTO have been deposited at the European Genome-phenome Archive (EGA) database (https://ega-archive.org), under accession number EGAS50000000525. For reproducibility purposes, the RNA sequencing data (fastq files) at baseline from NeoALTTO are deposited at the EGA database under accession number EGAS00001007563. The data for ALTTO and NeoALTTO on EGA can be obtained upon signature of data access agreements between the investigator requesting the access and Institut Jules Bordet (IJB), subject to applicable laws. For reproducibility purposes, the ALTTO and NeoALTTO clinical data at IJB can be obtained upon signature of data transfer agreements between the investigator and IJB, subject to applicable laws. Access to these data can be requested by contacting the corresponding Author Christos Sotiriou (christos.sotiriou@hubruxelles.be), with an approximate timeframe to reply of 4 weeks. For reproducibility purposes, the manuscript data can be used for a maximum of one year after its reception. For investigators aiming to perform original research, the ALTTO and NeoALTTO RNA sequencing data at baseline and the clinical data are available upon request after submission of a research project proposal (RPP) to the RPP's administrator (alttoresearchproposals@frontier-science.co.uk). In detail, access to data for research will be granted upon review of the RPP and its endorsement by the study Steering Committee, and after entering into an appropriate data access agreement between BIG, IJB, and the investigator, subject to applicable laws. More details and documents required can be found at https://bigagainstbreastcancer.org/clinical-trials/altto/ and https://bigagainstbreastcancer.org/clinical-trials/neoaltto/ under the section "Translational Research". The policy for access to residual biological samples and patients' data (clinical data, imaging data, any data collected in the study, and potential data generated by research proposals) in the ALTTO and NeoALTTO studies is a fair scientific review process set up to ensure precious biological samples or data collected in the studies are accessed appropriately, to avoid duplication of efforts and foster collaboration. The patients' data from the studies are not anonymized yet, only pseudonymized, therefore they are still considered identifiable, and cannot be made publicly available at this point. In order to ensure that they are shared in a way that preserves the privacy of patients and complies with the relevant laws and regulations including the European General Data Protection Regulation, researchers can only access the data after they sign the data transfer agreements mentioned above, either for reproducibility or for

original research purposes. For I-SPY2, gene-level expression data were obtained from GEO, GSE194040, and clinical data from ref. 35. For METABRIC[36], gene expression data were downloaded from cBioPortal (https://www.cbioportal.org/study/summary?id=brca_metabric), and clinical data obtained from ref. 39. TCGA[34] raw counts and clinical data were obtained using the R package *TCGAbiolinks*[71]. Survival data for TCGA were obtained from ref. 73. The TIL regional fraction was obtained from ref. 38, while PAM50 intrinsic subtypes were obtained from ref. 75. For SCAN-B[37], gene expression and clinical data were downloaded from https://data.mendeley.com/datasets/yzxtxn4nmd/3[77]. Clinicopathological data from ALTTO and NeoALTTO can be obtained as specified above. Other data supporting the findings of this study are available within the article, supplementary information files and supplementary data. Source data are provided in this paper. MSigDB data are available at https://www.gsea-msigdb.org/gsea/msigdb. Source data are provided with this paper.

Code availability

The procedure to apply the classifier for the HER2-positive molecular subtypes described in the present paper (starting from TPM/FPKM/microarray data) has been implemented in an R script available at https://github.com/BCTL-Bordet/HER2_subtypes. The R script allows to compute the molecular subtypes in HER2-positive breast cancer cohorts with RNA sequencing or microarray data available. Please note that our molecular classification was developed, applied and tested in HER2-positive breast cancer as defined by immunohistochemistry and fluorescence in situ hybridization criteria, including both hormone receptor-positive and negative tumors. Therefore, please select cohorts of HER2-positive breast cancer according to such criteria before computing the molecular subtypes with our classifier.

References

1. Slamon, D. J. et al. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* **235**, 177–182 (1987).
2. Cossetti, R. J. D., Tyldesley, S. K., Speers, C. H., Zheng, Y. & Gelmon, K. A. Comparison of breast cancer recurrence and outcome patterns between patients treated from 1986 to 1992 and from 2004 to 2008. *J. Clin. Oncol.* **33**, 65–73 (2015).
3. Martin, M. et al. Neratinib after trastuzumab-based adjuvant therapy in HER2-positive breast cancer (ExteNET): 5-year analysis of a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Oncol.* **18**, 1688–1700 (2017).
4. von Minckwitz, G. et al. Adjuvant pertuzumab and trastuzumab in early HER2-positive breast cancer. *N. Engl. J. Med.* **377**, 122–131 (2017).
5. Tolaney, S. M. et al. Seven-year follow-up analysis of adjuvant paclitaxel and trastuzumab trial for node-negative, human epidermal growth factor receptor 2-positive breast cancer. *J. Clin. Oncol.* **37**, 1868–1875 (2019).
6. Piccart-Gebhart, M. et al. Adjuvant lapatinib and trastuzumab for early human epidermal growth factor receptor 2-positive breast cancer: results from the randomized phase III adjuvant lapatinib and/or trastuzumab treatment optimization trial. *J. Clin. Oncol.* **34**, 1034–1042 (2016).
7. Cameron, D. et al. 11 years' follow-up of trastuzumab after adjuvant chemotherapy in HER2-positive early breast cancer: final analysis of the HERceptin Adjuvant (HERA) trial. *Lancet* **389**, 1195–1205 (2017).
8. Piccart, M. et al. Adjuvant pertuzumab and trastuzumab in early HER2-positive breast cancer in the APHINITY trial: 6 years' follow-up. *J. Clin. Oncol.* **39**, 1448–1457 (2021).
9. Baselga, J. et al. Lapatinib with trastuzumab for HER2-positive early breast cancer (NeoALTTO): a randomised, open-label, multicentre, phase 3 trial. *Lancet* **379**, 633–640 (2012).
10. 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).
11. Robidoux, A. et al. Lapatinib as a component of neoadjuvant therapy for HER2-positive operable breast cancer (NSABP protocol B-41): an open-label, randomised phase 3 trial. *Lancet Oncol.* **14**, 1183–1192 (2013).
12. 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).
13. 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).
14. von Minckwitz, G. et al. Trastuzumab emtansine for residual invasive HER2-positive breast cancer. *N. Engl. J. Med.* **380**, 617–628 (2019).
15. Swain, S. M., Shastry, M. & Hamilton, E. Targeting HER2-positive breast cancer: advances and future directions. *Nat. Rev. Drug Discov.* **22**, 101–126 (2023).
16. Prat, A. et al. Development and validation of the new HER2DX assay for predicting pathological response and survival outcome in early-stage HER2-positive breast cancer. *EBioMedicine* **75**, 103801 (2022).
17. Prat, A. et al. HER2-enriched subtype and ERBB2 expression in HER2-positive breast cancer treated with dual HER2 blockade. *J. Natl. Cancer Inst.* **112**, 46–54 (2020).
18. Venet, D. et al. Copy number aberration analysis to predict response to neoadjuvant Anti-HER2 therapy: results from the neoALTTO phase III clinical trial. *Clin. Cancer Res.* **27**, 5607–5618 (2021).
19. Fumagalli, D. et al. RNA sequencing to predict response to neoadjuvant anti-HER2 therapy: a secondary analysis of the NeoALTTO randomized clinical trial. *JAMA Oncol.* **3**, 227–234 (2017).
20. Fernandez-Martinez, A. et al. Survival, pathologic response, and genomics in CALGB 40601 (alliance), a neoadjuvant phase III Trial of paclitaxel-trastuzumab with or without lapatinib in HER2-positive breast cancer. *J. Clin. Oncol.* **38**, 4184–4193 (2020).
21. Guarneri, V. et al. Trastuzumab-lapatinib as neoadjuvant therapy for HER2-positive early breast cancer: survival analyses of the CHER-Lob trial. *Eur. J. Cancer* **153**, 133–141 (2021).
22. Salgado, R. et al. Tumor-infiltrating lymphocytes and associations with pathological complete response and event-free survival in HER2-positive early-stage breast cancer treated with lapatinib and trastuzumab: a secondary analysis of the NeoALTTO trial. *JAMA Oncol.* **1**, 448–454 (2015).
23. Denkert, C. et al. Tumour-infiltrating lymphocytes and prognosis in different subtypes of breast cancer: a pooled analysis of 3771 patients treated with neoadjuvant therapy. *Lancet Oncol.* **19**, 40–50 (2018).
24. Dieci, M. V. et al. Association of tumor-infiltrating lymphocytes with distant disease-free survival in the ShortHER randomized adjuvant trial for patients with early HER2+ breast cancer. *Ann. Oncol.* **30**, 418–423 (2019).
25. Fernandez-Martinez, A. et al. Prognostic and predictive value of immune-related gene expression signatures vs tumor-infiltrating lymphocytes in early-stage ERBB2/HER2-positive breast cancer: a correlative analysis of the CALGB 40601 and PAMELA trials. *JAMA Oncol.* **9**, 490–499 (2023).

26. Hanker, A. B. et al. Extracellular matrix/integrin signaling promotes resistance to combined inhibition of HER2 and PI3K in HER2+ breast cancer. *Cancer Res.* **77**, 3280–3292 (2017).
27. Sonnenblick, A. et al. Reactive stroma and trastuzumab resistance in HER2-positive early breast cancer. *Int. J. Cancer* **147**, 266–276 (2020).
28. Desmedt, C. et al. Characterization and clinical evaluation of CD10+ stroma cells in the breast cancer microenvironment. *Clin. Cancer Res.* **18**, 1004–1014 (2012).
29. Pernas, S. & Tolaney, S. M. Management of early-stage human epidermal growth factor receptor 2-positive breast cancer. *JCO Oncol. Pract.* **17**, 320–330 (2021).
30. Prat, A. et al. Molecular features and survival outcomes of the intrinsic subtypes within HER2-positive breast cancer. *J. Natl. Cancer Inst.* **106**, dju152 (2014).
31. Paquet, E. R. & Hallett, M. T. Absolute assignment of breast cancer intrinsic molecular subtype. *J. Natl. Cancer Inst.* **107**, 357 (2015).
32. Fan, C. et al. Building prognostic models for breast cancer patients using clinical variables and hundreds of gene expression signatures. *BMC Med. Genom.* **4**, 3 (2011).
33. Andersson, A. et al. Spatial deconvolution of HER2-positive breast cancer delineates tumor-associated cell type interactions. *Nat. Commun.* **12**, 6012 (2021).
34. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
35. Wolf, D. M. et al. Redefining breast cancer subtypes to guide treatment prioritization and maximize response: predictive biomarkers across 10 cancer therapies. *Cancer Cell* **40**, 609–623.e6 (2022).
36. Curtis, C. et al. The genomic and transcriptomic architecture of 2000 breast tumours reveals novel subgroups. *Nature* **486**, 346–352 (2012).
37. Staaf, J. et al. RNA sequencing-based single sample predictors of molecular subtype and risk of recurrence for clinical assessment of early-stage breast cancer. *NPJ Breast Cancer* **8**, 94 (2022).
38. Thorsson, V. et al. The immune landscape of cancer. *Immunity* **48**, 812–830.e14 (2018).
39. Rueda, O. M. et al. Dynamics of breast-cancer relapse reveal late-recurring ER-positive genomic subgroups. *Nature* **567**, 399–404 (2019).
40. Schettini, F. & Prat, A. Dissecting the biological heterogeneity of HER2-positive breast cancer. *Breast* **59**, 339–350 (2021).
41. Goutsouliak, K. et al. Towards personalized treatment for early stage HER2-positive breast cancer. *Nat. Rev. Clin. Oncol.* **17**, 233–250 (2020).
42. de Haas, S. L. et al. Tumor biomarkers and efficacy in patients treated with trastuzumab emtansine + pertuzumab versus standard of care in HER2-positive early breast cancer: an open-label, phase III study (KRISTINE). *Breast Cancer Res.* **25**, 2 (2023).
43. Griguolo, G., Pascual, T., Dieci, M. V., Guarneri, V. & Prat, A. Interaction of host immunity with HER2-targeted treatment and tumor heterogeneity in HER2-positive breast cancer. *J. Immunother. Cancer* **7**, 90 (2019).
44. Costa, R. L. B. & Czerniecki, B. J. Clinical development of immunotherapies for HER2+ breast cancer: a review of HER2-directed monoclonal antibodies and beyond. *NPJ Breast Cancer* **6**, 10 (2020).
45. Huober, J. et al. Atezolizumab with neoadjuvant anti-human epidermal growth factor receptor 2 therapy and chemotherapy in human epidermal growth factor receptor 2-positive early breast cancer: primary results of the randomized phase III IMpassion050 trial. *J. Clin. Oncol.* **40**, 2946–2956 (2022).
46. Fernandez-Martinez, A. et al. Tumor intrinsic subtypes and gene expression signatures in early-stage ERBB2/HER2-positive breast cancer: a pooled analysis of CALGB 40601, NeoALTTO, and NSABP B-41 Trials. *JAMA Oncol.* **10**, 603–611 (2024).
47. Del Mastro, L. et al. Extended therapy with letrozole as adjuvant treatment of postmenopausal patients with early-stage breast cancer: a multicentre, open-label, randomised, phase 3 trial. *Lancet Oncol.* **22**, 1458–1467 (2021).
48. 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).
49. 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).
50. Tolaney, S. M. et al. Adjuvant paclitaxel and trastuzumab for node-negative, HER2-positive breast cancer: final 10-year analysis of the open-label, single-arm, phase 2 APT trial. *Lancet Oncol.* **24**, 273–285 (2023).
51. Pérez-García, J. M. et al. Chemotherapy de-escalation using an 18F-FDG-PET-based pathological response-adapted strategy in patients with HER2-positive early breast cancer (PHERGain): a multicentre, randomised, open-label, non-comparative, phase 2 trial. *Lancet Oncol.* **22**, 858–871 (2021).
52. Tekpli, X. et al. An independent poor-prognosis subtype of breast cancer defined by a distinct tumor immune microenvironment. *Nat. Commun.* **10**, 5499 (2019).
53. Vernieri, C. et al. Resistance mechanisms to anti-HER2 therapies in HER2-positive breast cancer: current knowledge, new research directions and therapeutic perspectives. *Crit. Rev. Oncol. Hematol.* **139**, 53–66 (2019).
54. Wang, L., Zhang, S. & Wang, X. The metabolic mechanisms of breast cancer metastasis. *Front. Oncol.* **10**, 602416 (2020).
55. Koundouros, N. & Poulogiannis, G. Reprogramming of fatty acid metabolism in cancer. *Br. J Cancer* **122**, 4–22 (2020).
56. Vernieri, C. et al. Targeting Cancer Metabolism: Dietary and Pharmacologic Interventions. *Cancer Discov.* **6**, 1315–1333 (2016).
57. Creighton, C. J. et al. Insulin-like growth factor-I activates gene transcription programs strongly associated with poor breast cancer prognosis. *J. Clin. Oncol.* **26**, 4078–4085 (2008).
58. Prat, A. et al. A multivariable prognostic score to guide systemic therapy in early-stage HER2-positive breast cancer: a retrospective study with an external evaluation. *Lancet Oncol.* **21**, 1455–1464 (2020).
59. Papakonstantinou, A. et al. Efficacy and safety of tailored and dose-dense adjuvant chemotherapy and trastuzumab for resected HER2-positive breast cancer: Results from the phase 3 PANTHER trial. *Cancer* **126**, 1175–1182 (2020).
60. Wang, L. et al. Impact of dose-dense neoadjuvant chemotherapy on pathologic response and survival for HER2-positive breast cancer patients who receive trastuzumab. *NPJ Breast Cancer* **7**, 75 (2021).
61. Stine, Z. E., Schug, Z. T., Salvino, J. M. & Dang, C. V. Targeting cancer metabolism in the era of precision oncology. *Nat. Rev. Drug Discov.* **21**, 141–162 (2022).
62. Smith, A. E. et al. HER2 + breast cancers evade anti-HER2 therapy via a switch in driver pathway. *Nat. Commun.* **12**, 6667 (2021).
63. Valkenburg, K. C., de Groot, A. E. & Pienta, K. J. Targeting the tumour stroma to improve cancer therapy. *Nat. Rev. Clin. Oncol.* **15**, 366–381 (2018).
64. Xu, M., Zhang, T., Xia, R., Wei, Y. & Wei, X. Targeting the tumor stroma for cancer therapy. *Mol. Cancer* **21**, 208 (2022).
65. Vandereyken, K., Sifrim, A., Thienpont, B. & Voet, T. Methods and applications for single-cell and spatial multi-omics. *Nat. Rev. Genet.* **24**, 494–515 (2023).

66. Dillekås, H., Rogers, M. S. & Straume, O. Are 90% of deaths from cancer caused by metastases? *Cancer Med.* **8**, 5574–5576 (2019).
67. Moreno-Aspitia, A. et al. Updated results from the international phase III ALTTO trial (BIG 2-06/Alliance N063D). *Eur. J. Cancer* **148**, 287–296 (2021).
68. Huober, J. et al. Survival outcomes of the NeoALTTO study (BIG 1–06): updated results of a randomised multicenter phase III neoadjuvant clinical trial in patients with HER2-positive primary breast cancer. *Eur. J. Cancer* **118**, 169–177 (2019).
69. de Azambuja, E. et al. Lapatinib with trastuzumab for HER2-positive early breast cancer (NeoALTTO): survival outcomes of a randomised, open-label, multicentre, phase 3 trial and their association with pathological complete response. *Lancet Oncol.* **15**, 1137–1146 (2014).
70. Rediti, M. et al. Immunological and clinicopathological features predict HER2-positive breast cancer prognosis in the neoadjuvant NeoALTTO and CALGB 40601 randomized trials. *Nat Commun.* **14**, 7053 (2023).
71. Colaprico, A. et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* **44**, e71 (2016).
72. Zeng, D. et al. IOBR: multi-omics immuno-oncology biological research to decode tumor microenvironment and signatures. *Front. Immunol.* **12**, 687975 (2021).
73. Liu, J. et al. An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell* **173**, 400–416.e11 (2018).
74. Agostinetto, E. et al. HER2-low breast cancer: molecular characteristics and prognosis. *Cancers* **13**, 2824 (2021).
75. Berger, A. C. et al. A comprehensive pan-cancer molecular study of gynecologic and breast cancers. *Cancer Cell* **33**, 690–705.e9 (2018).
76. Saal, L. H. et al. The Sweden Cancerome Analysis Network - Breast (SCAN-B) Initiative: a large-scale multicenter infrastructure towards implementation of breast cancer genomic analyses in the clinical routine. *Genome Med.* **7**, 20 (2015).
77. Vallon-Christersson, J. RNA sequencing-based single sample predictors of molecular subtype and risk of recurrence for clinical assessment of early-stage breast cancer. **3**, 94 (2023).
78. Arora, S., Pattwell, S. S., Holland, E. C. & Bolouri, H. Variability in estimated gene expression among commonly used RNA-seq pipelines. *Sci. Rep.* **10**, 2734 (2020).
79. Wolff, A. C. et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J. Clin. Oncol.* **25**, 118–145 (2007).
80. Ho, D., Imai, K., King, G. & Stuart, E. A. MatchIt: nonparametric preprocessing for parametric causal inference. *J. Stat. Soft.* **42**, 1–28 (2011).
81. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
82. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
83. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Research* **4**, 1521 (2016).
84. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
85. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Research* **4**, 1521 (2015).
86. Fabregat, A. et al. The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* **44**, D481–487 (2016).
87. Liberzon, A. et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
88. Dolgalev, I. *Msigdbr: MSigDB Gene Sets for Multiple Organisms in a Tidy Data Format*. (2021).
89. Hänzelmann, S., Castelo, R. & Guinney, J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinform.* **14**, 7 (2013).
90. Lin, X. & Boutros, P. C. Optimization and expansion of non-negative matrix factorization. *BMC Bioinform.* **21**, 7 (2020).
91. Lee, D. D. & Seung, H. S. Learning the parts of objects by non-negative matrix factorization. *Nature* **401**, 788–791 (1999).
92. Stephens, M. False discovery rates: a new deal. *Biostatistics* **18**, 275–294 (2017).
93. Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P. SMOTE: synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **16**, 321–357 (2002).
94. Zhao, W. et al. Comparison of RNA-Seq by poly (A) capture, ribosomal RNA depletion, and DNA microarray for expression profiling. *BMC Genom.* **15**, 419 (2014).
95. Newton, Y. et al. Large scale, robust, and accurate whole transcriptome profiling from clinical formalin-fixed paraffin-embedded samples. *Sci. Rep.* **10**, 17597 (2020).
96. Lehmann, B. D. et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J. Clin. Investig.* **121**, 2750–2767 (2011).
97. Burstein, M. D. et al. Comprehensive genomic analysis identifies novel subtypes and targets of triple-negative breast cancer. *Clin. Cancer Res.* **21**, 1688–1698 (2015).
98. Yin, L., Duan, J.-J., Bian, X.-W. & Yu, S.-C. Triple-negative breast cancer molecular subtyping and treatment progress. *Breast Cancer Res.* **22**, 61 (2020).
99. Bareche, Y. et al. Unraveling triple-negative breast cancer tumor microenvironment heterogeneity: towards an optimized treatment approach. *J. Natl. Cancer Inst.* **112**, 708–719 (2020).
100. Bush, S. J., McCulloch, M. E. B., Summers, K. M., Hume, D. A. & Clark, E. L. Integration of quantitated expression estimates from polyA-selected and rRNA-depleted RNA-seq libraries. *BMC Bioinform.* **18**, 301 (2017).
101. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**, 57–63 (2009).
102. Patil, P., Bachant-Winner, P.-O., Haibe-Kains, B. & Leek, J. T. Test set bias affects reproducibility of gene signatures. *Bioinformatics* **31**, 2318–2323 (2015).
103. McShane, L. M. et al. REporting recommendations for tumor MARKer prognostic studies (REMARK). *Breast Cancer Res. Treat.* **100**, 229–235 (2006).
104. R: The R Project for Statistical Computing. https://www.r-project.org/.

Acknowledgements

ALTTO is a BIG and Alliance led trial. The conduct of the ALTTO trial was funded by GlaxoSmithKline and later Novartis, and NCI. Research reported in this publication was supported by the Fondation contre le cancer (grant number 2018-124), the Breast Cancer Research Foundation (BCRF, grant number 18-152), the Association Jules Bordet, and the Belgian Fonds National de la Recherche Scientifique (F.R.S-FNRS). NeoALTTO is a BIG and SOLTI led trial. The conduct of the NeoALTTO trial was funded by GlaxoSmithKline and later Novartis. The RNA sequencing in NeoALTTO, on which part of the analyses described in this manuscript are based, was funded by GlaxoSmithKline. For part of the analysis, computational resources have been provided by the Consortium des Équipements de Calcul Intensif (CÉCI), funded by the Fonds de la Recherche Scientifique de Belgique (F.R.S.-FNRS) under Grant No.

2.5020.11 and by the Walloon Region. M.R. was supported by Télévie and the Belgian Fonds National de la Recherche Scientifique (F.R.S.-FNRS), by Fondation Rose et Jean Hoguet, and by an AIRC fellowship for Italy. A.JG. was supported by Télévie and the Belgian Fonds National de la Recherche Scientifique (F.R.S.-FNRS), and by Fondation Rose et Jean Hoguet. S.L. is supported by the National Breast Cancer Foundation of Australia Endowed Chair and the Breast Cancer Research Foundation, New York. R.S. is supported by the Breast Cancer Research Foundation (BCRF, grant number 17-194).

Author contributions

M.R., D. Venet, F.R., C.S. conceived and designed the study. M.R., D. Venet, A.JG. contributed to develop the methodology. M.M., S.EA., M.I. provided logistical support. D. Vincent, S.M., G.V. contributed to sample preparation. S.DC., T.U., M.P., L.P., S.L., R.S., and G.V. contributed to sample and data acquisition and to the design of the study. M.R., D. Venet, C.S. analyzed and interpreted the data. M.R., D. Venet performed the analyses. F.R., C.S. supervised the analyses and the study. M.R., D. Venet, C.S. prepared the manuscript, which was reviewed and approved by all authors.

Competing interests

The authors declare the following competing interests. S.EA.: grants via the affiliation from Novartis for the submitted work and from Roche/Genentech, Astra Zeneca, Pfizer, and BCRF outside the submitted work. S.DC.: consultation fees from Pierre-Fabre, IQVIA, and Medica Scientia Innovation Research (MEDSIR); institutional grant from Fondazione Associazione Italiana Ricerca contro il Cancro (AIRC); and Cancer Can.Heal European EU4 Health Programme 101080009-European Commission. T.U.: honoraria from Astra Zeneca, Novartis Pharma K.K., Eisai Co., Ltd., Chugai Pharmaceutical Co. Ltd; research grant from Eli Lilly Japan K.K. M.I.: employee of Novartis, owner of Novartis shares. M.P.: invited speaker for AstraZeneca, Lilly, MSD, Novartis, Pfizer, Roche-Genentech; consultant for Roche-Genentech; advisory board for Frame Therapeutics, Gilead, Menarini, NBE Therapeutics, Odonate, Roche-Genentech, SeaGen, Seattle Genetics; member of boards of directors, scientific board for Oncolytics; research grants to her Institution from AstraZeneca, Lilly, Gilead; funding to her Institution from Menarini, MSD, Novartis, Pfizer, Radius, Roche-Genentech, Servier, Synthon. L.P.: 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. S.L.: research funding to institution from Novartis, Bristol Myers Squibb, MSD, Puma Biotechnology, Eli Lilly, Nektar Therapeutics, Astra Zeneca/Daiichi Sankyo and Seattle Genetics; consultant (not compensated) to Seattle Genetics, Novartis, Bristol Myers Squibb, MSD, AstraZeneca/Daiichi Sankyo, Eli Lilly, Pfizer, Gilead Therapeutics, Nektar Therapeutics, PUMA Biotechnologies, and Roche-Genentech; consultant (paid to institution) to Novartis, GlaxoSmithKline, Roche-Genentech, Astra Zeneca/Daiichi Sankyo, Pfizer, Gilead Therapeutics, Seattle Genetics, MSD, Tallac Therapeutics, Eli Lilly and Bristol Myers Squibb. R.S.: non-financial support from Merck, Case 45, and Bristol Myers Squibb (BMS); research support from Merck, Puma Biotechnology, and Roche; personal fees from Roche, BMS, Astra Zeneca, Daiichi Sankyo and Exact Sciences for advisory boards. G.V.: grants/research supports from Roche/Genentech, Ventana Medical Systems, Dako/Agilent Technologies; honoraria or consultation fees from Ventana, Dako/Agilent, Roche, MSD Oncology, AstraZeneca, Daiichi Sankyo, Pfizer, Gilead. C.S.: advisory board (receipt of honoraria or consultations fees) for Astellas, Cepheid, Vertex, Seattle genetics, Puma, Amgen, Exact Sciences, INC, Merck & Co; invited speaker for Eisai, Prime Oncology, Teva, Foundation Medicine, Exact Sciences; advisory equity from Signatur Biosciences; other support (travel, accommodation expenses) from Roche, Genentech, Pfizer. The remaining authors declare no non-financial or financial competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-54621-3.

Correspondence and requests for materials should be addressed to Christos Sotiriou.

Peer review information *Nature Communications* thanks Xavier Tekpli, who co-reviewed with Ilayda Altinönder, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.