

Classification of HER2-negative breast cancers by *ERBB2* copy number alteration status reveals molecular differences associated with chromosome 17 gene aberrations

Jui Wan Loh, Abner Herbert Lim, Jason Yongsheng Chan  and Yoon-Sim Yap

Ther Adv Med Oncol

2023, Vol. 15: 1–12

DOI: 10.1177/
17588359231206259

© The Author(s), 2023.
Article reuse guidelines:
[sagepub.com/journals-](https://sagepub.com/journals-permissions)
permissions

Abstract

Background: Recently, HER2-negative breast cancers have been reclassified by protein expression into ‘HER2-low’ and ‘HER2-zero’ subgroups, but the consideration of HER2-low breast cancer as a distinct biological subtype with differing prognoses remains controversial. By contrast, non-neutral *ERBB2* copy number alteration (CNA) status is associated with inferior survival outcomes compared to *ERBB2* CNA-neutral breast cancer, providing an alternative approach to classification.

Methods: Here, we investigated the molecular landscape of non-metastatic HER2-negative BCs in relation to *ERBB2* CNA status to elucidate biological differences. Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) and The Cancer Genome Atlas (TCGA) TCGA-BCA datasets ($n = 1875$) were analyzed.

Results: Nearly two-fifths of the cohort harbored *ERBB2* CNAs (39.4%), which were significantly enriched within hormone receptor-negative (56.1%) than within hormone receptor-positive BCs (35.5%; $p < 0.0001$). Globally, CNAs across the genome were significantly higher in *ERBB2* non-neutral compared to neutral cohorts ($p < 0.0001$). Notably, genetic aberrations on chromosome 17 – *BRCA1*, *NF1*, *TP53*, *MAP2K4*, and *NCOR1* – were widespread in the *ERBB2* non-neutral cases. While chromosome 17q arm-level alterations were largely in tandem with *ERBB2* CNA status, arm-level loss in chromosome 17p was prevalent regardless of *ERBB2* gain, amplification, or loss. Differential gene expression analysis demonstrated that pathways involved in the cell cycle, proteasome, and DNA replication were upregulated in *ERBB2* non-neutral cases.

Conclusion: Classification of HER2-negative BCs according to *ERBB2* CNA status reveals differences in the genomic landscape. The implications of concurrent aberrations in other genes on chromosome 17 merit further research in *ERBB2* non-neutral BCs.

Keywords: copy number alteration, genomics, HER2, METABRIC, TCGA

Received: 19 April 2023; revised manuscript accepted: 21 September 2023.

Background

The *ERBB2* gene encodes for human epidermal growth factor receptor 2 (HER2), a transmembrane receptor tyrosine kinase that is overexpressed in 10–30% of invasive breast cancers.¹ Patients with HER2-positive breast cancers, as defined by high protein expression on immunohistochemistry

(IHC) or gene amplification on in situ hybridization (ISH), derive significant benefit from HER2-directed targeted therapies.^{2,3} In recent years, a novel classification of HER2-negative breast cancers has emerged, defining a ‘HER2-low’ subtype based on low-moderate HER2 protein expression (IHC score of 1+ or 2+) and lack of gene

Correspondence to:
Jason Yongsheng Chan
Division of Medical
Oncology, National Cancer
Centre Singapore,
30 Hospital Blvd,
Singapore 168583

Cancer Discovery Hub,
National Cancer Centre
Singapore, 30 Hospital
Blvd, 168583

Duke-NUS Medical School,
8 College Rd, Singapore
169857

jason.chan.y.s@singhealth.com.sg

Yoon-Sim Yap
Division of Medical
Oncology, National Cancer
Centre Singapore, 30
Hospital Blvd, Singapore
168583

Duke-NUS Medical School,
8 College Rd, Singapore
169857

gmsyapys@nus.edu.sg

Jui Wan Loh
Abner Herbert Lim
Cancer Discovery Hub,
National Cancer Centre
Singapore, Singapore

amplification on ISH,⁴ in contrast to HER2-zero (IHC score of 0) breast cancers. Importantly, patients with HER2-low breast cancers may benefit from novel HER2 antibody–drug conjugates (ADCs), with progression-free survival and overall survival improvements, demonstrated over standard chemotherapy in a recent phase III trial.^{5–7}

However, it is contentious whether HER2-low breast cancers, as defined above, truly represent a distinct molecular subtype of breast cancer with unique biological characteristics.^{4,8–10} In particular, there are reproducibility issues in the scoring of HER2 protein expression on IHC, resulting in significant variability among pathologists and specialized centers.^{11,12} In addition, identification of the HER2 status may be complicated by intra-patient heterogeneity.¹³ Consequently, studies investigating survival outcomes in HER2-low *versus* HER2-zero breast cancers have demonstrated mixed results,^{8,14–20} leading to uncertainty as to whether the current definition of HER2-low breast cancer has any biological or prognostic implications.

In a recent study, we demonstrated that HER2-low breast cancers conferred a superior prognosis compared to HER2-zero cases in the non-metastatic setting, but absolute differences were modest. Compared with HER2-zero tumors, HER2-low cases had significantly better relapse-free survival [hazard ratio (HR) 0.90, $p < 0.001$] and overall survival (HR 0.86, $p < 0.001$).¹⁹ Specifically, improved survival outcomes were evident between patients with HER2 IHC 1+ and HER2-zero tumors, but not between those with HER2 IHC 2+ and HER2-zero tumors. In addition, we showed that *ERBB2* non-neutral copy number alteration (CNA) status (as defined by the presence of loss or gain/amplification in *ERBB2*) was associated with worse relapse-free survival compared to *ERBB2* neutral status in a combined Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) and TCGA cohort of HER2-negative non-metastatic breast cancers (HR 1.39, $p = 0.001$). These findings suggest that the classification of traditionally HER2-negative breast cancers may be improved using information about *ERBB2* CNA status.

Hence, in this paper, we examined the molecular landscape of non-metastatic HER2-negative breast cancers in relation to *ERBB2* CNA status to further elucidate biological differences and to discuss potential clinical implications in this context.

Methods

Study cohort

Cases diagnosed with non-metastatic HER2-negative breast cancers in the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC; $n = 1192$) and TCGA TCGA-BRCA ($n = 683$) datasets were extracted from cBioPortal for analysis.^{21–24} *ERBB2* CNA, genomic mutational data including single nucleotide variants, small indels, and copy number alterations (CNA), as well as clinical and pathological details including age, grade, stage, and hormone receptor status were retrieved. Discrete *ERBB2* CNA status was based on the Genomic Identification of Significant Targets in Cancer (GISTIC) method [–2, loss of both copies; –1, one copy loss; 0, neutral; 1, low-level gain (a few additional copies, often broad); 2, high-level amplification (more copies, often focal)].²⁵

Bioinformatic analysis

Differential mutational and CNA analysis was conducted by grouping the patients into two different cohorts: neutral and non-neutral *ERBB2* copy number. The entire cohort was first described with an oncoplot *via* maftools,²⁶ exhibiting 53 breast cancer-related genes that comprised variants with >25% of mean allele frequency.²⁷ Tumor mutation burden [i.e. the number of somatic coding variants per megabase (mt/Mb), with the capture size of 45 and 1 Mb, respectively, in TCGA and METABRIC datasets] was then calculated for each cohort. Furthermore, their association with ploidy status was investigated using Fisher's exact test. AbsCN-seq²⁸ was used to estimate ploidy based on the copy number data from METABRIC as well as copy number and single nucleotide variant data from TCGA. The top 10 differentially altered genes, sorted by odds ratio (OR), from the 53 genes shown in oncoplot were presented in a co-barplot, and the number of patients exhibiting these 53 genes in each cohort was tabulated in a table, with the respective OR and false discovery rate (FDR). To investigate the effect of hormone receptor status, we repeated the analysis in hormone receptor-positive and hormone receptor-negative cohorts, respectively, and the *ERBB2* non-neutral copy number cohort was further sub-categorized into gain/amplification (GISTIC = 1 or 2) and heterozygous loss (GISTIC = –1) to pinpoint the contributing subgroup in differentially altered genes.

Analysis of differentially expressed genes was performed in a similar approach. Since TCGA and METABRIC datasets were sequenced in different platforms (RNA-seq for TCGA and microarray for METABRIC), we employed DESeq2²⁹ and limma,³⁰ respectively, within TCGA and METABRIC datasets, to detect differentially expressed genes at $FDR < 0.01$ and $abs(\text{fold changes}) > 1$. Pathway enrichment analysis was then performed *via* Enrichr,^{31–33} using the KEGG database, to identify pathways that are significantly enriched in neutral and non-neutral cohorts, based on the upregulated genes that are shared between two datasets.

Statistical analysis

Comparisons of the frequencies of categorical variables were performed using Pearson's chi-squared tests. Box–Whisker plots were used to represent continuous variables and Mann–Whitney–Wilcoxon tests with Bonferroni correction were used to evaluate potential associations. All statistical calculations were performed assuming a two-sided test with a significance level of 0.05 unless otherwise stated. All tests were performed using MedCalc for Windows version 19.0.4 (MedCalc Software, Ostend, Belgium).

Results

Patient cohort

A total of 1875 patients diagnosed with stage I–III HER2-negative (i.e. IHC score of 0–2+ and HER2-negative on ISH if HER2 2+) breast cancers were extracted from the METABRIC ($n = 1192$) and TCGA ($n = 683$) datasets. These include hormone receptor-positive ($n = 1520$) and hormone receptor-negative ($n = 355$) cases. Nearly two-fifths of the cases harbored *ERBB2* CNAs or non-neutral *ERBB2* status ($n = 739$; 39.4%), which included amplification ($n = 57$), gain ($n = 242$), and heterozygous loss ($n = 440$). The remaining were *ERBB2* neutral ($n = 1136$; 60.6%).

In hormone receptor-negative breast cancer, heterozygous loss of *ERBB2* was most commonly observed (44.5%), followed by neutral (43.9%), gain (9.6%), and amplification (2.0%). In hormone receptor-positive breast cancer, *ERBB2* was most commonly copy number neutral (64.5%), while heterozygous loss (18.6%), gain

(13.7%), and amplification (3.3%) were observed in the rest (Table 1, Supplemental Figure 1). Altogether, non-neutral *ERBB2* CNA status was significantly more common within hormone receptor-negative (56.1%) than within hormone receptor-positive BCs (35.5%; $p < 0.0001$). We observed that there was a higher proportion of stage 1 cancer with neutral *ERBB2* CNA status (374/557; 67.1%), compared to 59.5% (649/1090) of stage II tumors and 49.6% (113/228) of stage III disease. Notably, grade I breast cancers also had proportionately more *ERBB2* neutral tumors (89/107; 83.2%), compared to 73.8% (360/488) of grade 2 tumors and 54.8% (304/555) of grade 3 tumors.

Somatic mutation and copy number landscape

The list of 53 breast cancer-related genes is shown in Figure 1(a). The contribution of each variant type for the top 10 differentially altered genes, among these 53 genes, was exhibited in a co-barplot between the *ERBB2* neutral and non-neutral cohorts. Alterations in *BRCA1*, *NF1*, *TP53*, *MAP2K4*, and *NCOR1* were most prevalent in the *ERBB2* non-neutral cohort [detected in $\geq 85\%$ of cases; Figure 1(b) and Supplemental Table 1]. *ERBB2*, *BRCA1*, and *NF1* lie on chromosome 17q, while *TP53*, *MAP2K4*, and *NCOR1* are located on chromosome 17p, suggesting a possible molecular pathologic event in *ERBB2* non-neutral cases within classical HER2-negative breast cancers.

In the whole cohort, the median TMB of *ERBB2* neutral and non-neutral tumors was similar at 3.00 mt/Mb and 2.63 mt/Mb, respectively ($p = 0.1254$). TMB was also not significantly different within hormone receptor-positive (neutral, 3.00 and non-neutral, 2.63; $p = 0.1158$) and hormone receptor-negative cohorts [neutral, 3.89 and non-neutral, 2.82; $p = 0.6720$; Figure 1(c)]. Globally, CNAs across the entire genome were significantly higher in the *ERBB2* non-neutral (whole cohort: 11,360; hormone receptor-negative cohort: 14,618; hormone receptor-positive cohort: 9318) compared to neutral (whole cohort: 4313; hormone receptor-negative cohort: 8350; hormone receptor-positive cohort: 4128) cohorts [$p < 0.0001$; Figure 1(d)]. An association analysis was performed between *ERBB2* copy number and estimated ploidy from AbsCN-seq, based on the copy number data from METABRIC and a combination of copy number and single

Table 1. Characteristics of HER2-negative breast cancers from TCGA and METABRIC according to ERBB2 CNA status.

Characteristics	Total (%)					p Value
	Neutral n = 1136	Heterozygous loss n = 440	Gain n = 242	Amplification n = 57	Overall n = 1875	
Hormone receptor status						
Positive	980 (86.3%)	282 (64.1%)	208 (86.0%)	50 (87.7%)	1520 (81.1%)	<0.0001
Negative	156 (13.7%)	158 (35.9%)	34 (14.0%)	7 (12.3%)	355 (18.9%)	
Age at diagnosis (years)						
Below 40	48 (4.2%)	43 (9.8%)	13 (5.4%)	4 (7.0%)	108 (5.8%)	0.0007
40–49	194 (17.1%)	79 (18.0%)	39 (16.1%)	12 (21.1%)	324 (17.3%)	
50–59	271 (23.9%)	121 (27.5%)	60 (24.8%)	8 (14.0%)	460 (24.5%)	
60–69	341 (30.0%)	104 (23.6%)	70 (28.9%)	22 (38.6%)	537 (28.6%)	
70–79	209 (18.4%)	59 (13.4%)	42 (17.4%)	11 (19.3%)	321 (17.1%)	
80 and over	73 (6.4%)	34 (7.7%)	18 (7.4%)	0 (0.0%)	125 (6.7%)	
Stage						
1	374 (32.9%)	107 (24.3%)	61 (25.2%)	15 (26.3%)	557 (29.7%)	0.0002
2	649 (57.1%)	260 (59.1%)	151 (62.4%)	30 (52.6%)	1090 (58.1%)	
3	113 (9.9%)	73 (16.6%)	30 (12.4%)	12 (21.1%)	228 (12.2%)	
Tumor status						
T1	120 (10.6%)	42 (9.5%)	33 (13.6%)	1 (1.8%)	196 (10.5%)	<0.0001
T2	183 (16.1%)	123 (28.0%)	83 (34.3%)	6 (10.5%)	395 (21.1%)	
T3	38 (3.3%)	20 (4.5%)	15 (6.2%)	3 (5.3%)	76 (4.1%)	
T4	5 (0.4%)	7 (1.6%)	4 (1.7%)	0 (0.0%)	16 (0.9%)	
Unknown	790 (69.5%)	248 (56.4%)	107 (44.2%)	47 (82.5%)	1192 (63.6%)	
Nodal status						
N0	175 (15.4%)	112 (25.5%)	66 (27.3%)	3 (5.3%)	356 (19.0%)	<0.0001
N1	119 (10.5%)	37 (8.4%)	48 (19.8%)	4 (7.0%)	208 (11.1%)	
N2	33 (2.9%)	28 (6.4%)	10 (4.1%)	2 (3.5%)	73 (3.9%)	
N3	17 (1.5%)	14 (3.2%)	8 (3.3%)	1 (1.8%)	40 (2.1%)	
Unknown	792 (69.7%)	249 (56.6%)	110 (45.5%)	47 (82.5%)	1198 (63.9%)	
Grade						
1	89 (7.8%)	12 (2.7%)	5 (2.1%)	1 (1.8%)	107 (5.7%)	<0.0001
2	360 (31.7%)	69 (15.7%)	45 (18.6%)	14 (24.6%)	488 (26.0%)	
3	304 (26.8%)	163 (37.0%)	56 (23.1%)	32 (56.1%)	555 (29.6%)	
Unknown	383 (33.7%)	196 (44.5%)	136 (56.2%)	10 (17.5%)	725 (38.7%)	

CNA, copy number alteration; HER2-, human epidermal growth factor receptor 2; METABRIC, Molecular Taxonomy of Breast Cancer International Consortium; TCGA, the cancer genome atlas.

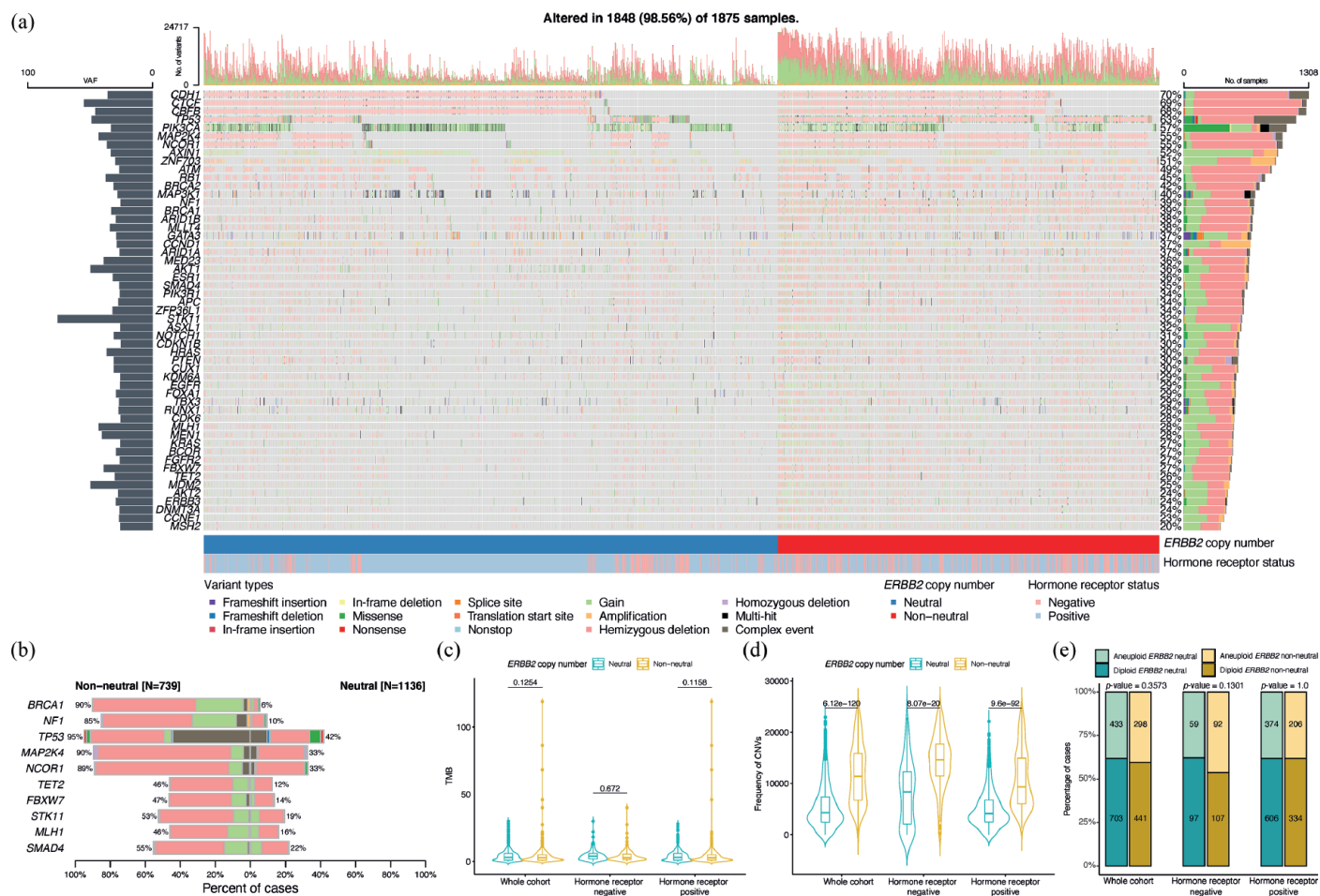


Figure 1. Genomic landscape of study cohort. Somatic mutational and copy number analysis was performed for 1875 cases from both METABRIC ($n = 1192$) and TCGA ($n = 683$) datasets. (a) An oncoplot is generated with the 53 breast cancer-related genes, which are comprised of variants with $>25\%$ of mean allele frequency. (b) The top 10 differentially altered genes from these 53 genes were shown in a co-barplot between the neutral and non-neutral cohorts. (c) Mann-Whitney-Wilcoxon tests, with Bonferroni correction, were performed for the TMB between patients harboring neutral and non-neutral *ERBB2* copy numbers, for the whole cohort as well as within hormone receptor-negative and hormone receptor-positive cohorts, respectively. There is no significant difference in TMB between neutral and non-neutral copy numbers for all three comparisons [median of whole cohort and neutral: 3.00, whole cohort and non-neutral: 2.73, hormone receptor-negative and neutral: 3.89, hormone receptor-positive and non-neutral: 2.82, hormone receptor-positive and neutral: 3.00, hormone receptor-negative and non-neutral: 2.63]. (d) As in (c), same statistical tests were performed on the frequency of CNA for all three different comparisons, and non-neutral cases have significantly more CNAs than those of neutral cases in whole cohort, hormone receptor-negative, and hormone receptor-positive cohorts (median of whole cohort and neutral: 4313, whole cohort and non-neutral: 11,360, hormone receptor-negative and neutral: 8350, hormone receptor-negative and non-neutral: 14,618, hormone receptor-positive and neutral: 4128, hormone receptor-positive and non-neutral: 9318). (e) An association analysis was performed between *ERBB2* copy number and estimated ploidy from AbsCN-seq, based on the copy number data from METABRIC and a combination of copy number and single nucleotide variant data from TCGA. No significant association was detected between copy number and ploidy for all three comparisons (p values are annotated on top of bar chart). CNA, copy number alteration; METABRIC, Molecular Taxonomy of Breast Cancer International Consortium; TCGA, The Cancer Genome Atlas; TMB, tumor mutation burden.

nucleotide variant data from TCGA. No significant association was detected between copy number and ploidy [Figure 1(e)].

Enrichment for alterations in seven genes (*BRCA1*, *NF1*, *TP53*, *MAP2K4*, *NCOR1*, *TET2*,

and *STK11*) was demonstrated in both *ERBB2* gain/amplification and heterozygous loss subgroups in the whole cohort, as well as in the hormone receptor-positive cohort (Supplemental Table 2 and Figure 2). Notably, *BRCA1* and *NF1* CNA were largely in congruence with *ERBB2*

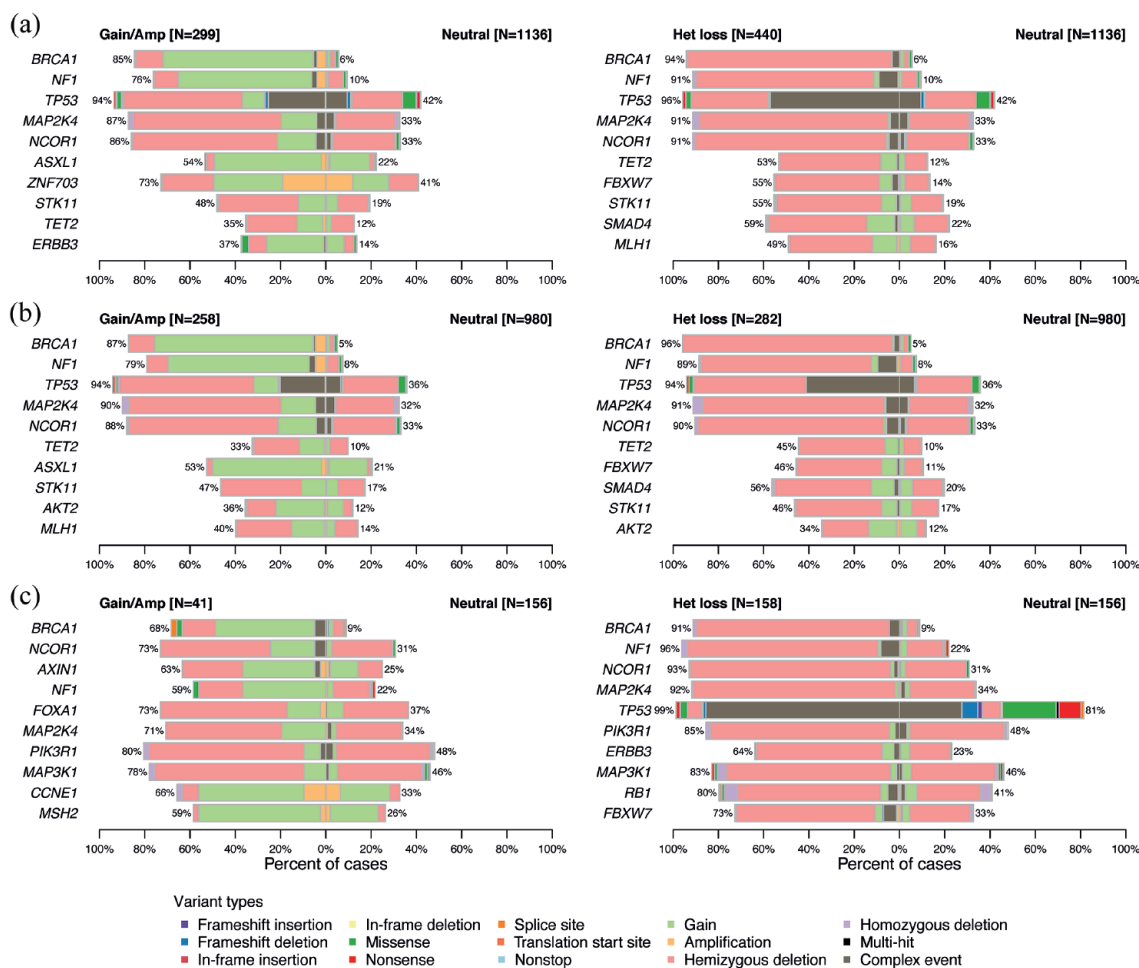


Figure 2. Genomic alteration landscape within *ERBB2* CNA subgroups. Additional differential mutational and CNA analysis was performed among three groups of copy number: neutral (GISTIC=0), gain/amplification (GISTIC=1 or 2), and heterozygous loss (GISTIC=-1) for (a) the whole cohort, (b) hormone receptor-positive, and (c) hormone receptor-negative cohorts of HER2-negative breast cancers, as depicted in co-barplots. Similar to Figure 1(b), the top 10 differentially altered genes among the 53 genes were shown for each comparison. Seven genes (*BRCA1*, *NF1*, *TP53*, *MAP2K4*, *NCOR1*, *TET2*, and *STK11*) were shared among *ERBB2* gain/amplification and heterozygous loss subgroups in the whole cohort, as well as in the hormone receptor-positive cohort. Amp, amplification; GISTIC, Genomic Identification of Significant Targets in Cancer; Het loss, heterozygous loss.

CNA status, while *TP53*, *MAP2K4*, and *NCOR1* showed frequent copy number loss regardless. In keeping with these findings from gene-level copy number analysis, arm-level loss in chromosome 17p was most prevalent regardless of *ERBB2* gain, amplification, or loss. Whereas chromosome 17q arm-level alterations were largely in tandem with *ERBB2* CNA status (Figure 3).

Gene expression analysis

Differential expression analysis between *ERBB2* copy number non-neutral and neutral cases

revealed a total of 11,600 and 4612 differentially expressed genes respectively within TCGA and METABRIC datasets, at $FDR < 0.01$ and $abs(\text{fold changes}) > 1$ [Figure 4(a)]. Among these genes, 1704 (upregulated in non-neutral cases) and 964 (upregulated in neutral cases) were shared between these two datasets, and they were used in the pathway enrichment analysis (with the KEGG database). Specifically, pathways involved in the cell cycle, proteasome, and DNA replication were upregulated in *ERBB2* non-neutral cases [Figure 4(b) and (c)]. Using the same filtering criteria of $FDR < 0.01$ and $abs(\text{fold changes}) > 1$, additional

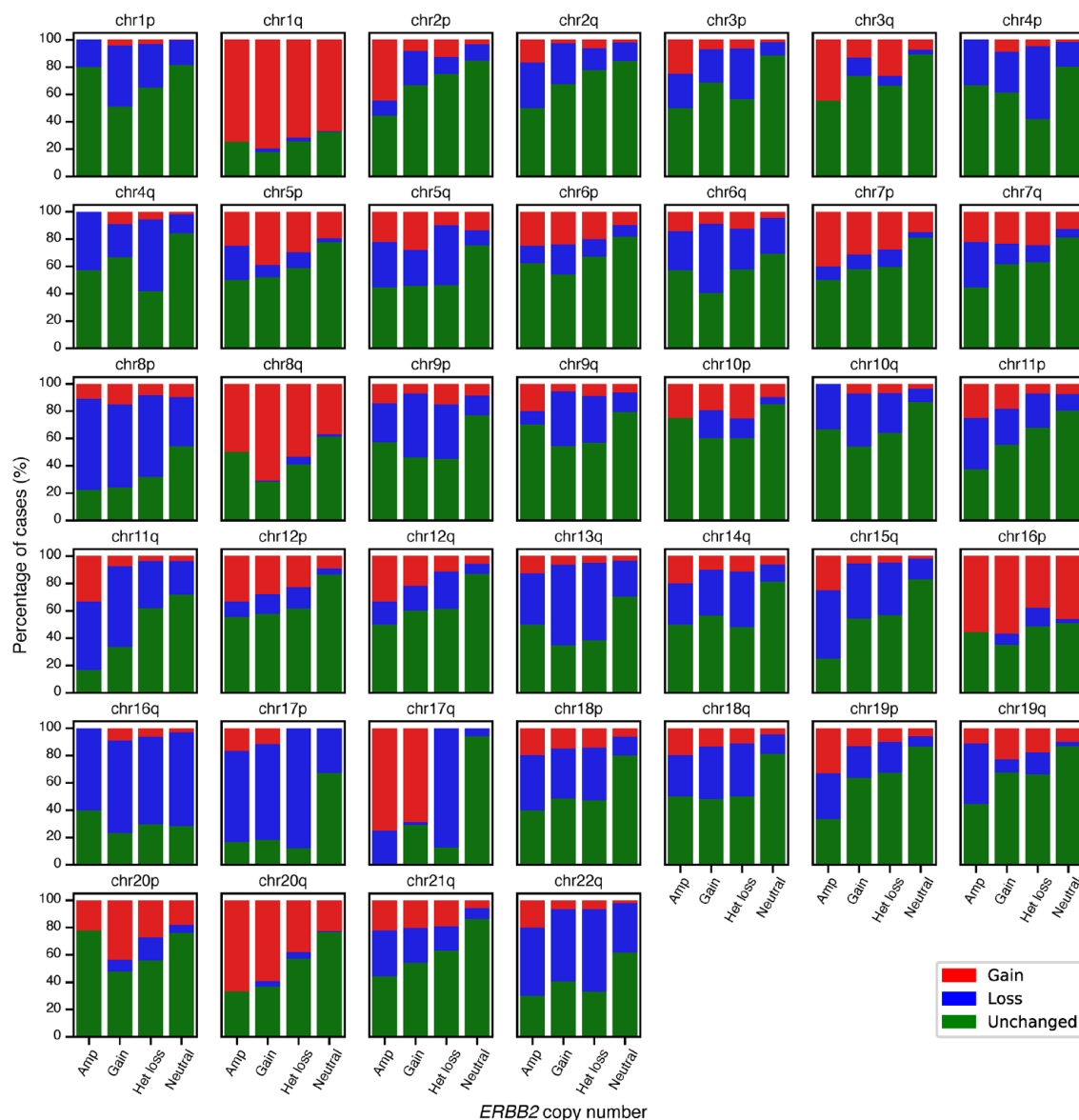


Figure 3. Distribution of TCGA arm-level data with respect to *ERBB2* copy number in each chromosome. In keeping with findings from gene-level CNA analysis, arm-level loss in chromosome 17p was most prevalent regardless of *ERBB2* gain, amplification, or loss. Whereas, chromosome 17q arm-level alterations were largely in tandem with *ERBB2* CNA. Only the TCGA dataset is included since no arm-level data are available for the METABRIC dataset.

ERBB2 copy number gain, loss, or neutral (unchanged).

CNA, copy number alteration; METABRIC, Molecular Taxonomy of Breast Cancer International Consortium; TCGA, the cancer genome atlas.

differential expression analysis was performed among three groups of *ERBB2* CNA status: *ERBB2* neutral (GISTIC=0), gain/amplification (GISTIC=1 or 2), and heterozygous loss (GISTIC=-1) for the whole cohort, hormone receptor-positive, and hormone receptor-negative cohorts of HER2-negative breast cancers from TCGA (Supplemental Table 3). Enrichment for

cell cycle, proteasome, and DNA replication pathways was noted in both *ERBB2* gain/amplification and heterozygous loss subgroups (Supplemental Figure 2A). These findings are mainly attributed to the hormone receptor-positive cohort (Supplemental Figure 2B and C). The METABRIC dataset was not utilized for this analysis since no differentially enriched pathways were

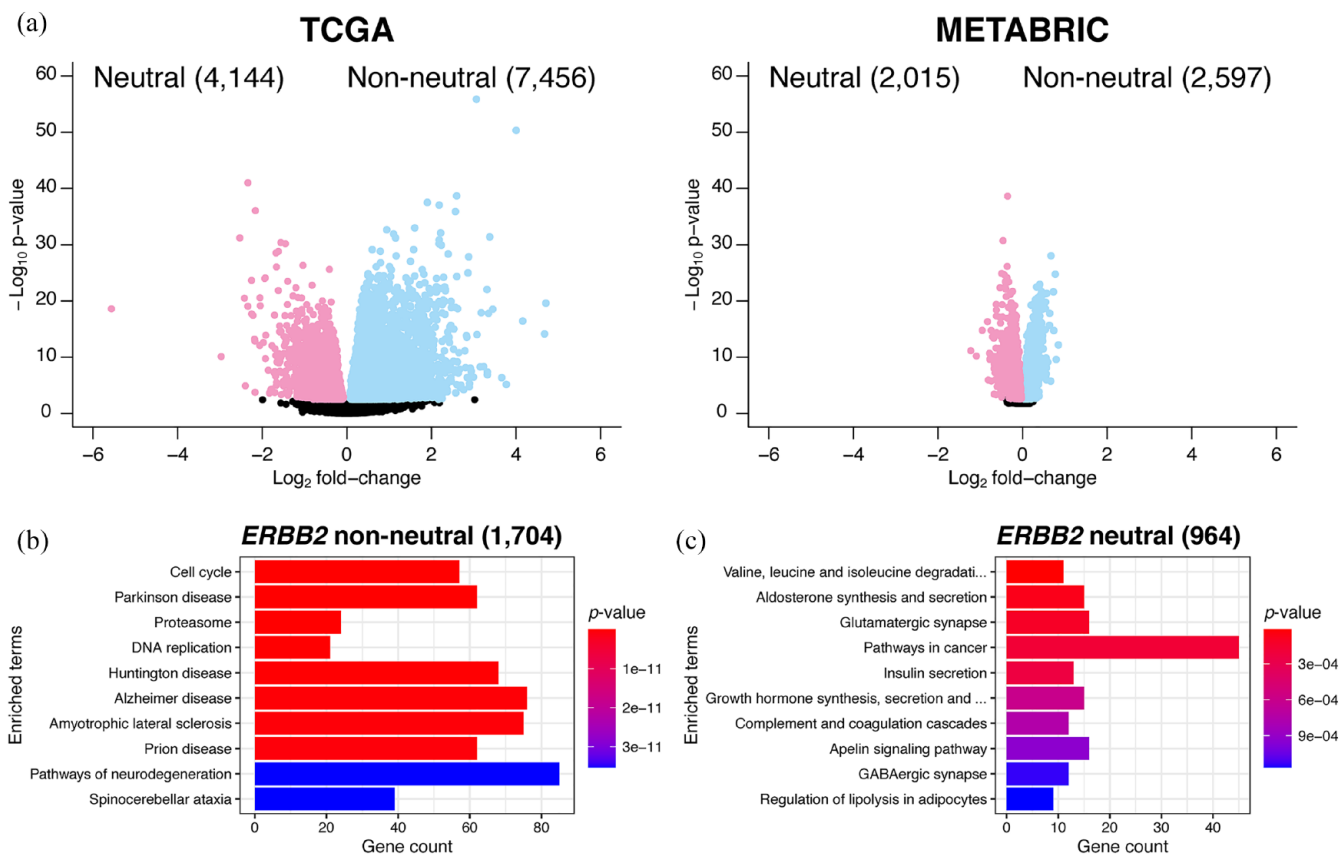


Figure 4. Differential expression analysis between *ERBB2* copy number non-neutral and neutral cases for TCGA and METABRIC datasets. (a) Volcano plots showing fold changes and *p* values of differentially expressed genes in respective datasets, with light blue color representing genes passing the filter of FDR < 0.01 and fold changes > 1 (i.e. genes that are upregulated in non-neutral cases), and light red color representing genes passing the filter of FDR < 0.01 and fold-changes < [-1; i.e. genes that are upregulated in neutral cases). There are 7456 and 2597 genes that are, respectively, upregulated in TCGA and METABRIC datasets of non-neutral cases, as well as 4144 and 2015 genes in neutral cases; 1704 and 964 genes are shared between these two datasets for non-neutral and neutral cases, respectively. These genes are later used in the pathway enrichment analysis of (b) non-neutral and (c) neutral cases with the KEGG database. FDR, false discovery rate; METABRIC, Molecular Taxonomy of Breast Cancer International Consortium; TCGA, The Cancer Genome Atlas.

significantly detected amongst the three sub-groups, due to the smaller number of differentially expressed genes observed with the microarray sequencing platform.

Discussion

The traditional classification of breast cancers based on HER2 status is grounded on protein expression on IHC and/or gene copy number status on ISH.^{2,3} Approximately 60% of ‘HER2-negative’ breast cancers display low immunohistochemical expression of HER2, including in both hormone receptor-positive and hormone receptor-negative subtypes. Although patients with this group of breast cancers do not classically

benefit from HER2-directed therapies, recent clinical trials demonstrated significant improvements in clinical outcomes with trastuzumab deruxtecan, an antibody–drug conjugate composed of an anti-HER2 antibody linked to a topoisomerase I inhibitor payload.⁷ In our previous study, tumors harboring HER2 IHC scores of 1+ showed improved prognosis, as did those with *ERBB2* CNA neutral status.¹⁹ *ERBB2* gene expression levels were correlated with both IHC and CNA scores in the TCGA-METABRIC cohort, with higher gene expression detected in patients with higher IHC and CNA scores, though no association with survival outcomes could be demonstrated based on RNA expression. These results highlight the complexities of

subclassifying HER2-negative breast cancers and may imply new approaches for therapeutic intervention.

Current assays for the assessment of HER2 status rely largely on conventional IHC testing, along with ISH testing when deemed equivocal (2+). The accuracy of these scoring methods, particularly for HER2 IHC in the low (0 and 1+) range, has been questioned. Data assessment from College of American Pathologists surveys showed a poor agreement in the evaluation of HER2 IHC scores of 0 and 1+, while a study on 18 pathologists demonstrated a concordance of only 26% between 0 and 1+ scores. Alternative quantitative methods for improved disease classification and optimal patient selection for therapy are therefore urgently warranted.¹¹ Our results suggest that *ERBB2* CNA estimation may offer a feasible approach for this purpose, although the ideal assay remains to be determined. Standard targeted FISH assays feature five ASCO-CAP groupings, with classic HER2-amplified breast cancer defined by HER2/CEP17 ratio ≥ 2 and mean HER2 copy number ≥ 4 . Classic HER2 non-amplified breast cancer is defined by HER2/CEP17 ratio < 2 and mean HER2 copy number < 4 , while the other four groups are defined as negative unless trumped by concurrent IHC scores.³⁴ The minimum cutoffs representing 'HER2-low' will require further studies.

Genome-wide assays such as array comparative genomic hybridization (aCGH), single nucleotide polymorphism arrays, and next-generation sequencing (NGS) methods offer yet additional options for obtaining *ERBB2* CNA status, along with other potentially actionable genomic information.^{35–38} To our knowledge, data comparing contemporary IHC and FISH with various genome-wide assays are not available currently. In addition, the impact of intratumoral heterogeneity will need to be carefully assessed. The question of whether *ERBB2* CNA status can predict benefit from novel HER2 ADCs also merits further investigation. *ERBB2* heterozygous loss might be associated with upfront resistance to trastuzumab–deruxtecan based on biomarker analyses from the DAISY trial.^{39,40} Perhaps with the increasing adoption of NGS technologies in the clinical setting, an NGS-based diagnostic approach to *ERBB2* CNA status assessment that complements conventional IHC/FISH testing might be feasible in the near future.

Breast cancer is known to have a generally low mutation burden, but is characterized by the high number of CNAs compared to other tumor types.⁴¹ From our analysis, nearly two-fifths of HER2-negative non-metastatic breast cancers harbored *ERBB2* CNAs, with heterozygous loss being most commonly observed. Non-neutral *ERBB2* CNA status was significantly more prevalent within hormone receptor-negative as compared to hormone receptor-positive breast cancers (56.1% versus 35.5%, respectively) and was previously shown to be an independent prognostic factor for worse relapse-free survival.¹⁹ Interestingly, non-neutral *ERBB2* status was associated with a significantly higher genome-wide burden of CNAs, although there were no statistically significant differences in TMB and ploidy. In terms of specific genes, somatic alterations (particularly CNA) in genes on chromosome 17 – *TP53*, *NF1*, *BRCA1*, *MAP2K4*, and *NCOR1* – were more prevalent in the *ERBB2* non-neutral group. This observation is related to chromosomal arm-level events and is consistent with previous studies reporting a high frequency of 17p (short arm) losses but with complex combinations of gains and losses within 17q (long arm).⁴² Furthermore, our results suggest that *ERBB2* CNA, particularly heterozygous loss, is accompanied almost always by concurrent loss of major tumor suppressor genes on chromosome 17.^{43–46} The implications of these genes as potential contributors driving the inferior prognosis in this group of breast cancers warrant further investigation. In keeping with the association of *ERBB2* CNA with worse patient survival outcomes,¹⁹ analysis of differential gene expression revealed upregulated oncogenic pathways such as those involved in cell cycle signaling and DNA replication in the *ERBB2* non-neutral group. Our findings are in keeping with a previous report demonstrating high somatic CNA burden with worse survival in patients with breast cancer,⁴⁷ as well as in prostate, endometrial, renal clear cell, thyroid, and colorectal cancer.⁴⁸ In other studies, higher somatic CNA burden was also associated with poorer responses to immunotherapy in melanoma,^{49,50} non-small-cell lung cancer,⁵¹ gastrointestinal cancer,⁵² and other metastatic cancers.⁴¹

The current findings are limited by the availability of datasets used for the *in silico* analysis. We focused on non-metastatic, stage I–III HER2-negative breast cancers to follow up on our

previous findings of inferior survival in *ERBB2* non-neutral breast cancer in the non-metastatic setting.¹⁹ Specific results pertaining to breast cancer subtypes, particularly hormone receptor-negative cases, will require validation in larger cohorts. Given that the new classification of HER2-low breast cancers impacts treatment decisions only in the metastatic setting currently, additional investigation in metastatic breast cancers may be required.

Conclusion

In conclusion, our work highlights the complexity of HER2-negative breast cancers in the context of their copy number status. Apart from differences in survival outcomes, the classification of HER2-negative breast cancers according to *ERBB2* CNA status reveals distinct patterns of genomic aberrations especially in chromosome 17 genes. The biological and therapeutic implications of CNA classification merit further research.

Declarations

Ethics approval and consent to participate

Ethics approval and consent to participate were obtained as part of TCGA and METABRIC studies.

Consent for publication

Consent to publication was obtained as part of TCGA and METABRIC studies.

Author contributions

Jui Wan Loh: Data curation; Formal analysis; Investigation; Writing – original draft.

Abner Herbert Lim: Data curation; Formal analysis; Investigation.

Jason Yongsheng Chan: Conceptualization; Investigation; Supervision; Writing – review & editing.

Yoon-Sim Yap: Conceptualization; Investigation; Supervision; Writing – review & editing.

Acknowledgements

None.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was

supported by Singapore National Medical Research Council (NMRC) Clinician Scientist Award (MOH-CSAINV18nov-0009) to Dr Yoon-Sim Yap and Transition Award (TA21jun-0005) to Dr Jason Yongsheng Chan.

Competing interests

Y-SY – Honoraria: Novartis, Pfizer, Lilly/DKSH, AstraZeneca, Eisai, MSD, Roche; Research funding: Merck Sharp & Dohme; Travel, accommodations, expenses: DKSH, AstraZeneca, Pfizer. The rest of authors report that they have no related competing interests or financial disclosures to declare.

Availability of data and materials

All data are available from the authors upon reasonable request.

ORCID iD

Jason Yongsheng Chan  <https://orcid.org/0000-0002-4801-3703>

Supplemental material

Supplemental material for this article is available online.

References

1. Slamon DJ, Godolphin W, Jones LA, *et al.* Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science* 1989; 244: 707–712.
2. Slamon DJ, Leyland-Jones B, Shak S, *et al.* Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med* 2001; 344: 783–792.
3. Baselga J, Cortés J, Kim SB, *et al.* Pertuzumab plus trastuzumab plus docetaxel for metastatic breast cancer. *N Engl J Med* 2012; 366: 109–119.
4. Tarantino P, Hamilton E, Tolane SM, *et al.* HER2-Low breast cancer: pathological and clinical landscape. *J Clin Oncol* 2020; 38: 1951–1962.
5. Banerji U, van Herpen CML, Saura C, *et al.* Trastuzumab duocarmazine in locally advanced and metastatic solid tumours and HER2-expressing breast cancer: a phase 1 dose-escalation and dose-expansion study. *Lancet Oncol* 2019; 20: 1124–1135.
6. Modi S, Park H, Murthy RK, *et al.* Antitumor activity and safety of trastuzumab deruxtecan in

- patients with HER2-low-expressing advanced breast cancer: results from a phase Ib study. *J Clin Oncol* 2020; 38: 1887–1896.
7. Modi S, Jacot W and Yamashita T. Trastuzumab deruxtecan in previously treated HER2-low advanced breast cancer. *N Engl J Med* 2022; 387: 9–20.
 8. Schettini F, Chic N, Brasó-Maristany F, *et al.* Clinical, pathological, and PAM50 gene expression features of HER2-low breast cancer. *NPJ Breast Cancer* 2021; 7: 1.
 9. Denkert C, Seither F, Schneeweiss A, *et al.* Clinical and molecular characteristics of HER2-low-positive breast cancer: pooled analysis of individual patient data from four prospective, neoadjuvant clinical trials. *Lancet Oncol* 2021; 22: 1151–1161.
 10. Tarantino P, Jin Q, Tayob N, *et al.* Prognostic and biologic significance of ERBB2-Low expression in early-stage breast cancer. *JAMA Oncol* 2022; 8: 1177–1183.
 11. Fernandez AI, Liu M, Bellizzi A, *et al.* Examination of low ERBB2 protein expression in breast cancer tissue. *JAMA Oncol* 2022; 8: 1–4.
 12. Allison KH and Wolff AC. ERBB2-Low breast Cancer-Is it a fact or fiction, and do we have the right assay? *JAMA Oncol* 2022; 8: 610–611.
 13. Geukens T, De Schepper M, Richard F, *et al.* Intra-patient and inter-metastasis heterogeneity of HER2-low status in metastatic breast cancer. *Eur J Cancer* 2023; 188: 152–160.
 14. Horisawa N, Adachi Y, Takatsuka D, *et al.* The frequency of low HER2 expression in breast cancer and a comparison of prognosis between patients with HER2-low and HER2-negative breast cancer by HR status. *Breast Cancer* 2022; 29: 234–241.
 15. Mutai R, Barkan T, Moore A, *et al.* Prognostic impact of HER2-low expression in hormone receptor positive early breast cancer. *Breast* 2021; 60: 62–69.
 16. Frenel J-S, Lusque A, Mailliez A, *et al.* 291P HER2-low metastatic breast cancer (MBC): management and prognosis of a new breast cancer entity in a real-world setting. *Ann Oncol* 2021; 32: S491.
 17. Hein A, Hartkopf AD, Emons J, *et al.* Prognostic effect of low-level HER2 expression in patients with clinically negative HER2 status. *Eur J Cancer* 2021; 155: 1–12.
 18. Gampenrieder SP, Rinnerthaler G, Tinchon C, *et al.* Landscape of HER2-low metastatic breast cancer (MBC): results from the Austrian AGMT_MBC-Registry. *Breast Cancer Res* 2021; 23: 112.
 19. Tan RSYC, Ong WS, Lee KH, *et al.* HER2 expression, copy number variation and survival outcomes in HER2-low non-metastatic breast cancer: an international multicentre cohort study and TCGA-METABRIC analysis. *BMC Med* 2022; 20: 105.
 20. Won HS, Ahn J, Kim Y, *et al.* Clinical significance of HER2-low expression in early breast cancer: a nationwide study from the Korean Breast Cancer Society. *Breast Cancer Res* 2022; 24: 22.
 21. The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nat* 2012; 490: 61–70.
 22. Curtis C, Shah SP, Chin SF, *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* 2012; 486: 346–352.
 23. Cerami E, Gao J, Dogrusoz U, *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2012; 2: 401–404.
 24. Gao J, Aksoy BA, Dogrusoz U, *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* 2013; 6: 1.
 25. Mermel CH, Schumacher SE, Hill B, *et al.* GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol* 2011; 12: R41.
 26. Mayakonda A, Lin DC, Assenov Y, *et al.* Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res* 2018; 28: 1747–1756.
 27. Nik-Zainal S, Davies H and Staaf J. Landscape of somatic mutations in 560 breast cancer whole genome sequences. *Nat* 2016; 534: 47–54.
 28. Bao L, Pu M and Messer K. AbsCN-seq: a statistical method to estimate tumor purity, ploidy and absolute copy numbers from next-generation sequencing data. *Bioinformatics* 2014; 30: 1056–1063.
 29. Love MI, Huber W and Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014; 15: 550.
 30. Ritchie ME, Phipson B, Wu D, *et al.* Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015; 43: e47–e47.

31. Chen EY, Tan CM, Kou Y, *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* 2013; 14: 128.
32. Kuleshov MV, Jones MR, Rouillard AD, *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* 2016; 44: W90–W97.
33. Xie Z, Bailey A, Kuleshov MV, *et al.* Gene set knowledge discovery with Enrichr. *Curr Protoc* 2021; 1: e90.
34. Wolff AC, Hammond MEH, Allison KH, *et al.* Human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Focused Update. *J Clin Oncol* 2018; 36: 2105–2122.
35. Yeh IT, Martin MA, Robetorye RS, *et al.* Clinical validation of an array CGH test for HER2 status in breast cancer reveals that polysomy 17 is a rare event. *Mod Pathol* 2009; 22: 1169–1175.
36. Staaf J, Jönsson G, Ringnér M, *et al.* Landscape of somatic allelic imbalances and copy number alterations in HER2-amplified breast cancer. *Breast Cancer Res* 2011; 13: R129.
37. Ross DS, Zehir A, Cheng DT, *et al.* Next-generation assessment of human epidermal growth factor receptor 2 (ERBB2) amplification status: clinical validation in the context of a hybrid Capture-Based, comprehensive solid tumor genomic profiling assay. *J Mol Diagn* 2017; 19: 244–254.
38. Lawrence MS, Stojanov P, Polak P, *et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 2013; 499: 214–218.
39. Mosele MF, Lusque A, Dieras VC, *et al.* LBA72 Unraveling the mechanism of action and resistance to trastuzumab deruxtecan (T-DXd): biomarker analyses from patients from DAISY trial. *Ann Oncol* 2022; 33: S1440–S869.
40. Mosele F, Deluche E, Lusque A, *et al.* Trastuzumab deruxtecan in metastatic breast cancer with variable HER2 expression: the phase 2 DAISY trial. *Nat Med* 2023; 29: 2110–2120.
41. Liu L, Bai X, Wang J, *et al.* Combination of TMB and CNA stratifies prognostic and predictive responses to immunotherapy across metastatic cancer. *Clin Cancer Res* 2019; 25: 7413–7423.
42. Orsetti B, Nugoli M, Cervera N, *et al.* Genomic and expression profiling of chromosome 17 in breast cancer reveals complex patterns of alterations and novel candidate genes. *Cancer Res* 2004; 64: 6453–6460.
43. Lakhani SR, Jacquemier J and Sloane JP. Multifactorial analysis of differences between sporadic breast cancers and cancers involving BRCA1 and BRCA2 mutations. *J Natl Cancer Inst* 1998; 90: 1138–1145.
44. Pharoah PD, Day NE and Caldas C. Somatic mutations in the p53 gene and prognosis in breast cancer: a meta-analysis. *Br J Cancer* 1999; 80: 1968–1973.
45. Griffith OL, Spies NC, Anurag M, *et al.* The prognostic effects of somatic mutations in ER-positive breast cancer. *Nat Commun* 2018; 9: 3476.
46. Noblejas-López MDM, Morcillo-García S, Nieto-Jiménez C, *et al.* Evaluation of transcriptionally regulated genes identifies NCOR1 in hormone receptor negative breast tumors and lung adenocarcinomas as a potential tumor suppressor gene. *PLoS One* 2018; 13: e0207776.
47. Zhang L, Feizi N, Chi C, *et al.* Association analysis of somatic copy number alteration burden with breast cancer survival. *Front Genet* 2018; 9: 421.
48. Hieronymus H, Murali R, Tin A, *et al.* Tumor copy number alteration burden is a pan-cancer prognostic factor associated with recurrence and death. *eLife* 2018; 7: e37294.
49. Roh W, Chen PL, Reuben A, *et al.* Integrated molecular analysis of tumor biopsies on sequential CTLA-4 and PD-1 blockade reveals markers of response and resistance. *Sci Transl Med* 2017; 9: eaah3560.
50. Davoli T, Uno H, Wooten EC, *et al.* Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science* 2017; 355: eaaf8399.
51. Kim HS, Cha H, Kim J, *et al.* Genomic scoring to determine clinical benefit of immunotherapy by targeted sequencing. *Eur J Cancer* 2019; 120: 65–74.
52. Lu Z, Chen H, Li S, *et al.* Tumor copy-number alterations predict response to immune-checkpoint-blockade in gastrointestinal cancer. *J Immunother Cancer* 2020; 8: e000374.