Original Studies

# Evaluation of therapeutic target gene expression based on residual cancer burden classification after neoadjuvant chemotherapy for HER2-negative breast cancer

Yuko Takahashi [1], Takayuki Iwamoto[2*], Yoko Suzuki [2], Yukiko Kajiwara [2], Minami Hatono [1], Takahiro Tsukioki [1], Kengo Kawada [1], Mariko Kochi [2], Hirokuni Ikeda [1], Tadahiko Shien [1], Naruto Taira [1], Junji Matsuoka [3], Hiroyoshi Doihara [1], and Shinichi Toyooka [1]

[1]Department of General Thoracic Surgery and Breast and Endocrinological Surgery, Graduate School of Medicine Dentistry and Pharmaceutical Sciences, Okayama University, Okayama, Japan, [2]Departments of Breast and Endocrine Surgery, Okayama University Hospital, Okayama, Japan, [3]Departments of Palliative and Supportive Medicine, Okayama University Hospital, Okayama, Japan

**Corresponding author:**

Takayuki Iwamoto, M.D., Ph D.

2-5-1, Shikata-cho, Kita-ku, Okayama 700-8558, Japan

E-mail : tiwamoto@md.okayama-u.ac.jp

**Key words:** Residual Tumor Burden; neoadjuvant chemotherapy; targeted therapy; hormone receptor positive; triple negative

**Short title**

Gene expression based on residual cancer burden classification for breast cancer

Conflict of Interest Disclosures: The authors declare no potential conflicts of interest.

**MicroAbstract**

Patients with residual diseases (RD) usually have poor prognoses after standard neoadjuvant chemotherapy for breast cancer. We explored novel therapeutic targets and potential additional adjuvant treatments for patients with RD after standard neoadjuvant chemotherapy, and found therapeutic targets by ER status.

Abstract

Purpose: Patients with residual diseases (RD) usually have poor prognoses after neoadjuvant chemotherapy for breast cancer. The aim of this study was to explore therapeutic targets and potential additional adjuvant treatments for patients with RD after standard neoadjuvant chemotherapy.

Materials and Methods: We retrieved publicly available cDNA microarray data from 399 human epidermal growth factor 2 negative primary breast cancer samples from patients undergone standard neoadjuvant chemotherapy. We analyzed the mRNA expression levels of key breast cancer markers and therapeutic target genes based on residual cancer burden (RCB) classification: RCB-0/I, RCB-II, and RCB-III.

Results: Among hormone receptor (HR) -positive samples, there were more luminal A tumors by PAM50 in RCB-III than in RCB-0/I and RCB-II ($P < 0.01$). The mRNA expressions of *ESR1* and *PGR* were significantly higher and that of *MKI67* was lower in RCB-II and RCB-III than in RCB-0/I. The mRNA expression of *cyclin D1* was upregulated in RCB-III and that of *CDKN2A* was down-regulated in RCB-III ($P = 0.027$ and $< 0.01$). Among triple negative (TN) samples, RCB-III had higher clinical Stage and more lymph node-positive samples than RCB-0/1 and RCB-II ($P < 0.01$). In both subtypes, *VEGF-C* expression was significantly higher in RCB-III than in RCB-0/I and RCB-II.

Conclusion: In HR-positive breast cancer, biological features such as luminal A were associated with RCB; this trend was not observed in TN breast cancer. Further, some targeted therapies should be tested as new strategies after standard neoadjuvant chemotherapy in future clinical trials.

## Introduction

Neoadjuvant chemotherapy has become a standard treatment option for breast cancer [1]. Apart from reducing tumor size and increasing the rate of breast-conserving surgeries, it also allows us to assess the response of the patient to systemic treatment. Furthermore, neoadjuvant therapy is an important strategy for drug development and identification of predictive biomarkers[2]. Clinical trials in neoadjuvant settings are increasingly being conducted for breast cancer; however, discovering new agents and novel therapeutic strategies in adjuvant settings could incur high costs and are highly time-consuming[3].

Neoadjuvant trials can help discover novel predictive biomarkers that can be validated by additional studies in neoadjuvant and adjuvant settings before being adopted for clinical practice [3]. Neoadjuvant trials can also be utilized to test the efficacy of biomarker-driven targeted therapies against different breast cancer subtypes [3].

The pathological complete response (pCR) after neoadjuvant chemotherapy is an important prognostic factor and a surrogate maker for long-term outcome in patients with primary breast cancer[4, 5]. Previous studies have shown that patients who achieved pCR after neoadjuvant chemotherapy had better prognoses in specific intrinsic subtypes, especially triple negative (TN) and human epidermal growth factor 2 (HER2)-positive breast cancer [5]. Patients with TN breast cancers who achieved pCR had excellent survival, while those who did not achieve pCR after neoadjuvant chemotherapy had significantly shorter overall and post-recurrence survival; more than 30% of recurrent patients with residual diseases had an overall survival (OS) of 3 years [6].

Previous meta-analyses have shown that pCR is not a suitable surrogate marker for long-term prognosis in hormone receptor (HR)-positive and HER2-negative subtypes; other potential surrogate markers should therefore be investigated for HR-positive breast cancer [7]. In cases that have not achieved pCR, the combination of residual tumor size, tumor cellularity and nodal status after neoadjuvant treatment is prognostic [8]. Symmans et al. introduced a measure called residual cancer burden (RCB), a continuous variable derived from the largest area and cellularity of residual invasive primary cancer, the number of involved lymph nodes, and the size of the largest metastasis [8]. They divided all tumors observed after neoadjuvant chemotherapy into four classes based on predefined cut points of 1.36 and 3.28 index scores:

pCR (stage yp-T0/is, ypN0; RCB = 0), minimal RCB (RCB-I), moderate RCB (RCB-II), and extensive RCB (RCB-III) [8, 9]. The pathological response to neoadjuvant chemotherapy based on RCB scores and classes is prognostic for all phenotypic subtypes of breast cancer, including HR-positive, HER2-positive, and TN [9]. For subgroups that have poorer prognoses and show residual breast cancer after standard neoadjuvant treatment, additional new treatment strategies are needed.

Several clinical trials have been conducted to evaluate the efficacy of adjuvant chemotherapy or targeted therapy for patients with residual breast cancer after standard neoadjuvant chemotherapy [10]. Masuda et al. reported the results of the Capecitabine for Residual Cancer as Adjuvant Therapy (CREATE-X) trial, a multi-center, open-label, randomized, phase-3 trial designed to evaluate the efficacy and safety of adjuvant capecitabine monotherapy in patients with HER2-negative primary breast cancer, who developed residual invasive diseases after standard neoadjuvant chemotherapy with anthracycline, taxane, or both [10]. Adjuvant capecitabine therapy was found to prolong disease-free survival (DFS) and OS among these patients. Capecitabine is therefore an effective adjuvant option in these patients [10]. It is still unclear whether there are other suitable candidates for additional adjuvant treatment agents.

The aim of this study was to explore therapeutic targets for patients with residual breast cancer based on RCB classes and to investigate other potential adjuvant treatments after standard neoadjuvant chemotherapy for subgroups with poor prognoses. We also performed functional analyses of the properties of RCB.

## Materials and methods

### Patients and cohort

We retrieved publicly available cDNA microarray data from 508 primary breast cancer samples (all patients with those samples had received anthracycline and taxane-based neoadjuvant chemotherapy from NCBI Gene Expression Omnibus [GEO] repository [https://www.ncbi.nlm.nih.gov/geo/] GSE 25066) [11].The cDNA microarray data were corrected from tumor biopsy samples prior to any systemic therapy[11]. ER and PR status were assessed using immunohistochemistry (IHC) (6F11; Novocastra Laboratories Ltd., Newcastle, UK). The cutoff for ER positivity and PR positivity was set at 10% positive tumor cells

after nuclear staining. Pathological HER2 positivity was defined as IHC (Dako North America Inc., Carpinteria, CA) staining of 3+ and/or amplification of HER2 gene copy number by fluorescent in situ hybridization (FISH) based on American Society of Clinical Oncology / College of American Pathologists guidelines [11, 12]. From the 508 samples retrieved, we excluded 109 HER2-positive samples, and the remaining 399 were analyzed. HR-positive breast cancer was defined as ER- and/or PR-positive, and HER2-negative and TN breast cancers were defined as ER-, PR-, and HER2-negative. All HR-positive patients received adjuvant hormone therapy. Clinical nodal status was determined before treatment through physical examination with or without axillary ultrasound and diagnostic fine-needle aspiration, as required [11]. The clinical and pathological characteristics for the cohort studied are shown in Online appendix 1.

This data was annotated using the Affymetrix Human Genome Array (Affymetrix Inc., Santa Clara, CA). All gene expression data were generated using Affymetrix gene chips and normalized using the Mas5 algorithm (http://www.bioconductor.org) with log 2 transformation.

All 399 breast cancer samples were stratified into three groups based on the RCB classification system: RCB-0/I, RCB-II, and RCB-III.

**Gene expression analysis**

First, we compared the mRNA expression levels of four well-established breast cancer markers: *ESR1*, *PGR*, *ERBB2*, and *MKI67*. To evaluate the associations between RCB classes and gene expression levels, we used the Wilcoxon rank sum test to distinguish between all the pairs of RCB groups and the Kruskal–Wallis rank sum test to distinguish between the three RCB groups.

Second, we selected 41 genes that are targeted by FDA-approved drugs or have been investigated with clinical trials as molecular target agents for different malignant tumors, including breast cancer, to explore additional novel adjuvant therapies for breast cancer patients with poor prognosis after neoadjuvant chemotherapy. The information on anticancer therapy drugs was obtained from National Cancer Institute (NCI) drug information [13], Drug@FDA [14], and Clinical Trials gov [15, 16].

Some of the 41 selected genes were associated with breast cancer (*AR*, *ERBB3*, and *p53*), DNA damage repair pathways and BRCA functions (*BRCA1*, *BRCA2*, *PARP1*, and *PARP2*), cyclin dependent kinase (CDK) pathway (*CDK2*, *CDK4*, *CDK6*, *CCND1*, *CDKN2A*, and *RB1*), vascular endothelial growth factor (VEGF) and VEGF receptor pathways (*VEGF-A*, *VEGF-B*, *VEGF-C*, *EGFR*, *PGF*, *KDR*, and *FTL4*), modulation of DNA methylation and histone acetylation (*HDAC1*, *HDAC2*, *HDAC3*, *DNMT1*, *DNMT3A*, and *DNMT3B*), immune responses (*PDCD1LG2*), and mTOR pathway (*mTOR*, *PIK3CA*). Others were used in FDA-approved drugs or under investigated for cancers other than breast cancer (*AKT1*, *ALK*, *RAF1*, *CTNNB1*, *MET*, *STK11*, *PTEN*, *NF1*, *ROS1*, *NOTCH1*, *ATM*, *KITL*, and *KRAS*).

We used the Wilcoxon rank sum test and the Kruskal-Wallis rank sum test to evaluate the associations between RCB classes and gene expression levels in the different breast cancer subtypes.

All statistical analyses were performed using the BRB Array Tools software (version 4.5.1; http://linus.nci.nih.gov/BRB-ArrayTools.html) and R software (version 3.4.1; http://www.r-project.org). Differences with two-sided P values ≤ 0.05 were considered statistically significant. This was retrospective study from the public database and there was no need for ethical approval by the institutional review board.

## Results

### Patients characteristics

We analyzed previously published gene expression data of 399 breast cancer samples obtained from Hatzis et al. (available at GSE 25066) and classified them based on their RCB classification [11]. Clinical and pathological characteristics of the samples are shown in Table 1 and Online appendix 1. Among the 399 samples, 206 (52%) were lymph node-positive, 168 (42%) were cStage III, 204 (51%) were histological grade III, and 253 (63%) were HR-positive. For the HR-positive samples, the different RCB subgroups had significantly different clinical nodal status, cStage before neoadjuvant chemotherapy, ER IHC status, and PAM50 classification ($P$ = 0.045, 0.0068 and <0.001, respectively). As shown in Table 1, among the HR-positive cases, RCB-0/I tumors were of significantly higher nuclear and histological grade than RCB-III tumors ($P$ = 0.0049); there were more luminal A tumors in RCB-III than in RCB-0/1 and

RCB-II (63% and 22%, respectively). The rate of clinical node-negative tumors was higher in RCB-III than in RCB-0/I and RCB-II ($P < 0.001$) (Table 1). Thus, tumors that are node-negative and show low proliferation rate may be associated with RCB-III and poor prognosis.

In contrast, for TN breast cancer samples, only clinical nodal status ($P = 0.0053$) and cStage ($P = 0.0029$) were significantly different among the RCB groups; as expected, RCB III had more cStage and lymph node-positive tumors than the other groups (Table 1).

**Gene expression analysis of four well-established breast cancer markers according to RCB subclasses**

To explore the associations between RCB subclasses and expression levels of well-established breast cancer markers, we deduced the mRNA gene expression levels of ESR1, PGR, ERBB2, and MKI67 separately by hormone receptor status.

**Hormone receptor-positive breast cancer**

The mRNA expression levels of ESR1 and PGR were significantly higher in RCB-II and RCB-III tumors than in RCB-0/I tumors ($P = 0.00053$ and $P = 0.0061$, respectively) (Figure 1). However, the expression level of MKI67 was lower in RCB-III than in RCB-0/I and RCB-II, while that of *ERBB2* was not significantly different between the RCB groups.

**Triple negative breast cancer**

The mRNA expression level of ESR1 was higher in RCB-III tumors than in RCB-I and RCB-II tumors, although the average expression level in all subgroups were lower than those in ER-positive tumors (Figure 1). In our previous study, we found ESR1 mRNA > 10.18 and ERBB2 mRNA > 12.54, and both were defined as ER- and HER2-positive [17]. The expression levels of PGR, ERBB2, and MKI67 were not significantly different between the RCB subgroups.

**Gene expression analysis of 41 molecular target markers according to RCB subclasses**

Next, we analyzed 41 molecular target markers according to RCB subclasses with the aim of seeking novel drug targets suitable for use in cases with poor prognosis after standard neoadjuvant chemotherapy. The 41 selected genes are listed in Online appendix 2.

**Hormone receptor-positive breast cancer**

We found that the mRNA expression level of cyclin D1 (CCND1) was significantly upregulated in RCB-III ($P = 0.027$), while those of CDKN2A ($P = 0.0047$) were significantly down-regulated in RCB-III. (Figure 2) We also observed upregulated VEGF-C expression in RCB-III, although other VEGF- and VEGF-R-related genes were not significantly upregulated in RCB-III (Figure 3).

In RCB-0/I, DNMT1 and DNMT3A mRNAs were significantly overexpressed ($P < 0.0001$ and $P = 0.014$, respectively). However, the expression levels of DNMT3B and HDAC family genes were not significantly different between the three RCB subgroups (Online appendix 3).

No other genes showed significant differences in their expression levels between RCB subgroups in the HR-positive samples.

**Triple negative breast cancer**

The expression level of VEGF-C was higher in RCB-III than in RCB-0/I and RCB-II ($P = 0.029$) (Figure 3). We found also that the mRNA expression level of CDKN2A ($P = 0.044$) were significantly down-regulated in RCB-III, however the difference was smaller than that shown in hormone receptor positive tumors (Figure2).DNMT3A was also upregulated in RCB-0/I than RCB-II ($P = 0.039$), although there was no difference in the expression levels of DNMT1 and DNMT3B between the RCB subgroups (Online appendix 3). No other genes showed significant differences in their expression levels between the different RCB subgroups in TN breast cancer samples.

**Discussions**

The RCB scoring system could be a reliable prognostic marker after neoadjuvant chemotherapy; RCB-III breast cancers after neoadjuvant chemotherapy show poor prognosis, regardless of hormone receptor

status [8, 9]. RCB-III breast cancers show poor outcomes despite favorable outcomes for most other HR-positive cancers [9]. In this study, we analyzed the relationship between poor prognoses and expression of candidate target genes; we also explored additional targeted therapies to improve prognosis after standard neoadjuvant chemotherapy.

First, we found that luminal A-like cases with higher clinical stages had poorer prognosis than luminal B, indicating that some luminal A-like cases might have poorer responses to chemotherapy. We next focused on pathological background features and the expression levels of four established breast cancer marker genes: ESR1, PGR, ERBB2, and MKI67. Among HR-positive samples, RCB-III contained more tumors with low proliferation characteristics, low expression of MKI67, high expression of PGR, luminal A by PAM 50 than RCB-0/I and RCB-II (Table 1). This was an interesting result, as Symmans et al. had shown that RCB score was a prognostic marker and RCB-III tumors have poor prognoses [8, 9], although previous studies had shown that luminal A-like tumors have excellent prognoses [18]. It is understood that luminal A tumors generally have better prognosis, while a small population of RCB-III tumors have poorer prognoses and may be resistant to chemotherapy and hormone therapy. These results were also supported by a previous study that used the same dataset (GSE25066), in which Symmans et al. showed that most RCB-III tumors were also resistant to hormone therapy [19]. Thus, our results suggest that RCB-III tumors show poor prognoses due to resistance to treatment, despite most of them being luminal-A tumors. Such HR-positive cases that show resistance to chemotherapy and hormone therapy should be tested with additional treatment strategies. In contrast, for TN breast cancers, there was no significant difference in the expression levels of the four marker genes between the RCB subgroups; however, there were more tumors of higher clinical stage and nodal metastasis in RCB-III than in RCB-I and RCB-II. (Table 1) Our results suggest that the clinical stage of the tumor influences the effectiveness of chemotherapy more than biological features (e.g., ER, PR, HER2 and Ki67) for the TN breast cancers, but not HR-positive ones.

HR-positive breast cancers show better prognoses than HER2-positive or TN ones. Our results suggest that cyclin-dependent kinase inhibitors may be suitable therapeutic targets for RCB-III HR-positive breast cancers after standard neoadjuvant chemotherapy. A cyclin-dependent kinase inhibitor, palbociclib, was recently approved around the world, based on the results of PALOMA-1, PALOMA-2, and PALOMA-3 trials for metastatic HR-positive and HER2-negative breast cancer. Based on a previous in vitro study, it

was hypothesized that the expression levels of Cyclin D1 and p16 are related to the response to palbociclib [20]. We found that the CCND1 (cyclin D1) mRNA was overexpressed in RCB-III tumors, while the CDKN2A mRNA was under-expressed (Figure 2). However, in the PALOMA-1 trial, a subset of patients with CCND1 amplification was assessed, and no difference was observed between patients with and without such amplification [21]. Further, in PALOMA-2, no differential benefit was observed in the palbociclib treatment group, whose tumors showed different expression levels of Cyclin D1 and p16 through immune-histochemical staining [22]. Nevertheless, it must be noted that immunohistochemistry is not accurate enough to assess the expression level of CCND1, and we suggest that assessing mRNA levels within tumors might provide more accurate perception. Therefore, the role of CCND1 amplification and/or loss of p16 in patient selection remains unclear and need to be studied further [23]. The German Breast Group and the International Collaborating Breast Cancer Group launched a study called A Study of Palbociclib in Addition to Standard Endocrine Treatment in Hormone Receptor Positive Her2 Normal Patients With Residual Disease After Neoadjuvant Chemotherapy and Surgery (PENELOPE-B). PENELOPE-B is designed to demonstrate that, with background standard endocrine therapy, palbociclib can provide superior invasive disease-free survival in pre- and postmenopausal women with HR-positive/HER2-negative early breast cancer who are at high risk of relapse after showing less than pCR to neoadjuvant therapy with taxane. We believe that the results of the PENELOPE-B trial will confirm the utility of palbociclib as adjuvant therapy for patients with residual tumor and poor prognosis after neoadjuvant chemotherapy.

Remarkably, the expression of CCND1 was not different between the RCB subgroups among TN breast cancers, although CDKN2A was over expressed in RCB-0/I compared to RCB-II and RCB-III (Figure 1). This was contrary to the results obtained for HR-positive tumors. The role of CDKN2A (p16) in TN breast cancer has been studied [24]. Arima et al. reported that a lack of p16 expression is associated with a reduced response of tumors to chemotherapy, possibly because of the acquisition of cancer stem cell-like properties; they also showed that downregulation of p16 expression is a marker for poor response to some chemotherapeutic agents and an aggressive phenotype in TN breast cancer [24]. These results support our observation of low CDKN2A expression in RCB-III (Figure 2).

In both HR-positive and TN breast cancers, VEGF-C was found to be overexpressed in RCB-III compared to RCB-0/I and RCB-II (Figure 2). VEGF-C has been identified as a multifaceted factor involved in the regulation of tumor angiogenesis and lymphangiogenesis [25]. Previous reports have shown that VEGF-C is overexpressed in breast cancer specimens compared to adjacent normal mammary glands, indicating a significant correlation with lymphatic vessel invasion and survival rate [26]. The binding partners of VEGF-C are two tyrosine kinase receptors, VEGFR2 (or KDR/FLK1) and VEGFR3 (or FLT4) [26].

Our results showed an overexpression of VEGF-C in RCB-III compared to RCB-0/I and RCB-II for both cancer subtypes, which is consistent with previous findings of associations between VEGF-C and lymphangiogenesis. Thus, our results indicate that VEGF-C may be an additional therapeutic target; however, monoclonal antibodies of VEGF-C are still relatively poorly studied. Multi-tyrosine kinase inhibitors, including monoclonal antibodies of VEGF-C, may be suitable therapeutic targets, although further studies are needed before practical use is possible. This would confirm the validity of our strategy of selecting candidate genes from deposited data after neoadjuvant chemotherapy [27].

We also found that DNMT1 and DNMT3A were significantly overexpressed in RCB-0/I (Online appendix 3). Among HR-positive breast cancer samples, we observed high expression levels of DNMT1 and DNMT3A in RCB-0/I; we also found more node-positive patients in RCB-0/I (51%) than in RCB-II (36%) and more IHC ER-positive tumors (64%) in RCB-0/I than in RCB-II or RCB-III (Table 1). These results are supported by previous studies that showed that DNMT1 is significantly correlated with lymph node metastasis and that DNMT1 and DNMT3A are correlated with promoter hypermethylation and reduced expression of ER $\alpha$[28-30]. Thus, for non-luminal A-like HR-positive RCB-0/1 tumors, anti-DNMT inhibitors may be suitable therapeutic candidates after neoadjuvant chemotherapy (Table 1). For TN cancers, DNMT1 and DNMT3A expression levels did not differ between RCB subgroups.

This study had several limitations, which could be overcome by performing additional tests. First, the sample sizes for both HR-positive and TN breast cancers were relatively small in each RCB subgroup. Second, we should perform variation testing for another cohort classified by RCB scores to confirm our results. Third, preparing different datasets may increase the reproducibility of our results. Fourth, RCB scores should be calculated after neoadjuvant chemotherapy, regardless of whether chemotherapy is

needed or not. Finally, for some HR-positive cases, receiving chemotherapy might have led to over-treatment with little or no benefit. Despite these limitations, our observations are consistent with the hypothesis that the expression levels of some target genes were higher in RCB-III after standard neoadjuvant therapy; our results also support those of previous clinical and preclinical studies. We have identified some therapeutic targets that may be candidates for additional treatment. Further studies will be needed to examine the efficacy of these additional therapies for tumors with low RCB scores in a clinical setting.

**Conclusion**

We found a relationship between mRNA gene expression pattern and the response of patients to chemotherapy by ER status. We have also identified some candidate targets that may be suitable for therapeutic strategies after standard neoadjuvant chemotherapy. We will explore therapeutic strategies for other situations using the same strategy in future studies.

**Clinical Practice Points**

- There was distinct gene expression pattern depending on degrees of efficacy after neoadjuvant chemotherapy by ER status.

- Distinct therapeutic targets by ER status may be candidate as additional adjuvant therapeutic options except one gene, VEGF-C.

- This study design could be a new strategy for exploring additional adjuvant therapeutic options.

**Acknowledgment**

This study was funded by subsidies for operating expenses at department of breast and endocrine surgery, Okayama university hospital.

**Ethical approval:** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

**Informed consent:** The study used only unidentifiable patient information, and no informed consent was required.

References

1. Kaufmann M, von Minckwitz G, Bear HD, et al. Recommendations from an international expert panel on the use of neoadjuvant (primary) systemic treatment of operable breast cancer: new perspectives 2006. Annals of oncology : official journal of the European Society for Medical Oncology. 2007;18:1927-1934.
2. Patel SA, DeMichele A. Adding adjuvant systemic treatment after neoadjuvant therapy in breast cancer: review of the data. Current oncology reports. 2017;19:56.
3. Bardia A, Baselga J. Neoadjuvant therapy as a platform for drug development and approval in breast cancer: AACR; 2013.
4. Kuerer HM, Newman LA, Smith TL, et al. Clinical course of breast cancer patients with complete pathologic primary tumor and axillary lymph node response to doxorubicin-based neoadjuvant chemotherapy. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 1999;17:460-469.
5. von Minckwitz G, Untch M, Blohmer JU, et al. Definition and impact of pathologic complete response on prognosis after neoadjuvant chemotherapy in various intrinsic breast cancer subtypes. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2012;30:1796-1804.
6. Liedtke C, Mazouni C, Hess KR, et al. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2008;26:1275-1281.
7. Houssami N, Macaskill P, von Minckwitz G, Marinovich ML, Mamounas E. Meta-analysis of the association of breast cancer subtype and pathologic complete response to neoadjuvant chemotherapy. European journal of cancer. 2012;48:3342-3354.
8. Symmans WF, Peintinger F, Hatzis C, et al. Measurement of residual breast cancer burden to predict survival after neoadjuvant chemotherapy. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2007;25:4414-4422.
9. Symmans WF, Wei C, Gould R, et al. Long-Term Prognostic Risk After Neoadjuvant Chemotherapy Associated With Residual Cancer Burden and Breast Cancer Subtype. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2017:JCO2015631010.

10. Masuda N, Lee S-J, Ohtani S, et al. Adjuvant capecitabine for breast cancer after preoperative chemotherapy. New England Journal of Medicine. 2017;376:2147-2159.

11. Hatzis C, Pusztai L, Valero V, et al. A genomic predictor of response and survival following taxane-anthracycline chemotherapy for invasive breast cancer. Jama. 2011;305:1873-1881.

12. Wolff AC, Hammond M, Schwartz JN, et al. College of American Pathologists American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2007;25:118-145.

13. Ali D, Verma A, Mujtaba F, Dwivedi A, Hans RK, Ray RS. UVB-induced apoptosis and DNA damaging potential of chrysene via reactive oxygen species in human keratinocytes. Toxicology letters. 2011;204:199-207.

14. Schulze-Makuch D, Haque S, de Sousa Antonio MR, et al. Microbial life in a liquid asphalt desert. Astrobiology. 2011;11:241-258.

15. Sylvain NJ, Brewster DL, Ali DW. Embryonic ethanol exposure alters synaptic properties at zebrafish neuromuscular junctions. Neurotoxicology and teratology. 2011;33:313-321.

16. Sun J, Wei Q, Zhou Y, Wang J, Liu Q, Xu H. A systematic analysis of FDA-approved anticancer drugs. BMC Syst Biol. 2017;11:87.

17. Iwamoto T, Bianchini G, Booser D, et al. Gene pathways associated with prognosis and chemotherapy sensitivity in molecular subtypes of breast cancer. Journal of the National Cancer Institute. 2011;103:264-272.

18. Uchida N, Suda T, Ishiguro K. Effect of chemotherapy for luminal a breast cancer. Yonago acta medica. 2013;56:51.

19. Symmans WF, Hatzis C, Sotiriou C, et al. Genomic index of sensitivity to endocrine therapy for breast cancer. Journal of clinical oncology. 2010;28:4111.

20. Spring LM, Zangardi ML, Moy B, Bardia A. Clinical Management of Potential Toxicities and Drug Interactions Related to Cyclin-Dependent Kinase 4/6 Inhibitors in Breast Cancer: Practical Considerations and Recommendations. The oncologist. 2017;22:1039-1048.

21. Finn RS, Crown JP, Lang I, et al. The cyclin-dependent kinase 4/6 inhibitor palbociclib in combination with letrozole versus letrozole alone as first-line treatment of oestrogen receptor-positive, HER2-negative, advanced breast cancer (PALOMA-1/TRIO-18): a randomised phase 2 study. The lancet oncology. 2015;16:25-35.

22. Finn R, Jiang Y, Rugo H, et al. Biomarker analyses from the phase 3 PALOMA-2 trial of palbociclib (P) with letrozole (L) compared with placebo (PLB) plus L in postmenopausal women with ER+/HER2–advanced breast cancer (ABC). Annals of Oncology. 2016;27.

23. Fang H, Huang D, Yang F, Guan X. Potential biomarkers of CDK4/6 inhibitors in hormone receptor-positive advanced breast cancer. Breast cancer research and treatment. 2018;168:287-297.

24. Arima Y, Hayashi N, Hayashi H, et al. Loss of p16 expression is associated with the stem cell characteristics of surface markers and therapeutic resistance in estrogen receptor‐negative breast cancer. International journal of cancer. 2012;130:2568-2579.

25. Chen JC, Chang YW, Hong CC, Yu YH, Su JL. The role of the VEGF-C/VEGFRs axis in tumor progression and therapy. Int J Mol Sci. 2012;14:88-107.

26. Kinoshita J, Kitamura K, Kabashima A, Saeki H, Tanaka S, Sugimachi K. Clinical significance of vascular endothelial growth factor-C (VEGF-C) in breast cancer. Breast cancer research and treatment. 2001;66:159-164.

27. Chen H, Li J, Cui T, Hu L. Adjuvant gonadotropin-releasing hormone analogues for the prevention of chemotherapy induced premature ovarian failure in premenopausal women. The Cochrane database of systematic reviews. 2011:CD008018.

28. Girault I, Tozlu S, Lidereau R, Bièche I. Expression analysis of DNA methyltransferases 1, 3A, and 3B in sporadic breast carcinomas. Clinical cancer research. 2003;9:4415-4422.

29. Gonzalez-Perez A, Jene-Sanz A, Lopez-Bigas N. The mutational landscape of chromatin regulatory factors across 4,623 tumor samples. Genome biology. 2013;14:r106.

30. Yu Z, Xiao Q, Zhao L, et al. DNA methyltransferase 1/3a overexpression in sporadic breast cancer is associated with reduced expression of estrogen receptor-alpha/breast cancer susceptibility gene 1 and poor prognosis. Molecular carcinogenesis. 2015;54:707-719.

**Figure legends**

Figure 1. mRNA expression level of known subtype specific prognostic biomarker according to residual cancer burden (RCB) score in ER and/or PR positive and triple negative breast cancer.

Box plots indicate 1st and 3rd quartiles and the bold line within the box represents the median value of log2-normalized mRNA expression levels. The outliers are defined by the R statistical package as data points that fall outside the 1st and 3rd quartiles by more than 1.5 times the interquartile range, and circles falling outside the box represent outliers. The p values were calculated from the Kruskal-Wallis nonparametric test or the Wilcox test comparing the differences between all three or two RCB subgroups, based on RCB scores. Boxplots of a; ESR1 (Probe 205225_at), b; PR (Probe 208305_at), c; ERBB2 (Probe 216836_s_at), d; MKi67 (Probe 212021_s_at).

Figure 2. mRNA expression levels of CDK4/6/cyclin D1 complex-related genes according to RCB scores in ER- and/or PR-positive and triple negative breast cancer.

The median ± S.D. values of log2-normalized mRNA expression levels are plotted on the y-axis, as described in Fig. 1 legend. The p values were calculated from the Kruskal-Wallis nonparametric test or the Wilcox test comparing the differences between all three or two RCB subgroups based on the RCB scores. Boxplots of a; CCND1 (cyclin D1) (Probe 208712_at), b; CDKN2A (Probe 209644_x_at).

Figure 3. mRNA expression levels of genes related to VEGF-C (Probe 209946_at) according to RCB scores in ER- and/or PR-positive and triple negative breast cancer. The median ± S.D. values of log2-normalized mRNA expression levels are plotted on the y-axis, as described in the legend for Fig. 1. The p values were calculated from the Kruskal-Wallis nonparametric test or the Wilcox test comparing the differences between all three or two RCB subgroups based on RCB scores.

Table1. Clinical and Pathologic Characteristics of all Patients. Fisher's exact test was used for between-group comparisons for this data.

| | Hormone receptor positive (N = 253) | | | | | | | | | Triple negative (N = 146) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| RCB Index | 0/I | | II | | III | | Total | | | 0/I | | II | | III | | Total | | |
| | Number | % | Number | % | Number | % | Number | % | *p* value | Number | % | Number | % | Number | % | Number | % | *p* value |
| | 53 | 21 | 135 | 53 | 65 | 26 | 253 | 63 | | 60 | 41 | 46 | 32 | 40 | 27 | 146 | 37 | |
| Age,years | | | | | | | | | | | | | | | | | | |
| <50 | 25 | 47 | 71 | 53 | 35 | 54 | 131 | 52 | 0.76 | 36 | 60 | 27 | 59 | 18 | 45 | 81 | 55 | 0.31 |
| ≥50 | 28 | 53 | 64 | 47 | 30 | 46 | 122 | 48 | | 24 | 40 | 19 | 41 | 22 | 55 | 65 | 45 | |
| median | 51 | | 50 | | 49 | | 50 | | | 48 | | 47.5 | | 51 | | 49 | | |
| clinical Tstage | | | | | | | | | | | | | | | | | | |
| cT0 | 1 | 21 | 2 | 1 | 0 | 0 | 3 | 1 | 0.15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.013 |
| cT1 | 4 | 8 | 10 | 7 | 2 | 3 | 16 | 6 | | 6 | 10 | 1 | 2 | 0 | 0 | 7 | 5 | |
| cT2 | 29 | 55 | 75 | 56 | 37 | 57 | 141 | 56 | | 36 | 60 | 21 | 46 | 15 | 38 | 72 | 49 | |
| cT3 | 13 | 25 | 30 | 22 | 8 | 12 | 51 | 2 | | 11 | 18 | 16 | 35 | 10 | 25 | 37 | 25 | |
| cT4 | 6 | 11 | 18 | 13 | 18 | 28 | 42 | 17 | | 7 | 12 | 8 | 17 | 15 | 38 | 30 | 21 | |
| Lymph node status | | | | | | | | | | | | | | | | | | |
| Positive | 27 | 51 | 53 | 39 | 12 | 18 | 92 | 36 | < 0.001 | 43 | 72 | 33 | 72 | 38 | 95 | 114 | 78 | 0.0053 |
| Negative | 26 | 49 | 82 | 61 | 53 | 82 | 161 | 64 | | 17 | 28 | 13 | 28 | 2 | 5 | 32 | 22 | |
| clinical AJCC Stage | | | | | | | | | | | | | | | | | | |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.045 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0029 |
| I | 2 | 47 | 2 | 1 | 0 | 0 | 4 | 2 | | 3 | 5 | 0 | 0 | 0 | 0 | 3 | 2 | |
| II | 32 | 6 | 86 | 64 | 30 | 46 | 148 | 58 | | 33 | 55 | 23 | 50 | 10 | 25 | 66 | 45 | |
| III | 19 | 36 | 47 | 35 | 35 | 54 | 101 | 40 | | 24 | 40 | 23 | 50 | 30 | 75 | 77 | 53 | |
| Grade | | | | | | | | | | | | | | | | | | |
| 1 | 2 | 4 | 16 | 12 | 8 | 12 | 26 | 10 | 0.0049 | 0 | 0 | 1 | 2 | 0 | 0 | 1 | 1 | 0.14 |
| 2 | 17 | 32 | 75 | 56 | 36 | 55 | 128 | 51 | | 4 | 7 | 8 | 17 | 6 | 15 | 18 | 12 | |
| 3 | 30 | 57 | 40 | 30 | 19 | 29 | 89 | 35 | | 52 | 87 | 35 | 76 | 28 | 70 | 115 | 79 | |
| NA | 4 | 8 | 4 | 3 | 2 | 3 | 10 | 4 | | 4 | 7 | 2 | 4 | 6 | 15 | 12 | 8 | |
| ER IHC status | | | | | | | | | | | | | | | | | | |
| Positive | 46 | 87 | 132 | 98 | 63 | 97 | 241 | 95 | 0.0068 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| Negative | 7 | 13 | 3 | 2 | 2 | 3 | 12 | 5 | | 60 | 1 | 46 | 1 | 40 | 1 | 146 | 1 | |
| PR IHC status | | | | | | | | | | | | | | | | | | |
| Positive | 34 | 64 | 106 | 79 | 50 | 77 | 190 | 75 | 0.12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| Negative | 19 | 36 | 29 | 21 | 15 | 23 | 63 | 25 | | 60 | 1 | 46 | 1 | 40 | 1 | 146 | 1 | |
| PAM50 | | | | | | | | | | | | | | | | | | |
| LuminalA | 12 | 23 | 74 | 55 | 41 | 63 | 127 | 50 | < 0.001 | 0 | 0 | 2 | 4 | 0 | 0 | 2 | 1 | 0.32 |
| LuminalB | 13 | 25 | 29 | 21 | 14 | 22 | 56 | 22 | | 0 | 0 | 2 | 4 | 1 | 3 | 3 | 2 | |
| HER2 | 5 | 9 | 10 | 7 | 4 | 6 | 19 | 8 | | 4 | 7 | 3 | 7 | 5 | 13 | 12 | 8 | |
| Basal | 17 | 32 | 11 | 8 | 4 | 6 | 32 | 13 | | 50 | 83 | 32 | 70 | 31 | 78 | 113 | 77 | |
| Normal | 6 | 11 | 11 | 8 | 2 | 3 | 19 | 8 | | 6 | 10 | 7 | 15 | 3 | 8 | 16 | 11 | |