A multifactorial ‘Consensus Signature’ by in silico analysis to predict response to neoadjuvant anthracycline-based chemotherapy in triple-negative breast cancer

Natalie Turner1,4, Mattia Forcato2,4, Simona Nuzzo2, Luca Malorni1,3, Silvio Bicciato2,5 and Angelo Di Leo1,5

BACKGROUND: Owing to the complex processes required for anthracycline-induced cytotoxicity, a prospectively defined multifactorial Consensus Signature (ConSig) might improve prediction of anthracycline response in triple-negative breast cancer (TNBC) patients, whose only standard systemic treatment option is chemotherapy.

AIMS: We aimed to construct and evaluate a multifactorial signature, comprising measures of each function required for anthracycline sensitivity in TNBC.

METHODS: ConSigs were constructed based on five steps required for anthracycline function: drug penetration, nuclear topoisomerase IIα (topolla) protein location, increased topolla messenger RNA (mRNA) expression, apoptosis induction, and immune activation measured by, respectively, HIF1α or SHARP1 signature, LAPTM4B mRNA, topolla mRNA, Minimal Gene signature or YWHAZ mRNA, and STAT1 signature. TNBC patients treated with neoadjuvant anthracycline-based chemotherapy without taxane were identified from publicly available gene expression data derived with Affymetrix HG-U133 arrays (training set). In silico analyses of correlation between gene expression data and pathological complete response (pCR) were performed using receiver-operating characteristic curves. To determine anthracycline specificity, ConSigs were assessed in patients treated with anthracycline plus taxane. Specificity, sensitivity, positive and negative predictive value, and odds ratio (OR) were calculated for ConSigs. Analyses were repeated in two validation gene expression data sets derived using different microarray platforms.

RESULTS: In the training set, 29 of 147 patients had pCR after anthracycline-based chemotherapy. Various combinations of components were evaluated, with the most powerful anthracycline response predictors being ConSig1: (STAT1+topolla mRNA +LAPTM4B) and ConSig2: (STAT1+topolla mRNA+HIF1α). ConSig1 demonstrated high negative predictive value (85%) and high OR for no pCR (3.18) and outperformed ConSig2 in validation sets for anthracycline specificity.

CONCLUSIONS: With further validation, ConSig1 may help refine selection of TNBC patients for anthracycline chemotherapy.

INTRODUCTION

Preferred breast cancer (neo)adjuvant chemotherapy regimens are generally anthracycline based, given the improved outcomes compared with cyclophosphamide/methotrexate/fluorouracil.1 However, across all anthracycline-treated patients, only a small percentage actually receives benefit while these agents are associated with significant toxicities. Breast cancer is well recognized as a heterogeneous disease and therefore treating all breast cancers with the same chemotherapeutic agents could be considered illogical. Of considerable use would be a predictive marker of response to distinguish patients likely to receive benefit from those who are not, sparing predicted ‘poor responders’ from associated toxicities. Unfortunately, suitable predictive biomarkers for chemotherapeutic agents have remained elusive to date.

Topoisomerase IIα gene (TOP2A) is a putative marker of anthracycline sensitivity, with its gene product being the direct target of anthracyclines. TOP2A amplification has been shown to predict increased sensitivity to anthracyclines in several studies,2-7 although this finding has not been entirely consistent.8,9 Indeed, a single biomarker may not be sufficient to predict anthracycline response, and a multifactorial approach using gene signatures might be required.7,10 In the TOP trial7 of 149 patients with estrogen receptor (ER)-negative early or locally advanced breast cancer treated with a single agent epirubicin, TOP2A amplification was significantly associated with pathological complete response (pCR). In addition, a multifactorial anthracycline sensitivity score (the A-score), comprised of three gene signatures, was evaluated, demonstrating a very high negative predictive value (NPV), although a much lower positive predictive value (PPV).7

Gene signatures are often defined through retrospective analyses of tumor tissue gene expression patterns correlated with patient outcomes. Although concordance in outcome prediction between signatures has been demonstrated,11 signatures derived in this manner often contain few genes in common as well as numerous genes of unknown function, making their clinical relevance less certain. It might be possible to improve clinical relevance of a signature by identifying the molecular processes required for a specific cellular function, such as anthracycline-
induced cytotoxicity, and construct a signature containing measures of each of these functions. On the basis of this hypothesis, we aimed to construct and evaluate a multifactorial Consensus Signature for predicting anthracycline sensitivity in triple-negative breast cancer (TNBC). The term Consensus Signature was chosen to reflect the concept that, by each selected component acting as a surrogate marker of the different steps required for anthracycline cytotoxicity, included components (the genes and gene signatures) would work synergistically to provide an overall measure of effective anthracycline function.

We focused specifically on TNBC for the following reasons. First, there has been substantial work performed already evaluating the predictive role of TOP2A in HER2+ breast cancer, due to the known relationship between TOP2A amplification and HER2 amplification. Anthracyclines are commonly used in TNBC and appear to have activity. We wanted to assess this without the confounding factor of HER2 overexpression. Second, with treatment options in TNBC limited to chemotherapy, more effective use of chemotherapy would be of considerable benefit. Finally, with growing understanding of breast cancer biological diversity, evaluation of a predictive biomarker within a specific subtype might be preferable, as positive results could otherwise be masked if evaluated across a heterogeneous combined cohort.

MATERIALS AND METHODS

Data set

The construction and evaluation of consensus signatures (ConSigs) were carried out using a retrospective cohort study design, with in silico analyses of previously collected genetic data, clinical characteristics, and responses. The data set comprised gene expression profiles of patients who had received neoadjuvant anthracycline-based chemotherapy without a taxane. As the ConSigs were designed to be specific for anthracyclines, taxane use was considered a confounding factor.

Data derived from Affymetrix (Santa Clara, CA, USA) gene expression arrays based on build 133 of UniGene database (HG-U133) were combined and evaluated as a single group, designated the ‘breast compendium’ (details on the construction and composition of the breast compendium are reported in Supplementary Table 1). The subset of samples treated with anthracycline-based neoadjuvant chemotherapy was used as training set to derive the ConSigs, whereas the cohort of patients treated with anthracycline plus taxane-based neoadjuvant chemotherapy served as a control group to assess ConSig specificity for anthracyclines. In addition, two cohorts of patients treated with anthracycline-based neoadjuvant chemotherapy with gene expression data derived from different microarray platforms served as validation sets. The European Organisation for Research and Treatment of Cancer (EORTC)/BIG00-01 data set included gene expression data obtained with the Affymetrix X3P array from patients with locally advanced, inflammatory, or large operable breast cancers treated with either fluorouracil/epirubicin/cyclophosphamide or docetaxel followed by docetaxel/epirubicin under the auspices of the EORTC 10994 trial. The data set was available in the Gene Expression Omnibus repository under accession number GSE14138, the Netherlands Cancer Institute (NKI) data set (accessible under accession number GSE43138) used the Illumina (San Diego, CA, USA) HumanWG 6 v3.0 expression beadchip for gene expression profiling and included patients with intermediate or high-risk ER-breast cancer treated with neoadjuvant dose-dense doxorubicin/cyclophosphamide. For this study, only TNBC patients in the training and control subsets of the compendium and in the validation sets were considered. Further details on the construction of the breast compendium and on the definition of TNBCs are reported in the Supplementary Methods.

Design of the consensus signature

In order for anthracyclines to be effective, we postulated that the following steps must occur: (1) penetration of the drug into the tumor bed, (2) location of the target (topoIIα protein) within the nucleus, (3) increased topoIIα messenger RNA (mRNA) expression above that related to proliferation alone, (4) induction of apoptosis, and (5) active immune/stromal function.

For each step, a representative gene or gene signature was selected, with gene signatures chosen over single genes where possible. In some cases, more than one potential genes/gene signatures was evaluated for each step, to compare their relative utility. Details of the marker(s) evaluated, their association with anthracycline response, and rationale for their selection are listed in Table 1.

Quantification of genes and gene signatures

SHARP1 signature, HIF1α hypoxia signature (HIF), and the Minimal Gene signature were quantified as previously described. Briefly, each signature was calculated by summarizing the standardized expression levels of the genes in the signature into a combined score with zero mean. AURKA, STAT1, and PLAU signatures were computed as previously described using geneR package. Briefly, for each signature, the expression was divided by the expression of each gene i included in the set of genes of interest and ωi is either +1 or −1 depending on the sign of the association under study. Gene expression lists for each signature are included in Supplementary Table 2.

LAPTM4B, AURKA, VWA4A, and topoIIα mRNA expression levels were calculated using the corresponding probe sets or the median expression if multiple probe sets were available for each gene. For the NKI data set, probes were filtered on the basis of their quality, keeping only probes classified as ‘perfect’ and ‘good’ in the Bioconductor Illumina Human v3.db annotation package.

Quantification of the Consensus Signature

The Consensus Signature score was calculated in a continuous form, consisting of a linear combination of the various components/signatures. Prior to combination, each component/signature was scaled to have the interquartile range equal to 1 and the median equal to 0. On the basis of the association with pCR (Table 1), we hypothesized that a high Consensus Signature score would predict for increased pCR rate, whereas a low score should predict anthracycline resistance.

Statistical analysis

All statistical analyses were performed in R version 2.15.1. Odds ratios (ORs) were used to compare pCR rates between groups defined by different clinical and molecular characteristics (stats R package). The area under the curve (AUC) was used to assess the prediction performance of any signature score (ROCR R package). AUC was estimated through the concordance index (survcomp R package) under the alternative hypothesis that AUC was greater than 0.5, as each signature score was designed to have positive AUC. Its significance and confidence interval were estimated assuming a symmetric normality. Because of the differences in array design and technology between the training and the two validation sets, the threshold ConSig score for each cohort was calculated using the score value that corresponded to the 75th percentile of the score distribution for that cohort. P values of < 0.05 were considered significant.

RESULTS

Data set

The training set was derived from a breast cancer data set originally consisting of 4,600 samples collected in 27 different studies. After exclusion of duplicate samples (n = 939) and adjusting for batch effect using ComBat, 1,069 samples (29%) were classified as TNBC by the SCM0D2 subtype clustering classifier (subtype clustering model) contained in geneR package. Among these samples, 491 had information about neoadjuvant chemotherapy. Eleven samples were from patients treated with taxane but not anthracycline and were excluded, whereas 147 and 333 samples were from patients treated with anthracycline-based therapy without taxane and anthracycline-based therapy with taxane, respectively (Figure 1).

Clinical and tumor characteristics for the patients in the training set (n = 147) treated with anthracycline-based chemotherapy are listed in Table 2. The samples were originally contained in four different data sets, and included 29 (19.7%) pCRs and 118 (80.3%) samples with residual disease. pCR was defined as ypT0/is, ypN0 in all included studies.
Clinical characteristics
All clinical variables were tested for their ability to predict pCR, with no significant association between any clinical characteristics and pCR found (data not shown).

Predictive power of single gene or gene signatures
Using receiver-operating characteristic curves, the ability of any single component (gene or gene signature) to discriminate patients with pCR from patients with residual disease in the training set was assessed. STAT1 was significantly associated with pCR status (Supplementary Table 3; Supplementary Figure 1a). All other components, when considered individually, were not significantly correlated with pCR. Topolla mRNA corrected for proliferation with either AURKA mRNA or AURKA signature showed no increased correlation with pCR compared with topolla mRNA alone (data not shown).

Predictive power of ConSigs
ConSigs was constructed using various combinations of components, with the starting point being components shown to have significant or near-significant predictive capability when
used alone, that is, STAT1, topoIIα, HIF, and LAPTM4B. Using a continuous score to quantify ConSig expression level, all combinations of core components demonstrated a significant correlation with pCR in patients in the training set treated with anthracycline-based chemotherapy without taxane. The two most predictive combinations were designated ConSig1: (STAT1 + topoIIα + LAPTM4B) with AUC 0.70 \( (P = 3.9 \times 10^{-5}) \), and ConSig2: (STAT1 + topoIIα + HIF) with AUC 0.71, \( P = 4.2 \times 10^{-6} \). High correlation with pCR was maintained with the addition of further component genes/gene signatures to either ConSig1 or ConSig2, but overall predictive power was not better than with three components (Table 3). The combination of STAT1 + PLAU, the components for TNBC of another multifactorial scoring signature, the A-score, was correlated with pCR, although less strongly than other combinations. Substituting topoIIα mRNA with topoIIα mRNA corrected for proliferation did not improve the performance of any of the ConSigs.

To assess specificity of ConSigs for anthracycline response compared with other chemotherapy regimens, we analyzed their respective performances in a control group of patients who received taxane in addition to anthracycline (n = 333), 299 of whom had information about response and 101 with pCR. For the ConSigs with the highest predictive power in the training set, i.e., ConSig1 and ConSig2, neither was correlated with pCR in this control group (Supplementary Figure 1c and d). Although (STAT1 + PLAU) had predicted response to anthracycline-based chemotherapy, it did not appear to be anthracycline specific, performing similarly in anthracycline+taxane-treated patients (Table 3).

**Classification performance of ConSigs**

A threshold score that could be used to classify a patient in the training set as a putative responder or as resistant was determined for each ConSig by selecting the score value corresponding to the 75th percentile of the score distribution. PPV and NPVs, specificity, and OR were then calculated. For ConSig1 NPV was high (85%) as was OR for lack of pCR \( (OR = 3.18, P = 0.008) \) (Table 4). PPV, however, was modest \( (PPV = 35\%) \). ConSig2 performed similarly, with high NPV and OR, but modest PPV. In the control group of 299 patients treated with anthracycline plus taxane, NPVs for ConSig1 and ConSig2 were lower, at 66 and 67\%, respectively, and ORs for lack of pCR were no longer statistically significant for either ConSig (Table 4), further supporting the specificity of these ConSigs for anthracycline response.

**Evaluation of ConSig1 and ConSig2 in two independent ‘validation’ data sets**

Two data sets, NKI and EORTC/BIG00-01, were selected as validation sets. The EORTC/BIG00-01 data set comprised 161 samples, 85 of which were TNBC, with 46 patients treated with anthracycline without taxane (18 pCRs) and 39 treated with anthracycline plus taxane (18 pCRs). The NKI data set included 178 ER-negative breast cancer samples, 52 of which were TNBC. Of these 52 patients, 24 had pCR. All patients received anthracycline-based neoadjuvant chemotherapy without taxane.

In the NKI data set, both ConSig1 and ConSig2 were significantly correlated with pCR for patients receiving anthra-
cyclophosphamide without taxane (Table 5). In the EORTC/BIG00-01 data set, 
ConSig1 remained specific for anthracyclines in the data set, whereas 
ConSig2 was correlated with pCR for both anthracycline-based and anthracycline+taxane-based neoadjuvant chemotherapy.

**Table 2.** Clinical and tumor characteristics of TNBC patients in the training set treated with anthracycline-based neoadjuvant chemotherapy without taxane (n = 147)

| Characteristic               | No. of patients (%) |
|------------------------------|---------------------|
| Age (years)                  |                     |
| < 40                         | 12 (8)              |
| 40–60                        | 40 (27)             |
| > 60                         | 7 (5)               |
| NA                           | 88 (60)             |
| Tumor size                   |                     |
| T1                           | 15 (10)             |
| T2                           | 90 (61)             |
| T3                           | 22 (15)             |
| T4                           | 19 (13)             |
| NA                           | 1 (1)               |
| Nodal status                 |                     |
| N0                           | 14 (9)              |
| N1                           | 13 (9)              |
| N2                           | 6 (4)               |
| NA                           | 114 (78)            |
| Histologic grade             |                     |
| G1                           | 2 (1)               |
| G2                           | 25 (17)             |
| G3                           | 112 (76)            |
| NA                           | 8 (6)               |
| pCR                          |                     |
| Yes                          | 29 (20)             |
| No                           | 118 (80)            |

Abbreviations: NA, not available; pCR, pathological complete response; TNBC, triple-negative breast cancer.

**Table 3.** Correlation with pCR for various combinations of ConSig components in the training set

| ConSig combination | Type of NAC | AUC | 95% CI | OR | 95% CI | P value |
|--------------------|-------------|-----|--------|----|--------|---------|
| STAT1+PLAU         | A           | 0.63| 0.52–0.73 | 3.11 | 1.2×10^-2 | 0.001 |
|                    | A+T         | 0.56| 0.50–0.63 | 7.6×10^-4 | 0.0001 | 0.53 |
| STAT1+topoll       | A           | 0.68| 0.57–0.78 | 3.9×10^-5 | 0.0025 | 0.43 |
|                    | A+T         | 0.60| 0.54–0.67 | 1.2×10^-9 | 0.0001 | 0.7 |
| STAT1+topoll+LAPTM4B (ConSig1) | A | 0.70| 0.60–0.80 | 6.2×10^-4 | 0.005 | 0.8 |
|                     | A+T         | 0.55| 0.48–0.61 | 1.2×10^-7 | 0.0001 | 0.42 |
| STAT1+topoll+HIF1 (ConSig2) | A | 0.71| 0.62–0.80 | 4.2×10^-6 | 0.0001 | 0.53 |
|                     | A+T         | 0.52| 0.45–0.58 | 6.8×10^-4 | 0.0001 | 0.8 |
| STAT1+topoll+YWHAZ  | A           | 0.66| 0.56–0.76 | 8.0×10^-4 | 0.0001 | 0.8 |
|                     | A+T         | 0.54| 0.48–0.61 | 1.3×10^-2 | 0.0001 | 0.8 |
| STAT1+topoll+LAPTM4B+YWHAZ | A | 0.66| 0.56–0.76 | 8.0×10^-4 | 0.0001 | 0.8 |
|                     | A+T         | 0.51| 0.45–0.58 | 1.3×10^-2 | 0.0001 | 0.8 |
| STAT1+topoll+LAPTM4B+MS+PLAU | A | 0.56| 0.49–0.62 | 4.2×10^-2 | 0.0001 | 0.8 |
|                     | A+T         | 0.56| 0.49–0.62 | 4.2×10^-2 | 0.0001 | 0.8 |

Abbreviations: AUC, area under the ROC curve; CI, confidence interval; ConSig, Consensus Signature; HIF, hypoxia-inducible factor; LAPTM4B, lysosomal-associated protein transmembrane 48 gene; MS, minimal gene signature; NAC, neoadjuvant chemotherapy; pCR, pathological complete response; ROC, receiver-operating characteristic.

**DISCUSSION**

Overall, the best performing combination of components was ConSig1 (STAT1+topollxLAPTM4B). Our results suggest that ConSig1 has excellent ability to predict anthracycline resistance within a cohort of anthracycline-treated TNBC patients. This is clinically relevant, as, if further validated, ConSig1 could be used to identify TNBC patients for whom the addition of anthracycline is likely to add toxicity without benefit, and thus for whom an alternate chemotherapy regimen might be selected. When evaluated in TNBC patients who received anthracycline and taxane, the predictive ability of ConSig1 was lost. Although not conclusive, a lack of predictive utility in patients who also received taxanes supports the anthracycline specificity of ConSig1. ConSig2 was most strongly correlated with pCR in the training set; however, in the validation set ConSig2 did not show discrimination in performance between patients treated with anthracycline-based and anthracycline plus taxane-based neoadjuvant chemotherapy.

Although we considered that five main processes should take place for anthracyclines to cause cell death, markers for two of these processes, that is, induction of apoptosis and hypoxia/drug penetration, did not appear to contribute significantly to the predictive power of ConSig1. Given that immune function (STAT1) was a powerful contributor to ConSig1 and ConSig2, we postulated that in the setting of a highly active immune response, intact apoptotic pathways as measured by Minimal Gene signature15 or

**Table 4.** Performance of ConSig1 and ConSig2 in the training set for predicting pathological complete response

| ConSig1 | NAC | NPV (%) | PPV (%) | OR | 95% CI | P value |
|---------|-----|---------|---------|----|--------|---------|
| A       | 85  | 32      | 6.20    | 4.2×10^-4 | 0.0001 | 0.8 |
| A+T     | 66  | 37      | 7.54    | 8.2×10^-3 | 0.0001 | 0.8 |

Abbreviations: A, anthracycline-based chemotherapy; A+T, anthracycline +taxane chemotherapy; CI, confidence interval; ConSig, Consensus Signature; NAC, neoadjuvant chemotherapy; NPV, negative predictive value; OR, odds ratio for lack of pathological complete response; PPV, positive predictive value.
Table 5. Performance of ConSig1 and ConSig2 in the validation sets for predicting pathological complete response

| Validation set | ConSig | NAC | AUC | 95% CI | \(P\) value | \(NPV\) (%) | \(PPV\) (%) | OR | 95% CI | \(P\) value |
|----------------|--------|-----|-----|--------|-------------|------------|------------|----|--------|-------------|
| EORTC/BIG00-01 | ConSig1 | A   | 0.70| 0.55–0.85| \(5.2 \times 10^{-3}\) | 74          | 75         | 8.33| 2.00–44.44| 6.0 \times 10^{-3}\ |
|                | ConSig2 | A+T | 0.60| 0.43–0.77| 0.12        | 62          | 70         | 3.81| 0.87–20.75| 0.09        |
|                |        | A+T | 0.65| 0.50–0.80| \(2.4 \times 10^{-2}\) | 62          | 42         | 1.15| 0.29–4.40 | 0.83        |
| NKI            | ConSig1 | A   | 0.74| 0.61–0.87| \(1.3 \times 10^{-4}\) | 64          | 77         | 5.95| 1.53–29.99| 1.6 \times 10^{-2}\ |
|                | ConSig2 | A   | 0.64| 0.49–0.78| \(3.1 \times 10^{-2}\) | 59          | 62         | 2.30| 0.65–8.86 | 0.20        |

Abbreviations: A, anthracycline-based chemotherapy; A+T, anthracycline+taxane chemotherapy; AUC, area under the ROC curve; CI, confidence interval; ConSig, Consensus Signature; EORTC, European Organisation for Research and Treatment of Cancer; NAC, neoadjuvant chemotherapy; NKI, Netherlands Cancer Institute; NPV, negative predictive value; OR, odds ratio for lack of pathological complete response; PPV, positive predictive value; ROC, receiver-operating characteristic.

YWHAZ\(^{23}\) might be less important. However, predictive power was not improved using either Minimal Gene signature or YWHAZ in absence of the immune/stromal components (STAT1 and PLAU) (data not shown). Although the inclusion of a component marker of hypoxia had the highest correlation with pCR in the training set (ConSig2), this combination did not show consistent anthracycline specificity, with similar performance in anthracycline- and anthracycline plus taxane-treated patients in the EORTC validation set. This might be because hypoxia is a critical factor in the function of other chemotherapy agents, not only anthracyclines. Interestingly, although topoIIα protein expression is known to relate to proliferation,\(^{26}\) correcting topoIIα mRNA expression level for proliferation made no difference to the predictive utility of the ConSigs. The lack of discriminating ability of the proliferative markers, AURKA or AURKA signature, might relate to the fact that the majority of tumors in our data set (76%) were grade 3 with most of the rest being grade 2, and thus were all moderately to highly proliferating. With no comparator group of low-proliferating tumors, the influence of a proliferation marker cannot easily be evaluated. Furthermore, when considering a cohort of TNBC, proliferation is characteristically high, making the ability to measure variability of topoIIα protein expression due to low versus high proliferation arguably less critical. Although the influence of proliferation itself in the results is uncertain, the specificity of ConSig1 for anthracycline response over and above that of proliferation is supported by the fact that it is not predictive in patients treated with anthracyclines plus taxanes, a situation where proliferation is still important.

Differently from the A-score, which evaluated ER-/HER2\(^{+}\) and ER-/HER2\(^{-}\) tumors, we focused specifically on TNBC, a group where treatment is limited to chemotherapy, and thus where optimization of chemotherapy regimen would be of considerable utility. Interestingly, the predictive ability of ConSig1 for TNBC patients treated with anthracycline-based chemotherapy without taxane appeared to be better than the combination of STAT1 +PLAU, the component gene signatures in the A-score for TNBC.

By constructing a biologically relevant multifactorial ConSig, we hypothesized that this should predict anthracycline response, as well as resistance; however, PPV of all ConSigs in the training set were lower than anticipated, and only around 35% for ConSig1. Conversely, ConSig1 predicts likelihood of lack of response (NPV), which is still a finding of considerable clinical utility if further validated. In general, high PPV for biomarkers has previously been shown to be hard to achieve. Considering two well-established biomarkers and the only two routinely used in breast cancer management, ER and HER2, positivity predicts treatment response in only around 50% of patients for endocrine therapy\(^{25}\) and less than 40% for trastuzumab as a single agent.\(^{26}\) However, it is worth noting that the small number of pCRs in the training set (29/147; 19.7%) might have contributed to less than stable results for PPV. Indeed, in the validation sets, where pCR rates are 39% (EORTC/BIG00-01) and 46% (NKI), PPV is higher than 75% (Table 5).

Finally, our sample size was relatively limited. The training set used contained all publicly available data on the same platform for TNBC treated with anthracycline-based chemotherapy. However, this incorporated only 147 TNBC patients. Thus, further validation of ConSig1 in independent TNBC cohorts is necessary.

In conclusion, this project demonstrated the feasibility of defining a multifactorial Consensus Signature for predicting anthracycline response and, when applied in a TNBC patient cohort, ConSig1 was highly predictive for anthracycline resistance. With further validation, this signature may provide clinicians with a useful tool for improved selection of TNBC patients for anthracyclines, potentially leading to better treatment tolerance and more effective therapy.

ACKNOWLEDGMENTS

We acknowledge the generous support of the Italian Association for Cancer Research (AIRC) that provided funding for this project under the auspices of the AIRC Special Program Molecular Clinical Oncology ‘5 per mille’ research initiative.

CONTRIBUTIONS

All the authors contributed to the conception and design of the manuscript and to the data analysis and interpretation. NT, MF, and SN contributed also to the collection and assembly of the data. All authors have written and approved the manuscript. ADL and SB are the manuscript’s guarantors.

COMPETING INTERESTS

The authors declare no conflict of interest.

FUNDING

The authors declare that no funding was received.

REFERENCES

1 Patel R, Davies C, Godwin J, Gray R, Pan HC, Clarke M et al. Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100,000 women in 123 randomised trials. Lancet 2012; 379: 432–444.

2 Di Leo A, Desmedt C, Bartlett JM, Piette F, Ejlersen B, Pritchard KI et al. HER2 and TOP2A as predictive markers for anthracycline-containing chemotherapy regimens as adjuvant treatment of breast cancer: a meta-analysis of individual patient data. Lancet Oncol 2011; 12: 1134–1142.

3 Di Leo A, Gancberg D, Larsimont D, Tanner M, Järvinen T, Rousaud G et al. HER-2 amplification and topoisomerase IIA gene aberrations as predictive markers in node-positive breast cancer patients randomly treated either with an anthracycline-based therapy or with cyclophosphamide, methotrexate, and 5-fluorouracil. Clin Cancer Res 2002; 8: 1107–1116.

4 Press MF, Sauter G, Buyse M, Bernstein L, Guzman R, Santiago A et al. Alteration of topoisomerase II-alpha gene in human breast cancer: association with responsiveness to anthracycline-based chemotherapy. J Clin Oncol 2011; 29: 859–867.
11. Fan C, Oh DS, Wessels L, Weigelt B, Nuyten DS, Nobel AB et al. Predictive molecular markers of anthracycline effectiveness in early breast cancer. *Euro J Clin Cancer* 2011; **Supp 9**: 16–21.

12. Farmer P, Bonnefoi H, Anderle P, Cameron D, Wirapati P, Becette V et al. A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer. *Nat Med* 2009; **15**: 68–74.

13. deRonde JJ, Lips EH, Mulder L, Vincent AD, Wesseling J, Nieuwland M et al. Topoisomerase II alpha amplification may predict benefit from adjuvant anthracyclines in HER2 positive early breast cancer. *Breast Cancer Res Treat* 2007; **106**: 181–189.

14. Montagner M, Eiermann W, Robert N, Pienkowski T, Martin M, Press M et al. Adjuvant trastuzumab in HER2-positive breast cancer. *N Eng J Med* 2011; **365**: 1273–1283.

15. Arriola E, Rodriguez-Pinilla SM, Lambros MB, Jones RL, James M, Savage K et al. Anthracyclines in HER2-positive early breast cancer. *Breast Cancer Res Treat* 2008; **106**: 5027–5035.

16. Martin M, Romero A, Cheang MC, López García-Asenjo JA, García-Saenz JA, Oliva B et al. Genomic predictors of response to doxorubicin versus docetaxel in primary breast cancer. *Breast Cancer Res Treat* 2011; **128**: 127–136.

17. di Leo A, Moretti E, Oakman C, Biganzoli L, Santarpia L. Predictive molecular markers of anthracycline effectiveness in early breast cancer. *Breast Cancer Res Treat* 2015; **128**: 163–173.

18. Nahta R, Esteva FJ. HER2 therapy: molecular mechanisms of trastuzumab resistance. *Breast Cancer Res Treat* 2011; **128**: 213–223.

19. Montagner M, Enzo E, Forcato M, Zanconato F, Parenti A, Rampazzo E et al. SHARP1 suppresses breast cancer metastasis by promoting degradation of hypoxia-inducible factors. *Nature* 2012; **487**: 380–384.

20. Adorno M, Cordenonsi M, Montagner M, Dupont S, Wong C, Hann B et al. A Mutant-p53/Smad complex opposes p63 to empower TGFbeta-induced metastasis. *Cell* 2009; **137**: 87–98.

21. Desmedt C, Haibe-Kains B, Wirapati P, Buyse M, Larsimont D, Bontempi G et al. Biological processes associated with breast cancer clinical outcome depend on the molecular subtypes. *Clin Cancer Res* 2008; **14**: 5158–5165.

22. Desmedt C, Di Leo A, de Azambuja E, Larsimont D, Haibe-Kains B, Selleslags J et al. Multifactorial approach to predicting resistance to anthracyclines. *J Clin Oncol* 2011; **29**: 1578–1586.

23. Bartlett JM, Munro A, Cameron DA, Thomas J, Prescott R, Twelves CJ. Type 1 receptor tyrosine kinase profiles identify patients with enhanced benefit from anthracyclines in the BR9601 adjuvant breast cancer chemotherapy trial. *J Clin Oncol* 2008; **26**: 5027–5035.

24. Martin M, Romero A, Cheang MC, López García-Asenjo JA, García-Saenz JA, Oliva B et al. Genomic predictors of response to doxorubicin versus docetaxel in primary breast cancer. *Breast Cancer Res Treat* 2011; **128**: 127–136.

25. Li Y, Zou L, Li Q, Haibe-Kains B, Tian R, Li Y et al. A genomic predictor of response and survival following taxane-anthracycline chemotherapy for invasive breast cancer. *JAMA* 2011; **305**: 1873–1881.

26. Tabchy A, Valero V, Vidaurre T, Lluch A, Gomez H, Martin M et al. Evaluation of a 30-gene paclitaxel, fluorouracil, doxorubicin, and cyclophosphamide chemotherapy response predictor in a multicenter randomized trial in breast cancer. *Clin Cancer Res* 2010; **16**: 5351–5361.

27. Arriola E, Rodriguez-Pinilla SM, Lambros MB, Jones RL, James M, Savage K et al. Anthracyclines in HER2-positive early breast cancer. *Breast Cancer Res Treat* 2008; **106**: 5027–5035.

28. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007; **8**: 118–127.

29. Wirapati P, Sotiriou C, Kunkel S, Farmer P, Pradervand S, Haibe-Kains B et al. Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures. *Breast Cancer Res* 2008; **10**: R65.

30. Haibe-Kains B, Schroeder M, Bontempi G, Sotiriou C, J Quackenbush genome: Relevant function for gene expression analysis, especially in breast cancer. *R* package version 1.6.1; http://www.bioconductor.org/packages/release/bioc/vignettes/genefu/inst/doc/genefu.pdf (Accessed on 31st October 2013).

31. Hatzis C, Pusztai L, Valero V, Booser DJ, Esserman L, Lluch A et al. A genomic predictor of response and survival following taxane-anthracycline chemotherapy for invasive breast cancer. *JAMA* 2011; **305**: 1873–1881.

32. Tabchy A, Valero V, Vidaurre T, Lluch A, Gomez H, Martin M et al. Evaluation of a 30-gene paclitaxel, fluorouracil, doxorubicin, and cyclophosphamide chemotherapy response predictor in a multicenter randomized trial in breast cancer. *Clin Cancer Res* 2010; **16**: 5351–5361.

33. Iwamoto T, Bianchini G, Booser D, Qi Y, Coutant C, Shiung CY et al. Gene pathways associated with prognosis and chemotherapy sensitivity in molecular subtypes of breast cancer. *J Natl Cancer Inst* 2011; **103**: 264–272.

34. Li Y, Zou L, Li Q, Haibe-Kains B, Tian R, Li Y et al. Amplification of LAPTMB4 and YWHAZ contributes to chemotherapy resistance and recurrence of breast cancer. *Nat Med* 2010; **16**: 214–218.

35. Baraghani A, Chiavitti A, di Leo A, Re-searching anthracycline therapy. *Breast Cancer Res Treat* 2010; **123**: 171–175.

36. Schiff R, Massarweh S, Shou J, Osborne CK. Breast cancer endocrine resistance: how growth factor signaling and estrogen receptor coregulators modulate response. *Clin Cancer Res* 2003; **9**: 4475–4545.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/4.0/

Supplementary Information accompanies the paper on the *npj Breast Cancer* website (http://www.nature.com/npjbcancer)