ESMO GOOD SCIENCE BETTER MEDICINE BEST PRACTICE

ANNALS OF ONCOLOGY driving innovation in oncology

ORIGINAL ARTICLE

Reconstructing tumor history in breast cancer: signatures of mutational processes and response to neoadjuvant chemotherapy☆

C. Denkert[1,2*†], M. Untch[3†], S. Benz[4], A. Schneeweiss[5], K. E. Weber[6], S. Schmatloch[7], C. Jackisch[8], H. P. Sinn[9,10], J. Golovato[4], T. Karn[11], F. Marmé[12], T. Link[13], J. Budczies[2,9,10], V. Nekljudova[6], W. D. Schmitt[2], E. Stickeler[14], V. Müller[15], P. Jank[1,2], R. Parulkar[4], E. Heinmöller[16], J. Z. Sanborn[4], C. Schem[17], B. V. Sinn[2], P. Soon-Shiong[4], M. van Mackelenbergh[18], P. A. Fasching[19], S. Rabizadeh[4] & S. Loibl[6,20]

[1]Institute of Pathology, Philipps-University Marburg and University Hospital Marburg (UK-GM), Marburg; [2]Charité — Universitätsmedizin Berlin, Institute of Pathology, Berlin; [3]Helios Klinikum Berlin-Buch, Department of Obstetrics and Gynaecology, Berlin, Germany; [4]NantOmics, LLC, Culver City, USA; [5]Nationales Centrum für Tumorerkrankungen, Universitätsklinikum und Deutsches Krebsforschungszentrum Heidelberg, Heidelberg; [6]German Breast Group (GBG), Neu-Isenburg; [7]Brustzentrum Kassel, Elisabeth Krankenhaus, Kassel; [8]Department of Obstetrics and Gynecology and Breast Cancer Center, Sana Klinikum Offenbach, Offenbach; [9]Institute of Pathology, University Hospital Heidelberg, Heidelberg; [10]German Cancer consortium (DKTK), Heidelberg; [11]Klinik für Frauenheilkunde und Geburtshilfe, Goethe Universität, Frankfurt; [12]Universitätsfrauenklinik Mannheim, Mannheim; [13]Department of Gynecology and Obstetrics, Technische Universität Dresden, Dresden; [14]Department of Gynecology, RWTH Aachen, Aachen; [15]Department of Gynecology, Universitätsklinikum Hamburg-Eppendorf, Hamburg; [16]Pathologie Nordhessen, Kassel; [17]Mammazentrum Hamburg am Krankenhaus Jerusalem, Hamburg; [18]Universitätsklinikum Schleswig-Holstein, Klinik für Gynäkologie und Geburtshilfe, Kiel; [19]Department of Gynecology and Obstetrics, University Hospital Erlangen, Comprehensive Cancer Center Erlangen-EMN, Erlangen; [20]University of Frankfurt, Frankfurt am Main, Germany

Available online 6 January 2021

Background: Different endogenous and exogenous mutational processes act over the evolutionary history of a malignant tumor, driven by abnormal DNA editing, mutagens or age-related DNA alterations, among others, to generate the specific mutational landscape of each individual tumor. The signatures of these mutational processes can be identified in large genomic datasets. We investigated the hypothesis that genomic patterns of mutational signatures are associated with the clinical behavior of breast cancer, in particular chemotherapy response and survival, with a particular focus on therapy-resistant disease.

Patients and methods: Whole exome sequencing was carried out in 405 pretherapeutic samples from the prospective neoadjuvant multicenter GeparSepto study. We analyzed 11 mutational signatures including biological processes such as APOBEC-mutagenesis, homologous recombination deficiency (HRD), mismatch repair deficiency and also age-related or tobacco-induced alterations.

Results: Different subgroups of breast carcinomas were defined mainly by differences in HRD-related and APOBEC-related mutational signatures and significant differences between hormone-receptor (HR)-negative and HR-positive tumors as well as correlations with age, Ki-67 and immunological parameters were observed. We could identify mutational processes that were linked to increased pathological complete response rates to neoadjuvant chemotherapy with high significance. In univariate analyses for HR-positive tumors signatures, S3 (HRD, $P < 0.001$) and S13 (APOBEC, $P = 0.001$) as well as exonic mutation rate ($P = 0.002$) were significantly correlated with increased pathological complete response rates. The signatures S3 (HRD, $P = 0.006$) and S4 (tobacco, $P = 0.011$) were prognostic for reduced disease-free survival of patients with chemotherapy-resistant tumors.

*Correspondence to: Prof. Carsten Denkert, UKGM (University Hospital of Gießen and Marburg) - Universitätsklinikum Marburg, Philipps-Universität Marburg, Baldingerstraße 1, D-35043 Marburg,
Germany. Tel: +49-06421-58-62270; Fax: +49-06421-58-62257
E-mail: carsten.denkert@uni-marburg.de (C. Denkert).

☆Note: This study was previously presented in part at the ASCO Meeting 2018, Chicago, IL, USA.

† Co-first authors.

Conclusion: The results of this investigation suggest that the clinical behavior of a tumor, in particular, response to neoadjuvant chemotherapy and disease-free survival of therapy-resistant tumors, could be predicted by the composition of mutational signatures as an indicator of the individual genomic history of a tumor. After additional validations, mutational signatures might be used to identify tumors with an increased response rate to neoadjuvant chemotherapy and to define therapy-resistant subgroups for future therapeutic interventions.
Key words: breast cancer, mutational signatures, neoadjuvant therapy, whole exome sequencing, response, prognosis

INTRODUCTION

Signatures of mutational processes have been described in several comprehensive analyses based on whole exome or whole genome sequencing.[1-3] These signatures are the consequence of the activity of endogenous and exogenous mutational processes acting in combination over long periods of time to generate the specific mutational landscape of each individual tumor.[4]

The molecular basis for identification of these mutational signatures is the observation that different mutational processes do not act completely randomly, but that each mutational process generates a specific pattern that can be identified by sequencing analysis. For example, specific patterns of mutations caused by UV light[5] (with strong prevalence in melanoma) or by tobacco carcinogens[6] (with strong prevalence in tumors of patients with a smoking history) have been identified.

The main challenge in identification of mutational signatures is that many different processes act over time to generate an individual pattern in each tumor. The theoretical model and computational framework to separate these different mutational signatures has been established by Alexandrov et al.[1,2] More than 30 different types of mutational signatures have been identified.[7-9] Based on these results, it is now possible to identify the contribution of different mutational processes to the genomic landscape and to reconstruct the history of each individual tumor.

For breast cancer, the predominant mutational processes include mutations induced by pathological activation of APOBEC enzymes,[10] mutations caused by alterations of *BRCA*-related[11] or mismatch repair (MMR)-related[12] DNA repair pathways and age-related mutational processes.[13]

The concept of mutational signatures provides a mechanistic biological framework explaining the complex interaction of different endogenous and exogenous carcinogenic pathways. The most important next step is the translation of this biological concept into clinical applications. Despite the progress in identifying oncogenic pathways and individual targetable mutations, the prediction of therapy response in solid tumors is still a major challenge. These tumors develop over a very long time by a combination of parallel mutational processes; therefore, there is a very strong hypothesis that the specific pattern of mutational signatures might also be relevant for the response of a tumor to a given therapy. Furthermore, these mutational processes do not act only on the cells of the clinically relevant tumor; they act on all cells in the patient's body. The majority of neoplastic cells or premalignant lesions are eliminated by cellular control processes or immunological processes, and progression into a clinically relevant tumor is a comparably rare event. We and others have provided evidence that immunological parameters are highly relevant for patient prognosis and success of therapy in breast cancer and other types of tumors.[14,15]

This elimination of premalignant lesions could be seen as a continuous training situation for the immune system, and this training may be different based on the different mutagenic processes that act over the lifetime of each human being. The mutational pattern observed in a clinically manifest tumor, therefore, may also allow us to look back into the history of all other premalignant lesions that have been successfully eliminated.

In this project, we have evaluated the hypothesis that signatures of mutational processes that have contributed to the training of the immune system and other antineoplastic processes are relevant for response to neoadjuvant chemotherapy and prognosis in breast cancer. To evaluate this hypothesis, we have carried out whole exome sequencing of a cohort of 405 pretherapeutic formalin-fixed, paraffin-embedded (FFPE) core biopsies from patients included in the neoadjuvant GeparSepto trial.

We investigated the prevalence of the different mutational signatures in different breast cancer subgroups as well as the role of the most predominant signatures for response to neoadjuvant chemotherapy.

PATIENTS AND METHODS

Patients and treatment

In the GeparSepto study (NCT01583426)[16,17] women with previously untreated, primary invasive breast cancer were included after written informed consent for study participation and biomaterial collection. This translational project was restricted to patients with human epidermal growth factor receptor 2 (HER2)-negative tumors ($n = 810$). Details on the clinical trial parameters are given in the Supplementary material, Methods section, available at https://doi.org/10.1016/j.annonc.2020.12.016.

Objectives and endpoints of the analysis

The current analysis is an exploratory retrospective analysis of a cohort from a prospective clinical trial. The primary objective was to investigate mutational signatures and mutational load in the tumor to predict response and

resistance to neoadjuvant therapy. The secondary objectives were to correlate mutational signatures with clinicopathological parameters. Predefined endpoints were pathological complete response (pCR) defined as ypT0ypN0 and disease-free survival (DFS). Predefined covariables for multivariate regression models were age (continuous), tumor stage (T1-2 versus T3-4), nodal stage (N0 versus N+), Ki-67 (continuous), hormone receptor (HR) status (negative versus positive) and treatment (nab-P versus P).

DNA isolation, whole exome sequencing and statistical analysis

A total of 703 FFPE core biopsies were selected based on the availability of sufficient tumor tissue after histopathological quality control and paired blood samples to control for germline alterations and single-nucleotide polymorphisms (SNPs). A total of 703 samples were shipped to NantOmics (Culver City, CA). Automated DNA extraction was carried out using QIAsymphony DSP DNA Kit (Qiagen, Germantown, MD) on the QIAsymphony. Quantification was done using the Invitrogen™ Quant-iT™ dsDNA Assay Kit, Broad Range (Thermofisher, Waltham, MA) and the BioTek Synergy HTX Multi-Mode Reader (BioTek, Winooski, VT). Samples were sequenced on an Illumina HiSeq sequencer (Illumina, San Diego, CA) according to standard NantOmics protocols. Two hundred and eight samples were not successfully sequenced due to failures during DNA extraction (<50 ng of extracted DNA) or failures during library preparation for sequencing which were established by analysis of electropherograms and yield post-library preparation. Of the 495 samples that were sequenced, 15 had to be included due to provenance fail (mismatch) and 75 due to contamination fail. Therefore, a total of 298 samples were excluded and successful whole exome sequencing was carried out for 405 samples after filtering for contamination and tumor/normal mismatches.

Mutational signatures were determined by the bioinformatical team at NantOmics LLC, using the method described by Alexandrov et al.,[1] which was reimplemented in Python via NMF in scikit learn, using the package versions scikit-learn = 0.14.1, scipy = 0.13.3 and numpy = 1.9.2 (https://scikit-learn.org/stable/). The matrix weights were used from the mutational signature website[18] using version v2 of mutational signatures. In addition, the exonic mutation rate (EMR) per Mb was calculated. Using the published algorithm for detection of mutational signatures, we identified the absolute integer number of mutations assigned to each signature in each tumor. In addition, we determined the presence of each mutational signature (with at least one mutation assigned to this signature) as a binary variable.

For a detailed evaluation, we selected only those 11 mutational signatures that were relevant for this dataset. The strategy for selection of mutational signature and the details of the statistical analysis are described in Supplementary methods, available at https://doi.org/10.1016/j.annonc.2020.12.016.

RESULTS

Baseline clinical data

The set of successfully sequenced patients ($n = 405$; Supplementary Figure S1, available at https://doi.org/10.1016/j.annonc.2020.12.016, consort diagram) did not significantly differ from the HER2-negative GeparSepto patients not sequenced with respect to tumor size, nodal status, grading, treatment arm and pCR rate (Supplementary Table S1, available at https://doi.org/10.1016/j.annonc.2020.12.016). In the sequenced cohort, there were fewer HR-negative tumors and fewer invasive-ductal tumors. The patients in the whole exome sequencing group were also slightly younger (mean age 49 years versus 50 years in the non-analyzed cohort; $P = 0.027$) and had slightly higher levels of stromal tumor-infiltrating lymphocytes (TILs) (mean 26% versus 23% in non-analyzed cohort; $P = 0.006$).

Frequency of mutational signatures in HR-positive and -negative tumors

In the complete dataset of 405 tumors, the median number of mutations was 153 [minimum: 26, maximum: 13 072; interquartile range (IQR) 159]. In the subset of 284 HR-positive/HER2-negative tumors, the median number of mutations was 142 (minimum: 27, maximum: 13 072; IQR 135). In the subset of 121 triple-negative tumors, the median number of mutations was 220 (minimum: 26, maximum: 1059; IQR 168).

We evaluated the number of mutations of each signature and the EMR in HR-positive and HR-negative tumors (Figure 1). Significant differences with higher numbers of mutations in HR-negative tumors are observed for S3 (homologous recombination deficiency, HRD; $P < 0.001$), S13 (APOBEC; $P < 0.001$), S6 (MMR; $P = 0.015$), S21 (MMR; $P = 0.015$), S4 (tobacco; $P < 0.001$) and in total (EMR; $P < 0.001$). Significantly higher mutation numbers in HR-positive tumors are observed for S16 (unknown process; $P < 0.001$) and S28 (unknown process; $P = 0.002$). One single tumor had a particularly high value for the MMR-related signatures S15, S21, S26 and a high EMR, which was validated by a reduced immunohistochemical expression of MSH2 and MSH6 indicating microsatellite-instability (MSI; Figure 1B).

Patterns of mutational signatures—HRD and APOBEC

To give an overview on the patterns of mutational signatures, an unsupervised cluster analysis in HR-positive and HR-negative tumors was carried out. In HR-positive tumors, two main clusters and four smaller clusters were observed (Figure 2A and B). One main cluster (C1) was characterized by the presence of signature S3 (HRD); the other main cluster (C4) showed an absence of S3 and a mixture of other signatures with comparably high levels of signatures with unknown role (S16, S28). The cluster C2 contained mainly APOBEC-related signatures, while the cluster C3 consisted of age-related mutations (S1) with low mutation numbers in most patients. A very small cluster (C5) included tumors

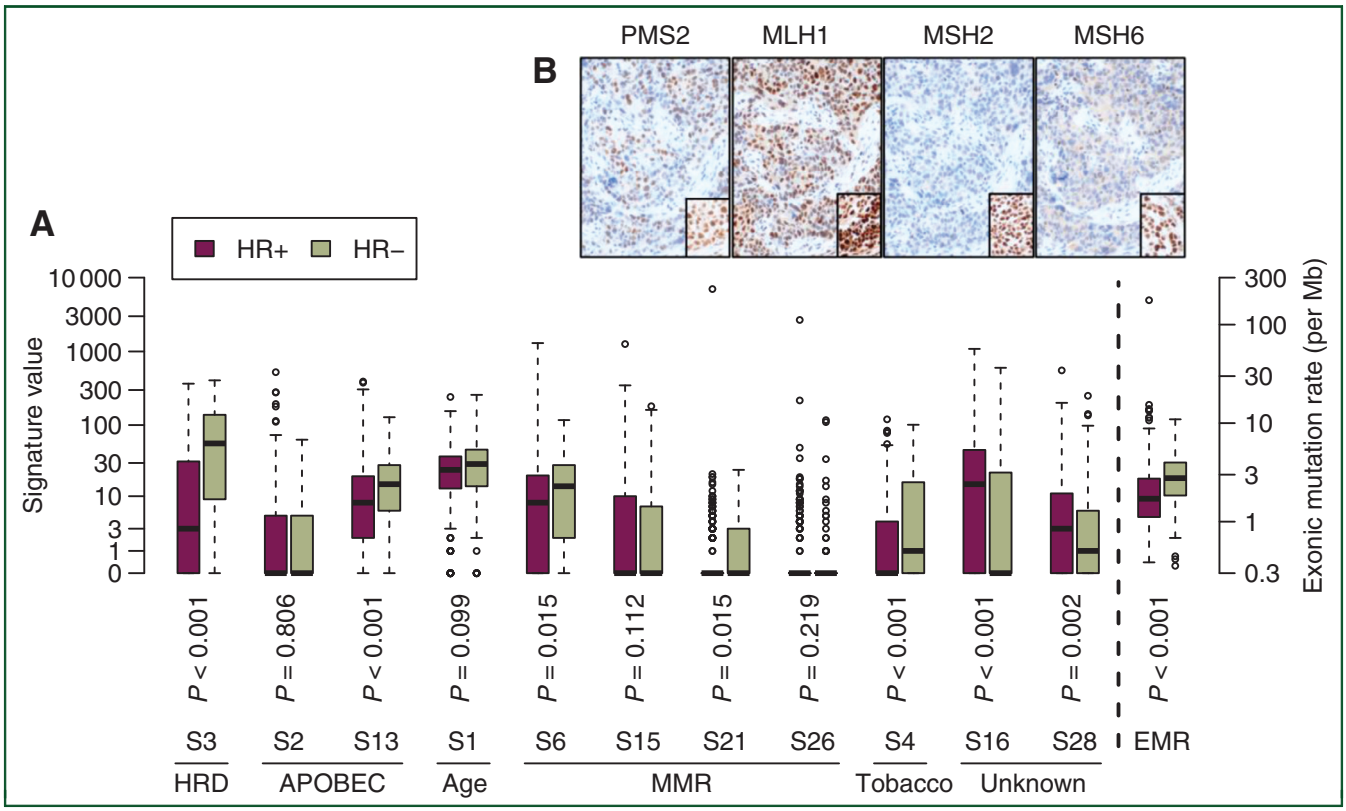

Figure 1. Comparison of mutational signatures in HR-positive and HR-negative breast carcinomas.
(A) Absolute number of mutations of each signature for HR-negative and HR-positive tumors. In addition, EMR is shown (right part, separate y-axis). *P* values were calculated using a two-sample Wilcoxon test. (B) Immunohistochemistry for microsatellite-instability (MSI) of one tumor with high EMR and high rates of MMR-related signatures S15, S21, S26. This tumor is marked with gray boxes in A. A reduced nuclear expression of MSH2 and MSH6 is observed (inserts: positive controls).
EMR, exonic mutation rate; HR, hormone receptor; HRD, homologous recombination deficiency; MMR, mismatch repair.

with high levels of the tobacco-related signature S4. The cluster C6 consisted of a small group of tumors with a higher contribution of other signatures. In HR-negative tumors (Figure 3A and B) the clusters C1-C5 were observed as well, and the majority of tumors showed S3 (HRD)-related mutations (C1).

Mutational signatures, therapy response and prognosis

As expected in breast cancer, the overall pCR rate of patients with HR-negative [triple-negative breast cancer (TNBC)] tumors was significantly higher than the pCR rate of patients with HR-positive tumors (38.8% versus 12.7%; $P < 0.001$).

Differences in pCR rate for different mutational signatures were observed in HR-positive tumors. For HR-positive tumors, the two main mutational signatures S3 (HRD, $P < 0.001$) and S13 (APOBEC, $P = 0.001$), evaluated as continuous variables, were significantly correlated with an increased pCR rate in univariable analysis (Figure 2C). In multivariable analysis, the correlation was highly significant for S13 (APOBEC, $P = 0.008$) and borderline significant for S3 (HRD, $P = 0.059$, Figure 4A). Tumors with S3 levels above the median had a pCR rate of 18%, compared with 8% for tumors with S3 below the median ($P = 0.012$, Figure 4B). Similarly, tumors with S13 above the median had a pCR rate of 22%, compared with 4% for tumors with low S13 levels ($P < 0.001$, Figure 4C).

For survival analysis, we focused on the clinically most relevant subgroup of therapy-resistant tumors not responding to the neoadjuvant therapy (non-pCR). In HR-positive tumors without a pCR, the signatures S3 (HRD, $P = 0.006$) and S4 (tobacco, $P = 0.011$) were linked to reduced DFS (Figure 2D) in univariable analysis, but not in multivariable analysis including the covariables age, cT, cN, Ki-67 and treatment (Supplementary Figure S2, available at https://doi.org/10.1016/j.annonc.2020.12.016). Signature S6 (MMR) was significant for reduced DFS only in multivariate analysis ($P = 0.038$). In Kaplan–Meier analysis using the median as an exploratory cut point (Figure 2E and F), a significant difference for S3 (HRD, $P = 0.034$) and a borderline significant difference for S4 (tobacco, $P = 0.053$) was observed.

In contrast, in HR-negative tumors there was no mutational signature that significantly predicted pCR (univariable: Figure 3C; multivariable: Supplementary Figure S3, available at https://doi.org/10.1016/j.annonc.2020.12.016). In therapy-resistant HR-negative tumors, S6 (MMR) had a borderline significance for reduced DFS (univariate: $P = 0.071$, Figure 3D; multivariate: $P = 0.044$, Supplementary Figure S4, available at https://doi.org/10.1016/j.annonc.2020.12.016), which was borderline significant in Kaplan–Meier analysis using the median as a cut point ($P = 0.083$, Figure 3E). The univariable and multivariable analysis for DFS in the complete cohort (including pCR- and non-pCR

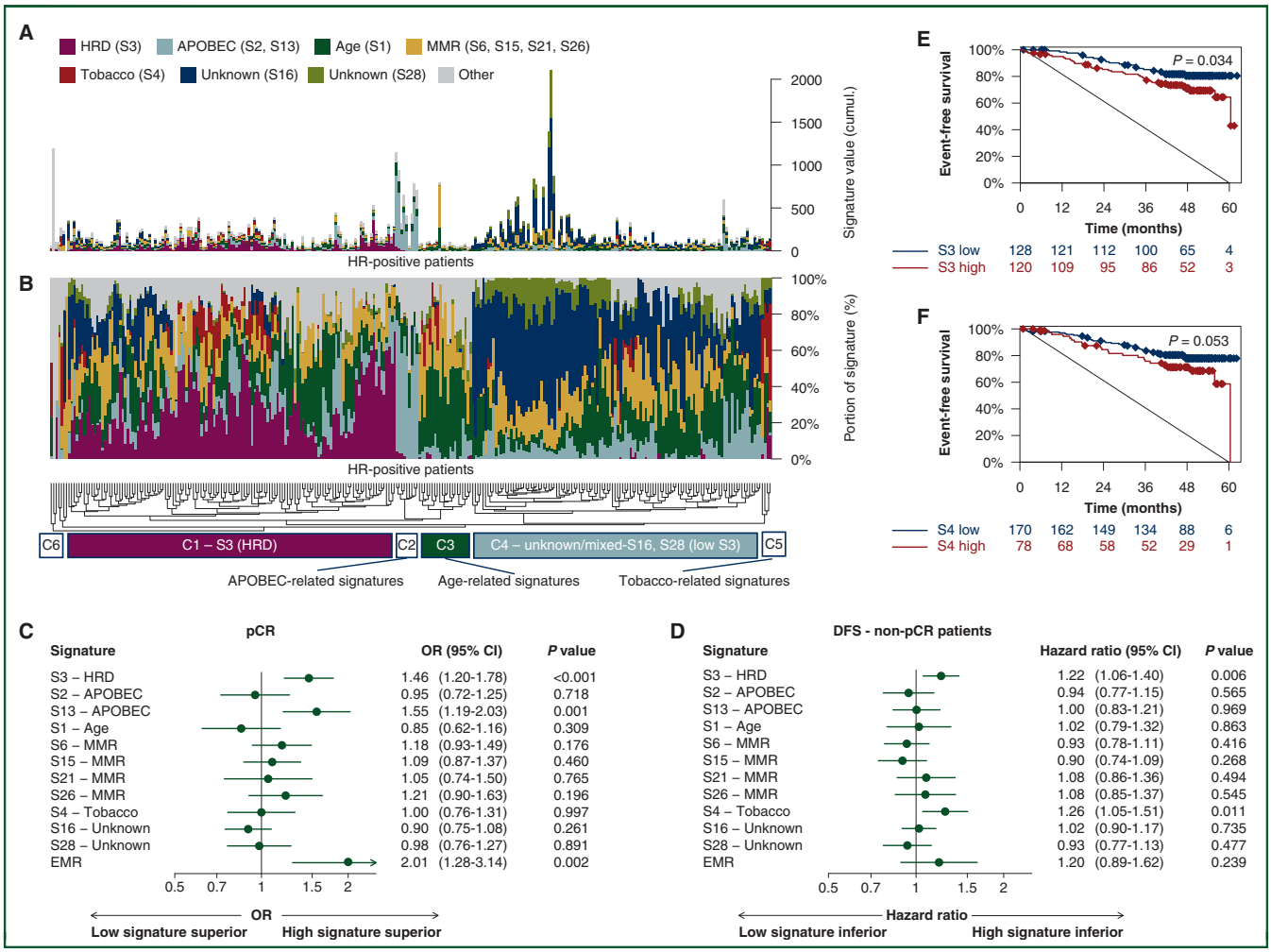

Figure 2. HR-positive tumors: patterns of mutational signatures and impact on therapy response as well as prognosis of therapy resistant tumors.
(A) Overview on absolute number of mutations of each signature in each sample. (In A, one single HR-positive tumor with extremely high mutation numbers has been removed to achieve a comparable scale of the x-axis for comparison with the HR-negative tumors in Figure 3). (B) Overview on the proportion of each signature in each sample. In A and B, mutational signatures are color-coded based on biological processes and patients are ordered by hierarchical clustering. (C) Mutational signatures and pCR to neoadjuvant chemotherapy in HR-positive tumors. Univariate logistic regression ($n = 284$). (D) Mutational signatures and DFS in therapy resistant HR-positive tumors (including only those patients with no pCR). Univariate ($n = 248$) Cox-regression for DFS in no-pCR patients. (E, F) Univariate Kaplan–Meier survival analysis in no-pCR patients for S3 (HRD, E) and S4 (tobacco, E) using the median as a cut point (P: log rank test).
APOBEC, apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like; CI, confidence interval; DFS, disease-free survival; HR, hormone receptor; HRD, homologous recombination deficiency; MMR, mismatch repair; OR, odds ratio; pCR, pathological complete response.

tumors) is shown in Supplementary Figure S5, available at https://doi.org/10.1016/j.annonc.2020.12.016 for HR-positive tumors and Supplementary Figure S6, available at https://doi.org/10.1016/j.annonc.2020.12.016. for HR-negative tumors.

EMR, therapy response and prognosis

In addition to the individual signatures, we also investigated the EMR as a measurement of total mutational burden. EMR correlated highly with increased pCR in HR-positive tumors (univariable $P = 0.002$, Figure 2C; multivariable $P = 0.005$, Figure 4A). HR-positive tumors with EMR levels above the median had a pCR rate of 20%, compared with 6% for tumors with EMR below the median ($P < 0.001$, Figure 4D). EMR did not predict DFS in the HR-positive subcohort (Figure 2D; Supplementary Figures S2 and S5, available at https://doi.org/10.1016/j.annonc.2020.12.016).

In HR-negative tumors, EMR was not significantly associated with pCR (Figure 3C and Supplementary Figure S3, available at https://doi.org/10.1016/j.annonc.2020.12.016) or DFS (Figure 3D; Supplementary Figures S4 and S6, available at https://doi.org/10.1016/j.annonc.2020.12.016).

Correlation with biological parameters—patient age, Ki-67, TILs

We evaluated the correlation between the number of mutations assigned to the different signatures and patient age, proliferation rate (Ki-67) and stromal TILs in HR-positive and HR-negative tumors (Figure 5). In HR-positive tumors, signature S1 had a highly significant correlation with increased patient age ($P < 0.001$, Figure 5A), which

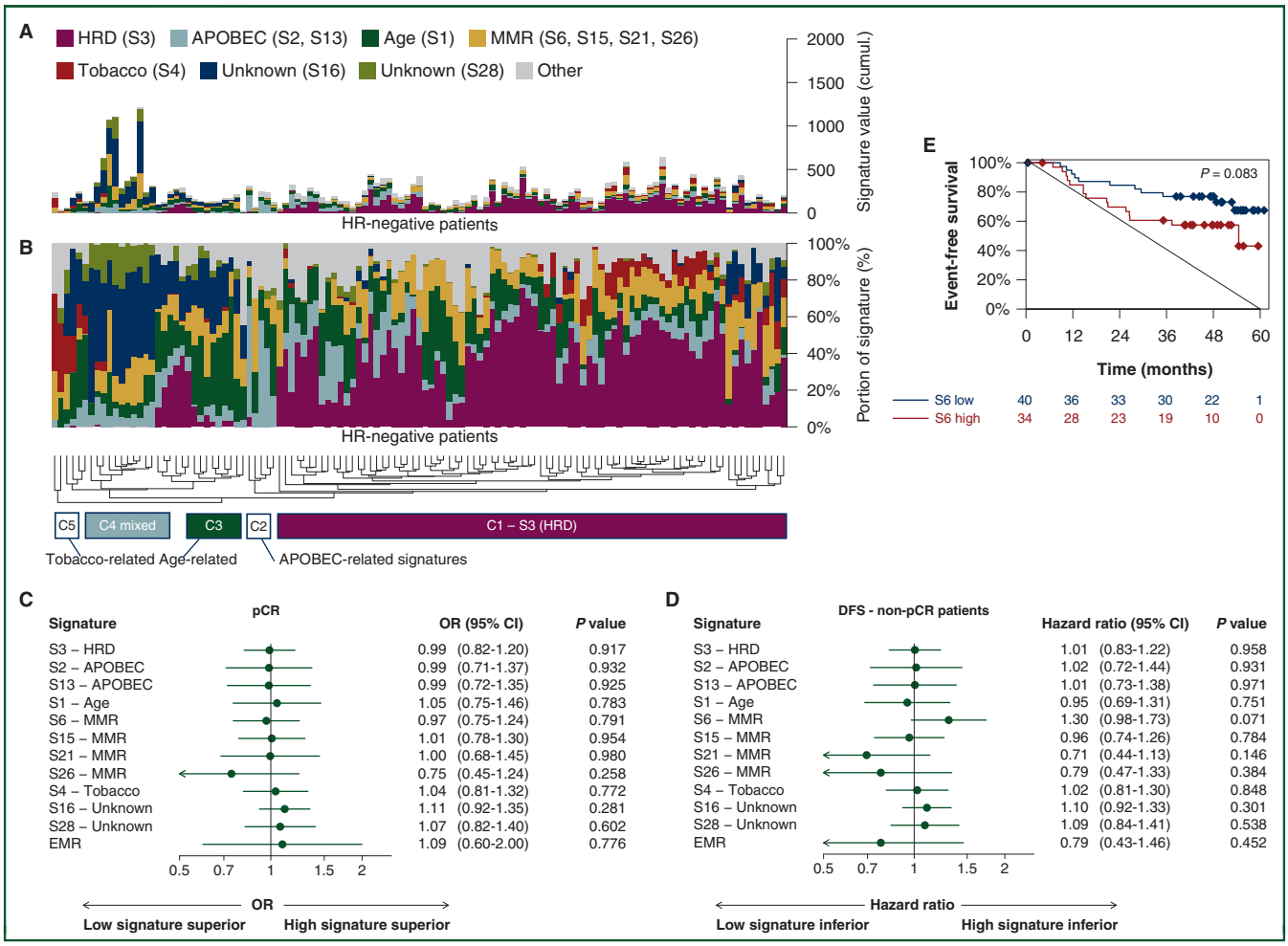

Figure 3. HR-negative (=TNBC) tumors: patterns of mutational signatures and impact on therapy response as well as prognosis of therapy resistant tumors.
(A) Overview on absolute number of mutations of each signature in each sample. (B) Overview on the proportion of each signature in each sample. In A and B, mutational signatures are color-coded based on biological processes and patients are ordered by hierarchical clustering. (C) Mutational signatures and pCR to neo-adjuvant chemotherapy in HR-negative tumors. Univariate logistic regression ($n = 121$). (D) Mutational signatures and DFS in therapy resistant HR-negative tumors (including only those patients with no pCR). Univariate ($n = 74$) Cox-regression for DFS in no-pCR patients. (E) Univariate Kaplan–Meier survival analysis in no-pCR patients for S6 (MMR) using the median as a cut point (*P*: log rank test).
APOBEC, apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like; CI, confidence interval; DFS, disease-free survival; HR, hormone receptor; HRD, homologous recombination deficiency; MMR, mismatch repair; OR, odds ratio; pCR, pathological complete response.

validates S1 as an age-related signature, as described before in other studies.[13]

In highly proliferating HR-positive tumors, significantly more mutations were found for signatures S3 (HRD, $P < 0.001$), S13 (APOBEC, $P < 0.001$) and S4 (tobacco, $P < 0.001$; Figure 5B). In immunologically active HR-positive tumors with high numbers of TILs, more mutations were found for S3 (HRD, $P < 0.001$) and S13 (APOBEC, $P < 0.001$; Figure 5C).

For HR-negative (TNBC) tumors (Figure 5A and B) there were no significant correlations of the mutational signatures with patient age and proliferation rate; for TILs, there were significant negative associations with S6 and S21 (both MMR, $P = 0.003$ and $P = 0.031$; Figure 5C).

In Figure 5, we have also evaluated the EMR as a parameter of tumor mutational burden in correlation with age, Ki-67 and TILs. All three correlations were positive and significant for HR-positive tumors (Figure 5). In this subtype, significantly higher mutation rates were observed in tumors with high TILs ($P < 0.001$), high proliferation ($P < 0.001$) and in tumors from older patients ($P = 0.046$). In HR-negative tumors, in contrast, EMR is significantly negatively correlated with high TILs ($P = 0.005$, Figure 5C). For HR-negative tumors there was also a significant positive correlation of EMR with age ($P = 0.042$).

In addition, we have also evaluated the correlation of mutational signatures and the categorical clinicopathological parameters grade, tumor size, nodal status as well as the therapy arm. The results are shown in Supplementary Tables S2 and S3, available at https://doi.org/10.1016/j.annonc.2020.12.016.

DISCUSSION

Signatures of mutational processes are indicators of multiple genetic events that act over years on all cells in the human body. Their biological relevance has been investigated in large institutional cohorts as well as tissue repositories of large international consortia. As a next step for

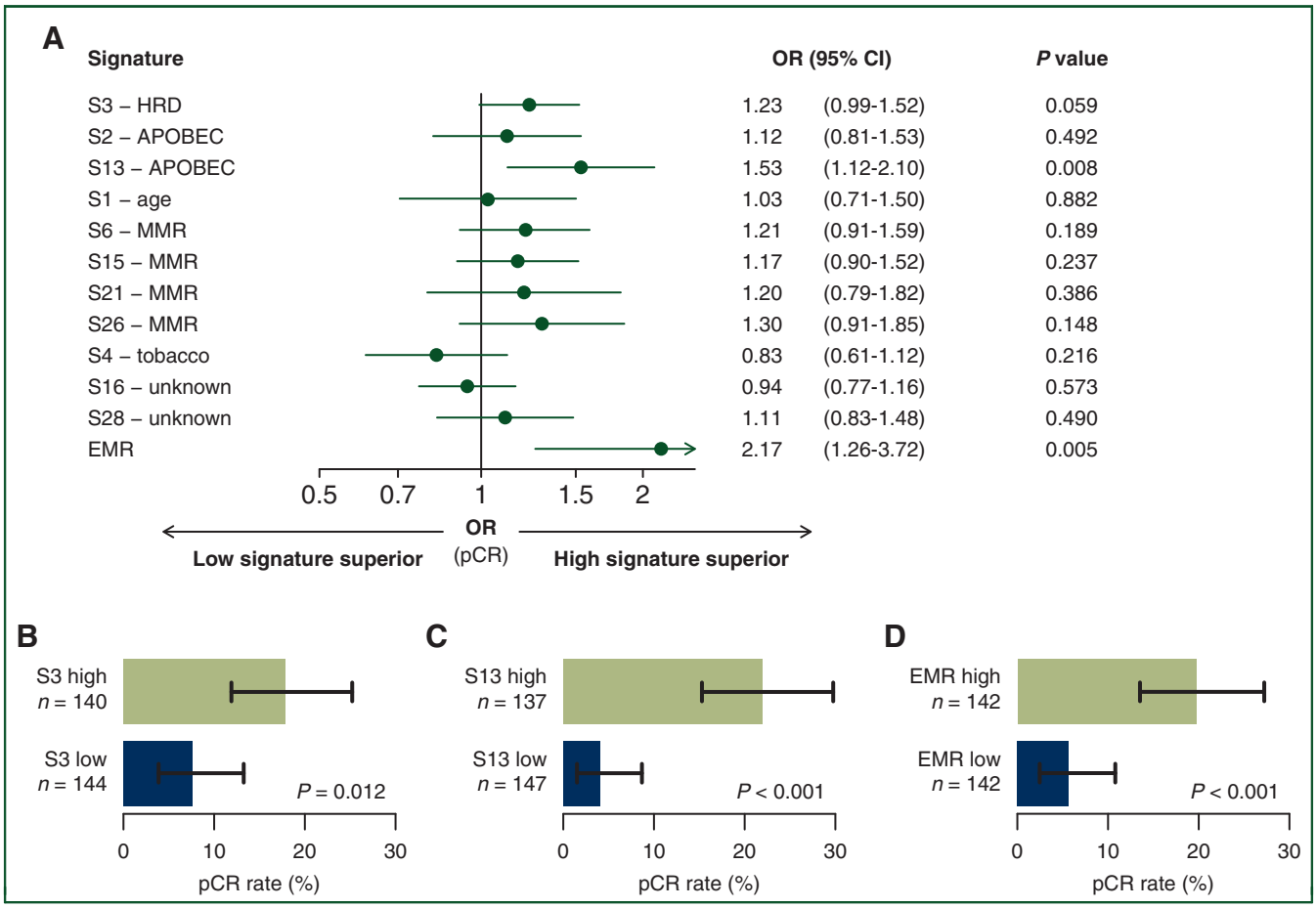

Figure 4. Pathological complete response to neoadjuvant chemotherapy in HR-positive tumors—multivariable analysis and differences in pCR rate.
(A) Multivariate logistic regression ($n = 274$). The covariables in the multivariable models are age, cT, cN, Ki-67, and treatment. (B, C, D) pCR rate for the variables with significance in the univariate model using the median as cut point.
APOBEC, apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like; CI, confidence interval; EMR, exonic mutation rate; HRD, homologous recombination deficiency; MMR, mismatch repair; OR, odds ratio; pCR, pathological complete response.

translation of these findings into the clinical setting, a detailed investigation of the clinical effects of different mutational signatures on therapy response and survival is necessary. In this study, we have used a large well-defined clinical trial cohort from a neoadjuvant clinical multicenter trial with known therapy response data and survival data, including HR-positive and HR-negative tumors.

The main result was that defined signatures could predict the clinical behavior of HR-positive tumors, in particular response to neoadjuvant chemotherapy and DFS of non-pCR patients. We could identify mutational signatures S3 (HRD) and S13 (APOBEC) that were linked to increased pCR rates to neoadjuvant chemotherapy with high significance in the HR-positive subcohort. The signatures S3 (HRD) and S4 (tobacco) were also associated with a reduced DFS in therapy-resistant (non-pCR) HR-positive tumors.

In addition, we investigated the association of S3 and S13 with tumor proliferation and tumor immune infiltrate. Both signatures were positively correlated with increased Ki-67 and also with increased levels of TILs, which might suggest involvement in the development of more aggressive subtypes of luminal tumors which also have a higher immunogenicity. These highly proliferating tumors respond better to chemotherapy, but may have a reduced prognosis overall, which is in line with the results of our investigation. The correlation with Ki-67 might also partly explain the non-significance in multivariable analysis.

A multitude of translational investigations have described TILs as one of the most relevant factors for chemotherapy response and prognosis. In our investigation, we have observed a possible explanation why some tumors have an accumulation of TILs. These tumors have been induced by mutational processes involving HRD-like alterations and the action of APOBEC signatures. Interestingly, there are comprehensive data from preclinical investigations that HRD and APOBEC might induce major immunological alterations. A molecular link between inhibition of PARP in *BRCA*-mutated tumor and immune activation via STING-pathway activation, which is triggered by cytosolic DNA, has been described for ovarian cancer and TNBC.[19,20] APOBEC enzymes are up-regulated by viral infections and physiologically inhibit retrovirus and transposon replication.[21] They also play an important role in mutagenesis in different cancer types, including breast cancer.[22]

In a previous investigation, we have shown that TILs are predictive for increased neoadjuvant response in HR-positive and HR-negative breast cancer, while for survival, TILs are a positive survival factor in TNBC and a negative

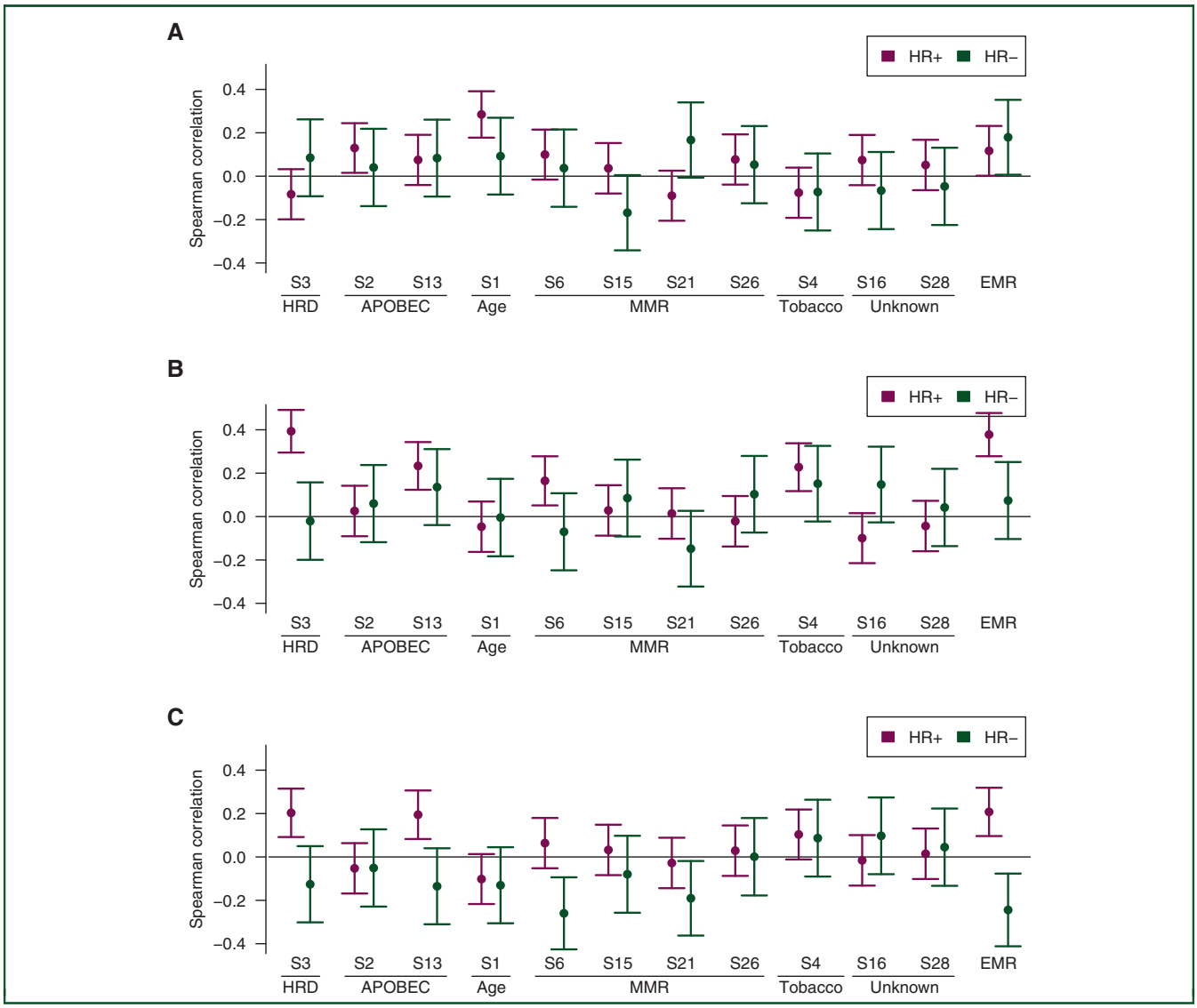

Figure 5. Correlation of absolute numbers of selected mutational signatures as well as EMR with the biological parameters.
(A) Patient age, (B) proliferation rate (Ki-67) and (C) immune cell infiltrate (stromal TILs). The Spearman correlation coefficient is shown (error bars: 95% CI).
APOBEC, apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like; EMR, exonic mutation rate; HR, hormone receptor; HRD, homologous recombination deficiency; MMR, mismatch repair; TILs, tumor-infiltrating lymphocytes.

survival factor in luminal/HER2-negative breast cancer.[14] The differences observed in the present comparison of mutational signatures between HR-positive and HR-negative tumors might at least partially explain this different prognostic effect of TILs. In luminal tumors, there is a high diversity of mutational processes and TILs are associated with those processes that are observed in more aggressive tumors, in particular S3 (HRD). Therefore, TILs are also linked to worse prognosis.

In TNBC, the repertoire of mutational signatures is more homogeneous, and there is no association with TILs or Ki-67. Therefore, in this subtype, tumors with different TIL levels are not distinguished by different mutational signatures. The higher homogeneity of the HR-negative tumors might be explained by the faster growth rate of these tumors and the shorter history of development for each individual tumor. For a fast growing tumor, immune escape might be more relevant compared with a slowly growing tumor, and only those fast growing tumor clones that survive are able to accumulate non-immunogenic mutations. This could explain the observation that high EMR is linked to reduced TILs in HR-negative (TNBC) tumors, which has also been observed in other studies.[23]

In our study, EMR and TILs, as an indicator of immunogenicity of a tumor, are positively correlated in HR-positive tumors, but negatively correlated in HR-negative tumors, which is in line with previous studies of TCGA data.[24]

In HR-positive breast cancer, this correlation is driven by two signatures, S3 and S13, which are the predominant signatures in HR-positive breast cancer with considerable intertumoral heterogeneity, resulting in a positive correlation of TILs and EMR and a highly significant association of EMR and chemotherapy response. In contrast, in HR-negative tumors, EMR is not related to chemotherapy

https://doi.org/10.1016/j.annonc.2020.12.016

response, and there is an inverse correlation of EMR and TILs. In contrast, in the GeparNuevo trial we observed an association between treatment response and TMB.[25]

For some of the signatures with a known biological function, we have been able to perform a validation in our dataset. We have validated signature S1 by its correlation with patient age and identified one tumor with MSI by the high EMR and the high levels of MMR-related signatures S15, S21 and S26.

It should be noted that there are relevant limitations of this project. As recently discussed,[26] the bioinformatical detection of mutational signatures is not always straightforward, and results vary depending on the methods used. In our analysis, some unknown signatures have been detected, particularly S16 and S28, which are not typical for breast cancer. This limits the interpretation of this finding, and we cannot exclude that the unexpected high prevalence of these unknown signatures might be related to artifacts included by the algorithms or the biomaterials used. In particular, we used small FFPE core biopsies for our study, which might limit the detection of some signatures and might induce some changes due to paraffin artifacts. We have carried out and published a previous analysis within our own laboratory indicating that limited bias is introduced from the FFPE process.[27]

In our study, we have decided to use a predefined standard algorithm for detection of signatures, to test its relevance in clinical trial cohorts. A further optimization of the bioinformatical methods was not the aim of our study. Nevertheless, a further standardization of the detection algorithms for clinical applications, as well as validation in other cohorts, is essential for further translation into clinical practice.

We would also like to emphasize that mutational signatures are quasi-continuous variables and therefore will not result in distinct, defined tumors types. From Figures 2 and 3 it is evident that some patterns can be delineated by clustering algorithms, but these should not be interpreted as new subtypes of breast cancer. For most tumors, the mutational signatures are heterogenous, suggesting that more than one mutational process is relevant for the development and progression of each individual tumor.

We would also like to emphasize that while mutational signatures in general provide a promising model for the development of malignant tumors, we cannot show causality in our present investigation. This analysis is based on a sample cohort from a clinical trial; therefore, it provides just correlative data of different markers, and no mechanistic and causal information. Many of the alterations are significant only in univariate, but not in multivariate analysis, which might be partly explained by the correlation with known biological parameters, in particular Ki-67. Therefore, these approaches will not replace classical diagnostic approaches that can be conducted without advanced sequencing technology. However, for validation of the mutational signatures, it is reassuring that they correlate with central parameters of tumor biology. Of course, in this study, we cannot draw conclusions of causality, but our results provide a strong hypothesis for further validation of genomic alterations, for example of HRD-related alterations, which have a high prevalence.

In this project, we aimed to look back into the individual history of breast carcinomas to identify the processes that lead to the accumulation of mutations in the development of the tumor. The setup of the GeparSepto cohort allowed us to study differences between fast growing triple-negative tumors and luminal/Her2-negative tumors. Our study underlines the fact that major differences exist between these different tumor types. In TNBC, the different mutational signatures are not associated with outcome parameters such as pCR and DFS, and most mutational signatures are comparably high in this tumor type. This might be explained by the fact that TNBC is a rapidly developing disease, and the rapid onset of the disease leaves not much time in which differences between individual tumors could develop. These tumors simply might not have a long history that could lead to differences in composition of mutational signatures.

In contrast, luminal tumors typically develop over several years, and it was possible to identify differences between individual tumors that suggest different roles for mutational processes in the development of individual tumors. The two main signatures involved are S3 (HRD) and S13 (APOBEC). The high number of HRD-associated mutations in subgroups of luminal tumors, as well as its association with therapy response, could be a basis for additional validations, which might open new therapeutic options for the HR-deficient subgroup of luminal tumors.

ACKNOWLEDGEMENTS

We thank all the patients who participated in the clinical study and the translational research and all investigators, pathologists and study personnel at the sites. We would like to thank Ines Koch, Britta Beyer, Peggy Wolkenstein, Barbara Meyer-Bartell and Silvia Handzik from Charité for their excellent technical assistance as well as Dr. Bärbel Felder and Stefanie Lettkemann from German Breast Group for the translational research organization.

FUNDING

This work was supported by the Translational Oncology Programme of the German Cancer Aid (Deutsche Krebshilfe), project TransLuminal B [grant number 111636] and Integrate -TN [grant number 70113450]. The sequencing analysis was carried out and funded by NantOmics (no grant number).

DISCLOSURE

CD reports personal fees from Novartis, personal fees and travel support from Roche, personal fees from MSD Oncology, from Daiichi Sankyo, grants from Myriad Genetics and GBG, other from Sividon Diagnostics/Myriad, outside the submitted work. In addition, CD has a patent EP18209672 pending, a patent EP20150702464 pending and a patent Software (VMscope digital pathology) pending.

MU reports personal fees and non-financial support from Abbvie, personal fees and non-financial support from Amgen GmbH, personal fees and non-financial support from AstraZeneca, personal fees from Bristol Myers Squibb (BMS), personal fees and non-financial support from Celgene GmbH, personal fees and non-financial support from Daiji Sankyo, personal fees and non-financial support from Eisai GmbH, personal fees from Lilly Deutschland, personal fees and non-financial support from Lilly Int., personal fees and non-financial support from MSD Merck, personal fees and non-financial support from Mundipharma, personal fees and non-financial support from Myriad Genetics, personal fees and non-financial support from Odonate, personal fees and non-financial support from Pfizer GmbH, personal fees from PUMA Biotechnology, personal fees and non-financial support from Roche Pharma AG, personal fees and non-financial support from Sanofi Aventis Deutschland GmbH, personal fees and non-financial support from TEVA Pharmaceuticals Ind Ltd, personal fees and non-financial support from Novartis, personal fees from Pierre Fabre, personal fees and non-financial support from Clovis Oncology, outside the submitted work. SB reports other from NantOmics, LLC, during the conduct of the study; other from NantOmics, LLC, outside the submitted work; and Employee of NantOmics, LLC with equity interest. AS reports grants from Celgene, grants from Roche, grants from AbbVie, grants from Molecular Partner, personal fees from Roche, personal fees from AstraZeneca, personal fees from Celgene, personal fees from Roche, personal fees from Celgene, personal fees from Pfizer, personal fees from AstraZeneca, personal fees from Novartis, personal fees from MSD, personal fees from Tesaro, personal fees from Lilly, other from Roche, outside the submitted work. CJ reports personal fees from Roche, personal fees from Celgene, personal fees from Amgen, during the conduct of the study. TL reports non-financial support from Pharma Mar, non-financial support from Daiichi Sankyo, personal fees and non-financial support from MSD, personal fees from Amgen, personal fees and non-financial support from Pfizer, personal fees from Novartis, personal fees from Teva, personal fees from Tesaro, personal fees and non-financial support from Roche, personal fees and non-financial support from Clovis, non-financial support from Celgene, outside the submitted work.WDS reports grants from German Breast Group, during the conduct of the study; personal fees from AstraZeneca, outside the submitted work. VM reports grants and personal fees from Roche, during the conduct of the study; personal fees from Amgen, Astra Zeneca, Daiichi-Sankyo, Eisai, Pfizer, MSD, Novartis, Roche, Teva, Seattle Genetics and consultancy honoraria from Genomic Health, Hexal, Roche, Pierre Fabre, Amgen, ClinSol, Novartis, MSD, Daiichi-Sankyo, Eisai, Lilly, Tesaro, Nektar, personal fees from Genomic Health, Hexal, Roche, Pierre Fabre, Amgen, ClinSol, Novartis, MSD, Daiichi-Sankyo, Eisai, Lilly, Tesaro and Nektar, other from I Novartis, Roche, Seattle Genetics, Genentech, outside the submitted work. PSS is the CEO of NantOmics and reports non-financial support from NantOmics, outside the submitted work. MvM reports honoraria from Roche, Amgen, Genomic Health, Astra Zeneca, as well as travel support from Lilly and Novartis. PAF reports grants from Novartis, grants from BioNTech, personal fees from Novartis, personal fees from Roche, personal fees from Pfizer, personal fees from Celgene, personal fees from Daiichi-Sankyo, personal fees from AstraZeneca, personal fees from Merck Sharp & Dohme, personal fees from Eisai, personal fees from Puma, grants from Cepheid, personal fees from Lilly, personal fees from Seattle Genetics, during the conduct of the study. SR is the Chief Scientific Officer for NantOmics and reports non-financial support from NantOmics, outside the submitted work. SL reports grants and other from Celgene, grants and other from Roche, during the conduct of the study; grants and other from Abbvie, grants and other from Amgen, grants and other from AstraZeneca, grants and other from Novartis, grants and other from Pfizer, other from Seattle Genetics, other from PriME/Medscape, personal fees from Chugai, grants from Teva, grants from Vifor, grants and other from Daiichi-Sankyo, other from Lilly, other from Samsung, other from Eirgenix, other from BMS, other from Puma, other from MSD, grants from Immunomedics, outside the submitted work. In addition, SL has a patent EP14153692.0 pending. All other authors have declared no conflicts of interest.

CD reports personal fees from Novartis, personal fees and travel support from Roche, personal fees from MSD Oncology, from Daiichi Sankyo, grants from Myriad Genetics and GBG, other from Sividon Diagnostics/Myriad, outside the submitted work. In addition, CD has a patent EP18209672 pending, a patent EP20150702464 pending and a patent Software (VMscope digital pathology) pending. MU reports personal fees and non-financial support from Abbvie, personal fees and non-financial support from Amgen GmbH, personal fees and non-financial support from AstraZeneca, personal fees from Bristol Myers Squibb (BMS), personal fees and non-financial support from Celgene GmbH, personal fees and non-financial support from Daiji Sankyo, personal fees and non-financial support from Eisai GmbH, personal fees from Lilly Deutschland, personal fees and non-financial support from Lilly Int., personal fees and non-financial support from MSD Merck, personal fees and non-financial support from Mundipharma, personal fees and non-financial support from Myriad Genetics, personal fees and non-financial support from Odonate, personal fees and non-financial support from Pfizer GmbH, personal fees from PUMA Biotechnology, personal fees and non-financial support from Roche Pharma AG, personal fees and non-financial support from Sanofi Aventis Deutschland GmbH, personal fees and non-financial support from TEVA Pharmaceuticals Ind Ltd, personal fees and non-financial support from Novartis, personal fees from Pierre Fabre, personal fees and non-financial support from Clovis Oncology, outside the submitted work. SB reports other from NantOmics, LLC, during the conduct of the study; other from NantOmics, LLC, outside the submitted work; and Employee of NantOmics, LLC with equity interest. AS reports grants from Celgene, grants from Roche, grants

from AbbVie, grants from Molecular Partner, personal fees from Roche, personal fees from AstraZeneca, personal fees from Celgene, personal fees from Roche, personal fees from Celgene, personal fees from Pfizer, personal fees from AstraZeneca, personal fees from Novartis, personal fees from MSD, personal fees from Tesaro, personal fees from Lilly, other from Roche, outside the submitted work. CJ reports personal fees from Roche, personal fees from Celgene, personal fees from Amgen, during the conduct of the study. TL reports non-financial support from Pharma Mar, non-financial support from Daiichi Sankyo, personal fees and non-financial support from MSD, personal fees from Amgen, personal fees and non-financial support from Pfizer, personal fees from Novartis, personal fees from Teva, personal fees from Tesaro, personal fees and non-financial support from Roche, personal fees and non-financial support from Clovis, non-financial support from Celgene, outside the submitted work.WDS reports grants from German Breast Group, during the conduct of the study; personal fees from AstraZeneca, outside the submitted work. VM reports grants and personal fees from Roche, during the conduct of the study; personal fees from Amgen, Astra Zeneca, Daiichi-Sankyo, Eisai, Pfizer, MSD, Novartis, Roche, Teva, Seattle Genetics and consultancy honoraria from Genomic Health, Hexal, Roche, Pierre Fabre, Amgen, ClinSol, Novartis, MSD, Daiichi-Sankyo, Eisai, Lilly, Tesaro, Nektar, personal fees from Genomic Health, Hexal, Roche, Pierre Fabre, Amgen, ClinSol, Novartis, MSD, Daiichi-Sankyo, Eisai, Lilly, Tesaro and Nektar, other from I Novartis, Roche, Seattle Genetics, Genentech, outside the submitted work. PSS is the CEO of NantOmics and reports non-financial support from NantOmics, outside the submitted work. MvM reports honoraria from Roche, Amgen, Genomic Health, Astra Zeneca, as well as travel support from Lilly and Novartis. PAF reports grants from Novartis, grants from BioNTech, personal fees from Novartis, personal fees from Roche, personal fees from Pfizer, personal fees from Celgene, personal fees from Daiichi-Sankyo, personal fees from AstraZeneca, personal fees from Merck Sharp & Dohme, personal fees from Eisai, personal fees from Puma, grants from Cepheid, personal fees from Lilly, personal fees from Seattle Genetics, during the conduct of the study. SR is the Chief Scientific Officer for NantOmics and reports non-financial support from NantOmics, outside the submitted work. SL reports grants and other from Celgene, grants and other from Roche, during the conduct of the study; grants and other from Abbvie, grants and other from Amgen, grants and other from AstraZeneca, grants and other from Novartis, grants and other from Pfizer, other from Seattle Genetics, other from PriME/Medscape, personal fees from Chugai, grants from Teva, grants from Vifor, grants and other from Daiichi-Sankyo, other from Lilly, other from Samsung, other from Eirgenix, other from BMS, other from Puma, other from MSD, grants from Immunomedics, outside the submitted work. In addition, SL has a patent EP14153692.0 pending. All other authors have declared no conflicts of interest.

REFERENCES

1. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. *Nature*. 2013;500(7463):415-421.
2. Nik-Zainal S, Davies H, Staaf J, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature*. 2016;534(7605):47-54.
3. Alexandrov LB, Kim J, Haradhvala NJ, et al. The repertoire of mutational signatures in human cancer. *Nature*. 2020;578(7793):94-101.
4. Harris RS. Cancer mutation signatures, DNA damage mechanisms, and potential clinical implications. *Genome Med*. 2013;5(9):87.
5. Hayward NK, Wilmott JS, Waddell N, et al. Whole-genome landscapes of major melanoma subtypes. *Nature*. 2017;545(7653):175-180.
6. Alexandrov LB, Ju YS, Haase K, et al. Mutational signatures associated with tobacco smoking in human cancer. *Science*. 2016;354(6312):618-622.
7. Alexandrov LB. Understanding the origins of human cancer. *Science*. 2015;350(6265):1175.
8. Alexandrov LB, Nik-Zainal S, Wedge DC, Campbell PJ, Stratton MR. Deciphering signatures of mutational processes operative in human cancer. *Cell Rep*. 2013;3(1):246-259.
9. Alexandrov LB, Stratton MR. Mutational signatures: the patterns of somatic mutations hidden in cancer genomes. *Curr Opin Genet Dev*. 2014;24:52-60.
10. Nik-Zainal S, Alexandrov LB, Wedge DC, et al. Mutational processes molding the genomes of 21 breast cancers. *Cell*. 2012;149(5):979-993.
11. Davies H, Glodzik D, Morganella S, et al. HRDetect is a predictor of BRCA1 and BRCA2 deficiency based on mutational signatures. *Nat Med*. 2017;23(4):517-525.
12. Davies H, Morganella S, Purdie CA, et al. Whole-genome sequencing reveals breast cancers with mismatch repair deficiency. *Cancer Res*. 2017;77(18):4755-4762.
13. Alexandrov LB, Jones PH, Wedge DC, et al. Clock-like mutational processes in human somatic cells. *Nat Genet*. 2015;47(12):1402-1407.
14. Denkert C, von Minckwitz G, Darb-Esfahani S, et al. Tumour-infiltrating lymphocytes and prognosis in different subtypes of breast cancer: a pooled analysis of 3771 patients treated with neoadjuvant therapy. *Lancet Oncol*. 2018;19(1):40-50.
15. Savas P, Salgado R, Denkert C, et al. Clinical relevance of host immunity in breast cancer: from TILs to the clinic. *Nat Rev Clin Oncol*. 2016;13(4):228-241.
16. Untch M, Jackisch C, Schneeweiss A, et al. NAB-paclitaxel improves disease-free survival in early breast cancer: GBG 69-GeparSepto. *J Clin Oncol*. 2019;37(25):2226-2234.
17. Untch M, Jackisch C, Schneeweiss A, et al. Nab-paclitaxel versus solvent-based paclitaxel in neoadjuvant chemotherapy for early breast cancer (GeparSepto-GBG 69): a randomised, phase 3 trial. *Lancet Oncol*. 2016;17(3):345-356.
18. Available at: https://cancer.sanger.ac.uk/cosmic/signatures_v2. Accessed June 12, 2020.
19. Pantelidou C, Sonzogni O, De Oliveria Taveira M, et al. PARP inhibitor efficacy depends on CD8(+) T-cell recruitment via intratumoral STING pathway activation in BRCA-deficient models of triple-negative breast cancer. *Cancer Discov*. 2019;9(6):722-737.
20. Ding L, Kim HJ, Wang Q, et al. PARP inhibition elicits STING-dependent antitumor immunity in Brca1-deficient ovarian cancer. *Cell Rep*. 2018;25(11):2972-2980.e5.
21. Venkatesan S, Rosenthal R, Kanu N, et al. Perspective: APOBEC mutagenesis in drug resistance and immune escape in HIV and cancer evolution. *Ann Oncol*. 2018;29(3):563-572.
22. Roberts SA, Lawrence MS, Klimczak LJ, et al. An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. *Nat Genet*. 2013;45(9):970-976.
23. Karn T, Jiang T, Hatzis C, et al. Association between genomic metrics and immune infiltration in triple-negative breast cancer. *JAMA Oncol*. 2017;3(12):1707-1711.

24. Safonov A, Jiang T, Bianchini G, et al. Immune gene expression is associated with genomic aberrations in breast cancer. *Cancer Res*. 2017;77(12):3317-3324.
25. Karn T, Denkert C, Weber KE, et al. Tumor mutational burden and immune infiltration as independent predictors of response to neoadjuvant immune checkpoint inhibition in early TNBC in GeparNuevo. *Ann Oncol*. 2020;31:1216-1222.
26. Nik-Zainal S, Morganella S. Mutational signatures in breast cancer: the problem at the DNA level. *Clin Cancer Res*. 2017;23(11):2617-2629.
27. Newton Y, Sedgewick AJ, Cisneros L, et al. Large scale, robust, and accurate whole transcriptome profiling from clinical formalin-fixed paraffin-embedded samples. *Sci Rep*. 2020;10(1):17597.