OncoImmunology

ISSN: (Print) 2162-402X (Online) Journal homepage: www.tandfonline.com/journals/koni20

Tumor mutational burden is a determinant of immune-mediated survival in breast cancer

Alexandra Thomas, Eric D. Routh, Ashok Pullikuth, Guangxu Jin, Jing Su, Jeff W. Chou, Katherine A. Hoadley, Cristin Print, Nick Knowlton, Michael A. Black, Sandra Demaria, Ena Wang, Davide Bedognetti, Wendell D. Jones, Gaurav A. Mehta, Michael L. Gatza, Charles M. Perou, David B. Page, Pierre Triozzi & Lance D. Miller

To cite this article: Alexandra Thomas, Eric D. Routh, Ashok Pullikuth, Guangxu Jin, Jing Su, Jeff W. Chou, Katherine A. Hoadley, Cristin Print, Nick Knowlton, Michael A. Black, Sandra Demaria, Ena Wang, Davide Bedognetti, Wendell D. Jones, Gaurav A. Mehta, Michael L. Gatza, Charles M. Perou, David B. Page, Pierre Triozzi & Lance D. Miller (2018) Tumor mutational burden is a determinant of immune-mediated survival in breast cancer, OncoImmunology, 7:10, e1490854, DOI: 10.1080/2162402X.2018.1490854

To link to this article: https://doi.org/10.1080/2162402X.2018.1490854

© 2018 The Author(s). Published by Taylor & Francis.

View supplementary material

Published online: 30 Jul 2018.

Submit your article to this journal

Article views: 6927

View related articles

View Crossmark data

Citing articles: 39 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=koni20

ONCOIMMUNOLOGY
2018, VOL. 7, NO. 10, e1490854 (12 pages)
https://doi.org/10.1080/2162402X.2018.1490854

Taylor & Francis
Taylor & Francis Group

ORIGINAL RESEARCH

OPEN ACCESS Check for updates

Tumor mutational burden is a determinant of immune-mediated survival in breast cancer

Alexandra Thomas[a,b], Eric D. Routh[c], Ashok Pullikuth[c], Guangxu Jin[b,c], Jing Su[d], Jeff W. Chou[b,e], Katherine A. Hoadley[f,g], Cristin Print [h], Nick Knowlton[h], Michael A. Black [i], Sandra Demaria[j], Ena Wang[k], Davide Bedognetti [k], Wendell D. Jones [l], Gaurav A. Mehta[m], Michael L. Gatza [m], Charles M. Perou[f,g], David B. Page[n], Pierre Triozzi [a,b], and Lance D. Miller [b,c]

[a]Section of Hematology and Oncology, Department of Internal Medicine, Wake Forest Baptist Medical Center, Winston-Salem, NC, USA; [b]Wake Forest Comprehensive Cancer Center, Winston-Salem, NC, USA; [c]Department of Cancer Biology, Wake Forest School of Medicine, Winston-Salem, NC, USA; [d]Division of Radiologic Sciences and Center for Bioinformatics and Systems Biology, Wake Forest School of Medicine, Medical Center Boulevard, Winston-Salem, NC, USA; [e]Department of Biostatistical Sciences, Wake Forest School of Medicine, Winston-Salem, NC, USA; [f]Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; [g]Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; [h]Department of Molecular Medicine and Pathology and Maurice Wilkins Institute, Faculty of Medical and Health Sciences, The University of Auckland, Auckland, New Zealand; [i]Department of Biochemistry, School of Biomedical Sciences, University of Otago, Dunedin, New Zealand; [j]Department of Radiation Oncology and Pathology and Laboratory Medicine, Weill Cornell Medical College, New York, NY, USA; [k]Department of Tumor Biology, Immunology and Therapy, Division of Translational Medicine, Sidra Medical and Research Center, Doha, Qatar; [l]EA Genomics, Division of Q Solutions, Morrisville, NC; [m]Department of Radiation Oncology, Rutgers Cancer Institute of New Jersey, New Brunswick, NJ, USA; [n]Department of Medicine, Providence Cancer Center, Earle A. Chiles Research Institute, Portland, OR, USA

ABSTRACT

Mounting evidence supports a role for the immune system in breast cancer outcomes. The ability to distinguish highly immunogenic tumors susceptible to anti-tumor immunity from weakly immunogenic or inherently immune-resistant tumors would guide development of therapeutic strategies in breast cancer. Genomic, transcriptomic and clinical data from The Cancer Genome Atlas (TCGA) and Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) breast cancer cohorts were used to examine statistical associations between tumor mutational burden (TMB) and the survival of patients whose tumors were assigned to previously-described prognostic immune subclasses reflecting favorable, weak or poor immune-infiltrate dispositions (FID, WID or PID, respectively). Tumor immune subclasses were associated with survival in patients with high TMB (TMB-Hi, $P < 0.001$) but not in those with low TMB (TMB-Lo, $P = 0.44$). This statistical relationship was confirmed in the METABRIC cohort (TMB-Hi, $P = 0.047$; TMB-Lo, $P = 0.39$), and also found to hold true in the more-indolent Luminal A tumor subtype (TMB-Hi, $P = 0.011$; TMB-Lo, $P = 0.91$). In TMB-Hi tumors, the FID subclass was associated with prolonged survival independent of tumor stage, molecular subtype, age and treatment. Copy number analysis revealed the reproducible, preferential amplification of chromosome 1q immune-regulatory genes in the PID immune subclass. These findings demonstrate a previously unappreciated role for TMB as a determinant of immune-mediated survival of breast cancer patients and identify candidate immune-regulatory mechanisms associated with immunologically cold tumors. Immune subtyping of breast cancers may offer opportunities for therapeutic stratification.

ARTICLE HISTORY
Received 12 March 2018
Revised 12 June 2018
Accepted 13 June 2018

KEYWORDS
breast cancer; mutational burden; immune subtypes; prognosis; survival

Introduction

The abundance of tumor infiltrating lymphocytes (TIL), a surrogate for activated anti-tumor immunity, is increasingly recognized as a positive prognostic and predictive marker in breast oncology.[1–5] Increasing tumor mutational burden (TMB) and the generation of neoepitopes have been associated with immunogenicity in several tumor types.[6,7] TMB is associated with clinical benefit to immune checkpoint blockade in patients with melanoma, lung, and colon cancer.-[8–11] The role of TMB in tumor immunogenicity is less clear in breast cancer. While TMB is higher in estrogen receptor (ER)-negative tumors compared with ER-positive tumors, consistent evidence that mutational burden itself significantly correlates with TIL levels is lacking.[12,13]

In recent years, the protective effects of endogenous anti-tumor immune responses in breast cancer have been elucidated by large-scale tumor expression profiling studies. These studies describe clusters of coordinately expressed genes that quantify the relative abundance and functions of tumor infiltrating immune cell populations and associate with patient outcomes.[14–18] In a recent report, we described three such prognostic immune signatures in breast tumors, each comprised of tens of genes with immune-specialized functions and expression patterns unique to immune cell populations.[19–21] These signatures, or metagenes, reflect the intratumoral

CONTACT Lance David Miller ldmiller@wakehealth.edu Department of Cancer Biology, Wake Forest Baptist Medical Center, Medical Center Blvd., Winston-Salem, NC 27157, USA

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/koni.
Supplemental data for this article can be accessed here.

© 2018 The Author(s). Published by Taylor & Francis.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

presence of different immune cell lineages: T cells/NK cells (T/NK metagene) comprised of genes with functional roles in cytotoxic T-cell and NK-cell activation, B cells/plasma cells (B/P metagene) marked by IgG antibody isotype-related genes, and Myeloid/Dendritic cells (M/D metagene) delineated by genes encoding myeloid-specific and MHC-class II antigen-presenting molecules. Multivariable analyses have shown that high metagene values (equating with abundant immune infiltrates) are significantly and independently associated with prolonged distant metastasis-free survival of breast cancer patients [19,21] as well as a greater likelihood of pathologic response to neoadjuvant chemotherapy.[20] These findings are consistent with functional roles for the metagene-associated leukocytes in protective anti-tumor immunity, namely cytolytic activity (T/NK), antigen presentation (M/D) and humoral response (B/P).

On this basis, we developed a breast tumor classification system that combines the prognostic attributes of the B/P, T/NK and M/D metagenes for partitioning tumors into one of three outcome-related immune-infiltrate subclasses (reviewed in eFigure 1 in Supplement 1). The tumor immune subclasses, FID, WID and PID, reflect favorable, weak and poor immune-infiltrate dispositions, respectively, that correspond with immunologically "hot/inflamed", "warm" and "cold/non-inflamed" tumor histologies.[21] In previous work we observed a strong prognostic association for the immune subclasses in some tumor types, but not others, indicating that effector immune infiltrates may impart protective anti-tumor immunity on a selective basis. We found that this selectivity depends on the proliferative capacity and intrinsic subtype of breast cancer, with tumors of high proliferative capacity and Luminal B (LumB), HER2-enriched (HER2E) or Basal-like subtype status defining a particularly immunogenic or "immune benefit-enabled" population whereby FID, WID and PID designations exhibit strong prognostic power.[21]

The proliferation rate and intrinsic subtype status of breast cancers have been correlated with TMB.[22,23] We hypothesized that TMB, as a quantifier of tumor antigenicity, may influence the prognostic power of the immune subclasses. We posited that a better understanding of the relationship between TMB and the prognostic potential of the immune subclasses could allow one to distinguish tumors inherently vulnerable to immune rejection from those that escape immune recognition by exclusion of immune cells and/or failure to elicit a protective immune response. In this report, we demonstrate that TMB discerns protective versus non-protective immune configurations in breast cancer.

Results

TMB determines immune-mediated survival of breast cancer patients

The TCGA Breast Cancer (BRCA) RNAseq expression data were used to classify 930 primary breast tumors according to the three immune subclasses (Figure 1A). A significant correlation between FID, WID and PID designations and decreasing histological estimates of presence of TIL was confirmed (eFigure 2 in Supplement 1). TMB was defined as the rate of nonsynonymous mutations per megabase (Mb) of sequenced DNA. The mutation rate for the BRCA cohort ranged from 0 to 115 mutations/Mb with a mean TMB of 1.63/Mb that was equivalent to the 80th percentile of the cohort's TMB distribution (Figure 1B). A comparison of the BRCA TMB distribution to that of other TCGA cancer cohorts is shown in eFigure 3 in Supplement 1. The mean TMB of the BRCA cohort was used as an initial threshold for assigning patients to low (below-mean) or high (above-mean) TMB categories (termed *TMB-Lo* and *TMB-Hi*, respectively) and verified as an acceptable threshold choice by alternative cutpoint testing (see Methods and eTable 1 in Supplement 1).

Patient overall survival (OS) was used to investigate the impact of TMB on the prognostic attributes of the immune subclasses. By Kaplan-Meier analysis, TMB was a strong determinant of the association between immune subclass and OS (Figure 1C). In TMB-Hi tumors, FID, WID and PID subclasses exhibited 5-year overall survival estimates of 100%, 76% and 60%, respectively. By contrast, the prognostic connotation for FID, WID and PID was absent in patients with TMB-Lo tumors, with the three subclasses displaying 5-year survival estimates between 81–86%. To test the reproducibility of this observation in an independent cohort, we utilized the METABRIC study consisting of > 2,000 breast tumors.[24] While the TCGA and METABRIC platforms for expression profiling and mutation analysis differed, the derivation of immune subclasses and TMB were determined to be moderately comparable between the two data sets (see Methods and eFigures 4–5 in Supplement 1). For computing TMB in the METABRIC dataset, the number of mutations/tumor (i.e., mutation count), but not the mutation rate, was available. Thus, we analyzed immune subclass-survival associations across a range of mutation count thresholds for calling tumors low or high TMB. Since the TCGA threshold was equivalent to the 80th percentile of the cohort's TMB distribution, we examined mutation-count thresholds in the METABRIC dataset that approximated the 70th, 80th and 90th percentiles of TMB (eFigure 6 in Supplement 1). At thresholds based on the 80th and 90th percentiles, the immune subclasses showed statistically significant survival differences in TMB-Hi tumors (≥ 80th percentile, $P = 0.047$; ≥ 90th percentile, $P = 0.015$), but not in TMB-low tumors (< 80th percentile, $P = 0.39$; < 90th percentile, $P = 0.41$), demonstrating the reproducibility of the effect of TMB on the prognostic power of the immune subclasses. Additionally, we examined TMB-Hi versus TMB-Lo survival differences within each immune subclass (eFigure 7 in Supplement 1). In the FID subclass, TMB-Hi tumors associated with significantly better OS, while in the PID subclass, this association was reversed, with TMB-Hi tumors associating with significantly worse OS. This observation is consistent with an opposing prognostic relationship between TMB and anti-tumor immunity, where increasing mutation rate associates with a more aggressive cancer phenotype in immunologically cold tumors (PID), yet provides an antigen-mediated immunological advantage in immune-inflamed tumors (FID) where immune-mediated tumor control predominates.

To better understand the immunological composition of the breast tumor immune subclasses, we used a genetic algorithm

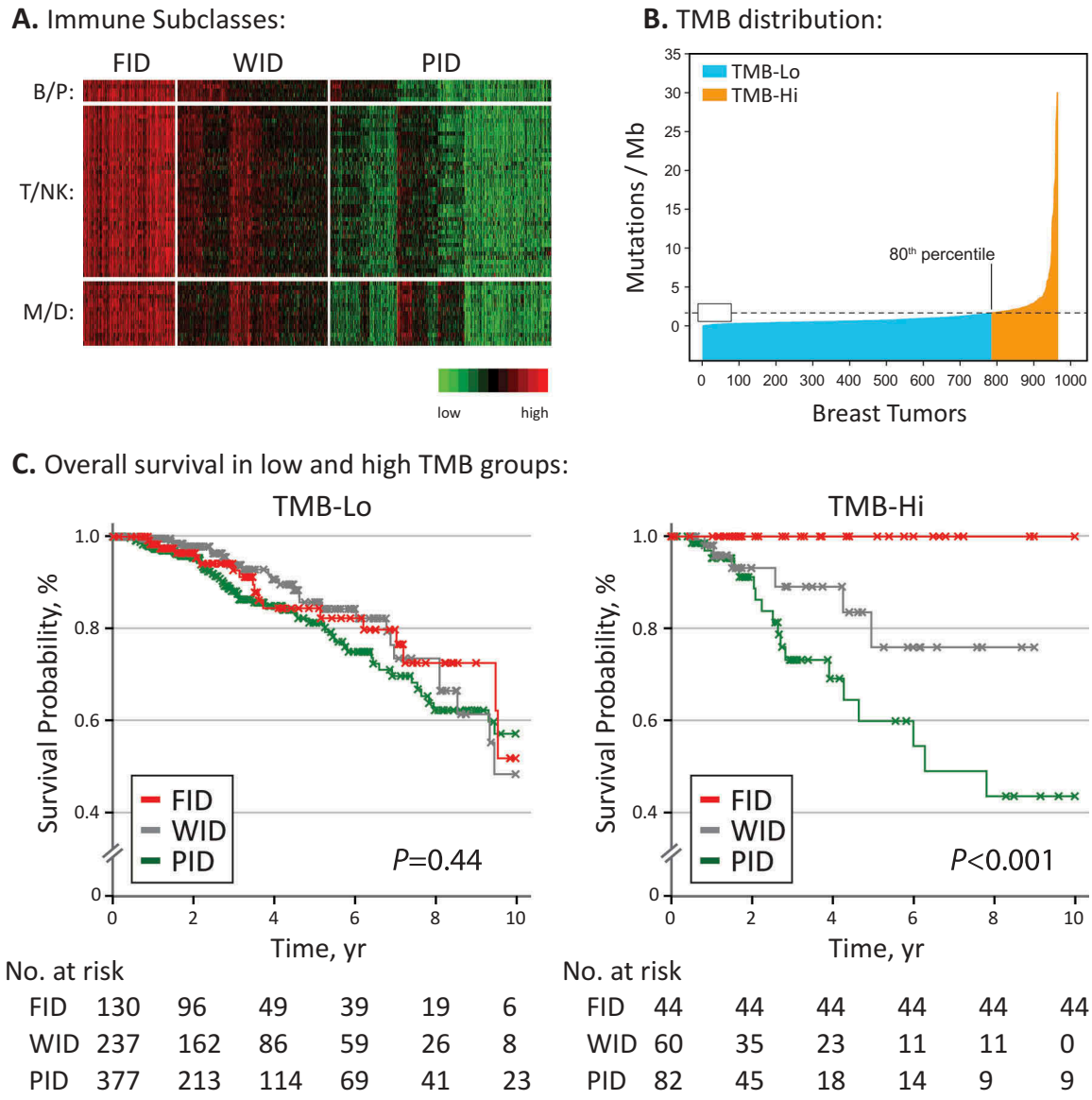

Figure 1. Tumor immune subclasses and survival associations by TMB status. (A) Derivation of the immune subclasses (FID, WID and PID) is illustrated in the gene expression heat map of tumors (columns) and genes comprising the immune metagenes (rows). Relative log2 expression levels are indicated by the color key (bottom right). **(B)** The TMB (mutation rate) distribution for the TCGA cohort is shown; tumors ordered by ascending mutation rate. The mean TMB (1.63/Mb) is marked by a dashed line and coincides with the 80th percentile; TMB-Lo and TMB-Hi tumors delineated by color. Mutation rates for three tumors (55/Mb, 92/Mb and 115/Mb) are omitted for scale. **(C)** Kaplan-Meier plots of overall survival (OS) by immune subclass according to TMB-Lo (left panel) and TMB-Hi (right panel). Log-rank test *P*-values are shown.

for estimating the abundance of different immune cell types within tumors.[25] We profiled the relative enrichment of 22 immune cell types across FID, WID and PID in TCGA TMB-Lo and TMB-Hi tumors (Figure 2). Irrespective of TMB, the majority of immune cell types showed differential enrichment between FID and PID subclasses. Notably, the greatest differences were observed among those annotated as *T.cells.CD8, T.cells.CD4.memory.resting, NK.cells.activated*, and *T.cells.gamma.delta* – all of which showed marked enrichment in FID tumors relative to WID and PID tumors. *Mast.cells.activated, Mast.cells.resting* and *Eosinophils* showed the most significant associations in the opposing direction, with enrichment in PID tumors relative to WID and FID. In TMB-Hi tumors, where FID and PID survival rates markedly differ, with the exception of *B.cells.naive* and *Macrophages.M0* all immune cell types showed significant differences between FID and PID subclasses. Within the immune subclasses, the most significant differences between TMB-Hi and TMB-Lo tumors was observed in the FID subclass, where TMB-Hi showed enrichment for *Macrohages.M1* and *T.cells.CD4.memory.activated* ($p < 0.001$, both comparisons). These findings suggest that robust immunological differences involving many immune cell types distinguish the immune subclasses, and FID and PID in particular, with minimal differences observed as a function of TMB.

Molecular subtypes: associations with TMB and immune subclasses

As mutational and immunological features of breast tumors vary across molecular subtypes, we performed subtype-specific analyses of TMB and immune subclass composition. Basal-like and Human Epidermal Growth Factor Receptor 2-

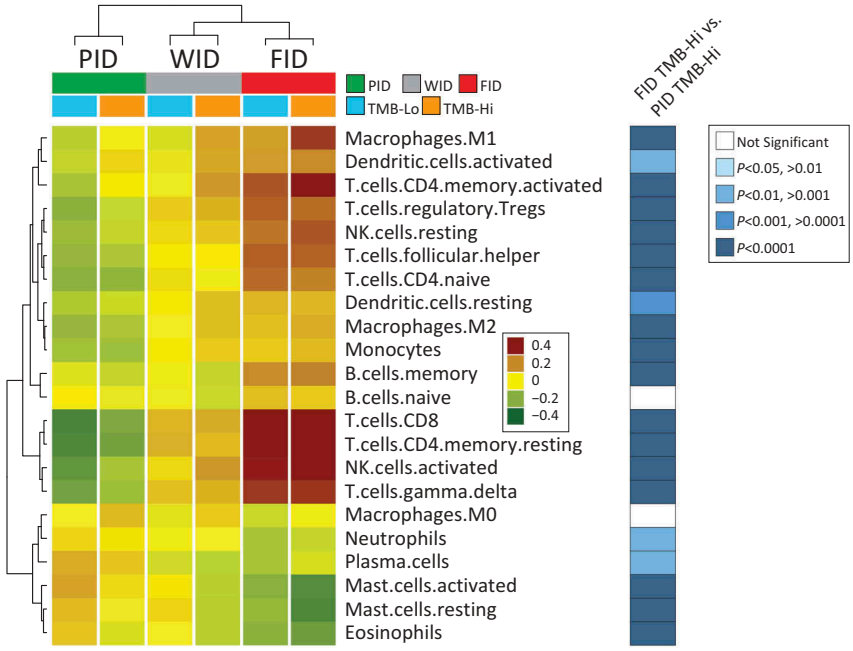

Figure 2. Transcriptomic features of anti-tumor immunity in breast cancer immune subclasses. A leukocyte gene signature matrix of 22 immune cell types was used as input for gene set variation analysis (GSVA). Mean enrichment scores for the immune subclasses, stratified by TMB status, are shown in the heat map; red indicates increasing relative enrichment, green indicates decreasing relative enrichment. Significant differences in enrichment scores between FID and PID subclasses in TMB-Hi tumors are shown (Mann-Whitney U test).

Enriched (HER2E) breast cancers had the highest TMB (TMB-Hi fraction, 37% and 38%, respectively) while Luminal A (LumA) tumors had the lowest TMB, with only 11.5% in the TMB-Hi fraction (Figure 3A) consistent with previous findings.[22] Among all tumors designated TMB-Hi, the FID subclass represented 44 women who exhibited a high OS rate of 100% at 10 years (Figure 1C); 66% of this excellent-outcome subclass was comprised of the more clinically aggressive subtypes: Basal-like, HER2E and Luminal B (LumB) (Figure 3B). Among the four major subtypes, LumA tumors comprised the majority of the TCGA cohort (53%; n = 495), with 88.5% of LumA tumors classifying with the TMB-Lo group. Thus, we examined the effect of the LumA subtype on the overall prognostic stratification of the immune subclasses (eFigure 8 in Supplement 1). In both LumA and non-LumA tumors, FID, WID and PID stratified patients to significantly different survival outcomes in the context of TMB-Hi, but not TMB-Lo. Strikingly, LumA tumors with high TMB status and belonging to the PID immune subclass displayed a poor overall survival rate of 42% at 5 years (eFigure 8D in Supplement 1). In TNBC tumors, high TMB associated with significantly better OS (as compared to TMB-Lo) in FID, but not WID or PID subclasses (eFigure 9A-C in Supplement 1). However, while the FID subclass trended

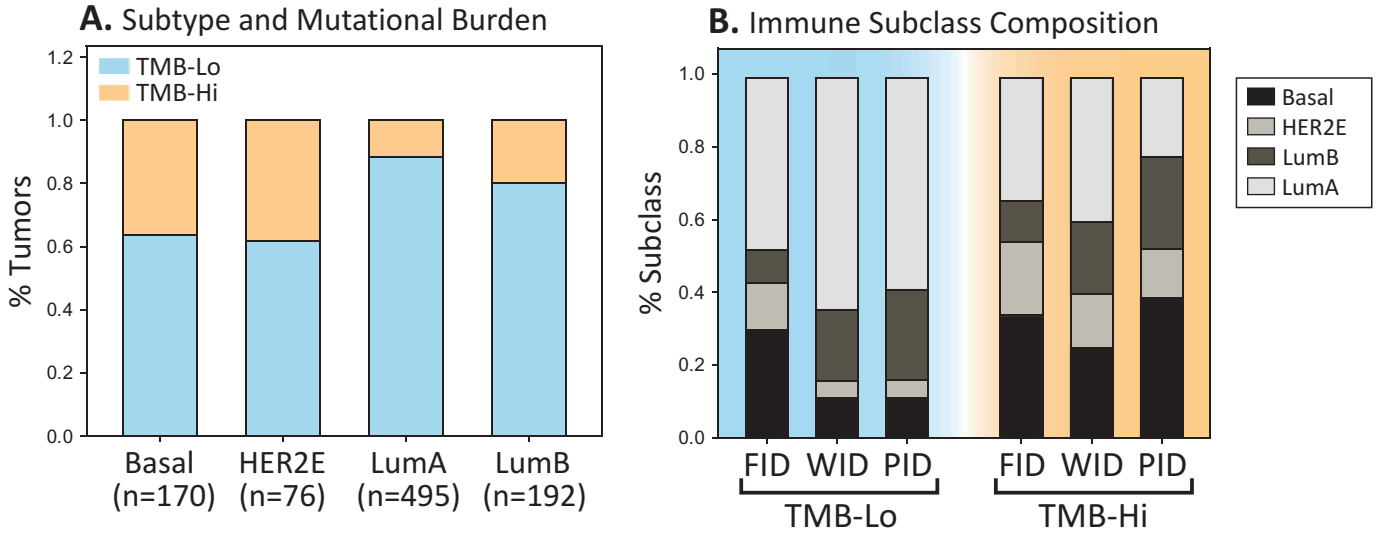

Figure 3. Molecular subtypes of breast cancer in low and high TMB and immune subclasses. (A) The fraction of TMB-Lo and TMB-Hi tumors comprising each molecular subtype is shown. **(B)** The fraction of molecular subtypes comprising each immune subclass according to low and high TMB is shown.

toward better OS (as compared to WID and PID) in the TNBC TMB-Hi tumors, it did not achieve significance likely owing in part to limited sample size (n = 69) and scant events (n = 6 deaths) (eFigure 9D,E in Supplement 1).

Comparative analysis of prognostic markers

To investigate the clinical relevance of the immune subclasses and TMB, multivariable Cox regression was used to model associations between clinical and molecular variables and survival time (Table 1). In the full cohort, variables that remained significant in the model included mutation rate, tumor stage (all stages vs. stage I), HER2E subtype (vs. LumA), unknown treatment (vs. all treatment groups) and patient age; all of which were associated with poor survival with the exception of treatments (eTable 2 in Supplement 1). In TMB-Lo tumors, stage (all stages vs. stage I), age and unknown treatment (vs. endocrine monotherapy) remained significant in the full model. In TMB-Hi tumors, the immune subclasses remained significant in the presence of all variables, together with stage III (vs. stage I) and unknown treatment (vs. endocrine monotherapy). Together, these observations demonstrate that the immune subclasses contribute additive prognostic power beyond that of stage, age, treatment and molecular subtypes in TMB-Hi tumors.

Genetic alterations underlie immune-survival associations

Somatic mutations can activate or suppress anti-tumor immunity. We therefore considered the possibility that genomic mutations may predispose tumors to different immune subclasses, or alternatively, arise as a consequence of genetic selection and clonal outgrowth driven by certain immune phenotypes. We examined the TCGA data set for recurrent nonsynonymous gene mutations and chromosomal copy number alterations (CNAs) that showed statistical evidence for immune subclass-specific enrichment. After false discovery rate (FDR) correction, only a single gene, TP53, was found to be significantly mutated in an immune subclass-dependent manner (FDR adjusted $P = 0.02$) with recurrent TP53 mutations observed in 46% of FID, 26% of WID and 29% of PID tumors, overall. No recurrently mutated genes showed significant enrichment after stratification to TMB-Hi and TMB-Lo. By contrast, several thousand genes exhibited chromosomal copy number changes significant after FDR adjustment, the majority of which reflected relatively small subclass-dependent differences. We therefore focused the analysis on the TMB-Hi tumors, where many genes displayed large copy number differences among the immune subclasses. We identified 133 genes amplified with significantly greater frequency in TMB-Hi PID tumors, and 6 genes deleted with significantly greater frequency in TMB-Hi PID tumors (Figure 4A, eWorksheet 2 in Supplement 2). Gene Ontology enrichment analysis of the genes amplified in PID tumors identified 5 significantly enriched gene ontologies inclusive of 26 genes: *Cytokine activity* ($P = 0.003$), *Interleukin-10 conserved site* ($P < 0.001$), *Selectin superfamily* ($P = 0.009$), *Extracellular space* ($P = 0.03$) and *Flavin monooxygenase* ($P < 0.001$) (eTable 3 in Supplement 1). A number of these genes are involved in immune regulatory processes and are highlighted in Figure 4A. Prominent among them was the immunosuppressive cytokine, IL10, which showed high-level amplification in 18% of PID, 14% of WID and 2% of FID tumors (Figure 4B). Similar amplification rates were observed for the

Table 1. Multivariable Cox regression for associations with overall survival.

	TMB-Lo Tumors (n = 744)			TMB-Hi Tumors (n = 186)		
Variable	Patients, No. (%)	HR (95% CI)	*P* value[a]	Patients, No. (%)	HR (95% CI)	*P* value
Immune Subclass[a]					2.67 (1.46–4.86)	0.001
FID	130 (17)	1 [Reference]		44 (24)	NA	
WID	237 (32)	0.91 (0.47–1.75)	0.77	60 (32)	NA	
PID	377 (51)	0.91 (0.51–1.63)	0.75	82 (44)	NA	
AJCC Stage						
I	132 (18)	1 [Reference]		25 (13)	1 [Reference]	
II	416 (56)	3.79 (1.59–9.01)	0.003	114 (61)	2.76 (0.50–15.1)	0.24
III	167 (22)	7.57 (3.06–18.7)	< 0.001	40 (22)	9.30 (1.70–51.0)	0.01
IV	14 (2)	15.7 (5.64–43.9)	< .001	1 (< 1)	NA[b]	NA
NA	15 (2)	6.22 (2.00–19.4)	0.002	6 (3)	1.06 (0.08–14.0)	0.96
PAM50 Subtype						
LumA	435 (58)	1 [Reference]		57 (31)	1 [Reference]	
LumB	154 (21)	1.29 (0.66–2.54)	0.45	38 (20)	0.38 (0.06–2.42)	0.30
HER2E	47 (6)	1.60 (0.66–3.87)	0.29	29 (16)	2.36 (0.50–11.3)	0.28
Basal-like	108 (15)	1.32 (0.52–3.32)	0.55	62 (33)	1.10 (0.13–9.31)	0.93
PAM50 Proliferation Score, median (IQR)	−0.2 (−0.6–0.1)	1.30 (0.57–2.99)	0.53	0.1 (−0.2–0.3)	1.13 (0.20–10.9)	0.92
Treatment						
Endo	147 (20)	1 [Reference]		25 (13)	1 [Reference]	
Chemo	155 (21)	1.54 (0.60–4.00)	0.37	50 (27)	0.14 (0.01–2.04)	0.15
Endo + Chemo	238 (32)	1.05 (0.43–2.56)	0.92	46 (25)	0.82 (0.11–5.98)	0.85
NA	204 (27)	3.04 (1.38–6.69)	0.006	65 (35)	6.29 (1.19–33.1)	0.03
Age, median (IQR), y	58 (48–67)	1.03 (1.01–1.05)	< 0.001	59 (49–68)	1.03 (0.99–1.06)	0.15

Abbreviations: CI, confidence interval; NA, not available; IQR, interquartile range; Endo, endocrine therapy; Chemo, chemotherapy

PAM50 Proliferation Score, Age and Immune Subclass (TMB-Hi cases) were entered as continuous variables, all others were entered as categorical variables

[a] Immune Subclass as a categorical variable resulted in a degenerate estimate (failed to converge) in the TMB-Hi population due to no deaths in the FID group, thus, we used Nelder-Mead optimization to derive a continuous score for immune subclass (see eMethods)

[b] not applicable, only 1 stage IV tumor in TMB-Hi population

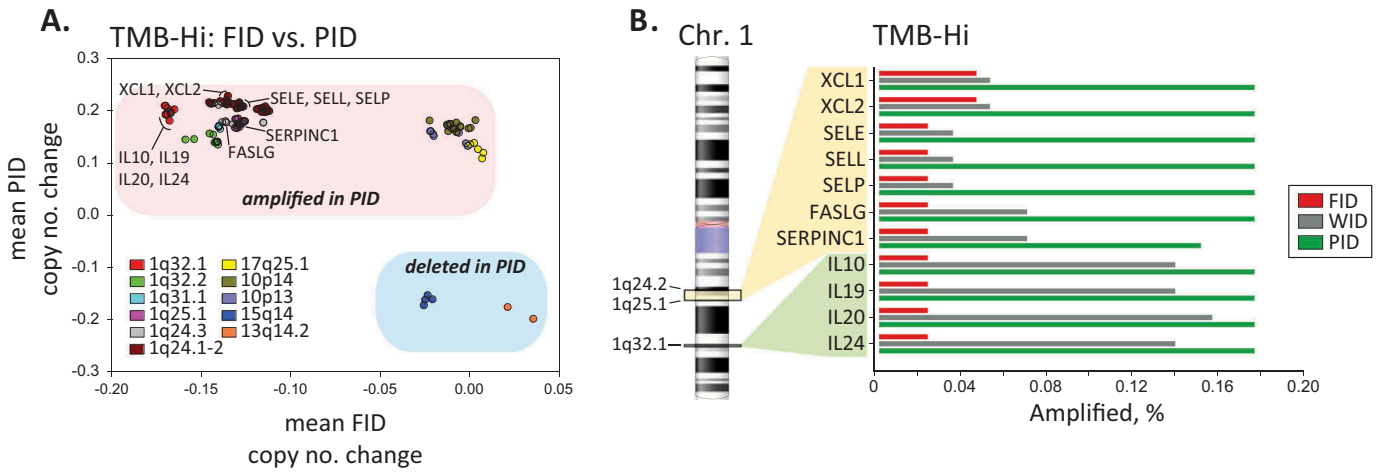

Figure 4. Copy number gains and losses associated with immune subclasses. Group means of gene-level copy number values were compared among FID, WID and PID subclasses via analysis of variance. **(A)** Genes with significant copy number variation (FDR adjusted $P < 0.1$) and that showed > 10% variation in high-level threshold amplification or deletion events (by GISTIC) between TMB-Hi FID and PID tumors are shown. 139 genes spanning 11 chromosomal loci are shown plotted as a function of the change in mean copy number between PID and all other tumors (y-axis) and FID and all other tumors (x-axis). 133 and 6 genes were identified as amplified in PID and deleted in PID, respectively. No genes were observed to be significantly amplified or deleted in FID relative to WID or PID. Select immune-related genes identified by significantly enriched ontology terms are highlighted. **(B)** The frequencies (x-axis) of GISTIC high-level threshold amplification events for select genes on chromosome 1 are shown for TMB-Hi FID, WID and PID subclasses.

genes IL19, IL20 and IL24, located adjacent to IL10 at chromosome 1q32.1. At locus 1q24.2-1q25.1, PID-enriched amplification was observed for genes encoding T-cell specific chemokines (XCL1, XCL2), leukocyte tethering molecules (SELE, SELL, SELP), a promotor of T and B-cell activation induced-cell death (FASLG) and a granzyme inhibitor (SERPINC1). For these genes, high-level amplification events ranged from 15–18% in PID, 4–7% in WID and 2–5% in FID tumors (Figure 4B). Notably, the significant enrichment for high-level amplification of these chromosome 1q immune genes in PID tumors was reproducible in the METABRIC cohort (eTable 4 in Supplement 1). In the METABRIC cohort, high-level amplification of the genes occurred at significantly higher frequencies in both TMB-Hi PID tumors and TMB-Lo PID tumors (relative to FID and WID), though the association in TMB-Lo tumors of the TCGA cohort was not evident (eTable 4 in Supplement 1).

Discussion

In this study we provide first evidence of the significant link between TMB and survival-associated protective immunity in breast cancer. Using the TCGA and METABRIC cohorts we demonstrate that approximately one-fifth of breast cancers, with a relatively high TMB (TMB-Hi), harbor immunologically-relevant numbers of mutations. The FID immune subclass, marked by elevated effector immune infiltrates (Figure 2), encompassed 20% of TCGA breast tumors overall, with 25% designated as TMB-Hi and 75% as TMB-Lo. In TMB-Hi, patients with FID tumors survived longer than those with WID or PID tumors, consistent with an intrinsic vulnerability of TMB-Hi FID tumors to antigen-stimulated immune control of tumor progression. Whether or not these tumors, given their high TMB and elevated effector immune infiltrates, would be more susceptible to immune checkpoint blockade is a point of clinical interest that warrants further investigation. Patients with TMB-Lo FID tumors, in contrast to those with TMB-Hi FID, did not experience a protective survival benefit as compared to WID and PID, suggesting a tolerant or tumor-promoting immune infiltrate that may facilitate immune evasion.[26] PID tumors, characterized by low level effector cell infiltrates, comprised the largest immune subclass (49% of tumors). This subclass comprised 44% of the TMB-Hi group where it was associated with worse survival consistent with a poor tumor-mediated immune response, despite high TMB. Whether or not these tumors would prove more resistant to checkpoint inhibition is a question of clinical merit.

Using the prognostic significance of the immune subclasses as a guide, we observed that 1.63 nonsynonymous mutations/Mb constituted a suitable threshold for discerning TMB low versus high breast tumors (eTable 1 in Supplement 1). In relation to other cancers, however, this TMB cutpoint is arguably low. In a recent report showing that high TMB predicts immunotherapy response in diverse cancers, where TMB was defined as the rate of nonsynonymous mut/Mb, the TMB categories were assigned as low (1–5 mut/Mb), intermediate (6–19 mut/Mb) and high (≥ 20 mut/Mb), and corresponded to approximately 50%, 40% and 10% of the sampled population, respectively.[8] Accordingly, this study would have assigned 83% of the TMB-Hi breast tumors (as defined by our TMB cutpoint) to their low TMB category, and only 7 breast tumors in the entire TCGA cohort would have qualified for their definition of high TMB. To better understand this, we used our TMB computation method to compare the TMB distribution of BRCA to that of other TCGA cancer types and immunotherapy-treated cancers, in particular (eFigure 3 in Supplement 1). Where a TMB of 1.63 mut/Mb equated with the 80th percentile of the BRCA cohort, this same TMB equated with only the 14th percentile of melanomas (SKCM) and the 19th percentiles of both lung adenocarcinomas (LUAD) and head/neck cancers (HNSC). Conversely, the

median TMBs for SKCM, LUAD and HNSC equated with the BRCA 98th, 97th and 92nd percentiles, respectively. Thus, as compared to more hypermutated cancers, even the most heavily mutated breast tumors can be said to exhibit a relatively low TMB. A simple explanation then as to why our low/high TMB assignment of breast tumors remained a strong determinant of immune subclass prognostic significance, may be that immune-mediated tumor control as it relates to TMB, varies from one cancer type to another, or varies in a manner that depends on the prevalence of other immuno-modulatory factors. In the case of breast cancer, the combination of subtle increments in TMB together with increased effector cell infiltrates (reflected by the FID subclass) may sum to a more protective immunogenicity than that afforded by either factor alone. Whether this prognostic synergy, which we demonstrate in both the TCGA and METABRIC cohorts, is unique to breast cancer or applicable to other cancer types remains to be seen. As a continuous variable in the BRCA cohort, TMB alone was found to associate with poor survival, consistent with previous observations in ER-positive breast cancer.[23] In TMB-Hi but not TMB-Lo tumors, the immune subclasses contributed significant additive prognostic information independent of AJCC stage, molecular subtype and other variables. These findings support the concept that integration of molecular and genomic factors that characterize immunological states of cancer will add value in risk-stratified treatment decision models.

The role of the immune system in the more indolent LumA breast cancers is poorly understood. While immune infiltrates have been associated with better survival in ER-negative and HER2-positive disease,[27,28] this relationship is less clear in ER-positive/HER2-negative breast cancers, where high TIL was recently reported to be associated with shorter OS in patients treated with neoadjuvant chemotherapy.[29] Here, we identified a rare population of high TMB LumA breast cancers (12% of LumA tumors) within which the PID immune subclass predicted poor survival. To our knowledge this is the first report of clinically-relevant immune subtypes of LumA breast cancer.

In other solid cancers, TMB as a reflection of neoantigen load is associated with cytolytic activity[26] and efficacy of immune therapies.[9,10,30–32] Although associated with efficacy, TMB has not yet been deemed sufficiently predictive to be of clinical use. Many tumors with high TMB are not responsive to immune checkpoint inhibitors. Some tumors with low TMB respond completely. Response to immunotherapy has also been associated with intratumoral gene-based markers of T cell activation and antigen presentation-[33,34] but such markers have not been introduced to practice. Triple negative breast cancer (TNBC), which comprises the majority of Basal-like tumors, is characterized by relatively high TMB.[35] The objective response rate in a recently reported phase Ib clinical trial of programmed-cell death protein 1 (PD-1) immune checkpoint inhibition in heavily pre-treated TNBC patients with PD-L1 expressing tumors was only 18.5%.[36] In our study, 36% of Basal-like tumors were classified as TMB-Hi, and 24% of these tumors (9% overall) belonged to the FID subclass, suggesting by extrapolation that only a relatively small fraction of these tumors may be particularly immunogenic. In TNBC tumors, we observed a protective effect for TMB-Hi in tumors of the FID immune subclass, but not in those of the WID and PID immune subclasses, though this sub-analysis was potentially limited by small sample size and sparsity of survival events. Whether this subgroup of TMB-Hi FID TNBC shows an enhanced responsiveness to immune checkpoint blockade is the focus of ongoing studies. Of note, mismatch repair deficiency which drives response to immune therapy in colon cancer and other cancers, is a relatively rare phenomenon in TNBC and other breast cancers.[37,38]

In a recent report by Pusztai and colleagues,[39] a lymphocyte-rich tumor subgroup of TNBC with good prognosis was shown to associate with reduced clonal heterogeneity and a reduced distribution of mutation count (compared to a poor prognosis TNBC subgroup). This finding suggests the selective culling of immunogenic TNBC cells by immune surveillance. To the extent that this phenomenon may be applicable to other subtypes, it is possible that tumors we defined as TMB-Hi and of the FID immune subclass may equate with reduced clonality but concurrent high mutational diversity within surviving clones. Thus, greater antigenic diversity could render these tumors prone to a continual immune sculpting that controls disease spread and mediates prolonged patient survival in TMB-Hi FID tumors.

Intriguingly, we identified significant and reproducible copy number gains in PID tumors involving genes on chromosome 1q associated with immune regulation. While amplification events at the gene-dense 1q locus occur in 50% or more of breast tumors,[40] "driver" genes with immune regulatory functions have not been proposed. In our analysis, amplified immune genes were identified by the significant enrichment of immune-related ontology terms. Interleukin-10 (encoded by the IL10 gene) is an immunosuppressive cytokine known to inhibit tumor-specific type 1 immune responses.[41] Fas ligand (FASLG) promotes activation-induced cell death (AICD) of activated T and B cells[42] and is believed to confer immune privilege to tumors by inducing apoptosis in tumor infiltrating lymphocytes.[43] Serpin Family C Member 1 (SERPINC1), also known as Antithrombin III, is a known inhibitor of T cell-derived Granzyme A.[44,45] The reproducible PID-enriched amplification of these genes suggests the possibility that amplification of one or more of these genes, alone or in combination, may play a functional role in immune escape that contributes to the immunologically cold phenotype of the PID immune subclass.

The findings presented here illuminate protective and non-protective immunological configurations of breast cancer with clinical implications. The results suggest that the mutational burden of breast cancer is a tumor-intrinsic determinant of immune-mediated patient survival, applicable even to Luminal A breast cancer. PID tumors that lack immune involvement and are associated with poor prognosis despite a high mutational burden, may be driven, in part, by immunosuppressive amplification events. A formal accounting of the genomic alterations and molecular pathways that underlie these immunological configurations could shed light on mechanisms of tumor immune escape

and reveal new opportunities for immunotherapeutic targeting.

Methods

Genomic data acquisition, annotation and sample selection

Transcriptomic, genomic and clinical data from The Cancer Genome Atlas (TCGA) and Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) breast cancer data sets were utilized in this study. *TCGA BRCA*: Clinical data, level 3-processed RNAseq, exome sequencing and Affymetrix (Human SNP Array 6.0) copy number data sets were downloaded from the Broad Genome Data Analysis Center's FireBrowse website (http://firebrowse.org/; TCGA data version 2016_01_28).[46] RNAseq data ('rnaseqv2') were generated via MapSplice alignment and RSEM quantitation. Affymetrix SNP6 copy number data were processed, normalized and segmented via the Broad's copy number inference pipeline described at (http://archive.broadinstitute.org/cancer/cga/sites/default/files/publications/SNP6_pipeline_application_notes_prerelease.pdf). Raw copy number data were further processed by GISTIC (v2.0.22)[47] to low- and high-level copy number thresholds. Exome sequence data were processed for somatic mutations by the Broad's Oncotator application (http://portals.broadinstitute.org/oncotator/) output as mutation annotation format (MAF) files. The 1,100 BRCA cases with RNAseq data were filtered to exclude male and gender "unknown" samples (n = 13), metastatic tissue samples (n = 7) and one errant skin cancer sample yielding 1,079 female primary breast tumor samples for downstream expression analysis. PAM50 intrinsic molecular subtype assignments and PAM50 proliferation scores were computed as previously described,[48,49] and are available in eWorksheet 1 in Supplement 2. The subtypes segregate as follows: Basal-like (n = 189), HER2E (n = 82), LumA (n = 559), LumB (n = 209) and Normal-like (n = 40). Tumor samples designated as Normal-like were omitted from downstream analyses due to evidence that they reflect normal tissue contamination.[50] The rate of nonsynonymous gene mutations as computed by the Broad GDAC using the MutSigCV algorithm (v0.9) [51] was downloaded from the FireBrowse website via the section on '*Distribution of Mutation Counts, Coverage, and Mutation Rates Across Samples*'. Mutation rate was available for 965 of the 1,079 samples analyzed by RNAseq. Patient overall survival data was available for 1,077 of the RNAseq-analyzed cases with average 10-year follow-up time of 3.18 years and 3.80 years for censored (time to last follow up) and uncensored (time to death) cases, respectively. In aggregate, 930 evaluable cases had corresponding RNAseq data, exome sequencing/mutation rate data and overall survival data (eWorksheet 1 in Supplement 2). Cases diagnosed clinically as TNBC were annotated according to Karn, et. al.[39] *METABRIC*: Normalized expression microarray and copy number data for 1,980 primary invasive breast cancer cases were obtained by Wake Forest IRB approval and controlled access from the European Genome-phenome Archive (https://www.ebi.ac.uk/ega/studies/EGAS00000000083). Mutation counts and updated clinical data were obtained from Pereira and colleagues.[52] GISTIC-processed copy number data were obtained from cBioPortal (http://www.cbioportal.org/).[53,54] In aggregate, 1,895 evaluable cases had corresponding gene expression, mutation count and overall survival data. Gene expression profiles were generated using the Illumina_Human_WG-v3 array platform[24] and normalized by quantile normalization with linear modeling batch correction, as described elsewhere.[55] Copy number levels were generated on the Affymetrix SNP Array 6.0 and normalized using the supervised normalization of microarrays (SNM) framework[56] as described by[55] and also using DNAcopy[54,57] to define low- and high-level copy number thresholds. A 173-gene exome sequencing panel was used to identify somatic gene mutations and generate measures of mutational burden (gene count).[52]

Immune subclass classification. The immune metagene model (IMM) is a survival-based prognostic classification system for the objective assignment of patients to the immune subclasses FID, WID or PID (reflecting favorable, weak or poor immunogenic dispositions, respectively) as previously described.[19,21] Sample classification relies on the relative expression levels (metagene scores) of the immune gene signatures (metagenes) which reflect the intratumoral presence of three immune cell types: B cells/plasma cells (B/P metagene) marked by IgG antibody isotype-related genes; T cells/NK cells (T/NK metagene) comprised mainly of genes with functional roles in Cytotoxic T-cell activation/helper T-cell type 1 (Th1) signaling, and Myeloid/Dendritic cells (M/D metagene) delineated by genes encoding myeloid-specific and MHC-class II antigen-presenting molecules. The genes comprising each metagene are defined in Additional File 6 of Nagalla and coworkers[19] and were mapped to the TCGA and METABRIC data sets by gene symbol. Metagene scores were computed for each tumor as the average of the $\log_2$-transformed normalized expression for the genes comprising each metagene. Then, for each metagene, scores were binned into lower, intermediate and upper tertiles based on the distribution of scores for each metagene across the entire breast tumor population. Tumors were assigned to immune subclasses according to the following tertile configurations: 1) FID: having metagene scores in the upper tertile of all three metagenes, simultaneously; 2) PID: having metagene scores in the lower tertile of one or more metagenes; and 3) WID: having metagene scores of both intermediate and upper metagene tertiles (but not classifying as FID or PID).[19] Immune subclass assignments are provided in eWorksheet 1 in Supplement 2.

Comparison of TCGA and METABRIC immune metagenes

The classification of tumors to immune subclasses FID, WID and PID depends on immune metagene tertile assignments and the relative scale of metagene values. The immune metagenes were originally defined according to the Affymetrix U133A expression profiling platform using a breast cancer meta-dataset (termed "MC1") of 1,954 tumor expression

profiles [19], then adapted to the TCGA RNAseq data set and the METABRIC cohort Illumina BeadChip data set. We compared metagene distributions among the 3 datasets (eFigure 4 in Supplement 1). Comparing each dataset across metagenes and tertile by tertile, we observed that the scale ranges of the metagenes significantly differed by dataset/expression profiling platform. These differences were associated with a common trend, with the TCGA RNAseq data set consistently showing the largest range of metagenes values, and the METABRIC BeadChip data set consistently showing the smallest range of values. This difference could reflect a greater precision for the TCGA data set to categorize tumors into tertiles (ie, due to greater dynamic range), relative to MC1 and METABRIC. A possible consequence is the more accurate classification of TCGA tumors to metagene tertiles, and a diminished accuracy for classifying METABRIC tumors to metagene tertiles and hence, immune subclasses. Thus, we conclude that this phenomenon could result in the diminished ability to observe a significant relationship between patient survival rates and immune subclasses in the METABRIC cohort, necessitating empirical assessment as performed in this study.

TCGA and METABRIC mutational burden estimates

Whole exome sequencing data from the TCGA breast cancer cohort was analyzed by the MutSigCV algorithm [51] to generate TMB estimates. TMB was defines as the rate of somatic nonsynonymous mutations per megabase of DNA sequenced. In the TCGA cohort, TMB values ranged from 0 to 115 mutations/Mb, and the mean TMB (1.63 mutations/Mb) was used as an initial threshold for assigning patients to low (below-mean, *TMB-Lo*) or high (above-mean, *TMB-Hi*) TMB categories. To determine the appropriateness of using the mean mutation rate threshold for discerning TMB-Lo from TMB-Hi, we considered the use of alternative TMB cutpoints. The mean TMB was found to be equivalent to the 80th percentile of the cohort's TMB distribution. Adjusting the threshold by steps of five-percentile increments either side of the 80th percentile, we re-assigned patients to TMB-Lo and TMB-Hi and assessed the statistical significance and hazard ratio of immune subclass survival differences by Cox regression (eTable 1 in Supplement 1). As compared to thresholds ranging from the 60th-95th percentiles, we observed that the original mean-based threshold yielded the most statistically significant association ($P < 0.0001$, likelihood ratio test) and the hazard ratio of greatest effect (HR = 0.27) suggesting the general suitability of the mean for this purpose.

By contrast, METABRIC samples were analyzed for somatic mutations using a 173-gene sequencing panel, and the number of mutations per tumor were computed as an estimate of TMB.[52] While the TCGA approach (mutation rate) accounts for sample-to-sample variation in tumor sequencing depth, the METABRIC approach (mutation count) does not. We therefore sought to determine if estimates of TMB derived from METABRIC differed significantly to that of TCGA. To address this question, we calculated the tumor mutation count for the TCGA cohort using only the 173 genes comprising the METABRIC panel and compared this mutation count to whole-exome mutation rate by correlation analysis (eFigure 5 in Supplement 1). We observed that gene-panel mutation count was highly significantly correlated ($r = 0.91$, $P < 0.001$) with whole-exome mutation rate in the TCGA cohort, suggesting that measures of TMB, as computed by the two methods, are reasonably comparable. However, other technical variables specific to the TCGA and METABRIC approaches could also hinder such comparisons. For example, the different variant calling and annotation pipelines used by the Broad Institute (for TCGA data) and Pereira, et. al.[52] (for METABRIC data) utilized different algorithms and mutation detection thresholds. Furthermore, while the TCGA approach utilized matched normal tissue for all somatic mutation calls, the METABRIC approach utilized a "pooled" normal strategy (ie, only a minor fraction of patient normal tissue specimens were sequenced), and thus, was reliant on a greater degree of inference for mutation calling.

Survival analyses

Cox proportional hazards regression and Kaplan-Meier analyses were performed using R (https://www.r-project.org/). All survival analyses examined patient overall survival defined as the time of diagnosis to last clinical follow up or death at 10 years. Variables were treated as categorical or continuous as described in the Table 1 footnotes. Wald test *P*-values were reported together with hazard ratios and 95% confidence intervals. In multivariable analysis of the TMB-Hi patient population, modeling immune subclass as a categorical variable (i.e., FID (reference) vs. WID and PID) gave a degenerate estimate resulting from no death events being present in the FID group. Thus, in Table 1 we used Nelder-Mead optimization to derive a continuous "meta"-score for immune subclass, where:

$$\text{Immune subclass meta-score} = -0.95\beta_{T/NK} - 0.25\beta_{B/P} + 1.03\beta_{M/D}$$

Alternatively, we considered maintaining FID, WID and PID as categorical variables in the model, but with recoding of the survival status of the longest-lived person in the FID group to an uncensored event (i.e., an artificial "death" event). While creating a small penalty related to explaining variance for FID, this avoids the need for representing the immune subclasses as a continuous variable. In the multivariable model using FID as the reference, the PID effect remained significant in the model ($P = 0.0073$, HR = 0.04, 95%CI = 0.005–0.43) while the WID effect did not ($P = 0.4898$, HR = 0.68, 95%CI = 0.22–2.05).

Transcriptomic correlates of immune cell abundance

To gain insight into the immunological composition of the FID, WID and PID subclasses, we examined the relative enrichment of immune cell types within tumors. The gene signature matrix of 22 immune cell types (LM-22) developed by Alizadeh and colleagues [25] was used as input for the Gene Set Variation Analysis (GSVA) algorithm of Hanzelmann and colleagues.[58] GSVA was implemented by R package "GSVA". Annotation and results from GSVA were visualized by R

package "heatmap". Subclass enrichment score distributions were compared by Mann-Whitney U test in Figure 2.

Copy number analysis

Both raw and GISTIC-processed CNA data were analyzed. Raw CNA values were adjusted for tumor purity using the Estimation of STromal and Immune cells in MAlignant Tumor tissues using Expression data (ESTIMATE) tool (http://dx.doi.org/10.1038/ncomms3612) via the R package (ESTIMATE). To assess the significance of gene copy number differences among immune subclasses, ANOVA (raw) or Chi-square statistics (GISTIC) were used. The Benjamini-Hochberg *P*-value adjustment method was used to control for false discoveries in the TCGA data set; unadjusted Chi-square test *P*-values were reported for the METABRIC validation (eTable 3 and eTable 4 in Supplement 1).

Other statistics

All statistical tests were performed using R (https://www.r-project.org/) or SigmaPlot 12.0. Significant differences between distributions were computed by the Mann-Whitney U test (2 groups) or Kruskal-Wallis test (> 2 groups). False Discovery Rate (FDR) was controlled using the Benjamini-Hochberg procedure. Hierarchical clustering of gene and tumor expression profiles, followed by heatmap-based visualization of data was performed using Eisen's Cluster (v2.11) and TreeView (v1.60) software.[59] Data were mean centered on genes and clustered using average linkage clustering and uncentered Pearson correlation. The Database for Annotation, Visualization and Integrated Discovery (DAVID) Bioinformatics Resource version 6.7[60] was used to study enrichment of Gene Ontologies associated with genes amplified in immune subclasses. FDR adjusted *P*-values are reported.

Acknowledgments

L.D.M. and A.T. conceived of and designed the study. All authors contributed to aspects of the data analysis and interpretation and the writing of the manuscript. We thank the Cancer Genome Atlas Research Network and the Molecular Taxonomy of Breast Cancer International Consortium for collecting, analyzing and sharing data used in this study.

COI Disclosure

The funders of this work had no role in study design; collection, analysis or interpretation of the data; or the preparation or submission of the manuscript for publication. C.M.P is an equity stock holder, consultant, and Board of Director Member of BioClassifier LLC and GeneCentric Diagnostics. C.M.P is also listed as an inventor on patent applications on the Breast PAM50 Subtyping assay.

Funding

This study was supported by funds from the American Cancer Society (RSG-12-198-01-TBG; L.D.M., E.D.R. and A.P.), the Mary Kirkpatrick Professorship for Breast Cancer Research (L.D.M.), the NCI Breast SPORE program (P50-CA58223-09A1; C.M.P.), RO1-CA195740-01 (C. M.P.), the Breast Cancer Research Foundation (C.M.P.), the Maurice Wilkins Centre, New Zealand (C. P. and N.K.) and Breast Cancer Cure NZ (C.P.). The Biostatistics Shared Resource and the Bioinformatics Shared Resource, supported by the Wake Forest Baptist Comprehensive Cancer Center's NCI Cancer Center Support Grant award number P30CA012197, also contributed to this work. The content is solely the responsibility of the authors and does not represent the official views of the National Cancer Institute.

ORCID

Cristin Print http://orcid.org/0000-0001-8345-7812
Michael A. Black http://orcid.org/0000-0003-1174-6054
Davide Bedognetti http://orcid.org/0000-0002-5857-773X
Wendell D. Jones http://orcid.org/0000-0002-9676-5387
Michael L. Gatza http://orcid.org/0000-0001-6796-7791
Pierre Triozzi http://orcid.org/0000-0003-2488-6136
Lance D. Miller http://orcid.org/0000-0003-3799-2528

References

1. Adams S, Gray RJ, Demaria S, Goldstein L, Perez EA, Shulman LN, Martino S, Wang M, Jones VE, Saphner TJ, et al. Prognostic value of tumor-infiltrating lymphocytes in triple-negative breast cancers from two phase III randomized adjuvant breast cancer trials: ECOG 2197 and ECOG 1199. J Clin Oncol. 2014;32:2959–2966. doi:10.1200/JCO.2013.55.0491.
2. Denkert C, Loibl S, Noske A, Roller M, Muller BM, Komor M, Budczies J, Darb-Esfahani S, Kronenwett R, Hanusch C, et al. Tumor-associated lymphocytes as an independent predictor of response to neoadjuvant chemotherapy in breast cancer. J Clin Oncol. 2010;28:105–113. doi:10.1200/JCO.2009.23.7370.
3. Loi S, Michiels S, Salgado R, Sirtaine N, Jose V, Fumagalli D, Kellokumpu-Lehtinen P-L, Bono P, Kataja V, Desmedt C, et al. Tumor infiltrating lymphocytes are prognostic in triple negative breast cancer and predictive for trastuzumab benefit in early breast cancer: results from the FinHER trial. Ann Oncol: Official Journal Eur Soc Med Oncology/ESMO. 2014;25:1544–1550. doi:10.1093/annonc/mdu112.
4. Loi S, Sirtaine N, Piette F, Salgado R, Viale G, Van Eenoo F, Rouas G, Francis P, Crown JPA, Hitre E, et al. Prognostic and predictive value of tumor-infiltrating lymphocytes in a phase III randomized adjuvant breast cancer trial in node-positive breast cancer comparing the addition of docetaxel to doxorubicin with doxorubicin-based chemotherapy: BIG 02-98. J Clin Oncol. 2013;31:860–867. doi:10.1200/JCO.2011.41.0902.
5. Salgado R, Denkert C, Campbell C, Savas P, Nuciforo P, Aura C, Aura C, De Azambuja E, Eidtmann H, Ellis CE, et al. Tumor-infiltrating lymphocytes and associations with pathological complete response and event-free survival in HER2-positive early-stage breast cancer treated with lapatinib and trastuzumab: a secondary analysis of the NeoALTTO Trial. JAMA Oncol. 2015;1:448–454. doi:10.1001/jamaoncol.2015.0830.
6. Brown SD, Warren RL, Gibb EA, Martin SD, Spinelli JJ, Nelson BH, Holt RA. Neo-antigens predicted by tumor genome meta-analysis correlate with increased patient survival. Genome Res. 2014;24:743–750.
7. Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, Xie M, Zhang Q, McMichael JF, Wyczalkowski MA, et al. Mutational landscape and significance across 12 major cancer types. Nature. 2013;502:333–339.
8. Goodman AM, Kato S, Bazhenova L, Patel SP, Frampton GM, Miller V, Stephens PJ, Daniels GA, Kurzrock R. Tumor mutational burden as an independent predictor of response to immunotherapy in diverse cancers. Mol Cancer Ther. 2017;16:2598–2608.
9. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, Lee W, Yuan J, Wong P, Ho TS, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. Science. 2015;348:124–128.

10. Snyder A, Makarov V, Merghoub T, Yuan J, Zaretsky JM, Desrichard A, Walsh LA, Postow MA, Wong P, Ho TS, et al. Genetic basis for clinical response to CTLA-4 blockade in melanoma. N Engl J Med. 2014;371:2189–2199. doi:10.1056/NEJMoa1406498.
11. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, Skora AD, Luber BS, Azad NS, Laheru D, et al. PD-1 blockade in tumors with mismatch-repair deficiency. N Engl J Med. 2015;372:2509–2520. doi:10.1056/NEJMoa1500596.
12. Luen S, Virassamy B, Savas P, Salgado R, Loi S. The genomic landscape of breast cancer and its interaction with host immunity. Breast. 2016;29:241–250. doi:10.1016/j.breast.2016.07.015.
13. Kotoula V, Lakis S, Vlachos IS, Giannoulatou E, Zagouri F, Alexopoulou Z, Gogas H, Pectasides D, Aravantinos G, Efstratiou I, et al. Tumor infiltrating lymphocytes affect the outcome of patients with operable triple-negative breast cancer in combination with mutated amino acid classes. PloS One. 2016;11:e0163138. doi:10.1371/journal.pone.0163138.
14. Rody A, Holtrich U, Pusztai L, Liedtke C, Gaetje R, Ruckhaeberle E, Solbach C, Hanker L, Ahr A, Metzler D, et al. T-cell metagene predicts a favorable prognosis in estrogen receptor-negative and HER2-positive breast cancers. Breast Cancer Res. 2009;11:R15. doi:10.1186/bcr2234.
15. Iglesia MD, Parker JS, Hoadley KA, Serody JS, Perou CM, Vincent BG. Genomic analysis of immune cell infiltrates across 11 tumor types. J Natl Cancer Inst. 2016;108(11). doi:10.1093/jnci/djw144.
16. Bense RD, Sotiriou C, Piccart-Gebhart MJ, Haanen JB, van Vugt MA, de Vries EG, Schröder CP, Fehrmann RS. Relevance of tumor-infiltrating immune cell composition and functionality for disease outcome in breast cancer. J Natl Cancer Inst. 2016 Oct 13;109(1). pii: djw192.
17. Bedognetti D, Hendrickx W, Marincola FM, Miller LD. Prognostic and predictive immune gene signatures in breast cancer. Curr Opin Oncol. 2015;27:433–444. doi:10.1097/CCO.0000000000000234.
18. Chifman J, Pullikuth A, Chou JW, Bedognetti D, Miller LD. Conservation of immune gene signatures in solid tumors and prognostic implications. BMC Cancer. 2016;16:911. doi:10.1186/s12885-016-2948-z.
19. Nagalla S, Chou JW, Willingham MC, Ruiz J, Vaughn JP, Dubey P, Lash TL, Hamilton-Dutoit SJ, Bergh J, Sotiriou C, et al. Interactions between immunity, proliferation and molecular subtype in breast cancer prognosis. Genome Biol. 2013;14:R34. doi:10.1186/gb-2013-14-4-r34.
20. Alistar A, Chou JW, Nagalla S, Black MA, D'Agostino R Jr., Miller LD. Dual roles for immune metagenes in breast cancer prognosis and therapy prediction. Genome Med. 2014;6:80. doi:10.1186/s13073-014-0080-8.
21. Miller LD, Chou JA, Black MA, Print C, Chifman J, Alistar A, Putti T, Zhou X, Bedognetti D, Hendrickx W, et al. Immunogenic subtypes of breast cancer delineated by gene classifiers of immune responsiveness. Cancer Immunol Res. 2016;4:600–610. doi:10.1158/2326-6066.CIR-15-0149.
22. Cancer Genome Atlas N. Comprehensive molecular portraits of human breast tumours. Nature. 2012;490:61–70. doi:10.1038/nature11412.
23. Haricharan S, Bainbridge MN, Scheet P, Brown PH. Somatic mutation load of estrogen receptor-positive breast tumors predicts overall survival: an analysis of genome sequence data. Breast Cancer Res Treat. 2014;146:211–220. doi:10.1007/s10549-014-2991-x.
24. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, Speed D, Lynch AG, Samarajiwa S, Yuan Y, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature. 2012;486:346–352. doi:10.1038/nature10983.
25. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12:453–457. doi:10.1038/nmeth.3337.
26. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. Cell. 2015;160:48–61. doi:10.1016/j.cell.2014.12.033.
27. Ali HR, Provenzano E, Dawson SJ, Blows FM, Liu B, Shah M, Earl HM, Poole CJ, Hiller L, Dunn JA, et al. Association between CD8 + T-cell infiltration and breast cancer survival in 12,439 patients. Ann Oncol: Official Journal Eur Soc Med Oncology/ESMO. 2014;25:1536–1543. doi:10.1093/annonc/mdu191.
28. Dieci MV, Mathieu MC, Guarneri V, Conte P, Delaloge S, Andre F, Goubar A. Prognostic and predictive value of tumor-infiltrating lymphocytes in two phase III randomized adjuvant breast cancer trials. Ann Oncol: Official Journal Eur Soc Med Oncology/ESMO. 2015;26:1698–1704. doi:10.1093/annonc/mdv239.
29. Denkert C, Von Minckwitz G, Darb-Esfahani S, Lederer B, Heppner BI, Weber KE, Budczies J, Huober J, Klauschen F, Furlanetto J, 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 Jan;19(1):40–50. doi:10.1016/S1470-2045(17)30904-X.
30. McGranahan N, Furness AJ, Rosenthal R, Ramskov S, Lyngaa R, Saini SK, Jamal-Hanjani M, Wilson GA, Birkbak NJ, Hiley CT, et al. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. Science. 2016;351:1463–1469. doi:10.1126/science.aaf1490.
31. Mehnert JM, Panda A, Zhong H, Hirshfield K, Damare S, Lane K, Sokol L, Stein MN, Rodriguez-Rodriquez L, Kaufman HL, et al. Immune activation and response to pembrolizumab in POLE-mutant endometrial cancer. J Clin Invest. 2016;126:2334–2340. doi:10.1172/JCI84940.
32. Rosenberg JE, Hoffman-Censits J, Powles T, Van Der Heijden MS, Balar AV, Necchi A, Dawson N, O'Donnell PH, Balmanoukian A, Loriot Y, et al. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. Lancet. 2016;387:1909–1920. doi:10.1016/S0140-6736(16)00561-4.
33. Chen PL, Roh W, Reuben A, Cooper ZA, Spencer CN, Prieto PA, Miller JP, Bassett RL, Gopalakrishnan V, Wani K, et al. Analysis of immune signatures in longitudinal tumor samples yields insight into biomarkers of response and mechanisms of resistance to immune checkpoint blockade. Cancer Discov. 2016;6:827–837. doi:10.1158/2159-8290.CD-15-1545.
34. Ulloa-Montoya F, Louahed J, Dizier B, Gruselle O, Spiessens B, Lehmann FF, Suciu S, Kruit WHJ, Eggermont AMM, Vansteenkiste J, et al. Predictive gene signature in MAGE-A3 antigen-specific cancer immunotherapy. J Clin Oncol. 2013;31:2388–2395. doi:10.1200/JCO.2012.44.3762.
35. Shah SP, Roth A, Goya R, Oloumi A, Ha G, Zhao Y, Turashvili G, Ding J, Tse K, Haffari G, et al. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. Nature. 2012;486:395–399. doi:10.1038/nature10933.
36. Nanda R, Chow LQ, Dees EC, Berger R, Gupta S, Geva R, Pusztai L, Pathiraja K, Aktan G, Cheng JD, et al. Pembrolizumab in patients with advanced triple-negative breast cancer: phase Ib KEYNOTE-012 study. J Clin Oncol. 2016;34:2460–2467. doi:10.1200/JCO.2015.64.8931.
37. Davies H, Morganella S, Purdie CA, Jang SJ, Borgen E, Russnes H, Glodzik D, Zou X, Viari A, Richardson AL, et al. Whole-genome sequencing reveals breast cancers with mismatch repair deficiency. Cancer Res. 2017;77:4755–4762. doi:10.1158/0008-5472.CAN-17-1083.
38. Wen YH, Brogi E, Zeng Z, Akram M, Catalano J, Paty PB, Norton L, Shia J. DNA mismatch repair deficiency in breast carcinoma: a pilot study of triple-negative and non-triple-negative tumors. Am J Surg Pathol. 2012;36:1700–1708. doi:10.1097/PAS.0b013e3182627787.
39. Karn T, Jiang T, Hatzis C, Sanger N, El-Balat A, Rody A, Holtrich U, Becker S, Bianchini G, Pusztai L. Association between genomic metrics and immune infiltration in triple-negative breast cancer. JAMA Oncol. 2017;3:1707–1711. doi:10.1001/jamaoncol.2017.2140.
40. Orsetti B, Nugoli M, Cervera N, Lasorsa L, Chuchana P, Rouge C, Ursule L, Nguyen C, Bibeau F, Rodriguez C, et al. Genetic

profiling of chromosome 1 in breast cancer: mapping of regions of gains and losses and identification of candidate genes on 1q. Br J Cancer. 2006;95:1439–1447. doi:10.1038/sj.bjc.6603433.
41. Halak BK, Maguire HC Jr., Lattime EC. Tumor-induced interleukin-10 inhibits type 1 immune responses directed at a tumor antigen as well as a non-tumor antigen present at the tumor site. Cancer Res. 1999;59:911–917.
42. Waring P, Mullbacher A. Cell death induced by the Fas/Fas ligand pathway and its role in pathology. Immunol Cell Biol. 1999;77:312–317. doi:10.1046/j.1440-1711.1999.00837.x.
43. O'Connell J, Houston A, Bennett MW, O'Sullivan GC, Shanahan F. Immune privilege or inflammation? Insights into the Fas ligand enigma. Nat Med. 2001;7:271–274. doi:10.1038/85395.
44. Masson D, Tschopp J. Inhibition of lymphocyte protease granzyme A by antithrombin III. Mol Immunol. 1988;25:1283–1289.
45. Kaiserman D, Bird PI. Control of granzymes by serpins. Cell Death Differ. 2010;17:586–595. doi:10.1038/cdd.2009.169.
46. Broad Institute TCGA Genome Data Analysis Center. Analysis-ready standardized TCGA data from Broad GDAC Firehose 2016_01_28 run. Broad Institute of MIT and Harvard. Dataset. 2016. doi:10.7908/C11G0KM9.
47. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol. 2011;12:R41. doi:10.1186/gb-2011-12-10-r102.
48. Ciriello G, Gatza ML, Beck AH, Wilkerson MD, Rhie SK, Pastore A, Zhang H, McLellan M, Yau C, Kandoth C, et al. Comprehensive molecular portraits of invasive lobular breast cancer. Cell. 2015;163:506–519. doi:10.1016/j.cell.2015.09.033.
49. Nielsen TO, Parker JS, Leung S, Voduc KD, Ebbert M, Vickery TL, Davies SR, Snider J, Stijleman IJ, Reed J, et al. A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor positive breast cancer. Clin Cancer Res. 2010;16:5222–5232. doi:10.1158/1078-0432.CCR-10-1282.
50. Prat A, Perou CM. Deconstructing the molecular portraits of breast cancer. Mol Oncol. 2011;5:5–23. doi:10.1016/j.molonc.2010.11.003.
51. Broad Institute TCGA Genome Data Analysis Center. Mutation Analysis (MutSigCV v0.9). Broad Institute of MIT and Harvard. 2016. doi:10.7908/C1PK0FH7.
52. Pereira B, Chin SF, Rueda OM, Vollan HK, Provenzano E, Bardwell HA, Pugh M, Jones L, Russell R, Sammut S-J, et al. The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. Nat Commun. 2016;7:11479. doi:10.1038/ncomms11479.
53. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012;2:401–404. doi:10.1158/2159-8290.CD-12-0095.
54. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013;6:pl1. doi:10.1126/scisignal.2004088.
55. Margolin AA, Bilal E, Huang E, Norman TC, Ottestad L, Mecham BH, Sauerwine B, Kellen MR, Mangravite LM, Furia MD, et al. Systematic analysis of challenge-driven improvements in molecular prognostic models for breast cancer. Sci Transl Med. 2013;5:181re1. doi:10.1126/scitranslmed.3006112.
56. Mecham BH, Nelson PS, Storey JD. Supervised normalization of microarrays. Bioinformatics. 2010;26:1308–1315. doi:10.1093/bioinformatics/btq118.
57. Venkatraman ES, Olshen AB. A faster circular binary segmentation algorithm for the analysis of array CGH data. Bioinformatics. 2007;23:657–663. doi:10.1093/bioinformatics/btl646.
58. Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics. 2013;14:7.
59. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A. 1998;95:14863–14868.
60. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4:44–57.