Ann Surg Oncol (2019) 26:2191–2199
https://doi.org/10.1245/s10434-019-07338-3

Annals of
SURGICAL ONCOLOGY
OFFICIAL JOURNAL OF THE SOCIETY OF SURGICAL ONCOLOGY

ORIGINAL ARTICLE – BREAST ONCOLOGY

Tumor Heterogeneity Correlates with Less Immune Response and Worse Survival in Breast Cancer Patients

Kerry-Ann McDonald, MD[1], Tsutomu Kawaguchi, MD, PhD[1,2], Qianya Qi[3], Xuan Peng[3], Mariko Asaoka, MD[1], Jessica Young, MD[1], Mateusz Opyrchal, MD, PhD[4], Li Yan, PhD[3], Santosh Patnaik, MD, PhD[5], Eigo Otsuji, MD, PhD[2], and Kazuaki Takabe, MD, PhD, FACS[1,6,7,8,9,10,11]

[1]Breast Surgery, Department of Surgical Oncology, Roswell Park Comprehensive Cancer Center, Buffalo, NY; [2]Department of Surgery, Kyoto Prefectural University of Medicine, Kyoto, Japan; [3]Department of Biostatistics and Bioinformatics, Roswell Park Comprehensive Cancer Center, Buffalo, NY; [4]Department of Medical Oncology, Roswell Park Comprehensive Cancer Center, Buffalo, NY; [5]Thoracic Surgery, Department of Surgical Oncology, Roswell Park Comprehensive Cancer Center, Buffalo, NY; [6]Department of Surgery, University at Buffalo Jacobs School of Medicine and Biomedical Sciences, The State University of New York, Buffalo, NY; [7]Department of Breast Surgery and Oncology, Tokyo Medical University, Tokyo, Japan; [8]Department of Surgery, Yokohama City University, Yokohama, Japan; [9]Department of Surgery, Niigata University Graduate School of Medical and Dental Sciences, Niigata, Japan; [10]Department of Breast Surgery and Oncology, Fukushima Medical University, Fukushima, Japan; [11]Breast Oncology and Surgery, Roswell Park Cancer Institute, Buffalo, NY

ABSTRACT

Background. Intratumor heterogeneity implies that subpopulations of cancer cells that differ in genetic, phenotypic, or behavioral characteristics coexist in a single tumor (Ma in Breast Cancer Res Treat 162(1):39–48, 2017; Martelotto in Breast Cancer Res 16(3):210, 2014). Tumor heterogeneity drives progression, metastasis and treatment resistance, but its relationship with tumor infiltrating immune cells is a matter of debate, where some argue that tumors with high heterogeneity may generate neoantigens that attract immune cells, and others claim that immune cells provide selection pressure that shapes tumor heterogeneity (McGranahan et al. in Science 351(6280): 1463–1469, 2016; McGranahan and Swanton in Cell 168(4):613–628, 2017). We sought to study the association between tumor heterogeneity and immune cells in a real-world cohort utilizing The Cancer Genome Atlas.

Methods. Mutant allele tumor heterogeneity (MATH) was calculated to estimate intratumoral heterogeneity, and immune cell compositions were estimated using CIBERSORT. Survival analyses were demonstrated using Kaplan–Meir curves.

Results. Tumors with high heterogeneity (high MATH) were associated with worse overall survival ($p = 0.049$), as well as estrogen receptor-positive ($p = 0.011$) and non-triple-negative tumors ($p = 0.01$). High MATH tumors were also associated with less infiltration of anti-tumor CD8 ($p < 0.013$) and CD4 T cells ($p < 0.00024$), more tumor-promoting regulatory T cells ($p < 4e{-}04$), lower expression of T-cell exhaustion markers, specifically PDL-1 ($p = 0.0031$), IDO2 ($p = 0.34$), ADORA2A ($p = 0.018$), VISTA ($p = 0.00013$), and CCR4 ($p < 0.00001$), lower expression of cytolytic enzymes granzyme A ($p = 0.0056$) and perforin 1 ($p = 0.053$), and low cytolytic activity score ($p = 0.0028$).

Conclusions. High heterogeneity tumors are associated with less immune cell infiltration, less activation of the immune response, and worse survival in breast cancer. Our results support the notion that tumor heterogeneity is shaped by selection pressure of tumor-infiltrating immune cells.

Kerry-Ann McDonald and Tsutomu Kawaguchi contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1245/s10434-019-07338-3) contains supplementary material, which is available to authorized users.

First Received: 8 August 2018;
Published Online: 8 April 2019

K. Takabe, MD, PhD, FACS
e-mail: Kazuaki.Takabe@roswellpark.org

Breast cancer is a heterogeneous disease with genetic and phenotypic variability. Two breast cancers that appear to be similar based on clinical, pathologic, and biomarker signatures can behave differently because of differences in underlying biology. It has recently been suggested that tumor biology, at least in part, may be determined by intratumor heterogeneity.[3] Intratumor heterogeneity implies that subpopulations of cancer cells that differ in their genetic, phenotypic, or behavioral characteristics coexist within the same tumor.[1,2] Tumor heterogeneity has been linked to cancer progression and therapeutic resistance.[1,2,4–6] Evaluation of intratumor heterogeneity of individual tumors, which impacts disease progression and efficacy of therapies, is essential to overcome treatment challenges of the primary tumors and subsequent metastasis in breast cancer.

Mutant-allele tumor heterogeneity (MATH) is a bioinformatic algorithm that provides a measurable and quantitative assessment of intratumor heterogeneity, which was generated from whole-exome sequencing of tumors and their matched normal DNA.[1] The prognostic impact of MATH has been studied in a variety of cancers, such as head and neck, colorectal, and breast cancer cohorts.[1,7,8] Mroz and Rocco analyzed next-generation sequencing results for 74 head and neck squamous cell cancers and found that MATH was higher in tumors with a mutated TP53 gene.[7] In a retrospective analysis of head and neck squamous cancers in The Cancer Genome Atlas (TCGA), a relationship was found between high MATH and worse overall survival (OS).[9] Rajput et al. used the next-generation sequencing approach to analyze mutations in samples of stage II and III colon cancer patients and found a strong correlation between higher MATH and the risk of metastases.[8] In addition, Ma et al. found that higher stage, triple-negative breast cancer (TNBC), and p53 mutations were associated with higher MATH in breast cancer patients.[1] Taken together, these studies suggest that intratumor heterogeneity is associated with worse cancer outcomes and has prognostic relevance.

Immune surveillance of cancer is important to suppress tumor growth, progression, and metastasis. Tumors evolve as a result of genomic instability leading to dominant mutations, which can drive cancer progression.[3,10] Genetic heterogeneity is suggested to generate neoantigens that attract immune cells. It has been shown that clonal neoantigens elicit T-cell immunoreactivity and sensitivity to immune checkpoint blockade.[3] PDL-1, the programmed death ligand, is an immune checkpoint protein for T cells. Specifically, CD8+ tumor-infiltrating lymphocytes (TILs) reactive to clonal neoantigens were identified in early-stage non-small cell lung cancer, expressed high levels of PD-1, and tumors that expressed more clonal neoantigens were more sensitive to PD-1 blockade.[3] Alternatively, there is evidence that TILs themselves function as a selective pressure, resulting in tumor clonality and intratumor heterogeneity.[3] In the first model, high heterogeneity tumors should attract a high number of TILs with high cytolytic activity, as opposed to the latter model, which should have less TILs with low cytolytic activity because high heterogeneity should be the consequence.

To our knowledge, the relationship between intratumor heterogeneity and the immunogenic landscape in breast cancer has not been fully studied in a large cohort of patients. The current study sought to identify the association of tumor heterogeneity and infiltrating immune cell compositions, as well as its cytolytic activity, immune response genes, and survival in breast cancer. We hypothesized that high intratumor heterogeneity calculated by the MATH algorithm was associated with low infiltrating immune cells, low cytolytic activity, and worse survival.

METHODS

Patient Cohort

TCGA, a project supervised by the National Cancer Institute and the National Human Genome Research Institute, is a publicly available database that includes clinical and genomic data on breast cancer patient samples collected worldwide.[11] Many data parameters are included in the TCGA dataset; however, grade is not included. Therefore, we discovered that a majority of the grade data of the patient tumors used in TCGA are available in the TIES database. We reviewed all pathology reports in TIES and associated the grade data with the tumor in TCGA. The primary endpoint was OS, defined as the time from the date of diagnosis to death by any cause. Overall, 1093 patients with messenger RNA (mRNA) expression from RNA sequencing were included in the breast cancer cohort (BRCA) of TCGA Provisional cohort, of whom OS data were available in 959 patients. The gene expression level quantification data (mRNA expression z-score from the RNA sequence) for TCGA cohort was downloaded through cBioPortal[12,13] and used as previously described.[14–22]

Mutant-Allele Tumor Heterogeneity (MATH)

The MATH level was calculated from the median absolute deviation (MAD) and the median of its mutant-allele fractions at tumor-specific mutated loci as previously described.[7,23]

Survival Analysis

The OS between MATH high and low tumors were analyzed using Kaplan–Meier curves with the log-rank test. Patients were divided based on the degree of heterogeneity as calculated by the MATH algorithm. The threshold of dichotomization of the MATH high and low groups was determined by comparing differences in the OS between the two groups at multiple candidate cut-off points within the range of MATH, which is similar to the method previously published.[24–27]

CIBERSORT

CIBERSORT, a bioinformatic algorithm to calculate immune cell composition from their gene expression profiles, was used to estimate tumor-infiltrating cell composition in tumors.[28] Immune cell fraction data were downloaded through TCIA (https://tcia.at/home).[29] Each immune cell fraction was compared between the MATH high and low tumors in TCGA cohort using the same cut-off of OS analysis.

Statistical Analysis

All statistical analyses were performed using R software (http://www.r-project.org/) and Bioconductor (http://bioconductor.org/). The Kaplan–Meier method with log-rank tests and Cox proportional hazard models were used to compare the MATH high and low groups. Pearson correlations were calculated based on expression levels of each interested gene, and then plotted. Gene expression comparison was analyzed using Student's t test, and immune cell fraction comparison was analyzed using the Wilcoxon signed-rank test. In all analyses, a two-sided p value < 0.05 was considered statistically significant.

RESULTS

Patient and Tumor Characteristics

Tumor characteristics of the low and high MATH groups were similar overall. The majority of patients in both groups were < 60 years of age, and had stage II, estrogen receptor (ER)-positive, progesterone receptor (PR)-positive, human epidermal growth factor receptor 2 (HER-2)-negative breast cancers. Factors found to be statistically significant included ER, PR, and TNBC status ($p \leq 0.0002$, $p < 0.0001$, and $p < 0.0003$, respectively), stage ($p = 0.0354$), tumor size ($p = 0.0001$), and grade ($p = 0.0002$). Specifically, ER-negative, PR-negative, and TNBCs were more associated with high tumor heterogeneity (Table 1). High MATH tumors were associated with T2 status and N2 nodal positivity, and were also found to have higher grades. The grade of breast carcinoma is a prognostic factor, and different grades have different profiles by proteomic and genomic analysis.[30]

TABLE 1 Patient demographics characterized by MATH level in TCGA breast cancer cohort

Factor	MATH		p value
	Low (n)	High (n)	
Age			
< 60	222	288	
> 60	184	253	0.658
ER			
Negative	66	140	
Positive	325	376	**0.0002**
PR			
Negative	99	195	
Positive	291	319	$<$ **0.0001**
HER2			
Negative	340	432	
Positive	58	92	0.2239
TNBC			
No	361	435	
Yes	46	108	**0.0003**
Stage			
I	86	76	
II	218	330	
III	94	122	
IV	6	8	**0.0354**
T-Stage			
1	134	118	
2	215	338	
3	54	65	
4	8	27	**0.0001**
N-stage			
Negative	187	268	
Positive	218	272	0.3238
Grade			
1	41	27	
2	116	113	**0.0002** (Chi square test) < 0.0001
3	69	128	(Cochran–Armitage trend test)

Statistically significant values are highlighted in bold

MATH mutant allele tumor heterogeneity, *TCGA* The Cancer Genome Atlas, *ER* estrogen receptor, *PR* progesterone receptor, *HER2* human epidermal growth factor receptor, *TNBC* triple-negative breast cancer

High Tumor Heterogeneity is Associated with Worse Survival in Hormone-Positive Breast Cancer

Previous studies demonstrated that patients with highly heterogeneous cancers have worse survival.[31] We sought to explore this association further in TCGA breast cancer cohort. We found that high MATH tumors had a higher mutation burden, higher KI-67 gene expression, and more intratumor heterogeneity (Fig. 1a). This correlated to high MATH tumors having worse OS compared with low MATH tumors (MATH high, $n = 548$; MATH low, $n = 411$; $p = 0.0499$) (Fig. 1b). Subgroup analysis revealed that this survival disadvantage in the high MATH group was statistically significant for ER-positive (MATH high, $n = 376$; MATH low, $n = 325$; $p = 0.011$), PR-positive (MATH high, $n = 319$; MATH low, $n = 291$; $p = 0.024$), and non-TNBCs (MATH high, $n = 435$; MATH low, $n = 361$; $p = 0.01$) (Fig. 1c). This result is in agreement with previous reports and validates that TCGA cohort is similar to previous cohorts.

Tumors with High Heterogeneity are Associated with Less Immune Cell Infiltration

We assessed immune cell composition in TCGA cohort and studied its association with tumor heterogeneity. Each immune cell fraction was compared between MATH high and low tumors using CIBERSORT. High MATH tumor was associated with significantly lower fractions of activated CD8+ T cells ($p = 0.013$), B cells ($p = 0.0016$), dendritic cells ($p = 0.0019$), activated memory CD4+ T cells ($p = 0.0002$), and mast cells ($p = 0.0007$), all of which play anti-tumor roles in breast cancer (Fig. 2). Interestingly, high MATH tumor was also associated with significantly higher fractions of immunosuppressive regulatory T cells (Tregs; $p \leq 0.0001$) compared with tumors

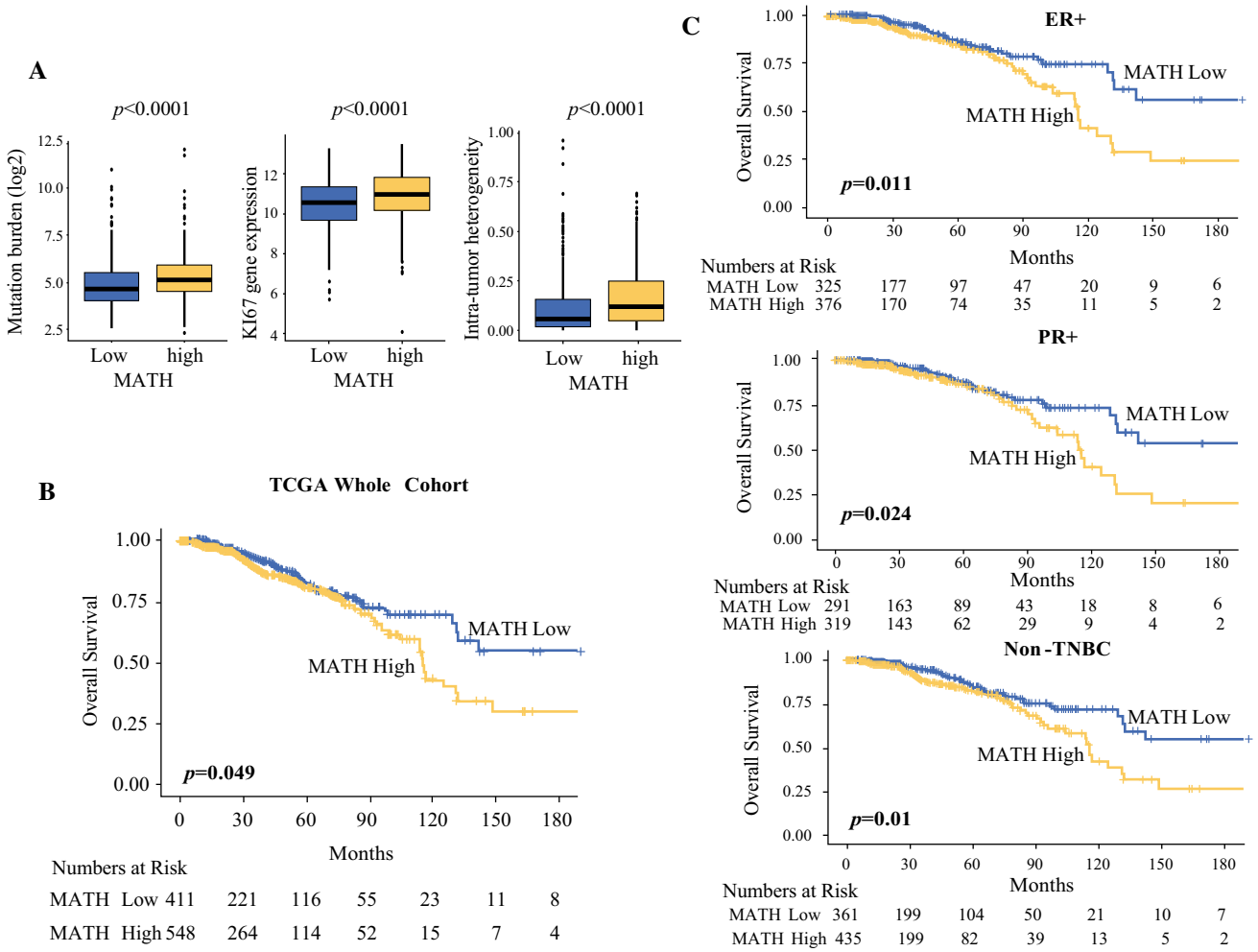

FIG. 1 MATH level and patient survival in TCGA breast cancer cohort. **a** Box plots correlating MATH to mutation burden, KI-67 gene expression, and intratumor heterogeneity. **b** MATH level and overall survival in the TCGA whole cohort. **c** MATH level and patient survival for ER-positive, PR-positive, and non-TNBCs. *MATH* mutant allele tumor heterogeneity, *TCGA* The Cancer Genome Atlas, *ER* estrogen receptor, *PR* progesterone positive, *TNBCs* triple-negative breast cancers

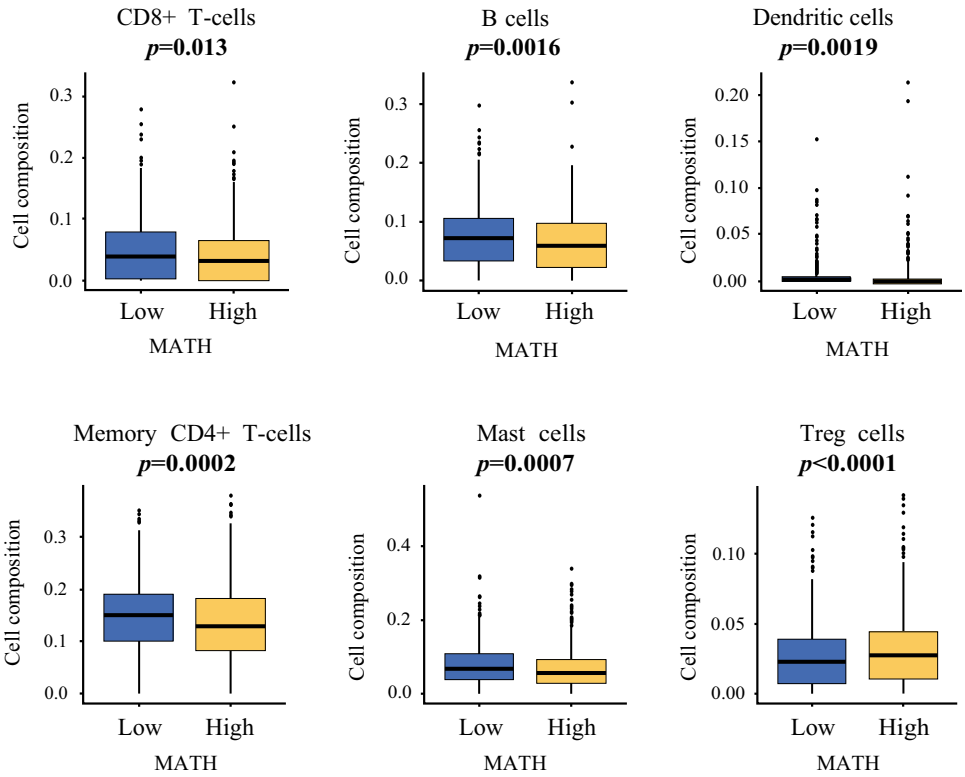

FIG. 2 MATH level and immune cell composition by CIBERSORT. Cell composition fraction comparison of activated memory CD8 T cells, B cells, dendritic cells, memory CD4 cells, mast cells, and Treg cells between MATH high and low expression in TCGA breast cancer cohort. *MATH* mutant allele tumor heterogeneity, *TCGA* The Cancer Genome Atlas, *Treg* regulatory T cells

with low MATH (Fig. 2). Our finding that tumors with high heterogeneity have not only less antitumor immune cells but also more immunosuppressive Tregs is in agreement with the notion that tumor heterogeneity is multifactorial, with immune cells being an important component.

Tumors with High Heterogeneity are Associated with Decreased Immune Response Genes

Immune checkpoint molecules are also known as T-cell exhaustion markers and reflect immunogenicity, i.e. their expression levels reflect the existence of humoral or cell-mediated immune responses. We found that high MATH tumor was associated with significantly lower mRNA expression of T-cell exhaustion markers, specifically PDL-1 ($p = 0.0031$), IDO2 (indoleamine 2,3-dioxygenase 2; $p = 0.34$), ADORA2A (adenosine A2a receptor; $p = 0.018$), VISTA [V-domain immunoglobulin (Ig)-containing suppressor of T-cell activation; $p = 0.00013$], and CCR4 (C–C chemokine receptor type 4; $p < 0.00001$) (Fig. 3). This result is consistent with the previous finding that tumors with high heterogeneity are associated with lower numbers of active immune cell infiltrations.

Tumors with High Heterogeneity are Associated with Decreased Immune Cytolytic Activity

The interplay of immune system activation and heterogeneity was further investigated by analyzing the expression of immune cytolysis genes, as well as measuring the cytolytic activity score (CYT). High MATH tumors were associated with significantly lower expression of GZMA (granzyme A; $p = 0.0056$) and PRF-1 (perforin 1; $p = 0.053$) in TCGA breast cancer cohort, both of which are key genes of immune cytolysis. Indeed, CYT was also significantly lower in high MATH tumors consistent with less cytolytic immune activity ($p = 0.0028$) (Fig. 4). This finding further supports the notion that when there are fewer active immune cell infiltrations and less cytolytic activity, the tumor is allowed to clonally evolve, and thus develops heterogeneity.

DISCUSSION

Tumor heterogeneity accounts for differences in tumor behavior among individual patients, and even within an individual tumor that may not be explained by clinical and pathologic features alone. Currently, the relationship between tumor heterogeneity and immune cells is a matter of debate, where some argue that tumor heterogeneity generates neoantigens that attract immune cells,[10] while others claim that immune cells function as a selective

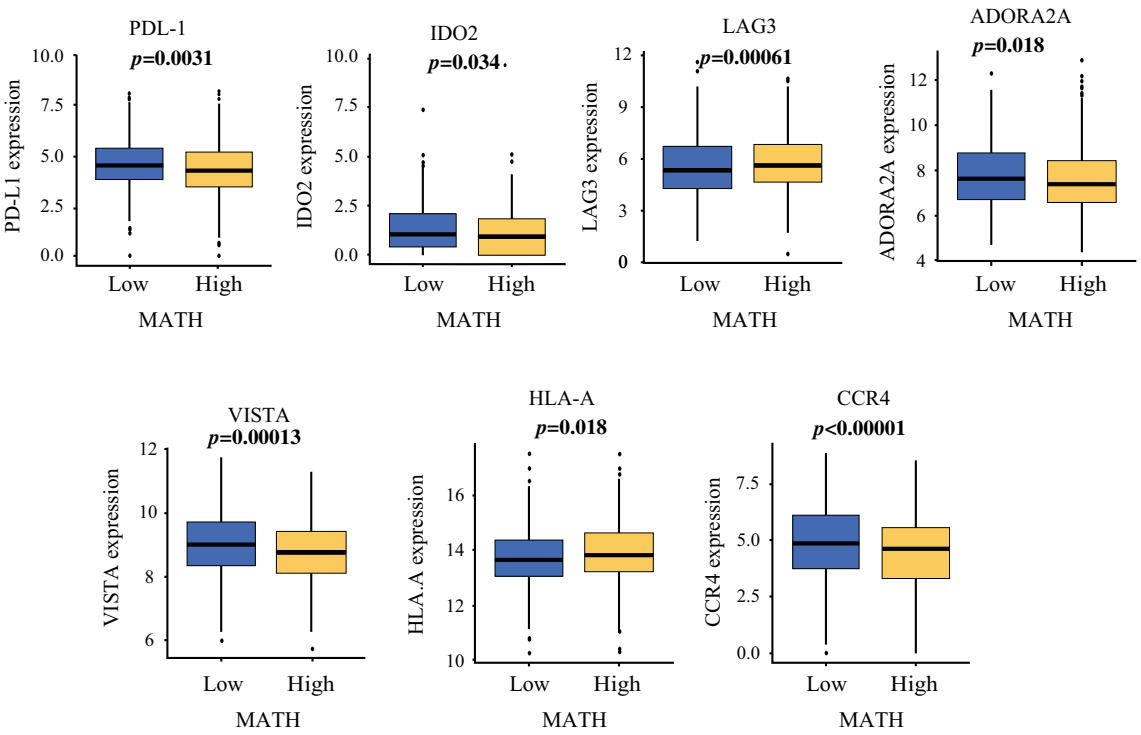

FIG. 3 MATH level and expression of immune response genes. mRNA expression of PDL-1, IDO2, LAG3, ADORA2A, VISTA, HLA-A, and CCR4 proteins in TCGA breast cancer cohort. *MATH* mutant allele tumor heterogeneity, *mRNA* messenger RNA, *PDL-1* programmed death ligand-1, *IDO2* indoleamine 2,3-dioxygenase 2, *LAG3* lymphocyte activation gene 3, *ADORA2A* adenosine A2a receptor, *VISTA* V-domain immunoglobulin (Ig)-containing suppressor of T-cell activation, *HLA-A* human leukocyte antigen-A, *CCR4* C-C chemokine receptor type 4, *TCGA* The Cancer Genome Atlas

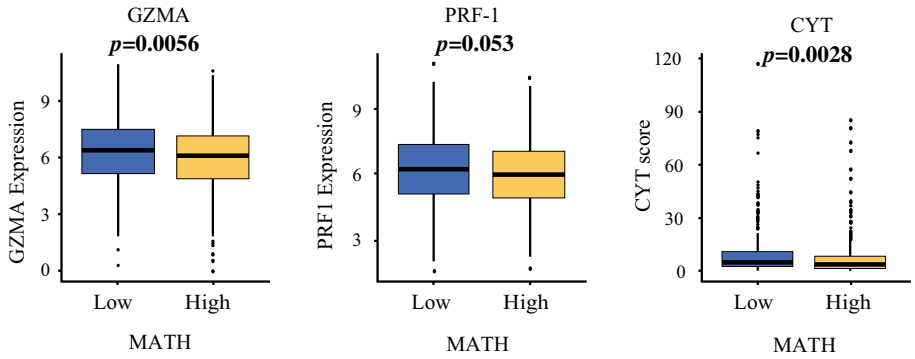

FIG. 4 MATH level and expression of cytolysis proteins. Box plot of mRNA expression of GZMA, PRF-1, and CYT protein in high and low MATH tumors in TCGA breast cancer cohort. *MATH* mutant allele tumor heterogeneity, *mRNA* messenger RNA, *GZMA* granzyme A, *PRF-1* perforin 1, *CYT* cytolysis, *TCGA* The Cancer Genome Atlas

pressure that controls tumor heterogeneity.[3,4] To this end, the current study was aimed at studying the association between tumor heterogeneity and immune cells and its impact on survival. Evaluation of breast tumor heterogeneity has advanced from whole-tumor analysis to single-cell analysis, allowing examination of more than 90% of the genome in a single cell.[32,33] MATH served to stratify patients into low- and high-risk groups based on a quantitative assessment of their tumor heterogeneity. Based on this study, high MATH tumors are synonymous with high heterogeneity. We showed that high heterogeneity correlates with worse prognosis, and that data suggesting that immune system interaction with the tumors plays an important role in shaping the cellular make-up of the tumors.

This study demonstrates that high MATH tumors are more mutation laden, have the ability to be more proliferative, and are more heterogeneous overall (Fig. 1a). The survival analysis in this study is similar to that reported by Ma et al., where hormone receptor-positive patients with high MATH showed a tendency toward worse OS.[1] In general, patients with ER-positive breast cancer are considered to have a better prognosis, confounded by the risk of recurrence occurring long after initial treatment.

Lindström et al. explored heterogeneity among hormone-positive tumors by examining the ER intensity level among 1780 postmenopausal lymph node-negative breast cancer patients with and without endocrine therapy.[34] They found that patients with low tumor heterogeneity of ER in the tamoxifen-treated arm had an excellent 25-year breast cancer-specific survival of 88.3%, while patients with high heterogeneity of ER had a 79.6% survival independent of other known tumor markers.[34] We have shown that high heterogeneity in ER-positive cancers correlates with worse prognosis.

We further explored tumor heterogeneity among HER2-positive and -negative breast cancers, because a subset of HER2-positive cancers may also have HER2-negative regions, and the subpopulations of cancer cells in these tumors may differ depending on many factors. For instance, tumors tend to lose biomarker expression as they become more poorly differentiated, but it may retain a mix of clones with positive or negative expression depending on distinct genetic alterations. We were unable to find a correlation between heterogeneity and prognosis, either in tumors that were HER2-positive or -negative by pathology or in the prediction analysis of microarray 50 (PAM50) subgroup (electronic supplementary Fig. S1).

However, cancer mutation calls may vary depending on the algorithm used for the determination. PyClone is one such algorithm that determines clonal populations in tumors.[35] Another statistical model, known as PhyloSub, has the ability to infer the clonal evolution of tumors from single nucleotide somatic mutations.[36] THetA (Tumor Heterogeneity Analysis) infers not only the number of clones and clonal alterations but also the most likely collection of genes in those clonal populations using high-throughput DNA sequencing data.[37] ABSOLUTE is a method that quantifies tumor heterogeneity directly from analysis of somatic DNA alterations.[38] Heterogeneity, as measured by MATH, can be affected by the accuracy of the mutation calls. Together, all of these individual algorithms can determine cellular clonality using genetic information. The application of more than one algorithm in this study may have yielded additional data supporting the association between MATH and survival. Ascertaining detailed information on subpopulations of clones can provide insight into a tumor metastatic potential and possible therapeutic resistance.

We have also shown that tumors with high heterogeneity associate with fewer numbers of immune cell infiltrations. Immune cells may serve as a selective pressure, and shape tumor heterogeneity. TILs in ER-positive, HER2-negative cancers are rare. Gu-Trantien et al. found that 75% of the leukocyte infiltration in breast cancer were T cells, 20% were B cells, $< 10\%$ were monocytes, and $< 5\%$ were natural killer cells.[39] Given its amount, the association between TILs and tumor heterogeneity was of interest because it could be driving the observed survival differences. Our finding is consistent with those of Morris et al., where breast cancers with high heterogeneity tended to have lower levels of immune cell infiltration.[31] We found that high MATH tumor was associated with increased tumor-promoting Tregs, and with less anti-tumor CD4+ and CD8+ T cells, as well as mast cells, B cells, and dendritic cells. This result is in agreement with the previous report that mast cells are associated with favorable prognosis in invasive breast cancer.[40] Since neoantigens are detected by immune cells such as dendritic cells and B cells that form the antibodies against those antigens,[3] our result does not support the notion that tumors with high heterogeneity provide high levels of neoantigens that attract immune cells. In fact, TILs have not been shown to be prognostic in other studies. Our data are different because tumors with high heterogeneity are associated with less infiltration of immune cells, less activation of the immune response, and worse OS in breast cancer, thus supporting the notion that tumor heterogeneity may be shaped by the selection pressure of tumor-infiltrating immune cells, which can ultimately influence survival.

Other authors have explored the correlation between heterogeneity and immune cells in TNBC. Some TNBCs have T-cell-related immune signature that is associated with less likelihood to have distant metastasis, and more likely to achieve a pathologic complete response to neoadjuvant chemotherapy.[39,41] We were unable to identify any other reports that demonstrated a direct correlation between TILs and heterogeneity in TNBCs. Several potentially actionable mutations have been found in TNBC, such as P13k/mTOR, RAS/REF/MEK, however none have been shown to definitively result in the TNBC phenotype.[39]

We have also found that tumors with high heterogeneity have less immune response genes. PDL-1 is an immune checkpoint protein for T cells; IDO2 is an enzyme that breaks down tryptophan, which results in immunosuppression;[42] and ADORA2A suppresses CD8+ T cells in an in vitro model of melanoma.[43] VISTA and CCR4 also play a role in immunosuppression in cancer.[44] Furthermore, the notion that tumor heterogeneity is shaped by selection pressure of immune cells was further supported by the demonstration of less cytolytic activity. Perforin/granzyme-induced apoptosis is the main pathway used by cytotoxic lymphocytes to eliminate cancer cells.[45] High MATH tumors were associated with decreased expression of GZMA and PRF-1, both cytolysis-related proteins, as well as less CYT and more cell death.

CONCLUSIONS

High MATH tumors have lower levels of effector T cells, lower expression of exhaustion markers, and lower levels of gene expression associated with T-cell cytolytic activity. These data are suggestive that high MATH tumors are associated with a lower, active anti-tumor response and worse OS in breast cancer. Our results are consistent with the notion that tumor heterogeneity is at least partially shaped by selection pressure of immune response to the tumor. Further understanding of the molecular, genetic, and cellular changes that influence tumor heterogeneity will allow us to better understand the changes leading to high heterogeneity, and develop methods to prevent it with resulting improvement of prognosis of patients with breast cancer.

AUTHOR CONTRIBUTIONS Conception and design: KAM, TK, and KT. Development of methodology: TK, LY, and KT. Acquisition of data (acquired and managed patients, provided facilities, etc.): TK, QQ, XP, MA, and LY. Analysis and interpretation of data (e.g. statistical analysis, biostatistics, computational analysis): TK, QQ, XP, MA, and LY. Writing, review, and/or revision of the manuscript: KAM, TK, MO, JY, SP, EO, and KT. Administrative, technical, or material support (i.e. reporting or organizing data, constructing databases): QQ, XP, and LY. Study supervision: KT and TK.

DISCLOSURES Kazuaki Takabe is funded by the United States National Institute of Health – National Cancer Institute (R01CA160688) and Susan G. Komen Foundation (investigator-initiated research Grant (IIR12222224). This work was also supported by National Cancer Institute Grant P30CA016056 involving the use of Roswell Park Cancer Institute's Bioinformatics and Biostatistics shared resources. Biospecimens or research pathology services for this study were provided by the Pathology Resource Network. Clinical Data Delivery and Honest Broker services for this study were provided by the Clinical Data Network. Kerry-Ann McDonald, Tsutomu Kawaguchi, Qianya Qi, Xuan Peng, Mariko Asaoka, Jessica Young, Mateusz Opyrchal, Li Yan, Santosh Patnaik, and Eigo Otsuji have no conflicts of interest to declare.

REFERENCES

1. Ma D, Jiang YZ, Liu XY, Liu YR, Shao ZM. Clinical and molecular relevance of mutant-allele tumor heterogeneity in breast cancer. *Breast Cancer Res Treat*. 2017;162(1):39–48.
2. Martelotto LG, Ng CK, Piscuoglio S, Weigelt B, Reis-Filho JS. Breast cancer intra-tumor heterogeneity. *Breast Cancer Res*. 2014;16(3):210.
3. McGranahan N, Furness AJ, Rosenthal R, et al. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science*. 2016;351(6280):1463–69.
4. McGranahan N, Swanton C. Clonal heterogeneity and tumor evolution: past, present, and the future. *Cell*. 2017;168(4):613–28.
5. Marusyk A, Almendro V, Polyak K. Intra-tumour heterogeneity: A looking glass for cancer? *Nat Rev Cancer*. 2012;12(5):323–34.
6. Andor N, Graham TA, Jansen M, et al. Pan-cancer analysis of the extent and consequences of intratumor heterogeneity. *Nat Med*. 2016;22(1):105–13.
7. Mroz EA, Rocco JW. MATH, a novel measure of intratumor genetic heterogeneity, is high in poor-outcome classes of head and neck squamous cell carcinoma. *Oral Oncol*. 2013;49(3):211–15.
8. Rajput A, Bocklage T, Greenbaum A, Lee JH, Ness SA. Mutant-allele tumor heterogeneity scores correlate with risk of metastases in colon cancer. *Clin Colorectal Cancer*. 2017;16(3):e165–e170.
9. PLOS Medicine Staff. Correction: intra-tumor genetic heterogeneity and mortality in head and neck cancer: Analysis of data from the cancer genome atlas. *PLoS Med*. 2015;12(6):e1001844.
10. 10. Burrell RA, McGranahan N, Bartek J, Swanton C. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature*. 2013;501(7467):338–45.
11. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature*. 2012;490(7418):61–70.
12. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov*. 2012;2(5):401–404.
13. Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*. 2013;6(269):pl1.
14. Ramanathan R, Raza A, Sturgill J, et al. Paradoxical association of postoperative plasma sphingosine-1-phosphate with breast cancer aggressiveness and chemotherapy. *Mediators Inflamm*. 2017;2017:5984819.
15. Ramanathan R, Olex AL, Dozmorov M, Bear HD, Fernandez LJ, Takabe K. Angiopoietin pathway gene expression associated with poor breast cancer survival. *Breast Cancer Res Treat*. 2017;162(1):191–98.
16. Kim SY, Kawaguchi T, Yan L, Young J, Qi Q, Takabe K. Clinical relevance of microRNA expressions in breast cancer validated using the cancer genome atlas (TCGA). *Ann Surg Oncol*. 2017;24(10):2943–49.
17. Narayanan S, Kawaguchi T, Yan L, Peng X, Qi Q, Takabe K. Cytolytic activity score to assess anticancer immunity in colorectal cancer. *Ann Surg Oncol*. 2018;25:2323–31
18. Moro K, Kawaguchi T, Tsuchida J, et al. Ceramide species are elevated in human breast cancer and are associated with less aggressiveness. *Oncotarget*. 2018;9(28):19874–890.
19. Kawaguchi T, Yan L, Qi Q, et al. Overexpression of suppressive microRNAs, miR-30a and miR-200c are associated with improved survival of breast cancer patients. *Sci Rep*. 2017;7(1):15945
20. Young J, Kawaguchi T, Yan L, Qi Q, Liu S, Takabe K. Tamoxifen sensitivity-related microRNA-342 is a useful biomarker for breast cancer survival. *Oncotarget*. 2017;8(59):99978–989.
21. Kawaguchi T, Yan L, Qi Q, et al. Novel MicroRNA-based risk score identified by integrated analyses to predict metastasis and poor prognosis in breast cancer. *Ann Surg Oncol*. 2018;25(13):4037–46.
22. Kawaguchi T, Narayanan S, Takabe K. ASO author reflections: "From computer to bedside": a new translational approach to immunogenomics. *Ann Surg Oncol*. 2018;25 Suppl 3:846–47.
23. Rocco JW. Mutant allele tumor heterogeneity (MATH) and head and neck squamous cell carcinoma. *Head Neck Pathol*. 2015;9(1):1–5.
24. Budczies J, Klauschen F, Sinn BV, et al. Cutoff finder: A comprehensive and straightforward web application enabling rapid biomarker cutoff optimization. *PLoS One*. 2012;7(12):e51862.
25. Chang C, Hsieh MK, Chang WY, Chiang AJ, Chen J. Determining the optimal number and location of cutoff points with application to data of cervical cancer. *PLoS ONE*. 2017;12(4):e0176231.
26. 26. Mazumdar M, Glassman JR. Categorizing a prognostic variable: Review of methods, code for easy implementation and

applications to decision-making about cancer treatments. *Stat Med*. 2000;19(1):113-132.
27. 27. Brondum L, Eriksen JG, Singers Sorensen B, et al. Plasma proteins as prognostic biomarkers in radiotherapy treated head and neck cancer patients. *Clin Transl Radiat Oncol*. 2017;2:46-52.
28. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12(5):453–57
29. Charoentong P, Finotello F, Angelova M, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep*. 2017;18(1):248–62.
30. Turashvili G, Brogi E. Tumor heterogeneity in breast cancer. *Front Med (Lausanne)*. 2017;4:227.
31. Morris LG, Riaz N, Desrichard A, et al. Pan-cancer analysis of intratumor heterogeneity as a prognostic determinant of survival. *Oncotarget*. 2016;7(9):10051–063.
32. Ellsworth RE, Blackburn HL, Shriver CD, Soon-Shiong P, Ellsworth DL. Molecular heterogeneity in breast cancer: state of the science and implications for patient care. *Semin Cell Dev Biol*. 2017;64:65–72.
33. Szulwach KE, Chen P, Wang X, et al. Single-cell genetic analysis using automated microfluidics to resolve somatic mosaicism. *PLoS One*. 2015;10(8):e0135007.
34. Lindström LS, Yau C, Czene K, et al. Intratumor heterogeneity of the estrogen receptor and the long-term risk of fatal breast cancer. *J Natl Cancer Inst*. 2018;110(7):726–33.
35. Roth A, Khattra J, Yap D, et al. PyClone: Statistical inference of clonal population structure in cancer. *Nat Methods*. 2014;11(4):396–98.
36. Jiao W, Vembu S, Deshwar AG, Stein L, Morris Q. Inferring clonal evolution of tumors from single nucleotide somatic mutations. *BMC Bioinformatics*. 2014;15:35–2105-15-35.
37. Oesper L, Mahmoody A, Raphael BJ. THetA: Inferring intra-tumor heterogeneity from high-throughput DNA sequencing data. *Genome Biol*. 2013;14(7):R80
38. Carter SL, Cibulskis K, Helman E, et al. Absolute quantification of somatic DNA alterations in human cancer. *Nat Biotechnol*. 2012;30(5):413–21.
39. Gu-Trantien C, Loi S, Garaud S, et al. CD4(+) follicular helper T cell infiltration predicts breast cancer survival. *J Clin Invest*. 2013;123(7):2873–92.
40. Rajput AB, Turbin DA, Cheang MC, et al. Stromal mast cells in invasive breast cancer are a marker of favourable prognosis: A study of 4,444 cases. *Breast Cancer Res Treat*. 2008;107(2):249–57.
41. Bianchini G, Balko JM, Mayer IA, Sanders ME, Gianni L. Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. *Nat Rev Clin Oncol*. 2016;13(11):674–90.
42. Munn DH, Mellor AL. Indoleamine 2,3 dioxygenase and metabolic control of immune responses. *Trends Immunol*. 2013;34(3):137–43.
43. Cekic C, Linden J. Adenosine A2A receptors intrinsically regulate CD8+ T cells in the tumor microenvironment. *Cancer Res*. 2014;74(24):7239–49.
44. Lines JL, Sempere LF, Broughton T, Wang L, Noelle R. VISTA is a novel broad-spectrum negative checkpoint regulator for cancer immunotherapy. *Cancer Immunol Res*. 2014;2(6):510–17.
45. Trapani JA, Smyth MJ. Functional significance of the perforin/granzyme cell death pathway. *Nat Rev Immunol*. 2002;2(10):735–47.