ESMO GOOD SCIENCE BETTER MEDICINE BEST PRACTICE

ESMO OPEN SCIENCE FOR OPTIMAL CANCER CARE

ORIGINAL ARTICLE

Linking tumor immune infiltration to enhanced longevity in recurrence-free breast cancer

L. Angelats[1,2,3†], L. Paré[4†], C. Rubio-Perez[1,2†], E. Sanfeliu[1,5], A. González[5,6], E. Seguí[1,2], G. Villacampa[7], M. Marín-Aguilera[4], S. Pernas[8,9], B. Conte[1,2], V. Albarrán-Fernández[1,2], O. Martínez-Sáez[1,2,3], Á. Aguirre[1,2], P. Galván[1,2,4], A. Fernandez-Martinez[10], S. Cobo[1,2], M. Rey[1,2], A. Martínez-Romero[1,2], B. Walbaum[1,2], F. Schettini[1,2], M. Vidal[1,2], W. Buckingham[4], M. Muñoz[1,2], B. Adamo[1,2], Y. Agrawal[11], S. Guedan[12], T. Pascual[1,2,3,13], J. Agudo[14,15,16,17], M. Grzelak[18], N. Borcherding[18], H. Heyn[18], A. Vivancos[19], J. S. Parker[11], P. Villagrasa[4], C. M. Perou[10,11], A. Prat[1,2,3,4,20‡] & F. Brasó-Maristany[1,2,4*‡]

[1]Translational Genomics and Targeted Therapies in Solid Tumors group, August Pi i Sunyer Biomedical Research Institute (IDIBAPS), Barcelona; [2]Institute of Cancer and Blood Diseases, Hospital Clinic of Barcelona, Barcelona; [3]Department of Medicine, University of Barcelona, Barcelona; [4]Reveal Genomics, Barcelona; [5]Department of Pathology, Hospital Clinic of Barcelona, Barcelona; [6]Immunogenetics and Immunotherapy in Autoinflammatory and Immune Responses, August Pi i Sunyer Biomedical Research Institute (IDIBAPS), Barcelona; [7]Statistics Unit, Vall d'Hebron Institute of Oncology, Barcelona; [8]Institut Catala d'Oncologia, IDIBELL, L'Hospitalet de Llobregat, Barcelona; [9]Institut Català d'Oncologia Medical Oncology, Breast Unit Barcelona, Barcelona, Spain; [10]Department of Genetics, University of North Carolina, Chapel Hill; [11]Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, USA; [12]Cellular Immunotherapies for Cancer Group, August Pi i Sunyer Biomedical Research Institute (IDIBAPS), Barcelona; [13]SOLTI Cancer Cooperative Group, Barcelona, Spain; [14]Department of Cancer Immunology and Virology, Dana-Farber Cancer Institute, Boston; [15]Department of Immunology, Harvard Medical School, Boston; [16]Parker Institute for Cancer Immunotherapy at Dana-Farber Cancer Institute, Boston; [17]New York Stem Cell Foundation, Robertson Investigator, New York; [18]Omniscope Inc., Palo Alto, USA; [19]Vall d'Hebron Institute of Oncology (VHIO), Cancer Genomics Group, Barcelona; [20]Institute of Oncology (IOB)—Hospital Quirónsalud, Barcelona, Spain

Available online 6 January 2025

Background: The infiltration of tumor-infiltrating B cells and plasma cells in early-stage breast cancer has been associated with a reduced risk of distant metastasis. However, the influence of B-cell tumor infiltration on overall patient survival remains unclear.

Materials and methods: This study explored the relationship between an antitumor immune response, measured by a 14-gene B-cell/immunoglobulin (IGG) signature, and mortality risk in 9638 breast cancer patients across three datasets. Associations with tumor subtype, stage, and age were examined. IGG was characterized using spatial GeoMx profiling and single-cell RNA sequencing, and its relationship with tertiary lymphoid structures (TLSs) was evaluated. The predictive value of each of the 14 IGG genes for B-cell receptor (BCR) and T-cell receptor (TCR) clonality and longevity was also assessed, along with its association with longevity in other cancer types.

Results: High IGG signature expression was significantly associated with a 41%-47% reduction in death risk in breast cancer survivors ($P < 0.001$), regardless of age, tumor stage, or subtype. Similar associations were observed in other cancers, including melanoma. In breast cancer, the IGG signature was significantly linked to overall survival without relapse in patients aged 41-70 years at diagnosis. Additionally, IGG expression correlated with the presence of TLSs and higher B- and T-cell polyclonality. A specific subset of seven IGG genes strongly correlated with BCR and TCR clonality, with predictive power for identifying clonality and improved longevity, especially when combining two of these genes.

Conclusions: This study uncovers a significant link between immune gene expression in tumors and extended longevity in breast cancer survivors, even in the absence of recurrence. The IGG signature, particularly its key gene subset, emerges as a powerful marker of sustained antitumor immunity and overall patient fitness. These findings pave the way for personalized treatment strategies that enhance both survival and long-term health outcomes.

Key words: breast cancer, immune gene expression, longevity, clonality

*Correspondence to: Dr Fara Brasó-Maristany, Institute of Cancer and Blood Diseases, Hospital Clinic of Barcelona, Barcelona, Spain. Tel: +0034 93 227 54 00 (ext 4801) E-mail: FBRASO@recerca.clinic.cat (F. Brasó-Maristany).

†These authors contributed equally.
‡These authors are senior authors.

INTRODUCTION

The 14-gene B-cell/immunoglobulin (IGG) signature has emerged as a robust prognostic marker in early-stage breast cancer.[1] It comprises genes involved in lymphocyte progenitor maturation, activation, differentiation, immunoglobulin production, chemotaxis, and lymphocyte activity regulation; it was first identified through unsupervised clustering of 550 node-negative breast tumors,[1] and has been externally validated across diverse clinical cohorts involving over 1000 patients.[2,3] This signature is incorporated into the 27-gene HER2DX genomic assay, designed for early-stage human epidermal growth factor receptor 2 (HER2)-positive disease.[1,2,4-9] The HER2DX assay integrates clinicopathological and genomic data to generate two distinct scores, which predict disease-free survival and the likelihood of achieving a pathological complete response (pCR) following neoadjuvant trastuzumab-based therapy,[2] and has been shown to have statistically significant association with both event-free survival and overall survival (OS).[5] Notably, within the HER2DX framework, a high IGG signature expression correlates with increased pCR rates and improved survival outcomes.[2-5] High IGG signature expression is also predictive of pCR, event-free survival, and OS in patients with early-stage triple-negative breast cancer (TNBC).[9,10]

Adaptive immunity, known for its high specificity and ability to provide long-term antigen-specific memory, might help explain the strong correlation between the IGG signature, breast cancer recurrence, and OS.[2,5] However, the predictive power of immune tumor infiltration for longevity, especially in the absence of disease recurrence, remains uncertain. Bridging this knowledge gap could substantially deepen our understanding of the immune system's persistent impact on patient outcomes, potentially leading to improved prognostic tools and therapeutic strategies.

MATERIALS AND METHODS

Patient datasets

The Molecular Taxonomy of Breast Cancer International Consortium (METABRIC)[11] ($n = 1904$) and The Cancer Genome Atlas (TCGA)[12] ($n = 1082$) breast cancer datasets and pan-cancer TCGA datasets [i.e. cervix ($n = 294$), head and neck ($n = 515$), lung ($n = 510$), melanoma ($n = 443$), and sarcoma ($n = 253$)] were sourced from the cBioPortal for Cancer Genomics (http://cbioportal.org),[11] while data from the most recent version of The Sweden Cancerome Analysis Network (SCAN-B)[13,14] ($n = 6652$) dataset were obtained from Mendeley Data (https://data.mendeley.com/datasets/yzxtxn4nmd). Except for the competing risk analyses, patients who experienced breast cancer recurrence were systematically excluded from both datasets. The final analysis comprised 1133 patients from METABRIC, 4470 patients from SCAN-B, and 936 patients with breast cancer from TCGA, ensuring a cohort that did not encounter recurrence during the follow-up period. The median follow-up in these patients was 96.9 months in SCAN-B, 185.0 months in METABRIC, and 130.0 months in TCGA. Gene expression was obtained from all datasets. Additionally, *TP53*, *PIK3CA*, and *RB1* somatic mutation data (i.e. missense, inframe, splicing, truncating) were downloaded for TCGA and METABRIC. Normal breast tissue gene expression data from 168 women were downloaded from GTEx (https://gtexportal.org/home/). RNA sequencing (RNA-seq) and B-cell receptor (BCR) and T-cell receptor (TCR) repertoire metrics from the CALGB-40601 neoadjuvant HER2-positive study were obtained from a previously published study.[15]

In silico gene expression analysis

Z-scores (METABRIC, TCGA) or log2-transformed (SCAN-B) normalized gene expression values were employed for single genes and for deriving the 14-gene IGG signature, depending on the available gene expression data for each dataset (refer to Supplementary Table S8, available at https://doi.org/10.1016/j.esmoop.2024.104109). The Prediction Analysis of Microarray 50 (PAM50) risk of recurrence (ROR) was also determined for METABRIC and TCGA.

Core IGG genes

Genes consistently associated with longevity across METABRIC, SCAN-B, and TCGA datasets were identified as core IGG genes (CIGGs). We further investigated whether the combined assessment of pairs of CIGGs enhanced longevity prediction compared with individual genes. Various mathematical operations, including addition, subtraction, multiplication, and ratio, were employed to combine the expressions of two CIGGs.

Clinical samples

To study the immune component of breast cancers with different levels of IGG signature, 132 formalin-fixed paraffin-embedded (FFPE) samples from patients with early-stage breast cancer treated at Hospital Clinic of Barcelona ($n = 122$) and Institut Català d'Oncologia ($n = 10$) were selected, including 35 hormone receptor-positive (HR+)/HER2-negative (HER2−), 84 HER2-positive, and 13 TNBC samples. The proportion of tumor-infiltrating lymphocytes (%TILs) was determined on hematoxylin−eosin (H&E) slides according to International TILs Working Group Guidelines.[16] The presence of tertiary lymphoid structures (TLSs) was defined as spatially organized, non-encapsulated areas of immune cell aggregates on H&E slides. Additionally, six biopsies from patients with untreated HER2-positive breast cancer were obtained to conduct single-cell RNA-seq.

Gene expression analysis of clinical samples

RNA was extracted from FFPE tumor diagnostic samples using the High Pure FFPET RNA Isolation Kit (Roche, Indianapolis, IN). One to five 10-μm FFPE slides depending on tumor cellularity were used for each tumor sample, and macrodissection was carried out, when needed, to avoid

normal tissue contamination. A minimum of 100 ng of total RNA was analyzed on the nCounter platform (NanoString Technologies, Seattle, WA) using a 192-gene custom panel, including 14 genes of the IGG signature, which was determined using R software v4.0.3 (R Foundation for Statistical Computing, Vienna, Austria).

GeoMx DSP data acquisition and analysis

FFPE tumor biopsies from 135 patients were stained with fluorescently labeled CD45 and PanCK antibodies, DAPI (4′,6-diamidino-2-phenylindole), and a multiplexed panel of protein antibodies that contained a photocleavable indexing oligonucleotide, enabling subsequent readouts. Regions of interest were selected on the GeoMx DSP platform (NanoString Technologies[17]), segmented according to CD45 and PanCK staining and illuminated using UV light. Released indexing oligonucleotides from each area of interest (AOI) were collected and deposited into designated wells on a microtiter plate, allowing for well indexing of each AOI during nCounter (NanoString Technologies) readout. For each tissue sample, counts for each marker were obtained from an average of 4.44 (range 1-12) CD45-positive (CD45+) AOIs. Raw protein counts for each marker in each AOI were generated using nCounter. The raw counts were normalized with the AOI surface area. All normalized counts were log2 transformed.

TCR and BCR sequencing

DNA from FFPE tumor samples with different IGG signature levels was purified with the QIAamp DNA FFPE Tissue Kit (QIAGEN, Hilden, Germany) according to manufacturer's instructions and BCR and TCR were sequenced from 23 and 33 DNA samples, respectively, using the immunoSEQ Assay (Adaptive Biotechnologies, Seattle, WA). The somatically rearranged *Homo sapiens* TCR locus complementarity-determining region 3 (CDR3) and BCR immunoglobulin heavy chain (IgH) locus CDR3 were amplified from DNA samples using a two-step, amplification bias-controlled multiplex PCR approach. CDR3 and reference gene libraries were sequenced on an Illumina (San Diego, CA) instrument according to the manufacturer's instructions. Raw sequence reads were demultiplexed according to Adaptive's proprietary barcode sequences. Demultiplexed reads were further processed to remove adapter and primer sequences, and identify and remove primer dimer, germline, and other contaminant sequences. The filtered data were clustered using both the relative frequency ratio between similar clones and a modified nearest-neighbor algorithm, to merge closely related sequences to correct for technical errors introduced through PCR and sequencing. The resulting sequences allowed annotation of the V, D, and J genes and the N1 and N2 regions constituting each unique CDR3 and the translation of the encoded CDR3 amino acid sequence. Gene definitions were based on annotation in accordance with the IMGT database (www.imgt.org). The set of observed biological TCR and BCR IgH CDR3 sequences was normalized to correct for residual multiplex PCR amplification bias and quantified against a set of synthetic CDR3 sequence analogues. Data were analyzed using the immunoSEQ Analyzer toolset.

Single-cell RNA sequencing

Fresh core biopsies obtained from six patients with HER2-positive breast cancer were collected by ultrasound-guided breast biopsy and the tissue ($\sim$3-5 mm^3) was supplied in 5 ml of RPMI medium. Tumor fragments were dissociated using enzymatic digestion (Human Tumor Dissociation Kit, Miltenyi, Bergisch Gladbach, Germany). Then, CD45+ cells were isolated using Human CD45 TIL MicroBeads (Miltenyi) and stored in phosphate-buffered saline containing 0.005% bovine serum albumin. CD45+ cells were processed with the 10× Genomics Chromium Controller (10X Genomics Inc.) and single-cell gene expression and TCR/BCR libraries were produced with the Chromium Single Cell 5′ Library assay (10X Genomics Inc., Newark, CA), sequenced on an Illumina NovaSeq6000 (100 cycle kit). Sequencing reads were aligned with CellRanger Single Cell Software (version 6.1.1, 10X Genomics Inc.) and mapped against the human GRCh38 reference genome (GENCODE v32/Ensembl 98). Quality control, cell annotation, and cluster identification were carried out based on rationale.[18]

Statistical analysis

Cumulative incidence functions were used to estimate death without recurrence in a competing risk setting using recurrence as a competing event. Fine and Gray competing risk regression was used to obtain a sub-distribution hazard ratio with 95% confidence interval (CI). In the sensitivity analysis of the subset of patients without recurrence, Kaplan—Meier curves were used to estimate survival outcomes and the log-rank test was used for statistical comparisons. Univariate and multivariable Cox regression and logistic regression models were employed, with the goodness of fit of each model assessed through the chi-square statistic obtained from a likelihood ratio test. The differences in variable distribution were explored by chi-square test. Correlations between two variables were evaluated using the Pearson method. Unpaired *t*-tests and analyses of variance were employed to ascertain the comparative analysis of numerical variables among different groups. A two-sided α of 0.05 was set as the significance level for all statistical analyses. No data imputation was carried out. All statistical analysis were carried out using the R software.

RESULTS

IGG, age, and longevity

We investigated the association between the IGG signature and OS in patients with early-stage breast cancer who have not experienced a relapse, a condition we defined as 'longevity' (Figure 1A). Leveraging comprehensive gene expression and clinical data from three distinct cohorts of

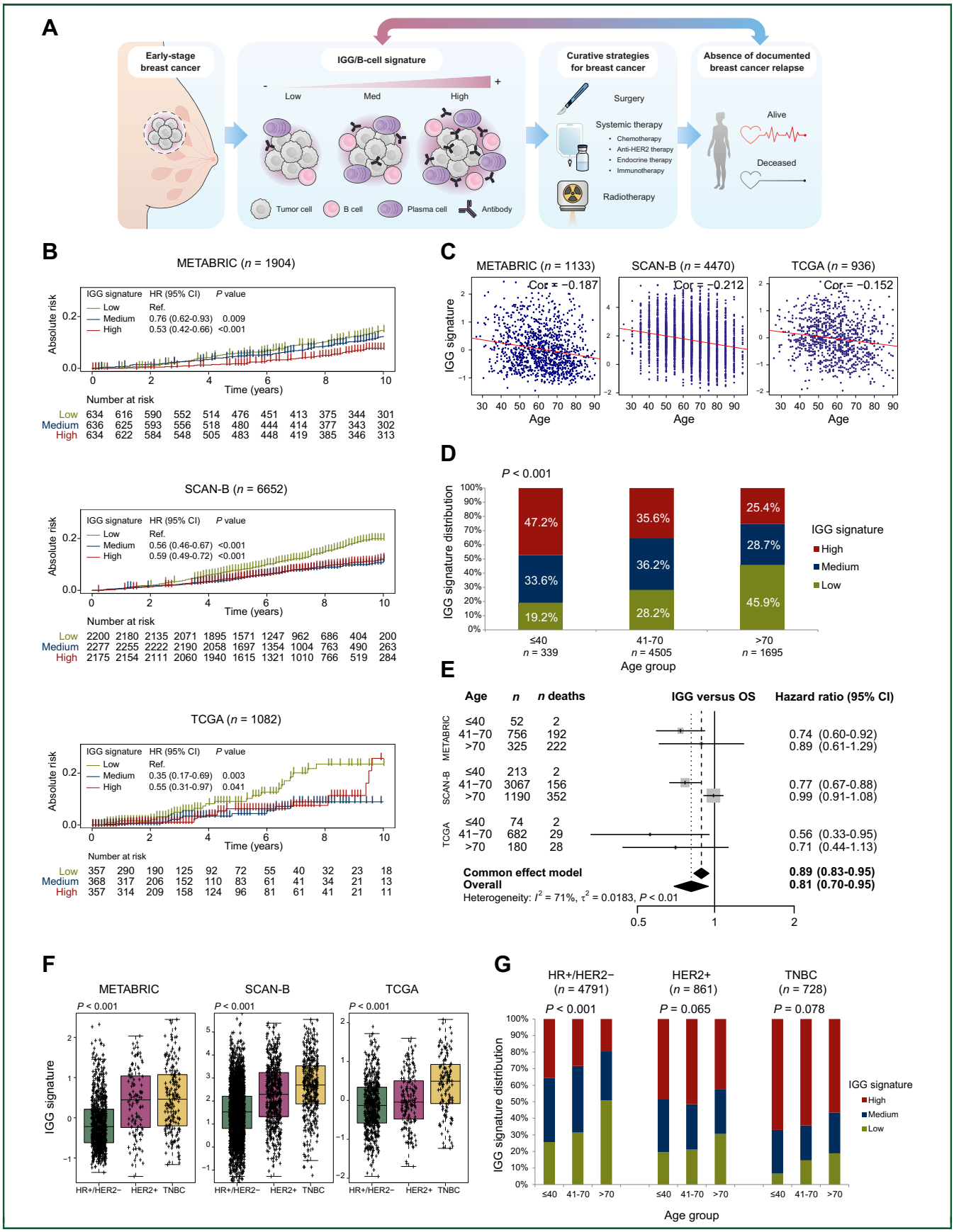

Figure 1. Assessment of the 14-gene IGG signature for predicting longevity in breast cancer survivors without a documented recurrence. (A) Schematic representation of the study approach. (B) Cumulative risk plots with censored recurrences illustrating the association of IGG expression with OS in all patients from METABRIC (n = 1940), SCAN-B (n = 6652), and TCGA datasets (n = 1082). (C) Correlation analysis depicting the relationship between the IGG signature and age in all patients without documented breast cancer recurrence during follow-up in METABRIC (n = 1133), SCAN-B (n = 4470), and TCGA datasets (n = 936).

9638 patients with early-stage breast cancer and long-term follow-up (i.e. METABRIC,[11] SCAN-B,[13,14] and TCGA,[12] Supplementary Table S1, available at https://doi.org/10.1016/j.esmoop.2024.104109), we observed a significant association between IGG expression and OS without recurrence both treating recurrences as a competing event (Figure 1B) and excluding recurrences (Supplementary Figure S1, available at https://doi.org/10.1016/j.esmoop.2024.104109), independently from breast cancer subtype, tumor stage, or nodal status (Supplementary Table S2, available at https://doi.org/10.1016/j.esmoop.2024.104109). This association persisted in patients with a very low risk of relapse (Supplementary Figure S1, available at https://doi.org/10.1016/j.esmoop.2024.104109), and patients with >5 years of follow-up without a documented recurrence. Moreover, we analyzed the relationship between the PAM50 ROR and the IGG signature in the TCGA and METABRIC datasets. Differences in PAM50 ROR distribution were observed across IGG groups, showing that IGG-high tumors are enriched in both ROR-high and ROR-low categories, with a lower proportion of ROR-medium, particularly within HR+/HER2− breast cancer (Supplementary Figure S2A, available at https://doi.org/10.1016/j.esmoop.2024.104109). From a longevity perspective, the PAM50 ROR was not significantly associated with OS in recurrence-free breast cancer survivors (Supplementary Figure S2B, available at https://doi.org/10.1016/j.esmoop.2024.104109), whereas the IGG signature remained independently associated with survival in this population, even after adjusting for ROR (Supplementary Table S3, available at https://doi.org/10.1016/j.esmoop.2024.104109).

A weak negative correlation between the IGG signature and age was observed (Figure 1C and D). However, the IGG signature was significantly associated with OS without a relapse in patients aged 41-70 years at diagnosis, which represented 66.7%-72.9% of all patients (Figure 1E). Across breast cancer subtypes, the expression of the IGG signature was higher in HER2-positive and TNBC than in HR+/HER2− (Figure 1F); however, 26.2% of HR+/HER2− tumors were IGG-high. Finally, IGG decreased with age across breast cancer subtypes, especially in patients >70 years of age with HR+/HER2− disease (Figure 1G). Additionally, to address this, we analyzed the association of *TP53*, *PIK3CA*, and *RB1* mutations with IGG groups across breast cancer subtypes in the TCGA and METABRIC datasets (Supplementary Figures S3 and S4, available at https://doi.org/10.1016/j.esmoop.2024.104109). Interestingly, *TP53* mutations were more frequently observed in IGG-high tumors. Despite this association, the IGG signature remained independently and significantly associated with OS in recurrence-free patients in a multivariable analysis (Supplementary Table S4, available at https://doi.org/10.1016/j.esmoop.2024.104109). In contrast, no significant differences were found in the frequency of *PIK3CA* or *RB1* mutations across IGG groups (Supplementary Figures S3 and S4, available at https://doi.org/10.1016/j.esmoop.2024.104109). Of note, the IGG signature assessed in normal breast tissue was not associated with age (Supplementary Figure S5, available at https://doi.org/10.1016/j.esmoop.2024.104109).

IGG, spatial profiling, and TLSs

The IGG signature is measured from bulk RNA obtained from a tumor sample.[2] To gain deeper insights of the tumor microenvironment (TME) and its relationship with the IGG signature, we carried out GeoMx spatial profiling of stroma from 132 patients with breast cancer, which had been previously categorized as IGG-high, IGG-medium, and IGG-low, according to pre-established cut-offs of the HER2DX genomic test (Figure 2A and Supplementary Figure S6, available at https://doi.org/10.1016/j.esmoop.2024.104109). We selected 458 CD45+ TME AOIs, and we measured 42 immune-related proteins (Supplementary Data S1, available at https://doi.org/10.1016/j.esmoop.2024.104109). Compared with IGG-medium and IGG-low tumors, IGG-high tumors showed increased expression of plasma cells (CD27), activated T cells, or antigen-presenting cells (programmed death-ligand 1 and CD40) and decreased expression of M2-like macrophages (CD163) and stromal fibronectin (Supplementary Figure S6, available at https://doi.org/10.1016/j.esmoop.2024.104109). Additionally, we measured the 42 immune-related proteins in 134 TLS-selected AOIs, and compared their expression to the TME AOIs. TLSs were defined as spatially organized, non-encapsulated areas of immune cell aggregates. We observed an association between the IGG signature expression and the presence of TLSs in the tumor stroma (Figure 2B). Specifically, 55.2% of tumors with TLSs were IGG-high compared with 25% of tumors without TLSs ($P < 0.001$) (Figure 2B). To better understand the biology associated with TLSs, we focused our attention on the expression of immune proteins in the 134 TLS regions using spatial profiling. Twenty-one proteins (51.2%, 21/41) were found differentially expressed between TLS and non-TLS TME regions (false discovery rate <5%); among them, CD20 (B cell), CD27 (plasma cell), CD11c (dendritic cells), and CD3 (T-cell related) were found more expressed in TLS compared with non-TLS regions, while smooth muscle actin, fibronectin, and CD163 (M2-like macrophages) were found downregulated in TLS compared with non-TLS regions (Supplementary Figure S7A, available at https://doi.org/10.1016/j.esmoop.2024.104109). Using the list of the

(D) Distribution of the IGG signature (low, med, high) across age groups in all patients without documented breast cancer recurrence. Chi-square *P* value <0.001. (E) Forest plot presenting the association of the IGG signature with OS in all patients without documented breast cancer recurrence across different age groups, and in the pooled analysis of the three datasets (in bold). (F) IGG signature expression across breast cancer clinical subtypes in all patients without documented breast cancer recurrence. (G) Distribution of the IGG signature (low, med, high) and age groups, stratified by breast cancer subtype in all patients without documented breast cancer recurrence.

CI, confidence interval; HER2, human epidermal growth factor receptor 2; HR, hazard ratio; HR+, hormone receptor-positive; IGG, 14-gene B-cell/immunoglobulin; med, medium; OS, overall survival.

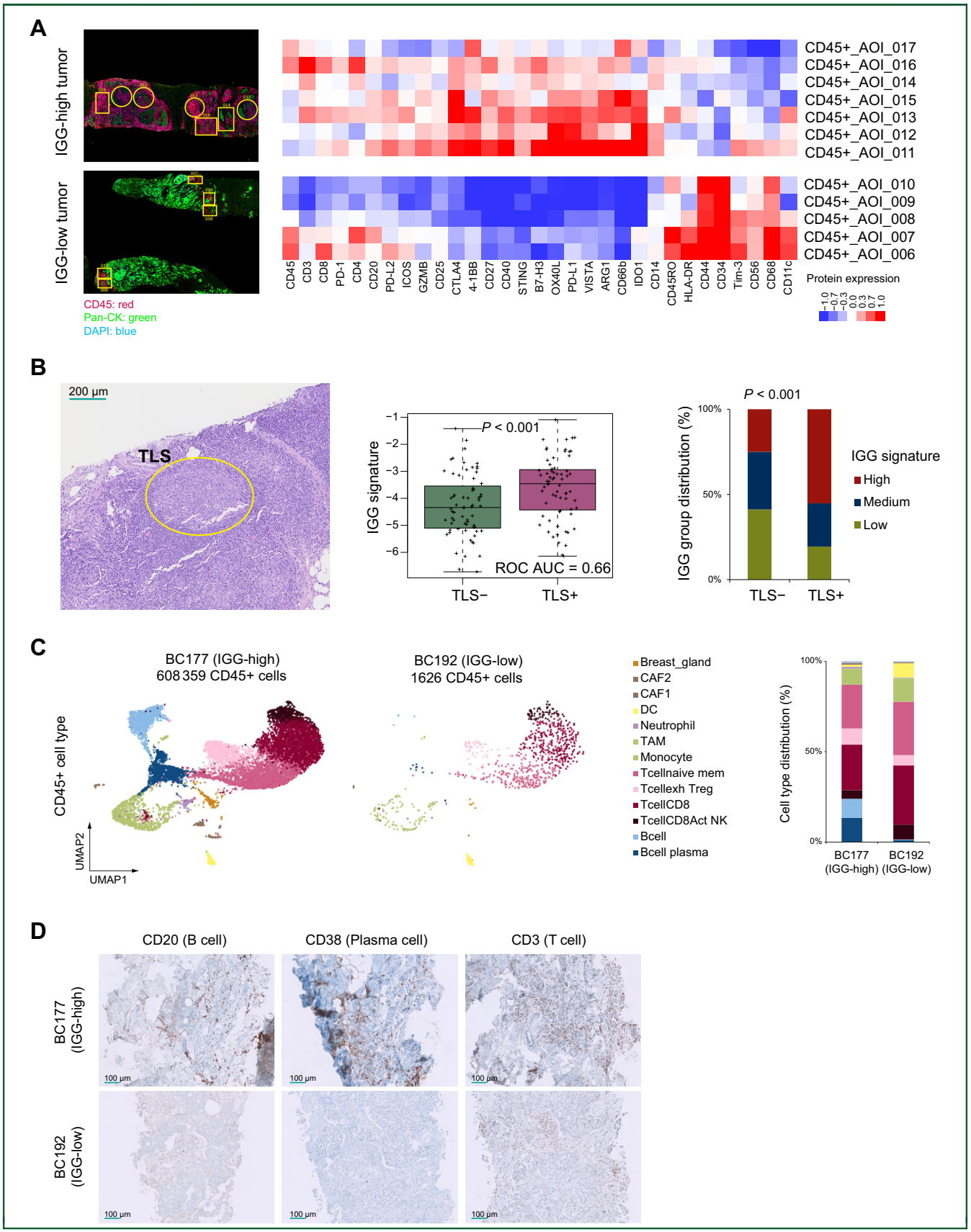

Figure 2. Characterization of the 14-gene IGG signature in breast cancer. (A) Exemplification of protein-based spatial profiling in an IGG-high tumor versus an IGG-low tumor using the GeoMx digital spatial profiling platform. Green denotes tumor cells, red represents immune cells, and blue signifies cell nuclei (DNA). Selected ROIs are highlighted in yellow. The heatmap represents the expression on immune proteins in CD45+ ROIs in each sample. (B) Hematoxylin–eosin staining image depicting a TLS breast tumor (left); IGG signature expression in tumors without TLSs ($n = 69$) versus tumors with TLSs ($n = 66$) (middle); and distribution of IGG signature high, medium, and low expression groups in TLS-negative and TLS-positive tumors (right). (C) UMAP plots showing cell types identified using single-cell RNA sequencing

significant proteins, we established a TLS-related protein signature (Supplementary Data S1, available at https://doi.org/10.1016/j.esmoop.2024.104109). We applied this signature to tumor samples that did not visually display TLSs. Notably, we observed a significantly increased expression of this signature in IGG-high tumors compared with those with IGG-medium/low levels (Supplementary Figure S7B, available at https://doi.org/10.1016/j.esmoop.2024.104109). This result underscores the IGG signature's capacity to reflect TLS-like biological processes in the tumor stroma, even in the absence of microscopically observable TLSs. Additionally, IGG is significantly associated with TILs (Supplementary Figure S8, available at https://doi.org/10.1016/j.esmoop.2024.104109).

IGG and CD45+ single-cell RNA sequencing

To explore the intricacies of the IGG signature at the cell level, single-cell RNA-seq of fresh CD45+ cells was carried out from six independent breast cancer biopsies with different IGG signature expression levels (Supplementary Figure S9, available at https://doi.org/10.1016/j.esmoop.2024.104109). In the tumor with the highest IGG signature score (BC177), 24.0% of CD45+ cells were identified as activated B cells and plasma cells, while these represented only 1.4% of all CD45+ cells in the sample with the lowest IGG signature score (BC192) (Figure 2C). The RNA-seq results observed in these two cases were confirmed by immunohistochemistry staining of CD20 (B cell), CD38 (plasma cell), and CD3 (T cell) (Figure 2D).

Interplay of IGG and B/T-cell clonality

Next, we analyzed BCR and TCR clonality in 23 and 33 breast tumors with known IGG expression values, respectively, using the immunoSEQ DNA-based assay. Overall, a higher polyclonality of B cells and T cells (i.e. lower Simpson clonality index and higher Shannon entropy) was observed in IGG-high tumors compared with IGG-medium/low tumors (Figure 3A and Supplementary Figure S10A, available at https://doi.org/10.1016/j.esmoop.2024.104109). Similar results were obtained from an external *in silico* validation breast cancer dataset ($n = 264$, CALGB-40601[13]) based on bulk RNA-seq data (Figure 3B).

Core IGG genes, longevity, and B/T-cell clonality

We then assessed the association of each IGG gene with longevity in the METABRIC, SCAN-B, and TCGA datasets. Seven genes [i.e. *TNFRSF17* (BCMA), *IL2RG*, *POU2AF1*, *CD27*, *IGJ* (JCHAIN), *CD79A*, and *PIM2*], defining the CIGGs, exhibited a statistically significant association with longevity across the three datasets (Figure 3C and Supplementary Table S5, available at https://doi.org/10.1016/j.esmoop.2024.104109). A moderate correlation coefficient (Cor = 0.63) among the CIGGs suggested that they may track slightly different immune components (Figure 3D). To evaluate if the combination of two CIGGs increases the ability to predict longevity over a single CIGG, we added the log base 2 expression values of two CIGGs. The combination of the two CIGGs into a single score provided a better association with longevity, when compared with single CIGG expression ($P < 0.001$) (Figure 3E and Supplementary Table S6, available at https://doi.org/10.1016/j.esmoop.2024.104109). Overall, 21 CIGG combinations increased the ability to predict longevity (Supplementary Figure S11A and Supplementary Table S7, available at https://doi.org/10.1016/j.esmoop.2024.104109), and the immunoglobulin J chain gene (i.e. *IGJ*, also known as *JCHAIN*) and the interleukin 2 receptor subunit gamma gene (i.e. *IL2RG*) emerged as one of the top gene combinations (Supplementary Figure S11B, available at https://doi.org/10.1016/j.esmoop.2024.104109).

Next, we explored the association of each individual CIGG, or any combinations of two CIGGs, with BCR/TCR entropy. The combination of two CIGGs had similar prediction of BCR and TCR Shannon entropy as the IGG signature (Figure 3F and Supplementary Data S2, available at https://doi.org/10.1016/j.esmoop.2024.104109). Overall, these results underscore the value of the IGG signature and the combination of CIGGs from bulk RNA to capture an effective adaptive B-cell and T-cell immune response.

IGG and longevity in other cancer types

In the pan-cancer TCGA dataset, we previously reported an association of the IGG signature with OS in cervical cancer ($n = 294$), head and neck cancer ($n = 515$), lung adenocarcinoma ($n = 510$), melanoma ($n = 443$), and sarcoma ($n = 253$).[19] Here we focused on the association of the IGG signature with longevity. Like breast cancer, the association of the IGG signature with OS was found to be independent of cancer relapse in a combined patient-level meta-analysis (hazard ratio 0.68, 95% CI 0.61-0.75, $P < 0.001$) and was consistent across the studied cancer types, except for cervical cancer (Supplementary Figure S12, available at https://doi.org/10.1016/j.esmoop.2024.104109).

DISCUSSION

To our knowledge, this study is the first to report an unexpected association between immune gene expression in tumors and longevity in cancer survivors in the absence of recurrence. Notably, high IGG expression, previously linked to reduced distant metastasis and extended survival,[1,2,4-6] emerges as a reliable marker for extended longevity, signifying its role in sustaining antitumor immunity in breast cancer survivors, and potentially other cancer types. B cells actively contribute to antitumor immunity by producing antibodies and fostering the formation of TLSs, as evidenced by our observed higher B-cell-related gene and

analysis in one IGG-high (BC177) and one IGG-low (BC192) breast cancer sample (left); bar plot distribution of all cell types in each sample (right). (D) IHC staining of CD20 (B-cell marker), CD38 (plasma cell marker), and CD3 (T-cell marker) in one IGG-high (BC177) and one IGG-low (BC192) breast cancer sample.
AUC, area under the curve; IGG, 14-gene B-cell/immunoglobulin; IHC, immunohistochemistry; ROC, receiver operating characteristic; ROIs, regions of interest; TLS, tertiary lymphoid structure; UMAP, uniform manifold approximation and projection.

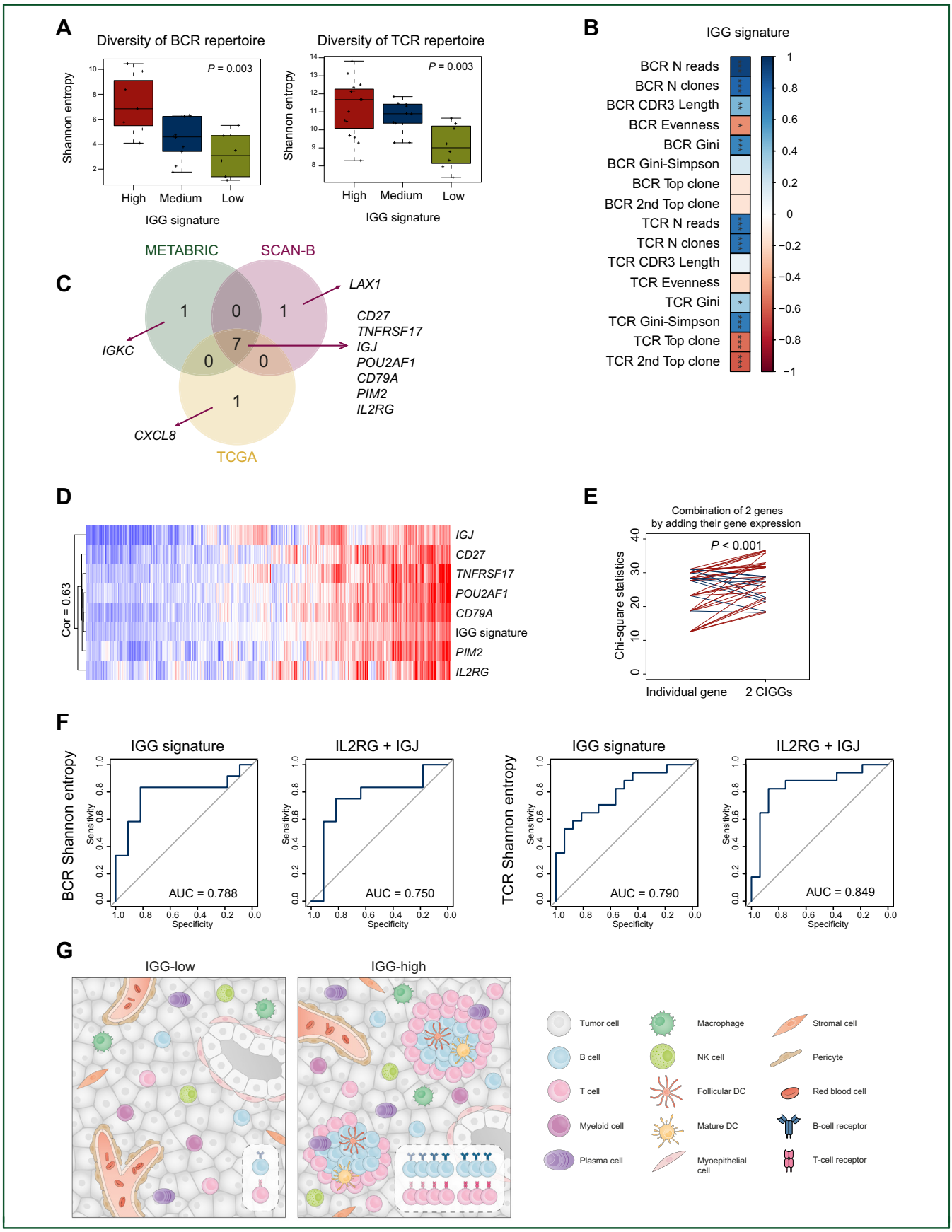

Figure 3. Association between IGG and BCR/TCR clonality and identification of CIGGs related to longevity in patients with breast cancer without documented recurrence. (A) BCR Shannon entropy across IGG-high ($n = 10$), IGG-medium ($n = 7$), and IGG-low ($n = 6$) tumors (left) and TCR Shannon entropy across IGG-high ($n = 18$), IGG-medium ($n = 7$), and IGG-low ($n = 8$) tumors assessed by immunoSEQ assay. (B) Pearson correlation between BCR and TCR clonality measures and IGG levels in the HER2-positive breast cancer CALGB-40601 cohort. (C) Venn diagram of the significant individual genes associated with longevity in univariate Cox regression

protein expression in IGG-high tumors with TLSs. These structures create an environment conducive to B-cell and plasma cell infiltration, promoting a robust and effective adaptive immune response against tumor cells. TLSs play a crucial role in enhancing antitumor immunity and are strongly associated with better clinical outcomes in patients with different solid tumors receiving immune checkpoint inhibitors,[20-22] suggesting their potential as a biomarker. The identification of TLSs in breast cancer has relied on histology, which only captures a limited two-dimensional section. The IGG signature might offer a more robust method for detecting TLSs, providing a comprehensive view of their presence and dynamics.

The observed polyclonality in B cells and T cells within IGG-high tumors emphasizes the richness and diversity of the immune response, potentially contributing to sustained antitumor immunity and reflecting the robust strength of the patient's immune system. Moreover, this heightened immune vigor suggests increased resilience not only against breast cancer but also against other life-threatening conditions, including other cancer types and infections. A recent study revealed that the composition of B cells in centenarians is shifted from naive to memory cells compared with younger individuals,[23] suggesting a change toward a more robust memory B-cell compartment, which is likely due to their lifetime exposure to a variety of antigens and possibly contributing to their longevity by providing faster and more efficient immune responses.

The identification of CIGGs implicated in various lymphocyte functions further advances our understanding of the molecular determinants of extended survival independently of chronological age. The interaction among the seven identified genes—*TNFRSF17* (BCMA), *IL2RG, POU2AF1, CD27, IGJ* (JCHAIN), *CD79A*, and *PIM2*—constitutes a complex network significantly influencing the longevity of patients with breast cancer. *TNFRSF17* (BCMA), crucial for B-cell maturation and survival, indicates its involvement in sustaining antitumor immunity.[24] *IL2RG* contributes to immune response modulation, potentially enhancing the persistence of activated B cells.[25] *POU2AF1,* a transcription factor, likely regulates the expression of genes crucial for B-cell function, further supporting the adaptive immune response.[26] *CD27*, a member of the tumor necrosis factor receptor superfamily, may contribute to prolonged B-cell activation and antitumor immune memory.[27] *IGJ* (JCHAIN) and *CD79A*, essential components of immunoglobulin production, underscore the significance of effective antibody responses in promoting extended survival.[28,29] Lastly, *PIM2*, a proto-oncogene, may play a role in regulating cell survival and immune cell function.[30] The collective action of these genes highlights the complex molecular mechanisms that contribute to sustained antitumor immunity. Notably, integrating two CIGGs significantly improved predictions of longevity and identified robust T-cell-mediated antitumor responses. Further investigation into their individual and synergistic roles is essential to fully understand their impact on patient outcomes and to guide the development of targeted therapeutic strategies.

Our study has limitations. The observed death events, despite lacking documented recurrence, may potentially be related to undetected recurrences, introducing a degree of uncertainty. Although we focused on patients with a low risk of relapse, the challenge of distinguishing between cancer-related and -unrelated mortality persists. In addition, the cause of death was not available in the datasets analyzed and follow-up time differed across datasets. Nonetheless, the similar observation found in three independent breast cancer datasets, and in other cancer types, decreases the likelihood of a false-positive finding. Finally, our study lacks a detailed mechanistic explanation for the observed associations.

In conclusion, our findings not only contribute to the evolving landscape of prognostic markers in breast cancer but also provide nuanced insights into the relationship between immune gene expression in tumors and longevity, paving the way for more personalized and effective therapeutic strategies that account for both recurrence and long-term outcomes.

ACKNOWLEDGEMENTS

LA received funding from Instituto de Salud Carlos III CM20/00091. FS is supported by a Rio Hortega clinical scientist contract from the Instituto de Salud Carlos III CM20/00073. CMP received funding from the Breast Cancer Research Foundation (BCRF-23-127) and the NCI Breast SPORE program (P50-CA058223). AP received funding from the Breast Cancer Research Foundation (BCRF-22-198 and BCRF-23-198), Beca Marta Santamaría, Fundación CRIS contra el Cancer PR_EX_2021-14, Agència de Gestó d'Ajuts Universitaris I de Recerca 2021 SGR 01156, Fundación Fero BECA ONCOXXI21, Asociación Cáncer de Mama Metastásico IV Premios M. Chiara Giorgetti, PI22/01017: funded by Instituto de Salud Carlos III (ISCIII) and co-funded by the European Union, and RESCUER: funded by European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 847912. FBM received funding from Fundación científica AECC Ayudas Investigador AECC 2021 (INVES21943BRAS) and Fundación Contigo Contra el Cancer de Mama de la Mujer (HIMALAIA). Figures 1A and 3G were produced by Antonio García, scientific illustrator from Bio-Graphics.

models in METABRIC ($n = 1133$), SCAN-B ($n = 4470$), and TCGA datasets ($n = 936$). (D) Expression of seven CIGGs and IGG signature in the METABRIC dataset. (E) Amount of variation explained in the prognosis of the expression of one single CIGG or the expression of the combination of two CIGGs (i.e. addition of the log base 2 expression values) as defined by chi-square statistics obtained from likelihood ratio testing. (F) ROC AUCs of IGG signature and the combination of the two CIGGs *IL2RG* + *IGJ* to predict BCR Shannon entropy across (left) and TCR Shannon entropy (right). (G) Schematic representation of the characteristics of IGG-low and IGG-high tumors. Spearman correlation: ***$P < 0.001$, **$P < 0.01$, *$P < 0.05$.

AUC, area under the curve; BCR, B-cell receptor; CIGG, core IGG gene; DC, dendritic cell; IGG, 14-gene B-cell/immunoglobulin; NK, natural killer; ROC, receiver operating characteristic; TCR, T-cell receptor.

FUNDING

This work was supported by Reveal Genomics (no grant number), Instituto de Salud Carlos III (grant PI22/01017), and Fundacion Contigo Contra el Cáncer de Mama de la Mujer (no grant number).

DISCLOSURE

Potential conflicts of interest are as follows: LP is an employee of Reveal Genomics and has patents filed: PCT/EP2021/070788, EP23382703, and EP23383369. GV has received a speaker's fee from MSD, Pfizer, GSK, and Pierre Fabre; has held an advisory role with AstraZeneca; and received consultant fees from Reveal Genomics. MMA is an employee of Reveal Genomics. OMS reports advisory/consulting fees from Reveal Genomics, Roche, and AstraZeneca; lecture fees from Daiichi Sankyo, Novartis, Pfizer, and Eisai; and travel expenses from Gilead and Novartis. PG reports part-time employment from Reveal Genomics. FS declares personal fees for educational events and/or materials from Gilead, Daiichi Sankyo, and Novartis; travel expenses from Gilead, Daiichi Sankyo, and Novartis; and advisory fees from Pfizer. HH is a co-founder and shareholder of Omniscope; advisor to NanoString and Mirxes; and consultant to Moderna and Singularity. MG and NB are employees of Omniscope. CMP reports stockholder and consulting fees from Reveal Genomics. AP reports advisory and consulting fees from AstraZeneca, Roche, Pfizer, Novartis, Daiichi Sankyo, and Peptomyc; lecture fees from AstraZeneca, Roche, Novartis, and Daiichi Sankyo; institutional financial interests from AstraZeneca, Novartis, Roche, and Daiichi Sankyo; stockholder and employee of Reveal Genomics; patents filed PCT/EP2016/080056, PCT/EP2022/086493, PCT/EP2023/060810, EP23382703, and EP23383369. FBM reports part-time employment from Reveal Genomics, and has patents filed: PCT/EP2022/086493, PCT/EP2023/060810, EP23382703, and EP23383369. All other authors have declared no conflicts of interest.

DATA SHARING

All genomic and clinical data used in this paper can be retrieved from public repositories. The METABRIC and TCGA data were retrieved from the cBioPortal for Cancer Genomics (http://cbioportal.org). Data from the last updated version of the SCAN-B dataset were retrieved from Mendeley Data (https://data.mendeley.com/datasets/yzxtxn4nmd). RNA-seq and BCR and TCR repertoire metrics from the CALGB-40601 neoadjuvant HER2-positive study were previously published.[15] The GeoMx raw data generated in this study and the expression of IGG can be found in Supplementary Data S1 and S2, available at https://doi.org/10.1016/j.esmoop.2024.104109.

REFERENCES

1. Fan C, Prat A, Parker JS, et al. Building prognostic models for breast cancer patients using clinical variables and hundreds of gene expression signatures. *BMC Med Genomics*. 2011;4:3.
2. Prat A, Guarneri V, Pascual T, et al. Development and validation of the new HER2DX assay for predicting pathological response and survival outcome in early-stage HER2-positive breast cancer. *eBioMedicine*. 2022;75:103801.
3. Tolaney SM, Tarantino P, Graham N, et al. Adjuvant paclitaxel and trastuzumab for node-negative, HER2-positive breast cancer: final 10-year analysis of the open-label, single-arm, phase 2 APT trial. *Lancet Oncol*. 2023;24(3):273-285.
4. Waks AG, Ogayo ER, Paré L, et al. Assessment of the HER2DX assay in patients with ERBB2-positive breast cancer treated with neoadjuvant paclitaxel, trastuzumab, and pertuzumab. *JAMA Oncol*. 2023;9(6):835-840.
5. Villacampa G, Parker JS, Perou CM, et al. Association of HER2DX with pathological complete response and survival outcomes in HER2-positive breast cancer. *Ann Oncol*. 2023;34(9):783-795.
6. Bueno-Muiño C, Echavarría I, López-Tarruella S, et al. Assessment of a genomic assay in patients with ERBB2-positive breast cancer following neoadjuvant trastuzumab-based chemotherapy with or without pertuzumab. *JAMA Oncol*. 2023;9(6):841-846.
7. Shepherd JH, Ballman K, Polley MC, et al. CALGB 40603 (Alliance): long-term outcomes and genomic correlates of response and survival after neoadjuvant chemotherapy with or without carboplatin and bevacizumab in triple-negative breast cancer. *J Clin Oncol*. 2022;40(12):1323-1334.
8. Fernandez-Martinez A, Pascual T, Singh B, et al. Prognostic and predictive value of immune-related gene expression signatures vs tumor-infiltrating lymphocytes in early-stage ERBB2/HER2-positive breast cancer: a correlative analysis of the CALGB 40601 and PAMELA trials. *JAMA Oncol*. 2023;9(4):490-499.
9. Conte B, Brasó-Maristany F, Hernández AR, et al. A 14-gene B-cell immune signature in early-stage triple-negative breast cancer (TNBC): a pooled analysis of seven studies. *eBioMedicine*. 2024;102:105043.
10. Martín M, Stecklein SR, Gluz O, et al. TNBC-DX genomic test in early-stage triple-negative breast cancer treated with neoadjuvant taxane-based therapy. *Ann Oncol*. 2024. https://doi.org/10.1016/j.annonc.2024.10.012.
11. Curtis C, Shah SP, Chin SF, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*. 2012;486(7403):346-352.
12. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature*. 2012;490(7418):61-70.
13. Staaf J, Häkkinen J, Hegardt C, et al. RNA sequencing-based single sample predictors of molecular subtype and risk of recurrence for clinical assessment of early-stage breast cancer. *NPJ Breast Cancer*. 2022;8(1):94.
14. Brueffer C, Gladchuk S, Winter C, et al. The mutational landscape of the SCAN-B real-world primary breast cancer transcriptome. *EMBO Mol Med*. 2020;12(10):e12118.
15. Rediti M, Fernandez-Martinez A, Venet D, et al. Immunological and clinicopathological features predict HER2-positive breast cancer prognosis in the neoadjuvant NeoALTTO and CALGB 40601 randomized trials. *Nat Commun*. 2023;14(1):7053.
16. Hendry S, Salgado R, Gevaert T, et al. Assessing tumor-infiltrating lymphocytes in solid tumors: a practical review for pathologists and proposal for a standardized method from the International Immunooncology Biomarkers Working Group: part 1: assessing the host immune response, TILs in invasive breast carcinoma and ductal carcinoma in situ, metastatic tumor deposits and areas for further research. *Adv Anat Pathol*. 2017;24(5):235-251.
17. Merritt CR, Ong GT, Church SE, et al. Multiplex digital spatial profiling of proteins and RNA in fixed tissue. *Nat Biotechnol*. 2020;38(5):586-599.
18. Pérez-Rubio C, Planas-Rigol E, Trincado JL, et al. Immune cell profiling of the cerebrospinal fluid enables the characterization of the brain metastasis microenvironment. *Nat Commun*. 2021;12(1):1503.
19. Schettini F, Brunet LP, Marín M, et al. 14-gene immunoglobulin (IGG) and proliferation signatures and association with overall survival across cancer-types. *J Clin Oncol*. 2022;40(suppl 16):2636.

20. Helmink BA, Reddy SM, Gao J, et al. B cells and tertiary lymphoid structures promote immunotherapy response. *Nature*. 2020;577(7791):549-555.
21. Vanhersecke L, Brunet M, Guégan JP, et al. Mature tertiary lymphoid structures predict immune checkpoint inhibitor efficacy in solid tumors independently of PD-L1 expression. *Nat Cancer*. 2021;2(8):794-802.
22. Meylan M, Petitprez F, Becht E, et al. Tertiary lymphoid structures generate and propagate anti-tumor antibody-producing plasma cells in renal cell cancer. *Immunity*. 2022;55(3):527-541.e525.
23. Karagiannis TT, Dowrey TW, Villacorta-Martin C, et al. Multi-modal profiling of peripheral blood cells across the human lifespan reveals distinct immune cell signatures of aging and longevity. *eBioMedicine*. 2023;90:104514.
24. Laabi Y, Gras MP, Brouet JC, Berger R, Larsen CJ, Tsapis A. The BCMA gene, preferentially expressed during B lymphoid maturation, is bidirectionally transcribed. *Nucleic Acids Res*. 1994;22(7):1147-1154.
25. Sharfe N, Shahar M, Roifman CM. An interleukin-2 receptor gamma chain mutation with normal thymus morphology. *J Clin Invest*. 1997;100(12):3036-3043.
26. Ponzoni M, Arrigoni G, Doglioni C. New transcription factors in diagnostic hematopathology. *Adv Anat Pathol*. 2007;14(1):25-35.
27. Buchan SL, Rogel A, Al-Shamkhani A. The immunobiology of CD27 and OX40 and their potential as targets for cancer immunotherapy. *Blood*. 2018;131(1):39-48.
28. Max EE, McBride OW, Morton CC, Robinson MA. Human J chain gene: chromosomal localization and associated restriction fragment length polymorphisms. *Proc Natl Acad Sci U S A*. 1986;83(15):5592-5596.
29. Flaswinkel H, Reth M. Dual role of the tyrosine activation motif of the Ig-alpha protein during signal transduction via the B cell antigen receptor. *EMBO J*. 1994;13(1):83-89.
30. Alvarado Y, Giles FJ, Swords RT. The PIM kinases in hematological cancers. *Expert Rev Hematol*. 2012;5(1):81-96.