Transcriptomic Profiling of Mouse Mammary Tumors Enables Prognostic and Predictive Biomarker Discovery for Human Breast Cancer

Matthew D. Sutcliffe[1,2,3], Kevin R. Mott[1,2], Tulay Yilmaz-Swenson[1,2], Brooke M. Felsheim[1,2,3], Alexander V. Lobanov[1,2,3], Anna R. Michmerhuizen[1,2], Patrick D. Rädler[1,2], Denis O. Okumu[1,4], Xiaping He[1,2], Adam D. Pfefferle[1,2], Stephanie Dance-Barnes[5], Christian Brueffer[6,7,8], Lao H. Saal[6,7,8], Michael P. East[4], Daniel P. Hollern[9], Timothy C. Elston[4], Gary L. Johnson[1,4], and Charles M. Perou[1,2]

ABSTRACT

The development and validation of prognostic and predictive biomarkers in breast cancer are limited by the availability of well-annotated datasets linking tumor molecular features to treatment response and survival outcomes. To address this need, we generated an extensive mouse model dataset comprised of 26 immunocompetent mammary tumor models spanning diverse genetic backgrounds, epithelial–mesenchymal states, the basal–luminal axis, and distinct immune microenvironments. For each model, survival was measured under no treatment, immune checkpoint inhibition (ICI), and carboplatin/paclitaxel chemotherapy, and RNA sequencing was performed on baseline tumors and on 7-day on-treatment samples for both regimens. Baseline murine tumor gene expression features were used to train a machine learning elastic net model that predicted survival outcomes on multiple human breast cancer datasets with performance comparable with that of existing prognostic assays. Models trained for ICI benefit, using either the untreated or 7-day ICI-treated samples, predicted ICI benefit on human ICI-treated datasets, with the 7-day treated tumor model showing better performance. A predictor of carboplatin/paclitaxel response developed from the murine mammary tumor data performed well in mice but did not generalize to human chemotherapy cohorts. Finally, comparison of multiple computational approaches, including XGBoost, random forests, and support vector regression, showed that all methods successfully predicted survival outcomes, with elastic net offering the best performance and interpretability. These results indicate conserved cancer biology between mouse and human tumors for prognosis and ICI response and establish a large preclinical dataset with linked phenotypic and genomic data as a resource for biomarker discovery.

Significance: The development of a genomically and phenotypically diverse murine tumor dataset with linked treatment outcomes establishes a robust translational resource to develop, test, and benchmark clinically relevant prognostic and therapeutic response biomarkers.

Introduction

Breast cancer is a heterogeneous disease with different clinical and molecular subtypes, immune contexts, and responsiveness to treatment. Triple-negative breast cancer (TNBC) remains a particularly aggressive and clinically challenging subtype that shows a high degree of genetic and phenotypic heterogeneity (1–3). For patients with early-stage TNBC, immune checkpoint inhibition (ICI) with pembrolizumab and four chemotherapeutic agents is the standard of care (4, 5). However, variable response rates and survival outcomes to this immunotherapy and chemotherapy regimen are seen and reinforce the need for biologically informed biomarkers that can guide treatment decisions. In addition, human clinical cohorts typically suffer from the logistical difficulty of obtaining serial samples during treatment to investigate the molecular features of response and drug-induced changes. Further, strict trial eligibility criteria, although important for trial design, can introduce selection bias and constrain the generalizability of observed biological and clinical associations.

Preclinical models are an important resource for addressing some of the deficiencies and challenges present in human translational research. Patient-derived xenografts (PDX) have widely been used because they preserve many human tumor features and enable drug screening of human tumors. Although ongoing efforts are being made to reconstitute a more complete immune system in PDXs (6) or in organoid cultures (7, 8), their lack of an intact adaptive immune system that can evolve remains a deficiency that can severely limit the ability to study drugs targeting adaptive immunity. In contrast, immunocompetent genetically engineered mouse models, including syngeneic transplant models (9–13), maintain a fully intact immune system and allow for therapeutic testing and mechanistic interrogation of tumor biology and response where the immune system is thought to be important.

[1]Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina. [2]Department of Genetics, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina. [3]Curriculum in Bioinformatics and Computational Biology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina. [4]Department of Pharmacology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina. [5]Department of Biological Sciences, DePaul University, Chicago, Illinois. [6]Division of Oncology, Department of Clinical Sciences, Lund University, Lund, Sweden. [7]Skåne University Hospital Comprehensive Cancer Center, Lund, Sweden. [8]Center for Interdisciplinary Research on Cancer and Equity in Women, Lund University, Lund, Sweden. [9]Salk Institute for Biological Studies, University of California San Diego (UCSD), La Jolla, California.

Corresponding Author: Charles M. Perou, Department of Genetics, University of North Carolina at Chapel Hill, Marsico Hall 5111, 125 Mason Farm Road, CB #7599, Chapel Hill, NC 27514. E-mail: chuck_perou@med.unc.edu

Cancer Res 2026;86:5337–53

doi: 10.1158/0008-5472.CAN-26-0855

This open access article is distributed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) license.

©2026 The Authors; Published by the American Association for Cancer Research

Prior studies have shown that mouse mammary tumor models can replicate many of the gene expression programs and histologic subtypes of human breast cancer (11, 14, 15). This disease-state conservation suggests that machine learning algorithms trained on high-quality preclinical data could be used to predict human disease-state outcomes. Nevertheless, many existing mouse transcriptomic datasets lack on-treatment measurements, particularly matched survival data, and almost always are limited to a handful of distinct preclinical models. Thus, there is a need for a comprehensive and well-annotated preclinical dataset that can connect murine tumor biology to human clinical response and with sufficient sample size for *de novo* model building.

Here, we present a comprehensive transcriptomic and phenotypic murine dataset comprising 331 bulk RNA sequencing (RNA-seq) samples from 26 genetically diverse and immunocompetent mouse mammary tumor models profiled longitudinally and under two relevant treatment regimens. This dataset enables a direct linkage between tumor states and both off- and on-treatment outcomes. Additionally, we demonstrate that machine learning algorithms trained on the mouse dataset can predict both disease outcomes and response to immunotherapy on independent human cancer cohorts, with performance metrics comparable with those of established clinical assays. This dataset can serve as a powerful resource for additional biomarker discovery and/or validation, which can also enable further hypothesis generation in a testable system.

Materials and Methods

Animal models

All genetically engineered mouse models have been previously published (see **Table 1**) except for the p53RB-426, p53RB-624, and p53RB-626 tumor models. These transplantable tumor lines were derived from three different spontaneous mammary tumors that developed in the mammary glands of K14-Cre (FVB-Tg(KRT14-cre)8Brn/Nci, RRID: MGI:3618476), $Trp53^{fl/fl}$ (FVB.129P2-$Trp53^{tm1Brn}$/Nci, RRID: MGI:3618506), and $Rb1^{fl/fl}$ (FVB;129-$Rb1^{tm2Brn}$/Nci, RRID: MGI:3618505) triple transgenic females in a Friend leukemia virus B (FVB) background. Final triple transgenics were generated by initial breeding of $Trp53^{fl/fl}$ mice with $Rb1^{fl/fl}$ mice to obtain dual transgenic homozygotes, which were then crossed with homozygous K14-Cre carriers to obtain heterozygous triple transgenic offspring that was subsequently bred and propagated into homozygosity for the $Trp53^{fl/fl}$ and $Rb1^{fl/fl}$ transgenes. Mechanistically, in K14-Cre; $Trp53^{fl/fl}$; $Rb1^{fl/fl}$ mice, Cre recombinase expression is driven by a keratin 14 (K14) promoter. In K14-expressing cells, Cre recombinase recognizes the loxP sites of the $Trp53^{fl/fl}$ and $Rb1^{fl/fl}$ alleles, leading to a recombination event that results in Trp53 and Rb1 null alleles. Single transgenics, $Trp53^{fl/fl}$, $Rb1^{fl/fl}$, and K14-Cre, were obtained from the National Cancer Institute (NCI) mouse repository (NCI repository strain numbers 01XC2, 01XC1, and 01XF1, respectively). The K14-TRT model (16) is a cell line derived from a spontaneous K14-Cre $Trp53^{fl/fl}$ $Brca1^{fl/fl}$ mouse as described in Hollern and colleagues (17); this cell line, developed in our lab, is grown in HuMEC Basal Serum-Free Medium (Gibco 12753018), supplemented with HuMEC supplements (Gibco 12755013), 1% penicillin–streptomycin (Gibco 15140122), and 5% fetal bovine serum (Sigma F2442).

The methods to develop the Trp53-deficient BALB/c models (2153, 2208, 2224, 2225, 2250, 2336, 9263, T11, and T12) are described in references (9, 10). The mutagenized lines (KPB25L-APOBEC, KPB25L-UV, T11-APOBEC, and T11-UV) are described in (17). The 4T1 cell line is described in (18). The E0771 cell line is

Table 1. Mouse tumor models used in this study.

Tumor model	Background	Genetics	Tumor source	Subtype	Molecular subtype
2153	BALB/c	$Trp53^{-/-}$	Syngeneic transplant	TNBC	Luminal like
2208	BALB/c	$Trp53^{-/-}$	Syngeneic transplant	TNBC	Luminal like
2224	BALB/c	$Trp53^{-/-}$	Syngeneic transplant	TNBC	Basal like
2225	BALB/c	$Trp53^{-/-}$	Syngeneic transplant	TNBC	Basal like
2250	BALB/c	$Trp53^{-/-}$	Syngeneic transplant	TNBC	Luminal like
2336	BALB/c	$Trp53^{-/-}$	Syngeneic transplant	TNBC	Basal like
9263	BALB/c	$Trp53^{-/-}$	Syngeneic transplant	TNBC	Basal like
4T1	BALB/c	Trp53 mut, Pik3cg mut	Syngeneic transplant	TNBC	Luminal like/claudin low
E0771	C57BL/6	Trp53 mut	Syngeneic transplant	TNBC	Luminal like/claudin low
EGAN-ILC	C57BL/6	Cdh1/Pik3ca mut	Autochthonous	TNBC	Normal like
K14-TRT	FVB	K14-Cre; $Trp53^{f/f}$ $Brca1^{f/f}$	Syngeneic transplant	TNBC	Claudin low
KPB1-A	FVB	K14-Cre; $Trp53^{f/f}$ $Brca1^{f/f}$	Syngeneic transplant	TNBC	Basal like/claudin low
KPB1-B	FVB	K14-Cre; $Trp53^{f/f}$ $Brca1^{f/f}$	Syngeneic transplant	TNBC	Basal like
KPB25L	FVB	K14-Cre; $Trp53^{f/f}$ $Brca1^{f/f}$	Syngeneic transplant	TNBC	Basal like
KPB25L-APOBEC	FVB	K14-Cre; $Trp53^{f/f}$ $Brca1^{f/f}$; Apobec3 overexpression	Syngeneic transplant	TNBC	Basal like
KPB25L-UV	FVB	K14-Cre; $Trp53^{f/f}$ $Brca1^{f/f}$; UV	Syngeneic transplant	TNBC	Basal like
p53RB-426	FVB	K14-Cre; $Trp53^{f/f}$ $RB^{f/f}$	Syngeneic transplant	TNBC	Claudin low
p53RB-624	FVB	K14-Cre; $Trp53^{f/f}$ $RB^{f/f}$	Syngeneic transplant	TNBC	Claudin low
p53RB-626	FVB	K14-Cre; $Trp53^{f/f}$ $RB^{f/f}$	Syngeneic transplant	TNBC	Claudin low
T11	BALB/c	$Trp53^{-/-}$	Syngeneic transplant	TNBC	Claudin low
T11-APOBEC	BALB/c	$Trp53^{-/-}$; Apobec3 overexpression	Syngeneic transplant	TNBC	Claudin low
T11-UV	BALB/c	$Trp53^{-/-}$; UV	Syngeneic transplant	TNBC	Claudin low
T12	BALB/c	$Trp53^{-/-}$	Syngeneic transplant	TNBC	Claudin low
C3-TAG	FVB	SV40 T-antigen	Autochthonous	TNBC	Luminal like
MMTV-PyMT	FVB	PyVT expression	Autochthonous	TNBC	Luminal like
MMTV-Neu	FVB	Her2/NEU expression	Autochthonous	HER2+/ER−	Luminal like

described in (19). The EGAN-ILC model is described in (20). The C3-TAG model is described in (21). The MMTV-PyMT model is described in (22). The MMTV-Neu model is described in (23).

All animal work was conducted in accordance with the Institutional Animal Care and Use Committee (IACUC) guidelines. All experiments were performed using female mice without blinding. For syngeneic transplant models, tumor propagation and transplantation was performed as previously described (17, 24). Recipient mice were matured to 12 weeks of age and then injected with 100,000 to 500,000 cells depending upon the model. Following implantation, mice were monitored for tumor growth 2 to 3 times per week by caliper measurement. Once the tumor diameter reached 5 mm in any dimension, mice were randomized into treatment or control groups.

Chemotherapy

Treated mice received carboplatin (50 mg/kg, Pfizer) and paclitaxel (10 mg/kg, Teva Pharmaceuticals) once per week by intraperitoneal injection. Mice were weighed twice weekly and monitored for adverse effects throughout treatment. Animals exhibiting signs of treatment-related toxicity were euthanized according to the IACUC guidelines and removed from the study.

Immunotherapy

Immunotherapy antibodies were administered intraperitoneally in 100 μL injection volumes twice per week. The following antibodies and doses were used: anti–PD-1 (250 μg, Bio X Cell, BE0146, RRID: AB_10949053), anti–CTLA-4 (125 μg, Bio X Cell, BE0164, RRID: AB_10949609), CD40AG (100 μg, Bio X Cell, BE0016-2, RRID: AB_1107647), IgG2a isotype control (125 μg, Bio X Cell, BE0089, RRID: AB_1107769), and IgG2b isotype control (125 μg, Bio X Cell, BE0086, RRID: AB_1107791). For the chemotherapy experiments, we also included controls that did not receive any IgG2a isotype control treatments. We compared the survival of IgG2a controls with that of no isotype antibody controls and observed no statistically significant survival differences; one set of controls was used for the survival ratio calculations for on-treatment survival times for both the chemotherapy and immunotherapy treatments.

Survival study

Baseline body weight was recorded prior to therapy initiation. Tumor size was measured by caliper instrumentation, and body condition score (BCS) was assessed every 2 to 3 days. Mice were humanely euthanized when tumors reached 20 mm in any dimension or when multiple tumors were noted. In the autochthonous models that typically develop multiple tumors (MMTV-PyMT and C3-TAG), mice were humanely euthanized when the total maximum tumor diameter reached 20 mm. Independent of tumor size, animals were also euthanized if body weight loss reached 20% or more of baseline or if the BCS declined to 2 or below. These criteria were applied uniformly across all treatment groups and any of these events was considered an event for our definition of "endpoint-free survival."

Mouse tumor bulk RNA-seq

mRNA sequencing libraries were generated from up to 1 μg of total RNA using a TruSeq Stranded mRNA kit (Illumina 20020594) following the manufacturer's protocol (Illumina 1000000040498). Paired-end sequencing was performed on either the Illumina HiSeq 2500 or Illumina NovaSeq 6000 system. FASTQ files were merged when necessary and aligned to the mm10 reference genome with Gencode M25 annotations (25) using STAR 2.7.11b (26), and quantification was performed using Salmon 1.10.3 (27). Next, Salmon abundance estimates were upper quartile–normalized (i.e., the 75th percentile of non-zero values was scaled to 1,000) and then $\log_2$-tranformed with the addition of 1 pseudocount. For differential expression analyses, Salmon abundance estimates were rounded to the nearest integer prior to analysis with DESeq2 (28).

Gene expression modules

Eight hundred ninety-four published gene expression signatures representing cell types, biological functions, and cell signaling pathways were previously identified based on published gene lists. Of these, 840 signatures were based on the per-sample median expression of a defined set of genes, 44 "special models" were calculated as the correlation with a predefined centroid, and another 10 "special models" are algorithms that have previously been published (29–37). Three of these modules were the basis for clinical prognostic signatures (Research Prosigna–"ROR_P_Model_JCO.2009_PMID.19204204," Research OncotypeDX–"GHI_RS_Model_NEJM.2004_PMID.15591335," and Research Mammaprint–"Pcorr_NKI70_Good_Correlation_Nature.2002_PMID.11823860"), and an additional three prognostic modules were previously published ("TNBC_Clinically_Relevant_Good.26_BCR.2011_PMID.21978456," "TNBC_Clinically_Relevant_Poor.230_BCR.2011_PMID.21978456," and "TNBC_Clinically_Relevant_Poor.26_BCR.2011_PMID.21978456"); thus, these features were excluded from all elastic net and other machine learning prognostic model building (34, 35, 38, 39). Details of these modules and their associated Entrez gene IDs are listed in Supplementary Table S1.

To compute gene expression module scores for each dataset, gene symbols or Entrez IDs were first mapped to Ensembl IDs. If present, duplicate Ensembl IDs were collapsed by summing counts. Upper quartile–normalized counts were then $\log_2$-transformed with a pseudocount of 1. To calculate median-based modules, genes with less than 70% zero values were removed. Next, rows (genes) were median-centered, and then columns (samples) were column-standardized. For each median-based module, module scores were calculated as the per-sample median expression value of the intersecting gene list. To calculate centroid-based modules, rows (genes) were first standardized. Depending on the module, the parametric Pearson correlation, nonparametric Spearman correlation, or Euclidean distance from each sample to the predefined centroid was computed using the intersecting gene lists. Special algorithm modules were computed using the best approximation based on the original methods (29–37). An R package to calculate all 894 gene expression module scores is available at GitHub: https://github.com/perou-lab/Breast-Cancer-Modules.

Human gene expression sets

The Swedish Cancerome Analysis Network – Breast (SCAN-B) is a multicenter prospective study including clinical and RNA-seq data (40). Clinical data for 6,329 breast cancer cases were obtained from a previous publication (41). FASTQ files from 6,329 breast tumor samples were obtained from collaborators at Lund University (42). The UNC-337 DNA microarray dataset [Gene Expression Omnibus (GEO): GSE18229] is a curated dataset of 337 distinct tumors with clinical follow-up that was previously published (36). For the NKI-295 dataset, which contains 295 distinct tumors, the gene expression and clinical data were obtained from a previous publication (43); the NKI-295 microarray $\log_{10}$ ratios were obtained from the "SeventyGeneData" R package (44). Probes with greater than 30%

missing values were removed. Probes were then mapped to gene symbols. Remaining missing values were imputed using the "impute" R package with k-nearest neighbors of k = 10. Duplicate genes were collapsed using the mean. Log_{10} ratios were converted to log_2 ratios.

For the Bassez and colleagues (45) single-cell RNA-seq (scRNA-seq) dataset, read count and phenotypic data were obtained from the original article. For each patient sample, counts from single cells were aggregated via summation to generate pseudobulk samples prior to normalization. For the Gide and colleagues (46) melanoma dataset, FASTQ files were obtained from the European Nucleotide Archive (47) under accession number PRJEB23709. Paired clinical data were obtained from the original article. The CALGB 40603 clinical trial data were obtained from the NCBI Database of Genotypes and Phenotypes (RRID: SCR_002709; RNA-seq data: phs001863.v1.p1; phenotypic data: phs003801.v1.p1), as previously described (48, 49). For all human RNA-seq data, FASTQ files were aligned to the hg38 reference genome with GENCODE (RRID: SCR_014966) v36 annotations (Gide melanoma) or v47 annotations (SCAN-B) using STAR (RRID: SCR_004463) and quantified with Salmon. Salmon abundance estimates were upper quartile–normalized and log_2-transformed with a pseudocount of 1.

Machine learning models

All machine learning models employed a bootstrap resampling strategy with 100 iterations. At each bootstrap iteration, mouse tumor models were sampled with replacement to create training sets of the same size (i.e., number of tumor models). Out-of-bag (OOB) sets consisted of tumor models not selected in the training set and therefore varied in size at each iteration. In all cases, biological replicates from each tumor model were restricted to either the training set or the OOB set to prevent overestimation of performance.

Elastic net regression

Elastic net models were implemented using the "glmnet" R package. A two-dimensional hyperparameter grid was explored with α values ranging from 0 to 1 in increments of 0.1 (mixing parameter between ridge and least absolute shrinkage and selection operator [LASSO] regularization), and λ values from 10^{-4} to 10^{4} on a logarithmic scale with increments of 0.1 log units. Models were trained to predict either the log-transformed median survival time for untreated models or the log-transformed survival time ratio (treated time/untreated time) for treatment response models. For each (α, λ) combination, models were fit on training data and evaluated on both training and OOB sets. Performance metrics included mean squared error (MSE) and concordance index (c-index) for survival prediction. Hyperparameter values were chosen based on the minimum MSE in the OOB sets.

Random forest

Random forest models were implemented using the "ranger" R package. Hyperparameter grids included the number of trees (500, 1,000, and 1,500), features per split (mtry: 10, 20, 40, and 80), minimum node size (5, 10, 15, and 20), maximum tree depth (5, 10, 15, and 20), and sample fraction (0.5, 0.632, and 0.7). Models were trained to predict log-transformed median survival time. Performance was evaluated using the root mean squared error (RMSE) and c-index on training and OOB sets.

XGBoost

XGBoost models were implemented using the "xgboost" R package. Hyperparameter grids included the learning rate (η: 10^{-2} to $10^{-0.5}$), maximum tree depth (1, 2, and 3), minimum split loss (λ: 0, 0.1, 0.5, and 1), L2 regularization (λ: 10^{-1} to 10^{2}), L1 regularization (α: 10^{-1} to 10^{2}), row sampling fraction (0.5, 0.8, and 1), and column sampling fraction (0.1, 0.2, and 0.5). Models were trained for 100 rounds. Performance was evaluated using the RMSE and c-index on training and OOB sets.

Support vector regression

Support vector regression (SVR) models were implemented using the "e1071" R package with ε-insensitive loss. Hyperparameter grids included kernel type (linear, radial basis function, and polynomial), regularization parameter C (0.01, 0.1, 1, 10, and 100), ε (0.01, 0.05, 0.1, and 0.2), kernel coefficient γ (0.001, 0.01, 0.1, and 1), and polynomial degree (2, 3, and 4) for polynomial kernels. Performance was evaluated using the RMSE and c-index on training and OOB sets.

Immunotherapy-based gene expression analyses

For immunotherapy response analysis, differential expression between anti–PD-1 + anti–CTLA-4–treated and untreated samples was performed using DESeq2 (28). Only samples from tumor models exhibiting a response to anti–PD-1 + anti–CTLA-4 (survival time ratio >1.25, "responders") were considered for differential expression analysis. A design matrix accounting for treatment and tumor model effects was used: design = ~ treatment + tumor model. Differential expression results were filtered for statistical significance (P value < 0.01) and effect direction (log_2 fold change > 0 for upregulated genes).

Linear regression was performed to identify genes associated with response to immunotherapy (as defined by the median survival time ratio) in untreated tumors. For each gene, a linear regression model was fit: log (treated time/untreated time) ~ expression. Genes with positive coefficients and a P value < 0.01 were considered biomarker candidates.

Genes were identified as biomarkers if they met the criteria for both the differential expression analysis and the correlation analysis. Additionally, unnamed mouse genes and retained intron genes were excluded from the final list. Gene set enrichment analysis was performed on biomarker gene lists using the Molecular Signatures Database (MSigDB; RRID: SCR_016863) Hallmark gene sets (50, 51). Enrichment results were filtered by the false discovery rate (FDR) q-value < 0.05.

Results

Mouse tumor model cohort, transcriptomic profiling, and baseline characterization

To establish a comprehensive resource for developing and testing gene expression–based prognostic and predictive computational models, we performed RNA-seq on fresh frozen tumors from 26 genetically distinct mouse mammary cancer models (**Table 1**); of these models, 22 are syngeneic transplants and 4 are autochthonous models. These mouse tumor models are all immunocompetent and include 25 TNBC models and 1 HER2+ model, yet represent a diversity of known genomic subtypes, including luminal, basal, and claudin low (11–13). These models also represent the most common somatic mutations in TNBC, including functional loss of TP53 and/or BRCA1 and/or RB1 (52). For each mouse tumor model, we collected bulk RNA-seq data from biological replicates of untreated tumors (n = 3–9 each) and separately monitored tumor growth over time using multiple untreated tumors ($n \geq 7$ each; **Fig. 1A**). From the tumor growth monitoring study, we defined survival time as the

time in days from an initial tumor diameter of 0.5 cm to an IACUC-mandated endpoint (including tumor >2 cm, multiple tumors, tumor ulceration, and poor overall animal health). The median survival times ranged widely across tumor models, from 9 days for p53Rb-426 to 43 days for KPB25L-APOBEC, representing substantial diversity in baseline tumor progression rates (**Fig. 1B**).

To improve interpretability and reduce dimensionality of the bulk RNA-seq data, we summarized the gene expression features into a curated set of 840 median expression gene signatures, 44 centroid-based models, and 10 special algorithms, which in total represent a diversity of key biological pathways, cellular states, and cell types (see Materials and Methods), similar to what is encompassed within the MSigDB (50). These gene expression "modules" have been extensively published by our laboratory (49, 53–55) and were compiled from independent publications, the MSigDB Hallmark gene sets (56), and literature relevant to cancer biology and the tumor microenvironment. We then used these 888 genomic features (excluding 6 known prognostic features) in a hierarchical clustering analysis of 196 untreated murine tumor RNA-seq samples (**Fig. 1C**). In almost all instances, the mouse model biological replicates clustered together, with mixing between tumor models only seen between a parental line and its mutagenized derivative (e.g., KPB25L and KPB25L-UV).

In breast cancer, a tumor's genomic subtype is a well-studied feature. Based upon previous publications (11, 17, 24), we assigned a molecular "intrinsic" subtype to each of the 26 models (**Table 1**). Subtype mapping using selected gene expression modules revealed alignment of mouse tumor models to human breast cancer molecular subtypes. Basal-like centroid correlations were predominantly highest among the p53-null models, including 2336, 2225, and 2153 (**Fig. 1D**), whereas luminal A centroid correlations were highest in the EGAN-ILC, MMTV-Neu, and MMTV-PyMT luminal models (**Fig. 1E**). In contrast, several models were highly enriched for the claudin-low mesenchymal phenotype, including the previously described T11, T12, and newly created $p53^{fl/fl}$ $Rb^{fl/fl}$ models (**Fig. 1F**). Finally, a gene expression signature representing tertiary lymphoid structure (TLS) scores, thus representing adaptive immunity, also varied across models, with higher TLS genomic scores indicating an immune-rich microenvironment (**Fig. 1G**). Together, these data establish a transcriptomic landscape capturing the basal–luminal axis, claudin-low/mesenchymal states, and the immune microenvironment across the mouse tumor model cohort.

Development of a mouse-derived elastic net prognostic model with cross-species validation

To evaluate the prognostic model building capability of the mouse tumor dataset, we trained an elastic net regression model using 888 gene expression features (after removing 6 known prognostic features; see Materials and Methods) and the $\log_2$ of the tumor model–specific median survival time as the "clinical" outcome for each untreated sample. We applied bootstrap resampling with OOB evaluation across tumor models to identify optimal values for the α (ridge/LASSO mixing term) and λ (regularization weight) hyperparameters (**Fig. 2A**). In each bootstrap simulation, all tumor biological replicates per model were restricted to either the training set or the OOB set to prevent overfitting (see Materials and Methods). Following hyperparameter selection of $\alpha = 0.1$ and $\lambda = 0.126$, we then refit the final model on the entire mouse dataset. The final model identified a set of 106 features of varying weights (40 most highly weighted features are shown in **Fig. 2B**; see Supplementary Table S5 for the full list) representing a diverse set of biological features. Top-weighted features included basal (poor outcome) or luminal (good outcome) gene expression modules, as well as immune signatures (good outcome). We then applied the final model to the median expression profile of the untreated mouse tumors and stratified endpoint-free survival into tertiles, with a c-index of 0.939 indicating strong agreement between predicted and observed survival (**Fig. 2C**).

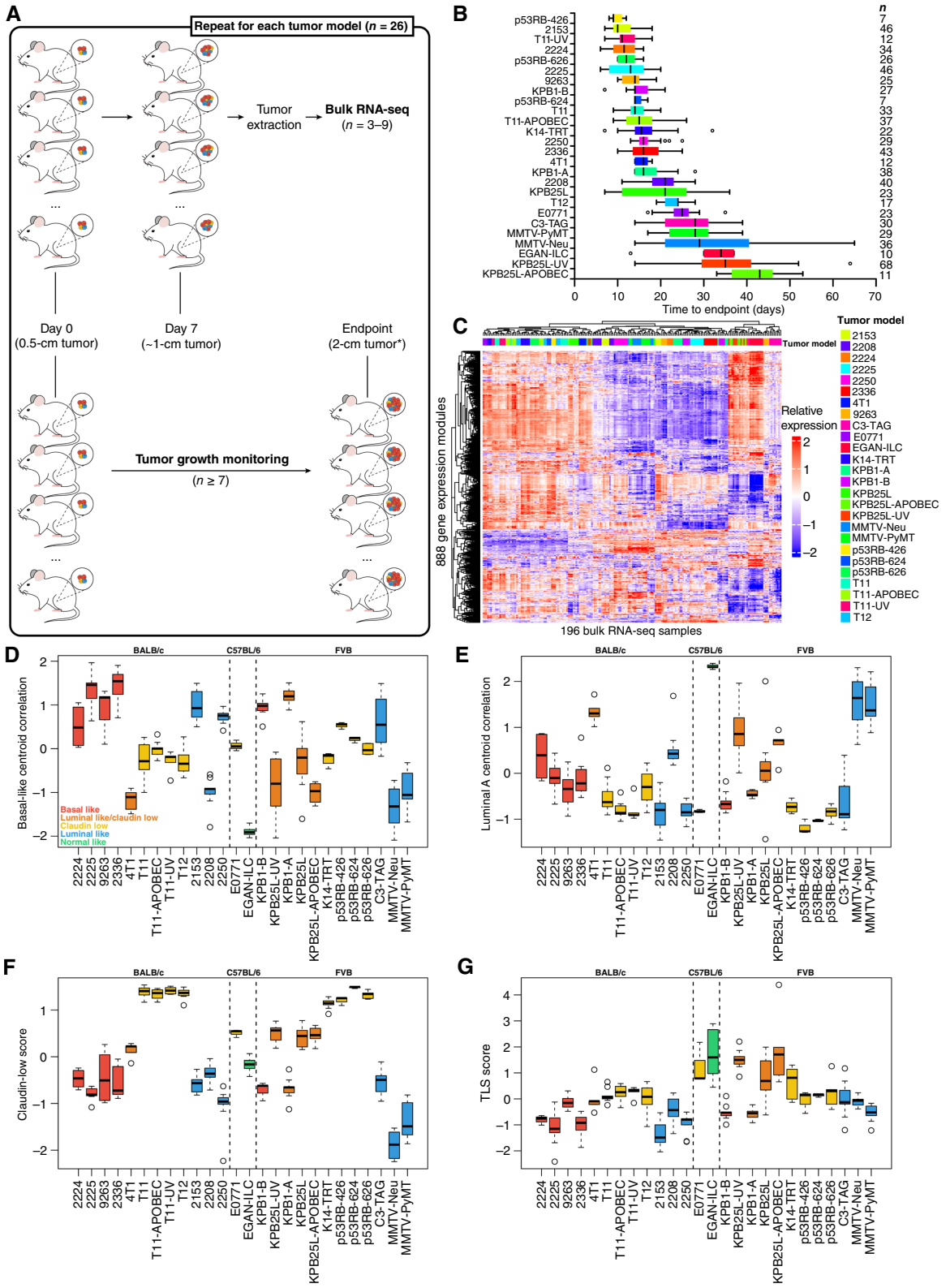
Next, we assessed the translational performance of this mouse-derived elastic net prognostic model on three different human datasets. We calculated the 888 gene expression module scores on three human breast cancer datasets: SCAN-B (RNA-seq based; refs. 40, 41), UNC-337 (DNA microarray; ref. 36), and NKI-295 (DNA microarray; refs. 39, 57). In SCAN-B, the model successfully stratified according to relapse-free survival with an overall performance c-index of 0.657 (**Fig. 2D**). Similar performance was observed in the UNC-337 cohort (c-index = 0.706; **Fig. 2E**) and the NKI-295 cohort (c-index = 0.629; **Fig. 2F**), noting that the NKI cohort is similar to our murine training set in that these patients received no systemic adjuvant therapies (39, 57). For comparison, we benchmarked the mouse-derived risk predictor against established clinical prognostic assays by applying analogous tertile-based stratifications in SCAN-B using research versions of Prosigna (34), Oncotype DX (35), and MammaPrint (**Fig. 2G–I**; ref. 39). Interestingly, the human-derived prognostic models (c-index values ranging from 0.662 to 0.700) showed similar performance to that of the mouse-derived elastic net model, supporting the utility of this mouse dataset in identifying clinically relevant genomic features that might serve as a biomarker. Furthermore, we compared the sample-level predictions from our elastic net model with the prognostic model scores and found that they were moderately correlated ($R^2 = 0.291$–0.485), with each model stratifying human samples along a basal–luminal axis (Supplementary Fig. S1). Lastly, we examined within-clinical subtype performance (i.e., ER+/HER2−, HER2+, and TNBC), and we observed that the murine-derived model performed similarly well in HER2+ and ER+/HER2− subsets, but not in TNBC, again reflecting the performance of the human prognostic models (Supplementary Fig. S2).

To further evaluate the elastic net modeling approach in our mouse tumor dataset, we also performed bootstrapping on the aggregated mouse tumor biological replicates by establishing a single median gene expression profile per tumor model (Supplementary Fig. S3A). With an optimal $\alpha = 0.1$ and $\lambda = 0.158$, we fit an elastic net model with 90 selected features (Supplementary Fig. S3B). We compared these feature weights with that of the individual sample-derived elastic net model described above and found that although some features were only present in one of these model versions, the top-weighted features were consistent between the two and feature weights were correlated (Spearman correlation = 0.489; Supplementary Fig. S3C). We then applied this aggregated sample-derived elastic net model to the mouse training dataset (c-index = 0.987; Supplementary Fig. S3D) and the SCAN-B human test set (c-index = 0.613; Supplementary Fig. S3E) and found that this model performed better on the mouse training dataset, but slightly worse on the human validation dataset, compared with our original model above.

Predictive modeling of response to ICI

Building upon the prognostic modeling framework, we aimed to build a predictive model for survival on ICI. To do this, we tested 23 of the mouse tumor models (excluding EGAN-ILC, K14-TRT, and KPB25L-APOBEC) with combined anti–PD-1 + anti–CTLA-4

therapy (ICI), which is used in human patients with melanoma and currently being tested in patients with TNBC (58, 59). In parallel to the untreated data, we collected at least three biological replicates per model, following 7 days of ICI treatment for bulk RNA-seq processing. In addition, we evaluated endpoint-free survival times on continuous ICI treatment (administered twice weekly). To



quantify the response to ICI therapy, we defined the median survival time ratio in treated versus untreated mice for each model (**Fig. 3A**). Although many of the tumor models did not benefit from ICI treatment, several tumor models were exquisitely sensitive to ICI therapy, including E0771, T11-APOBEC, and T11-UV. Using the same set of 888 gene expression modules described earlier, we fit an elastic net regression model on the 7 day–treated murine tumors to predict the survival time ratio and identified 87 selected features (25 most highly weighted features are shown in **Fig. 3B**; see Supplementary Table S5 for the full list). Among the features with positive weights (i.e., higher expression predicting sensitivity to ICI) were many T-cell signatures, with the top-weighted feature being a T regulatory feature, which is the target of ICI.

We applied this elastic net model to the untreated mouse tumors to evaluate pretreatment biomarker activity, using the median expression profiles of the tumor biological replicates as input to rank-order the sensitivity predictions (**Fig. 3C**). When the underlying survival data were divided into those tumor models with below-median predicted sensitivities and above-median predicted sensitivities (**Fig. 3D** and **E**), we observed strong predictive capabilities. To test this elastic net model on external data, we used two human immunotherapy datasets. The first set is Bassez and colleagues (45), which comprises an scRNA-seq dataset collected before and during anti–PD-1 treatment for which the outcome of interest was T-cell expansion (as measured by T-cell receptor sequencing) following pembrolizumab treatment in 40 patients with breast cancer. We computed gene expression module scores using the pseudobulk counts of each scRNA-seq human sample. We then separately tested the murine ICI elastic net model on the pre- and on-treatment samples. We observed strong predictive capability for the T-cell expansion endpoint in both the pretreatment (**Fig. 3F**) and on-treatment (**Fig. 3G**) samples. Next, we tested the Gide and colleagues (46) dataset, in which patients with melanoma received anti–PD-1 monotherapy or anti–PD-1 + anti–CTLA-4 combination therapy. Similarly, we observed strong predictive capability for tumor response as defined by the RECIST criteria (complete or partial response vs. stable or progressive disease) using both pretreatment ($n = 72$; **Fig. 3H**) and posttreatment ($n = 18$; **Fig. 3I**) bulk RNA-seq samples. Additionally, the elastic net model successfully predicted progression-free survival (PFS) using both pretreatment (**Fig. 3J**) and posttreatment (**Fig. 3K**) samples.

Finally, we also constructed an elastic net model using the untreated mouse tumor bulk RNA-seq samples instead of the 7-day ICI-treated samples. Here, the model selected fewer immune signatures (Supplementary Fig. S4A), yet like the previous analysis, we observed good stratification of mouse tumors based on pretreatment median gene expression (Supplementary Fig. S4B–S4D). This model successfully predicted T-cell expansion on the Bassez breast tumor dataset (Supplementary Fig. S4E and S4F) but was less robust in predicting melanoma ICI response (Supplementary Fig. S4G–S4J).

ICI treatment biological discovery

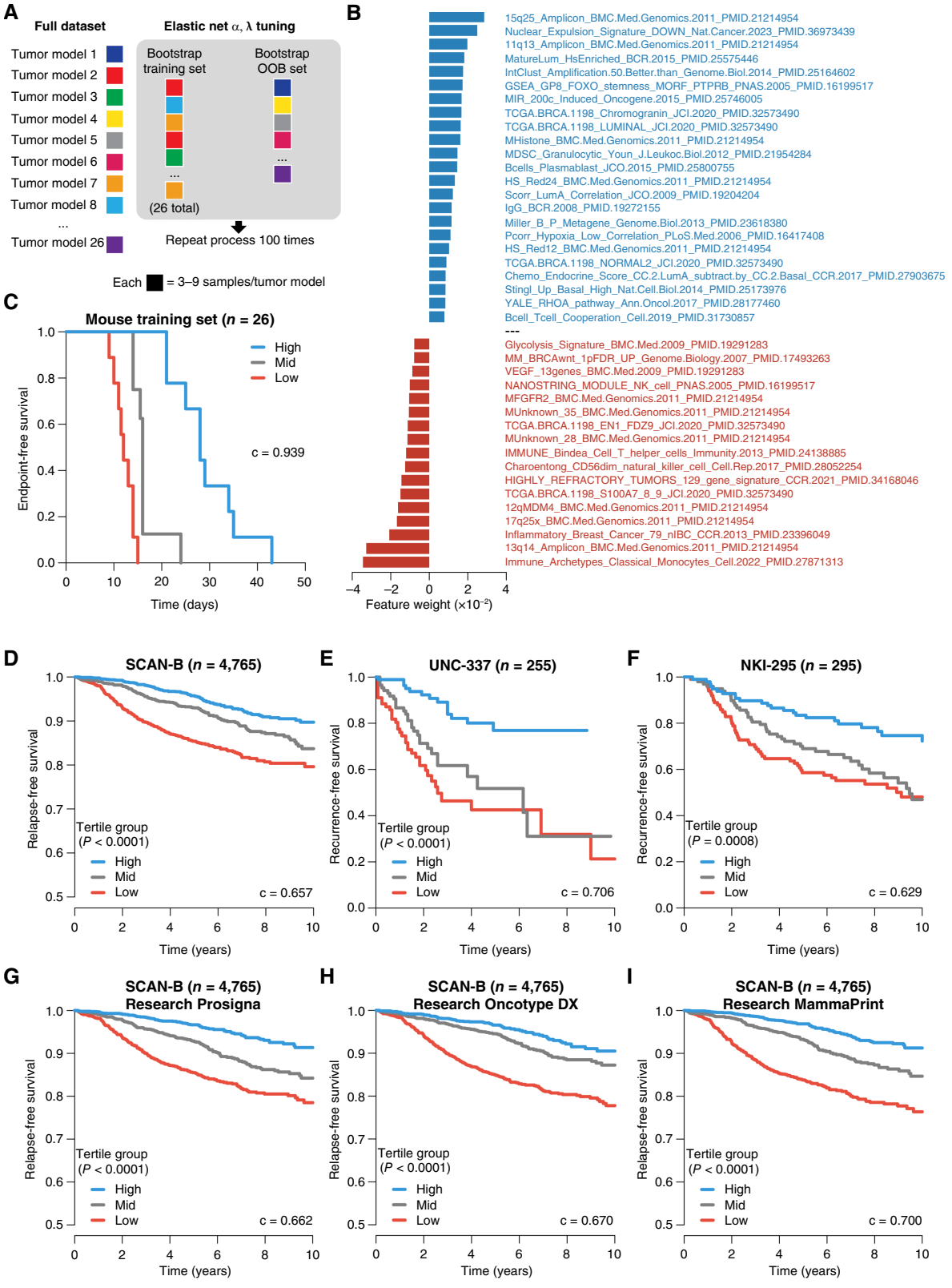
Because our ICI sensitivity elastic net prediction model is based upon previously published gene expression modules as features, we sought to perform exploratory analyses using the original (individual gene-level) bulk RNA-seq data to identify finer grained biological features. We first divided the mouse tumor models into two groups, namely, ICI "responders" and "nonresponders" based on the survival time ratios (**Fig. 3A**, dotted line). Next, we identified genes that were differentially expressed in the responders in the untreated versus 7 day treated tumors. In addition, we fit a linear regression model on the untreated tumors to identify a "sensitivity coefficient" for each gene. When combining the differential expression analysis with the sensitivity coefficient, we identified a set of 246 genes that both increased following ICI treatment in the responders and were individually predictive of sensitivity to ICI (**Fig. 4A**). This gene list included many T-cell (e.g., *Cd3[d, e, g]* and *Cd8*) and B-cell (e.g., *Cd22* and *Cd40*) genes and identified immune-related Hallmark gene sets, including IFNγ response and allograft rejection as the most highly enriched gene sets (**Fig. 4B**).

We next created a new gene expression signature based upon this set of 246 genes, with the per-sample median expression as the "immune signature" score. We calculated immune signature scores in the median expression untreated and ICI-treated murine samples and observed an increase following treatment in many of the tumor models (**Fig. 4C**), as would be expected. Furthermore, we also considered this immune signature score as a biomarker of ICI treatment and split the mouse models into above-median and below-median expressing tumors before treatment. Here, we observe that those with below-median immune signature scores include the tumor models that are resistant to ICI (**Fig. 4D**), whereas those with above-median immune signature scores displayed sensitivity to ICI (**Fig. 4E**). Notably, compared with the elastic net model described previously, which has 87 modules based on 3,020 unique genes, this much smaller list is similarly capable of stratifying mouse tumors by ICI response. To further evaluate this signature's performance, we again utilized Bassez and colleagues (45) and Gide and colleagues (46) as test sets (**Fig. 4F–K**) and observed qualitatively similar results to those of the ICI-treated elastic net model.

Additionally, we compared the gene list that arose from our biomarker analysis with the findings of the Bassez and colleagues (45) and Gide and colleagues test sets. Bassez and colleagues (45) reported that high pretreatment expression levels of PD-1 in T cells and PD-L1 in macrophages or low pretreatment expression levels of TCF7 in T cells correlated with subsequent T-cell expansion following anti–PD-1 treatment. Concordantly, expression of *Pdcd1* (PD1) and *Cd274* (PD-L1) was significantly higher at baseline in mouse tumors that respond to ICI (Supplementary Fig. S5A and S5B), and *Tcf7* expression was significantly higher at baseline in mouse tumors that do not respond (Supplementary Fig. S5C).

Figure 1.

Transcriptomic and phenotypic profiling of 26 mouse mammary tumor models. **A,** Summary of untreated bulk RNA-seq samples and survival endpoint measurements for each tumor model; 3–9 biological replicates per model were collected and processed for bulk RNA-seq. **B,** Boxplots of time-to-endpoint measurements for 26 mouse mammary tumor models. Center line represents median, boxes represent the 25th to 75th percentile, and whiskers represent minimum or maximum values within 1.5 times the interquartile range from the 25th or 75th percentile. **C,** Heatmap of 196 untreated bulk RNA-seq samples across 26 tumor models (columns). Rows represent row-standardized scores for literature-derived gene expression modules (see Materials and Methods). **D,** Correlation of each sample to the basal-like subtype centroid. Mouse tumors are grouped by genetic background and colored by molecular subtype. **E,** Correlation of each sample to the luminal-like subtype centroid. Mouse tumors are grouped by genetic background and colored by molecular subtype. **F,** Inverse distance of each sample to the claudin-low centroid. Mouse tumors are grouped by genetic background and colored by molecular subtype. **G,** Median module score for TLS. Mouse tumors are grouped by genetic background and colored by molecular subtype.


A
Full dataset
Elastic net α, λ tuning
Tumor model 1
Tumor model 2
Tumor model 3
Tumor model 4
Tumor model 5
Tumor model 6
Tumor model 7
Tumor model 8
...
Tumor model 26
Bootstrap training set
Bootstrap OOB set
(26 total)
Repeat process 100 times
Each = 3–9 samples/tumor model
B
15q25_Amplicon_BMC.Med.Genomics.2011_PMID.21214954
Nuclear_Expulsion_Signature_DOWN_Nat.Cancer.2023_PMID.36973439
11q13_Amplicon_BMC.Med.Genomics.2011_PMID.21214954
MatureLum_HsEnriched_BCR.2015_PMID.25575446
IntClust_Amplification.50.Better.than_Genome.Biol.2014_PMID.25164602
GSEA_GP8_FOXO_stemness_MORF_PTPRB_PNAS.2005_PMID.16199517
MIR_200c_Induced_Oncogene.2015_PMID.25746005
TCGA.BRCA.1198_Chromogranin_JCI.2020_PMID.32573490
TCGA.BRCA.1198_LUMINAL_JCI.2020_PMID.32573490
MHistone_BMC.Med.Genomics.2011_PMID.21214954
MDSC_Granulocytic_Youn_J.Leukoc.Biol.2012_PMID.21954284
Bcells_Plasmablast_JCO.2015_PMID.25800755
HS_Red24_BMC.Med.Genomics.2011_PMID.21214954
Scorr_LumA_Correlation_JCO.2009_PMID.19204204
IgG_BCR.2008_PMID.19272155
Miller_B_P_Metagene_Genome.Biol.2013_PMID.23618380
Pcorr_Hypoxia_Low_Correlation_PLoS.Med.2006_PMID.16417408
HS_Red12_BMC.Med.Genomics.2011_PMID.21214954
TCGA.BRCA.1198_NORMAL2_JCI.2020_PMID.32573490
Chemo_Endocrine_Score_CC.2.LumA_subtract.by_CC.2.Basal_CCR.2017_PMID.27903675
Stingl_Up_Basal_High_Nat.Cell.Biol.2014_PMID.25173976
YALE_RHOA_pathway_Ann.Oncol.2017_PMID.28177460
Bcell_Tcell_Cooperation_Cell.2019_PMID.31730857

Glycolysis_Signature_BMC.Med.2009_PMID.19291283
MM_BRCAwnt_1pFDR_UP_Genome.Biology.2007_PMID.17493263
VEGF_13genes_BMC.Med.2009_PMID.19291283
NANOSTRING_MODULE_NK_cell_PNAS.2005_PMID.16199517
MFGFR2_BMC.Med.Genomics.2011_PMID.21214954
MUnknown_35_BMC.Med.Genomics.2011_PMID.21214954
TCGA.BRCA.1198_EN1_FDZ9_JCI.2020_PMID.32573490
MUnknown_28_BMC.Med.Genomics.2011_PMID.21214954
IMMUNE_Bindea_Cell_T_helper_cells_Immunity.2013_PMID.24138885
Charoentong_CD56dim_natural_killer_cell_Cell.Rep.2017_PMID.28052254
HIGHLY_REFRACTORY_TUMORS_129_gene_signature_CCR.2021_PMID.34168046
TCGA.BRCA.1198_S100A7_8_9_JCI.2020_PMID.32573490
12qMDM4_BMC.Med.Genomics.2011_PMID.21214954
17q25x_BMC.Med.Genomics.2011_PMID.21214954
Inflammatory_Breast_Cancer_79_nIBC_CCR.2013_PMID.23396049
13q14_Amplicon_BMC.Med.Genomics.2011_PMID.21214954
Immune_Archetypes_Classical_Monocytes_Cell.2022_PMID.27871313
Feature weight (×10⁻²)
C
Mouse training set (n = 26)
High
Mid
Low
c = 0.939
Endpoint-free survival
Time (days)
D
SCAN-B (n = 4,765)
Relapse-free survival
Tertile group
(P < 0.0001)
c = 0.657
Time (years)
E
UNC-337 (n = 255)
Recurrence-free survival
Tertile group
(P < 0.0001)
c = 0.706
F
NKI-295 (n = 295)
Tertile group
(P = 0.0008)
c = 0.629
G
SCAN-B (n = 4,765)
Research Prosigna
c = 0.662
H
SCAN-B (n = 4,765)
Research Oncotype DX
c = 0.670
I
SCAN-B (n = 4,765)
Research MammaPrint
c = 0.700

Similarly, Gide and colleagues reported that a subset of EOMS+/CD69$^+$/CD45$^+$ T cells was higher at baseline in patients with longer PFS on immunotherapy. Concordantly, we found that higher baseline expression of *Eomes*, *Cd69*, and *Ptprc* (CD45) was associated with ICI response (Supplementary Fig. S5D–S5F). These results indicate that murine-derived biomarkers of ICI response were consistent with biological or clinical outcomes in human immunotherapy trials.

Given the translatability of these murine computational models onto human tumor datasets, we next searched the list of 246 genes for possible drug targets and observed that both the CD40 receptor and CD40 ligand were among the 246 genes in the immune signature. The CD40 receptor is a central regulator of immune response, primarily expressed on antigen-presenting cells, including dendritic cells and B cells. T cells can produce CD40 ligand and thus engage and activate CD40$^+$ cells, leading to activated dendritic cells and B cells (60–62). The CD40 ligand, often referred to as CD40AG, is a monoclonal antibody therapy that is being tested on multiple human tumors including pancreatic tumors and melanoma (63–65). Therefore, we tested a murine compatible version of a CD40AG antibody on ICI-sensitive and ICI-resistant models, alone or in combination with ICI. We first tested an ICI-sensitive model (E0771) and observed that CD40AG alone, and in combination with anti-PD1, extended survival and exhibited similar benefit to the anti–PD-1 + anti–CTLA-4 combination (**Fig. 5A**). Next, we tested the anti–PD-1 + CD40AG combination on five p53-null tumor models that were resistant to the anti–PD-1 + anti–CTLA-4 combination. We found that two tumor models significantly benefited from anti–PD-1 + CD40AG therapy [2208 (**Fig. 5B**) and 2224 (**Fig. 5C**)] and three did not (2153, 2336, and 2225; **Fig. 5D–F**). We were curious about what distinguished CD40AG-sensitive tumors from insensitive tumors, specifically those treated with combination anti-PD1 + CD40AG. To investigate, we performed a differential expression analysis of baseline gene expression using DESeq2 and identified 2,557 genes elevated at baseline in CD40AG-sensitive tumors and 2,139 genes elevated at baseline in CD40AG-insensitive tumors (**Fig. 5G**). We performed gene set enrichment on the CD40AG sensitivity– and CD40AG resistance–associated genes and found that several biological features distinguished these groups. Tumors sensitive to CD40AG had higher baseline levels of immune Hallmark signatures (allograft rejection, IFNα/γ response, and inflammatory response) and insensitive tumors had higher baseline levels of E2F target genes (**Fig. 5H**).

Predictive modeling of response to carboplatin/paclitaxel chemotherapy

Because most patients with early-stage TNBC receive multiagent neoadjuvant chemotherapy as part of standard-of-care treatments, we were interested to see if the mouse tumor panel could be used to derive a chemotherapy benefit predictor. As seen with untreated and ICI-treated mouse models, we observed varying sensitivities to carboplatin/paclitaxel treatment across our 24 models tested (Supplementary Fig. S6A), noting the three of the four most sensitive models were all derived from our Trp53$^{fl/fl}$/Brca1$^{fl/fl}$ model (13). Next, we trained an elastic net model to predict carboplatin/paclitaxel treatment survival benefit using 7 day–treated mouse tumor bulk RNA-seq samples and treated median survival time ratios. Using the same elastic net modeling approach, the resulting model identified several PARP inhibitor sensitivity modules among the most highly weighted features, the high expression of which predicted response (Supplementary Fig. S6B). Notably, when the elastic net model is applied to the median expression profiles of the untreated mouse tumors, all BRCA-deficient tumor models (KPB1-A, KPB1-B, KPB25L, and KPB25L-UV) were predicted to have above-average sensitivity to carboplatin/paclitaxel (Supplementary Fig. S6C). Additionally, when the underlying survival data were split into above- and below-median sensitivity predictions, the elastic net model successfully predicts carboplatin/paclitaxel response compared with untreated samples (Supplementary Fig. S6D and S6E). However, when applied to human chemotherapy-treated datasets, the elastic net model failed to robustly predict response or survival. From the SCAN-B cohort, we subset to the chemotherapy-treated patients and evaluated predicted chemotherapy sensitivities; the elastic net had a modest prognostic effect in the SCAN-B chemotherapy subset (c-index = 0.58; $P = 0.0009$; Supplementary Fig. S6F); however, the high-score patients had worse outcomes. Additionally, when the elastic net model is applied to samples from CALGB 40603 (48), a TNBC phase II trial testing the addition of carboplatin or bevacizumab to anthracycline- and taxane-based chemotherapy, the model failed to predict event-free survival or pathologic complete response (Supplementary Fig. S6G and S6H).

Comparative analysis of machine learning approaches

To assess whether alternative machine learning algorithms might improve computational model performance in our untreated mouse prognostic modeling, we trained XGBoost (66), random forest, and SVR models using the murine dataset and tested on a human dataset. Similar to the elastic net regression training, we used bootstrapping to optimize hyperparameter selection (see Materials and Methods for details). With each of these approaches, we fit a model using the same mouse untreated tumor expression data and untreated endpoint-free survival times and evaluated these additional models on the full SCAN-B cohort (**Table 2**). All methods performed similarly, with elastic net achieving the best c-index on the entire SCAN-B cohort. When analyzing clinical subtype-specific performance, similar patterns were observed with

Figure 2.

Elastic net prognostic modeling of untreated mouse tumors and application to human datasets. **A,** Bootstrapping approach to select optimal combination of α and λ (see Materials and Methods). **B,** Top 40 features from final fitted elastic net model. Higher expression of features with positive weights (blue) predicts longer survival times, and higher expression of features with negative weights (red) predicts shorter survival times. **C,** Mouse training set model performance on aggregated mouse tumor samples ($n = 26$), with predicted scores split into equal tertiles. **D,** Mouse-derived elastic net model applied to the human SCAN-B dataset. Predictions were split into tertiles prior to plotting. Tertile *P* values using the log-rank test. **E,** Mouse-derived elastic net model applied to the human UNC-337 microarray dataset. Predictions were split into tertiles prior to plotting. Tertile *P* values using the log-rank test. **F,** Mouse-derived elastic net model applied to the human NKI-295 microarray dataset. Predictions were split into tertiles prior to plotting. Tertile *P* values using the log-rank test. **G,** Research version of the Prosigna PAM50 ROR-P prognostic model applied to the human SCAN-B dataset. Predictions were split into tertiles prior to plotting. Tertile *P* values using the log-rank test. **H,** Research version of the Oncotype DX prognostic model applied to the human SCAN-B dataset. Predictions were split into tertiles prior to plotting. Tertile *P* values using the log-rank test. **I,** Research version of the MammaPrint prognostic model applied to the human SCAN-B dataset. Predictions were split into tertiles prior to plotting. Tertile *P* values using the log-rank test.

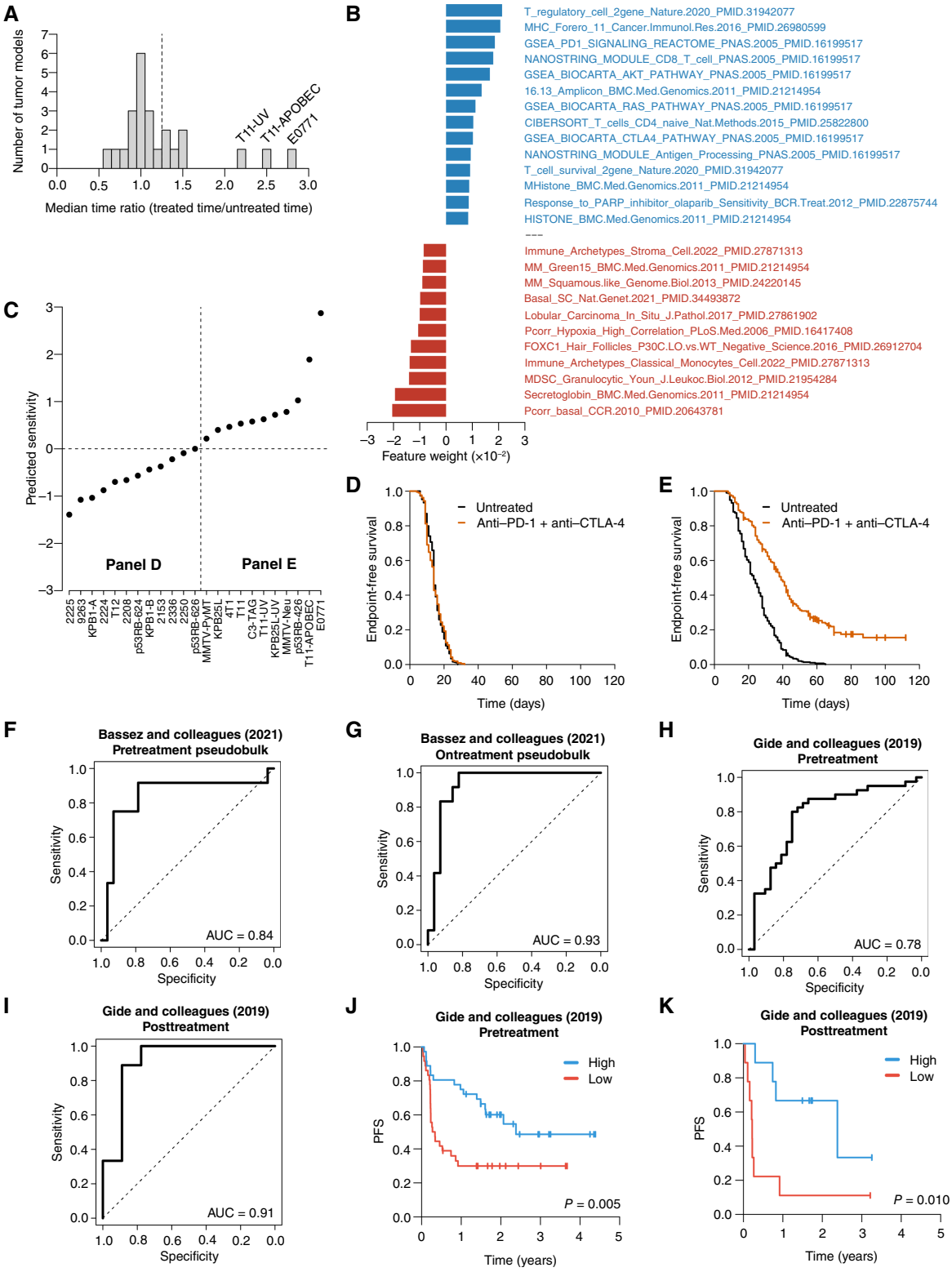


Figure 3.

Elastic net predictive modeling of ICI response in untreated mouse tumors and validation in human datasets. **A,** Histogram of the mouse tumor model response to ICI. A time ratio of 1 indicates no difference in median survival time in ICI-treated vs. untreated tumors. **B,** Top 25 weighted features in final elastic net ICI response predictor. Positive weights (blue) predict sensitivity to ICI, and negative weights (red) predict tumor resistance. (*Continued on the following page.*)

consistent performance for all methods on the ER+/HER2− and HER2+ subsets, and all showed reduced performance in TNBC. Given these similar performance metrics, we favor the elastic net modeling method due to its improved interpretability of selected features.

Discussion

In this study, we created and used a phenotypically and genetically diverse panel of 26 mouse models of mammary cancer to develop prognostic and predictive transcriptomic profiling–based machine learning models of survival and therapeutic response. We found that this large and diverse mouse tumor model dataset captured much of the molecular heterogeneity of human breast cancers, particularly the spectrum ranging from differentiated luminal phenotypes to the more aggressive and less differentiated basal-like and claudin-low phenotypes. The conservation of the basal–luminal axis in murine models is known (11, 67) and well represented in this cohort, which allows for the interrogation of these biological programs and new hypotheses in a controlled, immunocompetent setting that may better allow for the translation of some findings from mouse models into human clinical breast cancer research. We believe that a key advance here was to use the mouse in a manner more similar to what is done in human studies, namely, to base findings upon a phenotypically diverse population of "individuals," instead of relying on one or two individuals and then extrapolating to a larger population. In addition, this murine dataset provides a unique resource often only seen with cell lines, in which multiple treatments are performed on the same model and on tens of models, thus providing multiple individually determined treatment responses on the same model.

To efficiently characterize the high-dimensional bulk tumor RNA-seq data generated here, we employed a curated set of 894 gene expression modules; we now provide an R package that can run these 840 median gene signatures, 44 centroid-based modules, and 10 special algorithms (https://github.com/perou-lab/Breast-Cancer-Modules). This approach addresses a common challenge in cross-species genomics: the noise associated with individual genes and the precise linking of gene-by-gene orthologs. By summarizing the median expression of functionally related and/or co-expressed genes into modules, and by using previously rigorously identified special algorithms (i.e., Oncotype DX, MammaPrint, and Prosigna), this analysis emphasizes biological processes and buffers the noise coming from the expression of individual genes. We have previously used these gene expression modules to identify prognostic and predictive features for patients with early-stage HER2+ breast cancer and TNBC (48, 49, 54, 55, 68), and here we show that they can be used in mouse datasets to predict murine model survival times and murine model response to immunotherapy. Unlike dataset-specific principal component analysis or metagene analysis, in which clusters of genes are used as input features in machine learning analyses, the gene sets presented here are independent of the original datasets and thus much more readily used across a variety of technologies, as indicated by our RNA-seq models working well on decades-old DNA microarray data.

Using these gene expression features, we evaluated multiple machine learning approaches and determined that elastic net regression had comparable performance with more complex algorithms, including XGBoost (66), random forest (69), and SVR (70). This finding is consistent with recent analyses evaluating the risk of recurrence models in hormone receptor–positive breast cancer samples (71). This is also consistent with similar studies using DNA methylation–based cancer features, in which elastic net was compared with decision-tree and support vector–based methods and found to be superior (72). Furthermore, the elastic net regression approach selects a subset of all input features and returns easily interpretable feature weights that explain the relative contribution and the directionality of each selected feature. Although importance metrics exist for tree-based algorithms, the relationships between features and predicted outcomes are nonlinear and not necessarily monotonic, requiring cut points that vary greatly between decision trees. Additionally, we compared methods within our elastic net framework for aggregating the gene expression profiles of biological replicates versus treating them as independent observations with the same median outcome. Our analysis showed that the latter was associated with a lower c-index on the training dataset but a higher c-index on the human SCAN-B test dataset, which would be consistent with greater generalizability of the complete dataset approach (i.e., treating individual samples separately vs. averaging together).

Using the elastic net approach, we first developed a survival model entirely trained on our murine tumor dataset of 26 models represented by 196 RNA-seq samples, and it achieved performance comparable with established clinical tools (research versions) tested in the laboratory setting, including Prosigna, Oncotype DX, and MammaPrint, when evaluated on the same publicly available patient datasets (SCAN-B, NKI-295, and UNC-337). Although another prognostic predictor in breast cancer is not clinically needed, our results further inform tumor biology and suggest that existing prognostic models may be improved by including new immune system features. Importantly, these findings support the idea that the underlying tumor biology in our syngeneic transplant and autochthonous mouse models is conserved across species and that we have captured much of the molecular diversity of human breast cancer in our 26 mouse model dataset. Interestingly, we did not observe a strong contribution of proliferation-related signatures in the top-weighted features, likely because most of the mouse tumors grow rapidly, and thus proliferation may not be a distinguishing prognostic feature when fitting the model. However, we found that features related to molecular subtype, particularly luminal and basal

(*Continued*.) **C,** Elastic net model predictions for the median expression profiles of untreated mouse tumors. **D,** Pooled ICI-treated and untreated tumor survival data for mouse tumor models with below-median predicted sensitivity ($n = 378$ untreated mice; $n = 213$ ICI-treated mice). **E,** Pooled ICI-treated and untreated tumor survival data for mouse tumor models with above-median predicted sensitivity ($n = 310$ untreated mice; $n = 217$ ICI-treated mice). **F,** Receiver operating characteristic (ROC) curve of elastic net model predictions for Bassez pseudobulk samples ($n = 40$) used to predict T-cell expansion in pretreatment samples. **G,** ROC curve of elastic net model predictions for Bassez pseudobulk samples ($n = 40$) used to predict T-cell expansion in on-treatment samples. **H,** ROC curve of elastic net model predictions for Gide samples used to predict tumor response (partial/complete response vs. stable/progressive disease) using the pretreatment samples ($n = 72$). **I,** ROC curve of elastic net model predictions for Gide samples used to predict tumor response (partial/complete response vs. stable/progressive disease) using the posttreatment samples ($n = 18$). **J,** Kaplan–Meier plots for elastic net model predictions for Gide pretreatment samples ($n = 72$), with predictions split into above- and below-median groups. **K,** Kaplan–Meier plots for elastic net model predictions for Gide posttreatment samples ($n = 18$), with predictions split into above- and below-median groups.

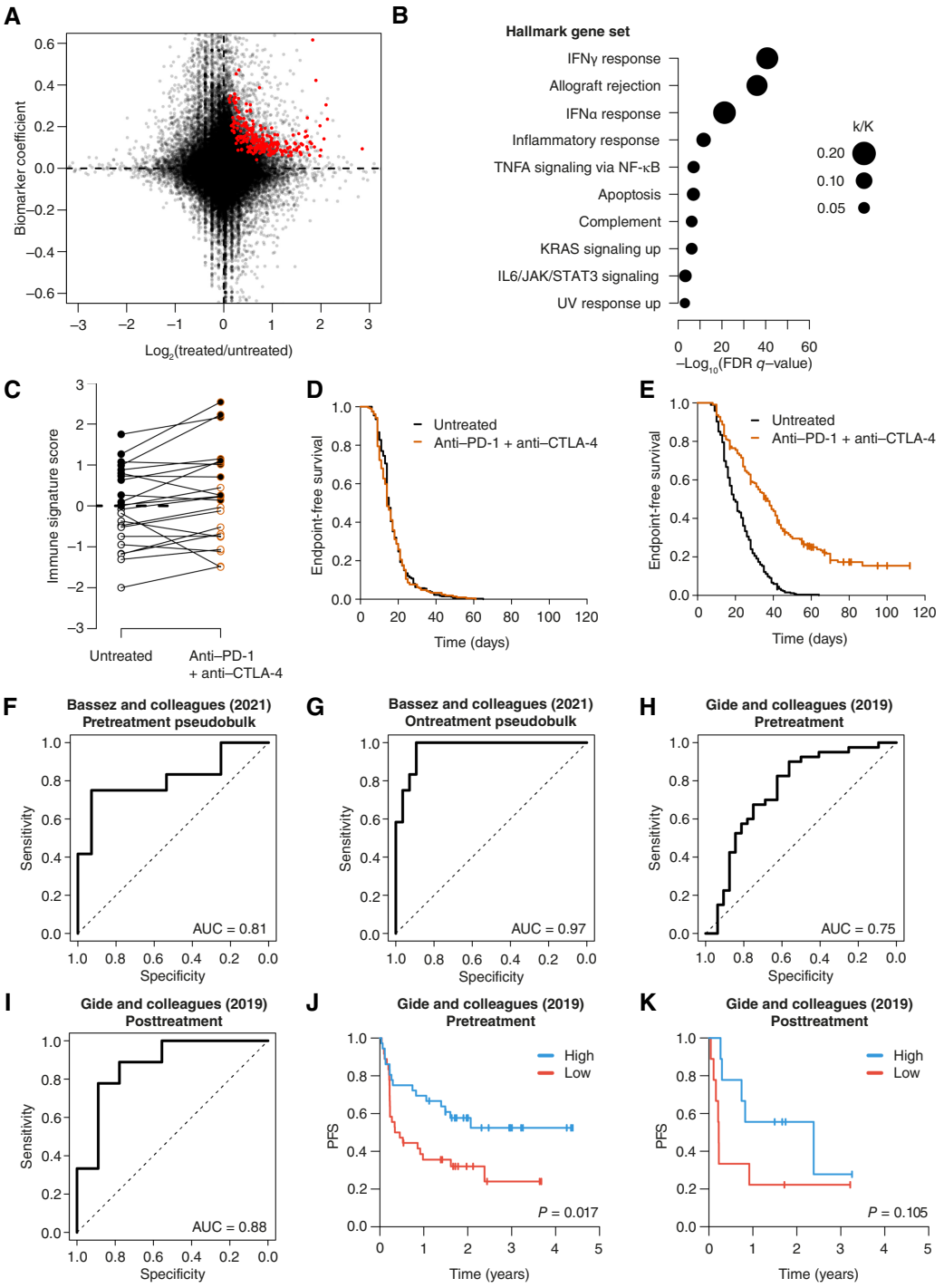


Figure 4.

Agnostic ICI biomarker identification in ICI-treated mouse tumors. **A,** Scatter plot of gene-based differential expression and biomarker analyses. *X*-axis represents fold change in ICI-treated vs. untreated mouse tumors categorized as responders. *Y*-axis represents the slope of the best fit line when the median survival time ratio is predicted using untreated expression levels. Red, two hundred forty-six genes that were significantly increased in ICI-treated tumors and significantly correlated with survival outcomes. **B,** Gene set enrichment analysis of 246 significant genes from **A**. Top 10 Hallmark pathways. The size of circles represents the fraction of overlapping genes (k) over the size of the Hallmark gene set (K). **C,** Immune signature score calculated for median expression profiles of untreated and ICI-treated tumor models. Black, tumor models with above-median immune signature scores; white, tumor models with below-median immune signature scores. **D,** Pooled ICI-treated and untreated tumor survival data for mouse tumor models with below-median immune signature scores ($n = 400$ untreated mice; $n = 209$ ICI-treated mice). **E,** Pooled ICI-treated and untreated tumor survival data for mouse tumor models with above-median immune signature scores ($n = 288$ untreated mice; $n = 221$ ICI-treated mice). **F,** ROC curve for Bassez pseudobulk (*Continued on the following page.*)

features, were well represented in the top-weighted features. This is further highlighted by the correlation between the elastic net model predictions and the three prognostic models used for comparison, which all pick up on a well-understood spectrum of basal (more aggressive) to luminal (less aggressive) biological features. Despite this, the elastic net model failed to predict survival in human TNBCs, performing on par with the clinical tools.

One of the more striking findings in this study was the success of our mouse model–derived ICI predictor, which worked well in two different human datasets spanning two different tumor types and "rediscovered" the ICI drug target (i.e., T regulatory cells) as the top-weighted elastic net model feature. It is interesting to note, however, that although three mouse tumor models were highly sensitive to ICI (T11-UV, T11-APOBEC, and E0771), only two tumor models (T11-APOBEC and E0771) had very high predicted sensitivities. This may indicate that there are additional biological features that confer sensitivity to ICI that are not well represented in our modules or that these features do not generalize to all mouse tumors. Our finding that the on-treatment–derived elastic net model performed better than the pretreatment-derived elastic net model, even when applied to pretreatment tumors, was also interesting and emphasizes the utility of short-term treatment sample data. However, it is important to note that this finding may be specific to the 7-day timepoint and the ICI treatment context examined here. In addition, when we performed more conventional exploratory analyses to identify treatment-induced features associated with response to ICI, we identified a rich set of adaptive immunity genes that performed well as a biomarker of response on human ICI-treated datasets, even when using only 246 genes. Given its simplicity as a single feature median-based signature, further testing through retrospective analysis of existing ICI randomized trials in TNBC might provide support for clinical implementation. In addition, this gene set includes many genes of mechanistic importance, including *IFNγ* and *IL21*. Using one of these 26 models (KPB25L-UV), we previously showed that administering neutralizing antibodies against IL21 at the same time as anti–PD-1 + anti–CTLA-4 antibodies inhibited the responsiveness of this model to ICI (17). Another mechanistically and possibly therapeutically important finding from this analysis was the identification of the CD40–CD40 ligand axis, which is a known immune cell cross-talk network often resulting in dendritic cell and/or B-cell activation (73). Therefore, we performed functional experiments using CD40AG antibodies in multiple different mouse models and saw single-agent activity in one of our ICI responsive models (E0771) and activity with anti–PD-1 in two of five ICI-resistant models tested (**Fig. 5**). Our biomarker analysis highlighted adaptive immune features, including many T-cell marker genes, as being significantly elevated at baseline in anti–PD-1 + CD40AG–responsive models, which is similar to our findings in the anti–PD-1 + anti–CTLA-4 biomarker analysis; more work will be needed to mechanistically evaluate the differential response to these two combination treatments. These promising preclinical model results suggest testing CD40AG antibodies in patients with TNBC either alone or in combination with anti–PD-1 antibodies. Additionally, we compared the results of our biomarker analysis with the original analyses done in the ICI test sets [Bassez and colleagues (45) and Gide and colleagues (46)]. Although we were unable to evaluate specific cell-type compositions in our bulk RNA-seq analyses, untreated gene expression levels in ICI responders and nonresponders were largely consistent with T-cell expansion biomarkers in breast cancer [Bassez and colleagues (45)] and ICI on-treatment outcomes in melanoma [Gide and colleagues (46)].

Given the success of our murine-derived prognostic and ICI-specific predictor, it was surprising to determine that a similarly trained carboplatin/paclitaxel chemotherapy predictor largely failed when applied to human data (although it performed well on the mouse training dataset). One possible explanation for this result is that there are important biological differences in response to chemotherapy between mice and humans, which are not captured in our untreated cohort or 7 day–treated chemotherapy cohort. However, we suspect that the more likely explanation is also responsible for the lack of an existing clinically used biomarker in human TNBCs; thus, it is quite challenging to identify such a biomarker using any model system. A second explanation for this failure, and for the lack of an established human TNBC chemotherapy biomarker, is that patients with TNBC typically receive three or more chemotherapeutics in the early-stage setting, each with a distinct mechanisms of action, which we recapitulated here on our mouse models by giving carboplatin (DNA damage) and paclitaxel (microtubule inhibitor). In human TNBC, it is even more complex because of the inclusion of additional chemotherapeutics (i.e., cyclophosphamide and doxorubicin; ref. 74). Given this complexity of drugs, and its failure to generalize to human datasets, we would not advocate clinical testing of our murine-derived chemotherapy predictor. Finally, we did not perform murine experiments of chemotherapy plus ICI and an attempt to build a predictor of this, which is the treatment regimen used for human patients with TNBC [i.e., KEYNOTE-522 regimen; ref 4], because we did not want to preface an ICI predictor on the foundation of a failed chemotherapy predictor.

Our study has several limitations, including that the biological replicate tumor samples used for bulk RNA-seq were different from those used for the survival studies. Intratumoral heterogeneity and stochastic biological variation between tumor replicates could introduce noise that affects the predictive accuracy of our models. However, through hierarchical clustering of the bulk RNA-seq data, we showed that the transcriptomic states of mouse tumor biological replicates are stable and reproducible for the models we tested. Additionally, we mitigated this limitation by using tumor model– and treatment-specific median survival times when modeling prognosis and response to treatment to reduce the impact of individual mouse variability. Our aggregated tumor model analysis yielded slightly worse performance on the human SCAN-B dataset when compared with our analysis using individual mouse biological replicates. Although these findings are consistent with less overfitting to the training dataset, additional

(*Continued*.) samples, in which immune signature scores are used to predict T-cell expansion in pretreatment samples ($n = 40$). **G,** ROC curve for Bassez pseudobulk samples, in which immune signature scores are used to predict T-cell expansion in on-treatment samples ($n = 40$). **H,** ROC curve for Gide samples, in which immune signature scores are used to predict tumor response (partial/complete response vs. stable/progressive disease) using the pretreatment samples ($n = 72$). **I,** ROC curve for Gide samples, in which immune signature scores are used to predict tumor response (partial/complete response vs. stable/progressive disease) using the posttreatment samples ($n = 18$). **J,** Kaplan–Meier plots for immune signature scores for Gide pretreatment samples ($n = 72$), with predictions split into above- and below-median groups. **K,** Kaplan–Meier plots for immune signature scores for Gide posttreatment samples ($n = 18$), with predictions split into above- and below-median groups.

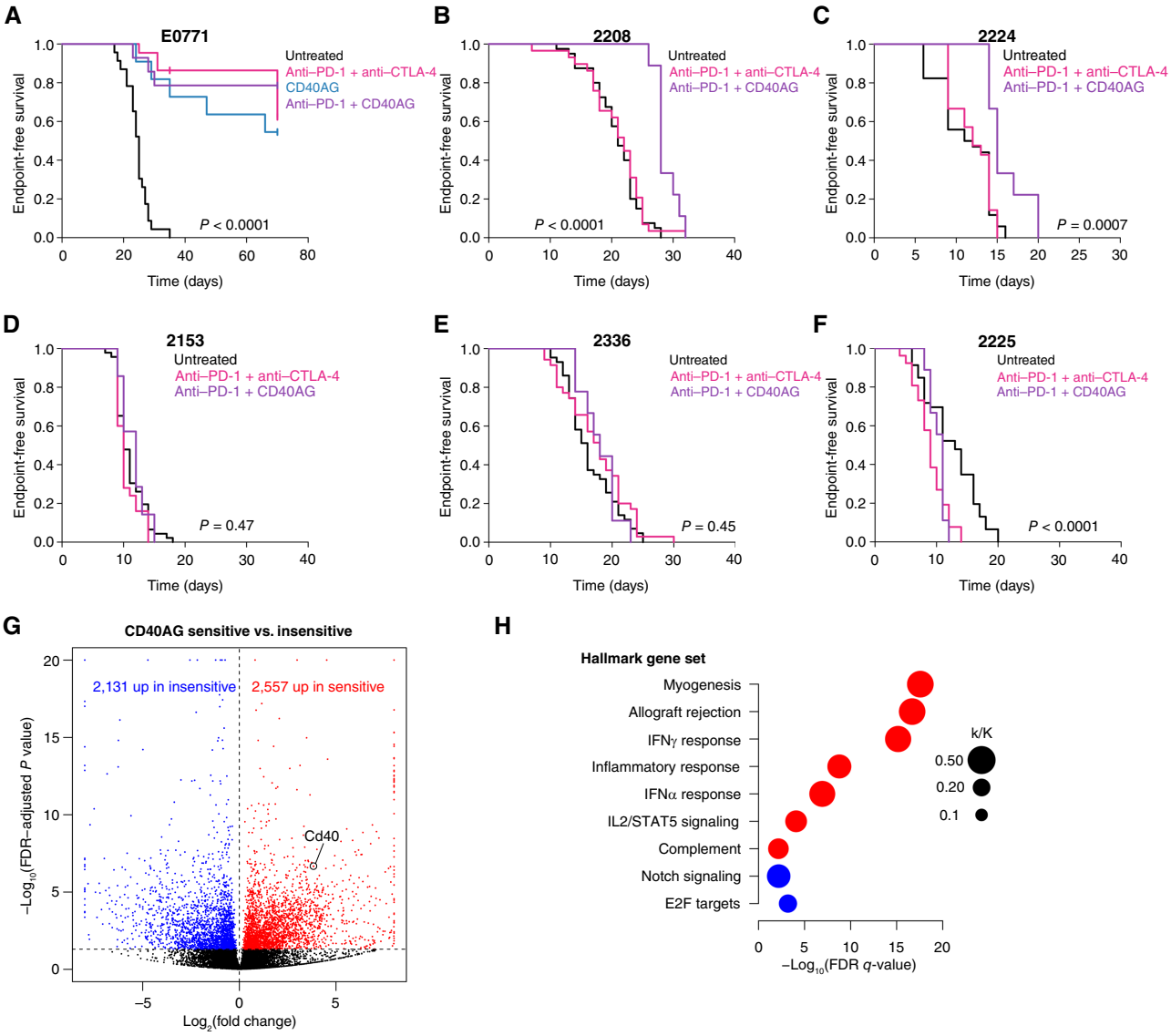


Figure 5.

CD40 agonist treatment in ICI-sensitive and ICI-resistant tumor models. **A,** Kaplan–Meier plots for E0771 ($n = 23$ untreated mice; $n = 22$ anti–PD-1 + anti–CTLA-4–treated mice; $n = 11$ CD40AG-treated mice; $n = 14$ anti–PD-1 + CD40AG–treated mice). P value using the log-rank test. **B,** Kaplan–Meier plots for 2208 ($n = 40$ untreated mice; $n = 29$ anti–PD-1 + anti–CTLA-4–treated mice; $n = 9$ anti–PD-1 + CD40AG–treated mice). P value using the log-rank test. **C,** Kaplan–Meier plots for 2224 ($n = 34$ untreated mice; $n = 21$ anti–PD-1 + anti–CTLA-4–treated mice; $n = 9$ anti–PD-1 + CD40AG–treated mice). P value using the log-rank test. **D,** Kaplan–Meier plot for 2153 ($n = 46$ untreated mice; $n = 25$ anti–PD-1 + anti–CTLA-4–treated mice; $n = 7$ anti–PD-1 + CD40AG–treated mice). P value using the log-rank test. **E,** Kaplan–Meier plot for 2336 ($n = 43$ untreated mice; $n = 35$ anti–PD-1 + anti–CTLA-4–treated mice; $n = 9$ anti–PD-1 + CD40AG–treated mice). P value using the log-rank test. **F,** Kaplan–Meier plot for 2225 ($n = 46$ untreated mice; $n = 26$ anti–PD-1 + anti–CTLA-4–treated mice; $n = 9$ anti–PD-1 + CD40AG–treated mice). P value using the log-rank test. **G,** Volcano plot of untreated CD40AG-sensitive mouse tumors (E0771, 2208, and 2224) vs. untreated CD40AG-insensitive mouse tumors (2153, 2336, and 2225). Highlighted genes are those with an FDR-adjusted P value < 0.05. **H,** MSigDB gene set enrichment analysis of 2,557 genes related to CD40AG sensitivity and 2,131 genes related to CD40AG resistance from **G**. The size of the circles indicates the proportion of overlapping genes between the genes (k) and the gene sets (K).

studies with larger numbers of independent tumor models would be needed to determine the optimal training strategy and to assess the robustness of the selected biological features. A second limitation is that we only sampled tumors for RNA-seq after 7 days of treatment rather than at multiple timepoints. We chose this timepoint to characterize the early treatment response, which we believe to be most relevant for biomarker prediction; however, this limits the biological insight into longer-term adaptive resistance that could be captured in a different experimental design. A third limitation is that we separately tested ICI and chemotherapy, whereas in clinical settings, most patients with TNBC receive both chemotherapy and ICI. A fourth limitation is that our mouse model set did not contain any estrogen receptor–positive, hormonally driven models, which is the most common clinical human breast cancer subtype; however, our mouse

Table 2. Comparison of machine learning methods for the untreated prognostic model trained on mouse tumors and tested on human SCAN-B.

SCAN-B set	Elastic net regression	XGBoost	Random forest	SVR
All	0.657	0.619	0.633	0.651
HR+/HER2−	0.614	0.607	0.622	0.602
HER2+	0.667	0.605	0.611	0.631
TNBC	0.530	0.571	0.566	0.526

models contained "luminal" murine tumors, and the basal–luminal axis was prognostic and translated to prognostic differences across human tumors. These results demonstrate that many clinically relevant features are conserved across species, although some are not. Although the mouse may not accurately model all aspects of human breast cancers, it successfully models some, including baseline prognosis and many aspects of adaptive immunity (75).

In summary, our diverse and "population-based" mouse mammary tumor dataset has both prognostic and predictive capabilities and extends these capabilities to human datasets. Additionally, the gene expression module features utilized in this study are publicly available and conveniently accessible as an R package. These modules can readily be extended with new or custom gene expression signatures and applied to new datasets to summarize and interpret the biology of mouse mammary or human breast cancer datasets. This murine data resource is a rich dataset for new discoveries and validation of existing biomarkers and should continue to interface with similar human datasets to test new hypotheses and computational methods.

Data Availability

The underlying mouse tumor gene expression is available at GSE223630, GSE124821, and GSE304115. See Supplementary Table S1 for the list of gene expression modules and their associated Entrez IDs. See Supplementary Table S2 for GEO sample metadata. Mouse survival data can be found in Supplementary Table S3. Median survival ratios and associated log-rank test statistics can be found in Supplementary Table S4. Elastic net models with weights and immune response signature genes can be found in Supplementary Table S5. All other raw data are available upon request to the corresponding author.

Code Availability

The code to calculate the 894 gene expression modules is available as an R package at GitHub: https://github.com/perou-lab/Breast-Cancer-Modules. The code used to generate the machine learning models for this study is available at https://github.com/perou-lab/mouse-mammary-tumor-modeling.

Authors' Disclosures

C. Brueffer reports personal fees and other support from InSilico Consulting AB and SAGA Diagnostics US, Inc. outside the submitted work, as well as a patent for US patent number 18/240,445 pending. L.H. Saal reports grants from Mrs. Berta Kamprad Foundation, The Swedish Cancer Society, Swedish Research Council, and Swedish Research Council for Health, Working Life and Welfare (FORTE) during the conduct of the study, as well as other support from SAGA Diagnostics US, Inc., grants from Gilead Sciences, and personal fees from DoMore Diagnostics outside the submitted work. C.M. Perou reports grants from NCI Breast Specialized Program of Research Excellence (SPORE) program P50-CA58223, NCI U01CA238475, Breast Cancer Research Foundation (BCRF-23-127), Susan G. Komen (SAC-160074), National Institute of General Medical Sciences (5T32 GM067553), UNC Integrated Translational Oncology Program (iTOP) (T32-CA24415), and UNC Lineberger Triple Negative Breast Cancer Center during the conduct of the study; personal fees from Bioclassifier LLC outside the submitted work; and a patent for US patent number 12,995,459 issued, licensed, and with royalties paid from Bioclassifier LLC. No disclosures were reported by the other authors.

Authors' Contributions

M.D. Sutcliffe: Conceptualization, data curation, formal analysis, investigation, visualization, methodology, writing–original draft, writing–review and editing. **K.R. Mott:** Conceptualization, resources, data curation, supervision, funding acquisition, methodology, project administration, writing–review and editing. **T. Yilmaz-Swenson:** Data curation, software, writing–review and editing. **B.M. Felsheim:** Conceptualization, resources, software, supervision, funding acquisition, methodology, project administration, writing–review and editing. **A.V. Lobanov:** Data curation, software, funding acquisition, writing–review and editing. **A.R. Michmerhuizen:** Data curation, software, writing–review and editing. **P.D. Rädler:** Data curation, software, funding acquisition, methodology, writing–review and editing. **D.O. Okumu:** Data curation, software, funding acquisition, methodology, writing–review and editing. **X. He:** Data curation, funding acquisition, writing–review and editing. **A.D. Pfefferle:** Data curation, software, methodology, writing–review and editing. **S. Dance-Barnes:** Data curation, funding acquisition, writing–review and editing. **C. Brueffer:** Resources, data curation, writing–review and editing. **L.H. Saal:** Resources, data curation, writing–review and editing. **M.P. East:** Data curation, writing–review and editing. **D.P. Hollern:** Conceptualization, data curation, software, funding acquisition, writing–review and editing. **T.C. Elston:** Data curation, software, funding acquisition, writing–review and editing. **G.L. Johnson:** Conceptualization, resources, data curation, software, supervision, funding acquisition, methodology, project administration, writing–review and editing. **C.M. Perou:** Conceptualization, resources, data curation, software, supervision, funding acquisition, methodology, project administration, writing–review and editing.

Acknowledgments

This research was in part supported by funds from the NCI U01CA238475, the Breast SPORE program (P50-CA058223), the Breast Cancer Research Foundation (BCRF-23-127), the Susan G. Komen (SAC-160074), the National Institute of General Medical Sciences (5T32 GM067553), UNC iTOP (T32-CA24415), and the UNC Lineberger Triple Negative Breast Cancer Center. We would also like to thank the UNC Lineberger Comprehensive Cancer Center (LCCC) Office of Genomics Research, the UNC LCCC Translational Genomics Lab (RRID: SCR_025231), and the UNC High-Throughput Sequencing Facility.

Note

Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Received March 3, 2026; revised June 3, 2026; accepted August 12, 2026; posted first August 14, 2026.

References

1. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. J Clin Invest 2011;121:2750–67. https://doi.org/10.1172/JCI45014.
2. Prat A, Lluch A, Albanell J, Barry WT, Fan C, Chacón JI, et al. Predicting response and survival in chemotherapy-treated triple-negative breast cancer. Br J Cancer 2014;111:1532–41. https://doi.org/10.1038/bjc.2014.444.

3. Gao R, Davis A, McDonald TO, Sei E, Shi X, Wang Y, et al. Punctuated copy number evolution and clonal stasis in triple-negative breast cancer. Nat Genet 2016;48:1119–30. https://doi.org/10.1038/ng.3641.
4. Schmid P, Cortes J, Pusztai L, McArthur H, Kümmel S, Bergh J, et al. Pembrolizumab for early triple-negative breast cancer. N Engl J Med 2020;382:810–21. https://doi.org/10.1056/NEJMoa1910549.
5. Riaz F, Gruber JJ, Telli ML. New treatment approaches for triple-negative breast cancer. Am Soc Clin Oncol Educ Book 2025;45:e481154. https://doi.org/10.1200/EDBK-25-481154.
6. Chuprin J, Buettner H, Seedhom MO, Greiner DL, Keck JG, Ishikawa F, et al. Humanized mouse models for immuno-oncology research. Nat Rev Clin Oncol 2023;20:192–206. https://doi.org/10.1038/s41571-022-00721-2.
7. Magré L, Verstegen MMA, Buschow S, van der Laan LJW, Peppelenbosch M, Desai J. Emerging organoid-immune co-culture models for cancer research: from oncoimmunology to personalized immunotherapies. J Immunother Cancer 2023;11:e006290. https://doi.org/10.1136/jitc-2022-006290.
8. Wang J, Tao X, Zhu J, Dai Z, Du Y, Xie Y, et al. Tumor organoid-immune co-culture models: exploring a new perspective of tumor immunity. Cell Death Discov 2025;11:195. https://doi.org/10.1038/s41420-025-02407-x.
9. Jerry DJ, Kittrell FS, Kuperwasser C, Laucirica R, Dickinson ES, Bonilla PJ, et al. A mammary-specific model demonstrates the role of the p53 tumor suppressor gene in tumor development. Oncogene 2000;19:1052–8. https://doi.org/10.1038/sj.onc.1203270.
10. Herschkowitz JI, Zhao W, Zhang M, Usary J, Murrow G, Edwards D, et al. Comparative oncogenomics identifies breast tumors enriched in functional tumor-initiating cells. Proc Natl Acad Sci U S A 2012;109:2778–83. https://doi.org/10.1073/pnas.1018862108.
11. Pfefferle AD, Herschkowitz JI, Usary J, Harrell JC, Spike BT, Adams JR, et al. Transcriptomic classification of genetically engineered mouse models of breast cancer identifies human subtype counterparts. Genome Biol 2013;14:R125. https://doi.org/10.1186/gb-2013-14-11-r125.
12. Pfefferle AD, Agrawal YN, Koboldt DC, Kanchi KL, Herschkowitz JI, Mardis ER, et al. Genomic profiling of murine mammary tumors identifies potential personalized drug targets for p53-deficient mammary cancers. Dis Model Mech 2016;9:749–57. https://doi.org/10.1242/dmm.025239.
13. Hollern DP, Contreras CM, Dance-Barnes S, Silva GO, Pfefferle AD, Xiong J, et al. A mouse model featuring tissue-specific deletion of p53 and Brca1 gives rise to mammary tumors with genomic and transcriptomic similarities to human basal-like breast cancer. Breast Cancer Res Treat 2019;174:143–55. https://doi.org/10.1007/s10549-018-5061-y.
14. Hollern DP, Andrechek ER. A genomic analysis of mouse models of breast cancer reveals molecular features of mouse models and relationships to human breast cancer. Breast Cancer Res 2014;16:R59. https://doi.org/10.1186/bcr3672.
15. Rennhack J, To B, Wermuth H, Andrechek ER. Mouse models of breast cancer share amplification and deletion events with human breast cancer. J Mammary Gland Biol Neoplasia 2017;22:71–84. https://doi.org/10.1007/s10911-017-9374-y.
16. Dance-Barnes ST, Evans E, Steward S, He X, Usary J, Perou C. Abstract A24: characterization of a novel p53; BRCA1-deficient claudin-low mammary tumor cell line. Cancer Res 2016;76(3_Supplement):A24–4. https://doi.org/10.1158/1538-7445.FBCR15-A24.
17. Hollern DP, Xu N, Thennavan A, Glodowski C, Garcia-Recio S, Mott KR, et al. B cells and T follicular helper cells mediate response to checkpoint inhibitors in high mutation burden mouse models of breast cancer. Cell 2019;179:1191–206.e21. https://doi.org/10.1016/j.cell.2019.10.028.
18. Pulaski BA, Ostrand-Rosenberg S. Mouse 4T1 breast tumor model. Curr Protoc Immunol 2001 May;Chapter 20:Unit 20.2. https://doi.org/10.1002/0471142735.im2002s39.
19. Casey AE, Laster WR, Ross GL. Sustained enhanced growth of carcinoma EO771 in C57 black mice. Proc Soc Exp Biol Med 1951;77:358–62. https://doi.org/10.3181/00379727-77-18779.
20. An Y, Adams JR, Hollern DP, Zhao A, Chang SG, Gams MS, et al. Cdh1 and Pik3ca mutations cooperate to induce immune-related invasive lobular carcinoma of the breast. Cell Rep 2018;25:702–14.e6. https://doi.org/10.1016/j.celrep.2018.09.056.
21. Maroulakou IG, Anver M, Garrett L, Green JE. Prostate and mammary adenocarcinoma in transgenic mice carrying a rat C3(1) simian virus 40 large tumor antigen fusion gene. Proc Natl Acad Sci U S A 1994;91:11236–40. https://doi.org/10.1073/pnas.91.23.11236.
22. Guy CT, Cardiff RD, Muller WJ. Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease. Mol Cell Biol 1992;12:954–61. https://doi.org/10.1128/mcb.12.3.954-961.1992.
23. Muller WJ, Arteaga CL, Muthuswamy SK, Siegel PM, Webster MA, Cardiff RD, et al. Synergistic interaction of the Neu proto-oncogene product and transforming growth factor alpha in the mammary epithelium of transgenic mice. Mol Cell Biol 1996;16:5726–36. https://doi.org/10.1128/MCB.16.10.5726.
24. Lim C, Hwang D, Yazdimamaghani M, Atkins HM, Hyun H, Shin Y, et al. High-dose paclitaxel and its combination with CSF1R inhibitor in polymeric micelles for chemoimmunotherapy of triple negative breast cancer. Nano Today 2023;51:101884. https://doi.org/10.1016/j.nantod.2023.101884.
25. Frankish A, Diekhans M, Ferreira AM, Johnson R, Jungreis I, Loveland J, et al. GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res 2019;47:D766–73. https://doi.org/10.1093/nar/gky955.
26. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15–21. https://doi.org/10.1093/bioinformatics/bts635.
27. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 2017;14:417–9. https://doi.org/10.1038/nmeth.4197.
28. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550. https://doi.org/10.1186/s13059-014-0550-8.
29. McGrail DJ, Lin CCJ, Dai H, Mo W, Li Y, Stephan C, et al. Defective replication stress response is inherently linked to the cancer stem cell phenotype. Cell Rep 2018;23:2095–106. https://doi.org/10.1016/j.celrep.2018.04.068.
30. Prat A, Lluch A, Turnbull AK, Dunbier AK, Calvo L, Albanell J, et al. A PAM50-Based chemoendocrine score for hormone receptor-positive breast cancer with an intermediate risk of relapse. Clin Cancer Res 2017;23:3035–44. https://doi.org/10.1158/1078-0432.CCR-16-2092.
31. Broz ML, Binnewies M, Boldajipour B, Nelson AE, Pollack JL, Erle DJ, et al. Dissecting the tumor myeloid compartment reveals rare activating antigen-presenting cells critical for T cell immunity. Cancer Cell 2014;26:638–52. https://doi.org/10.1016/j.ccell.2014.09.007.
32. Cui Y, Li B, Pollom EL, Horst KC, Li R. Integrating radiosensitivity and immune gene signatures for predicting benefit of radiotherapy in breast cancer. Clin Cancer Res 2018;24:4754–62. https://doi.org/10.1158/1078-0432.CCR-18-0825.
33. Bill R, Wirapati P, Messemaker M, Roh W, Zitti B, Duval F, et al. CXCL9: SPP1 macrophage polarity identifies a network of cellular programs that control human cancers. Science 2023;381:515–24. https://doi.org/10.1126/science.ade2292.
34. Parker JS, Mullins M, Cheang MCU, Leung S, Voduc D, Vickery T, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol 2009;27:1160–7. https://doi.org/10.1200/JCO.2008.18.1370.
35. Paik S, Shak S, Tang G, Kim C, Baker J, Maureen C, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 2004;351:2817–26. https://doi.org/10.1056/NEJMoa041588.
36. Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res 2010;12:R68. https://doi.org/10.1186/bcr2635.
37. Garcia-Recio S, Thennavan A, East MP, Parker JS, Cejalvo JM, Garay JP, et al. FGFR4 regulates tumor subtype differentiation in luminal breast cancer and metastatic disease. J Clin Invest 2020;130:4871–87. https://doi.org/10.1172/JCI130323.
38. Rody A, Karn T, Liedtke C, Pusztai L, Ruckhaeberle E, Hanker L, et al. A clinically relevant gene signature in triple negative and basal-like breast cancer. Breast Cancer Res 2011;13:R97. https://doi.org/10.1186/bcr3035.
39. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AAM, Mao M, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002;415:530–6. https://doi.org/10.1038/415530a.
40. Saal LH, Vallon-Christersson J, Häkkinen J, Hegardt C, Grabau D, Winter C, et al. The Sweden Cancerome Analysis Network - Breast (SCAN-B) initiative: a large-scale multicenter infrastructure towards implementation of breast cancer genomic analyses in the clinical routine. Genome Med 2015;7:20. https://doi.org/10.1186/s13073-015-0131-9.
41. Staaf J, Häkkinen J, Hegardt C, Saal LH, Kimbung S, Hedenfalk I, 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;8:94. https://doi.org/10.1038/s41523-022-00465-3.

42. Cha YJ, O'Connell CE, Calhoun BC, Felsheim BM, Fernandez-Martinez A, Fan C, et al. Genomic characteristics related to histology-based immune features in breast cancer. Mod Pathol 202538:100736. https://doi.org/10.1016/j.modpat.2025.100736.
43. Harrell JC, Prat A, Parker JS, Fan C, He X, Carey L, et al. Genomic analysis identifies unique signatures predictive of brain, lung, and liver relapse. Breast Cancer Res Treat 2012;132:523–35. https://doi.org/10.1007/s10549-011-1619-7.
44. Marchionni L, Afsari B, Geman D, Leek JT. A simple and reproducible breast cancer prognostic test. BMC Genomics 2013;14:336. https://doi.org/10.1186/1471-2164-14-336.
45. Bassez A, Vos H, Van Dyck L, Floris G, Arijs I, Desmedt C, et al. A single-cell map of intratumoral changes during anti-PD1 treatment of patients with breast cancer. Nat Med 2021;27:820–32. https://doi.org/10.1038/s41591-021-01323-8.
46. Gide TN, Quek C, Menzies AM, Tasker AT, Shang P, Holst J, et al. Distinct immune cell populations define response to anti-PD-1 monotherapy and anti-PD-1/anti-CTLA-4 combined therapy. Cancer Cell 2019;35:238–55.e6. https://doi.org/10.1016/j.ccell.2019.01.003.
47. Leinonen R, Akhtar R, Birney E, Bower L, Cerdeno-Tárraga A, Cheng Y, et al. The European nucleotide archive. Nucleic Acids Res 2011;39:D28–31. https://doi.org/10.1093/nar/gkq967.
48. Shepherd JH, Ballman K, Polley MYC, Campbell JD, Fan C, Selitsky S, 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;40:1323–34. https://doi.org/10.1200/JCO.21.01506.
49. Felsheim BM, Fernandez-Martinez A, Fan C, Pfefferle AD, Hayward MC, Hoadley KA, et al. Prognostic and molecular multi-platform analysis of CALGB 40603 (alliance) and public triple-negative breast cancer datasets. NPJ Breast Cancer 2025;11:24. https://doi.org/10.1038/s41523-025-00740-z.
50. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545–50. https://doi.org/10.1073/pnas.0506580102.
51. Castanza AS, Recla JM, Eby D, Thorvaldsdóttir H, Bult CJ, Mesirov JP. Extending support for mouse data in the molecular signatures database (MSigDB). Nat Methods 2023;20:1619–20. https://doi.org/10.1038/s41592-023-02014-7.
52. Koboldt DC, Fulton RS, McLellan MD, Schmidt H, Kalicki-Veizer J, McMichael JF, et al. Comprehensive molecular portraits of human breast tumours. Nature 2012;490:7418. https://doi.org/10.1038/nature11412.
53. Fernandez-Martinez A, Pascual T, Singh B, Nuciforo P, Rashid NU, Ballman KV, 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;9:490–9. https://doi.org/10.1001/jamaoncol.2022.6288.
54. Fan C, Prat A, Parker JS, Liu Y, Carey LA, Troester MA, et al. Building prognostic models for breast cancer patients using clinical variables and hundreds of gene expression signatures. BMC Med Genomics 2011;4:3. https://doi.org/10.1186/1755-8794-4-3.
55. Fernandez-Martinez A, Rediti M, Tang G, Pascual T, Hoadley KA, Venet D, et al. Tumor intrinsic subtypes and gene expression signatures in early-stage ERBB2/HER2-Positive breast cancer: a pooled analysis of CALGB 40601, NeoALTTO, and NSABP B-41 trials. JAMA Oncol 2024;10:603–11. https://doi.org/10.1001/jamaoncol.2023.7304.
56. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database (MSigDB) hallmark gene set collection. Cell Syst 2015;1:417–25. https://doi.org/10.1016/j.cels.2015.12.004.
57. van de Vijver MJ, He YD, van 't Veer LJ, Dai H, Hart AAM, Voskuil DW, et al. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 2002;347:1999–2009. https://doi.org/10.1056/NEJMoa021967.
58. Roussos Torres ET, Ho WJ, Danilova L, Tandurella JA, Leatherman J, Rafie C, et al. Entinostat, nivolumab and ipilimumab for women with advanced HER2-negative breast cancer: a phase Ib trial. Nat Cancer 2024;5:866–79. https://doi.org/10.1038/s43018-024-00729-w.
59. Nederlof I, Isaeva OI, de Graaf M, Gielen RCAM, Bakker NAM, Rolfes AL, et al. Neoadjuvant nivolumab or nivolumab plus ipilimumab in early-stage triple-negative breast cancer: a phase 2 adaptive trial. Nat Med 2024;30:3223–35. https://doi.org/10.1038/s41591-024-03249-3.
60. Elgueta R, Benson MJ, de Vries VC, Wasiuk A, Guo Y, Noelle RJ. Molecular mechanism and function of CD40/CD40L engagement in the immune system. Immunol Rev 2009;229:152–72. https://doi.org/10.1111/j.1600-065X.2009.00782.x.
61. van Kooten C, Banchereau J. Functions of CD40 on B cells, dendritic cells and other cells. Curr Opin Immunol 1997;9:330–7. https://doi.org/10.1016/s0952-7915(97)80078-7.
62. King C. New insights into the differentiation and function of T follicular helper cells. Nat Rev Immunol 2009;9:757–66. https://doi.org/10.1038/nri2644.
63. Byrne KT, Betts CB, Mick R, Sivagnanam S, Bajor DL, Laheru DA, et al. Neoadjuvant selicrelumab, an agonist CD40 antibody, induces changes in the tumor microenvironment in patients with resectable pancreatic cancer. Clin Cancer Res 2021;27:4574–86. https://doi.org/10.1158/1078-0432.CCR-21-1047.
64. Van Laethem JL, Borbath I, Prenen H, Geboes KP, Lambert A, Mitry E, et al. Combining CD40 agonist mitazalimab with mFOLFIRINOX in previously untreated metastatic pancreatic ductal adenocarcinoma (OPTIMIZE-1): a single-arm, multicentre phase 1b/2 study. Lancet Oncol 2024;25:853–64. https://doi.org/10.1016/S1470-2045(24)00263-8.
65. Osorio JC, Knorr DA, Weitzenfeld P, Blanchard L, Yao N, Baez M, et al. Fc-optimized CD40 agonistic antibody elicits tertiary lymphoid structure formation and systemic antitumor immunity in metastatic cancer. Cancer Cell 2025;43:1902–16.e9. https://doi.org/10.1016/j.ccell.2025.07.013.
66. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining [Internet]. San Francisco (CA). Association for Computing Machinery; 2016. p. 785–94. https://doi.org/10.1145/2939672.2939785.
67. Herschkowitz JI, Simin K, Weigman VJ, Mikaelian I, Usary J, Hu Z, et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. Genome Biol 2007;8:R76. https://doi.org/10.1186/gb-2007-8-5-r76.
68. Fernandez-Martinez A, Krop IE, Hillman DW, Polley MY, Parker JS, Huebner L, et al. Survival, pathologic response, and genomics in CALGB 40601 (alliance), a neoadjuvant phase III trial of paclitaxel-trastuzumab with or without lapatinib in HER2-Positive breast cancer. J Clin Oncol 2020;38:4184–93. https://doi.org/10.1200/JCO.20.01276.
69. Breiman L. Random forests. Mach Learn 2001;45:5–32. https://doi.org/10.1023/A:1010933404324.
70. Drucker H, Burges CJC, Kaufman L, Smola A, Vapnik V. Support vector regression machines. In: Proceedings of the 10th international conference on neural information processing systems. Denver (CO): MIT Press; 1996. p. 155–61. (NIPS'96).
71. Howard FM, Fasching PA, Santa-Maria CA, Lim E, Sparano JA, Lustberg MB, et al. Machine learning-based prediction of distant recurrence risk and ribociclib treatment effect in HR+/HER2− early breast cancer using real-world and NATALEE data. Clin Cancer Res 2026;32:428–37. https://doi.org/10.1158/1078-0432.CCR-25-1946.
72. Maros ME, Capper D, Jones DTW, Hovestadt V, von Deimling A, Pfister SM, et al. Machine learning workflows to estimate class probabilities for precision cancer diagnostics on DNA methylation microarray data. Nat Protoc 2020;15:479–512. https://doi.org/10.1038/s41596-019-0251-6.
73. Khalil M, Vonderheide RH. Anti-CD40 agonist antibodies: preclinical and clinical experience. Update Cancer Ther 2007;2:61–5. https://doi.org/10.1016/j.uct.2007.06.001.
74. Sikov WM, Berry DA, Perou CM, Singh B, Cirrincione CT, Tolaney SM, et al. Impact of the addition of carboplatin and/or bevacizumab to neoadjuvant once-per-week paclitaxel followed by dose-dense doxorubicin and cyclophosphamide on pathologic complete response rates in stage II to III triple-negative breast cancer: CALGB 40603 (alliance). J Clin Oncol 2015;33:13–21. https://doi.org/10.1200/JCO.2014.57.0572.
75. Leach DR, Krummel MF, Allison JP. Enhancement of antitumor immunity by CTLA-4 blockade. Science 1996;271:1734–6. https://doi.org/10.1126/science.271.5256.1734.