Breast tumors educate stromal tissue with individualized but coordinated proteomic signatures

Xuya Wang[1,2,3,*], **Arshag D. Mooradian**[4,*], **Petra Erdmann-Gilmore**[4], **Qiang Zhang**[4], **Rosa Viner**[9], **Sherri R. Davies**[4], **Kuan-lin Huang**[4], **Ryan Bomgarden**[10], **Brian A. Van Tine**[4,5], **Jieya Shao**[4,5], **Li Ding**[4,5,8], **Shunqiang Li**[4,5], **Matthew J. Ellis**[11], **John C. Rogers**[10], **R. Reid Townsend**[4,5], **David Fenyö**[1,2,3,#], and **Jason M. Held**[4,5,7,#]

[1]Institute for Systems Genetics, New York University, New York, NY 10016, USA

[2]Center for Health Informatics and Bioinformatics, New York University, New York, NY 10016, USA

[3]Department of Biochemistry and Molecular Pharmacology, New York University, New York, NY 10016, USA

[4]Deparment of Medicine, Washington University in Saint Louis Medical School, St. Louis, MO 63110, USA

[5]Siteman Cancer Center, Washington University in Saint Louis Medical School, St. Louis, MO 63110, USA

[6]Department of Genetics, Washington University in Saint Louis Medical School, St. Louis, MO 63110, USA

[7]Department of Anesthesiology, Washington University in Saint Louis Medical School, St. Louis, MO 63110, USA

[8]McDonnell Genome Institute, Washington University in Saint Louis Medical School, St. Louis, MO 63110, USA

[9]Thermo Fisher Scientific, San Jose, CA 95134, USA

[10]Thermo Fisher Scientific, Rockford, IL 61101, USA

[11]Lester and Sue Smith Breast Care Center, Department of Oncology, and Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX 77030, USA

Abstract

[#]Corresponding author. David@FenyoLab.org (D.F.); jheld@wustl.edu (J.M.H.).
[*]These authors contributed equally to this work

Cancer forms specialized microenvironmental niches that promote local invasion and colonization. Engrafted patient-derived xenografts (PDXs) locally invade and colonize naïve stroma, while enabling unambiguous molecular discrimination of human proteins in the tumor from mouse proteins in the microenvironment. To characterize how patient breast tumors form a niche and educate naïve stroma, subcutaneous breast cancer PDXs were globally profiled using species-specific quantitative proteomics. Regulation of PDX stromal proteins by breast tumors was extensive, with thirty-five percent of the stromal proteome consistently altered by tumors across different animals and passages. Differentially regulated proteins in the stroma clustered into six signatures that included both known and novel contributors to tumor invasion and colonization. Stromal proteomes were coordinately regulated, though the sets of proteins altered by each tumor were highly distinct. Integrated analysis of tumor and stromal proteins, a comparison possible in xenograft models, indicated that the known hallmarks of cancer contribute pleiotropically to establishing and maintaining the tumor's microenvironmental niche. Tumor education of the stroma is therefore an intrinsic property of breast tumors that is highly individualized, yet proceeds by consistent, non-random and defined tumor-promoting molecular alterations.

Introduction

Tumors typically encounter a naïve microenvironment twice: during local invasion of the primary tumor and upon metastatic dissemination and acclimatization to foreign tissues. Pernicious tumors are adept at remodeling the microenvironment to co-evolve tumor-promoting niches. Understanding how primary tumors acquire the ability to direct modification of normal tissue remains a rate-limiting step in our ability to constrain tumor growth and metastasis.

In breast cancer, the stroma plays an important role in progression (1) and resistance to therapy (2, 3). Genomic analysis shows close clonal relationships between primary tumors and their metastases, and no metastasis-specific mutations have been identified to date (4, 5). Together, this has led to the prevailing notion that a key mechanisms underlying tumor growth and metastasis is interaction of a tumor with the neighboring stroma to educate its microenvironment (6–8) and that this occurs via tumor-intrinsic changes in the transcriptome, epigenome, or proteome. However, the lack of experimental models with sufficient biological complexity to both model the multicellular tumor microenvironment and to allow unambiguous delineation of the tumor from the naïve stroma have limited molecular characterization of stromal education by tumors. Therefore, it is unclear what component of a tumor's education of the stroma is cell-autonomous and tumor intrinsic, versus an adaptation to the host's cues or resources in the naïve or tumor-associated microenvironment. Proteomic investigation of the tumor-stroma interface is therefore a much needed step to clarify the regulatory programs that enable development of the cellularly heterogeneous tumor microenvironmental niche.

Patient-derived xenograft (PDX) models provide the most reproducible experimental approximation of primary human tumors for cancer research (9). Like primary tumors, PDXs are complex tissues composed of multiple distinct cell types that heterotypically interact. PDXs model the architecture of the original tumor and prove to be valuable as

preclinical models for drug treatment studies (10, 11). Tumors evolve through a multi-step progression in which they acquire a succession of metastatic and survival capabilities. Serially passaging PDXs serves as a basic model of both local invasion and metastasis in which cancer cells disseminate from a supportive microenvironment and need to form a new niche in a new environment in order to survive (12, 13). PDXs are rather stable genetically and cytogenetically during serial passage (14) as well as in proteomic and phosphoproteomic analyses (15).

Breast cancer stroma are extensively studied at the mRNA level, and analysis of large primary tumor datasets finds that stromal gene expression signatures can predict clinical outcome in both breast (16) and colorectal (17) cancer. Proteome-level characterization of the stroma is limited to targeted analysis with specific antibodies used for immunohistochemistry, or proteomic characterization of isolated stromal subfractions such as the matrisome (18) or decellularized ECM (19). Several mRNA studies focus on the stroma in PDX models and confirm stable gene expression patterns (13, 20). However, PDXs present a unique opportunity to differentiate proteins in patient-derived tumors from the naïve, murine-derived microenvironment by using species-specific differences in amino acid sequence, thus eliminating the need for physical or biochemical separation. Though large-scale unbiased proteomic characterization of stromal remodeling has not been performed, it holds tremendous potential to clarify the tumor-stroma interaction, as many important factors in the microenvironment are regulated at the protein level, including enzymatic remodeling and export of regulatory proteins into the extracellular matrix (ECM).

The persistence, prevalence, heterogeneity, and coordination of stromal education by individual patient tumors remain largely unexamined, perhaps because analysis of a single patient tumor sample, without the availability of biological replicates, cannot distinguish whether molecular changes in the microenvironment are due to normal variation, sample preparation variability, or tumor-specific regulation. We sought to investigate the process by which breast tumors establish a tumor-promoting niche using cross-species proteomic profiling of breast cancer PDXs. As a starting point to investigate tumor education of their microenvironments, we used subcutaneously-implanted breast cancer PDXs, a common model of tumor-host interactions. Our goal was to determine if there is heterogeneity in patient breast tumors' stromal education and to delineate which components of the stromal proteome are susceptible to regulation by breast tumors. We found that regulation of PDX stromal proteins by tumors was persistent, individualized, and biologically organized.

Results

Species-specific proteomic characterization of breast cancer PDXs reveals widespread, tumor-specific stromal regulation

The seven subcutaneous breast cancer PDXs that were proteomically profiled (workflow in fig. S1) were previously described (15) and represent basal (n=3), HER2-enriched (HER2E; n=1), luminal B (n=2), and claudin-low (n=1) breast cancer subtypes. No PDX was analyzed prior to three passages in mice (table S1) to allow full exchange of the tumor's stroma with mouse elements (21). Three biological replicates of each patient tumor representing multiple PDX passages were analyzed. All PDXs were passaged in female mice except for six basal

tumors, which are estrogen receptor-negative and were passaged in male mice (table S1). All breast tumors used to generate PDXs in this study metastasized and three of the PDXs (WHIMs 2, 12 and 20) often formed metastases in mice (table S1); therefore, the PDX models investigated have metastatic capacity. The engrafted PDX tumors were carefully dissected away from normal mouse tissues once they reached 1500 mm^3 and proteomic analysis was performed to identify species-unique proteins, genes, and peptides from each species (Fig. 1A). Only species-unique peptides were used to accurately quantify and distinguish the tumor proteome (with human-unique peptide sequences) from the murine-derived proteins of the naïve tumor microenvironment (fig. S1). Protein products of 2,038 human and 415 mouse genes were detected with gene-specific peptides across all samples after filtering to remove abundant murine plasma proteins and erythrocyte markers (fig. S2; see also Materials and Methods).

Relative quantification was performed with TMT10 isobaric tagging (table S2, 3). Relative protein abundance in process replicates was highly correlated with an average R^2 of 0.75 (fig. S3A), indicating high technical reproducibility in the proteomic workflow. A high correlation of human protein abundance was also observed in biological, serially-passaged replicates of the same PDX (fig. S3), suggesting that the proteomes of the engrafted tumors remained stable. Relative changes in the mouse proteome were also highly correlated within biological replicates of individual PDX tumor models (fig. S3A). Median correlation of species-unique proteins and peptides within biological replicates of the same PDX (intra-PDX) were more than 10 fold higher than across PDX lines (inter-PDX). Median intra-PDX R^2 values were 0.53 and 0.10 for human and mouse proteins, respectively (fig. S3B). These data suggested that tumors are intrinsically capable of differentially educating stromal proteomes, which we further investigated here.

Individual breast cancer PDXs induce unique and biologically coordinated stromal protein signatures

One-way analysis of variance (ANOVA) was performed to identify specific proteins in the tumor and the PDX stroma that were differentially regulated across the subcutaneous PDX models. 82% and 35% of human and mouse proteins representing the tumor and stroma, respectively, exhibited significant changes in protein abundance that were consistent across biological replicates of a PDX. Median intra-PDX R^2 values were 0.63 and 0.44 for human and mouse proteins after ANOVA filtering, respectively (Fig. 1B). Biological replicates of each PDX co-clustered based on both human and mouse protein abundance (Fig. 1C). Hierarchical clustering segregated mouse stromal proteins into 6 clusters (Fig. 1C). All differentially abundant stromal proteins in each of the 6 clusters are listed in table S4. Despite the small sample size, all PDXs clustered adjacently by subtype on the basis of human tumor protein abundance, whereas nearly all PDXs clustered by subtype on the basis of mouse stromal protein abundance (Fig. 1C). PCA was performed to orthogonally assess whether patterns in the proteomes, especially those of the stroma, were consistent within biological replicates of individual PDX tumors. Again, the biological replicates of each PDX were closely grouped on the basis of both human tumor- and mouse stroma-unique protein abundance (Fig. 1D). Principal component analysis (PCA) of ANOVA-filtered proteins showed no clustering by gender for greater than 80% of the data's cumulative variance,

suggesting that the use of male mice for basal tumors was not a source of significant variation in the data (fig. S4). Together, these data demonstrated that patient breast tumors consistently induce unique stromal proteomic changes when placed in a naïve microenvironment. In addition, these results revealed that the magnitude of tumor-specific stromal education is substantial, encompassing over one-third of the detected mouse proteins. These data also established that tumor education of the stroma is prevalent among breast tumors, given that it occurred in all seven breast tumors analyzed. Therefore, stromal regulation during colonization of a naïve microenvironment is predominantly an intrinsic property of a patient tumor that persists across biological replicates and passages, with regulatory components that are stably ingrained.

To examine whether the differentially regulated PDX stromal proteins share a common biological role, unbiased annotation by the Molecular Signatures Database (22) was performed using the Hallmark, KEGG, and Reactome gene sets. Each of the 6 stromal clusters had significant enrichment of specific and unique molecular signatures (Fig. 2A). Specific components of the extracellular matrix (ECM), including laminins, collagens, and elastic fibers, as well as actin-cytoskeleton binding proteins, were each enriched in single clusters, including Clusters I, II, and III (Fig. 2A). Cluster II associated with cancer-associated fibroblast (CAF) activity since it contains the CAF-secreted factors POSTN, FBN1, COL1A1, COL12A1, and BGN (23). Cluster IV corresponded to members of the complement system which are known to facilitate tumor survival (24, 25). Clusters V and VI were annotated to multiple molecular signatures with high statistical significance, but lacked a clear biological organizing principle. This is a limitation of even large databases of curated gene sets as no analysis can fully and precisely encompass all biological processes. The annotation of each cluster, especially Clusters V and VI, was therefore refined using literature analysis (Fig. 2B). Cluster V contained numerous markers of myeloid-derived suppressor cells (7, 26), potent immunosuppressors that can be recruited to the invasive edge of tumors by tumor-derived chemoattractants (27) to promote tumor survival (7). Cluster V also contained three glycolytic enzymes, PKM, GPI, and LDHA, as well as transketolase (TKT) in the pentose phosphate shunt, consistent with tumor regulation of extracellular metabolites (28–31). Cluster VI contained several canonical markers that are signatures of M2 type tumor-associated macrophages (TAMs) (32–34), in addition to ER processing and stress as suggested by the MSigDB analysis. M2 type TAMs are also recruited to the microenvironment and secrete proinflammatory factors that support tumor survival (27, 35).

To assess whether stromal regulation by individual PDXs occurs by the co-regulation of multiple proteins with shared biological roles, we quantified all stromal proteins within each molecular signature across biological replicates for each PDX. The stromal protein changes within each signature were highly coordinated (Fig. 2B). Tumor-specific education of stromal proteomes was also unique and individualized, as demonstrated by highly significant differences in each protein signature across PDX tumors (Fig. 2C). All seven PDXs coordinately regulated at least one stromal cluster with statistical significance. Taken together, breast PDXs individually educate their microenvironment with highly coordinated stromal proteomic changes that encompass well-defined pro-tumorigenic signatures.

PDX stromal protein signatures are common in patient breast tumors and separate patients by subtype and stage

Despite PDX's burgeoning position as the leading experimental model to approximate the biology and genetics of human tumors (9), the use of PDXs to investigate the biology of tumor niche formation by unambiguous delineation of tumor and stromal-derived murine proteins remains undeveloped. Limited adoption is partly because PDXs lack the original stromal complement of the primary tumor and NSG mice that PDXs are engrafted onto do not produce cells of the lymphoid lineage. Notably, PDX models do contain murine myeloid-derived cells that play important roles in tumor growth, chemotaxis, and invasion in the microenvironment (36).

We therefore sought to validate that the stromal protein signatures seen in the subcutaneous PDX models were consistent with regulation of the microenvironmental by primary tumors. The breast TCGA transcriptomic (37) and proteomic datasets (38) were examined to determine if primary patient tumors similarly co-regulate proteins within the stromal clusters at the mRNA and protein levels. Since the TCGA data cannot be used to differentiate tumor from stroma, we focused on Clusters IV-VI that represent stromal proteins that are not often found in the tumor itself and represent the stromal microenvironment. Only 3 human homologs of the 24 mouse proteins in Clusters IV-VI were consistently detected in all seven PDXs, and only 6 human homologs were detected in any PDX, confirming that these proteins are primarily stromal-derived. Since the complement system is almost entirely regulated at the protein level (39), only protein results were shown for Cluster IV. We found highly correlated gene products within stromal clusters IV-VI at both the mRNA and protein level for the 1,095 and 77 publicly available TCGA breast cancer patient samples, respectively (Fig. 3A). Monte Carlo simulation was used to determine the statistical significance of co-regulation within each stromal cluster. All stromal clusters were significantly correlated in patient tumors at both the mRNA and protein levels (Fig. 3B). Examples of individual primary breast tumors co-regulating mRNA or protein from multiple genes within stromal clusters IV-VI, similar to the PDX stromal protein clusters in Fig. 2B, are also shown (fig. S5).

The biological pathway annotations assigned to the stromal protein signatures are well known to play a role in breast cancer tumorigenicity and colonization. For example, a stromal-based serpin-enriched cluster correlates with better prognosis (3). In addition, TAM markers, as determined through mRNA analysis of the stroma from 53 primary tumors, associate with poor outcome (16). We therefore assessed whether stromal protein clusters IV-VI correlated with three clinically-relevant parameters in the TCGA set: overall survival, intrinsic subtype, and tumor stage. There was no association between stromal clusters and patient survival. Because stromal proteomic signatures separated all but one PDX by subtype in the seven PDXs analyzed (Fig. 1C), we next determined whether stromal clusters IV-VI stratified patient tumors by subtype in the much larger TCGA proteomic dataset. Strikingly, all but one of the 24 genes in the three clusters had significantly increased mRNA levels in basal versus non-basal TCGA tumors (Fig. 3C). This result was consistent with the hypothesis that basal, predominantly triple-negative, tumors are more immunogenic than other subtypes and recruit myeloid-derived cells (40, 41).

All members of stromal clusters V and VI had higher protein quantities in Stage I tumors (Fig. 3D), but not Cluster IV. Stromal proteomic clusters can therefore differentiate patient tumor populations, perhaps because Stage I tumors, like PDXs, are in the initial stages of setting up the microenvironmental niche. The TCGA proteomic data contains only a single Stage IV metastatic tumor, which was excluded from the analysis.

Together, the stromal proteomic regulation by patient tumors observed in PDX models was consistent with stromal changes induced by primary tumors and was associated with clinically-relevant parameters including intrinsic subtype and tumor stage. No correlation by stage or subtype was seen at the mRNA level, suggesting that stage and subtype regulation of the stromal clusters occur at the protein level.

Integrating tumor and stromal proteome analysis reveals that cancer hallmarks are highly correlated with stromal regulation

It has been hypothesized that oncogenic signaling observed in cancer cells is selected for pleiotropic pro-tumorigenic roles that include regulation of the microenvironment, not simply promotion of proliferation and transformation (42). These data were consistent with this hypothesis since tumor-intrinsic factors are responsible for widespread, individualized changes in the stromal proteomes (Figs. 1 and 2). This hypothesis also suggests that the tumor proteome encodes the educational programs relevant to stromal regulation, which is in agreement with studies indicating that epigenetic factors are primarily responsible for colonization (4, 5). We therefore performed an integrated analysis of the tumor and stromal proteomes to discover novel candidates in the tumor responsible for stromal education in subcutaneous PDX models.

To assess the relationship between the tumor and stromal proteomes, the protein abundance of each mouse gene in the six stromal clusters was correlated to protein abundance of each human gene across all 21 samples in the dataset. To broadly determine the biological roles of the proteins in the tumor that were associated with each stromal protein signature, human proteins were ranked by their maximum absolute correlation to each stromal cluster and analyzed by Gene Set Enrichment Analysis (GSEA), utilizing only the small, highly curated MSigDB 'Hallmarks' gene set (43). The MSigDB Hallmarks gene set contains 50 refined sets that broadly describes many important biological processes with minimal redundancy (43). 23 of the 50 hallmarks were enriched in at least one stromal cluster using a stringent cutoff of <10% false discovery rate (FDR) (Fig. 4A). Nearly all of the gene sets identified in the MSigDB hallmarks are implicated in Hanahan and Weinberg's hallmarks of cancer (42) (Fig. 4A), consistent with the hypothesis that the hallmarks of cancer are pleiotropically responsible for educating the microenvironment in addition to tumor intrinsic biology. The MSigDB hallmarks gene sets found in the tumor were highly associated with specific stromal clusters associated with microenvironmental regulation including EMT, apical junctions, hypoxia signaling, as well as interferon alpha and gamma responses.

Local invasion and metastatic colonization are complex, multi-step processes and the proteins involved remains incomplete. One unanswered question is what fraction of a tumor's proteome is dedicated to regulatory crosstalk with the microenvironment. To provide one estimate using integrated PDX tumor-stroma analysis, we determined how many

stromal proteins were significantly correlated with each tumor protein. These results (Fig. 4B) demonstrated that a large portion of the tumor proteome was co-regulated with the stroma. 19 proteins in the tumor were significantly correlated with at least 40% of proteins in one or more of the 6 stromal clusters, and 44 were significantly correlated with at least 30% of stromal proteins in a single cluster. Among the proteins in the tumor that were most correlated, both positively and negatively, with the stromal proteome (Table 1) are many known to regulate the tumor-stroma interface. PEBP1, also known as RKIP, a metastasis suppressor that decreases infiltration of tumors by myeloid cells (44), was significantly negatively correlated with 7 proteins in stromal Cluster V which contains MDSC markers. Nardilysin (NRD1), a metalloprotease that can remodel the ECM, was significantly negatively correlated with 9 proteins in the ECM-related Cluster II. Notably, the list of tumor proteins highly correlated with stromal protein clusters included many with unknown roles in the tumor microenvironment. The full list of tumor-stroma correlation values (Data file S1) serves as a set of candidates responsible for stromal remodeling.

Discussion

Tumors reprogram their microenvironment to form a tumor-associated niche (6), but our current conceptualization of tumor-stroma remains static and superficial, with limited understanding of the dynamic, progressive process of stromal education by tumors. Many of the key processes involved in remodeling of a naïve microenvironment are depicted using cell lines and xenografts, but it is unclear which of these are relevant to, and used by, patient tumors. In addition, molecular knowledge of microenvironmental signaling interactions remains far from complete. The phenotypic plasticity of cancer cells coupled with the potential selection of myriad possible cellular subclones of tumors intimates that development of the microenvironmental niche may be highly mutable, dynamic, and context-dependent. We used species-specific proteomic analysis of serially passaged subcutaneous breast cancer PDXs to address three major questions. First, can unbiased cross-species proteomic profiling provide sufficient depth and repeatability to improve our conceptualization of the tumor-stroma interface in breast cancer? Second, how do patient breast tumors educate a naive microenvironment during colonization and what is the molecular persistence and heterogeneity of this process across tumors? Third, do subcutaneous PDXs provide molecular insights into stromal education that are consistent with what occurs in primary patient breast tumors?

Holistically, we found that tumor-specific education of the stroma was highly prevalent, individualized, and molecularly coordinated in breast cancer. The reproducibilityof a tumor's education of stromal proteins is consistent with the notion that patient breast tumors are locked into a predetermined set of instructions due to metabolic, energetic, or other deficiencies that are rigid and specific (45). This is likely selected for due to the numerous obstacles that normal cells have to overcome to become tumorigenic. Alternatively, tumors harbor cancer stem cells (46), which may be enriched in PDXs during passage and leads to the limited stromal proteomic heterogeneity within a single tumor line observed across multiple passages for each PDX. The lack of observed proteomic heterogeneity within a single PDX tumor line suggests that renewed emphasis on drugging the stroma as an antitumor strategy may be highly effective.

These results were enabled by the high repeatability, multiplexing, and depth of coverage of the cross-species quantitative proteomics method employed in this study, and serve as an initial characterization of proteomic stromal education by patient tumors. A previously published study used species-specific peptide sequences in an effort to identify proteins exported to the ECM (47); our study here was a large-scale analysis of the tumor and microenvironment in PDXs using species-specific proteomics and yielded a 10-fold increase in proteomic coverage of the microenvironment over the previous report (47), including detection and quantitation of 4,784 human and 1,721 mouse genes at the proteome-level with species- and gene-unique peptides. 51% and 52% of peptides and genes, respectively, were filtered out due to lack of species specificity, and an additional 3% of protein identifications were filtered out due lack of gene specificity. Our study focused on proteomic variability of PDXs across biological replicates and passages, as occurs during routine PDX propagation, which was consistent with a median R^2 value of 0.47. This robust cross-species proteomic platform has the potential to characterize perturbations in PDXs and the tumor-stroma interface due to genomic heterogeneity, across subtypes, and in response to drug treatment to uncover drug resistance mechanisms. PDXs complement syngeneic mouse models since PDXs are derived from patient tumors and also enable unambiguous molecular characterization of the tumor versus stroma. Future studies using PDXs of other cancer origins are needed to determine whether heterogeneity of the proteomes of tumor microenvironments is specific to breast cancer.

Delineating the molecular crosstalk between tumor and stromal cells in the tumor microenvironment remains an important challenge in the effort to understand tumor biology (48–50), especially due to its correlation with anticancer response (51–55) and metastasis (6, 56). Our study characterizes proteins in the stroma that are susceptible to differential induction by the tumor (table S4) and identifies tumor-specific proteins that are highly correlated with stromal protein signatures (table S5). This is an important step toward identification of microenvironmental targets that inhibit tumor growth, or even return tumor cells to normal, which has been described as the next frontier in anticancer agents (57). These results were consistent with the prevailing view that tumors colonize their microenvironment through regulation of the ECM and recruitment of cells and factors, given that we observed changes in the ECM (Clusters II-III) and regulation of the complement system (Cluster IV), MDSCs (Cluster V) and macrophages (Cluster VI). Many of the differentially quantified stromal proteins identified correlate with breast cancer prognosis and treatment. For example, nine stromal genes with differential protein quantities, including *S100A9* and *LCP1*, have mRNA quantities in the stroma that corresponds to patient outcome (16). Considering the high co-regulation of stromal protein signatures in both PDX and patient tumors in the TCGA, these observations demonstrated that PDX models largely recapitulate how primary tumors interact with the microenvironment and are good models to study the tumor-stroma interface despite their lack of a lymphoid system, which also limits the utility of PDX models to study immunotherapy and other T-cell driven biology. Notably, this regulation occurs in PDXs grown subcutaneously, in a non-orthotopic location. Because many of the identified biological pathways are druggable, these results suggest that proteomics can guide therapeutic studies targeting the stroma. Furthermore, proteomic characterization of PDX models may provide insight into how seemingly dormant

micrometastases are prompted to grow by tumor removal, an important consideration in treatment decisions.

Tumors have been shown to harness the stroma for their benefit, but the extent of stromal education by patient tumors at the level of the proteome has not been reported. Our analysis suggested that stromal education is far more pervasive, coordinated, and individualized at the protein level than previously anticipated. Despite each tumor educating a unique repertoire of stromal proteins, the shared regulation of stromal protein signatures suggests that they play a critical role in tumorigenesis. The high consistency and molecular rigidity of stromal education by patient tumors suggests that elucidating how individual tumors alter their stromal environment will yield significant insights into this important and complex aspect of tumor biology.

Materials and Methods

Materials

Materials were purchased from Sigma Aldrich except for dithiothreitol (DTT), trifluoroacetic acid (TFA), and NaCl (ThermoFisher), and Tris and Glycine (BioRad). LC-MS solvents were purchased pre-mixed from Honeywell. High pH reversed phase fractionation columns and TMT10 reagents were from ThermoFisher.

Generation and preparation of xenograft tumor samples for proteomic analysis

The source and generation of all PDXs in this study were previously reported(15). All human tissues for these experiments were processed in compliance with NIH regulations and institutional guidelines approved by the Institutional Review Board at Washington University. All animal procedures were reviewed and approved by the Institutional Animal Care and Use Committee at Washington University in St. Louis, MO. PDX tumors from established basal (WHIM2, 4, 14), luminal (WHIM16, 20), claudin-low (WHIM12) and HER2E (WHIM11) breast cancer subtypes were raised subcutaneously in 8-week-old NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ mice (Jackson Labs, Bar Harbor, ME) as previously described (15, 58). Tumors from each animal were harvested by surgical excision at ~1.5 cm^3, rapidly divided into four pieces, and snap-frozen by immersion in a liquid nitrogen bath immediately post-excision.

Tissue Lysate Preparation

Samples were cryopulverized into powder using a Covaris CP02 CryoPrep system and solubilized in lysis buffer with a Covaris S220X sonicator (Peak incident power: 100 Watt, 500 cycles per burst, 10% duty factor, 4°C, 4 minutes) as in (58). The lysis buffer (50 mM HEPES, pH 7.5) contained the following: 150 mM NaCl, 0.5% Triton X-100, 1 mM EDTA, 1 mM EGTA, 10 mM NaF, 2.5 mM NaVO$_4$, 1X Protease Inhibitor Cocktail (Roche), and phosphatase inhibitor cocktails 2 and 3 (Sigma). Lysates were centrifuged at $22,000 \times g$ for 10 minutes to pellet any debris and then filtered through a 0.45 uM filter (Millipore Ultrafree-MC-HV) to further remove insoluble protein. Protein concentrations were determined (Advanced protein assay, Cytoskeleton) to prepare aliquots with concentrations of ~ 5 mg/mL) storage at -80°C and subsequent digestion.

Sample Digestion and TMT labeling

The PDX lysates (0.8–2 mg) were thawed on ice and the protein was precipitated using2-D clean-up kit (GE Healthcare) and resolubilized in 8M urea with 100 mM TEAB (pH 8.5). The samples were reduced with 5 mM TCEP for 30 min at 25°C, alkylated with 40 mM iodoacetamide for 30 min at 25°C in the dark, and quenched with 20 mM dithiothreitol for 15 min at 25°C. Samples were diluted to 2M urea with 100 mM TEAB prior to addition of 5μg trypsin (Fluka, analytical grade). Samples were incubated at 37°C for 16 hr. Another 5 μg trypsin was added and incubated at 37°C for 4 hr. Samples were filtered through a 30 kD filter (Millipore Ultracell YM-30) and labeled with TMT10 reagents per manufacturer's instructions. Samples were desalted with C4 and graphitized carbon tips (Glygen), loading using 1% ACN, 1% formic acid and elution with 60% ACN, 1% FA, and combined prior to high pH reversed phase fractionation.

High pH Reversed phase fractionation

TMT10 labeled peptides were separated by high pH reversed- phase spin columns (ThermoFisher Scientific, catalog #84868). After conditioning columns twice with 300 μL ACN and three times with 300 μL water, samples were loaded into the column. Columns were centrifuged at $3,000 \times g$ for 1 minute, followed by a wash with 300 μL 0.1% triethylamine. 300 μL of each elution buffer (5%, 7.5%, 10%, 12.5%, 15%, 17.5%, 20%, and 50% ACN in 0.1% TFA) were added in sequence to the column followed by centrifugation at $3,000 \times g$ for 1 min to collect each eluted fraction. Elutions from 5% and 7.5% ACN were pooled following elution. Samples were evaporated to dryness and resuspended in 0.1% FA for LC-MS analysis.

Pooling TMT-labeled samples

An initial set of 6 PDX tumors, 1 biological replicate each, was pooled together to make Pool A (Supp. Fig S6A). The sample set was first used to evaluate assay repeatability (Supp. Fig S6B, TMT10 plex #1) that is shown in Supplemental Fig. S3 as well as determine relative protein abundance in these PDXs. A set of 15 additional PDXs were pooled into Pool B (Supp. Fig S6A) that was used only to determine relative protein abundance in each PDX in TMT10 plex #2 and plex #3 (Supp. Fig S6B). Pool A was included as a TMT10 channel in all analyses to stitch all three TMT10 plexes together. The Pearson's correlation coefficients of all protein abundance ratios between Pool A and B, including both human and mouse proteins, were very high, averaging 0.994 across analyses (Supp. Fig. S7A). In addition, the raw intensity of individual spectral matches in each pool was very linear (average slope = 0.981, Supp. Fig. S7B) and similar (average linear regression r^2 = 0.963, Supp. Fig. S7B). We also analyzed 1 sample (WHIM 12.2, Supp. Fig. S6B) in both TMT10 plex #2 and #3 to validate that all relative protein abundance across the two TMT10 plexes with Pool B (Supp. Fig. S6B) resulted in similar protein profiles. The TMT10 labeling schematic for all samples analyzed is shown in table S2.

LC-MS data acquisition of TMT10-labeled samples for quantitative protein profiling on Thermo Elite LC-MS

TMT10 plex #1 was analyzed on a Thermo Elite LC-MS. LC was performed using two 75μm × 15cm ChromXP C18 columns in tandem with a 4 hr LC gradient from 5% to 30% ACN over 180 min and 30–45% over 25 min. MS analysis was performed on an Orbitrap Elite with a TOP15 method. MS1 scan range was 380–16000 *m/z* at a resolution of 120k for 10 ms. MS2 spectra were generated by HCD fragmentation and acquired with 30k resolution and scans starting at 110 *m/z*.

LC-MS data acquisition of TMT10-labeled samples for quantitative protein profiling on Thermo Q Exactive LC-MS

For each high pH-RP fraction, two μL of sample was loaded onto a 75 μm i.d. × 25 cm Acclaim PepMap 100 RP column (Thermo-Fisher Scientific). Peptide separations were started with 95% mobile phase A (0.1% FA) for 5 min and increased to 25% mobile phase B (100% ACN, 0.1% FA) over 95 min, followed by a 10-min gradient to 40% B, a 5-min gradient to 90% B and wash at 90% B for 5 min, with a flow rate of 300 nL/min. Full-scan mass spectra were acquired by the Orbitrap mass analyzer in the mass-to-charge ratio (m/z) of 375 to 1400 with a mass resolving power of 70,000. Fifteen data-dependent high-energy collisional dissociations (HCD) were performed with a mass resolving power set to 35,000, an *m/z* range starting from 100 up to a maximum determined by the instrument, an isolation width of 0.7 *m/z*, and normalized collision energy setting of 32. The maximum injection time was 50 ms for parent-ion analysis and 105 ms for product-ion analysis. Target ions already selected for MS/MS were dynamically excluded for 30 sec. An automatic gain control (AGC) target value of 3e6 ions was used for full MS scans and 1e5 ions for MS/MS scans. Peptide ions with charge states of one, or greater than six, were excluded from MS/MS interrogation.

Peptide identification

Protein and peptide identification as well as relative quantitation were performed with Proteome Discoverer software (version 1.4.0.288; Thermo Fisher Scientific) using Mascot (v. 2.4.1) as the search engine. MS/MS spectra were searched against a concatenated NCBI Reference Sequence (RefSeq version July, 2013) database of human (36,380 entries) and mouse (24,821 entries), and cRAP(59) (version 1.0 Jan. 1, 2012, 116 entries). The search parameters included digestion with trypsin/P (2 or 4 missed trypsin cleavages with Elite versus QE data, respectively), static modification of cysteine by carbamidomethylation, TMT10 labeling of lysine and peptide N-termini, and variable modifications included methione sulfoxide, deamidation of asparagine, deamidation of glutamine, acetylation of protein N-terminus, TMT 10-plex derivatization of peptide N-termini, and S-carbamoylmethylcysteine cyclization of N-terminal cysteine. Mass tolerances were 10 ppm mass accuracy on precursors and 0.02 Da on fragment ions. Peptide FDR was calculated by target-decoy searching against a reversed dataset and peptides were filtered at 1% FDR. A peak integration tolerance of 20 ppm was used for extracting TMT10 reporter ion intensities.

Quantitation, data processing and normalization

The resulting peptide sequences were mapped onto both human and mouse genes using the PGx software package (version 1.0, (60)) to identify species- and gene-unique peptides (fig. S2). When a peptide sequence was observed in both human and mouse sequences, it was defined as 'species-shared'. Otherwise peptide sequences were defined as 'species-unique'. The species-specific peptides were further characterized as gene-unique for their respective species. When a peptide was matched to only one gene that was represented by only one database sequence entry or multiple database sequence entries, the peptide was defined as 'gene-unique'. Otherwise, peptides were considered as 'gene-shared'. To quantify proteins at the gene level, only species- and gene-unique peptides were used.

Relative quantification of protein abundance was performed using the reporter ion signals from the TMT-10 multiplex experiments (fig. S2). All peptide spectrum matches (PSMs) with an FDR 1% were 'rolled up' to the gene level by summing the peak heights for each gene-and species-specific PSM, and quantified proteins were reported by their respective gene symbols. Genes were excluded from further analysis if their summed peak heights from the reporter ions were zero or had missing values in any TMT channel. For each gene, summed peak heights were then normalized to the internal reference pool included in each TMT10 plex. Mouse genes were further filtered by removing plasma (61) and abundant erythrocyte proteins identified by proteomics (62) prior to downstream analysis. Relative intensities were $\log_2$ transformed, subtracted with $\log_2$ scale median and then divided by $\log_2$ scale standard deviation sample-wise.

Statistical Analysis

The coefficients of determination (R^2) were represented in boxplots (Fig. 1B). In each boxplot, the lower and higher whiskers represent first and third quartiles respectively, the horizontal line indicates the position of median, and dots outside of whiskers are outliers. GSEA (version 2.2.1) analysis was performed using default parameters and the software package generated statistical analysis. Hierarchical clustering (heatmap.2 in R) was performed using Spearman's correlation coefficients and Ward's minimum variance method. Color gradient of heatmap was set as red being the highest and blue being the lowest protein abundance. The human and mouse data were statistically analyzed using one-way ANOVA by considering PDX samples as multiple groups, respectively. Differences were regarded as significant if the adjusted p-value (Benjamini-Hochberg method) was less than 0.05. For PCA, the differentially regulated human and mouse proteins were used to calculate principal components, and the data points projected onto first three principal components were visualized in three-dimensional space.

To explore whether the use of male and female mice in the basal subtype PDXs affects our species-specific proteomics results, we performed principal component analysis across all 21 tumors using the proteins with significantly altered abundance that were identified by ANOVA. Principal components with up to 80% variance were kept for visualization (fig. S4).

To identify interaction patterns between tumor (human) and stromal (mouse) clusters, correlation analysis was performed based on significantly differentially quantified human and mouse proteins (Fig. 4). Specifically, mouse and human proteins after applying ANOVA were utilized to calculate Spearman's correlation coefficients between species, and p-values from correlation test were further adjusted by Benjamini-Hochberg method. For each human protein, the correlation with maximum absolute value in a stromal cluster was used to represent the tumor and stromal interaction. The output dataset served as an input for GSEA analysis (Fig. 4A). The numbers of significantly correlated human-mouse protein pairs were also recorded for each human protein split by mouse clusters when set threshold for adjusted p value as 0.05 (Fig. 4B). Blue and red lines indicate the number of negative and positive significant correlation, respectively.

Data from breast cancer TCGA (Fig. 3) were used to determine the extent to which the clusters in the mouse stroma are regulated in human patients. RNA-Seq data from 1,095 primary breast tumors from the TCGA (37) (TCGA RNA-SeqV2 pipeline) were analyzed for coordination at the transcript level. Global iTRAQ proteomic data from 105 patients in the CPTAC analysis of the TCGA(38) were analyzed for coordinate protein level regulation. For the gene sets found by hierarchical clustering of the mouse stromal data (Fig. 2B), Spearman's correlation coefficients were used to determine the strength of coordination in each data set. P-values (Fig. 3B) for the correlation matrices were determined using Monte Carlo simulation by sampling 10,000 randomized sets of equal size to the test clusters, and ranking the sums of the Spearman's correlation coefficients. For Fig. 3, C and D, TCGA proteomics data were parsed by TCGA-assigned PAM50 subtypes and pathologic stage, respectively. Statistical significance was assessed using Student's t-tests. Data visualization was implemented using R (Version 3.1.2) unless indicated otherwise.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References and Notes

1. Conklin MW, Keely PJ. Why the stroma matters in breast cancer. Cell Adhesion & Migration. 2014; 6:249–260.

2. Farmer P, Bonnefoi H, Anderle P, Cameron D, Wirapati P, Becette V, André S, Piccart M, Campone M, Brain E, MacGrogan G, Petit T, Jassem J, Bibeau F, Blot E, Bogaerts J, Aguet M, Bergh J, Iggo R, Delorenzi M. A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer. Nat Med. 2009; 15:68–74. [PubMed: 19122658]

3. Bergamaschi A, Tagliabue E, Sørlie T, Naume B, Triulzi T, Orlandi R, Russnes HG, Nesland JM, Tammi R, Auvinen P, Kosma VM, Menard S, Børresen-Dale AL. Extracellular matrix signature identifies breast cancer subgroups with different clinical outcome. J Pathol. 2007; 214:357–367.

4. Vanharanta S, Massagué J. Origins of Metastatic Traits. Cancer Cell. 2013; 24:410–421. [PubMed: 24135279]

5. Naxerova K, Jain RK. Using tumour phylogenetics to identify the roots of metastasis in humans. Nat Rev Clin Oncol. 2015; 12:258–272. [PubMed: 25601447]

6. Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. Nat Med. 2013; 19:1423–1437. [PubMed: 24202395]

7. Gabrilovich DI, Ostrand-Rosenberg S, Bronte V. Coordinated regulation of myeloid cells by tumours. Nat Rev Immunol. 2012; 12:253–268. [PubMed: 22437938]

8. Malanchi I, Santamaria-Martínez A, Susanto E, Peng H, Lehr HA, Delaloye JF, Huelsken J. Interactions between cancer stem cells and their niche govern metastatic colonization. Nature. 2011; 481:85–89. [PubMed: 22158103]

9. Hidalgo M, Amant F, Biankin AV, Budinská E, Byrne AT, Caldas C, Clarke RB, de Jong S, Jonkers J, Mælandsmo GM, Roman-Roman S, Seoane J, Trusolino L, Villanueva A, FTE Consortium. Patient-Derived Xenograft Models: An Emerging Platform for Translational Cancer Research. Cancer Discov. 2014; 4:998–1013. [PubMed: 25185190]

10. Choi SYC, Lin D, Gout PW, Collins CC, Xu Y, Wang Y. Lessons from patient-derived xenografts for better in vitro modeling of human cancer. Adv Drug Deliv Rev. 2014; 79–80:222–237.

11. Whittle JR, Lewis MT, Lindeman GJ, Visvader JE. Patient-derived xenograft models of breast cancer and their predictive power. Breast Cancer Res. 2015; 17:1. [PubMed: 25567532]

12. Garrido-Laguna I, Uson M, Rajeshkumar NV, Tan AC, de Oliveira E, Karikari C, Villaroel MC, Salomon A, Taylor G, Sharma R, Hruban RH, Maitra A, Laheru D, Rubio-Viqueira B, Jimeno A, Hidalgo M. Tumor engraftment in nude mice and enrichment in stroma- related gene pathways predict poor survival and resistance to gemcitabine in patients with pancreatic cancer. Clin Cancer Res. 2011; 17:5793–5800. [PubMed: 21742805]

13. Martinez-Garcia R, Juan D, Rausell A, Muñoz M, Baños N, Menéndez C, Lopez-Casas PP, Rico D, Valencia A, Hidalgo M. Transcriptional dissection of pancreatic tumors engrafted in mice. Genome Med. 2014; 6:27. [PubMed: 24739241]

14. Ding L, Ellis MJ, Li S, Larson DE, Chen K, Wallis JW, Harris CC, McLellan MD, Fulton RS, Fulton LL, Abbott RM, Hoog J, Dooling DJ, Koboldt DC, Schmidt H, Kalicki J, Zhang Q, Chen L, Lin L, Wendl MC, McMichael JF, Magrini VJ, Cook L, McGrath SD, Vickery TL, Appelbaum E, DeSchryver K, Davies S, Guintoli T, Lin L, Crowder R, Tao Y, Snider JE, Smith SM, Dukes AF, Sanderson GE, Pohl CS, Delehaunty KD, Fronick CC, Pape KA, Reed JS, Robinson JS, Hodges JS, Schierding W, Dees ND, Shen D, Locke DP, Wiechert ME, Eldred JM, Peck JB, Oberkfell BJ, Lolofie JT, Du F, Hawkins AE, O'Laughlin MD, Bernard KE, Cunningham M, Elliott G, Mason MD, Thompson DM, Ivanovich JL, Goodfellow PJ, Perou CM, Weinstock GM, Aft R, Watson M, Ley TJ, Wilson RK, Mardis ER. Genome remodelling in a basal-like breast cancer metastasis and xenograft. Nature. 2010; 464:999–1005. [PubMed: 20393555]

15. Li S, Shen D, Shao J, Crowder R, Liu W, Prat A, He X, Liu S, Hoog J, Lu C, Ding L, Griffith OL, Miller C, Larson D, Fulton RS, Harrison M, Mooney T, McMichael JF, Luo J, Tao Y, Goncalves R, Schlosberg C, Hiken JF, Saied L, Sanchez C, Giuntoli T, Bumb C, Cooper C, Kitchens RT, Lin A, Phommaly C, Davies SR, Zhang J, Kavuri MS, McEachern D, Dong YY, Ma C, Pluard T, Naughton M, Bose R, Suresh R, McDowell R, Michel L, Aft R, Gillanders W, DeSchryver K, Wilson RK, Wang S, Mills GB, Gonzalez-Angulo A, Edwards JR, Maher C, Perou CM, Mardis ER, Ellis MJ. Endocrine-therapy-resistant ESR1 variants revealed by genomic characterization of breast-cancer-derived xenografts. Cell Reports. 2013; 4:1116–1130. [PubMed: 24055055]

16. Finak G, Bertos N, Pepin F, Sadekova S, Souleimanova M, Zhao H, Chen H, Omeroglu G, Meterissian S, Omeroglu A, Hallett M, Park M. Stromal gene expression predicts clinical outcome in breast cancer. Nat Med. 2008; 14:518–527. [PubMed: 18438415]

17. Calon A, Lonardo E, Berenguer-Llergo A, Espinet E, Hernando-Momblona X, Iglesias M, Sevillano M, Palomo-Ponce S, Tauriello DVF, Byrom D, Cortina C, Morral C, Barceló C, Tosi S, Riera A, Attolini CSO, Rossell D, Sancho E, Batlle E. Stromal gene expression defines poor-prognosis subtypes in colorectal cancer. Nat Genet. 2015; 47:320–329. [PubMed: 25706628]

18. Naba A, Clauser KR, Hoersch S, Liu H, Carr SA, Hynes RO. The matrisome: in silico definition and in vivo characterization by proteomics of normal and tumor extracellular matrices. Molecular & Cellular Proteomics. 2012; 11 M111.014647.

19. Hill RC, Calle EA, Dzieciatkowska M, Niklason LE, Hansen KC. Quantification of extracellular matrix proteins from a rat lung scaffold to provide a molecular readout for tissue engineering. Molecular & Cellular Proteomics. 2015; 14:961–973. [PubMed: 25660013]

20. Peng S, Creighton CJ, Zhang Y, Sen B, Mazumdar T, Myers JN, Woolfson A, Lorenzi MV, Bell D, Williams MD, Johnson FM. Tumor grafts derived from patients with head and neck squamous carcinoma authentically maintain the molecular and histologic characteristics of human cancers. Journal of Translational Medicine 2013 11:1. 2013; 11:1.

21. Moro M, Bertolini G, Tortoreto M, Pastorino U, Sozzi G, Roz L, Moro M, Bertolini G, Tortoreto M, Pastorino U, Sozzi G, Roz L. Patient-Derived Xenografts of Non Small Cell Lung Cancer: Resurgence of an Old Model for Investigation of Modern Concepts of Tailored Therapy and Cancer Stem Cells. BioMed Research International 2012. 2012:1–11.

22. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. Bioinformatics. 2011; 27:1739–1740. [PubMed: 21546393]

23. Torres S, Bartolomé RA, Mendes M, Barderas R, Fernandez-Aceñero MJ, Peláez-García A, Peña C, Lopez-Lucendo M, Villar-Vázquez R, de Herreros AG, Bonilla F, Casal JI. Proteome Profiling of Cancer-Associated Fibroblasts Identifies Novel Proinflammatory Signatures and Prognostic Markers for Colorectal Cancer. Clin Cancer Res. 2013; 19:6006–6019. [PubMed: 24025712]

24. Rutkowski MJ, Sughrue ME, Kane AJ, Mills SA, Parsa AT. Cancer and the complement cascade. Molecular Cancer Research. 2010; 8:1453–1465. [PubMed: 20870736]

25. Kikuchi Y, Kashima TG, Nishiyama T, Shimazu K, Morishita Y, Shimazaki M, Kii I, Horie H, Nagai H, Kudo A, Fukayama M. Periostin is expressed in pericryptal fibroblasts and cancer-associated fibroblasts in the colon. Journal of Histochemistry and Cytochemistry. 2008; 56:753–764. [PubMed: 18443362]

26. Yu J, Du W, Yan F, WANG Y, Li H, Cao S, Yu W, Shen C, Liu J, Ren X. Myeloid-Derived Suppressor Cells Suppress Antitumor Immune Responses through IDO Expression and Correlate with Lymph Node Metastasis in Patients with Breast Cancer. J Immunol. 2013; 190:3783–3797. [PubMed: 23440412]

27. Condeelis J, Pollard JW. Macrophages: obligate partners for tumor cell migration, invasion, and metastasis. Cell. 2006; 124:263–266. [PubMed: 16439202]

28. Martinez-Outschoorn UE, Lisanti MP, Sotgia F. Catabolic cancer-associated fibroblasts transfer energy and biomass to anabolic cancer cells, fueling tumor growth. Seminars in Cancer Biology. 2014; 25:47–60. [PubMed: 24486645]

29. Pavlides S, Whitaker-Menezes D, Castello-Cros R, Flomenberg N, Witkiewicz AK, Frank PG, Casimiro MC, Wang C, Fortina P, Addya S, Pestell RG, Martinez-Outschoorn UE, Sotgia F, Lisanti MP. The reverse Warburg effect: aerobic glycolysis in cancer associated fibroblasts and the tumor stroma. Cell Cycle. 2009; 8:3984–4001. [PubMed: 19923890]

30. Casazza A, Di Conza G, Wenes M, Finisguerra V, Deschoemaeker S, Mazzone M. Tumor stroma: a complexity dictated by the hypoxic tumor microenvironment. Oncogene. 2013; 33:1743–1754. [PubMed: 23604130]

31. Fong MY, Zhou W, Liu L, Alontaga AY, Chandra M, Ashby J, Chow A, O'Connor STF, Li S, Chin AR, Somlo G, Palomares M, Li Z, Tremblay JR, Tsuyada A, Sun G, Reid MA, Wu X, Swiderski P, Ren X, Shi Y, Kong M, Zhong W, Chen Y, Wang SE. Breast-cancer-secreted miR-122 reprograms glucose metabolism in premetastatic niche to promote metastasis. Nat Cell Biol. 2015; 17:183–194. [PubMed: 25621950]

32. Sica A, Mantovani A. Macrophage plasticity and polarization: in vivo veritas. J Clin Invest. 2012; 122:787–795. [PubMed: 22378047]

33. Luo Y, Zhou H, Krueger J, Kaplan C, Lee SH, Dolman C, Markowitz D, Wu W, Liu C, Reisfeld RA, Xiang R. Targeting tumor-associated macrophages as a novel strategy against breast cancer. J Clin Invest. 2006; 116:2132–2141. [PubMed: 16862213]

34. Zhang F, Liu H, Jiang G, Wang H, Wang X, Wang H, Fang R, Cai S, Du J, Krijgsveld J. Changes in the proteomic profile during the differential polarization status of the human monocyte-derived macrophage THP-1 cell line. Proteomics. 2015; 15:773–786. [PubMed: 25411139]

35. Noy R, Pollard JW. Tumor-Associated Macrophages: From Mechanisms to Therapy. Immunity. 2014; 41:49–61. [PubMed: 25035953]

36. Condeelis J, Pollard JW. Macrophages: Obligate Partners for Tumor Cell Migration, Invasion, and Metastasis. Cell. 2006; 124:263–266. [PubMed: 16439202]

37. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature. 2012; 490:61–70. [PubMed: 23000897]

38. Mertins P, Mani DR, Ruggles KV, Gillette MA, Clauser KR, Wang P, Wang X, Qiao JW, Cao S, Petralia F, Kawaler E, Mundt F, Krug K, Tu Z, Lei JT, Gatza ML, Wilkerson M, Perou CM, Yellapantula V, Huang KL, Lin C, McLellan MD, Yan P, Davies SR, Townsend RR, Skates SJ, Wang J, Zhang B, Kinsinger CR, Mesri M, Rodriguez H, Ding L, Paulovich AG, Fenyo D, Ellis MJ, Carr SA. Proteogenomics connects somatic mutations to signalling in breast cancer. Nature. 2016; 534:55–62. [PubMed: 27251275]

39. Zipfel PF, Skerka C. Complement regulators and inhibitory proteins. Nat Rev Immunol. 2009; 344:1058.

40. Medrek C, Pontén F, Jirström K, Leandersson K. The presence of tumor associated macrophages in tumor stroma as a prognostic marker for breast cancer patients. BMC Cancer. 2012; 12:306. [PubMed: 22824040]

41. Teschendorff AE, Miremadi A, Pinder SE, Ellis IO, Caldas C. An immune response gene expression module identifies a good prognosis subtype in estrogen receptor negative breast cancer. Genome Biol. 2007; 8:R157. [PubMed: 17683518]

42. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011; 144:646–674. [PubMed: 21376230]

43. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. 2015; 1:417–425. [PubMed: 26771021]

44. Frankenberger C, Rabe D, Bainer R, Sankarasharma D, Chada K, Krausz T, Gilad Y, Becker L, Rosner MR. Metastasis Suppressors Regulate the Tumor Microenvironment by Blocking Recruitment of Prometastatic Tumor-Associated Macrophages. Cancer Res. 2015; 75:4063–4073. [PubMed: 26238785]

45. Fong MY, Zhou W, Liu L, Alontaga AY, Chandra M, Ashby J, Chow A, O'Connor STF, Li S, Chin AR, Somlo G, Palomares M, Li Z, Tremblay JR, Tsuyada A, Sun G, Reid MA, Wu X, Swiderski P, Ren X, Shi Y, Kong M, Zhong W, Chen Y, Wang SE. Breast-cancer-secreted miR-122 reprograms glucose metabolism in premetastatic niche to promote metastasis. Nat Cell Biol. 2015; 17:183–194. [PubMed: 25621950]

46. Hasmim M, Bruno S, Azzi S, Gallerne C, Michel JG, Chiabotto G, Lecoz V, Romei C, Spaggiari GM, Pezzolo A, Pistoia V, Angevin E, Gad S, Ferlicot S, Messai Y, Kieda C, Clay D, Sabatini F, Escudier B, Camussi G, Eid P, Azzarone B, Chouaib S. Isolation and characterization of renal cancer stem cells from patient-derived xenografts. Oncotarget. 2015; 7:15507–15524.

47. Naba A, Clauser KR, Lamar JM, Carr SA, Hynes RO. Extracellular matrix signatures of human mammary carcinoma identify novel metastasis promoters. Elife. 2014; 3:e01308. [PubMed: 24618895]

48. Suzuki HI, Katsura A, Matsuyama H, Miyazono K. MicroRNA regulons in tumor microenvironment. Oncogene. 2014

49. Hanash S, Schliekelman M. Proteomic profiling of the tumor microenvironment: recent insights and the search for biomarkers. Genome Med. 2014; 6:12. [PubMed: 24713112]

50. Bolouri H. Network dynamics in the tumor microenvironment. Seminars in Cancer Biology. 2015; 30:52–59. [PubMed: 24582766]

51. Hirata E, Girotti MR, Viros A, Hooper S, Spencer-Dene B, Matsuda M, Larkin J, Marais R, Sahai E. Intravital imaging reveals how BRAF inhibition generates drug-tolerant microenvironments with high integrin β1/FAK signaling. Cancer Cell. 2015; 27:574–588. [PubMed: 25873177]

52. Koti M, Siu A, Clément I, Bidarimath M, Turashvili G, Edwards A, Rahimi K, Mes-Masson AM, Masson AMM, Squire JA. A distinct pre-existing inflammatory tumour microenvironment is associated with chemotherapy resistance in high-grade serous epithelial ovarian cancer. Br J Cancer. 2015; 112:1215–1222. [PubMed: 25826225]

53. Junttila MR, de Sauvage FJ. Influence of tumour micro-environment heterogeneity on therapeutic response. Nature. 2013; 501:346–354. [PubMed: 24048067]

54. Burger JA. Targeting the microenvironment in chronic lymphocytic leukemia is changing the therapeutic landscape. Curr Opin Oncol. 2012; 24:643–649. [PubMed: 22960555]

55. Liang W, Kujawski M, Wu J, Lu J, Herrmann A, Loera S, Yen Y, Lee F, Yu H, Wen W, Jove R. Antitumor activity of targeting SRC kinases in endothelial and myeloid cell compartments of the tumor microenvironment. Clin Cancer Res. 2010; 16:924–935. [PubMed: 20103658]

56. Rohan TE, Xue X, Lin H-M, D'Alfonso TM, Ginter PS, Oktay MH, Robinson BD, Ginsberg M, Gertler FB, Glass AG, Sparano JA, Condeelis JS, Jones JG. Tumor microenvironment of metastasis and risk of distant metastasis of breast cancer. J Natl Cancer Inst. 2014; 106:dju136–dju136. [PubMed: 24895374]

57. McMillin DW, Negri JM, Mitsiades CS. The role of tumour–stromal interactions in modifying drug response: challenges and opportunities. Nat Rev Drug Discov. 2013; 12:217–228. [PubMed: 23449307]

58. Mertins P, Yang F, Liu T, Mani DR, Petyuk VA, Gillette MA, Clauser KR, Qiao JW, Gritsenko MA, Moore RJ, Levine DA, Townsend R, Erdmann-Gilmore P, Snider JE, Davies SR, Ruggles KV, Fenyo D, Kitchens RT, Li S, Olvera N, Dao F, Rodriguez H, Chan DW, Liebler D, White F, Rodland KD, Mills GB, Smith RD, Paulovich AG, Ellis M, Carr SA. Ischemia in tumors induces early and sustained phosphorylation changes in stress kinase pathways but does not affect global protein levels. Mol Cell Proteomics. 2014; 13:1690–1704. [PubMed: 24719451]

59. Mellacheruvu D, Wright Z, Couzens AL, Lambert J-P, St-Denis NA, Li T, Miteva YV, Hauri S, Sardiu ME, Low TY, Halim VA, Bagshaw RD, Hubner NC, al-Hakim A, Bouchard A, Faubert D, Fermin D, Dunham WH, Goudreault M, Lin Z-Y, Badillo BG, Pawson T, Durocher D, Coulombe B, Aebersold R, Superti-Furga G, Colinge J, Heck AJR, Choi H, Gstaiger M, Mohammed S, Cristea IM, Bennett KL, Washburn MP, Raught B, Ewing RM, Gingras A-C, Nesvizhskii AI. The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. - PubMed - NCBI. Nat Meth. 2013; 10:730–736.

60. Askenazi M, Ruggles KV, Fenyo D. PGx: Putting Peptides to BED. J Proteome Res. 2016; 15:795–799. [PubMed: 26638927]

61. Tu C, Rudnick PA, Martinez MY, Cheek KL, Stein SE, Slebos RJC, Liebler DC. Depletion of Abundant Plasma Proteins and Limitations of Plasma Proteomics. J Proteome Res. 2010; 9:4982–4991. [PubMed: 20677825]

62. Kakhniashvili DG, Bulla LA, Goodman SR. The human erythrocyte proteome: analysis by ion trap mass spectrometry. Mol Cell Proteomics. 2004; 3:501–509. [PubMed: 14963112]
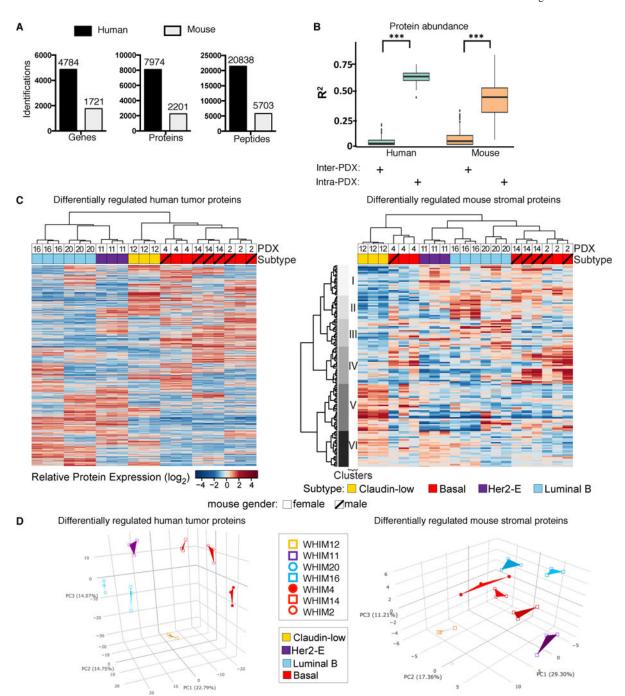
Fig. 1. Cross-species proteomic profiling finds that proteome of breast cancer PDXs share common protein abundance changes in both the tumor and microenvironment across biological replicates

A) Number of genes, proteins and peptides identified from species- and gene-unique peptides sequences. **B)** Boxplot presentation of protein abundance correlation for all seven PDXs (N=3) within biological replicates of the same PDX (intra-PDX) compared to other PDXs (inter-PDX) for the tumor (human) and naïve microenvironment (mouse). The P-value of the correlation differences between intra- and inter-PDX mouse proteins was $2.7e^{-7}$. **C)** Clustered heatmaps of differentially regulated human and mouse proteins. Each column

represents a biological replicate of each PDX. Male mice used for six of the basal subtype PDX samples are denoted (hash). Values are median normalized by column and $\log_2$ transformed. **D)** PCA of differentially regulated human and mouse proteins with each point representing one of three biological replicates for each PDX.
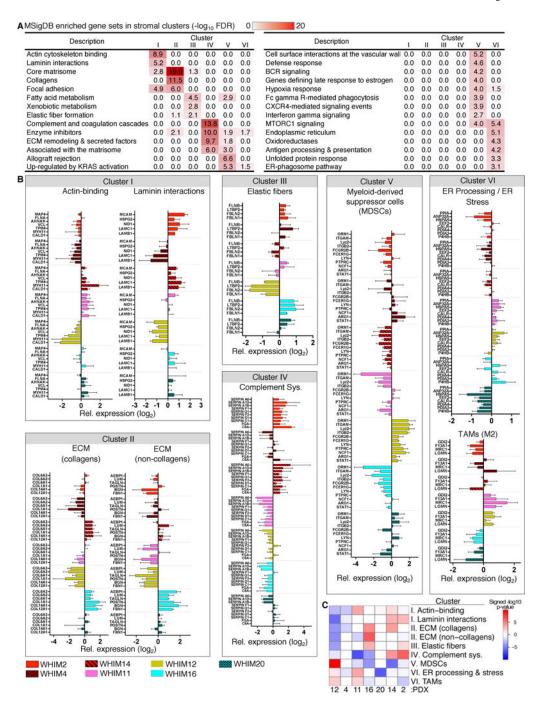
Fig. 2. Individual breast PDXs have unique microenvironmental protein signatures
A) Unbiased MSigDB annotation of proteins (by gene symbol) found within each stromal cluster. **B)** Bar plots of relative protein abundance for specific molecular signatures including the extracellular matrix (ECM), complement system (Complement sys.), endoplasmic reticulum (ER) biology and tumor-associated macrophages (TAMs). Each PDX has a unique signature consisting of multiple proteins within each category. Error bars are the standard deviation of biological replicates (N=3). **C)** Differential patterns of mouse protein signatures were observed between individual tumors. Gene sets annotated in Fig 2A

were used to compare with the baseline using a Student's *t*-test for each combination of gene set and tumor. The intensity of color key is proportional to the magnitude of −log10 adjusted p-value (Benjamini-Hochberg method). Color choices are assigned according to the directionality of deviation from baseline protein abundance (red, up; blue, down). Three biological replicates were analyzed for each PDX.
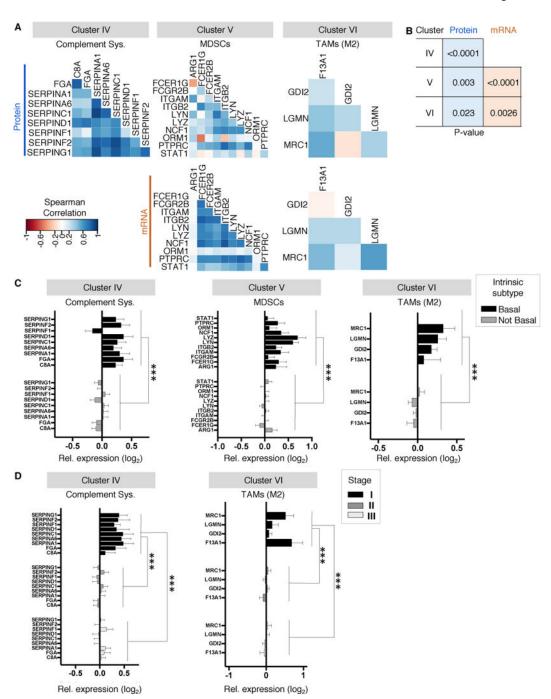
Fig. 3. Stromal proteomic signatures separate patient breast tumors by Stage and Subtype
A) Spearman's correlation coefficients for both the mRNA (N=1,095) and protein (N=77) data sets (by gene symbol) of stromal Clusters IV-VI in the TCGA. Because the complement system is regulated at the protein level (39), only protein results are shown for Cluster IV. **B)** P-values for the correlation matrices determined using Monte Carlo simulation by sampling 10,000 randomized sets of equal size to the test clusters, and ranking the sums of the Spearman's correlation coefficients. **C)** Boxplot of TCGA proteomics data (38) for proteins in Clusters IV and VI by intrinsic subtype (N=77). Error bars are s.e.m. P-value determined

by Student's t-test. **D)** Boxplot of TCGA proteomics data (38) for proteins in Clusters IV and VI by tumor stage (N=77). Error bars are s.e.m. P-value determined by Student's t-test.

A

MSigDB Hallmark	Stromal Cluster						Biological Process	Cancer Hallmark
	I	II	III	IV	V	VI		
Apical junction	-1.52	-2.11	-1.32	1.71	1.91	0.78	Cellular component	Activating invasion and metastasis
Adipogenesis	-1.15	1.67	1.93	-1.76	-1.10	-1.46	Development	Deregulating cellular energetics
Angiogenesis	-1.80	-1.48	-0.63	1.15	1.32	1.02	Development	Inducing angiogenesis
Epithelial-mesenchymal transition	-2.60	-2.61	-1.79	1.82	2.72	1.79	Development	Activating invasion and metastasis
Coagulation	-1.15	0.78	1.32	0.80	2.55	1.12	Immune	Inducing angiogenesis
Complement	-1.61	-1.32	1.57	-0.53	1.72	-0.74	Immune	Tumor-promoting inflammation
Interferon alpha response	-1.40	1.66	0.83	-1.64	1.57	2.43	Immune	Tumor-promoting inflammation
Interferon gamma response	-1.59	0.90	-0.71	-1.74	1.99	3.09	Immune	Tumor-promoting inflammation
Fatty acid metabolism	0.65	1.91	0.96	-1.76	-1.50	1.02	Metabolic	
Heme metabolism	-1.39	0.86	-1.12	0.85	1.78	1.39	Metabolic	
Oxidative Phosphorylation	1.05	2.04	1.74	-2.21	-2.72	-1.28	Metabolic	Deregulating cellular energetics
Xenobiotic metabolism	-0.94	1.57	1.88	-1.23	-0.46	-1.23	Metabolic	
Hypoxia	-1.77	-1.55	-0.89	1.32	2.43	1.13	Pathway	Inducing angiogenesis, Reprogramming
Protein secretion	-1.76	0.78	1.86	-1.31	1.20	0.86	Pathway	Proliferative signaling, Inducing angiogenesis
Unfolded protein response	-1.73	-1.24	-0.92	0.87	1.23	-1.25	Pathway	Resisting cell death
E2F Targets	2.70	-0.68	-2.91	0.66	-1.69	2.33	Proliferation	Proliferative signaling, Evading growth suppressors
G2M checkpoint	2.10	-1.20	-2.25	0.79	-1.90	1.33	Proliferation	Proliferative signaling, Evading growth suppressors
Mitotic spindle	-0.81	-2.02	-1.50	0.95	1.02	1.42	Proliferation	Resisting cell death
IL2 & STAT5 signaling	-2.12	-0.97	0.60	1.55	2.24	0.86	Signaling	Tumor-promoting inflammation
Genes downregulated by activated KRAS	1.43	1.81	-1.71	0.71	-0.92	-0.59	Signaling	Proliferative signaling
MTORC1	-2.03	-1.01	1.59	0.88	1.69	-1.05	Signaling	Resisting cell death
PI3K_AKT_MTOR signaling	-0.89	-2.08	-0.68	0.45	-0.93	1.03	Signaling	Proliferative signaling, Resisting cell death
TNFA signaling via NFKB	-2.01	-1.43	-1.70	1.64	2.25	1.21	Signaling	Tumor-promoting inflammation

B
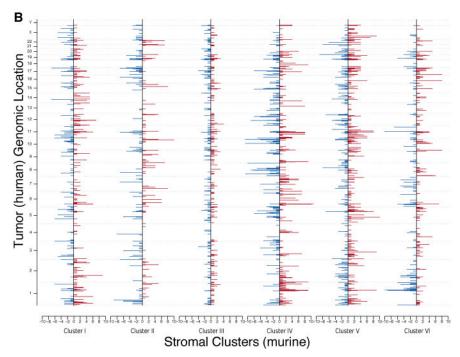
Fig. 4. Tumor proteins that are highly correlated with stromal protein clusters

A) Unbiased MSigDB annotation of proteins found within each stromal (mouse) cluster and assignment to Hanahan and Weinberg's Hallmarks of cancer (42) (N=21 PDX tumors). FDR was calculated using default parameters in GSEA. **B)** Correlation of proteins of individual tumor (human) proteins to stromal (mouse) protein clusters (N=21 PDX tumors). Numbers of significant (Benjamini-Hochberg adjusted p-value <0.05) positive (red) and negative (blue) correlations between human proteins and mouse protein clusters are indicated on the x-axis. Human proteins are displayed by the genome coordinates of their genes based on UCSC hg19 annotation.

Table 1

Tumor proteinss most correlated with stromal proteomic signatures

Human proteins (by gene symbol) most correlated to specific stromal protein signatures. The median Spearman's correlation coefficients (ρ) is highlighted as red (positive) or blue (negative) in breast PDX tumors (N=21).

Stromal Cluster	Human Protein	# Correlated Stromal Proteins (% Proteins in Cluster)	ρ
I	ZNF638	9 (41%)	0.71
	CS	7 (32%)	0.71
	PIP4K2C	7 (32%)	0.73
	ZNF326	6 (27%)	0.74
	MDC1	6 (27%)	0.73
	ALDH4A1	6 (27%)	0.74
	HSPA1B	6 (27%)	0.78
	SF3B2	6 (27%)	0.69
	CFL2	7 (32%)	−0.70
	LGALS3BP	7 (32%)	−0.73
	ITGA2	6 (27%)	−0.75
	LRPAP1	6 (27%)	−0.72
	DEGS1	6 (27%)	−0.76
	RCN1	6 (27%)	−0.75
	RNH1	6 (27%)	−0.71
II	PDCD4	10 (63%)	0.76
	ANP32A	8 (50%)	0.74
	GGCT	8 (50%)	0.78
	ERMP1	7 (44%)	0.79
	HMGN1	7 (44%)	0.73
	ST13	7 (44%)	0.73
	NRD1	9 (56%)	−0.72
	SEPT11	8 (50%)	−0.74
	HDLBP	7 (44%)	−0.74
	PLOD3	7 (44%)	−0.75
	HOOK3	7 (44%)	−0.74
	FTH1	7 (44%)	−0.77
	CYB5B	7 (44%)	−0.78
	HN1	7 (44%)	−0.72
III	ARHGDIB	4 (20%)	0.73
	IARS2	3 (15%)	0.73
	ECI1	3 (15%)	0.73
	MRI1	3 (15%)	0.76
	SORD	3 (15%)	0.69
	PSMD5	3 (15%)	0.69

Stromal Cluster	Human Protein	# Correlated Stromal Proteins (% Proteins in Cluster)	ρ
	SLC9A3R2	3 (15%)	0.82
	STARD10	3 (15%)	0.69
	TPD52	3 (15%)	0.75
	ATP1B3	4 (20%)	−0.71
	LRRC59	4 (20%)	−0.72
	TRIP10	3 (15%)	−0.71
	PFN1	3 (15%)	−0.69
	FLNA	3 (15%)	−0.71
	DPYSL3	3 (15%)	−0.73
	TPD52L2	3 (15%)	−0.72
	AHNAK2	3 (15%)	−0.72
	NXF1	3 (15%)	−0.70
	CUTA	12 (44%)	0.78
	ABCF1	10 (37%)	0.77
	PXDN	9 (33%)	0.80
	APOA1BP	9 (33%)	0.76
	DSG2	9 (33%)	0.75
	IGSF8	9 (33%)	0.75
	PSIP1	9 (33%)	0.74
	ADK	12 (44%)	−0.73
	GSR	11 (41%)	−0.74
IV	PRDX3	11 (41%)	−0.73
	GSDMD	10 (37%)	−0.75
	LACTB2	9 (33%)	−0.81
	CD63	9 (33%)	−0.71
	IVD	8 (30%)	−0.80
	PNPO	8 (30%)	−0.73
	ARFGEF1	8 (30%)	−0.71
	REEP5	8 (30%)	−0.78
	RMDN1	8 (30%)	−0.76
	SLC25A3	8 (30%)	−0.75
	MAP1B	10 (29%)	0.73
	PEA15	9 (26%)	0.73
	GBE1	9 (26%)	0.72
	USP9X	9 (26%)	0.71
	ARL2	8 (24%)	0.73
V	CFL1	8 (24%)	0.72
	HINT1	8 (24%)	0.71
	SERPINH1	8 (24%)	0.74
	TRIM28	9 (26%)	−0.78
	C21ORF33	8 (24%)	−0.76

Stromal Cluster	Human Protein	# Correlated Stromal Proteins (% Proteins in Cluster)	ρ
	PEBP1	7 (21%)	−0.75
	RAD50	7 (21%)	−0.72
VI	MVP	8 (31%)	0.75
	ENO3	8 (31%)	0.73
	PFKP	8 (27%)	0.72
	HLA.A	7 (27%)	0.72
	PSMB8	7 (27%)	0.78
	MSRA	6 (23%)	0.71
	TSTD1	10 (38%)	−0.77
	DHCR7	10 (38%)	−0.74
	APOA1BP	9 (35%)	−0.78
	PEX19	8 (31%)	−0.75
	HYOU1	7 (27%)	−0.72