HHS Public Access

Author manuscript

Clin Cancer Res. Author manuscript; available in PMC 2017 February 01.

Published in final edited form as:

Clin Cancer Res. 2016 February 1; 22(3): 609–620. doi:10.1158/1078-0432.CCR-15-0876.

A patient-derived, pan-cancer EMT signature identifies global molecular alterations and immune target enrichment following epithelial to mesenchymal transition

Milena P. Mak[1,5,*], **Pan Tong**[2,*], **Lixia Diao**[2], **Robert J. Cardnell**[1], **Don L. Gibbons**[1,3], **William N. William**[1], **Ferdinandos Skoulidis**[1], **Edwin R. Parra**[4], **Jaime Rodriguez-Canales**[4], **Ignacio I. Wistuba**[4], **John V. Heymach**[1], **John N. Weinstein**[2], **Kevin R. Coombes**[6], **Jing Wang**[2,#], and **Lauren Averett Byers**[1,#]

[1]Department of Thoracic/Head and Neck Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

[2]Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

[3]Department of Molecular and Cellular Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

[4]Department of Translational and Molecular Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

[5]Medical Oncology, Instituto do Cancer do Estado de Sao Paulo, Faculdade de Medicina da Universidade de Sao Paulo, Sao Paulo, Brazil

[6]Department of Biomedical Informatics, Ohio State University, Columbus, OH, USA

Abstract

Purpose—We previously demonstrated the association between epithelial-to-mesenchymal transition (EMT) and drug response in lung cancer using an EMT signature derived in cancer cell lines. Given the contribution of tumor microenvironments to EMT, we extended our investigation of EMT to patient tumors from 11 cancer types to develop a pan-cancer EMT signature.

Experimental Design—Using the pan-cancer EMT signature, we conducted an integrated, global analysis of genomic and proteomic profiles associated with EMT across 1,934 tumors including breast, lung, colon, ovarian, and bladder cancers. Differences in outcome and in vitro

[#]Corresponding authors: Lauren Averett Byers, Department of Thoracic/Head and Neck Medical Oncology, The University of Texas MD Anderson Cancer Center, Phone: 713-792-6363, Fax: 713-792-1220, lbyers@mdanderson.org. Jing Wang, Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Phone: 713-794-4190, Fax: 713-563-4242.

[*]Co-first authors

Author Contributions

M.P.M., P.T, L.D., J.W. and L.A.B. designed the study. P.T., L.D., F.S., E.R.P., J.R., I.I.W., K.R.C. and J.W. designed and performed the bioinformatics. D.L.G., W.N.W., J.V.H, and J.N.W provided data and contributed to analysis. M.P.M., P.T., and L.A.B. wrote the manuscript. R.J.C. edited the manuscript. L.A.B. supervised the project.

Conflicts of Interest: WNW receives research support from Astellas.

drug response corresponding to expression of the pan-cancer EMT signature were also investigated.

Results—Compared to the lung cancer EMT signature, the patient-derived, pan-cancer EMT signature encompasses a set of core EMT genes that correlate even more strongly with known EMT markers across diverse tumor types and identifies differences in drug sensitivity and global molecular alterations at the DNA, RNA, and protein levels. Among those changes associated with EMT, pathway analysis revealed a strong correlation between EMT and immune activation. Further supervised analysis demonstrated high expression of immune checkpoints and other druggable immune targets such as PD1, PD-L1, CTLA4, OX40L, and PDL2, in tumors with the most mesenchymal EMT scores. Elevated PD-L1 protein expression in mesenchymal tumors was confirmed by immunohistochemistry in an independent lung cancer cohort.

Conclusions—This new signature provides a novel, patient-based, histology-independent tool for the investigation of EMT and offers insights into potential novel therapeutic targets for mesenchymal tumors, independent of cancer type, including immune checkpoints.

Keywords

EMT; Pan-cancer; Immune checkpoints

Introduction

Over the past decade, multiple lines of evidence have suggested that epithelial cancers can transform into a more mesenchymal phenotype, a process known as “epithelial-to-mesenchymal transition” (EMT). Studies have shown that EMT plays an important biological role in cancer progression, metastasis, and drug resistance(1–3). Tools that facilitate the study of EMT, therefore, will provide new insights into molecular regulation and evolution during oncogenesis and may help to improve treatments for mesenchymal cancers. The development of EMT signatures or other molecular markers to identify whether a cancer has undergone EMT is an area of active research. Most studies in this area, however, have focused on a single tumor type and/or on preclinical models(3–7).

We previously developed a robust, platform-independent EMT signature based on a set of 54 lung cancer cell lines (hereafter called the lung cancer EMT signature). *In vitro*, this lung cancer EMT signature predicted resistance to EGFR and PI3K/Akt inhibitors and identified AXL as a potential therapeutic target for overcoming resistance to EGFR inhibitors (a common treatment for non-small cell lung cancer (NSCLC)). We and others observed significantly greater resistance to EGFR inhibitors in lung cancers that had undergone EMT, as determined by either their baseline EMT signature(3) or the expression of specific EMT markers after the development of acquired EGFR inhibitor resistance(8).

Although it is appealing to apply the lung cancer EMT signature to diverse tumor types, this approach may be limited by tumor type specific differences in EMT or discrepancies between cell line models (the starting point of the lung cancer EMT signature(3)) and clinical patient samples—especially in regard to the interplay between EMT, tumor microenvironments, and immune response at different disease sites(9). Therefore, we built

upon our previous approach to derive a pan-cancer EMT signature by leveraging molecular datasets from 11 tumor types (n = 1,934 tumors overall; see Supplementary Table S1) using clinical patient samples from The Cancer Genome Atlas (TCGA). The datasets were obtained from primarily epithelial malignancies such as breast, colorectal, and endometrial cancer, and two cohorts of lung cancer (adenocarcinoma and squamous cell carcinoma).

We tested the performance of the pan-cancer EMT signature by comparing its association with established, but independent, EMT markers at the proteomic, microRNA (miRNA), and mRNA level and then applied the signature as a tool for exploring cross-platform molecular and clinical features associated with EMT across multiple cancer types. Given the previously described association of EMT with drug response (both sensitivity and resistance), we then further evaluated the therapeutic relevance of the pan-cancer EMT signature in preclinical models using two publicly available drug sensitivity databases: the Cancer Cell Line Encyclopedia (CCLE)(10) and the Genomics of Drug Sensitivity in Cancer (GDSC)(11).

Materials and Methods

Datasets

We downloaded cell line drug sensitivity databases from the Cancer Cell Line Encyclopedia (CCLE)(10) and Genomics of Drug Sensitivity in Cancer (GDSC)(11). This included gene expression data (Affymetrix U133 Plus 2 array from CCLE and Affymetrix HT HG U133A array from GDSC) and drug sensitivity data (IC_{50} values). The CCLE data included 1,035 cell lines and the GDSC data included 654 cell lines. In total, 425 cell lines were present in both drug sensitivity databases. Twenty-four targeted drugs were profiled in CCLE and 138 drugs (both targeted and cytotoxic) were profiled in GDSC with a cell viability assay. The GDSC confirmed identity of cell lines by STR analysis and the CCLE by SNP fingerprinting at multiple steps, profiles were compared to existing profiles (10, 11). We used level-3 TCGA pan-cancer data(12, 13), including RNAseqV2, reverse phase protein array (RPPA), miR, copy number, mutation, and clinical data. A summary of the TCGA data can be found in Supplemental Table S1. The PROSPECT (Profiling of Resistance Patterns and Oncogenic Signaling Pathways in Evaluation of Cancers of the Thorax and Therapeutic Target Identification) dataset of surgically resected NSCLC has been previously described(9). Array-based expression profiling of PROSPECT tumors was performed using the Illumina Human WG-6 v3 BeadChip according to the manufacturer's protocol and gene expression data have been previously deposited in the GEO repository (GSE42127).

Developing the pan-cancer EMT signature

We adopted an approach similar to that used in Byers *et al*(3) to derive the pan-cancer EMT signature. We used four established EMT markers, namely *CDH1* (epithelial marker, E type), *CDH2* (mesenchymal marker, M type), *VIM* (M type) and *FN1* (M type) as seeds to derive the pan-cancer EMT signature on the basis of TCGA pan-cancer RNAseq data. In particular, we computed correlations (Pearson's correlation, r) between all mRNAs in the RNAseq data and each of the established EMT markers for each individual tumor type.

Identification of EMT-associated genomic features

We computed Pearson's correlation between the EMT score and individual genomic features since these features are normally distributed. The genomic features per gene included the mRNA expression, RPPA expression (either total protein or phosphorylated protein) and miRNA expression levels (5p and 3p mature strands).

Association between EMT score and other covariates

We applied the log-rank test to assess the association between the dichotomized EMT score (with a cutoff of 0) and overall survival. We used the ANOVA test to assess the association between EMT score and tumor grade.

Pathway analysis

We performed a functional annotation of the pan-cancer EMT signature as well as pathway enrichment analysis using QIAGEN's Ingenuity Pathway Analysis®(14).

Quantitative immunohistochemistry

4μ-thick tissue sections were cut from formalin-fixed, paraffin-embedded blocks containing representative tumor and processed for immunohistochemistry (IHC), using an automated staining system (Leica Bond Max, Leica microsystems, Vista, CA, USA). For assessment of PD-L1 expression, the PD-L1 (E1L3N®)XP® rabbit monoclonal antibody (Cell Signaling Technology) was applied at 1:100 dilution followed by detection using the Leica Bond Polymer Refine detection kit (Leica Microsystems), DAB staining and hematoxylin counterstaining. For quantification, stained slides were digitally scanned using the Aperio® ScanScope Turbo slide scanner (Leica Microsystems). Captured images (200x magnification) were visualized using ImageScope™ software (Leica Microsystems,) and analyzed using Aperio Image Toolbox (Aperio, Leica Microsystems). For PD-L1 analysis in tumor and non-tumor cells, 5 randomly selected square regions (1 mm^2) in the core of each tumor were evaluated. Analysis of PD-L1 expression specifically in tumor cells was based on assessment of the whole section. The cell membrane staining algorithm was used to obtain the PD-L1 H-score (0–300) which is computed on the basis of both extent and intensity of PD-L1 staining.

Association with drug sensitivity data

For each drug sensitivity database, we calculated the EMT scores for the cell lines and associated them with IC_{50} values using the Spearman rank correlation. Cell lines were obtained from commercial vendors and authentication was confirmed by STR analysis matched to existing STR profiles, as reported previously (10, 11).

Additional more detailed methods are included in the Supplemental Methods and Materials.

Results

Development and evaluation of the pan-cancer EMT signature

We derived the pan-cancer EMT signature using an approach similar to that for the lung cancer EMT signature, as shown in the Fig. 1A (3). We selected candidate genes from the mRNAs that best correlated with four established EMT markers "*the seed*", E-cadherin (*CDH1*), vimentin (*VIM*), fibronectin (*FN1*) and N-cadherin (*CDH2*), across 9 distinct tumors types (see Fig. 1 and Supplementary Table S1 for TCGA tumor types and sample sizes). For this initial process, we excluded kidney renal clear cell carcinoma (KIRC) and rectal adenocarcinoma (READ) tumors because they contained predominantly mesenchymal (KIRC) or epithelial (READ) tumors (i.e., low dynamic range of EMT; see Methods). Using this approach, we identified a set of 77 unique genes as the pan-cancer EMT signature (see Methods and Supplementary Table S2). Fig. 1B shows the correlation between these 77 signature genes and the seed component (one of the four established markers) that identified them. We observed strong and consistent correlations with the seed across different tumor types. This suggests that the pan-cancer EMT signature encompasses core EMT markers that function across different tumors.

We observed a wide dynamic range of EMT scores calculated from the pan-cancer EMT signature across the 1,934 tumors that represent 11 distinct tumor types (Fig. 1C). As expected, EMT scores (calculated from the resulting pan-cancer EMT signature) successfully identified KIRC as mostly mesenchymal and READ as mostly epithelial. With the exception of KIRC (90.1% mesenchymal), READ (87.2% epithelial), and colon adenocarcinoma (COAD; 94.4% epithelial), each tumor type included a significant number of both epithelial (EMT score < 0) and mesenchymal (EMT score > 0) samples.

Fourteen genes overlapped between the original lung cancer EMT signature and the new pan-cancer EMT signature (Supplementary Fig. S1A). EMT scores computed from the pan-cancer EMT signature were highly correlated with scores calculated from the lung cancer EMT signature for individual tumor types ($r = 0.76–0.95$; overall $r = 0.82$) (Fig. 1D). The most notable outlier when comparing the two signatures was KIRC, which had the lowest correlation ($r = 0.759$), possibly due to a relatively narrow range of EMT scores due to its mesenchymal tissue of origin.

To investigate whether the pan-cancer EMT signature performed better than the original lung cancer EMT signature, we correlated the EMT scores computed from each signature with six putative EMT markers. The markers included two proteins (E-cadherin and fibronectin) quantified by reverse-phase protein array (RPPA), two microRNAs (miRNAs) (miR-200a, miR-200b; well established as regulators of EMT(15)) as measured by RNAseq, and two EMT-associated transcription factors (*TWIST1, TWIST2*) represented by mRNA expression levels (Fig. 1D). These six putative markers were selected on the basis of their established roles in EMT(1, 2, 16). They served as an independent validation set since they (1) include data from different (non-signature) platforms (e.g., protein, miRNA) and/or (2) were not components of either mRNA-based EMT signature. Fig. 1D shows that correlations with the three putative mesenchymal markers, fibronectin, TWIST1 and TWIST2 (positive correlation with EMT scores), were significantly stronger for the pan-

cancer EMT signature (one-sided t-test, $p < 0.01$); whereas correlations with the three epithelial markers, E-cadherin, miR-200a, and miR-200b, were similar for both signatures (one-sided t-test, $p > 0.05$) (see also Supplementary Fig. S1B).

We then performed a pathway analysis to investigate the biological function of the genes identified in the new pan-cancer EMT signature (Fig. 2A). Genes in the pan-cancer EMT signature were significantly associated with several molecular functions such as cellular movement, morphology, growth and proliferation, cell-to-cell signaling, and cellular assembly and organization. Another enriched canonical pathway was leukocyte extravasation signaling, suggesting immune activation.

An integrated molecular analysis of EMT across tumor types: gene expression and potential therapeutic targets

Using the pan-cancer EMT signature, we evaluated the impact of EMT on the mutational landscape, copy number alterations (CNAs), mRNA, miRNA and protein expression to identify events associated with EMT in patient tumors. To assess how EMT impacts the overall transcriptome, we correlated EMT scores with mRNA expression data for each tumor type (see Supplementary File S1). Genes expressed at higher levels in mesenchymal tumors were highly conserved across multiple tumor types, indicating greater post-EMT biological homogeneity; whereas those expressed at higher levels in epithelial tumors tended to be more specific to individual tumor types (Supplemental Fig. S2A–C).

Given our previously reported association of EMT with drug resistance(3, 4), individual genes or pathways overexpressed in mesenchymal tumors may have clinical utility as therapeutic targets. For example, we previously showed that the receptor tyrosine kinase *AXL* is highly expressed in lung cancers that have undergone EMT and, therefore, is a top candidate drug target in mesenchymal lung cancers(3). In the pan-cancer EMT signature derived here, *AXL* was again identified as a signature gene based on its strong correlation with established EMT markers. *AXL*, therefore, represent a hallmark gene that is positively correlated with EMT scores across all tumor types (r = 0.53–0.95, median = 0.674, Supplementary Fig. S2D). Other potential therapeutic targets that were highly expressed in mesenchymal tumors include the alpha-type platelet-derived growth factor receptor (PDGFRα) (r = 0.22–0.78, median = 0.65), beta-type platelet-derived growth factor receptor (PDGFRβ) (r = 0.59–0.92, median=0.81) and discoidin domain-containing receptor 2 (DDR2) (r = 0.29–0.84, median = 0.63), inhibitors of which are in clinical studies or have been approved for cancer treatment(17–20).

Immune checkpoint expression in mesenchymal tumors

To better understand pathways globally dysregulated in the setting of EMT, we next performed a pathway analysis of genes that correlated with the pan-cancer EMT signature in all 11 tumor datasets, using a cutoff of an absolute r>0.3. In addition to functions related to EMT regulation (e.g., cellular movement, cell-cell signaling), the top activated pathways were related to immune cell signaling (Fig. 2B). Given recently published data from our group that suggests a role of EMT in regulating immune escape in lung cancer(9) and the strong therapeutic potential of emerging immunotherapies, we performed a more focused

analysis to investigate the association between EMT and key genes involved in the immune response across the 11 cancer types.

The expression of 20 potentially targetable immune checkpoint genes [based on current drug inhibitors that are in preclinical development, clinical trials, or which have been approved by the U.S. Food and Drug Administration for specific cancer types (Table 1)] were correlated with the EMT scores for each cancer type. We observed a consistent and strong positive correlation between EMT score and the expression levels of the immune checkpoint genes across all cancer types, with more mesenchymal tumors expressing higher levels of these immune targets (Fig. 2C). Notably, high expression of programmed cell death 1 (*PD1*), cytotoxic T-lymphocyte-associated protein 4 (*CTLA4*) and tumor necrosis factor receptor superfamily, member 4 (*OX40L*, also known as *CD134*) were associated with high EMT score regardless of tumor type (*PD1* median r = 0.33, median $P = 2.1*10^{-4}$; *CTLA4* median r = 0.36, median $P = 1.3*10^{-5}$; *OX40L* median r = 0.65, median $P = 2.4*10^{-13}$). It has been reported that OX40L regulates T cell response, which has led to the study of OX40L inhibition in conjunction with other checkpoint blockades(21, 22).

Looking at individual tumor types, samples with higher EMT scores expressed the immune checkpoint genes in a coordinated way, even in predominantly epithelial tumors such as READ (Fig. 3A). This enrichment of immune target expression in tumors with more mesenchymal characteristics corroborated our recent findings in lung cancer, namely that lung adenocarcinomas with higher lung cancer EMT signature scores expressed higher levels of *PD-L1*, which is a target of miR200 (a suppressor of EMT and metastasis)(9). Clinically, our findings are consistent with results of immunotherapy trials that have shown the greatest activity of immune checkpoint inhibitors in cancer types with the strongest mesenchymal phenotypes, such as melanoma(23–25).

Finally, we validated the association between the mesenchymal phenotype and elevated expression of PD-L1 immunohistochemically in chemotherapy-naïve lung adenocarcinomas and squamous lung cancers that were surgically resected at the University of Texas MD Anderson Cancer Center (PROSPECT dataset). For this analysis, we compared tumors in the top and bottom tertiles for EMT score, using a PD-L1 Histo-score (H-score) ranging from 0 to 300. In agreement with their gene expression profiles that confirmed enrichment for immune-related genes (Fig 3B and 3C, top panels), mesenchymal adenocarcinomas from PROSPECT exhibited higher PD-L1 H-score than epithelial tumors (p=0.003, Mann-Whitney U test) (Fig 3B bottom, left panel). Significantly, this finding was upheld when the analysis was limited to tumor cells (p=0.002, Mann-Whitney U test) (Fig 3B, bottom right panel). We further confirmed the significant, positive correlation between the EMT score and PD-L1 H-score in the entire cohort of adenocarcinomas from PROSPECT, using Spearman's rank correlation co-efficient (p=0.002 for tumor and non-tumor cells and p=0.011 when the analysis was limited to tumor cells).

Similar results were noted in squamous lung cancers, with higher PD-L1 H-score recorded in mesenchymal tumors (p=0.019, Mann-Whitney U test) (Fig 3C, bottom left panel) and a significant, positive correlation noted between the EMT and PD-L1 H –score (p=0.011). Comparison of tumor cell PD-L1 H-score in tumors with EMT scores in the top versus

bottom tertiles did not reach statistical significance in squamous tumors (p=0.110, Mann-Whitney U test, Fig 3C bottom right panel), likely due to a limited number of tumors included in this analysis. However, the correlation between the EMT score and PD-L1 H-score (using all squamous tumors) remained significant (p=0.028).

MicroRNA expression patterns in mesenchymal tumors

To better understand potential factors contributing to the EMT phenotype, we then investigated changes in other data types (miRNA, protein, mutation, copy number) that corresponded with differences in EMT score across the pan-cancer set. Because several miRNAs are known to regulate EMT, we identified miRNAs that were significantly correlated with EMT scores, either positively or negatively, across the pan-cancer sets. Fig. 4A and B show the miRNAs that strongly correlate with the EMT score (absolute r > 0.3 in four or more tumor types). As expected, high expression of miR-200 family members (miR-200a/b/c, miR-141, miR-429), which are known to suppress *ZEB1* and thereby maintain E-cadherin expression and other hallmarks of an epithelial phenotype, was observed among tumors with the lowest (most epithelial) EMT scores (Fig. 4A and B). We observed a positive correlation between the EMT score and numerous other miRNAs, with the miR-199 family (miR-199a/b) having the greatest expression in mesenchymal tumors (Supplementary Fig. S3A). The miR-199 family has been shown to regulate skeletal development and to play a role in melanoma invasion and metastasis, but has been rarely studied in EMT(26, 27). Despite the tumor-specific association observed between the expression of some miRNAs and the EMT score, positively correlated miRNAs (i.e., those expressed at higher levels in mesenchymal tumors) were more frequent and homogenous in this analysis than negatively correlated miRNAs (those higher in epithelial tumors) (Fig. 4A and B and Supplementary Figure S3A), suggesting that miRNA expression may play an important role in regulating EMT.

To investigate the functional roles of miRNAs identified by this analysis, we next investigated whether expression of EMT-associated miRNA target genes also correlated with EMT scores. For the EMT-associated miRNAs in Fig. 4A and B, we downloaded their experimentally validated targets using miRWalk(28) and observed a significant overall association between target expression levels and EMT scores ($p < 0.001$) (Fig 4C). Taking the miR-200 family as an example, Supplementary Fig. S3B shows the strong associations between the miRNA targets and the EMT score. Collectively, these findings support an intricate regulation of EMT that is conserved across multiple tumor types by miRNAs beyond the miR-200 family and which merit further investigation.

Protein expression and EMT

Next, we evaluated the correlation between the EMT score and protein expression using Reverse Phase Protein Array (RPPA) data. RPPA is a highly quantitative platform for the measurement of key total and phosphorylated proteins. In this analysis, we included 181 proteins for which data were available across all 11 tumor types. Notably, some proteins relevant to EMT, such as AXL, are excluded from the analysis because they were evaluated in only a subset of the TCGA pan-cancer tumor types included here.

We found higher levels of several protein markers such as E-cadherin, claudin-7 and Her-3 in epithelial tumors across most tumor types (≥ 8); with higher levels of PAI1 and p21 in mesenchymal tumors (Fig. 4D). Fibronectin-1 was the only marker with consistently higher expression in all 11 cancer types, possibly due to the use of fibronectin (*FN1*) mRNA levels in deriving the pan-cancer signature (one of the seed genes). In contrast, other proteins such as protein kinase C-alpha (PKC-α) were associated with EMT in a subset of tumor types (4 of 11), reflecting tumor-specific EMT differences.

Mutational landscape and copy number alterations

To determine the mutational landscape of mesenchymal tumors, we investigated whether a mutation was associated with the EMT score independent of tumor origin. Despite a significant association, the results can still be primarily attributed to mutations specific to certain tumor types (Supplementary Fig. S4A). For example, a *VHL* mutation primarily observed in KIRC was strongly associated with higher EMT score. Therefore, after adjusting for tumor type and multiple testing, we performed another analysis (see Methods), and found no significant association between mutation and EMT score (Supplementary Fig. S4A). We also explored whether mesenchymal status would influence tumor mutation and CNA burden by quantifying the mutation rate and percentage of CNAs in each sample per tumor type, as described by Ciriello *et al*(29). Using this approach, we found no correlation between CNAs versus mutation distribution and EMT score (Supplementary Fig. S4C), indicating that enrichment of CNAs or mutations is also primarily driven by the tumor tissue of origin.

EMT score and clinical outcomes

In some cancers, EMT may be associated with worse clinical outcomes. Therefore, we assessed whether pan-cancer EMT scores were associated with clinical covariates such as overall survival and tumor grade. We did not observe a significant association between EMT score and overall survival, although we found a trend toward shorter survival times in patients with mesenchymal ovarian carcinomas (OVCA) (Hazard Ratio (HR) 1.32; 95% confidence interval (CI) 0.96–1.81, $P = 0.08$) and uterine corpus endometrial carcinomas (UCEC) (HR 2.07; 95% CI 0.93–4.63, $P = 0.09$). We also found higher EMT scores in KIRC and head and neck squamous cell carcinomas (HNSC) with higher tumor grade ($P < 0.0001$ and $P = 0.01$, respectively) (Supplementary Fig. S5).

Drug sensitivity in mesenchymal tumors

Finally, we tested the ability of the pan-cancer EMT signature to identify associations between EMT and patterns of drug response across 1,689 preclinical models, representing 18 tumor types (Supplementary Table S3). We curated gene expression data from two publicly available databases (CCLE(10) and GDSC(11)) to compute an EMT score for each cell line, which we subsequently correlated with drug sensitivity (based on IC_{50} values). Among 1,035 cell lines in CCLE and 654 cell lines in GDSC with gene expression data, 425 cell lines were common to both sets, allowing us to compare the concordance of EMT scores derived from independently generated expression datasets. Despite the gene expression data having been generated from different versions of the Affymetrix array platforms, EMT

scores calculated for individual cell lines were highly concordant, suggesting a robust performance of our approach across platforms (Fig. 5A). Furthermore, the distribution of EMT scores for the cell line models of each disease type closely mirrored those observed across patient tumors (Fig. 1B and 5B). For example (and as expected), bone sarcoma, glioblastoma multiforme, and melanoma were predominantly mesenchymal (Fig. 5B). To determine patterns of drug sensitivity in mesenchymal tumors, we correlated the EMT score from each cell line and IC_{50} values for targeted drugs tested in the CCLE (24 drugs) and GDSC (138 drugs) databases. As previously demonstrated with the lung cancer EMT signature(3), we found increased resistance to EGFR inhibitors (e.g., erlotinib). Mesenchymal models also demonstrated greater resistance to other drugs that target the ErbB family, including the dual EGFR/VEGF inhibitor vandetanib (ZD-6474) and the EGFR/Her2 inhibitor lapatinib (Fig. 5C and D). In contrast to ErbB inhibitors, but consistent with higher expression of FGFR1 and PDGFR in mesenchymal cancers, we observed increased sensitivity to the FGFR1 and PDGFR multikinase inhibitor dovitinib (TKI258) in cell lines with higher EMT scores (Fig. 5C). To evaluate potential therapeutic targets in the GDSC dataset, we performed an exploratory analysis of the *in vitro* effect of drugs against the same targets. From this target analysis (Fig. 5D and Supplementary Fig. S6), we observed greater sensitivity to PDGFR inhibitors in mesenchymal tumors. Another interesting finding was increased sensitivity in the mesenchymal cell lines to the inhibition of the serine/threonine protein kinases GSK3 and TBK1(30).

Discussion

EMT has long been recognized as playing a major role in cancer progression and metastasis(1, 2). Despite extensive research in the field, the best way to characterize this phenomenon, especially in the clinical setting, is still under debate(2). Here, through the use of a gene expression signature developed from multiple tumor types, we were able to evaluate global molecular alterations associated with EMT, as well as potential therapeutic targets. Comparing our original lung cancer EMT signature with the pan-cancer EMT signature, we observed a better correlation of the pan-cancer EMT signature with known markers of EMT across all tumor types evaluated, verifying the strength of this approach. Interestingly, commonly expressed genes correlated with the pan-cancer EMT signature in all 11 tumor types included a number of potential therapeutic targets, such as the receptor tyrosine kinase AXL, which was also an important target in mesenchymal lung cancers identified our lung cancer EMT signature(3). With AXL inhibitors entering clinical trials, the finding of high *AXL* expression across multiple cancer types in the setting of EMT has important clinical implications(31). Another major finding of this analysis was the significant enrichment of multiple immune targets in cancers that have undergone EMT, which has potentially important implications for identifying cancer patients who may benefit from immunotherapies. Specifically, we found a high correlation of expression of 20 druggable immune targets and mesenchymal status, as defined by the pan-cancer EMT signature and were able to validate these findings by mRNA expression analysis and immunohistochemistry in an independent lung cancer cohort.

Several reports have shown the importance of miRNA regulation in EMT(2, 15, 32). As expected, the miR-200 family was highly expressed in epithelial tumors. However, a number

of novel miRNAs were found to be highly expressed in mesenchymal tumors, regardless of cancer type, including the miR-199 family. Previously implicated in skeletal development and fibrosis(27, 33), miR-199a seems to have a dual role in cancer, both as a tumor suppressor and a promoter, depending on the tissue of origin(27). Although its role in EMT is still unknown, the correlation we found between miR-199 and EMT warrants further investigation.

While mesenchymal tumors had similar patterns of gene and miRNA expression, the overall mutational landscape and alterations in copy number seem to be mostly determined by tumor type. In contrast, protein expression and drug sensitivity patterns appear to be driven by both tumor type and EMT status. From this analysis, we identified several drugs as potentially more active in the mesenchymal subgroup of each tumor type. However, some limitations of this study include different numbers of cell lines for each drug tested within each tumor type and multiple targets (both inhibitory and promoter) for many of the drugs included in the analysis. Nevertheless, we identified drugs that may be more active in mesenchymal cancers, which are often highly resistant to established targeted therapies (e.g., ErbB2 family inhibitors used in breast, head and neck, and lung cancers). Further validation is necessary to determine the clinical relevance of these drugs in the context of EMT.

In summary, the new pan-cancer EMT signature introduced here identifies global molecular alterations across multiple cancer types that are associated with EMT. Furthermore, the identification of the association between pan-cancer EMT signature scores and expression of candidate drug targets suggests that EMT may be a clinically useful marker for selecting patients most likely to respond to specific cancer-targeting approaches. The finding that mesenchymal tumors express higher levels of immune checkpoint targets is of particular clinical relevance given the significant clinical activity of immunotherapies such as PD1 and PD-L1 inhibitors in lung cancer(34, 35) and other malignancies, but current lack of validated biomarkers to select patients most likely to benefit from these targeted immunotherapies. The results described here suggest a possible role for using a tumor’s EMT phenotype to identify cancers that may be more sensitive to immune targeting strategies.

Except for one previous report(4), other existing EMT signatures were developed from a single tumor type(3, 5–7), which is likely to limit their application across diverse cancer types. Furthermore, many of these were developed in cell line models, which do not reflect the contribution of the tumor microenvironment and immune response, which play important roles in therapeutic response, especially with the emergence of the field of immunotherapy. By deriving the pan-cancer EMT signature from patients’ tumors, we overcame some of these limitations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Financial Support:

This project was partially supported by the National Institutes of Health through the TCGA (5-U24-CA143883-05), the Lung SPORE (P50 CA070907), DoD PROSPECT grant W81XWH-07-1-0306, and the Cancer Center Support Grant (CCSG CA016672), by the Mary K. Chapman Foundation, and through generous philanthropic contributions to The University of Texas MD Anderson Lung Cancer Moon Shot Program.

D.L.G. was supported by Rexanna's Foundation for Fighting Lung Cancer, The Stading Lung Cancer Research Fund, a MD Anderson Cancer Center Physician Scientist Award, and the National Institutes of Health (K08-CA151651); W.N.W was supported by a Conquer Cancer Foundation Career Development Award (4546); I.I.W was supported by the UT Lung Specialized Programs of Research Excellence grant (P50CA70907), and MD Anderson's Institutional Tissue Bank (ITB), Award Number 2P30CA016672 from the NIH National Cancer Institute. J.V.H. was supported by NIH/NCI (1 R01 CA168484-04, and P50 CA070907) the LUNGevity foundation and The V Foundation for Cancer Research; J.N.W. was supported by NIH/NCI (5 U24 CA143883 04, 2P50 CA100632-11, 2 UL1RR00037106-A1, and 2 P30 CA016672 38), the Cancer Prevention & Research Institute of Texas (RP130397), and The Michael and Susan Dell Foundation; L.A.B. was supported by a MD Anderson Cancer Center Physician Scientist Award, NCI Cancer Clinical Investigator Team leadership Award (P30 CA016672), the TCGA (5-U24-CA143883-05), the LUNGevity Foundation, the North Carolina Chapter of the National Lung Cancer Partnership, Uniting Against Lung Cancer, The Sidney Kimmel Foundation for Cancer Research, and the Sheikh Khalifa Bin Zayed Al Nahyan Institute for the Personalized Cancer Therapy's (IPCT's) Center for Professional Education and Training; D.L.G. and L.A.B are R. Lee Clark Fellows of The University of Texas MD Anderson Cancer Center, supported by the Jeane F. Shelby Scholarship Fund.

References

1. Singh A, Settleman J. EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer. Oncogene. 2010; 29:4741–51. [PubMed: 20531305]
2. De Craene B, Berx G. Regulatory networks defining EMT during cancer initiation and progression. Nat Rev Cancer. 2013; 13:97–110. [PubMed: 23344542]
3. Byers LA, Diao L, Wang J, Saintigny P, Girard L, Peyton M, et al. An epithelial-mesenchymal transition gene signature predicts resistance to EGFR and PI3K inhibitors and identifies Axl as a therapeutic target for overcoming EGFR inhibitor resistance. Clin Cancer Res. 2013; 19:279–90. [PubMed: 23091115]
4. Tan TZ, Miow QH, Miki Y, Noda T, Mori S, Huang RY, et al. Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients. EMBO Mol Med. 2014; 6:1279–93. [PubMed: 25214461]
5. Loboda A, Nebozhyn MV, Watters JW, Buser CA, Shaw PM, Huang PS, et al. EMT is the dominant program in human colon cancer. BMC Med Genomics. 2011; 4:9. [PubMed: 21251323]
6. Chung CH, Parker JS, Ely K, Carter J, Yi Y, Murphy BA, et al. Gene expression profiles identify epithelial-to-mesenchymal transition and activation of nuclear factor-{kappa}B signaling as characteristics of a high-risk head and neck squamous cell carcinoma. Cancer Res. 2006; 66:8210–8. [PubMed: 16912200]
7. Alonso SR, Tracey L, Ortiz P, Perez-Gomez B, Palacios J, Pollan M, et al. A high-throughput study in melanoma identifies epithelial-mesenchymal transition as a major determinant of metastasis. Cancer Res. 2007; 67:3450–60. [PubMed: 17409456]
8. Zhang Z, Lee JC, Lin L, Olivas V, Au V, Laframboise T, et al. Activation of the AXL kinase causes resistance to EGFR-targeted therapy in lung cancer. Nat Genet. 2012; 44:852–60. [PubMed: 22751098]
9. Chen L, Gibbons DL, Goswami S, Cortez MA, Ahn YH, Byers LA, et al. Metastasis is regulated via microRNA-200/ZEB1 axis control of tumour cell PD-L1 expression and intratumoral immunosuppression. Nat Commun. 2014; 5:5241. [PubMed: 25348003]
10. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature. 2012; 483:603–7. [PubMed: 22460905]
11. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res. 2013; 41:D955–61. [PubMed: 23180760]
12. Akbani R, Ng PK, Werner HM, Shahmoradgoli M, Zhang F, Ju Z, et al. A pan-cancer proteomic perspective on The Cancer Genome Atlas. Nat Commun. 2014; 5:3887. [PubMed: 24871328]

13. Hoadley KA, Yau C, Wolf DM, Cherniack AD, Tamborero D, Ng S, et al. Multiplatform Analysis of 12 Cancer Types Reveals Molecular Classification within and across Tissues of Origin. Cell. 2014; 158:929–44. [PubMed: 25109877]
14. http://www.ingenuity.com/
15. Gregory PA, Bert AG, Paterson EL, Barry SC, Tsykin A, Farshid G, et al. The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. Nat Cell Biol. 2008; 10:593–601. [PubMed: 18376396]
16. Yang J, Mani SA, Donaher JL, Ramaswamy S, Itzykson RA, Come C, et al. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. Cell. 2004; 117:927–39. [PubMed: 15210113]
17. Brunner AM, Costa DB, Heist RS, Garcia E, Lindeman NI, Sholl LM, et al. Treatment-related toxicities in a phase II trial of dasatinib in patients with squamous cell carcinoma of the lung. J Thorac Oncol. 2013; 8:1434–7. [PubMed: 24128713]
18. Lewis NL, Lewis LD, Eder JP, Reddy NJ, Guo F, Pierce KJ, et al. Phase I study of the safety, tolerability, and pharmacokinetics of oral CP-868,596, a highly specific platelet-derived growth factor receptor tyrosine kinase inhibitor in patients with advanced cancers. J Clin Oncol. 2009; 27:5262–9. [PubMed: 19738123]
19. Sternberg CN, Davis ID, Mardiak J, Szczylik C, Lee E, Wagstaff J, et al. Pazopanib in locally advanced or metastatic renal cell carcinoma: results of a randomized phase III trial. J Clin Oncol. 2010; 28:1061–8. [PubMed: 20100962]
20. van der Graaf WT, Blay JY, Chawla SP, Kim DW, Bui-Nguyen B, Casali PG, et al. Pazopanib for metastatic soft-tissue sarcoma (PALETTE): a randomised, double-blind, placebo-controlled phase 3 trial. Lancet. 2012; 379:1879–86. [PubMed: 22595799]
21. Redmond WL, Linch SN, Kasiewicz MJ. Combined targeting of costimulatory (OX40) and coinhibitory (CTLA-4) pathways elicits potent effector T cells capable of driving robust antitumor immunity. Cancer Immunol Res. 2014; 2:142–53. [PubMed: 24778278]
22. Guo Z, Wang X, Cheng D, Xia Z, Luan M, Zhang S. PD-1 blockade and OX40 triggering synergistically protects against tumor growth in a murine model of ovarian cancer. PLoS One. 2014; 9:e89350. [PubMed: 24586709]
23. Robert C, Thomas L, Bondarenko I, O'Day S, MDJ, Garbe C, et al. Ipilimumab plus dacarbazine for previously untreated metastatic melanoma. N Engl J Med. 2011; 364:2517–26. [PubMed: 21639810]
24. Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, Haanen JB, et al. Improved survival with ipilimumab in patients with metastatic melanoma. N Engl J Med. 2010; 363:711–23. [PubMed: 20525992]
25. Hamid O, Robert C, Daud A, Hodi FS, Hwu WJ, Kefford R, et al. Safety and tumor responses with lambrolizumab (anti-PD-1) in melanoma. N Engl J Med. 2013; 369:134–44. [PubMed: 23724846]
26. Pencheva N, Tran H, Buss C, Huh D, Drobnjak M, Busam K, et al. Convergent multi-miRNA targeting of ApoE drives LRP1/LRP8-dependent melanoma metastasis and angiogenesis. Cell. 2012; 151:1068–82. [PubMed: 23142051]
27. Gu S, Chan WY. Flexible and Versatile as a Chameleon-Sophisticated Functions of microRNA-199a. Int J Mol Sci. 2012; 13:8449–66. [PubMed: 22942713]
28. Dweep H, Sticht C, Pandey P, Gretz N. miRWalk--database: prediction of possible miRNA binding sites by "walking" the genes of three genomes. J Biomed Inform. 2011; 44:839–47. [PubMed: 21605702]
29. Ciriello G, Miller ML, Aksoy BA, Senbabaoglu Y, Schultz N, Sander C. Emerging landscape of oncogenic signatures across human cancers. Nat Genet. 2013; 45:1127–33. [PubMed: 24071851]
30. Wei C, Cao Y, Yang X, Zheng Z, Guan K, Wang Q, et al. Elevated expression of TANK-binding kinase 1 enhances tamoxifen resistance in breast cancer. Proc Natl Acad Sci U S A. 2014; 111:E601–10. [PubMed: 24449872]
31. Sheridan C. First Axl inhibitor enters clinical trials. Nat Biotechnol. 2013; 31:775–6. [PubMed: 24022140]

32. Schliekelman MJ, Gibbons DL, Faca VM, Creighton CJ, Rizvi ZH, Zhang Q, et al. Targets of the tumor suppressor miR-200 in regulation of the epithelial-mesenchymal transition in cancer. Cancer Res. 2011; 71:7670–82. [PubMed: 21987723]

33. Lino Cardenas CL, Henaoui IS, Courcot E, Roderburg C, Cauffiez C, Aubert S, et al. miR-199a-5p Is upregulated during fibrogenic response to tissue injury and mediates TGFbeta-induced lung fibroblast activation by targeting caveolin-1. PLoS Genet. 2013; 9:e1003291. [PubMed: 23459460]

34. Brahmer J, Reckamp KL, Baas P, Crino L, Eberhardt WE, Poddubskaya E, et al. Nivolumab versus Docetaxel in Advanced Squamous-Cell Non-Small-Cell Lung Cancer. N Engl J Med. 2015; 373:123–35. [PubMed: 26028407]

35. Gettinger SN, Horn L, Gandhi L, Spigel DR, Antonia SJ, Rizvi NA, et al. Overall Survival and Long-Term Safety of Nivolumab (Anti-Programmed Death 1 Antibody, BMS-936558, ONO-4538) in Patients With Previously Treated Advanced Non-Small-Cell Lung Cancer. J Clin Oncol. 2015; 33:2004–12. [PubMed: 25897158]

Statement of translational relevance

Epithelial-to-mesenchymal transition (EMT) is associated with resistance to many approved drugs and with tumor progression. Here, we sought to characterize the common biology of EMT across multiple tumor types and identify potential therapeutic vulnerabilities in mesenchymal tumors. A patient-derived, pan-cancer EMT signature was developed using 11 distinct tumor types from The Cancer Genome Atlas. Mesenchymal tumors had similar patterns of gene, protein and microRNA expression independent of cancer type. Tumors with mesenchymal EMT scores not only had a higher expression of the receptor tyrosine kinase Axl (previously implicated with EMT and associated with response to Axl inhibitors), but also expressed high levels of multiple immune checkpoints including PD-L1, PD1, CTLA4, OX40L, and PDL2. This novel finding, which was validated in an independent patient cohort, highlights the possibility of utilizing EMT status--independent of cancer type--as an additional selection tool to select patients who may benefit from immune checkpoint blockade.

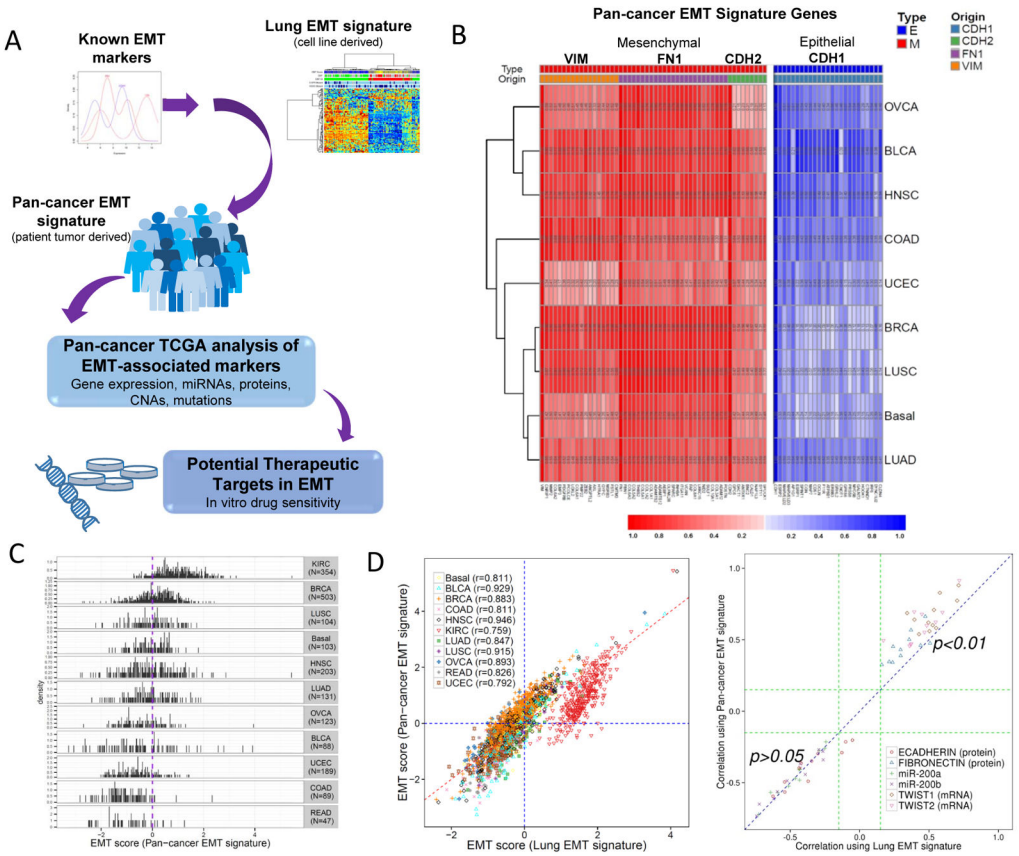

Figure 1. Development and application of the pan-cancer EMT signature

(A) Schematic flow of how the pan-cancer EMT score was derived from known EMT markers (CDH1, VIM, FN1 and CDH2) using patient tumor-derived samples from TCGA and subsequently utilized to identify molecular markers and potential therapeutic targets (B) Genes in the pan-cancer EMT signature preserved strong correlation with known EMT markers across different tumors. “Type:” indicates whether the signature gene was epithelial-like (E) or mesenchymal-like (M) and “Origin:” indicates which seed gene was used to identify the signature gene. E markers (red) had positive correlation with *CDH1* (Pearson’s correlation mean=0.44, median=0.43); M markers (blue) had positive correlation with *VIM*, *FN1* or *C*DH2 (Pearson’s correlation mean=0.67, median=0.71). (OVCA: Ovarian cancer (n=123); BLCA: Bladder urothelial carcinoma (n=88); HNSC: Head and neck squamous cell carcinoma (n=203); COAD: Colon adenocarcinoma (n-89); UCEC: Uterine corpus endometrial carcinoma (n=189); BRCA: Breast invasive carcinoma (n=503); LUSC: Lung squamous cell carcinoma (n=104); Basal: Basal-like breast cancer (n=103); LUAD: Lung adenocarcinoma (n=131). (C) TCGA pan-cancer tumor types exhibit a range of epithelial (score <0) and mesenchymal (score > 0) samples. (D) Comparison of lung cancer and pan-cancer EMT signatures. The EMT scores from each signature were highly concordant, except for KIRC. Using TWIST1/TWIST2 mRNA, miR-200a/b and E-cadherin/fibronectin RPPA data as external validation, the pan-cancer EMT signature performed significantly better ($P < 0.01$) for fibronectin and TWIST1/2 and similarly well for miR-200a/b and E-cadherin.

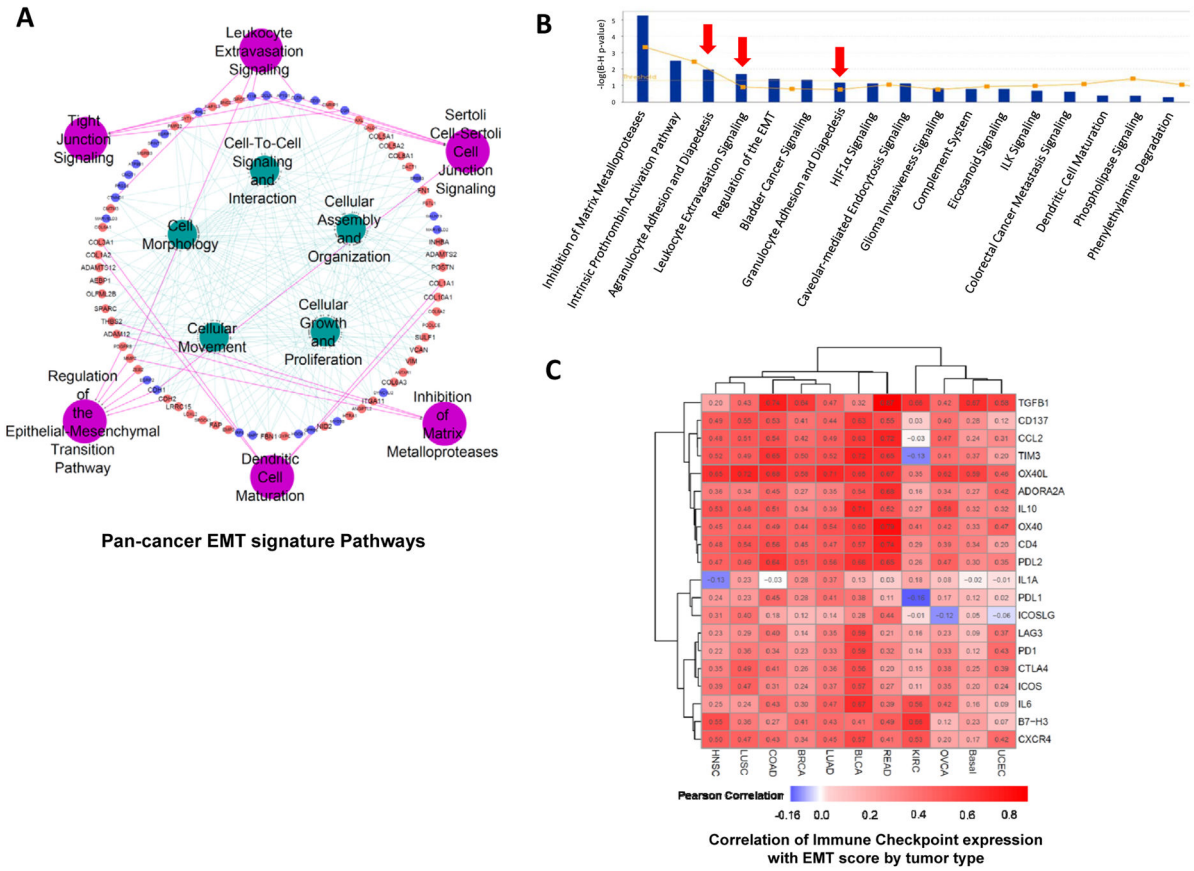

Figure 2. The Pan-cancer EMT signature is enriched with different molecular functions and pathways

(A) Functional annotation of the pan-cancer EMT signature using QIAGEN's Ingenuity Pathway Analysis®(14) found five molecular functions (cyan) and six pathways (purple) were significantly enriched in the pan-cancer EMT signature. Signature gene names are shown in red (genes associated with mesenchymal phenotype) and blue (genes associated with epithelial phenotype) circles, where text size represents average correlation to EMT signature score across tumors. (B) Enriched pathway for genes (using RNAseq data) strongly correlating with EMT score (Ingenuity Pathway Analysis). Immune pathways ranked top among the enriched pathways derived from commonly expressed genes across different mesenchymal tumor types (red arrows). (C) Correlation between immune checkpoint genes and EMT score across different tumor types.

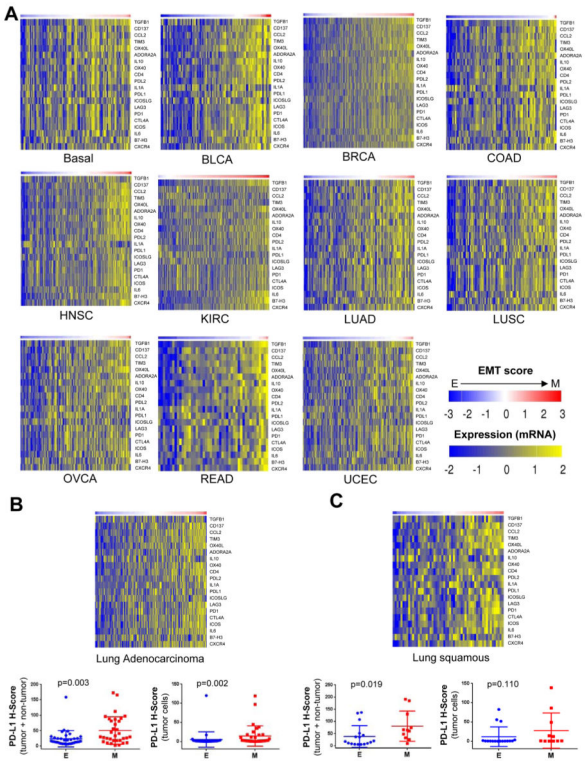

Figure 3. Immune target enrichment in mesenchymal tumors

(A) Expression of immune checkpoint genes was found in the Pan-caner tumor types. Tumor samples within each cancer type are ordered by EMT score. Mesenchymal tumors have higher expression than epithelial tumors of immune checkpoint genes. (B, C) Heatmap representation of mRNA expression levels of selected immune checkpoint genes in 151 lung adenocarcinomas (B) and 57 lung squamous carcinomas (C) from the PROSPECT study, rank ordered on the basis of their EMT score (upper panels). Comparison of PD-L1 Histo-score (H-score) in epithelial (E) and mesenchymal (M) tumors that were included in the PROSPECT tumor microarray. Automated quantification of extent and intensity of PD-L1 staining was performed in tumor and non-tumor cells or limited to tumor cells (B&C, lower panels). Lung adenocarcinomas and squamous carcinomas with EMT score in the top/bottom tertile were included in this analysis. Statistical comparisons are based on the Mann-Whitney U test. Asterisks denote significance at the $p \leq 0.05$ (*) and $p \leq 0.01$ (**) level.

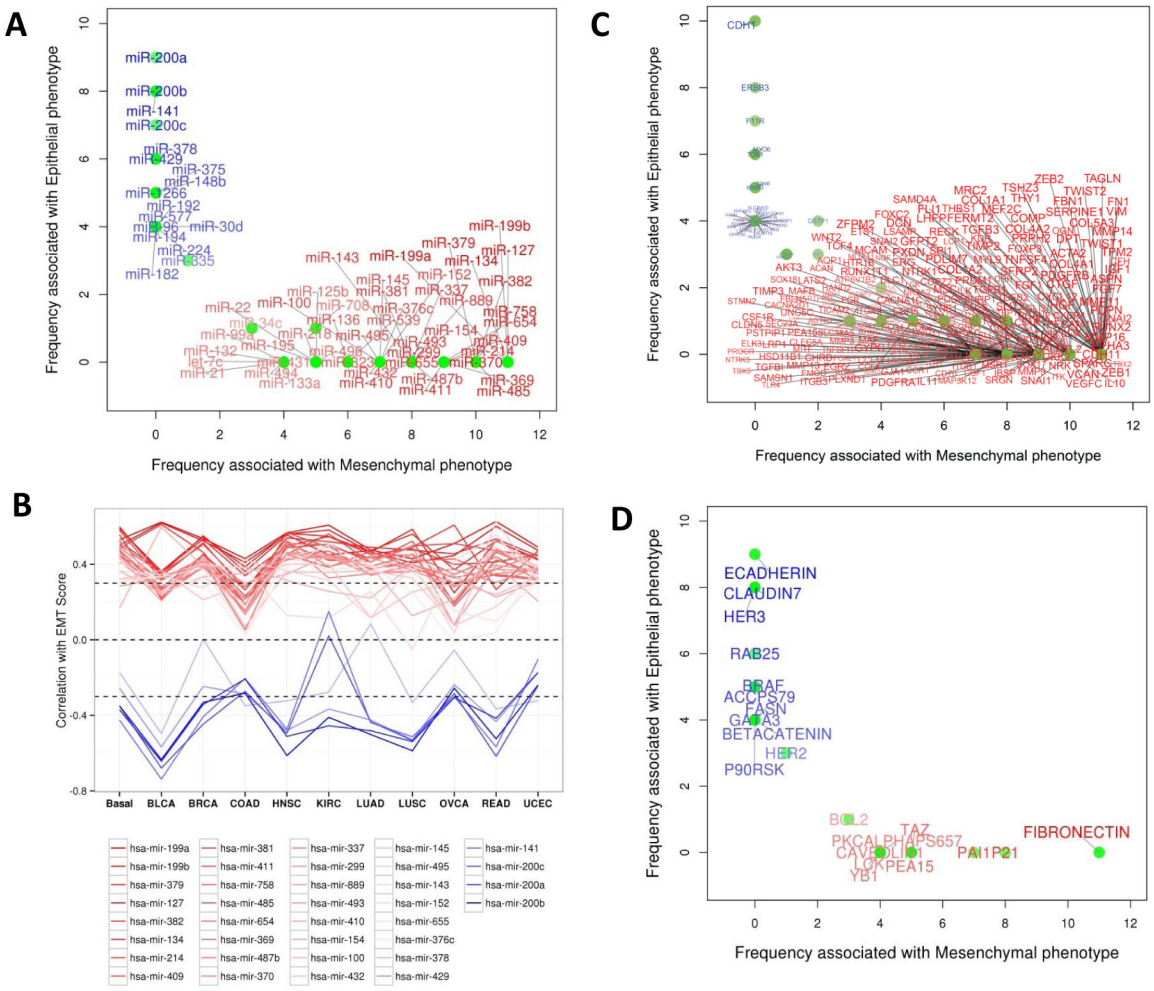

Figure 4. MicroRNA expression is highly correlated with EMT status

(A) Axes show the number of tumor types for which a miRNA was correlated with EMT score (x-axis: miRNA expressed at higher levels in mesenchymal tumors (red); y-axis: miRNA higher in epithelial tumors (blue) at a cutoff of $p \geq 0.3$). miR-200 family (200c/b/c and 141) is anti-correlated with EMT (blue), whereas numerous miRNAs are positively correlated with EMT (red). (B) Correlation between miRNA and EMT score for each cancer type. (C) Targets of EMT-associated miRNAs (from 4A and 4B) were identified by miRWalk. Among these, we observed a statistically significant enrichment of targets whose expression correlated with EMT score (Kolmogorov–Smirnov test $p = 1.8 \times 10^{-28}$). Top miRNA targets associated with EMT score are shown based on an absolute correlation >0.3 in at least four tumors (full list of targets and their correlation with EMT score for each cancer type is provided in Supplemental File 2. (D) Proteins correlated with EMT score. Similar to the finding at gene expression level, epithelial status correlates with expression of HER3 protein, among others. Proteins with an absolute correlation coefficient higher than 0.25 in more than four tumor types are shown (blue, epithelial; red, mesenchymal).

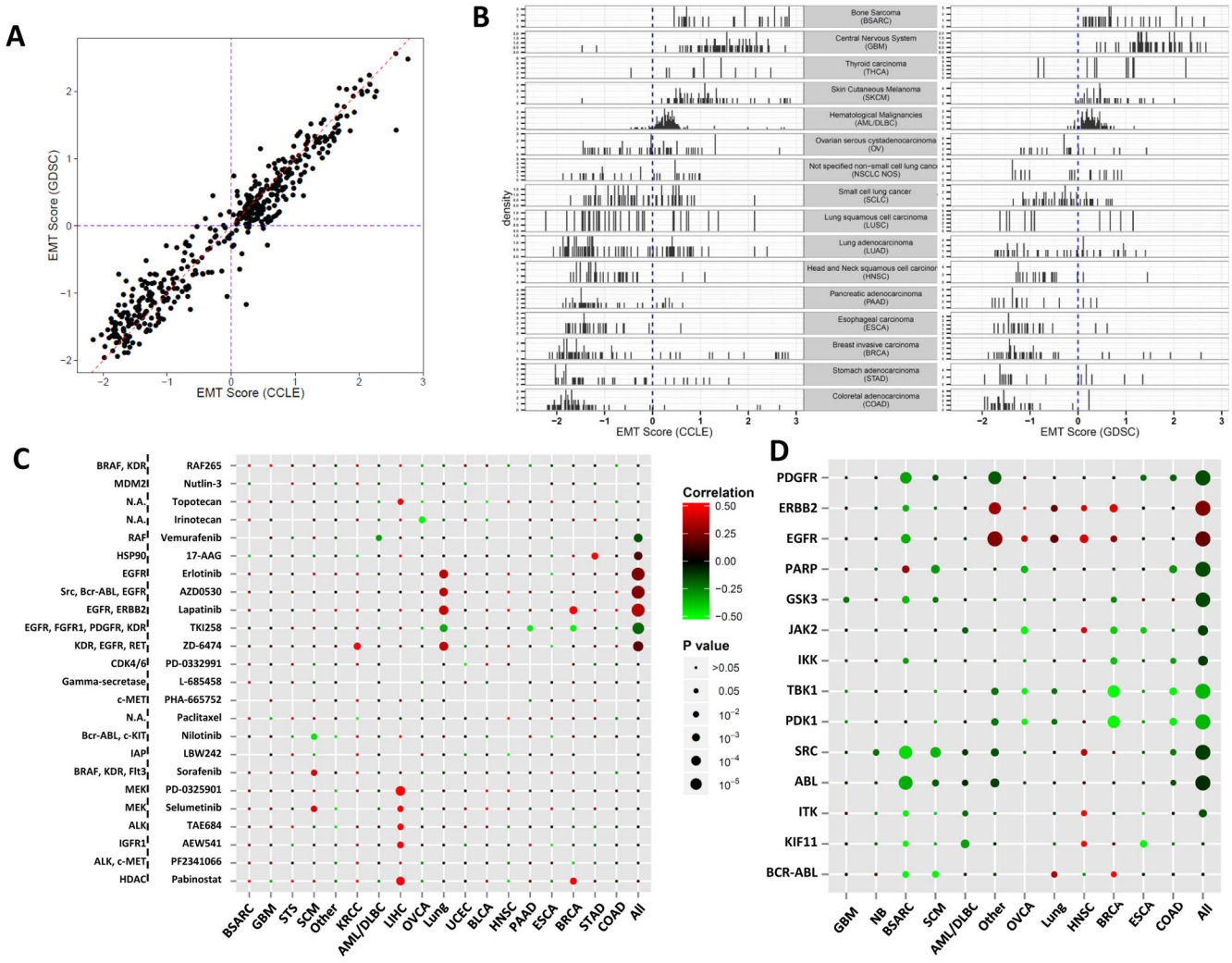

Figure 5. Potential therapeutic targets in mesenchymal cell lines

(A) EMT scores computed from different datasets are highly concordant. Although gene expression data were generated from different laboratories utilizing distinct platforms, high concordance (Pearson's correlation = 0.97, concordance correlation coefficient = 0.96) of the EMT score is found in shared cell lines from both datasets (GDSC, y-axis; CCLE, x-axis), reinforcing the robustness of the signature. (B) Spectra of EMT scores across different tumor types in CCLE and GDSC. Similar to results from patient-derived tumor samples, cell lines from different tumor types exhibited different spectra of EMT scores. (C) Correlation between IC_{50} and EMT score in CCLE. Dot size is proportional to the *P* value of correlation; color indicates magnitude of correlation. (D) Correlation between IC_{50} and EMT score in GDSC. Drugs associated with resistance (red) include EGFR and ERBB2 inhibitors; whereas PDGFR inhibitors have lower IC_{50}s in mesenchymal samples (green).

Table 1

Clinical development of enriched immune targets in mesenchymal tumors

Target	Basal		BLCA		BRCA		COAD		HNSC		KIRC		LUAD		LUSC		OVCA		READ		UCEC		Drugs	Clinical Development
	r	*P*	r	*P*	r	*P*	r	*P*	r	*P*	r	*P*	r	*P*	r	*P*	r	*P*	r	*P*	r	*P*		
ADORA2A	0.27	0.006	0.54	<0.001	0.27	<0.001	0.45	<0.001	0.36	<0.001	0.16	0.002	0.35	<0.001	0.34	0.001	0.34	<0.001	0.68	<0.001	0.42	<0.001	ZM 241385	Pre-clinical
CCL2	0.24	0.016	0.63	<0.001	0.42	<0.001	0.54	<0.001	0.48	<0.001	−0.03	0.625	0.49	<0.001	0.51	<0.001	0.47	<0.001	0.72	<0.001	0.31	<0.001	Carlumab	Phase II (prostate)
CD274 (PDL1)	0.12	0.228	0.38	<0.001	0.28	<0.001	0.45	<0.001	0.24	0.001	−0.16	0.002	0.41	<0.001	0.23	0.017	0.17	0.066	0.11	0.480	0.02	0.762	Pembrolizumab (MK3475)[a], MPDL3280A, MDX-1105, BMS-936559, MSB0010718C	Phase III
CD276 (B7-H3)	0.23	0.021	0.41	<0.001	0.41	<0.001	0.27	0.010	0.55	<0.001	0.66	<0.001	0.43	<0.001	0.36	<0.001	0.12	0.173	0.49	<0.001	0.07	0.340	MGA-271, 8H9	Phase I
CD4	0.34	<0.001	0.57	<0.001	0.45	<0.001	0.56	<0.001	0.48	<0.001	0.29	<0.001	0.47	<0.001	0.54	<0.001	0.39	<0.001	0.74	<0.001	0.20	0.005	Zanolimumab	Phase II
CXCR4	0.17	0.083	0.57	<0.001	0.34	<0.001	0.43	<0.001	0.50	<0.001	0.53	<0.001	0.45	<0.001	0.47	<0.001	0.20	0.028	0.41	0.005	0.42	<0.001	BMS-936564, MSX-122	Phase I
CTLA4	0.25	0.010	0.56	<0.001	0.26	<0.001	0.41	<0.001	0.35	<0.001	0.15	0.005	0.36	<0.001	0.49	<0.001	0.38	<0.001	0.20	0.175	0.39	<0.001	Ipilimumab[a], Tremelimumab	Phase III
HAVCR2 (TIM3)	0.37	<0.001	0.72	<0.001	0.50	<0.001	0.65	<0.001	0.52	<0.001	−0.13	0.014	0.52	<0.001	0.49	<0.001	0.41	<0.001	0.65	<0.001	0.20	0.005	MoAb	Pre-clinical
ICOS	0.20	0.048	0.57	<0.001	0.24	<0.001	0.31	0.003	0.39	<0.001	0.11	0.038	0.37	<0.001	0.47	<0.001	0.35	<0.001	0.27	0.065	0.24	0.001	MoAb[b]	Pre-clinical
ICOSLG	0.05	0.609	0.28	0.008	0.12	0.006	0.18	0.094	0.31	<0.001	−0.01	0.912	0.14	0.103	0.40	<0.001	−0.12	0.172	0.44	0.002	−0.06	0.402	MoAb	Pre-clinical
IL1A	−0.02	0.834	0.13	0.236	0.28	<0.001	−0.03	0.756	−0.13	0.063	0.18	0.001	0.37	<0.001	0.23	0.021	0.08	0.374	0.03	0.844	−0.01	0.888	Anakinra	Phase I
IL6	0.16	0.097	0.67	<0.001	0.30	<0.001	0.43	<0.001	0.00	<0.001	0.56	<0.001	0.47	<0.001	0.24	0.016	0.42	<0.001	0.39	0.007	0.09	0.204	Tocilizumab[a]	Phase I/II (ovarian cancer)
IL10	0.32	0.001	0.71	<0.001	0.34	<0.001	0.51	<0.001	0.53	<0.001	0.27	<0.001	0.39	<0.001	0.48	<0.001	0.58	<0.001	0.52	<0.001	0.32	<0.001	MoAb	Pre-clinical
LAG3	0.09	0.380	0.59	<0.001	0.14	0.001	0.40	<0.001	0.23	0.001	0.16	0.003	0.35	<0.001	0.29	0.003	0.23	0.011	0.21	0.152	0.37	<0.001	IMP321, BMS-986016, MoAb	Phase III (breast), Phase I
PDCD1 (PD1)	0.12	0.226	0.59	<0.001	0.23	<0.001	0.34	0.001	0.22	0.002	0.14	0.008	0.33	<0.001	0.36	<0.001	0.33	<0.001	0.32	0.028	0.43	<0.001	Nivolumab[a], CT011, AMP224	Phase III
PDCD1LG2 (PDL2)	0.30	0.002	0.66	<0.001	0.51	<0.001	0.64	<0.001	0.47	<0.001	0.26	<0.001	0.56	<0.001	0.49	<0.001	0.47	<0.001	0.65	<0.001	0.35	<0.001	Nivolumab[a,b], Pembrolizumab (MK3475), CT011, AMP224	Phase III
TGFB1	0.67	<0.001	0.32	0.002	0.64	<0.001	0.74	<0.001	0.20	0.005	0.66	<0.001	0.47	<0.001	0.43	<0.001	0.42	<0.001	0.87	<0.001	0.58	<0.001	SB-431542, GC 1008	Phase I
TNFRSF4 (OX40)	0.33	0.001	0.60	<0.001	0.44	<0.001	0.49	<0.001	0.45	<0.001	0.41	<0.001	0.54	<0.001	0.44	<0.001	0.42	<0.001	0.79	<0.001	0.47	<0.001	Anti-OX40 MoAb	Phase I
TNFSF4 (OX40L)	0.59	<0.001	0.65	<0.001	0.58	<0.001	0.68	<0.001	0.65	<0.001	0.35	<0.001	0.71	<0.001	0.72	<0.001	0.62	<0.001	0.67	<0.001	0.46	<0.001	Anti-OX40 MoAb[b]	Phase I
TNFRSF9 (CD137)	0.28	0.004	0.63	<0.001	0.41	<0.001	0.53	<0.001	0.49	<0.001	0.03	0.564	0.44	<0.001	0.55	<0.001	0.40	<0.001	0.55	<0.001	0.12	0.102	Anti-CD137 MoAb, BMS-663513 (Urelumab)	Phase I, II (NSCLC CRT)

r – Pearson correlation; MoAb – monoclonal antibody; NSCLC – non-small cell lung cancer

[a]Approved by the U.S. Food and Drug Administration

[b] Indirect targeting