Computational and Structural Biotechnology Journal 19 (2021) 3842–3851

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

# Analysis of immune subtypes across the epithelial-mesenchymal plasticity spectrum

Priyanka Chakraborty [a], Emily L. Chen [b], Isabelle McMullen [b], Andrew J. Armstrong [b,c,d], Mohit Kumar Jolly [a,*], Jason A. Somarelli [b,c,*]

[a] Centre for BioSystems Science and Engineering, Indian Institute of Science, Bangalore 560012, India
[b] Department of Medicine, Durham, NC, United Kingdom
[c] Duke Cancer Institute Center for Prostate and Urologic Cancers, Durham, NC, United Kingdom
[d] Department of Pharmacology and Cancer Biology, Duke University, Durham, NC, United Kingdom

ARTICLE INFO

*Article history:*
Received 23 March 2021
Received in revised form 14 June 2021
Accepted 15 June 2021
Available online 17 June 2021

*Keywords:*
Gene expression signature
Immune subtypes
Tumor heterogeneity
Immune invasion
Hybrid epithelial/mesenchymal cells

ABSTRACT

Epithelial-mesenchymal plasticity plays a critical role in many solid tumor types as a mediator of metastatic dissemination and treatment resistance. In addition, there is also a growing appreciation that the epithelial/mesenchymal status of a tumor plays a role in immune evasion and immune suppression. A deeper understanding of the immunological features of different tumor types has been facilitated by the availability of large gene expression datasets and the development of methods to deconvolute bulk RNA-Seq data. These resources have generated powerful new ways of characterizing tumors, including classification of immune subtypes based on differential expression of immunological genes. In the present work, we combine scoring algorithms to quantify epithelial-mesenchymal plasticity with immune subtype analysis to understand the relationship between epithelial plasticity and immune subtype across cancers. We find heterogeneity of epithelial-mesenchymal transition (EMT) status both within and between cancer types, with greater heterogeneity in the expression of EMT-related factors than of MET-related factors. We also find that specific immune subtypes have associated EMT scores and differential expression of immune checkpoint markers.



## 1. Introduction

Epithelial-to-mesenchymal transition (EMT) is a cellular process in which epithelial cells lose epithelial characteristics, such as apical-basal polarity and tight cell–cell adhesions, and gain mesenchymal features, such as anterior-posterior polarization, focal cell contacts, and enhanced motility and invasiveness [1–3]. Initially characterized in the field of embryology as a feature of normal development, EMT is now also known to play fundamental roles in multiple cellular processes, such as wound healing, fibrosis, and cancer [4]. The regulation of EMT is complex and controlled by the combined action of core EMT transcription factors (EMT-TFs, such as SNAI1/SNAI2, Twist1, and ZEB1/2) [2,3], epithelial factors such as GRHL2 [5]; Jason A. [6] and OVOL1/2 [7], post-transcriptional regulation, including microRNAs [8,9] and alternative splicing [10–12], epigenetic modifications [13], and post-translational regulation [14]. In addition to these cell-intrinsic mechanisms, features of the tumor microenvironment, such as hypoxia [15–18] and interactions between cancer cells and stromal [19–22] and immune cells [23–25] also promote EMT.

Early work on EMT in cancer focused on EMT as a driver of metastatic dissemination. The reversion of cancer cells from a post-EMT-like state back to an epithelial-like phenotype by mesenchymal-to-epithelial transition (MET) was shown to be important for metastatic colonization subsequent to dissemination and seeding [6,26–29]. In addition to metastatic potential, however, EMT has also been shown to function in several other key cancer processes, including tumor-propagating/stemness-like phenotypes [30–34] and treatment resistance [35,36]. Consistent with the role of EMT in metastatic dissemination and therapy resistance the number and EMT phenotype of circulating tumor cells (CTCs) varies depending on a patient's treatment response status, with treatment-refractory patients having more overall CTCs with a higher proportion of mesenchymal-type CTCs, and treatment responders having fewer overall CTCs with a higher proportion of epithelial-type CTCs [37].

* Corresponding author at: Department of Medicine, Durham, NC, United Kingdom (J.A. Somarelli).
*E-mail addresses:* mkjolly@iisc.ac.in (M. Kumar Jolly), jason.somarelli@duke.edu (J.A. Somarelli).

https://doi.org/10.1016/j.csbj.2021.06.023

Although early studies on EMT and MET viewed these phenotypic transitions as binary states, as this field of study has matured the simplified view of EMT/MET dynamics and roles in cancer progression has been shown to be more complex than originally described [38,2]. EMT, in some contexts, has now been shown to be dispensable for metastatic dissemination [35,2,36]. Similarly, MET has also been demonstrated to be dispensable for metastatic colonization, with two distinct paths to metastatic colonization [1,6] and predicted by [39]. Likewise, while the importance of these processes is more complex, the dynamics of this gene regulatory program are also more nuanced than originally appreciated. Rather than a binary switch between states, EMT/MET is now viewed as a spectrum along which cells can have various hybrid E/M phenotypes [38,1,2]. In fact, hybrid E/M phenotypes may be marks of increased cancer aggressiveness and metastatic ability [40]. This has been observed not only in cancer cell lines, but also in clinical CTC samples [41,40,37].

Just as EMT promotes chemoresistance in multiple cancer types, the EMT/MET status of cancers has also been linked to resistance to novel immunotherapies. Immune checkpoint inhibitors have revolutionized the treatment landscape and vastly improved patient outcomes of several cancer subtypes, including non-small cell lung cancer (NSCLC) [42], melanoma [43], and renal cell carcinoma (RCC) [44,43]. However, resistance to these agents is common, and there are several cancer subtypes that do not respond well to immune checkpoint blockade.

One possible explanation for poor immunotherapy response may be the relationship between the EMT status of a tumor and the expression of immune checkpoint molecules [45]. For example, mesenchymal-like cancer cells have been shown to be capable of immunosuppression via interactions with stromal immune cells in the tumor microenvironment. In particular, SNAI1, a core EMT-TF, upregulates expression of CXCL2, a major mediator of myeloid-derived suppressor cell (MDSC) infiltration [46]. MDSCs are major drivers of immunosuppression in the tumor microenvironment. MDSC infiltration renders cancer cells less susceptible to attack by CD8 + T cells and NK cells and leads to an increased ratio of immunosuppressive CD4 + Foxp3 + Treg-like cells, which facilitates tumor growth [47]. SNAI1 and vimentin, a mesenchymal cytoskeletal marker, are positively correlated with programmed death-ligand 1 (PD-L1) scores [48]. Furthermore, PD-L1 transcription is regulated by SNAI1 and the miR200/ZEB1 axis [49]. In contrast, E-cadherin (an epithelial marker) is negatively correlated with PD-L1 scores and epithelial-like cancer cells have higher numbers of infiltrating M1 macrophages and CD8 + T cells in the tumor microenvironment, which allows for greater susceptibility to immune checkpoint inhibitors [50]. It has also been proposed that the regulation between EMT and PD-L1 is actually bidirectional [50,51]. This is supported by the observation that PD-L1 upregulates the EMT-TF and Twist [52]. In addition to this bidirectional regulation, there may be common inducers of EMT and immune evasion, including chronic inflammation, hypoxia, and metabolic reprogramming. There may also be non-cell-autonomous factors contributing to immune responses; for instance, EMT factors, such as ZEB1, driving an M2 macrophage-enriched environment by inducing both PD-L1 and CD47, the latter of which drives M2 polarization [53]. OVOL1, on the other hand, can enable M1 macrophage enrichment by controlling the transcription of IL-10 [54].

Our understanding of EMT, immune evasion, and other facets of cancer biology has been greatly aided by large, publicly-available gene expression datasets [55,56]. Analyses of these datasets have allowed researchers to characterize cancer types to an unprecedented level of detail. The coupling of these large genomics data sets with innovative algorithms that can uncouple bulk RNA-Seq data has also enabled inference of immune subtypes based on differential expression of immune-related genes within tumors [57–59]. These immune subtypes have prognostic value and can be used for survival stratification. Differences in outcome exist within and between cancers when stratified by immune subtypes.

In this study, we combined three novel EMT scoring metrics [60] with immune subtype analysis from the iATLAS algorithm [59–61] to understand the relationships between EMT and immune signature within and between cancer types. Our analyses demonstrate heterogeneity in both E/M phenotype and immune signature within a single cancer type. By comparing known drivers of epithelial and mesenchymal lineages across cancers, we reveal consistency in gene expression of epithelial factors across epithelial tumors, but cancer type-specific expression of a subset of mesenchymal drivers, suggesting that epithelial-derived cancers have convergence of key epithelial driver genes while mesenchymal-derived cancers may be driven by heterogeneous expression of one or more EMT-TFs. These drivers of E/M phenotype are also associated with distinct immune subtypes, with enrichment for specific immune subtypes across the EMT spectrum, illustrating the relationships between E/M status, immune subtype, with potential implications for patient response to immunotherapy.

## 2. Results

### *2.1. The EMT status of cancers is heterogeneous within and between cancer types*

The EMT status of specific tumor types has been quantified using a variety of previously-established signatures, most of which use a limited set of molecules or functional traits and/or individual algorithms to measure the extent of EMT on a continuum [62–68]. Here, to provide a robust comparison of EMT-like status across cancer types, we calculated the EMT scores with available RNA-Seq data from the Cancer Cell Encyclopedia (CCLE) using three distinct EMT scoring algorithms [69,40,70]. Each of these three metrics – KS, MLR and GS76 - score the extent of EMT on a continuum, based on the expression of EMT-specific genes identified by various groups. These three methods use different gene lists and scoring methods: the GS76 method uses a weighted sum of expression levels of 76 genes, the KS method compares the cumulative distribution functions of epithelial and mesenchymal signatures, and the MLR method uses a multinomial logistic regression to calculate the probability of a sample to belong to varying EMT categories. KS and MLR score samples on a scale of [-1, 1] and [0,2], respectively, while the GS76 metric has no pre-defined scale. Higher MLR or KS scores represent more mesenchymal samples while this is the inverse for GS76 scores (lower = more mesenchymal). Thus, KS and MLR scores of samples in a given dataset correlate positively with one another, and both KS and MLR correlate negatively with GS76 scores, as observed across multiple datasets [60].

Consistent with their lineages of origin, carcinoma cell lines, such as colorectal, breast, stomach and prostate lines have lower median KS and MLR scores and higher GS76 scores as compared to mesenchymally-derived cancer cell lines, such as sarcoma, melanomas and glioma (Fig. 1**A, S1A, S2A**). It is also noteworthy that the variance in EMT scores is higher in cell lines of epithelial lineages as compared to mesenchymally-derived cell lines (Fig. 1**A, S1A, S2A**).

We next investigated the distribution of EMT scores across tumor types from TCGA. Similar to the CCLE data, epithelial tumors (adenocarcinomas of the colon, rectum, stomach, prostate, etc.) have lower mean value of KS and MLR scores and a higher mean value GS76 score as compared to tumors derived from mesenchymal lineages (glioblastoma, glioma, sarcoma) (Fig. 1**B, S1B, S2B**). Also consistent with CCLE data, mesenchymally-derived tumors

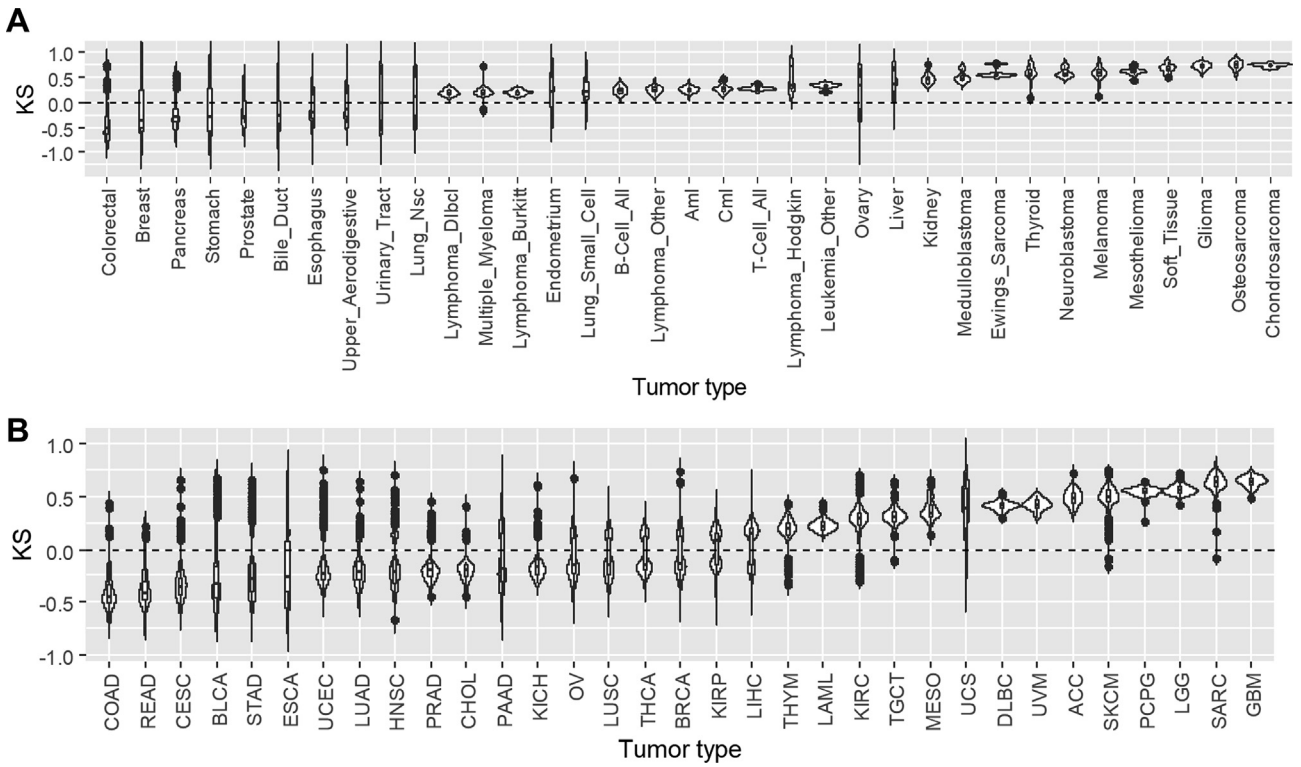


**Fig. 1. Carcinomas have more epithelial EMT scores**. The KS algorithm scores samples on a scale of [-1, 1], with more positive scores representing more mesenchymal samples. **(A)** KS algorithm applied to RNA-Seq data from the Cancer Cell Encyclopedia (CCLE). **(B)** KS algorithm applied to RNA-Seq data from the Cancer Genome Atlas (TCGA); COAD = Colon adenocarcinoma, READ = Rectum adenocarcinoma, CESC = Cervical squamous cell carcinoma and endocervical adenocarcinoma, BLCA = Bladder urothelial carcinoma, STAD = stomach adenocarcinoma, ESCA = esophageal carcinoma, UCEC = uterine corpus endometrial carcinoma, LUAD = lung adenocarcinoma, HNSC = head and neck squamous cell carcinoma, PRAD = prostate adenocarcinoma, CHOL = cholangiocarcinoma, PAAD = pancreatic adenocarcinoma, KICH = kidney chromophobe, OV = ovarian serous cystadenocarcinoma, LUSC = lung squamous cell carcinoma, THCA = thyroid carcinoma, BRCA = breast invasive carcinoma, KIRP = kidney renal papillary cell carcinoma, LIHC = liver hepatocellular carcinoma, THYM = thymoma, LAML = acute myeloid leukemia, KIRC = kidney renal clear cell carcinoma, TGCT = testicular germ cell tumors, MESO = mesothelioma, UCS = uterine carcinosarcoma, DLBC = lymphoid neoplasm diffuse large B-cell lymphoma, UVM = uveal melanoma, SKCM = skin cutaneous melanoma, PCPG = pheochromocytoma and paraganglioma, LGG = brain lower grade glioma, SARC = sarcoma, GBM = glioblastoma multiforme. Violin plots show the median and interquartile range.

have a lower variance in scores whereas carcinomas display higher variability in EMT score (Fig. 1**B, S1B, S2B**), suggesting less heterogeneity in the EMT status of mesenchymally-derived cancers. To account for any potential bias in EMT score calculation that may be due to excess stromal contamination we applied a previously-reported combination of five tumor purity prediction methods [71]. EMT score was weakly correlated with tumor purity (**Figure S3A**). The ESTIMATE algorithm was the most highly correlated to EMT score across cancers, with correlations of − 0.47, −0.23, and 0.17 to KS, MLR, and GS76, respectively (**Figure S3A**). Other metrics of tumor purity, such as immunohistochemistry and the LUMP, ABSOLUTE, and CPE algorithms, had a maximum correlation of − 0.23 (CPE vs. KS; **Figure S3A**). We also re-analyzed TCGA samples in the 50th and 75th percentiles of tumor purity. While samples with higher tumor purity tended to have lower variance, samples with higher purity were not markedly different in EMT score compared to all samples in a given cancer type (**Figure S3B**). Together, these analyses suggest that, while single samples may be influenced to a moderate extent by tumor purity, EMT score for a given cancer type is driven predominantly by tumor cell lineage and not a consequence of high stromal contamination.

### *2.2. EMT-TFs display heterogeneous patterns across cancers while MET factors are more highly correlated*

Given the heterogeneity in EMT scores within and between carcinomas, we asked if this heterogeneity was correlated with differences in EMT-TFs and MET-associated factors. To address this question we assessed the levels of a panel of five well-studied EMT-inducing or EMT-associated factors – SNAI1, SNAI2, ZEB1, ZEB2, and TWIST1 [72–75] – and five MET-inducing or MET-associated factors – GRHL2, OVOL1, OVOL2, ESRP1, and ESRP2 [11,76–79] – across tumor types. These factors are well-studied regulators of EMT and MET *in vitro* and *in vivo*, and many of them are involved through double negative feedback loops enabling multiple phenotypes; for instance ZEB1 forms such loops with OVOL1/2, GRHL2 and ESRP1 [11,7,80]. Such feedback loops drive decision-making during embryonic development and cancer progression [81]. The breakdown of these feedback loops can restrict cellular plasticity, with consequent impact on metastatic potential [82,83]. While GRHL2 and OVOL1/2 are transcription factors, ESRP1 and ESRP2 are splicing regulatory players associated with epithelial splicing programs in development and in cancer [10,84,85].

From this analysis of EMT and MET factors, we noted the emergence of two large clusters: one set of tumors comprised predominantly of carcinomas, with higher levels of MET factors and low levels of EMT factors (top cluster, Fig. 2**A**), and another set of tumors with the opposite trend, comprised mostly of mesenchymally-derived cancers (glioblastoma, glioma, and sarcoma). We also noted a distinct difference in the relationship between MET factors and EMT factors in these groups: In the carcinoma subset MET factors are all highly expressed, while this consistency across EMT factors is not observed in the predominantly-mesenchymal cancers. Instead, mesenchymal tumors are characterized by upregulation of one or two predominant EMT-factors, such as SNAI2 for uveal melanoma and TWIST1

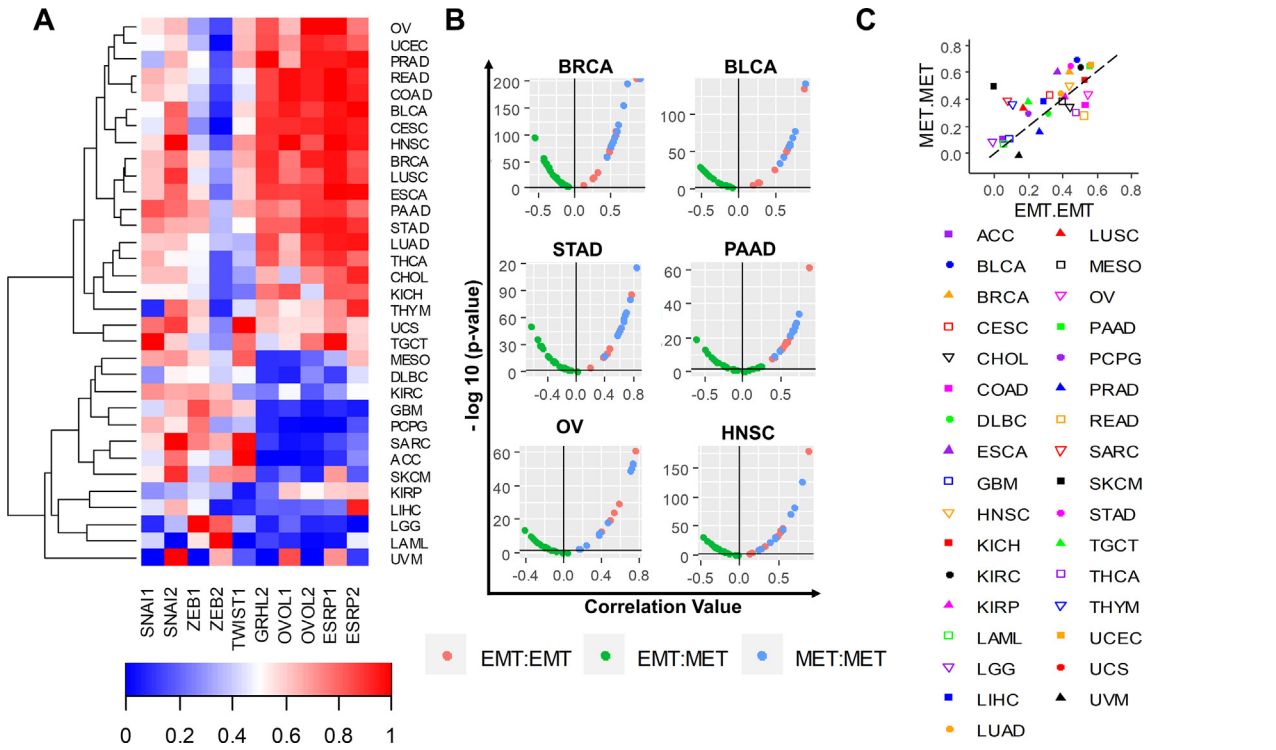


**Fig. 2. EMT factors are more heterogeneous across cancers than MET factors.** EMT and MET marker expression across TCGA tumor types; COAD = Colon adenocarcinoma, READ = Rectum adenocarcinoma, CESC = Cervical squamous cell carcinoma and endocervical adenocarcinoma, BLCA = Bladder urothelial carcinoma, STAD = stomach adenocarcinoma, ESCA = esophageal carcinoma, UCEC = uterine corpus endometrial carcinoma, LUAD = lung adenocarcinoma, HNSC = head and neck squamous cell carcinoma, PRAD = prostate adenocarcinoma, CHOL = cholangiocarcinoma, PAAD = pancreatic adenocarcinoma, KICH = kidney chromophobe, OV = ovarian serous cystadenocarcinoma, LUSC = lung squamous cell carcinoma, THCA = thyroid carcinoma, BRCA = breast invasive carcinoma, KIRP = kidney renal papillary cell carcinoma, LIHC = liver hepatocellular carcinoma, THYM = thymoma, LAML = acute myeloid leukemia, KIRC = kidney renal clear cell carcinoma, TGCT = testicular germ cell tumors, MESO = mesothelioma, UCS = uterine carcinosarcoma, DLBC = lymphoid neoplasm diffuse large B-cell lymphoma, UVM = uveal melanoma, SKCM = skin cutaneous melanoma, PCPG = pheochromocytoma and paraganglioma, LGG = brain lower grade glioma, SARC = sarcoma, GBM = glioblastoma multiforme **(A)** Normalized EMT and MET marker expression across all TCGA tumor types. **(B)** Pairwise correlations between expression values of all EMT-EMT, MET-MET, and EMT-MET factor pairs across a subset of TCGA tumor types. **(C)** Plot of mean pairwise correlation coefficients for all EMT-EMT factor pairs (x-axis) versus mean pairwise correlation coefficients for all MET-MET factor pairs (y-axis) across all TCGA tumor types.

and SNAI2 for sarcoma (Fig. 2**A**). Conversely, while the mesenchymal-like cluster displays consistently low levels of MET factors, the epithelial-like cluster displays high expression of EMT factors in several cases. Principal component analysis indicated that expression of MET factors is the predominant contributor to the formation of the epithelial-like and mesenchymal-like clusters, while the EMT factors alone are unable to segregate clusters by epithelial/mesenchymal lineage (**Figure S4**).

To further analyse this trend quantitatively, we calculated the pairwise correlations between expression values of all 10 EMT and MET factors across a subset of tumor types. Consistent with our qualitative observations above, the five MET factors are significantly positively correlated with each other, while the EMT and MET factors are significantly negatively correlated. However, unlike the MET factors, the correlation among different EMT factors is less consistent and often not statistically significant (Fig. 2**B**). We also quantified the mean of the correlation coefficients for all pairwise correlations between any two EMT factors or for any two MET factors. These values are shown as x- and y-axes on a scatter plot, where each dot represents a tumor type. The majority of the tumors are above the $\times = y$ line, signifying that the average correlation between any two MET factors is greater than that between two EMT factors (Fig. 2**C**). Such correlation is reminiscent of recent observations of groups of factors engaged in various feedback loops, driving phenotypic plasticity and heterogeneity in many cancers [86,87]. Similarly, visualization of the variance of the pairwise correlation coefficients for EMT and MET factors revealed a lower variability in MET factors as compared to EMT factors (**Figure S5**). These trends are also observed in CCLE cancer types (**Figures S6, S7**). Together, these analyses suggest that more heterogeneity in gene expression exists for EMT factors as compared to MET factors.

### *2.3. EMT score is associated with specific immune subtypes in cancer*

Activation of a partial or complete EMT has been associated with immune evasion across multiple cancers [88,89]. A recent approach characterized immune tumor microenvironment across 33 cancer types in TCGA and defined six major immune subtypes spanning cancer types and molecular subtypes – C1 (wound healing), C2 (IFN-γ dominant), C3 (inflammatory), C4 (lymphocyte depleted), C5 (immunologically quiet), and C6 (TGF-β dominant) [61]. Given the relationship between EMT and immune checkpoint molecules [49,48,88] as well as immunomodulatory cytokines, we sought to understand the relationship between EMT status and immune subtype within and across cancers. To do this, we calculated EMT scores for all TCGA samples for which immune subtypes have been assigned. Overall, cancers with the C1 wound healing subtype and C2 IFN-γ dominant subtype tended to have more epithelial-like scores, while all other subtypes were more mesenchymal (Fig. 3**A, S8A, C**). In particular, the C5 quiescence subtype is the most mesenchymal subtype (Fig. 3**A-B, S8**). This observation is consistent with its sample composition, as C5 contained almost exclusively low-grade glioma samples [61]. A leave-one-out analysis further demonstrated these trends, with the wound healing (C1) and IFN-γ (C2) signatures more epithelial than all other cancer

samples, and the other immune signatures more mesenchymal than other signatures/samples (Fig. 3**B, S8B, D**). We applied Singscore, a recently-developed rank-based single sample method to quantify the epithelial and mesenchymal scores separately [90] (Fig. 3**C**). Singscore is a non-parametric method that calculates scores for ensembles of gene sets corresponding to epithelial and mesenchymal phenotypes. Using this method, the C5 samples showed high mesenchymal scores and very low epithelial scores. Interestingly, the mesenchymal scores for all six immune subtypes do not vary much in terms of fold-change (despite being statistically significantly different) while the epithelial score is the lowest for C5 (Fig. 3**D**). The C6 TGF-β dominant immune subtype shows a high score for both the axes, which may suggest that cancers with this immune signature would have a more hybrid epithelial/mesenchymal (E/M) phenotype (Fig. 3**D**).

Given the marked heterogeneity in EMT scores and immune subtypes, we sought to understand whether particular cancer types are enriched in specific EMT scores and immune subtypes. Analysis of the upper and lower quartiles of EMT score across immune subtypes revealed distinct cancer types within each of the immune subtypes. For example, tumors in the lower quartile of the C1 wound healing subtype are enriched in colorectal cancer specimens as compared to the lower quartile of C1 (**Figure S9A**). Conversely, the composition of cancer types between upper quartile of EMT scores within the IFNγ-dominant C2 subtype are spread across multiple cancer types of diverse lineages, including, but not limited to, breast cancer, head and neck cancer, sarcomas, melanomas, and testicular germ cell tumors (**Figure S9A, B**). Likewise, the cancer types in upper and lower quartiles of EMT scores within the inflammatory (C3) subtype also differ substantially. More mesenchymal-like tumors within the C3 subtype (inflammatory) are enriched in renal clear cell carcinomas as compared to the lower quartile of C3 tumors, which is comprised of lung, prostate, breast, and thyroid cancers (**Figure S9A, B**). Analysis of tumor microenvironment cell populations within each immune subtype using the microenvironment cell populations (MCP)-counter method revealed enrichment of gene expression signatures consistent with fibroblasts, endothelial cells, and cells of a monocyte lineage (**Figure S10**). The lymphocyte-depleted (C4) and immunologically quiescent (C5) subtypes are the most depleted of immune subtypes, while the IFN-γ dominant subtype (C2) has the highest enrichment in T cell signatures and the TGF-β-dominant subtype (C6) has the greatest enrichment in signatures of myeloid-derived cells (**Figure S10**).

To further understand how EMT and MET factors may be associated with specific immune signatures we performed separate analyses of EMT factors for each immune subtype. EMT factors SNAI1, SNAI2, and Twist are most highly expressed in the C6 TGF-β dominant immune signature, with ZEB1 and ZEB2 most upregulated in the C5 immunologically quiescent signature (Fig. 4**-A-E**). Other immune subtypes have low to moderate levels of EMT factors, with wide distributions in expression across these subtypes. MET factors are predominantly expressed in C1-C3 signatures, with a bimodal distribution of expression in C4, low expression in C5, and relatively high expression in C6 (Fig. 4**F-J**). Such co-expression of EMT and MET factors in C6 samples suggests that samples in the TGF-β dominant C6 immune subtype correspond to a more hybrid epithelial/ mesenchymal phenotype(s).

We also calculated the pairwise correlations between the EMT scores from the different scoring metrics (GS76, KS, MLR) and expression of EMT and MET factors from samples in each immune subtype. Across all immune signatures, these pairwise correlations showed consistent trends: GS76 scores correlated positively with levels of MET-factors and negatively with those of EMT-factors;

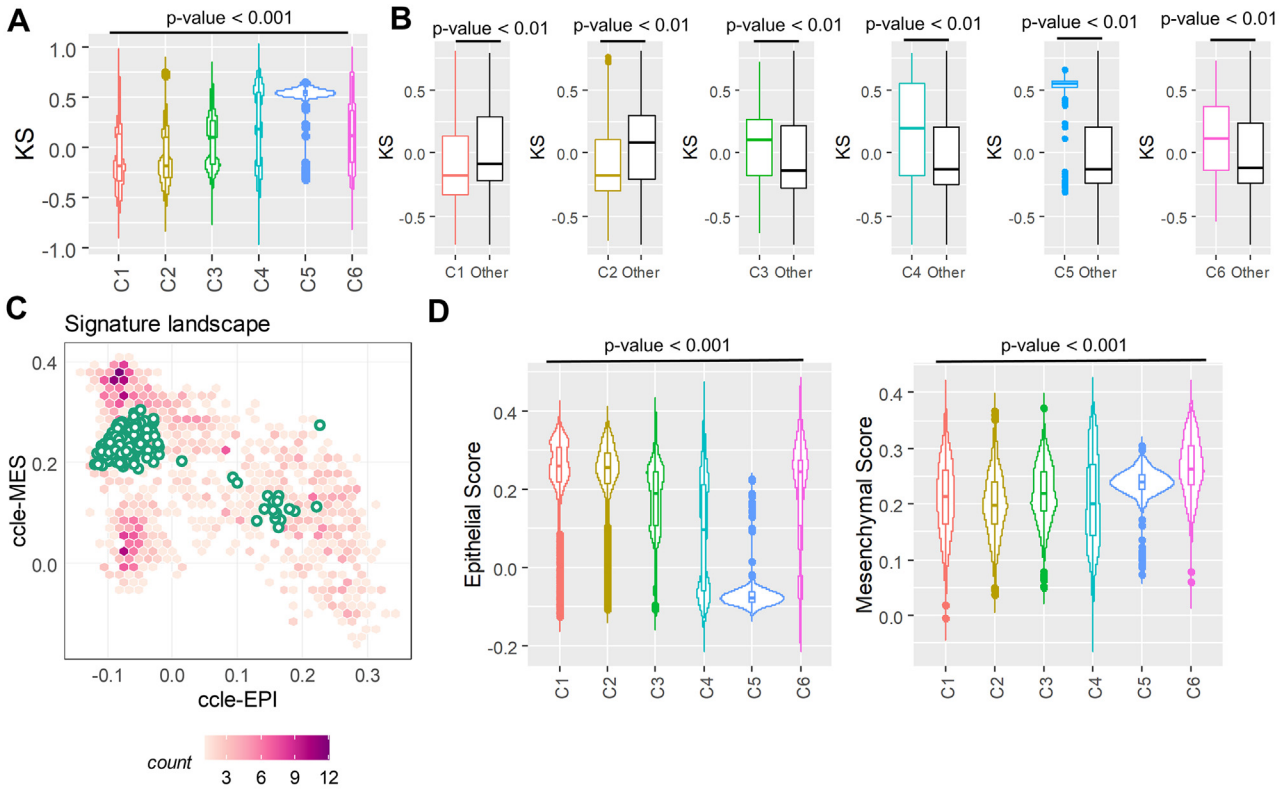


**Fig. 3. Immune subtypes are associated with EMT scores.** EMT scores across TCGA immune subtypes; C1 = wound healing, C2 = IFN-γ dominant, C3 = inflammatory, C4 = lymphocyte depleted, C5 = immunologically quiet, C6 = TGF-β dominant. **(A)** Plot of calculated KS score across all immune subtypes. **(B)** Leave-one-out-analysis: pairwise comparison of each cancer immune subtype's KS score to the KS scores of all other immune subtypes. **(C)** Application of Singscore to C5 immune subtype samples to quantify epithelial and mesenchymal scores separately. **(D)** Singscore epithelial and mesenchymal scores across immune subtypes.

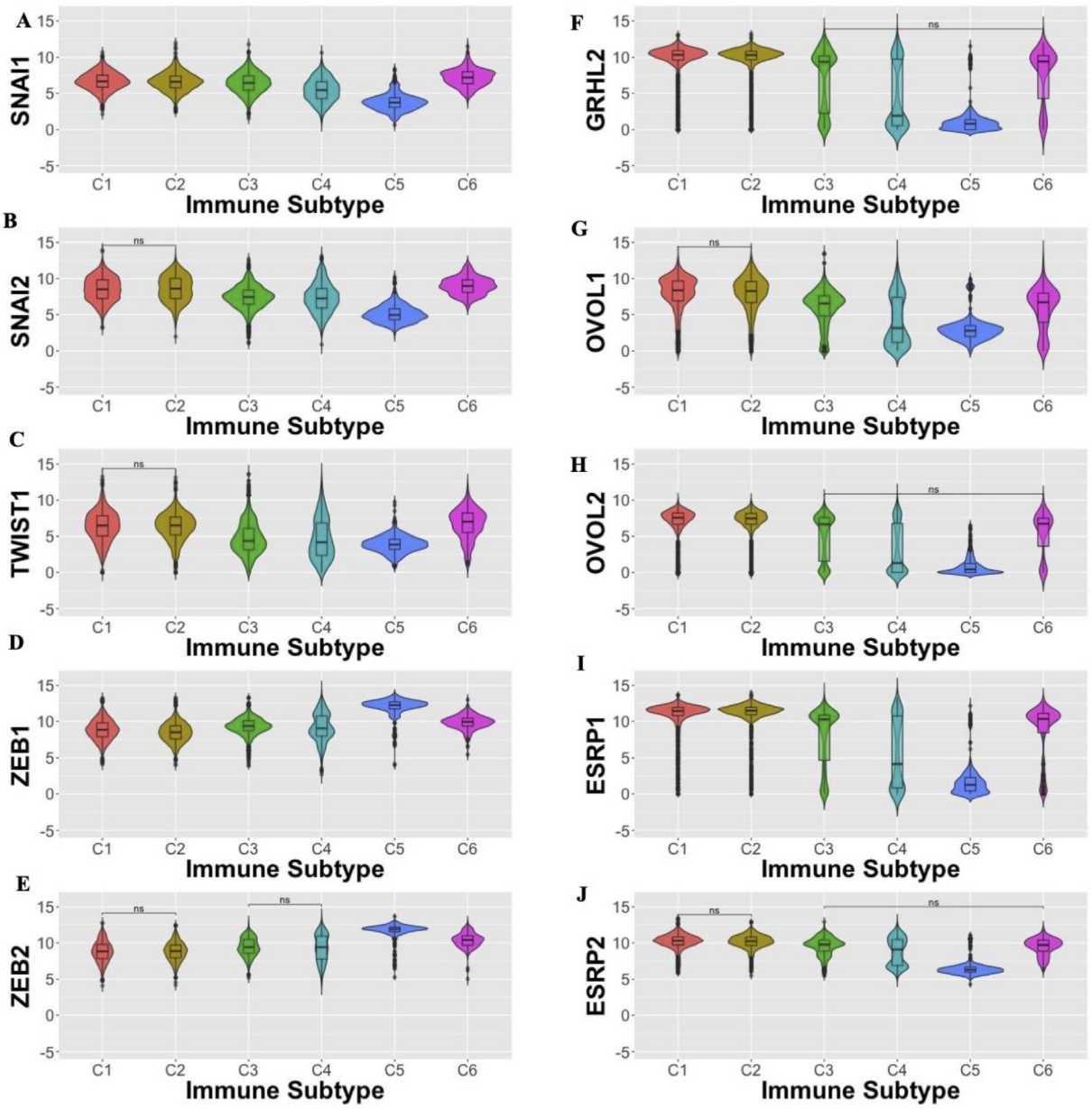


**Fig. 4. EMT and MET factor gene expression varies across immune subtypes. (A-E)** Expression of EMT factors across immune subtypes; **(F-J)** Expression of MET factors across immune subtypes; **(A)** SNAI1, **(B)** SNAI2, **(C)** TWIST1, **(D)** ZEB1, **(E)** ZEB2, **(F)** GRHL2, **(G)** OVOL1, **(H)** OVOL2, **(I)** ESRP1, and **(J)** ESRP2 expression. Unless specified by n.s. (not significant), all comparisons are statistically reliable ($p < 0.05$; Wilcoxon rank means test with Benjamini-Hochberg adjustment for multiple comparisons).

KS and MLR scores followed the inverse trend (**Figure S11**). Among the EMT factors, ZEB1 and ZEB2 had the highest correlations with all EMT scoring metrics (**Fig S11**). The upregulation of ZEB1 and ZEB2 in C5, and the strong correlations of ZEB1 and ZEB2 with EMT scoring metrics further supports observations about the key roles of these two EMT-TFs in maintenance of a mesenchymal phenotype [91].

### *2.4. Immune checkpoint markers across immune subtypes*

We next investigated the levels of different immune checkpoint markers across the immune subtypes: CTLA-4 (cytotoxic T lymphocyte antigen-4), CD274 (PD-L1; programmed death-ligand 1) [92], LAG3 (lymphocyte activation gene 3) [93], CD276 (B7-H3) ([94], CD47 (cluster of differentiation 47) [95], and HAVCR2 (TIM-3) and its ligand LGASL9 (galectin 9) [96] (Fig. 5**A-G**). While expression of most immune checkpoint molecules varies across all immune subtypes, the immunologically quiescent C5 immune subtype has the lowest levels of CTLA-4, CD274, CD276, and LAG3, but with HAVCR2 and CD47 expression similar to other immune subtypes (Fig. 5). The TGF-β-dominant subtype (C6) displays elevated HAVCR2, LGALS9, and CD276 (Fig. 5). The expression levels of these immune checkpoint markers are positively correlated with one another as well as with the single-sample GSEA scores for EMT and partial EMT [65] signatures (Fig. 5**H**). These results suggest that common signalling pathways implicated in EMT may be associated with changes in levels of various immune checkpoint markers.

## 3. Discussion

Applying three distinct RNA-based EMT scoring algorithms – KS, MLR, and GS76 – we characterized the diversity and heterogeneity of EMT status both within and across cancer types. These

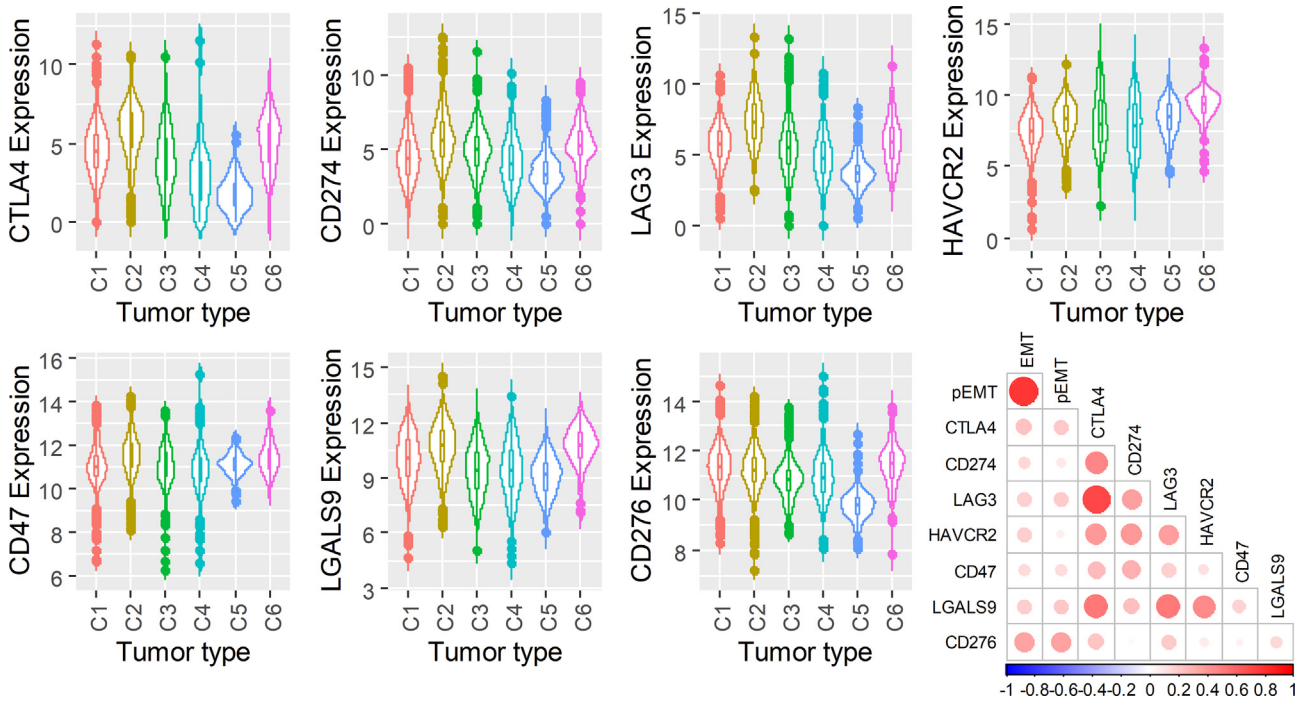


**Fig. 5.** Immune checkpoint expression varies across immune subtypes. **(A)** CTLA4, **(B)** CD274, **(C)** LAG3, **(D)** HAVCR2, **(E)** CD47, **(F)** LGALS9, and **(G)** CD276 expression across immune subtypes. **(H)** Correlation matrix of EMT and pseudo-EMT scores with immune checkpoint markers.

analyses revealed a higher variance in EMT score among cancers of an epithelial lineage as compared to mesenchymally-derived tumors. Similarly, analysis of EMT- and MET-associated factors revealed more heterogeneity in gene expression for EMT factors as compared to MET factors. While these data suggest that the EMT program is more heterogeneous than the MET program across cancers, the reasons for this are not clear. One possible explanation for this may be that the selected MET factors, such as GRHL2, OVOL1, and OVOL2 are indicative of epithelial lineages [97,98,99], while the selected EMT factors may be more unique to a specific cellular lineage. It is possible that greater heterogeneity in the MET program would perhaps be more accurately reflected by analysing tissue-specific cadherins and keratins that mark specific lineages [100]. It is also possible that consideration of other aspects of the cellular response (post-transcriptional, translational/ post-translational, epigenetic [6], metabolic [101,102], etc. could reveal additional layers of regulation and perhaps more conserved circuitry across the MET program.

Another possible explanation may be that the MET program represents a more fixed, derived phenotype while the EMT program is more plastic in nature. This notion is supported by the observations that the EMT program is coupled to cancer stemness-like pathways [103,104,34]. Although it is possible that EMT scores may be skewed by differences in tumor:stromal ratios across samples and cancer types [21,71], our analyses revealed low correlations between EMT score and tumor purity scores [71], suggesting that the EMT scores were not substantially skewed by high stromal content. In addition, the scoring metrics were also consistent when applied to both TCGA and CCLE data, suggesting that these scoring metrics can be applied to both potentially-heterogeneous cancer samples and more homogeneous cancer cell lines. The remarkable heterogeneity in EMT and MET scores and EMT/MET factors across cancers underscores that EMT/MET dynamics are context-specific and more of a spectrum of states instead of a binary classification [2,3].

The present work sheds further light on our collective understanding of the potential cross-talk between cancer cells and immune subsets across the EMT spectrum. Both EMT scores and known EMT/MET factors were associated with certain immune subtypes. For example, cancers with C1 (wound healing) and C2 (IFN-γ dominant) subtypes tended to have more epithelial scores with high expression of MET factors, C5 (quiescent) was dominated by a high mesenchymal score with upregulation of ZEB1 and ZEB2, and C6 (TGF-β dominant) was enriched for a more hybrid epithelial/mesenchymal score. Consistent with this, the TGF-β dominant subtype shows upregulation of ZEB1 and ZEB2 as well as GRHL2, ESRP1, and ESRP2 (Fig. 4). The hybrid E/M state plays an important role in therapeutic resistance and formation of highly tumorigenic CTC clusters [105,98]. TGF-β has been shown to activate ZEB1 through DNA methylation of the ZEB1 repressive microRNA, miR-200 [106]. Our prior work indicated that both GRHL2 and miR200s are necessary to induce MET in sarcomas [6], and it is possible that the TGF-β-mediated subtype additional signals to drive the phenotype toward a more complete EMT. One limitation of this study is that the analysis of TCGA data was from almost exclusively primary tumor sites rather than metastases. Metastatic samples comprise just 3.4% of all TCGA solid tumor samples (https://portal.gdc.cancer.gov/). Given the considerable differences in EMT biology between primary tumors and metastases [27,29,107,108], it is likely that the biology of the immune subsets within metastatic microenvironments also differs.

Analysis of seven immune checkpoint molecules – CTLA-4, CD274/PD-L1, LAG3, CD276/B7-H3, CD47, HAVCR2/TIM-3, and LGAL29/galectin-9 – across the immune subtypes revealed the highest levels among C2 (IFN-γ dominant) and C6 (TGF-β dominant) subtypes and lowest in the C5 (quiescent) subtype. IFN-γ is a potent inducer of PD-L1 expression [109], which may explain the high levels of immune checkpoint molecules in the C2 subtype despite its more epithelial-like score. In addition, epithelial-like cancer cells are known to have higher numbers of infiltrating M1 macrophages and CD8 + T cells in the tumor microenvironment [50], which is consistent with the observation that the C2 (IFN-γ dominant) subtype had the highest M1/M2 macrophage

polarization and strong CD8 signal among all six subtypes [61]. The TGF-β dominant subtype (C6) displayed increased levels of HAVCR2 (also known as Tim-3), and LGALS2 (galectin-9). Interestingly, HAVCR2 expression is upregulated by TGF-β [110,111], and galectin-9 promotes TGF-β-induced signalling and subsequent conversion of CD4 + CD25- T cells into regulatory T-cells [112,113]. These observations suggest that patients enriched for the C6 subtype may benefit from TGF-β inhibitors in combination with HAVCR2 or galectin-9 suppression.

## 4. Methods

All analyses were performed in R (version 3.4.4), and data was plotted using the ggplot2 package. Heatmaps were plotted using the gplots R package.

**TCGA datasets**: TCGA gene expression datasets were obtained from https://xenabrowser.net/datapages/.

**RNA Seq data pre-processing:** Normalized gene expression counts of all the tumor datasets from TCGA website were further pre-processed to ensure a robust approximation of EMT scores. To do this, a regression analysis was used to convert all RNA-Seq count data to a standard EMT metric score based on previous microarray data. The regression parameters to transform RNASeq to microarray data were estimated as described previously [114].

**CCLE dataset**: CCLE gene expression data was downloaded from https://portals.broadinstitute.org/ccle/data

**EMT scoring:** Three different EMT scoring methods – KS, MLR, GS76 – were used to score samples separately in all the datasets as previously described [60]. Unless noted otherwise, the KS score was used for analyses of immune subtypes and EMT/MET factors.

**ssGSEA analysis:** ssGSEA analysis for various different gene sets were performed using GSVA R Bioconductor package with "ssgsea" option for method argument [115].

**Statistical analysis:** All the pairwise comparison significance was tested using student's *t*-test and the multiple group comparisons significance was tested using ANOVA. A Wilcoxon rank means test with Benjamini-Hochberg adjustment for multiple comparisons was used for the analysis of immune subtypes in Fig. 4.

**Min-max standardization**: The gene expression values and EMT scores were standardized in the range of 0 to 1 as following:

Where, $X_{scaled}$ is the min–max standardized value of a gene *X*.

**Principal component analysis:** Principal component analysis was used to visualize the gene expression data of multiple variables (5 EMT and/or 5 MET factors). multidimensional gene expression and simulation data. To determine the correlation between variables and the representation of variables by the principal components, a correlation circle with squared cosines was plotted.

## CRediT authorship contribution statement

**Priyanka Chakraborty:** Conceptualization, Data curation, Formal analysis, Visualization, Writing - original draft, Writing - review & editing. **Emily L. Chen:** Writing - original draft, Writing - review & editing. **Isabelle McMullen:** Conceptualization, Data curation, Formal analysis, Visualization, Writing - original draft, Writing - review & editing. **Andrew J. Armstrong:** Conceptualization, Writing - original draft, Writing - review & editing. **Mohit Kumar Jolly:** Conceptualization, Data curation, Formal analysis, Visualization, Supervision, Writing - original draft, Writing - review & editing. **Jason A. Somarelli:** Conceptualization, Data curation, Formal analysis, Visualization, Supervision, Writing - original draft, Writing - review & editing.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

JAS is supported by the Department of Defense (W81XWH-18-1-0189). AJA and JAS acknowledge funding support from NCI 1R01CA233585-02. MKJ was supported by Ramanujan Fellowship awarded by Science and Engineering Research Board, Department of Science and Technology, Government of India (SB/S2/RJN-049/2018). The results published here are based, in part, upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2021.06.023.

## References

[1] Jolly MK, Ware KE, Gilja S, Somarelli JA, Levine H. EMT and MET: necessary or permissive for metastasis? Mol Oncol 2017;11(7):755–69.
[2] Nieto MA, Huang R-Y-J, Jackson RA, Thiery JP. EMT: 2016. Cell 2016;166(1):21–45.
[3] Yang, J., Antin, P., Berx, G., Blanpain, C., Brabletz, T., Bronner, M., Campbell, K., Cano, A., Casanova, J., Christofori, G., Dedhar, S., Derynck, R., Ford, H. L., Fuxe, J., García de Herreros, A., Goodall, G. J., Hadjantonakis, A.-K., Huang, R. J. Y., Kalcheim, C., ... EMT International Association (TEMTIA). (2020). Guidelines and definitions for research on epithelial-mesenchymal transition. Nature Reviews. Molecular Cell Biology, 21(6), 341–352.
[4] Stone RC, Pastar I, Ojeh N, Chen V, Liu S, Garzon KI, et al. Epithelial-mesenchymal transition in tissue repair and fibrosis. Cell Tissue Res 2016;365(3):495–506.
[5] Gao X, Bali AS, Randell SH, Hogan BLM. GRHL2 coordinates regeneration of a polarized mucociliary epithelium from basal stem cells. The Journal of Cell Biology 2015;211(3):669–82.
[6] Somarelli JA, Shetler S, Jolly MK, Wang X, Bartholf Dewitt S, Hish AJ, et al. Mesenchymal-Epithelial Transition in Sarcomas Is Controlled by the Combinatorial Expression of MicroRNA 200s and GRHL2. Mol Cell Biol 2016;36(19):2503–13.
[7] Roca H, Hernandez J, Weidner S, McEachin RC, Fuller D, Sud S, et al. Transcription factors OVOL1 and OVOL2 induce the mesenchymal to epithelial transition in human cancer. PLoS ONE 2013;8(10):e76773.
[8] Gregory PA, Bracken CP, Bert AG, Goodall GJ. MicroRNAs as regulators of epithelial-mesenchymal transition. Cell Cycle 2008;7(20):3112–8.
[9] Paterson EL, Kolesnikoff N, Gregory PA, Bert AG, Khew-Goodall Y, Goodall GJ. The microRNA-200 family regulates epithelial to mesenchymal transition. TheScientificWorldJournal 2008;8:901–4.
[10] Bebee TW, Park JW, Sheridan KI, Warzecha CC, Cieply BW, Rohacek AM, et al. The splicing regulators Esrp1 and Esrp2 direct an epithelial splicing program essential for mammalian development.. eLife 2015:4. https://doi.org/10.7554/eLife.08954.
[11] Jolly MK, Preca B-T, Tripathi SC, Jia D, George JT, Hanash SM, et al. Interconnected feedback loops among ESRP1, HAS2, and CD44 regulate epithelial-mesenchymal plasticity in cancer. APL Bioengineering 2018;2(3):031908.
[12] Warzecha CC, Sato TK, Nabet B, Hogenesch JB, Carstens RP. ESRP1 and ESRP2 are epithelial cell-type-specific regulators of FGFR2 splicing. Mol Cell 2009;33(5):591–601.
[13] Skrypek N, Goossens S, De Smedt E, Vandamme N, Berx G. Epithelial-to-Mesenchymal Transition: Epigenetic Reprogramming Driving Cellular Plasticity. Trends in Genetics: TIG 2017;33(12):943–59.
[14] Serrano-Gomez SJ, Maziveyi M, Alahari SK. Regulation of epithelial-mesenchymal transition through epigenetic and post-translational modifications. Molecular Cancer 2016;15:18.
[15] Cannito S, Novo E, Compagnone A, Valfrè di Bonzo L, Busletta C, Zamara E, et al. Redox mechanisms switch on hypoxia-dependent epithelial-mesenchymal transition in cancer cells. Carcinogenesis 2008;29(12):2267–78.
[16] Choi B-J, Park S-A, Lee S-Y, Cha YN, Surh Y-J. Hypoxia induces epithelial-mesenchymal transition in colorectal cancer cells through ubiquitin-specific protease 47-mediated stabilization of Snail: A potential role of Sox9. Sci Rep 2017;7(1):15918.

[17] Saxena K, Jolly MK. Acute vs. Chronic vs. Cyclic Hypoxia: Their Differential Dynamics, Molecular Mechanisms, and Effects on Tumor Progression. Biomolecules 2019;9(8). https://doi.org/10.3390/biom9080339.
[18] Yang M-H, Wu M-Z, Chiou S-H, Chen P-M, Chang S-Y, Liu C-J, et al. Direct regulation of TWIST by HIF-1alpha promotes metastasis. Nat Cell Biol 2008;10(3):295–305.
[19] Gao M-Q, Kim BG, Kang S, Choi YP, Park H, Kang KS, et al. Stromal fibroblasts from the interface zone of human breast carcinomas induce an epithelial-mesenchymal transition-like state in breast cancer cells in vitro. J Cell Sci 2010;123(Pt 20):3507–14.
[20] Giannoni E, Bianchini F, Masieri L, Serni S, Torre E, Calorini L, et al. Reciprocal activation of prostate cancer cells and cancer-associated fibroblasts stimulates epithelial-mesenchymal transition and cancer stemness. Cancer Res 2010;70(17):6945–56.
[21] Wang L, Saci A, Szabo PM, Chasalow SD, Castillo-Martin M, Domingo-Domenech J, et al. EMT- and stroma-related gene expression and resistance to PD-1 blockade in urothelial cancer. Nat Commun 2018;9(1):3503.
[22] Wang Y, Lan W, Xu M, Song J, Mao J, Li C, et al. Cancer-associated fibroblast-derived SDF-1 induces epithelial-mesenchymal transition of lung adenocarcinoma via CXCR4/β-catenin/PPARδ signalling. Cell Death Dis 2021;12(2):214.
[23] Fedele V, Melisi D. Permissive State of EMT: The Role of Immune Cell Compartment. Front Oncol 2020;10:587.
[24] Romeo E, Caserta CA, Rumio C, Marcucci F. The Vicious Cross-Talk between Tumor Cells with an EMT Phenotype and Cells of the Immune System. Cells 2019;8(5). https://doi.org/10.3390/cells8050460.
[25] Toh B, Wang X, Keeble J, Sim WJ, Khoo K, Wong W-C, et al. Mesenchymal transition and dissemination of cancer cells is driven by myeloid-derived suppressor cells infiltrating the primary tumor. PLoS Biol 2011;9(9):e1001162.
[26] Esposito M, Mondal N, Greco TM, Wei Y, Spadazzi C, Lin S-C, et al. Bone vascular niche E-selectin induces mesenchymal-epithelial transition and Wnt activation in cancer cells to promote bone metastasis. Nat Cell Biol 2019;21(5):627–39.
[27] Ocaña OH, Córcoles R, Fabra A, Moreno-Bueno G, Acloque H, Vega S, et al. Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer Prrx1. Cancer Cell 2012;22(6):709–24.
[28] Stankic M, Pavlovic S, Chin Y, Brogi E, Padua D, Norton L, et al. TGF-β-Id1 signaling opposes Twist1 and promotes metastatic colonization via a mesenchymal-to-epithelial transition. Cell Reports 2013;5(5):1228–42.
[29] Tsai JH, Donaher JL, Murphy DA, Chau S, Yang J. Spatiotemporal regulation of epithelial-mesenchymal transition is essential for squamous cell carcinoma metastasis. Cancer Cell 2012;22(6):725–36.
[30] Guen VJ, Chavarria TE, Kröger C, Ye X, Weinberg RA, Lees JA. EMT programs promote basal mammary stem cell and tumor-initiating cell stemness by inducing primary ciliogenesis and Hedgehog signaling. PNAS 2017;114(49):E10532–9.
[31] Guo D, Xu B-L, Zhang X-H, Dong M-M. Cancer stem-like side population cells in the human nasopharyngeal carcinoma cell line cne-2 possess epithelial mesenchymal transition properties in association with metastasis. Oncol Rep 2012;28(1):241–7.
[32] Li Y, Wang L, Pappan L, Galliher-Beckley A, Shi J. IL-1β promotes stemness and invasiveness of colon cancer cells through Zeb1 activation. Molecular Cancer 2012;11:87.
[33] Pasani S, Sahoo S, Jolly MK. Hybrid E/M Phenotype(s) and Stemness: A Mechanistic Connection Embedded in Network Topology. J Clin Med Res 2020;10(1). https://doi.org/10.3390/jcm10010060.
[34] Wellner U, Schubert J, Burk UC, Schmalhofer O, Zhu F, Sonntag A, et al. The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs. Nat Cell Biol 2009;11(12):1487–95.
[35] Fischer KR, Durrans A, Lee S, Sheng J, Li F, Wong STC, et al. Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. Nature 2015;527(7579):472–6.
[36] Zheng X, Carstens JL, Kim J, Scheible M, Kaye J, Sugimoto H, et al. Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. Nature 2015;527(7579):525–30.
[37] Yu M, Bardia A, Wittner BS, Stott SL, Smas ME, Ting DT, et al. Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition. Science 2013;339(6119):580–4.
[38] Biswas K, Jolly MK, Ghosh A. Stability and mean residence times for hybrid epithelial/mesenchymal phenotype. Phys Biol 2019;16(2):025003.
[39] Brabletz T. To differentiate or not–routes towards metastasis. Nat Rev Cancer 2012;12(6):425–36.
[40] George JT, Jolly MK, Xu S, Somarelli JA, Levine H. Survival Outcomes in Cancer Patients Predicted by a Partial EMT Gene Expression Scoring Metric. Cancer Res 2017;77(22):6415–28.
[41] Armstrong AJ, Marengo MS, Oltean S, Kemeny G, Bitting RL, Turnbull JD, et al. Circulating tumor cells from patients with advanced prostate and breast cancer display both epithelial and mesenchymal markers. Molecular Cancer Research: MCR 2011;9(8):997–1007.
[42] Brahmer JR, Govindan R, Anders RA, Antonia SJ, Sagorsky S, Davies MJ, et al. The Society for Immunotherapy of Cancer consensus statement on immunotherapy for the treatment of non-small cell lung cancer (NSCLC). J ImmunoTher Cancer 2018;6(1):75.
[43] Weiss SA, Wolchok JD, Sznol M. Immunotherapy of Melanoma: Facts and Hopes. Clinical Cancer Research: An Official Journal of the American Association for Cancer Research 2019;25(17):5191–201.
[44] Rini BI, Battle D, Figlin RA, George DJ, Hammers H, Hutson T, et al. The society for immunotherapy of cancer consensus statement on immunotherapy for the treatment of advanced renal cell carcinoma (RCC). J ImmunoTher Cancer 2019;7(1):354.
[45] Cao Y, Zhang L, Kamimura Y, Ritprajak P, Hashiguchi M, Hirose S, et al. B7–H1 overexpression regulates epithelial-mesenchymal transition and accelerates carcinogenesis in skin. Cancer Res 2011;71(4):1235–43.
[46] Taki M, Abiko K, Baba T, Hamanishi J, Yamaguchi K, Murakami R, et al. Snail promotes ovarian cancer progression by recruiting myeloid-derived suppressor cells via CXCR2 ligand upregulation. Nat Commun 2018;9(1):1685.
[47] Qian L, Liu Y, Wang S, Gong W, Jia X, Liu L, et al. NKG2D ligand RAE1ε induces generation and enhances the inhibitor function of myeloid-derived suppressor cells in mice. J Cell Mol Med 2017;21(9):2046–54.
[48] Kim S, Koh J, Kim M-Y, Kwon D, Go H, Kim YA, et al. PD-L1 expression is associated with epithelial-to-mesenchymal transition in adenocarcinoma of the lung. Hum Pathol 2016;58:7–14.
[49] Chen L, Gibbons DL, Goswami S, Cortez MA, Ahn Y-H, 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.
[50] Dongre A, Rashidian M, Reinhardt F, Bagnato A, Keckesova Z, Ploegh HL, et al. Epithelial-to-Mesenchymal Transition Contributes to Immunosuppression in Breast Carcinomas. Cancer Res 2017;77(15):3982–9.
[51] Jiang Y, Zhan H. Communication between EMT and PD-L1 signaling: New insights into tumor immune evasion. Cancer Lett 2020;468:72–81.
[52] Cao Y, Zhang L, Ritprajak P, Tsushima F, Youngnak-Piboonratanakit P, Kamimura Y, et al. Immunoregulatory molecule B7–H1 (CD274) contributes to skin carcinogenesis. Cancer Res 2011;71(14):4737–41.
[53] Guo Y, Lu X, Chen Y, Rendon B, Mitchell RA, Cuatrecasas M, et al. Zeb1 induces immune checkpoints to form an immunosuppressive envelope around invading cancer cells. *Science*. Advances 2021;7(21). https://doi.org/10.1126/sciadv.abd7455.
[54] OVOL2 inhibits macrophage M2 polarization by regulating IL-10 transcription, and thus inhibits the tumor metastasis by modulating the tumor microenvironment. (2021). Immunology Letters. https://doi.org/10.1016/j.imlet.2021.05.004
[55] Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. Nature 2011;474(7353):609–15.
[56] Ding, L., Bailey, M. H., Porta-Pardo, E., Thorsson, V., Colaprico, A., Bertrand, D., Gibbs, D. L., Weerasinghe, A., Huang, K.-L., Tokheim, C., Cortés-Ciriano, I., Jayasinghe, R., Chen, F., Yu, L., Sun, S., Olsen, C., Kim, J., Taylor, A. M., Cherniack, A. D., … Cancer Genome Atlas Research Network. (2018). Perspective on Oncogenic Processes at the End of the Beginning of Cancer Genomics. Cell, 173(2), 305–320.e10.
[57] Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol 2017;18(1):220.
[58] Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling Tumor Infiltrating Immune Cells with CIBERSORT. Methods Mol Biol 2018;1711:243–59.
[59] Eddy JA, Thorsson V, Lamb AE, Gibbs DL, Heimann C, Yu JX, et al. CRI iAtlas: an interactive portal for immuno-oncology research. F1000Research 2020;9:1028.
[60] Chakraborty P, George JT, Tripathi S, Levine H, Jolly MK. Comparative Study of Transcriptomics-Based Scoring Metrics for the Epithelial-Hybrid-Mesenchymal Spectrum. Front Bioeng Biotechnol 2020;8:220.
[61] Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang T-H, et al. The Immune Landscape of Cancer. Immunity 2018;48(4):812–830.e14.
[62] Bhatia S, Monkman J, Blick T, Pinto C, Waltham M, Nagaraj SH, et al. Interrogation of Phenotypic Plasticity between Epithelial and Mesenchymal States in Breast Cancer. J Clin Med Res 2019;8(6). https://doi.org/10.3390/jcm8060893.
[63] Devaraj V, Bose B. Morphological State Transition Dynamics in EGF-Induced Epithelial to Mesenchymal Transition. J Clin Med Res 2019;8(7). https://doi.org/10.3390/jcm8070911.
[64] Mandal M, Ghosh B, Anura A, Mitra P, Pathak T, Chatterjee J. Modeling continuum of epithelial mesenchymal transition plasticity. Integrative Biology: Quantitative Biosciences from Nano to Macro 2016;8(2):167–76.
[65] Puram SV, Tirosh I, Parikh AS, Patel AP, Yizhak K, Gillespie S, et al. Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. Cell 2017;171(7):1611–1624.e24.
[66] Ruscetti M, Dadashian EL, Guo W, Quach B, Mulholland DJ, Park JW, et al. HDAC inhibition impedes epithelial-mesenchymal plasticity and suppresses metastatic, castration-resistant prostate cancer. Oncogene 2016;35(29):3781–95.
[67] Schliekelman MJ, Taguchi A, Zhu J, Dai X, Rodriguez J, Celiktas M, et al. Molecular portraits of epithelial, mesenchymal, and hybrid States in lung adenocarcinoma and their relevance to survival. Cancer Res 2015;75(9):1789–800.
[68] Tripathi S, Chakraborty P, Levine H, Jolly MK. A mechanism for epithelial-mesenchymal heterogeneity in a population of cancer cells. PLoS Comput Biol 2020;16(2):e1007619.

[69] 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. Clinical Cancer Research: An Official Journal of the American Association for Cancer Research 2013;19(1):279–90.
[70] Tan TZ, Miow QH, Miki Y, Noda T, Mori S, Huang R-Y-J, et al. Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients. EMBO Mol Med 2014;6(10):1279–93.
[71] Aran D, Sirota M, Butte AJ. Systematic pan-cancer analysis of tumour purity. Nat Commun 2015;6:8971.
[72] Drápela S, Bouchal J, Jolly MK, Culig Z, Souček K. ZEB1: A Critical Regulator of Cell Plasticity, DNA Damage Response, and Therapy Resistance. Frontiers in Molecular Biosciences 2020;7:36.
[73] Subbalakshmi AR, Sahoo S, Biswas K, Jolly MK. A Computational Systems Biology Approach Identifies SLUG as a Mediator of Partial Epithelial-Mesenchymal Transition (EMT). Tissues, Organs: Cells; 2021. p. 1–14.
[74] Vandewalle C. SIP1/ZEB2 induces EMT by repressing genes of different epithelial cell-cell junctions. *Nucleic Acids Res* (Vol. 2005;33(20):6566–78. https://doi.org/10.1093/nar/gki965.
[75] 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(7):927–39.
[76] Cieply B, Riley P, Pifer PM, Widmeyer J, Addison JB. Suppression of the epithelial–mesenchymal transition by Grainyhead-like-2. Cancer Res 2012. , https://cancerres.aacrjournals.org/content/72/9/2440.short.
[77] Fici P, Gallerani G, Morel A-P, Mercatali L, Ibrahim T, Scarpi E, et al. Splicing factor ratio as an index of epithelial-mesenchymal transition and tumor aggressiveness in breast cancer. Oncotarget 2017;8(2):2423–36.
[78] Jeong HM, Han J, Lee SH, Park HJ, Lee HJ, Choi JS, et al. ESRP1 is overexpressed in ovarian cancer and promotes switching from mesenchymal to epithelial phenotype in ovarian cancer cellsThis article has been corrected since Advance Online Publication and an erratum is also printed in this issue. Oncogenesis 2017;6(10):e389.
[79] Saxena K, Srikrishnan S, Celia-Terrassa T, Jolly MK. OVOL1/2: Drivers of Epithelial Differentiation in Development, Disease, and Reprogramming. Tissues, Organs: Cells; 2020. p. 1–10.
[80] Mooney SM, Talebian V, Jolly MK, Jia D, Gromala M, Levine H, et al. The GRHL2/ZEB Feedback Loop-A Key Axis in the Regulation of EMT in Breast Cancer. J Cell Biochem 2017;118(9):2559–70.
[81] Zhou JX, Huang S. Understanding gene circuits at cell-fate branch points for rational cell reprogramming. Trends in Genetics: TIG 2011;27(2):55–62.
[82] Hari K, Sabuwala B, Subramani BV, La Porta CAM, Zapperi S, Font-Clos F, et al. Identifying inhibitors of epithelial-mesenchymal plasticity using a network topology-based approach. npj Syst Biol Appl 2020;6(1):15.
[83] Celià-Terrassa T, Bastian C, Liu DD, Ell B, Aiello NM, Wei Y, et al. Hysteresis control of epithelial-mesenchymal transition dynamics conveys a distinct program with enhanced metastatic ability. Nat Commun 2018;9(1):5005.
[84] Warzecha CC, Shen S, Xing Y, Carstens RP. The epithelial splicing factors ESRP1 and ESRP2 positively and negatively regulate diverse types of alternative splicing events. RNA Biol 2009;6(5):546–62.
[85] Ishii H, Saitoh M, Sakamoto K, Kondo T, Katoh R, Tanaka S, et al. Epithelial splicing regulatory proteins 1 (ESRP1) and 2 (ESRP2) suppress cancer cell motility via different mechanisms. The Journal of Biological Chemistry 2014;289(40):27386–99.
[86] Pillai M, Jolly MK. Systems-level network modeling deciphers the master regulators of phenotypic plasticity and heterogeneity in melanoma. bioRxiv 2021. https://doi.org/10.1101/2021.03.11.434533 (p. 2021.03.11.434533).
[87] Chauhan, L., Ram, U., Hari, K., & Jolly, M. K. (n.d.). 2021. Topological signatures in regulatory network enable phenotypic heterogeneity in small cell lung cancer. eLife. https://doi.org/10.1101/2020.10.30.362228
[88] Terry S, Savagner P, Ortiz-Cuaran S, Mahjoubi L, Saintigny P, Thiery J-P, et al. New insights into the role of EMT in tumor immune escape. Mol Oncol 2017;11(7):824–46.
[89] Tripathi SC, Peters HL, Taguchi A, Katayama H, Wang H, Momin A, et al. Immunoproteasome deficiency is a feature of non-small cell lung cancer with a mesenchymal phenotype and is associated with a poor outcome. PNAS 2016;113(11):E1555–64.
[90] Foroutan M, Bhuva DD, Lyu R, Horan K, Cursons J, Davis MJ. Single sample scoring of molecular phenotypes. BMC Bioinf 2018;19(1):404.
[91] Addison JB, Voronkova MA, Fugett JH, Lin C-C, Linville NC, Trinh B, et al. Functional Hierarchy and Cooperation of EMT Master Transcription Factors in Breast Cancer Metastasis. MCR Molecular Cancer Research 2021. https://doi.org/10.1158/1541-7786.MCR-20-0532.
[92] Wei SC, Anang N, Sharma R. Combination anti–CTLA-4 plus anti–PD-1 checkpoint blockade utilizes cellular mechanisms partially distinct from monotherapies. Proceedings of the, 2019.
[93] Solinas, C., Migliori, E., De Silva, P., & Willard-Gallo, K. (2019). LAG3: The Biological Processes That Motivate Targeting This Immune Checkpoint Molecule in Human Cancer. Cancers, 11(8). https://doi.org/10.3390/cancers11081213
[94] Yang S, Wei W, Zhao Q. B7–H3, a checkpoint molecule, as a target for cancer immunotherapy. *International Journal of Biological Sciences* (Vol. 2020;16 (11):1767–73. https://doi.org/10.7150/ijbs.41105.
[95] Liu X, Pu Y, Cron K, Deng L, Kline J, Frazier WA, et al. CD47 blockade triggers T cell–mediated destruction of immunogenic tumors. *Nat Med* (Vol. 2015;21 (10):1209–15. https://doi.org/10.1038/nm.3931.
[96] Holderried TAW, de Vos L, Bawden EG, Vogt TJ, Dietrich J, Zarbl R, et al. Molecular and immune correlates of TIM-3 (HAVCR2) and galectin 9 (LGALS9) mRNA expression and DNA methylation in melanoma. In *Clinical*. Epigenetics 2019;Vol. 11, Issue 1. https://doi.org/10.1186/s13148-019-0752-8.
[97] Hong T, Watanabe K, Ta CH, Villarreal-Ponce A, Nie Q, Dai X. An Ovol2-Zeb1 Mutual Inhibitory Circuit Governs Bidirectional and Multi-step Transition between Epithelial and Mesenchymal States. PLoS Comput Biol 2015;11(11): e1004569.
[98] Jolly MK, Tripathi SC, Jia D, Mooney SM, Celiktas M, Hanash SM, et al. Stability of the hybrid epithelial/mesenchymal phenotype. Oncotarget 2016;7 (19):27067–84.
[99] Lee B, Villarreal-Ponce A, Fallahi M, Ovadia J, Sun P, Yu Q-C, et al. Transcriptional mechanisms link epithelial plasticity to adhesion and differentiation of epidermal progenitor cells. Dev Cell 2014;29(1): 47–58.
[100] Wahl GM, Spike BT. Cell state plasticity, stem cells, EMT, and the generation of intra-tumoral heterogeneity. npj Breast Cancer 2017;3:14.
[101] Giudetti AM, De Domenico S, Ragusa A, Lunetti P, Gaballo A, Franck J, et al. A specific lipid metabolic profile is associated with the epithelial mesenchymal transition program. Biochimica et Biophysica Acta, Molecular and Cell Biology of Lipids 2019;1864(3):344–57.
[102] Ochoa S, de Anda-Jáuregui G, Hernández-Lemus E. Multi-Omic Regulation of the PAM50 Gene Signature in Breast Cancer Molecular Subtypes. Front Oncol 2020;10:845.
[103] Guo W, Keckesova Z, Donaher JL, Shibue T, Tischler V, Reinhardt F, et al. Slug and Sox9 cooperatively determine the mammary stem cell state. Cell 2012;148(5):1015–28.
[104] Mani SA, Guo W, Liao M-J, Eaton EN, Ayyanan A, Zhou AY, et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell 2008;133(4):704–15.
[105] Jolly MK. Implications of the Hybrid Epithelial/Mesenchymal Phenotype in Metastasis. In *Frontiers in Oncology* 2015;Vol. 5. https://doi.org/10.3389/fonc.2015.00155.
[106] Gregory PA, Bracken CP, Smith E, Bert AG, Wright JA, Roslan S, et al. An autocrine TGF-beta/ZEB/miR-200 signaling network regulates establishment and maintenance of epithelial-mesenchymal transition. Mol Biol Cell 2011;22(10):1686–98.
[107] Bonnomet A, Syne L, Brysse A, Feyereisen E, Thompson EW, Noël A, et al. A dynamic in vivo model of epithelial-to-mesenchymal transitions in circulating tumor cells and metastases of breast cancer. Oncogene 2012;31 (33):3741–53.
[108] Somarelli JA, Schaeffer D, Marengo MS, Bepler T, Rouse D, Ware KE, et al. Distinct routes to metastasis: plasticity-dependent and plasticity-independent pathways. Oncogene 2016;35(33):4302–11.
[109] Zou W, Wolchok JD, Chen L. PD-L1 (B7–H1) and PD-1 pathway blockade for cancer therapy: Mechanisms, response biomarkers, and combinations. Sci Transl Med 2016;8(328):328rv4.
[110] Wiener Z, Kohalmi B, Pocza P, Jeager J, Tolgyesi G, Toth S, et al. TIM-3 is expressed in melanoma cells and is upregulated in TGF-beta stimulated mast cells. J Invest Dermatol 2007;127(4):906–14.
[111] Yan W, Liu X, Ma H, Zhang H, Song X, Gao L, et al. Tim-3 fosters HCC development by enhancing TGF-β-mediated alternative activation of macrophages. Gut 2015;64(10):1593–604.
[112] Lv K, Zhang Y, Zhang M, Zhong M, Suo Q. Galectin-9 promotes TGF-β1-dependent induction of regulatory T cells via the TGF-β/Smad signaling pathway. Mol Med Rep 2013;7(1):205–10.
[113] Wu C, Thalhamer T, Franca RF, Xiao S, Wang C, Hotta C, et al. Galectin-9-CD44 interaction enhances stability and function of adaptive regulatory T cells. Immunity 2014;41(2):270–82.
[114] Zhao S, Fung-Leung W-P, Bittner A, Ngo K, Liu X. Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. PLoS ONE 2014;9 (1):e78644.
[115] Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* (Vol. 2009;462(7269):108–12. https://doi.org/10.1038/nature08460.