

# Quantifying the patterns of metabolic plasticity and heterogeneity along the epithelial-hybrid-mesenchymal spectrum in cancer

Running title: EMT-metabolism interconnection

Srinath Muralidharan<sup>1</sup>, Sarthak Sahoo<sup>2</sup>, Aryamaan Saha<sup>1,#</sup>, Sanjay Chandran<sup>1,#</sup>,  
Sauma Suvra Majumdar<sup>3,#</sup>, Herbert Levine<sup>4,\*</sup>, Mohit Kumar Jolly<sup>2,\*</sup>

<sup>1</sup> Department of Biotechnology, Indian Institute of Technology Madras, Chennai, India

<sup>2</sup> Centre for BioSystems Science and Engineering, Indian Institute of Science, Bangalore, India

<sup>3</sup> Department of Biotechnology, National Institute of Technology, Durgapur, India

<sup>4</sup> Departments of Physics and Bioengineering, and Centre for Theoretical Biological Physics, Northeastern University, Boston, USA

#These authors contributed equally.

\*Authors to whom correspondence should be addressed:

[mkjolly@iisc.ac.in](mailto:mkjolly@iisc.ac.in) (M.K.J.), [h.levine@northeastern.edu](mailto:h.levine@northeastern.edu) (H.L.)

## Abstract

Cancer metastasis is the leading cause of cancer-related mortality and the process of Epithelial to Mesenchymal Transition (EMT) is crucial for cancer metastasis. Either a partial or complete EMT have been reported to influence the metabolic plasticity of cancer cells in terms of switching among oxidative phosphorylation, fatty acid oxidation and glycolysis pathways. However, a comprehensive analysis of these major metabolic pathways their associations with EMT across different cancers is lacking. Here, we analyse more than 180 cancer cell datasets and show diverse associations of these metabolic pathways with the EMT status of cancer cells. Our bulk data analysis shows that EMT generally positively correlates with glycolysis but negatively with oxidative phosphorylation and fatty acid metabolism. These correlations are also consistent at the level of their molecular master regulators, namely AMPK and HIF1 $\alpha$ . Yet, these associations are shown to not be universal. Analysis of single-cell data of EMT induction shows dynamic changes along the different axes of metabolic pathways, consistent with general trends seen in bulk samples. Together, our results reveal underlying patterns of metabolic plasticity and heterogeneity as cancer cells traverse through the epithelial-hybrid-mesenchymal spectrum of states.

**Keywords** Cancer metabolism, Epithelial Mesenchymal Transition, Glycolysis, Oxidative phosphorylation, Fatty acid metabolism, HIF1 $\alpha$ , AMPK

## Introduction

The Epithelial to Mesenchymal transition (EMT) is a cellular programme that gives rise to a loss of epithelial phenotypes (reduced expression of cadherins involved in cell-cell attachment, loss of apical-basal polarity) with a concomitant gain in mesenchymal traits such as migration and invasion [1]. This programme has long been known to be involved in embryogenesis and wound healing in adults [2]. Cancer cells are known to activate EMT during the metastatic progression of a tumour, which allows the cells to invade and establish secondary tumours. EMT is not a binary process; instead, cells can acquire and maintain one or more hybrid epithelial/mesenchymal (E/M) states [3]. Plasticity of cancer cells along the epithelial-hybrid-mesenchymal landscape is highly dynamic, complex and multi-dimensional in nature [4]. Often, other relevant biological traits such as immune evasion, stemness, anoikis resistance and therapy resistance are coupled with this dynamic nature of cancer cells [5–10].

The metabolic status of cancer cells has also been reported to be coupled with EMT. Specifically, many EMT-inducing transcription factors (EMT-TFs) regulate the expression of various metabolic genes involved in glucose, lipid, glutamine, amino acid and nucleotide metabolism [11–13]. Furthermore, a change in metabolic state of cancer cells can induce a change in their EMT status [13]. The exact modalities by which these two processes are related are, however, relatively unclear, and have only begun to be investigated. For instance, activation of glycolytic enzymes by EMT has been reported in breast and prostate cancer cells [14]. A similar study reports EMT-driven activation of glycolysis in non-small cell lung cancer cells (NSCLC) by transcriptional activation of glucose transporter 3 (GLUT3) [15]. In the reverse direction, upregulation of glycolysis has been shown to promote stemness and EMT in pancreatic cancer cells [16]. On the other hand, TGF- $\beta$  induced EMT has been shown to inhibit glycolysis and instead activate oxidative phosphorylation (OXPHOS) via the repression of pyruvate dehydrogenase kinase 4 (PDK4) [17], an enzyme that prevents conversion of pyruvate to acetyl-CoA. Overexpressing PDK4 inhibited EMT [17], thus demonstrating mutual regulation of these two axes of plasticity. Similarly, TGF- $\beta$  induced EMT in colon cancer cells was shown to suppress glycolysis by nuclear translocation of pyruvate kinase M2 (PKM2) [18], a cytosolic enzyme required for pyruvate formation. Finally, expression of SNAI1, an EMT-TF, was shown to repress another glycolytic enzyme, fructose-1,6-bisphosphatase (FBP1) [19]. Similar context-dependent trends were observed in the case of the association of lipid metabolism with EMT. While TNF $\alpha$  or TGF- $\beta$  induced EMT activation promoted the lipid synthesis pathway in prostate cancer cells [20], overexpression of SNAI1 can also lead to inactivation of lipid synthesis enzymes [21]. Overall, these studies suggest context-dependency of EMT mediated coupling to metabolic networks.

Additionally, several lipid metabolism enzymes such as acetyl-CoA synthetases (ACSL1 or ACSL4, steroyl-CoA desaturase (SCD) etc.) can activate EMT [22–24]. Notably, two pathways with whose links with EMT have been particularly well studied are glycolysis and mitochondrial metabolism. For instance, the glycolytic enzyme phosphoglucose isomerase (PGI) could also act as a cytokine and activate EMT via ZEB1 and ZEB2 stabilization in breast cancer cells [25]. But, again, this trend may not be ubiquitous across cancer subtypes and different microenvironments. The glycolytic enzyme FBP1, for instance, blocks the induction of SNAI1-driven EMT in breast cancer cells and the loss of this enzyme favours EMT, as shown *in vitro* [26]. Downregulation of several mitochondrial metabolic genes and mutations in TCA cycle enzymes have also been associated with EMT activation. Mutations in fumarate hydratase, an enzyme that converts fumarate to malate in TCA cycle, can induce EMT by inhibiting the activity of miR-200 [27]. Similarly, mutations in the TCA enzymes succinate dehydrogenase (SDH) and isocitrate dehydrogenase (IDH) also induce EMT via epigenetic suppression of miR-200, leading to alterations in the miR200-ZEB1 axis [27,28] that regulates the

EMT status of cells [29]. Moreover, silencing of another TCA cycle enzyme, citrate synthase (CS), induces EMT-like cellular changes *in vitro* and promotes metastasis *in vivo* [30]. However, more recent studies reveal CS to be upregulated in several other tumour types and that its inactivation impedes EMT programme in tumour cell lines [31]. While these studies point towards a causal link between EMT and the metabolic pathways of glycolysis, fatty acid oxidation and oxidative phosphorylation, the overall interconnection landscape among these pathways is quite confounding, thereby necessitating further research.

Here, we sought to analyse the association of three main aspects of cellular metabolism – glycolysis, oxidative phosphorylation and fatty acid synthesis – with the process of EMT in more than 180 publicly available microarray/RNA-seq datasets comprising cell lines and patient tumors. We found that oxidative phosphorylation is predominantly negatively correlated with the process of EMT, and its primary regulator AMPK is primarily correlated (positively) specifically with an epithelial programme. Conversely, glycolysis and its key regulator HIF1 $\alpha$  predominantly positively correlated with a mesenchymal programme and the induction of EMT. However, glycolysis also showed a positive correlation with the epithelial programme in many datasets, highlighting its complex interaction with the EMT programme. Fatty acid oxidation was correlated negatively with acquisition of a mesenchymal phenotype and positively with the epithelial nature of cancer cells. However, alternative modalities of association of metabolic axes with the EMT programme were also observed. Analysis of EMT induction in single cell RNA-seq data showed largely consistent trends with the generic patterns seen in the analysis of bulk samples.

## Results

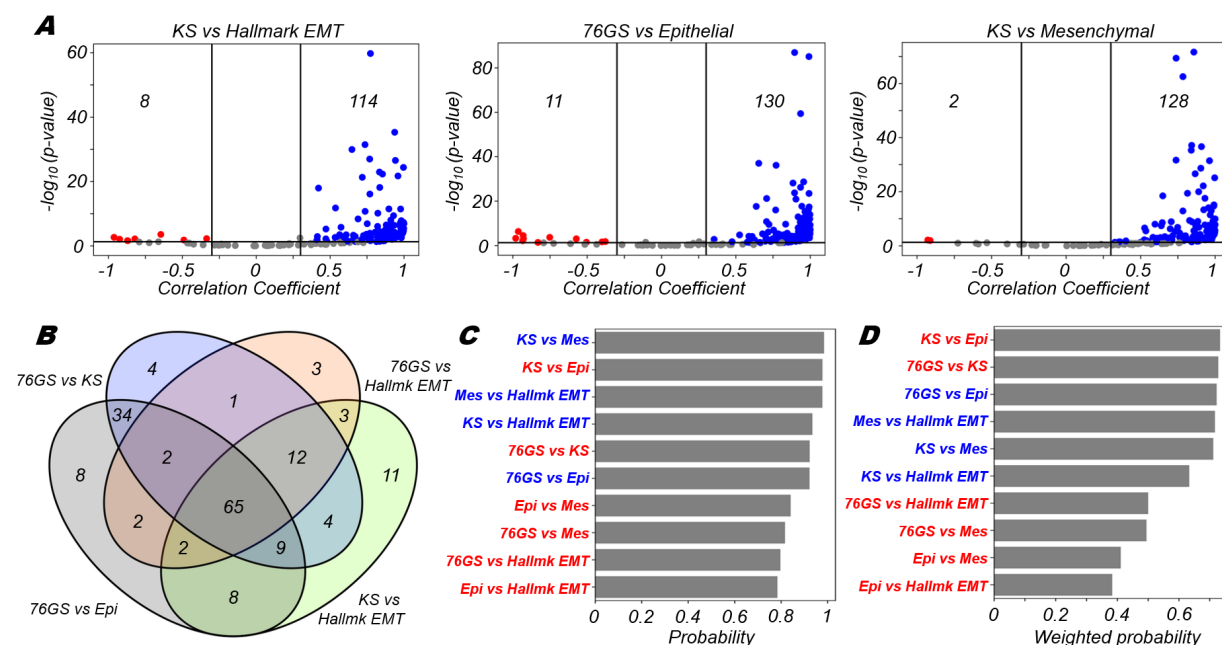
### EMT scoring metrics are largely consistent across datasets

Multiple transcriptomic-based scoring metrics have been used to quantify the EMT status of biological samples [32]. We used 5 different approaches to quantify the EMT status of biological samples in a set of 182 datasets. The 76GS [33,34] and KS [35] EMT scoring methods use two different sets of gene lists (including epithelial and/or mesenchymal genes) to score a sample along the epithelial-hybrid-mesenchymal spectrum. We further used ssGSEA (single-sample Gene Set Enrichment Analysis) and/or Singscore to calculate the activity of epithelial and mesenchymal gene lists (**see Methods**) to estimate the epithelial and mesenchymal nature of the samples respectively.

For 114 out of 182 datasets, the KS score (a higher KS score implies a more mesenchymal state) is positively correlated with the enrichment of Hallmark EMT signature; in only 8 datasets, this correlation is significantly negative (**Fig 1A, left**). On the other hand, the 76GS EMT score (a higher 76GS score indicates a more epithelial state) largely correlates with epithelial signature (**Fig 1A, middle**). The KS metric was also positively correlated with an independent mesenchymal signature (**Fig 1A, right**). A comparison of 4 representative pairs of metrics show that in most datasets we looked at, these EMT scores were correlated significantly and consistently, including a negative correlation between the 76GS and KS scores, as expected (**Fig 1B**). These results show that all these EMT scoring metrics are largely consistent with one another.

Next, we wanted to quantify the consistency of pairs of metrics if they were significantly correlated. To quantify that trend, we computed a “probability” score for a given pair of metrics by considering the number of datasets correlated significantly ( $p < 0.05$ ) either positively ( $r > 0.3$ ) or negatively ( $r < -0.3$ ).

We computed a ratio of number of positively or negatively correlated datasets (depending on the trend seen) to the total number of datasets that showed a significant ( $p < 0.05$ ) association (irrespective of the direction of association). The higher this ratio is, the better the concordance between these two metrics in a given direction. We found that a) KS score vs. Mesenchymal, b) KS score vs. Epithelial and c) Mesenchymal signature vs. Hallmark EMT signatures were most consistent with one another in positive, negative and positive directions respectively (**Fig 1C**). When these probabilities were further weighted by a fraction of significant cases out of all datasets considered, we saw that a) KS score vs. Epithelial, b) 76GS score vs. KS score and c) 76GS vs. Epithelial were most consistent with each other in negative, negative and positive directions respectively, as expected (**Fig 1D**). KS score vs. Mesenchymal and KS score vs. Hallmark EMT correlations also maintained their trends as seen in earlier scenario (compare **Fig 1D** with **Fig 1C**). Put all together, these results show that the EMT metrics considered here are highly consistent with one another in a majority of the datasets.



**Fig 1: Consistency between different EMT scoring metrics. (A)** Volcano plots depicting the Pearson correlation coefficient and the  $-\log_{10}(p\text{-value})$  for 3 pairs of EMT scoring metrics – KS vs Hallmark EMT, 76GS vs epithelial and KS vs mesenchymal. Vertical boundaries are set at correlation coefficient -0.3 and 0.3. The cut-off for  $p\text{-value}$  is set at 0.05. **(B)** 4-way Venn diagram for comparison of 4 representative pairs of EMT scoring metrics. **(C)** Probability of a dataset having a positive (blue) or a negative (red) correlation (correlation coefficient  $> 0.3$ ) given that it is significant ( $p\text{-value} < 0.05$ ) for different pairs of EMT scoring metrics. **(D)** Probability of a dataset having a positive (blue) or a negative (red) correlation (correlation coefficient  $> 0.3$ ) given that it is significant ( $p\text{-value} < 0.05$ ) weighted by the fraction of significant cases for different pairs of EMT scoring metrics.

## **OXPPOS is more likely to be negatively correlated with a mesenchymal program and positively with an epithelial one**

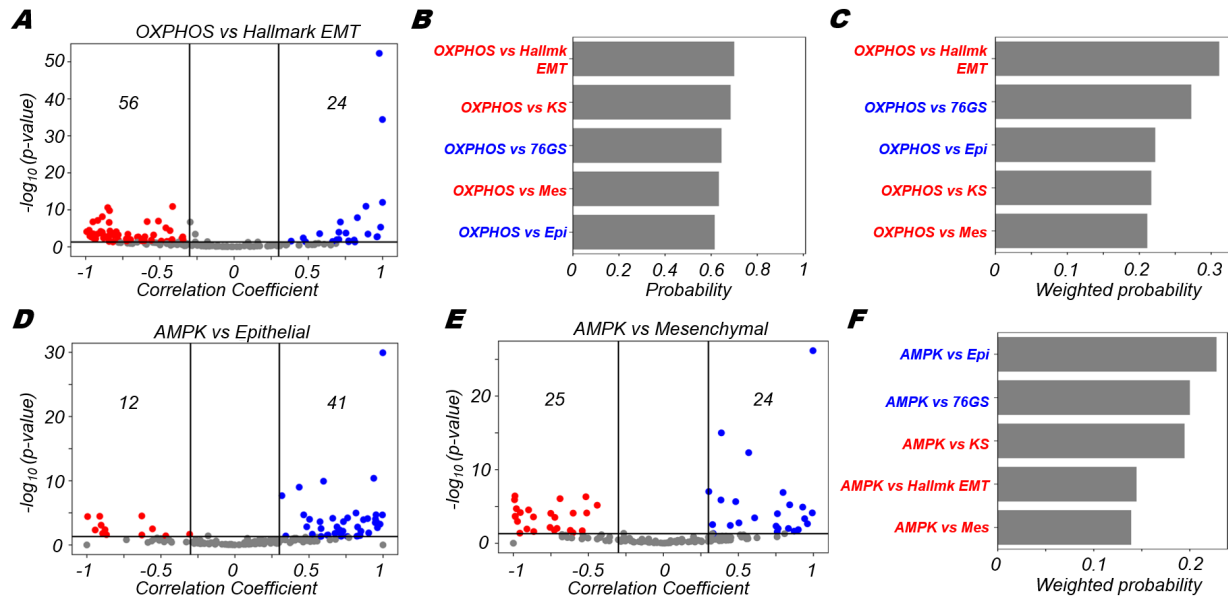
Having assessed the consistency of different EMT scoring metrics among themselves on a cohort of datasets, we next wanted to understand how the different axes of metabolism associated with EMT metrics. Oxidative phosphorylation is a predominant way by which cells generate energy for survival. To study how the biological process of oxidative phosphorylation associates with different EMT metrics, we correlated the ssGSEA activity scores calculated for the Hallmark oxidative phosphorylation gene set with different EMT metrics.

Upon correlating the OXPPOS signature with the Hallmark EMT scores, we found that although there were significant correlations both in the positive and negative directions, there were many more datasets correlated negatively (56 vs. 24) with Hallmark EMT than were correlated positively (**Fig 2A**). This overall observation is in accordance with many experimental studies that point towards a negative association between oxidative phosphorylation and EMT [36–39]. However, this relationship is not exclusive, i.e. these quantities could be positively correlated in a variety of contexts, as reported in other experimental studies [40–42]. Among all pairs of correlations between OXPPOS signature and EMT metrics, the OXPPOS-Hallmark EMT pair showed the strongest propensity of negative association with one another, given that the correlation was significant in either direction (**Fig 2B**). Furthermore, the OXPPOS-Hallmark EMT pair was also the top scoring pair when weighted with the fraction of significant cases (**Fig 2C**), further highlighting the finding that the acquisition of the mesenchymal features was more likely to result in the decline in activity of the OXPPOS gene set.

Oxidative phosphorylation in cells has been reported to be positively regulated by AMPK activity levels in cells. To assess AMPK activity, we considered a list of AMPK target genes that have been used as a proxy for the activity of the phosphorylated active form of AMPK (**see Methods**). We find that the AMPK signature is more likely to be positively correlated with epithelial signature (41 datasets showed positive correlations vs. 12 showed negative correlations) (**Fig 2D**). However, the AMPK signature did not show a strong skew towards being either positively or negatively correlated with the separate mesenchymal signature (**Fig 2E**). Together, these trends could indicate towards the fact that the active form of AMPK is likely more strongly correlated with the presence of an epithelial signature rather than with the absence of a mesenchymal one, especially if we deconvolute EMT into two-dimensional process where loss of epithelial traits and gain of mesenchymal traits can be at treated semi-independently.

Quantifying the trends of association of the AMPK signature with various EMT metrics, we noticed that the probability of positive correlation between AMPK and epithelial metrics (Epithelial signature, 76GS scores) was higher than that of a negative correlation between AMPK and mesenchymal ones (KS score, Hallmark EMT, Mesenchymal signature) (**Fig 2F**). These observations suggest that AMPK is strongly coupled with epithelial traits of cells, rather than with their mesenchymal ones. However, we noticed OXPPOS is strongly negatively correlated with both Hallmark EMT signature as well as the Mesenchymal signature (**Fig 2A, 2B**). This difference seen between trends of AMPK and OXPPOS can be due to additional context-specific factors, apart from AMPK, that might also mediate the crosstalk between EMT and OXPPOS [43], thus leading to an overall stronger negative association of OXPPOS with the hallmark EMT program.





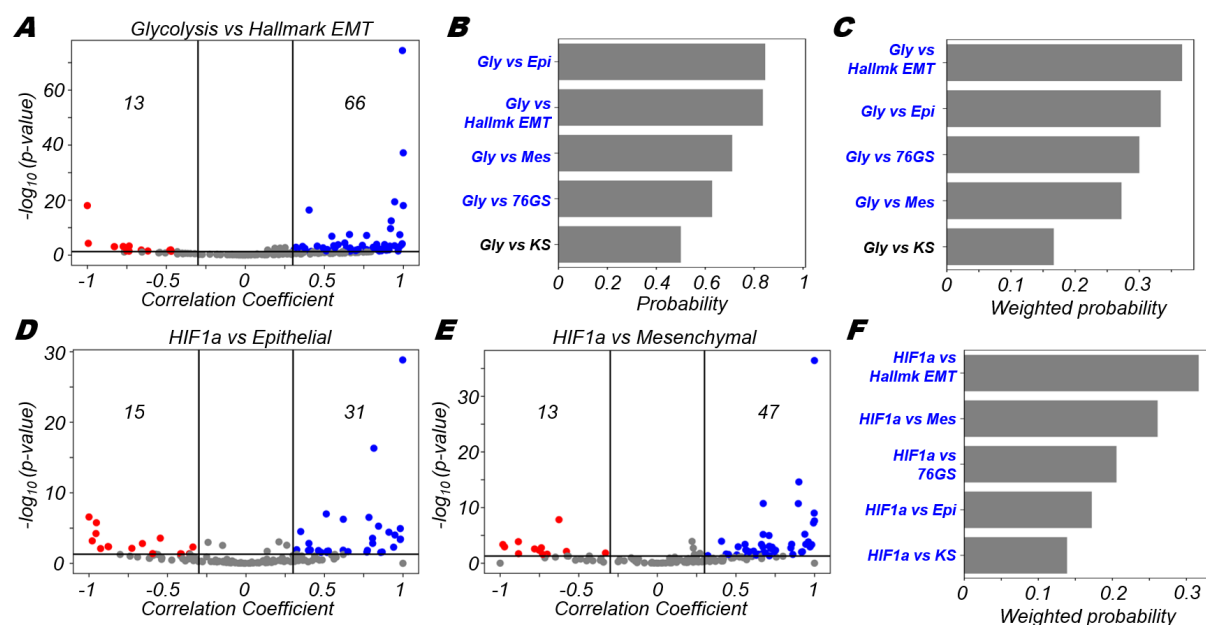
**Fig 2: OXPPOS is more likely to correlate negatively with EMT.** (A) Volcano plots depicting the Pearson correlation coefficient and the  $-\log_{10}(p\text{-value})$  for hallmark OXPPOS and hallmark EMT signatures. (B) Probability of a dataset having a positive (blue) or a negative (red) correlation (correlation coefficient  $> 0.3$ ) given that it is significant ( $p\text{-value} < 0.05$ ) for OXPPOS and different EMT scoring metrics. (C) Probability of a dataset having a positive (blue) or a negative (red) correlation (correlation coefficient  $> 0.3$ ) given that it is significant ( $p\text{-value} < 0.05$ ) weighted by the fraction of significant cases for OXPPOS and different EMT scoring metrics. Volcano plots depicting the Pearson correlation coefficient and the  $-\log_{10}(p\text{-value})$  for (D) AMPK signature and epithelial signature, (E) AMPK signature and mesenchymal signature. (F) Probability of a dataset having a positive (blue) or a negative (red) correlation (correlation coefficient  $> 0.3$ ) given that it is significant ( $p\text{-value} < 0.05$ ) weighted by the fraction of significant cases for AMPK signature and different EMT scoring metrics. Vertical boundaries for volcano plots are set at correlation coefficient  $-0.3$  and  $0.3$ . The cut-off for  $p\text{-value}$  is set at  $0.05$ .

## Glycolysis is more likely to be positively correlated with a (partial) EMT programme.

Next, we wanted to check how the glycolytic process was associated with the EMT programme in the datasets we had considered. To assess this association, we correlated the enrichment (ssGSEA) scores for hallmark EMT and hallmark glycolysis signatures, across our datasets. We observed that glycolysis was more likely to be significantly positively correlated with EMT than being significantly negatively correlated (66 vs. 13 respectively) (**Fig 3A**). One would therefore expect that glycolysis should be negatively correlated with the epithelial programme or with the 76GS EMT scoring metric that assign epithelial samples a higher score. This is, however, not what we observed. Instead, glycolysis was also found to be positively correlated with scoring metrics that report an enriched epithelial program (Epithelial gene list, as well as 76GS scores), to a comparable extent with which it correlated with a mesenchymal program (**Fig 3B**). Similar trends are also seen when the association probability values were further weighted by number of significant datasets in which a given trend was observed (**Fig 3C**). Here, the association of glycolysis with Hallmark EMT programme was stronger, albeit not

to a large degree, than that seen for the epithelial gene list and 76GS score, suggesting its putative association with a partial EMT state and/or other context-specific factors not included in our analysis.

HIF1 $\alpha$  is a known mediator of the glycolytic pathway [44] and in modulating the EMT status of cells [45]. Thus, next we probed how the HIF1 $\alpha$  signature associated with epithelial and mesenchymal programmes. Intriguingly, we found that both the volcano plots showed a skew towards the positive side (**Fig 3D-E**), suggesting that HIF1 $\alpha$  activation may associate with a partial EMT state exhibiting both epithelial and mesenchymal features [46]. It should be noted that in the case of mesenchymal programme, the HIF1 $\alpha$  signature was somewhat more strongly skewed towards the positive side in comparison to the positive skew present in the case of epithelial programme (47 out of 60 datasets vs. 31 datasets out of 46 respectively) (**Fig 3D-E**). The positive association of Glycolysis as well as its known regulator HIF1 $\alpha$  with both the epithelial and mesenchymal axes may indicate that glycolysis is a hallmark feature of hybrid E/M states. Recent observations about glycolysis accompanying collective cell migration endorse this association of glycolytic shift in partial EMT state(s) [47,48]; however, how metabolic heterogeneity maps on to leader-follower dynamic switching remains to be investigated in more detail [49–51]. Nevertheless, stronger trends as measured by weighted probability scores for HIF1 $\alpha$  vs. Hallmark EMT and HIF1 $\alpha$  vs. Mesenchymal compared to HIF1 $\alpha$  vs. 76GS or HIF1 $\alpha$  vs. Epithelial indicates enrichment of HIF1 $\alpha$  in being associated with a relatively more mesenchymal phenotype (**Fig 3F**). The degree of coupling of gain of mesenchymal with loss of epithelial traits in a given scenario [52] may play a key role in associating HIF1 $\alpha$  with a partial or complete EMT.

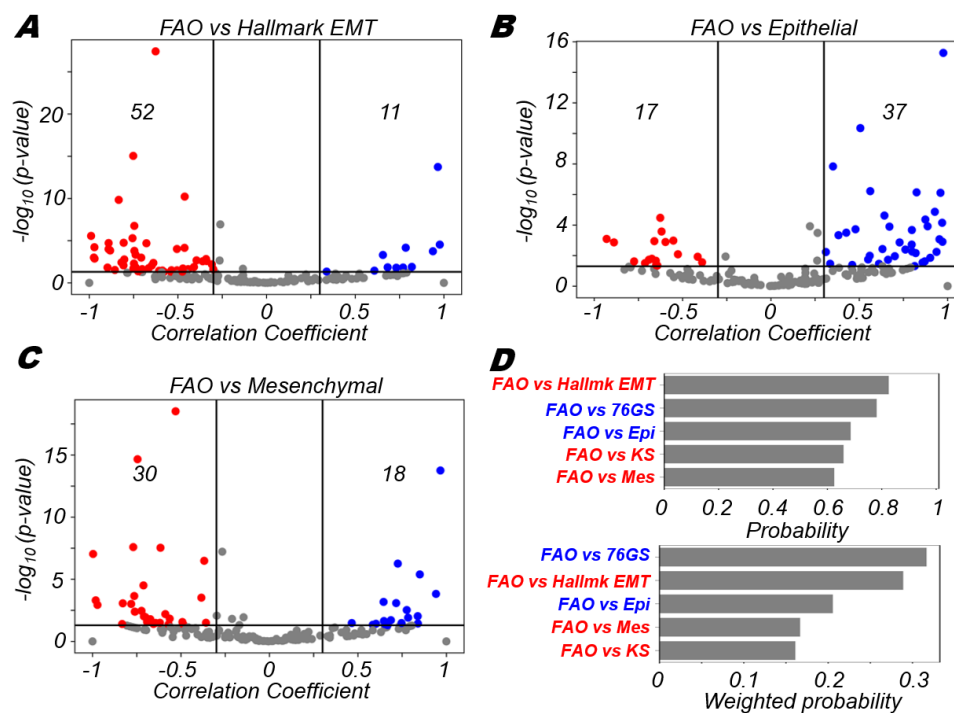


**Fig 3: Glycolysis is more likely to correlate positively with EMT.** (A) Volcano plots depicting the Pearson correlation coefficient and the  $-\log_{10}(p\text{-value})$  for hallmark glycolysis and hallmark EMT signatures. (B) Probability of a dataset having a positive (blue) or a negative (red) correlation (correlation coefficient > 0.3) given that it is significant ( $p\text{-value} < 0.05$ ) for glycolysis and different EMT scoring metrics. (C) Probability of a dataset having a positive (blue) or a negative (red) correlation (correlation coefficient > 0.3) given that it is significant ( $p\text{-value} < 0.05$ ) weighted by the fraction of significant cases for glycolysis and different EMT scoring metrics. Volcano plots depicting the Pearson correlation coefficient and the  $-\log_{10}(p\text{-value})$  for (D) HIF1a signature and epithelial signature, (E) HIF1a signature and mesenchymal signature.

**(F)** Probability of a dataset having a positive (blue) or a negative (red) correlation (correlation coefficient > 0.3) given that it is significant ( $p$ -value < 0.05) weighted by the fraction of significant cases for HIF1a signature and different EMT scoring metrics. Vertical boundaries for volcano plots are set at correlation coefficient -0.3 and 0.3. The cut-off for  $p$ -value is set at 0.05.

### FAO is more likely to positively correlate with an epithelial program and negatively with a mesenchymal program

Fatty acid oxidation (FAO) is a catabolic process in which fatty acids are broken down and is another key mechanism by which cancer cells can generate energy for survival [53]. Genes involved in fatty acid oxidation have been characterised previously [53]. We used one such gene set as a proxy for the activity of the FAO pathway in our datasets (see **Methods**). We found that as with OXPHOS, FAO was most likely to be negatively correlated with the Hallmark EMT programme (52 significantly negative vs. 11 significantly positive cases) (**Fig 4A**). The epithelial programme alone was more likely to be correlated positively (37 positive vs. 17 negative) (**Fig 4B**) while the mesenchymal programme alone was likely to correlate negatively (30 negative vs. 18 positive) (**Fig 4C**). These results show that FAO more likely associates negatively with acquisition of a mesenchymal phenotype. Upon calculation of the probability of positive/negative correlations given the correlation is significant as well as the overall weighted probability, we observed that the mesenchymal metrics (hallmark EMT, mesenchymal signature and KS) were all skewed towards the negative side, while the more epithelial metrics (76GS and epithelial signature) were positively correlated with FAO (**Fig 4D**). These results collectively show that while OXPHOS and FAO are more likely to be associated negatively with the mesenchymal programme and the process of EMT, glycolysis is more likely to be positively associated with the mesenchymal characteristics of cells. These associations, at least in part, are supported by the activity of molecular regulators such as AMPK and HIF1 $\alpha$ .



**Fig 4: Fatty acid oxidation is more likely to correlate negatively with EMT. (A)** Volcano plots depicting the Pearson correlation coefficient and the  $-\log_{10}(p\text{-value})$  for fatty acid oxidation (FAO) and hallmark EMT



signatures. **(B)** Probability of a dataset having a positive (blue) or a negative (red) correlation (correlation coefficient  $> 0.3$ ) given that it is significant ( $p$ -value  $< 0.05$ ) for FAO and different EMT scoring metrics. **(C)** Probability of a dataset having a positive (blue) or a negative (red) correlation (correlation coefficient  $> 0.3$ ) given that it is significant ( $p$ -value  $< 0.05$ ) weighted by the fraction of significant cases for FAO and different EMT scoring metrics. Volcano plots depicting the Pearson correlation coefficient and the  $-\log_{10}(p\text{-value})$  for **(D)** Probability of a dataset having a positive (blue) or a negative (red) correlation (correlation coefficient  $> 0.3$ ) given that it is significant ( $p$ -value  $< 0.05$ ) (top panel) and weighted by the fraction of significant cases (bottom panel) for FAO signature and different EMT scoring metrics. Vertical boundaries for volcano plots are set at correlation coefficient  $-0.3$  and  $0.3$ . The cut-off for  $p$ -value is set at  $0.05$ .

## Different modalities of association between pairs of metabolic pathways and EMT

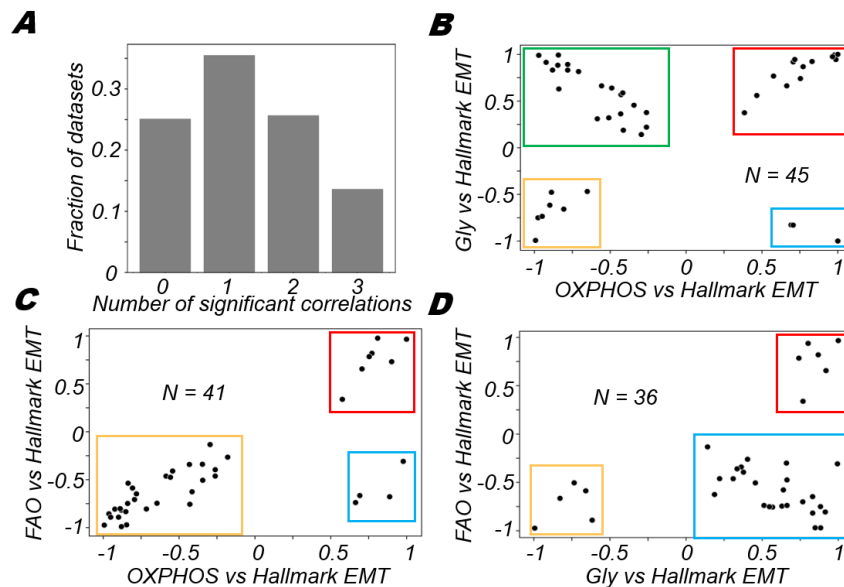
After exploring the three major metabolic axes (OXPHOS, glycolysis and FAO) independently in terms of their association with EMT, we wanted to investigate the pair wise associations between the metabolic pathways in the context of EMT. For datasets under consideration, we first computed the fractions of datasets that had none, one, two or all three axes of metabolism associated significantly with the hallmark EMT signature (**Fig 5A**). Most datasets (~60%) had a maximum of one axis of metabolism correlated with the Hallmark EMT programme (**Fig 5A**). In about 25% of the datasets, hallmark EMT was not correlated with any of the metabolic axes, probably indicative of biological scenarios where these metabolic axes are not coupled directly with the EMT spectrum. In the remaining 40% of datasets, where two or more than two axes were significantly correlated with the hallmark EMT signature, we investigated if certain combinations of associations were more likely than others in context of their correlations with EMT. To answer this question, we first plotted all 45 datasets that had significant correlations with the EMT axis and either OXPHOS or glycolysis (**Fig 5B**). Among those, 21 (46.67%) datasets showed a positive correlation between glycolysis and hallmark EMT programme, while OXPHOS was negatively correlated with the hallmark EMT programme (**Fig 5B, green box**). This co-occurring association of OXPHOS and glycolysis with EMT in inverse directions has been reported earlier experimentally [54,55].

However, this co-occurrence is not the only mode of association between these three axes. The next most predominant modality of association is the scenario where both glycolysis and OXPHOS are positively correlated with EMT – this trend is shown in 14 (31.11%) datasets (**Fig 5B, red box**). This could be indicative of the EMT associating positively with a hybrid metabolic state in which both OXPHOS and glycolysis are high [54]. The other two case – both OXPHOS and glycolysis correlating negatively with EMT (**Fig 5B, yellow box**) and OXPHOS being positively associated while glycolysis being negatively associated (**Fig 5B, blue box**) – were 15.55% and 6.67% respectively. Collectively, this analysis shows that besides predominant modalities of association between OXPHOS and glycolysis in the context of EMT, other modalities also exist although less frequently.

Next, we wanted to explore how FAO and OXPHOS are associated with each other in the context of their correlations with EMT. Upon plotting scatter plots, similar to what we had done for glycolysis and OXPHOS, we find that only 3 quadrants are populated with different propensities (**Fig 5C**). The most predominant modality was the scenario in which EMT was negatively correlated with both OXPHOS and FAO in 30 (73.17%) datasets (**Fig 5C, yellow box**). In 7 (17.07%) datasets, OXPHOS and FAO were both positively correlated with the hallmark EMT signature (**Fig 5C, red box**). In the remaining four datasets, OXPHOS was positively correlated with EMT while FAO was significantly negatively

correlated with EMT (**Fig 5C, blue box**). These results show that FAO and OXPHOS are more likely to be coordinated in a similar manner – either positively or negatively correlated to EMT.

Similarly, when we compared glycolysis with FAO, we found that the major modality of action was the scenario where glycolysis was positively correlated with EMT while FAO was negatively correlated with the EMT programme (**Fig 5D, blue box**). This association was observed in 25 (69.44%) datasets. The other two observed modalities of regulation were the cases where both FAO and glycolysis were both positively (**Fig 5D, red box**) or both negatively (**Fig 5D, yellow box**) correlated with EMT. Overall, our analysis uncovers the different possibilities and propensities by which these three axes of metabolism might associate with one another in terms of their connection with EMT.

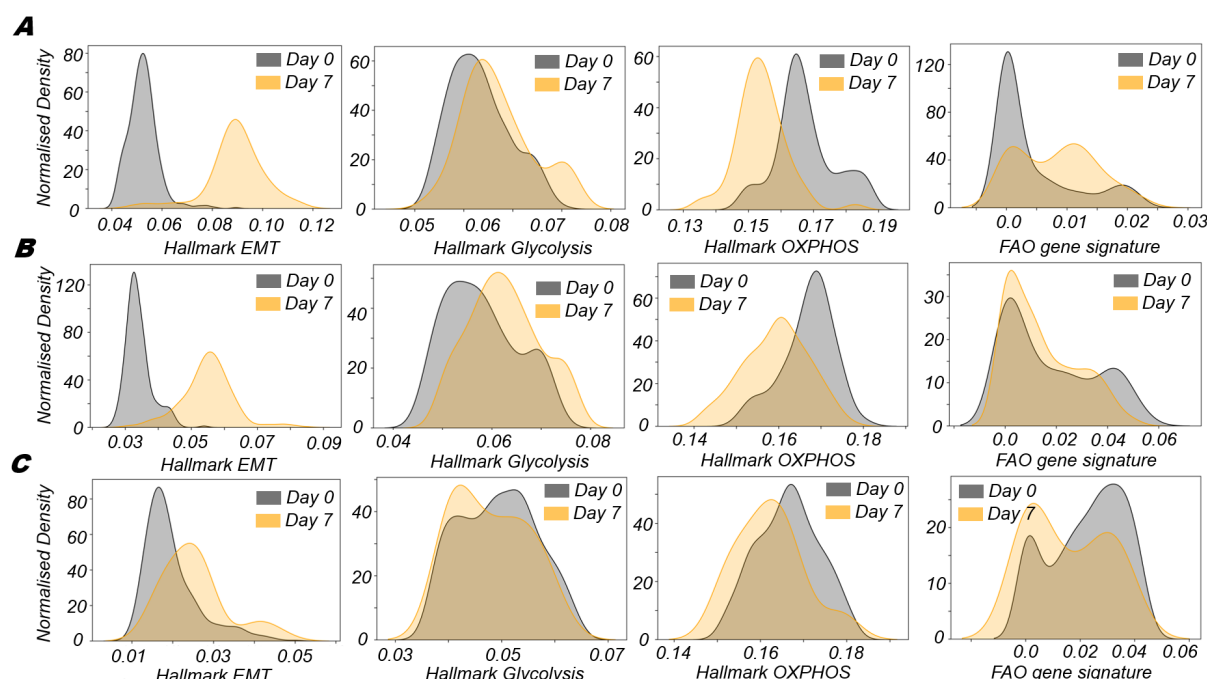


**Fig 5: Varied associations between different metabolic axes with EMT.** (A) Proportion of datasets that have a given number of metabolic axis significantly correlated with the Hallmark EMT signature ( $p$ -value < 0.05). Scatter plot of correlation coefficients of (B) OXPHOS with EMT and glycolysis with EMT (C) hallmarks glycolysis with EMT and glycolysis with EMT (D) glycolysis with EMT and fatty acid oxidation with EMT.

## Heterogeneity in associations between different axes of metabolism in relation to EMT is also reflected in single cell RNA-seq data

Until now, our analysis was focused on bulk samples. We next examined whether there was evidence for the observed heterogeneity of association of metabolic axes with EMT present in single-cell data as well. To do so, we analysed single-cell data across different cell lines induced to undergo EMT by TGF $\beta$  [56], to ask a) if there was a shift along any of the metabolic axis upon induction of EMT, and b) whether the different modalities of associations were seen across different biological conditions. For the cell line A549, as EMT was induced for 7 days, there was a distinct rise in the hallmarks EMT scores of the cells at day 7 compared to day 0 (**Fig 6A**). Concomitantly, there was a significant rise in the levels of glycolysis and FAO, while there was a drop in the levels of OXPHOS (**Fig 6A**). In the case of DU145 cells, as the hallmarks EMT scores of cells increased at day 7 compared to day 0, a significant increase in the level of glycolysis, but a significant shift towards reduced levels of OXPHOS as well as FAO was noted, as expected from the most dominant trend seen in bulk datasets (**Fig 6B**). In the case

of MCF7 cells, there was no significant change in glycolysis levels at day 0 and day 7. However, there was significant drop in the levels of both OXPHOS and FAO (**Fig 6C**). These results demonstrate that some level of heterogeneity in the modalities by which pairs of metabolic axes associate contingent upon the biological context, in this case the cell line considered and extent to which EMT was induced.



**Fig 6: Metabolic signatures in single cells RNA-seq data upon EMT induction.** Kernel density estimate plots of hallmark EMT and metabolic signatures – glycolysis, OXPHOS and fatty acid oxidation – at day 0 (untreated) and day 7 (upon TGFβ treatment) for different cell lines (A) A549 (B) DU145 (C) MCF7

## Discussion

Metabolic reprogramming in cancer cells is a key step in the adaptation and survival of cancer cells in the changing milieu of the tumour microenvironment. Metabolic reactions are more likely to act on a smaller time scale compared to transcriptional and translational processes. This difference in time scales makes metabolic remodelling an attractive mode for instantaneous adaptation for cancer cells. However, changes to the metabolic programmes in cells also happens at a time scale longer than such immediate adaptations. Long term changes in metabolic programmes can happen due to cross talk with other dynamic biological process in cells. In the context of EMT, cells switch from an epithelial to a more mesenchymal phenotype through multiple intermediate states, which facilitate metastasis. Changes in cellular motility and stresses arising from new microenvironments during metastasis may require more energy to adapt and survive and thus necessitate altered metabolism. Such metabolic alterations that affect the overall energy balance in the cell ultimately determine its fitness. Thus, it is not surprising that EMT and metabolism have been shown to influence one another [54,57].

In this study, we have focused on three major energy producing metabolic processes – glycolysis, oxidative phosphorylation (OXPHOS) and fatty acid oxidation (FAO). Cancer cells typically facilitate glycolysis as their primary energy source, irrespective of the presence of oxygen [58]. This process is

referred to as Warburg effect or aerobic glycolysis. Conversely, the role of oxidative phosphorylation has also been observed to be important in cancer cells and cannot be ignored [59,60]. EMT induced metabolic alterations have been an active field of research with the identification of numerous mechanisms by which the metabolic state of cells is altered. Several EMT-inducing signals and EMT-TFs have been shown to activate glycolysis [15,19,61]. Further, glycolysis has been shown to promote EMT in turn, thus forming a positive feedback loop [62–65]. In addition, several studies have shown that EMT-TFs can also inhibit mitochondrial respiration and oxidative phosphorylation [66,67]. Thus, EMT has been consistently shown to be associated with activation of glycolysis and inhibition of OXPHOS, as also noted in scenarios of EMT induction [36]. However, in contrast to these studies, other pieces of evidence suggest that cancer cells with activated EMT may also have increased levels of OXPHOS in some cases [21,60]. Such conflicting findings [21,60] regarding glycolysis vs. OXPHOS may be due to differing tumor microenvironments in these studies or due to differences in cell lines/patient samples used. Another possible explanation is that cancer cells may exhibit a hybrid metabolic phenotype [68] where both glycolysis and OXPHOS states may co-exist which allows high metabolic adaptability.

Consistent with the existing literature, our analysis of more than 180 gene expression datasets reveals that the process of EMT is more likely to be positively associated with the glycolytic process and negatively with OXPHOS programme. We also report here that FAO is also likely to be negatively associated with EMT progression, similar to OXPHOS. These broad trends are, however, not binding, but rather probabilistic in nature given different biological contexts. In other words, the absence of a universal rule indicates that cancer cells may favour glycolysis or oxidative metabolism depending on factors present in the tumor microenvironment (TME) such as availability of glucose, hypoxia, reactive oxygen species, etc. Such an ability to shift the metabolic balance dynamically may provide an advantage amidst shifting energy demands inherent in an evolving TME. Upon analysis of pairs of metabolic pathways in the context of their associations with EMT we observed that while glycolysis and OXPHOS are more likely to be antagonistic in their associations with EMT, OXPHOS and FAO were more likely to be both associated negatively with EMT. The other modalities of associations were also observed albeit in lower propensities. The observed heterogeneities were also seen in single cell RNA-seq data upon the induction of EMT.

Transcriptomic analysis of metabolic genes for given pathways, as done here, is one of the ways to estimate the level of activity of a pathway and its corresponding associations with the EMT programme. However, analysis of metabolomics data would give a more precise picture of the actual metabolic state of cells. Furthermore, there is a need to better characterise the molecular players and associated mechanisms that could allow for heterogeneity in the various modalities of associations between the different metabolic pathways and if there exist feedback loops/networks that might allow a switch from one modality of association to another. Identifying such mechanistic basis would be important to develop this understanding for any therapeutic strategies. Based on the current study, we cannot comment if the observed associations are likely to hold in the context of EMT induction only or also hold in the context of mesenchymal to epithelial transition (MET). EMT/MET dynamics has been shown to be hysteretic (non-symmetric) in nature [69–71]; whether that feature extends to metabolic reprogramming remains to be seen.

Despite these limitations, our work sheds light upon underlying patterns in terms of metabolic plasticity and heterogeneity along the epithelial-hybrid-mesenchymal spectrum in cancer cells. Understanding this coupling between EMT/MET and metabolic plasticity will enable effective targeting of cells in heterogeneous tumor populations.

## Materials and Methods

### **Software and Datasets**

Python (version 3.6) and R (version 4.0.2) were used for conducting all computational and statistical analyses. Microarray datasets were downloaded from NCBI GEO using *GEOquery* R Bioconductor package. FASTQ files for RNA sequencing datasets were downloaded from the ENA (European Nucleotide Archive) database. A complete list of datasets used, and the associated metadata has been provided in Supplementary Table 1.

### **Pre-processing of Microarray datasets**

Gene wise expression for each sample was obtained after appropriate pre-processing of microarray datasets. Probe wise expression matrices downloaded using *GEOquery* were log2 normalised and annotation files corresponding to microarray platforms were utilized for mapping the probes to respective genes. In cases where more than one probe mapped on to a single gene, the mean of expression values of all these probes was used for such genes.

### **Pre-processing of RNA-seq datasets**

Adapter contamination and overall quality of sequences were inspected using FastQC. Sequences were aligned with the hg38 human (or mm10 mouse) reference genome using the STAR alignment software. Finally, the raw counts for each gene were calculated with these aligned sequences using htseq-count. These raw counts were then normalised for gene length and transformed to TPM (transcripts per million) values which were then log2 normalised to obtain the final values.

### **EMT scoring methods**

EMT scores were calculated using five different methods for each dataset. Each method requires gene expression data as input. Each method uses a distinct gene set or a distinct algorithm.

**76GS:** The 76-gene EMT scoring method (76GS) was developed using transcriptomic data from NSCLC cell lines and patient samples [33,34]. As the name suggests, it utilizes 76 gene signatures. The weighted sum of gene expression values of 76 genes was calculated for each sample, where the weight factors are correlation coefficients with CDH1 levels. The values obtained through this method have no specific range. EMT score for each sample was subtracted by the mean obtained from all samples such that the resultant mean score was zero. As per this new scale, negative scores indicate a M phenotype, and positive scores indicate an E phenotype.

**KS:** KS method uses the two-sample Kolmogorov Smirnov test (KS) to score EMT for cell lines and tumor samples [35]. It uses 218 gene signatures for cell line samples and 315 gene signatures for tumor samples. Briefly, cumulative distribution functions (CDFs) are obtained for each of the two signatures (E and M) and the maximum distance between these CDFs is used as the test statistic for a two-sample KS test to obtain the EMT scores. The final EMT scores lie in the range [-1, 1]. Positive & negative scores represent mesenchymal and epithelial phenotypes respectively.

**Hallmark EMT:** This method uses the hallmark geneset for EMT available (Supplementary Table 2) in the MSigDB [72] repository. For each sample, ssGSEA (single sample gene set enrichment analysis) [73] analysis was performed using this geneset to obtain the normalized enrichment score (NES). All calculations were done using the GSEAPY python library.



*Epithelial and Mesenchymal scores:* These metrics use the KS epithelial and mesenchymal gene signatures to quantify E & M status separately. A rank-based single sample scoring method called Singscore [74] was used for quantifying the enrichment level of these gene sets in a given sample. The final value obtained from this method has a range of [-1, 1]. For the epithelial score, a higher value indicates a more epithelial phenotype. The mesenchymal score also operates in a similar manner.

*Metabolic pathways scoring methods:* ssGSEA scores were calculated using the hallmark oxidative phosphorylation and glycolysis gene sets (MSigDB) to obtain OXPHOS and glycolysis signatures respectively (Supplementary Table 2). AMPK and HIF-1 signatures were quantified using expression levels of their downstream target genes as previously reported [68]. A total of 33 downstream genes for AMPK and 23 downstream genes for HIF-1 were used. The final scores were obtained using the Singscore method [74] performed on these gene sets. The FAO scores were calculated based on equations previously reported [75] which uses expression levels of 14 FAO enzyme genes.

**Conflicts of Interest** The authors declare no conflict of interest.

**Acknowledgements** This work was supported by Ramanujan Fellowship (SB/S2/RJN-049/2018) awarded to MKJ by Science and Engineering Research Board (SERB), Department of Science and Technology, Government of India. HL was supported by National Science Foundation sponsored Center for Theoretical Biological Physics – award PHY-2019745, and by PHY-1605817. We are thankful to Prof. Karthik Raman (IIT Madras) for helpful discussions.

**Author contributions** MKJ and HL conceived research; MKJ supervised research; SM, SS, AS, SC and SSM performed research. SM, SS, MKJ and HL analysed data and wrote the manuscript. All authors have read and approved this version of the manuscript.

## Supplementary Table Legends

**Table S1:** Details of 184 transcriptomic datasets included in the analysis obtained from the NCBI GEO repository.

**Table S2:** List of various gene signatures used in the study for quantifying activities of the EMT program and the metabolic pathways.

## References

1. Kalluri, R.; Weinberg, R. A. The basics of epithelial-mesenchymal transition. *J. Clin. Invest.* **2009**, *119*, 1420–1428.
2. Thiery, J. P.; Acloque, H.; Huang, R. Y. J.; Nieto, M. A. Epithelial-mesenchymal transitions in development and disease. *Cell* **2009**, *139*, 871–90.
3. Celià-Terrassa, T.; Jolly, M. K. Cancer Stem Cells and Epithelial-to-Mesenchymal Transition in Cancer Metastasis. *Cold Spring Harb. Perspect. Med.* **2020**, *10*, a036905.
4. Jolly, M. K.; Ware, K. E.; Gilja, S.; Somarelli, J. A.; Levine, H. EMT and MET : necessary or permissive for metastasis ? *Mol. Oncol.* **2017**, *11*, 755–769.
5. Sahoo, S.; Nayak, S. P.; Hari, K.; Purkait, P.; Mandal, S.; Kishore, A.; Levine, H.; Jolly, M. K. Immunosuppressive traits of the hybrid epithelial/mesenchymal phenotype. *Front. Immunol.* **2021**, in press.
6. Huang, R. Y.-J.; Wong, M. K.; Tan, T. Z.; Kuay, K. T.; Ng, a H. C.; Chung, V. Y.; Chu, Y.-S.; Matsumura, N.; Lai, H.-C.; Lee, Y. F.; Sim, W.-J.; Chai, C.; Pietschmann, E.; Mori, S.; Low, J. J. H.; Choolani, M.; Thiery, J. P. An EMT spectrum defines an anoikis-resistant and spheroidogenic intermediate mesenchymal state that is sensitive to e-cadherin restoration by a src-kinase inhibitor, saracatinib (AZD0530). *Cell Death Dis.* **2013**, *4*, e915.
7. Sahoo, S.; Mishra, A.; Kaur, H.; Hari, K.; Muralidharan, S.; Mandal, S.; Jolly, M. K. A mechanistic model captures the emergence and implications of non-genetic heterogeneity and reversible drug resistance in ER+ breast cancer cells. *NAR Cancer* **2021**, *3*, zcab027.
8. Kröger, C.; Afeyan, A.; Mraz, J.; Eaton, E. N.; Reinhardt, F.; Khodor, Y. L.; Thiru, P.; Bieri, B.; Ye, X.; Burge, C. B.; Weinberg, R. A. Acquisition of a hybrid E / M state is essential for tumorigenicity of basal breast cancer cells. *Proc Natl Acad Sci U S A* **2019**, *116*, 7353–7362.
9. Terry, S.; Savagner, P.; Ortiz-Cuaran, S.; Mahjoubi, L.; Saintigny, P.; Thiery, J. P.; Chouaib, S. New insights into the role of EMT in tumor immune escape. *Mol. Oncol.* **2017**, *11*, 824–846.
10. Bocci, F.; Levine, H.; Onuchic, J. N.; Jolly, M. K. Deciphering the Dynamics of Epithelial-Mesenchymal Transition and Cancer Stem Cells in Tumor Progression. *Curr. Stem Cell Reports* **2019**, *5*, 11–21.
11. Payen, V. L.; Porporato, P. E.; Baselet, B.; Sonveaux, P. Metabolic changes associated with tumor metastasis, part 1: Tumor pH, glycolysis and the pentose phosphate pathway. *Cell. Mol. Life Sci.* **2016**, *73*, 1333–48.
12. Porporato, P. E.; Payen, V. L.; Baselet, B.; Sonveaux, P. Metabolic changes associated with tumor metastasis, part 2: Mitochondria, lipid and amino acid metabolism. *Cell. Mol. Life Sci.* **2016**, *73*, 1349–63.
13. Sciacovelli, M.; Frezza, C. Metabolic reprogramming and epithelial-to-mesenchymal transition in cancer. *FEBS J.* **2017**, *284*, 3132–44.
14. Shiraishi, T.; Verdone, J. E.; Huang, J.; Kahlert, U. D.; Hernandez, J. R.; Torga, G.; Zarif, J. C.; Epstein, T.; Gatenby, R.; McCartney, A.; Elisseeff, J. H.; Mooney, S. M.; An, S. S.; Pienta, K. J. Glycolysis is the primary bioenergetic pathway for cell motility and cytoskeletal remodeling in human prostate and breast cancer cells. *Oncotarget* **2015**, *6*, 130–43.
15. Masin, M.; Vazquez, J.; Rossi, S.; Groeneveld, S.; Samson, N.; Schwalie, P. C.; Deplancke, B.; Frawley, L. E.; Gouttenoire, J.; Moradpour, D.; Oliver, T. G.; Meylan, E. GLUT3 is induced during epithelial-mesenchymal transition and promotes tumor cell proliferation in non-small cell lung cancer. *Cancer Metab.* **2014**, *2*, 11.
16. Zhao, H.; Duan, Q.; Zhang, Z.; Li, H.; Wu, H.; Shen, Q.; Wang, C.; Yin, T. Up-regulation of glycolysis promotes the stemness and EMT phenotypes in gemcitabine-resistant pancreatic cancer cells. *J. Cell. Mol. Med.* **2017**, *21*, 2055–2067.

17. Sun, Y.; Daemen, A.; Hatzivassiliou, G.; Arnott, D.; Wilson, C.; Zhuang, G.; Gao, M.; Liu, P.; Boudreau, A.; Johnson, L.; Settleman, J. Metabolic and transcriptional profiling reveals pyruvate dehydrogenase kinase 4 as a mediator of epithelial-mesenchymal transition and drug resistance in tumor cells. *Cancer Metab* **2014**, 2, 20.
18. Hamabe, A.; Konno, M.; Tanuma, N.; Shima, H.; Tsunekuni, K.; Kawamoto, K.; Nishida, N.; Koseki, J.; Mimori, K.; Gotoh, N.; Yamamoto, H.; Doki, Y.; Mori, M.; Ishii, H. Role of pyruvate kinase M2 in transcriptional regulation leading to epithelial-mesenchymal transition. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, 111, 15526–31.
19. Dong, C.; Yuan, T.; Wu, Y.; Wang, Y.; Fan, T. W. M.; Miriyala, S.; Lin, Y.; Yao, J.; Shi, J.; Kang, T.; Lorkiewicz, P.; St Clair, D.; Hung, M. C.; Evers, B. M.; Zhou, B. P. Loss of FBP1 by snail-mediated repression provides metabolic advantages in basal-like breast cancer. *Cancer Cell* **2013**, 23, 316–331.
20. Dalmau, N.; Jaumot, J.; Tauler, R.; Bedia, C. Epithelial-to-mesenchymal transition involves triacylglycerol accumulation in DU145 prostate cancer cells. *Mol. Biosyst.* **2015**, 11, 3397–406.
21. Jiang, L.; Xiao, L.; Sugiura, H.; Huang, X.; Ali, A.; Kuro-O, M.; DeBerardinis, R. J.; Boothman, D. A. Metabolic reprogramming during TGF $\beta$ 1-induced epithelial-to-mesenchymal transition. *Oncogene* **2015**, 34, 3908–16.
22. Sánchez-Martínez, R.; Cruz-Gil, S.; de Cedrón, M. G.; Álvarez-Fernández, M.; Vargas, T.; Molina, S.; García, B.; Herranz, J.; Moreno-Rubio, J.; Reglero, G.; Pérez-Moreno, M.; Feliu, J.; Malumbres, M.; de Molina, A. R. A link between lipid metabolism and epithelial-mesenchymal transition provides a target for colon cancer therapy. *Oncotarget* **2015**, 6, 38719–36.
23. Hanai, J. I.; Doro, N.; Sasaki, A. T.; Kobayashi, S.; Cantley, L. C.; Seth, P.; Sukhatme, V. P. Inhibition of lung cancer growth: ATP citrate lyase knockdown and statin treatment leads to dual blockade of mitogen-activated protein Kinase (MAPK) and Phosphatidylinositol-3-kinase (PI3K)/AKT pathways. *J. Cell. Physiol.* **2012**, 227, 1709–20.
24. Xu, Y.; Huang, J.; Xin, W.; Chen, L.; Zhao, X.; Lv, Z.; Liu, Y.; Wan, Q. Lipid accumulation is ahead of epithelial-to-mesenchymal transition and therapeutic intervention by acetyl-CoA carboxylase 2 silence in diabetic nephropathy. *Metabolism*. **2014**, 63, 716–26.
25. Ahmad, A.; Aboukameel, A.; Kong, D.; Wang, Z.; Sethi, S.; Chen, W.; Sarkar, F. H.; Raz, A. Phosphoglucose isomerase/autocrine motility factor mediates epithelial-mesenchymal transition regulated by miR-200 in breast cancer cells. *Cancer Res.* **2011**, 71, 3400–9.
26. Li, J.; Wang, Y.; Li, Q. G.; Xue, J. J.; Wang, Z.; Yuan, X.; Tong, J. D.; Xu, L. C. Downregulation of FBP1 promotes tumor metastasis and indicates poor prognosis in gastric cancer via regulating epithelial-mesenchymal transition. *PLoS One* **2016**, 11, e0167857.
27. Sciacovelli, M.; Gonçalves, E.; Johnson, T. I.; Zecchini, V. R.; Da Costa, A. S. H.; Gaude, E.; Drubbel, A. V.; Theobald, S. J.; Abbo, S. R.; Tran, M. G. B.; Rajeeve, V.; Cardaci, S.; Foster, S.; Yun, H.; Cutillas, P.; Warren, A.; Gnanapragasam, V.; Gottlieb, E.; Franze, K.; Huntly, B.; Maher, E. R.; Maxwell, P. H.; Saez-Rodriguez, J.; Frezza, C. Fumarate is an epigenetic modifier that elicits epithelial-to-mesenchymal transition. *Nature* **2016**, 537, 544–547.
28. Grassian, A. R.; Lin, F.; Barrett, R.; Liu, Y.; Jiang, W.; Korpai, M.; Astley, H.; Gitterman, D.; Henley, T.; Howes, R.; Levell, J.; Korn, J. M.; Pagliarini, R. Isocitrate dehydrogenase (IDH) mutations promote a reversible ZEB1/MicroRNA (miR)-200-dependent epithelial-mesenchymal transition (EMT). *J. Biol. Chem.* **2012**, 287, 42180–94.
29. Drápela, S.; Bouchal, J.; Jolly, M. K.; Culig, Z.; Souček, K. ZEB1: A Critical Regulator of Cell Plasticity, DNA Damage Response, and Therapy Resistance. *Front. Mol. Biosci.* **2020**, 7, 36.
30. Lin, C. C.; Cheng, T. L.; Tsai, W. H.; Tsai, H. J.; Hu, K. H.; Chang, H. C.; Yeh, C. W.; Chen, Y. C.; Liao, C. C.; Chang, W. T. Loss of the respiratory enzyme citrate synthase directly links the Warburg effect to tumor malignancy. *Sci. Rep.* **2012**, 2, 785.

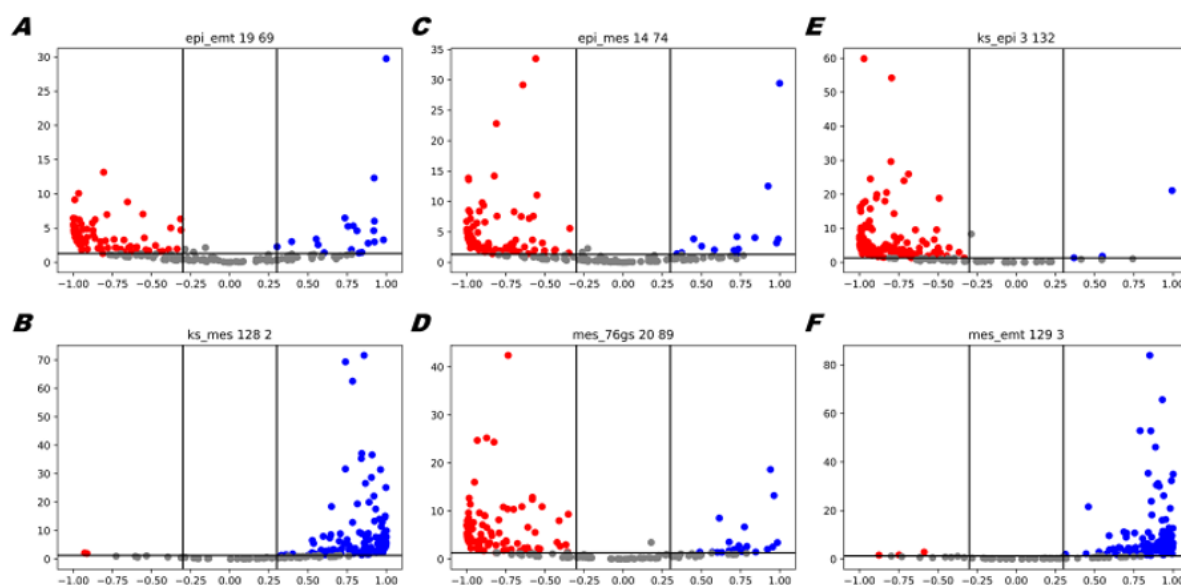
31. Chen, L.; Liu, T.; Zhou, J.; Wang, Y.; Wang, X.; Di, W.; Zhang, S. Citrate synthase expression affects tumor phenotype and drug resistance in human ovarian carcinoma. *PLoS One* **2014**, *9*, e115708.
32. Chakraborty, P.; George, J. T.; Tripathi, S.; Levine, H.; Jolly, M. K. Comparative study of transcriptomics-based scoring metrics for the epithelial-hybrid-mesenchymal spectrum. *Front. Bioeng. Biotechnol.* **2020**, *8*, 220.
33. Guo, C. C.; Majewski, T.; Zhang, L.; Yao, H.; Bondaruk, J.; Wang, Y.; Zhang, S.; Wang, Z.; Lee, J. G.; Lee, S.; Cogdell, D.; Zhang, M.; Wei, P.; Grossman, H. B.; Kamat, A.; Duplisea, J. J.; Ferguson, J. E.; Huang, H.; Dadhania, V.; Gao, J.; Dinney, C.; Weinstein, J. N.; Baggerly, K.; McConkey, D.; Czerniak, B. Dysregulation of EMT Drives the Progression to Clinically Aggressive Sarcomatoid Bladder Cancer. *Cell Rep.* **2019**, *27*, 1781-1793.e4.
34. Byers, L. A.; Diao, L.; Wang, J.; Saintigny, P.; Girard, L.; Peyton, M.; Shen, L.; Fan, Y.; Giri, U.; Tumula, P. K.; Nilsson, M. B.; Gudikote, J.; Tran, H.; Cardnell, R. J. G.; Bearss, D. J.; Warner, S. L.; Foulks, J. M.; Kanner, S. B.; Gandhi, V.; Krett, N.; Rosen, S. T.; Kim, E. S.; Herbst, R. S.; Blumenschein, G. R.; Lee, J. J.; Lippman, S. M.; Ang, K. K.; Mills, G. B.; Hong, W. K.; Weinstein, J. N.; Wistuba, I. I.; Coombes, K. R.; Minna, J. D.; Heymach, J. V. 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*, 279–290.
35. Tan, T. Z.; Miow, Q. H.; Miki, Y.; Noda, T.; Mori, S.; Huang, R. Y.; Thiery, J. P. Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients. *EMBO Mol. Med.* **2014**, *6*, 1279–1293.
36. Sun, H.; Yang, X.; Liang, L.; Zhang, M.; Li, Y.; Chen, J.; Wang, F.; Yang, T.; Meng, F.; Lai, X.; Li, C.; He, J.; He, M.; Xu, Q.; Li, Q.; Lin, L.; Pei, D.; Zheng, H. Metabolic switch and epithelial–mesenchymal transition cooperate to regulate pluripotency. *EMBO J.* **2020**, *39*, e102961.
37. Abd-El-Raouf, R.; Ouf, S. A.; Gabr, M. M.; Zakaria, M. M.; El-Yasergy, K. F.; Ali-El-Dein, B. Escherichia coli foster bladder cancer cell line progression via epithelial mesenchymal transition, stemness and metabolic reprogramming. *Sci. Rep.* **2020**, *10*, 18024.
38. Zhao, S.; Zhang, X.; Shi, Y.; Cheng, L.; Song, T.; Wu, B.; Li, J.; Yang, H. MIEF2 over-expression promotes tumor growth and metastasis through reprogramming of glucose metabolism in ovarian cancer. *J. Exp. Clin. Cancer Res.* **2020**, *39*, 286.
39. Shu, D. Y.; Butcher, E. R.; Saint-Geniez, M. Suppression of pgc-1 $\alpha$  drives metabolic dysfunction in tgfb2-induced emt of retinal pigment epithelial cells. *Int. J. Mol. Sci.* **2021**, *22*, 4701.
40. Liu, Q. Q.; Huo, H. Y.; Ao, S.; Liu, T.; Yang, L.; Fei, Z. Y.; Zhang, Z. Q.; Ding, L.; Cui, Q. H.; Lin, J.; Yu, M.; Xiong, W. TGF- $\beta$ 1-induced epithelial-mesenchymal transition increases fatty acid oxidation and OXPHOS activity via the p-AMPK pathway in breast cancer cells. *Oncol. Rep.* **2020**, *44*, 1206–1215.
41. Hu, Y.; Xu, W.; Zeng, H.; He, Z.; Lu, X.; Zuo, D.; Qin, G.; Chen, W. OXPHOS-dependent metabolic reprogramming prompts metastatic potential of breast cancer cells under osteogenic differentiation. *Br. J. Cancer* **2020**, *123*, 1644–1655.
42. Soukupova, J.; Malfettone, A.; Bertran, E.; Hernández-Alvarez, M. I.; Peñuelas-Haro, I.; Dituri, F.; Giannelli, G.; Zorzano, A.; Fabregat, I. Epithelial–mesenchymal transition (Emt) induced by tgfb in hepatocellular carcinoma cells reprograms lipid metabolism. *Int. J. Mol. Sci.* **2021**, *22*, 5543.
43. Bhattacharya, D.; Scimè, A. Metabolic Regulation of Epithelial to Mesenchymal Transition: Implications for Endocrine Cancer. *Front. Endocrinol. (Lausanne)*. **2019**, *10*, 773.
44. Semenza, G. L. HIF-1 mediates metabolic responses to intratumoral hypoxia and oncogenic mutations. *J. Clin. Invest.* **2013**, *123*, 3664–3671.
45. Saxena, K.; Jolly, M. K.; Balamurugan, K.; Signaling, D. Hypoxia , partial EMT and collective migration : emerging culprits in metastasis. *Transl. Oncol.* **2020**, *13*, 100845.

46. Saxena, K.; Jolly, M. K. Acute vs. Chronic vs. cyclic hypoxia: Their differential dynamics, molecular mechanisms, and effects on tumor progression. *Biomolecules* **2019**, *9*, 339.
47. DeCamp, S. J.; Tsuda, V. M. K.; Ferruzzi, J.; Koehler, S. A.; Giblin, J. T.; Roblyer, D.; Zaman, M. H.; Weiss, S. T.; Kiliç, A.; De Marzio, M.; Park, C. Y.; Ogassavara, N. C.; Mitchel, J. A.; Butler, J. P.; Fredberg, J. J. Epithelial layer unjamming shifts energy metabolism toward glycolysis. *Sci. Rep.* **2020**, *10*, 18302.
48. Bhattacharya, D.; Azambuja, A. P.; Simoes-Costa, M. Metabolic Reprogramming Promotes Neural Crest Migration via Yap/Tead Signaling. *Dev. Cell* **2020**, *53*, 199-211.e6.
49. Zhang, J.; Goliwas, K. F.; Wang, W.; Taufalele, P. V.; Bordeleau, F.; Reinhart-King, C. A. Energetic regulation of coordinated leader–follower dynamics during collective invasion of breast cancer cells. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 7867–7872.
50. Commander, R.; Wei, C.; Sharma, A.; Mouw, J. K.; Burton, L. J.; Summerbell, E.; Mahboubi, D.; Peterson, R. J.; Konen, J.; Zhou, W.; Du, Y.; Fu, H.; Shanmugam, M.; Marcus, A. I. Subpopulation targeting of pyruvate dehydrogenase and GLUT1 decouples metabolic heterogeneity during collective cancer cell invasion. *Nat. Commun.* **2020**, *11*, 1533.
51. Kondo, H.; Ratcliffe, C. D. H.; Hooper, S.; Ellis, J.; MacRae, J. I.; Hennequart, M.; Dunsby, C. W.; Anderson, K. I.; Sahai, E. Single-cell resolved imaging reveals intra-tumor heterogeneity in glycolysis, transitions between metabolic states, and their regulatory mechanisms. *Cell Rep.* **2021**, *34*, 108750.
52. Sahoo, S.; Duddu, A. S.; Biddle, A.; Jolly, M. K. Interconnected High-Dimensional Landscapes of Epithelial-Mesenchymal Plasticity and Stemness. *Preprints* **2021**, 2021040415.
53. Currie, E.; Schulze, A.; Zechner, R.; Walther, T. C.; Farese, R. V. Cellular fatty acid metabolism and cancer. *Cell Metab.* **2013**, *18*, 153–61.
54. Jia, D.; Park, J. H.; Kaur, H.; Jung, K. H.; Yang, S.; Tripathi, S.; Galbraith, M.; Deng, Y.; Jolly, M. K.; Kaiparettu, B. A.; Onuchic, J. N.; Levine, H. Towards decoding the coupled decision-making of metabolism and epithelial-to-mesenchymal transition in cancer. *Br. J. Cancer* **2021**, *124*, 1902–1911.
55. Walsh, H. R.; Cruickshank, B. M.; Brown, J. M.; Marcato, P. The flick of a switch: Conferring survival advantage to breast cancer stem cells through metabolic plasticity. *Front. Oncol.* **2019**, *9*, 753.
56. Cook, D. P.; Vanderhyden, B. C. Context specificity of the EMT transcriptional response. *Nat. Commun.* **2020**, *11*, 2142.
57. Shaul, Y. D.; Freinkman, E.; Comb, W. C.; Cantor, J. R.; Tam, W. L.; Thiru, P.; Kim, D.; Kanarek, N.; Pacold, M. E.; Chen, W. W.; Bieri, B.; Possemato, R.; Reinhardt, F.; Weinberg, R. A.; Yaffe, M. B.; Sabatini, D. M. Dihydropyrimidine Accumulation Is Required for the Epithelial-Mesenchymal Transition. *Cell* **2014**, *158*, 1094–1109.
58. Gatenby, R. A.; Gillies, R. J. Why do cancers have high aerobic glycolysis? *Nat. Rev. Cancer* **2004**, *4*, 891–9.
59. Viale, A.; Corti, D.; Draetta, G. F. Tumors and mitochondrial respiration: a neglected connection. *Cancer Res.* **2015**, *75*, 3685–6.
60. LeBleu, V. S.; O'Connell, J. T.; Gonzalez Herrera, K. N.; Wikman, H.; Pantel, K.; Haigis, M. C.; de Carvalho, F. M.; Damascena, A.; Domingos Chinen, L. T.; Rocha, R. M.; Asara, J. M.; Kalluri, R. PGC-1 $\alpha$  mediates mitochondrial biogenesis and oxidative phosphorylation in cancer cells to promote metastasis. *Nat. Cell Biol.* **2014**, *16*, 992–1003.
61. Yalcin, A.; Solakoglu, T. H.; Ozcan, S. C.; Guzel, S.; Peker, S.; Celikler, S.; Balaban, B. D.; Sevinc, E.; Gurpinar, Y.; Chesney, J. A. 6-phosphofructo-2-kinase/fructose 2,6-bisphosphatase-3 is required for transforming growth factor  $\beta$ 1-enhanced invasion of Panc1 cells in vitro. *Biochem. Biophys. Res. Commun.* **2017**, *484*, 687–693.

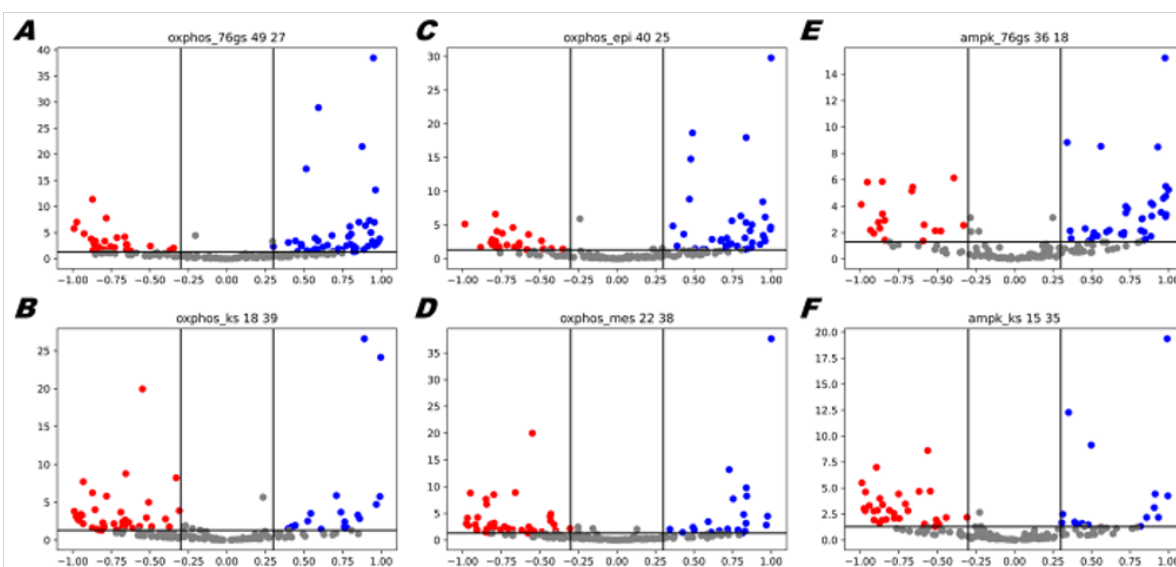


62. Ito, S.; Fukusato, T.; Nemoto, T.; Sekihara, H.; Seyama, Y.; Kubota, S. Coexpression of glucose transporter 1 and matrix metalloproteinase-2 in human cancers. *J. Natl. Cancer Inst.* **2002**, *94*, 1080–91.
63. Anderson, M.; Marayati, R.; Moffitt, R.; Yeh, J. J. Hexokinase 2 promotes tumor growth and metastasis by regulating lactate production in pancreatic cancer. *Oncotarget* **2017**, *8*, 56081–94.
64. Wu, N.; He, C.; Zhu, B.; Jiang, J.; Chen, Y.; Ma, T. 3-phosphoinositide dependent protein kinase-1 (PDK-1) promotes migration and invasion in gastric cancer cells through activating the NF- $\kappa$ B pathway. *Oncol. Res.* **2017**, *25*, 1153–1159.
65. Zhang, Y.; Lin, S.; Chen, Y.; Yang, F.; Liu, S. LDH-A promotes epithelial-mesenchymal transition by upregulating ZEB2 in intestinal-type gastric cancer. *Onco. Targets. Ther.* **2018**, *11*, 2363–2373.
66. Gaude, E.; Frezza, C. Tissue-specific and convergent metabolic transformation of cancer correlates with metastatic potential and patient survival. *Nat. Commun.* **2016**, *7*, 13041.
67. Røslund, G. V.; Dyrstad, S. E.; Tusubira, D.; Helwa, R.; Tan, T. Z.; Lotsberg, M. L.; Pettersen, I. K. N.; Berg, A.; Kindt, C.; Hoel, F.; Jacobsen, K.; Arason, A. J.; Engelsens, A. S. T.; Ditzel, H. J.; Lønning, P. E.; Krakstad, C.; Thiery, J. P.; Lorens, J. B.; Knappskog, S.; Tronstad, K. J. Epithelial to mesenchymal transition (EMT) is associated with attenuation of succinate dehydrogenase (SDH) in breast cancer through reduced expression of SDHC. *Cancer Metab.* **2019**, *7*, 6.
68. Yu, L.; Lu, M.; Jia, D.; Ma, J.; Ben-Jacob, E.; Levine, H.; Kaiparettu, B. A.; Onuchic, J. N. Modeling the Genetic Regulation of Cancer Metabolism: Interplay Between Glycolysis and Oxidative Phosphorylation. *Cancer Res.* **2017**, *77*, 1564–74.
69. Tripathi, S.; Levine, H.; Jolly, M. K. The Physics of Cellular Decision-Making during Epithelial-Mesenchymal Transition. *Annu. Rev. Biophys.* **2020**, *49*, 1–18.
70. Karacosta, L. G.; Anchang, B.; Ignatiadis, N.; Kimmey, S. C.; Benson, J. A.; Shrager, J. B.; Tibshirani, R.; Bendall, S. C.; Plevritis, S. K. Mapping Lung Cancer Epithelial-Mesenchymal Transition States and Trajectories with Single-Cell Resolution. *Nat. Commun.* **2019**, *10*, 5587.
71. Celià-Terrassa, T.; Bastian, C.; Liu, D. D.; Ell, B.; Aiello, N. M.; Wei, Y.; Zamalloa, J.; Blanco, A. M.; Hang, X.; Kunisky, D.; Li, W.; Williams, E. D.; Rabitz, H.; Kang, Y. Hysteresis control of epithelial-mesenchymal transition dynamics conveys a distinct program with enhanced metastatic ability. *Nat. Commun.* **2018**, *9*, 5005.
72. Liberzon, A.; Subramanian, A.; Pinchback, R.; Thorvaldsdóttir, H.; Tamayo, P.; Mesirov, J. P. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **2011**, *27*, 1739–1740.
73. Subramanian, A.; Tamayo, P.; Mootha, V. K.; Mukherjee, S.; Ebert, B. L.; Gillette, M. A.; Paulovich, A.; Pomeroy, S. L.; Golub, T. R.; Lander, E. S.; Mesirov, J. P. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **2005**, *102*, 15545–15550.
74. Foroutan, M.; Bhuva, D. D.; Lyu, R.; Horan, K.; Cursons, J.; Davis, M. J. Single sample scoring of molecular phenotypes. *BMC Bioinformatics* **2018**, *19*, 404.
75. Jia, D.; Paudel, B. B.; Hayford, C. E.; Hardeman, K. N.; Levine, H.; Onuchic, J. N.; Quaranta, V. Drug-Tolerant Idling Melanoma Cells Exhibit Theory-Predicted Metabolic Low-Low Phenotype. *Front. Oncol.* **2020**, *10*, 1426.

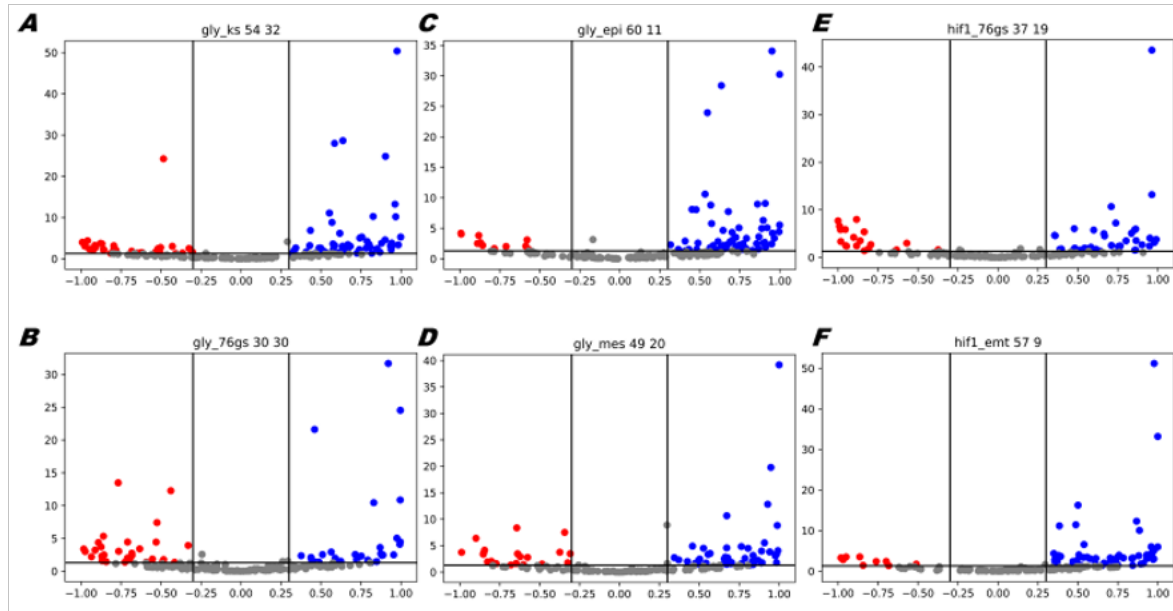
## Supplementary figures



**Fig S1: Consistency between different EMT scoring metrics.** Volcano plots depicting the Pearson correlation coefficient and the  $-\log_{10}(p\text{-value})$  for (A) Epithelial vs Hallmark EMT (B) KS vs Mesenchymal (C) Epithelial vs Mesenchymal (D) Mesenchymal vs 76GS (E) KS vs Epithelial (F) Mesenchymal vs Hallmark EMT. Vertical boundaries are set at correlation coefficient -0.3 and 0.3. The cut-off for p-value is set at 0.05.



**Fig S2: OXPHOS and its regulator AMPK is more likely to correlate negatively with EMT.** Volcano plots depicting the Pearson correlation coefficient and the  $-\log_{10}(p\text{-value})$  for (A) OXPHOS vs 76GS (B) OXPHOS vs KS (C) OXPHOS vs Epithelial (D) OXPHOS vs Mesenchymal (E) AMPK vs 76GS (F) AMPK vs KS. Vertical boundaries are set at correlation coefficient -0.3 and 0.3. The cut-off for p-value is set at 0.05.



**Fig S3: Glycolysis and its regulator HIF1a is more likely to correlate negatively with EMT.** Volcano plots depicting the Pearson correlation coefficient and the  $-\log_{10}(p\text{-value})$  for (A) Glycolysis vs KS (B) Glycolysis vs 76GS (C) Glycolysis vs Epithelial (D) Glycolysis vs Mesenchymal (E) HIF1a vs 76GS (F) HIF1a vs Hallmark EMT. Vertical boundaries are set at correlation coefficient -0.3 and 0.3. The cut-off for  $p\text{-value}$  is set at 0.05.