A Rational Drug Combination Design Proves to Inhibit Epithelial-Mesenchymal Transition in a Three-Dimensional Microenvironment

Farnaz Barneh[1,2], Mehdi Mirzaie[3], Payman Nickchi[2,4], Mehran Piran[2], Mona Salimi[5], Fatemeh Goshadrou[1], Amir Reza Aref[6*], Mohieddin Jafari[2*]

[1] Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

[2] Drug Design and Bioinformatics Unit, Medical Biotechnology Department, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran.

[3] Department of Applied Mathematics, Faculty of Mathematical Sciences, Tarbiat Modares University, Tehran, Iran.

[4] Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, BC, Canada.

[5] Department of Physiology and Pharmacology, Pasteur Institute of Iran, Tehran, Iran.

[6] Belfer Center for Applied Cancer Science, Dana-Farber Cancer Institute, Harvard Medical School, Boston, USA.

Correspondence authors
Mohieddin Jafari: m_jafari@pasteur.ac.ir
Amir Reza Aref: Amir_Aref@hms.harvard.edu

Running title: Drug Combination Design to inhibit EMT
Character count of main manuscript including spaces: 36831

Abstract

Epithelial-Mesenchymal Transition (EMT) is the major player of tumor invasiveness whose inhibition is challenged by redundancy of multiple inducing factors in tumor microenvironment. Here, we applied a systems pharmacology approach by integrating network-based analyses with multiple bioinformatic resources to propose a combinatorial drug regimen that reverses EMT phenotype in aggressive cancers. The results demonstrated that histone deacetylases were central targets to relay tumor-host interactions by tuning expression of multiple epithelial versus mesenchymal genes. Moreover, SRC and IKBK served as the principal intracellular kinases regulating information flow from tumor microenvironment to multiple signaling pathways associated with EMT. To experimentally validate these predictions, we inhibited the pinpointed targets with already prescribed drugs and observed that whereas low dose mono-therapy failed to limit cell dispersion in a three-dimensional (3D) microfluidic device as a metric of EMT, the combinatorial regimen fully inhibited dispersion and invasion of cancer cells in collagen spheroids toward co-cultured endothelial cells while maintaining their survival. Taking into account Food and Drug Administered (FDA) - approval as well as safety profiles of the proposed drugs, we offer a combinatorial regimen with immediate translational implications to inhibit EMT in aggressive cancers. Besides, our study provides a grasp on the interdisciplinary methodologies behind plethora of the introduced biologist-friendly bioinformatic tools for bench and bed cancer scientists to help them effectively plan their experiments.

Key words: Cancer/ Combinatorial drug therapy/ Epithelial-mesenchymal transition/ Metastasis/ Systems pharmacology.

Background

Epithelial_ Mesenchymal Transition (EMT) is a complex reprogramming process through which tumor cells confront with their limitations in their microenvironment to facilitate their dissemination(Gao et al, 2012; Yang & Weinberg, 2008). Despite recognition of EMT as the key machinery to evade environmental barriers in tumors, full understanding of the process is yet far from complete (Pasquier et al, 2015; Singh & Settleman, 2010; Thiery & Sleeman, 2006). From the therapeutic point of view, the dynamic and complex nature of EMT necessitates a rational polypharmacology paradigm to hit multiple targets but the challenge then would be selecting which of many potential combinatorial drugs testable in a relevant setting (Said et al, 2013). To address the aforementioned problems, we believe that elaborate bioinformatic and statistical analyses required for interpreting such complex data should be passed to biologists and clinicians through a pipeline of user-friendly tools, so that they can perceive principal mechanisms of EMT process as a whole (Beck et al, 2014). Here, we aimed to set a stage for rational assimilation of multiple bioinformatic resources to propose a multi-target shotgun against of EMT from currently published signatures (**Figure1**).

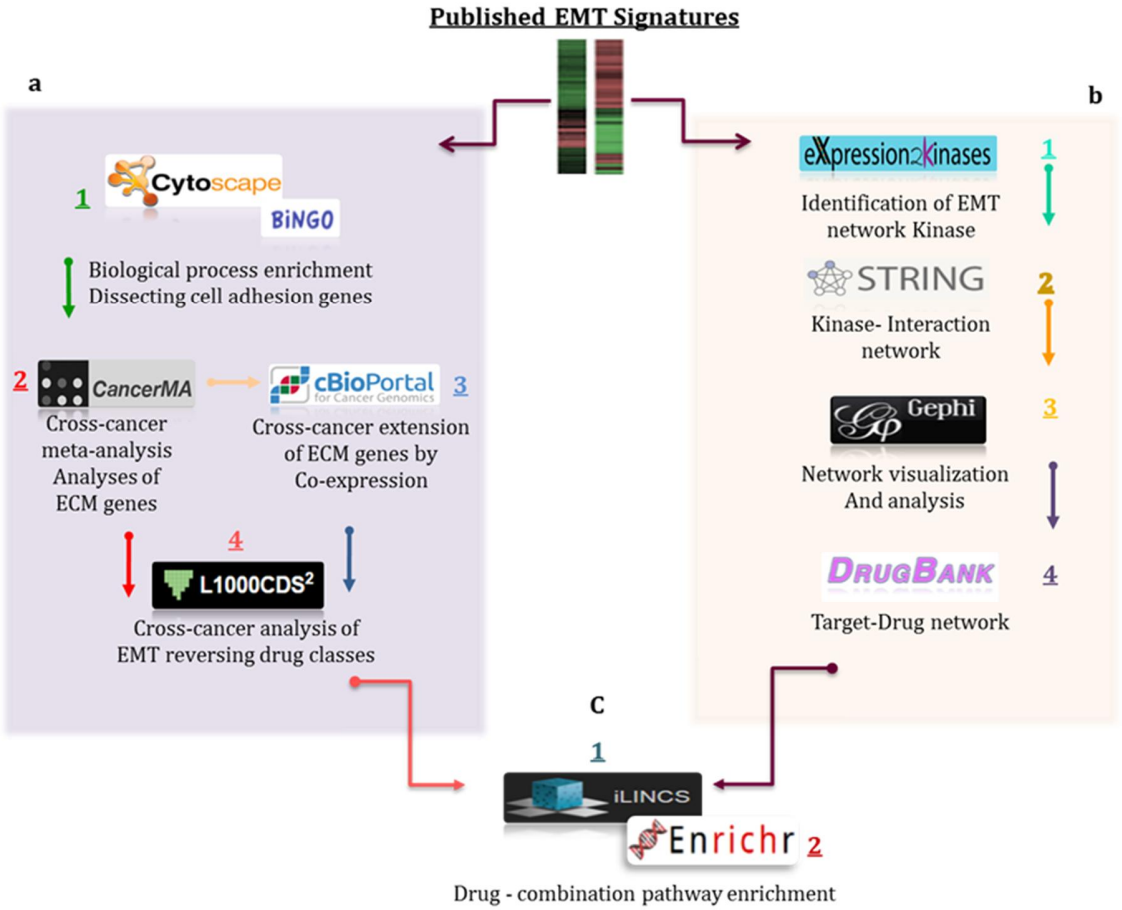

Figure 1. A systems-pharmacology workflow for generating insights into key players of EMT in cancer.

A From published EMT signatures, genes associated with cell adhesion were extracted (***BiNGO*** App in Cytoscape), converted to up/down signature in various solid cancers (***CancerMA***) and extended by co-expression (***cBioPortal***) to identify drugs (***L1000CDS²***) that induce expression of epithelial-related genes while down-regulate mesenchymal associated genes.

B Other important druggable target classes in published EMT signatures were identified to be kinases (***PHAROS***), thus EMT-associated kinome were identified by enrichment analysis (***Expression2Kinases***) and were used to create a kinase- Interaction Network in EMT (***STRING***). Network analysis (***Gephi***) was performed to identify important kinases and their inhibitors (***DrugBank***).

C The signatures of the identified drugs were extracted (***iLINCS***) and subjected to pathway enrichment analysis (***Enrichr***) to estimate drug combination efficacy.

Methods

Study design

In order to gain insight into therapeutic targets inhibiting EMT during tumor invasion and metastasis, we compiled EMT-related gene lists from four published studies reporting pan - cancer EMT signatures and generated a library of 962 genes (Byers et al, 2013), (Tan et al, 2014), (Gröger et al, 2012) and (Liang et al, 2016). Taking into account that an important aspect of EMT initiation is deregulation of cell adhesion –related genes, Gene Ontology-based enrichment analysis was performed on the 962-EMT gene list and 93 genes associated with cell adhesion were dissected. The gene list was then converted into an up-regulated (regarded as mesenchymal-related genes) and down- regulated (regarded as epithelial-related genes) mode for ten individual solid tumors by means of meta-analysis executed in ***"CancerMA"*** database.

To identify drugs which reverse the expression of cell adhesion-related genes with aberrant expression in EMT, the up/down gene lists for each cancer was queried into **"L1000CDS^2 "** search engine to prioritize small molecule inhibitors based on characteristics direction method. In order to take into account other genes in patients' samples which are co-expressed and thus cooperative to execute EMT reprogramming, co-expression toolbox in "***cBioPortal***" database was used to access TCGA data and select highly correlated genes (Pearson's correlation coefficient >0.8) with ECM-related gene sets in each cancer. The extended lists were then resubmitted in ***"L1000CDS^2 "*** for reversing drugs.

To determine other druggable targets in EMT, target families of proteins exploited in pharmaceutical industry were extracted from ***"PHAROS"*** database (Nguyen et al, 2016) and were mapped onto the 962- EMT gene library. Taking into account the importance of kinases in regulating signaling pathways in EMT, we were then promoted to elucidate EMT-related druggable kinome whose inhibition may offer synergism to maximize EMT reversal effects of HDACIs. We thus performed kinase- substrates enrichment analysis using ***"Expresion2Kinases"*** software, and identified 83 kinases to be associated with EMT gene

sets. From this kinase list, we then constructed EMT-associated kinase- interaction network by submitting the kinase list into ***"STRING 10.0"*** database. Elucidating the relationships among various kinases as a network then enabled us to apply centrality metrics of graph theory to prioritize the principal kinases.

We confirmed our findings based on previously published literature and validated the anti-EMT effect of proposed drug combination in a three-dimensional microfluidic device in which A549 lung adenocarcinoma collagen spheres co-cultured with HUVEC cells.

Statistical analysis

Statistical significance of enrichment analyses were adjusted for multiple by False Discovery Rate (FDR) <5% method embedded in the bioinformatic tools used in this study, and is designated as q-value in the figures. For meta-analysis results from CancerMA database, -1.5 <Fold Change <+ 1.5 was considered significant (Dalman et al, 2012). For co-expression analyses using cBioPortal, Pearson correlation co-efficient >0.8 was considered significant. In cell-based experiments for validation of drug effects on cell dispersion from the spheroids, comparisons were made across time-points or group of drugs, Two-way ANOVA with Tukey's multiple comparison or two tailed t-test with statistical significance of was level < 0.05 were performed in GraphPad Prism 6. Data with error bars represent mean± SEM for indicated number of replicates. Quantification of cell-based experiments was performed with MATLAB 2014b.

Results and Discussion

Among the 93 gene subset, 82 were found to be deregulated among multiple cancers however, no common gene was found to be deregulated in all cancer types in the study. Assessment of similarity extent for these genes among ten solid tumors showed that Fibrillin (MFAP) was upregulated maximally in seven cancer types while a metalloproteinase (ADAM12) and Fibronectin (FN1) were upregulated only in six cancer types. Dermatopontin (DPT) and stromal cell-derived factor-1 (CXCL12) were down regulated only in six and five cancer types respectively **(Figure 2)**. Interestingly, we noticed that even though each cancer exhibited unique set of expression patterns for cell-

adhesion related genes, Histone Deacetylase Inhibitors (HDACIs) including Trichostatin A, Vorinostat, Panobinostat and Pracinostat were repeatedly observed to reverse the expression of various cell adhesion-related genes across all caner types in the study **(Table 1)**. Likewise, HDACIs were among the top ranked reversing drugs for the EMT related co-expressed genes in TCGA data. These results underscore the significance of overcoming epigenetic barriers to switch from epithelial traits into mesenchymal features independent of cancer type (Bedi et al, 2014; Tanaka & Ogishima, 2015). Confirming our results, Tang et al. performed a chemical screening strategy to identify drugs that could induce expression of E-cadherin as a measure of EMT inhibition. They observed that among multiple drug classes, HDAC inhibitors could restore E-cadherin expression with low cytotoxicity (Tang et al, 2016). Javaid et al. demonstrated that next generation of HDAC inhibitors including Panobinonstat completely abolished SNAIL-induced EMT (Javaid et al, 2013). Epigenetic modifications have been also identified as key regulatory mechanisms in reactivation of EMT-related transcription factors (McDonald et al, 2011; Serrano-Gomez et al, 2016; Tam & Weinberg, 2013; West & Johnstone, 2014). We can conclude from this part that regardless of different phenotypes a single tumor cell may acquire in a changing microenvironment, targeting its adaptability machinery by HDACI drugs would be a smart approach to inhibit tumor cross-talk with its microenvironment.

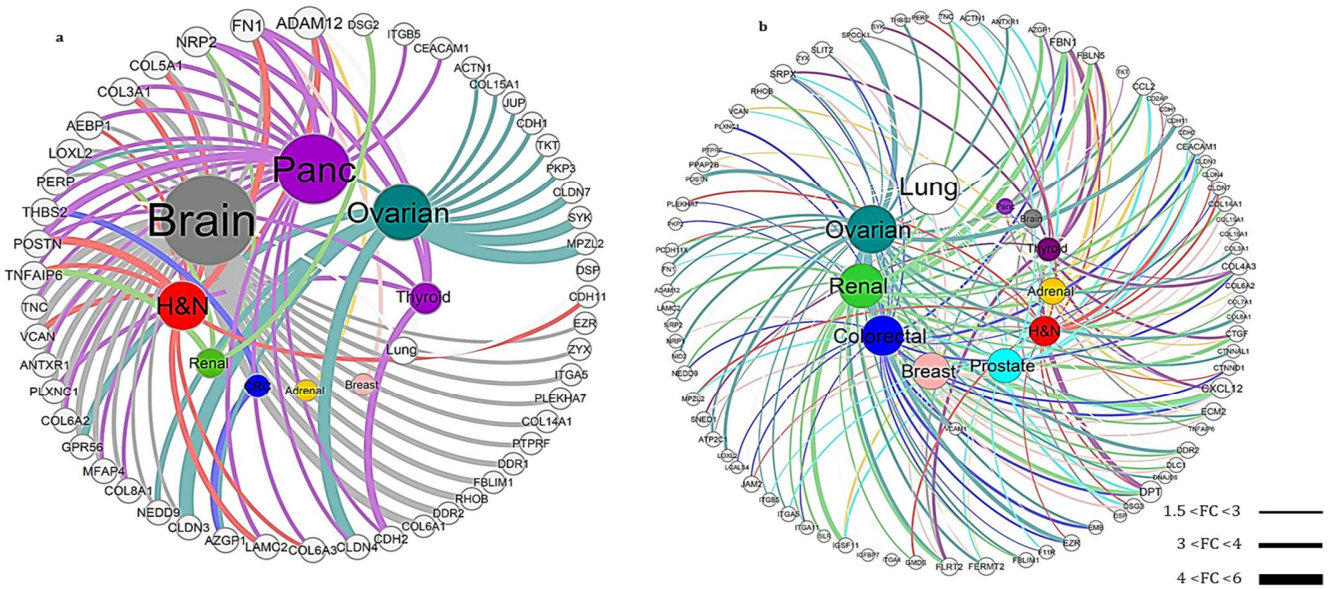

Figure 2. **Meta-analysis results on the expression status of cell adhesion-related genes across various solid cancers**.

A Pan-cancer cell adhesion related genes which were dissected from collated EMT signatures and their expression were increased in meta-analysis performed by CancerMA database. These upregulated genes were regarded as genes associated with mesenchymal features.

B Pan-cancer downregulated cell-adhesion related from collated EMT signature genes in various cancers obtained by meta-analysis in CancerMA. These downregulated genes were assumed to be associated with epithelial characteristics.

The outer layer shows de-regulated genes and the inner layer represents cancer types. Edge thickness in the networks signifies folds change (FC) range.

Among other druggable targets in EMT, kinases constituted the main class of proteins in collated EMT gene set. Thus the EMT-associated kinome were identified to construct a kinase interaction network **(Figure3)**. With this network-based identification of appropriate druggable targets, all neighboring interactions were taken into account so that influential kinases whose inhibition will impose fragility to the whole network could be identified (Korcsmáros et al, 2007). We observed that proto-oncogene SRC with 45 interactions was the hub node in the largest module of the network as well as in the whole network. SRC also owned the highest betweenness (0.27) and closeness (0.65) values which imply the role of SRC to be at the crossroad of multiple pathways. These results suggest that interrupting SRC kinase activity may disrupt information flow from large array of extracellular signals to various cellular pathways involved in initiation and maintenance of EMT. With closer inspection of kinase- interaction network, we observed that by targeting SRC kinase alone, only three modules of the network (shown with blue, orange and green nodes) would be manipulated while an important module of the network associated with cytokine-cytokine receptor interaction and TGF-beta signaling (red nodes) would be unaffected . In search for a druggable co-target for SRC, we found IKBK, which regulates activation of NF-kB transcription factor, as the hub kinase of green module associated with inflammatory pathways such as TNF-alpha, a known contributor to EMT. IKBK was also directly connected to SRC, TGF-beta, MAPK14 and CDK7 which were distributed among four other modules. IKBK owned the second rank for betweenness

(0.15) and third rank for closeness (0.56) centralities in the network confirming its deliberate position in the network. These results suggest that by dual targeting SRC and IKBK, the whole kinase- interaction network associated with EMT would be affected (**Figure 3B**).

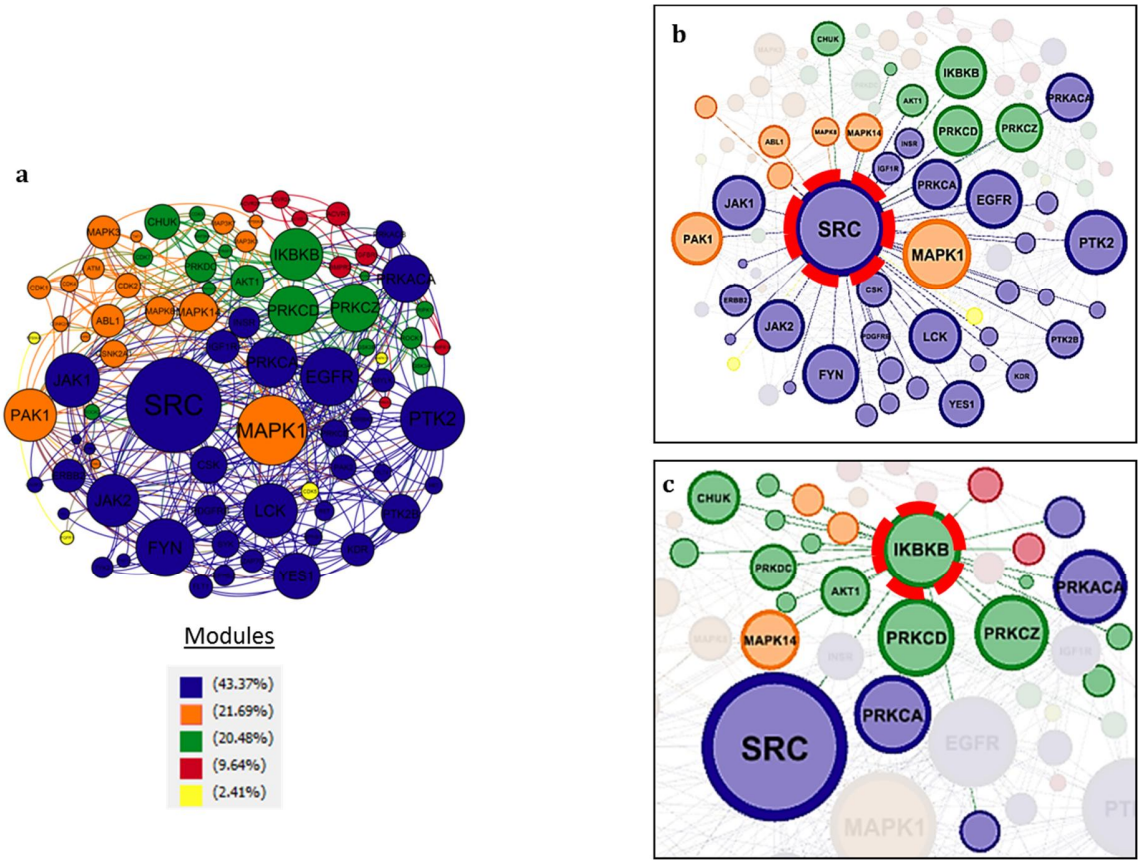

Figure 3. Functional Kinase -Interaction network in EMT

Important kinases associated with EMT signatures were identified by Expression2Kinases enrichment tool and converted to an interaction network using STRING database and visualized in Gephi software.

A Overall view of Kinase- Interaction Network with its constituent modules.

B, C First-neighbor kinases functionally connected to SRC and IKBK kinases along with currently available FDA-approved kinase inhibitors for SRC and IKBK are shown. Co-targeting SRC and IKBK results in manipulation of all modules in the network indicated by colors.

To confirm the relevance of SRC and IKBK kinases in inducing EMT, ***"CREEDS database"*** of gene perturbation signatures (Wang et al, 2016) was queried to identify differentially expressed genes (GEGs) in conditions where SRC and IKBK are overexpressed or constitutively activated. Gene ontology analysis of DEGs showed that signatures of SRC overexpression were associated with cell adhesion and actin cytoskeleton organization. In IKBK over-activation signatures, epithelial cell differentiation, extracellular matrix organization, cell junction organization, regulation of inflammatory response were observed (adjusted P-value<0.05), all of which are evidently associated processes in EMT. Consistently, the overlap between DEGs in the SRC and IKBK perturbation signatures and the 962-EMT gene set was shown to be significant (P-value< 0.01) in hypergeometric test.

We then proceeded to identify Food and Drug Administered (FDA)-approved drugs that inhibit the aforementioned kinases. EMT-associated kinases were thus mapped onto Drug-Target Network (Barneh et al, 2015; Wishart & Wu, 2016). We found four FDA-approved drugs including Dasatinib, Bosutinib, Nintedanib and Ponatinib to inhibit SRC kinase. Aspirin, Sulfasalazine, Mesalazine currently used for treatment of inflammatory diseases were also found to inhibit IKBK

It was hypothesized that combination of these drugs with HDAC inhibitors may be effective for limiting EMT consequences in cancer. To primarily confirm such prediction, gene expression changes induced by three selected drug classes including Vorinostat (HDACI inhibitor), Dasatinib (multi-kinase SRC inhibitor) and Sulfasalazine (as IKBK inhibitor) were extracted from ***"iLINCS, the integrative LIKNCS"*** database for two cancer cell lines including MCF-7 epithelial cancer cell line and A549 mesenchymal lung cancer cell line. Several pathways involved in EMT were enriched following collation of DEGs from three drugs based on cell line **(Figure 4)**. In MCF-7, a hormone responsive breast cancer cell line with epithelial traits, cell cycle, proteoglycans and Wnt signaling were enriched in the combination signature while in A549, a lung adenocarcinoma derived cell line with reversible mesenchymal traits, the combinatorial drug signature was associated with manipulation of several EMT-related pathways including adherens junction and regulation of actin cytoskeleton, FoxO signaling, chemokine and NF-kB signaling. Other pathways

including mismatch repair and cell cycle were also affected by drug combinations which are underlying mechanisms of EMT-induced drug resistance.

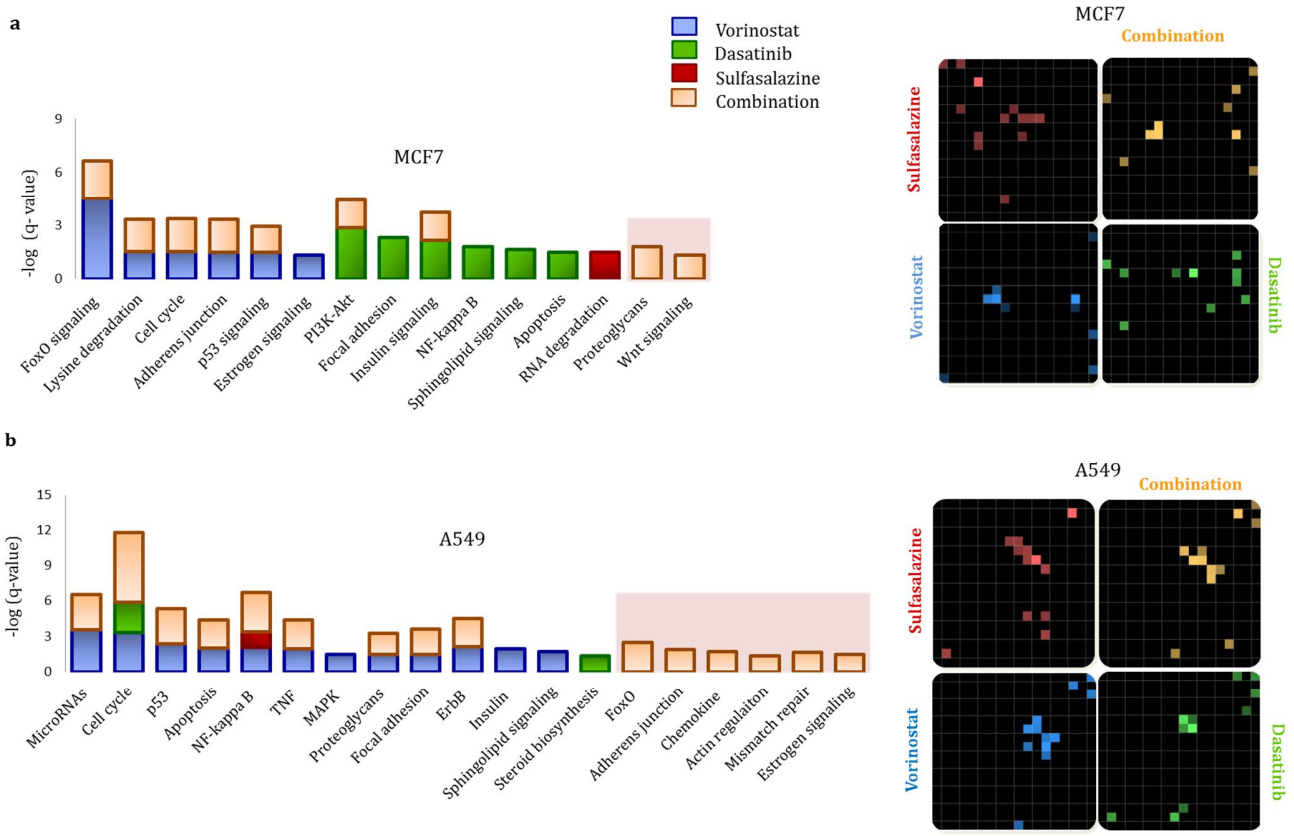

Figure 4. Pathway analysis of genes that are altered by drugs and their combinations.

A Enriched biological pathways (left) and their perturbation status in a grid view representation of cells (right) by each predicated drug and the combination in MCF-7 breast cancer cells

B Enriched biological pathways (left) and their perturbation status in a grid view representation of cells (right) by each predicated drug and the combination in lung adenocarcinoma cell line A549.

The adjusted *p*-value (q-value) <0.05 is considered to be significant. Color codes represent each drug and the combination. Pathways specifically enriched by combination are highlighted in purple.

Finally, for validating the efficacy of predicted EMT reversing drug combination in vitro, we implemented a 3-dimensional (3D) microfluidic device with collagen- embedded A549 lung cancer spheroids and co-cultured HUVEC endothelial cells to induce EMT to account for interaction of tumor cells with extracellular matrix in EMT. As previously shown by Aref et al, dissociation of cancer cells from the spheroids is highly correlated with expression of EMT markers E-cadherin and vimentin (Aref et al, 2013).

Despite support of literature given that each of the selected drugs manifest anti-EMT properties in 2D experiments, we found that in 3D environment, low dose treatments (1 μM) with each of the drugs including Vorinostat, Dasatinib or Aspirin alone did not inhibit cell dispersion from the initial aggregates however; we observed full blockade of dispersion when three drugs were combined **(Figure 5)**.

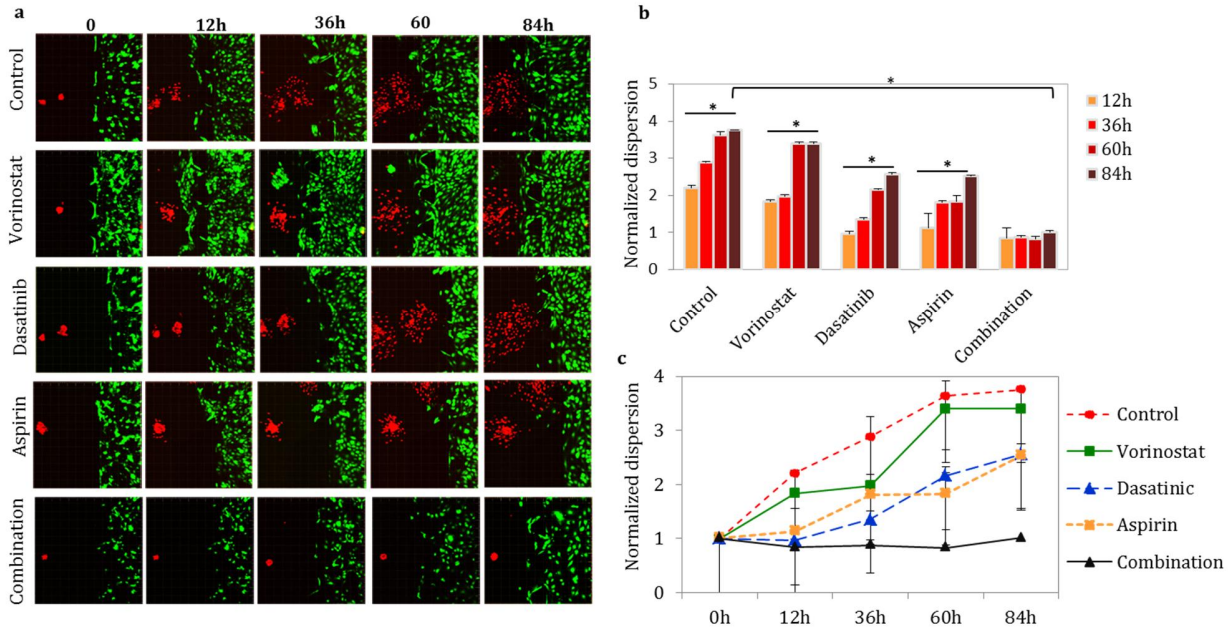

Figure 5. Real-time analysis of Human A549 lung cancer cell dispersion from collagen aggregates when co-cultured with HUVEC cell layer in presence and absence of predicted drugs or their combinations.

A For each drug, 1uM concentration was used and live cells emitting red (A549 cells) and green (HUVEC cells) fluorescence were monitored with time-lapse confocal microscope in a period of 84 hours.

B Quantification of normalized A549 cell dispersion by pixel measurement in MATLAB 2014b following treatment with hypothesized drugs or their combinations. Two-way ANOVA was performed at statistical significance $p < 0.05$*. Normalization was performed by dividing dispersion distance at each time-point to 0h. Data are represented as mean± SEM and the error bars represent dispersion measurements in three independent spheroids.

C Time-course representation of A549 dissociation from the spheroids during 84 h.

As shown, the three-drug combination also confined invasion of endothelial cell towards spheroids which may correlate with anti-angiogenic effects in vivo as another consequence of EMT which incites future confirmatory experiments. Consistent with our observation, Gopal et al. showed that aggressive cancer cells undergoing EMT induce angiogenesis by secreting exosomes (Gopal et al, 2016). These results imply that the combinatorial regimen not only inhibits cancer cells response to the microenvironmental signals (indicated by dispersion from the aggregates), but may also alter secreted factors from cancer cells per se and limit their communication with the host. It is also grossly evident from the images that the triple-combination treatment maintained survival of co-cultured endothelial cells which denotes tolerability of proposed combination regimen for normal cells.

In order to exclude the dependency of synergistic effects on drugs themselves rather than their mechanism of action, we also tested efficacy of Sulfasalazine (another IKBK inhibitor) on inhibition of EMT **(Figure 6)**. We observed that in Sulfasalazine treated group, a significant delay in dispersion was observed between 0h and 84h compared to control, however, when Dasatinib was added to Sulfasalazine, it decreased the delay in dissociation which is observed when Sulfasalazine is used alone. These results indicate that at least two drugs with different mechanism of actions are required to efficiently inhibit cell dispersion as a measure of EMT when complex microenvironmental signals are taken into account. However, this combination was not sufficient to inhibit tumor cross-talk with its microenvironment since as shown**,** invasion of endothelial cells towards spheroids was not inhibited.

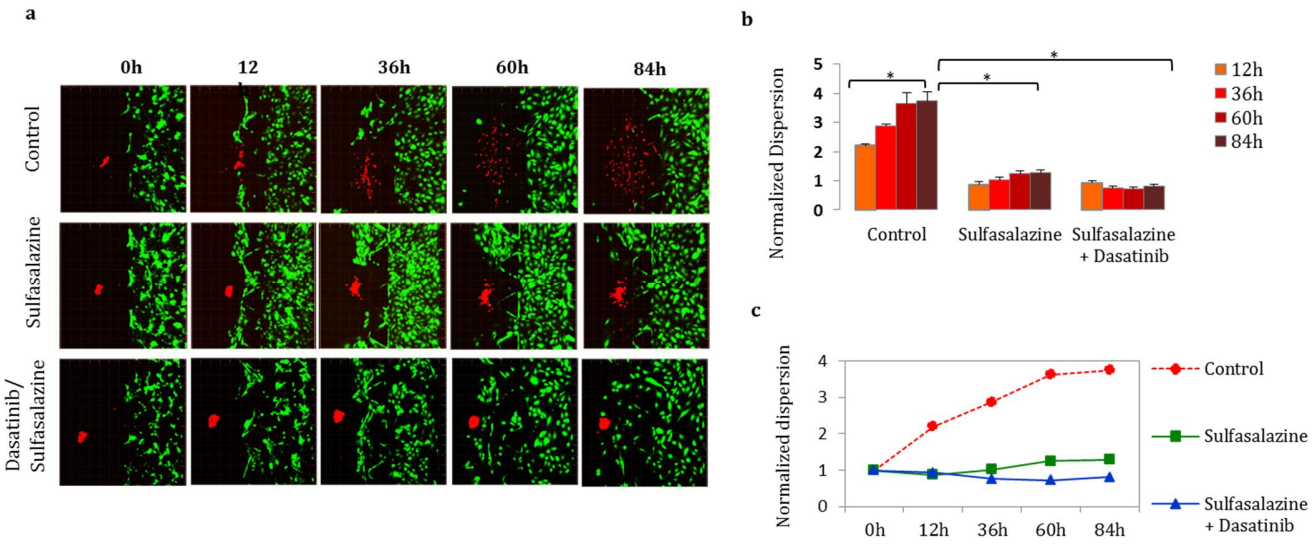

Figure 6. Inhibition of A549 cell dispersion in presence and absence of 1 μM Sulfasalazine and in combination with Dasatinib.

A Live cell imaging of control and treated groups obtained in a period of 84 hours.

B Quantification of normalized A549 cell dispersion from collagen aggregates by mean distance of measured pixels in MATLAB 2014b following treatment with Sulfasalazine or its combination with Dasatinib. Two-tailed t-test with p <0.05 was used to compare dispersion difference between control versus Sulfasalazine or control versus Sulasalazine/ Dasatinib at 84h. Data represent mean± SEM and the error bars indicate dispersion measurements in two independent spheroids.

C Time-point view of dissociation from the spheroids in control and Sulfasalazine or sulfasalazine plus Dasatinib.

Conclusion

Here we took a rational polypharmacology approach and proposed a drug combination regimen to inhibit EMT in aggressive tumors. The highlighted points of our study are: first, practicing a systems pharmacology methodology by elucidating how rational integration of currently available bioinformatics databases and software with network analysis can translate the inconsistent genome-wide expression signatures into a predictive pipeline for optimization of drug combination design. Second, implementing a 3-D co-culture system to simulate tumor microenvironment in which, real-time cell behaviors are tracked for

evaluating drug combination effects. Third, use of clinically achievable doses of already prescribed drugs consistent with their IC50 values. Taking into account these points, we observed agreement between in-silico and in-vitro results to confirm the relevance of proposed combinatorial regimen in stopping a process such as EMT. Moreover, given the safety of the proposed drugs in this study, as well as their FDA- approval for other indications, the outcomes presented here, are of direct translatability to clinical implications.

Author contributions:

FB Conceived the project, designed the experiments and performed computational/ bioinformatic tests, analyzed data and prepared the manuscript. **MM** Supervised bioinformatic analyses and **PN** performed the statistical experiments. **MP** assisted in performing statistical tests. **MS** together with **FG** assisted in interpretation of drug effects on EMT and helped in evaluation of the results. **ARA** performed the co- culture experiments with microfluidic device, evaluated the manuscript and corresponded the manuscript. **MJ** Supervised the whole process from experimental design to data analysis and evaluation of the work, also corresponded the manuscript. All authors reviewed the manuscript.

Conflict of Interest:

Authors declare no confliction in the interests.

References

Aref AR, Huang RY-J, Yu W, Chua K-N, Sun W, Tu T-Y, Bai J, Sim W-J, Zervantonakis IK, Thiery JP (2013) Screening therapeutic EMT blocking agents in a three-dimensional microenvironment. *Integr Biol* **5:** 381-389

Barneh F, Jafari M, Mirzaie M (2015) Updates on drug–target network; facilitating polypharmacology and data integration by growth of DrugBank database. *Brief Bioinform* **17:** 1070-1080

Beck TN, Chikwem AJ, Solanki NR, Golemis EA (2014) Bioinformatic approaches to augment study of epithelial-to-mesenchymal transition in lung cancer. *Physiol Genomics* **46:** 699-724

Bedi U, Mishra V, Wasilewski D, Scheel C, Johnsen S (2014) Epigenetic plasticity: a central regulator of epithelial-to-mesenchymal transition in cancer. *Oncotarget* **5:** 2016-2029

Byers LA, Diao L, Wang J, Saintigny P, Girard L, Peyton M, Shen L, Fan Y, Giri U, Tumula PK (2013) 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* **19:** 279-290

Gao D, Vahdat LT, Wong S, Chang JC, Mittal V (2012) Microenvironmental regulation of epithelial–mesenchymal transitions in cancer. *Cancer Res* **72:** 4883-4889

Gopal SK, Greening DW, Hanssen EG, Zhu H-J, Simpson RJ, Mathias RA (2016) Oncogenic epithelial cell-derived exosomes containing Rac1 and PAK2 induce angiogenesis in recipient endothelial cells. *Oncotarget* **7:** 19709-19722

Gröger CJ, Grubinger M, Waldhör T, Vierlinger K, Mikulits W (2012) Meta-analysis of gene expression signatures defining the epithelial to mesenchymal transition during cancer progression. *PloS One* **7:** e51136

Javaid S, Zhang J, Anderssen E, Black JC, Wittner BS, Tajima K, Ting DT, Smolen GA, Zubrowski M, Desai R (2013) Dynamic chromatin modification sustains epithelial-mesenchymal transition following inducible expression of Snail-1. *Cell Rep* **5:** 1679-1689

Korcsmáros T, Szalay MS, Böde C, Kovács IA, Csermely P (2007) How to design multi-target drugs: target search options in cellular networks. *Expert Opin Drug Discov* **2:** 799-808

Liang L, Sun H, Zhang W, Zhang M, Yang X, Kuang R, Zheng H (2016) Meta-Analysis of EMT Datasets Reveals Different Types of EMT. *PloS One* **11:** e0156839

McDonald OG, Wu H, Timp W, Doi A, Feinberg AP (2011) Genome-scale epigenetic reprogramming during epithelial-to-mesenchymal transition. *Nat Struct Mol Biol* **18:** 867-874

Nguyen D-T, Mathias S, Bologa C, Brunak S, Fernandez N, Gaulton A, Hersey A, Holmes J, Jensen LJ, Karlsson A (2016) Pharos: Collating protein information to shed light on the druggable genome. *Nucleic Acids Res***:** gkw1072

Pasquier J, Abu-Kaoud N, Al Thani H, Rafii A (2015) Epithelial to mesenchymal transition in a clinical perspective. *J Oncol* **2015**

Said NAB, Simpson KJ, Williams ED (2013) Strategies and challenges for systematically mapping biologically significant molecular pathways regulating carcinoma epithelial-mesenchymal transition. *Cells Tissues Organs* **197:** 424-434

Serrano-Gomez SJ, Maziveyi M, Alahari SK (2016) Regulation of epithelial-mesenchymal transition through epigenetic and post-translational modifications. *Mol Cancer* **15:** 1

Singh A, Settleman J (2010) EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer. *Oncogene* **29:** 4741-4751

Tam WL, Weinberg RA (2013) The epigenetics of epithelial-mesenchymal plasticity in cancer. *Nat Med* **19:** 1438-1449

Tan TZ, Miow QH, Miki Y, Noda T, Mori S, Huang RYJ, Thiery JP (2014) Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients. *EMBO Mol Med***:** e201404208

Tanaka H, Ogishima S (2015) Network biology approach to epithelial–mesenchymal transition in cancer metastasis: three stage theory. *J Mol Cell Biol* **7:** 253-266

Tang H, Kuay K, Koh P, Asad M, Tan T, Chung V, Lee S, Thiery J, Huang R-J (2016) An epithelial marker promoter induction screen identifies histone deacetylase inhibitors to restore epithelial differentiation and abolishes anchorage independence growth in cancers. *Cell Death Dis* **2**

Thiery JP, Sleeman JP (2006) Complex networks orchestrate epithelial–mesenchymal transitions. *Nat Rev Mol Cell Biol* **7:** 131-142

Wang Z, Monteiro CD, Jagodnik KM, Fernandez NF, Gundersen GW, Rouillard AD, Jenkins SL, Feldmann AS, Hu KS, McDermott MG (2016) Extraction and analysis of signatures from the Gene Expression Omnibus by the crowd. *Nat Commun* **7**

West AC, Johnstone RW (2014) New and emerging HDAC inhibitors for cancer treatment. *J Clin Invest* **124:** 30-39

Wishart DS, Wu A (2016) Using DrugBank for In Silico Drug Exploration and Discovery. *Curr Protoc Bioinformatics***:** 14.14. 11-14.14. 31

Yang J, Weinberg RA (2008) Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. *Dev Cell* **14:** 818-829

Table-1. Predicted drugs to reverse EMT gene sets in various cancers.

Cancer type	Drugs	Mechanism	Targets
Lung	Digoxigenin	Cardiovascular, Inotropic	ADAM12, DSP, THBS2.
	Vorinostat	**HDACI***	ADAM12, COL4A3, FBLN5, COL4A3, MFAP4,
	Sunitinib	Multi-kinase inhibitor*	DSP, THBS2, DLC1.
Colorectal	Trametinib, Saracatinib, Linifanib	Multi-kinase inhibitor*	COL6A3, THBS2, AZGP1, COL6A3, COL4A3, CXCL12.
	Vorinostat, Pracinostat	**HDACI***	AZGP1, COL6A3, THBS2, COL6A3, EZR, MFAP4, CEACAM1.
	Albendazole	Anthelmintic*	THBS2, CEACAM1, EZR.
Breast	Itraconazole	Antifungal*	COL14A1, MFAP4, SRPX.
	Finasteride	5-alpha reductase inhibitor	COL14A1, DPT, PPAP2B.
	Topotecan	Anticancer*	COL14A1, CXCL12, DPT.
	Granisetron	antiemetic	FN1, CXCL12, SRPX.
	Regorafenib, Dasatinib	Multi-kinase inhibitor*	FN1, PPAP2B, SRPX, CXCL12, MFAP4.
	Mocetinostat	**HDACI***	CXCL12, PPAP2B, SRPX.
	Azacitidine	Anticancer	CXCL12, MFAP4, PPAP2B.
Prostate	Vorinostat	**HDACI**	SPOCK1, CEACAM1.
Pancreatic	Vorinostat	**HDACI***	AEBP1, COL3A1, COL5A1, COL6A2, COL6A3, ITGB5, PERP, POSTN, THBS2, TNFAIP6, ADAM12, ITGB5.
	Nintedanib	Multi-kinase inhibitor*	COL3A1, COL5A1, COL6A2, COL6A3, POSTN, THBS2, TNC, TNFAIP6.
Head and neck	Vorinostat	**HDACI**	ADAM12, CDH11, COL3A1, COL6A3, POSTN, TNFAIP6, VCAN, AEBP1.
	Nintedanib	Multi-kinase inhibitor	CDH11, COL3A1, COL5A1, COL6A3, POSTN TNFAIP6, VCAN.
	Cabergoline	Dopamine agonist	CDH11, COL5A1, POSTN, TNFAIP6, MPZL2.
	ALBENDAZOLE	Anthelmintic	AEBP1, CDH11, COL3A1, VCAN, CEACAM1.
	Ruxolitinib	JAK 1 and 2 inhibitor	COL3A1, COL5A1, COL6A3, FN1, POSTN.
Brain (GBM)	Vorinostat	**HDACI***	COL13A1, COL5A1, COL6A1, COL6A2, VCAN, GPR56, AEBP1, PERP.

	Saracatinib	Multi-kinase inhibitor*	ADAM12, COL5A1, FN1, GPR56, NEDD9, RHOB, FLRT3.
	Cyclosporin A	Calcineurin inhibitor ,Immunosuppressant	AEBP1, COL14A1, COL5A1, NEDD9, TNC.
	Parthenolide		AEBP1, COL3A1, COL5A1, FN1, NEDD9, RHOB.
	IVERMECTIN	Anthelmintic	ADAM12, COL5A1, NEDD9, PLXNC1, RHOB.
Adrenal	Vorinostat	**HDACI**	ADAM12, CXCL12, MFAP4.
	Trifluoperazine	Antipsychotic	LOXL2, AZGP1, CXCL12, FBLN5.
Renal	Vorinostat	**HDACI***	COL4A3, CTGF, FBLN5, RHOB.
	Topotecan	Topoisomerase inhibitor, Anticancer	AZGP1, COL14A1, CXCL12, DPT.
Ovarian	Vorinostat, Panobinostat	**HDACI**	ACTN1, FN1, POSTN, SRPX, THBS2.
	Itraconazole	Antifungal	CDH11, MFAP4, SLIT2, SRPX, THBS2.
	Doxorubicin*		NEDD9, PPAP2B, CDH11, COL3A1, FLRT2 , NID2, SRPX, THBS2.
Thyroid	Vorinostat	**HDACI***	POSTN, CDH2, CTGF, FBLN5, SRPX, COL4A3, CXCL12, MFAP4.
	Timolol	Antihypertensive	CDH2, FN1, DPT.
	Doxepin	Tricyclic antidepressant	CDH2, DPT, FLRT2.
	DIGOXIN	Cardiovascular, Inotropic	FN1, POSTN, CTGF.
	Trifluoperazine	Antipsychotic	FN1, CXCL12, FBLN5.
	Sildenafil	Vasoactive agent	CDH2, FN1, FLRT2.
	Azathioprine	Immunosuppressive	CDH2, CXCL12, DPT.
	Itraconazole	Antifungal	CTGF, MFAP4, SRPX.

*Enriched drugs in the extended ECM list are shown by asterisk.