NRF2 activates a partial Epithelial-Mesenchymal Transition and is maximally present in a hybrid Epithelial/Mesenchymal phenotype

Running title: NRF2 is a hallmark of a hybrid E/M phenotype

Federico Bocci*[,1,2], Satyendra C Tripathi*[,6], Jason T George[1,3,7], Julian P Casabar[6], Samuel A Vilchez Mercedes[8], Pak K Wong[8,9], Samir M Hanash[6], Herbert Levine[1,2,3,4,#], José N Onuchic[1,2,4,5#], Mohit Kumar Jolly[1,#]

[1]Center for Theoretical Biological Physics, Department of [2]Chemistry, [3]Bioengineering, [4]Physics and Astronomy, [5]Biosciences, Rice University, Houston, TX, USA;
[6]Department of Clinical Cancer Prevention, UT MD Anderson Cancer Center, Houston, TX, USA;
[7]Medical Scientist Training Program, Baylor College of Medicine, Houston, TX, USA;
[8]Department of Biomedical Engineering, The Pennsylvania State University, University Park, PA, USA; [9]Department of Mechanical Engineering and Department of Surgery, The Pennsylvania State University, University Park, PA, USA

*These authors equally contributed to this work.

[#]corresponding authors:
Herbert Levine: herbert.levine@rice.edu
José N Onuchic: jonuchic@rice.edu
Mohit Kumar Jolly: mkjolly.15@gmail.com

Keywords: NRF2, epithelial-mesenchymal transition, hybrid epithelial/mesenchymal (E/M) phenotype

Abstract

Epithelial-Mesenchymal Transition (EMT) is a key process implicated in cancer metastasis and therapy resistance. Recent studies have emphasized that cells can undergo partial EMT to attain a hybrid epithelial/mesenchymal (E/M) phenotype - a cornerstone of tumour aggressiveness and poor prognosis. These cells can have enhanced tumour-initiation potential as compared to purely epithelial or mesenchymal ones, and can integrate the properties of cell-cell adhesion and motility that facilitates collective cell migration in the form of clusters of Circulating Tumour Cells (CTCs) - the prevalent mode of metastasis. Thus, identifying the molecular players that can enable cells to maintain a hybrid E/M phenotype stably is crucial to curb the metastatic load. Here, using an integrated computational-experimental approach, we show that the transcription factor NRF2 can prevent a complete EMT and stabilize a hybrid E/M phenotype instead. Knockdown of NRF2 in hybrid E/M non-small cell lung cancer cells H1975 drove progression to a complete EMT at molecular and morphological levels. Further, we predict that NRF2 is maximally expressed in hybrid E/M phenotype(s), and demonstrate that this biphasic dynamics arises from the interconnections among NRF2 and the EMT regulatory circuit. Finally, clinical data from multiple datasets suggests a correlation between a hybrid E/M phenotype, high levels of NRF2 and its targets and poor survival, further strengthening the emerging notion that hybrid E/M phenotype(s) may occupy the 'metastatic sweet spot'.

Introduction

Epithelial-Mesenchymal Transition (EMT) is a hallmark of physiological and pathological processes such as embryonic development, wound healing and tumour progression [1]. During EMT, cells typically lose cell-cell adhesion and gain the traits of migration and invasion. However, EMT is not a binary process; instead, cells can also attain one or more hybrid epithelial/mesenchymal (E/M) phenotypes that may be more stem-like, metastatic and aggressive than those in completely epithelial or mesenchymal ones [1, 2]. Cells in hybrid E/M phenotype(s) display a mix of epithelial and mesenchymal traits, thereby enabling collective cell migration leading to the formation of clusters of Circulating Tumour Cells (CTCs) that can form a large majority of metastases, and correlate with poor clinical outcome [3].

Mechanisms enabling cancer cells to stably maintain hybrid E/M phenotype(s), however, remain largely unclear. Recent experimental and computational studies have suggested that a set of 'phenotypic stability factors' (PSFs) such as GRHL2, OVOL2, ΔNp63α and NUMB can help cells in stably maintaining a hybrid E/M phenotype by preventing 'cells that have gained partial plasticity' from undergoing a full EMT [4–9] and thus enabling collective cell migration. Knockdown of these PSFs drove the cells to a complete EMT *in vitro* and *in vivo,* while their exogenous overexpression can drive Mesenchymal-Epithelial Transition (MET) – the reverse of EMT [10]. Thus, the relative levels of PSFs and canonical EMT-inducing transcription factors (EMT-TFs) modulate the epithelial-mesenchymal plasticity of cells. A recent study highlighted the role of the transcription factor nuclear factor E2-related factor 2 (NFE2L2) – commonly referred to as NRF2 – in preventing cells from undergoing a full EMT during injury-induced wound healing [11]. Hence, we hypothesized that NRF2 can also act as a PSF.

NRF2 connects with EMT circuitry through the following links: a) NRF2 can inhibit EMT-TF Snail [12], b) members of miR-200 family – gatekeepers of an epithelial phenotype [13] – can increase NRF2 levels by targeting KEAP1 for ubiquitination and hence preventing the cytoplasmic degradation of NRF2 [14], and c) nuclear accumulation and transcriptional activity of NRF2 can be restricted by E-cadherin [15] (Fig. 1A); thus, it is possible that ZEB1 – a key EMT-TF that represses E-cadherin directly [16] – can effectively increase the nuclear levels and activity of NRF2.

Elucidating the emergent dynamics of the crosstalk among EMT and NRF2 signalling pathways requires constructing a mechanism-based quantitative model. Here, we incorporate the abovementioned connections of NRF2 with our previously developed mathematical framework for EMT signalling [17]. Our model predicts that NRF2 can act as a PSF for a hybrid E/M phenotype. To validate this prediction, we knocked down NRF2 in H1975 non-small cell lung cancer cells (NSCLC) that display a hybrid E/M phenotype stably over multiple passages *in vitro* [5]. Knockdown of NRF2 drove H1975 cells towards a mesenchymal phenotype at a molecular and morphological level. Furthermore, we suggest that NRF2 exhibits a biphasic response during EMT progression and reaches its maximal levels when the cell is in a hybrid E/M phenotype. This biphasic dynamics emerges from the topology of the network formed between NRF2 and core EMT circuit. Finally, we found correlations among elevated expression levels of NRF2 and its transcriptional targets, a hybrid E/M phenotype and poor patient survival in multiple clinical datasets, hence strengthening the emerging notion that a hybrid E/M phenotype, but not necessarily a completely mesenchymal one, can aggravate tumour progression.

Results

Mathematical modelling predicts that high NRF2 levels can stabilize and are a hallmark of a hybrid E/M phenotype

To elucidate the role of NRF2 in mediating EMT, we first investigate the crosstalk between NRF2 and a core network that regulates EMT (Fig. 1A). This core EMT network includes two families of EMT-TFs SNAIL and ZEB, and two families of microRNAs that inhibit EMT – miR-200 and miR-34 [17] (Fig. 1A). The intracellular dynamics of NRF2-EMT circuit is described via a system of coupled ordinary differential equations (ODEs) (see Methods) to investigate how NRF2 affects the response of a cell when it is exposed to an external EMT-inducing signal.

First, we calculate the bifurcation diagram of cellular EMT phenotypes in response to different levels of an external EMT-inducing signal (I_{ext}). For increasing values of I_{ext}, cells switch from an epithelial phenotype to a hybrid E/M phenotype to a mesenchymal phenotype, as displayed the decreasing levels of miR-200 (Fig. 1B) and increasing levels of EMT-TFs (Fig. S1A). In the presence of NRF2, a cell requires a much stronger stimulus to undergo a complete EMT and instead maintains a hybrid E/M phenotype over a much broader range of the external signal I_{ext}, as compared to the control case – i.e. NRF2 absent (Fig. 1B, compare the red and black curves). This result is quite robust upon variation of the coupling strength among NRF2 and the EMT circuit (Fig. S1B-C). Intriguingly, the model predicts that cells in a hybrid E/M phenotype have higher NRF2 levels as compared to those in epithelial or mesenchymal ones (Fig. 1C). Such behaviour emerges because both miR-200, an EMT-inhibiting microRNA, and ZEB, an EMT-TF, inhibit molecular players that, in turn, inhibit NRF2 (KEAP1 and E-Cadherin, respectively). Therefore, NRF2 receives an indirect activation from both mir-200 and ZEB, thereby leading to higher levels in a hybrid E/M phenotype.

Next, we investigate the temporal dynamics of a cell in response to the EMT-inducing signal (I_{ext}). In absence of NRF2, a cell starting from an epithelial state (i.e. high miR-200, low Zeb) first undergoes a partial EMT (i.e. transition to a hybrid E/M phenotype) and later attains a mesenchymal phenotype (Fig 1D, black curve) signal. The presence of NRF2, however, delays the transition to a fully mesenchymal state and maintains the hybrid E/M phenotype much longer (Fig. 1D, red curve). Concurrently, NRF2 increases during partial EMT and reduces once the cell attains a complete mesenchymal phenotype (Fig. 1E).

Our model therefore suggests that NRF2 prevents the transition from a hybrid E/M to a M phenotype, but not the transition from an E state to a hybrid E/M state. To further investigate this hypothesis, we perform multiple simulations exposing the cell to random levels of EMT-inducing signal I_{ext} and count the frequency of attaining E, E/M and M phenotypes, upon full equilibration of the system. We found an increased fraction of hybrid E/M cells and decreased fraction of M cells in presence of NRF2, while the frequency of the E phenotype remains unaltered (Fig. 1F), further strengthening our hypothesis.

Finally, to identify the robustness of the observed behaviour, we perform a sensitivity analysis where we varied the parameters of the model by $\pm$ 10% – one parameter at a time – and measured the corresponding change in the range of values of I_{ext} enabling a stable hybrid E/M phenotype (i.e. interval of x-axis where the hybrid E/M state is stable in Fig. 1B, as shown by the red solid middle branch). For most parameters, the relative change in this range of values of I_{ext} was small, indicating that the model is largely robust, except a strong sensitivity to the parameters related to NRF2 production and degradation, and those describing the inhibition of SNAIL by NRF2 (Fig. 1G). All these parameters control the overall effect of NRF2 on EMT, thus, the range of stability of hybrid E/M phenotype decreases. This decrease, however, is much weaker as compared to the control case (dotted line on Fig. 1G, i.e. NRF2 absent).

Thus, overall, these results suggest that NRF2 can act as a 'phenotypic stability factor' (PSF) for a hybrid E/M phenotype and its enhanced levels can be considered as a hallmark of the hybrid E/M phenotype.

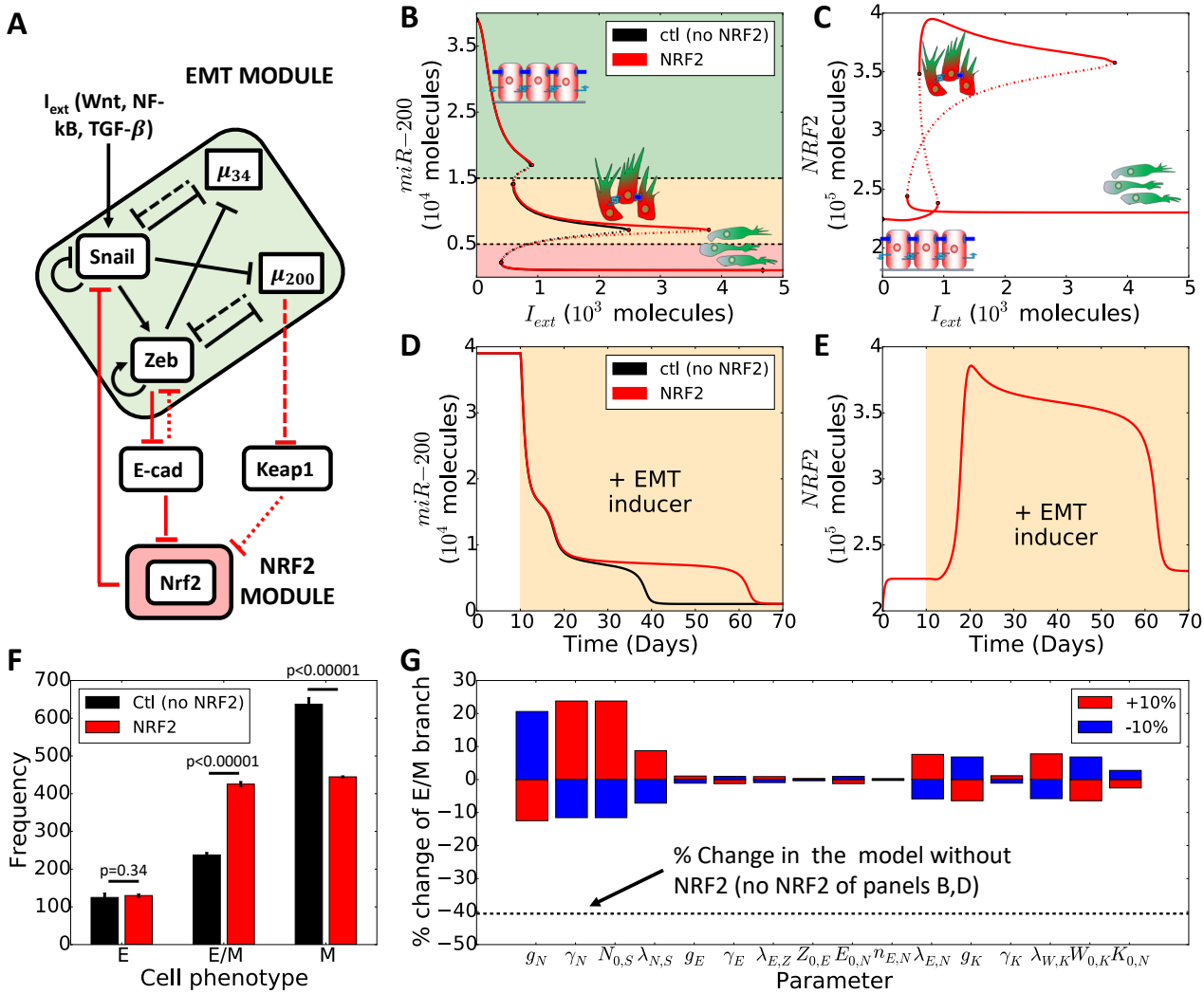


Figure 1. ***NRF2 stabilizes a hybrid E/M phenotype in presence of an external EMT-inducing signal. (A)*** *Schematic of EMT-NRF2 intracellular coupled regulatory network. Red arrows represent NRF2-related interactions that are not considered in the control case (i.e. NRF2 absent). Solid arrows represent transcriptional activation and solid bars stand for transcriptional inhibition. Dotted lines depict micro-RNA mediated translational inhibition. KEAP1 inhibits NRF2 by facilitating degradation while E-cadherin inhibits ZEB via β-catenin (fine-dotted arrows).* ***(B)*** *Bifurcation diagram of miR-200 levels as a function of I_{ext}. Curves denote the value of miR-200 upon equilibration, where continuous lines stand for stable steady states and dotted lines represent unstable solutions (black for the control case, red for NRF2-dependent case). Cartoons and colour shadings depict epithelial (green), hybrid E/M (yellow) and mesenchymal (red) phenotypes.* ***(C)*** *Bifurcation diagram of NRF2 levels as a function of I_{ext}.* ***(D)*** *Temporal dynamics of miR-200 levels in a cell starting in an epithelial phenotype, when exposed to a high level of $I_{ext} = 5000$ molecules (orange-shaded region).* ***(E)*** *Temporal dynamics of NRF2 levels, for case shown in D).* ***(F)*** *Fraction of E, hybrid E/M, and M cells for multiple simulations of a cell starting from random initial conditions and exposed to varied level of I_{ext} chosen randomly in $[0, 5000]$ molecules. Error bars indicate the standard deviation of three independent simulation sets (n=1000 each) for the control case (black) and the case with NRF2 (red). p-values were computed using a two-tailed paired t-test.* ***(G)*** *Variation of the interval of stability of the hybrid E/M phenotype in response to a 10% variation of the model parameters. Black dotted line indicates the relative change in the stability of hybrid E/M phenotype for the control case (NRF2 absent).*

NRF2 knockdown in H1975 NSCLC cells destabilizes a hybrid E/M phenotype *in vitro*

To investigate our prediction that NRF2 can act as a PSF for a hybrid E/M phenotype, we knocked down NRF2 using siRNAs in non-small cell lung cancer (NSCLC) H1975 cells that display a hybrid E/M phenotype stably *in vitro* over multiple passages [5].

We observed that knockdown of NRF2 increases the EMT-TFs ZEB1 and SNAIL and mesenchymal marker vimentin at both mRNA and protein levels, and decreased the mRNA levels of E-cadherin (in particular siNRF2 #2). The protein levels of E-cadherin, however, did not change significantly, potentially due to its long half-life [18] (Fig. 2A,B). These observations suggest that NRF2 knockdown can drive cells from a hybrid E/M phenotype to a mesenchymal phenotype, as predicted by the mathematical modelling (Fig. 2C). Consistently, bright-field microscopy revealed a cellular morphology characterized by more spindle-shaped traits characteristic of mesenchymal cells upon knockdown of NRF2 (Fig. 2D). Further, H1975 cells migrated individually, as opposed to a collectively, after NRF2 knockdown in transwell migration assays, thus reflecting a post-EMT transition state (Fig. 2E). Similarly, NRF2 knockdown compromised the cells' ability to move collectively and close the wound. While control cells almost exclusively moved collectively while maintaining physical cell-cell contacts, a subpopulation of KD-NRF2 cells also migrated solitarily, thus validating our hypothesis that a hybrid E/M cellular phenotype is being destabilized and cells progress towards a completely mesenchymal state (Fig. 2F).

Overall, NRF2 knockdown shows a progression to a complete EMT both at molecular and morphological levels, thus confirming that NRF2 can act as a PSF for a hybrid E/M phenotype.

Enhanced levels of NRF2 in a hybrid E/M phenotype emerge from the underlying network topology

To understand the detailed origin of the connection between NRF2 and EMT, we applied a recently developed computational analysis method - Random Circuit Perturbation (RACIPE) [19]. RACIPE takes the topology of a regulatory network as its input, and generates an ensemble of mathematical models in which parameters are randomly chosen. For every set of randomly chosen parameters, the algorithm returns the levels of all the considered variables upon full equilibration. Thus, it identifies robust gene expression profiles that can be expected from a given network topology. For the circuit we considered, the RACIPE output is an array containing the levels of (miR-34, miR-200, ZEB, SNAIL, E-Cadherin, KEAP1, NRF2). These levels are presented by the algorithm at a logarithmic scale and can therefore assume negative values.

After evaluating the NRF2-EMT circuit for 10000 randomly chosen sets of parameters, we performed a clustering of the obtained solutions via a K-means algorithm for K=3. Importantly, NRF2 was not considered for the clustering to allow an unbiased evaluation of the association of NRF2 levels with various clusters. Among the three clusters, one cluster (or group of solutions) is characterized by high expression of microRNAs and E-cadherin and low expression of EMT-TFs, and could thus be associated to an epithelial phenotype (Fig. 3A, bottom). The second cluster exhibits high expression of EMT-TFs and low microRNAs – signature for a mesenchymal phenotype (Fig. 3A, top). Finally, the third group has intermediate to high levels of both microRNAs and EMT-TFs and is thus interpreted as a hybrid E/M phenotype (Fig. 3A, center). Importantly, when plotted on miR-200 and ZEB axes, these clusters segregate elegantly and occupy different regions, suggesting that miR-200 and ZEB dynamics may play crucial roles in driving the emergence of these three phenotypic clusters (Fig. 3B) [17].

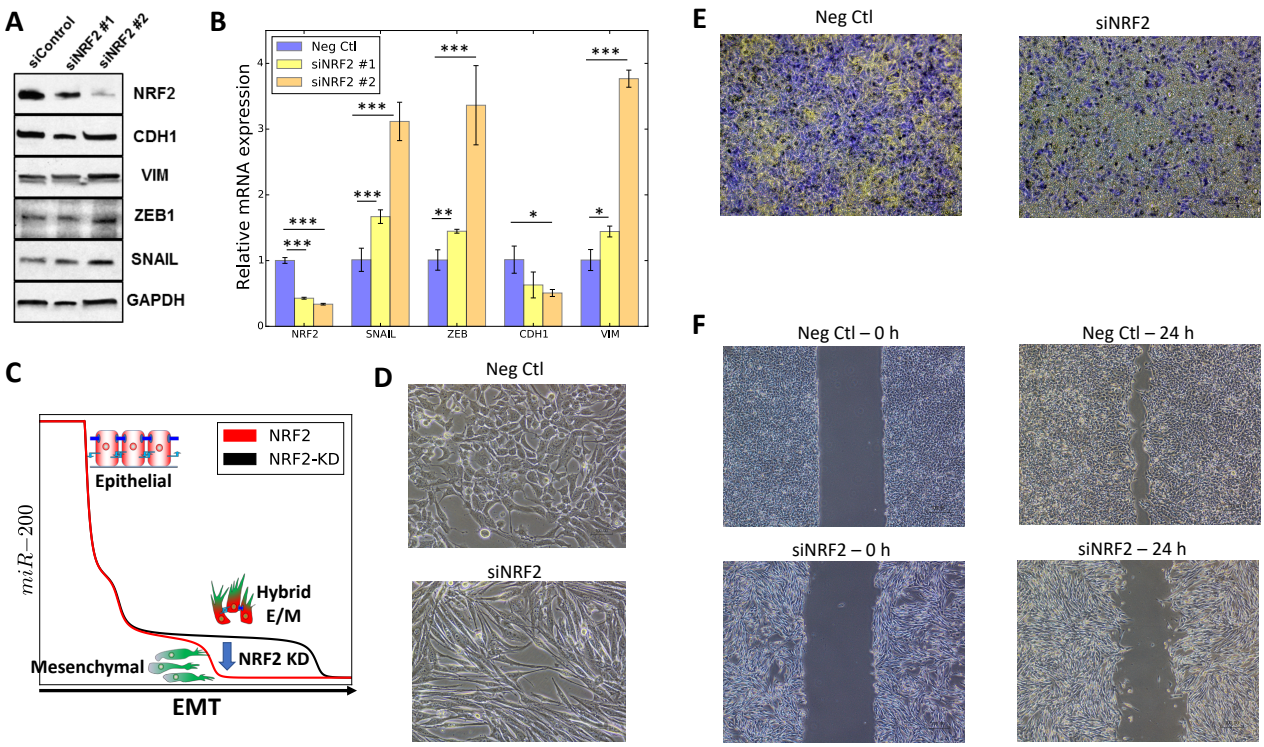


Figure 2. ***Knockdown of NRF2 induces a progression towards a complete EMT in H1975 cells. (A)*** *Western blot for NRF2, CDH1, VIM, ZEB1, SNAIL in H1975 cells. GAPDH is used as a loading control. Left column represents the control siRNA, second and third columns represent two independent siRNAs against NRF2.* ***(B)*** *mRNA levels of NRF2, SNAIL, ZEB, CDH1 (E-cadherin) and VIM (vimentin) in cells treated with siRNA against NRF2. *p < 0.05, **p<0.005, ***p<0.001 using two-tailed paired t-test.* **(C)** Temporal profile of miR-200 in response to an EMT-inducing signal in presence of NRF2 (red, same as 'NRF2' case Fig. 1D) and upon NRF2-KD (black) as predicted by mathematical modelling. NRF2 knockdown is modelled as a 90% reduction in the production rate constant of NRF2, and compromises the temporal stability of the hybrid E/M phenotype as depicted by the blue arrow. ***(D)*** *Bright-field morphology images for control (left) and NRF2-knockdown (right) H1975 cells.* ***(E)*** *Transwell migration assay showing trends of collective migration (left, Neg ctl) and individual migration (right, siNRF2). Violet colour represents cells while yellow colour represents membrane* ***(F)*** *Wound healing/scratch assay for control (neg ctl, top) and NRF2-KD (siNRF2, bottom) H1975 cells.*

Next, we computed average levels of NRF2 in these clusters, and showed that the hybrid E/M cluster expresses higher levels of NRF2 (Fig. 3C and S2A) as compared to epithelial or mesenchymal clusters, hence supporting our model's prediction (Fig. 3D). This analysis highlights that the maximal level of NRF2 is not an artefact enabled by a specific set of parameters, but instead emerges from the connections, or architecture of the network of interactions among NRF2 and the EMT module.

To better understand the underlying topological features enabling maximal levels of NRF2 in hybrid E/M state, we knocked down *in silico* either the inhibition of NRF2 by KEAP1 (KEAP1-NRF2 KD), or that by E-cadherin (Ecad-NRF2 KD). KEAP1-NRF2 KD maintained high NRF2 levels in the hybrid E/M cluster, but Ecad-NRF2 KD resulted in higher NRF2 in the epithelial cluster (Fig. S2B-C). Consistently, reducing the effect of KEAP1-mediated inhibition of NRF2 in our EMT-NRF2 mathematical model (shown in Fig 1A) maintained high NRF2 in the hybrid E/M state, while reducing E-Cadherin-mediated inhibition of NRF2 reduced the difference in NRF2 levels between

the E and hybrid E/M phenotypes (Fig. S3). This asymmetry can be perhaps attributed to a larger number of connections among ZEB, SNAIL, E-cadherin and NRF2.

Because EMT-NRF2 interplay involves modulation of NRF2 at post-transcriptional levels, further, we compared the protein levels of NRF2 in a panel of lung adenocarcinoma cell lines that have been characterized as epithelial, mesenchymal, and hybrid E/M, based on morphological and proteomic characteristics [20] (Supplementary Table 2). This analysis confirmed that high NRF2 levels tends to correlate with a higher probability to express a hybrid E/M phenotype (Fig. S4A-B). Additionally, we investigated NRF2 downstream target genes in a panel of 60 diverse human cancer cell lines (NCI-60) [21]. These cell lines have been categorized into epithelial, hybrid E/M, and mesenchymal, based on protein levels of E-cadherin and vimentin [13]. We observed that most NRF2 downstream targets, such as ABCB6, FTL or TALDO1 [22], were highly expressed in hybrid E/M cell lines as compared to epithelial or mesenchymal ones (Fig. S4C).

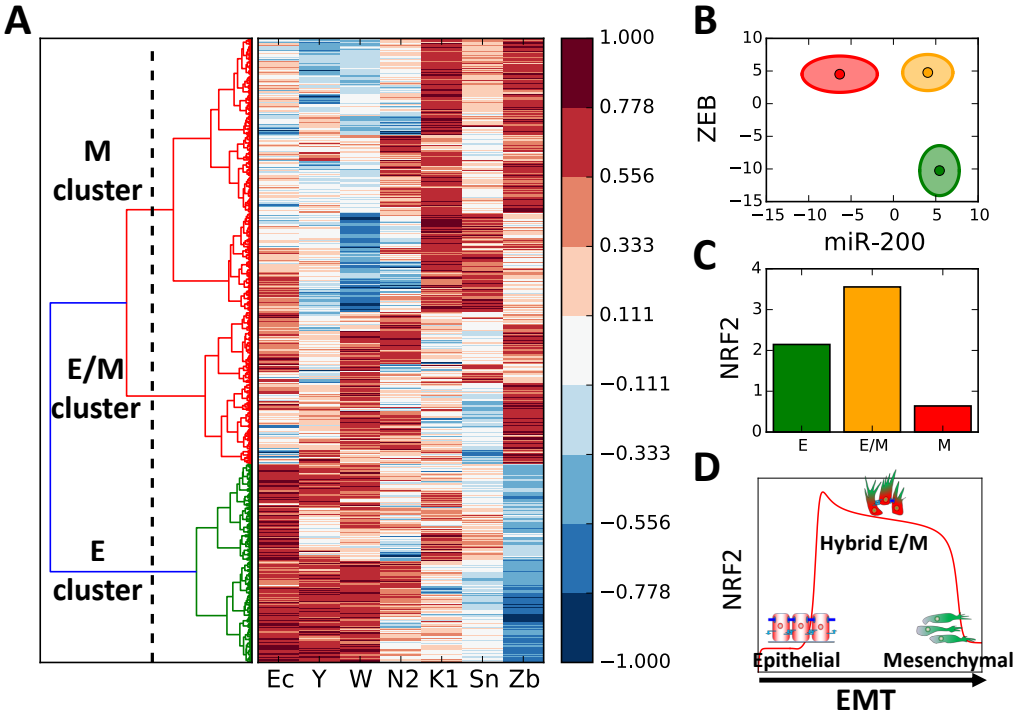


Figure 3. ***Random Circuit Perturbation of the EMT-NRF2 circuit.*** ***(A)*** *Right: gene expression profile of the states predicted by RACIPE. Each row represents a solution obtained with a different randomly chosen combination of model's parameters presented here as an array of the model's variables (Ec: E-cadherin, Y: miR-34, W: miR-200, N2: NRF2, K1: Keap1, Sn: SNAIL, Zb: ZEB). All gene expression levels are normalized between -1 and 1 to help eye comparison. Left: clustering of the solutions and threshold for K-means clustering (vertical dotted line leaving n=3 groups of solutions).* ***(B)*** *Representation of the three clusters in the miR-200/ZEB plane. The centers of the ellipses represent the cluster average while the axes quantify the cluster's standard deviation. Solutions belonging to the epithelial, hybrid E/M and mesenchymal clusters are plotted in green, yellow, and red, respectively.* ***(C)*** *Average level of NRF2 in the epithelial, hybrid E/M and mesenchymal clusters.* ***(D)*** *Temporal profile of NRF2 in response to an EMT-inducing signal as predicted by mathematical modelling (Fig. 1E) with relevant cartoons added.*

Clinical data suggests correlation between worse survival, high NRF2 levels and a hybrid E/M phenotype

The association of NRF2 with hybrid E/M phenotype(s) suggests addressing the role of NRF2 in a clinical context. Recent studies have emphasized that a hybrid E/M phenotype can be considered as a hallmark of tumour aggressiveness [1, 2] due to multiple reasons - it enables CTC dissemination as cluster instead of individual cells (typically mesenchymal), and it possesses enhanced tumour-initiation potential. Consistently, several previously discovered PSFs correlate with a poor clinical outcome. Among other examples, GRHL2 was associated with poor survival in different breast cancer subtypes [23] and Numb/Numbl correlated with a worse clinical outcome in lung and ovarian cancer [7].

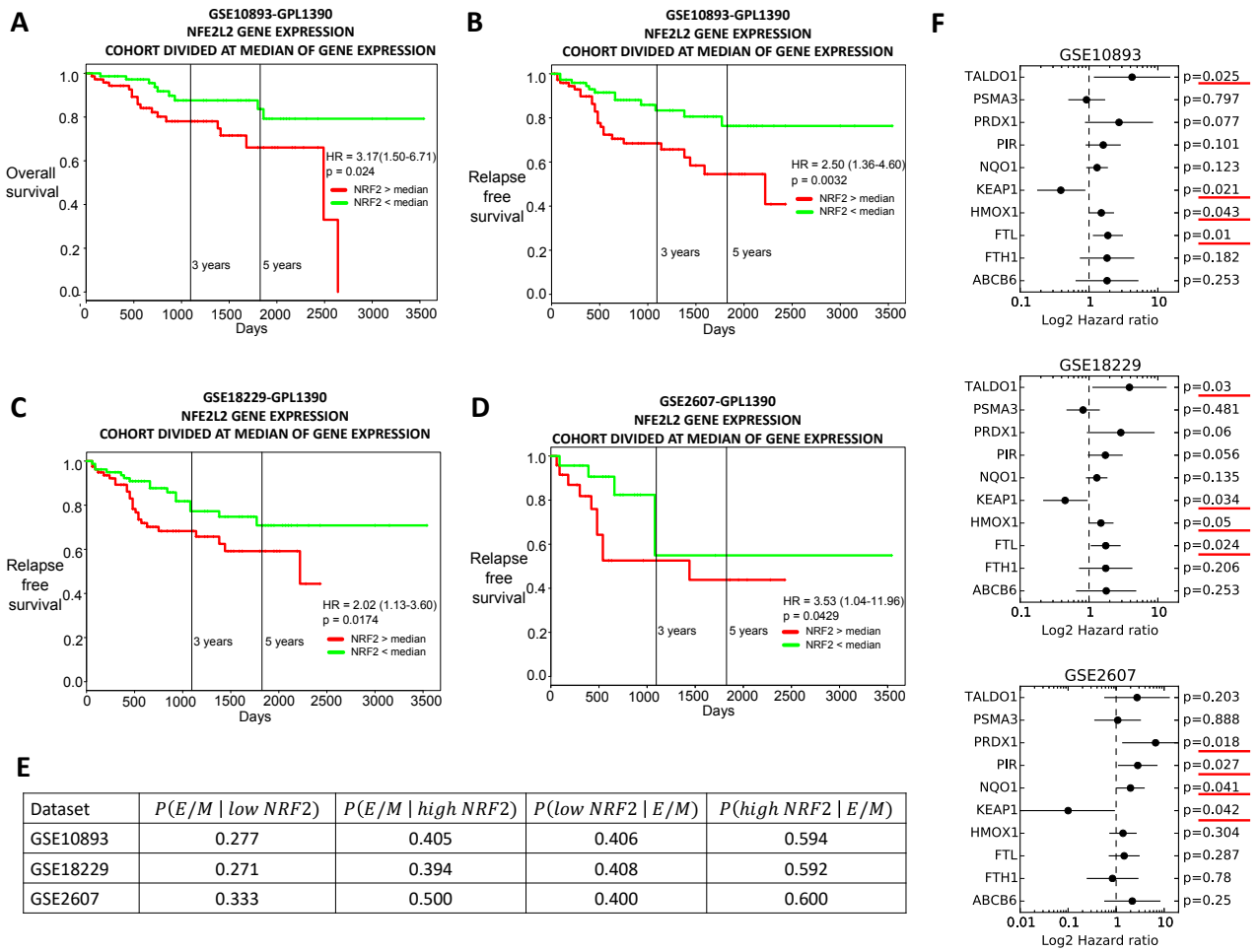
We analysed NRF2 expression levels in patients from multiple independent datasets, and observed that NRF2 correlates with a worse clinical outcome both in terms of overall survival (Fig. 4A) and relapse-free survival (Fig. 4B-D), when these samples are divided by median. We further calculated the 'EMT score' for all these patients' gene expression profiles using an inferential parameter developed by George *et al*. [24] based on the expression level of several key EMT regulators. These scores are calculated on a scale of 0 (highly epithelial) to 2 (highly mesenchymal), with the following categorization: epithelial (0<score<0.5), hybrid E/M (0.5<score<1.5) or mesenchymal (1.5<score<2). An EMT score distribution is thus reconstructed within each dataset for the 'low NRF2' group and the 'high NRF2' group (Fig. S5). We next calculated the conditional probabilities of patients belonging to 'high NRF2' or 'low NRF2' groups to have different EMT scores. This analysis revealed that a patient sample belonging to the 'high NRF2' group has a higher probability to display a hybrid E/M state (0.5< EMT score <1.5), as compared to exhibiting either an E or M state (Fig. 4E). Furthermore, a patient sample characterized as displaying a hybrid E/M state (0.5< EMT score <1.5) has a higher probability of belonging to the 'high NRF2' category as compared to belonging to the 'low NRF2' one (Fig. 4E). Both these observations are consistent across the three datasets where we observed a significant association between high NRF2 levels and poor patient survival.

Finally, due to a prominent regulation of NRF2 at post-transcriptional level by KEAP1 [14], we investigated the expression levels of NRF2 downstream targets [22]. High levels of many NRF2 downstream targets identified correlate with worse survival (hazard ratio > 1) significantly in these three considered datasets (Fig. 4F). Among those targets, only KEAP1 shows an opposite trend, i.e. its higher levels associated with better survival (Fig. 4F). This apparently counterintuitive finding can be rationalized because KEAP1 also acts as an inhibitor of NRF2, as also included in our mathematical model. Thus, put together, these results suggest a strong correlation among a hybrid E/M phenotype, high expression levels of NRF2 and its downstream targets, and a worse prognosis.

Discussion

EMT is a hallmark of many physiological and pathological processes contributing to morphogenesis, wound healing and cancer metastases [1]. With recently growing acknowledgement that EMT is not an 'all-or-none' process[1], much focus has moved to identifying 'phenotypic stability factors' (PSFs), i.e. players that can maintain a hybrid E/M phenotype stably, such as GRHL2, OVOL2, ΔNp63α and NUMB [5, 7–9] as well as understanding the unique traits of intermediate phenotypes [2, 25].

Here, we developed a mathematical framework to investigate the role of the transcription factor NRF2 (nuclear factor erythroid 2 related factor 2) in mediating EMT. NRF2 has been well-studied in the context of cellular redox status, where it acts as a regulator in response to diverse forms of stress [26], for instance by activating antioxidant response within the cell to respond to oxidative stress, thus promoting cell survival [27]. However, its role in mediating clustered cell migration – a proposed outcome of partial EMT[1] – has been only recently explored [11, 28].



E

Dataset	$P(E/M \mid low\ NRF2)$	$P(E/M \mid high\ NRF2)$	$P(low\ NRF2 \mid E/M)$	$P(high\ NRF2 \mid E/M)$
GSE10893	0.277	0.405	0.406	0.594
GSE18229	0.271	0.394	0.408	0.592
GSE2607	0.333	0.500	0.400	0.600

Fig. 4. ***NRF2 correlates with a worse clinical outcome in cancer patients. (A)*** *Overall survival for GSE10893. The group of patients with low NRF2 (green curve) performed better than the group with high NRF2 (red curve).* ***(B)*** *Relapse-free survival for GSE10893.* ***(C)*** *Relapse-free survival for GSE18229.* ***(D)*** *Relapse-free survival for GSE2607.* ***(E)*** *Left: conditional probability of exhibiting a hybrid E/M EMT score given a low NRF2 expression (second column) or a high NRF2 expression (third column). Right: conditional probability of a low NRF2 level (forth column) or a high NRF2 level (fifth column) given a hybrid E/M EMT score.* ***(F)*** *Hazard ratio, 95% confidence interval (error bars) and p value for several NRF2 downstream target genes for GSE10893 (top), GSE18229 (center) and GSE207 (bottom).*

These recent observations drove the development of our computational model which predicts that NRF2 can act as a PSF for the hybrid E/M phenotype. This result agrees well with the proposed role of NRF2 in mediating collective cell migration during wound healing [11], and its implications in restricting a complete EMT in pulmonary fibrosis [12]. A hybrid E/M or a partial EMT phenotype has been recently suggested to be potentially more aggressive than a complete EMT[1, 2]. Thus, the association of high levels of NRF2 with accelerated metastasis [29–31] and poor patient survival further [32] lends support to the idea of NRF2 acting as a PSF.

Our model further predicts that NRF2 follows a biphasic dynamics during EMT, i.e. its level increases upon EMT induction, it is maximally expressed in a hybrid E/M phenotype and decreases on completion of EMT. Therefore, NRF2 activation is proposed here as a hallmark of hybrid E/M phenotypes. While a direct and rigorous experimental validation of this prediction is still lacking, preliminary evidence on NRF2 dynamics during injury-induced collective cell migration [11], and the association of NRF2 with enhanced drug resistance [33] argue in favour of this

predicted dynamics, given the functional correlation between a hybrid E/M phenotype and acquired drug resistance [1]. Multi-drug resistance proteins have been identified as NRF2 target genes [34], and stable overexpression of NRF2 led to enhanced resistance of cancer cells to chemotherapeutic agents such as cisplatin, doxorubicin, and etoposide[35]. Our mathematical model predicts that such biphasic behaviour is conserved upon knockdown of NRF2 inhibition by KEAP1, but not upon knockdown of NRF2 inhibition by E-Cadherin. Thus, even in cases of mutations in KEAP1 and/or NRF2 that can impair NRF2-KEAP1 interactions, NRF2 is likely to be maximally expressed in and represent a highly aggressive hybrid E/M phenotype. This asymmetry can perhaps offer an explanation for why KEAP1 and/or NRF2 mutations still result in poor prognosis [36, 37].

Our earlier attempt at quantifying the association between EMT and stemness have predicted that cells in hybrid E/M states are more inclined to gain stem-like properties [25, 38], a prediction that has been recently validated *in vitro* and *in vivo* [2, 39, 40]. Reinforcing these predictions, NRF2 has been related to cancer stem-like traits and chemo-resistance [33]. Remarkably, a recent study reported a strong NRF2-mediated antioxidant response in hybrid E/M-like breast cancer stem cells (BCSC), and a significant association of NRF2 with poor prognosis [41]. These experimental observations resonate well with the worse clinical outcome and an association of a hybrid E/M phenotype with high NRF2 expression, and support the increasingly accepted notion that a hybrid E/M phenotype, rather than a completely mesenchymal one, may correlate with a higher metastatic potential [1].

To further investigate the implications of NRF2 – or any other PSF – in promoting clustered cell migration, future modelling efforts should be devoted toward integrating mechanics of cell motility with the regulatory biochemical signalling pathways. Such efforts will offer a more reliable quantitative relationship between NRF2 signalling axis and the aggressiveness of the disease.

Methods

Mathematical model formulation

The level of a protein, mRNA or micro-RNA (X) is described via a chemical rate equation that assumes the generic form:

$$\frac{dX}{dt} = \Gamma_X H^S(A, A_0, n, \lambda) - \gamma_X X \qquad (1)$$

In (eq. 1), the first term on the right-hand side ($\Gamma_X H^S(A, A_0, n, \lambda)$) represents the production rate of X expressed as a product of a basal production rate constant (Γ_X) and a function ($H^S(A, A_0, n, \lambda)$) that represents transcriptional/translational/post-translational regulations emerging from interactions with other components of the system. The second term ($-\gamma_X X$) is the degradation rate of X given by the product of a degradation rate constant (γ_X) and the number of molecules (X) (assuming first-order kinetics).

Specifically, an activator/inhibitor (A) of X alters the production rate constant according to a shifted Hill function[17], defined as:

$$H^S(A, A_0, n, \lambda) = \frac{1}{1 + \left(\frac{A}{A_0}\right)^n} + \lambda \frac{\left(\frac{A}{A_0}\right)^n}{1 + \left(\frac{A}{A_0}\right)^n} \qquad (2)$$

where the first argument (A) of the Hill function is the number of molecules of the inhibitor/activator while the second argument (A_0) is the half-maximal concentration parameter. Additionally, the Hill function is characterized by the Hill coefficient n that relates to how steep the regulation is, and the fold change λ ($\lambda < 1$ for inhibition, $\lambda > 1$ for activation). If multiple components modulate X, the production rate is a product of Hill functions. The NRF2-EMT model describes the dynamics of the circuit depicted in Fig. 1A (SI section 1) by extending the EMT circuit model proposed by Lu et al[17].

The newly-introduced parameters for this model are taken from relevant references [12, 14, 15, 42–44]. A qualitative rationale for parameter estimation is provided in SI section 2.

Analysis of clinical data
For all the examined datasets, the group of patients was divided into two sub-groups based on median NRF2 level. Kaplan-Meier analysis, i.e. hazard ratios and confidence intervals for NRF2 and its downstream targets for the two subgroups, was obtained with PROGgeneV2 software [45].

EMT score quantification
The EMT Metric [24] assigns to each sample categorical probabilities of membership in {E, E/M, M} by performing logistic regression against a collection of EMT-relevant predictor transcripts and cross-platform normalizer transcripts. The output of sample i may be represented by an ordered triple of probabilities $S_i = (P_E,\ P_{E/M}, P_M)$. Categorization is assigned based on the maximal value of this ordered triple. S_i is then projected onto $[0,2]$ by use of the EMT metric. The metric places epithelial and mesenchymal samples close to 0 and 2, respectively; while maximally hybrid E/M samples are assigned values close to 1.

Cell culture and transfection
H1975 cells were cultured in RPMI 1640 medium containing 10% fetal bovine serum and 1% penicillin/streptomycin cocktail (Thermo Fisher Scientific, Waltham, MA). Cells were transfected at a final concentration of 50 nM siRNA using Lipofectamine RNAiMAX (Thermo Fisher Scientific) according to the manufacturer's instructions using following siRNAs: siControl (Silencer Select Negative Control No. 1, Thermo Fisher Scientific), siNRF2 #1 (n290469, Silencer Select siRNA, Invitrogen), siNRF2 #2 (n290470, Silencer Select siRNA, Invitrogen).

RT-PCR
Total RNA was isolated following manufacturer's instructions using RNAeasy kit (Qiagen). cDNA was prepared using iScript gDNA clear cDNA synthesis kit (Bio-Rad). A TaqMan PCR assay was performed with a 7500 Fast Real-Time PCR System using TaqMan PCR master mix, commercially available primers, and FAM™-labeled probes for CDH1, VIM, ZEB1, NRF2, SNAIL, and VIC™-labeled probes for 18S, according to the manufacturer's instructions (Life Technologies). Each sample was run in triplicate. Ct values for each gene were calculated and normalized to Ct values for 18S (ΔCt). The ΔΔCt values were then calculated by normalization to the ΔCt value for control.

Western Blotting Analysis and Immunofluorescence
Cells were lysed in RIPA lysis assay buffer (Pierce) supplemented with protease and phosphatase inhibitor. The samples were separated on a 4–15% SDS-polyacrylamide gel (Biorad). After transfer to PVDF membrane, probing was carried out with primary antibodies and subsequent secondary antibodies. Primary antibodies were purchased from the following commercial sources: anti-CDH1 (1:1000; Cell Signaling Technology), anti-vimentin (1:1000; Cell Signaling Technology), anti-Zeb1 (1:1000; Cell Signaling Technology), anti-SNAIL (1:1000; Cell Signaling Technology), anti-NRF2 (1:1000; Abcam), and anti-GAPDH (1:10,000; Abcam). Membranes were exposed using the ECL method (GE Healthcare) according to the manufacturer's instructions.

Wound-Healing assay
Scratch wound-healing assay was performed to determine cell migration using confluent cultures (80%−90% confluence). Briefly, cells (1 × 105 cells/ml) were seeded in 6-well tissue culture plate and grown confluent. Cells were starved for 24 hours using 0.2% serum in growth media to minimize cell proliferation. Next day, the confluent monolayer was scratched with a sterile p200 pipet tip and media replaced by complete growth media. Images were acquired at 0 and 16 hours; the assay was performed at least twice per cell line.

Trans-well migration assay
H1975 cells were grown in 6-well plates and treated with siRNAs for NRF2 for 24 hours. After 48hrs of siRNA knockdown, cell monolayers were harvested, counted and cell concentration was

adjusted to 2×10^4 viable cells/200 µl of serum free medium. The cell suspension was seeded on top of a Transwell insert with 0.8 µm pore diameter (Millipore) and placed on a 24 well cell culture plate. At the bottom of the Transwell insert, 10% fetal bovine serum was added as chemo-attractant and plate was incubated at 37°C for 18 hours. Medium was aspirated from assay plate and Transwell inserts. Cells that did not migrate were removed by gently swabbing the inside of each insert using cotton swabs. After 3 washes with PBS, cells were fixed and stained with a 0.5% crystal violet solution for 10 minutes. After the cells have been stained, the inserts were washed thoroughly with water until the water runs clear. Inserts were dried completely before visualizing with a microscope.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Author's contribution
F. Bocci and M.K. Jolly conceived the research; F. Bocci (mathematical model, RACIPE simulation), S.C. Tripathi (NRF2 knockdown) developed the methodology; F. Bocci, J.T. George and M.K. Jolly analysed clinical datasets; all authors discussed results and participated in the manuscript preparation.

Acknowledgements
Federico Bocci thanks Dongya Jia and Bin Huang for helpful conversations about RACIPE simulations.

Grant support
This work was sponsored by the National Science Foundation NSF grants PHY-1427654 (Center for Theoretical Biological Physics), PHY-1605817, CHE- 1614101, MCB-1241332 and by the Cancer Prevention and Research Institute of Texas (CPRIT-grant R1110). FB is also supported by the Hasselman fellowship for excellence in Chemistry. JTG is also supported by the National Cancer Institute of the National Institutes of Health (F30CA213878). MKJ is also supported by a training fellowship from the Gulf Coast Consortia, on the Computational Cancer Biology Training Program (CPRIT Grant No. RP170593).

References

1. Jolly MK, Boareto M, Huang B, Jia D, Lu M, Ben-Jacob E, Onuchic JN, Levine H (2015) Implications of the hybrid epithelial/mesenchymal phenotype in metastasis. Front Oncol 5: . doi: 10.3389/fonc.2015.00155
2. Pastushenko I, Brisebarre A, Sifrim A, Fioramonti M, Revenco T, Boumahdi S, Van Keymeulen A, Brown D, Moers V, Lemaire S, De Clercq S, Minguijón E, Balsat C, Sokolow Y, Dubois C, De Cock F, Scozzaro S, Sopena F, Lanas A, D'Haene N, Salmon I, Marine J-C, Voet T, Sotiropoulou PA, Blanpain C (2018) Identification of the tumour transition states occurring during EMT. Nature 556:463–468 . doi: 10.1038/s41586-018-0040-3
3. Cheung KJ, Ewald AJ (2016) A collective route to metastasis: Seeding by tumor cell clusters. Science 352:167–169 . doi: 10.1126/science.aaf6546
4. Watanabe K, Villarreal-Ponce A, Sun P, Salmans ML, Fallahi M, Andersen B, Dai X (2014) Mammary morphogenesis and regeneration require the inhibition of EMT at terminal end buds by Ovol2 transcriptional repressor. Dev Cell 29:59–74 . doi: 10.1016/j.devcel.2014.03.006
5. Jolly MK, Tripathi SC, Jia D, Mooney SM, Celiktas M, Hanash SM, Mani SA, Pienta KJ, Ben-Jacob E, Levine H (2016) Stability of the hybrid epithelial/mesenchymal phenotype. Oncotarget 7:27067–84 . doi: 10.18632/oncotarget.8166
6. Dang TT, Esparza MA, Maine EA, Westcott JM, Pearson GW (2015) ΔNp63α promotes breast cancer cell motility through the selective activation of components of the epithelial-to-mesenchymal transition program. Cancer Res 75:3925–3935 . doi: 10.1158/0008-5472.CAN-14-3363
7. Bocci F, Jolly MK, Tripathi SC, Aguilar M, Onuchic N, Hanash SM, Levine H, Levine H (2017) Numb prevents a complete epithelial – mesenchymal transition by modulating Notch signalling. J R Soc Interface 14:20170512
8. Hong T, Watanabe K, Ta CH, Villarreal-Ponce A, Nie Q, Dai X (2015) An Ovol2-Zeb1 mutual inhibitory circuit governs bidirectional and multi-step transition between epithelial and mesenchymal states. PLoS Comput Biol 11:
9. Jolly MK, Boareto M, Debeb BG, Aceto N, Farach-Carson MC, Woodward WA, Levine H (2017) Inflammatory breast cancer: a model for investigating cluster-based dissemination. npj Breast Cancer 3:21 . doi: 10.1038/s41523-017-0023-9
10. Roca H, Hernandez J, Weidner S, McEachin RC, Fuller D, Sud S, Schumann T, Wilkinson JE, Zaslavsky A, Li H, Maher CA, Daignault-Newton S, Healy PN, Pienta KJ (2013) Transcription Factors OVOL1 and OVOL2 Induce the Mesenchymal to Epithelial Transition in Human Cancer. PLoS One 8:e76773 . doi: 10.1371/journal.pone.0076773
11. Riahi R, Long M, Yang Y, Dean Z, Zhang DD, Slepian MJ, Wong PK (2014) Single cell gene expression analysis in injury-induced collective cell migration. Integr Biol 6:192–202 . doi: 10.1039/C3IB40095F
12. Zhou W, Mo X, Cui W, Zhang Z, Li D, Li L, Xu L, Yao H, Gao J (2016) Nrf2 inhibits epithelial-mesenchymal transition by suppressing snail expression during pulmonary fibrosis. Sci Rep 6:38646 . doi: 10.1038/srep38646
13. Park S-M, Gaur AB, Lengyel E, Peter ME (2008) The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. Genes Dev 22:894–907
14. Eades G, Yang M, Yao Y, Zhang Y, Zhou Q (2011) miR-200a regulates Nrf2 activation by targeting Keap1 mRNA in breast cancer cells. J Biol Chem 286:40725–40733 . doi: 10.1074/jbc.M111.275495
15. Kim WD, Kim YW, Cho IJ, Lee CH, Kim SG (2012) E-cadherin inhibits nuclear accumulation of Nrf2: implications for chemoresistance of cancer cells. J Cell Sci 125:1284–1295 . doi: 10.1242/jcs.095422
16. Sánchez-Tilló E, Lázaro A, Torrent R, Cuatrecasas M, Vaquero EC, Castells A, Engel P, Postigo A (2010) ZEB1 represses E-cadherin and induces an EMT by recruiting the SWI/SNF chromatin-remodeling protein BRG1. Oncogene 29:3490–3500 . doi: 10.1038/onc.2010.102
17. Lu M, Jolly MK, Levine H, Onuchic JN, Ben-Jacob E (2013) MicroRNA-based regulation of

epithelial-hybrid-mesenchymal fate determination. Proc Natl Acad Sci U S A 110:18174–9 . doi: 10.1073/pnas.1318192110
18. de Beco S, Gueudry C, Amblard F, Coscoy S (2009) Endocytosis is required for E-cadherin redistribution at mature adherens junctions. Proc Natl Acad Sci 106:7010–7015 . doi: 10.1073/pnas.0811253106
19. Huang B, Lu M, Jia D, Ben-Jacob E, Levine H, Onuchic JN (2017) Interrogating the topological robustness of gene regulatory circuits by randomization. PLoS Comput Biol 13:1–21 . doi: 10.1371/journal.pcbi.1005456
20. Schliekelman MJ, Taguchi A, Zhu J, Dai X, Rodriguez J, Celiktas M, Zhang Q, Chin A, Wong C-H, Wang H, McFerrin L, Selemat S, Yang C, Kroh E, Garg K, Behrens C, Gazdar A, Laird-Offringa J, Tewari M, Wistuba J, Thiery JP, Hanash SM, others (2015) Molecular portraits of epithelial, mesenchymal, and hybrid States in lung adenocarcinoma and their relevance to survival. Cancer Res 75:1789–1800 . doi: 10.1158/0008-5472.CAN-14-2535
21. Shankavaram UT, Varma S, Kane D, Sunshine M, Chary KK, Reinhold WC, Pommier Y, Weinstein JN (2009) CellMiner: A relational database and query tool for the NCI-60 cancer cell lines. BMC Genomics 10: . doi: 10.1186/1471-2164-10-277
22. Chorley BN, Campbell MR, Wang X, Karaca M, Sambandan D, Bangura F, Xue P, Pi J, Kleeberger SR, Bell DA (2012) Identification of novel NRF2-regulated genes by ChiP-Seq: Influence on retinoid X receptor alpha. Nucleic Acids Res 40:7416–7429 . doi: 10.1093/nar/gks409
23. Mooney SM, Talebian V, Jolly MK, Jia D, Gromala M, Levine H, McConkey BJ (2017) The GRHL2/ZEB Feedback Loop—A Key Axis in the Regulation of EMT in Breast Cancer. J Cell Biochem 118:2559–2570 . doi: 10.1002/jcb.25974
24. George JT, Jolly MK, Xu S, Somarelli JA, Levine H (2017) Survival Outcomes in Cancer Patients Predicted by a Partial EMT Gene Expression Scoring Metric. Cancer Res 77:6415–6428 . doi: 10.1158/0008-5472.CAN-16-3521
25. Jolly MK, Huang B, Lu M, Mani SA, Levine H, Ben-Jacob E (2014) Towards elucidating the connection between epithelial--mesenchymal transitions and stemness. J R Soc Interface 11:20140962
26. D.Hayes J, T.Dinkova-Kostova A (2014) The Nrf2 regulatory network provides an interface between redox and intermediary metabolism. Trends Biochem Sci 39:199–218 . doi: https://doi.org/10.1016/j.tibs.2014.02.002
27. Jaramillo MC, Zhang DD (2013) The emerging role of the Nrf2 – Keap1 signaling pathway in cancer. Genes Dev 27:2179–2191 . doi: 10.1101/gad.225680.113.
28. Pan H, Wang H, Zhu L, Mao L, Qiao L, Su X (2013) The Role of Nrf2 in Migration and Invasion of Human Glioma Cell U251 Nrf2 in U251. World Neurosurg 80:363–370 . doi: 10.1016/j.wneu.2011.06.063
29. Wang H, Liu X, Long M, Huang Y, Zhang L, Zhang R, Zheng Y, Liao X, Wang Y, Liao Q, Li W, Tang Z, Tong Q, Wang X, Fang F, Vega D, Ouyang Q, Zhang DD, Yu S, Zheng H (2016) NRF2 activation by antioxidant antidiabetic agents accelerates tumor metastasis. Sci Transl 8:1–13
30. Zhang C, Wang H, Bao Q, Wang L, Guo T (2016) NRF2 promotes breast cancer cell proliferation and metastasis by increasing RhoA / ROCK pathway signal transduction. Oncotarget 7:73593–73606
31. Chang L, Fan C, Tseng W, Chein H, Hsieh T, Chen J, Hwang C, Hua C (2016) The Ratio of Hmox1 / Nrf2 mRNA Level in the Tumor Tissue Is a Predictor of Distant Metastasis in Colorectal Cancer. Dis Markers. doi: http://dx.doi.org/10.1155/2016/8143465
32. Zhang M, Zhang C, Zhang L, Yang Q, Zhou S, Wen Q, Wang J (2015) Nrf2 is a potential prognostic marker and promotes proliferation and invasion in human hepatocellular carcinoma. BMC Cancer 15:531 . doi: 10.1186/s12885-015-1541-1
33. Ryoo IG, Lee SH, Kwak MK (2016) Redox Modulating NRF2: A Potential Mediator of Cancer Stem Cell Resistance. Oxid Med Cell Longev. doi: 10.1155/2016/2428153
34. Rachakonda G, Sekhar KR, Jowhar D, Samson PC, Wikswo JP, Beauchamp RD, Datta PK, Freeman ML (2010) A promising target towards lung cancer prevention and therapeutics. Oncogene 29:3703–3714 . doi: 10.1038/onc.2010.118
35. Wang X, Sun Z, Villeneuve NF, Zhao F, Li Y, Chen W, Yi X, Zheng W, Wondrak GT, Wong

PK, Zhang DD (2008) Nrf2 enhances resistance of cancer cells to chemotherapeutic drugs , the dark side of Nrf2. Carcinogenesis 29:1235–1243 . doi: 10.1093/carcin/bgn095
36. Tong Y-H, Zhang B, Fan Y, Lin N-M (2015) Keap1–Nrf2 pathway: A promising target towards lung cancer prevention and therapeutics. Chronic Dis Transl Med 1:175–186 . doi: 10.1016/j.cdtm.2015.09.002
37. Hast BE, Cloer EW, Goldfarb D, Li H, Siesser PF, Yan F, Walter V, Zheng N, Hayes DN, Major MB (2014) Cancer-derived mutations in KEAP1 impair NRF2 degradation but not ubiquitination. Cancer Res 74:808–817 . doi: 10.1158/0008-5472.CAN-13-1655
38. Bocci F, Jolly MK, George JT, Levine H, Onuchic JN (2018) A mechanism-based computational model to capture the interconnections among epithelial-mesenchymal transition, cancer stem cells and Notch-Jagged signaling. Oncotarget 9: . doi: 10.18632/oncotarget.25692
39. Grosse-Wilde A, d'Hérouël AF, McIntosh E, Ertaylan G, Skupin A, Kuestner RE, del Sol A, Walters K-A, Huang S (2015) Stemness of the hybrid epithelial/mesenchymal state in breast cancer and its association with poor survival. PLoS One 10:e0126522
40. Hojo N, Huisken AL, Wang H, Chirshev E, Kim NS, Nguyen SM, Campos H (2018) Snail knockdown reverses stemness and inhibits tumour growth in ovarian cancer. Sci Rep 1–10 . doi: 10.1038/s41598-018-27021-z
41. Luo M, Shang L, Brooks MD, Jiagge E, Zhu Y, Buschhaus JM, Conley S, Fath MA, Davis A, Gheordunescu E, Wang Y, Harouaka R, Lozier A, Triner D, McDermott S, Merajver SD, Luker GD, Spitz DR, Wicha MS (2018) Targeting Breast Cancer Stem Cell State Equilibrium through Modulation of Redox Signaling. Cell Metab 28:69–86
42. Mooney SM, Jolly MK, Levine H, Kulkarni P (2016) Phenotypic plasticity in prostate cancer: role of intrinsically disordered proteins. Asian J Androl 18:704 . doi: 10.4103/1008-682X.183570
43. Zeiler M, Straube WL, Lundberg E, Uhlen M, Mann M (2012) A Protein Epitope Signature Tag (PrEST) Library Allows SILAC-based Absolute Quantification and Multiplexed Determination of Protein Copy Numbers in Cell Lines. Mol Cell Proteomics 11:1–13 . doi: 10.1074/mcp.O111.009613
44. Eden E, Geva-zatorsky N, Issaeva I, Cohen A, Dekel E, Cohen L, Mayo A, Alon U, Eden E, Issaeva I, Cohen A, Dekel E, Danon T (2011) Proteome Half-Life Dynamics in Living Human Cells. Science 331:764–768
45. Goswami CP, Nakshatri H (2014) PROGgeneV2 : enhancements on the existing database. BMC Cancer 14: