

<https://doi.org/10.1038/s41540-024-00378-w>

# Comprehensive molecular interaction map of TGF $\beta$ induced epithelial to mesenchymal transition in breast cancer

Check for updates

Sai Bhavani Gottumukkala <sup>1</sup>, Trivadi Sundaram Ganesan <sup>2</sup> & Anbumathi Palanisamy <sup>1</sup>

Breast cancer is one of the prevailing cancers globally, with a high mortality rate. Metastatic breast cancer (MBC) is an advanced stage of cancer, characterised by a highly nonlinear, heterogeneous process involving numerous signalling pathways and regulatory interactions. Epithelial–mesenchymal transition (EMT) emerges as a key mechanism exploited by cancer cells. Transforming Growth Factor- $\beta$  (TGF $\beta$ )-dependent signalling is attributed to promote EMT in advanced stages of breast cancer. A comprehensive regulatory map of TGF $\beta$  induced EMT was developed through an extensive literature survey. The network assembled comprises of 312 distinct species (proteins, genes, RNAs, complexes), and 426 reactions (state transitions, nuclear translocations, complex associations, and dissociations). The map was developed by following Systems Biology Graphical Notation (SBGN) using Cell Designer and made publicly available using MINERVA ([http://35.174.227.105:8080/minerva/?id=Metastatic\\_Breast\\_Cancer\\_1](http://35.174.227.105:8080/minerva/?id=Metastatic_Breast_Cancer_1)). While the complete molecular mechanism of MBC is still not known, the map captures the elaborate signalling interplay of TGF $\beta$  induced EMT-promoting MBC. Subsequently, the disease map assembled was translated into a Boolean model utilising CaSQ and analysed using Cell Collective. Simulations of these have captured the known experimental outcomes of TGF $\beta$  induced EMT in MBC. Hub regulators of the assembled map were identified, and their transcriptome-based analysis confirmed their role in cancer metastasis. Elaborate analysis of this map may help in gaining additional insights into the development and progression of metastatic breast cancer.

Metastatic breast cancer (MBC) is a lethal form of breast cancer with high incidence and mortality rates<sup>1,2</sup>. MBC is a heterogeneous disease characterised by differences between and within tumours. It is primarily distinguished based on a variety of characteristic markers such as ER<sup>+/-</sup>, PR<sup>+/-</sup>, HER<sup>+/-</sup>, Ki-67. MBC arises as a multistep process with cancer cells spreading to various organs such as bone, brain, lung, and liver<sup>1,3,4</sup>. Epithelial to mesenchymal transition (EMT) is a critical development process that significantly contributes to the initiation and progression of metastasis depending upon the microenvironmental cues<sup>5</sup>. During EMT, non-invasive breast cancer transforms into invasive breast cancer by losing their polarised epithelial assets and acquiring invasive migratory properties of mesenchymal stem cells<sup>6-8</sup>. The progression of EMT is characterised by the status of tight junctions, cell polarity, expression of various epithelial (E-Cadherin,  $\beta$ -Catenin, p-120, few

miRNAs) and mesenchymal markers (SNAIL, ZEB, TWIST, N-Cadherin) regulated through multiple signalling pathways.

Transforming growth factor  $\beta$  (TGF $\beta$ ) and its family of cytokines are well known for their stage-dependent dual role as tumour suppressors in the early stages of cancer and as promoters of proliferation at a later stage<sup>9-21</sup>. The role of TGF $\beta$  is recognised in many cancer types, if not all<sup>12,16,19,22-27</sup>. It regulates various physiological and biological processes such as cell proliferation, apoptosis, differentiation, and migration<sup>11,19,21,28-31</sup>. Experimental studies have identified the significant role of TGF $\beta$  signalling in promoting EMT in breast cancer epithelial cells leading to MBC<sup>14,32-37</sup>. This regulation emerges because of the activation of its serine/threonine receptors through phosphorylation. This consequently activates a signalling cascade downstream including SMAD-dependent and SMAD-independent pathways which are involved in modulating the process of EMT in breast cancer

<sup>1</sup>Department of Biotechnology, National Institute of Technology Warangal, Warangal, India. <sup>2</sup>Department of Medical Oncology, Sri Ramachandra Institute of Higher Education and Research, Chennai, India. e-mail: [anbu@nitw.ac.in](mailto:anbu@nitw.ac.in)

cells<sup>11,19,28,38–41</sup>. Thus, the initiation and progression of TGF $\beta$  induced metastasis in breast cancer are regulated through multiple signalling molecules and mechanisms.

Genetic heterogeneity plays a pivotal role in driving metastasis and therapeutic relapse in cancer. This is influenced by the spatio-temporal expression of a multitude of genes in response to external stress<sup>42</sup>. This vast molecular heterogeneity and cellular characteristics exhibited by MBC across different patients present a significant challenge in identifying a reliable biomarker for EMT. The current study provides a comprehensive understanding of TGF $\beta$  induced EMT in MBC. This work emphasises cancer as a network-level disease that involves multiple signalling pathways and their intricate molecular interactions. Pathway databases like IMEx consortium, Signor, IntAct, OMNIpath Cancer Genome Atlas provide diagrammatic representations of the mechanisms implicated in complex signalling pathways associated with various diseases<sup>43–45</sup>. The ever-increasing need to understand the molecular players, mechanisms in complex diseases such as cancer underscores the growing need for disease maps<sup>46</sup>. The rise in the number of disease maps in the past two decades indicates the need for such maps which may guide the computational pipelines, develop experimental hypothesis and support drug-related studies. There are consortiums that focus on the development, quality, standards and periodical update of such maps. The present study specifically focuses on the functional interactions between downstream effectors of TGF- $\beta$  signaling that contribute to the development of EMT-driven MBC. The molecular interaction mechanisms were explored in detail. The map developed may serve the community to develop hypotheses and may assist in exploring them experimentally or theoretically.

## Results

### Pathway map of TGF $\beta$ induced EMT in MBC

The comprehensive map developed for TGF $\beta$  induced EMT signalling in MBC is shown in Fig. 1. The map was assembled based on an extensive literature search (Fig. 10). The map captures the heterogeneous molecular processes and regulators implicated in TGF $\beta$  mediated signalling, particularly in the context of EMT and metastasis of breast cancer (Supplementary Table 1 and Supplementary Table 3). The MBC map comprises 312 distinct species interconnected through 426 reactions across two compartments (cytoplasm and nucleus). It also encompasses 160 proteins, 30 genes, 48 RNAs, 57 complexes, 11 degradations, and 5 unknowns altogether resulting in one end phenotype of epithelial–mesenchymal transition (EMT). A range of molecular processes including 309 state transitions (activation and inhibition), 48 complex associations, 2 dissociations, 10 transports, 22 unknown transitions, 29 logical activations were categorised among the species. The “unknown transitions” in the map indicate instances where regulatory interactions are acknowledged, but the precise underlying mechanisms require additional experimental exploration for a comprehensive understanding. “Unknown” regulators represent intermediates in these processes that warrant further investigation.

The map was constructed using Cell Designer tool V 4.4 (<http://celldesigner.org/>). This map strictly adheres to the Systems Biology Graphics Notation (SBGN), ensuring an accurate representation of receptors, proteins, genes, RNAs, their modifications, as well as the complex associations and dissociations. The map portrays these entities as species, while their interactions in regulating the process of EMT are illustrated as reactions. Further, the storage of the map was implemented using Systems Biology Markup Language (SBML), an XML-based format widely used for communication<sup>47</sup>. All species names used in this study, their supporting literature and their corresponding HUGO names are shown (Supplementary Table 1). The map was further made available online using the MINERVA platform for active exploration, analysis and management (Fig. 2) and is available at: [http://35.174.227.105:8080/minerva/?id=Metastatic\\_Breast\\_Cancer\\_1](http://35.174.227.105:8080/minerva/?id=Metastatic_Breast_Cancer_1).

### Architecture of TGF $\beta$ induced EMT in MBC

In the developed map (Fig. 1), TGF $\beta$  stimulus acts as the input triggering the signalling cascade through activation of their serine/threonine kinase

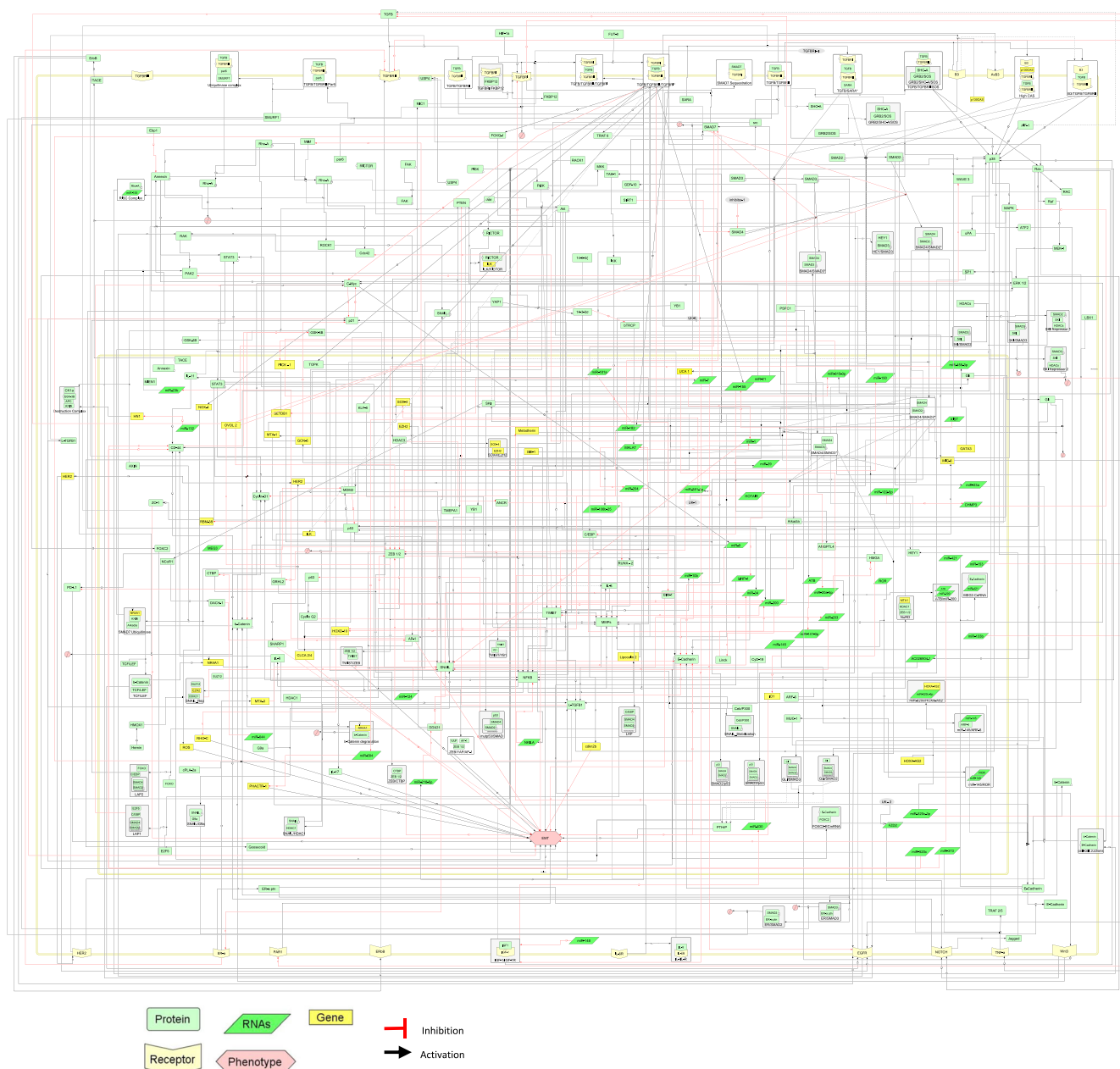
receptors. This subsequently regulates the SMAD-dependent as well as SMAD-independent signalling pathways such as RHO, RAS, P38, MAPK, STAT3, Wnt and EGFR. The map developed was further illustrated elaborately as zones (Supplementary Figs. 1–6). TGF $\beta$  induced SMAD-dependent and independent signalling pathways were observed to be integrating into the major regulators namely SNAIL, ZEB, E-Cadherin, N-Cadherin,  $\beta$ -Catenin, NF- $\kappa$ B, Twist, MMPs (Supplementary Figs. 7–10) in orchestrating the process of EMT and promoting metastasis of breast cancer. Supplementary Figs. 7–10 provide simplified illustrations around the major EMT regulators offering a clearer visual representation of the interactions and relationships between these key regulators. During the assembly process, efforts were made to maintain the closeness of components related to the SMAD-dependent and independent cascades. SMAD-independent pathway branches into multiple signalling pathways involving common regulators and interconnectedness between them. The regulators shared in both pathways makes the overall map highly nonlinear.

### SMAD-dependent signalling-based modulation of EMT in MBC

Mechanistically, TGF $\beta$  and its isoforms stimulate the SMAD signalling through the activation of its RTKs. TGF $\beta$  binds with its high-affinity TGF $\beta$ RII receptor, causing a conformation change that creates a high-affinity binding site for TGF $\beta$ RI<sup>48</sup>. This binding results in the activation of receptors through phosphorylation of TGF $\beta$ RI in the juxta membrane subdomain by TGF $\beta$ RII<sup>49</sup>. The activation of TGF $\beta$ RI depends on the dissociation of FKBP12, a negative regulator of TGF $\beta$ RI through the allosteric change in receptor confirmation<sup>50,51</sup>. Thus, the active form of the complex includes TGF $\beta$  ligand, and its phosphorylated type I and type II receptors. Upon activation, TGF $\beta$ RI releases the SARA sequestered R-SMADs ensuring the phosphorylation of R-SMADs, particularly SMAD2 and SMAD3. The phosphorylated R-SMADs form a trimeric complex with Co-SMAD (SMAD4), which translocates into the nucleus modulating gene expression.

Subsequently, the translocated SMAD complex (R-SMAD/Co-SMAD) associates with co-factors to induce the effects of TGF $\beta$  stimulus either as tumour suppressor or as a promoter by regulating cell proliferation and/or migration. SMAD complex in association with FOXO, LAP (an inhibitory isoform of c/EBP $\beta$ ) activates the tumour suppressor role of TGF $\beta$ , while LIP is an isoform of c/EBP $\beta$  which endorses the tumour promoting role of TGF $\beta$  by inhibiting LAP<sup>52</sup>. Similarly, 14-3-3 $\zeta$  was known to be overexpressed in malignant cells shifting the role of TGF $\beta$  from tumour suppressor to tumour activator through the activation of transcription factor Gli. Gli acts as a decisive partner of SMAD regulating PTHrP in fostering bone metastasis of breast cancer<sup>53</sup>. Elevated levels of PSPC1 in cancer cells result in the tumour-promoting role of TGF $\beta$  by its association with R-SMADs further promoting autocrine signalling through the activation of L-TGF $\beta$ <sup>54</sup>. Furthermore, GDF-10 activated by TGF $\beta$ /SMAD signalling, promotes the expression of SMAD7 and regulates the anti-proliferative effects suppressing EMT in TNBC cell lines<sup>55</sup>. SMAD complex associates with transcription factor AP-1 regulating the expression of genes like IL-11, CTGF, PTHrP, CXCR4 and thus promotes metastasis<sup>33,56,57</sup>. Mutant p53 as a cofactor of SMAD complex, regulates the role of SHARP, Cyclin G2 through p63 moderating metastasis<sup>58</sup>. ZO-1 an epithelial marker was known to be inhibited by TGF $\beta$ /SMAD signalling in promoting EMT<sup>59</sup>. MMPs regulated by TGF $\beta$  play a crucial role in maintaining the integrity of epithelial cells<sup>60</sup>. EMT was also observed to be altered by various negative feedback mechanisms regulated through TGF $\beta$ /SMAD-induced TMEPA1, SKI and Arkadia<sup>37,42,61,62</sup>. Thus, depending on the co-factors associated with SMAD-dependent signalling can support the anti-proliferative and proliferative role of TGF $\beta$ .

In addition to the direct interactions and feedback mechanisms discussed above, experimental evidence from TGF $\beta$ /SMAD signalling shows that the mutations in SMAD2 and SMAD4 are uncommon in breast cancer thus making SMAD3 a key player in the process of EMT<sup>33,56</sup>. However, the specific R-SMAD involved in the TGF $\beta$ /SMAD signalling is unknown. Thus, TGF $\beta$ /SMAD signalling modulates the process of EMT directly or indirectly in the presence of appropriate co-factors.



**Fig. 1 | Molecular regulatory map of metastatic breast cancer.** The map represents proteins in green, RNAs in lime green, genes in canary, and phenotypes in pink. Interactions among regulators are displayed in black and inhibitions are displayed in red. Compartments are distinguished as bounding boxes. The SBGN-compliant map consists of 312 species, 426 interactions built using Cell Designer V 4.4. The

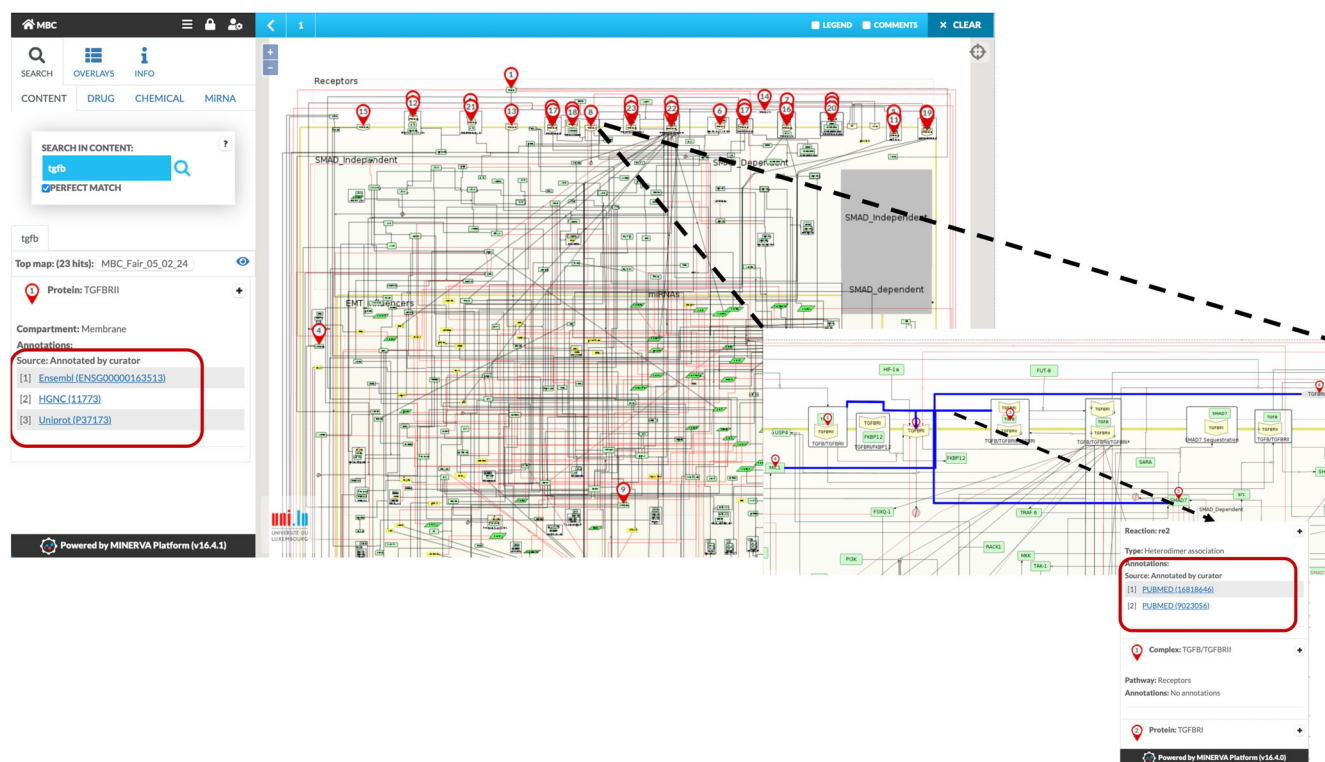
map captures the TGFβ induced epithelial-to-mesenchymal transition signalling network in metastatic breast cancer. TGFβ parallelly regulates other signalling pathways like TNF-α, integrins, EGFR, in modulating the regulators associated with mesenchymal phenotype and epithelial phenotype further moderating EM transition.

**SMAD-independent signalling-based Modulation of EMT in MBC**

Numerous experimental studies have demonstrated the role of TGFβ-TGFβRII-TGFβRI complex in employing the SMAD-independent pathways in regulating EMT. These include the activation of PI3K/Akt<sup>63-65</sup>, STAT3<sup>66,67</sup>, c-Myc<sup>52,68-70</sup>, TNF-α/TRAF/TAK<sup>36</sup>, Annexin-2<sup>66,67,71-75</sup>-based signalling pathways. Phosphorylation of TGFβRII at Tyr residues by Src leads to the activation of Shc-A. This, in turn, facilitates the assembly of GRB2-SHC-A-SOS complex, which plays a crucial role in transducing downstream signalling through the p38/MAPK axis directly. Additionally, this pathway can also be influenced indirectly by β3 integrin signalling<sup>76,77</sup>. The functioning of TGFβ signalling in the participation of cytoskeletal rearrangement, cell polarity and migration by Rho GTPases can rapidly promote EMT in a SMAD2/3-independent manner<sup>78</sup>. Crosstalk between

IGF1/PI3K signalling induced LTGFβ results in activation of TGFβ signalling promoting EMT in breast cancer<sup>64</sup>. This diverse upstream signalling modulated by TGFβ in SMAD-dependent and independent manner converges at regulators that are majorly associated with invasive properties of EMT in MBCs. These include SNAIL/SLUG, ZEB1/2, TWIST1/2, NF-kB, MMPs which are regulating the adherent junctions (E-Cadherin/β-Catenin) and further the process of EMT. This experimental evidence supports the critical role of TGFβ signalling in EMT.

SNAIL, a family of zinc finger transcription motifs representing SNAI1 (SNAIL), SNAI2 (SLUG) and SNAI3 are key regulators of TGFβ induced EMT in breast cancer cells<sup>79-83</sup>. TGFβ modulates the transcriptional activity of SNAIL through many paths (Supplementary Fig. 7). SNAIL/SLUG primarily functions as a repressor complex targeting the E-Boxes of



**Fig. 2 | MBC Map in the MINERVA platform.** Users can search for their regulator of interest from the search box. The results are shown as pins. The annotations of the corresponding regulators like HGNC ID, Uniport ID, Ensembl identifiers are displayed on the left upon selecting any element along with the PUBMED identifiers.

Further users can also navigate interactions starting from a molecule of interest tracking the signal from TGF $\beta$  to the end phenotype. MAP is available at: [http://35.174.227.105:8080/minerva/?id=Metastatic\\_Breast\\_Cancer\\_1](http://35.174.227.105:8080/minerva/?id=Metastatic_Breast_Cancer_1).

E-Cadherin affecting its gene expression<sup>79,84,85</sup>. E-cadherin is a major epithelial marker that regulates the tight junctions when complexed with  $\beta$ -catenin. SNAIL promotes cancer cell invasion when it forms a complex with Sin3A-HDAC1/2 through SNAG domain repressing E-Cadherin. Conversely, SNAIL curbs invasion in the presence of p53, MDM2 in wild type<sup>86–88</sup>. The association of SNAIL with mutant p53 has been found to stabilise SNAIL, thereby promoting EMT<sup>88,89</sup>. However, when SNAIL is coupled with G9a, it facilitates the degradation of E-cadherin, leading to activation of invasive switch<sup>90</sup>. Additionally, SNAIL activity is amplified through the activation of NF-kB signalling regulated by TGF $\beta$  induced c-Myc, TRAF/TAK, TWIST, STAT3, ANXA2 and AKT<sup>54,59,81,88,91–99</sup>. SNAIL is also regulated through TGF $\beta$  induced ERK signalling in promoting EMT<sup>86,100,101</sup>. GSK-3 $\beta$  suppresses EMT through SNAIL degradation, while AKT stabilises SNAIL from GSK-3 $\beta$  degradation promoting EMT. The combination of paths leading from TGF $\beta$  to SNAIL with varying signs, including both positive and negative interactions, is a result of the intricate network architecture inferred from experimental literature. The overall activation of SNAIL is determined by the collective effect of these combined interactions. Thus, SNAIL was observed to be a prominent player of EMT in MBC.

Similarly, ZEB, also a Zinc finger transcription motif comprising of ZEB1 and ZEB2 are important regulators of TGF $\beta$  induced EMT in breast cancer cells (Supplementary Fig. 8). Primarily regulated by SNAIL, ZEB was known to play a crucial role in maintaining the mesenchymal phenotype during the process of EMT<sup>91,102–104</sup>. ZEB recruits either YAP/AP-1 or CTBP along with other co-factors, which act as activators or repressors of mesenchymal or epithelial genes, respectively<sup>105</sup>. ZEB interacts with factors like OVOL2, GRHL2, in a mutually inhibitory loop resulting in multiple hybrid states<sup>104,106,107</sup>. Further, the presence of mutual inhibition between ZEB and GRHL2, as well as ZEB and miR-145, results in a hybrid with enhanced stability. Together, SNAIL and ZEB regulate the expression of the immune modulatory gene PDL1 in moderating the process of EMT<sup>91</sup>. In

addition to SNAIL and ZEB families, the TWIST family of transcription factors plays an important role in the process of EMT and metastasis (Supplementary Fig. 8). Regulated through several SMAD-independent signalling pathways, TWIST is induced by AKT, NF-kB, SNAIL, ZEB. TWIST in turn activates FOXC2, AKT (though double positive feedback loop), LTGF $\beta$ , RHOC in regulating the expression of E-Cadherin thus modulating EMT. Thus, all the signalling upstream assists the key transcriptional factors in fine-tuning the process of EMT by either repressing epithelial phenotype-associated genes or by promoting the mesenchymal phenotype-associated genes.

Furthermore, various studies support that the augmentation of tumour invasion and metastasis is centred around autocrine TGF $\beta$  signalling mechanism<sup>14,32,103,108</sup>. TGF $\beta$ RIII (beta glycan receptor family) a ubiquitously expressed coreceptor of TGF $\beta$  regulates the breast cancer progression and metastasis by either sequestering TGF $\beta$  or by interfering with PTEN inhibition of LTGF $\beta$ <sup>32,109</sup>. Above all, TGF $\beta$  signalling upstream also regulates the expression of matrix metalloproteinases (MMPs) (Supplementary Fig. 9)<sup>64,110,111</sup> which are mainly involved in the cleaving of the E-Cadherin/ $\beta$ -Catenin tight junctions initiating the migratory and invasive phenotype of breast cancer cells as shown in Fig. 1. The progression of metastatic breast cancer in TGF $\beta$  simulated cells has also been linked to the expressions of stemness markers CD44<sup>+</sup>/CD24<sup>-</sup> during the early stages of breast cancer<sup>112</sup>. Cells with an increased expression of CD44<sup>+</sup>/CD24<sup>-</sup> further express high levels of EMT-associated genes thus promoting cell invasion and metastasis.

### RNAs in TGF $\beta$ induced MBC

TGF $\beta$  signalling regulates multiple RNAs through SMAD-dependent and SMAD-independent pathways during the emergence of MBC (Supplementary Fig. 10). Specifically, TGF $\beta$  regulates miR-190, miR-23a, miR-182, miR-106b in SMAD-dependent manner, promoting invasion and metastasis of breast cancer<sup>113–116</sup>. Regulation of miR-181a, miR-155, miR-21, miR-615, miR-1, by TGF $\beta$  suppresses EMT by altering the activity of

SMADs<sup>35,117–119</sup>. Additionally, p53 plays an important role in regulating the expression of various miRNAs in maintaining the expression of epithelial phenotype<sup>120</sup>. Some of these miRNAs (miR-200, miR-34, miR-203, miR-204) are part of mutually inhibitory feedback loops with major mesenchymal markers ZEB1/2, SNAIL/SLUG which suppress EMT by preserving the E-Cadherin expression. Mechanistically TGF $\beta$  downregulates these miRNAs through SNAIL/SLUG, ZEB1/2 which actively target the genes involved in cell motility and invasion (Supplementary Fig. 10).

Although the majority of miRNAs are downregulated during malignant transformation, some miRNAs such as miR-520c, miR-133b, miR-153, miR-145 are observed to be elevated in the presence of TGF $\beta$  signalling which can negatively affect its signalling at the receptor level. On the contrary, TWIST-induced miR-10b has a context-dependent role in inducing EMT in different breast cancer cells<sup>121–124</sup>. Similarly, several other RNAs including miRNAs, mRNAs, LncRNAs were observed to be expressed in TGF $\beta$  induced EMT in breast cancer. Despite the evidence of several RNAs being expressed in TGF $\beta$  induced EMT of breast cancer cells, the mechanism of how several RNAs are regulated remains unknown.

### Tangled crosstalk and downstream response

The molecular classification of breast cancer cells is mainly based on the expression of oestrogen receptor (ER<sup>+/−</sup>), Progesterone receptor (PR<sup>+/−</sup>) and HER (+/−) receptors<sup>125</sup>. Stimulation of TGF $\beta$  triggers downstream signalling cascades, regulating HER status in breast cancer cells and enhancing the process of EMT, as depicted in Fig. 1. Wang et al.<sup>126</sup>, have reported the activation of ErBb3(HER3) signalling through TGF $\beta$  induced regulator TACE(ADAM17) which promotes EMT through PI3K signalling<sup>126</sup>. HER2 signalling through TGF $\beta$  induced TWIST and YB1 further regulates the expression of NF $\kappa$ B through STAT3 in modulating EMT<sup>127–129</sup>. In normal cells, ER- $\alpha$  exerts inhibitory effects on TGF- $\beta$ RII and SNAIL through MTA1, contributing to the maintenance of epithelial structures. However, in malignant cells, TGF- $\beta$  regulated SNAIL exhibits a feedback mechanism and inhibits ER- $\alpha$ , promoting the invasiveness of breast cancer cells<sup>81,99,130</sup>. Activation of TNF- $\alpha$  by p38 acts in parallel with TGF $\beta$  regulated TRAF in promoting transcription factor NF- $\kappa$ B regulating the process of EMT<sup>101</sup>. TGF $\beta$  signalling works collectively with other signalling axis in augmenting the process of EMT. These include  $\alpha$ v $\beta$ 6 integrin,  $\beta$ 3 integrin, UPA, IL-6, Wnt, EGFR (Supplementary Table 1). Additionally, TGF $\beta$  signalling is also observed to regulate several proinvasive genes (IL-8, IL-17, IL-11), cell cycle regulators (Cyclin G2, p21, p27, p15, p53, p63) and stemness markers (CD24, CD44) in both SMAD-dependent and independent pathways. Recent studies have revealed a positive correlation between EMT and Androgen Receptor (AR) signalling influenced by ER signalling, specifically within the Triple-Negative Breast Cancer (TNBC) subtype<sup>131</sup>. Consequently, the induction of EMT in breast cancer cell lines by TGF $\beta$  and the cross-regulation observed underscores the distinct molecular composition of TNBC. Thus, TGF $\beta$  works in concert with other regulators and signalling pathways in moderating the EMT in breast cancer.

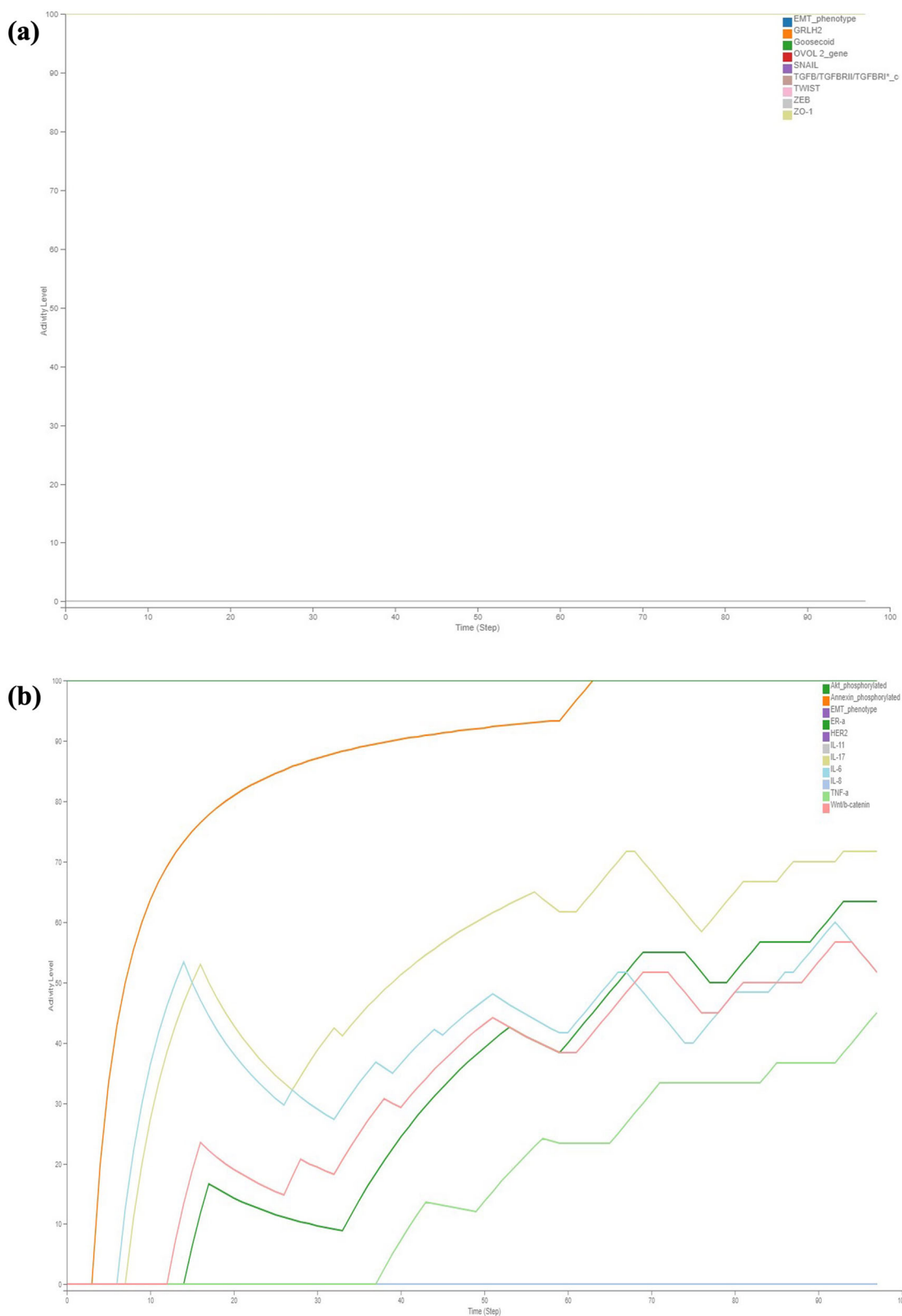
### Dynamic analysis of pathway Map of TGF $\beta$ induced EMT in MBC

The pathway map of TGF $\beta$  induced EMT in MBC constructed using Cell Designer V 4.4 was stored as SBML file. The resulting SBML-Process Description file was converted into a Boolean inference model (SBML-*Qual*) utilising CaSQ. The executable SBML-*Qual* file obtained from CaSQ was imported into Cell Collective to capture the system dynamics during the process of MBC (Supplementary Information). To evaluate the dynamics, an epithelial state of the model was initialised with TGF $\beta$ <sup>−</sup>. The phenotypic characterisation was performed by the status of the nodes SNAIL, ZEB, TWIST, ZO-1, Goosecoid and EMT. These nodes are controlled by various epithelial and mesenchymal regulators like miR-200, miR-34, p53, PI3K, OVOL2, GRHL2, NF- $\kappa$ B, HER and ER towards maintaining the adherent cell-cell junctions (E-Cadherin/ $\beta$ -Catenin) upon TGF $\beta$  stimulation. Hence the regulation of these molecules results in the epithelial or mesenchymal or hybrid phenotypes. In the absence of TGF $\beta$  (Fig. 3a), the activity levels of

most EMT markers and their regulators do not show any change in their activity, i.e., all the epithelial markers are at maximum activity levels and mesenchymal markers at minimum activity levels. However, some pathway regulators like PI3K, Akt Signalling, MTDH, NF- $\kappa$ B, Proinvasive genes like IL-11, IL-17, IL-6, IL-8, TNF- $\alpha$ , Wnt that stimulate the activity of EMT markers were observed to show low to moderate activity levels (Figs. 3b and 4a). E-Cadherin and  $\beta$ -catenin membrane levels were observed to be at their low activity suggesting the integrity of cell-cell junctions in the absence of TGF $\beta$ . From Fig. 3b it can be observed that other signalling pathways HER, ER were not regulated in the absence of TGF $\beta$ . Similarly, all the epithelial miRNAs involved were observed to be at their maximum activity and mesenchymal miRNAs at their minimum activity with the exemption of a few miRNAs (Fig. 4b). These findings shed light on the complex regulatory dynamics underlying the EMT process and provide valuable insights into the behaviour of key molecular components in the absence of TGF $\beta$  signalling.

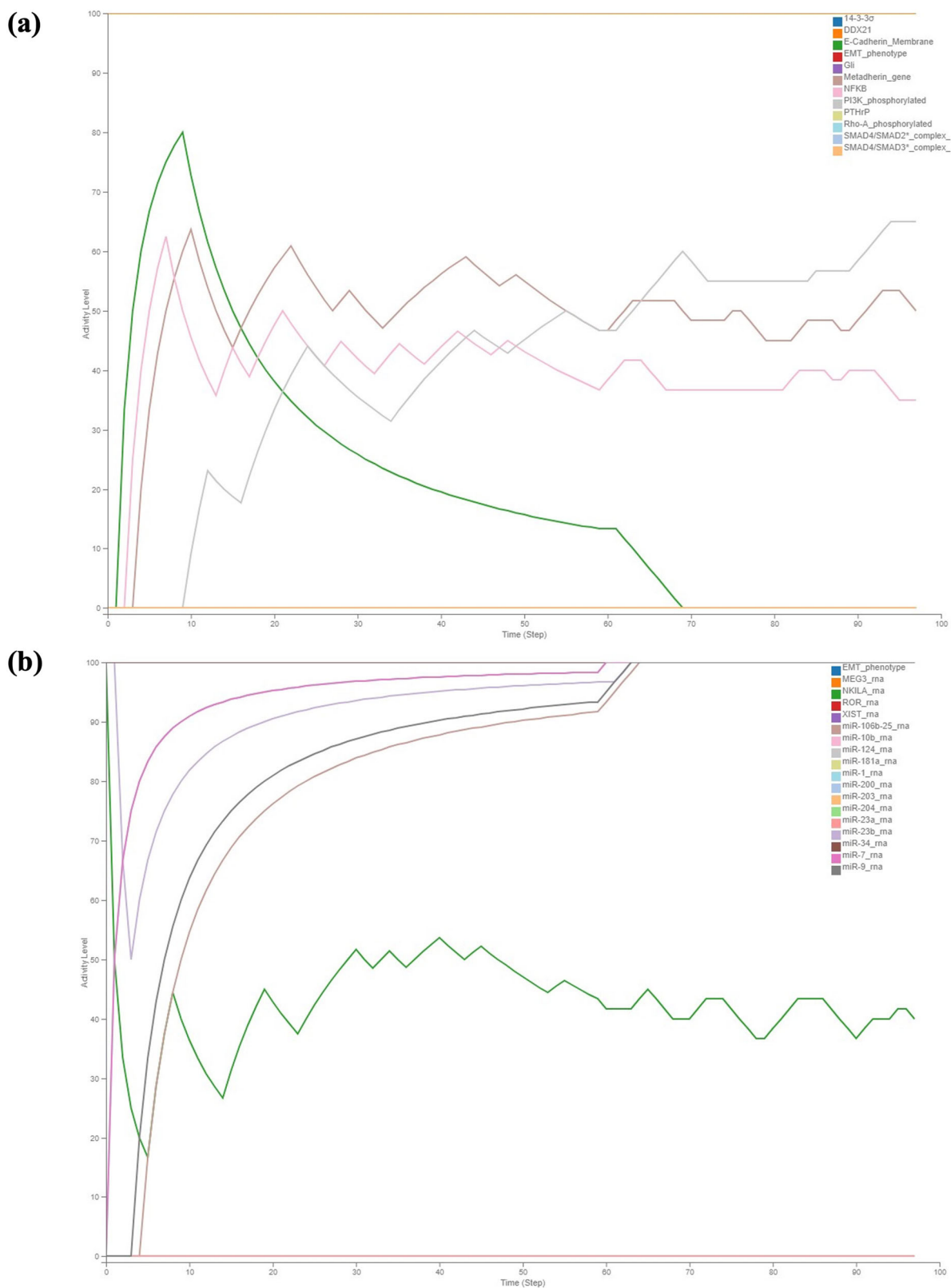
When TGF $\beta$  is active, all the regulators involved in modulating the EMT were observed to be expressed at their maximum activity levels (Figs. 5 and 6). This indicates that the presence of active TGF $\beta$  signalling induces a coordinated upregulation of the regulatory components that orchestrate the EMT process. These pathway components, exhibit peak activity levels, leading to the promotion of EMT and subsequent changes in cell behaviour, such as increased cell motility and invasion. Stimulation of TGF $\beta$  by setting its input to 100% activated its receptor complex, initiating downstream signalling through both the SMAD signalling pathway and SMAD-independent pathways, as shown in Figs. 5a and 6a. The activation of these pathways resulted in an increase in the activity levels of E-Cadherin<sub>membrane</sub> and  $\beta$ -Catenin<sub>membrane</sub>, indicating the disruption of cell-cell junctions (Fig. 5a). Similarly, all the epithelial miRNAs involved were observed to be at their minimum activity and mesenchymal miRNAs at their maximum activity with the exception of few miRNAs (Fig. 6b). The LncRNA NIKLA, in conjunction with NF- $\kappa$ B, was observed to exhibit oscillatory behaviour due to the presence of a negative feedback loop between these two nodes. This negative feedback mechanism leads to fluctuations in the activity levels of both LncRNA NIKLA and NF- $\kappa$ B, resulting in oscillations in their expression patterns over time. The oscillatory nature of this regulatory circuit may play a significant role in modulating cellular responses to TGF $\beta$  stimulus (Fig. 6b). The collective activity of all the regulators, including the TGF $\beta$  receptor complex, SMAD signalling pathway and SMAD-independent pathways modulate the activity of key EMT regulators. These regulators include SNAIL/SLUG, ZEB1/2, TWIST1, Goosecoid and ZO-1, as well as phenotype stability factors such as OVOL2 and GRHL2. (Fig. 5a). From Figs. 5a, 6a it is evident that the onset of EMT occurs when both EMT regulators and EMT markers reach their maximum activity levels. This observation suggests that the activation of EMT is tightly regulated and requires the simultaneous upregulation of key EMT regulators and markers. From Figs. 5, 6 coexistence of epithelial and mesenchymal markers was observed for a period of time indicating the hybrid phenotype. This suggests that the process of EMT is not advocated by a single regulator but by a group of regulators in action.

Supplementary Fig. 11 depicts the activity network of the comprehensive map assembled in the absence (Supplementary Fig. 11a) and presence (Supplementary Fig. 11b) of the stimulus TGF $\beta$ . The activity network provides a visual representation of the regulatory interactions and signalling pathways that are activated or suppressed under these two conditions. When TGF $\beta$  is absent (Supplementary Fig. 11a), certain regulators and pathways show low to moderate activity levels, while in the presence of active TGF $\beta$  (Supplementary Fig. 11b), all the regulators involved in modulating the EMT process exhibit their maximum activity levels. This comparison highlights the dynamic changes in the network and the impact of TGF $\beta$  signalling on the regulation of EMT-related processes. Thus, the activation of TGF $\beta$  initiates a cascade of molecular events, ensuring a synchronised and potent induction of EMT, contributing to cancer metastasis.



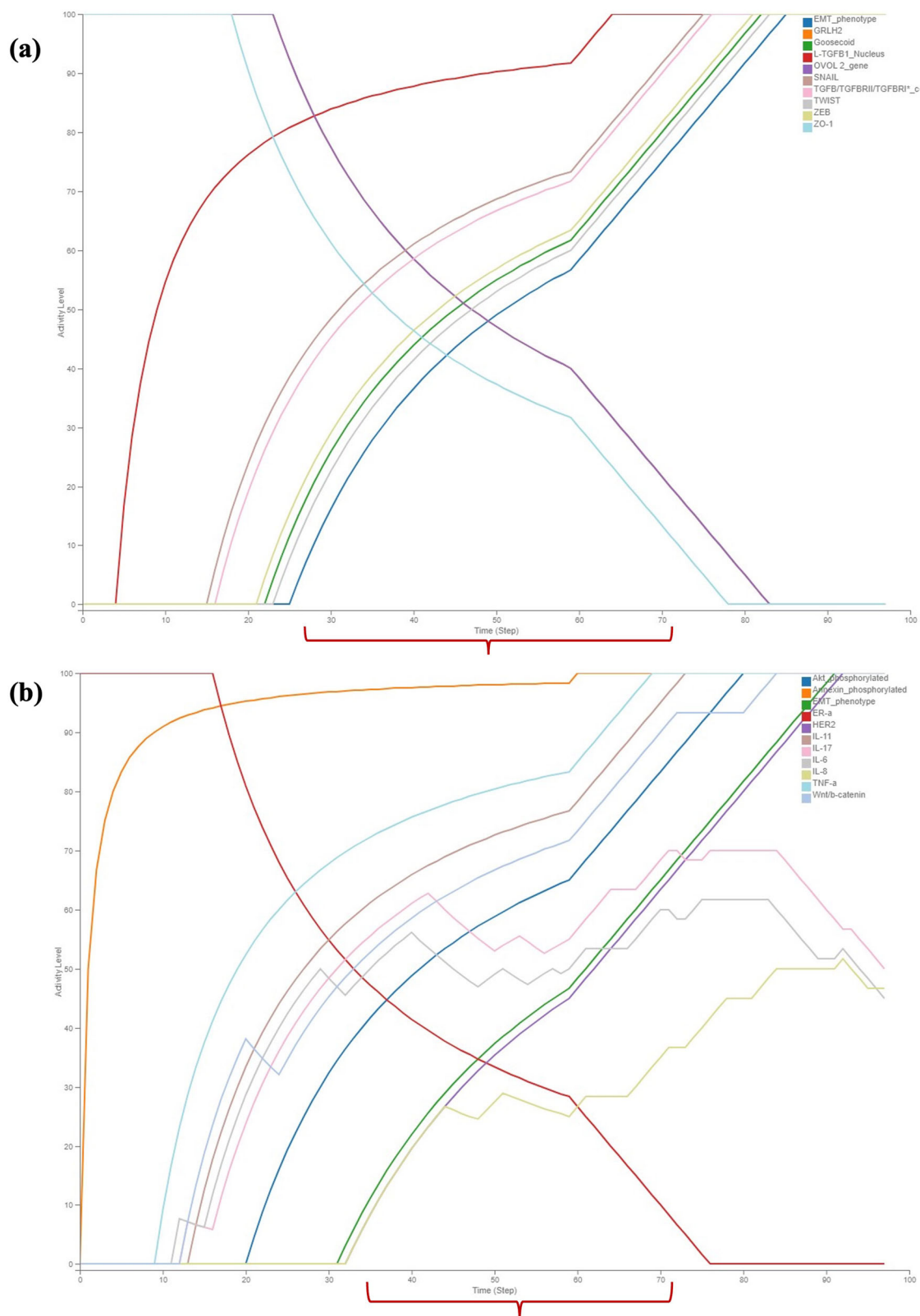
**Fig. 3 | Discrete dynamic modelling and analysis of the assembled MBC map illustrating the activity of various EMT factors and receptors regulating them using Cell Collective in the absence of stimulus TGFβ. a** The activity of all the mesenchymal markers was observed to be at their minimal levels and the epithelial markers were observed to be at maximal levels, **b** The activity of the receptors

associated with epithelial phenotype remained at their maximum while the receptors associated with mesenchymal phenotype remain at their minimal level. The observation of moderate activity in Akt signaling and proinvasive genes suggests the presence of cross-regulation among the signaling molecules.



**Fig. 4 | Discrete dynamic modelling and analysis of the assembled MBC map illustrating the activity of various key signaling pathways of EMT and miRNAs using Cell Collective in the absence of stimulus TGFβ. a** The regulators involved in modulating the expressions of EMT markers exhibit moderate to minimum activity

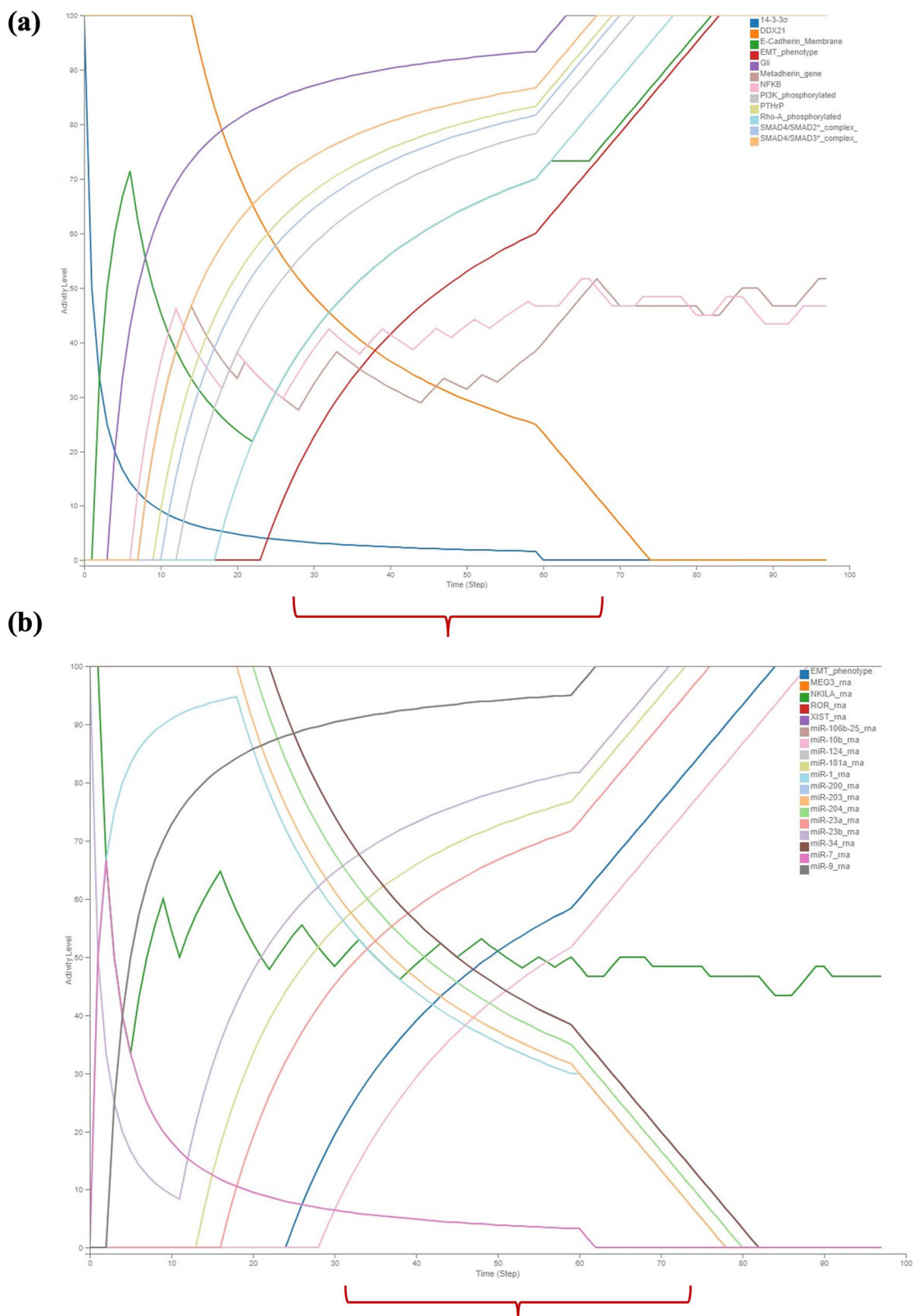
levels over time, with the notable exceptions of NF-κB, PI3K, and Metadherin regulators, **b** Epithelial miRNAs show maximum activity levels, while miRNAs regulating the mesenchymal phenotype were observed to exhibit minimum activity.



**Fig. 5 | Discrete dynamic modelling and analysis of the assembled MBC map activity of various EMT factors and receptors regulating them using Cell Collective in the presence of stimulus TGFβ. a** Shows maximum activation of core transcription factors of EMT like ZEB, SNAIL, and TWIST along with the upregulation of the EMT marker Goosecoid in regulating EMT. Simultaneously, the

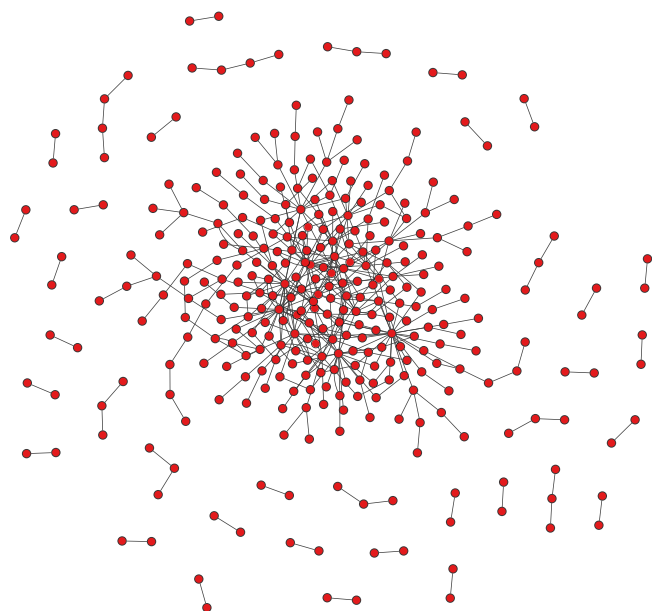
activity of the epithelial gene ZO-1 was observed to be reduced. **b** Induction of the other signaling pathways HER, TNF-α, Wnt, Akt and proinvasive genes along with the concurrent loss of ER-α was observed. Notably, the coexistence of epithelial and mesenchymal regulators (i.e., hybrid phenotype) was observed in the presence of TGFβ.





**Fig. 6 | Discrete dynamic modelling and analysis of the assembled MBC map illustrating the activity of various key signaling pathways of EMT and miRNAs using Cell Collective in the presence of stimulus TGFβ. a** The regulators responsible for maintaining the epithelial phenotype were observed to exhibit reduced activity, while those governing the mesenchymal phenotype were observed

to display increased activity, driving EMT. **b** The activity of miRNAs governing epithelial phenotype was observed to decrease, while that of mesenchymal miRNAs was increased, with a few exceptions; Notably, the coexistence of epithelial and mesenchymal regulators (i.e., hybrid phenotype) was observed in the presence of TGFβ.



**Fig. 7 | Visualisation of the assembled map using Cytoscape.** The assembled map of metastatic breast cancer (MBC) visualised using the Compound Spring Embedder (CoSE) layout within the Cytoscape. This visualisation highlights the intricate relationships within the MBC network. The network is visualised as a complex structure comprising a single centrally connected network (graph) and 37 separate subgraphs.

### Network analysis using Cytoscape

Basic network analysis was performed to explore the properties of the map developed. XML and CSV tabular formats of MBC map were imported into the Cytoscape V3.10 and analysed further using the inbuilt plugin Network analyser for topological metrics. The XML file imported comprises of 733 nodes and 1023 interactions connected as one core component (Supplementary Fig. 12). Whereas the CSV file imported into Cytoscape has 340 nodes and 426 interactions which was utilised as input for further analysis (Fig. 7).

Topological analysis using the built-in plugin Network analyser of Cytoscape revealed that the network consists of 38 connected components corresponding to connecting subgraph with a core subgraph and 37 smaller subgraphs (Fig. 7). It was observed that all node degree distributions follow the power law, showing that the networks are scale-free (Supplementary Fig. 13). Topological parameters obtained for both directed and undirected graph of the MBC network assembled in this study were tabulated (Table 1). Each node in MBC network has an average of ~2.9 neighbours. The network density is 0.015 for the directed graph which implies that the network is dense. Directed graphs indicate the flow of information thus consider both the presence of edges and their direction. The clustering coefficient of the MBC network was 0.037 which implies that there are very few links between the neighbours of the node in the network.

There are 15 multi-edge node pairs that provide information about the frequency of multiple connections between neighbouring nodes in the network. The longest shortest path between any two nodes, i.e., the network diameter was 14 suggesting that the signal originating from TGF $\beta$  ligand-receptor complexes in the membrane can reach a significant portion of the network within 14 steps. The characteristic path length of the network, which signifies the average distance between connected nodes, was approximately 4.7. This indicates that the response to a signal and its propagation can occur with moderate efficiency within the network. The highly connected nodes with a large number of incoming and outgoing edges indicate the network's tendency to have hub nodes (These are referred to as 'multi-edge node pairs' in directed graphs and 'network heterogeneity' in undirected graphs).

**Table 1 | Simple topological parameters obtained with the network analyser tool of Cytoscape for the assembled map of MBC**

Property	Directed	Undirected
Number of nodes	340	340
Number of edges	426	426
Avg. number of neighbours	2.418	2.847
Network diameter	14	14
Network radius	1	7
Characteristic path length	5.375	4.786
Clustering coefficient	0.015	0.037
Network density	0.004	0.011
Connected Components	38	38
Multi-edge node pairs/network heterogeneity	15	1.143
Number of Self Loops/network centralisation	0	0.092
Analysis time (sec)	1,693,824,729.066	0.061

Cytoscape plug-in Cytohubba-based analysis of the converted network (Fig. 7) resulted in the identification of the 25 most influential genes (Fig. 8) of the MBC map assembled (Fig. 1). The gradient from yellow to red implies significant to the highly significant genes. SNAIL, NF- $\kappa$ B, MMPs, SMAD4/SMAD2, SMAD4/SMAD3 complexes, TGF $\beta$ \_TGFBR1\_TGFBR2 complex, ZEB1/2, TWIST, p38, E-Cadherin, MDM2, STAT3, RAS, C-Myc, EGFR, ERK,  $\beta$ -Catenin, p53, L-TGF $\beta$ 1, CD44, miR-200, Annexin, Cyclin D1, PAK2, miR-7 and p21 were these 25 most influential genes. These Hub genes represent key players within the MBC map. All the hub genes identified were found to be driver/ oncogenes/tumour suppressor genes promoting MBC<sup>58,69,86,100,132</sup>. Most of the identified hub genes were observed to contribute to invasiveness (Supplementary Table 2). Supplementary Table 2 provides a detailed list of identified hub genes. Some of these genes are also part of interesting network motif architecture that plays a crucial role in driving emergent behaviour.

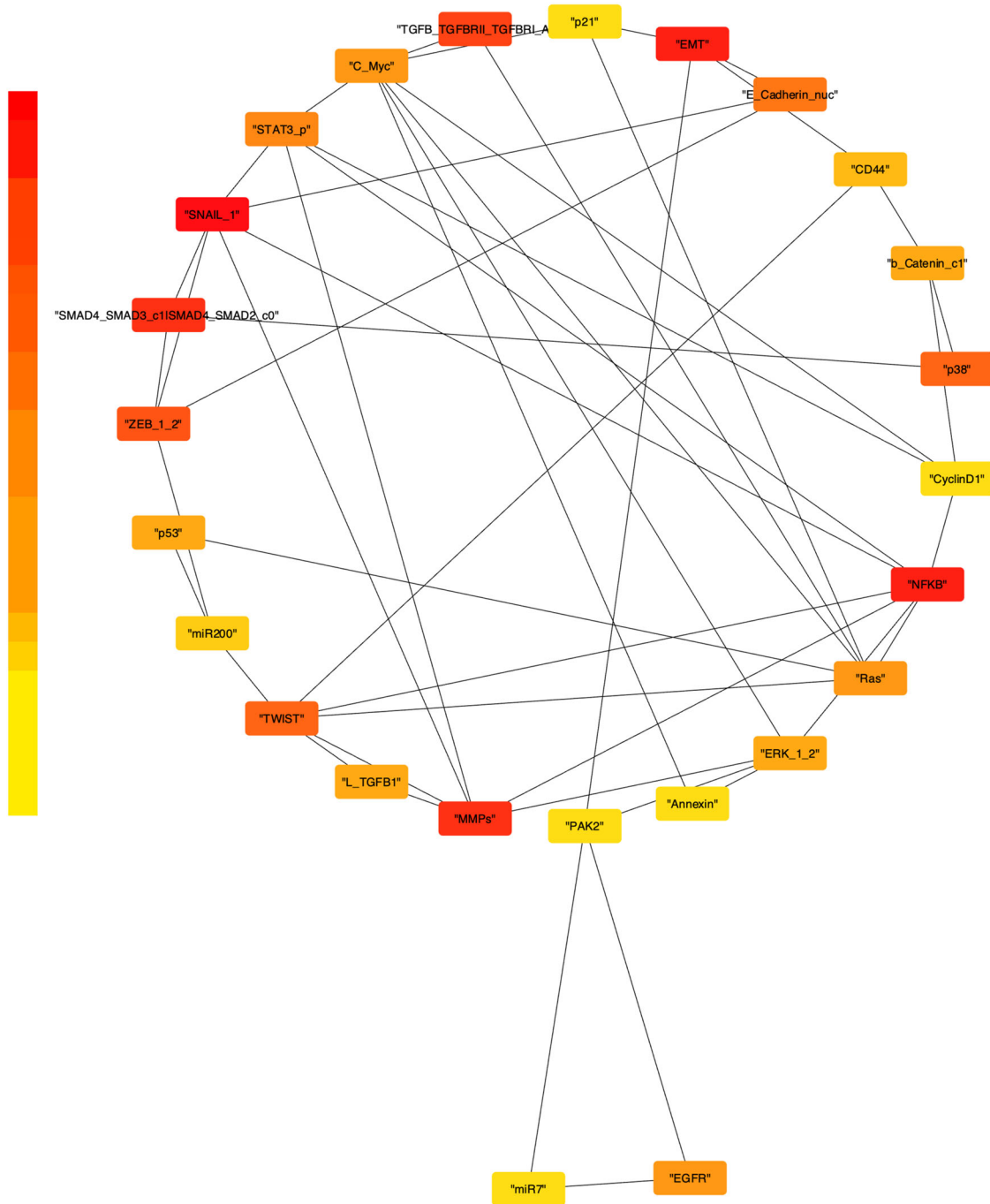
### Transcriptome-based validation

The prognostic significance of identified hub genes in MBC and their clinical relevance was further assessed through survival analysis. For survival analysis, the invasive breast cancer patients (BRCA) from TCGA and GTEx databases were evaluated using the GEPIA database (Supplementary Figs. 17 and 18). SNAIL1, NF- $\kappa$ B2, KRAS, MMP7, MAPK1, LTBP1, EGFR, TP53, PAK2, and CDKN1A (p21) exhibited significant prognostic differences in their overall survival with  $p < 0.05$  (Table 2). Additionally, LTBP1, SNAIL1, STAT3 were observed to be significant in terms of disease-free survival. Notably, low expressions of CDKN1A, TP53 are associated with better overall survival, while high expression of LTBP1 is associated with better overall survival respectively. Moreover, High expressions of LTBP1, SNAIL1 are observed to be associated with better disease-free survival.

Further, the significance of these prognostically relevant genes was evaluated by comparing the RNA-seq expression patterns between normal, tumour, and metastatic samples. TNMplot database was utilised for this analysis and the statistical significance was determined based on Kruskal–Wallis test (Fig. 9). The expressions of KRAS, STAT3, CDKN1A, PAK2, NF $\kappa$ B2, SNAIL1, were observed to be significantly higher in metastatic samples compared to tumour and normal samples. However, the expression of MAPK1, MMP7, EGFR, LTBP1, TP53 was observed to be lower in the metastatic samples compared to normal and tumour samples. This highlights the distinct gene expression states of prognostically relevant genes in metastatic samples compared to both tumour and normal samples.

### Discussions

In this work, a comprehensive map illustrating the TGF $\beta$  induced migration of breast cancer cells leading to metastatic breast cancer was compiled. The



**Fig. 8 | Hub genes identified using Cytohubba.** Top 25 influential genes in the assembled MBC as identified by the cytohubba tool employing the MCC (Maximal Clique Centrality) method. Hub genes are those genes that are members of the

largest cliques within the network. The level of importance for these hubs was visually represented using a colour scale from highly significant to significant genes (red to yellow).

regulatory map illuminates the role of numerous individual regulators and their associated pathways activated because of TGFβ. The map comprises of both SMAD-dependent and SMAD-independent pathways regulating EMT in MBC. These diverse regulatory components and their interconnected pathways highlight the complexity of the signalling mechanisms involved in promoting MBC. This network perspective allows viewing the complex phenomenon of EMT as a function of all the regulators and their interplay involved which provides a unique view of the emergence of MBC.

TGFβ plays a crucial role in orchestrating various regulators and regulatory networks, triggering a cascade of molecular events contributing to cell migration, invasion, and ultimately metastasis<sup>11,21</sup>. Several experimental

studies have identified the role of TGFβ induced regulators and their mechanistic roles in regulating EMT and cancer metastasis. Activation of the developmental process of EMT is fundamental for several cellular functions<sup>5,133</sup>. Within the spectrum of EMT, Type I and type II EMT are associated with the physiological functions of a cell, including organ development, embryogenesis, organ fibrosis, and tissue development, whereas type III EMT is implicated in the pathophysiological functions of a cell, contributing to the progression of neoplasia and metastasis<sup>6,133</sup>. The scope of the comprehensive map assembled in this study was to enumerate such experimental information focusing particularly on the process of TGFβ induced EMT in metastatic breast cancer.

**Table 2 | Transcriptome analysis of the hub genes**

Gene	Overall survival	Disease-free survival
TGFB1	0.36	0.26
TGFB2	0.52	0.78
TGFB1	0.34	0.77
LTBP1	0.069*	0.059*
SNAI1	0.017*	0.062*
NFKB1	0.96	0.84
NFKB2	0.042*	0.32
SMAD2	0.33	0.55
SMAD3	0.97	0.14
SMAD4	0.42	0.57
MMP2	0.32	0.82
MMP3	0.93	0.5
MMP7	0.041*	0.82
MMP9	0.2	0.8
ZEB1	0.44	0.99
ZEB2	0.27	0.37
TWIST1	0.16	0.61
TWIST2	0.82	0.7
MAPK14	0.66	0.43
CDH1	0.4	0.99
STAT3	0.85	0.035*
KRAS	0.031*	0.78
HRAS	0.28	0.55
MYC	0.61	0.48
EGFR	0.091*	0.61
MAPK3	0.37	0.12
MAPK1	0.029*	0.95
TP53	0.073*	0.67
CD44	0.61	0.43
PAK2	0.075*	0.34
CDKN1A	0.058*	0.9
ANXA2	0.14	0.55
CTNNB1	0.31	0.27
CCND1	0.75	0.16

Expression levels of the hub genes and their correlation with patient survival in invasive breast carcinoma (BRCA) obtained from GEPIA database along with *p*-value is shown. Genes with significant *p*-values ( $p < 0.05$ ) were indicated by \*. They indicate the genes that exhibit significant associations with patient survival.

Similar comprehensive maps are available in literature focusing on the underlying complex molecular processes presenting a consensus view on the network-level processes and diseases. Examples of such focused studies include an illustration of cell cycle regulatory events<sup>134</sup>, EGFR signalling<sup>135</sup>, RB/E2F signalling<sup>136</sup>, G-protein coupled receptor signalling architecture<sup>137</sup>, Alzheimer's disease map<sup>138</sup>, cancer signalling<sup>139</sup>, COVID-19 map<sup>46,140</sup>, atlas of inflammation resolution<sup>141</sup>, rheumatoid arthritis map<sup>142</sup>. Systems biology-based studies of molecular interaction networks of numerous biological processes have led to unique experimental studies that have revealed various molecular mechanisms<sup>143–147</sup>. This demonstrates the potential value of such disease maps and future prospects for such maps in designing experiments. The true potential of disease maps is yet to be realised.

The systems-level elucidation of TGF $\beta$  pathway map (Fig. 1) introduced in this work adds to this collection and provides systematic

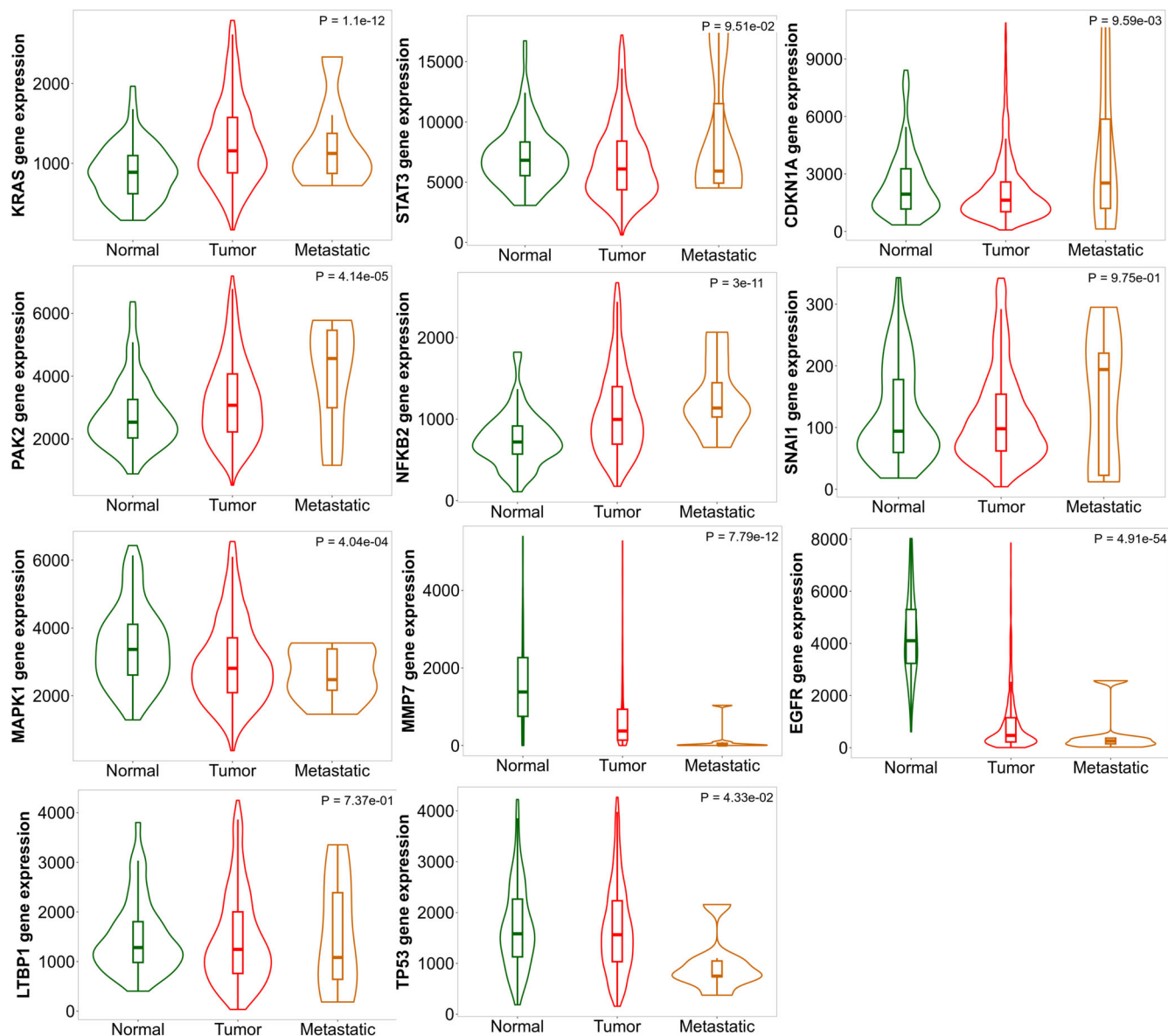
information of TGF $\beta$  induced EMT in the emergence of MBC. Lurking in these intricate signalling networks are key interactions and motifs that orchestrate the TGF $\beta$  stimulus in inducing EMT and cancer metastasis<sup>104,148–151</sup>. Exploring these underlying motifs is essential in understanding the contribution of TGF $\beta$  signalling activation of downstream effectors and transcription factors that drive the phenotypic changes associated with EMT and cancer metastasis.

Network modelling and analysis can explain the diverse contexts in which the TGF $\beta$  signal governs the process of EMT. Outcomes from the discrete dynamic modelling and simulation studies using Cell Collective have demonstrated that the model developed can reproduce the known biological dynamics of TGF $\beta$  signalling in metastatic breast cancer. The constitutive presence of TGF $\beta$  was sufficient to regulate the SMAD signalling, HER signalling, TNF signalling, Wnt signalling, ER- $\alpha$  signalling thereby regulating EMT-associated regulators and markers. These include SNAI1, ZEB, TWIST, ZO-1, E-Cadherin\_Membrane, L-TGF $\beta$  which further orchestrate EMT. It should be noted that these regulators and signalling pathways have been individually implicated in EMT (Supplementary Table 1). The systems-level impact of TGF $\beta$  in promoting EMT is not attributable to TGF $\beta$  alone but is dependent on the intricate interplay within the entire system that TGF $\beta$  signal participates with (Steinway et al.,<sup>152</sup>). Gaining insights into the dynamic behaviour of the TGF $\beta$  signalling pathway is crucial in identifying molecular events that drive cancer progression and metastasis.

The topological analysis of the disease map has shown scale free nature of the MBC map assembled. The topological properties of the cancer networks, including one connected component, moderate connectivity, network density, and efficient signal propagation, suggest a highly interconnected and information-rich network architecture. Thus, illustrating the complex structure and its potential to facilitate the dynamic signal transmission in deciphering complex disease mechanisms of EMT in MBC. Further, the topological analysis also assists in identifying the significantly influential nodes (hubs). 25 hubs genes of the MBC map thus identified as seen in Fig. 8 were characterised by their implications in the disease. Survival analysis using GEPIA, and expression validation in metastatic samples using TNMplot highlighted the prognostic relevance of TP53, LTBP1, NF- $\kappa$ B2, CDKN1A (p21), MMP7, KRAS, PAK2 based on the transcriptome data. The set of transcriptome analysis performed on the hub genes of the assembled map, validates their role in the context of cancer data available in public databases like TCGA, GTEx. Specific molecular biology-based experiments could assist in finding the mechanistic roles of these regulators and their impact in altering the pathways during the breast cancer progression.

Many publicly available pathway databases like Reactome<sup>153</sup>, KEGG<sup>154</sup>, STRING<sup>155</sup>, IPA<sup>156</sup>, GeneMANIA<sup>157</sup> collate data from multiple experimental and omics studies. These databases provide insights into different aspects of cellular interactome at various levels. However, one of the key challenges lies in ensuring the exclusiveness of the specific molecular interactions within pathways. We have observed that the pathways obtained from STRING, Reactome gives an overview of the gene regulatory network for a given query, they may be a shortfall in various information including exclusive molecular mechanism, source of information, cross-regulation between the signalling molecules (Supplementary Figs. 14–16). The map (Fig. 1) derived from an extensive literature survey addresses various inconsistencies encountered within these pathway databases thereby offering a more accurate and reliable representation of the molecular process under study. This rigorous curation process along with the validation through logical model-based validation illustrates the authenticity and usefulness of the map by answering specific questions for further research and understanding of cancer metastasis. This comprehensive map of TGF $\beta$  induced EMT allows exploration of underlying mechanistic processes contributing towards the emergence of MBC.

In summary, through an extensive literature survey, a comprehensive regulatory network encompassing both TGF $\beta$  induced SMAD-dependent and SMAD-independent signalling pathways involved in MBC has been



**Fig. 9 | Transcriptome-based analysis of the significant hub genes.** Violin plots illustrating the expression patterns of prognostically significant hub genes identified in invasive breast carcinoma. These plots compare the gene expression across tumour, normal, and metastatic RNA-seq data. The bars within the plots represent

the proportions of samples, and the statistical analysis was conducted based on Kruskal–Wallis test for the regulators with  $\log p < 0.05$  from survival analysis (Table 2).

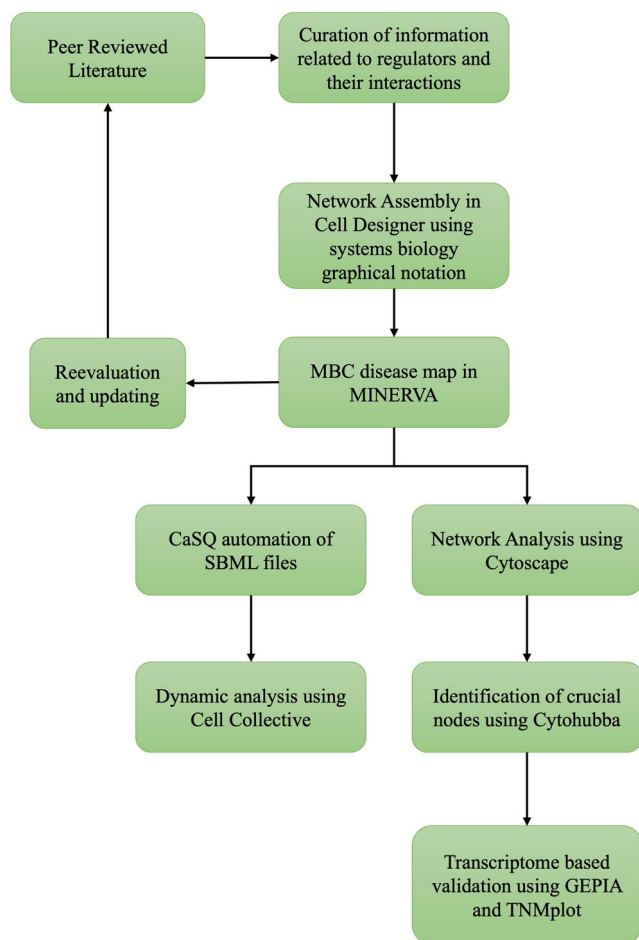
constructed. This has led to the identification and integration of a wide range of regulators and their interactions involved in the malignant transformation of epithelial cells in MBC. The map developed ([http://35.174.227.105:8080/minerva/?id=Metastatic\\_Breast\\_Cancer\\_1](http://35.174.227.105:8080/minerva/?id=Metastatic_Breast_Cancer_1)) serves as a valuable knowledgebase, consolidating information from various sources and facilitating a deeper understanding of the complex regulatory mechanisms underlying MBC. While the complete molecular mechanism of MBC remains incompletely understood, this work provides valuable insights into the regulatory networks governing TGFβ induced MBC. Further to validate, the map assembled was translated into a logical model and simulations of which have captured the known experimental outcomes of TGFβ induced Epithelial to Mesenchymal Transition in MBC. The identified hub regulators of the map and their transcriptome-based analysis have confirmed their role in breast cancer metastasis. By shedding light on the complex interplay of signalling pathways in MBC, this work contributes to advancing the knowledge and potential therapeutic approaches for combating MBC. Ultimately, the findings of this study have the potential to guide further

research, promote data-driven analysis, and inspire new strategies for personalised cancer therapy.

## Methods

### Curation of EMT network

Metastatic breast cancer (MBC) is modulated through various pathways and diverse regulators stimulated by TGFβ signalling. Several in vitro and in silico studies have been performed previously to identify the regulators and their mechanisms involved in MBCs. The overall approach followed for developing the map (Fig. 1) is shown in Fig. 10. Disease map of TGFβ induced EMT in breast cancer was constructed by performing an extensive literature survey that accounts for various regulators involved in disease emergence. The survey primarily focused on the experimental studies that identify the regulators involved in TGFβ induced EMT. Specifically, genes, proteins, RNAs other receptor signalling pathways regulated by TGFβ or regulating TGFβ and their interactions between them in inducing EMT were explored. The focus was majorly on TGFβ and its association with



**Fig. 10 | Workflow for the development and analysis of a comprehensive disease map of TGF $\beta$  induced EMT in MBC.** This figure illustrates the workflow for developing a comprehensive disease map of the signaling pathways involved in TGF $\beta$  induced epithelial-to-mesenchymal transition (EMT) in metastatic breast cancer (MBC). The development of a map is an iterative process involving the integration of information from various experimental studies found in the scientific literature. The final map is made publicly available through the MINERVA platform and is subjected to further validation through various analyses.

EMT in metastatic breast cancer. The information gathered largely consists of studies from the human cell lines-based experiments with a few exceptions based on mouse model-based experiments. The literature obtained is thoroughly cross-checked multiple times for the role of the regulators and their specific mechanism of action. HUGO nomenclature was acquired (<https://www.genenames.org>) for all the regulators involved in the disease map to avoid uncertainty in terminology.

The MBC regulatory network map was assembled using Cell Designer V 4.4 (<http://celldesigner.org/>) and follows the graphical notation system proposed by Kitano et al.<sup>158</sup>. Cell Designer supports the systems biology graphical notation (SBGN) and the constructed map can be exported into systems biology markup language (SBML) format for preferred computational analysis. MINERVA platform<sup>159–161</sup> was utilised to create a navigation-friendly web version of the assembled MBC map. It allows for visual exploration, analysis and management of assembled MBC Map (Fig. 2). It also provides automated annotation (Pubmed, Ensembl, Uniprot, RefSeq, miRbase, HGNC ID) of species and reactions of the networks.

### Insilco simulations

To further capture the dynamic behaviour of TGF $\beta$  induced EMT in metastatic breast cancer Insilco, Boolean network modelling and analysis were performed. Boolean models are well suited to handling large sizes of

data described in molecular interaction maps where there is a lack of kinetic data<sup>162</sup>. Boolean formalisms are scale-free, simplest forms of logical models where nodes represent the species (Protein, Genes, RNA, Complex) and edges (Activation, Inhibition, Association, Dissociation) represent the interactions between them. Each node is associated with a binary variable determining its qualitative level (0 False or 1 True) and its ability to influence its target (0 Inactive or 1 Active). The state of each node (Regulator) is determined by the states of neighbouring nodes (Regulators) represented by a set of regulatory functions that updates with time. Regulatory functions for each node (Regulator) were formulated depending on their upstream nodes (regulators) using logical operators AND, OR, NOT and are known to produce recognisable biological outcomes as obtained from experimental literature. The state of each node remains the same or changes depending on the regulatory function that updates with time. Thus, the state of a regulator follows a dynamic trajectory based on discrete variables 0 through 1. The updating of the rules can be synchronous i.e., all the nodes are updated at the same time, or asynchronous i.e., where only one node can be updated every time<sup>163–165</sup>.

To perform the logical modelling and analysis of the TGF $\beta$  induced EMT in MBC, the SBML file of the MBC interaction map obtained from the Cell Designer was converted into an SBML-*Qual* file using CaSQ, (<https://github.com/soli/casq>) to obtain a Boolean model. SBML-*Qual* is an extension of Systems Biology Markup Language (SBML) Level 3 Standard, represented for the qualitative models of biological networks<sup>166</sup>. The SBML-*Qual* file obtained is compatible for further analysis with the modelling tool Cell Collective (<https://cellcollective.org/>). The SBML-*Qual* file was imported into Cell Collective, a web-based modelling platform to perform real-time stimulations. Regulators of the interaction map are inferred as internal and external components by the Cell Collective under ‘Model’ tab along with a panel for regulatory expression for each selected regulator. The dynamics of the TGF $\beta$  induced MBC were performed using the ‘Simulations’ tab on Cell Collective using asynchronous update under epithelial initial environment conditions. The outcomes are measured based on the activity of EMT along with several epithelial and mesenchymal regulators using the ‘simulation graph’ panel under ‘Simulations’ tab.

### Topological analysis

The MBC map assembled (Fig. 1) was imported into Cytoscape for topological analysis using the built-in plugin Network Analyzer<sup>167</sup>. Cytoscape is an open-source platform for visualising and integrating molecular interaction networks with annotations, gene expression profiles and other data types. The MBC network was first analysed as an undirected network to obtain the overall degree distribution, then as a directed network to obtain the indegree, outdegree distribution and other topological properties. Subsequently, to obtain the influential genes of the MBC map Cytoscape plugin Cytohubba was utilised<sup>168</sup>. Specifically maximal clique centrality (MCC) of Cytohubba was used to identify the potential hub genes involved and to obtain the subnetwork.

### Transcriptome-based analysis and validation

To validate the prognostic relevance of identified hub genes, survival analysis was performed with invasive breast cancer (BRCA) datasets using GEPIA database<sup>169</sup>. GEPIA performs the Kaplan–Meier survival analysis by analysing the relative expression of hub genes. GEPIA comprises the data of the normal (112) and tumour (1085) samples from the TCGA and GTEx databases. For the analysis all the parameters were set to default values, with quartile cutoff. Overall survival and disease-free survival were evaluated using the Mantel–Cox test with a 95% confidence interval and Cox proportional hazardous ratio. A log *p*-value of less than 0.05 was considered as statistically relevant to identify the prognostically relevant genes. Further, the validation of the expression of prognostically relevant genes between tumour, normal and metastatic samples was analysed using the web platform TNMplot that integrates the transcriptome data from NCBI-GEO, TCGA, TARGET, GTEx repositories<sup>170</sup>. Specifically, RNA-seq data of invasive breast carcinoma was opted to compare the datasets to perform

Kruskal–Wallis test. The combination of survival analysis using GEPIA and DE analysis using TNM plots could further help in coarse-graining and validating the prognostically relevant genes of the MBC map.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

### Data availability

All the files generated and analysed during the current study are available in the GitHub repository, <https://github.com/gsb-sai/Metastatic-Breast-Cancer>.

Received: 20 October 2023; Accepted: 29 April 2024;

Published online: 17 May 2024

### References

- Desantis, C. E. et al. Breast cancer statistics, 2019. *CA: A Cancer J. Clin.* **69**, 438–451 (2019).
- Ferlay, J. et al. *Global Cancer Observatory: Cancer Today*. <Available from: <https://gco.iarc.fr/today> (International Agency for Research on Cancer, Lyon, France, 2020) accessed [24 February 2022].
- Wei, S. & Siegal, G. P. Metastatic organotropism: an intrinsic property of breast cancer molecular subtypes. *Adv. Anat. Pathol.* **24**, 78–81 (2017).
- Blanco, M. A. & Kang, Y. Signaling pathways in breast cancer metastasis - novel insights from functional genomics. *Breast Cancer Res.* **13**, 206 (2011).
- Vella, V. et al. Microenvironmental determinants of breast cancer metastasis: focus on the crucial interplay between estrogen and insulin/insulin-like growth factor signaling. *Front. Cell Dev. Biol.* 1458 <https://doi.org/10.3389/fcell.2020.608412> (2020).
- Brabletz, S., Schuhwerk, H., Brabletz, T. & Stemmler, M. P. Dynamic EMT: a multi-tool for tumor progression. *EMBO J* **40** <https://doi.org/10.15252/embj.2021108647> (2021).
- Zavadil, J. & Bottlinger, E. P. TGF-beta and epithelial-to-mesenchymal transitions. *Oncogene* **24**, 5764–5774 (2005).
- Thiery, J. P., Acloque, H., Huang, R. Y. & Nieto, M. A. Epithelial-mesenchymal transitions in development and disease. *Cell* **139**, 871–890 (2009).
- Miettinen, P. J., Ebner, R., Lopez, A. R. & Derynck, R. TGF-beta induced transdifferentiation of mammary epithelial cells to mesenchymal cells: involvement of type I receptors. *J. Cell Biol.* **127**, 2021–2036 (1994).
- Hay, E. D. An overview of epithelio-mesenchymal transformation. *Acta Anat. (Basel)* **154**, 8–20 (1995).
- Attisano, L. & Wrana, J. L. Signal transduction by members of the transforming growth factor-beta superfamily. *Cytokine Growth Factor Rev.* **7**, 327–339 (1996).
- Attisano, L. & Wrana, J. L. Mads and Smads in TGF beta signalling. *Curr. Opin. Cell Biol.* **10**, 188–194 (1998).
- Miyazono, K. Positive and negative regulation of TGF-beta signaling. *J. Cell Sci.* **113**, 1101–1109 (2000).
- Tobin, S. W. et al. Consequences of altered TGF-β expression and responsiveness in breast cancer: evidence for autocrine and paracrine effects. *Oncogene* **21**, 108–118 (2002).
- Bierie, B. & Moses, H. L. Tumour microenvironment: TGFbeta: the molecular Jekyll and Hyde of cancer. *Nat. Rev. Cancer* **6**, 506–520 (2006).
- Matsuzaki, K. Modulation of TGF-beta signaling during progression of chronic liver diseases. *Front. Biosci. (Landmark Ed.)* **14**, 2923–2934 (2009).
- Cellière, G., Fengos, G., Hervé, M. & Iber, D. plasticity of TGF-β signaling. *BMC Syst. Biol.* **5**, 184 (2011).
- Laise, P., Fanelli, D., Lió, P. & Arcangeli, A. Modeling TGF-β signaling pathway in epithelial-mesenchymal transition. *AIP Adv.* **2**, 011201 (2012).
- Massagué, J. TGFbeta signalling in context. *Nat. Rev. Mol. Cell Biol.* **13**, 616–630 (2014).
- Suzuki, H. I. MicroRNA control of TGF-beta signaling. *Int. J. Mol. Sci.* **19**, 1901–1901 (2018).
- Massagué, J. TGF-beta signal transduction. *Annu. Rev. Biochem.* **67**, 753–791 (1998).
- Dumont, N. & Arteaga, C. L. Targeting the TGF beta signaling network in human neoplasia. *Cancer Cell* **3**, 531–536 (2003).
- Xu, J., Lamouille, S. & Derynck, R. TGF-beta-induced epithelial to mesenchymal transition. *Cell Res.* **19**, 156–172 (2009).
- Reichl, P., Haider, C., Grubinger, M. & Mikulits, W. TGF-beta in epithelial to mesenchymal transition and metastasis of liver carcinoma. *Curr. Pharm. Des.* **18**, 4135–4147 (2012).
- Syed, V. TGF-beta signaling in cancer. *J. Cell Biochem.* **117**, 1279–1287 (2016).
- Hao, Y., Baker, D. & Ten Dijke, P. TGF-β-mediated epithelial-mesenchymal transition and cancer metastasis. *Int. J. Mol. Sci.* **20**, 2767 (2019).
- Tzavlaki, K. & Moustakas, A. TGF-β signaling. *Biomolecules* **10**, 487 (2020).
- Attisano, L. & Wrana, J. L. Signal transduction by the TGF-beta superfamily. *Science* **296**, 1646–1647 (2002).
- Massague, J. & Gomis, R. R. The logic of TGFbeta signaling. *FEBS Lett.* **580**, 2811–2820 (2006).
- David, C. J. et al. TGF-beta tumor Suppression through a lethal EMT. *Cell* **164**, 1015–1030 (2016).
- Massague, J. TGFbeta signalling in context. *Nat. Rev. Mol. Cell Biol.* **13**, 616–630 (2012).
- Lei, X., Bandyopadhyay, A., Le, T. & Sun, L. Autocrine TGFbeta supports growth and survival of human breast cancer MDA-MB-231 cells. *Oncogene* **21**, 7514–7523 (2002).
- Kang, Y. et al. Breast cancer bone metastasis mediated by the Smad tumor suppressor pathway. *Proc. Natl Acad. Sci. USA* **102**, 13909–13914 (2005).
- Foroni, C., Broggin, M., Generali, D. & Damia, G. Epithelial-mesenchymal transition and breast cancer: role, molecular mechanisms and clinical impact. *Cancer Treat. Rev.* **38**, 689–697 (2012).
- Taylor, M. A., Sossey-Alaoui, K., Thompson, C. L., Danielpour, D. & Schieman, W. P. TGF-beta upregulates miR-181a expression to promote breast cancer metastasis. *J. Clin. Invest.* **123**, 150–163 (2013).
- Liao, S. J. et al. TGF-beta1 and TNF-alpha synergistically induce epithelial to mesenchymal transition of breast cancer cells by enhancing TAK1 activation. *J. Cell Commun. Signal.* **13**, 369–380 (2019).
- Wardhani, B. W., Louisa, M., Watanabe, Y., Setiabudy, R. & Kato, M. TGF-β-induced TMEPAI promotes epithelial–mesenchymal transition in doxorubicin-treated triple-negative breast cancer cells via SMAD3 and PI3K/AKT pathway alteration. *Breast Cancer Targets Ther.* **13**, 529–538 (2021).
- Derynck, R. et al. Synthesis of messenger RNAs for transforming growth factors alpha and beta and the epidermal growth factor receptor by human tumors. *Cancer Res.* **47**, 707–712 (1987).
- Derynck, R. & Zhang, Y. E. Smad-dependent and Smad-independent pathways in TGF-beta family signalling. *Nature* **425**, 577–584 (2003).
- Moustakas, A. & Heldin, C. H. Non-Smad TGF-beta signals. *J. Cell Sci.* **118**, 3573–3584 (2005).
- Zhang, Y. E. Non-Smad pathways in TGF-beta signaling. *Cell Res.* **19**, 128–139 (2009).

42. Tecalco-Cruz, A. C., Rios-Lopez, D. G., Vazquez-Victorio, G., Rosales-Alvarez, R. E. & Macias-Silva, M. Transcriptional cofactors Ski and SnoN are major regulators of the TGF-beta/Smad signaling pathway in health and disease. *Signal. Transduct. Target. Ther.* **3**, 15 (2018).
43. Türei, D., Korcsmáros, T. & Saez-Rodriguez, J. OmniPath: guidelines and gateway for literature-curated signaling pathway resources. *Nat. Methods* **13**, 966–967 (2016).
44. Del Toro, N. et al. The IntAct database: efficient access to fine-grained molecular interaction data. *Nucleic Acids Res.* **50**, D648–D653 (2022).
45. Lo Surdo, P. et al. SIGNOR3.0, the SIGNaling network open resource 3.0: 2022 update. *Nucleic Acids Res.* **51**, D631–D637 (2023).
46. Ostaszewski, M. et al. COVID-19 disease map, building a computational repository of SARS-CoV-2 virus-host interaction mechanisms. *Sci. Data* **7**, 1–4 (2020).
47. Dubitzky, W., Wolkenhauer, O., Cho, K.-H. & Yokota, H. *Encyclopedia of Systems Biology*. Vol. 402 (Springer New York, 2013).
48. Groppe, J. et al. Cooperative assembly of TGF-β superfamily signaling complexes is mediated by two disparate mechanisms and distinct modes of receptor binding. *Mol. Cell* **29**, 157–168 (2008).
49. Wrana, J. L., Attisano, L., Wieser, R., Ventura, F. & Massagué, J. Mechanism of activation of the TGF-beta receptor. *Nature* **370**, 341–347 (1994).
50. Wang, T. et al. The immunophilin FKBP12 functions as a common inhibitor of the TGFβ family type I receptors. *Cell* **86**, 435–444 (1996).
51. Huse, M., Chen, Y.-G., Massagué, J. & Kuriyan, J. Crystal structure of the cytoplasmic domain of the type I TGF β receptor in complex with FKBP12. *Cell* **96**, 425–436 (1999).
52. Gomis, R. R., Alarcon, C., Nadal, C., Van Poznak, C. & Massagué, J. C/EBPβ at the core of the TGFβ cytostatic response and its evasion in metastatic breast cancer cells. *Cancer Cell* **10**, 203–214 (2006).
53. Xu, J. et al. 14-3-3ζ turns TGF-β's function from tumor suppressor to metastasis promoter in breast cancer by contextual changes of Smad partners from p53 to Gli2. *Cancer Cell* **27**, 177–192 (2015).
54. Yeh, H. W. et al. PSPC1 mediates TGF-beta1 autocrine signalling and Smad2/3 target switching to promote EMT, stemness and metastasis. *Nat. Cell Biol.* **20**, 479–491 (2018).
55. Zhou, T. et al. GDF10 inhibits proliferation and epithelial-mesenchymal transition in triple-negative breast cancer via upregulation of Smad7. *Aging (Albany NY)* **11**, 3298–3314 (2019).
56. Kang, Y. et al. A multigenic program mediating breast cancer metastasis to bone. *Cancer Cell* **3**, 537–549 (2003).
57. Sundqvist, A. et al. Specific interactions between Smad proteins and AP-1 components determine TGFβ-induced breast cancer cell invasion. *Oncogene* **32**, 3606–3615 (2013).
58. Adorno, M. et al. A mutant-p53/Smad complex opposes p63 to empower TGFβ-induced metastasis. *Cell* **137**, 87–98 (2009).
59. Wu, J. et al. RBM38 is involved in TGF-beta-induced epithelial-to-mesenchymal transition by stabilising zonula occludens-1 mRNA in breast cancer. *Br. J. Cancer* **117**, 675–684 (2017).
60. Fu, J. et al. The natural compound codonolactone attenuates TGF-beta1-mediated epithelial-to-mesenchymal transition and motility of breast cancer cells. *Oncol. Rep.* **35**, 117–126 (2016).
61. Le Scolan, E. et al. Transforming growth factor-beta suppresses the ability of Ski to inhibit tumor metastasis by inducing its degradation. *Cancer Res.* **68**, 3277–3285 (2008).
62. Watanabe, Y. et al. TMEPAI, a transmembrane TGF-β-inducible protein, sequesters smad proteins from active participation in TGF-β signaling. *Mol. Cell* **37**, 123–134 (2010).
63. Basho, R. K. et al. Targeting the PI3K/AKT/mTOR pathway for the treatment of mesenchymal triple-negative breast cancer: evidence from a phase 1 trial of mTOR inhibition in combination with liposomal doxorubicin and bevacizumab. *JAMA Oncol.* **3**, 509–515 (2017).
64. Walsh, L. A. & Damjanovski, S. IGF-1 increases invasive potential of MCF 7 breast cancer cells and induces activation of latent TGF-beta1 resulting in epithelial to mesenchymal transition. *Cell Commun. Signal.* **9**, 10 (2011).
65. Xue, G. et al. Akt/PKB-mediated phosphorylation of Twist1 promotes tumor metastasis via mediating cross-talk between PI3K/Akt and TGF-beta signaling axes. *Cancer Discov.* **2**, 248–259 (2012).
66. Wu, B. et al. Up-regulation of Anxa2 gene promotes proliferation and invasion of breast cancer MCF-7 cells. *Cell Prolif.* **45**, 189–198 (2012).
67. Wang, T. et al. Anxa2 binds to STAT3 and promotes epithelial to mesenchymal transition in breast cancer cells. *Oncotarget* **6**, 30975–30992 (2015).
68. Massagué, J., Blain, S. W. & Lo, R. S. TGFβ signaling in growth control, cancer, and heritable disorders. *Cell* **103**, 295–309 (2000).
69. Liu, M. et al. p21CIP1 attenuates Ras- and c-Myc-dependent breast tumor epithelial mesenchymal transition and cancer stem cell-like gene expression in vivo. *Proc. Natl Acad. Sci. USA* **106**, 19035–19039 (2009).
70. Ma, L. et al. miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis. *Nat. Cell Biol.* **12**, 247–256 (2010).
71. Sharma, M. R., Koltowski, L., Ownbey, R. T., Tuszynski, G. P. & Sharma, M. C. Angiogenesis-associated protein annexin II in breast cancer: selective expression in invasive breast cancer and contribution to tumor invasion and progression. *Exp. Mol. Pathol.* **81**, 146–156 (2006).
72. Zhang, F. et al. P-glycoprotein associates with Anxa2 and promotes invasion in multidrug resistant breast cancer cells. *Biochem Pharm.* **87**, 292–302 (2014).
73. Kpetemey, M. et al. MIEN1, a novel interactor of Annexin A2, promotes tumor cell migration by enhancing AnxA2 cell surface expression. *Mol. Cancer* **14**, 156 (2015).
74. Yuan, J. et al. Tyr23 phosphorylation of Anxa2 enhances STAT3 activation and promotes proliferation and invasion of breast cancer cells. *Breast Cancer Res. Treat.* **164**, 327–340 (2017).
75. Zhao, C. et al. CCL18 promotes the invasion and metastasis of breast cancer through Annexin A2. *Oncol. Rep.* <https://doi.org/10.3892/or.2019.7426> (2019).
76. Galliher, A. J. & Schiemann, W. P. β3Integrin and Src facilitate transforming growth factor-β mediated induction of epithelial-mesenchymal transition in mammary epithelial cells. *Breast Cancer Res.* **8** <https://doi.org/10.1186/bcr1524> (2006).
77. Galliher, A. J. & Schiemann, W. P. Src phosphorylates Tyr284 in TGF-beta type II receptor and regulates TGF-beta stimulation of p38 MAPK during breast cancer cell proliferation and invasion. *Cancer Res.* **67**, 3752–3758 (2007).
78. Bhowmick, N. A. et al. Transforming growth factor-beta1 mediates epithelial to mesenchymal transdifferentiation through a RhoA-dependent mechanism. *Mol. Biol. Cell* **12**, 27–36 (2001).
79. Nieto, M. A. The snail superfamily of zinc-finger transcription factors. *Nat. Rev. Mol. Cell Biol.* **3**, 155–166 (2002).
80. De Herreros, A. G., Peiró, S., Nassour, M. & Savagner, P. Snail family regulation and epithelial mesenchymal transitions in breast cancer progression. *J. Mammary Gland Biol. Neoplasia* **15**, 135–147 (2010).
81. Dhasarathy, A., Phadke, D., Mav, D., Shah, R. R. & Wade, P. A. The transcription factors Snail and Slug activate the transforming growth factor-beta signaling pathway in breast cancer. *PLoS ONE* **6**, e26514 (2011).
82. Wang, Y., Shi, J., Chai, K., Ying, X. & Zhou, B. The role of snail in EMT and tumorigenesis. *Curr. Cancer Drug Targets* **13**, 963–972 (2013).
83. Naber, H. P., Drabsch, Y., Snaar-Jagalska, B. E., ten Dijke, P. & van Laar, T. Snail and Slug, key regulators of TGF-beta-induced EMT,



- are sufficient for the induction of single-cell invasion. *Biochem. Biophys. Res. Commun.* **435**, 58–63 (2013).
84. Hajra, K. M., Chen, D. Y. & Fearon, E. R. The SLUG zinc-finger protein represses E-cadherin in breast cancer. *Cancer Res.* **62**, 1613–1618 (2002).
  85. Bolos, V. et al. The transcription factor Slug represses E-cadherin expression and induces epithelial to mesenchymal transitions: a comparison with Snail and E47 repressors. *J. Cell Sci.* **116**, 499–511 (2003).
  86. Wang, S. P. et al. p53 controls cancer cell invasion by inducing the MDM2-mediated degradation of Slug. *Nat. Cell Biol.* **11**, 694–704 (2009).
  87. Shi, D. & Gu, W. Dual roles of MDM2 in the regulation of p53: ubiquitination dependent and ubiquitination independent mechanisms of MDM2 repression of p53 activity. *Genes Cancer* **3**, 240–248 (2012).
  88. Araki, S. et al. TGF-beta1-induced expression of human Mdm2 correlates with late-stage metastatic breast cancer. *J. Clin. Invest.* **120**, 290–302 (2010).
  89. Lu, X. et al. Mouse double minute 2 (MDM2) upregulates Snail expression and induces epithelial-to-mesenchymal transition in breast cancer cells in vitro and in vivo. *Oncotarget* **7**, 37177–37191 (2016).
  90. Dong, C. et al. G9a interacts with Snail and is critical for Snail-mediated E-cadherin repression in human breast cancer. *J. Clin. Investig.* **122**, 1469–1486 (2012).
  91. Noman, M. Z. et al. The immune checkpoint ligand PD-L1 is upregulated in EMT-activated human breast cancer cells by a mechanism involving ZEB-1 and miR-200. *Oncoimmunology* **6**, e1263412 (2017).
  92. Ding, X., Park, S. I., McCauley, L. K. & Wang, C. Y. Signaling between transforming growth factor beta (TGF-beta) and transcription factor SNAI2 represses expression of microRNA miR-203 to promote epithelial-mesenchymal transition and tumor metastasis. *J. Biol. Chem.* **288**, 10241–10253 (2013).
  93. Fujita, N. et al. MTA3, a Mi-2/NuRD complex subunit, regulates an invasive growth pathway in breast cancer. *Cell* **113**, 207–219 (2003).
  94. Wu, W. et al. LncRNA NKILA suppresses TGF-beta-induced epithelial-mesenchymal transition by blocking NF-kappaB signaling in breast cancer. *Int. J. Cancer* **143**, 2213–2224 (2018).
  95. Ni, T. et al. Snail1-dependent p53 repression regulates expansion and activity of tumour-initiating cells in breast cancer. *Nat. Cell Biol.* **18**, 1221–1232 (2016).
  96. Li, H. M. et al. A potent CBP/p300-Snail interaction inhibitor suppresses tumor growth and metastasis in wild-type p53-expressing cancer. *Sci. Adv.* **6**, eaaw8500 (2020).
  97. Thuault, S. et al. Transforming growth factor-beta employs HMGA2 to elicit epithelial-mesenchymal transition. *J. Cell Biol.* **174**, 175–183 (2006).
  98. Lee, Y. J., Park, J. H. & Oh, S. M. Activation of NF-kappaB by TOPK upregulates Snail/Slug expression in TGF-beta1 signaling to induce epithelial-mesenchymal transition and invasion of breast cancer cells. *Biochem. Biophys. Res. Commun.* **530**, 122–129 (2020).
  99. Bai, J. W. et al. The zinc-finger transcriptional factor Slug transcriptionally downregulates ERalpha by recruiting lysine-specific demethylase 1 in human breast cancer. *Oncogenesis* **6**, e330 (2017).
  100. Shin, I., Kim, S., Song, H., Kim, H. R. & Moon, A. H-Ras-specific activation of Rac-MKK3/6-p38 pathway: its critical role in invasion and migration of breast epithelial cells. *J. Biol. Chem.* **280**, 14675–14683 (2005).
  101. Lee, J. 3,3'-diindolylmethane inhibits TNF-alpha- and TGF-beta-induced epithelial-mesenchymal transition in breast cancer cells. *Nutr. Cancer* **71**, 992–1006 (2019).
  102. Bracken, C. P. et al. A double-negative feedback loop between ZEB1-SIP1 and the microRNA-200 family regulates epithelial-mesenchymal transition. *Cancer Res.* **68**, 7846–7854 (2008).
  103. Gregory, P. A. et al. An autocrine TGF-beta/ZEB/miR-200 signaling network regulates establishment and maintenance of epithelial-mesenchymal transition. *Mol. Biol. Cell* **22**, 1686–1698 (2011).
  104. Mooney, S. M. et al. The GRHL2/ZEB feedback loop—a key axis in the regulation of EMT in breast cancer. *J. Cell Biochem.* **118**, 2559–2570 (2017).
  105. Feldker, N. et al. Genome-wide cooperation of EMT transcription factor ZEB1 with YAP and AP-1 in breast cancer. *EMBO J.* **39**, e103209 (2020).
  106. Wu, R. S. et al. OVOL2 antagonizes TGF-beta signaling to regulate epithelial to mesenchymal transition during mammary tumor metastasis. *Oncotarget* **8**, 39401–39416 (2017).
  107. Xiang, X. et al. Grhl2 determines the epithelial phenotype of breast cancers and promotes tumor progression. *PLoS ONE* **7**, e50781 (2012).
  108. Scheel, C. et al. Paracrine and autocrine signals induce and maintain mesenchymal and stem cell states in the breast. *Cell* **145**, 926–940 (2011).
  109. Dong, M. et al. The type III TGF-beta receptor suppresses breast cancer progression. *J. Clin. Investig.* **117**, 206–217 (2007).
  110. Leng, X. et al. Inhibition of Lipocalin 2 Impairs Breast Tumorigenesis and Metastasis. *Cancer Res.* **69**, 8579–8584 (2009).
  111. Wang, X. et al. KLF8 promotes human breast cancer cell invasion and metastasis by transcriptional activation of MMP9. *Oncogene* **30**, 1901–1911 (2011).
  112. Sheridan, C. et al. CD44+/CD24- breast cancer cells exhibit enhanced invasive properties: an early step necessary for metastasis. *Breast Cancer Res.* **8**, R59 (2006).
  113. Yu, Y. et al. miR-190 suppresses breast cancer metastasis by regulation of TGF-beta-induced epithelial-mesenchymal transition. *Mol. Cancer* **17**, 70 (2018).
  114. Ma, F. et al. MiR-23a promotes TGF-beta1-induced EMT and tumor metastasis in breast cancer cells by directly targeting CDH1 and activating Wnt/beta-catenin signaling. *Oncotarget* **8**, 69538–69550 (2017).
  115. Yu, J. et al. MicroRNA-182 targets SMAD7 to potentiate TGFbeta-induced epithelial-mesenchymal transition and metastasis of cancer cells. *Nat. Commun.* **7**, 13884 (2016).
  116. Smith, A. L. et al. The miR-106b-25 cluster targets Smad7, activates TGF-beta signaling, and induces EMT and tumor initiating cell characteristics downstream of Six1 in human breast cancer. *Oncogene* **31**, 5162–5171 (2012).
  117. Kong, W. et al. MicroRNA-155 is regulated by the transforming growth factor beta/Smad pathway and contributes to epithelial cell plasticity by targeting RhoA. *Mol. Cell Biol.* **28**, 6773–6784 (2008).
  118. Lei, B. et al. miR-615-3p promotes the epithelial-mesenchymal transition and metastasis of breast cancer by targeting PICK1/TGFBRI axis. *J. Exp. Clin. Cancer Res.* **39**, 71 (2020).
  119. Chou, J. et al. MALAT1 induced migration and invasion of human breast cancer cells by competitively binding miR-1 with cdc42. *Biochem. Biophys. Res. Commun.* **472**, 262–269 (2016).
  120. Chang, C.-J. et al. p53 regulates epithelial-mesenchymal transition and stem cell properties through modulating miRNAs. *Nat. Cell Biol.* **13**, 317–323 (2011).
  121. Ma, L., Teruya-Feldstein, J. & Weinberg, R. A. Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. *Nature* **449**, 682–688 (2007).
  122. Ma, L. et al. Therapeutic silencing of miR-10b inhibits metastasis in a mouse mammary tumor model. *Nat. Biotechnol.* **28**, 341–347 (2010).
  123. Liu, Y. et al. MicroRNA-10b targets E-cadherin and modulates breast cancer metastasis. *Med. Sci. Monit.* **18**, BR299–BR308 (2012).

124. Han, X. et al. Critical role of miR-10b in transforming growth factor-beta1-induced epithelial-mesenchymal transition in breast cancer. *Cancer Gene Ther.* **21**, 60–67 (2014).
125. De Ruijter, T. C., Veeck, J., De Hoon, J. P. J., Van Engeland, M. & Tjan-Heijnen, V. C. Characteristics of triple-negative breast cancer. *J. Cancer Res. Clin. Oncol.* **137**, 183–192 (2011).
126. Wang, S. E. et al. Transforming growth factor beta engages TACE and ErbB3 to activate phosphatidylinositol-3 kinase/Akt in ErbB2-overexpressing breast cancer and desensitizes cells to trastuzumab. *Mol. Cell Biol.* **28**, 5605–5620 (2008).
127. Gupta, P. & Srivastava, S. K. HER2 mediated de novo production of TGFbeta leads to SNAIL driven epithelial-to-mesenchymal transition and metastasis of breast cancer. *Mol. Oncol.* **8**, 1532–1547 (2014).
128. Burnett, J. P. et al. Trastuzumab resistance induces EMT to transform HER2(+) PTEN(-) to a triple negative breast cancer that requires unique treatment options. *Sci. Rep.* **5**, 15821 (2015).
129. Wu, Y. et al. A83-01 inhibits TGF-beta-induced upregulation of Wnt3 and epithelial to mesenchymal transition in HER2-overexpressing breast cancer cells. *Breast Cancer Res. Treat.* **163**, 449–460 (2017).
130. Ito, I. et al. Estrogen inhibits transforming growth factor beta signaling by promoting Smad2/3 degradation. *J. Biol. Chem.* **285**, 14747–14755 (2010).
131. Rajarajan, S. et al. An androgen receptor regulated gene score is associated with epithelial to mesenchymal transition features in triple negative breast cancers. *Transl. Oncol.* **37**, 101761 (2023).
132. Yu, L., Hebert, M. C. & Zhang, Y. E. TGF-beta receptor-activated p38 MAP kinase mediates Smad-independent TGF-beta responses. *EMBO J.* **21**, 3749–3759 (2002).
133. Kalluri, R. & Weinberg, R. A. The basics of epithelial-mesenchymal transition. *J. Clin. Invest.* **119**, 1420–1428 (2009).
134. Kohn, K. W. Molecular interaction map of the mammalian cell cycle control and DNA repair systems. *Mol. Biol. Cell* **10**, 2703–2734 (1999).
135. Oda, K., Matsuoka, Y., Funahashi, A. & Kitano, H. A comprehensive pathway map of epidermal growth factor receptor signaling. *Mol. Syst. Biol.* **1**, 2005 0010 (2005).
136. Calzone, L., Gelay, A., Zinovyev, A., Radvanyi, F. & Barillot, E. A comprehensive modular map of molecular interactions in RB/E2F pathway. *Mol. Syst. Biol.* **4**, 173 (2008).
137. Polouliakh, N., Nock, R., Nielsen, F. & Kitano, H. G-protein coupled receptor signaling architecture of mammalian immune cells. *PLoS ONE* **4**, e4189 (2009).
138. Mizuno, S. et al. AlzPathway: a comprehensive map of signaling pathways of Alzheimer's disease. *BMC Syst. Biol.* **6**, 52 (2012).
139. Kuperstein, I. et al. Atlas of cancer signalling network: a systems biology resource for integrative analysis of cancer data with Google Maps. *Oncogenesis* **4**, e160 (2015).
140. Ostaszewski, M. et al. COVID19 disease map, a computational knowledge repository of virus-host interaction mechanisms. *Mol. Syst. Biol.* **17**, e10387 (2021).
141. Serhan, C. N. et al. The atlas of inflammation resolution (AIR). *Mol. Asp. Med.* **74**, 100894 (2020).
142. Singh, V. et al. RA-map: building a state-of-the-art interactive knowledge base for rheumatoid arthritis. *Database* **2020**, baaa017 (2020).
143. Steinway, S. N. et al. Combinatorial interventions inhibit TGFbeta-driven epithelial-to-mesenchymal transition and support hybrid cellular phenotypes. *NPJ Syst. Biol. Appl.* **1**, 15014 (2015).
144. Kocharczyk, M. et al. Relaxation oscillations and hierarchy of feedbacks in MAPK signaling. *Sci. Rep.* **7**, 38244 (2017).
145. Sible, J. C. & Tyson, J. J. Mathematical modeling as a tool for investigating cell cycle control networks. *Methods* **41**, 238–247 (2007).
146. Pomerening, J. R., Sontag, E. D. & Ferrell, J. E. Building a cell cycle oscillator: hysteresis and bistability in the activation of Cdc2. *Nat. Cell Biol.* **5**, 346–351 (2003).
147. Zhang, J. et al. TGF-beta-induced epithelial-to-mesenchymal transition proceeds through stepwise activation of multiple feedback loops. *Sci. Signal.* **7**, ra91 (2014).
148. Bhavani, G. S. & Palanisamy, A. SNAIL driven by a feed forward loop motif promotes tgfbeta induced epithelial to mesenchymal transition. *Biomed. Phys. Eng. Express* **8** <https://doi.org/10.1088/2057-1976/ac7896> (2022).
149. Sai Bhavani, G. & Palanisamy, A. Network motifs and hypermotifs in TGFbeta-induced epithelial to mesenchymal transition and metastasis. *Front. Syst. Biol.* **3** <https://doi.org/10.3389/fsysb.2023.1099951> (2023).
150. Lu, M., Jolly, M. K., Levine, H., Onuchic, J. N. & Ben-Jacob, E. MicroRNA-based regulation of epithelial-hybrid-mesenchymal fate determination. *Proc. Natl Acad. Sci. USA* **110**, 18144–18149 (2013).
151. Lu, M. et al. Tristability in cancer-associated microRNA-TF chimera toggle switch. *J. Phys. Chem. B* **117**, 13164–13174 (2013).
152. Steinway, S. N. et al. Network modeling of TGFbeta signaling in hepatocellular carcinoma epithelial-to-mesenchymal transition reveals joint sonic hedgehog and Wnt pathway activation. *Cancer Res.* **74**, 5963–5977 (2014).
153. Gillespie, M. et al. The reactome pathway knowledgebase 2022. *Nucleic Acids Res.* **50**, D687–D692 (2022).
154. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
155. Szklarczyk, D. et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* **49**, D605–D612 (2021).
156. Krämer, A., Green, J., Pollard, J. & Tugendreich, S. Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics* **30**, 523–530 (2014).
157. Warde-Farley, D. et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* **38**, W214–W220 (2010).
158. Kitano, H., Funahashi, A., Matsuoka, Y. & Oda, K. Using process diagrams for the graphical representation of biological networks. *Nat. Biotechnol.* **23**, 961–966 (2005).
159. Gawron, P. et al. MINERVA—a platform for visualization and curation of molecular interaction networks. *NPJ Syst. Biol. Appl.* **2**, 16020 (2016).
160. Hoksza, D., Gawron, P., Ostaszewski, M., Smula, E. & Schneider, R. MINERVA API and plugins: opening molecular network analysis and visualization to the community. *Bioinformatics* **35**, 4496–4498 (2019).
161. Hoksza, D., Gawron, P., Ostaszewski, M., Hasenauer, J. & Schneider, R. Closing the gap between formats for storing layout information in systems biology. *Brief. Bioinform.* **21**, 1249–1260 (2020).
162. Sherekar, S. & Viswanathan, G. A. Boolean dynamic modeling of cancer signaling networks: prognosis, progression, and therapeutics. *Comput. Syst. Oncol.* **1** <https://doi.org/10.1002/cso2.1017> (2021).
163. Kauffman, S. A. Metabolic stability and epigenesis in randomly constructed genetic nets. *J. Theor. Biol.* **22**, 437–467 (1969).
164. Glass, L. & Kauffman, S. A. The logical analysis of continuous, non-linear biochemical control networks. *J. Theor. Biol.* **39**, 103–129 (1973).
165. Thomas, R. Boolean formalization of genetic control circuits. *J. Theor. Biol.* **42**, 563–585 (1973).
166. Chaouiya, C. et al. SBML qualitative models: a model representation format and infrastructure to foster interactions between qualitative modelling formalisms and tools. *BMC Syst. Biol.* **7**, 135 (2013).
167. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).

168. Chin, C. H. et al. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol.* **8**, S11 (2014).
169. Tang, Z., Li, C., Kang, B., Gao, G. & Zhang, Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* **45**, W98–W102 (2017).
170. Bartha, Á. & Gyórfy, B. TNMplot.com: a web tool for the comparison of gene expression in normal, tumor and metastatic tissues. *Int. J. Mol. Sci.* **22** <https://doi.org/10.3390/ijms22052622> (2021).

## Acknowledgements

Our research is supported by facilities and scholarship provided by NITW. GSB is supported by an NITW Institute scholarship. All the authors acknowledge the facilities and support extended by NIT Warangal. We thank Dr. Anna Niarakis, University of Toulouse for introducing CaSQ and Cell Collective. We thank Dr. Marek Ostaszewski, University of Luxembourg for help with the MINERVA web application. We thank Prof. KV Venkatesh, IIT Bombay, Founder & Director MetFlux, Karan Lamore and Rohith Rahman, MetFlux for their help and support with the public web upload of the developed MBC map.

## Author contributions

A.P. and G.S.B. designed the work; G.S.B. developed the map and performed the modelling and analysis study. T.S.G. periodically reviewed the work and has provided valuable suggestions. A.P. and G.S.B. wrote the article. All the authors have reviewed the article.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41540-024-00378-w>.

**Correspondence** and requests for materials should be addressed to Anbumathi Palanisamy.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024