Epithelial/mesenchymal plasticity: how have quantitative mathematical models helped improve our understanding?

Mohit Kumar Jolly1, Satyendra C. Tripathi2, Jason A. Somarelli3, Samir M. Hanash2 and Herbert Levine1

1 Center for Theoretical Biological Physics, Rice University, Houston, TX, USA
2 Department of Clinical Cancer Prevention, UT MD Anderson Cancer Center, Houston, TX, USA
3 Department of Medicine, Duke Cancer Institute, Duke University, Durham, NC, USA

Keywords
- circulating tumor cells; collective cell migration; epithelial-mesenchymal transition; hybrid epithelial/mesenchymal; mathematical modeling; stemness

Correspondence
H. Levine, Center for Theoretical Biological Physics, Rice University, Houston, TX 77005, USA
E-mail: herbert.levine@rice.edu

(Received 26 January 2017, revised 11 May 2017, accepted 18 May 2017, available online 19 June 2017)

doi:10.1002/1878-0261.12084

1. Introduction

A remarkable feature that cancer cells use to evade therapy, metastasize, and drive tumor progression is phenotypic plasticity, that is, the ability of cells to switch back and forth among multiple phenotypes in response to varied internal or external signals (Holzel et al., 2012). Phenotypic plasticity is usually tightly controlled during adult homeostasis. It comes into play only when needed, such as during tissue repair, when resident stem cells give rise to cells that need to be replenished. However, during tumor progression, many of the molecular brakes against phenotypic plasticity are deregulated, enabling cancer cells to behave as ‘moving targets’ that can play ‘hide-and-seek’ with multiple therapeutic regimes (Roesch, 2015; Varga et al., 2014). In addition, these phenotypic conversions can facilitate adaptation by enabling genetically identical cells to exhibit a diverse set of phenotypes and may also help fuel genetic evolution of cancer cells (Brooks et al., 2015; Mooney et al., 2016; Yadav et al., 2016).

A canonical example of such phenotypic plasticity that contributes significantly to both metastasis and drug resistance is epithelial/mesenchymal (E/M) transition. The ability of cells to reversibly alter their phenotypes in response to signals, presents a significant clinical challenge to treating solid tumors. Tumor cells utilize phenotypic plasticity to evade therapies, metastasize, and colonize distant organs. As a result, phenotypic plasticity can accelerate tumor progression. A well-studied example of phenotypic plasticity is the bidirectional conversions among epithelial, mesenchymal, and hybrid epithelial/mesenchymal (E/M) phenotype(s). These conversions can alter a repertoire of cellular traits associated with multiple hallmarks of cancer, such as metabolism, immune evasion, invasion, and metastasis. To tackle the complexity and heterogeneity of these transitions, mathematical models have been developed that seek to capture the experimentally verified molecular mechanisms and act as ‘hypothesis-generating machines’. Here, we discuss how these quantitative mathematical models have helped us explain existing experimental data, guided further experiments, and provided an improved conceptual framework for understanding how multiple intracellular and extracellular signals can drive E/M plasticity at both the single-cell and population levels. We also discuss the implications of this plasticity in driving multiple aggressive facets of tumor progression.

Abbreviations
- CSCs, cancer stem cells; CTCs, circulating tumor cells; E/M, epithelial/mesenchymal; ECM, extracellular matrix; EMT, epithelial-mesenchymal transition; MET, mesenchymal-epithelial transition; NSCLC, non-small-cell lung cancer; PSFs, phenotypic stability factors; VIM, vimentin.
plasticity, that is, the ability of cells to undergo a partial or full epithelial-mesenchymal transition (EMT) and its reverse mesenchymal-epithelial transition (MET) (Diepenbruck and Christofori, 2016). Interestingly, emerging evidence strongly suggests that these transitions are rarely ‘all-or-none’. Rather, cancer cells can often display a hybrid E/M phenotype by combining various epithelial and mesenchymal morphological and/or molecular features (Jolly et al., 2015; Nieto, 2013; Nieto et al., 2016). Cells in this (these) hybrid state(s) can be much more tumorigenic and drug resistant as compared to those that are more fixed in a strongly epithelial or mesenchymal state (Biddle et al., 2016; Grosse-Wilde et al., 2015; Jolly et al., 2015). Thus, elucidating the underlying principles of these dynamic transitions is of foundational importance for countering the yet insuperable clinical aspects of cancer – metastasis and drug resistance.

Recent progress in dissecting the molecular mechanisms underlying these phenotypic transitions has enabled the development of quantitative mathematical models that can be used as hypothesis-generating tools to guide further experiments. In this review, we highlight how an integrative theoretical-experimental approach has helped us better characterize E/M plasticity. For instance, mathematical models capturing the dynamics of core EMT signaling network have predicted that cells can maintain a hybrid E/M phenotype stably and that cells with same genetic background (cell lines) can contain admixtures of epithelial, hybrid E/M, and mesenchymal subpopulations. These predictions have been validated by experimental observations showing different cell lines can contain subpopulations of different phenotypes in varying ratios.

2. Why develop quantitative mathematical models?

Quantitative mathematical models offer us a powerful conceptual framework to elucidate underlying biological mechanisms and to propose new sets of experiments by generating falsifiable hypotheses. They can help interpret or explain the existing experimental data, confirm or reject alternate hypotheses, predict cellular behaviour, and eventually guide further experiments (Mobius and Laan, 2015). They can decode the emergent dynamics of various regulatory networks and biological phenomena, and enable the experimental biologists to think more quantitatively in terms of regulatory dynamics. Mathematical models can also help unravel the principles that govern cancer progression, from the molecular scale all the way to the population level (Anderson and Quaranta, 2008). Thus, these models can aid in guiding optimal treatment modalities and can contribute to improved risk prognoses (Altrock et al., 2015).

3. What is a quantitative mathematical model?

A model of any system is a replica that captures the system’s essential features and can thus be used to predict how the ‘original’ system would behave in a variety of conditions. Each model has its own assumptions, strengths, and limitations and is therefore suitable to answer a specific set of questions. In biology, we often use various preclinical models (e.g., cell lines, mouse models, patient-derived xenografts) to investigate different phenomenon relevant to human biology, with an implicit expectation that lessons learned in these preclinical models can provide useful insights into the functioning of the human system. Broadly speaking, these biological models can be in vitro or in vivo. Similar to these model systems, a quantitative mathematical model is an in silico representation of the ‘original’ system, where a set of equations captures the essence of biological phenomenon through terms representing different objects involved in a phenomenon and interactions among them that govern that phenomenon (Fig. 1A). A bidirectional communication among mathematical and experimental biologists has been fruitful in teasing out the mechanistic aspects of many biological processes such as timing and patterning of developmental events (Lewis, 2008; Oates et al., 2009; Shaya and Sprinzak, 2011).

Just like biological models, mathematical models differ in scope and purpose (Mobius and Laan, 2015). For instance, different mathematical models developed to understand E/M plasticity have focused on different questions: (a) ‘How do a set of transcription factors and microRNAs (miR) regulate the intracellular dynamics of a partial or full EMT/MET and modulate phenotypic heterogeneity in an isogenic population (Lu et al., 2013; Steinway et al., 2014; Tian et al., 2013)?’; (b) ‘How does cell–cell communication affect the spatial arrangement of epithelial, mesenchymal and hybrid E/M cells (Boareto et al., 2016)?’; and (c) ‘How do cells alter their morphological and motility traits during EMT?’ As one may suspect, developing mathematical models to answer each of these questions requires quite different experimental data. Therefore, often times the scope of the model is decided by the data that are available; for example, whether longitudinal data are available either in discrete time points or in a more continuous fashion, whether data are available at a population vs. single-cell level, or whether the available data are merely for altered protein and
transcript levels vs. the data also includes morphology and motility aspects too. In this review, we will focus on a set of mathematical models that can be compared extensively against the existing experimental data.

4. How does one develop a quantitative mathematical model?

As discussed earlier, the first step in developing a mathematical model entails being clear both about the biological question that the model should be able to answer, and the experimental data available with which to construct, calibrate, and compare the model. Second, one must realize the implicit assumptions of different modeling frameworks and decide whether operating under those assumptions enables a reasonable replica of the ‘original’ biological system. These assumptions should always be judged in light of the question/phenomenon of interest. Third, one should strive to accurately incorporate multiple key features of a phenomenon in one’s model. Finally, the model should be validated by comparing the predictions of the model in cases where robust experimental data are available a priori. Subsequent to model validation, one can generate predictions that can be tested experimentally and confirmed or falsified (Fig. 1B).

Generating falsifiable predictions is the most useful application of developing mathematical models. Therefore, simply fitting experimental data to a model does little to contribute to new knowledge. Rather, one should seek to ‘stick the model’s neck out after it is fitted and try to falsify it’ (Gunawardena, 2014) by predicting how the ‘original’ system (often, the biological model system being studied) would behave under altered conditions, such as by introducing genetic mutations or overexpressing a specific gene.

What happens if there is a mismatch between the prediction of the mathematical model and the experimental results generated? This mismatch can occur due to multiple reasons, such as (a) underlying assumptions of the model are not entirely valid; (b) the model is not robust, that is, relatively small changes to the model or its parameters dramatically change the behavior of the model; and/or (c) technical inaccuracies in running experiments and/or model simulations. Once the underlying reason(s) is (are) identified, and predictions of the mathematical model score well with experimental results, this iterative cycle can continue to identify the next set of exciting research directions to be answered using the same or a different mathematical and/or biological model(s), as applicable.
5. How can epithelial/mesenchymal plasticity be represented by a set of mathematical equations?

An exemplary biological phenomenon in which mathematical modeling has helped provide useful biological insights is that of E/M plasticity. This plasticity arises via a gene regulatory network that controls reversible switches between phenotypes, and has implications for numerous key biological processes in normal and disease states. For example, in the context of cancer, phenotypic switching between epithelial and mesenchymal states via EMT and MET drives cancer progression, metastasis, and therapy resistance. These epithelial and mesenchymal cells have distinct morphological and molecular features. For instance, epithelial cells have E-cadherin (CDH1) localized at the cell membrane, which contributes to adherens junctions. Conversely, mesenchymal cells lack E-cadherin and typically have higher levels of vimentin (VIM), N-cadherin (CDH2), and αSMA (smooth muscle actin). Thus, EMT and MET typically involve widespread changes in gene expression, micro-RNAs, and epigenetic profiles, as well as cytoskeletal reprogramming (De Craene and Berx, 2013). An understanding of the set of molecular players of interest and the interactions among them can facilitate development of a mathematical model that can trace these changes during EMT and MET, and potentially highlight novel areas of susceptibility to therapeutic targeting.

The first set of mathematical models developed for EMT/MET focused on a specific question: Can the underlying EMT/MET regulatory network enable the existence of a stable hybrid E/M phenotype, and if so, what is the molecular signature of this hybrid E/M phenotype (Lu et al., 2013; Tian et al., 2013)? These efforts at addressing this question modeled the interactions among two sets of microRNAs and two sets of transcription factors that were reported to govern EMT/MET in multiple cell lines – miR-34, miR-200, ZEB, and SNAIL (Bracken et al., 2008; Gregory et al., 2008; Kim et al., 2011; Park et al., 2008; Siemens et al., 2011) (Fig. 2A). The models predicted
that under certain conditions, a hybrid E/M phenotype can be stable and that in isogenic populations, multiple phenotypes can co-exist. In other words, a clonal population may harbor more than one phenotypic subpopulation, owing to the nonlinear and highly interconnected feedback loops among a set of core EMT players (see three solid blue lines in Fig. 2B), each of which represents a distinct phenotype – E (low ZEB1), hybrid E/M (medium ZEB1), and M (high ZEB1) – as illustrated by cartoons drawn alongside. These predictions were later validated by experiments demonstrating subpopulations of E, hybrid E/M, and M phenotypes in varying ratios in cell lines across multiple cancer types, as assessed by flow cytometry and immunofluorescence (Andriani et al., 2016; Grosse-Wilde et al., 2015; Jolly et al., 2016b; Ruscetti et al., 2016).

Such co-existing phenotypes, as also observed experimentally in H2291 and DU145 cells (Fig. 2B), may enable dynamic switching among cells in different phenotypes (Ruscetti et al., 2016). This heterogeneity does not eliminate the possibility that under certain scenarios (i.e., in some cell lines), most, if not all, isogenic cells display the same phenotype. For instance, the model predicted regions corresponding to solely epithelial (SNAIL < 180K molecules in Fig. 2B) and solely mesenchymal (SNAIL > 230K molecules in Fig. 2B) states, as validated experimentally by H820 and H1792 cells, respectively (Fig. 2B). It should be noted that the baseline models predicted such homogeneous regions only for epithelial and mesenchymal phenotypes, but not for a hybrid E/M phenotype.

More importantly, these models motivated the investigation of behavior of a set of non-small-cell lung cancer (NSCLC) cell lines that were categorized as ‘hybrid’ based on population-based measurements (Schliekelman et al., 2015). At a single-cell level, these ‘hybrid’ cell lines contained subpopulations of epithelial and mesenchymal cells (H2291; Fig. 2B) and/or individual cells co-expressing epithelial and mesenchymal markers, such as CDH1 and VIM (H1975). H1975 cells exhibited a hybrid E/M phenotype at the single-cell level over multiple passages (Fig. 3A), strongly suggesting that a hybrid E/M state can be a stable phenotype (Jolly et al., 2016b). As compared to epithelial cells (H820) and mesenchymal cells (H1299) (Schliekelman et al., 2015), H1975 cells also stained for nuclear ZEB1 (Jia et al., 2017), thus confirming the prediction made by the mathematical model developed by Lu et al. (2013) (Fig. 3B).

Observations in H1975 cells serve as a remarkable example of the power of leveraging an integrated theoretical-experimental framework. Although the mathematical models predicted regions where a hybrid E/M state may exist as a stable phenotype (see solid blue line corresponding to 200 < ZEB1 mRNA levels < 600 molecules in Fig. 2B), as already noted, a parameter region enabling a hybrid E/M state alone was not observed. Consequently, that led to a search for potential ‘phenotypic stability factors’ (PSFs) – molecular players that can enable a monostable {E/M} region. Incorporating two proteins OVOL2 and GRHL2 that were reported to form mutually inhibitory loops with ZEB (Cieply et al., 2012, 2013; Roca et al., 2013) – in the mathematical model – predicted the existence of a desired {E/M} region (Hong et al., 2015; Jia et al., 2015; Jolly et al., 2016b) (Fig. 3A). The role of OVOL2 and GRHL2 as PSFs was validated by experiments showing that knockdown of either of these proteins in H1975 drove the cells toward a stable hybrid E/M state to a fully mesenchymal phenotype (Jolly et al., 2016b).

Similar results in developmental EMT contexts strengthened the notion that these PSFs can act as ‘molecular brakes’ on EMT that can prevent cells ‘that have gained partial plasticity’ from undergoing a complete EMT (Watanabe et al., 2014; Werner et al., 2013). Furthermore, the mathematical model suggested that overexpression of PSFs can drive an MET, a prediction already verified in breast and prostate cancer cell lines (Roca et al., 2013; Werner et al., 2013), and kidney cells (Aue et al., 2015), thereby indicating that such models can behave as ‘semiquantitative predictive paradigms’ to predict the cellular behavior pertinent to EMT regulation in multiple cell lines.

These mathematical models also proposed certain network motifs that can be used to identify further PSFs, one of which is that a potential PSF typically forms a double negative feedback loop with ZEB (Jolly et al., 2016b). Given that E-cadherin is a transcriptional target of ZEB, and E-cadherin can sequester β-catenin on the cell membrane, thus inhibiting transcriptional activation of ZEB via Wnt/β-catenin pathway (Mooney et al., 2016), E-cadherin and ZEB seem to repress one another. Thus, E-cadherin can be thought of as a potential PSF. However, detailed mechanism-based models need to be constructed to investigate that possibility comprehensively.

Despite the utility of these models, it is important to note that we neither claim that these particular models can accurately predict EMT regulation for all cell lines nor that they can necessarily predict responses to all perturbations that can alter EMT status in a given cell line. For instance, overexpression of GRHL2 did not drive MET in the RD and 143B human sarcoma cell lines (Somarelli et al., 2016). Further experiments indicated that in RD and 143B, GRHL2 coupled to
miR-200 and ZEB1 in a different topology as compared to that in multiple (adenoc)carcinoma cell lines. Therefore, GRHL2 did not seem to couple with miR-200/ZEB feedback loop in one of the topologies proposed to identify potential PSFs (Jolly et al., 2016b). Consistently, in sarcoma cells, GRHL2 had no effect on ZEB1 levels. Instead, GRHL2-induced changes were only observed when ZEB1 was knocked down. These findings led to the development of a revised mathematical model that captured these newly revealed interactions. The revised model was able to reproduce robustly the key features of experiments in sarcoma cells, such as the synergistic induction of E-cadherin levels upon overexpression of both GRHL2 and miR-200 (Somarelli et al., 2016), and predicted how epigenetic regulation of GRHL2 can modulate MET. Therefore, ‘no one size fits all’; no model – either biological or mathematical – fits all different biological contexts; carcinoma cell lines may not be reliable biological models to understand sarcoma biology, and similarly, networks that work well for predicting carcinoma cell line behavior need not be the same for sarcoma.

Notwithstanding the complexity and heterogeneity in the gene regulatory networks that drive EMT and MET in different contexts, mathematical models can be constructed to help rationalize existing experimental results and to guide further experiments, by making certain approximations or estimations about the model parameters. Each of the mathematical models developed above has multiple variables – ZEB, miR-200, GRHL2, etc. – and each variable is represented by an equation tracing their levels over time. Each equation has terms representing the innate production and degradation rates for those species that can be estimated from their half-lives and/or typical number of molecules in a cell (Milo et al., 2010). Similarly, each equation contains terms pertaining to regulation of the respective species by one another, for instance, inhibition of ZEB by miR-200. The quantitative parameters
describing these interactions, such as the number of binding sites and the fold change in levels upon over-expression or inhibition, can also be gained from relevant experimental data. For example, whereas miR-200s can bind up to eight to nine binding sites on Zeb mRNA and reduces the protein levels by 90% (Gregory et al., 2011), miR-34 binds to two binding sites on Snail mRNA and reduces the protein levels only by 50% (Kim et al., 2011). Upon estimating a relevant range of parameter variation, the sensitivity of these mathematical models to different parameters can be tested. For instance, the range of levels of SNAIL for which a hybrid E/M phenotype is observed is largely robust to ±20% variation in parameters (Jia et al., 2015). Thus, one need not know the exact value of each parameter in the mathematical model for every cell line. Instead, estimating their typical range from the experimental data can be a good first approximation. This approximation is good because it can be often impossible to perform all experiments to measure every single parameter for every single-cell line, and these measurements can themselves be subject to uncertainty (Azeloglu and Iyengar, 2015; Kirk et al., 2015).

Deriving mathematical models to represent biological systems is rarely straightforward (Kirk et al., 2015). Thus, a key to justifiably use mathematical models is to state the assumptions and uncertainty in the model structure and/or parameters clearly. If one believes the assumptions of the model, one must also believe its conclusions (Gunawardena, 2014) – and this applies to both mathematical and biological models. For instance, in models of the EMT/MET regulatory network described above (Lu et al., 2013; Tian et al., 2013), more than one family member of a protein or microRNA are lumped into one variable, for the sake of simplicity. So, an implicit assumption of these mathematical models is that, for instance, both ZEB1 and ZEB2 – two members of the ZEB family – behave identically, which need not be true in all contexts. Similarly, in the context of biological models, an underlying assumption in in vitro cell culture is that the observed behavior of cells in a two-dimensional setup plated on plastic recapitulates the ‘true’ behavior of cells in vivo.

6. How can mathematical models be used to study changes in other cellular traits connected with EMT/MET?

EMT and MET are considered as the motors of cellular plasticity due to their coupling with other cellular traits such as metabolism, tumor-initiating potential, genome plasticity, drug resistance, immunosuppression, and cell–cell communication (Brabletz et al., 2011; Chen et al., 2014; Fischer et al., 2015; Lu et al., 2014; Mani et al., 2008; Morel et al., 2008; Tripathi et al., 2016; Wellner et al., 2010; Zheng et al., 2015). By using mathematical models similar to those described above, one can investigate the interplay of EMT/MET with any one or more of these traits.

For instance, mathematical models have helped reconcile apparently contradictory results with regard to the interplay between EMT/MET and ‘stemness’ or tumor-initiating potential. Initially, EMT was proposed to promote a gain of stem-like properties (Mani et al., 2008; Morel et al., 2008). However, later studies suggested that cells locked in a mesenchymal phenotype often lose their stem-like traits (Celíà-Terrassa et al., 2012; Tran et al., 2014) and that both epithelial-like and mesenchymal-like stem-like subpopulations may exist (Liu et al., 2014) (Fig. 4Ai–iii). To provide a unifying schema to explain these apparently conflicting results, a mathematical model was developed to connect core EMT players, miR-200 and ZEB, with the master regulators of stemness, LIN28 and let-7 (Yang et al., 2010). This model proposed that cells in a hybrid E/M phenotype can be more likely to gain stemness as compared to those in either a fully epithelial or mesenchymal state (Jolly et al., 2014) (Fig. 4Bi–ii). Follow-up experiments in breast cancer cells demonstrated that hybrid E/M cells – cells co-expressing canonical epithelial and mesenchymal genes to a similar level – can form up to 10 times more mammospheres as compared to strongly epithelial or mesenchymal cells, thus validating the prediction of the model (Grosse-Wilde et al., 2015) (Fig. 4Biii). Hybrid E/M cells also drove aggressive tumor growth in vivo (Goldman et al., 2015). Moreover, enhanced or acquired drug resistance of breast cancer and oral squamous carcinoma cells in a hybrid E/M phenotype further substantiate the proposed correlation between a hybrid E/M phenotype and ‘stemness’ (Biddle et al., 2016; Brown et al., 2016; Goldman et al., 2015). Despite initial promising validations, further research is needed to evaluate how well the hypothesis holds that the hybrid E/M state is more stem-like (Celíà-Terrassa and Kang, 2016). Moreover, the positioning of a ‘stemness window’ need not be fixed midway on the EMT axis, but could instead be much more dynamic and subtype- and/or patient-specific (Jolly et al., 2016a).

Similarly, in a study demonstrating that a mesenchymal phenotype correlates with immune evasion via reduced expression of the immunoproteasome (a proteolytic machinery that plays a key role in immunity and
homeostasis), a mathematical model was developed to capture an underlying mechanism of immunoproteasome regulation that involved STAT3, STAT1, and miR-200s (Tripathi et al., 2016). The model predicted that inhibiting the activation of STAT3 can increase the levels of immunoproteasome subunits PSMB8 and PSMB9 in mesenchymal NSCLC cell lines. Indeed, inhibition of STAT3 using rapamycin led to enhanced levels of PSMB8 and PSMB9 via an activated STAT1 pathway.

Another specific question where mathematical models may prove to be crucial to decode the underlying dynamics is the epigenetic reprogramming accompanying EMT/MET (Tam and Weinberg, 2013). The ‘poised’ chromatin state of ZEB1 in which the ZEB1 promoter simultaneously displays epigenetic marks of both active and repressed chromatin may enhance cellular plasticity among cancer stem cells (CSCs) and non-CSCs and consequently spike tumorigenic potential (Chaffer et al., 2013). Similarly, epigenetic
differences can modulate MET induction in sarcomas (Somarelli et al., 2016). Finally, these epigenetic interactions could possibly modulate the transition rates among epithelial, mesenchymal, and hybrid E/M phenotypes in specific cell lines by controlling genome-wide chromatin marks. A quantitative comparison of transition rates as measured using various reporter systems (Somarelli et al., 2013; Toneff et al., 2016) and those predicted by modeling of the underlying regulatory networks (Li et al., 2016) can bridge the gaps in our understanding of E/M plasticity. Similarly, existing theoretical frameworks to investigate epigenetic regulation (Steffen et al., 2012) can be integrated with mathematical models incorporating interconversion among CSCs and non-CSCs (Li and Wang, 2015; Yang et al., 2012; Zhou et al., 2013) and temporal mapping of epigenetic changes during EMT/MET (Kao et al., 2016) to identify the epigenetic marks that can be targeted to constrain cellular plasticity and thus abate metastatic and therapy-resistant progression.

7. How can mathematical models connect signaling aspects to cellular motility associated with EMT/MET?

Altered cellular motility and cellular morphology traits are considered to be the primary consequence of EMT/MET. During EMT, cells typically have reduced adhesion with their neighbors, and migrate collectively or individually depending on their intercellular adhesion and spatial confinement (Boekhorst et al., 2016). For instance, during embryonic development, neural crest cells undergoing a partial or complete EMT can migrate as either a multicellular stream or individually, in order to reach distant tissues. Similarly, during gastrulation, both these modes of migration are observed at different spatiotemporal coordinates (Scarpa and Mayor, 2016). Typically, collective migration is associated with a partial EMT or hybrid E/M phenotype (Kuriyama et al., 2014; Sarioglu et al., 2015), whereas fully mesenchymal cells tend to migrate alone. Depending on cell–matrix adhesion, the migrating cells can also reversibly switch to an amoeboid migration mode, where cells migrate individually and predominantly via squeezing through the gaps in extracellular matrix (ECM) (Pankova et al., 2010; Wolf et al., 2007). Similar to EMT/MET, the choice between mesenchymal and amoeboid modalities need not be a binary process and cells can exhibit signatures of both mesenchymal and amoeboid motility – lamellipodia and bleb-like protrusions, respectively (Bergert et al., 2012; Yoshida and Soldati, 2006). Preliminary mathematical models of some of the underlying signaling mechanisms governing these transitions have been developed (Huang et al., 2014, 2015), but a detailed analysis of how these molecules impinge upon changes in cytoskeletal reorganization, cell shape, cell–cell adhesion, cellular contractility, and cell–ECM mechanics and consequently drive different migration modes remains to be accomplished.

Multiple existing theoretical approaches for cell motility models focus on these key mechanical aspects. Most frameworks for single-cell migration have focused on fish keratocytes (Holmes and Edelstein-Keshet, 2012; Ziebert et al., 2012). For instance, Shao et al. (2012) illustrate how cell morphology is determined by collective effects of myosin contraction, actin polymerization, and adhesion site dynamics. This type of approach could actually make contact with the time-course data correlating cell shape with EMT states (Mandal et al., 2016; Sarkar et al., 2016). In contrast to these single-cell models, other frameworks have concentrated on tissue-level dynamics by constructing models for adhesive cell clusters and monolayers (Basan et al., 2013; Bi et al., 2016; Harris et al., 2012; Kabla, 2012; Zimmermann et al., 2016) (Fig. 5A, B). In addition to actomyosin dynamics, these models can incorporate intercellular forces, cell density, substrate properties, and contact inhibition of locomotion (CIL) – a fundamental feature of collective cell migration that promotes the formation of protrusions in a direction away from their contacts with the follower cells, thereby propelling the migration by leader cells (Fig. 5C) (Mayor and Etienne-Manneville, 2016). With an emerging understanding of mechanochemical coupling regulating the determination of leader and follower cells (Riahi et al., 2015), the models described above focusing on tissue dynamics can elucidate how different signaling aspects crosstalk with cell and tissue mechanics during collective cell migration.

In terms of its application to cancer, a form of collective cell migration where multicellular clusters of tumor cells can bud off the primary lesions and enter circulation, has been observed to be the predominant way of successful colonization (Aceto et al., 2014; Cheung et al., 2016). These clusters of circulating tumor cells (CTCs) retain their epithelial traits, at least partially, and act as primary harbingers of metastasis (Cheung and Ewald, 2016; Grigore et al., 2016; Jolly et al., 2015). Differential gene expression signatures of leader vs. follower cells in collective migration and invasion during metastasis has highlighted JAG1 as a key player (Cheung et al., 2016; Jolly et al., 2017), thereby reminiscent of the involvement of Notch signaling in regulating leader vs. follower phenotypes in multiple contexts of
collective cell migration (Blanco and Gerhardt, 2013; Boareto et al., 2015; Riahi et al., 2015).

This connection between Notch signaling and collective migration motivated a recently developed mathematical model that incorporated the coupling between the EMT circuit and the Notch signaling pathway based on existing experimental data (de Antonellis et al., 2011; Brabletz et al., 2011; Bu et al., 2013; Niessen et al., 2008; Sahlgren et al., 2008). This model predicted that Notch-Jagged signaling, but not Notch-Delta signaling, can enable both increased numbers and spatial proximity of hybrid E/M cells that, owing to their ability to both adhere and migrate, may lead to the formation of clusters of CTCs (Boareto et al., 2016). This prediction provides mechanistic insights into why JAG1 may be crucial for mediating clustered migration (Cheung et al., 2016), and is consistent with the evidence that JAG1 is related to drug resistance (Boareto et al., 2016; Guo et al., 2010), if we refer to the earlier claim that hybrid E/M cells are more likely to exhibit stemness. Yet, it remains to be rigorously and extensively tested experimentally whether knockdown of JAG1 can reduce the frequency of clustered migration and thereby curtail metastasis.

For a comprehensive characterization of collective cell migration in cancer, such signaling mechanism-based models need to be tied to previously described models of cell motility in multiple ways, for instance, by incorporating the effect of cellular stress on the activation of Notch signaling (Riahi et al., 2015); integrating how matrix stiffness can drive EMT through TWIST1-GP3B2 pathway (Wei et al., 2015); including how matrix density can alter the levels of membranous E-cadherin and affect the EMT status of cells (Kumar et al., 2014); and considering that ZEB1-mediated collagen deposition and stabilization (Peng et al., 2016) can increase matrix density. Developing such mechanochemical models can reveal how phenotypic transitions are coupled to the repertoire of mechanical signals that cancer cells experience and generate (Przybyla et al., 2016).

8. What other open questions in the regulation of EMT/MET can benefit from mathematical models?

Multiple open questions related to EMT/MET furnish exciting opportunities for cross-pollination of ideas among experimental and computational biologists, including (a) ‘How many intermediate states can cells attain en route to EMT and MET’; (b) ‘What is the
genomic, proteomic, and epigenetic signature of these states?; (c) 'How symmetric are the dynamics of EMT and MET, and do cells display hysteresis (i.e., cellular memory)?'; and (d) 'What is the relative stability and relative 'stemness' possessed by each of these states?'

As expected, mathematical models encompassing a larger number of EMT/MET regulatory players than considered in the initial models (Lu et al., 2013; Tian et al., 2013) have suggested multiple intermediate states (Hong et al., 2015; Huang et al., 2017; Steinway et al., 2015), but these predictions remain to be experimentally verified, thus providing impetus for many collaborative efforts.

Furthermore, E/M plasticity of cancer cells has also been linked to metabolic shifts (Dong et al., 2013; Kondaveeti et al., 2015; LeBleu et al., 2014) – another hallmark of cancer (Hanahan and Weinberg, 2011). Mathematical models that calculate metabolic fluxes by considering mass balance of various intracellular metabolites is a standard technique to analyze metabolic signatures (Markert and Vazquez, 2015). Such models are being increasingly implemented to quantify metabolic changes in tumor cells (Achreja et al., 2017). Constructing mathematical modeling frameworks that integrate these flux-balance models with models for the dynamics of signaling networks can help investigate the coupling of metabolic networks with signaling pathways that regulate E/M plasticity and stemness (Menendez and Alarcón, 2014; Peiris-pagès et al., 2016). These new frameworks can offer novel insights into the emergent consequences of bidirectional crosstalk among these networks driving these different hallmarks of cancer.

In addition to discerning this intracellular crosstalk, mathematical models can infer the dynamics of stromal cells as well as intercellular tumor–stroma signaling and act as in silico coculture systems. For instance, mechanism-based mathematical models can explain how macrophages can exhibit an intermediate polarization status between M1 and M2 (Italiani and Boraschi, 2014). Further, models capturing the crosstalk between differentially polarized macrophages and cancer cells (Yang et al., 2016) at an intracellular decision-making level as well as at a population level (i.e., multiscale models) can help visualize how cancer cells can engineer their microenvironment to their benefit and drive tumor progression, and hence propose strategies to restrict it.

9. Conclusion

As discussed above, an integrated theoretical-experimental approach has been instrumental in characterizing E/M plasticity and cellular traits associated with this plasticity. Concomitant with the renewed understanding that cancer can be viewed as an ecosystem unto itself (Yang et al., 2014), mathematical models capturing the interplay between tumor cells and multiple components of the tumor microenvironment can decode underlying organizing principles that manifest as myriad phenotypic complexities (Hanahan and Weinberg, 2011). Therefore, an iterative crosstalk between theory and experiment can help propel the hope that cancer biology and treatment will become a science with a conceptual structure and logical coherence that rivals that of chemistry or physics (Hanahan and Weinberg, 2000) into reality.

Acknowledgements

This work was supported by the National Science Foundation (NSF) Center for Theoretical Biological Physics (NSF PHY-1427654, NSF PHY-1605817, and NSF DMS-1361411. HL was also supported as a CPRIT (Cancer Prevention and Research Institute of Texas) Scholar in Cancer Research of the State of Texas at Rice University. SH was supported by the Rubenstein Family Foundation and the Canary Foundation. JAS acknowledges support from the Duke Cancer Institute, the Duke Genitourinary Oncology Laboratory, the Department of Orthopaedics, and the Triangle Center for Evolutionary Medicine (TriCEM). MKJ was supported by a training fellowship from the Gulf Coast Consortia on the Computational Cancer Biology Training Program (CPRIT Grant No. RP170593).

References


