The competing effects of the immune system and EMT on epithelial cancers

Daniel R. Bergman[1], Matthew Karikomi[1], Qing Nie[1,2,*] and Adam L. MacLean[3,*]

[1]Department of Mathematics, University of California, Irvine, Irvine, CA 92697, USA
[2]Department of Cell and Developmental Biology, University of California, Irvine, Irvine, CA 92697, USA
[3]Department of Biological Sciences, University of Southern California, Los Angeles, CA 90089, USA
*Correspondence: qnie@uci.edu (QN); macleana@usc.edu (ALM).

Abstract

Preceding and during cancer progression, the immune system is engaged in complex sets of interactions with epithelia that affect tumor incidence and prognosis. Inflammation also has significant effects on gastrointestinal and pancreatic tumors. Moreover, epithelial cell plasticity can significantly alter disease trajectories (for better or worse) via epithelial-to-mesenchymal transition (EMT). Several of the same pathways that regulate EMT are involved in tumor-immune interactions, yet little is known about the mechanisms and consequences of these regulatory processes. Here we introduce a multiscale evolutionary model to describe the interplay between tumor-immune-EMT interactions and their impact on epithelial tissues. Through *in silico* analysis of large patient cohorts, we found controllable regimes that maximize cancer-free survival. Delaying tumorigenesis depended crucially on properties of the mesenchymal phenotype: its growth rate and its immune-evasiveness. We performed analysis of pancreatic cancer data from The Cancer Genome Atlas to compare clinical measurements with model predictions. We found, in broad agreement with the model, that association with EMT worsened the survival probabilities of inflammation-associated cancer. These results offer novel means to delay disease onset by regulating properties of EMT, and demonstrate the importance of studying cancer-immune interactions in the light of EMT.

1   Introduction

Cancer and the immune system interact in myriad ways. Beyond direct tumor-immune cell interactions, the immune system modulates the tumor microenvironment (TME): responding to

local inflammatory signals, targeting mutated cells, and eradicating tumor cells to potentially restore homeostasis in a local context [1]. Recent breakthroughs in the field of immunotherapy are beginning to produce therapies that have had a significant impact on patient health and survival [2, 3].

Tumor cells interact with the immune system by initiating a sequence of events that results in the activation of the adaptive immune response. This is accomplished by the presentation of antigens recognized by innate immune cells that are transported to lymph nodes where T cells (and other components) can be activated [4]. The tumor also engages in cellular processes that indirectly shape the TME, for example by releasing transforming growth factor beta (TGF-β), thus shifting the TME towards a tumor-supportive environment by enhancing immunosuppression via activation of T regulatory cells (Tregs) [4].

The effects of the immune system on a tumor can be broadly summarized into two modules. The cytotoxic branch of the immune system, such as natural killer cells (NKs) and cytotoxic T cells (CTLs), seek out and lyse tumor cells. Upon carrying out their effector functions, these cytotoxic cells lose efficacy or deactivate entirely [5]. The regulatory branch of the immune system (Tregs and other factors), inhibits the effective functioning of the cytotoxic branch [6]. In addition, the immune system plays important regulatory roles in pre-cancerous tissue, where inflammation can increase the probability of tumorigenesis (and/or decrease the time to cancer), particularly in tumors originating in gastrointestinal and pancreatic tissues [7, 8]. Recent work raises some questions about the relationship between inflammation and cancer, suggesting that under certain conditions inflammation may not be oncogenic but rather onco-protective [9].

Epithelial-to-mesenchymal transition (EMT) describes a reversible process by which cells with an epithelial phenotype transition into cells with a mesenchymal phenotype. While epithelial cells are in part defined by their adhesion to neighbors, mesenchymal cells show much greater ranges of motility, and may display stem-like properties [10], although controversy regarding 'stemness' and EMT remains [11, 12]. Recent work has shown that – rather than being a binary process – at least two stable intermediate states exist on the EMT axis [13, 14]. Ongoing investigations into the plasticity and stability of EMT overlap with discussions elsewhere, e.g. of discrete vs. continuous processes during cell differentiation [15]. Intermediate states have emerged as a central mechanism by which cell fates (and the noise inherent within them) can be controlled [16–18].

While EMT has for some time been assumed to impact metastasis, more recently EMT-related phenotypes have been linked to other aspects of both tumor initiation and progression [10]. TGF-β is a master regulator of EMT [19], linking it to the signaling networks involved in tumor-mediated immune responses, since Tregs release TGF-β upon arriving at the tumor site [20]. Thus, even by considering only a single signaling factor, we find these three components (the tumor, the immune system, and EMT) to be linked. It therefore strikes us as a priority to develop models to understand how interactions between each of these three components affect cancer incidence and progression.

Two features of the mesenchymal phenotype are of particular relevance in the context of cancer and the immune system: i) mesenchymal cells are less susceptible to immune clearance [20]; and ii) mesenchymal cells proliferate on average at a slower rate than epithelial cells. The immune evasiveness of mesenchymal cells is well-established: as a cell is targeted by cytotoxic immune cells for clearance, a physical connection between the two cells must be established before the lysing event. This immunological synapse is dependent upon the surface proteins,

such as the T-cell receptors of the major histocompatibility complex, of the target cell, and in the case of mesenchymal cells there is preliminary evidence that these surface markers are down-regulated in such a way that the immune synapse is more difficult to form [20]. Below we refer to this phenotype as mesenchymal immune evasion (MIE).

The other facet of mesenchymal cells is their slower proliferation rate, and this is tied to the observation that mesenchymal cells have been speculate to exhibit a more "stemness" phenotype. While many cell behaviors are associated with stemness, here we target the specific reduction in proliferation rate [21] (which, under certain conditions can be called quiescence, though we avoid this term as it can be controversial, and is not a necessary assumption to the model developed below). Cancer cells are of course characterized by some degree of unregulated proliferation, thus it is appealing to explore the effects of differential proliferation rates within the cancer cell population. In our model, we refer to the reduced proliferative capacity associated with a mesenchymal cell phenotype as mesenchymal growth arrest (MGA).

Mathematical oncology, that is, mathematical models of cancer incidence, progression, and treatment, has become a well-developed field; many models have offered insight into the cellular interactions underlying cancer and its interplay with the immune system, including older [22, 23] and more recent works [24–29]. These studies have increased our understanding of how tumors grow in the presence of various immune components, and how treatment regimes can be designed to maximize the efficacy of cytotoxicity while minimizing risks to the patient. However, to our knowledge no models have addressed how the effects of EMT alter interactions between the immune system and cancer, and the subsequent implications for treatment.

Here we develop a model with the goal of studying interactions between the tumor, the immune system and EMT. We seek to describe a set of crucial molecular and cellular interactions in epithelial tissue cells, including effects due to DNA damage and mutation, to investigate the probability that cancer will occur and, if so, when. A recent model of cancer-immune interactions [9] described the effects of the TME on the risk of cancer, and we build on the core cell cycle component of this model, adding significant new interactions to the immune component of the model (which was previously modeled by a single interaction), as well as adding the effects of EMT. We seek to understand whether this more complex immune module will change our understanding of inflammatory effects on the tumor, and how the epithelial-mesenchymal axis influences these.

In the next section we develop the model and give explanation of the intuition behind each component. We then go on to analyze its behavior. Global sensitivity analysis identifies parameters that are crucial for tumorigenesis. We then study these in more depth, focusing on the competing effects of EMT and of the immune system on tumorigenesis: we find that EMT intricately regulates tumorigenesis and that under certain regimes a careful balance of EMT- and immune-driven processes can significantly prolong cancer-free survival. To test these predictions, we analyze data from the Cancer Genome Atlas (TCGA), and find evidence for the synergistic effects predicted by the model for patients with pancreatic cancer.

2   Methods

We develop an agent-based model to describe the relationships between cancer, the immune system, and EMT, building on cell-cycle and tissue-cell components of [9]. The agents in the

model are the tissue cells with tumorigenic potential. These can either be mutation-free or, resulting from DNA damage during the cell cycle, can have any combination of three possible pathway mutations (shown in Fig. 1A). If the pathway is mutated, the component is colored red and is marked with 'X'. A single mutation is sufficient for the cell to be considered as a mutant cell.

We model immune cells as continuous variables, appropriate since at the level of immune cells it is not necessary to keep track of single cell effects. The cytokine TGF-β is continuous within the TME, i.e. the system is well-mixed. Tissue cells can take on either epithelial or mesenchymal phenotypes in a plastic manner: these phenotypes depend on both the TME and cell-intrinsic factors. While the score is continuous, a threshold determines if a given cell is labeled as epithelial or mesenchymal. In Fig. 1A, the EMT labeling is given by the color of the cell. Even though each cell does have an EMT score on a spectrum of values, only the labeling influences its interactions in the model.

2.1 Tissue Cells

The tissue cells at all times have associated with them two important quantities: which combination of mutations they carry and an EMT score. The model has three pathways that have the potential to mutate over the course of a simulation. There is a proliferation pathway that when mutated increases the probability of the cell proliferating. There is an apoptosis pathway that when mutated decreases the probability of the cell undergoing apoptosis. There is an immune evasion pathway that when mutated decreases the probability that an immune cell can effectively clear the mutated cell. The EMT score will be exposited in Section 2.3.

2.2 The Immune System

The immune system is modeled as having three active immune cell types: NKs, CTLs, and Tregs. The NKs and CTLs act on the system by recognizing mutated cells and clearing them. Upon successful clearance, they themselves are then deactivated and removed from the immune population. Tregs act on the system by suppressing the effector functions of NKs and CTLs, making it less likely they can clear mutated cells. In addition to this, Tregs release TGF-β which further shapes the TME by pushing tissues cells more towards a mesenchymal phenotype.

2.3 EMT

Each tissue cell has an EMT score between 0 and 1. When the score is above a fixed threshold value, the cell is labeled as mesenchymal. Otherwise, it is below the threshold and labeled as epithelial. For the purposes of the model, a cell in an epithelial state is considered in the base state, and one in a mesenchymal state will have some of its parameters updated. In particular, mesenchymal cells benefit from increased immune evasion but suffer from decreased proliferation. Note, that because the immune system can only clear mutated cells, the increased immune evasion only benefits mutated mesenchymal cells.

2.4 Model Simulation

2.4.1 Initial conditions

Every model run starts with a fixed amount of cells, N_0. However, different parameter values will lead to different steady states for the population of tissue cells. In light of this, a warmup period of 1000 cell cycles is used in which time no mutations are allowed to happen. Thus, the only immune cell present is NKs. After the warmup cycles are complete, a mutagenic event is simulated in which all the tissue cells experience a random increase in their probability to mutate. From then on, a cell which proliferates also either mutates or increases its probability of mutating later on.

2.4.2 Tissue cell fate

During each cell cycle, every cell randomly is assigned a cell fate from the following options:

- proliferation
- apoptosis
- immune clearance (by NKs or CTLs)
- rest in G_0

For each cell, a weight is chosen for each option and these are normalized to probabilities which then are used to randomly determine what each cell does during the cell cycle.

Proliferation   There are four factors that contribute to the weight of a cell to proliferate. The first is a base proliferation rate that all cells have, p. Second, if the cell has a mutation in the proliferation pathway ($\delta_P = 1$), then the weight for proliferation is proportionally increased by Δ_P. Third, if the cell is mesenchymal ($\zeta = 1$), then the weight for proliferation is proportionally decreased by Δ_{MGA}, which stands for mesenchymal growth arrest. Fourth, there is a negative feedback of the cells on their own proliferation which is quantified by a Hill factor as a function of the tissue cell population, N_C, with EC50 term K_0. In total, the weight for proliferation is given by

$$\rho_P = p(1 + \delta_P\Delta_P)(1 - \zeta\Delta_{\text{MGA}})\frac{K_0}{K_0 + N_C} \tag{2.1}$$

Apoptosis   There are two factors that contribute to a cell's weight for undergoing apoptosis. There is a basal apoptosis rate that all cells experience, d_C for death. Second, if the cell has a mutation in the apoptosis pathway ($\delta_A = 1$), then the weight for undergoing apoptosis is proportionally decreased by Δ_A. In total, the weight for apoptosis is given by

$$\rho_A = d_C(1 - \delta_A\Delta_A) \tag{2.2}$$

Immune Clearance   For both NK clearance and CTL clearance, the weights are built with the same factors but have different parameter values for NK and CTLs. First of all, the cell needs to be mutated ($\delta_{\text{MUT}} = 1$). Second, there is a Hill factor that captures the probability of an

immune cell finding and interacting with the given tissue cell with EC50 term K_1. Third, NKs and CTLs have their own efficacy parameters, E_{NK} and E_{CTL}, which can be understood as the rate of immune clearance given an immune cell has found the mutated cell. Fourth, there is a decreasing Hill factor based on the number of Treg cells present with EC50 term K_2. Finally, there are two factors that proportionally decrease the weight of immune clearance depending on if the cell has an immune evasion mutation ($\delta_{\text{IE}} = 1$) or if it is mesenchymal ($\zeta = 1$) with respective decreases Δ_{IE} and Δ_{MIE}. In total, the weight of NK clearance is given by

$$\rho_{\text{NK}} = \delta_{\text{MUT}} \frac{N_{\text{NK}}}{N_C/K_1 + N_{\text{NK}}} \frac{E_{\text{NK}}}{1 + N_{\text{Treg}}/K_2} (1 - \delta_{\text{IE}}\Delta_{\text{IE}})(1 - \zeta\Delta_{\text{MIE}}) \tag{2.3}$$

A similar formula holds for CTLs with only the number of CTLs and their efficacy being different from the above equation.

Rest in G_0 The weight associated with quiescence is taken as 1 except in the case of mesenchymal cells. For mesenchymal cells, the weight lost to proliferation is added to the weight of quiescence. Hence, the weight of quiescence is given by

$$\rho_R = 1 + \zeta p(1 + \delta_P \Delta_P)\Delta_{\text{MGA}} \frac{K_0}{K_0 + N_C} \tag{2.4}$$

The reason for adding that term is due to the understanding that overall mesenchymal cells proliferate less as individual cells rest longer in the G_0 phase.

2.4.3 Completing the Cell Cycle

After the cell fates are determined and the results reflected in the system, there are a few things that happen before the system moves on to a new cell cycle. First, the NK and CTL populations are reduced by the number of mutated cells they cleared. This represents the fact that individual immune cells lose efficacy as they carry out their effector functions. Second, all proliferating cells have a cell-specific probability of undergoing a pathway mutation. If they do, one is randomly chosen among the three pathways and the pathway in that cell mutates. If the cell does not undergo a mutation, then its probability of mutation during subsequent cell cycles increases. Finally, the EMT values for each cell is updated. This depends on the cells current EMT score and how much TGF-β is currently in the system. Each cell receives an amount of TGF-β given by

$$\tau_i = \frac{\tau_{\text{max}}}{N_C} \frac{1}{1 + K_3/\tau} + X_i, \quad X_i \sim N(0, \sigma^2) \tag{2.5}$$

The summand on the left deterministically divides up the total TGF-β among all N cells evenly and a Hill function on this quantity determines how much TGF-β each cell thus receives. Then, some randomness is introduced by the independent and identically distributed variables X_i where σ is a parameter in the model. If this quantity is large enough, relative to the EMT score of the cell, then the cell will undergo EMT and its EMT score will increase. If not, then the cell undergoes the reverse process, MET, and its EMT score will decrease. Each cell then is relabeled as either epithelial or mesenchymal depending on its new EMT score and whether it is below or above the mesenchymal threshold. Thus, there are two main factors that determine

if a cell will end a cell cycle as mesenchymal: concentration of TGF-β in the system and the current EMT score of the cell.

Next, the amount of TGF-β for the next cell cycle is determined by the number of mutated cells, N_{MUT}, and the number of Treg cells, N_{Treg}, each one producing a fixed amount of TGF-β. It is given by

$$\tau = \tau_{\text{MUT}} N_{\text{MUT}} + \tau_{\text{Treg}} N_{\text{Treg}} \tag{2.6}$$

Finally, the immune populations are updated. For the NKs, they obey the following differential equation:

$$N'_{\text{NK}} = \sigma_{\text{NK}} - d_{\text{NK}} N_{\text{NK}} \tag{2.7}$$

which is discretized to

$$N_{\text{NK}}(k+1) = \left(N_{\text{NK}}(k) - \frac{\sigma_{\text{NK}}}{d_{\text{NK}}} \right) \exp(-d_{\text{NK}}\Delta t) + \frac{\sigma_{\text{NK}}}{d_{\text{NK}}} \tag{2.8}$$

For CTLs and Tregs, they rely on mutant cells being cleared before they can be activated. Let $N^*_{\text{MUT}}(k)$ represent the number of mutant cells cleared by the immune system during cell cycle k. In addition, Treg recruitment is upregulated by TGF-β, which will be incorporated via a Hill function with EC50 term K_4. We choose the following differential equations to govern the CTL and Treg populations:

$$
\begin{aligned}
N'_{\text{CTL}} &= \sigma_{\text{CTL}} N^*_{\text{MUT}} - d_{\text{CTL}} N_{\text{CTL}} \\
\text{Treg}' &= \sigma_{\text{Treg}} N^*_{\text{MUT}} \frac{\tau}{1 + \tau/K_4} - d_{\text{Treg}} N_{\text{Treg}}
\end{aligned}
\tag{2.9}
$$

Discretized, these are:

$$
\begin{aligned}
N_{\text{CTL}}(k+1) &= (N_{\text{CTL}}(k) - \sigma_{\text{CTL}} N^*_{\text{MUT}}(k)/d_{\text{CTL}}) \exp(-d_{\text{CTL}}\Delta t) + \sigma_{\text{CTL}} N^*_{\text{MUT}}(k)/d_{\text{CTL}} \\
N_{\text{Treg}}(k+1) &= \left(N_{\text{Treg}}(k) - \frac{\sigma_{\text{Treg}} N^*_{\text{MUT}}(k)}{d_{\text{Treg}}} \frac{\tau(k)}{1 + \tau(k)/K_4} \right) \exp(-d_{\text{Treg}}\Delta t) \\
&\quad + \frac{\sigma_{\text{Treg}} N^*_{\text{MUT}}(k)}{d_{\text{Treg}}} \frac{\tau(k)}{1 + \tau(k)/K_4}
\end{aligned}
\tag{2.10}
$$

Name	Description
p	proliferation rate of tissue cells
d_C	death rate of tissue cells
Δ_{MIE}	mesenchymal immune evasion
Δ_{MGA}	mesenchymal growth arrest
Δ_A	mutant cells decreased apoptosis
Δ_{IE}	mutant cells increased immune evasion
Δ_P	mutant cells increased proliferation
K_0	EC50 term for negative feedback of tissue cells on own proliferation
K_1	EC50 term for probability of NK cell finding mutant cell
K_2	EC50 term for Treg inhibition of cytotoxic functions
K_3	EC50 term for how much TGF-β each cell has
E_{NK}	rate of NKs clearing mutants
E_{CTL}	rate of CTLs clearing mutants
σ_{NK}	NK source rate
σ_{CTL}	CTL source rate per cleared mutant cell
σ_{Treg}	Treg source rate per cleared mutant cell
k_{EMT}	EMT/MET rate
σ	standard deviation of noise in TGF-β each cell receives
τ_{max}	max amount of TGF-β any cell can receive
τ_{MUT}	rate of TGF-β production by mutant cells
τ_{Treg}	rate of TGF-β production by Treg

Table 1: *The model parameter names and descriptions. Note that many of these values are affected by the inflammation state of the system.*

2.4.4 Mutation and tumorigenesis

At the end of each cell cycle, the proportion of tissue cells that are mutated is calculated, and if it is above a certain threshold, a tumorigenic event is recorded and the patient is determined to have cancer. The Time to Cancer is given as the time from the start of the simulation, not including warmup, until tumorigenesis. Simulations run until either a tumorigenic event is recorded or until the maximum number of cycles has been reached.

2.5 Parameter Sensitivity Analysis

To study parameter sensitivity, we implemented the Morris method, a one-step-at-a-time global sensitivity analysis algorithm. The algorithm runs the model numerous times by picking several base points and then varying one parameter at a time from each base point and tracking the difference [30, 31]. Each model run involved 1000 patients and the value returned was the area under the survival curve. We ran the algorithm by sampling at 30 points in parameter space and then changing each parameter one time at each point. In the implementation by Sohier, they recommend using at least 10 points [31].

To determine which points in parameter space to sample at, we had to create distributions on all parameter values. To do this, we derived many parameters from existing literature and the rest we made reasoned predictions. For example, many previous studies have established the parameters of immune population dynamics such as recruitment and death rates [32]. Because our model only works with on the order of 10^2 tissue cells whereas an actual tumor would have on the order of 10^9 cells [32], we scaled the parameters so the immune populations would also be proportionately reduced. That is, for instance, how we ended up with NK cells having a steady state population of 10 cells in the model.

Once we had these parameter values, we created Gaussian distributions centered at these values with large variances and truncated at 0. Some of the parameters were also truncated to the right because they represented a proportion, such as MIE and MGA, or because we simply wanted to keep the values within a reasonable range, such as the inflammation low and high duration parameters. The variances were chosen as twice the mean to get a large spread of values.

The algorithm computes the average of the absolute change of the output, which in our model is the area under the survival curve. This output out the algorithm, μ^*, is the sensitivity plotted in Fig. 2.

2.6   Analysis of patient survival data from TCGA database

We obtain primary tumor bulk mRNA sequencing and censored survival data for all individuals monitored in the relevant project from the Cancer Genome Atlas (TCGA) from the Genomic Data Commons portal in R, using the package TCGABiolinks. Given $n >= 1$ gene set keywords (e.g. "inflammatory" and "emt"), the symbolic names of all msigdb gene sets are searched for matches to these keywords, and matched gene sets are grouped by keyword. For each element in the product S of all keyword groups, the following analysis is performed.

1. The first principle component of the expression across the gene set is obtained for each element of the n-tuple of gene sets.
2. Two clusters of n-dimensional patient vectors are obtained by k-means clustering.
3. The patients are separated by their cluster identity and Kaplan-Meier curves are fit to their corresponding survival data.
4. Under the null hypothesis that there is no difference between the survival of the two groups, a log-rank test is performed.
5. The log-rank test p-value for each element in S is placed in an ordered list, the lowest of which defines the element of S whose composite gene sets are most predictive of patient survival in the given TCGA project or tumor type.

Below we analyze data gathered for pancreatic cancer, denoted by "PAAD" in TCGA.

2.7   A multiscale agent-based model of EMT-immune-tissue cell interactions to study tumorigenesis

We begin by investigating general features of the model to establish baseline conditions and to assess the impact of various model components on the key measured outcomes: the probability of cancer, and the Time to Cancer. Within the cell cycle, cell fate is determined via a set of
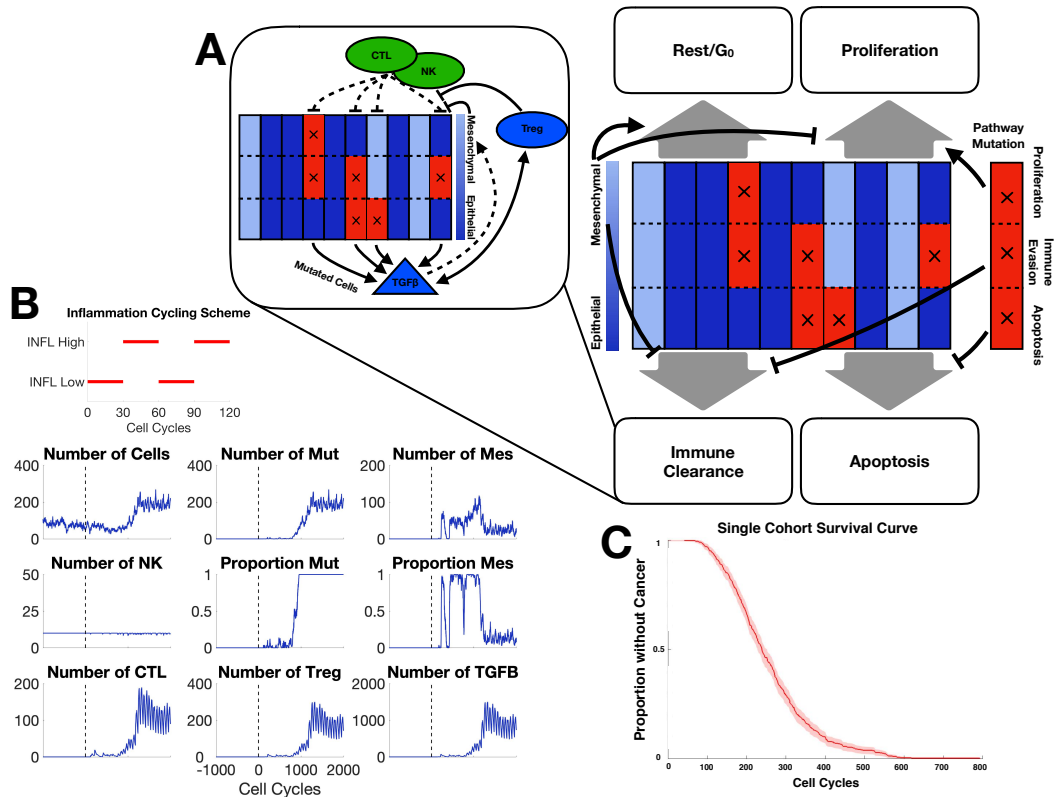
Figure 1: *A. Schematic depiction of agent-based model components; blue/red denote tissue cells with/without mutations. Black arrows depict regulation of the cell fate in each cell cycle. Inset depicts major interactions between the immune system and tissue cells. B. A representative simulation of one patient. The model parameter values used can be found in Supplementary Table 2. The inflammation cycling scheme is represented above the patient dynamics. The vertical dashed line denotes the end of the warmup period. Mut: mutated cells; Mes: mesenchymal cells. C. Survival curve for one cohort of patients with the parameter values given in Supplementary Table 2.*

rules that are influenced by EMT and immune interactions (Fig. 1A). For example, if a cell undergoes EMT, the probability that it will proliferate is reduced; if a cell gains a mutation in the apoptosis pathway, the probability that it undergoes apoptosis is greatly reduced. The different means by which the immune system acts on tissue cells are also shown (1A Inset). NK cells and CTLs attempt to clear mutated tissue cells, and deactivate upon successfully carrying out their cytotoxic function. Tregs inhibit cytotoxic activity. Tregs also release TGF-β which promotes the recruitment of Tregs, as well as increasing the rate of EMT.

The inflammation cycling scheme for a typical *in silico* patient consists of alternating high and low regimes (Fig. 1B); the inflammation schemes modeled will be discussed in detail in the next section. For this patient, after the warmup period, cell mutations are observed at a rate low enough that they are cleared by cytotoxic cells before being able to establish a tumor. However, after approximately 1000 cell cycles (750 days), one mutated cell exhibits clonal growth, and a tumor is established. At this point, large numbers of cells from the adaptive immune system (CTLs and Tregs) are being heavily recruited, and a peak in TGF-β expression is observed. The number of mesenchymal cells increases initially with increasing TGF-β, but decreases after the clone takes over the tissue. This shift back towards an epithelial tissue phenotype occurs even though the expression of TGF-β remains high. After 841 cell cycles, the proportion of mutated cells reaches 50%: this is modeled as the threshold defining tumorigenesis. Thus this patient has a Time to Cancer of 841 cell cycles, or 631 days. After cancer incidence, the number of mutated cells continues to grow rapidly and soon makes up 100% of the cell population. A peak in mesenchymal phenotypes is observed following cancer incidence, before most transition back to an epithelial state.

Considering the immune system dynamics, we see that the NK population is approximately constant, while the adaptive populations (CTL and Treg cells) grow quickly and dwarf NK cells in number following the accumulation of oncogenic mutations. The adaptive immune populations also appear to exhibit oscillatory behavior, however note that this is not due to intrinsic dynamics but rather due to the inflammation scheme that the patient is undergoing: alternating between 30 cell cycles of high inflammation and 60 cell cycles of low inflammation. Within each of these periods, adaptive immune populations increase or decrease rapidly in accordance with the inflammation state.

The *in silico* patient described here is given for the purpose of illustrating features of the model. Given the multiple sources of noise in the model, in order to quantify patient dynamics and cancer-free survival rates, below we will simulate large cohorts of patients. In Fig. 1C we simulate the survival curve for a cohort of 500 patients: we see that all the patients survive for approximately 100 cell cycles (75 days). Subsequently, in the approximate range of $T = [100, 300]$, the cancer onset rate is roughly constant, and after 600 cell cycles, no patients remain cancer free.

2.8 Identification of key regulatory parameters via Morris global sensitivity analysis

To assess the sensitivity of the model parameters and thus identify those that are most important in determining cancer-free survival times, we applied Morris one-step-at-a-time (OAT) sensitivity analysis (see Methods) to the model. The results of Morris OAT on the 31 model parameters are shown in Figure 2: a subset of parameters demonstrate much higher levels of sensitivity
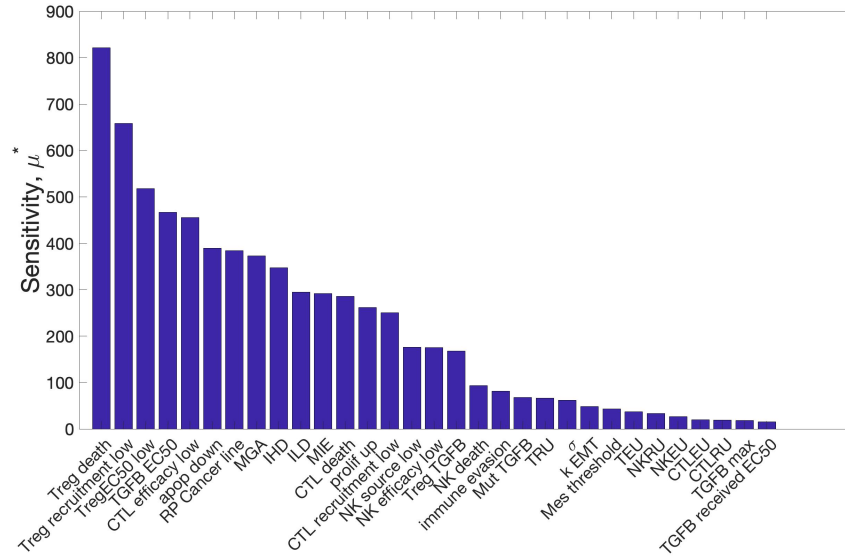
Figure 2: *Sensitivity analysis of model parameters, determined via the global Morris one-step-at-a-time method. μ^* denotes the average absolute change in Time to Cancer when the parameter is varied.*

than others. The two most influential according to this analysis are the death rate and the recruitment rate of Treg cells, this is most likely due to the dual roles Treg cells play in both suppressing the cytotoxic effects of other immune cells and secreting TGF-β, which drives EMT. This ties Treg cells to all three components of the model. Since we seek to separate the effects of different model components, we do not choose the parameters influencing Treg cells for detailed analysis below.

Many parameters in the model change depending on the inflammation state. For the sake of a naming convention, any parameter which includes "low" in the name represents the parameter value when inflammation is low. For when the inflammation is high, these "low" parameters get scaled by parameters which include "up" in the name. The ILD and IHD parameters determine the duration of low and high inflammation, respectively. A cell with mutated pathways will be affected by the Δ_A, Δ_P, and Δ_{IE} parameters, depending on whether the apoptosis, proliferation, or immune evasion pathways are mutated. The mesenchymal cells are influenced by the MGA and MIE parameters. Finally, σ, k_{EMT}, τ_{max}, and K_3 all control how cells transition between the epithelial and mesenchymal states.

Since one goal of our analysis is to assess the specific effects of EMT on immune-cancer dynamics, the parameters MIE and MGA are of particular interest. In addition, inflammation parameters dictating the cycling scheme are of interest because they are both highly influential on Time to Cancer and capable of being targeted by therapeutic treatments. In terms of Treg cells, their secretion of TGF-β is highly significant and we will also study it further.
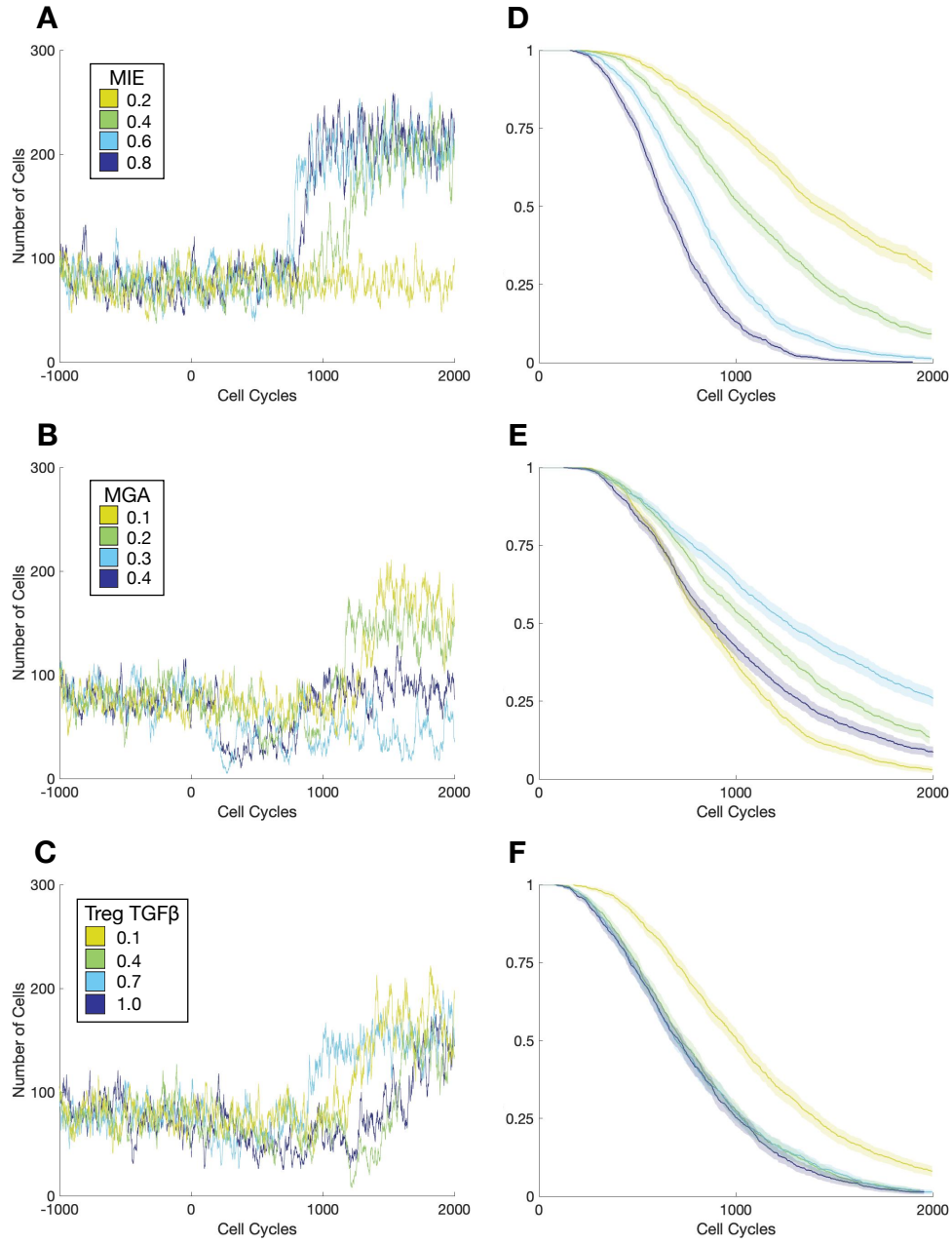
Figure 3: *Effects of mesencymal cell properites on the time to cancer. Trajectories of one patient per cohort from warmup period ([-1000,0]) to 2000 cell cycles, for Δ_{MIE} (A); Δ_{MGA} (B); and τ_{Treg} (C). D. Survival curves corresponding to changes in MIE (A) for a patient cohort of size 1000; shaded region represents the 95% confidence interval for the evaluated function. E. Survival curves for patient cohort corresponding to changes in MGA (B). F. Survival curves for patient cohort corresponding to changes in Treg TGF-β (C).*

2.9 Mesenchymal phenotypic properties dramatically change cancer-free survival times

When a cell transitions from an epithelial to a mesenchymal state, two phenotypic cell characteristics change: mesenchymal immune evasion (MIE) and mesenchymal growth arrest (MGA). Both parameters are defined as proportions, and lie in $[0, 1]$, where higher values indicate more mesenchymal-like properties. The parameter Δ_{MIE} is the proportional reduction in the probability that a mutated mesenchymal cell will be subject to immune clearance in a given cycle. The parameter Δ_{MGA} is the proportional reduction in the probability that a mesenchymal cell will proliferate in a given cell cycle, thus is equivalent (strictly in terms of its impact on the cell cycle) to an increased proportion of time spent resting in the G_0 phase. The cytokine TGF-β is also involved in the EMT process (by increasing the probability of EMT). All three of these parameters were found to have large effects on the model by the Morris OAT analysis presented in Section 2.8.

As MIE is increased, the Time to Cancer decreases (Fig. 3A) under all sets of parameters studied: as this subpopulation of potentially mutated cells becomes more resistant to immune clearance, the tumor as a whole grows more resilient and thus will grow faster. These results are summarized by Fig. 3A.

Second, as MGA increases, Time to Cancer increases, i.e. lower proliferation rates for mesenchymal cells slow down cancer growth (Fig 3B). This is not immediately intuitive, since the decreased proliferation rate affects both mutated and non-mutated tissue cells.

Third, TGF-β can be varied in two ways: the production by mesenchymal cells and the production by Treg cells. In Fig 3C, the results of varying Treg TGF-β production are shown, indicating that an increased Treg TGF-β production leads to a shorter Time to Cancer. The two main ways in which TGF-β influences the system is in recruitment of Treg cells and in pushing tissue cells to a mesenchymal phenotype. Treg cells are modeled as tumor-protective and thus increasing their number will naturally decrease Time to Cancer. Mesenchymal cells are more likely to evade the immune system, so pushing the system towards an overall more mesenchymal phenotype will better protect the cancer and decrease the Time to Cancer.

2.10 A key EMT regime maximizes cancer-free survival time under chronic inflammation

We explored the effects of varying the inflammation state of the patient on cancer-free survival, to investigate competing interactions within the TME and their effect on EMT. Patient cohorts were simulated under different inflammation regimes: permanently low inflammation; permanently high inflammation; or variable inflammation. For patients drawn from cohorts in a permanently high inflammatory state, the relationship between mesenchymal parameters and the Time to Cancer is monotonic, i.e. increasing either MIE (Fig. 4A-B) or MGA (Fig. 4C-D) decreases the Time to Cancer. However, under regimes with either temporary or permanent periods of low inflammation, different relationships emerge: a local maximum for the Time to Cancer is found with respect to MGA ($\Delta_{\text{MGA}} = 0.3$). This is seen both for intermittent or low inflammation (Fig. 4D).

These striking differences in the mean Time to Cancer – extended by up to one year by fine-grained optimization of mesenchymal growth arrest – have clear therapeutic implications.
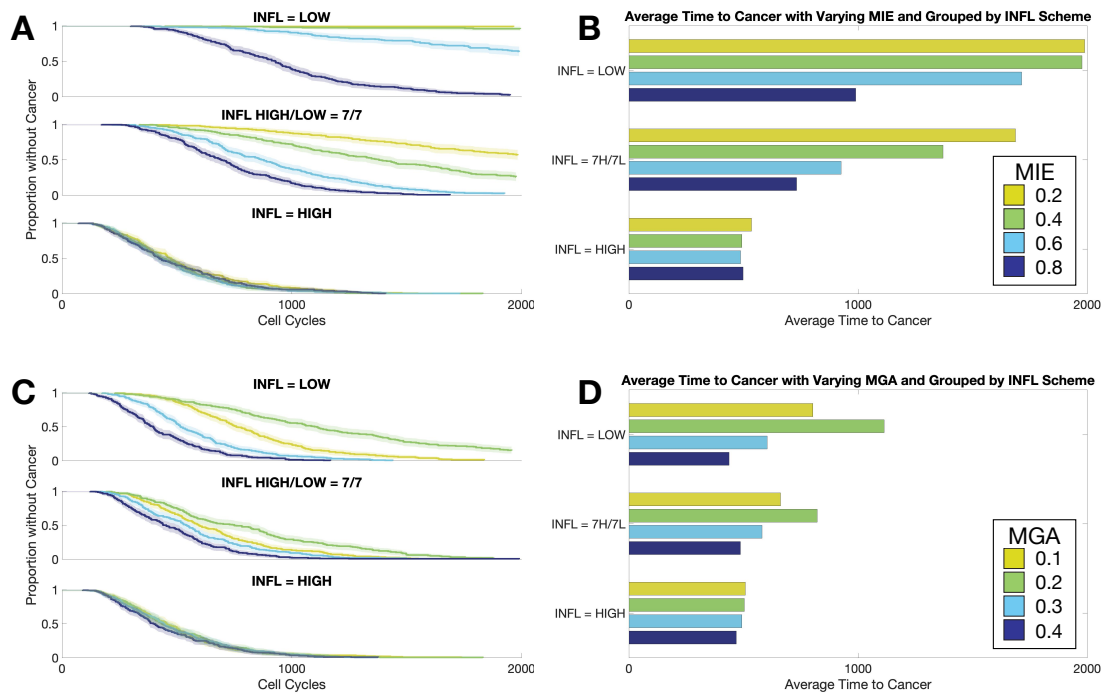
Figure 4: *Effects of inflammation on the time to cancer under different cycling schemes. A-B. As MIE varies, survival curves (each of 200 patients) and corresponding bar plots to summarize the mean Time to Cancer for each cohort are shown. C-D. As MGA varies, survival curves and corresponding bar plots to summarize the mean time to cancer for each cohort are shown.*
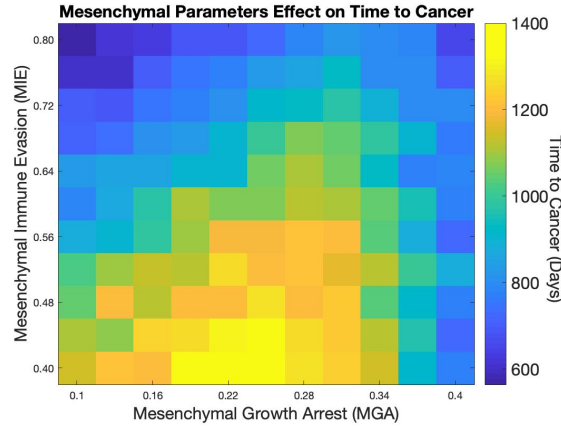
Figure 5: *Summary of the contrasting effects of MIE and MGA on the Time to Cancer.*

Our model predicts that reducing the inflammatory state of the TME in three weeks out of every nine would be provide means to prolong cancer-free survival in combination with mesenchymal cell therapies to make the control of mesenchymal proliferation an effective anti-cancer strategy.

In contrast, when MIE is varied for different inflammation cycling schemes, regardless of the cycling scheme, increasing the MIE leads to decreases in the Time to Cancer (worse survival probabilities), although in the case where the inflammatory state is permanently high, MIE has little effect on the Time to Cancer. Thus under any inflammation regime with periods of low inflammation, any reductions in the rate of mesenchymal immune evasion will lead to improved patient outcomes.

2.11 Analysis of TCGA data in comparison with model predictions suggests that mesenchymal phenotypes reduce cancer-free survival probability

To further assess the effects on tumorigenesis of the mesenchymal properties of immune evasion and growth arrest, we study the joint density plot for each of these parameters against the Time to Cancer (Fig. 5). We found that over the full range of values of Δ_{MGA} considered, increasing Δ_{MIE} decreases the Time to Cancer. However, for any given value of mesenchymal immune evasion, there is a value of mesenchymal growth arrest that maximizes the Time to Cancer. Moreover, this optimal value increases with increasing MIE. Together these results show that while reducing the ability of a mutated cell to evade the immune system by any amount will improve the probability of cancer-free survival, an optimal value of MGA will prolong survival.

To compare these model predictions with experimental studies, we analyzed data from The Cancer Genome Atlas (TCGA) database to study the effects of immune and EMT interactions on prognosis of cancers for which inflammation is known to play an important role, such as colonic or pancreatic cancers [7, 8]. Using gene ontologies as a measure of the of effects of inflammatory or EMT processes on cancer-free survival, we investigate pancreatic cancer. We note that the comparison between simulation and data here is indirect since the model studies processes occurring before incidence, while the data address subsequent events, but analogous processes are at play during both the pre-cancerous clonal dynamics addressed by the model, and the progression of disease that follows cancer incidence.
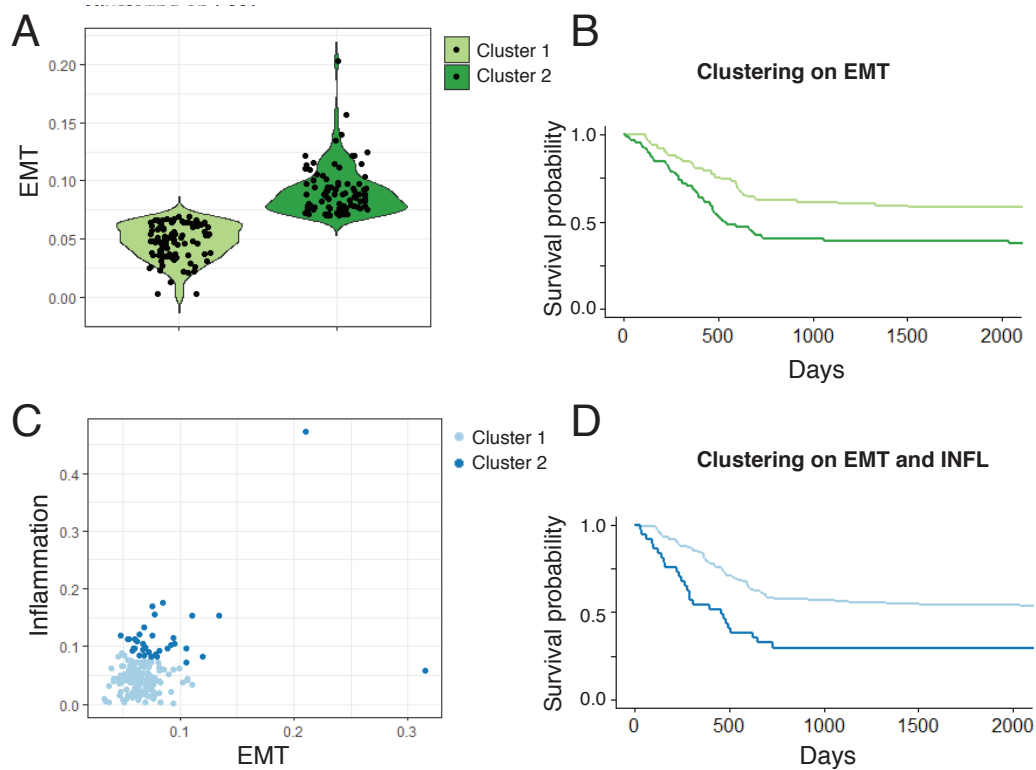
Figure 6: *A. K-means clustering of pancreatic cancers using gene ontology terms indicative of an EMT signature ($k = 2$). B. Survival plots corresponding to the clustering on EMT. C. K-means clustering of pancreatic cancers using gene ontology terms indicative of EMT and Inflammation signatures ($k = 2$). D. Survival plots corresponding to the clustering on EMT and inflammation.*

For a cohort of pancreatic cancer patients, we cluster the patients via k-means ($n = 2$) either on gene ontologies relating only to EMT (Fig. 6A) or on gene ontologies relating to EMT and inflammation (Fig. 6C). We also plot the corresponding survival curves for each of the clusters obtained (Fig. 6B, D). We see that in both cases, survival is affected by the gene ontology signature, but the presence of both EMT and inflammatory signatures has a greater impact on survival than the effects of EMT alone. This suggests that when studying interactions between cancer and the immune system, is it important to consider EMT; the effects of which may impact the dynamics and should not be overlooked.

3   Discussion

Despite the intense interest in cancer and the immune system, and in the effects of EMT on cancer, there has not previously, to the best of our knowledge, been a model developed that combines these components: cancer, the immune system, and EMT. We saw this as an even more pressing need given the shared factors that influence all of these, such as TGF-β. We used an individual cell-based model framework to describe the multiscale processes that can lead to cancer: DNA damage occurs during the cell cycle and this can lead to mutations in

pathways that affect the fitness of the cell, which in turn affects the cell population dynamics, which are also influenced by the intrinsic state of the cell (EMT), and extrinsic immune factors (inflammation, cytotoxic cells, etc).

We found that this model recapitulated cancer-free survival dynamics, and via global parameter sensitivity analysis, we identified parameters exerting key control over model behavior. Focusing on these led us to identify that: increasing mesenchymal immune evasion, decreasing mesenchymal growth arrest, or increasing Treg TGF-β production all lead to shorter Times to Cancer. However varying the level of inflammation leads to paradoxical effects: under regimes with periods of low inflammation, an optimal level of mesenchymal growth arrest can improve outcomes and maximize the Time to Cancer.

To test these predictions, we performed unsupervised analysis of pancreatic cancer data from The Cancer Genome Atlas, and looked at survival across two groups: we found that the presence of an EMT signature increased differences in survival between groups in combination with inflammation. We chose to summarize *in silico* patient simulations with a single parameter: the Time to Cancer, which we found captured the essential characteristics of the model. There are, of course, many tissue trajectories that all result in cancer, and it is our opinion that analysis of the transient cell dynamics in pre-cancerous tissues is a pressing need to shed insight into cellular biomarkers of cancer.

These results represent promising steps in understanding the competing roles of the immune system and EMT during (pre-) development of epithelial cancers, yet much remains to be done. Further development of the inflammation module of this model is important given the large and sometimes paradoxical roles that the inflammatory state exerts on the epithelia and cancer-free survival (Figs. 2 and 4B, D). Currently, inflammation is modeled as independently cycling between high and low schemes, however many of the agents considered in the model actively contribute to the inflammatory state, thus it would be interesting to consider connecting these components (for example by letting the level of inflammation be conditional on the number of mutant cells in the tissue). Another layer of complexity is revealed by the natural anti-inflammatory role of Tregs. One consequence of the current model is that decreasing the number of Tregs increases the cancer-free survival. Clearly, there exists a trade-off to be accounted for, and adding to the model the main effector function of Tregs could remedy this and add depth to our understanding of the various roles that Tregs play in and around the tumor.

The role that TGF-β plays throughout the tumor and its microenvironment also requires further research. We found that, below a threshold, reduction of TGF-β increases the Time to Cancer (Fig. 3E), thus reducing expression levels of TGF-β in the tumor microenvironment benefits survival. However, TGF-β is a master regulator implicated in numerous cellular signaling processes, and changing the concentrations of TGF-β even in a local tumor microenvironment could have large off-target effects. Thus it might be more fruitful to target signaling factors downstream of TGF-β that still have the ability to modulate epithelial cell dynamics: this is an area of active and future investigation, we are currently developing a larger TGF-β signaling pathway module with appropriate crosstalks to epithelial/mesenchymal/immune cell functions to be incorporated into the model.

A further goal for future work is to explore (and exploit) the heterogeneity of tumor evolution in greater depth: this heterogeneity aids the evasion of the tumor from immune effects. Studying the consequences of decanalization [33] during cancer progression is too-often

sidelined, despite evidence supporting its prominence [34–36]. Yet despite these challenges, for which the complexity of the disease may be often in large part responsible, great progress has been and continues to be made. As we approach a new generation of immunotherapies, it is these very complexities that we must better understand in order to control or eradicate the disease.

Acknowledgements

References

[1] De Visser KE, Eichten A, Coussens LM. Paradoxical roles of the immune system during cancer development. Nature reviews cancer. 2006;6(1):24.

[2] Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. Nature Reviews Cancer. 2012;12(4):252.

[3] Restifo NP, Dudley ME, Rosenberg SA. Adoptive immunotherapy for cancer: harnessing the T cell response. Nature Reviews Immunology. 2012;12(4):269.

[4] Schreiber RD, Old LJ, Smyth MJ. Cancer Immunoediting: Integrating Immunity's Roles in Cancer Suppression and Promotion. Science. 2011;331(6024):1565–1570. doi:10.1126/science.1203486.

[5] Finn OJ. Immuno-oncology: understanding the function and dysfunction of the immune system in cancer. Annals of Oncology. 2012;23(Suppl 8):viii6–viii9. doi:10.1093/annonc/mds256.

[6] Ruffell B, DeNardo DG, Affara NI, Coussens LM. Lymphocytes in cancer development: polarization towards pro-tumor immunity. Cytokine & growth factor reviews. 2010;21(1):3–10.

[7] Hu B, Elinav E, Huber S, Booth CJ, Strowig T, Jin C, et al. Inflammation-induced tumorigenesis in the colon is regulated by caspase-1 and NLRC4. Proceedings of the National Academy of Sciences. 2010;107(50):21635–21640. doi:10.1073/pnas.1016814108.

[8] Balkwill F, Mantovani A. Inflammation and cancer: back to Virchow? The Lancet. 2001;357(9255):539–545. doi:10.1016/S0140-6736(00)04046-0.

[9] Guo Y, Nie Q, MacLean AL, Li Y, Lei J, Li S. Multiscale modeling of inflammation-induced tumorigenesis reveals competing oncogenic and onco-protective roles for inflammation. Cancer research. 2017; p. canres–1662.

[10] Nieto MA, Huang RYJ, Jackson RA, Thiery JP. EMT: 2016. Cell. 2016;166(1):21–45.

[11] Nie Q. Stem cells: a window of opportunity in low-dimensional EMT space. Oncotarget. 2018;9(61). doi:10.18632/oncotarget.25852.

[12] Sha Y, Haensel D, Gutierrez G, Du H, Dai X, Nie Q. Intermediate cell states in epithelial-to-mesenchymal transition. Physical Biology. 2019;16(2):021001. doi:10.1088/1478-3975/aaf928.

[13] Hong T, Watanabe K, Ta CH, Villarreal-Ponce A, Nie Q, Dai X. An Ovol2-Zeb1 mutual inhibitory circuit governs bidirectional and multi-step transition between epithelial and mesenchymal states. PLoS computational biology. 2015;11(11):e1004569.

[14] Jolly MK, Jia D, Boareto M, Mani SA, Pienta KJ, Ben-Jacob E, et al. Coupling the modules of EMT and stemness: A tunable 'stemness window' model. Oncotarget. 2015;6(28):25161–25174.

[15] Moris N, Pina C, Martinez Arias A. Transition states and cell fate decisions in epigenetic landscapes. Nature Reviews Genetics. 2016;17(11):693–703. doi:10.1038/nrg.2016.98.

[16] MacLean AL, Hong T, Nie Q. Exploring intermediate cell states through the lens of single cells. Current Opinion in Systems Biology. 2018;9:32–41. doi:10.1016/j.coisb.2018.02.009.

[17] Ta CH, Nie Q, Hong T. Controlling stochasticity in epithelial-mesenchymal transition through multiple intermediate cellular states. Discrete and Continuous Dynamical Systems - Series B. 2016;21(7):2275–2291. doi:10.3934/dcdsb.2016047.

[18] Rackauckas C, Schilling TF, Nie Q. Mean-Independent Noise Control of Cell Fates via Intermediate States. iScience. 2018;0(0). doi:10.1016/j.isci.2018.04.002.

[19] Lim J, Thiery JP. Epithelial-mesenchymal transitions: insights from development. Development. 2012;139(19):3471–3486.

[20] Terry S, Savagner P, Ortiz-Cuaran S, Mahjoubi L, Saintigny P, Thiery JP, et al. New insights into the role of EMT in tumor immune escape. Molecular oncology. 2017;11(7):824–846.

[21] Woods K, Pasam A, Jayachandran A, Andrews MC, Cebon J. Effects of epithelial to mesenchymal transition on T cell targeting of melanoma cells. Frontiers in oncology. 2014;4:367.

[22] Sherratt Jonathan A , Nowak Martin A . Oncogenes, anti-oncogenes and the immune response to cancer : a mathematical model. Proceedings of the Royal Society of London Series B: Biological Sciences. 1992;248(1323):261–271. doi:10.1098/rspb.1992.0071.

[23] Pillis LGd, Radunskaya AE, Wiseman CL. A Validated Mathematical Model of Cell-Mediated Immune Response to Tumor Growth. Cancer Research. 2005;65(17):7950–7958. doi:10.1158/0008-5472.CAN-05-0564.

[24] An G, Kulkarni S. An agent-based modeling framework linking inflammation and cancer using evolutionary principles: Description of a generative hierarchy for the hallmarks of cancer and developing a bridge between mechanism and epidemiological data. Mathematical Biosciences. 2015;260:16–24. doi:10.1016/j.mbs.2014.07.009.

[25] Serre R, Benzekry S, Padovani L, Meille C, André N, Ciccolini J, et al. Mathematical Modeling of Cancer Immunotherapy and Its Synergy with Radiotherapy. Cancer Research. 2016;76(17):4931–4940. doi:10.1158/0008-5472.CAN-15-3567.

[26] Louzoun Y, Xue C, Lesinski GB, Friedman A. A mathematical model for pancreatic cancer growth and treatments. Journal of Theoretical Biology. 2014;351:74–82. doi:10.1016/j.jtbi.2014.02.028.

[27] Benzekry S, Lamont C, Barbolosi D, Hlatky L, Hahnfeldt P. Mathematical Modeling of Tumor–Tumor Distant Interactions Supports a Systemic Control of Tumor Growth. Cancer Research. 2017;77(18):5183–5193. doi:10.1158/0008-5472.CAN-17-0564.

[28] West J, Robertson-Tessi M, Luddy K, Park DS, Williamson DFK, Harmon C, et al. The Immune Checkpoint Kick Start: Optimization of Neoadjuvant Combination Therapy Using Game Theory. JCO Clinical Cancer Informatics. 2019;3:1–12. doi:10.1200/CCI.18.00078.

[29] Owen MR, Stamper IJ, Muthana M, Richardson GW, Dobson J, Lewis CE, et al. Mathematical Modeling Predicts Synergistic Antitumor Effects of Combining a Macrophage-Based, Hypoxia-Targeted Gene Therapy with Chemotherapy. Cancer Research. 2011;71(8):2826–2837. doi:10.1158/0008-5472.CAN-10-2834.

[30] Morris MD. Factorial sampling plans for preliminary computational experiments. Technometrics. 1991;33(2):161–174.

[31] Sohier H, Farges JL, Piet-Lahanier H. Improvement of the Representativity of the Morris Method for Air-Launch-to-Orbit Separation. IFAC Proceedings Volumes. 2014;47(3):7954–7959.

[32] de Pillis LG, Radunskaya AE. Modeling tumor–immune dynamics. In: Mathematical Models of Tumor-Immune System Dynamics. Springer; 2014. p. 59–108.

[33] Gibson G. Decanalization and the origin of complex disease. Nature Reviews Genetics. 2009;10(2):134–140. doi:10.1038/nrg2502.

[34] Cyll K, Ersvær E, Vlatkovic L, Pradhan M, Kildal W, Kjær MA, et al. Tumour heterogeneity poses a significant challenge to cancer biomarker research. British Journal of Cancer. 2017;117(3):367–375. doi:10.1038/bjc.2017.171.

[35] Punt CJA, Koopman M, Vermeulen L. From tumour heterogeneity to advances in precision treatment of colorectal cancer. Nature Reviews Clinical Oncology. 2017;14(4):235–246. doi:10.1038/nrclinonc.2016.171.

[36] Dagogo-Jack I, Shaw AT. Tumour heterogeneity and resistance to cancer therapies. Nature Reviews Clinical Oncology. 2018;15(2):81–94. doi:10.1038/nrclinonc.2017.166.