

Supplementary Materials for Modeling the competing effects of the immune system and EMT on epithelial cancers

Daniel R. Bergman¹, Matthew Karikomi¹, Qing Nie^{1,2,*} and
Adam L. MacLean^{3,*}

¹Department of Mathematics, University of California, Irvine, Irvine,
CA 92697, USA

²Department of Cell and Developmental Biology, University of
California, Irvine, Irvine, CA 92697, USA

³Department of Biological Sciences, University of Southern California,
Los Angeles, CA 90089, USA

*Correspondence: qnie@uci.edu (Q.N.); macleana@usc.edu (A.L.M.)

Contents

| | | |
|----------|--|----------|
| 1 | Model Description | 2 |
| 1.1 | Tissue cell fate | 2 |
| 1.1.1 | Proliferation | 2 |
| 1.1.2 | Apoptosis | 3 |
| 1.1.3 | Immune Clearance | 3 |
| 1.1.4 | Rest in G_0 | 3 |
| 1.1.5 | Completing the Cell Cycle | 4 |
| 2 | Definition of parameters specifying the model | 6 |
| 3 | Parameter values used for simulation | 7 |

List of Tables

| | | |
|----|--|---|
| S1 | Parameter names and descriptions | 6 |
| S2 | Parameter values | 8 |

List of Figures

| | | |
|----|---|---|
| S1 | Fig. 1BC without targeting premalignant cells | 9 |
|----|---|---|

| | | |
|----|---|----|
| S2 | Fig. 2 without targeting premalignant cells | 9 |
| S3 | Fig. 3 without targeting premalignant cells | 10 |
| S4 | Fig. 4 without targeting premalignant cells | 10 |
| S5 | Fig. 5 without targeting premalignant cells | 11 |
| S6 | SKCM TCGA Data Analysis | 12 |
| S7 | LIHC TCGA Data Analysis | 13 |
| S8 | LUAD TCGA Data Analysis | 14 |

1 Model Description

1.1 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.

1.1.1 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. This lost proliferation for mesenchymal cells will later be used to increase their chance of resting. 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

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

1.1.2 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

$$\rho_A = d_C(1 - \delta_A \Delta_A) \quad (2.2)$$

1.1.3 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 malignant ($\delta_{\text{MUT}} = 1$). 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

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

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

1.1.4 Rest in G_0

The weight associated with rest is taken as 1 except in the case of mesenchymal cells. Recall that mesenchymal cells had their proliferation rate decreased by $1 - \zeta \Delta_{\text{MGA}}$ (see Eq. 2.1). The biological assumption here is that mesenchymal cells instead of proliferating will instead rest, so this lost proliferation weight is added to the resting weight. Hence, the weight of rest is given by

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

Again, 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.

1.1.5 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 driver mutation in one of the three pathways. If they do, one is randomly chosen among the three pathways and the pathway in that cell becomes altered. 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. The amount of TGF- β absorbed by all cells is given by an increasing Hill function in terms of the TGF- β in the TME. The saturation effect is to limit the amount of TGF- β a cell can absorb in a given time interval. This quantity is then divided up randomly among the N_C living cells via a normally distributed noise term to determine how much exogenous TGF- β each cell receives during this cell cycle. Should this value, τ_i in Eq. 2.5, be negative, we interpret this as the cell losing TGF- β to the TME and thus being more likely to undergo MET.

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

We then combine τ_i with the current EMT score of the cell, as a proxy for the endogenous TGF- β . Finally, if this quantity is large enough, the EMT score of the cell increases towards 1; otherwise, it decreases towards 0. 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

$$\tau = \tau_{\text{MUT}} N_{\text{MUT}} + \tau_{\text{Treg}} N_{\text{Treg}} \quad (2.6)$$

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

$$N'_{\text{NK}} = \sigma_{\text{NK}} - d_{\text{NK}} N_{\text{NK}} \quad (2.7)$$

which is discretized to

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

For CTLs and Tregs, they rely on malignant cells being cleared before they can be activated. Let $N_{\text{MUT}}^*(k)$ represent the number of malignant 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:

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

Discretized, these are:

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

2 Definition of parameters specifying the model

| 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 |
| K_4 | EC50 term for TGF- β activation of Tregs |
| 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 |
| d_{NK} | NK death rate |
| d_{CTL} | CTL death rate |
| d_{Treg} | Treg death rate |
| 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 S1:]

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

3 Parameter values used for simulation

| Name | Description | INFL Low Value | INFL High Value |
|-----------------------|---|---|----------------------|
| p | weight of proliferation for tissue cells | 0.28 | |
| d_C | weight of apoptosis for tissue cells | 0.14 | |
| Δ_{MIE} | MIE | 0.6 | |
| Δ_{MGA} | MGA | 0.2 | |
| Δ_A | proportional decrease to weight of apoptosis for cells with mutated apoptosis pathway | 0.3 | |
| Δ_{IE} | proportional increase to weight of immune evasion for cells with mutated immune evasion pathway | 0.48 | |
| Δ_P | proportional increase to weight of proliferation for cells with mutated proliferation pathway | 0.36 | |
| K_0 | EC50 term for negative feedback of tissue cells on own proliferation | 80 cells | |
| K_1 | EC50 term for probability of NK cell finding mutant cell | 8 cells | |
| K_2 | EC50 term for Treg inhibition of cytotoxic functions | 5 cells / volume | 0.025 cells / volume |
| K_3 | EC50 term for cumulative absorption of TGF- β | 200 amount / volume | |
| K_4 | EC50 term for TGF- β activation of Tregs | 50 amount / volume | |
| E_{NK} | weight of NKs clearing mutants | 10 | 30 |
| E_{CTL} | weight of CTLs clearing mutants | 200 | 600 |
| σ_{NK} | NK source rate | 1.3 cells / cycle | |
| σ_{CTL} | CTL source rate per cleared mutant cell | 100 cells / (cleared mutants \times cycles) | |

| | | | |
|------------------------|--|--|--|
| σ_{Treg} | Treg source rate per cleared mutant cell | 200 cells / (cleared mutants \times concentration of TGF- β \times cycles) | |
| d_{NK} | NK death rate | 0.13 / cycle | |
| d_{CTL} | CTL death rate | 0.0260 / cycle | |
| d_{Treg} | Treg death rate | 0.0260 / cycle | |
| k_{EMT} | EMT/MET rate | 0.01 / concentration of TGF- β | |
| σ | standard deviation of noise in TGF- β each cell receives | 6 concentration of TGF- β | |
| τ_{max} | max amount of TGF- β any cell can receive | 500 concentration of TGF- β | |
| τ_{MUT} | rate of TGF- β production by mutant cells | 0.05 concentration of TGF- β / cell / cycle | |
| τ_{Treg} | rate of TGF- β production by Treg | 0.5 concentration of TGF- β / cell / cycle | |
| | RP Cancer Line | 0.5 | |
| | INFL High Duration | 30 cycles | |
| | INFL Low Duration | 30 cycles | |
| | Mes Threshold | 0.7 | |
| | maximum initial mutation damage after warmup | 0.01 | |
| | increase in probability to mutate for non-mutating proliferating cells | 0.0001 | |

Table S2: The model parameter names, descriptions, and values during both low and high inflammation. Parameters with only one value do not change with the inflammatory state.

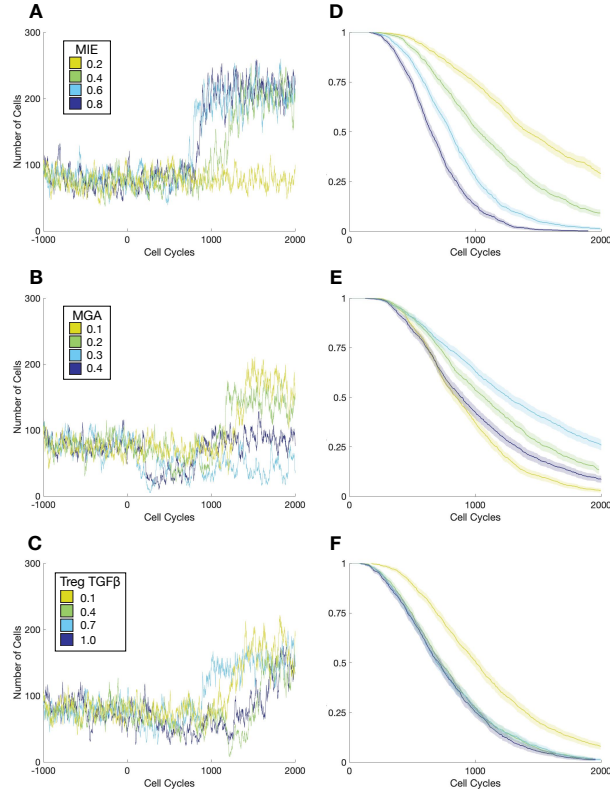


Figure S3: Effects of mesenchymal tumor cell properties on the Time to Cancer without immune cells targeting premalignant cells. Compare to Fig. 3.

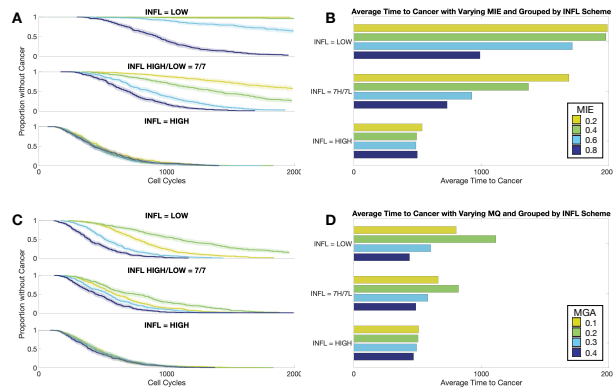


Figure S4: Effects of inflammation on the Time to Cancer without immune cells targeting premalignant cells. Compare to Fig. 4.

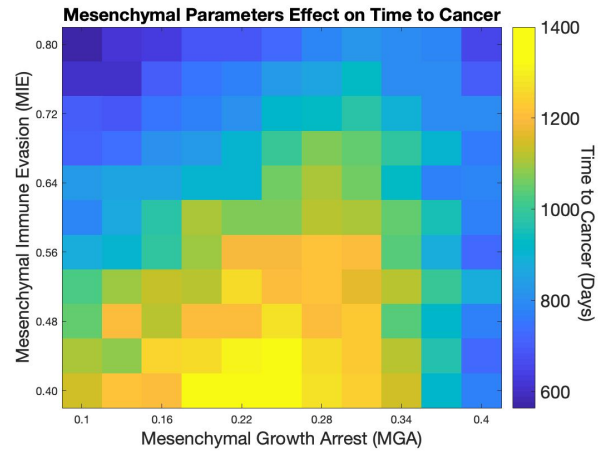


Figure S5: Summary of the contrasting effects of MIE and MGA on Time to Cancer without immune cells targeting premalignant cells. Compare to Fig. 5.

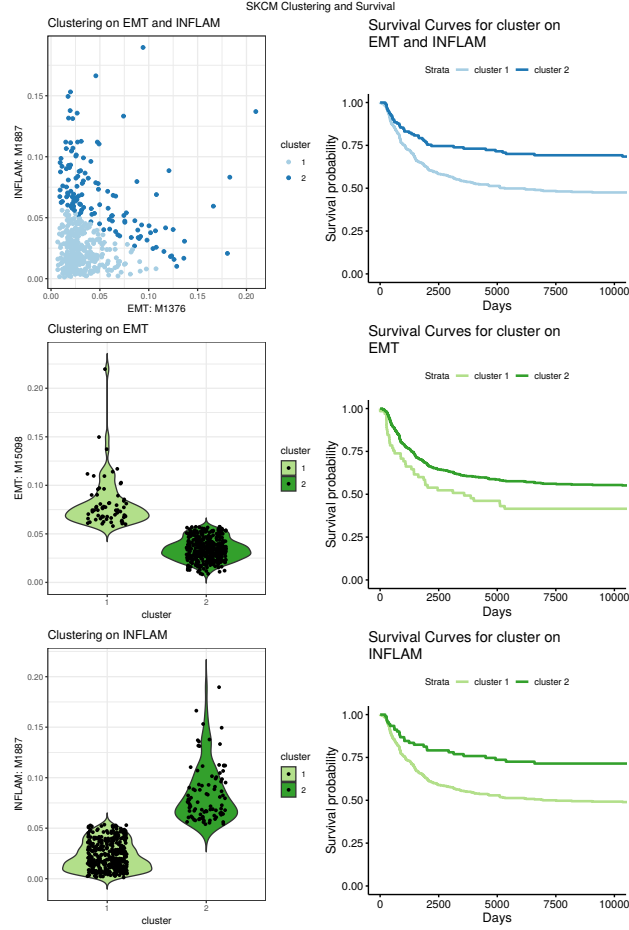


Figure S6: A. K-means clustering of SKCM using gene ontology terms indicative of EMT and inflammation signatures ($k = 2$). B. Survival plots corresponding to the clustering on EMT and inflammation. C. K-means clustering of SKCM using gene ontology terms indicative of an EMT signature ($k = 2$). D. Survival plots corresponding to the clustering on EMT. E. K-means clustering of SKCM using gene ontology terms indicative of inflammation ($k = 2$). F. Survival plots corresponding to the clustering on inflammation.

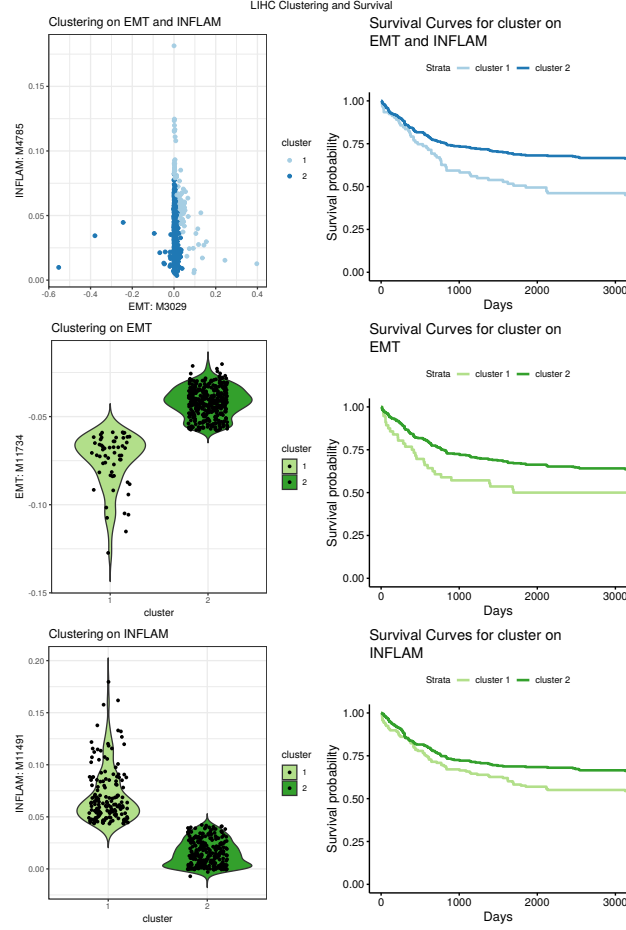


Figure S7:]

A. K-means clustering of LIHC using gene ontology terms indicative of EMT and inflammation signatures ($k = 2$). B. Survival plots corresponding to the clustering on EMT and inflammation. C. K-means clustering of LIHC using gene ontology terms indicative of an EMT signature ($k = 2$). D. Survival plots corresponding to the clustering on EMT. E. K-means clustering of LIHC using gene ontology terms indicative of inflammation ($k = 2$). F. Survival plots corresponding to the clustering on inflammation.

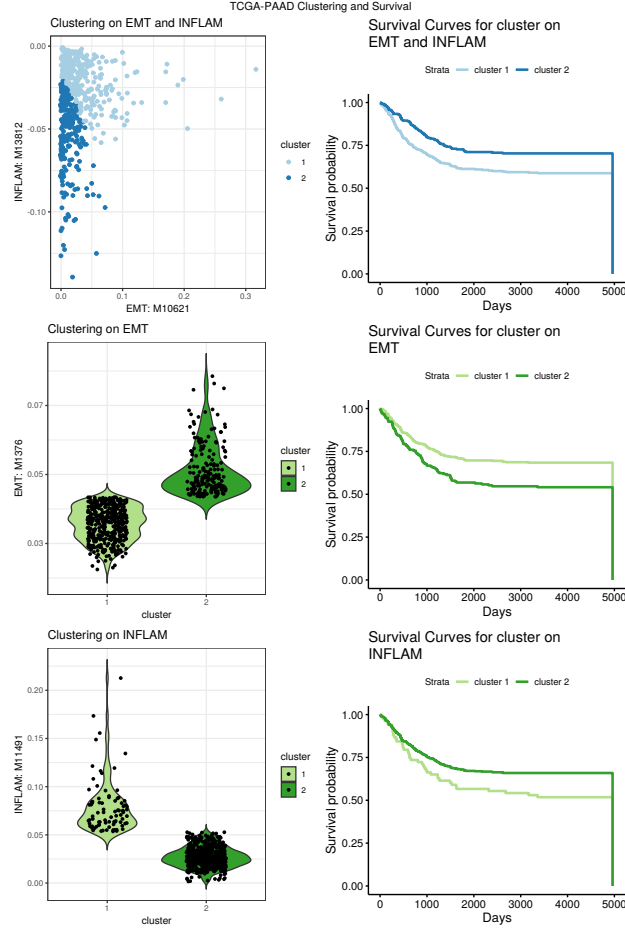


Figure S8: A. K-means clustering of LUAD using gene ontology terms indicative of EMT and inflammation signatures ($k = 2$). B. Survival plots corresponding to the clustering on EMT and inflammation. C. K-means clustering of LUAD using gene ontology terms indicative of an EMT signature ($k = 2$). D. Survival plots corresponding to the clustering on EMT. E. K-means clustering of LUAD using gene ontology terms indicative of inflammation ($k = 2$). F. Survival plots corresponding to the clustering on inflammation.