Modeling the evolutionary advantage of Tregs in spatially structured environments

Burcu Tepekule1\*, Roger D. Kouyos2,3,a, C. Jessica E. Metcalf4,a

1 Dept of Mucosal Immunology, University of Bern, Bern, Switzerland.

2 Dept of Infectious Diseases and Hospital Epidemiology, University Hospital Zurich, Zurich, Switzerland.

3 Institute of Medical Virology, University of Zurich, Zurich, Switzerland.

3 Dept of Ecology and Evolutionary Biology, Princeton University, Princeton, USA.

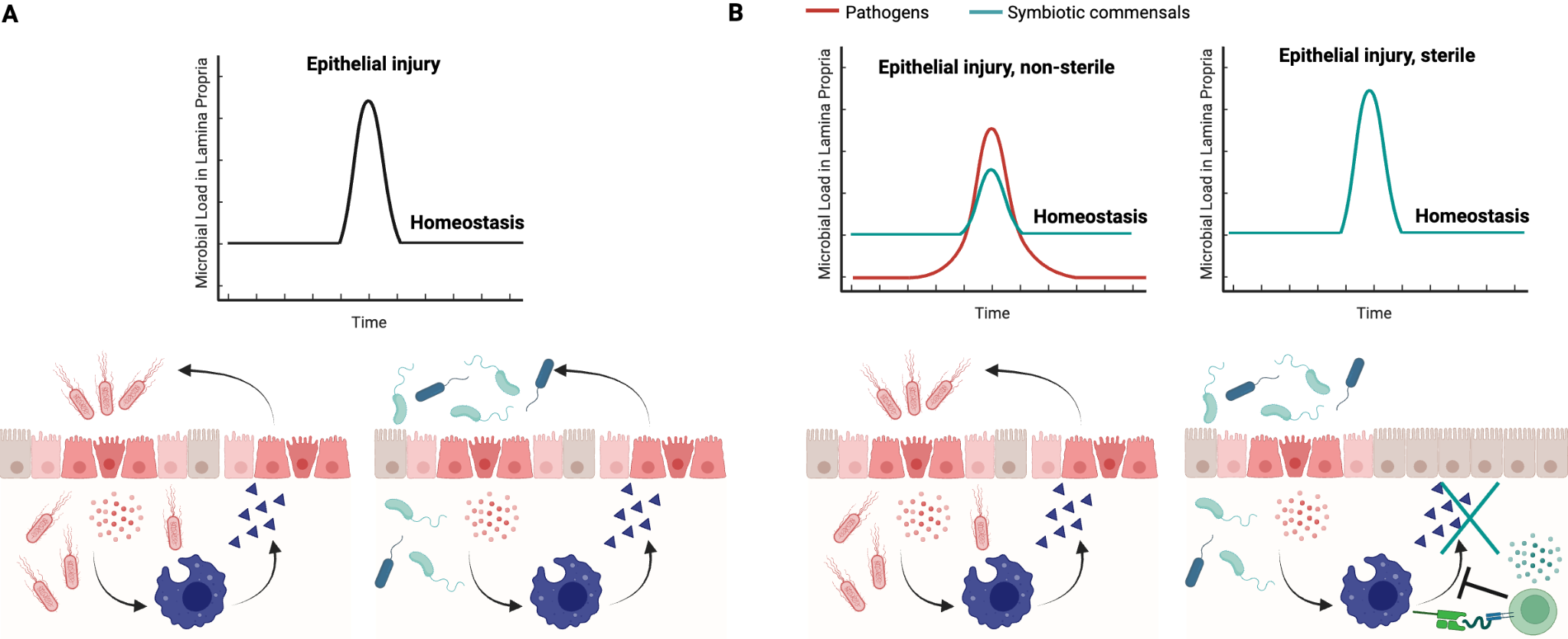
\*Corresponding author. Emails: [burcutepekule@gmail.com](mailto:burcutepekule@gmail.com)

a RDK and CJEM contributed equally to this work

**Abstract:**

**Introduction:**

**What evolutionary problem did the first regulatory T cells (Tregs) solve?** We hypothesize that it is to differentiate the type of tissue injury so that collateral damage is minimized (as opposed to) allowing for a more complex microbiome — but it allowed for adaptive immunity to bank on it and the microbiome got more “complex” afterwards - what is complex? probably the ones between a pathogen and a commensal, 2) suppressing autoimmune responses due to the self reactive clones that can rise by adaptive immune system - but that problem came after Tregs were invented, so that adaptive immunity can bank on it - turns into a chicken and an egg problem). The immune system responds to breaks in homeostasis, but these breaks can arise from different types of perturbations that require different responses—and the ability to distinguish between them and respond appropriately confers an evolutionary advantage by minimizing collateral damage. This can be seen as a primitive approach to avoid the collateral damage we see today with developed vertebrate adaptive immune systems and Tregs role in controlling autoimmunity.



**Fig 1. The immune system responds to a perturbation in homeostasis - but the type of perturbation matters.** **a)** Pathogenic and sterile injury without Tregs and cycle of tissue destruction, **b)** injuries with Tregs. Red dots are DAMPs released by the epithelial cells upon injury. Triangles are the reactive oxygen species (ROS) and nitric oxide (NO) that have antibacterial activity but along the way damage the epithelium (source of collateral damage). When Tregs sense that the macrophage is dominantly engulfing commensals, it signals it to block the secretion of ROS and NO, preventing further epithelial damage.

**Why is this a good strategy?** Danger signals can be released even when there is no actual threat to eliminate, which is why it makes sense for the immune system to learn what is safe—not necessarily what is self. When you have a sterile (pathogen-free) injury, you don't want to destroy your tissue even further, you should go into repair mode. Differentiating sterile from non-sterile injury gives you an evolutionary advantage, especially at the mucosal sites which are prone to having both types of injuries frequently.

**Why couldn't the epithelium do it?** Why couldn't the basolateral side of the epithelium evolve to use the RAG gene and differentiate what is leaking underneath and stop the DAMPs? Well that's hard, when you are ripped apart, there is no way of avoiding your DNA coming out, which is a very ancient danger signal.

**Why couldn't the phagocyte do it?** Because it needs to be activated when there are DAMPs. For it to do this job, it needs an internal memory that keeps track of who is a commensal vs pathogen, and has to combine that information with DAMPs. Not impossible maybe (through epigenetics?) but quite complicated.

**How does this align with evidence from modern Tregs?** They are good at wound healing! + look into MAIT cells. Belkaid had a paper on this?

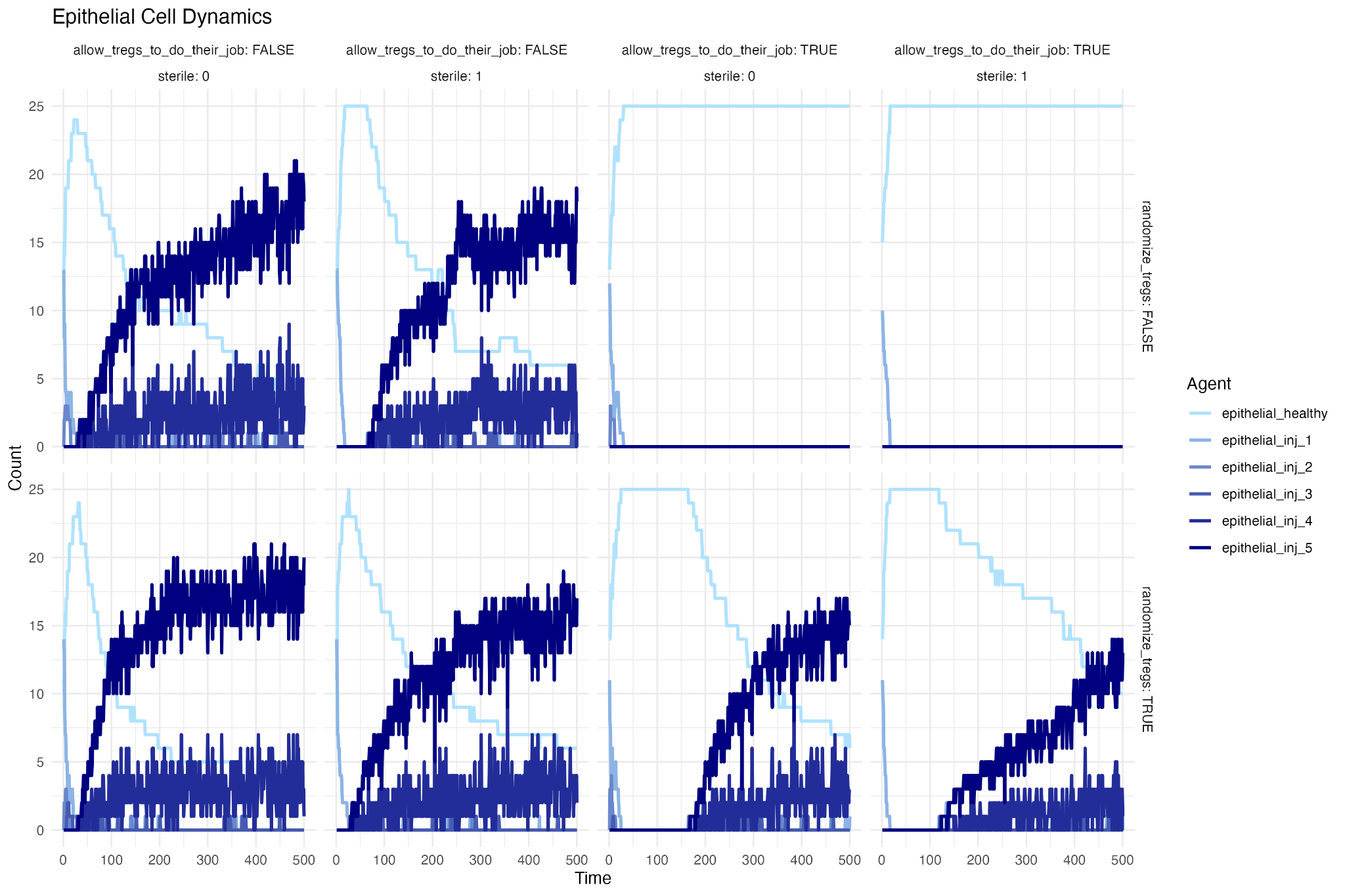
**It also aligns conceptually with the Danger model** - since it proposes that Tregs are yet another effector class that eliminates pathogens in a more silent way at tissues that can't afford Th1 destruction. So it goes back to minimizing tissue damage.

**Tregs being yet another type of effector cell:** This is based on their promotion of IgA and changing the “class” of the response (matzinger). IgA response is the least damaging among all, but is not enough on its own to eliminate pathogens (as seen from IgA deficient vs IgG deficient models). So why is it produced in massive amounts, if not that good at protecting from pathogens? Maybe that’s where the microbiome can get more “complex”. Commensals that are really not invasive do not need to be kept away from the epithelium, but the ones in the middle might need some “management” to not trigger immune responses constantly. That can make the microbiome more complex. Honestly IgA might help responses to pathogens, but that might be a side effect.

**Danger model and Treg differentiation:** It gets a bit confusing here - signal 0 needs to be present for any APC to be activated. So some level of stress (sterile or pathogenic) is necessary for any T cell differentiation. If Tregs are yet another class of T effector cells, that means their differentiation will require some signal 0 to be activated so that they can produce signal 2 for the T cell and tolerogenic signal 3 will help determine the class to be a Treg. A perfectly calm environment cannot do that. But this also means that the level of signal 0 required to activate a professional APC is very low? It's almost like it doesn't really matter, and everything boils down to signal 3? If the threshold of activation by signal 0 is that low, then APCs are always active? Also - do immature DCs render T cells anergic (as danger model suggests) or help Treg differentiation (as some literature suggests)? -> This actually gets into early life - if no immunostimulatory LPS, no Tregs as well (Vatanen paper: complete absence of any stimulation leads to immunological ignorance, not tolerance. Active tolerance requires some level of engagement.). If the signal 0 threshold is so low that it's always crossed, then it becomes meaningless as a danger detector. It starts to look more like **signal 0 is required for basic APC function**, not danger discrimination. After the APC is activated, most of the job is on Signal 3 to determine the class, which includes the microenvironmental cues.

**Why is it not a bad strategy to dampen the response during the end of an infection?** Decision by the Treg cannot be made based on “only” presenting commensal antigens. It will always be density dependent (more commensals than pathogens, how much more?). Therefore phagocytes will also be shut down when pathogen load is relatively small compared to commensals. This might seem like a bad strategy: you dampen the response toward the end of an infection – or at the beginning – when the pathogenic load is low – wouldn't that give the pathogen another advantage to rise again? Not really thanks to spatial heterogeneity! It is not even neutral, it is useful. Because as the epithelium heals, it lets less pathogens leak into lamina propria.

**How did the evolutionary path continue from here?** Now you have the type of cell that the information can be dumped onto by sampling the environment minimally during homeostasis. Consider a sterile organ like the liver. During apoptosis - without any infection related inflammation - apoptotic cells will be digested by macrophages. They will present these antigens in the absence of danger signals - same as commensals. So why would the initial location for Tregs to emerge be the gut? By this logic it can be anywhere? I guess (not a definitive answer) but the high frequency and dual nature of immune challenges in the gut (both sterile and pathogenic) likely made it the selective pressure for Treg evolution.



**When Tregs Don't Help?**

**How to test this experimentally?** Test whether lamprey MPL-L+ VLRA+ cells can suppress macrophage activation in co-culture, especially when the stimulus is commensals vs. pathogens.

Think.

**Code and Data availability:** <https://github.com/burcutepekule/Treg_problem>

**Supplementary Materials**

Agent-Based Model Simulation Rules

**1. General Parameters & Environment**

* **Grid Size:** The simulation takes place on a 25×25 grid.
* **Time Steps:** The simulation runs for a maximum of 1000 time steps ().
* **Injury Site**: An initial injury is defined as a percentage of the grid width (default 60%) in the first row, representing the epithelial layer.
* **Stochasticity:** Since the same parameter set can yield different outcomes depending on the random seed, each set is simulated 10 times. A parameter set is assigned to a class—Tregs helping, harming, or being irrelevant to epithelial health or pathogen control—if the corresponding outcome is observed more than half the time across the realizations (more than 5 runs in 10 runs per parameter set).

**2. Epithelial Cells**

* **Location:** Epithelial cells are located at y=0 (effectively y=1 in the 1-indexed grid for interaction purposes).
* **Injury Levels**: Epithelial cells have 6 states:
  + 0: Healthy
  + 1 to 5: Injured, with 5 being the maximum injury level.
* **Initial Injury:** The simulation begins with a percentage of epithelial cells at injury level 1 within the injury\_site.
* **Injury Progression:**
  + Microbe-induced: Injury level of the epithelial cells increase based on the presence of pathogens. Pathogens are virulent and injure epithelial cells. Commensals are not virulent and do not directly injure epithelial cells in the same way, but contribute to DAMPs upon detection.
  + ROS-induced: If the mean ROS level in the vicinity of an epithelial cell exceeds a threshold (th\_ROS\_epith\_recover = 0.15), its injury level increases by 1.
  + Maximum Injury: Injury level cannot exceed 5 (max\_level\_injury).
* **Recovery:** Injured epithelial cells recover stochastically with a certain probability at each time step to reduce their injury level by 1 (stochastic recovery).
* **DAMP Release:**
  + Epithelial cells release DAMPs proportional to their level\_injury.
  + Basolateral PRR stimulation by microbes (both commensals and pathogens) also triggers DAMP release from the epithelium. The epithelium cannot differentiate between commensals and pathogens based on this initial PRR stimulation.
* **Microbe Leakage:** The more injured an epithelial cell is, the more commensals and pathogens are leaked underneath it into the lamina propria.

**3. Microbes**

* **Types:** Commensals and Pathogens.
* **Movement:** Microbes move randomly (one step in any of 8 directions or stay still). If at y=1 (epithelium layer), they prioritize moving deeper into the tissue (y can only increase or stay the same).
* **Leakage into Lamina Propria:**
  + Pathogens: Leak into the lamina propria from injured epithelial cells, with the rate rate\_leak\_pathogen\_injury proportional to the average level\_injury and length(injury\_site\_updated). If sterile = 1, no new pathogens leak.
  + Commensals: Leak at a baseline rate (rate\_leak\_commensal\_baseline) across the entire epithelium and at an increased rate (rate\_leak\_commensal\_injury) from injured epithelial cells.
* **ROS-induced Death:** Microbes die if the local ROS concentration at their location exceeds a threshold (th\_ROS\_microbe).

**4. Phagocytes (Macrophages/Dendritic Cells)**

* **Population:** A fixed number of phagocytes (n\_phagocytes = 200).
* **Movement:** Phagocytes move towards areas of higher DAMP concentration. If no DAMP gradient exists, they move randomly.
* **Phenotypes:** Phagocytes can have three phenotypes:
  + **M0 (Resting):** Default state.
  + **M1 (Pro-inflammatory):** Activated by DAMPs predominantly.
  + **M2 (Anti-inflammatory/Resolving):** Activated by SAMPs predominantly.
* **Activation:**
  + M0 phagocytes differentiate into M1 if local DAMPs (avg\_DAMPs) exceed activation\_threshold\_DAMPs and avg\_DAMPs > avg\_SAMPs.
  + M0 phagocytes differentiate into M2 if local SAMPs (avg\_SAMPs) exceed activation\_threshold\_SAMPs and avg\_SAMPs > avg\_DAMPs.
* **Digestion & Activity:**
  + **Microbe Registry:** Each phagocyte maintains a phagocyte\_bacteria\_registry, which is a matrix tracking microbes it has recently engulfed.
  + **Digestion Time:** Every digestion\_time steps (default 1), the oldest entry in the registry is "digested" (removed), and new entries are added for newly engulfed microbes. This means the bacteria\_count (sum of current entries in the registry) reflects the microbes currently being processed.
  + **Activity Adjustment:** The bacteria\_count directly influences the phagocyte's activity levels:
    - **ROS Production (activity\_ROS):** M1 phagocytes produce ROS. Their activity increases proportionally with the bacteria\_count (up to cc\_phagocyte, which is the maximum number of bacteria a phagocyte can effectively process at one time). M0 and M2 phagocytes have minimal or no ROS production.
    - **Engulfment (activity\_engulf):** All phenotypes can engulf microbes. M1 and M2 phagocytes have higher baseline engulfment activity than M0. Their engulfment activity also increases proportionally with the bacteria\_count they are currently processing.
* **Phenotype Change & Deactivation (from Active State):** After active\_age\_limit steps, active phagocytes (M1 or M2) reassess their environment:
  + If avg\_DAMPs >= activation\_threshold\_DAMPs and avg\_DAMPs > avg\_SAMPs, the phagocyte becomes (or remains) M1.
  + If avg\_SAMPs >= activation\_threshold\_SAMPs and avg\_SAMPs > avg\_DAMPs, the phagocyte becomes (or remains) M2. This means an M1 phagocyte can differentiate into an M2 if SAMPs become the dominant signal.
  + If both avg\_SAMPs < activation\_threshold\_SAMPs and avg\_DAMPs < activation\_threshold\_DAMPs, the phagocyte returns to the M0 (resting) state.

**5. Tregs (Regulatory T cells)**

* **Population:** A fixed number of Tregs (n\_tregs = 200).
* **Movement:** Tregs move towards areas of higher DAMP concentration. If no DAMP gradient exists, they move randomly.
* **Phenotypes:** Tregs can be:
  + 0 (Resting)
  + 1 (Activated)
* **Activation:**
  + If allow\_tregs\_to\_do\_their\_job is TRUE: Tregs activate if they are in the vicinity (treg\_vicinity\_effect = 1) of an M1 or M2 phagocyte that has engulfed a high ratio of commensal to pathogenic antigens (rat\_com\_pat > rat\_com\_pat\_threshold).
  + Activated Tregs immediately start producing SAMPs.
* **Deactivation:** Activated Tregs return to a resting state after active\_age\_limit steps, stopping SAMP production.

**6. Signaling Molecules (DAMPs, SAMPs, ROS)**

* **Generation:**
  + **DAMPs (Damage-Associated Molecular Patterns):** Released by injured epithelial cells, by epithelial cells stimulated by microbes (both commensals and pathogens), and at the location of the pathogens themselves, assuming that they are harming other cell types within the lamina propria—such as stromal cells—leading to the release of similar intracellular components to those from epithelial injury, including ATP, mitochondrial DNA, and nuclear proteins.
  + **SAMPs (Suppression-Associated Molecular Patterns):** Released by activated Tregs.
  + **ROS (Reactive Oxygen Species):** Released by M1 phagocytes.
* **Diffusion:** All signaling molecules diffuse across the grid (diffuse\_matrix function) at specific diffusion\_speed rates. The maximum value for any cell is 1.
* **Decay:** All signaling molecules decay over time (ros\_decay, DAMPs\_decay, SAMPs\_decay).

**7. Metrics Recorded (Longitudinal Data)**

At each time step, the following data is recorded:

* **Epithelium:** Number of cells at each injury level (0-5).
* **Phagocytes:**
  + Number of M0 phagocytes.
  + Number of M1 phagocytes at each active age level (0-5).
  + Number of M2 phagocytes at each active age level (0-5).
* **Microbes:** Total number of commensals and pathogens.
* **Tregs:** Number of resting and active Tregs.
* **Microbe Cumulative Death:** Counts of commensals and pathogens killed by ROS, M0 phagocytes, M1 phagocytes, and M2 phagocytes.