

High-Performance Interaction-Based Simulation of Gut Immunopathologies with ENteric Immunity Simulator (ENISI)

Keith Bisset*, Md. Maksudul Alam*, Josep Bassaganya-Riera*, Adria Carbo*, Stephen Eubank*, Raquel Hontecillas*, Stefan Hoops*, Yongguo Mei*, Katherine Wendelsdorf*, Dawen Xie*, Jae-Seung Yeom*[†], and Madhav Marathe*[†]

*Virginia Bioinformatics Institute

[†]Department of Computer Science

Virginia Tech, Blacksburg VA 24061

Email: {kbisset,maksud,jbassaga,acarbo,seubank,rmagarzo,shoops,yimei,dawenx,wkath83,jyeom,mmarathe}@vbi.vt.edu

Abstract—Here we present the ENteric Immunity Simulator (ENISI), a modeling system for the inflammatory and regulatory immune pathways triggered by microbe-immune cell interactions in the gut. With ENISI, immunologists and infectious disease experts can test and generate hypotheses for enteric disease pathology and propose interventions through experimental infection of an *in silico* gut. ENISI is an agent based simulator, in which individual cells move through the simulated tissues, and engage in context-dependent interactions with the other cells with which they are in contact. The scale of ENISI is unprecedented in this domain, with the ability to simulate 10^7 cells for 250 simulated days on 576 cores in one and a half hours, with the potential to scale to even larger hardware and problem sizes.

In this paper we describe the ENISI simulator for modeling mucosal immune responses to gastrointestinal pathogens. We then demonstrate the utility of ENISI by recreating an experimental infection of a mouse with *Helicobacter pylori* 26695. The results identify specific processes by which bacterial virulence factors do and do not contribute to pathogenesis associated with *H. pylori* strain 26695. These modeling results inform general intervention strategies by indicating immunomodulatory mechanisms such as those used in inflammatory bowel disease may be more appropriate therapeutically than directly targeting specific microbial populations through vaccination or by using antimicrobials.

Keywords-Computational Immunology; Parallel Efficiency and Scalability; Agent Based Simulation; BioComputing

I. INTRODUCTION

Enteric diseases are diseases of the gastrointestinal (GI) tract often caused by ingestion of microbes in food and water. Upon microbe entry, immune cells in the GI tract mount an inflammatory response that eliminates the microbe, but may also cause tissue damage. This collateral damage is often the basis disease pathogenesis.

As the GI tract is constantly exposed to foreign antigens, mostly innocuous, this inherent inflammatory response must be regulated so that the system does not remain in a constant state of tissue-damaging hyper-inflammation. Immune regulation is carried out by the *regulatory*, or anti-inflammatory, immune response triggered by factors such as host tissue damage or commensal gut microflora. The current picture of the gut mucosa is one in which immune cells

of a dynamic balance between regulatory and inflammatory responses, with regulatory phenotypes generally predominating [1], [2]. Understanding which components of these immune pathways contribute to microbial persistence and severity of symptoms is necessary to devise treatments and infection prevention strategies against gut pathogens such as pathogenic strains of *Escherichia coli* and *Helicobacter pylori*.

Here we present the ENteric Immunity Simulator (ENISI) [3], a modeling environment for studying the inflammatory and regulatory immune pathways initiated by microbe-immune cell interactions in the gut. ENISI is an interaction-based model where individual cells are modeled, along with their movement through different tissues, and the probabilistic outcomes of cell-cell interaction. ENISI has the ability to simulate at least 10^8 individual cells. With ENISI, mucosal immunologists can test and generate hypotheses for enteric disease pathology and propose interventions through experimental infection of an *in silico* gut. This is done by using a simple scripting language to assign parameter values that conform to one's knowledge and assumptions of the experimental scenario they wish to simulate. Simulation outcomes given different experimental conditions allow observation of *in silico* behaviors that are not readily seen through *in vitro* and *in vivo* techniques. This information can then be used to better understand immunological mechanisms and to generate novel treatment strategies that can be tested in the laboratory.

In the next section, we discuss the significance of modeling mucosal immune responses. Section III defines the biological model we are studying and formally defines the problem. Section IV discusses the implementation, including extensions to the EpiSimdemics simulator to support ENISI and Section V presents the scalability of our simulation. In Section VI we present a study that demonstrates the utility of ENISI. Finally, we discuss where ENISI is headed.

II. SIGNIFICANCE AND RELATED WORK

Aspects of the presented inflammatory and regulatory immune pathways have been represented in previous models

of mucosal infection [4]–[7] that have provided insight on mechanisms of clinical symptoms as well as pathogen persistence. The ENISI model is unique in its scope and approach. The model incorporates regulatory mechanisms of both adaptive and innate immunity, multi-location migration of cells, and cross talk between antigen presenting cells and T-cells. In addition, it is mechanism-based – explicitly representing each participating cell of the immune pathway. This facilitates mapping of model parameter specifications and predictions to laboratory techniques that manipulate specific cell populations.

We previously implemented a larger scale version of the model, encompassing these aspects, as a system of differential equations. Simulations based on this initial version identified a relationship between the effector CD4+ T helper cells (Th) and classically activated M1 macrophage concentrations in the LP and chronic epithelial damage [8]. However, ordinary differential equations (ODEs) can only capture the dynamics of each cell population as a whole. Hence, this work identified a relationship between M1 and Th levels and epithelial damage, but the ODE representation did not allow us to identify the specific pathways in which T cells induce epithelial damage after being stimulated by M1 macrophages. An additional drawback of the ODE representation is that it assumes deterministic, average behavior by each individual cell. However, biological systems are known to act stochastically due to attributes, such as cytokine secretion and association time with stimulating factors, that vary widely across individual cells in a population. Additionally, the randomness introduced by cell movement leads to non-uniform distribution across single tissue sites. Due to these assumptions of determinism and homogeneity, that are surely violated by the system in reality, dynamics predicted by an ODE model may not accurately reflect those seen in nature.

The ENISI model can be viewed as an extension of the interacting state machine models or agent-based models. A key aspect of these models is a procedural and interactive (a.k.a. mechanistic, algorithmic, executable) view of the underlying systems. In this view, components of the system interact locally with other components and the behavior of individual objects is described procedurally as a function of the internal state and the local interactions. This agent-based approach allows incorporation of spatial effects and randomness of cell-cell and cell-bacteria contact. In the case of colonic inflammation spawned by a small number of pathogens, such randomness is believed to significantly affect the outcome of the system and, therefore, an agent-based model is an appropriate representation [9]. This also creates a foundation for encompassing emergent properties such as bacterial strain evolution and changes in microflora demographics as the model is elaborated and the simulator extended. However, the drawback to such methods is that they are often not scalable due to limitations of computation

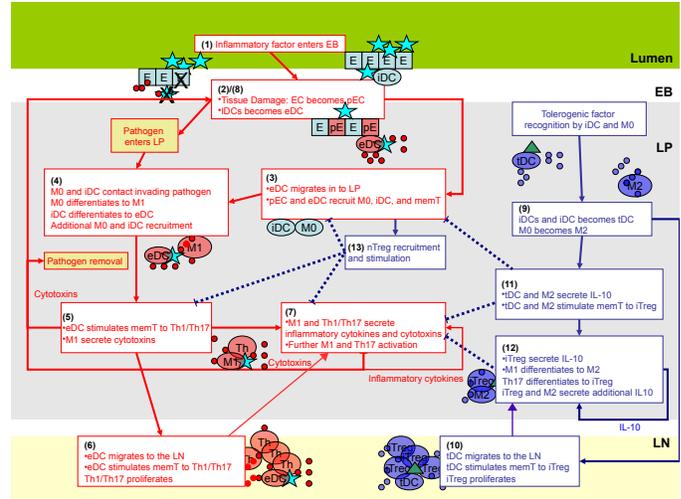


Figure 1. Illustration of sequential events in the inflammatory (red arrows) and regulatory (blue arrows) pathways described in the text. Dashed lines indicate events that inhibit the occurrence of another event.

power.

Scalability is highly relevant when seeking to reproduce emergent tissue-level phenomena by simulating individual cell interactions. Larger scale models are necessary as the purpose of immune simulators is to reproduce dynamics in a true *in vivo* system where immune cell concentrations can reach $10^8/mL$ [10]. It may not be sufficient to simulate the dynamics of a small sample and extrapolate results to the entire organ. To do so is to ignore the non-linear and complex nature of the cell interactions and dynamics and make the assumption of uniform mixing which defeats the purpose of an agent-based approach.

There are various general, agent-based biological simulator tools publicly available including Rhapsody [11], [12], NFSim [13], BIS [14], and that developed in [15] that translate graphical models in to executable code to run simulations. These simulators place an emphasis on rules governing cell-cell contacts and signaling interactions allowing one to enter complicated functions for these mechanisms. They, therefore, provide the useful capability of incorporating complex mathematical models for receptor-ligand interactions and phenotype differentiation in to cell contact networks. However, the scalability of these implementation algorithms in terms of system complexity and the number of cells in a network is unclear. For example, Rhapsody has been shown to simulate up to 10^4 individuals efficiently [11], [12].

ENISI is an interaction-based modeling environment of immunological processes at the cellular level. The resolution at which ENISI works is one of its unique features – we can simulate 10^7 cells for three simulated months in $1\frac{1}{2}$ hours. The model is represented by a domain specific language in a textual form that is given as input to the

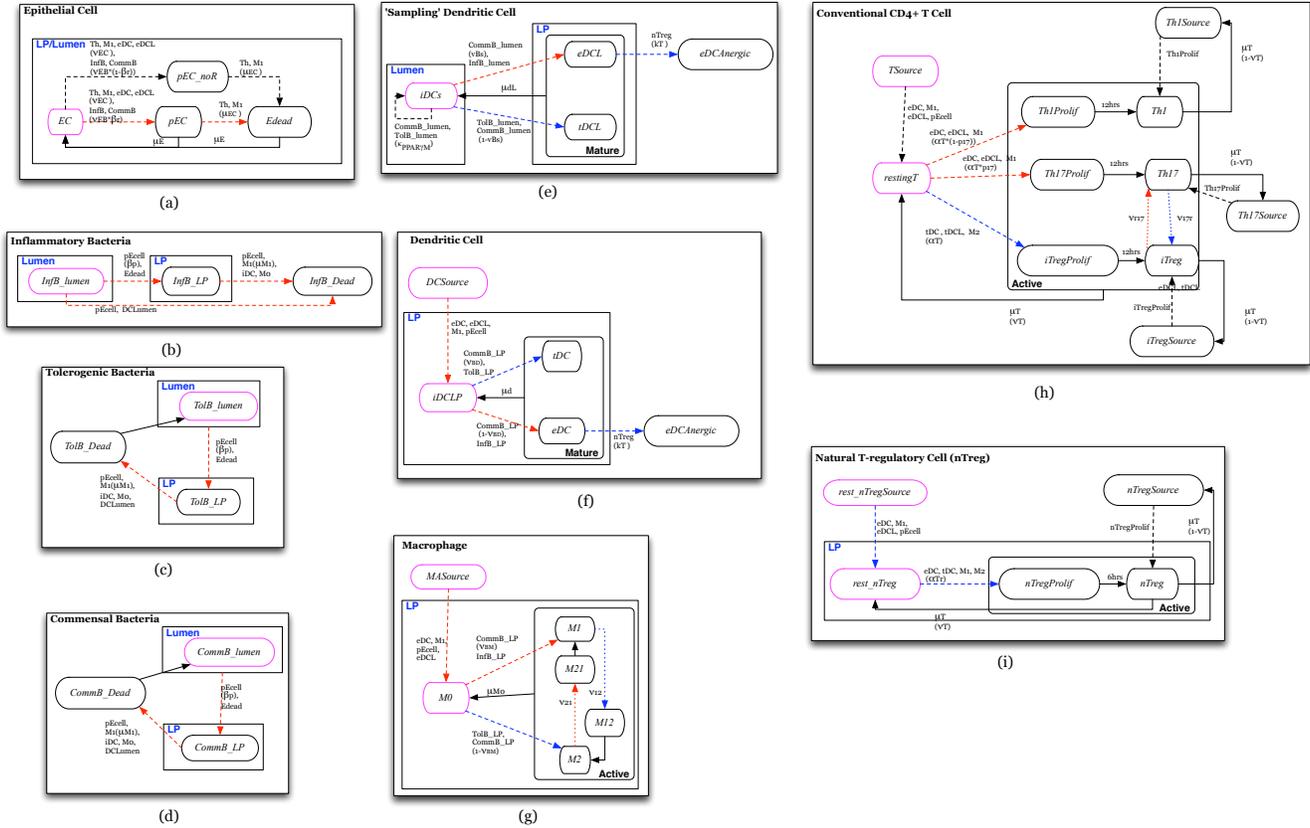


Figure 2. State transition functions for each of the eight automaton that represent specific cell-types. Each cell is represented by one of the eight automaton.

model. Although this specification is complex, it does not require extensive programming knowledge to modify the interaction rules. In addition, a graphical user interface will be publicly available at <http://www.modelingimmunity.org/models/enisi-helicobacter-pylori/> that will allow Immunologists, with minimal training, to parameterize, run, and analyze ENISI experiments. This is part of a larger effort to increase the visibility of *in silico* experimentation and systems biology approaches in immunology research and discovery.

The work reported in this paper advances our previous work in three areas. First, the model has been improved in several areas, as described in Section IV-D. Second, several extensions have been made to the underlying simulator in order to produce ENISI, as described in Section IV-C. Finally, the scalability of the system is several orders of magnitude greater than other similar models, as described in Section V.

III. PROBLEM DESCRIPTION

The specific inflammatory and regulatory immune pathways encoded in ENISI are shown in Figure 1. A complete description can be found in [16]. There are 4 compartments

represented in ENISI, which are defined on the basis of function and anatomy: *i*) the *Lumen*, which has a direct connection to the external environment, is the entry site for ingested food and foreign microbes, and houses the gut microflora, *ii*) the *Lamina propria* (LP), tissue separated from the lumen by an epithelial monolayer that is occupied by resting immune cells, *iii*) the epithelial barrier (EB), a monolayer of columnar epithelial cells, that divides the lumen and LP, and *iv*) the gastric or mesenteric lymph node (LN), the primary site of T cell activation. Also known as inductive sites of the mucosal immune system (i.e., where immune responses are initiated).

In Figure 2 we show the state transition functions for each of the automata corresponding to a specific cell-type in a state chart like formalism [17]. Red arrows indicate transitions that represent events in the inflammatory pathway depicted in Figure 1 and blue arrows indicate transitions that represent those of the regulatory pathway. Ovals represent states of the automaton. Solid arrows represent time-dependent transitions labeled with the time in one state before transitioning to another. The dashed arrows represent single contact-dependent transitions, labeled with the set of *Interactor* states necessary to induce state transition and,

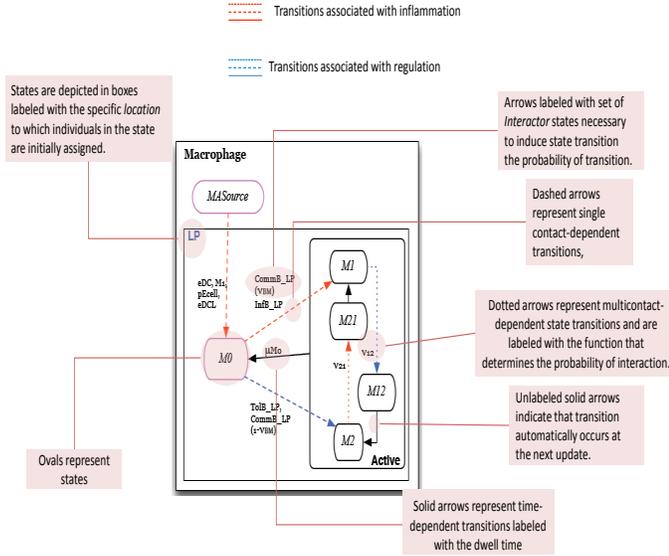


Figure 3. **State chart description.** A labeled state chart depiction of the **Macrophage** automaton.

in parenthesis, the probability of transition upon interaction. The default probability is 1. Dotted arrows represent multicontact-dependent state transitions and are labeled with the function that determines the probability of interaction. Unlabeled solid arrows indicate that transition automatically occurs at the next update. States outlined in pink indicate the initial state that determines which automaton a cell will be. States are depicted in boxes labeled with blue text that indicate the specific *Location* to which individuals in the state are initially assigned. An illustrative description of this state chart like formalism is given in Figure 3.

Parameter values were assigned according to measurements published in literature when available and default model assumptions explained in [18]. There one can also find a detailed explanation of each automaton and semantic approximations of the model.

A. Formalization of ENISI

The formal mathematical model is a co-evolving graphical discrete dynamical system (CGDDS) framework that captures the co-evolution of emergent inflammation/regulation dynamics, the interaction network and individual cellular behavior. The CGDDS formalism is described in [19]. We will modify this mathematical model so that it captures cellular interactions where in the cells can move and thus cause a change in the interaction network. Due to lack of space, we provide an informal description; a complete description can be found in [?].

An extended **Co-evolving Graphical Discrete Dynamical System** (CGDDS) *CGDDS* over a given domain \mathbb{D} of state values is a triple (G_t, \mathcal{F}, W) where G_t is a time varying graph that is the cell contact network, \mathcal{F} is a set

- 1: **for** $t = 0$ **to** T **do**
- 2: Compute the interaction graph G_t using function g .
- 3: **for** for each vertex v **do**
- 4: Compute $s^t(v)$ using state update function f
- 5: **end for**
- 6: **end for**

Figure 4. Pseudo-code describing ENISI execution as formalized using CGDDS.

of functions that describes state transitions and W is an update scheme. In the rest of the paper, we will assume that nodes are updates synchronously. The components can be described as follows:

The graph $G_t(V_t, E_t)$ represents the dynamic network – the set of vertices and edges evolve in time. We will often omit the time index for sake of simplicity and the usage will be clear from context. $V = \{v_1, v_2, \dots, v_n\}$ represents the set of vertices in the graph, denoting the set of cells or bacteria. Also let $S = \{S_1, S_2, \dots, S_m\} \in \mathbb{D}$ be a set of all possible states (phenotypes) that the cells can take. For any vertex (cell or a bacteria) $v \in V$, we define $s^t(v) \in S$ be the state of vertex v at time t .

Cells/Bacteria move through the tissue and this movement is the basis of network change. Recall that the edges in the network are based on spatial proximity. As cells move through the tissue, new edges are formed and some of the old edges are deleted. The details of the movement model are described subsequently. For each vertex (cell) v , let g denotes the edge modification function that takes as input the state $s^t(v)$ (including its position) of node v at time t and returns the set of edges that v will be adjacent to in a given time period. In other words, g captures the time varying edges at $t + 1$ in the graph resulting from the movement of the cell.

We now discuss how cell states are updated. For each vertex $v \in V$ let f be the state transition function. The function f maps the state of vertex v at time t to its state at time $t + 1$; and the state transition depends on states of other vertices and the edges incident on v . If vertices v_1, v_2, \dots, v_k are adjacent to vertex v we can represent the function as:

$$s^{t+1}(u) = f(s^t(u), s^t(v_1), s^t(v_2), \dots, s^t(v_k))$$

This function may also be stochastic. As stated, the local transition function depends on the state of all nodes that are adjacent. In the following discussion, we will describe a few approximations that render this computation more efficient.

Figure 4 describes the dynamics of ENISI as formalized using modified CGDDS. in the extended CGDDS. Note that the state update for each cell is performed synchronously. That is, the network and state updates are *realized* only when all nodes have finished executing the relevant functions.

B. Modeling Immune System using CGDDS

We now discuss how we can model the immune system and its interaction with the bacteria in a gut using extended CGDDS described above.

The dynamic network As mentioned earlier the vertices of our dynamic network represent cells as well as bacteria (we will call them agents). Each cell can be in one of many states (phenotypes). A partial list cell phenotypes are given in the Table II. A complete list can be found in [16].

Vertices are connected by edges – these are abstractions of the fact that cells that are spatially proximal. The length scale – maximum distance between cells that allows for the likelihood of interaction is parameterized to capture the relevant biology; as we will see later in our model, cells within a sub-location can interact with each other. Spatially proximal cells can interact with each other – these interactions change the state of of the cells. The dynamic graph is a result of these individual agents moving through the tissue as discussed earlier.

In the current implementation, the movement model is simple: there are three tissue sites: (i) Lumen, (ii) LP and (iii) the lymphoid. These are called locations in our language. Each of the tissue is divided into small spatial patches – called sublocations. Agents move randomly through these patches. All agents within a patch are deemed close enough so as to be able to interact. All cell-types and bacteria *except epithelial cells* change sublocations every 30 real minutes. The movement of the cells from one sublocation to another sublocation is random. Epithelial cells do not change sublocations once they are assigned a sublocation. On the other hand, location for each agent is governed by their current state and the time that has elapsed in that state. Location changes can happen only after 6 hours of real time. System updates itself synchronously every 6 hours and this is described further in subsequent sections.

Interactions and Update Function E_t represents set of contacts between any two agents at time t . Cells interact with other cells and bacteria – this leads to effector or regulatory responses depending on the execution of the pathway. For simplicity, certain interactions are best represented as modification of cell states as it is affected by concentration of certain chemicals. This view allows in effect a “mass action” like interaction. In general there are four kinds of interactions – (i) interaction between a cell and a bacteria, (ii) interaction between two cells, (iii) interaction between a cell and a group of cells.

A convenient way to represent a local update function is to use an appropriate automaton. Here we use probabilistic timed transition systems (PTTS) to represent the time evolution of the cell states. A PTTS is a set of states. Each state has an id, a set of attributes values, a dwell time distribution, and one or more labeled sets of weighted transitions to other states. The label on the transition sets is used to select

the appropriate set of transitions. The attributes of a state describe the features possessed by a cell that is in that state. Once a cell enters a state, the amount of time that it will remain in that state is drawn from the dwell time distribution.

Each cell in the system consists of a set of automata, a set of attributes, a set of schedules (one of which is active at any time), and a scenario, represented by a set of triggers and associated actions. For efficiency, there is only one copy of each automata, and each cell records its current state and its next automatic transition time. In our implementation we consider two types of interactions: *i) pairwise interactions* including cell-cell and cell-bacteria interactions and *ii) group-agent* interaction. Each such interaction results in a potential state change and is further described below.

Pairwise interaction: In cell-cell interaction the outcome is determined by a pairwise state transition function between two cells. Though a cell may come to contact with many adjacent cells, only one will interact with it. If a cell $u \in V$ interacts with cells $v_1, v_2 \dots v_k$, then there are k pairwise transition functions. In our implementation we choose one transition with a prespecified probability and apply the PTTS to get the final outcome. Future extensions will consider more sophisticated schemes inspired work on chemical reactions and Gillispe’s algorithm and its extensions. The automata in Figure 3 provide pictorial details of the possible cell-cell interactions and the resulting change of cell states.

Group-agent interaction: In group-agent interaction the outcome of a cell is dependent on all the interacting neighbor cells. There are two group of cells that can interact with a cell. One group of cells promote state transition named *activators*. Other group inhibits state transition called *inhibitors*. The probability that a transition will take place is defined by the probability:

$$p = \left(\frac{aA}{aA + iI} \right)^y \quad (1)$$

A is the total number of neighbors in a state that induces a state change (activators) and I is the total number of neighbors in a state that inhibits a state change (inhibitors). The variables a , i , and y are model parameters.

During a single 6 hour period, an agent can undergo no more than one state change. This can be caused by either a pairwise interaction or a group-agent interaction. Earliest interaction gets precedence in the event more than one such interactions are feasible. Note that the interaction may not lead to a state change; this depends on the current state that the agent is in. The rationale for this decision is discussed in the later sections and this issue needs further investigation.

IV. IMPLEMENTATION

The ENISI simulator is based on a previous simulator called EpiSimdemics [?], [20], [21], designed for general purpose contagion diffusion problems [22], [23]. It is implemented in C++ and uses the Message Passing Interface

(MPI) for interprocess communication. EpiSimdemics required several general purpose extensions to support the ENISI model. We will first briefly describe the EpiSimdemics algorithm, and then cover the extensions.

A. EpiSimdemics

The computation structure of implementation consists of three main components: cells, locations, and message brokers. We assume a parallel system consisting of N cores, or processing elements (PEs). Processing proceeds in the following manner:

Partitioning: Cells and locations are partitioned into N groups denoted by C_1, C_2, \dots, C_N and L_1, L_2, \dots, L_N respectively. Currently the distribution is done in a round-robin fashion to allow even load balancing and simpler data management. More complex data distribution and load balancing schemes are being considered, Each PE also creates a copy of the message broker, denoted by MB_1, MB_2, \dots, MB_N . Each PE then executes the ENISI algorithm, shown in Figure 5 on its local data set (C_i, L_i) .

Computing Visit Data: The first phase of the algorithm consists of computing a set of locations to visit for each cell, c_i , for the iteration according to the assigned schedule. A light-weight “copy” of each cell (called a *visit message*) is then sent to each location (which may be on a different PE) via the local message broker. The message broker coalesces multiple small messages destined for a particular remote process into one large message in order to reduce network traffic. The maximum size of the coalesced message can be controlled with the buffer size parameter.

Computing Interactions: Each location receives the visit messages and forms a serial discrete event simulation (DES) by collecting the messages into a time-ordered list of arrive and depart events. Using this data, each location computes interactions for each cell at that location.

Whether a cell u interacts with other cells when co-located in time and space in a sublocation is determined by a probability calculated by one of two types functions: *i) cell-cell* interaction *ii) group-cell* interaction.

Collecting Interaction Messages: At the end of each iteration, interaction messages for each cell on a PE are processed and the resulting state of each cell automaton is updated according to its type-specific transition function. If a cell c_j received a message it then probabilistically transitions from its current state to one of the next states in the PTTS, described below.

All the PEs in the system are synchronized after each simulation phase above. This guarantees that each location has received all the data required to form a DES and each cell has all the data needed to compute its new state.

B. ENISI implementation and semantic approximations

We discuss important model approximations that we made to improve the overall efficiency of ENISI. We also discuss

```

1: initialize();                                ▷ partition data across PEs
2: for  $t = 0$  to  $T$  increasing by  $\Delta t$  do
3:   foreach cell  $p_j \in P_i$  do                ▷ send visits to location PEs
4:      $Visits_j \leftarrow \text{computeVisits}(j, t \text{ to } t + \Delta t)$ ;
5:      $MB_i.\text{sendVisits}(Visits_j)$ ;
6:   end for
7:    $Visits \leftarrow MB_i.\text{retrieveMessages}()$ ;
8:   synchronize();
9:   foreach location  $l_k \in L_i$  do          ▷ compose a serial DES
10:     $\text{makeEvents}(k, Visits)$ ;             ▷ turn visit data into events
11:     $Outcomes \leftarrow \text{computeInteractions}(k)$ ;    ▷ Process
Events
12:     $MB_i.\text{sendOutcomes}(Outcomes)$ ;
13:  end for
14:   $Interactions \leftarrow MB_i.\text{retrieveMessages}()$ ;
15:  synchronize();
16:  foreach  $j \in P_i$  do                      ▷ combine outcomes of interactions
17:     $\text{updateState}(Interactions_j)$ ;
18:  end for
19: end for

```

Figure 5. Parallel version of the EpiSimdemics algorithm

the extensions we had to make to EpiSimdemics to represent problem specific interactions and semantics.

Semantic Approximations The contact-dependency of state transitions in the graphical framework, as well as the need for computational efficiency, require a number of approximations to the biological model. The CGDDS model stipulates that for a state change in one cell to be induced by another cell, the cells must be co-located. Hence, the model cannot explicitly include induction of state-transitions across location barriers as may occur when cytokines secreted by a cell in the LP influence migration of cells in the blood. To reduce complexity, cells are not newly created or removed from the contact network G following the start of the simulation. Rather, biological processes that require these functions are either not included or represented in an indirect fashion. For example, the model does not include the constitutive flow of resting immune cells in and out of tissue. Nor do we represent bacterial replication. The latter approximation can be interpreted as the assumption that each bacterium in contact with the epithelial barrier will be rapidly removed by immune cells before it is able to replicate.

Additionally, the synchronous update at the end of each iteration implies that any changes in behavior that result from the state transition do not take place until the next iteration. As each iteration represents six simulated hours, the synchronous update assumes a six hour delay between a cell receiving the signal to differentiate and actual expression of cytokines or movement-mediating factors, such as integrins, that will affect subsequent movement, contacts, and effects on neighboring cells. The length of the iteration is a trade-off between computational efficiency and model granularity, and can be tuned as any other model parameter.

Given these model approximations, we describe how the following biological functions are represented in the ENISI

implementation.

Bacterial death: As scaling is a constraint, only those bacteria in contact with the epithelial border are represented. Given these simplifications, bacteria in the lumen does not explicitly 'die', but rather it is assumed that when one commensal bacterium is removed by phagocytosis, another bacterium, immediately takes its place due to the high concentration in the outer lumen.

Cross-barrier recruitment: A key function of pro-inflammatory epithelial cells, M1, and eDC is secretion of MCP-1, a factor that recruits resting T-cells as well as resting DC and macrophage precursors, called monocytes, from the blood to the inflamed LP tissue. The model stipulates that any state transition dependent on the state of another cell be contact-dependent and defined as an explicit interaction. Hence the function of recruitment of monocyte and memory T cells in the blood by M1 and eDC in the LP is represented as follows: Cells in the *M1* or *eDC* state briefly migrate to a sublocation in the *Blood* location where they contact cells in the *MASource*, *DCSource*, or *memT* states. This induces the contacted monocyte or memory T cell to transition to an *M0*, *iDC*, or *memT* in the *LP*. Upon this transition, the cells are assigned a new schedule with locations in the *LP*.

T cell death: In the true mucosa, when T cells are no longer active a fraction revert to a resting memory T cell state and the rest undergo programmed cell death. To conserve the number of represented cells in the model, when individual T cells undergo programmed cell death they do not enter a *dead* state. Rather, they replenish the *ThSource* and *iTregSource* population pools.

C. ENISI specific extensions to EpiSimdemics

ENISI is based on the EpiSimdemics simulator, which was originally designed in the context of modeling the spread of infectious disease over large social contact networks. EpiSimdemics has proven to be very general and able to simulate a variety of other contagion diffusion problems, including wireless malware in computer systems [24], and now cell based interactions. Several extensions to the original EpiSimdemics implementation were required to support ENISI.

The iteration length for EpiSimdemics was fixed at 24 hours, and each person had a corresponding 24 hour schedule that repeated daily. In order to support the faster dynamics of cell-cell interaction, the iteration length was changed to be an input parameter (six hours for the work reported in this paper). The iteration length is a trade off between efficiency and simulation granularity. The schedule length was made independent of the iteration length, so that multi-iteration schedules are allowed, with each cell potentially having a different length schedule. When the end of a schedule is reached, it continues again from the start.

EpiSimdemics has the concept of locations and sublocations (i.e., rooms) within a location. A schedule may

specify a location and either a fixed sublocation or a random sublocation within a location. Random sublocations are chosen each time a person visits a location. This restricts all of the sublocations for a given location to the same PE. In the initial ENISI implementation, tissues were represented by locations, with sublocations representing a completely mixed portion of the tissue. This led to scaling problems, as the maximum number of PEs used was limited by the number of tissue types. To remedy this, the concept of randomly chosen locations was added. Each tissue is comprised of multiple locations, each of which has multiple sublocations. A schedule may specify a random location within a particular tissue, and a fixed or random sublocation within that location. The locations that make up a tissue type may be distributed across multiple PEs, allowing the model to scale. In addition, in order for system dynamics to be evaluated, the ability to collect counts of cell type per location was added.

EpiSimdemics has a single interaction function for person-person interactions. While this is sufficient to model some of the cell level interactions, an additional type of group interaction is needed. In this interaction, there are three classes of interactors that are in contact in a sublocation: activators, inhibitors, and targets. These groups need not be mutually exclusive. Each activator increases the likelihood that an interaction will take place, each inhibitor reduces that likelihood, and targets are the potential recipients of the interaction. The total activators and inhibitors present is determined, and the probability of interaction calculated by Equation 1. Then a Bernoulli trial is conducted on each target to determine if they are the recipient of the interaction.

All of these features are intended to be of general use, and are being incorporated into EpiSimdemics as well. They are proving to be of use as we extend the complexity of the behavior and public policy modeling.

D. Model Extensions

The following extensions and refinements were made to the ENISI model presented in [16] for the current version:

- 1) *Three general species of bacteria:* A **tolerogenic bacteria** that induces tDC and M2 phenotypes in antigen-presenting cells, an **inflammatory bacteria** that represents a general pathogen and a general **commensal bacteria** that may have inflammatory and tolerogenic properties, depending on model parameterization.
- 2) *Impaired epithelial cell state:* Epithelial cells may occupy an impaired state, *pEC_noR* in which they do not secrete cell recruiting chemokines. This may be relevant when simulating infection with bacteria that have demonstrated the ability to inhibit this function.
- 3) Cell proliferation and recruitment are localized. If a cell in sublocation *x* proliferates or recruits other immune cells, the nascent daughter cells and newly arrived cells will appear in sublocation *x*.

Table I
DATASET SIZES AND SIMULATION EXECUTION TIMES FOR 250 SIMULATION DAYS. THE EXECUTION TIME (HH:MM:SS) AND NUMBER OF NODES IS GIVEN FOR THE FASTEST RUNNING CONFIGURATION FOR EACH PROBLEM SIZE.

Problem Size	N_{cells}	Execution Time	N_{nodes}	N_{cores}	Message Coalescing Buffer Size	Memory Usage of PE_0 (MB)	Parallel Efficiency	Communication Overhead
10^5	100,038	00:00:41	16	192	10	161	0.24	22.6%
10^6	1,000,594	00:02:26	32	384	10	168	0.73	22.2%
10^7	10,000,110	01:26:48	48	576	1,000	340	0.94	6.5%
10^8	100,001,314	226:39:10	56	672	1,000	814	NA	50.1%

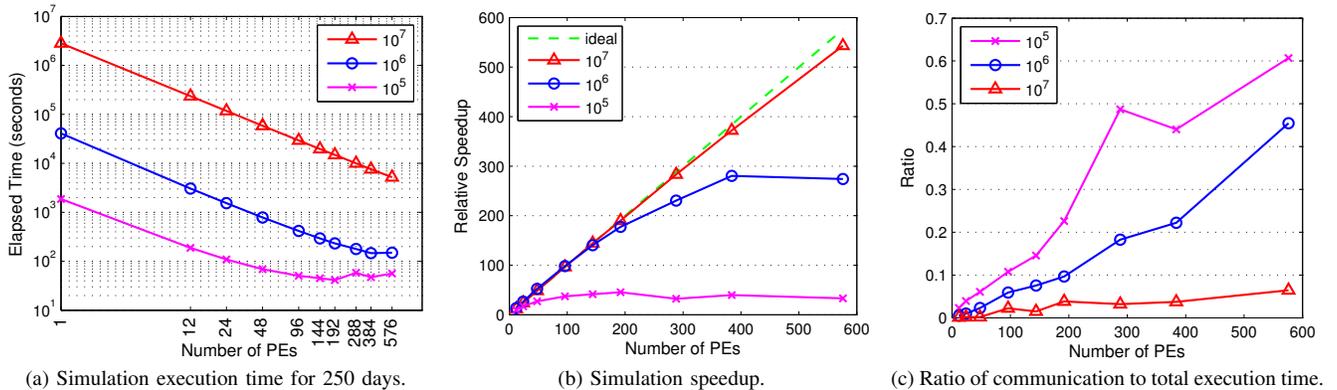


Figure 6. Scalability of ENISI on different sized problems.

- 4) Inflammatory T cells may be either a Th1 or Th17 phenotype, instead of a composite population of Th1 and Th17.
- 5) Cells may migrate between tissue locations independent of state transitions.
- 6) A lymphoid tissue site was added. Previous versions only represented cells in the LP and Lumen.

V. MODEL SCALABILITY

We evaluate the scalability of our model’s implementation against multiple problem sizes ranging from 10^5 cells to 10^7 cells. We simulate 250 days for each test. The high performance computing cluster we use consists of total 62 compute nodes. Each node is comprised of two six-core Intel Xeon X5670 processors and 48 GB of 1333 MHz DDR3 memory. We use MVAPICH2 MPI over a Mellanox 40 Gb/s dual-port QDR InfiniBand interconnect.

Table I lists the dataset sizes used and the fastest simulation execution time achieved for each of the datasets. Figure 6a plots the simulation execution time for different number of processing elements (PEs) for each of the datasets used. Figure 6b shows the speedup achieved for each of the datasets. For the small sized problem with 10^5 cells, the simulation scales only up to 192 PEs, and gains a speedup of up to 45. However, for a medium size problem with 10^6 cells, the scalability extends to more PEs (up to 384 PEs), gaining a speedup of 280. For the large size problem with 10^7 cells, it scales with all the 576 PEs that we used, and achieves a speedup of up to 543 which is 94% of the ideal

speedup, finishing in just 1¹/₂ hours. We believe that the 10^5 and 10^6 problem sizes have too little data at larger core counts, leading to poor parallel efficiency. For an extremely large problem with 10^8 cells, we were unable to evaluate the scaling on the test machine as the network resource became saturated, but can successfully run the simulation at this unprecedented scale. From these results, we expect to be able to continue to scale both the problem size and number of PEs on much larger machines. In addition, there are still several optimizations available in both the ENISI model and the simulator itself that will improve these numbers.

VI. EXAMPLE APPLICATION: PATHOGENICITY FUNCTIONS OF *H. pylori* 26695 CPI

Helicobacter pylori is a bacteria that resides in the mammalian gastric mucosa and infects a large number of the world’s population. In 85% of *H. pylori* infections the bacteria persists as a commensal bacteria in the gastric mucosa with no negative impact on the host. However, in 15% of cases the infection is associated with gastric disease including ulcer formation as persistent inflammation induced lesions in the gastric epithelium. *H. pylori* is a genotypically diverse species and disease is associated with the expression of several virulence factors that participate in various immune modulation mechanisms [25], [26].

The CagA pathogenicity island (cpi) encodes numerous proteins that participate in various immune modulation mechanisms as reviewed in [26]. These include *i*) inhibition of microbicide secretion by macrophages by the protein

arginase RocF, *ii*) an ability to invade the mucous layer gaining direct access to the epithelial barrier through flagellar motility as well as expression of various adhesion factors, *iii*) disruption of epithelial barrier through damage of cell-cell junctions by proteins VacA and HtrA that allows entry in to the LP triggering inflammatory signals, *iv*) triggering of NF- κ B pathway in epithelial cells by protein CagA that leads to expression of chemoattractants and microbicides, as well as *v*) inhibition of inflammatory phenotypes in antigen presenting cells (dendritic cells and macrophages). Prevailing theories state that individuals with different health outcomes are infected by *H. pylori* strains that carry out different combinations of these immunomodulatory capabilities.

In this demonstration of the ENISI tool we set model conditions to recreate experimental infection of individual mice with the cpi-containing *H. pylori* strain 26695 carried out by experimental collaborators. Upon observing immunopathogenesis *in silico* we analyze the contact network to identify specific mechanisms of pathogenicity. The results identify specific processes associated with the infection that do and do not contribute to this simulated pathogenesis.

The *H. pylori* strain 26695 was represented by assigning functions to **commensal bacteria** representing a combination of the immunomodulatory mechanisms mentioned. As default it was assumed to have all possible effects on epithelial cells. That is, we set $\nu_{EB} = \beta_r = \beta_d = \beta_p = \beta_c = 1.0$. We also assume an effector response in antigen presenting cells ($\nu_{BD} = \nu_{BM} = \nu_{Bs} = 1.0$) and set the parameter p_{17} to 0 based on experimental observations of a Th1-dominant response (not shown).

The initial number of individuals in each state is shown in Table II. These were assigned according to experimental measurements gathered by collaborators and to represent a 1 mm³ sample of the mouse gastric mucosa, roughly 1% of the total volume as estimated from images in [27].

With these assumptions and the initial population sizes in Table II simulations with *H. pylori*-representing **commensal bacteria** ('infected') and with out **commensal bacteria** ('control') were carried out in replicates of seven. Specifically, infection was simulated by adding **commensal bacteria** on day 2 and following the state changes and migration of cells over 63 days. The model was fit to qualitative trends observed among tissue samples of mice experimentally infected with *H. pylori* strain 26695, compared to the control group, gathered by collaborators. In these experiments both LP and GLN tissue samples were taken from infected and control mice on days 7, 14, 30, and 60 post-infection (p.i.) and the count of cells of each regulatory/inflammatory phenotype was determined through flow cytometry. Figure 7 depicts the dynamics of specific cell populations over the course of seven simulated 63 day infections that reproduce experimental observations. Specifically *i*) there is a nearly undetectable level of *H. pylori*-specific immune response before day 30 p.i., when regulatory phenotypes, iTreg and

Table II
INITIAL POPULATIONS IN GASTRIC MUCOSA

State	Description	Initial Number
Phenotypes		
<i>restingT</i>	Resting conventional CD4+ T cell	$1 \cdot 10^3$
<i>iDCs</i>	Immature 'sampling' DC in the superficial LP with access to the Lumen	1000
<i>iDCLP</i>	Immature dendritic cell in the LP	1000
<i>M0</i>	Undifferentiated macrophage	$1 \cdot 10^3$
<i>EC</i>	Healthy epithelial cell	10^5
<i>MASource</i>	monocytes: MA precursor	10^5
<i>DCSource</i>	monocytes:DC precursor	10^5
<i>TSource</i>	resting T cell in blood	10^4
<i>Th1Source</i>	Potential child cell from a proliferating Th1	$5 \cdot 10^5$
<i>Th17Source</i>	Potential child cell from a proliferating Th17	$5 \cdot 10^5$
<i>iTregSource</i>	Potential child cell from a proliferating iTreg	$5 \cdot 10^5$
Locations		
<i>CommB_lumen</i>	Commensal bacterium in the lumen	10
<i>InfB_lumen</i>	Inflammatory bacterium in the lumen	0
<i>TolB_lumen</i>	Tolerogenic bacterium in the lumen	1000

M2 dominate the system *ii*) by day 60, not prior, there is a statistically significant increase in M2 macrophages in the infected group over the control group, but not an increase in M1, *iii*) by day 60 p.i., but not prior, there is a statistically significant increase in both effector and tolerogenic dendritic cells in the GLN, *iv*) by day 60 there is a significant increase in active T cells in the LP in the infected group, with Th1 dominating Th17 on average. One can see that this rise in T cells occurs in conjunction with a rise in effector dendritic cell levels in the LP (Figure 7b) and a low level of M1 emergence (Figure 7c). This increase in immune activity is associated with mounting epithelial damage, (Figure 7d) represented by transition of **epithelial cell** automata from the *EC* state to the *pEC* state and from the *pEC* state to the *E* state. To identify the pathways by which this mounting immune response is associated with tissue damage we focus on the simulation replicate, replicate 5, that resulted in the greatest epithelial damage and identified the states of neighbors that induced key health state-defining state changes.

It was found that, in replicate 5, for all individuals that undergo the transition *E* cell \rightarrow *pEC* cell, the transition is primarily induced by *H. pylori* directly in the earlier stages of infection (days 1-30) as shown in Figure 8a. However, it can be clearly seen that in the chronic stage (days 30-63 post-infection), when pathogenesis is seen, epithelial damage is occurring through contact with IFN γ -secreting Th1, with less contribution by IFN γ -secreting Th17 (Figure 8a). Indeed, for all individuals that undergo the transition *pEC* cell \rightarrow *E* state, this occurred only upon

contact with neighbors in the *Th1* state (not shown). Notably macrophages and dendritic cells are not contributing directly to epithelial damage in this first two months of infection.

T cells may be stimulated by macrophages, dendritic cells in the LP, or ‘sampling’ dendritic cells in the lumen. It was found that all individuals that stimulated T cells to a Th1 phenotype were in the *eDCL* state, effector ‘sampling’ dendritic cells, indicating ‘sampling’ dendritic cells are solely responsible for T cell stimulation over the entire 63 day infection in this case. As the count of individuals in the *eDCL* state remains relatively constant over the course of infection (Figure 7b), the only explanation for the continued rise in Th1 (Figure 7a) is an increase in resting T cells in the LP that are being recruited to the infection site and subsequently stimulated. To identify which cells were responsible for this recruitment Figure 8b depicts the number of individuals in each state that induce the transitions of *TSource* \rightarrow *restingT*. It can be seen that pro-inflammatory epithelial cells, stimulated by *H. pylori* and Th1 to secrete chemoattractants and effector dendritic cells also stimulated by *H. pylori*, are equal contributors to resting T cell recruitment. This identifies a positive feedback loop that is fed by *H. pylori* presence in which *H. pylori*-damaged EC secrete chemoattractants that recruit resting T cells. These are subsequently stimulated to Th1, which contribute to further epithelial damage leading to increased chemoattractant secretion and T cell recruitment. The indication is that ulcerative inflammation may be treated with chemicals such as glycerol monolaurate known to reduce secretion of lymphocyte chemoattracts such as MIP3 α and has been proposed for intervention in other inflammation-enhanced diseases such as SIV [28].

Though *H. pylori* appears to be contributing to epithelial damage and the parameter set maximizes its ability to degrade epithelial tight junctions ($\beta_p = 1.0$), in replicate 5 no individuals representing *H. pylori* (**commensal bacteria** automata) are seen in the LP. This was explained by finding that the majority of **commensal bacteria** automata that transition to the *CommB_Dead* state, transition when in contact with neighbors in the microbicide-secreting *pECell* state. This indicates that, in this *in silico* scenario, even though *H. pylori* is effective in degrading epithelial tight junctions, the fact that it also induces secretion of microbicides, such as defensin through the $\text{NF}\kappa\beta$ pathway, ensure that it is eliminated prior to migration in to the LP canceling out any benefit that may come from this cpi-encoded function.

Of the processes associated with the cpi, this analysis of simulated *H. pylori* 26695 infection indicates that pathogenesis seen in the first 2 months of infection is not due to inhibition of microbicide secretion by macrophages by the protein arginase RocF as M1 has no effect on removing *H. pylori*. Rather *H. pylori* is removed only by defensin-secreting epithelial cells and phagocytic ‘sam-

pling’ dendritic cells in the lumen. Indeed, in the replicate with the *least* amount of epithelial damage, replicate 7, no individuals enter the *M1* state further demonstrating a lack of significance of macrophage activity is decreasing pathogenicity. With regard to the various modifications of epithelial cells mediated by cpi-encoded proteins, induction of chemoattractant secretion through the $\text{NF}\kappa\beta$ by CagA protein is a key component of the immunopathological feedback loop identified. Where as induction of secretion of inflammatory cytokines IL6, IL21, $\text{TNF}\alpha$ or IL23 seems to have a minor effect as the primary role is to induce Th17 and M1 phenotypes, which are not significant contributors to epithelial damage in this stage of infection. The relevance of direct contact between *H. pylori* and epithelial cells through invasion of the mucous layer and expression of various adhesion factors is not clear from this analysis. Though contact is necessary to induce chemoattractant secretion by epithelial cells, factors secreted by effector dendritic cells may be sufficient to start the positive feedback loop of resting T cell recruitment and epithelial damage. Disruption of the epithelial barrier through damage of cell-cell junctions by proteins VacA and HtrA that potentially allows entry in to the LP does not appear significant as *H. pylori* does not enter the LP prior to elimination by defensin secreted from epithelial cells. Hence, the primary mechanism of pathogenesis observed in this sample is the induction of the inflammatory pathway in dendritic cells in the lumen. The direct induction of chemoattractant secretion by epithelial cells may also be necessary, but further simulations would need to be carried out to demonstrate necessity as T cell recruitment by effector dendritic cells may be sufficient. What is clear is that *inhibition* of inflammatory phenotypes in antigen presenting cells (dendritic cells and macrophages) does not likely account for pathogenesis associated with *H. pylori* strain 26695.

In conclusion, this demonstrates how the ability to analyze an infection recreated *in silico* can inform hypothesis for the sequence of events occurring in a real infection. In this manner we provide an explanation for the delayed immunopathogenesis observed experimentally.

Table III
INTERACTIONS CALCULATED BY CLINICAL OUTCOME

No infection	5004.8	0 (488,368.56)
Infected, least inflammation	7534.92	399.8 (515,723.04)
Infected, most inflammation	13521.28	2922.92 (625,092)
Infected, middle inflammation	10327.45	978 (578,316)

VII. RELEVANCE AND FUTURE DIRECTIONS

Aspects of the presented inflammatory and regulatory immune pathways have been represented in previous models of mucosal infection [4]–[7]. The ENISI model is unique in its scope and approach. The model incorporates regulatory

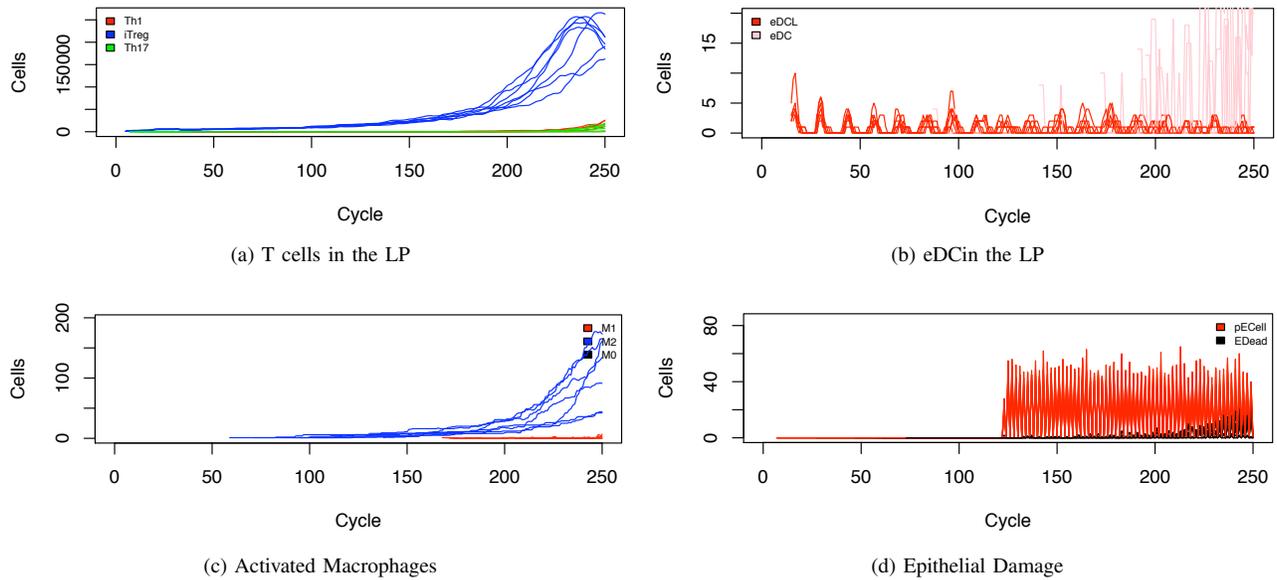


Figure 7. *H. pylori* 26695 infection dynamics Dynamics of cell populations over a period of 62 days in the presence of *H. pylori* strain 26695

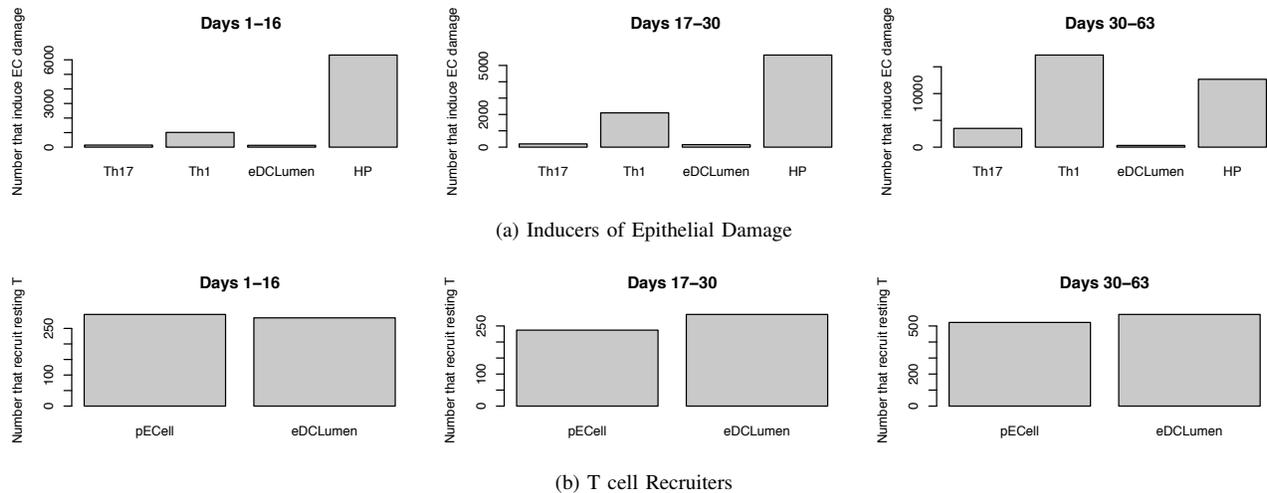


Figure 8. **Epithelial damage and T cell recruitment.** (a) the counts of individuals in each state that induce epithelial damage over simulated infection period. (b) the counts of individuals in each that recruit a T cell to the LP during the course of the simulation infection.

mechanisms of both adaptive and innate immunity, multi-location migration of cells, and cross talk between antigen presenting cells and T cells.

The current ENISI modeling system is really a first step towards a system that will be useful for immunologists, bioinformaticists and infectious disease experts in their everyday work. We are currently investigating the following issues to make ENISI usable by subject matter experts: (i) increasing the breadth and depth of the ENISI model, (ii) increasing the performance by optimizing both the model and the underlying simulator, (iii) integrating the modeling system into an easy to use, web based cyberinfrastructure to allow access to detailed models, high performance com-

puting and storage systems, (iv) a digital library of past simulation results in a platform that promotes collaboration among researchers and (v) sensitivity analysis and quantifying uncertainty. Future efforts also will refine immune cell movement in the lamina propria and provide visualizations of lesion formation and healing at the site of infection.

ACKNOWLEDGMENT

We thank our external collaborators and members of the Network Dynamics and Simulation Science Laboratory (NDSSL) and the Nutritional Immunology and Molecular Medicine Laboratory (NIMM) for their suggestions and comments. This work has been partially supported by DTRA - R&D (HPC) Award No. HDTRA1-09-1-0017, DTRA - Validation Award No.

HDTRA1-11-1-0016, DTRA - CNIMS Award No. HDRTA1-11-D-0016-0001, NSF PetaApps Grant OCI-0904844, NSF Netse Grant CNS-1011769, NSF SDCI Grant OCI-1032677, NIH MIDAS project 2U01GM070694-7 and NIAID & NIH project HHSN272201000056C.

REFERENCES

- [1] S. Gordon and P. R. Taylor, "Monocyte and macrophage heterogeneity," *Nat Rev Immunol*, vol. 5, no. 12, pp. 953–64, 2005.
- [2] A. Iwasaki, "Mucosal dendritic cells," *Annual Review of Immunology*, vol. 25, pp. 381–418, 2007.
- [3] (2012) Modeling h. pylori using enisi and cell designer. [Online]. Available: <http://www.modelingimmunity.org/models/enisi-helicobacter-pylori/>
- [4] J. C. Arciero, G. B. Ermentrout, J. S. Upperman, Y. Vodovotz, and J. E. Rubin, "Using a mathematical model to analyze the role of probiotics and inflammation in necrotizing enterocolitis," *PLoS One*, vol. 5, no. 4, p. e10066, 2010.
- [5] J. E. Wigginton and D. Kirschner, "A model to predict cell-mediated immune regulatory mechanisms during human infection with mycobacterium tuberculosis," *J Immunol*, vol. 166, no. 3, pp. 1951–67, 2001.
- [6] M. J. Blaser and D. Kirschner, "Dynamics of helicobacter pylori colonization in relation to the host response," *Proc Natl Acad Sci USA*, vol. 96, no. 15, pp. 8359–64, 1999.
- [7] M. Blaser and D. Kirschner, "The equilibria that allow bacterial persistence in human hosts," *Nature*, vol. 449, no. 7164, pp. 843–9, 2007.
- [8] K. Wendelsdorf, J. Bassaganya-Riera, R. Hontecillas, and S. Eubank, "Model of colonic inflammation: Immune modulatory mechanisms in inflammatory bowel disease," *Journal of Theoretical Biology*, 2010.
- [9] P. A. Bauer A., Beauchemin C., "Agent-based modeling in host-pathogen systems: The successes and challenges," *Information Sciences*, vol. 179, pp. 1379–1389, 2009.
- [10] A. T. Haase, "Population biology of HIV-1 infection: Viral and CD4(+) t cell demographics and dynamics in lymphatic tissues," *Annual Review of Immunology*, vol. 17, pp. 625–656, 1999.
- [11] C. I. Efroni S., Harel D., "Reactive animation: Realistic modeling of complex dynamic systems," *Computer*, pp. 38–47, 2005.
- [12] N. Swerdlin, I. R. Cohen, and D. Harel, "The lymph node b cell immune response: Dynamic analysis in-silico," *Proceedings of the IEEE*, vol. 96, no. 8, pp. 1421–1443, 2008.
- [13] M. W. Sneddon, J. R. Faeder, and T. Emonet, "Efficient modeling, simulation and coarse-graining of biological complexity with nfm," *Nature Methods*, vol. 8, no. 2, pp. 177–U112, 2011.
- [14] V. A. Folcik, G. C. An, and C. G. Orosz, "The basic immune simulator: an agent-based model to study the interactions between innate and adaptive immunity," *Theor Biol Med Model*, vol. 4, p. 39, 2007.
- [15] T. Sutterlin, S. Huber, H. Dickhaus, and N. Grabe, "Modeling multi-cellular behavior in epidermal tissue homeostasis via finite state machines in multi-agent systems," *Bioinformatics*, vol. 25, no. 16, pp. 2057–2063, 2009.
- [16] K. Wendelsdorf, J. Bassaganya-Riera, K. Bisset, S. Eubank, R. Hontecillas, and M. Marathe, "Enteric immunity simulator: A tool for in silico study of gut immunopathologies," *Proceedings of the IEEE International Conference Bioinformatics and Biomedicine*, 2011.
- [17] D. Gery and E. Harel, "Executable object modeling with statecharts," *Computer*, vol. 30, no. 7, pp. 31–42, 1997.
- [18] K. Wendelsdorf, J. Bassaganya-Riera, K. Bisset, S. Eubank, R. Hontecillas, and M. Marathe, "Enteric immunity simulator version 0.9: A tool for in silico study of gut immunopathologies," NDSSSL, Tech. Rep., 2011. [Online]. Available: <http://modelingimmunity.vbi.vt.edu/modeling/enisi/>
- [19] C. L. Barrett, K. Bisset, J. Chen, S. Eubank, B. Lewis, V. S. A. Kumar, M. V. Marathe, and H. S. Mortveit, "Interactions among human behavior, social networks, and societal infrastructures: A case study in computational epidemiology," in *Fundamental Problems in Computing*, S. S. Ravi and S. K. Shukla, Eds. Springer Netherlands, 2009, pp. 477–507.
- [20] C. L. Barrett, K. R. Bisset, S. G. Eubank, X. Feng, and M. V. Marathe, "Episimdemics: an efficient algorithm for simulating the spread of infectious disease over large realistic social networks," in *SC '08: ACM/IEEE Conference on Supercomputing*. Piscataway, NJ, USA: IEEE Press, 2008, pp. 1–12.
- [21] K. Bisset, X. Feng, M. Marathe, and S. Yardi, "Modeling interaction between individuals, social networks and public policy to support public health epidemiology," in *Proceedings of the 2009 Winter Simulation Conference*, Dec. 2009, pp. 2020–2031.
- [22] C. Barrett, K. Bisset, S. Eubank, M. Marathe, V. S. A. Kumar, and H. Mortveit, "Modeling and simulation of large biological, information and socio-technical systems: An interaction-based approach," in *Short Course on Modeling and Simulation of Biological Networks*, R. Laubenbacher, Ed. AMS, Jan. 2007, pp. 101–147.
- [23] C. L. Barrett, K. R. Bisset, J. Leidig, A. Marathe, and M. V. Marathe, "An integrated modeling environment to study the co-evolution of networks, individual behavior and epidemics," *AI Magazine*, vol. 31, no. 1, pp. 75–87, 2010.
- [24] K. Channakeshava, D. Chafekar, K. R. Bisset, V. S. A. Kumar, and M. V. Marathe, "EpiNet: A simulation framework to study the spread of malware in wireless networks," in *2nd International Conference on Simulation Tools and Techniques for Communications, Networks and Systems, (SimuTools) 2009*, O. Dalle, G. A. Wainer, L. F. Perrone, and G. Stea, Eds. ICST Press, Mar. 2–6 2009.
- [25] D. B. Polk and R. M. Peek, "Helicobacter pylori: gastric cancer and beyond," *Nature Reviews Cancer*, vol. 10, no. 8, pp. 593–593, 2010.
- [26] S. Backert, N. Tegtmeyer, and S. Wessler, "Role of the cag-pathogenicity island encoded type iv secretion system in helicobacter pylori pathogenesis," *Febs Journal*, vol. 278, no. 8, pp. 1190–1202, 2011.
- [27] K. A. Eaton, M. J. Radin, and S. Krakowka, "An animal model of gastric ulcer due to bacterial gastritis in mice," *Vet Pathol*, vol. 32, no. 5, pp. 489–97, 1995.
- [28] Q. Li, J. D. Estes, P. M. Schlievert, L. Duan, A. J. Brosnahan, P. J. Southern, C. S. Reilly, M. L. Peterson, N. Schultz-Darken, K. G. Brunner, K. R. Nephew, S. Pambuccian, J. D. Lifson, J. V. Carlis, and A. T. Haase, "Glycerol monolaurate prevents mucosal SIV transmission," *Nature*, 2009.