

<b>Manuscript Number:</b>	GIGA-D-18-00435R2	
<b>Full Title:</b>	High-Resolution Computational Modeling of Immune Responses in the Gut	
<b>Article Type:</b>	Research	
<b>Funding Information:</b>	Defense Threat Reduction Agency (HDTRA1-18-1-0008)	Dr. Josep Bassaganya-Riera Dr. Raquel Hontecillas
<b>Abstract:</b>	<p>Background: <i>Helicobacter pylori</i> causes gastric cancer in 1-2% of cases, but is also beneficial for protection against allergies and gastroesophageal diseases. An estimated 85% of <i>H. pylori</i>-colonized individuals do not present any detrimental effects. To study the mechanisms promoting host tolerance to the bacterium in the gastrointestinal mucosa and systemic regulatory effects, we investigated the dynamics of immunoregulatory mechanisms triggered by <i>H. pylori</i> using a high-performance computing driven ENteric Immunity Simulator multiscale model. Immune responses were simulated by integrating an agent-based model, ordinary and partial differential equations.</p> <p>Results: The outputs were analyzed using two sequential stages: the first used a partial rank correlation coefficient regression-based and the second employed a metamodel-based global sensitivity analysis. The influential parameters screened from the first stage were selected to be varied for the second stage. The outputs from both stages were combined as a training dataset to build a spatiotemporal metamodel. The Sobol' indices measured time-varying impact of input parameters during initiation, peak and chronic phases of infection. The study identified epithelial cell proliferation and epithelial cell death as key parameters that control infection outcomes. In-silico validation showed that colonization with <i>H. pylori</i> decreased with a decrease in epithelial cell proliferation, which was linked to regulatory macrophages and tolerogenic dendritic cells.</p> <p>Conclusion: The hybrid model of <i>H. pylori</i> infection identified epithelial cell proliferation as a key factor for successful colonization of the gastric niche and highlighted the role of tolerogenic dendritic cells and regulatory macrophages in modulating the host responses and shaping infection outcomes.</p>	
<b>Corresponding Author:</b>	Raquel Hontecillas, Ph.D. Virginia Polytechnic Institute and State University Blacksburg, Virginia UNITED STATES	
<b>Corresponding Author Secondary Information:</b>		
<b>Corresponding Author's Institution:</b>	Virginia Polytechnic Institute and State University	
<b>Corresponding Author's Secondary Institution:</b>		
<b>First Author:</b>	Meghna Verma, M.S.	
<b>First Author Secondary Information:</b>		
<b>Order of Authors:</b>	Meghna Verma, M.S.	
	Josep Bassaganya-Riera, Ph.D.	
	Andrew Leber, Ph.D.	
	Nuria Tubau-Juni, M.S.	
	Stefan Hoops, Ph.D.	
	Vida Abedi, Ph.D.	
	Xi Chen, Ph.D.	
	Raquel Hontecillas, Ph.D.	

<b>Order of Authors Secondary Information:</b>	
<b>Response to Reviewers:</b>	<p>Point by point response to the Reviewer reports</p> <p>We would like to thank the reviewers and editors for taking time to review our manuscript entitled “High-Resolution Computational Modeling of Immune Responses in the Gut” and for providing valuable and constructive criticism. The review process has been helpful in the improvement of our submission. We have considered the comments that were made and have prepared the following point-by-point response. We hope that the revised version of the manuscript can now be accepted for publication. Thanks in advance.</p> <p>Reviewer #1: The unit of sizes of the model grid can't be right (e.g. grid is 30nm x 10 nm). Animal cells should have measurements in the order of micrometres instead of nanometres. Please check if these are just typos, or do these errors affect any aspect of the simulation, such as diffusion.  Response: We thank the reviewer for pointing this out. We fixed the typos and the unit size of the model grid are 30 µm x 10 µm. These typos do not affect any aspect of the simulations as these units are only annotations and the model takes the numbers as input. We updated the manuscript and fixed the typos throughout the manuscript. Please refer to L120 – L121, and L216, L220-L221.</p> <p>Reviewer #2: The authors have made significant improvements to the manuscript and thoroughly responded to reviewer comments. One major concern remains surrounding the authors' response to questions around the grid dimensions. The dimensions for the entire grid are given in nm which is smaller than a single cell. Furthermore they state that there are no limits to cell(agent) occupancy per grid compartment. This is rather confusing and calls into question how much spatial information is really contained in this model (e.g. if cytokines are diffusing over the 30nm grid what does that mean for the concentrations that individual cells (measured in micrometers)are seeing?). Based on the author responses it appears that the model is a multi-compartment model with well-mixed discrete agents in each compartment rather than a spatio-temporal model as they claim.  Response: We thank the reviewer for their comment.  We thank the reviewer for pointing out the concern regarding the dimensions of the grid. The correct dimensions of the grid are 30 µm x 10 µm. We updated the manuscript and fixed the typos. Please refer to L120 – L121, and L216, L220-L221. The mention regarding no limits to cell (agent) occupancy refers to the cells (agents) having no physical size. Further, once a cell (agent) dies it is removed from the simulation to minimize the computational costs of agents that do not contribute to the biology.  The model output contains information about the x and y co-ordinate of the agents at every time point. The cytokines and internal signaling pathways that drive functional fates of cells are well mixed within a cell, i.e., we have only temporal resolution within the cell during a time step. However, the production, degradation, and diffusions are cell specific thus the cytokine concentration results are also spatio-temporal. Since, the model is capable of providing information regarding spatial co-ordinates over time, we claim the model to be a spatio-temporal model. We updated the manuscript, please refer to L163-L170.</p> <p>Please also ensure that your revised manuscript conforms to the journal style, which can be found in the Instructions for Authors on the journal homepage.  Response: The revised manuscript conforms to the journal style.</p>
<b>Additional Information:</b>	
<b>Question</b>	<b>Response</b>
Are you submitting this manuscript to a special series or article collection?	No
<b>Experimental design and statistics</b>	Yes
Full details of the experimental design and	

<p>statistical methods used should be given in the Methods section, as detailed in our <a href="#">Minimum Standards Reporting Checklist</a>. Information essential to interpreting the data presented should be made available in the figure legends.</p> <p>Have you included all the information requested in your manuscript?</p>	
<p><b>Resources</b></p> <p>A description of all resources used, including antibodies, cell lines, animals and software tools, with enough information to allow them to be uniquely identified, should be included in the Methods section. Authors are strongly encouraged to cite <a href="#">Research Resource Identifiers</a> (RRIDs) for antibodies, model organisms and tools, where possible.</p> <p>Have you included the information requested as detailed in our <a href="#">Minimum Standards Reporting Checklist</a>?</p>	<p>Yes</p>
<p><b>Availability of data and materials</b></p> <p>All datasets and code on which the conclusions of the paper rely must be either included in your submission or deposited in <a href="#">publicly available repositories</a> (where available and ethically appropriate), referencing such data using a unique identifier in the references and in the “Availability of Data and Materials” section of your manuscript.</p> <p>Have you have met the above requirement as detailed in our <a href="#">Minimum Standards Reporting Checklist</a>?</p>	<p>Yes</p>

[Click here to view linked References](#)

## High-Resolution Computational Modeling of Immune Responses in the Gut

Meghna Verma<sup>1,2</sup>, Josep Bassaganya-Riera<sup>1</sup>, Andrew Leber<sup>1</sup>, Nuria Tubau-Juni<sup>1</sup>,  
Stefan Hoops<sup>1</sup>, Vida Abedi<sup>1</sup>, Xi Chen<sup>3</sup>, Raquel Hontecillas<sup>1,\*</sup>

<sup>1</sup>Nutritional Immunology and Molecular Medicine Laboratory, Biocomplexity Institute of Virginia Tech, Blacksburg, VA 24060, USA.

<sup>2</sup>Graduate Program in Translational Biology, Medicine and Health, Virginia Tech, Blacksburg, VA, 24061, USA.

<sup>3</sup>Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, VA, USA.

### \* Correspondence:

Dr. Raquel Hontecillas

Email: [rmagarzo@vt.edu](mailto:rmagarzo@vt.edu)

**Keywords:** agent-based model, ordinary differential equation, Gaussian process, *Helicobacter pylori*, high-performance computing, metamodel, sensitivity analysis, immune system, dendritic cells, macrophages.

Email address for all authors:

Meghna Verma: [meghna89@vt.edu](mailto:meghna89@vt.edu)

Josep Bassaganya-Riera: [jbassaga@vt.edu](mailto:jbassaga@vt.edu)

Andrew Leber: [ajleber@vt.edu](mailto:ajleber@vt.edu)

Nuria Tubau-Juni: [nuriaj@vt.edu](mailto:nuriaj@vt.edu)

Stefan Hoops: [shoops@vt.edu](mailto:shoops@vt.edu)

Vida Abedi: [vidaabedi@gmail.com](mailto:vidaabedi@gmail.com)

Xi Chen: [xchen6@vt.edu](mailto:xchen6@vt.edu)

Raquel Hontecillas: [rmagarzo@vt.edu](mailto:rmagarzo@vt.edu)

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1     **Abstract**

2     **Background:** *Helicobacter pylori* causes gastric cancer in 1-2% of cases, but is  
3     also beneficial for protection against allergies and gastroesophageal diseases.  
4     An estimated 85% of *H. pylori*-colonized individuals do not present any  
5     detrimental effects. To study the mechanisms promoting host tolerance to the  
6     bacterium in the gastrointestinal mucosa and systemic regulatory effects, we  
7     investigated the dynamics of immunoregulatory mechanisms triggered by *H.*  
8     *pylori* using a high-performance computing driven **EN**teric **I**mmunity **S**imulator  
9     multiscale model. Immune responses were simulated by integrating an agent-  
10    based model, ordinary and partial differential equations.

11    **Results:** The outputs were analyzed using two sequential stages: the first used  
12    a partial rank correlation coefficient regression-based and the second employed  
13    a metamodel-based global sensitivity analysis. The influential parameters  
14    screened from the first stage were selected to be varied for the second stage.  
15    The outputs from both stages were combined as a training dataset to build a  
16    spatiotemporal metamodel. The Sobol' indices measured time-varying impact of  
17    input parameters during initiation, peak and chronic phases of infection. The  
18    study identified epithelial cell proliferation and epithelial cell death as key  
19    parameters that control infection outcomes. *In-silico* validation showed that  
20    colonization with *H. pylori* decreased with a decrease in epithelial cell  
21    proliferation, which was linked to regulatory macrophages and tolerogenic  
22    dendritic cells.

23    **Conclusion:** The hybrid model of *H. pylori* infection identified epithelial cell  
24    proliferation as a key factor for successful colonization of the gastric niche and

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

25 highlighted the role of tolerogenic dendritic cells and regulatory macrophages in  
26 modulating the host responses and shaping infection outcomes.

## 27 **1. Background**

28 Computational modeling of the immune response dynamics can provide  
29 novel insights and facilitate the systems level understanding of the interactions  
30 at the gastric mucosa during infection. Ordinary differential equation (ODE-  
31 based methods are deterministic and based on the average response of cells  
32 over time. Dynamical models are used in immunology for system-level analyses  
33 of CD4+ T cell differentiation [1], macrophage differentiation [2], immune  
34 responses elicited by *Clostridium difficile* infection [3], co-infections [4], and in  
35 cancer and immunotherapy [5]. However, ODE-based models lack the spatial  
36 aspects and the features to study the organ and immune cell topology over time.  
37 Agent-based models (ABM) employ a bottom-up approach that focuses on the  
38 spatial and temporal aspects of individual immune cells, unlike the ODE-based  
39 methods. This rule-based method includes agents that act as local entities which  
40 interact locally with other agents, move in space, and follow set of rules  
41 representing their role in a given system and contribute towards generating an  
42 emergent behavior. Since, the immune system is a complex dynamical system  
43 [6] wherein the components *i.e.*, the immune cells move in space and time  
44 changing their location, ABMs are useful tools that can be employed to  
45 understand biological mechanisms and the hidden insights.

46 *Helicobacter pylori* is a gram-negative bacterium that has persistently  
47 colonized the human stomach since early evolution [7] [8] and is currently found  
48 in over 50% [9] of the global population. *H. pylori* has co-evolved with humans  
49 for thousands of years, such that an estimated 85% of the *H. pylori*-colonized

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

individuals, do not present any detrimental effects. Thus, the vast majority of carriers (*i.e.*, up to 75%) remain asymptomatic, while only 15% develop ulcers, and less than 3% develop cancer. Further, growing and sometimes contradictory evidence from recent experimental, clinical studies and epidemiological studies suggest that *H. pylori* might provide protection against obesity-related inflammation and type 2 diabetes [10], esophageal, cardiac pathologies, childhood asthma and allergies [11] and autoimmune diseases. In this context, it is crucial to understand the mechanisms that promote host tolerance to the bacterium in the gastrointestinal mucosa and its systemic regulatory effects since these have been linked to the beneficial commensal aspects of *H. pylori*-human host interaction. Computational models provide a cost-effective and predictive way to study the complex and dynamic immune system interactions and form a non-intuitive novel hypothesis. Solving the complex puzzle of immunoregulatory mechanisms that include large spatiotemporal scales ranging from cellular, intracellular, tissue and organ level scales is a major unsolved challenge that requires applying computational modeling and data analytics.

An advanced hybrid model used to study the mucosal immune response during gut inflammation highlighted the mechanisms by which effector CD4+ T cell responses, contributed to tissue damage in the gut mucosa following immune dysregulation [12]. Other hybrid models with the integration of ABM, ODE, and PDE technologies, were developed to understand the dynamics of tumor development [13] and tumor growth models [14]. These combined techniques have been used to develop multi-organ models in various situations, including the study of granuloma formation [15] and pressure-driven ulcer formation in post spinal cord injury patients [16]. The summary of different agent-based simulators with immunology related applications are discussed and

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

76 summarized in [17, 18]. The comparison between different multiscale modeling  
77 tools and agent-based immune simulators, are discussed in [12, 19].

78 In this study, we utilize a high-resolution **EN**teric Immunity **SI**mulator (ENISI)-  
79 based model of the stomach for simulating the mucosal immune responses to *H*  
80 *pylori* infection. The advanced hybrid multiscale modeling platform ENISI  
81 multiscale model (MSM) is capable of scaling up to  $10^{12}$  agents [20]. The host  
82 immune responses initiated during *H. pylori* infection and the underlying  
83 immunoregulatory mechanisms are captured using the ENISI multiscale hybrid  
84 model. The underlying intracellular mechanisms that control cytokine production,  
85 signaling and differentiation of macrophages and T cells are modeled by using  
86 ODEs, the diffusion of cytokine values is modeled using PDEs and the location  
87 and interactions among the immune cells, bacteria and epithelial cells are  
88 modeled by using ABMs. The hybrid model thereby represents a high-  
89 performance computing (HPC)-driven large-scale simulation of the massively  
90 interacting cells and molecules in the immune system, integrating the multiple  
91 modeling technologies from molecules to systems across multiple  
92 spatiotemporal scales.

93 To understand the dynamics and emergent immunological patterns  
94 described by this hybrid model, we employed sensitivity analysis (SA), an  
95 important part of the model analysis used to explore the influence of varying  
96 model parameters on the simulation outputs. The influence of the effects of  
97 changes in parameter values on the model output explains the model dynamics  
98 that underlay the outputs [21, 22]. Furthermore, SA examines the robustness of  
99 the model output at a different range of parameter values that correspond to a  
100 range of different assumptions. We employed global SA and conducted a two-  
101 stage spatiotemporal global SA approach. First, we used a regression-based



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

method such as the partial rank correlation coefficient (PRCC) and screened the important input parameters that were shown to have the most influence on the output cell populations obtained from the hybrid model. Second, the screened input parameters from the first stage were varied to build a second stage parameter design matrix, and the computer simulations were again run using the hybrid ENISI model. The outputs from both analytics stages were combined and used as a ‘training dataset’ to build a spatiotemporal Gaussian process based metamodel. Finally, variance-based decomposition global SA was used to compute the Sobol’ indices and the most influential parameters over the course of infection were identified. The data analytics methods conducted on the hybrid model identified the epithelial cell parameters such as epithelial cell proliferation as the most influential ones, required for the successful colonization of *H. pylori* in the gastric microenvironment.

## 2. Methods

### 2.1 Hybrid multiscale *Helicobacter pylori* infection model

We developed a multi-compartment, high-resolution, hybrid ABM/ODE/PDE model to capture the dynamics of the immune response during *H. pylori* colonization of the gastric mucosa. The model has a spatial discretization such that the dimension of the entire (two-dimensional, (2D)) grid is 30  $\mu\text{m}$  x 10  $\mu\text{m}$ . An individual grid cell for our simulation is 1  $\mu\text{m}$  x 1  $\mu\text{m}$ , however, this is a configurable run parameter and can be changed without modifying the model. An individual grid cell is a unit wherein all the agents located within that location have the same cytokine environment, *i.e.*, for all the agents in that location, ENISI-MSM would send the same concentration of the cytokines to COPASI.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

126 The entire grid is divided within into four functionally and anatomically distinct  
127 sized compartments: lumen, epithelium, lamina propria and gastric lymph node.  
128 In the model, there are multiple cells and cell types (i.e., agents) within this  
129 dimensional grid. At the beginning of each simulation cycle, the cells (agents)  
130 are randomly placed within the within the 2D grid. The separation of different  
131 types of agents, corresponding to different cell types, into compartments within  
132 the grid is based on the conceptual framework that underlines the model, which  
133 is based on author's expertise and available information. Currently the individual  
134 agents do not have any physical size meaning such that there is no limit of agents  
135 within each individual spatial grid. The model is initialized with the concentration  
136 of different cell types (*i.e.* agents for *e.g.* macrophages) at the beginning of the  
137 simulation by the user.

138 The use of a border implementation permits the migration of agents (cells)  
139 across compartments and facilitates the unidirectional and bidirectional  
140 movement of the agents. At the cellular scale, ENISI MSM, simulated epithelial  
141 cells, macrophages, dendritic cells (DC), CD4+ T cells and bacteria that are  
142 implemented as agents in the model. At the intracellular scale, calibrated ODE-  
143 based models of T cells [23] and macrophages [2] were used to represent the  
144 intracellular pathways controlling cytokine production. The CD4+ T cell ODE  
145 model was calibrated using the experimental data provided in the Table S1 of  
146 [23]. The Particle Swarm algorithm implemented in COPASI was used to  
147 determine unknown model parameter values and fully calibrate the CD4+T cell  
148 ODE model, the details are described in [23]. The intracellular macrophage ODE  
149 model was calibrated using a combination of sourced and new data generated  
150 from *in vitro* macrophage differentiation studies, that were compiled into a  
151 dataset provided within S2 file of [2]. The parameter values are specified within

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

152 the previously published manuscripts - CD4+ T cell ODE model (Carbo,  
153 Hontecillas et al. 2013) and macrophages [2]. The parameters of the calibrated  
154 ODEs were kept unchanged, and the ABM parameters were calibrated by  
155 approximating the output simulations such that they qualitatively resembled the  
156 patterns observed in a mouse model of *H. pylori* infection [24], also described in  
157 detail in section 3.1. Cytokines secreted by immune cells and their change in  
158 concentration were modeled by PDE. The degradation value of the cytokines  
159 and the diffusion constant determines the spread of the cytokine value of one  
160 grid cell to its neighboring grid cell similar to as described in [12]. The features  
161 of ABM, ODE, and PDE were combined to create a multiscale modeling  
162 environment which spanned across different orders of spatiotemporal scales.  
163 The model output contains information about the x and y co-ordinate of the  
164 agents at every time point. The cytokines and internal signaling pathways that  
165 drive functional fates of cells are well mixed within a cell, i.e., we have only  
166 temporal resolution within the cell during a time step. However, the production,  
167 degradation, and diffusions are cell-specific thus the cytokine concentration  
168 results are also spatio-temporal. Since, the model is capable of providing  
169 information regarding spatial co-ordinates over time, we claim the model to be a  
170 spatio-temporal model.

171  
172 The code for the hybrid model is freely accessible and can be downloaded  
173 at <https://github.com/NIMML/ENISI-MSM>. The detailed instructions for the  
174 usability, instructions on 'how to run a simulation' and codes for creating specific  
175 examples presented here are presented in Additional file S1. The SciCrunch.org  
176 database assigned research identification initiative ID (RRID) for ENISI-MSM is  
177 RRID:SCR\_016918. The design of the implementation of the code structure is

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

178 depicted in the Additional file **Fig S1**. The hybrid model is implemented in C++  
179 and utilized the Repast HPC library ([https://repast.github.io/repast\\_hpc.html](https://repast.github.io/repast_hpc.html))  
180 [25]. For the ODEs, we utilized COPASI [26], an ODE-based modeling tool used  
181 in computational biology. The rules in the model that described the interaction of  
182 *H. pylori* with the gastric mucosa and the immune responses resulting from the  
183 infection are derived from the findings in our previously published studies [1, 2].  
184 Specifically, this hybrid model reproduced the immune responses generated by  
185 the interaction *H. pylori* and the resident macrophages as shown in the mouse  
186 model of *H. pylori* infection [24]. The rules for each cell type in the *H. pylori*  
187 infection are summarized in **Table 1**. A pictorial representation of the rules is  
188 depicted in *Fig 1*. These cell types represented as agents, act according to the  
189 rules (as in **Table 1**) that are updated at discrete simulation cycle.

190  
191  
192 **Fig 1. *Helicobacter pylori* infection schematic diagram of the hybrid ABM**  
193 **ODE model**

194 *The model comprises four compartments, i) the lumen that contains H. pylori and*  
195 *bacteria, ii) epithelium that contains epithelial cells and dendritic cells, iii) lamina*  
196 *propria that contains variety of immune cells including the infiltrating effector*  
197 *(eDCs) and tolerogenic (tDCs) dendritic cells, monocytes, regulatory*  
198 *macrophages (both resident and monocyte-derived macrophages), T helper*  
199 *cells and naïve CD4+ T cells (nT), Th1, iTreg, Th17, Tr cells. and iv) gastric*  
200 *lymph node compartment that contains eDCs, tDCs, Th1, Th17, iTreg and nT.*  
201 *The Tr cells in the lamina propria are the type 1 regulatory (Tr1) T cells with*  
202 *regulatory function whose expansion is largely dependent on environmental IL-*  
203 *10. These are different than iTreg which are T cells differentiated from naïve T*

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

204 cell in presence of tolerogenic dendritic cells and TGF- $\beta$  cytokine The two  
205 calibrated ODEs for T cells and regulatory macrophages are integrated as the  
206 ODE components in the hybrid model. The cellular agents are simulated in a  
207 two-dimensional grid space with their behavior defined by a set of rules during a  
208 course of *H. pylori* infection.

**Model description**

211 ENISI MSM is a multiscale agent-based modeling platform for computational  
212 immunology which was built on our previous works, ENISI-MSM [12] that  
213 integrated COPASI, the ODE solver, ENISI, an agent based simulator.

**Spatial discretization**

215 The model has a spatial discretization such that the dimension of the entire (two  
216 dimensional) grid is 30  $\mu\text{m}$  x 10  $\mu\text{m}$ . An individual grid cell is 1  $\mu\text{m}$  x 1  $\mu\text{m}$ ,  
217 however, this is a configurable run parameter and can be changed without  
218 modifying the model. The four functionally and anatomically distinct sized  
219 compartments are separated by border implementation such that the dimensions  
220 of the four compartments are lumen (2  $\mu\text{m}$ ), epithelium (1  $\mu\text{m}$ ), lamina propria (5  
221  $\mu\text{m}$ ) and gastric lymph node (2  $\mu\text{m}$ ). The following compartments are adjacent to  
222 each other: lumen – epithelium, epithelium - lamina propria and lamina propria –  
223 gastric lymph node. A figure describing the spatial discretization is shown in the  
224 Additional file **Fig S2**.

225 The parameters that define the initial concentration of the agents and the  
226 diffusivity of cytokines are obtained from a properties file (*model.props* in the  
227 Howtorunasimulation folder in the GitHub repository). All the value of the  
228 parameters as listed in *Table S1*. The detailed mechanism that each parameter

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

229 corresponds to is described in the second column, *parameter description*, of  
230 *Table S1*. We demonstrate below how we obtain a count of thousands resident  
231 macrophages. For *e.g.*, if the initial concentration of resident macrophages in the  
232 lamina propria is 30, the total number of these resident macrophages can be  
233 calculated by the equation described below -

$$\begin{aligned} 234 & \\ 235 & n(\text{resident macrophages}) = \text{size}_{\text{compartment}}(\text{lamina propria}) \times \text{concentration}_{\text{initial}} \\ 236 & (\text{resident macrophages}) \\ 237 & n(\text{resident macrophages}) = (30 \times 5) \times 30 = 4500. \end{aligned}$$

239 ***Time Step size***

240 The time step size is 1 tick ~ 1 day which was obtained during the process of  
241 qualitatively comparing the output to the results from the mouse model of *H.*  
242 *pylori* infection. For *e.g.*, the peak of resident macrophages in lamina propria  
243 (refer *Fig 2b, d*) is observed at ~21 days which is similar to the results obtained  
244 in *Fig 2A* described in [24] (also described in detail in section 3.1).

246 ***Updating***

247 Each agent has an ‘act’ function within the code that describes the rules  
248 implemented for each of the agent groups. At every simulation cycle, each agent  
249 inspects its location and updates its state. If the agents were T cells and  
250 macrophages, they obtained the cytokine concentration from the ValueLayers,  
251 sent that information to COPASI that calculated the differentiation subtype of the  
252 agent and cytokines to be secreted that into the environment [12]. The input to  
253 the ODEs were the cytokine values at the agent’s location. Thus, the intracellular  
254 ODE models were utilized to determine and update the state. Each agent

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

255 proliferated, died, changed its state and moved across the compartment,  
256 following the set of rules defined for them.

257 The COPASI setup for the solver used the LSODA (Livermore Solver for  
258 Ordinary Differential Equations) differential equation solver. The default values  
259 for the setup such as the - relative tolerance (1e-6), absolute tolerance (1e-12)  
260 and maximum internal steps of 10000 were maintained. The ENISI MSM sends  
261 the current concentrations of the cytokines  
262 to COPASI. COPASI uses those values to integrate the deterministic  
263 model for one tick, i.e., 1 day. The resulting time series of cytokine  
264 concentrations are used to update the cytokine value in the ABM/PDE  
265 system. COPASI simulates different model for each relevant cell type.

266 The ENISI MSM PDE solver uses a simple numerical scheme to solve the PDEs  
267 (<https://github.com/NIMML/ENISI-MSM/tree/master/src/diffuser>) and process  
268 distributed value layer (<https://github.com/NIMML/ENISI-MSM/blob/master/src/grid/ValueLayer.h>). The ValueLayer stores the value for a  
269 grid space and provides methods to change the values of individual grid cells.  
270 The Diffuser is used to diffuse the values of the ValueLayer using diffusion (d)  
271 and degradation (delta) constants as described in [12]. The diffusion constant  
272 determines the migration of values of a grid cell to its neighboring grid cells. As  
273 implemented in ValueLayer library, the diffusion of cytokines follows the equation  
274 shown below also described in Mei et al, 2015. Here,  $v_n$  is the value of the grid  
275 cell itself at step n. The values of  $C_{\text{delta}}$  and  $C_d$  are degradation and diffusion  
276 constant respectively.  
277

$$278 \quad v_n = v_{n-1} + C_{\text{delta}} * [ \sum ( C_d^{\text{neighbor}} * v_{n-1}^{\text{neighbor}} ) - 6.0 * v_{n-1} ]$$

0.3	1.2	0.3
1.2	-6.0	1.2
0.3	1.2	0.3

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16 279

17  
18  
19  
20 280  
21  
22 281  
23  
24 282  
25  
26 283  
27  
28  
29 284

The PDE solver uses the above number scheme  $c_d^{neighbor}$  for the diffusion process. The step size  $c_{delta}$  is automatically adjusted at the beginning of the simulation based on the degradation and diffusion constants to avoid underflow errors, *i.e.*, multiple PDE steps are in general executed per tick. The grid size is the identical with the spatial discretization for the agents.

30  
31  
32 285

***Movement***

33  
34 286  
35  
36  
37 287  
38  
39 288  
40  
41 289  
42  
43 290  
44  
45 291  
46  
47 292  
48  
49 293  
50  
51 294  
52  
53 295  
54  
55 296  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

The cells and bacteria agents presented in the model have Brownian motion and move randomly within the compartment. Brownian movement is an inherent property of a cell. Depending on cell phenotypes the movement can vary, but all cells with the same phenotype exhibit similar movements. Additionally, chemokine-driven movement is dependent on chemokine concentration in a tissue site. The capability of chemokine-driven movement exists in ENISI-MSM if the right chemokines are represented in the model. However, the focus of this model was to investigate changes in cell phenotype and not chemokine-driven movement of cells. Thus, the chemokines driving the movement are not represented in the current model. Cell migration is implemented in the code as the *move()* function for each of the cells and agents, which call the



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

297 `moveRandom()` function from the ([https://github.com/NIMML/ENISI-](https://github.com/NIMML/ENISI-MSM/src/compartiment/Compartiment.cpp)  
298 [MSM/src/compartiment/Compartiment.cpp](https://github.com/NIMML/ENISI-MSM/src/compartiment/Compartiment.cpp)) file.

299

300 The hybrid model simulations were run on an Ivy Bridge-EX E7-4890 v2 2.80  
301 GHz (3.40 GHz Turbo) quad processor nodes. The code was parallelized such  
302 that the simulation time on a single node with four parallel tasks, varied between  
303 9-10 minutes. This runtime was based on the model parameters at the initiation  
304 stage, which included the number of immune cell, bacteria, epithelial cells,  
305 number of time steps, and size of the two-dimensional grid. To facilitate the  
306 investigation of the mechanisms underlying host responses during *H. pylori*  
307 infection, anatomical and functional compartments were spatially linked such that  
308 the agents had both unidirectional and bidirectional movement. All the agents  
309 worked in a synchronous format wherein the two agent populations  
310 (macrophages and T cells) made function calls to their respective ODE models  
311 [2] [23]. These agents used the varying cytokine concentration (*i.e.*, environment  
312 variable) in their grid spaces as inputs to the ODE model, and these models were  
313 run using COPASI [26]. **Table 2** shows information on the agents and the states  
314 that they can acquire.

315

1  
2  
3  
4 316  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28 317  
29  
30 318  
31  
32  
33 319  
34  
35 320  
36  
37 321  
38  
39 322  
40  
41 323  
42  
43  
44 324  
45  
46 325  
47  
48 326  
49  
50 327  
51  
52 328  
53  
54  
55 329  
56  
57 330  
58  
59 331  
60  
61  
62  
63  
64  
65

Name of agents	States it can acquire	Name of the states in the hybrid model
<i>Helicobacter pylori</i>	0	<i>H. pylori</i>
Macrophages	0	Monocyte
	1	Resident
	2	Regulatory
	3	Inflammatory
Dendritic cells	0	Immature
	1	Effector
	2	Tolerogenic
T cell	0	Naïve
	1	Th1
	2	Th17
	3	iTreg
	4	Tr
Epithelial	0	Healthy
	1	Damaged
Bacteria	1	Infectious
	2	Tolerogenic

**Table 2. List of all the agents and the states they can acquire.**

All the agents can acquire at least 1 and at the most 5 states. The names chosen for the acquired states are closely related to their functional properties based on the underlying “rules”.

## 2.2 Global sensitivity analysis

To conduct the global SA, we determined a list of 38 parameters to be varied that were selected based on the calibration process (wherein the parameters that did not show a lot of variation were not included). A range of values (maximum and minimum) was specified for each of the parameters (refer to Additional file **Table S1**) by expert judgment, summarized by bounded intervals. The practice of using expert judgment is known in the SA field as supported in [27]. As discussed in [28], one of the challenges encountered using ABM is the

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

332 process of determining the parameter values, for *e.g.* this may include the lack  
333 of the availability of experimental techniques to measure such parameters. The  
334 values of the parameters for the model presented here are obtained via the best  
335 guess based on the qualitative comparison of the computer model outputs with  
336 that of the experimental results obtained from the mouse model of *H. pylori*  
337 infection (Viladomiu, Bassaganya-Riera et al. 2017) (as described in detail in  
338 Section 3.1). Since, the source of the parameters is not known we estimated the  
339 values to fit the data obtained from the mouse model of infection.

340 The values of these parameters were normalized within the range of 0 and 1  
341 for SA purposes. We employed a two-stage metamodeling methodology to  
342 determine the influence of each input parameter to the model output, in a high  
343 dimensional screening setting inspired by [29]. The step-wise procedure is  
344 described in the Additional file, *Fig S3*. All the files for global SA are freely  
345 accessible and can be downloaded at [https://github.com/NIMML/Sensitivity-](https://github.com/NIMML/Sensitivity-Analysis)  
346 [Analysis](https://github.com/NIMML/Sensitivity-Analysis).

347 The two-stage global SA is described in detail in the below section. To  
348 summarize, for the first stage the input parameter matrix was designed using the  
349 method described in Moon, Dean et al. 2012 and simulations were run using the  
350 hybrid computer model. The simulation output from the first stage was analyzed  
351 using PRCC as it was computationally efficient, and the active inputs (significant  
352 effect) were screened to reduce the input parameter space. Second, the active  
353 parameters were varied whereas the inactive parameters from the first stage  
354 were maintained at a nominal value for the input parameter matrix design to be  
355 employed for the second stage. Third, the simulation outputs from both stages  
356 were combined and used as a training dataset to fit a spatio-temporal  
357 metamodel. Fourth, the unknown model parameters for the spatio-temporal

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

358 metamodel were estimated using the maximum log-likelihood function. The  
359 spatio-temporal metamodel was used as a substitute for the hybrid computer  
360 model, and the variance-decomposition method was used to compute the Sobol'  
361 total and first-order indices. Overall, we employed both approaches, PRCC  
362 based (for screening) and Sobol' indices calculation to perform a complete global  
363 SA of the hybrid computer model. The following sections, describe a detailed  
364 step by step explanation of the procedure.

366 Design of two-stage experiments and analysis

367  
368 The input for the hybrid model are varying parameter values obtained from  
369 the design matrix and the output are the number of cells (agents) that vary over  
370 time. The first stage experiment was focused on the screening of the input  
371 variables to reduce the number of input parameters to vary for the SA and to limit  
372 the computational cost. Computational costs are often a limiting factor that play  
373 an important role in the inclusion of model parameters in the SA [21]. For the  
374 design, we assumed the total number of input parameters under consideration  
375 as  $d$  (in our case, 38). With an assumption of a maximum of 50% active inputs  
376 that is aimed to improve the screening performance, the number of runs for stage  
377 1, was fixed to  $n_1 = 4d$ , such that  $n_1 > 5 * d * 0.5 = 2.5d$  as in [29]. To construct a  $n_1$   
378 \*  $(n_1 - 1)$  preliminary input parameter design matrix,  $X^*$ , needed to be constructed  
379 ([29]). The input parameter design matrix for first stage sampling was drawn from  
380  $X^*$ .

381 The algorithm for the first stage design generated a design matrix  $X^{(1)}$  that  
382 satisfied the below three listed properties as in [29]

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

- 383 i) The columns of  $X^*$  were uncorrelated thereby facilitating the independent
- 384 assessments of the effects due to the input parameters.
- 385 ii) The maximum and minimum value in each input parameter column were
- 386 ensured to be 0 and 1 respectively, thereby preventing any input values
- 387 with larger values to have a larger influence on the response, induced by
- 388 the design.
- 389 iii) The designs defined by  $X^*$  had “space-filling” properties such that all the
- 390 regions of the input space were exhaustively explored.

391  
392 First stage sampling plan:

393 The first stage input parameter design matrix  $X^{(1)}$  was obtained by selecting

394 the first  $d$  columns of  $X^*$ , i.e.  $X^{(1)} = (\xi_1, \dots, \xi_d)$ . The hybrid computer model was

395 run and the simulation outputs at these  $n_1$  design points were obtained.

396 In our case, the model comprised of  $d = 38$  input variables. The total number of

397 distinct input parameter design points obtained using the above procedure was

398  $n_1 = 152$  ( $4 \cdot d = 4 \cdot 38$ ). To account for the variability in the output, we run 20

399 replicates ( $r$ ). Thus, the total number of simulations run using the hybrid model

400 computer simulator with  $X^{(1)}$  as input parameter design matrix, were  $r \times n_1 = 20$

401  $\times 152 = 3040$ .

402  
403 First stage analysis

404 We analyzed the outputs from first stage analysis and screened the active

405 inputs from using PRCC. To measure the effect of input parameter on output, we

406 performed both PRCC and the spearman rank correlation coefficient (SRCC)

407 analysis. PRCC and SRCC were chosen because they were computationally

efficient (accounting for the low computational budget). A correlation analysis provides a measure of the strength of linear association between input and output variable [30]. A correlation coefficient between  $x_j$  and  $y$  is calculated as follows:

$$r_{x_j y} = \frac{Cov(x_j, y)}{\sqrt{Var(x_j)Var(y)}} = \frac{\sum_{i=1}^N (x_{ij} - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_{ij} - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}}$$

$$j = 1, 2, \dots, k.$$

where  $Cov(x_j, y)$  stands for the covariance between  $x_j$  and  $y$ , and  $Var(x_j)$  and  $Var(y)$  are the variance of  $x_j$  and  $y$  respectively.

PRCC is performed when i) a non-linear but monotonic relation exists between the input and outputs, and ii) when little or no correlation exists between the input variables (which is guaranteed by the property (i) of our input parameter matrix,  $X^{(1)}$  described above). As described in Marino, Hogue et al. 2008, the PRCC between rank transformed  $x_j$  and  $y$  is the CC between the two residuals  $(x_j - \widehat{x}_j)$  and  $(y_j - \widehat{y}_j)$  where  $\widehat{x}_j$  and  $\widehat{y}_j$  are rank transformed and follow the linear regression models as follows:

$$\widehat{x}_j = c_o + \sum_{\substack{p=j \\ p \neq j}}^k c_p x_p \text{ and } \widehat{y}_j = c_o + \sum_{\substack{p=j \\ p \neq j}}^k c_p x_p .$$

We performed the PRCC analysis on the outputs obtained from the hybrid computer model with  $X^{(1)}$  as an input, using 'epi.prcc' package in R (<https://cran.r-project.org/web/packages/epiR/epiR.pdf>). The significance test evaluated the strength of influence each input parameters and assessed if the PRCC

1  
2  
3  
4 428 coefficients were significantly different than zero [30]. We run the PRCC analysis  
5  
6 429 for 13 output cell populations (*Fig 3* shows data for two output populations and  
7  
8  
9 430 the rest of the data not shown) and identified the active input parameters using  
10  
11 431 the significance test. PRCC and SRCC produced identical outputs, hence results  
12  
13 432 from SRCC are not shown here. If an input parameter was shown to be  
14  
15 433 significant ( $P < 0.05$ ) in one of the 13 output cell populations, it was considered  
16  
17 434 as an active input for the second stage input parameter design matrix.  
18  
19  
20 435 Additionally, domain expert knowledge was employed to include additional  
21  
22 436 parameters, based on the biological significance, that were otherwise shown to  
23  
24 437 be non-significant. In all, based on the PRCC analysis performed on the outputs  
25  
26  
27 438 obtained from the first stage simulations and domain expert knowledge, we  
28  
29 439 chose 23 input parameters as active inputs for the second stage (see Additional  
30  
31 440 file *Fig S4*). Thus, PRCC screened inputs at significance level  $p < 0.05$  and inputs  
32  
33 441 based on expert knowledge were selected as active inputs to be varied for the  
34  
35 442 second stage sampling plan.

36  
37  
38  
39 443 *Second stage sampling plan:*

40  
41  
42 444 The number of active inputs obtained from the first stage analysis amounted  
43  
44 445 to 23 parameters out of the initial set of 38 parameters. We followed the design  
45  
46  
47 446 described in [29] for the second stage and the number of design points amounted  
48  
49 447 to,  $n_2 = 100\% * 5 * a$  where 'a' stands for the number of active inputs from the first  
50  
51 448 stage. This resulted into  $n_2 = 23 * 5 = 115$  parameters combinations for the second  
52  
53  
54 449 stage input parameter design matrix. Since outputs from both stages are to be  
55  
56 450 combined for second stage analysis, per [29], the design for the second stage  
57  
58 451 was chosen to build on top of  $X^{(1)}$ . The sampling phase design algorithm ensured

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

452 that the columns satisfied the properties (i) (uncorrelated design points) and (ii)  
453 (between values 0 and 1) as listed in the previous section. We constructed the  
454 115 x 38 (115 parameter setting and 38 parameters) design matrix for the second  
455 stage that incorporated the 23 active inputs obtained from the PRCC screening  
456 in the first stage output analysis. After combining the design points from both the  
457 stages, the parameter design matrix  $X$  with space filling properties contained 267  
458 (152 from the first stage and 115 from the second stage) design points.

459 Second stage analysis

460 We run the computer code for the hybrid model with the second stage input  
461 parameter design matrix (with 115 ( $n_2$ ) design points), for 20 ( $r$ ) replicates, which  
462 amounted to 115 x 20 (2300) runs. The outputs from the first stage (152 x 20  
463 runs) and second stage (115 x 20 runs) were combined to provide the training  
464 data to build a spatio-temporal metamodel. For the second stage analyses, we  
465 utilized a metamodeling-based approach. Metamodels are surrogate models that  
466 can be used as a substitute for the simulation model [31]. The use of metamodels  
467 reduces the computational budget, cost of analysis, and are useful options in  
468 cases when the simulation model is expensive to run (in our case 9-10 minutes  
469 for 1 design point) [31]. The various metamodeling techniques used to build  
470 surrogates for a computer model output include linear regression models, neural  
471 networks, high dimensional model representation methods, Gaussian process  
472 (GP) regression models, polynomial chaos expansion and more that are  
473 discussed in length in [32, 33]. Amongst these, GPs are one of the most popular  
474 emulators as it allows modeling of fairly complex functional forms. The GPs not  
475 only provide prediction at a new point but also an estimate of the uncertainty in



1  
2  
3  
4 476 that prediction [32]. A GP is a stochastic process for which any finite set of  $y$ -  
5  
6 477 variables has a joint multivariate Gaussian distribution [34] [32]. Suppose,  $y_j(w)$ ,  
7  
8  
9 478 the simulation response obtained on the  $j$ th simulation replicate, at a design point  
10  
11 479  $w = (X^T, t)^T \in \mathcal{X} \times \mathbb{T}$ , it can be described as follows:  
12  
13

$$14 \quad y_j(w) = Y(w) + \varepsilon_j(w) = \beta_0 + M(w) + \varepsilon_j(w), \quad (1)$$

15  
16  
17  
18 481 where  $Y(w)$  represents the mean function of  $y_j(w)$ , the quantity of interest that  
19  
20 482 we intend to estimate at any design point  $w$ . The  $\beta_0$  is a constant trend term and  
21  
22  
23 483 is assumed to be unknown. The input parameter  $X \in \mathcal{X} \subset \mathbb{R}^d$  and the time  $t \in$   
24  
25 484  $\mathbb{T} \subset \mathbb{R}_+$ ; and  $X$  is independent of  $t$ . The  $\varepsilon_j(w)$  are represents the sampling  
26  
27  
28 485 variability inherent in a stochastic simulation, that are that are assumed to be  
29  
30 486 independent and identically distributed across the replications at any given  
31  
32 487 design point [35].  
33  
34

35 488 The term  $M(w)$  represents a stationary Gaussian process with mean = 0 and  
36  
37  
38 489 covariance between any points was modeled as the Gaussian covariance  
39  
40 490 defined in [36]. Thus, the covariance between any design points  $w_a =$   
41  
42 491  $(X_a^T, t_a)^T$  and  $w_b = (X_b^T, t_b)^T$  in the random field can be modeled as-  
43  
44  
45

$$46 \quad Cov(M(w_a), M(w_b)) = \Gamma^2 \exp(-\sum_{r=1}^d \theta_r (X_{ar} - X_{br})^2 R(t_a - t_b; \gamma)), \quad (2)$$

47  
48  
49  
50 493 wherein,  $\exp(-\sum_{r=1}^d \theta_r (X_{ar} - X_{br})^2)$  models the spatial correlation between two  
51  
52 494 input design points  $X_a$  and  $X_b$  in the input parameter space, whereas  
53  
54 495  $R(t_a - t_b; \gamma)$  also given by  $\exp(-\sum_{r=1}^d \gamma_r (t_{ar} - t_{br})^2)$  models the temporal  
55  
56  
57 496 correlation between time points  $t_a$  and  $t_b$ . The parameters  $\theta$  and  $\gamma$  represents  
58  
59 497 the rate at which i) spatial correlation decreases as the points move farther in  
60  
61  
62  
63  
64  
65

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

498 space with the same time index, and ii) temporal correlation decreases as the  
499 time points are farther apart in time at the same input vector, respectively. Both  
500 the spatial correlation and temporal correlation are modeled using the Gaussian  
501 covariance. The parameter  $\Gamma^2$  can be interpreted as the variance of  $M(w)$  for all  
502  $w$ . The input parameter design consists of  $((w_a, n_i)_{i=1}^k)$  design points to run  
503 independent simulations with replicates applied to each of the design points. Let,  
504  $k \times 1$  denote a vector of sample averages of simulation responses given by  $\bar{y} =$   
505  $(\bar{y}(w_1), \bar{y}(w_2), \dots, \bar{y}(w_k))^T$ , where in  $\bar{y}(w_i)$  is the resulting estimate of  
506 performance measure obtained at design point  $w_i$  and  $\bar{\varepsilon}(w_i)$  is the sampling  
507 variability inherent in a stochastic simulation (Ankenman, Nelson, & Staum,  
508 2010). The equations associated with  $\bar{y}(w_i)$  and  $\bar{\varepsilon}(w_i)$  are described below in  
509 equation (3):

$$\bar{y}(w_i) = \frac{1}{n_i} \sum_{j=1}^{n_i} y_j(w_i) = Y(w_i) + \bar{\varepsilon}(w_i) \text{ and } \bar{\varepsilon}(w_i) = \frac{1}{n_i} \sum_{j=1}^{n_i} \varepsilon_j(w_i), i = 1, 2, \dots, k.$$

(3)

512 Similar as in (Ankenman, Nelson, & Staum, 2010), shown below in equation (4),  
513 let  $\Sigma_M$  be the  $k \times k$  covariance matrix across all design points and let  $\Sigma_M(w_o, \cdot)$  be  
514 the  $k \times 1$  vector,  $(\text{Cov}[M(w_o, w_1)], \text{Cov}[M(w_o, w_2)], \dots, \text{Cov}[M(w_o, w_k)]^T$  that contains  
515 spatial covariance between the  $k$  design points and a given prediction point  $w_o$ .  
516 Also, let  $\Sigma_\varepsilon$  be the  $k \times k$  covariance matrix of the vector of simulation errors  
517 associated with the vector of point estimates  $\bar{y}$ , across all design points. As  
518 described in [35], the best linear predictor  $Y(w_o)$  that has the minimum mean  
519 squared error (MSE) among all linear predictors at a given point  $w_o =$   
520  $(X_o^T, t_o)^T$  can be given by equation (4):

$$\hat{Y}(w_o) = \hat{\beta}_o + \sum_M(w_o, \cdot)^T [\sum_M + \sum_\varepsilon]^{-1}(\bar{y} - 1_k \hat{\beta}_o), \quad (4)$$

where,  $1_k$  is the  $k \times 1$  vector of ones and  $\hat{\beta}_o$  is estimated to be 1. The corresponding optimal MSE as in [35] is given by equation (5):

$$MSE(\hat{Y}(w_o)) = \sum_M X_{o, w_o} - \sum_M(w_o, \cdot)^T [\sum_M + \sum_\varepsilon]^{-1} \sum_M(w_o, \cdot) \quad (5).$$

To implement the metamodeling approach as described above, the unknown model parameters are estimated through maximizing the log-likelihood function. The underlying standard assumption is that  $(Y(w_o), \bar{y}^T)^T$  follows a multivariate normal distribution, for e.g., see [35] and [37]. The function implemented in the *mlegp* package in R [38] is used for the estimation of the parameters. Once the parameters are estimated the prediction then follows equations (4) and (5).

### Sensitivity index calculation

To determine the effect of input variables on the output, we employed the variance decomposition method. These methods involve the decomposition of the variance of the output as a sum of the variance produced by each input parameter [34].

We independently generated 10,000 x 38 sampling matrices, such that the parameter combinations are generated via Latin Hypercube sampling and as described in [39]. Simulations were performed using the GP spatio-temporal model as described in the previous section, and the Sobol' indices were computed as described in [40] [39]. The Sobol' method quantitatively measured the contribution of each input parameter by computing the first order and total

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

544 order index [39]. For output  $Y$ , input parameter matrix  $X_i$  where,  $i$  is the input  
545 parameters of the model, the Sobol' indices are computed as follows:

$$SI_1^{Xi} = \frac{V[E(Y|X_i)]}{V(Y)},$$

and

$$SI_{tot}^{Xi} = \frac{V[E(Y|X_{\sim i})]}{V(Y)}.$$

551  
552 The Sobol' first order sensitivity index  $SI_1^{Xi}$  measures the impact of one single  
553 parameter on the model output, whereas the Sobol' total order index measures  
554 the influence of  $X_i$  including all the interactions with other parameters. The First-  
555 order indices were computed using the Sobol-Saltelli's method as described in  
556 [39] [41] whereas, the total order indices were computed using Sobol-Jansen as  
557 in [39, 42].

### 3. Results

#### 3. 1 Hybrid model simulations produce similar immune response dynamics observed in previously published experimental data

561 We first aimed to simulate the findings observed in previous gut models [24]  
562 to ensure that we obtained similar response dynamics from the hybrid ENISI  
563 model of *H. pylori* infection. As in [24], to demonstrate that the gastric mucosa

1  
2  
3  
4 567 harbors a system of macrophages that contribute to the outcome of *H. pylori*  
5  
6 568 infection, we created an *in-silico* Peroxisome proliferator-activated receptor  
7  
8  
9 569 gamma (PPAR $\gamma$ ) macrophage-specific knockout (KO) model. PPAR $\gamma$  is an  
10  
11 570 important transcription factor that controls the expression of genes that  
12  
13 571 contribute to the inflammatory response once this is initiated. To disrupt the  
14  
15 572 downregulation of pro-inflammatory responses, we simulated a PPAR $\gamma$  KO  
16  
17  
18 573 system in either macrophage or T cell populations and compared the response  
19  
20 574 to a wild-type system. In the model, we created three different macrophage  
21  
22 575 populations, comprised of, “resident” macrophage agents that mimic the  
23  
24 576 properties of the F4/80hi CD11b+ CD64+ CXCR1+ macrophages reported in  
25  
26  
27 577 [24], monocyte-derived (infiltrating) and macrophage populations with regulatory  
28  
29 578 (M2, or alternatively activated) and pro-inflammatory function (M1 or classically  
30  
31 579 activated) (see *Table 2*).

32  
33 580 We simulated an *in-silico* *H. pylori* infection by creating four groups, i) a control -  
34  
35  
36 581 WT (representing a wild-type group), ii) CD4Cre (T cell specific PPAR $\gamma$  KO-lacks  
37  
38 582 PPAR $\gamma$  gene in all CD4 T cells), iii) LysMCre (Myeloid cell specific PPAR $\gamma$  KO-  
39  
40 583 lacks PPAR $\gamma$  gene in all macrophages) and clodronate group (simulating the  
41  
42 584 removal of macrophages by chemical depletion via clodronate treatment). To  
43  
44  
45 585 simulate the CD4Cre group, the probabilities of a naive T cell transitioning to an  
46  
47 586 iTreg cell ( $p_{nTtoiTreg}$ ) and Th17 cell differentiating to iTreg ( $p_{Th17toiTreg}$ )  
48  
49 587 were reduced to 5% and 10% of the control value, respectively (refer to Table  
50  
51 588 S1). As described in [23], to simulate the LysMCre experimental conditions, the  
52  
53  
54 589 probabilities of i) a monocyte transitioning to a regulatory macrophage  
55  
56 590 ( $p_{Mregdiff}$ ) and ii) immature dendritic cells switching to tolerogenic dendritic  
57  
58 591 cells ( $p_{iDCtotDC}$ ) were reduced approximately to 60% and 30% of the control  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

592 value, respectively (refer to Table S1). A complete set of parameter for each of  
593 the biological KOs are included as separate columns in Table S1. Lastly, the  
594 removal of macrophages by clodronate were simulated by decreasing the initial  
595 numbers of the macrophage population including the resident macrophages. The  
596 rationale to include the clodronate group (macrophage removal) was to evaluate  
597 if depletion of phagocytic cells (terminology with respect to model, *i.e.*,  
598 monocytes, resident, monocyte-derived macrophages and inflammatory  
599 macrophages) would affect *H. pylori* colonization levels, as we have previously  
600 reported in an *in vivo* model [24]. Further, to simulate the myeloid cell PPAR $\gamma$  KO  
601 system, the initial population of resident macrophages were also reduced.

602 All the groups were initialized with equal loads of *H. pylori* agents. Ten  
603 replicates of the simulations were performed for each of the input parameter  
604 settings specific to each group. The outputs were averaged, and standard error  
605 of the means were plotted as ribbons (shaded regions) across the graphs. After  
606 running the ten replicates of the time series *in-silico* simulation, the hybrid model  
607 showed significantly ( $p < 0.05$ ) higher levels of *H. pylori* in the WT and CD4Cre  
608 groups as compared to LysMCre KO and macrophage-depleted groups (*Fig. 2*,  
609 panel a and d).

610  
611  
612 ***Fig 2. Time course simulations representing the immune response during***  
613 ***Helicobacter pylori infection.***

614 *The upper half of the plot in both the panels shows the dynamics of the*  
615 *population cells*  
616 *over time representing the number of cells (y-axis) versus time (x-axis) in a WT*  
617 *(black), CD4Cre (green), clodronate (red) and LysMCre (blue) simulated in-silico*

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

618 groups during *H. pylori* infection. The cell populations include - a) *H. pylori*; b)  
619 the resident macrophages and, c) monocyte-derived macrophages in the lamina  
620 propria compartment. The figures in the lower half (d-f) of both the panels, show  
621 the results for statistical comparison between the groups using ANOVA with the  
622 post-hoc analysis. The letters 'a', 'ab' and 'b' represent statistically significant  
623 differences ( $P < 0.05$ ) between the groups obtained after running the Tukey's  
624 Honestly Significant Difference. A side by side comparison with the bacterial load  
625 and macrophage population as observed in the mouse model of *H. pylori*  
626 infection are also included.

627  
628 In addition to the increase in *H. pylori*, WT and CD4Cre *in-silico* experimental  
629 groups had a higher resident as well as monocyte-derived regulatory  
630 macrophages as compared to clodronate (macrophage depleted) and LysMCre  
631 groups (Fig. 2b-c, e-f). The results in the mouse model indicated that between  
632 weeks 2 and 3 post-infection a decrease in bacterial burden in the stomach of  
633 LysMcre mice was observed as shown in Fig 1A of Viladomiu, Bassaganya-Riera  
634 et al. 2017. The decrease in bacterial burden led to a significant and sustained  
635 lower colonization levels when compared to WT and CD4Cre. Similar to the  
636 results observed in the mouse model, we observed a significant decrease (Fig  
637 2a, d) in the bacterial burden in the simulated LysMcre group as compared to  
638 the simulated WT and CD4cre groups. Furthermore, the results from the mouse  
639 model indicated that a significant increase in numbers of F4/80hiCD11b+ CD64+  
640 CX3CR1+ cells (here referred to as resident macrophages in this paper), was  
641 observed in WT mice in comparison with LysMcre mice as shown in Fig. 2A, 2E  
642 of Viladomiu, Bassaganya-Riera et al. 2017. These cells accumulated in the  
643 stomach mucosa starting on day 14 post-infection in the WT mice but not in the

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

644 LysMcre mice. We observed a similar increase (*Fig 2b,e and Fig 2c,f*) in the  
645 number of resident macrophages as well as monocyte derived macrophages in  
646 the simulated WT groups in comparison to the simulated LysMcre group. We  
647 estimated the parameter values to fit the data obtained from the mouse model of  
648 *H. pylori* infection. Thus, the observations were qualitatively similar to the  
649 findings in [24], where the stomach of WT mice was enriched in a population of  
650 F4/80+CD11b+CD64+ myeloid cells, compared to LysMCre mice.

651 Overall, with the results in *Fig 2*, we showed the ability of the hybrid model to  
652 replicate the experimental results in [24], and this preliminary data was used as  
653 a base calibration setting for SA and other *in-silico* findings.

654 **3.2 Partial correlation coefficient analysis screened the influential**  
655 **parameters**

656  
657 To reduce the computational complexity of varying an input parameter space  
658 of 38 parameters, we divided the SA process in two stages. For first-stage  
659 analysis, we utilized the PRCC regression-based SA method to screen the  
660 influential inputs and used it for the second stage design of the experiments (refer  
661 Methods 2.2). Using PRCC, we determined the impact of the input parameters  
662 on the output cell populations in the model. The parameters with significant  
663 correlation with *H. pylori* in the gastric lamina propria compartment and resident  
664 macrophages are shown in *Fig 3*, along with their PRCC values. The bars in  
665 blue, highlight the parameters that are significantly different than 0, at  $P < 0.05$   
666 compared to grey bars which are not significant. It is important to note that at this  
667 stage the analysis using PRCC was non-temporal.



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

668 The SA from first stage results showed that the epithelial damage due to  
669 infectious bacteria (*epiinfbctdam*) with a coefficient value of (~0.2), was positively  
670 correlated with the colonization of *H. pylori* in the lamina propria compartment,  
671 indicating the important role of epithelial cell damage during the course of  
672 infection, similar to our findings obtained in [43]. Another parameter included the  
673 probability of the release of IL-6 (*IL6*) with a coefficient value within the range  
674 (0.3-0.4).

675 Next, the epithelial cell damage parameters (*epiinfbctdam* = (0.2-0.3),  
676 *epiTh17dam* = 0-0.2) were shown to have positive influence on the resident  
677 macrophage cells whereas, the T cell type transition parameters (*p\_iTregtoTh17*  
678 = (0.3 - 0.4) and *p\_Th17toiTreg* = (0.1 - 0.2)) showed a negative impact on the  
679 resident macrophages. Similarly, we performed the PRCC analysis for all the cell  
680 populations under consideration during the infection (not shown).

681  
682

**Fig 3. Bar plots for the partial rank correlation coefficients.**

684 *The magnitude of the bar-plot indicates the value of the partial rank correlation*  
685 *coefficient. The blue bar indicated the input parameters shown to be significantly*  
686 *different than 0, at P <0.05 as influential whereas the grey bars indicate the non-*  
687 *influential parameters on a) H. pylori and b) resident macrophages, in the lamina*  
688 *propria compartment.*

689  
690  
691  
692

The significant parameters (marked in blue bars) obtained from the SA of the  
output from first stage design of experiments (152 parameter settings with 20  
replicates, refer Methods 2.2), were selected to be varied for the second stage

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

693 design. All the selected inputs are shown in Additional file *Fig S4*. In all, we  
694 obtained 23 active inputs from the first stage.

### 3.3 Metamodel based spatio-temporal sensitivity analysis

696 The outputs obtained after running the first (152 x 20 runs) and second (115  
697 x 20 runs) stage simulations, wherein x20 denotes the 20 replicates, were  
698 combined to be used as a training dataset. The combined output was utilized to  
699 build a Gaussian process based spatiotemporal metamodel (refer Methods 2.2),  
700 using *mlegp* package in R [38].

701  
702 The outputs from the training dataset were sub-divided into 6 datasets,  
703 corresponding to six time periods (Days 1-14, 15-21, 22-30, 31-42, 43-90, 91-  
704 201) and averaged across these periods. The sub-division of output across the  
705 time periods, aided the temporal analysis over the initiation (Day 1-14), peak of  
706 infection (Days 15-30) and chronic phase (post Day 31) stages as in [24]. We  
707 then fit a Gaussian process model (with nugget) and evaluated the performance  
708 of the fitting of the metamodel for *H. pylori*, resident macrophages, and  
709 monocyte-derived macrophages in *lamina propria* compartment, and tolerogenic  
710 DC in the gastric lymph node, using the diagnostic plots (see figures in Additional  
711 file, Fig S5). After fitting the models, we performed variance based global SA by  
712 computing the Sobol' total order and first order sensitivity index (refer Methods  
713 2.2). The estimates of the Sobol' total order indices for the input parameters  
714 calculated over the six time periods are shown in *Fig 4* (a-d).

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

**Fig 4. Heat-maps of Sobol' total order index for the input parameters across different output populations.**

The values in the heat-map indicate the Sobol' total order sensitivity index obtained from the metamodel, for the 38 input parameters with respect to the cell populations. The values with darker color indicate a stronger influence on the cell population as compared to the ones with lighter shade that indicate non-influential parameters for the cell populations - a) *H. pylori*, b) monocyte-derived macrophages, c) resident macrophages, in the lamina propria compartment and d) tolerogenic DCs, in the gastric lymph node compartment. The indexes are calculated over six time points ranging across the three stages of infection, including initiation (Day 1-14), peak (Days 15-42) and recovery stages (Days 43-201).

As shown in Fig 4a, the metamodel based global SA showed that the input parameters, epithelial cell proliferation (*EpiProlifer*) and epithelial cell death (*EpicellDeath*) had the strongest impact on the population of *H. pylori* in lamina propria compartment. As time progressed from initiation of the infection (Days 1-14), through peak (Days 15-30), the epithelial cell proliferation had a continued impact on the colonization of *H. pylori*. Next, the influence of the probability of epithelial cell death decreased over the course of infection. Further, Fig 4b highlighted the impact of epithelial cell proliferation (*EpiProlifer*) and epithelial cell death (*EpicellDeath*) on the monocyte-derived macrophages.

For the resident macrophage population in the lamina propria, that have emergent properties similar to the one characterized in [24], we observed that the resident macrophage replication parameter (*ResmMacRep*) has an impact during the initiation and peak stages of the infection which indicates that these

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

743 subsets of macrophages replicate during the course of *H. pylori* infection. This  
744 result highlights the reliability of the two-staged global SA method used here, as  
745 these findings are consistent with the ones in [24] wherein we observed that  
746 these subsets of macrophages expand in the gastric stomach lamina propria  
747 during the course of *H. pylori* infection.

748 Finally, for the tolerogenic DCs in *Fig 4d*, we observed that the epithelial cell  
749 death (*Epicelldeath*) seemed to have an impact. Another parameter that stands  
750 for the probability of naive T cell transitioning to iTreg cell (*nTtoiTreg*) was shown  
751 to have an impact on the tolerogenic dendritic cells. Tolerogenic dendritic cells  
752 are involved in the rule that transitions the naive T cells to iTreg cells in the gastric  
753 lymph node, and the stronger impact of the *nTtoiTreg* during the initiation and  
754 peak stages of the infection highlights the role of the tolerogenic dendritic cells  
755 during the course of infection.

756 The global SA data suggested that the main contributors of the chronic  
757 colonization of *H. pylori* in the lamina propria are the epithelial cells, specifically  
758 the epithelial cell proliferation parameter.

### 3.4 Effect of different ranges of epithelial cell proliferation

760 An interesting prediction derived from the metamodel based global SA is that  
761 epithelial cell proliferation is one of the parameters that has a strong impact on  
762 the size of *H. pylori* population. The biological hypothesis derived from this  
763 prediction is that the epithelial cell proliferation is responsible for the higher  
764 colonization of *H. pylori*. Prior to conducting any experimental studies, we wanted  
765 to explore the hypothesis using our hybrid computer model *in silico* and study  
766 the model outputs obtained after we changed the epithelial cell proliferation  
767 parameter. Thus, we varied the epithelial cell proliferation parameter across

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

768 different ranges (0.1-0.9, with 0.6 being the value for baseline conditions) and  
769 ran the simulations using the hybrid model and studied its effect on the different  
770 output cell population (obtained after running the simulations). These outputs  
771 were the ones obtained after running the simulation using the hybrid computer  
772 model, as we varied the epithelial cell proliferation parameter. We analyzed the  
773 outputs from the hybrid computer model and interestingly, observed that upon  
774 decreasing the *Epiprolifer* from a range of values 0.9-0.1, the output cell  
775 populations with regulatory function, namely regulatory macrophages and  
776 tolerogenic dendritic cells were found to vary. We observed a decreasing effect  
777 (*Fig 5a-d*) on *H. pylori*, monocyte-derived macrophages, resident macrophages  
778 in the lamina propria compartment and tolerogenic dendritic cells in gastric lymph  
779 node. Overall, these cell populations varied due to the variation in the epithelial  
780 cell proliferation parameter.

781 For clarification, such connection was not embedded in the mechanisms  
782 included in Table 1 but it represents an emergent behavior from the simulations  
783 predicting the involvement of regulatory and tolerogenic dendritic cells in the  
784 mechanisms of immunoregulation during *H. pylori* infection. Finally, the  
785 simulations targeting the epithelial cell proliferation resulted in changes in  
786 regulatory and tolerogenic dendritic cell populations. This shows that the  
787 simulations indirectly targeted the regulatory and tolerogenic dendritic cell  
788 population. Thus, we hypothesize that epithelial cell proliferation might be  
789 responsible for the higher colonization of *H. pylori* through an immunoregulatory  
790 mechanism that involves regulatory macrophages and tolerogenic cells. This is  
791 in line with our own conclusions drawn from a previous paper [24] where we  
792 show that the presence of cells with regulatory phenotype favor higher levels of  
793 *H. pylori* colonization. The results from the sensitivity analysis presented in this

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

794 paper suggest that epithelial proliferation might be a crucial part of the  
795 mechanisms by which these regulatory responses are induced and that there is  
796 a link between these parameters. The exact biological process however cannot  
797 be inferred from the current model and it will be investigated in follow-up in vivo  
798 studies.

**Fig 5. *In silico* study of the effect of epithelial cell proliferation parameter on the cell populations.**

*The plots show the effect of varying epithelial cell proliferation ( $p_{\text{Epi}}\text{prolifer}$ ) parameter (with values 0.1, 0.5, 0.6(WT), and 0.9) on the output cell population of a) *H. pylori*, b) tolerogenic dendritic cells, c) resident macrophages and d) monocyte-derived macrophages. The parameter has a decreasing effect on the cellular populations under consideration, wherein a decrease in the parameter value, decreases the abundance of the cells over time. The lower half of the figures (a-d), show the results for statistical comparison between the groups using ANOVA with the post-hoc analysis. The letters 'a', 'b', 'c', and 'bc' represent statistically significant differences ( $P < 0.05$ ) between the groups obtained after running the Tukey's Honestly Significant Difference.*

The *in silico* findings suggested the involvement of regulatory macrophages (both resident as well as monocyte-derived) and tolerogenic DC on the colonization of *H. pylori* in the gastric lamina propria. This highlighted and validated the role of epithelial cell proliferation as one of the main factor affecting *H. pylori* levels in the gastric niche.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

820

821 **4. Discussion**

822 *H. pylori* is the dominant indigenous bacterium of the gastric microbiota. In the  
823 majority of individuals, *H. pylori* colonizes the stomach without causing adverse  
824 effects, with little to no activation of inflammatory pathways. However, certain  
825 members of the population lose immune tolerance to the bacterium thereby  
826 contributing to the development of chronic gastric diseases. The immunological  
827 mechanisms underlying its ability to persist in a harsh acidic gastric environment and  
828 its dual role as a pathogen and beneficial organism remain unknown. A subset of  
829 macrophages helps create a regulatory microenvironment that promotes the chronic  
830 colonization of *H. pylori* [24]. However, the immune regulatory mechanisms are  
831 incompletely understood. Computational models of the immune system featuring  
832 immune responses are powerful tools for testing the different ‘what-if’ scenarios.  
833 Multiscale models of the immune response are attractive in terms of modeling the  
834 responses at different spatiotemporal scales [44].

835 In this study, we developed a HPC-driven hybrid, high-resolution, multiscale  
836 model to simulate the complex immunoregulatory mechanisms during *H. pylori*  
837 infection. The hybrid model was integrated with two intracellular ODEs capturing the  
838 dynamics of CD4+ T cells and regulatory macrophages. The inputs to the hybrid  
839 model are the set of parameters whose variation governs the immune system  
840 dynamics during infection. The obtained outputs were emergent patterns of different  
841 cell types, cytokines, and bacterial levels for instance the levels of *H. pylori*, and that  
842 qualitatively matched the patterns observed in an *in vivo* infection model [1, 24]. We  
843 presented an *in-silico* framework that evaluated the global SA of the hybrid model  
844 and studied how the variation in the biological parameters affected the simulation

1  
2  
3  
4 845 outputs. The two-stage global SA indicated that epithelial cell parameters,  
5  
6 846 specifically, the proliferation of epithelial cells affected the colonization of *H. pylori* in  
7  
8 847 the gastric mucosa. These results were validated *in silico*, and highlighted the  
9  
10 848 involvement of regulatory macrophages and tolerogenic DC in facilitating *H. pylori*  
11  
12 849 colonization of the gastric mucosa. Previous studies highlighted *H. pylori* inhabits the  
13  
14 850 apical surfaces of the epithelial cells and maintains a persistent infection [45].  
15  
16

17 851 Further, Mimuro et al. demonstrated that *H. pylori* promotes epithelial gastric cell  
18  
19 852 survival by attenuating apoptosis. These events showed how *H. pylori* regulated the  
20  
21 853 gastric niche and utilized epithelial cells to facilitate its persistence within the  
22  
23 854 stomach [46] [47]. Thus, the findings in the current study are in line with the literature  
24  
25 855 that suggests epithelial cell proliferation favor the colonization of *H. pylori* in the  
26  
27 856 stomach.  
28  
29  
30

31 857 Our group also showed another mechanism used by *H. pylori* to create a gut  
32  
33 858 microenvironment that involved the induction of IL-10-driven regulatory mechanism  
34  
35 859 mediated by CD11b<sup>+</sup>F4/80<sup>hi</sup>CD64<sup>+</sup>CX<sub>3</sub>CR1<sup>+</sup> mononuclear phagocytes, which  
36  
37 860 facilitated bacterial colonization [24]. Additionally, in this paper, we reported that  
38  
39 861 regulatory macrophages were involved in the process of colonization with *H. pylori*  
40  
41 862 when we varied the epithelial cell proliferation parameter *in-silico*. Zhang et al.,  
42  
43 863 demonstrated that *H. pylori* directed active tolerogenic programming of DCs that  
44  
45 864 favored chronic bacterial colonization, by altering the balance of Th17/Treg cells [48].  
46  
47 865 Rizzuti, Ang et al., demonstrated *H. pylori*-mediated IL-10 release caused the  
48  
49 866 activation of signal transducer and activator of transcription 3 (STAT3) in DC. This  
50  
51 867 activation of STAT3 via IL-10 release was shown to induce the production of  
52  
53 868 tolerogenic DC phenotype [49]. The findings from this paper also indicated the  
54  
55 869 involvement of tolerogenic DCs in affecting the mucosal levels of *H. pylori*.  
56  
57  
58 870 Therefore, the literature combined with the results from this study, collectively  
59  
60  
61  
62  
63  
64  
65



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

871 suggest that during *H. pylori* infection, the epithelial cell favors the colonization of *H.*  
872 *pylori* by creating a regulatory microenvironment. This process is mediated by the  
873 regulatory macrophages and tolerogenic programming of DC. Based on the results  
874 from this paper and findings from the literature, this leads us to propose that the  
875 induction of IL-10 by the regulatory macrophages is potentially involved in directing  
876 the tolerogenic programming of DC. All experimental evidence combined with our  
877 model prediction suggest the action of an underlying biological mechanism that links  
878 the presence of *H. pylori* in the gastric mucosa with changes in the rates of epithelial  
879 cell proliferation which ultimately affects the levels of colonization. Our prediction  
880 points towards a link between epithelial cell proliferation and the action of tolerogenic  
881 dendritic cells and regulatory macrophages. The exact cellular mechanism induced  
882 during this process however cannot be inferred from the current model and it will be  
883 investigated in follow-up in vivo studies.

884

885 At its current stage, the hybrid ENISI model reproduces the overall immune  
886 system dynamics observed during an *H. pylori* infection. The parameters of  
887 calibrated ODEs were kept unchanged, whereas the ABM parameters were  
888 calibrated by qualitatively matching the patterns of the output simulations as  
889 observed in an *in vivo* model of *H. pylori* infection [24]. For ABM, its calibration and  
890 validation remain the major key issues, discussed elsewhere [21] [50] [51]. Further,  
891 developing targeted methods of SA have been identified as an important challenge  
892 in the field [21, 52, 53]. In this paper, we highlighted the use of SA methods with a  
893 two-stage global SA framework comprised of first, screening the input parameters  
894 (using PRCC) and second, building of a surrogate model (using GP) of the hybrid  
895 model, to understand the emergent behavior of the represented system. It is  
896 important to note that each SA method known, has its own merits and produces

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

897 useful information however none provide a complete picture of the emergent model  
898 behavior [21]. First, we employed PRCC methods as the initial step in our two staged  
899 SA that aided the screening of active inputs and reduced the parameter space. The  
900 choice of PRCC was advantageous and justified by the low computational cost and  
901 low complexity in the computation of the coefficients. Another advantage of the  
902 regression-based PRCC method is that the complex output from our hybrid model  
903 was condensed into a descriptive relationship that can be described by statistical  
904 measures such as  $R^2$  [21]. As described in [21] the results from PRCC are good  
905 descriptors of the outputs produced if the regression function constitutes a good fit  
906 to the output [21]. However, if the function does not yield a good fit, the regression-  
907 based SA are proven to be useful in screening the influential parameters for further  
908 analysis [21], as described in our analysis.

909 Further, the interaction effects between the parameters are not considered in  
910 regression-based methods, and hence it was followed by the use of variance-based  
911 methods in later stage analysis. Second, we employed metamodeling-based  
912 approach and Sobol' method as they provided information on the interaction  
913 between the input variable and the use of metamodels allowed to compute the  
914 sensitivity indices. One of the advantages of the Sobol' method is that it is model-  
915 free and no fitting functions are used to decompose the output variance [31]. It  
916 considers the averaged effect of parameters over the whole parameter space but  
917 fails to explore the different patterns within the space [21]. Further, the method is not  
918 suitable for quantification of output variability if the output distributions deviate from  
919 a normal distribution [21]. The detailed comparison of different SA methods used for  
920 the global SA of ABMs are described in detail in [21]. Thus, we performed both the  
921 PRCC and computation of Sobol' indices approaches to evaluate the influence of

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

922 the input parameter variation and identified the parameters involved in the successful  
923 colonization of the gastric niche by *H. pylori*.

924 Some limitations of the model include implementation through a two-dimensional  
925 grid system and including all cells of the same size. Although we parallelize the  
926 computation of the hybrid model output, the large number of simulations required for  
927 the global SA compensates for the benefits of parallelization. To improve the  
928 calibration process and overall usability of the model, the data required for model  
929 calibration would include tissue biopsies from people infected with *H. pylori* that can  
930 be used to quantify the cells and take into account their spatial arrangement. The  
931 current version is also limited in terms of the interactions that are based on epithelial  
932 cells and DC as they are strictly rule-based. The building of ODE models for these  
933 cells and integrating them with the ABM model will help capture the dynamics of  
934 epithelial cells and DC more in-depth. Overall the immunoregulatory mechanisms  
935 underlying the chronic colonization of *H. pylori* and the predictive capacity of the  
936 model can be further improved by incorporating cell-specific models for epithelial  
937 cells and DC.

938 In summary, a high-resolution, hybrid, multiscale spatiotemporal stochastic  
939 model of *H. pylori* infection was built and global SA was performed. The results from  
940 the global SA highlight the key role played by epithelial cells in affecting the levels of  
941 *H. pylori* colonization. The *in-silico* validation of varying the epithelial cell proliferation  
942 parameter demonstrated the involvement of regulatory macrophages and the  
943 tolerogenic DC. The next steps aimed to enrich the model will involve the validation  
944 of the findings *in vivo* to study the underlying mechanisms involved in the successful  
945 immune evasion by *H. pylori*. The computational modeling predictions will be further  
946 validated experimentally and clinically.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

## 5. Potential Implications

The computational model of the gut contains high-resolution information processing representations of immune responses that are generalizable for other infectious and autoimmune diseases. Complex diseases such as autoimmune disorders, infectious diseases, and cancer all require integration of the multiscale level data, information and knowledge, ranging from genes, proteins, cells, tissue to organ level. The ENISI model of the gut presented here can be generalized to other diseases by implementing the agents and rules specific to that disease, plus recalibrating the model based on data that are specific to the new indication. Since ABMs have modular architectures, an addition of new agent-types and modification of rules can be done without restructuring the entire simulation setup [19]. The use of ABM in such hybrid models not only facilitates the implementation of already known mechanisms but also helps validate and predict any unforeseen new mechanisms using data analytics methods such as global SA to analyze emerging behaviors at the systems level. The finer details regarding intracellular and intercellular interactions that contribute towards the nonlinear and complex behavior of the gut can also be studied by integrating the intracellular ODE models as implemented here.

### Tables

Name of Agent	Agent Type	Rules
<i>Helicobacter pylori</i>	<i>H. pylori</i>	- Moves across the epithelial cell border if near damaged epithelial layer - Proliferates in the lumen and lamina propria - Dies (removed from the simulation) in lamina propria and in the lumen due to the damage of epithelial cells by Th1 or Th17 cells
Macrophages	Monocyte	- Proliferates in presence the of effector dendritic cells or damaged epithelial cells

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

		<ul style="list-style-type: none"> <li>- Proliferates in the lamina propria</li> <li>- Differentiates to regulatory macrophage in based on the output from the Macrophage ODE</li> <li>- Differentiates to inflammatory macrophages in presence of IFN-<math>\gamma</math></li> <li>- Dies naturally (removed from the model)</li> </ul>
	Resident	<ul style="list-style-type: none"> <li>- Proliferates in the presence of <i>H. pylori</i></li> <li>- Secretes IL10</li> <li>- Dies naturally</li> <li>- Dies due to Th1 and Tr cells</li> </ul>
	Regulatory	<ul style="list-style-type: none"> <li>- Proliferates and removes bacteria</li> <li>- Dies</li> <li>- Secretes IL10</li> </ul>
	Inflammatory	<ul style="list-style-type: none"> <li>-Proliferates in the presence of damaged epithelial cell</li> <li>-Dies naturally</li> </ul>
Dendritics	Immature	<ul style="list-style-type: none"> <li>-Moves from lamina propria to epithelium compartment and from the epithelium to the lamina propria</li> <li>- Differentiates to tolerogenic dendritic cell in the presence of tolerogenic bacteria, both in epithelium and lamina propria</li> <li>- Differentiates to effector dendritic cell in the presence of <i>H. pylori</i></li> <li>- Proliferates in lamina propria and gastric lymph node</li> <li>- Dies naturally</li> </ul>
	Effector	<ul style="list-style-type: none"> <li>- Moves from lamina propria to gastric lymph node</li> <li>- Moves form epithelium to lamina propria</li> <li>- Secretes IL6 and IL12</li> <li>- Dies naturally</li> </ul>
	Tolerogenic	<ul style="list-style-type: none"> <li>- Moves from lamina propria to gastric lymph node</li> <li>- Moves from epithelium to lamina propria</li> <li>- Secretes TGF-<math>\beta</math></li> <li>- Dies naturally</li> </ul>
T cells	Naïve	<p>In the presence of effector dendritic cells:</p> <ul style="list-style-type: none"> <li>- Differentiates to Th1 in the presence of IFN-<math>\gamma</math> or IL12</li> <li>- Differentiates to Th17 in the presences of IL6 or TGF-<math>\beta</math></li> </ul> <p>In the presence of tolerogenic dendritic cells:</p> <ul style="list-style-type: none"> <li>- Differentiates to iTreg in the presence of TGF-<math>\beta</math></li> <li>- Differentiates to Tr in the presences of IL10</li> <li>- Dies naturally</li> </ul>
	Th1	<ul style="list-style-type: none"> <li>- Secretes IFN-<math>\gamma</math></li> <li>- Moves from gastric lymph node to lamina propria</li> <li>- Proliferates in lamina propria and gastric lymph node</li> <li>- Dies naturally</li> </ul>
	Th17	<ul style="list-style-type: none"> <li>- Secretes IL17</li> <li>- In the presence of tolerogenic dendritic cell, transition to iTreg cells</li> <li>- Moves from gastric lymph node to lamina propria</li> <li>- Proliferates in lamina propria and gastric lymph node</li> <li>- Dies naturally</li> </ul>
	iTreg	<ul style="list-style-type: none"> <li>- Secretes IL10</li> </ul>

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

		<ul style="list-style-type: none"> <li>- In the presence of tolerogenic dendritic cell, transition to iTreg cells</li> <li>- Moves from gastric lymph node to lamina propria</li> <li>- Proliferates in lamina propria and gastric lymph node</li> <li>- Dies naturally</li> </ul>
	Tr	<ul style="list-style-type: none"> <li>- Secretes IL10</li> <li>- Dies naturally</li> <li>- Proliferates in the lamina propria</li> </ul>
Epithelial	Healthy	<ul style="list-style-type: none"> <li>- Damaged due to infectious bacteria</li> <li>- Damaged due to Th1 and Th17 cells</li> <li>- Proliferates</li> <li>- Secretes IL6 and IL12</li> <li>- Dies naturally</li> </ul>
	Damaged	<ul style="list-style-type: none"> <li>- Transitions to healthy state in the presence of IL10</li> <li>- Dies naturally</li> </ul>
Bacteria	Infectious	<ul style="list-style-type: none"> <li>- Dies due to Th1 or Th17 or inflammatory macrophages or damaged epithelial cells</li> <li>- Dies naturally</li> <li>- Proliferates in the lamina propria</li> </ul>
	Tolerogenic	<ul style="list-style-type: none"> <li>- Moves from lumen to the epithelium in the presence of damaged epithelial cells</li> <li>- Becomes infectious if moves in the lamina propria compartment</li> <li>- Proliferates in lumen and lamina propria</li> <li>- Dies naturally</li> </ul>

968

969 ***Table 1. A list of rules for all the agent types implemented in the hybrid model***

970

971

972 **Additional Files**

973 File S1

974 Fig S1

975 Table S1

976 Fig S2

977 Fig S3

978 Fig S4

979 Fig S5

1  
2  
3  
4 980  
5  
6  
7  
8 981  
9  
10 982  
11  
12  
13 983  
14  
15  
16  
17 984  
18  
19 985  
20  
21 986  
22  
23 987  
24  
25 988  
26  
27 989  
28  
29 990  
30  
31  
32 991  
33  
34 992  
35  
36  
37 993  
38  
39 994  
40  
41 995  
42  
43  
44 996  
45  
46 997  
47  
48 998  
49  
50 999  
51  
52 1000  
53  
54 1001  
55  
56 1002  
57  
58 1003  
59  
60  
61  
62  
63  
64  
65

*File S1 – The detailed instruction to Install ENISI MSM (Step I), Run a simulation (Step II) and Conduct Sensitivity Analysis (Step III) are described.*

**Fig S1. Design implementation of the hybrid multiscale model used to simulate *Helicobacter pylori* infection**

*The figure shows the class structure used in the ENISI MSM hybrid agent based-ODE model. Each group consists of an act() function that includes the implemented rule for each agent. The previously published ODE models for T cells and Macrophage are used to integrate in the ABM code.*

**Table S1** *Table describing the input parameters used in the sensitivity analysis and their ranges used.*

**Fig S2. A pictorial representation of the spatial discretization of the 2D grid.**

**Fig S3. Flowchart for the two-staged global sensitivity analysis.**

**Fig S4. The active and inactive inputs selected from the stage 1 analysis**

*The rows represent the input parameters and columns represent the output cell populations. The green boxes highlight the ‘active’ input parameters (row) that are shown to have a significant influence (calculated based on the results obtained from partial correlation coefficient analysis), on an output cell (columns) under consideration.*

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

**Fig S5. Diagnostic and residual plots obtained for the Gaussian processes fitted metamodels**

The upper panel represents the diagnostic Q-Q plots where the open circles represent the cross-validated predictions; solid black lines represent observed response. The “observed simulations” data in the first half of the lower panel, refer to the observed output values of the simulations obtained after running the hybrid computer model, whereas the y axis refers to the predicted simulation values obtained from the Cross-validated model. Each point represents 1 output point obtained as an output from the simulation. The second half of the lower panel, refers to the standard residual plot wherein the x-axis represents the observed simulation values obtained from the simulation and the y-axis refers to the residual error ((error (predicted values – observed values) / standard deviation (error))) obtained. The diagnostic plots denote the black circles which are the cross-validated prediction. Cross-validation is in the sense that for predictions made at design point  $x$ , all observations at design point  $x$  are removed from the training set. The lower panel represents the residual plots for the cell populations –(a) *Helicobacter pylori*; (b) Resident macrophages; (c) Monocyte-derived macrophages in the Lamina propria and (d) Tolerogenic dendritic cells in the Gastric lymph node compartment.

**Data and materials**

The data sets and files supporting the results of this article are available in the ENISI-MSM GitHub repository, RRID: SCR\_016918 <https://github.com/NIMML/ENISI-MSM>.

**Availability of source code and requirements**



- 1  
2  
3  
4 1029 • Project Name: ENISI MSM  
5  
6 1030 • Project homepage: <https://github.com/NIMML/ENISI-MSM>  
7  
8  
9 1031 • Operating system(s): Linux, Mac OSX  
10  
11 1032 • Programming language: C++, R, MATLAB  
12  
13 1033 • Other requirements: CMake 3.7.2,  
14  
15 ENISI Dependencies <https://github.com/NIMML/ENISI-Dependencies>  
16 1034  
17  
18 1035 • License: Apache License 2.0  
19  
20 1036 • RRID: SCR\_016918  
21  
22

23  
24 1037 **Declarations**

25  
26 1038 **List of abbreviations**

- 27  
28 1039 ABM – Agent based model  
29  
30 1040 DC – Dendritic cells  
31  
32 1041 ENISI MSM – Enteric Immunity Simulator Multi-scale Modeling  
33  
34 1042 GLN – gastric lymph node  
35  
36 1043 GP - Gaussian process  
37  
38 1044 *H. pylori* – *Helicobacter pylori*  
39  
40 1045 HPC – High performance computing  
41  
42 1046 LP – Lamina propria  
43  
44 1047 ODE – Ordinary Differential Equation  
45  
46 1048 PDE – Partial Differential Equation  
47  
48 1049 SA – Sensitivity analysis  
49  
50 1050 PRCC - Partial rank correlation coefficient  
51  
52

53 1051 **Consent for publication**

54  
55 1052 Not applicable.  
56  
57

58  
59 1053 **Competing interests**  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 1054 The author(s) declare that they have no competing interests.  
5

6 1055 **Funding**  
7

8 1056 This work was supported by the Defense Threat Reduction Agency (DTRA) grant  
9  
10 1057 HDTRA1-18-1-0008 to JBR and RH and funds from the Nutritional Immunology and  
11  
12 1058 Molecular Medicine Laboratory ([www.nimml.org](http://www.nimml.org)). The funding body had no role in  
13  
14 1059 the design of the study, data collection, analysis, interpretation of data and writing of  
15  
16 1060 the manuscript.  
17  
18

19 1061 **Authors' contributions**  
20

21 1062 MV, RH and JBR formulated the model, implemented, performed the simulations,  
22  
23 1063 analyzed model-generate outputs, made the figures and wrote the manuscript. MV,  
24  
25 1064 AL, JBR, RH, and SH formulated the model. SH, AL and VA implemented the code  
26  
27 1065 architecture and benchmarked the parallel version of the hybrid model. XC and MV  
28  
29 1066 wrote the codes for global sensitivity analysis and generated the design matrices.  
30  
31 1067 NTJ generated macrophage and *H. pylori* experimental data. JBR, VA, and RH  
32  
33 1068 supervised the project. JBR and RH edited the manuscript. JBR, AL, NTJ, SH, VA,  
34  
35 1069 XC and RH participated in discussions on the model and results. All authors provided  
36  
37 1070 critical feedback on the project.  
38  
39  
40

41 1071 **Acknowledgements**  
42

43 1072 Not applicable.  
44  
45

46 1073  
47

48 1074 **References**  
49

- 50  
51 1075 1. Carbo A, Bassaganya-Riera J, Pedragosa M, Viladomiu M, Marathe M, Eubank S, et al.  
52  
53 1076 Predictive computational modeling of the mucosal immune responses during *Helicobacter pylori*  
54  
55 1077 infection. PLoS One. 2013;8(9):e73365. doi: 10.1371/journal.pone.0073365.  
56  
57  
58 1078 2. Leber A, Bassaganya-Riera J, Tubau-Juni N, Zoccoli-Rodriguez V, Viladomiu M, Abedi V,  
59  
60 1079 et al. Modeling the Role of Lanthionine Synthetase C-Like 2 (LANCL2) in the Modulation of  
61  
62

- 1  
2  
3  
41080 Immune Responses to Helicobacter pylori Infection. PloS one. 2016;11(12):e0167440. Epub  
5  
61081 2016/12/10. doi: 10.1371/journal.pone.0167440. PubMed PMID: 27936058; PubMed Central  
7  
81082 PMCID: PMCPMC5147901.
- 10  
111083 3. Leber A, Viladomiu M, Hontecillas R, Abedi V, Philipson C, Hoops S, et al. Systems  
12  
131084 Modeling of Interactions between Mucosal Immunity and the Gut Microbiome during Clostridium  
14  
151085 difficile Infection. PloS one. 2015;10(7):e0134849. Epub 2015/08/01. doi:  
16  
171086 10.1371/journal.pone.0134849. PubMed PMID: 26230099; PubMed Central PMCID:  
18  
191087 PMCPMC4521955.
- 21  
221088 4. Verma M, Erwin S, Abedi V, Hontecillas R, Hoops S, Leber A, et al. Modeling the  
23  
241089 Mechanisms by Which HIV-Associated Immunosuppression Influences HPV Persistence at the  
25  
261090 Oral Mucosa. PloS one. 2017;12(1):e0168133. Epub 2017/01/07. doi:  
27  
281091 10.1371/journal.pone.0168133. PubMed PMID: 28060843; PubMed Central PMCID:  
29  
301092 PMCPMC5218576.
- 32  
331093 5. Qomlaqi M, Bahrami F, Ajami M, Hajati J. An extended mathematical model of tumor  
34  
351094 growth and its interaction with the immune system, to be used for developing an optimized  
36  
371095 immunotherapy treatment protocol. Mathematical biosciences. 2017;292:1-9. Epub 2017/07/18.  
38  
391096 doi: 10.1016/j.mbs.2017.07.006. PubMed PMID: 28713023.
- 41  
421097 6. Vodovotz Y, Xia A, Read EL, Bassaganya-Riera J, Hafler DA, Sontag E, et al. Solving  
43  
441098 Immunology? Trends in immunology. 2017;38(2):116-27. doi: 10.1016/j.it.2016.11.006. PubMed  
45  
461099 PMID: PMC5695553.
- 48  
491100 7. Kusters JG, van Vliet AH, Kuipers EJ. Pathogenesis of Helicobacter pylori infection.  
50  
511101 Clinical microbiology reviews. 2006;19(3):449-90. Epub 2006/07/19. doi: 10.1128/cmr.00054-05.  
52  
531102 PubMed PMID: 16847081; PubMed Central PMCID: PMCPMC1539101.
- 54  
551103 8. Mane S, Dominguez-Bello M, Blaser M, Sobral B, Hontecillas R, Skoneczka J, et al. Host-  
56  
571104 interactive genes in Amerindian Helicobacter pylori diverge from their Old World homologs and  
58  
591105 mediate inflammatory responses. Journal of bacteriology. 2010;192(12):3078-92.

1  
2  
3  
4 1106 9. Cover TL, Blaser MJ. Helicobacter pylori in health and disease. Gastroenterology.  
5  
6 1107 2009;136(6):1863-73. doi: 10.1053/j.gastro.2009.01.073. PubMed PMID: PMC3644425.  
7  
8  
9 1108 10. Bassaganya-Riera J, Dominguez-Bello MG, Kronsteiner B, Carbo A, Lu P, Viladomiu M,  
10  
11 1109 et al. Helicobacter pylori colonization ameliorates glucose homeostasis in mice through a PPAR  
12  
13 1110  $\gamma$ -dependent mechanism. PLoS One. 2012;7(11):e50069. doi: 10.1371/journal.pone.0050069.  
14  
15 1111 11. Oertli M, Sundquist M, Hitzler I, Engler DB, Arnold IC, Reuter S, et al. DC-derived IL-18  
16  
17 1112 drives Treg differentiation, murine Helicobacter pylori-specific immune tolerance, and asthma  
18  
19 1113 protection. The Journal of clinical investigation. 2012;122(3):1082-96. Epub 2012/02/07. doi:  
20  
21 10.1172/jci61029. PubMed PMID: 22307326; PubMed Central PMCID: PMC3287234.  
22  
23  
24 1115 12. Mei Y, Abedi V, Carbo A, Zhang X, Lu P, Philipson C, et al. Multiscale modeling of mucosal  
25  
26 1116 immune responses. BMC bioinformatics. 2015;16 Suppl 12:S2. Epub 2015/09/04. doi:  
27  
28 10.1186/1471-2105-16-s12-s2. PubMed PMID: 26329787; PubMed Central PMCID:  
29  
30 PMCPMC4705510.  
31  
32  
33 1119 13. Gong C, Milberg O, Wang B, Vicini P, Narwal R, Roskos L, et al. A computational  
34  
35 1120 multiscale agent-based model for simulating spatio-temporal tumour immune response to PD1  
36  
37 and PDL1 inhibition. Journal of the Royal Society, Interface. 2017;14(134). Epub 2017/09/22. doi:  
38  
39 10.1098/rsif.2017.0320. PubMed PMID: 28931635; PubMed Central PMCID: PMC5636269.  
40  
41  
42 1123 14. Wang Z, Birch CM, Sagotsky J, Deisboeck TS. Cross-scale, cross-pathway evaluation  
43  
44 1124 using an agent-based non-small cell lung cancer model. Bioinformatics (Oxford, England).  
45  
46 1125 2009;25(18):2389-96. Epub 2009/07/07. doi: 10.1093/bioinformatics/btp416. PubMed PMID:  
47  
48 19578172; PubMed Central PMCID: PMC2735669.  
49  
50  
51 1127 15. Marino S, El-Kebir M, Kirschner D. A hybrid multi-compartment model of granuloma  
52  
53 1128 formation and T cell priming in tuberculosis. Journal of theoretical biology. 2011;280(1):50-62.  
54  
55 1129 Epub 2011/03/30. doi: 10.1016/j.jtbi.2011.03.022. PubMed PMID: 21443879; PubMed Central  
56  
57 PMCID: PMC3740747.  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 1131 16. Solovyev A, Mi Q, Tzen YT, Brienza D, Vodovotz Y. Hybrid equation/agent-based model  
5  
6 1132 of ischemia-induced hyperemia and pressure ulcer formation predicts greater propensity to  
7  
8 1133 ulcerate in subjects with spinal cord injury. PLoS computational biology. 2013;9(5):e1003070.  
9  
10 Epub 2013/05/23. doi: 10.1371/journal.pcbi.1003070. PubMed PMID: 23696726; PubMed Central  
11 1134  
12 PMCID: PMCPMC3656105.  
13 1135  
14  
15 1136 17. Cappuccio A, Tieri P, Castiglione F. Multiscale modelling in immunology: a review.  
16  
17 Briefings in bioinformatics. 2016;17(3):408-18. Epub 2015/03/27. doi: 10.1093/bib/bbv012.  
18 1137  
19 PubMed PMID: 25810307.  
20 1138  
21  
22 1139 18. Bassaganya-Riera J. Computational Immunology: Models and Tools: Academic Press;  
23  
24 1140 2015.  
25  
26 1141 19. An G, Mi Q, Dutta-Moscato J, Vodovotz Y. Agent-based models in translational systems  
27  
28 biology. Wiley interdisciplinary reviews Systems biology and medicine. 2009;1(2):159-71. Epub  
29 1142  
30 2010/09/14. doi: 10.1002/wsbm.45. PubMed PMID: 20835989; PubMed Central PMCID:  
31 1143  
32 PMCPMC3640333.  
33 1144  
34  
35 1145 20. Abedi V, Hontecillas R, Hoops S, Liles N, Carbo A, Lu P, et al., editors. ENISI multiscale  
36  
37 modeling of mucosal immune responses driven by high performance computing. 2015 IEEE  
38 1146  
39 International Conference on Bioinformatics and Biomedicine (BIBM); 2015 9-12 Nov. 2015.  
40 1147  
41  
42 1148 21. Ten Broeke G, Van Voorn G, Ligtenberg A. Which sensitivity analysis method should I use  
43  
44 1149 for my agent-based model? Journal of Artificial Societies and Social Simulation. 2016;19(1):5.  
45  
46 1150 22. Ligmann-Zielinska A, Kramer DB, Spence Cheruvellil K, Soranno PA. Using uncertainty  
47  
48 and sensitivity analyses in socioecological agent-based models to improve their analytical  
49 1151  
50 performance and policy relevance. PLoS One. 2014;9(10):e109779. Epub 2014/10/24. doi:  
51 1152  
52 10.1371/journal.pone.0109779. PubMed PMID: 25340764; PubMed Central PMCID:  
53 1153  
54 PMCPMC4207681.  
55 1154  
56  
57 1155 23. Carbo A, Hontecillas R, Kronsteiner B, Viladomiu M, Pedragosa M, Lu P, et al. Systems  
58  
59 modeling of molecular mechanisms controlling cytokine-driven CD4+ T cell differentiation and  
60 1156  
61  
62  
63  
64  
65

1  
2  
3  
4 1157 phenotype plasticity. PLoS computational biology. 2013;9(4):e1003027. Epub 2013/04/18. doi:  
5  
6 1158 10.1371/journal.pcbi.1003027. PubMed PMID: 23592971; PubMed Central PMCID:  
7  
8 1159 PMCPMC3617204.

10  
11 1160 24. Viladomiu M, Bassaganya-Riera J, Tubau-Juni N, Kronsteiner B, Leber A, Philipson CW,  
12  
13 1161 et al. Cooperation of Gastric Mononuclear Phagocytes with Helicobacter pylori during  
14  
15 1162 Colonization. Journal of immunology (Baltimore, Md : 1950). 2017;198(8):3195-204. Epub  
16  
17 1163 2017/03/08. doi: 10.4049/jimmunol.1601902. PubMed PMID: 28264969; PubMed Central  
18  
19 1164 PMCID: PMCPMC5380565.

21  
22 1165 25. Collier N, North M. Repast HPC: A platform for large-scale agentbased modeling. Large-  
23  
24 1166 Scale Computing Techniques for Complex System Simulations. 2011:81-110.

26  
27 1167 26. Hoops S, Sahle S, Gauges R, Lee C, Pahle J, Simus N, et al. COPASI—a complex  
28  
29 1168 pathway simulator. Bioinformatics (Oxford, England). 2006;22(24):3067-74.

30  
31 1169 27. Saltelli A, Tarantola S, Campolongo F. Sensitivity analysis as an ingredient of modeling.  
32  
33 1170 Statistical Science. 2000;15(4):377-95.

35  
36 1171 28. Thorne BC, Bailey AM, Peirce SM. Combining experiments with multi-cell agent-based  
37  
38 1172 modeling to study biological tissue patterning. Briefings in Bioinformatics. 2007;8(4):245-57. doi:  
39  
40 1173 10.1093/bib/bbm024.

41  
42 1174 29. Moon H, Dean AM, Santner TJ. Two-stage sensitivity-based group screening in computer  
43  
44 1175 experiments. Technometrics. 2012;54(4):376-87.

46  
47 1176 30. Marino S, Hogue IB, Ray CJ, Kirschner DE. A methodology for performing global  
48  
49 1177 uncertainty and sensitivity analysis in systems biology. Journal of theoretical biology.  
50  
51 1178 2008;254(1):178-96. Epub 2008/06/24. doi: 10.1016/j.jtbi.2008.04.011. PubMed PMID:  
52  
53 1179 18572196; PubMed Central PMCID: PMCPMC2570191.

55  
56 1180 31. Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, et al. Global sensitivity  
57  
58 1181 analysis: the primer: John Wiley & Sons; 2008.

- 1  
2  
3  
4 1182 32. Rasmussen CE, Williams CK. Gaussian processes for machine learning. 2006. The MIT  
5  
6 1183 Press, Cambridge, MA, USA. 2006;38:715-9.  
7  
8 1184 33. Santner TJ, Williams BJ, Notz WI. The design and analysis of computer experiments:  
9  
10 Springer Science & Business Media; 2013.  
11 1185  
12  
13 1186 34. Thiele JC, Kurth W, Grimm V. Facilitating parameter estimation and sensitivity analysis of  
14  
15 1187 agent-based models: A cookbook using NetLogo and R. Journal of Artificial Societies and Social  
16  
17 1188 Simulation. 2014;17(3):11.  
18  
19 1189 35. Ankenman B, Nelson BL, Staum J. Stochastic kriging for simulation metamodeling.  
20  
21 Operations research. 2010;58(2):371-82.  
22 1190  
23  
24 1191 36. Lamoureux B, Mechbal N, Massé J-R. A combined sensitivity analysis and kriging  
25  
26 1192 surrogate modeling for early validation of health indicators. Reliability Engineering & System  
27  
28 1193 Safety. 2014;130:12-26.  
29  
30  
31 1194 37. Chen X, Kim K-K. Stochastic kriging with biased sample estimates. ACM Trans Model  
32  
33 1195 Comput Simul. 2014;24(2):1-23. doi: 10.1145/2567893.  
34  
35 1196 38. Dancik GM, Dorman KS. mlegp: statistical analysis for computer models of biological  
36  
37 1197 systems using R. Bioinformatics (Oxford, England). 2008;24(17):1966-7.  
38  
39  
40 1198 39. Saltelli A, Annoni P, Azzini I, Campolongo F, Ratto M, Tarantola S. Variance based  
41  
42 1199 sensitivity analysis of model output. Design and estimator for the total sensitivity index. Computer  
43  
44 1200 Physics Communications. 2010;181(2):259-70.  
45  
46  
47 1201 40. Sobol IM. Sensitivity estimates for nonlinear mathematical models. Mathematical  
48  
49 1202 modelling and computational experiments. 1993;1(4):407-14.  
50  
51 1203 41. Sobol' IM, Tarantola S, Gatelli D, Kucherenko SS, Mauntz W. Estimating the  
52  
53 1204 approximation error when fixing unessential factors in global sensitivity analysis. Reliability  
54  
55 1205 Engineering & System Safety. 2007;92(7):957-60. doi: <https://doi.org/10.1016/j.ress.2006.07.001>.  
56  
57  
58 1206 42. Jansen MJ. Analysis of variance designs for model output. Computer Physics  
59  
60 1207 Communications. 1999;117(1-2):35-43.  
61  
62  
63  
64  
65

- 1  
2  
3  
4 1208 43. Alam M, Deng X, Philipson C, Bassaganya-Riera J, Bisset K, Carbo A, et al. Sensitivity  
5  
6 1209 Analysis of an ENteric Immunity Simulator (ENISI)-Based Model of Immune Responses to  
7  
8 1210 *Helicobacter pylori* Infection. *PLoS one*. 2015;10(9):e0136139. Epub 2015/09/04. doi:  
9  
10 10.1371/journal.pone.0136139. PubMed PMID: 26327290; PubMed Central PMCID:  
11 1211 PMCPMC4556515.  
12  
13 1212  
14  
15 1213 44. Heiner M, Gilbert D. BioModel engineering for multiscale Systems Biology. *Progress in*  
16  
17 1214 *Biophysics and Molecular Biology*. 2013;111(2):119-28. doi:  
18  
19 <https://doi.org/10.1016/j.pbiomolbio.2012.10.001>.  
20 1215  
21  
22 1216 45. Alzahrani S, Lina TT, Gonzalez J, Pinchuk IV, Beswick EJ, Reyes VE. Effect of  
23  
24 1217 *Helicobacter pylori* on gastric epithelial cells. *World Journal of Gastroenterology : WJG*.  
25  
26 1218 2014;20(36):12767-80. doi: 10.3748/wjg.v20.i36.12767. PubMed PMID: PMC4177462.  
27  
28  
29 1219 46. Mimuro H, Suzuki T, Nagai S, Rieder G, Suzuki M, Nagai T, et al. *Helicobacter pylori*  
30  
31 1220 dampens gut epithelial self-renewal by inhibiting apoptosis, a bacterial strategy to enhance  
32  
33 1221 colonization of the stomach. *Cell host & microbe*. 2007;2(4):250-63. Epub 2007/11/17. doi:  
34  
35 1222 10.1016/j.chom.2007.09.005. PubMed PMID: 18005743.  
36  
37  
38 1223 47. Wroblewski LE, Peek RM. Orchestration of Dysregulated Epithelial Turnover by a  
39  
40 1224 Manipulative Pathogen. *Cell Host & Microbe*. 2007;2(4):209-11. doi:  
41  
42 1225 <https://doi.org/10.1016/j.chom.2007.09.011>.  
43  
44  
45 1226 48. Zhang M, Liu M, Luther J, Kao JY. *Helicobacter pylori* directs tolerogenic programming of  
46  
47 1227 dendritic cells. *Gut microbes*. 2010;1(5):325-9. Epub 2011/02/18. doi: 10.4161/gmic.1.5.13052.  
48  
49 1228 PubMed PMID: 21327041; PubMed Central PMCID: PMCPMC3023617.  
50  
51 1229 49. Rizzuti D, Ang M, Sokollik C, Wu T, Abdullah M, Greenfield L, et al. *Helicobacter pylori*  
52  
53 1230 inhibits dendritic cell maturation via interleukin-10-mediated activation of the signal transducer  
54  
55 and activator of transcription 3 pathway. *Journal of innate immunity*. 2015;7(2):199-211. Epub  
56 1231 2014/11/22. doi: 10.1159/000368232. PubMed PMID: 25412627.  
57  
58 1232  
59  
60  
61  
62  
63  
64  
65



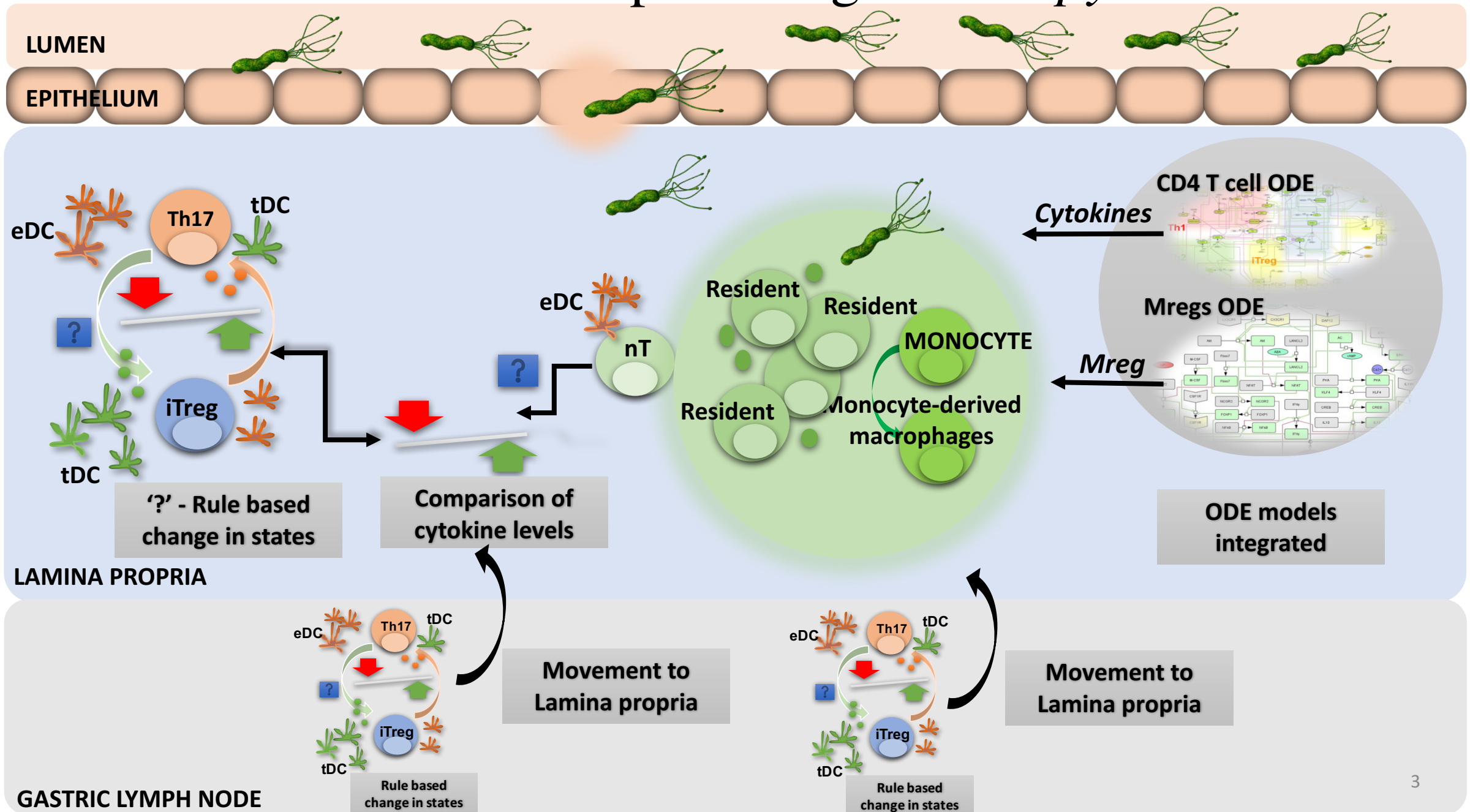
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1233 50. Windrum P, Fagiolo G, Moneta A. Empirical validation of agent-based models:  
1234 Alternatives and prospects. *Journal of Artificial Societies and Social Simulation*. 2007;10(2):8.  
1235 51. Fagiolo G, Moneta A, Windrum P. A critical guide to empirical validation of agent-based  
1236 models in economics: Methodologies, procedures, and open problems. *Computational*  
1237 *Economics*. 2007;30(3):195-226.  
1238 52. Crooks A, Castle C, Batty M. Key challenges in agent-based modelling for geo-spatial  
1239 simulation. *Computers, Environment and Urban Systems*. 2008;32(6):417-30.  
1240 53. Filatova T, Verburg PH, Parker DC, Stannard CA. Spatial agent-based models for socio-  
1241 ecological systems: Challenges and prospects. *Environmental modelling & software*. 2013;45:1-  
1242 7.

Name of Agent	Agent Type	Rules
<i>Helicobacter pylori</i>	<i>H. pylori</i>	<ul style="list-style-type: none"> <li>- Moves across the epithelial cell border if near damaged epithelial layer</li> <li>- Proliferates in the lumen and lamina propria</li> <li>- Dies (removed from the simulation) in lamina propria and in the lumen due to the damage of epithelial cells by Th1 or Th17 cells</li> </ul>
Macrophages	Monocyte	<ul style="list-style-type: none"> <li>- Proliferates in presence the of effector dendritic cells or damaged epithelial cells</li> <li>- Proliferates in the lamina propria</li> <li>- Differentiates to regulatory macrophage in based on the output from the Macrophage ODE</li> <li>- Differentiates to inflammatory macrophages in presence of IFN-<math>\gamma</math></li> <li>- Dies naturally (removed from the model)</li> </ul>
	Resident	<ul style="list-style-type: none"> <li>- Proliferates in the presence of <i>H. pylori</i></li> <li>- Secretes IL10</li> <li>- Dies naturally</li> <li>- Dies due to Th1 and Tr cells</li> </ul>
	Regulatory	<ul style="list-style-type: none"> <li>- Proliferates and removes bacteria</li> <li>- Dies</li> <li>- Secretes IL10</li> </ul>
	Inflammatory	<ul style="list-style-type: none"> <li>-Proliferates in the presence of damaged epithelial cell</li> <li>-Dies naturally</li> </ul>
Dendritics	Immature	<ul style="list-style-type: none"> <li>-Moves from lamina propria to epithelium compartment and from the epithelium to the lamina propria</li> <li>- Differentiates to tolerogenic dendritic cell in the presence of tolerogenic bacteria, both in epithelium and lamina propria</li> <li>- Differentiates to effector dendritic cell in the presence of <i>H. pylori</i></li> <li>- Proliferates in lamina propria and gastric lymph node</li> <li>- Dies naturally</li> </ul>
	Effector	<ul style="list-style-type: none"> <li>- Moves from lamina propria to gastric lymph node</li> <li>- Moves form epithelium to lamina propria</li> <li>- Secretes IL6 and IL12</li> <li>- Dies naturally</li> </ul>
	Tolerogenic	<ul style="list-style-type: none"> <li>- Moves from lamina propria to gastric lymph node</li> <li>- Moves from epithelium to lamina propria</li> <li>- Secretes TGF-<math>\beta</math></li> <li>- Dies naturally</li> </ul>
T cells	Naïve	<p>In the presence of effector dendritic cells:</p> <ul style="list-style-type: none"> <li>- Differentiates to Th1 in the presence of IFN-<math>\gamma</math> or IL12</li> <li>- Differentiates to Th17 in the presences of IL6 or TGF-<math>\beta</math></li> </ul> <p>In the presence of tolerogenic dendritic cells:</p> <ul style="list-style-type: none"> <li>- Differentiates to iTreg in the presence of TGF-<math>\beta</math></li> </ul>

		<ul style="list-style-type: none"> <li>- Differentiates to Tr in the presences of IL10</li> <li>- Dies naturally</li> </ul>
	Th1	<ul style="list-style-type: none"> <li>- Secretes IFN-<math>\gamma</math></li> <li>- Moves from gastric lymph node to lamina propria</li> <li>- Proliferates in lamina propria and gastric lymph node</li> <li>- Dies naturally</li> </ul>
	Th17	<ul style="list-style-type: none"> <li>- Secretes IL17</li> <li>- <u>i</u>n the presence of tolerogenic dendritic cell, transition to iTreg cells</li> <li>- Moves from gastric lymph node to lamina propria</li> <li>- Proliferates in lamina propria and gastric lymph node</li> <li>- Dies naturally</li> </ul>
	iTreg	<ul style="list-style-type: none"> <li>- Secretes IL10</li> <li>- In the presence of tolerogenic dendritic cell, transition to iTreg cells</li> <li>- Moves from gastric lymph node to lamina propria</li> <li>- Proliferates in lamina propria and gastric lymph node</li> <li>- Dies naturally</li> </ul>
	Tr	<ul style="list-style-type: none"> <li>- Secretes IL10</li> <li>- Dies naturally</li> <li>- Proliferates in the lamina propria</li> </ul>
Epithelial	Healthy	<ul style="list-style-type: none"> <li>- Damaged due to infectious bacteria</li> <li>- Damaged due to Th1 and Th17 cells</li> <li>- Proliferates</li> <li>- Secretes IL6 and IL12</li> <li>- Dies naturally</li> </ul>
	Damaged	<ul style="list-style-type: none"> <li>- <u>Transitions</u> to healthy state in the presence of IL10</li> <li>- Dies naturally</li> </ul>
Bacteria	Infectious	<ul style="list-style-type: none"> <li>- Dies due to Th1 or Th17 or inflammatory macrophages or damaged epithelial cells</li> <li>- Dies naturally</li> <li>- Proliferates in the lamina propria</li> </ul>
	Tolerogenic	<ul style="list-style-type: none"> <li>- Moves from lumen to the epithelium in the presence of damaged epithelial cells</li> <li>- Becomes infectious if moves in the lamina propria compartment</li> <li>- Proliferates in lumen and lamina propria</li> <li>- Dies naturally</li> </ul>

# Immune responses against *H. pylori*



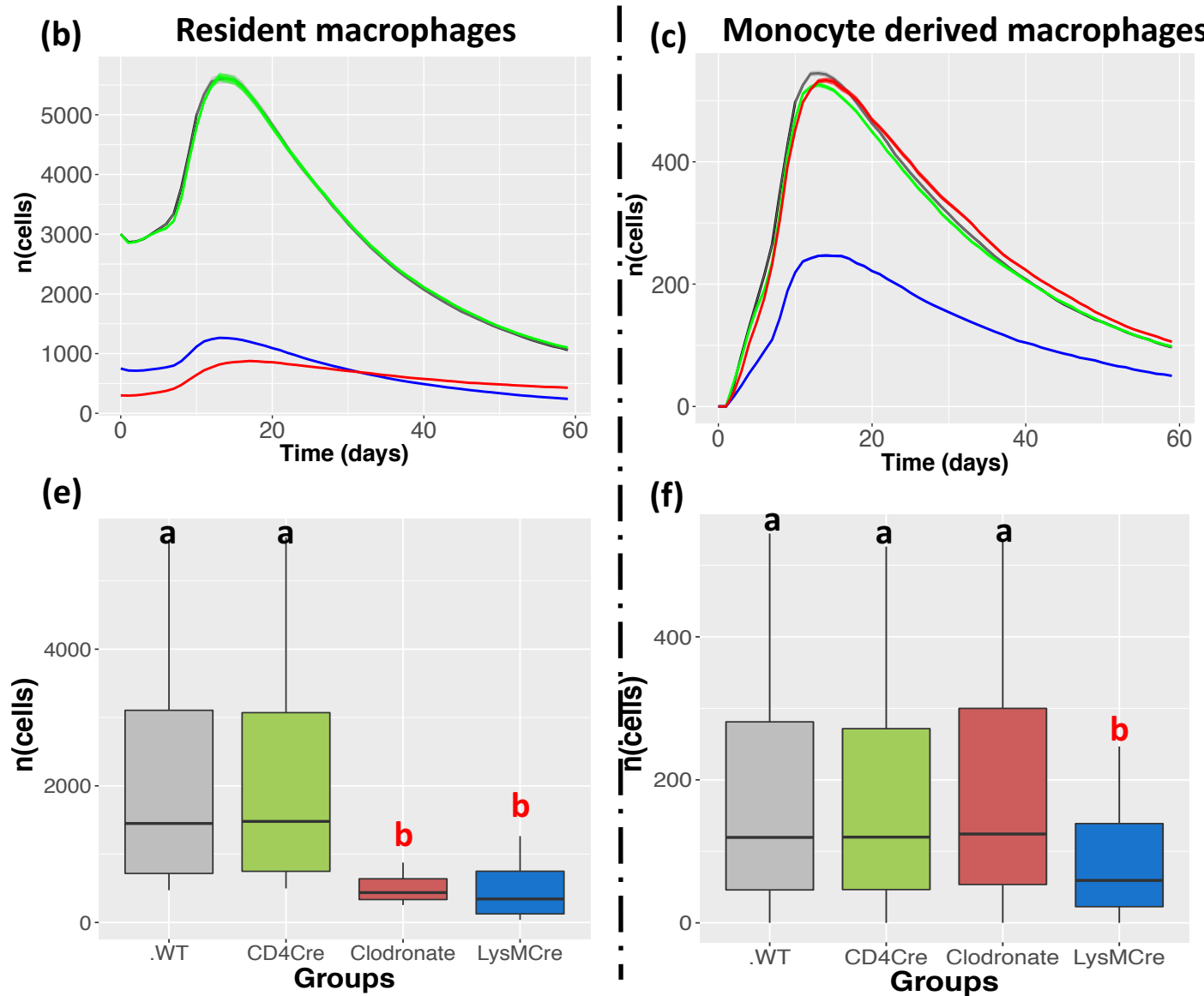
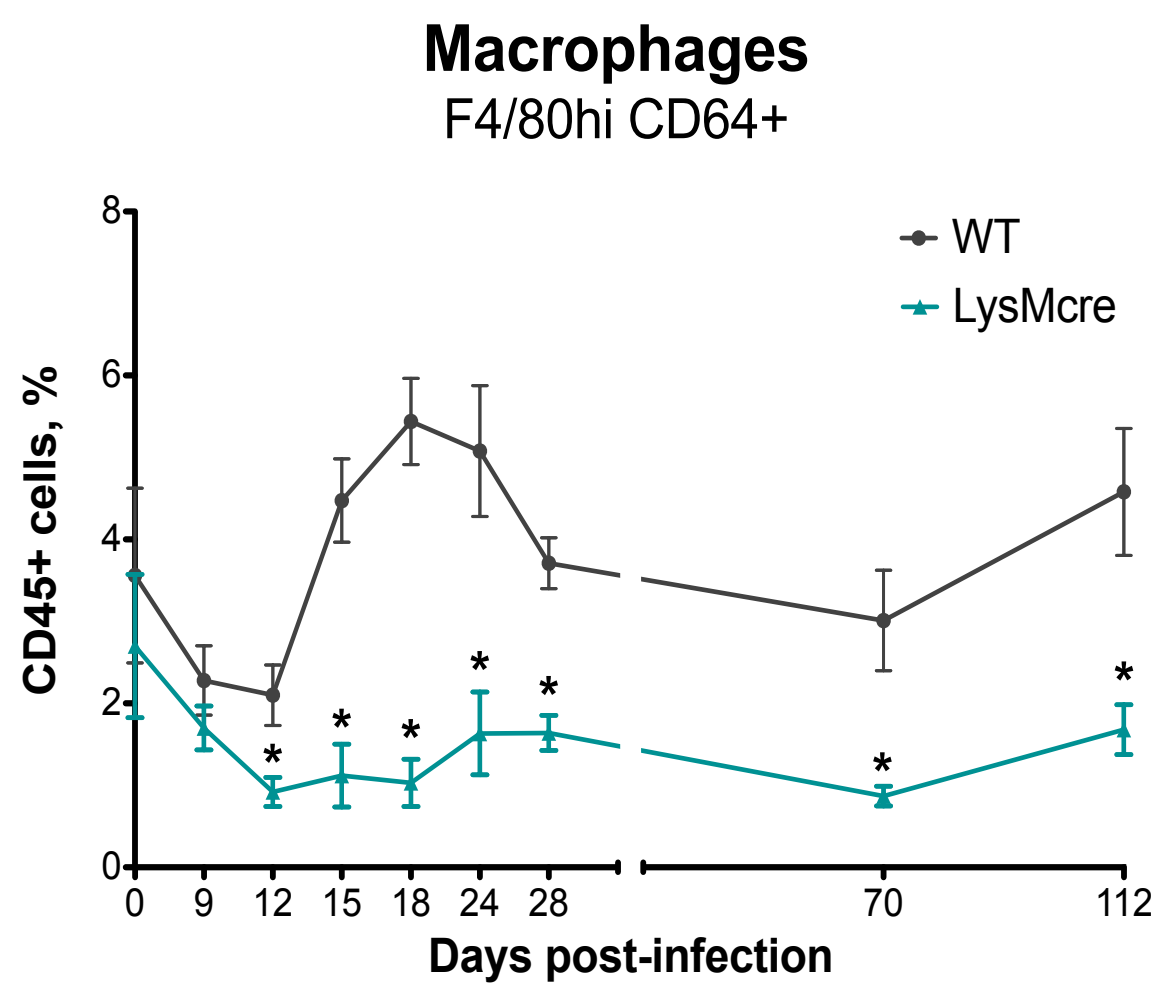
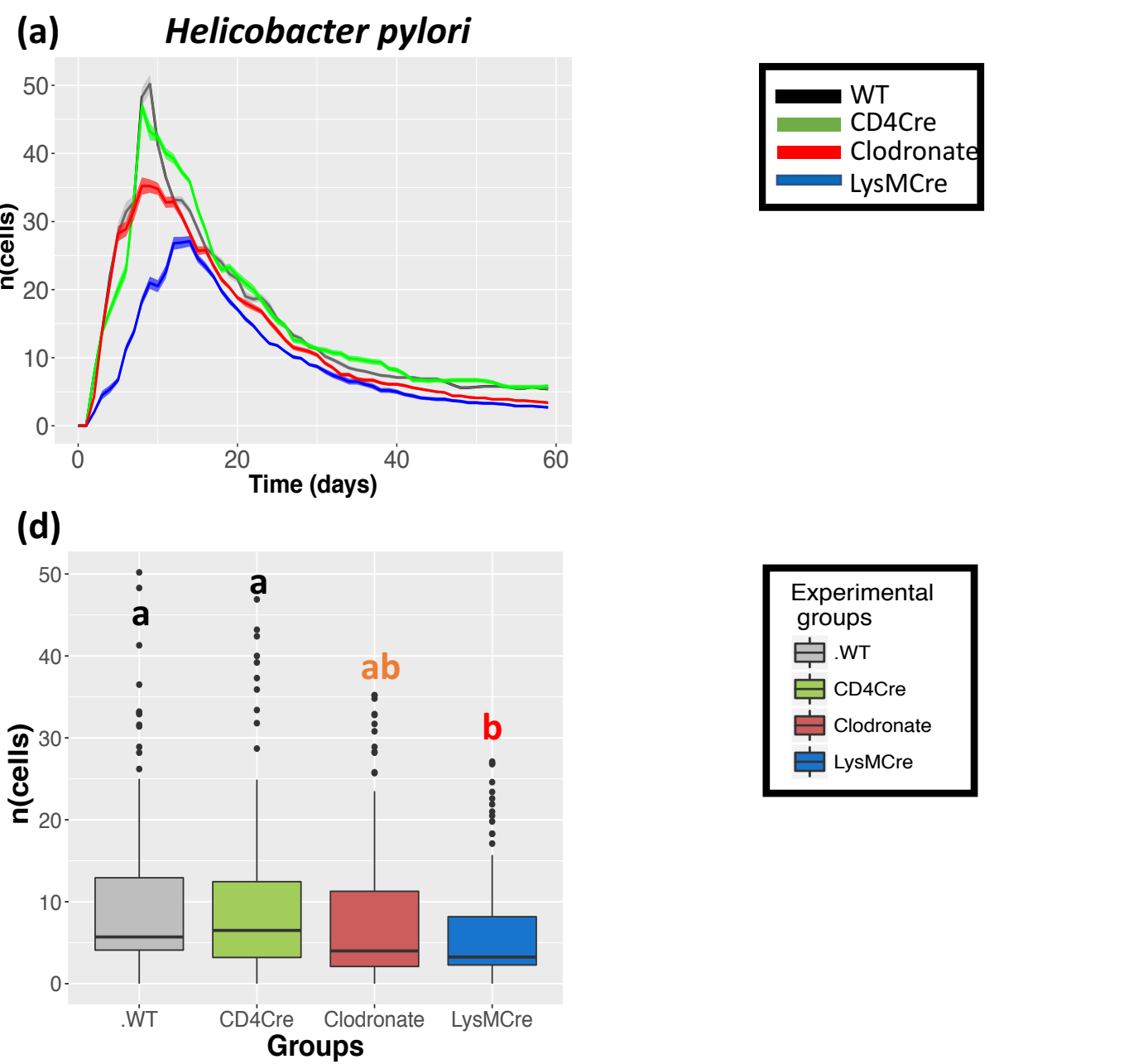
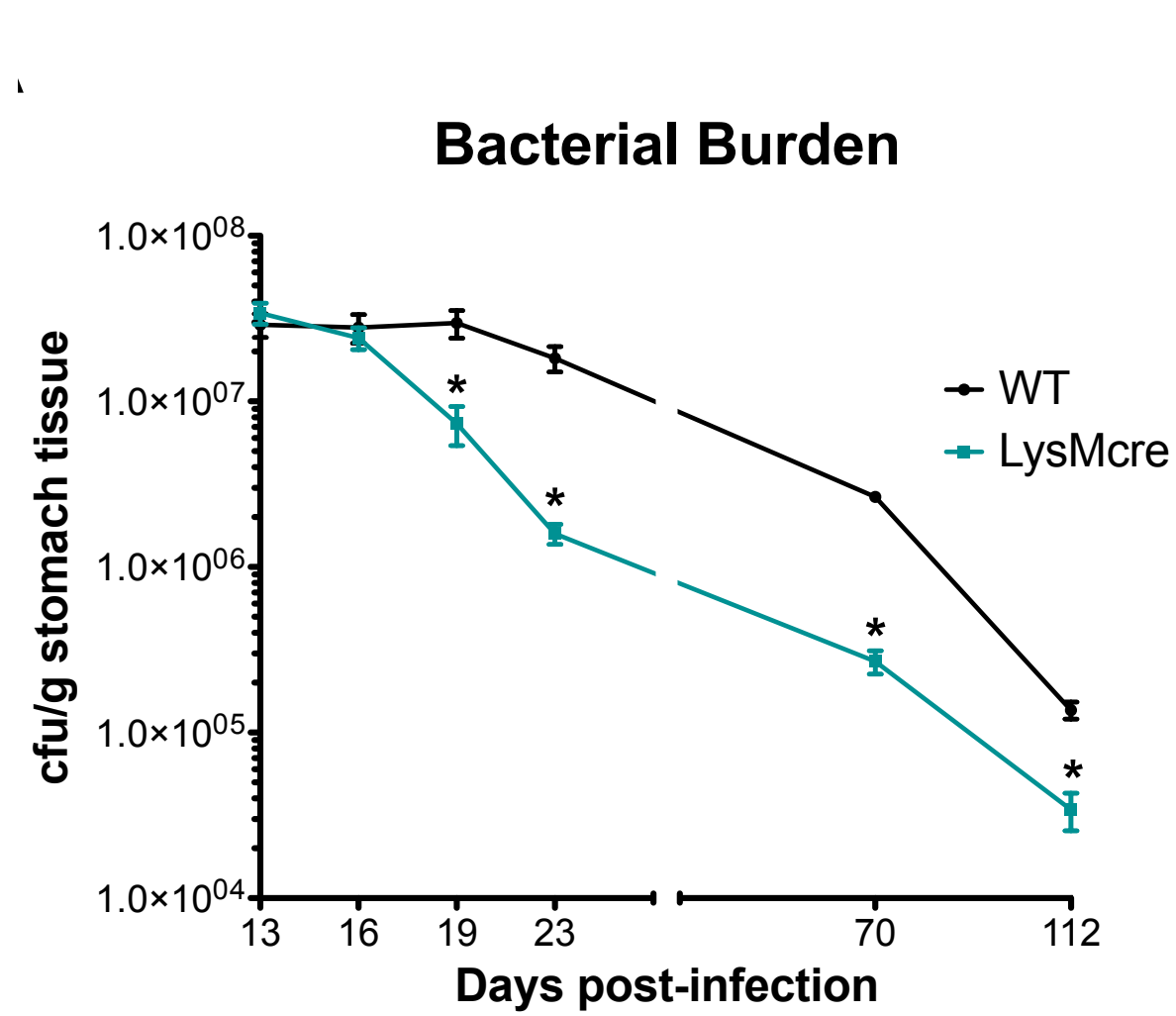
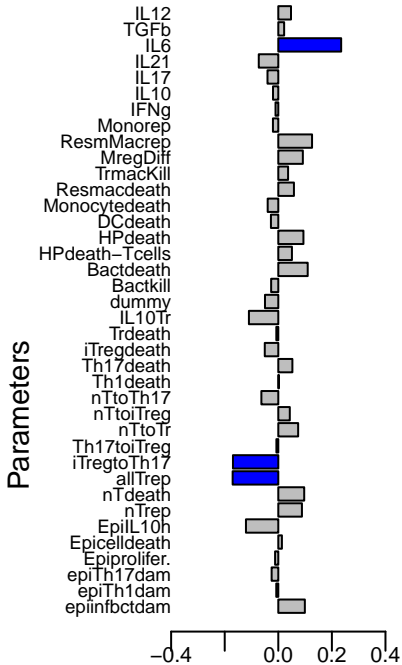


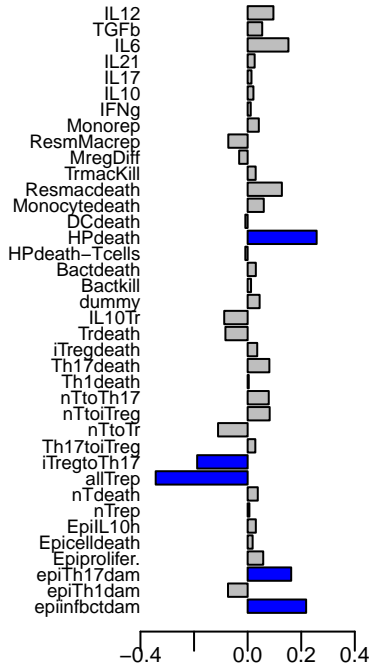
Figure 3

**Helicobacter pylori in Lamina propria**

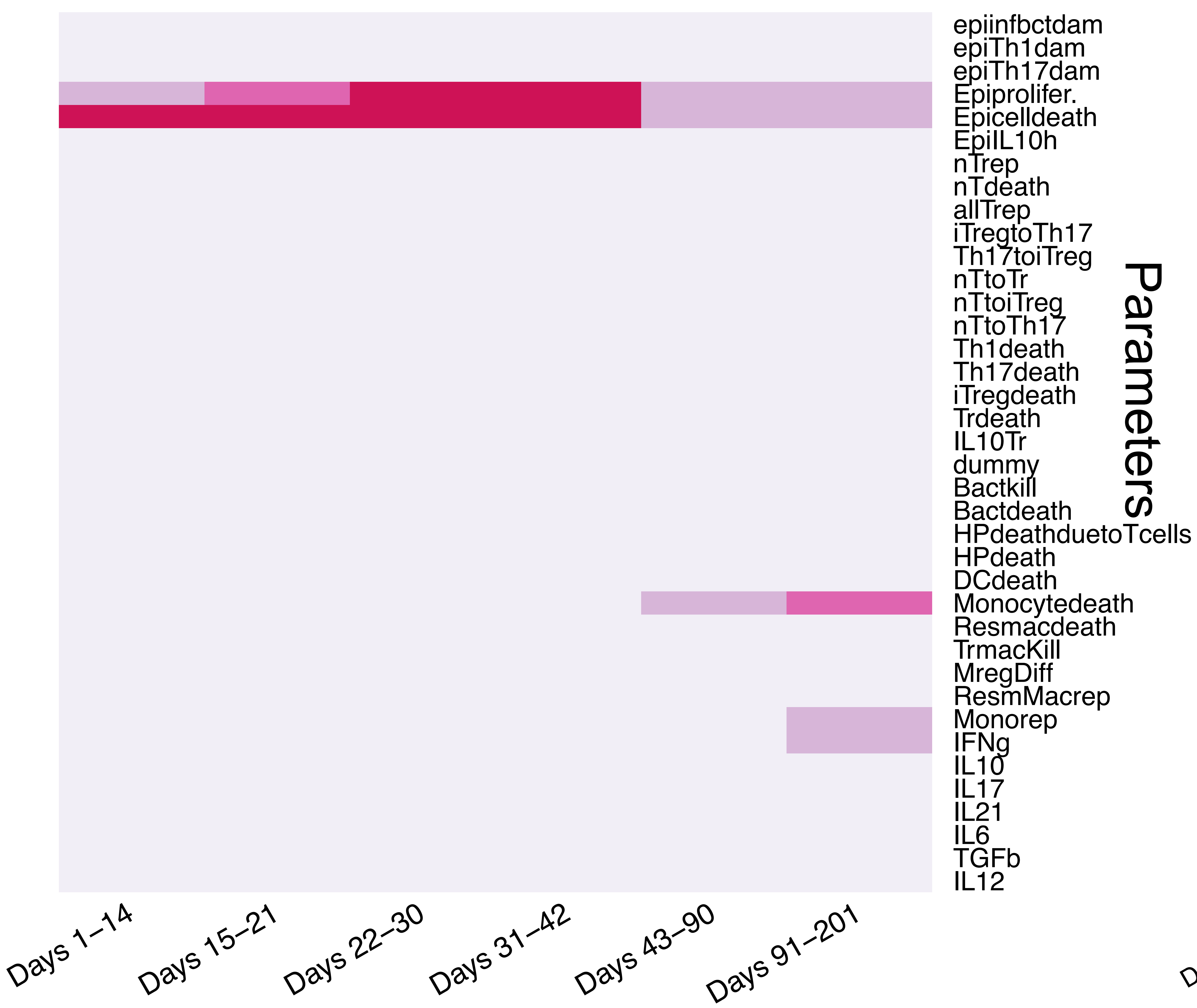


[Click here to access/download; Figures; Figure9.pdf](#)

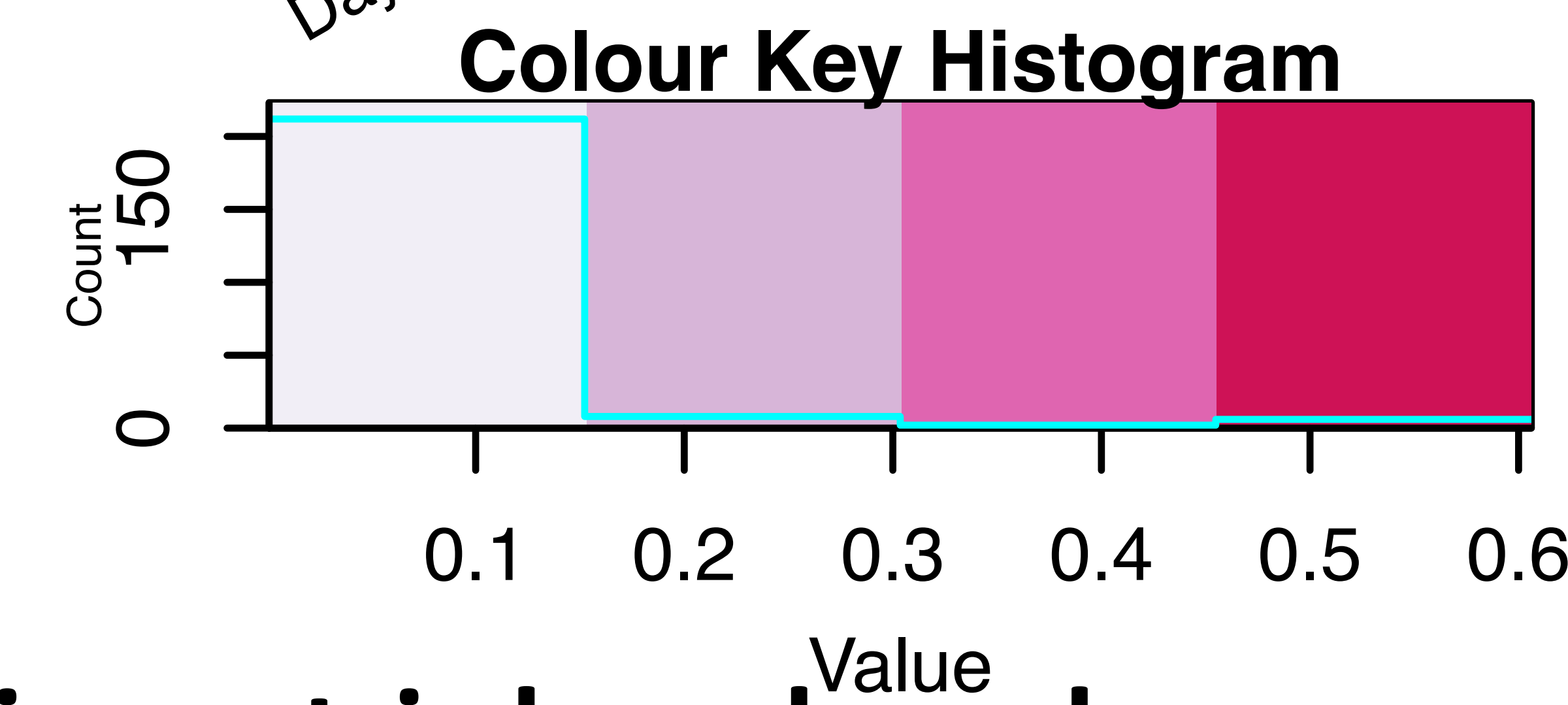
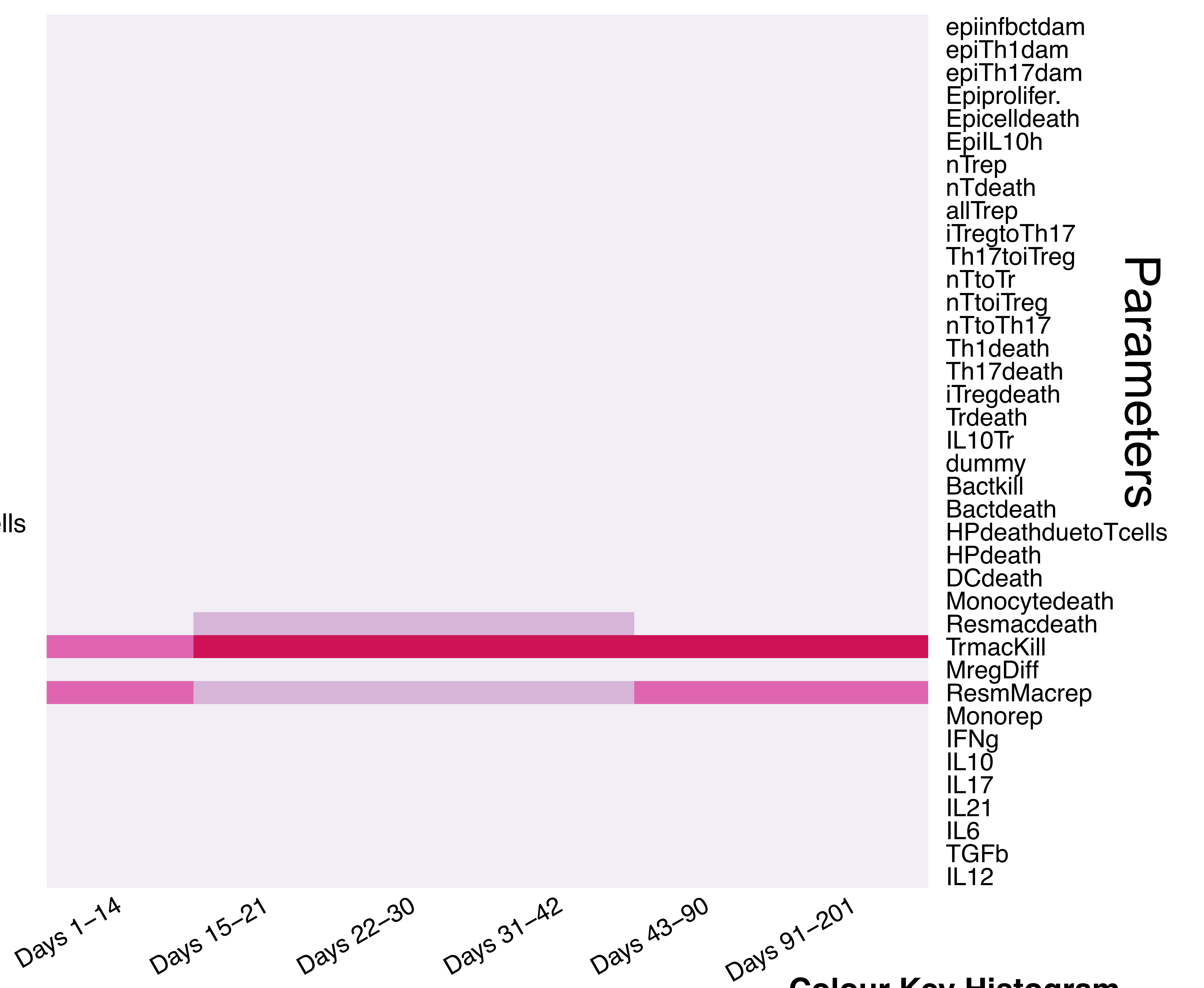
**Resident macrophages in Lamina propria**



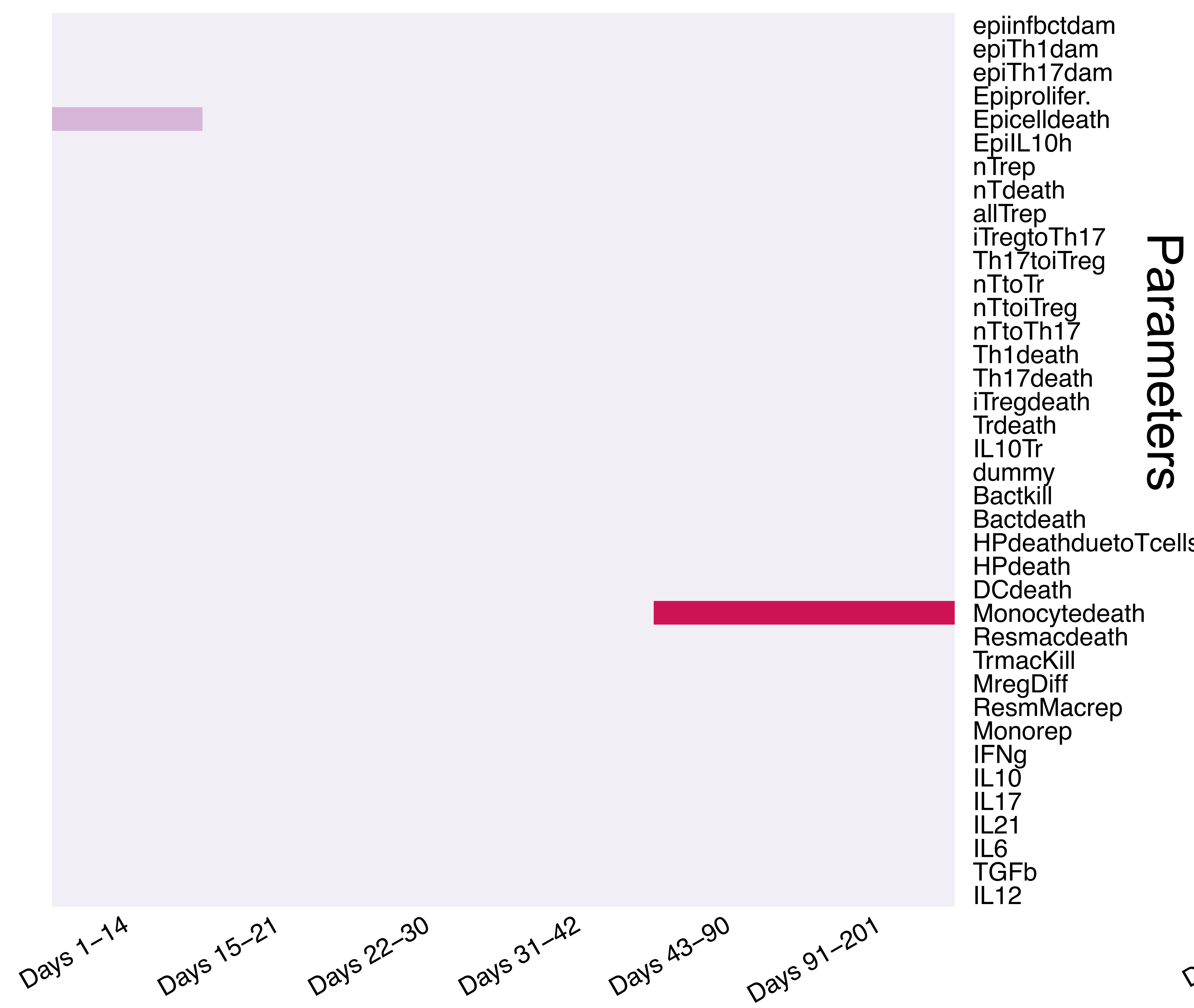
**(a) *Helicobacter pylori* in lamina propria**



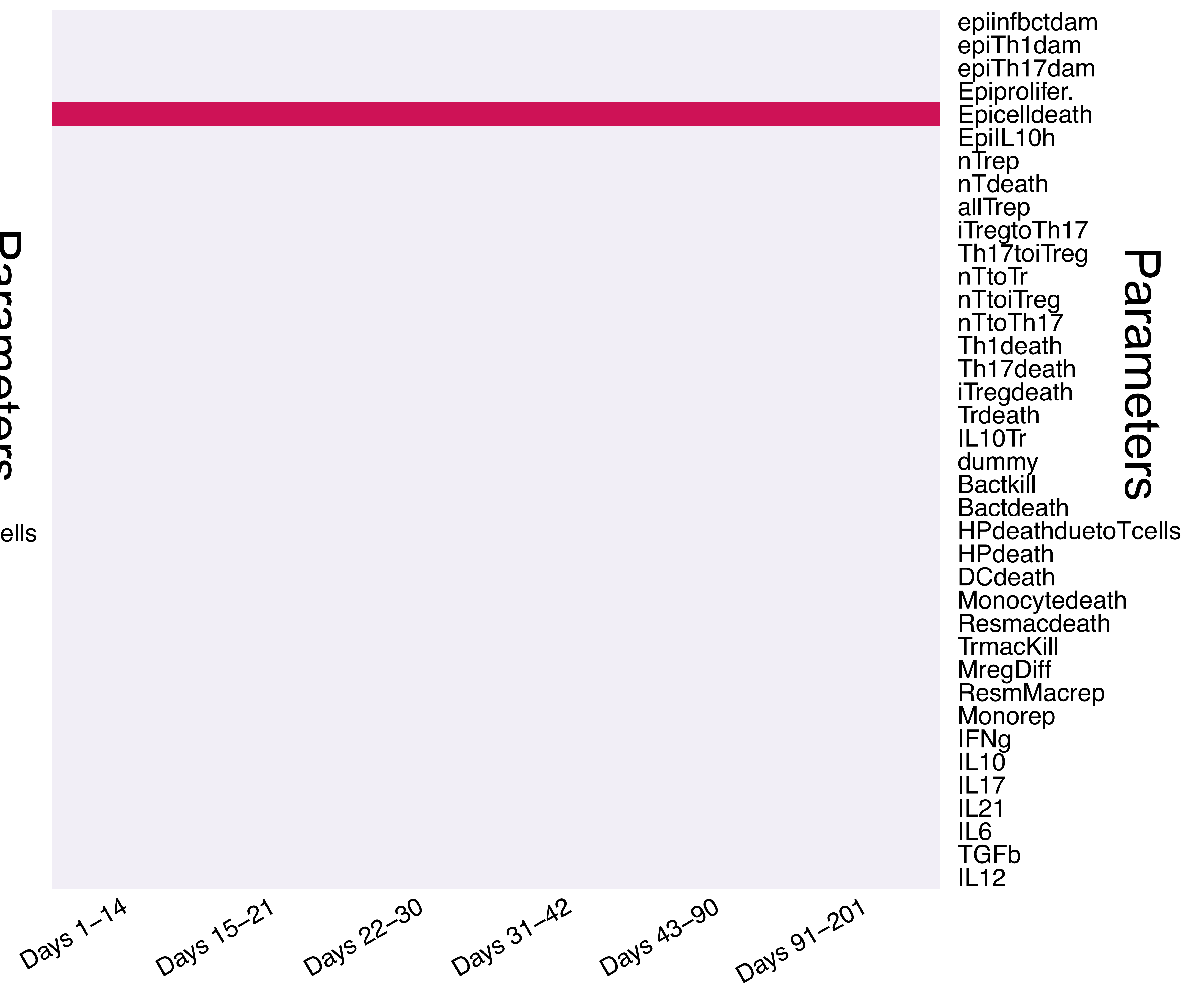
**(b) Resident macrophages in lamina propria**

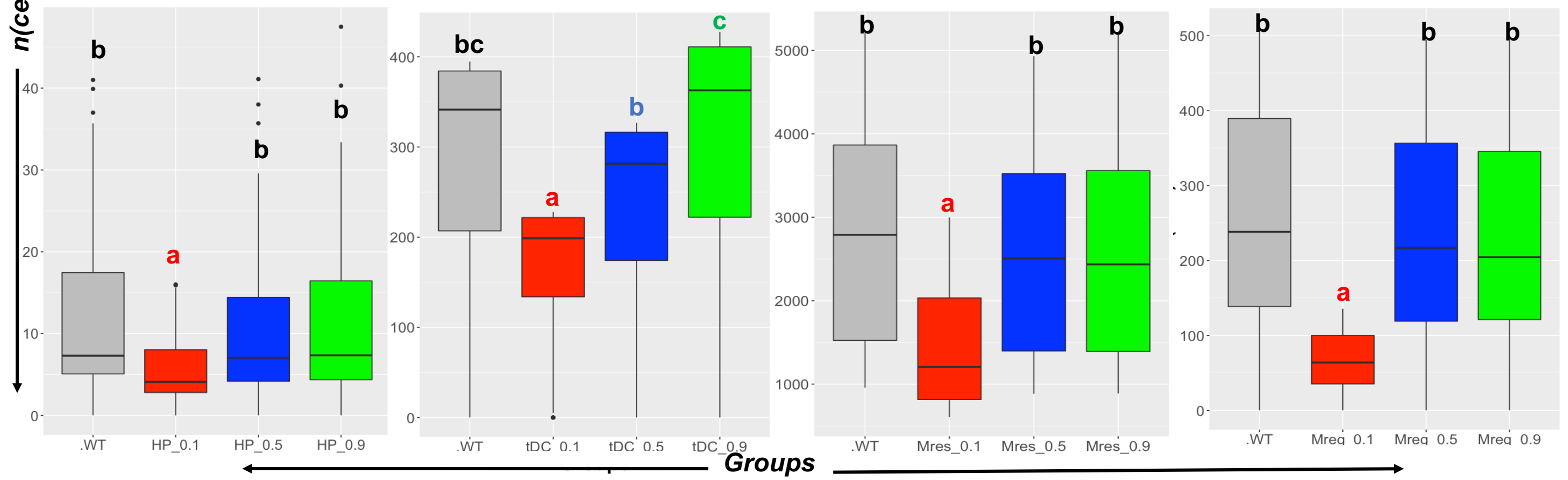
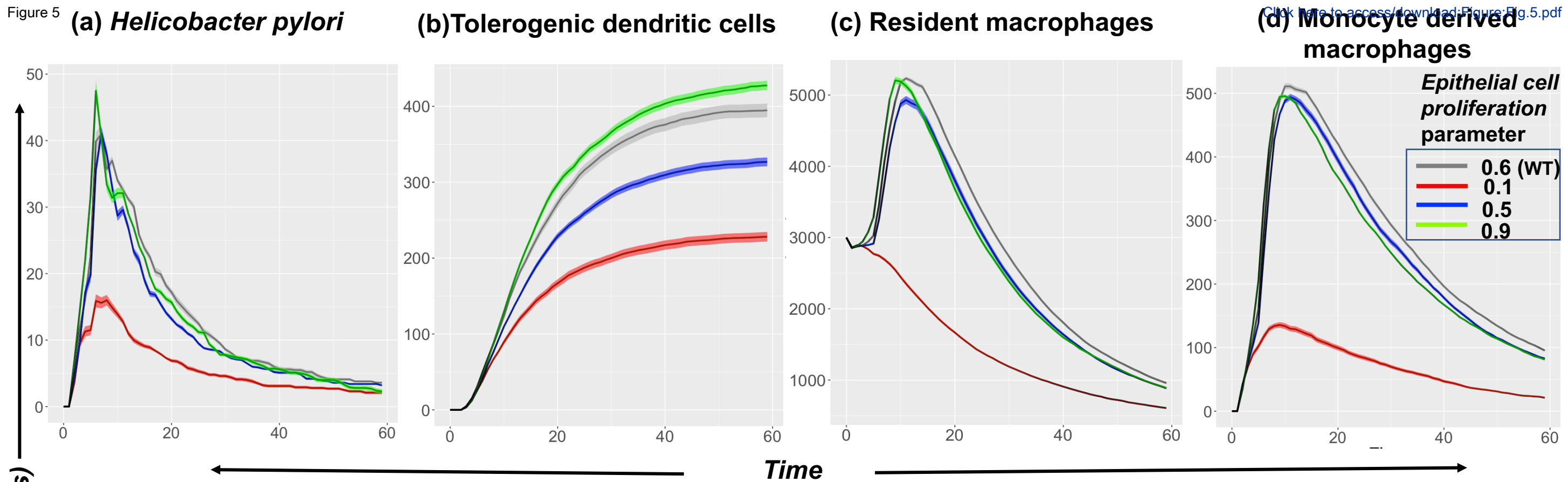


**(c) Monocyte derived macrophages in lamina propria**




**(d) Tolerogenic dendritic cells in gastric lymph node**












Click here to access/download  
**Supplementary Material**  
FileS1.pdf






Click here to access/download  
**Supplementary Material**  
FigS1.png

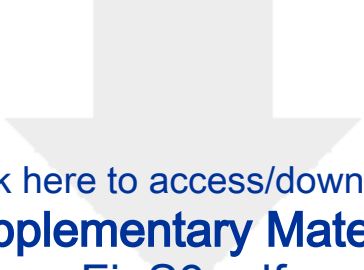


Click here to access/download  
**Supplementary Material**  
Table\_S1\_revised.docx

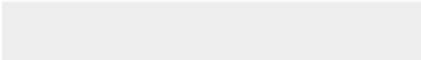





Click here to access/download  
**Supplementary Material**  
FigS2\_revised.pdf



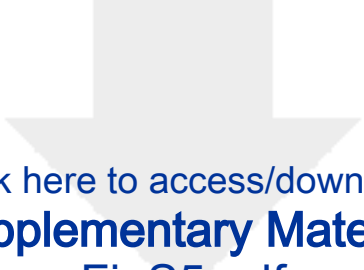
Click here to access/download  
**Supplementary Material**  
FigS3.pdf



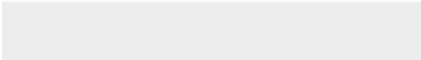



Click here to access/download  
**Supplementary Material**  
FigS4.docx





Click here to access/download  
**Supplementary Material**  
FigS5.pdf



### Point by point response to the Reviewer reports

We would like to thank the reviewers and editors for taking time to review our manuscript entitled “*High-Resolution Computational Modeling of Immune Responses in the Gut*” and for providing valuable and constructive criticism. The review process has been helpful in the improvement of our submission. We have considered the comments that were made and have prepared the following point-by-point response. We hope that the revised version of the manuscript can now be accepted for publication. Thanks in advance.

**Reviewer #1:** The unit of sizes of the model grid can't be right (e.g. grid is 30nm x 10 nm). Animal cells should have measurements in the order of micrometres instead of nanometres. Please check if these are just typos, or do these errors affect any aspect of the simulation, such as diffusion.

**Response:** *We thank the reviewer for pointing this out. We fixed the typos and the unit size of the model grid are 30  $\mu\text{m}$  x 10  $\mu\text{m}$ . These typos do not affect any aspect of the simulations as these units are only annotations and the model takes the numbers as input. We updated the manuscript and fixed the typos throughout the manuscript. Please refer to L120 – L121, and L216, L220-L221.*

**Reviewer #2:** The authors have made significant improvements to the manuscript and thoroughly responded to reviewer comments. One major concern remains surrounding the authors' response to questions around the grid dimensions. The dimensions for the entire grid are given in nm which is smaller than a single cell. Furthermore they state that there are no limits to cell(agent) occupancy per grid compartment. This is rather confusing and calls into question how much spatial information is really contained in this model (e.g. if cytokines are diffusing over the 30nm grid what does that mean for the concentrations that individual cells (measured in micrometers)are seeing?). Based on the author responses it appears that the model is a multi-compartment model with well-mixed discrete agents in each compartment rather than a spatio-temporal model as they claim.

**Response:** *We thank the reviewer for their comment.*

*We thank the reviewer for pointing out the concern regarding the dimensions of the grid. The correct dimensions of the grid are 30  $\mu\text{m}$  x 10  $\mu\text{m}$ . We updated the manuscript and fixed the typos. Please refer to L120 – L121, and L216, L220-L221.*

*The mention regarding no limits to cell (agent) occupancy refers to the cells (agents) having no physical size. Further, once a cell (agent) dies it is removed from the simulation to minimize the computational costs of agents that do not contribute to the biology.*

*The model output contains information about the x and y co-ordinate of the agents at every time point. The cytokines and internal signaling pathways that drive functional fates of cells are well mixed within a cell, i.e., we have only temporal resolution within the cell during a time step. However, the production, degradation, and diffusions are cell specific thus the cytokine concentration results are also spatio-temporal.*

*Since, the model is capable of providing information regarding spatial co-ordinates over time, we claim the model to be a spatio-temporal model. We updated the manuscript, please refer to L163-L170.*

Please also ensure that your revised manuscript conforms to the journal style, which can be found in the Instructions for Authors on the journal homepage.

**Response:** *The revised manuscript conforms to the journal style.*