Neural Network Models for Classifying Immune Cell Subsets

Yongguo Mei, Raquel Hontecillas, Xiaoying Zhang, Adria Carbo, Josep Bassaganya-Riera
The Center for Modeling Immunity to Enteric Pathogens
Nutritional Immunology and Molecular Medicine Laboratory
Virginia Bioinformatics Institute, Virginia Tech, Blacksburg, VA 24061, USA
{ymei, rmagarzo, xzhang, acarbo, jbassaga}@vbi.vt.edu

***Abstract*— The immune system is composed of heterogeneous cell populations and it includes several hundreds of distinct cells types such as neutrohils, eosinophils, basophils, macrophages, dendritic cells, CD4+ and CD8+ T cells, $\gamma\delta$ T cells, mast cells, and B cells and each main cell type can be further differentiated into subsets with unique and overlapping functions. For example CD4+ T cells can be differentiated into T helper (Th)1, Th2, Th17, and regulatory T cell (Treg) subsets. To study molecular mechanisms of cell differentiation, Systems Biology Markup Language (SBML) based Ordinary Differential Equation (ODE) models can be used for representing such processes. These intracellular signaling models often require many equations to accurately represent intracellular pathways and biochemical reactions. Another challenge in studying the immune system and immune responses is the need for integration of complex processes that occur at different time and space scales (i. e., populations, whole organism, tissue level, cellular and molecular) through multi-scale modeling.**

This study presents two novel neural network models for modeling CD4+ T cell differentiation and immune cell subset classification. The first model reduces the complex ODE intracellular model by focusing on the input and output cytokines and the second model establishes an automated subset classification based on molecular patterns expressed in immune cells. After training, the first model achieves small prediction errors of cytokine concentrations and the second model achieves 98% prediction rate for subset classification. Neural network algorithm and models have been widely used for many data mining tasks such as classification and pattern recognition. However, to the best of our knowledge this study is the first one applying the neural network algorithm for immune cell differentiation and subset classification. In addition, these novel neural network models can be computationally efficiently integrated into multi-scale models with limited computational costs.

I. Introduction

A. Immune cell differentiation and modeling

The process of immune cell differentiation plays a central role in orchestrating immune responses. This process is based on the differentiation of nave immune cells that will activate its transcriptional machinery through a variety of signaling cascades to become phenotypically and functionally different entities capable of responding to viruses, bacteria, parasites, cancer cells, etc. Functionally, immune cells have been classified in either regulatory or effector subsets. For instance naive CD4+ T cells differentiate into effector phenotypes such as Th1, Th17 or Th2 when the cytokine milieu is rich in IFNγ or IL-12 (for Th1), IL-6, IL-1β and TGF-β (for Th17) or IL-4 (for Th2). The external cytokine tissue environment, therefore, is a key and decisive factor for delineating the cell differentiation outcomes, although selective factors are also important in this process.

The cell differentiation process involves a series of sequential and complex biochemical reactions within the intracellular compartment of each cell. The SBML is XML-based and human readable, and is widely used to represent models of biological processes. SBML allows the biochemical reactions to be represented by master equations characterized to be first- order ODE. Of note, ODE models are extensively used to model biological processes such as cell differentiation, immune responses towards specific pathogens or intracellular activation of specific pathways. Often, these models require several equations to adequately represent the process they intend to model, being either at the tissue level or at the intracellular level. In one of our previous studies, Carbo et. al. [4] published the first comprehensive ODE model of CD4+ T cell differentiation that encompassed the Th1, Th2, Th17 and Treg phenotypes. The CD4+ T cell differentiation model is composed of 60 ODEs and built upon the current paradigms of molecular interactions that occur in CD4+ T cells. The model is intended to help researchers to elucidate the regulatory mechanisms underlying CD4+ T cell differentiation, identify novel putative CD4+ T cell subsets, and study the dynamics of Th cell differentiation.

B. Multi-scale modeling and model reduction

In one of our previous studies, Mei et. al. [26] presented Enteric Immunity Simulator (ENISI) Visual, an agent-based simulator for modeling mucosal immune responses to enteric pathogens. ENISI Visual can simulate cells, cytokines, cell movement and cell-cell interactions. ENISI Visual is rule-based and either inter-cellular or tissue-level. To be able to model finer-grained intracellular behaviors, a multi-scale modeling approach, which embeds intra-cellular models into the inter-cellular tissue level models is needed. While the fine-grained ODE models of intracellular pathways controlling immune cell differentiation are adequate for studying mechanisms of cell differentiation, it can also be highly complex and expensive from a computational stand-point, especially when embedded within large-scale agent-based simulations. ENISI Visual models a large number of cells and microbes in the gastroeintestinal mucosa. If each agent is represented by 60 ODEs, as an example, the simulation will hardly scale up. To be able to develop efficient agent-based multi-scale models, model reduction is a necessary step in order to obtain a workable model able to compute appropriately. In addition, multi-scale models usually do not require all the internal details of intra-cellular models to have predictive value and this provides a great avenue to apply novel automated model reduction strategies to reduce molecular models before integrating them into large-scale agent-based tissue-level model, thus fulfilling the multi-scalability requirements.

C. Neural network algorithms and its applications

Artificial neural networks (ANN) algorithms, inspired by the biological neural systems, are powerful modeling and data mining tools based upon the theory of connectionism. Neurons are connected to each other through synapses. A neuron receives inputs from multiple neurons and outputs one value based upon the activation function. The network structures and the parameters of the activation function are important factors when developing neural network models. Feedforward neural networks are frequently used structures in modeling. There are effective learning algorithms for the parameters once the structures are set in the feedforward ANNs. Neural network algorithms are widely used for data mining tasks such as classification and pattern recognition. Neural network algorithms are especially effective in modeling non-linear relationships which makes them ideal candidates for differentiation processes and automated classification of phenotypes. Importantly, this process is scalable.

To the best of our knowledge, this study is the first applying neural network algorithms into studying the immune cell differentiation, cell heterogeneity and subset classification. More specifically, the neural network algorithms are used in the following two models:

- Model reduction for T cell differentiation. We use neural network algorithms to reduce the intracellular CD4+ T cell differentiation ODE model into a neural network model with 4 inputs, 5 outputs, and one middle layer of 6 nodes. The 4 input nodes represent the 4 external cytokines that regulate the cell differentiation; the 5 output nodes represent the 5 cytokines that are externalized and secreted by the T cell subsets.
- T cell subset classification. We developed a neural network model for subset classification with 4 input nodes representing the 4 external cytokines as in the first model and 4 output nodes representing four possible subset classification outcomes. This model also has 6 intermediate nodes.

After training, the first model achieves high accuracy in predicting the concentrations of output cytokines. For example, the output is in the range of [0, 1] and the largest average prediction error is 0.0253 for IL17. The second model achieves 98% classification accuracy based upon the rules set by expert immunologists.

The remaining of this paper is organized as follows. We first review the literature of related previous work. The problem of modeling T cell differentiation and subset classification is presented next. Neural network algorithm and models for solving important immunological problems are presented thereafter. At the end, the paper provides some concluding remarks, including discussions of future proposed work.

II. Related Work

The immune system protects the human body from pathogens by recognizing, containing, and destroying non-self or foreign substance [2]. At the highest level, immune systems can be divided into innate and adaptive. The innate immune systems [1], involving cells such as macrophages, epithelial cells, neutrophils, respond quickly but non-specifically to stimuli. On the contrary, adaptive immune systems [16] involving T cells and B cells respond more specifically to antigens. Immune cells are activated and differentiated into ever-growing numbers of cell subsets [10] [27] [28] regulated by different cytokines in their micro-environment. Using CD4+ T cells as an example, when both transforming growth factor-b (TGFβ) and interleukin-6 (IL-6) are present in the environment, native CD4+ T cells differentiate

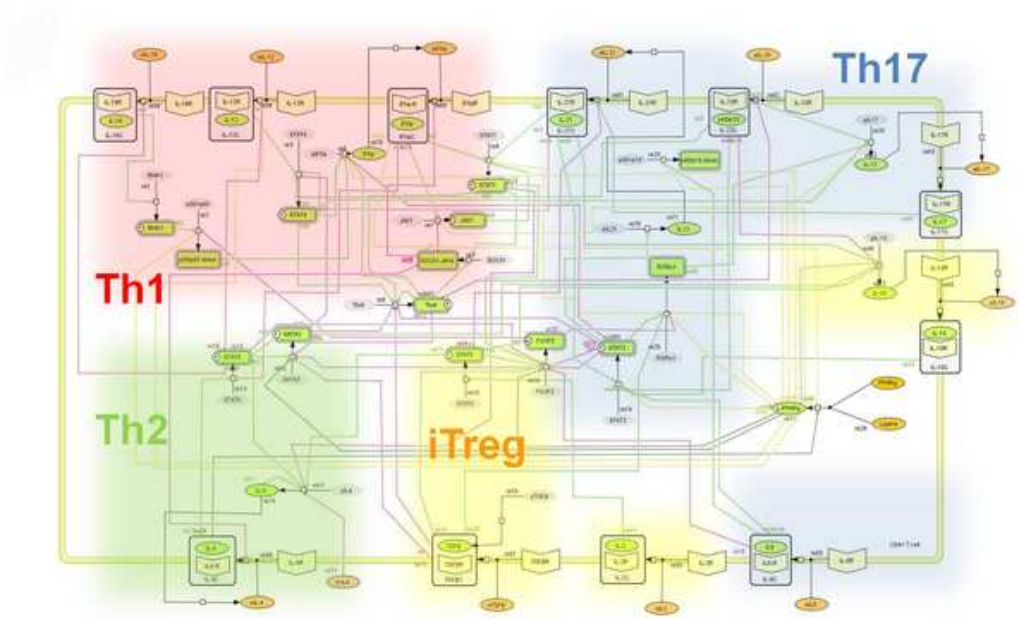

Fig. 1. The networked ODE model for T Cell differentiations [4]. The model has four regions of pathways that regulates the cell differentiation into Th1, Th2, Treg, and Th17.

into Th17 [25], [20]. When TGFβ alone presents in the environment, CD4+ T Cells differentiate into Treg [14]. When both IFNγ and IL12 are present, T cell differentiate into Th1 [19].

Systems biology [17] has become an important paradigm of life science research, using mathematical and computational models to synthesize and mine exiting knowledge, and discover new knowledge from big data. Biological systems and processes can be modeled using networks [18] where nodes and edges represent biological agents such as cells and their interactions. Furthermore, mathematical or computational dynamics can be applied to the networked models so that *in silico* simulations [8] can be performed. SBML [15] is a XML-based file format used to represent computational models of biological processes. There are many types of models [24] used for modeling biological processes such as Bayesian networks, ordinary differential equations (ODE), agent-based models. For metabolic and signaling networks, the reactions are biochemical reactions. As biochemical reactions can be represented first-order ODEs [7], metabolic and signaling networks often are modeled by ODEs [9].

In one of our previous studies, Carbo et. al. [4] published the first ODE model of CD4+ T Cell differentiation, which comprises of 60 ODEs. The model as shown in Figure 1 represents the intracellular pathways that are critical for T cell differentiation. This model has been well calibrated using experimental data. In another our previous study, Mei et. al. [26] presented ENISI Visual, an agent-based simulator for modeling enteric immunity. One example modeling scenario is shown in Figure 2.

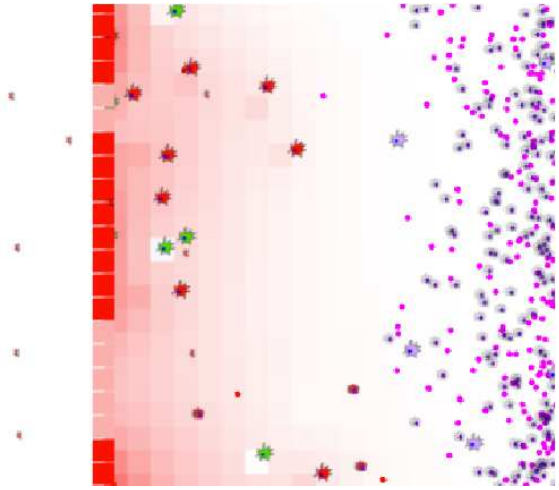

Fig. 2. ENISI Visual [26], an agent-based model for enteric immunity. The immune cells are represented by icons with different colors representing different subset state. The background color represent the concentrations of cytokines.

Neural network algorithms have been widely used in data mining tasks [5][23]. Neural network algorithms

have been also used in medical applications [6]. Snow et al. [29] developed neural networks for prostate cancer diagnosis and prognosis. Lek et al. [22] introduced using neural networks in ecological modeling. Brusic et al. [3] used neural networks for predicting HC binding peptides. Learning is an important research topic in neural networks. White [30] presented neural network learning algorithms from the statistical perspective. Hagan et al. [12] presented an effective learning algorithm called back-propagation for training feedforward networks. In addition to modeling and predictions, Neural network has also been used for solving ordinary and partial differential equations [21].

III. The problem: modeling immune cell differentiation and subset classification

To define the problem, we make the following assumptions. There are m input cytokines that regulate immune cell differentiation: C_{i1}, C_{i2}, ... , C_{im}. The n output cytokines secreted by immune cells are: C_{o1}, C_{o2}, ... , C_{on}. The immune cell's subset type is S with l possible subsets: S_1, S_2, ..., S_l. The cytokine concentrations are positive continuous values. The cell subsets are categorized.

The problem of modeling immune cell differentiation and subset classification is to develop two models for the following two functional relationships:

$$\{C_{o1}, C_{o2}, ..., C_{on}\} = F_c(C_{i1}, C_{i2}, ..., C_{im}) \quad (1)$$

$$\{S|S_1, S_2, ..., S_l\} = F_s(C_{i1}, C_{i2}, ..., C_{im}) \quad (2)$$

The first model is to predict the output cytokine concentrations giving concentrations of input cytokines. The second model is to predict the cell subset type also giving concentrations of input cytokines.

A. T cell differentiation and subsets

This study will focus on the T cell differentiation. However, the techniques and algorithms developed herein can be applied to differentiations of other types of immune cells, such as macrophages, dendritic cells, B cells, etc. The input cytokines are internalized by the naive T cells and regulate the T cell differentiation. The output cytokines are cytokines externalized and secreted by the T cells. For the T Cell differentiation, the four input cytokines are IFNγ, IL12, IL6, and TGFβ. The five output cytokines are IL17, RORγt, IFNγ, Tbet, and FOXP3. We focus on three T cell subsets, Th1, Th17, and Treg.

B. Data for training and testing models

The data for modeling the relationship from the input and output cytokines can be derived from the ODE T Cell differentiation model [4] since the model was calibrated using data from biological experiments. We change the concentrations of the input cytokines and then we calculate the steady state of the ODE model and extract the values of the output cytokines. All the data is normalized to the range of [0, 1]. We choose two ways to set values for input cytokines. The first method is equal-distance sampling. For each input cytokine, we choose five values 0, 0.25, 0.5, 0.75, and 1. The second method is randomizing. A random value of each cytokine is independently generated following an even distribution between [0, 1].

Based upon the output values of cytokines, we asked immunologists to give us the rules to classify the subsets. The rules include the following conditional rules applied in order. Rule1: if ($IL17 > 0.6$) and ($ROR\gamma t > 0.7$), the subset is Th17; Rule 2: if ($FOXP3 > 0.9$), the subset is Treg; Rule 3: if ($Tbet > 0.6$) and ($IFN\gamma_o > 0.9$), the subset is Th1. If none of the above rules applies, the subset is unclassified.

We use COPASI [13] to do the ODE modeling and it provides parameter scans both randomly and predefinedly. We write a script to automatically extract the wanted values from the COPASI result file. Table I shows some example data that can be used for training and testing the model. Some of the data will be used in training and some of the data will be used for testing/predicting purpose. Furthermore, these files are tab-separated files and data in table format can be readily read into R for further data processing.

IV. Neural network models

A. Linear regression model

Before we develop neural network models, we try linear regression model first. For the linear regression model, the function in equation (1) is linear. R has a linear regression module that can be readily used for this study. The result of the linear regression is essentially a linear transformation from the input cytokines to the output cytokines.

$$\begin{bmatrix} FOXP3 \\ IFN\gamma_i \\ IL17 \\ ROR\gamma t \\ Tbet \end{bmatrix} = M_{Tran} \times \begin{bmatrix} 1 \\ IFN\gamma_o \\ IL12 \\ IL6 \\ TGF\beta \end{bmatrix} \quad (3)$$

TABLE I

EXAMPLE DATA USED FOR TRAINING AND TESTING OF THE MODELS. THE LEFT FOUR COLUMNS ARE VALUES OF INPUT CYTOKINES; THE MIDDLE FIVE CYTOKINES ARE VALUES OF OUTPUT CYTOKINES; AND THE RIGHT COLUMN IS THE SUBSET CATEGORY. THE FIRST 6 ROWS OF DATA ARE GENERATED FROM EQUAL-DISTANCE SAMPLING METHOD AND THE LOWER 6 DATA ROWS ARE GENERATED FROM RANDOMIZED SAMPLING METHOD.

IFNγ_i	IL12	IL6	TGFβ	IL17	RORγt	IFNγ_o	Tbet	FOXP3	Subset
0	0.25	0	0.5	0.09	0.01	0.90	0.50	0.99	Treg
0	0.25	0.25	0.25	0.89	0.94	0.31	0.33	0.20	Th17
0	0.5	0	0.75	0.09	0.01	0.89	0.47	0.99	Treg
0	0.75	1	1	0.92	0.99	0.27	0.26	0.16	Th17
0.25	0	0.75	0.75	0.91	0.98	0.10	0.27	0.17	Th17
0.25	0.5	0.25	0.25	0.89	0.93	0.31	0.33	0.20	Th17
0.44	0.65	0.80	0.60	0.91	0.99	0.28	0.28	0.17	Th17
0.92	0.69	0.97	0.12	0.92	0.99	0.30	0.35	0.17	Th17
0.04	0.48	0.67	0.46	0.91	0.98	0.28	0.29	0.17	Th17
0.76	0.44	0.63	0.57	0.91	0.98	0.28	0.29	0.18	Th17
0.29	0.40	0.57	0.51	0.91	0.98	0.29	0.29	0.18	Th17
0.34	0.14	0.35	0.42	0.90	0.96	0.30	0.31	0.19	Th17

Where the transformation matrix M_{Tran} is

$$\begin{bmatrix} 0.439 & 0.0145 & -0.0118 & -0.541 & 0.0699 \\ 0.443 & 0.0313 & 0.164 & -0.361 & 0.0363 \\ 0.479 & 0.0169 & 0.0422 & 0.539 & 0.00752 \\ 0.466 & 0.00758 & 0.0337 & 0.666 & 0.00722 \\ 0.425 & 0.0471 & 0.0456 & -0.145 & -0.135 \end{bmatrix}$$

The fitting error, the average absolute difference between the model predictions and real outputs from the data, of the linear model is shown in table II.

TABLE II

FITTING ERROR OF THE LINEAR REGRESSION MODEL

IL17	RORγt	IFNγ	Tbet	FOXP3
0.216	0.243	0.163	0.078	0.233

Considering the data are normalized within [0, 1], the fitting errors are obviously large. This indicate that the T cell differentiation is highly non-linear and linear regression is not a good tool to model this process.

B. Neural network model for cytokine concentrations

Neural network models can be used to model non-linear relationships. We use a package in R named neuralNet [11] to develop a model for T cell differentiation as shown in Figure 3.

This model has 4 nodes for inputs, 5 nodes for outputs, and one middle layer of 6 intermediate nodes. It fit the training data very well with a small number 0.4797 as the sum of the squared errors. It took 37984 steps to converge. The learning algorithm used is back propagation.

Since the fitting error is very small, we use this model to further predict the output cytokines with randomized input cytokine concentrations. The average prediction errors in terms of average absolute difference between the predictions and the outputs of the ODE model are shown in table III. The maximum error is 0.0253 that is low.

TABLE III

PREDICTION ERROR OF THE NEURAL NETWORK MODEL

IL17	RORγt	IFNγ	Tbet	FOXP3
0.0253	0.0119	0.0104	0.0101	0.0404

C. Neural network model for subset classification

With the success of neural network model for predicting concentrations of output cytokines, we further develop another neural network model for classifying cell subset. In the context of *Helicobacter pylori* pathogen-host immune responses simulation, Th1, Th17, and Treg are the three subsets of interests.

We asked the immunologists to classify the cells into subsets according to the concentrations of output cytokines. We use part of the data to train the model and part of the data to test the model. This model has four outputs corresponding to the three subsets, Th1, Th17, and Treg, and one unclassified. For the training

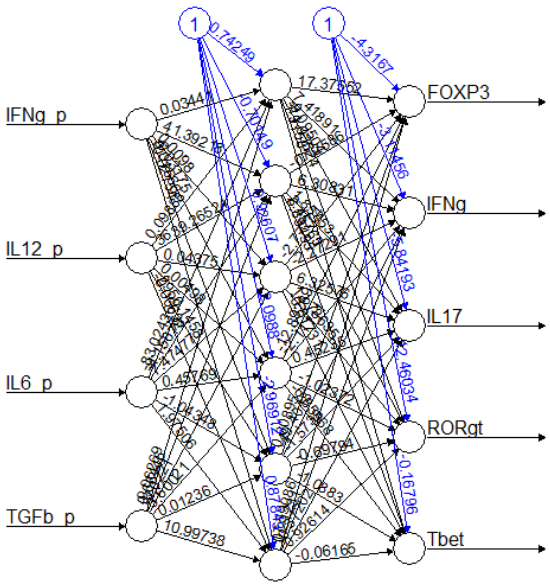

Fig. 3. A neural network model for T Cell differentiation. The inputs are cytokines that regulate the T cell differentiation and the outputs corresponding to the cytokines that secreted by the immune cells. There is a middle layer of 6 nodes in the feedforward network. The training algorithm used is back propagate.

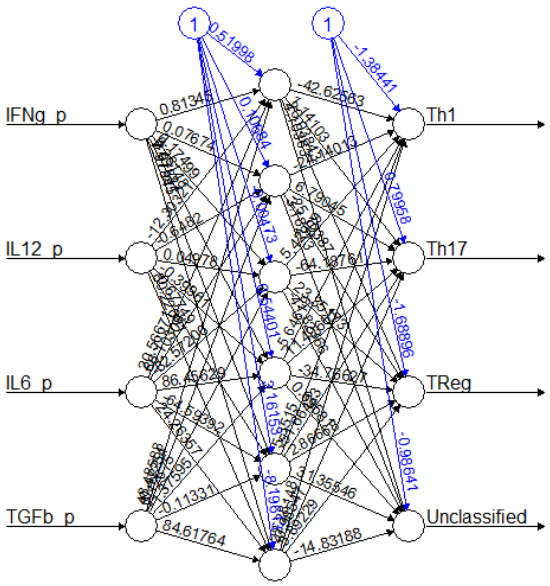

Fig. 4. A neural network model for CD4+ T Cell subset classification. The inputs are cytokines that regulate the T cell differentiation and the outputs corresponding to four categories: three classified T cell subsets, Th1, Th17, and Treg, and one unclassified. There is a middle layer of 6 nodes in the feedforward network. The training algorithm used is back propagate.

and testing data, the four outputs will have one 1 and three 0s. The cell is classified as the output subset if it value is 1.

Presented in Figure 4, this model takes 859 steps to converge and the sum of squared errors is a small number 0.518. Remember the training data has outputs of 1s and 0s; but the model outputs are numbers between 0 and 1. We use this model to predict another 100 data of randomized input cytokines. The predicted subset is the subset corresponding the maximum output. The prediction accuracy is 98% with 2 wrong predictions and 98 correct predictions.

V. Conclusions

In this study, we presented two neural network models for CD4+ T Cell differentiation and subset classification. Immune cell differentiation and subset classification are important immunological processes that are not fully characterized. Based upon our previous studies on the ODE model of CD4+ T cell differentiation and agent-based modeling for enteric immunity, developing multi-scale models requires significant reduction of the intracellular ODE model before integrating them into the inter-cellular agent-based models.

Immune cell differentiation is a non-linear process and linear regression models are not capable of fitting well the data. To address this problem, a feed-forward neural network model has been developed, focusing on modeling the relationship between the input external cytokines regulating the cell differentiation and the output cytokines secreted and externalized by the immune cell subsets. After training using back propagate algorithm, this neural network model accurately predicts the concentrations of the output cytokines. Furthermore we developed another neural network model for automatically classifying T cell subset. This model achieves 98% classification accuracy. These two models significantly reduce the ODE model complexity by focusing on the needs of multi-scale models and provide great modeling and prediction accuracy. These approaches are scalable and can be integrated into future multi-scale modeling efforts.

To our best knowledge, this is the first study using neural networks to model immune cell differentiation and subset classification. For future work, we will extend this study in the following three major areas:

- Validating the model directly using experimental

data instead of currently relying on the calibrated ODE model and thus being validated indirectly by experimental data

- Developing more neural network models for the differentiation and subset classification of other types of immune cells such as macrophage cells and B cells
- Integrating the neural network models into the agent-based models we developed using ENISI Visual for multi-scale models to study more integrated, broader and deeper scopes of immunological processes

ACKNOWLEDGEMENT

This work was supported in part by NIAID Contract No. HHSN272201000056C to JBR and funds from the Nutritional Immunology and Molecular Medicine Laboratory (URL: www.nimml.org).

REFERENCES

[1] S. Akira, S. Uematsu, and O. Takeuchi. Pathogen recognition and innate immunity. *Cell*, 124(4):783–801, 2006.

[2] W. C. Boyd et al. Fundamentals of immunology. *Fundamentals of Immunology.*, 1946.

[3] V. Brusic, G. Rudy, and L. C. Harrison. Prediction of mhc binding peptides using artificial neural networks. *Complex systems: mechanism of adaptation*, pages 253–260, 1994.

[4] A. Carbo, R. Hontecillas, B. Kronsteiner, M. Viladomiu, M. Pedragosa, P. Lu, C. W. Philipson, S. Hoops, M. Marathe, S. Eubank, K. Bisset, K. Wendelsdorf, A. Jarrah, Y. Mei, and J. Bassaganya-Riera. Systems modeling of molecular mechanisms controlling cytokine-driven cd4+ t cell differentiation and phenotype plasticity. *PLoS computational biology*, 9(4):e1003027, 2013.

[5] M. W. Craven and J. W. Shavlik. Using neural networks for data mining. *Future generation computer systems*, 13(2):211–229, 1997.

[6] J. E. Dayhoff and J. M. DeLeo. Artificial neural networks. *Cancer*, 91(S8):1615–1635, 2001.

[7] D. T. Gillespie. A rigorous derivation of the chemical master equation. *Physica A: Statistical Mechanics and its Applications*, 188(1):404–425, 1992.

[8] C. Goble, S. Pettifer, R. Stevens, and C. Greenhalgh. Knowledge integration: In silico experiments in bioinformatics. *The Grid: Blueprint for a New Computing Infrastructure*, 2, 2003.

[9] I. Goryanin, T. C. Hodgman, and E. Selkov. Mathematical simulation and analysis of cellular metabolism and regulation. *Bioinformatics*, 15(9):749–758, 1999.

[10] H. Groux, A. O'Garra, M. Bigler, M. Rouleau, S. Antonenko, J. E. de Vries, and M. G. Roncarolo. A cd4+ t-cell subset inhibits antigen-specific t-cell responses and prevents colitis. *Nature*, 389(6652):737–742, 1997.

[11] F. Günther and S. Fritsch. neuralnet: Training of neural networks. *The R Journal*, 2(1):30–38, 2010.

[12] M. T. Hagan and M. B. Menhaj. Training feedforward networks with the marquardt algorithm. *Neural Networks, IEEE Transactions on*, 5(6):989–993, 1994.

[13] S. Hoops, S. Sahle, R. Gauges, C. Lee, J. Pahle, N. Simus, M. Singhal, L. Xu, P. Mendes, and U. Kummer. COPASI–a COmplex PAthway SImulator. *Bioinformatics*, 22(24):3067–3074, Dec 2006.

[14] D. A. Horwitz, S. G. Zheng, J. Wang, and J. D. Gray. Critical role of il-2 and tgf-β in generation, function and stabilization of foxp3+ cd4+ treg. *European journal of immunology*, 38(4):912–915, 2008.

[15] M. Hucka, A. Finney, H. M. Sauro, H. Bolouri, J. C. Doyle, H. Kitano, A. P. Arkin, B. J. Bornstein, D. Bray, A. Cornish-Bowden, A. A. Cuellar, S. Dronov, E. D. Gilles, M. Ginkel, V. Gor, I. I. Goryanin, W. J. Hedley, T. C. Hodgman, J. H. Hofmeyr, P. J. Hunter, N. S. Juty, J. L. Kasberger, A. Kremling, U. Kummer, N. Le Novere, L. M. Loew, D. Lucio, P. Mendes, E. Minch, E. D. Mjolsness, Y. Nakayama, M. R. Nelson, P. F. Nielsen, T. Sakurada, J. C. Schaff, B. E. Shapiro, T. S. Shimizu, H. D. Spence, J. Stelling, K. Takahashi, M. Tomita, J. Wagner, and J. Wang. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics*, 19(4):524–531, Mar 2003.

[16] C. Janeway. Immunobiology: the immune system in health and disease. 2005.

[17] H. Kitano. Computational systems biology. *Nature*, 420(6912):206–210, 2002.

[18] H. Kitano, A. Funahashi, Y. Matsuoka, and K. Oda. Using process diagrams for the graphical representation of biological networks. *Nature biotechnology*, 23(8):961–966, 2005.

[19] K. Kohno, J. Kataoka, T. Ohtsuki, Y. Suemoto, I. Okamoto, M. Usui, M. Ikeda, and M. Kurimoto. Ifn-gamma-inducing factor (igif) is a costimulatory factor on the activation of th1 but not th2 cells and exerts its effect independently of il-12. *The Journal of Immunology*, 158(4):1541–1550, 1997.

[20] T. Korn, M. Mitsdoerffer, A. L. Croxford, A. Awasthi, V. A. Dardalhon, G. Galileos, P. Vollmar, G. L. Stritesky, M. H. Kaplan, A. Waisman, et al. Il-6 controls th17 immunity in vivo by inhibiting the conversion of conventional t cells into foxp3+ regulatory t cells. *Proceedings of the National Academy of Sciences*, 105(47):18460–18465, 2008.

[21] I. E. Lagaris, A. Likas, and D. I. Fotiadis. Artificial neural networks for solving ordinary and partial differential equations. *Neural Networks, IEEE Transactions on*, 9(5):987–1000, 1998.

[22] S. Lek and J.-F. Guégan. Artificial neural networks as a tool in ecological modelling, an introduction. *Ecological modelling*, 120(2):65–73, 1999.

[23] H. Lu, R. Setiono, and H. Liu. Effective data mining using neural networks. *Knowledge and Data Engineering, IEEE Transactions on*, 8(6):957–961, 1996.

[24] D. Machado, R. S. Costa, M. Rocha, E. C. Ferreira, B. Tidor, and I. Rocha. Modeling formalisms in systems biology. *AMB Express*, 1(1):1–14, 2011.

[25] P. R. Mangan, L. E. Harrington, D. B. O'Quinn, W. S. Helms, D. C. Bullard, C. O. Elson, R. D. Hatton, S. M. Wahl, T. R. Schoeb, and C. T. Weaver. Transforming growth factor-β induces development of the th17 lineage. *Nature*, 441(7090):231–234, 2006.

[26] Y. Mei, R. Hontecillas, X. Zhang, K. Bisset, S. Eubank, S. Hoops, M. Marathe, and J. Bassaganya-Riera. Enisi visual, an agent-based simulator for modeling gut immunity. In *Bioinformatics and Biomedicine (BIBM), 2012 IEEE International Conference on*, pages 1–5. IEEE, 2012.

[27] T. R. Mosmann and S. Sad. The expanding universe of t-cell subsets: Th1, th2 and more. *Immunology today*, 17(3):138–146, 1996.

[28] P. J. Murray and T. A. Wynn. Protective and pathogenic functions of macrophage subsets. *NATURE REVIEWS— IMMUNOLOGY*, 11:723, 2011.

[29] P. B. Snow, D. S. Smith, and W. J. Catalona. Artificial neural networks in the diagnosis and prognosis of prostate cancer: a pilot study. *The Journal of urology*, 152(5 Pt 2):1923–1926, 1994.

[30] H. White. Learning in artificial neural networks: A statistical perspective. *Neural computation*, 1(4):425–464, 1989.