

HHS Public Access

Trends Immunol. Author manuscript; available in PMC 2020 February 21.

Published in final edited form as:

Author manuscript

Trends Immunol. 2019 August ; 40(8): 665–668. doi:10.1016/j.it.2019.06.006.

Computational and Systems Immunology: A Student's Perspective

Zinaida Good^{1,2,3,4,7,*}, Jacob Glanville^{1,3,4,8}, Marvin H. Gee^{1,5,9}, Mark M. Davis³, Purvesh Khatri^{6,*}

¹Computational and Systems Immunology track, PhD Program in Immunology, Stanford University, Stanford, CA, USA

²Baxter Laboratory in Stem Cell Biology, Stanford University, Stanford, CA, USA

³Department of Microbiology and Immunology, Stanford University, Stanford, CA, USA

⁴Department of Pathology, Stanford University, Stanford, CA, USA

⁵Department of Molecular and Cellular Physiology, Stanford University, Stanford, CA, USA

⁶Department of Medicine, Center for Biomedical Research, Stanford University, Stanford, CA, USA

⁷Current address: Stanford Cancer Institute and Department of Biomedical Data Science, Stanford, CA, USA

⁸Current address: Distributed Bio, South San Francisco, CA, USA

⁹Current address: 3T Biosciences, Menlo Park, CA, USA

Abstract

The big data revolution has transformed the landscape of immunology research. As inaugural students of Stanford's new Computational and Systems Immunology PhD track, we share our experiences and advice with other institutions considering a similar program.

An Evolving Field of Immunology

Technological revolution in immunology is producing ever-increasing amounts of data [1]. Translating these data into insights requires advanced statistics, data mining, and machine learning skills. It also requires a strong background in immunology to ask suitable questions, design appropriate experiments, and recognize discoveries in data trends. The paucity of researchers with a combination of these skills has created a bottleneck in systems-level immunology research. Training a new generation of computational immunologists will provide opportunities for discoveries that simply have not been possible with traditional approaches [2,3].

^{*}Correspondence: zinaida@stanford.edu (Z. Good) and pkhatri@stanford.edu (P. Khatri).

As immunology PhD students at Stanford University, we praise our faculty for anticipating this need and establishing a Computational and Systems Immunology (CSI) track within the PhD Program in Immunology in 2012. Initiated by Mark Davis and Atul Butte (the latter now at University of California, San Francisco), the CSI track runs in parallel with the Molecular, Cellular, and Translational Immunology (MCTI) track. To our knowledge, Stanford University remains the only academic institution to offer a formal graduate program in computational immunology.

As CSI students, we are grateful for the skills we received and the opportunities that this training has opened. Here, we describe our program and enthusiastically recommend that other institutions offer a similar opportunity to their trainees. In collaboration with the CSI faculty and Immunology leadership, we further provide recommendations on creating and continuously refining a successful computational immunology program.

The Need for a Computational Immunology Program

As novel multiplexed technologies accelerate high-throughput interrogation of immune phenomena, these data accumulate in large public repositories, such as the NCBI Gene Expression Omnibus, NIAID-funded ImmPort, and ImmGen. An increasing number of studies are utilizing data from these repositories to identify hitherto unknown immunology [4–6]. A broad spectrum of problems that require computational immunology skills include immune repertoire analysis [7,8], structural basis of antigen recognition [9,10], single-cell differentiation trajectory construction [11,12], novel cell subset detection [13], simulation of immune processes [14], and clinical outcome prediction [4–6,15]. Thus, the field needs computational immunologists who are able to integrate their understanding of technology platforms and databases with expertise in immunology and data analysis.

Currently, computationally savvy immunologists are often self-taught, making it easy to cultivate bad habits and be swayed by a 'method of the week'. Collaborations with bioinformaticians are important, but without an expert with hybrid training, there can be delays, mix-ups, and failures to recognize data inconsistencies or unanticipated discoveries. Thus, immunology trainees would benefit from a structured curriculum designed to master computational immunology problems.

Learning Objectives in Stanford's CSI Track

The Stanford PhD program in Immunology offers two tracks: MCTI and CSI. Our CSI program accommodates incoming students skilled in either immunology or programming and provides an opportunity to cover the remaining material in year 1 (Figure 1A). In the first year, students in both MCTI and CSI tracks rotate in research laboratories and take many of the same courses. In addition to traditional MCTI course work that builds expertise in immunology (with a reduction of one core course and one MCTI elective), CSI students obtain computational skills through courses in computer science, statistics, and bioinformatics. As we do not select a track until the end of the first year, undecided students are able to try CSI courses and a computational rotation without committing to the track. In the second year, all students define a thesis project, take a qualifying exam, and serve as

Trends Immunol. Author manuscript; available in PMC 2020 February 21.

Good et al.

teaching assistants. CSI students also complete the remaining core courses, including a newly designed series focused specifically on computational immunology problems (Figure 1B). The series consists of seminars and courses focused on understanding, practicing, and solving problems specific to computational immunology (https://med.stanford.edu/immunol/phd-program/resources/curriculum.html). After building a strong foundation through coursework and defining a thesis project, CSI students take two electives in years 3+ to refine project-specific skills, which typically include courses in advanced statistics, bioinformatics, or machine learning.

Since 2012, 11 of 57 students began this training, and three (J.G., M.H.G., and Z.G.) have graduated. At this time, the CSI track takes 4–5 years (mean 4.6 years); comparable, or perhaps slightly less than a graduation mean of 5.7 years for the Stanford Immunology PhD program. In addition, students focusing on computational immunology prior to the formation of the CSI track also graduated in 4–6 years. These limited data suggest that students in both MCTI and CSI tracks graduate in a similar timeframe despite the intensive CSI coursework. Finally, computational immunology training has resulted in innovative published work [7–10,12,15].

Recommendations to Future Computational Immunology Programs

Based on Stanford's program, in this Scientific Life article, we advise other institutions to foster a culture of continuous learning through coursework (if not already incorporated). Specifically, creating a formal computational immunology program enables students to take a heavy course load without taking time away from research. We posit that the time invested in immunology, computer science, and CSI courses has so far yielded a high return by accelerating thesis research. Other interdisciplinary PhD programs (e.g., biomedical informatics) have also shown that combining coursework requirements across multiple subject areas is certainly feasible.

Next, we recommend developing a rigorous core curriculum. The core course work should emphasize advanced immunology, statistics, computer science, and dynamic modeling. A set of dedicated courses should then cover specialized problems, data formats, and methodologies specific to computational immunology (e.g., CSI core; Figure 1B). Flexible electives can maximize the relevance of coursework to student theses. Thus, in addition to fostering support for students taking courses, developing an effective track requires faculty to dedicate time to developing and teaching these new courses.

To reduce the burden on existing faculty and to provide sufficient guidance to students, institutions should be encouraged to recruit additional faculty and postdocs working in computational immunology. Another possibility is a joint mentorship of a given student by both experimental and computational faculty members. For example, of the immunology students who started their PhD program between 2012 and 2017, 8/11 (73%) CSI students were co-advised, in contrast to 7/46 (15%) MCTI students. Although co-mentorship offers numerous benefits to a student (advisor support, access to laboratory expertise and resources, and more independence), arguably it has some drawbacks (twice the laboratory-related activities, and potentially conflicting expectations from co-advisors). Thus, although

Trends Immunol. Author manuscript; available in PMC 2020 February 21.

Good et al.

comentoring can be beneficial, a decision as to whether a student should be comentored needs to be carefully considered on a case-by-case basis.

A successful program relies on the self-motivation of graduate students, but also requires providing access to advising, as needed (Figure 1C). Program advisors might potentially tailor an individual student's curriculum to ensure that incoming PhD students are challenged at just the right level, learn skills directly applicable to their work, and hence, be able to graduate in a reasonable time. Finally, a continuous two-way dialog between students and faculty is key to ensuring collaborative and continuous improvements in the curriculum and training.

Building a Computational Immunology Community

Aside from the course work, building a scientific community within Stanford's CSI track was a key to its early success (Figure 1C). In the program, senior students advise junior students through formal quarterly advising dinners, one-onone mentoring programs, and informally. Community support ensures that nobody 'falls through the cracks', and individual stories help break the 'it's too late for me to learn how to code' barrier. Postdocs, faculty, and administrators set norms, advise students, and provide the necessary resources and support. Through CSI seminars, students learn about the emerging computational immunology concepts and obtain career advice during informal dinner discussions with each speaker.

Concluding Remarks

Within the past decade, big (voluminous), deep (high-dimensional), and multiomics data have become commonplace in immunology and other areas of biomedical research, such as neuroscience, cancer biology, and developmental biology. As a new generation of scientists, we are expected to build and utilize these resources to pose and answer outstanding questions in immunology. We are thus grateful to programs such as Stanford's new CSI track, preparing us for the future. We strongly encourage other immunology graduate programs to offer a similar option to their students. The concepts presented here may be applicable to designing extended education programs for clinical fellows, postdocs, and other trainees.

Acknowledgments

We thank David R. Glass, Geoffrey T. Ivison, Aditya M. Rao, Graham L. Barlow, Erika Bongen, Sean C. Bendall, Garry P. Nolan, Patricia Jones, Olivia M. Martinez, Nikesh Kotecha, Parag Mallick, Dara M. Strauss-Albee, Matthew H. Spitzer, Michael H. Birnbaum, Bryan J. Xie, Eden Maloney, Sarah Kongpachith, Daniel Lu, Thomas Keller, Dmitry Tebaykin, Otavio B. Good, Jonathan E. Wosen, Maureen Panganiban, Andrew J. Gentles, and Robert Tibshirani for insightful discussions and comments on the earlier versions of this manuscript. We also acknowledge Atul Butte, Patricia Jones, Olivia M. Martinez, Nikesh Kotecha, and Parag Mallick for their roles in helping to develop Stanford's CSI track. The PhD Program in Immunology at Stanford Immunology is supported by NIH training grant 5T32 AI007290-34. Z.G. and M.M.D are members of the Parker Institute for Cancer Immunotherapy, which supported this work.

References

- 1. Davis MM et al. (2017) Systems immunology: just getting started. Nat. Immunol 18, 725–732 [PubMed: 28632713]
- 2. Spreafico R et al. (2015) Training the 21st century immunologist. Trends Immunol. 36, 283–285 [PubMed: 25911462]
- Schultze JL (2015) Teaching 'big data' analysis to young immunologists. Nat. Immunol 16, 902– 905 [PubMed: 26287590]
- Andres-Terre M et al. (2015) Integrated, multi-cohort analysis identifies conserved transcriptional signatures across multiple respiratory viruses. Immunity 43, 1199–1211 [PubMed: 26682989]
- 5. Chowdhury R et al. (2018) A multi-cohort study of the immune factors associated with M. tuberculosis infection outcomes. Nature 560, 644–648 [PubMed: 30135583]
- Scott MKD et al. (2019) Increased monocyte count as a cellular biomarker for poor outcomes in fibrotic diseases: a retrospective, multicenter cohort study. Lancet Respir. Med 7, 497–508 [PubMed: 30935881]
- Glanville J et al. (2017) Identifying specificity groups in the T cell receptor repertoire. Nature 547, 94–98 [PubMed: 28636589]
- Han A et al. (2014) Linking T-cell receptor sequence to functional phenotype at the single-cell level. Nat. Biotechnol 32, 684–692 [PubMed: 24952902]
- 9. Gee MH et al. (2018) Antigen identification for orphan T cell receptors expressed on tumorinfiltrating lymphocytes. Cell 172, 549–563 e516 [PubMed: 29275860]
- Gee MH et al. (2018) Stress-testing the relationship between T cell receptor/peptide-MHC affinity and cross-reactivity using peptide velcro. Proc. Natl. Acad. U. S. A 115, E7369–E7378
- 11. Schiebinger G et al. (2019) optimal-transport analysis of single-cell gene expression identifies developmental trajectories in reprogramming. Cell 176, 928–943 e922 [PubMed: 30712874]
- 12. Good Z et al. (2019) Proliferation tracing with single-cell mass cytometry optimizes generation of stem cell memory-like T cells. Nat. Biotechnol 37, 259–266 [PubMed: 30742126]
- Villani AC et al. (2017) Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. Science 356 Published online 4 21, 2017 10.1126/science.aah4573
- An G et al. (2017) Optimization and control of agent-based models in biology: a perspective. Bull. Math. Biol 79, 63–87 [PubMed: 27826879]
- Good Z et al. (2018) Single-cell developmental classification of B cell precursor acute lymphoblastic leukemia at diagnosis reveals predictors of relapse. Nat. Med 24, 474–483 [PubMed: 29505032]

(A)

Computational and systems immunology curriculum

	Emphasis	Immunology core			CS and statistics core		
Prerequisites (or year 1)	Research experience	Chemistry of biological processes (3) Molecular and cellular immunology (4)			Programming met Programming abs	hodology (5) tractions - accelerated (5)	
Year 1	Rotations	Foundations in experimental biology (6) Molecular and cellular immunology literature review (1) The responsible conduct of research (1) Advanced immunology I (3) Advanced immunology II (3) Immunology journal club (1) Seminar in immunology (1)			Biostatistics (5) Introduction to probability for computer scientists (5) Design and analysis of algorithms (5) CSI Core Introduction to applied comp. Tools in immunology (2) Seminar in computational and systems immunology (1)		
Year 2	Thesis research Qualifying exam	TeachingRepresentationTeaching in immunology (2 courses)IntrodEssertAdvart		Representation Introduction to Essential meth Advanced com	sentations and algorithms for computational molecular biology (4) uction to biomedical informatics research methodology (3) itial methods in computational and systems immunology (3) aced computational and systems immunology (3)		
Years 3–5	Committee meetings Thesis defense		Electives - Ar - M - B		anced statistics chine learning nformatics	- Statistical learning - R programming	

(B) Specialized computational and systems immunology courses

Course	Credits	Goal
Introduction to applied computational tools in immunology	2	To introduce students to diverse problems, technologies, and tools specific to computational immunology. Includes a grant proposal.
Essential methods in computational and systems immunology	3	To provide students with practice working on published computational immunology datasets. Includes assignments on data analysis.
Advanced computational and systems immunology	3	To enable students to develop novel tools and algorithms for their computational immunology project. Includes preparing a manuscript.
Seminar in computational and systems immunology	1	To expose students to ongoing computational immunology research and leading scientists. Includes dinner discussions with each speaker.

(C)

Critical components of success

Mentorship

Student quarterly advising dinners Student 1x1 mentor program Faculty mentor guidance Program and administrative support

Curriculum evolution Continuous dialog among faculty and students Tailor coursework to student's project needs Subcommittee on curriculum for the track Town hall meetings

Community

Seminar series (Immunology, CSI) Journal clubs (student-led) Social events (with faculty, postdocs, staff) Outreach events (with industry, startups, schools)

Trends in Immunology

Figure 1. Stanford's Computational and Systems Immunology (CSI) Curriculum and Critical Components.

(A) CSI track timeline, research emphasis, and coursework, as of 2018–2019. The number of credits for each course are shown in parentheses. This curriculum has evolved since 2012 and may be developed further (see Curriculum Evolution). (B) Specialized CSI courses. Full descriptions are available on https://med.stanford.edu/immunol/phd-program/resources/ curriculum.html. (C) Critical components that have enabled the success of the CSI track at Stanford University. Abbreviations: CS, computer science.