

SPONSOR AND CO-SPONSOR STATEMENTS

Justin D. Silverman, M.D., Ph.D.

Research Support Available

Source	ID	Title	PD/PI	Start	End	Amount
NIH	1R01GM148972-01	Addressing Measurement Limitations for Sequence Count Data	Silverman, JD	9/2022	8/2025	\$599,945
PNNL	1R01GM149650-01	A Novel high resolution MS platform for high-throughput screening of G protein- coupled receptors	Jacobs, J & Rajagopal S	6/2023	3/2027	\$86,000
Duke Univ	2R01DK116187-06A1	Dietary plant diversity and the human gut microbiome	David, LA	5/2024	5/2028	\$168,134

Beyond the funding listed above and other proposal that are currently submitted or planned, the PI has over \$400,000 in startup funds from The College of Information Science and Technology (IST) at the Pennsylvania State University which can be used to support Mr. Sugarman's research and training.

Sponsor's Previous Fellows/Trainees

I have trained and mentored two doctoral students (both graduated with Ph.D. degrees) and two masters students (both graduated with M.S. degrees) and am currently training or mentoring 3 doctoral students and 1 Assistant Research Professor. Of the two doctoral students that have graduated from my lab, one currently holds a senior position in industry and the other is pursuing postdoctoral education.

Name	Position in Silverman Lab	Current Position and Institution
Kimberly Roche, Ph.D.	Predoctoral	Senior Translational Scientist at Tempus Labs
Emily Van Syoc, Ph.D.	Predoctoral	Postdoctoral Research, PSU
Farhani Momotaz, M.S.	Masters Student	Research Associate, PSU
Zhao Ma, M.S.	Masters Student	Research Assistant, Zhejiang University

Keith C. Cheng, M.D., Ph.D.

Research Support Available

My lab also has substantial research support that will provide Andrew with a surplus of the resources needed to carry out this proposal. In addition to our track record of NIH funding, our group maintains active collaboration with synchrotron beamlines such as the Lawrence Berkeley National Laboratory where Andrew has collected preliminary data and will continue to carry out experiments to support the proposed work.

Source	ID	Title	PD/PI	Start	End	Amount
NIH	1R24OD18559	Groundwork for a Synchrotron MicroCT Imaging Resource for Biology	Cheng, KC	8/2015	7/2024 (NCE)	\$2,680,046
NIH	R24OD18559	Renovation Supplement to Groundwork for a Synchrotron MicroCT Imaging Resource for Biology	Cheng, KC	8/2022	7/2024	\$231,489
DOE/ LBNL	ALS-11922	X-ray histotomography applications of new wide-field, submicron resolution lens and camera systems	Cheng, KC	7/2022	6/2025	Synchrotron Imaging Time
NIH	1R24OD035407-01A1	Building a Wide-field, High-resolution Histotomography Resource for Biology	Cheng, KC	5/2024	5/2028	\$3,862,968 (score 20) Council 1/24

Co-Sponsor's Previous Fellows/Trainees

I train and mentor doctoral (including MD, Veterinary, PhD, and post-graduate) students in multiple fields, including Genetics, Biomedical Sciences, Pathology, Bioinformatics and Genomics, and Anatomy, including several MD/PhD students. As an active participant in intercollege and interdisciplinary programs including the Huck Institute for the Life Sciences and Institute for Computational and Data Sciences involving undergraduate and graduate students and faculty on multiple Penn State's campuses, and have served on graduate committees for students in Information Sciences and Technology and Computer Sciences, I interact with students at multiple levels of training

and fields. I am currently training or mentoring 3 doctoral students, a postdoctoral student, a post-baccalaureate student, and 1 Assistant Research Professor. My mentees include the following:

Name	Current Position and Institution
Rebecca Lamason, Ph.D. (postbaccalaureate)	Associate Professor of Biology, MIT
Darin Clark, Ph.D. (postbaccalaureate)	Assistant Professor of Radiology, Duke Univ
Amogh Adishesha, Ph.D. (IST)	Applied Scientist, Captions (captions.ai)
William Zinnanti, M.D., Ph.D. (Biomed Sci)	Private Practice for Child and Adult Neurology
Brian Canada, Ph.D. (Bioinformatics/Genomics)	Chair, Department of Computer Science, Univ of South Carolina
Yifu Ding, M.D. Ph.D (Biomed Sci)	Resident Physician in Radiation Oncology, Emory University
Spencer Katz, M.D. Ph.D (Biomed Sci)	Resident Physician in Pediatric Medical Genetics, Cincinnati Children's Hospital
Maksim Yakovlev, Ph.D (Biomed Sci)	Postdoctoral Researcher, Argonne National Laboratory

Sponsor Statement

Training Plan, Environment, Research Facilities

Training Plan: The goal of this F30 application is to formally train Andrew in basic and translational sciences. He will continue to receive rigorous and comprehensive training from a variety of sources including coursework, one-on-one mentorship, discussion with other graduate students, and group meetings with Dr. Silverman and Dr. Cheng and Synchrotron and University Imaging Groups.

Andrew has completed and excelled in all formal coursework requirements for the Bioinformatics and Genomics PhD program and has passed his comprehensive exam. He will supplement this coursework to advance his skills in statistics and to support aims 2 and 3 of this proposal. In addition to coursework, Andrew has made strides as a researcher over the first year and a half of his PhD. He has studied the fundamentals of probability theory in the Silverman lab, and developed expertise in both sample-preparation, micro-CT physics, and image processing in the Cheng lab. His knowledge of statistics and machine learning and experience with synchrotron microCT allowed him to contribute as an author to a recent publication in *eLife* that constitutes part of the foundation of this proposal. With regard to sample preparation and micro-CT, Andrew has absorbed knowledge from prior students in the Cheng lab and now occupies a leadership role in synchrotron trips.

As Andrew continues to conduct experiments in support of the goals of this proposal, his training will focus on two gaps in his development as a trainee. First, Andrew will receive further one-on-one training in statistical methods under the direction of Dr. Silverman. He will work through the “Matrix Algebra from a Statistician’s Perspective” textbook by David A. Harville, and in addition will take the next course in Topological Data Analysis offered at Penn State to enhance his knowledge of persistent homology and develop his expertise in support of Aim 3.

Clinical Training Plan: Even though both of us work full-time on basic science research, both Dr. Cheng and I have completed clinical training of our own and understand the importance of continuing to cultivate knowledge and patient-centered clinical skills during the graduate studies. Investing in clinical training during Andrew’s graduate years will prepare him for his transition back to the third year of medical school and for his long term goal of becoming an independent physician scientist. In Andrew’s case, we view that this will actually enhance the value of his graduate studies, given that Andrew draws substantial motivation from what he has observed in clinic and has initiated this proposal with the hope it will be able to contribute to problems that affect cancer patients.

Andrew participates in the Clinical Exposure Program (CEP) since he has completed his comprehensive exam. CEP is a tool for MD/PhD students to prepare for the clinical portions of medical training while they complete their PhDs. He has chosen to work with Dr. Lilia Reyes in the Pediatric Emergency Department to focus on becoming a well-rounded medical student and ultimately a versatile physician scientist. To supplement this work, he will also continue to shadow Dr. Raymond Hohl, an attending in hematology/oncology at Hershey and the director of the Penn State Cancer Institute, in addition to serving as one of Andrew’s thesis committee members. Committee member Dr. Joshua Warrick, Director of Anatomic Pathology at the College of Medicine, brings invaluable perspective gained from a decade of supervising residency training, and therefore work with Dr. Cheng (who also is Board-certified in Anatomic Pathology) to ensure that the projects Andrew pursues are of maximal scientific and clinical value.

Professional Development Plan: Andrew will have career development opportunities through presenting at multiple conferences, seminars, workshops, and participating in unique research focused lab trips. He has already presented his work at the MD/PhD National Student Conference and the Bioinformatics and Genomics program retreat. Presentation skills are essential components of a successful career in science and Andrew will have ample opportunities to develop strong scientific communication abilities in both of our labs.

We will be sure that Andrew is also able to attend the MD/PhD retreat twice a year, where he will observe and interact with keynote speakers and colleagues that will also contribute mentorship to his development as a physician-scientist. Wherever possible, Andrew will also engage in other presentations and seminars on campus such as the Graduate Student Research Forum, the Bioinformatics and Genomics Student Seminar, and the Experimental Pathology Colloquium.

Environment: Andrew is a graduate student in my laboratory and is part of our regular lab meetings. These meetings are devoted to the development and application of statistical methods for complex biomedical data. A full time assistant research professor (Dr. Michelle Nixon, PhD in Statistics) is employed by the lab and is available to assist graduate students in their research. Lab meetings are held for 2 hours every other week during which time one student or faculty member presents their work with ample time for questions and discussion. Presenters at lab meeting rotate and each student or faculty member presents approximately once every other month. Beyond lab meetings, I meet with Andrew individually for between 1-2 hours weekly to discuss his progress and provide training and/or career mentorship. These meetings are held more frequently during key periods such as during manuscript preparation, preparation for committee meetings, or preparation for conference presentations. We have established Andrew's thesis committee which meets annually to discuss his progress. Beyond these annual meetings Andrew meets regularly with each committee member to discuss aspects of his research that intersect with their own expertise.

Andrew has access and regularly interacts with a wide range of extramurally funded researchers with expertise directly applicable to the proposed work and to Andrew's larger career goals. I am an assistant professor in the College of Information Science and Technology (IST), the Department of Statistics, and the Department of Medicine. I have formal training in both statistics (PhD) and medicine (MD) which gives me a keen appreciation of both the translational context of key biomedical questions as well as the statistical challenges associated with answering those questions. I am currently the PI on an NIH R01 award which focuses on developing key theory and tools for scale reliant inference that are complementary yet non-overlapping with the proposed work (1R01GM148972-01). Other extramurally funded faculty that Andrew regularly interacts with include Dr. Francesca Chiaromonte (Statistics, Focus on Bioinformatics and Genomics), Dr. Matthew Reimherr (Statistics, Focus on Functional Methods for Biostatistics), and Dr. Vasant Hanovar (IST, Focus on Statistical Inference and Machine Learning for Biomedical Data).

Outside of the lab and his thesis committee, there are numerous venues on campus that regularly feature current work from extra and intra-mural researchers. These include the Bioinformatics and Genomics Colloquium Series, The Statistics Seminar Series, and the Microbiome Center Seminar Series.

Research Facilities: Andrew will have access to the facilities and resources necessary to undertake and complete the proposed work. Andrew has dedicated office space in my lab located in the Westgate building at the Pennsylvania State University. This building is modern (built in 2004 and recently renovated), centrally located on campus, and within a 5 minute walk from the Statistics department and the Huck Life Sciences Institute. Andrew's office space is just down the hall from my own. Andrew has access to numerous high performance computing clusters both through the College of IST as well as through the Institute for Computational and Data Science where I have a dedicated allocation for my lab on the Roar cluster. The Roar cluster contains over 36,500 computing cores and 25 PB of storage. Andrew will have access to this entire cluster as well as having priority (<1min wait time) access to my personal allocation which includes one high-memory node (40 cores, 1TB RAM), two standard-memory nodes (each with 40 cores and 256GB RAM), 10TB Active Group Storage, and 20TB Nearline/Archive Storage. To supplement these capabilities the Roar cluster also has a help-desk service (i-Ask) which assists users with all aspects of cluster computing from debugging job execution to systems administration and database maintenance. Beyond computing resources, the Penn State University Libraries rank among the top 10 North American research libraries based on the Association of Research Libraries Library Investment Index Rankings. The library system consists of 36 libraries at 24 locations throughout the Commonwealth of Pennsylvania. The University Libraries house a collection of nearly 6 million items, with annual additions of roughly

100,000 volumes. The libraries have access to 579 online databases and other e-resources and subscribe to nearly 118,000 online, full-text journals.

Number of Fellows/Trainees To Be Supervised During the Fellowship

Four, in addition to Andrew:

- Tinghua Chen, Graduate Student (Informatics)
- Kyle McGovern, Graduate Student (Bioinformatics and Genomics)
- Won Gu, Graduate Student (Statistics)
- Maxwell Konnaris, Graduate Student (Bioinformatics and Genomics)

Applicant's Qualifications and Potential For A Research Career

While I have only formally advised Andrew for the past year and a half (since he started in the PhD portion of his training), it is important to note that I have been working with Andrew for almost three and a half years. In that time, I have watched Andrew combine his intellect, creativity, and grit to great effect. For example, Andrew learned the foundations of probability theory, Frequentist statistics, and topological data analysis in just over two months after he decided he wanted to learn how to rigorously model shape in 3D images. This is a remarkable feat and, in my experience, tantamount to crushing a wall that other students would simply walk away from. Andrew displays incredible intellectual curiosity and has the grit and intellect to follow that curiosity. Combined with his long-term interests in cancer, I have no doubt that Andrew will ultimately lead a independent research program which will improve human health.

Co-Sponsor Statement:

It is important to note that the development of a 3D computational phenomics for cancer is at its start, highly technical, and involves a degree of interdisciplinary collaboration that is challenging and non-traditional. The first reason this is important is that the increasing complexity of science means that trainees like Andrew who are trained to handle this complexity, will likely be enabled to make unique and important contributions that strictly focused scientists cannot.

Since I particularly enjoy the pursuit of science enabled by interdisciplinary collaboration that cannot be accomplished by one individual alone, I culture a learning environment that encourages learning habits of mind that allow one to productively steer through the details of complexity, and to be flexible in dealing with different personality types and disciplinary vocabularies to make significant interdisciplinary advances. Andrew is a particularly gifted student in that only a fraction of students have the dedication to science and the smarts to understand that we are addressing a large gap, and want to. Then only a fraction of those students can actually do so. In this case, Andrew has understood enough of how pathology and genetics work to understand biological and medical questions, understand enough chemical and biophysics to understand the basis of sample preparation, learn enough micro-CT physics using both cone-beam based local and parallel-beam based synchrotron x-ray sources, image processing, machine learning, and interacting in an multidisciplinary environment - a strength of the Cheng lab.

Andrew caught on quickly in terms of all these areas, being aware of, but never shying away from the complexity. He showed leadership skills in first learning how synchrotron micro-CT imaging trips work, and then being a primary manager of the most recent experimental trip to the LBNL. He led discussion of phase-based, unstained sample microCT based on readings from the literature as personally guided by Dr. Cheng. The striking preliminary data inspired and are foundational to this proposal; its extensions across tissue and sample types will contribute to multiple publications. Please also note that not only has Andrew already participated productively in other publications in the lab, he is also contributing to the projects of others in the lab in a way that will be mutually beneficial and result in multiple collaborative publications. I have mentored numerous PhD and MD/PhD students, and have significant experience and understanding of the various types of qualities that can drive success in interdisciplinary science. Andrew possesses a wonderful set of qualities, giving me every faith that he will become an excellent physician-scientist and make impactful contributions to cancer research.