Hoyin Chu

https://hoyinchu.github.io/

Email : hoyin_chu@dfci.harvard.edu

Mobile : 646-715-2873

EDUCATION

Northeastern University Boston, MA

- *B.S. in Computer Science and Mathematics* *Sep. 2016 – May. 2021*
Magna Cum Laude

RESEARCH EXPERIENCES

- **Computational Biologist, Dana-Farber Cancer Institute** Boston, MA
Advisors: Saud AlDubayan, Eliezer Van Allen *Jun 2021 - Present*

  - Independently led multiple projects over a range of topics in cancer genomics, including uncovering the germline genetic architecture of breast cancer in understudied populations, understanding the influence of germline elements on disease presentation, identifying predisposition genes in rare cancers, associating tumor's genetic characteristic with immunotherapy response and more. These findings have been summarized in multiple peer-reviewed manuscripts, with many having first or co-first authorships.

  - Performed large-scale genomic analysis using cloud computing tools (Terra, Google Cloud) and developed more than 50 workflows for common genomic tasks including sequence alignment, variant calling (short and structural variants), VCF manipulation, genotype phasing and imputation, genome and transcriptome wide association analysis, polygenic risk score calculation, statistical analysis and developed novel methods for data mining. The workflows are extensively documented and are shared amongst team members, which improved their analysis and resulted in co-authorships in multiple peer-reviewed papers.

  - Acquired additional data to improve the statistical power of existing projects through initiating conversations with internal and external collaborators as well as curating data from publicly available databases such as the sequencing read archive. Presented research progress and learnings from conferences in multiple lab meetings and assisted with onboarding new members.

- **Researcher Assistant, Northeastern University** Boston, MA
Advisor: Predrag Radivojac *Mar 2020 - Present*

  - Identified population subgroups at high-risk of developing adverse pregnancy outcomes using association rules mining. A non-linear interactions between exercise, genetic predisposition as quantified by polygenic risk scores, and risk for developing gestational diabetes was discovered. In addition, racial-ethnic disparities in high-risk subgroups were quantified using Gini coefficient and potential effects of policy interventions were investigated. Findings from the project were summarized in multiple first or co-first authorship manuscript, including a publication in JAMA Open Network and an oral presentation at the Pacific Symposium on Biocomputing.

  - Constructed protein networks from multiple publicly available databases and utilized graph mining tools to discover protein complexes. Evaluated the performance of various machine learning models such as distance metric, SVM, neural network ensembles on protein-protein interaction prediction using measures such as area under ROC curve, precision-recall curve, and obtained confidence intervals via bootstrapping.

  - Collaborated with computer scientists, clinicians, and biologist from multiple institutes and communicated findings through presentations to an diverse audience including clinicians and researchers.

- **Research Experience for Undergraduate, Northeastern University** Boston, MA
Advisor: Jonier Autunes *May 2019 - June 2019*

  - Modeled biological networks with a death-birth update rule using graph libraries and reduced simulation run-time by 40% by implementing a custom data structure and optimizing the existing algorithm

  - Visualized experimental results using matplotlib and gave presentation on evolutionary dynamics to an audience of faculties in the mathematics department

Published/Accepted in Peer-reviewed Journals/Conferences: (* - Equal Contribution)

1. **Hoyin Chu**, Rashika Ramola, Shantanu Jain, David M. Haas, Sriraam Natarajan, Predrag Radivojac, *Using Association Rules to Understand the Risk of Adverse Pregnancy Outcomes in a Diverse Population*, Pacific Symposium on Biocomputing, accepted for publication and oral presentation, (January 2023)

2. **Hoyin Chu**, Abdullah Al-Sulaiman, Mohammed Al-Jumaan, Seunghun Han, Sabrina Y. Camp, Riaz Gillani, Yousef Al Marzooq, Fatmah Almulhim, Chitti Babu Vatte, Arij Alnimer, Afnan Almuhanna, Eliezer Van Allen, Amein Al-Ali, Saud H AlDubayan, *Multi-modal characterization of ultra-rare germline genetic variants driving breast cancer risk in the indigenous Arab population*, Abstract, Poster Presentation, the Annual Meeting of The American Society of Human Genetics , (October 2022)

3. Kymberleigh A. Pagel*, **Hoyin Chu***, Rashika Ramola*, Rafael F. Guerrero, Judith H. Chung, Samuel Parry, Uma M. Reddy, Robert M. Silver, Jonathan G. Steller, Lynn M. Yee, Ronald J. Wapner, Matthew W. Hahn, Sriraam Natarajan, David M. Haas*, Predrag Radivojac*, *The influence of genetic predisposition and physical activity on risk of Gestational Diabetes Mellitus in the nuMoM2b cohort*, JAMA Network Open, doi:10.1001/jamanetworkopen.2022.29158 (July 2022)

4. Riaz Gillani, Sabrina Y. Camp, Seunghun Han, Jill K. Jones, **Hoyin Chu**, Schuyler O'Brien, Erin L. Young, Lucy Hayes, Gareth Mitchell, Trent Fowler, Alexander Gusev, Junne Kamihara, Katherine A. Janeway, Joshua D. Schiffman, Brian D. Crompton*, Saud H. AlDubayan*, Eliezer M. Van Allen*, *Germline predisposition to pediatric Ewing sarcoma is characterized by inherited pathogenic variants in DNA damage repair genes*, The American Journal of Human Genetics, https://doi.org/10.1016/j.ajhg.2022.04.007 (May 2022)

Submitted Manuscripts:

5. Mohammed Al-Jumaan*, **Hoyin Chu***, Abdullah Al-Sulaiman, Yousef Al Marzooq, Fatmah Almulhim, Chitti Babu Vatte, Arij Alnimer, Afnan Almuhanna, Amein Al-Ali*, Saud H AlDubayan*, *Low-pass Whole Genome Imputation Enables the Detection of Polygenic Breast Cancer Risk in Understudied Populations*, Under Review (October 2022)

6. Mark Leick, Seunghun Han, Kathleen Gallagher, Harrison Silva, Grace Martin, Sarbina Camp, **Hoyin Chu**, Riaz Gillani, Michael Kann, Bryan Choi, Rebecca Larson, Merle Phillips, Tamina Kienka, Stefanie Bailey, Charlotte Graham, Christopher Reilly, Max Jan, Elba Gonzalez, Nora Horick, Justin Budka, Simone Filosto, Rhine Shen, Eliezer Van Allen, Saud AlDubayan, Marcela Maus, *Patients with Deleterious Germline Variants in STXBP2 Develop Toxicity after CAR-T Cell Therapy with Axicabtagene Ciloleucel*, Abstract, Under Review (August 2022)

7. Kiyomi Mashima, Benjamin Lampson, Aditi Gupta, Nicholas Moore, Anna Petrackova, Samantha Shupe, Stacey Fernandes, Amaro Taylor-Weiner, Riaz Gillani, **Hoyin Chu**, Seunghun Han, Sabrina Camp, Eric Kofman, Gad Getz, Catherine Wu, Eliezer Mendel Van Allen, Saud AlDubayan, Jennifer Brown, *Characteristics of rare germline ATM variants in chronic lymphocytic leukemia*, Abstract, Under Review (July 2022)

8. Romualdo Barroso-Sousa, Tianyu Li, Sangeetha Reddy, Leisha A. emens, Saud H. AlDubayan, **Hoyin Chu**, Beth Overmoyer, Edward T. Richardson III, Paulina Lange, Molly K. DiLullo, Juelie Kasparian, Melissa E. Hughes, Victoria Attaya, Ameer Basta, Nancy U. Lin, Nabihah Tayob, Elizabeth A. Mittendorf, Sara M. Tolaney, *NIMBUS: A phase II study of nivolumab plus low dose ipilimumab in hypermutated HER2-negative metastatic breast cancer*, Under Review (July 2022)

9. Abdullah Al-Sulaiman*, **Hoyin Chu***, Mohammed Al-Jumaan, Mohammed Alyahya, Yousef Al Marzooq, Fatmah Almulhim, Chitti Babu Vatte, Arij Alnimer, Afnan Almuhanna, Amein Al-Ali*, Saud H AlDubayan*, *Molecular landscape and clinical actionability of pharmacogenomic variants in unselected breast cancer patients from the indigenous Arab population*, Under Review (June 2022)

10. Sabrina Y. Camp, Stephanie A. Cohen, Kelsey Moriarty, Seyram Doe-Tetteh, Eric R Kofman, Seunghun Han, Riaz Gillani, Brendan Reardon, Jett Crowdis, **Hoyin Chu**, Tina Alano, Darren Feldman, Anne Marie Noronha, Christopher J. Sweeney, David B. Solit, Saud H. AlDubayan, Eliezer M. Van Allen, *Germline loss of function variants in CHEK2 and susceptibility to Ovarian Germ Cell Tumors*, Under Review (June 2022)

11. **The Critical Assessment of Genome Interpretation Consortium**, *CAGI, the Critical Assessment of Genome Interpretation, establishes progress and prospects for computational genetic variant interpretation methods*, arXiv http://arxiv.org/abs/2205.05897, Under Review (May 2022)

Teaching Experience

Technology Teaching Assistant   May - Jun. 2018
MATH 3081: Probability and Statistics

- Tutored students and created programming homework assignments in R.

Honors and Awards

Dean's List, Khoury College of Computer and Information Sciences   *2016-2021*
Civic Digital Fellowship, Coding it Forward   *2020*
Mi10 Award, Chewy   *2019*
Best Beginner's Hack, HackWITus   *2018*
Amelia Peabody Scholarship, Northeastern University   *2017*

Work Experiences

National Institute of Allergy and Infectious Diseases   (Remote) Bethesda, MD
Data Scientists Fellow   *Oct 2020 - Dec 2020*

- Developed web scraping, data storage, data cleaning and visualization pipeline for the data usage of dbGaP
- Maintained project repository with standard open-source practices such as version control, documentation, issue tracking and continuous integration
- Presented package to senior institute leaders and demonstrated its capability to help answer research questions

Chewy   Boston, MA
Software Engineer Co-op   *July 2019 - Dec 2019*

- Developed new features on company product display pages using Vue components and maintained the legacy Stripes framework using JSP fragments, Java and Kotlin
- Implemented Google Analytics tagging across the site and tracked user impressions and clicks on page elements by using Javascript and the observer design pattern
- Analyzed 2019 e-commerce conversion rate by retrieving session data using SQL queries, segmenting sessions based on attributes such as device type and traffic channel, and forecasting conversion rate using linear regression

BlockTEST   Cambridge, MA
Front-end Engineer Co-op   *July 2018 - Nov 2018*

- Designed and implemented 20+ webpages with HTML, CSS, Javascript and dynamically populated Bootstrap templates by making REST calls
- Created tutorials on deploying smart contracts on the Ethereum Ropsten Testnet using Solidity
- Organized healthcare themed hackathon with emphasis on blockchain technology and hosted webinars.

Miscellaneous

Sub-reviewer
ASM-BCB, 2020-2022
ISMB/ECCB, 2021-2022