Andy Jinseok Lee

Email: andyjinseoklee@gmail.com Website: https://andyjslee.com
ORCID ID: 0000-0003-4120-8708

HIGHLIGHT

I am experienced in applying various analytical methods to massive volumes of high-throughput sequencing data of cancer exomes/genomes, epigenomes, and transcriptomes. I have regularly employed the high-performance computing system to run my computational analyses on Next-generation sequencing (NGS) data. I am proficient in python, R, C/C++, and Java programming languages. I have closely collaborated with other labs in a multidisciplinary manner and have co-first authored two publications on novel computational methods in the field of cancer genomics research. My interdisciplinary background with a strong foundation in computer science, graduate education in health informatics, and cancer genomics research experience are a rare mix. I have also significantly contributed to writing research grant proposals and creating presentation materials for my principal investigator. I have lived and studied in South Korea, Indonesia, Singapore, and the United States, and thrive in multicultural settings. I am a bilingual who is fluent in English and Korean.

EDUCATION

SEP 2015 - APR 2017 **University of Michigan**, Ann Arbor, MI, United States
Master of Health Informatics

SEP 2010 - DEC 2014 **University of Michigan**, Ann Arbor, MI, United States
Bachelor of Science in Engineering in Computer Science (minor in Economics)

PUBLICATIONS

Co-first authored (*)

1. Kim, H.*, **Lee, A.J.***, Lee, J., Chun, H., Ju, Y.S., and Hong, D. (2019). FIREVAT: finding reliable variants without artifacts in human cancer samples using etiologically relevant mutational signatures. **Genome Med** 11, 81. PMID: 31847917.
2. Lee, J.*, **Lee, A.J.***, Lee, J.K.*, Park, J., Kwon, Y., Park, S., Chun, H., Ju, Y.S., and Hong, D. (2018). Mutalisk: a web-based somatic MUTation AnaLyIS toolKit for genomic, transcriptional and epigenomic signatures. **Nucleic Acids Res** *46*, W102-W108. PMID: 29790943.

Co-authored

1. Seewald, N.J., Smith, S.N., **Lee, A.J.**, Klasnja, P., and Murphy, S.A. (2019). Practical Considerations for Data Collection and Management in Mobile Health Micro-randomized Trials. **Stat Biosci** *11*, 355-370. PMID: 31462937.
2. Klasnja, P., Smith, S., Seewald, N.J., **Lee, A.**, Hall, K., Luers, B., Hekler, E.B., and Murphy, S.A. (2019). Efficacy of Contextually Tailored Suggestions for Physical Activity: A Micro-randomized Optimization Trial of HeartSteps. **Ann Behav Med** *53*, 573-582. PMID: 30192907.
3. Rabbi, M., Philyaw-Kotov, M., **Lee, J.**, Mansour, A., Dent, L., Wang, X., Cunningham, R., Bonar, E., Nahum-Shani, I., Klasnja, P.*, et al.* (2017). SARA: A Mobile App to Engage Users in Health Data Collection. **Proc ACM Int Conf Ubiquitous Comput *2017***, 781-789. PMID: 29503985.

ABSTRACTS

1. Hong, D., Lee, J., Park, J., Chae, J.S., **Lee, A.J.**, Kim, H., Eho, S., Lee, R., Kim, Y., Hwangbo, Y., Choi, G., Kong, S., Park, J.B., Park, S., Lee, E.S. An introduction of ARCHON (Advanced poRtal of Clinical History and multi-Omics iNformation) database [abstract]. In: AACR-KCA Joint Workshop on Precision Medicine in conjunction with the 23rd KCA Fall Symposium; 2019 Nov 14-15; Seoul, Korea. **AACR-KCA**; 2019. Abstract SP6-1.
2. **Lee, A.J.**, Fu, S., and Vydiswaran, V.G.V. Supervised Learning Approach to Link Prediction in FDA Adverse Event Reporting System (FAERS) Database Network [abstract]. In: AMIA 2018 Annual Symposium; 2018 Nov 3-7; San Francisco, California (CA). **AMIA**; 2018. Podium Abstract S55.

BOOK CHAPTER

1. Smith, S.N., **Lee, A.J.**, Hall, K., Seewald, N.J., Boruvka, A., Murphy, S.A., and Klasnja, P. (2017). Design Lessons from a Micro-Randomized Pilot Study in Mobile Health. In Mobile Health - Sensors, Analytic Methods, and Applications, J.M. Rehg, S.A. Murphy, and S. Kumar, eds. (Springer), pp. 59-82.

RESEARCH EXPERIENCE

JUN 2017 – JUN 2020 **Researcher at National Cancer Center Korea**
Bioinformatics Analysis Team, Goyang, South Korea
(Substitution for national service as a technical research personnel)
Supervised by Drs. Dongwan Hong and Yuhseog Jung

1. Mutalisk Project
   - Developed a web-based mutation analytics toolkit called Mutalisk, which enables comprehensive analysis of somatic DNA mutations with genome regulatory elements (enrichments) and DNA sequence contexts (mutational signature). Published in *Nucleic Acids Research* (2018) as a co-first author.
2. FIREVAT Project
   - Designed, developed, and validated a novel somatic variant refinement method called FIREVAT (R software package), which uses mutational signatures to remove low-quality variants called in human cancer samples. Manuscript accepted by *Genome Medicine* (2019) as a co-first author.
3. Prostate Cancer NGS Project
   - Developed a computational pipeline to analyze evolution of mutational processes (somatic point mutations and copy number alterations) in prostate cancer exomes and genomes.
4. Head and Neck Cancer NGS Project
   - Research on identifying novel biomarkers of prognosis in whole genomes and transcriptomes of head and neck tumors.

MAY 2016 - SEP 2016 **Research Intern at Microsoft Research**
Medical Devices Group, Redmond, WA, United States
Supervised by Drs. Dan Morris, Miah Wander, and Sumit Basu

1. Independently designed a pilot study to investigate the feasibility of monitoring additional physiological signals from the medical device under research and development for cardiovascular disease.
2. Developed the necessary software (python and MATLAB) and hardware (Arduino, custom sensors, and analog circuit boards) experiment apparatus for the pilot/feasibility study and conducted the study on human participants.
3. Created and delivered a software program that interfaces with Microsoft Health API to fetch Microsoft Band data for external research partners.

SEP 2015 - APR 2017 **Graduate Student Research Assistant at University of Michigan**
School of Information, Ann Arbor, MI, United States
Supervised by Professors Predrag Klasnja, Ambuj Tewari, and Susan Murphy

1. HeartSteps Project
   - Research (R01 grant from the National Institutes of Health) on developing an online reinforcement learning system for maintaining physical activity levels of post-cardiac patients entering phase III rehabilitation.
   - Led the software development of HeartSteps; developed cloud backend, Android and iOS applications that together seamlessly collect multiple sources of sensor data, including wearables, to inform treatment policies for the reinforcement learning system.
   - Contributed to designing user experience and behavioral intentions of HeartSteps to increase user engagement for better data collection.
2. SARA Project
   - Led a group of graduate and undergraduate students to develop a research platform for the National Institute on Drug Abuse (NIDA) challenge (Addiction Research: There's an App for that). Placed 2nd in the challenge.

MAY 2014 - JUL 2015 **Research Assistant at University of Michigan**
School of Information, Ann Arbor, MI, United States
Supervised by Professor Predrag Klasnja

1. Designed and developed the prototype version of HeartSteps as an Android application with a cloud backend.
2. Assisted with conducting a pilot study with 40+ human participants to test the hypothesis on whether contextualized messages delivered via a smartphone application can help increase physical activity levels of healthy individuals.

TRAINING & CERTIFICATION

JUL 2019 **NVIDIA Deep Learning Institute**, Seoul, South Korea
- Certificate of competency in *Fundamental of Deep Learning for Natural Language Processing*

JUL 2017 **Asian Institute in Statistical Genetics and Genomics**, Seoul, South Korea
- Certification of course completion in *Cancer Genomics*
- Certification of course completion in *Statistical Genomics – Methods and Analyses for Omics Data*
- Certification of course completion in *Biostatistics using R*

DEC 2009 **British Council and United World College of South East Asia**, Singapore
- Certification of course completion in *Teaching English to Speakers of Other Languages (TESOL)*

AWARD/PRIZE

AUG 2016 2nd place in the 2016 National Institute on Drug Abuse (NIDA)
"Addiction Research: There's an App for that" challenge. Served as the student group leader of the SARA (Substance Abuse Research Assistant) project.

WORK EXPERIENCE

DEC 2016 - DEC 2016 **Software Developer at University of Southern California**
Keck School of Medicine, remote from Ann Arbor, MI, United States
Supervised by Dr. Jimi Huh
- Developed and delivered an Android mobile application and a cloud application for a pilot research study investigating pediatric sun protective behaviors.

AUG 2015 - AUG 2015 **Software Developer at University of Michigan Health System**
Internal Medicine, Ann Arbor, MI, United States
Supervised by Dr. Jennifer Meddings
- Developed and delivered cloud backed Android and iOS mobile application prototypes for a research study that investigates the misuse of catheters in clinical settings.

ENTREPRENEURIAL EXPERIENCE

SEP 2015 - APR 2017 **Co-founder and Lead Developer at StepFor, LLC**
Ann Arbor, MI, United States
- Developed the cloud backend and the Android/iOS mobile applications for a crowdfunding corporate donation marketing platform that empowers users to help charities with their steps.
- Raised grant funds from startup competitions: 2016 optiMize Social Innovation Challenge ($5,000 as a finalist) and 2016 Michigan Business Challenge Social Impact Track ($7,500 as 2nd place winner).
- Organized a fundraising campaign for the American Cancer Society with the Relay For Life of University of Michigan.

JUN 2012 - AUG 2013 **Director of Technology at FriendsLearn, Inc.**
San Jose, CA, United States; Chennai, India; Ann Arbor, MI, United States
Supervised by Bhargav Sri Prakash
- Directed a team of five developers to create and launch Fooya on iOS, an educational mobile game about nutrition and healthy eating habits for kids.
- Organized and executed a successful Kickstarter campaign that raised over $50,000.

EXTRACURRICULAR ACTIVITIES

SEP 2011 - APR 2013 **Vice President at Davis United World College Scholars Society**
University of Michigan, Ann Arbor, MI, United States

- Served as one of the student representatives of the University of Michigan chapter which consisted of students from diverse ethnicities and backgrounds.
- Helped organize and host social events for the incoming freshmen and the chapter members.

JAN 2011 - MAY 2012 **Student Engineer at Michigan Hybrid Racing**
University of Michigan, Ann Arbor, MI, United States

- Contributed to the development of our team's 2012 hybrid race car Hyperion.
- Delivered the team's business presentation at the 2012 SAE Formula Hybrid International Competition to senior engineers from the automotive industry.

RELEVANT SKILLS

Programming Languages and Environments
Python, Jupyter/IPython, R, MATLAB, UNIX/Bash, Java, C, C++, C#, SQL (SQLite, MySQL, PostgreSQL), JavaScript, PHP, CSS/HTML, Swift, XML.

Computer Science, Data Science, and Software Engineering
OOP, algorithm development, operating systems (parallelization and threading), machine learning (scikit-learn), deep learning (TensorFlow), natural language processing (NLTK), high-performance computing system, data visualization (ggplot2 and ggpubr), containerization (Docker), version control (Git), desktop application development (Electron), mobile application development (Android and iOS), web development (Django), cloud application development (AWS and Google App Engine), Arduino.

Bioinformatics and Cancer Genomics
BWA-MEM, SAMtools, GATK, Picard, mpileup, IGV, SRA, Bioconductor, MuTect2, Muse, Varscan2, SomaticSniper, Strelka2, DToxoG, MutSigCV, Oncotator, DELLY2, LUMPY, deconstructSigs, WTSI, MutationalPatterns, Sequenza, CNVkit, FACETS, GISTIC2, TitanCNA, STAR, STAR-Fusion, RSEM, DESeq2, CIRCexplorer2.