

Soo Min Lee

Email: lsm46147570@gmail.com | Telephone: (+82) 10-3606-3204
Homepage: <https://soominll.github.io/soominlee>

EDUCATION

Korea Advanced Institute of Science and Technology (KAIST)
M.S., Major in Bio and Brain Engineering

Daejeon, Korea
Sep 2020 – August 2022

Handong Global University (HGU)
B.S., Double major in *Computer Science* and *Life Science*
Summa cum laude, Early graduation

Pohang, Korea
Feb 2017 – Jun 2020

Allendale Columbia School
Studied abroad for high school in the United States

Rochester, NY, U.S.A.
Jun 2016

RESEARCH EXPERIENCE

Korea Advanced Institute of Science and Technology (KAIST)
Researcher, SBiE laboratory (prof. Kwanghyun Cho)

Daejeon, Korea
July 2020 – Feb 2023

1. ‘Deep learning untangles the resistance mechanism of p53 reactivator in lung cancer cells’
(*Master Thesis & Journal publication & KOR/US patent*)

- **Problem:** Though TP53 mutations in lung cancer make p53 activator treatments essential, drug resistance limits their effectiveness.
- **Contributions:** Developed a novel computational approach using Generative adversarial networks and Graph neural networks to identify drug resistance mechanisms and therapeutic targets. Conducted *in vitro* validation, confirming the therapeutic potential of the identified gene targets.
- **Results:** Identified a key biomarker linked to p53 resistance, leading to a Korean/US patent application and publication in *Cell iScience*.

2. ‘A systems biological study to investigate the fundamental principle and overcoming method of adaptive resistance to targeted anti-cancer therapy’

- **Problem:** Overcoming the limitations of colon cancer treatment requires understanding adaptive resistance to targeted therapies, like MEK inhibitors, and developing strategies to combat it.
- **Contributions:** Performed drug concentration and timepoint screenings, established MEK inhibitor-resistant colon cancer cell lines, analyzed gene expression changes in key pathways, and produced RNA sequencing data.
- **Results:** Discovered significant resistance-related genes, supporting model construction and contributing to securing \$300K in funding from the *National Research Foundation of Korea*.

Handong Global University (HGU)
Researcher, Biodata laboratory (prof. Taejin Ahn)

Pohang, Korea
Jan 2019 – Jun 2020

1. ‘Identification of Lifestyle Factors correlated with Gastroesophageal Reflux Disease’ (*Bachelor Thesis*)

- **Problem:** The increasing global incidence of Gastroesophageal Reflux Disease (GERD) is associated with unhealthy lifestyles, making it crucial to identify contributing lifestyle factors.
- **Contributions:** Collaborated with a physician from Seoul National University Hospital (SNUH) to analyze GERD risk factors using logistic regression, hierarchical clustering, and subgroup analysis.
- **Results:** Identified lifestyle factors contributing to GERD risk, which were presented in my bachelor’s thesis.

2. ‘Diabetic Status Monitoring and Prediction using Electronic Health Record and Clinical Data’

- **Problem:** Continuous monitoring of diabetes is essential, but challenging for patients to manage on their own, making the development of predictive tools important.
- **Contributions:** Developed predictive models for short-term and long-term blood glucose levels using deep neural networks and various regression methods (ridge, lasso, and stepwise multivariate). Analyzed clinical data from SNUH and the GWAS catalog.
- **Results:** Revealed key variables influencing blood glucose prediction, leading to a scholarship award through the Capstone project submission.

Seoul National University (SNU)
Research intern, Bioinformatics & Biostatistics laboratory (prof. Taesung Park)

Seoul, Korea
Jan 2020 – Feb 2020

1. ‘Bio-synergy: Metabolites Data Analysis’

- **Problem:** Metabolites are crucial for type 2 diabetic status prediction, but accurate predictions depend on effective data preprocessing and appropriate model selection.

- **Contributions:** Preprocessed LC/MS-based metabolite data from the Korea Basic Science Institute using Combat and quantile normalization to reduce non-biological variation. Improved a diabetes prediction model by applying elastic net regression.
- **Results:** Enhanced prediction accuracy and gained experience in handling metabolite data.

INDUSTRY EXPERIENCE

Inocras (Headquarter in San Diego, CA)

Computational Genomics center – Pipeline Development team

Daejeon, Korea
March 2023 – Present

- **Integrated and refined The Clinical Knowledgebase (CKB) from Jackson Laboratory into the somatic pipeline,** enhancing data accessibility for actionable findings, precise on-label drug recommendations, and clinical trial matching based on patients' genetic variations. Collaborated with oncologists to develop a **Disease Checklist** for accurate on-label drug recommendations.
- **Improved Structural Variation (SV) detection** in somatic and germline pipelines by optimizing SV calling methods and post-processing, resulting in more accurate variant identification (e.g., detecting missing fusion variants, filtering false positives, and reducing short inversions, which decreased SV calls by up to 78%) and speeding up the process.
- **Contributed to an AI project** by developing an AI-driven deep mutation caller for small variants, enhancing the accuracy and efficiency of mutation detection.

PUBLICATIONS & PATENTS

Publications

- Lee SM, Han Y, Cho KH. Deep learning untangles the resistance mechanism of p53 reactivator in lung cancer cells. *iScience*. 2023 Nov 1;26(12):108377. doi: 10.1016/j.isci.2023.108377. PMID: 38034356; PMCID: PMC10682260.

Patents

- **USES OF THBS1 INHIBITOR FOR OVERCOMING DRUG RESISTANCE OF CANCER**
 - Application: 10-2023-0038078, 2023 March 23 (Republic of Korea), PCT/KR2023/003905, 2023 March 23 (PCT), 18/840,605, 2024 August 22 (USA)

AWARDS & OTHER EXPERIENCES

Awards

- Top honor in Life Science from Handong Global University (1%)

2020
- Academic Excellence Scholarship

2018, 2019, 2020
- Software Mileage Scholarship

2020
- Excellence prize at Software Festival held by Handong University; Convergence Research Contest & SW Problem-Solving Idea Contest.

2019

Other Experiences

- Compassion Mate Translation Volunteer

Sep 2016 – Dec 2020
- Machine Learning with Real World Data Camp, C++ Camp

Jul 2019
- Student Government – Scribe, Director, Treasurer

Feb 2019 – Jun 2019
- Translation Society – Co-Chair

Feb 2018 – Dec 2018
- National Undergraduate Symposium on Biology – Topic: Autoimmune Disease

Jun 2018 – Sep 2018
- World Friends ICT Volunteers Program in Kenya (National Information Society Agency)

Jul 2017 – Aug 2017

SKILLS & OTHERS

Languages: Native fluency in Korean, fluent in English

Programming Languages: Python, R, Java, C, C++

Laboratory Skills: Cell culture, DNA/RNA extraction and purification, recombinant DNA technology, PCR, qRT-PCR, Western blotting, gel electrophoresis, transfection, viral transduction

Coursework: Applied Genomics, Biodata Analysis, Molecular Biology, Machine Learning, Graph Machine Learning and Mining, Data Mining, Genome Bioinformatics

Test Scores: TOEIC (990/990), TOEFL (111/120, My Best Score)

Qualifications: Big Data Analyst (Issued by Ministry of Science and ICT / Statistics Korea)

Interests: Personalized Medicine, Machine Learning, Bioinformatics, Multi-omics, Drug Discovery, Cancer