Kwak Yejin

ng4812@pusan.ac.kr
+82 10-5027-4118
 https://github.com/Kwakyejin
 linkedin.com/in/yejin-kwak-000962316
 0009-0006-0892-9211

Education

2020.03 – 2023.08   B.S., Biomedical Convergence Engineering, Data Science Major
Pusan National University, Yangsan
(GPA: 3.95 / 4.5) *(Magna cum laude, Dean's list)*

2023.09 – 2025.02   M.Sc., Information Convergence Engineering, Biomedical Engineering Major
Pusan National University, Yangsan
(GPA: 4.5 / 4.5)

Research Experience

2025.03 – present   **Associate Researcher**
Computational Omics Laboratory (Advisor: Prof. Dr. Jeongbin Park)

2022.06 – 2025.02   **M.Sc. Candidate (via combined B.Sc. - M.Sc. course)**
Computational Omics Laboratory (Advisor: Prof. Dr. Jeongbin Park)

2022.06-2023.08: Undergraduate Researcher

2023.09-2025.02: Master's student

Web Development Projects
1) HCA:Organoid (Human Cell Atlas Organoid) Data Portal
- Converted organoid sequencing data from the HCA:Organoid into Zarr format.
- Developed a WebGL-powered portal for rapid scatter plot visualization of organoid data.
2) READRetro Web Portal
Developed a website to support visualization and execution of AI-based multi-step retrosynthesis predictions for plant metabolites.

Data Analysis Using SSAM
1) MERSCOPE Alzheimer's Disease Mouse Models Analysis
Identified key differentially expressed genes (DEGs) within the hippocampus (ROI).
2) MERSCOPE Duchenne Muscular Dystrophy (DMD) Patient Data Analysis
Highlighted genes with high Moran's I values were examined to confirm their alignment with lesion areas.

SSAM2 Development
- Extended the SSAM framework to support Visium HD.
- Added GPU-based preprocessing steps(sum normalization, average convolution) to correct striped artifacts in Visium HD.
*SSAM: A cell segmentation-free method for inferring cell types from spatial transcriptomics.

2023.10 – 2023.12   **Visiting Researcher**
Division of Statistical Genomics and Systems Genetics at the German Cancer Research Center (DKFZ), Germany
SSAM2 Development
Enabled SSAM to read and write SpatialData format, improving interoperability with multi-modal spatial omics tools.

HCA:Organoid (Human Cell Atlas Organoid) Data Portal
Enhanced visualization capabilities by incorporating additional interactive components to improve user experience.

2024.07 – 2025.01   **Visiting Researcher**
Center for Genome Engineering at the Institute for Basic Science (IBS), Korea
SCON DB Portal
Developed a portal providing comprehensive details on SCON(Short Conditional intrON: inserted into a specific SCONable exon for precise gene editing).

Skills

Programming Language	Python, R, MATLAB, ECMAScript(JavaScript), C/C++, HTML, Java, C#, Kotlin
Deep learning / ML	Pytorch, Numpy, Pandas, Scikit-learn, Tensorflow, Scipy
Bioinformatics	Single-cell and Spatial data analysis, with both scverse tools (Scanpy, SpatialData, Squidpy etc.) and Seurat
Language	Korean(native), English(TOEFL IBT: 92, TOEIC: 845), German(basic), Japanese(basic)
Others	Linux, Git, Bash

Poster Presentation

2023.01   KSBMB(Korean Society for Biochemistry and Molecular Biology), "HCA|Organoid portal: An easy-to-use web portal provides comprehensive organoid data", Hongcheon-gun, Gangwon-do, Korea

2023.05   KSBMB, "HCA:Organoid data portal - a comprehensive web database for organoids", Busan, Korea

2024.01   SCSOK(Single Cell & Spatial Omics Korea), "Standardizing data format for cell segmentation-free in situ spatial transcriptomics data analysis framework (SSAM)", Yangsan, Gyeongsangnam-do, Korea

2024.09   scverse, "Standardizing data format for cell segmentation-free in situ spatial transcriptomics data analysis framework (SSAM)", Munich, Germany

2025.06   HCA(Human Cell Atlas) General Meeting, "HCA|Organoid data portal: a comprehensive web-based database for organoids", Singapore

Scholarships

2020 – 2025   School Scholarships, Pusan National University

2020   School Leader Scholarships, Pusan National University

2021 – 2022   Mentoring Scholarships, Samsung Welfare Foundation

2022   Student Scholarships, PNU Medical AI Human Resource Training Project

Awards

2023.08   **Busan ICT Convergence Hackathon - 3rd place**
https://github.com/Kwakyejin/daw_app
Create an AI-powered app that recommends travel routes by applying Dijkstra's algorithm

2023.02   **Bio-Health Medical Entrepreneurship Contest – 2nd place**
https://github.com/Kwakyejin/memorygame
Web-based platform for memory training to slow cognitive decline in dementia.

2022.12   **Sports data analysis/utilization contest – 3rd place**
https://github.com/Kwakyejin/Sports-data-analysis-contest
AI model that recommends exercise for disabled individuals by matching with healthy counterparts using propensity scores.

2022.01   **AI Hackathon – 3rd place**
https://github.com/Kwakyejin/gongbang_app
Plant prediction app using EfficientNetV2 to identify native plants of Jeju Island.

2021.12   **Mini Kaggle Contest - Kong counting night – 2nd place**
https://github.com/Kwakyejin/Team_countingkong
Bean counting contest using OpenCV to estimate bean quantities from multi-angle images.

2021.12   **Try anything contest – 2nd place**
Developed EEG/EMG devices with OpenBCI and 3D printing; created EMG-based game; awarded Best Poster & Video.

Publication

- Kim, T., Lee, S., **Kwak, Y.**, Choi, M.-S., Park, J., Hwang, S.J. and Kim, S.-G. (2024), READRetro: natural product biosynthesis predicting with retrieval-augmented dual-view retrosynthesis. New Phytol, 243: 2512-2527. https://doi.org/10.1111/nph.20012

- Eun Young Jeon, **Yejin Kwak**, Hyeji Kang, Se Young Jin, Soojin Park, Ryeo Kim, Dayoung Ko, Jae-Kyung Won, Anna Cho, Inkyung Jung, Chul-Hwan Lee, Jeongbin Park, Hyun-Young Kim, Jong-Hee Chae, Murim Choi, **Inhibiting EZH2 complements steroid effects in Duchenne muscular dystrophy**. Sci. Adv.11,eadr4443(2025). DOI:10.1126/sciadv.adr4443

- **Yejin Kwak**, Taein Kim, Sang-Gyu Kim, Jeongbin Park, READRetro Web: A User-Friendly Platform for Predicting Plant Natural Product Biosynthesis, Molecules and Cells, 2025, 100235, ISSN 1016-8478, https://doi.org/10.1016/j.mocell.2025.100235

- **SCON**, Gene editing technique with database (submitted)

- Single Cell RNA-Seq Analysis Reveals State Heterogeneity Dynamics of Germline Stem Cells in vitro (in preparation)

- **HCA:Organoid data portal**, a comprehensive web database for organoids (in preparation)

- **SSAM2 Development** (in preparation)

Edited 2025.07.08