HEEWON SEO, MSc, PhD

3330 Hospital Drive NW, Calgary, AB T2N 4N1
Heewon.Seo@ucalgary.ca

SUMMARY

Accomplished bioinformatician with over a decade of experience in spatial transcriptomics, pharmacogenomics, and multi-omics integration. Demonstrated leadership in developing robust analytical pipelines, deploying scalable tools, and leading cross-functional collaborations. Authored 20+ peer-reviewed publications and contributed to translational insights that have informed therapeutic strategies in oncology and precision medicine.

AREAS OF EXPERTISE

- Understanding the biomedical data life cycle in an era of continuous data generation
- Development of innovative tools and methods for the analysis of biomedical big data
- Rapid construction and optimization of complex computational pipelines
- Experience working in high-performance computing environments (on-premise and cloud)

EDUCATION

PhD in Medical Science (Biomedical Informatics) **Sep 2012–Dec 2017**
Seoul National University College of Medicine, Seoul, Korea
Thesis: **Methods for Variant- and Gene-based Analysis for Pharmacogenomics Research.**

MSc in Medicine (Biomedical Informatics) **Sep 2010–Aug 2012**
Seoul National University College of Medicine, Seoul, Korea
Thesis: **Loss of Function Gene-set Analysis of Personal Genome using Pathway-disease Similarity.**

BS in Computer Science **Mar 2004–Feb 2010**
Sejong University, Seoul, Korea
GPA: 4.37/4.5, ***Summa Cum Laude,*** achieved the early graduation of excellent students within 3 years

RESEARCH/WORK EXPERIENCE

Lead Bioinformatician **Jan 2024–*present***
Applied Spatial Omics Centre, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

- Developed and standardized quality control (QC) and analytical pipelines for spatial transcriptomic profiling using GeoMx DSP and CosMx SMI (NanoString), as well as Visium HD and Xenium In Situ (10x Genomics) platforms
- Designed and implemented an automated reporting tool to generate diagnostic visualizations and perform rigorous QC and in-depth spatial transcriptomic analyses
- Encapsulated spatial transcriptomics data processing workflows into Dockerized applications to promote research reproducibility and computational replicability
- Created and maintained a public-facing web portal for dissemination of analysis results and resources: https://ASOC.ucalgary.ca/

RESEARCH/WORK EXPERIENCE (CONT'D)

- Organized and led the Hands-on Bioinformatics Workshop for principal investigators and trainees, demonstrating spatial omics workflows and Docker applications using public datasets: https://ASOC.ucalgary.ca/HBW
- Developed interactive Shiny applications enabling researchers to visualize and explore gene expression patterns within spatial transcriptomic datasets: https://shinyapps.ucalgary.ca/
- Established an image analysis pipeline utilizing H&E and DAPI-stained images to accurately segment nuclei and enhance spatial data quality
- Actively contributed to the research community by sharing code and analysis pipelines via GitHub.
- Completed specialized training in Adobe Illustrator and Photoshop through Continuing Education at the University of Calgary to enhance scientific visualization capabilities

Bioinformatician **Nov 2020–Dec 2023**
Arnie Charbonneau Cancer Institute, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

- Performed spatial transcriptomic analyses to elucidate the complex interactions between tumor cells and the tumour microenvironment (TME), revealing mechanisms underlying tumor progression, invasion, and resistance to therapy in glioblastoma
- Utilized immunofluorescence-labeled tissue microarrays in conjunction with nearest-neighbor spatial analysis to investigate tumour cell adaptation within the local TME, identifying spatial micro-neighborhoods associated with patient prognosis
- Developed a novel correlation-based integrative method for multi-omics analysis, enabling the identification of biologically meaningful ranks from non-negative matrix factorization outputs
- Identified potential chimeric antigen receptor (CAR) T-cell targets by mining compendium-scale RNA-seq datasets for genes encoding tumour cell surface proteins in glioblastoma and sarcoma
- Curated and prioritized CAR-T targets by systematically mining cell surface protein-coding genes across multiple data repositories and literature sources, incorporating levels of supporting evidence
- Analyzed CRISPR screening datasets to identify genes significantly enriched in T-cell-exposed tumour cells compared to non-exposed controls, elucidating potential immune evasion mechanisms
- Designed and implemented a pipeline for generating consensus BAM files from shallow-depth and FFPE-derived exome data using unique molecular identifiers, enhancing data recovery and variant calling accuracy
- Developed a spatial image analysis workflow for neighborhood mapping that integrates with spatial transcriptomic data to quantify tumour–TME architecture and its biological implications

Visiting Researcher **Jan 2018–Oct 2020**
Ontario Institute for Cancer Research (OICR), Toronto, Canada

- Performed integrative analyses of large-scale genomic and clinical datasets using robust statistical methodologies to identify biomarkers associated with gemcitabine resistance in pancreatic cancer cohorts
- Identified and prioritized gemcitabine resistance biomarkers to inform the selection of sensitizing therapeutic agents, e.g., targeted inhibitors, with findings validated through survival analysis
- Conducted meta-analyses of survival outcomes across multiple independent pancreatic cancer cohorts to validate biomarker relevance and support their potential clinical utility
- Contributed to the Molecular Tumor Board of the COMPASS (Changes and Characteristics of Genes in Patients With Pancreatic Cancer for Better Treatment Selection) clinical trial, providing analytical insights to inform personalized treatment strategies

RESEARCH/WORK EXPERIENCE (CONT'D)

Postdoctoral Researcher **Jan 2018–Oct 2020**

Princess Margaret Cancer Centre, University Health Network, Toronto, Canada

- Developed SYNERGxDB, the largest integrated database of high-throughput drug combination studies, incorporating molecular profiles of corresponding preclinical models to facilitate the discovery of synergistic drug combinations and predictive biomarkers across various cancer types and cell line panels
- Identified two synergistic drug combinations and four expression-based biomarkers with potential predictive value for drug response using integrative analyses within SYNERGxDB
- Established a translational framework for precision oncology by integrating multimodal pharmacogenomic datasets from preclinical models with clinical patient cohort data to support biomarker-driven therapy selection
- Developed a semi-automated analysis pipeline on Microsoft Azure utilizing Jupyter Notebooks for efficient exploration and interrogation of large-scale pharmacogenomics datasets
- Created a web-based tool for semi-automated reporting of significant biomarkers, enabling multi-resolution data visualization and streamlined investigation of molecular determinants of drug response
- Performed meta-analysis across pancreatic cancer cell line datasets to identify gene expression biomarkers predictive of gemcitabine response, with the goal of informing more effective treatment strategies
- Conducted integrative analyses of drug response and molecular data from various preclinical models, including patient-derived cell lines and organoids, under single-agent and combinatorial drug testing conditions
- Discovered two candidate biomarkers associated with gemcitabine resistance in pancreatic cancer, with potential therapeutic implications for combination strategies involving targeted inhibitors

Doctoral Researcher **Sep 2012–Dec 2017**

Seoul National University College of Medicine, Seoul, Korea

- Identified pharmacogenes and genetic variants associated with adverse drug reactions, including mercaptopurine-induced neutropenia and busulfan-induced hepatotoxicity in pediatric cancer patients, as well as ritodrine-induced pulmonary edema in pregnant women
- Discovered prognostic genetic markers in donor exomes predictive of survival outcomes in patients undergoing allogeneic hematopoietic stem cell transplantation
- Independently designed and implemented a comprehensive pharmacogenomics analysis platform to interpret patient genomes and exomes, facilitating the identification of drug-response-associated genes and variants linked to adverse effects
- Developed a novel computational framework for gene-level aggregation of heterogeneous variant effects, improving the functional interpretation of high-throughput sequencing data
- Conducted integrative analyses of whole-genome, whole-exome, targeted panel, transcriptome (RNA-seq), small RNA, and microarray datasets across multiple sequencing platforms, including Ion Proton, Complete Genomics, and Illumina in conjunction with detailed clinical metadata
- Performed in-depth inspection of variant calls and read alignment artifacts to refine variant calling algorithms and enhance analytical accuracy

RESEARCH/WORK EXPERIENCE (CONT’D)

- Curated and localized large-scale public omics datasets, including The Cancer Genome Atlas (TCGA), 1000 Genomes Project (1KGP), Alzheimer’s Disease Sequencing Project (ADSP), and Simons Foundation Autism Research Initiative (SFARI), for integrative and comparative analyses
- Developed a targeted diagnostic sequencing panel for pharmacogenetic profiling in childhood rare cancers, such as acute myeloid leukemia, to support precision medicine initiatives
- Contributed significantly to the grant proposal titled “Precision Medicine and Clinical Evaluation Technologies for Childhood Rare Cancers,” which received approval from the Korea Food & Drug Administration in 2016
- Served concurrently as the server administrator, overseeing the configuration and maintenance of high-performance computing infrastructure, including 472 CPU cores, 4.25 TB RAM, and 756 TB of storage using Network, Lustre, and Fraunhofer file systems

PUBLICATIONS

† : (co-) First author, *: (co-) Corresponding author

1. **Evaluating Gene Representation in Spatial Transcriptomics across Pre-Designed Panels.** Heewon Seo*, Roman Krawetz *bioRxiv 2025*

2. **Impaired Parvalbumin Interneurons in the Retrosplenial Cortex as the Cause of Sex-dependent Vulnerability in Alzheimer’s Disease.** Dylan J. Terstege†, Yi Ren, Bo Young Ahn, Heewon Seo, Alzheimer’s Disease Neuroimaging Initiative, Liisa A. M. Galea, Derya Sargin, Jonathan R. Epp* *Sci Adv 2025;11(18):eadt8976*

3. **CD73 Inhibits cGAS-STING and Cooperates with CD39 to Promote Pancreatic Cancer.** Célia Jacoberger-Foissac†, Isabelle Cousineau†, Yacine Bareche†, David Allard, Pavel Chrobak, Bertrand Allard, Sandra Pommey, Nouredin Messaoudi, Geneviève Soucy, Secil Koseoglu, Ricard Masia, Andrew C. Lake, Heewon Seo, Christopher B. Eeles, Neha Rohatgi, Simon C. Robson, Simon Turcotte, Benjamin Haibe-Kains, John Stagg* *Cancer Immunol Res 2023;11(1):56-71*

4. **Orchestrating and Sharing Large Multimodal Data for Transparent and Reproducible Research.** Anthony Mammoliti†, Petr Smirnov, Minoru Nakano, Zhaleh Safikhani, Christopher Eeles, Heewon Seo, Sisira Kadambat Nair, Ian Smith, Chantal Ho, Gangesh Beri, Marc Hafner, Benjamin Haibe-Kains* *Nat Commun 2021;12(1):5797*

5. **CaReAl: Capturing Read Alignments in a BAM file Rapidly and Conveniently.** Yoomi Park†, Heewon Seo†, Kyunghun Yoo, and Ju Han Kim* *J Big Data 2021;8:23*

6. **Identifying Genetic Variants Associated with Ritodrine-induced Pulmonary Edema.** Seung Mi Lee†, Yoomi Park†, Young Ju Kim, Han-Sung Hwang, Heewon Seo, Byung-Joo Min, Kye Hwa Lee, So Yeon Kim, Young Mi Jung, Suehyun Lee, Chan-Wook Park, Ju Han Kim*, and Joong Shin Park* *PLoS One 2020;15(11):e0241215*

7. **Homozygote CRIM1 Variant is Associated with Thiopurine-induced Neutropenia in Leukemic Patients with both Wildtype NUDT15 and TPMT.** Yoomi Park†, Hyery Kim†, Heewon Seo, Jung Yoon Choi, Youngeun Ma, Sunmin Yun, Byung-Joo Min, Myung-Eui Seo, Keon Hee Yoo, Hyoung Jin Kang, Ho Joon Im, and Ju Han Kim* *J TRANSL MED* 2020;18(1):265

8. **Gene-wise Variant Burden and Genomic Characterization of Nearly Every Gene.** Yoomi Park†, Heewon Seo, Brian Ryu, and Ju Han Kim* *Pharmacogenomics* 2020;21(12):827-840

PUBLICATIONS (CONT'D)

9. **SYNERGxDB: an Integrative Pharmacogenomic Portal to Identify Synergistic Drug Combinations for Precision Oncology.** Heewon Seo†, Denis Tkachuk, Chantal Ho, Anthony Mammoliti, Aria Rezaie, Seyed Ali Madani Tonekaboni, and Benjamin Haibe-Kains* *Nucleic Acids Res* 2020;48(W1):W494-W501

10. **ToxicoDB: an Integrated Database to Mine and Visualize Large-scale Toxicogenomic Datasets.** Sisira Kadambat Nair†, Christopher Eeles, Chantal Ho, Gangesh Beri, Esther Yoo, Denis Tkachuk, Amy Tang, Parwaiz Nijrabi, Petr Smirnov, Heewon Seo, Danyel Jennen, and Benjamin Haibe-Kains* *Nucleic Acids Res* 2020;48(W1):W455-W462

11. **Discovery of Donor Genotype Associated with Long-term Survival of Patients with Hematopoietic Stem Cell Transplantation in Refractory Acute Myeloid Leukemia.** Chan-Young Ock†, Heewon Seo†, Dae-Yoon Kim, Byung Joo Min, Yoomi Park, Hyun Sub Cheong, Eun-Young Song, Inho Kim, Sung-Soo Yoon, Ju Han Kim*, and Youngill Koh* *Leuk Lymphoma* 2018;60(7):1775-1781

12. **Deleterious Genetic Variants in Ciliopathy Genes Increase Risk of Ritodrine-induced Cardiac and Pulmonary Side Effects.** Heewon Seo†, Eun Jin Kwon†, Young-Ah You, Yoomi Park, Byung Joo Min, Kyunghun Yoo, Han Sung Hwang, Ju Han Kim*, and Young Ju Kim* *BMC Med Genomics* 2018;11(1):4

13. **APEX1 Polymorphism and Mercaptopurine-related Early Onset Neutropenia in Pediatric Acute Lymphoblastic Leukemia.** Hyery Kim†, Heewon Seo†, Yoomi Park, Byung Joo Min, Myung Eui Seo, Kyung Duk Park, Hee Young Shin, Ju Han Kim*, and Hyoung Jin Kang* *Cancer Res Treat* 2018;50(3):823-834

14. **Idiopathic Hypereosinophilia Is Clonal Disorder? Clonality Identified by Targeted Sequencing.** Jee-Soo Lee†, Heewon Seo, Kyongok Im, Si Nae Park, Sung-Min Kim, Jung-Ah Kim, Seon Young Kim, Joon-hee Lee, Sunghoon Kwon, Miyoung Kim, Insong Koh, Seungwoo Hwang, Heung-Woo Park, Ju Han Kim, and Dong Soon Lee* *PLoS One* 2017;12(10):e0185602

15. **Evaluation of Exome Variants using the Ion Proton Platform to Sequence Error-Prone Regions.** Heewon Seo†, Yoomi Park†, Byung Joo Min, Myung Eui Seo, and Ju Han Kim* *PLoS One* 2017;12(7):e0181304

16. **Posttranslational control of T-cell development by the ESCRT protein CHMP5.** Stanley Adoro†, Kwang H Park, Sarah E Bettigole, Raphael Lis, Hee Rae Shin, Heewon Seo, Ju Han Kim, Klaus-Peter Knobeloch, Jae-Hyuck Shim*, Laurie H Glimcher* *Nat Immunol* 2017;18(7):780-790

17. **Markers of Disease and Steroid Responsiveness in Paediatric Idiopathic Nephrotic Syndrome: Whole-transcriptome Sequencing of Peripheral Blood Mononuclear Cells.** Hee Gyung Kang†, Heewon Seo†, Jae Hyun Lim, Jong Il Kim, Kyoung Hee Han, Hey Won Park, Ja Wook Koo, Kee Hyuck Kim, Ju Han Kim*, Hae Il Cheong, and Il-Soo Ha* *J Int Med Res* 2017;45(3):948-963

18. **Gastrointestinal Tuberculosis is not Associated with Proton Pump Inhibitors: A Retrospective Cohort Study.** Kyoung Sup Hong†, Seung Joo Kang, Jong Kyoung Choi, Ju Han Kim, Heewon Seo, Suehyun Lee, Jae-Woo Jung, Hye-Ryun Kang, Sang-Heon Cho, and Joo Sung Kim* *World J Gastroenterol* 2013;19(2):258-264

19. **Development of Korean Rare Disease Knowledge Base.** Heewon Seo†, Dokyoon Kim, Jong-Hee Chae, Hee Gyung Kang, Buyng Chan Lim, Hae Il Cheong, and Ju Han Kim* *Healthc Inform Res* 2012;18(04):272-278

SOFTWARE

- **Pancreatic Ductal Adenocarcinoma Toolkit** (PDATK, https://doi.org/doi:10.18129/B9.bioc.PDATK): A patient stratification tool that integrates molecular profiling and survival meta-analyses to predict patient outcomes in pancreatic cancer.

- **SYNERGxDB** (https://SYNERGxDB.ca/): A comprehensive web-based resource integrating the largest collection of drug combination screening data with molecular profiles to enable the discovery of synergistic therapies and predictive biomarkers.

- **CaReAl** [kæriːəl] (Capturing Read Alignments, https://github.com/lootpiz/CaReAl): A high-performance tool for visualizing read alignment patterns and extracting read-level data to assess variant calls and uncover technical biases in sequencing analyses.

- **VVA** (Variant Visualization and Annotation, https://github.com/lootpiz/VVA): A gene- and variant-centric visualization tool optimized for exome sequencing data to intuitively display variant distributions within genes for rapid interpretation.

- **KRDK** (Korean Rare Disease Knowledge base, http://www.snubi.org/software/raredisease/): A comprehensive web-based research platform integrating clinical, genetic, and biobanking resources to advance rare disease research and patient care.

TECHNICAL SKILLS

- High-level languages: C/C++
- Scripting languages: R, Python, HTML, PHP, and javascript
- Structured query languages (SQL): MySQL, MariaDB, and NoSQL (MongoDB)
- Operating system: Good working knowledge of Linux and MacOS
- Adobe products: Illustrator, Photoshop, and Premiere Pro