HEEWON SEO, MSc, PhD

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

SUMMARY

Senior bioinformatics scientist with more than a decade of experience developing reproducible analytical workflows for spatial transcriptomics, pharmacogenomics, metagenomics, and multi-omics integration. Demonstrated expertise in designing scalable computational pipelines, deploying software tools, leading cross-functional collaborations, and translating complex biomedical datasets into actionable biological and clinical insights. Author of 20+ peer-reviewed publications and developer of multiple research platforms supporting oncology, precision medicine, spatial omics, and microbiome research.

AREAS OF EXPERTISE

- Multi-omics integration of transcriptomic, genomic, pharmacogenomic, CRISPR screening, image-derived, and clinical datasets
- Development, optimization, and deployment of reproducible bioinformatics pipelines
- Long-read metagenomic sequencing analysis using Oxford Nanopore data, including POD5 basecalling and downstream taxonomic profiling
- Spatial transcriptomics analysis across GeoMx DSP, CosMx SMI, Visium HD, and Xenium In Situ platforms
- High-performance computing and cloud-based analytical environments, including on-premise HPC systems, Amazon Web Services, and Microsoft Azure
- Containerized workflow development using Docker to support reproducibility, portability, and computational replicability
- Automated reporting, quality control frameworks, and diagnostic visualization for large-scale biomedical data
- Statistical modeling, survival analysis, meta-analysis, non-negative matrix factorization, and biomarker discovery
- Translational oncology, precision medicine, pharmacogenomics, tumour microenvironment analysis, and biomarker-driven therapeutic prioritization

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***
Graduated early through an accelerated excellent-student track within three years

RESEARCH/WORK EXPERIENCE

Senior Bioinformatics Scientist **Sep 2025–*present***

Snyder Institute for Chronic Diseases, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

- Developed and maintained Microbial Clinical Atlas (MCA), an open-source platform and curated knowledge resource for structured Taxon Passports that summarize clinically relevant biology, ecology, and evidence-linked associations of human-associated microorganisms
- Performed Oxford Nanopore long-read metagenomic sequencing analyses to characterize microbial community composition across humanized gnotobiotic mouse models, elucidating donor-recipient engraftment dynamics and inter-individual variability
- Developed and maintained BASEN (Base-level Abundance estimation with Species-assigned Evidence using Nanopore), an open-source bioinformatics workflow/toolkit supporting reproducible processing and analysis of sequencing data within microbiome and metagenomic research contexts
- Established an end-to-end metagenomic analysis pipeline spanning POD5 basecalling, host-read decontamination, quality filtering, and taxonomic classification to generate relative-abundance profiles.

Lead Bioinformatician **Jan 2024–Aug 2025**

Applied Spatial Omics Centre, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

- Developed and standardized quality-control and analytical pipelines for spatial transcriptomic profiling across GeoMx DSP and CosMx SMI platforms from NanoString, as well as Visium HD and Xenium In Situ platforms from 10x Genomics
- Developed an integrative analysis framework, Spatial Omics Toolkit, to leverage spatial transcriptomics data for cross-platform comparison and multi-modal biological interpretation
- Designed and implemented an automated reporting tool to generate diagnostic visualizations and perform rigorous quality control and in-depth spatial transcriptomic analyses
- Encapsulated spatial transcriptomics data-processing workflows into Dockerized applications to promote reproducibility, portability, and computational replicability
- 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
- Established an image-analysis pipeline using H&E- and DAPI-stained images to support accurate nuclei segmentation and improve spatial data quality
- Completed specialized training in Adobe Illustrator and Photoshop through Continuing Education at the University of Calgary to strengthen scientific visualization capabilities

Bioinformatician **Nov 2020–Dec 2023**

Arnie Charbonneau Cancer Institute, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

- Led the bioinformatic identification and prioritization of GPNMB as a highly expressed tumour cell-surface target for CAR-T therapy, integrating transcriptomic and surfaceome evidence to support preclinical and translational development of GPNMB-directed CAR-T therapy for fusion-driven solid tumours
- Curated and prioritized CAR-T targets through systematic mining of cell-surface protein-coding genes across multiple data repositories and literature sources, incorporating levels of supporting evidence

RESEARCH/WORK EXPERIENCE (CONT'D)

- Performed spatial transcriptomic analyses to elucidate interactions between tumour cells and the tumour microenvironment, revealing mechanisms underlying tumour progression, invasion, and resistance to therapy in glioblastoma
- Used immunofluorescence-labelled tissue microarrays together with nearest-neighbour spatial analysis to investigate tumour-cell adaptation within the local tumour microenvironment
- Developed a novel correlation-based integrative method for multi-omics analysis, enabling identification of biologically meaningful ranks from non-negative matrix factorization outputs
- Identified potential chimeric antigen receptor T-cell targets by mining compendium-scale RNA-seq datasets for genes encoding tumour cell-surface proteins in glioblastoma and sarcoma
- Analyzed CRISPR screening datasets to identify genes significantly enriched in T-cell-exposed tumour cells compared with non-exposed controls, supporting investigation of potential immune-evasion mechanisms

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, including targeted inhibitors
- Conducted pharmacogenomic profiling of patient-derived pancreatic cancer organoids to characterize drug-response patterns and support prediction of effective combination therapies, including FOLFIRINOX-based treatment strategies
- Performed meta-analysis across pancreatic cancer cell-line datasets to identify gene-expression biomarkers predictive of gemcitabine response
- Validated biomarker findings through survival analysis across pancreatic cancer patient cohorts
- Conducted meta-analyses of survival outcomes across multiple independent pancreatic cancer cohorts to validate biomarker relevance and support potential clinical utility
- Contributed to the Molecular Tumor Board of the COMPASS clinical trial, providing analytical insights to inform personalized treatment strategies

Postdoctoral Researcher **Jan 2018–Oct 2020**

Princess Margaret Cancer Centre, University Health Network (UHN), Toronto, Canada

- Developed SYNERGxDB, a large integrated database of high-throughput drug-combination studies incorporating molecular profiles of corresponding preclinical models to facilitate discovery of synergistic drug combinations and predictive biomarkers across 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 using Jupyter Notebooks for efficient exploration and interrogation of large-scale pharmacogenomics datasets

RESEARCH/WORK EXPERIENCE (CONT'D)

- Conducted integrative analyses of drug-response and molecular data from diverse preclinical models, including patient-derived cell lines and organoids, under single-agent and combinatorial drug-testing conditions
- 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

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 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 functional interpretation of high-throughput sequencing data
- Conducted integrative analyses of whole-genome, whole-exome, targeted panel, transcriptome, 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
- Curated and localized large-scale public omics datasets, including The Cancer Genome Atlas, 1000 Genomes Project, Alzheimer's Disease Sequencing Project, and Simons Foundation Autism Research Initiative, for integrative and comparative analyses
- Developed a targeted diagnostic sequencing panel for pharmacogenetic profiling in childhood rare cancers, including 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 was approved by the Korea Food & Drug Administration in 2016
- Served concurrently as server administrator, overseeing configuration and maintenance of high-performance computing infrastructure comprising 472 CPU cores, 4.25 TB RAM, and 756 TB storage using Network, Lustre, and Fraunhofer file systems

PUBLICATIONS

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

1. **ON-Time enables rapid microbiome sequencing and analysis for precision medicine.** Colin MacKenzie†, Heewon Seo, Jared Schlechte, Aydin Herik, Ish Bains, Ian-Ling Yu, Kathy D McCoy, Christina S Thornton, Braedon McDonald* *Cell Rep Methods* 2026

2. **Reproducible transcriptional modules define glioblastoma ecosystems across independent cohorts.** Heewon Seo *bioRxiv* 2026

3. **Dual platform spatial transcriptomics reveals parvalbumin interneuron subtype vulnerability in mouse models of Alzheimer's disease.** Heewon Seo†, Dylan J. Terstege, Shiying Liu, Kimberly-Ann Ruth Goring, Bo Young Ahn, Jonathan R. Epp* *Nat Commun* 2026;17(1):6668

4. **Spatial transcriptomics identifies IL-32 as a lipid droplet-associated cytokine linked to tubular injury in human diabetic kidney disease**. Kieran Meadows†, Hyunjae Chung, Son Vo, Aysa Imanzadeh, Heewon Seo, Sisay Getie Belay, Asha Swamy, Wulin Teo, Kevin Chapman, Graciela Andonegui, Hallgrimur Benediktsson, Peter K. Stys, Thang Pham, Daniel A. Muruve, Justin Chun* *Inflamm Res* 2026;75(1):33

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

6. **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

7. **Spatiotemporal Modeling Reveals High-resolution Invasion States in Glioblastoma.** Varsha Thoppey Manoharan†, Aly Abdelkareem, Gurveer Gill, Samuel Brown, Aaron Gillmor, Courtney Hall, Heewon Seo, Kiran Narta, Sean Grewal, Ngoc Ha Dang, Bo Young Ahn, Kata Osz, Xueqing Lun, Laura Mah, Franz Zemp, Douglas Mahoney, Donna L Senger*, Jennifer A Chan*, A Sorana Morrissy* *Genome Biol* 2024;25(1):264

8. **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

9. **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

10. **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

11. **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

PUBLICATIONS (CONT'D)

12. **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

13. **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

14. **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

15. **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

16. **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

17. **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

18. **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

19. **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

20. **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

21. **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

22. **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

PUBLICATIONS (CONT'D)

23. **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

24. **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

- **Microbial Clinical Atlas** (https://MCA.thebiohub.ca/): A curated knowledge base of Taxon Passports: structured records that summarize clinically relevant biology, ecology, and evidence-linked associations of human-associated microorganisms.

- **Spatial Omics Toolkit 2** (https://github.com/Snyder-Institute/sotk2): An R package that provides a comprehensive suite of functions for identifying biologically meaningful modules from spatial transcriptomics data.

- **Pancreatic Ductal Adenocarcinoma Toolkit** (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 drug-combination screening data with molecular profiles to enable 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.