HEEWON SEO, MS, PhD

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

SUMMARY

Skilled bioinformatician in the field of genomics with experience analyzing thousands of whole-genomes, exomes, and (spatial) transcriptomes to discover molecular biomarkers for precision therapeutics and novel synergistic drug combinations for determining best treatment strategies via advanced statistical analysis and predictive modelling methods.

AREAS OF EXPERTISE

- Understanding the data life cycle in the age of constant data generation
- Development of new tools and advanced statistical models for biomedical big data analysis
- Rapid construction of complex computational pipelines and their optimization
- Experience in working in high performance computing environments (on-premise and cloud)

EDUCATION

PhD in Medical Science (Biomedical Informatics) *Sep 2012 - Feb 2018*
Seoul National University College of Medicine, Seoul, Korea
Thesis: **Methods for Variant- and Gene-based Analysis for Pharmacogenomics Research.**

MS 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 *Nov 2020 - present*
Arnie Charbonneau Cancer Institute, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

- Analyzed spatial transcriptomics to characterize interplays between tumour cells and tumour microenvironment (TME) which promote tumour progression, invasion, and resistance to therapy
- Leveraged immunofluorescence labelled tissue microarrays to understand tumour cell adaptation to the local TME using nearest neighbour analysis and identified micro-neighbourhoods that affect patients' prognosis
- Developed a correlation-based method, Solution Neighbour Space (SNS), to integrate multiomics datasets and SNS can be used to find the best rank(s) from non-negative matrix factorization driven by biological features
- Discovered CAR (chimeric antigen receptor)-T cell targets that are highly expressed in glioblastoma or sarcoma tumour cell surface using compendium RNA-seq datasets

RESEARCH/WORK EXPERIENCE (CONT'D)

- Mined and enumerated cell surface proteins (genes) from multiple data/knowledge-base and literature along with the level of evidence to prioritize CAR-T targets
- Identified genes that are highly enriched to the T-cell exposure compared to the non-exposure group in CRISPR screening
- Devised a pipeline to generate a consensus BAM using UMI (unique molecular identifiers) to rescue reads in shallow-depth/FFPE exome
- Developed an image analysis workflow for spatial neighborhood analysis that can be coupled with spatial transcriptomics to assess the correspondence between tumor-TME organization

Postdoctoral Fellow *Jan 2018 - Oct 2020*

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

- Developed the largest integrated database in high-throughput drug combination studies, SYNERGxDB (http://SYNERGxDB.ca/), and integrated high-dimensional molecular profiles of the corresponding preclinical models
- Discovered two synergistic drug combinations and four meaningful expression-based biomarkers via statistical analysis in personalized medicine
- Created a semi-automatic analysis tool on Microsoft Azure using Jupyter Notebook for exploring large pharmacogenomics datasets
- Built a translational framework for precision cancer medicine using a compendium of preclinical multimodal pharmacogenomics datasets and large clinical patients cohorts
- Devised an online tool to report significant biomarkers using a semi-automation system for better data representation and easier data investigation with multiple levels of granularity
- Performed meta-analysis of pancreatic cancer cell lines to discover expression-based biomarkers predictive of drug responses in order to improve current treatment strategies
- Analyzed multiple preclinical models including patient-derived cell lines and organoids in single or pairwise drug testing and their molecular profiles (feature space exceeding 50K)
- Identified two promising drug resistance biomarkers that may help patients overcome drug resistance in combination with inhibitors in pancreatic cancer

Visiting Researcher *Jan 2018 - Oct 2020*

Ontario Institute for Cancer Research (OICR), Toronto, ON, Canada

- Analyzed large scale genomic and clinical data with appropriate statistical methods in order to identify gemcitabine resistance biomarkers in pancreatic cancer cohorts
- Discovered gemcitabine resistance biomarkers to select drugs (i.e. inhibitors) that sensitize gemcitabine resistance in pancreatic cancer and validated in patient-derived organoids
- Participated in the molecular tumour board of the COMPASS (Changes and Characteristics of Genes in Patients With Pancreatic Cancer for Better Treatment Selection) trial

Doctoral Researcher *Sep 2012 - Dec 2017*

Seoul National University College of Medicine, Seoul, Korea

- Successfully found pharmaco-genes and -variants that have an association with mercaptopurine-induced neutropenia or busulfan-induced hepatotoxicity in paediatric cancer patients and ritodrine-induced pulmonary oedema in pregnant women

RESEARCH/WORK EXPERIENCE (CONT'D)

- Identified favourable prognostic markers of survival in allogeneic hematopoietic stem cell transplantation patients in donor exomes
- Solely developed a pharmacogenomics analysis platform for the analysis and interpretation of patient genomes and exomes in order to identify drug side effect associated genes and variants
- Built a computational gene-level approach to aggregate the impact of heterogeneous variants in sequencing data
- Analyzed whole-genome, whole-exome, targeted panels, whole-transcriptome, small RNAs, and microarray data with clinical information across platforms (i.e., Ion Proton, Complete Genomics, and Illumina)
- Developed a diagnostic sequencing panel for pharmacotyping in childhood rare cancers (i.e. *acute myeloid leukemia*)
- Localized big open public omics datasets such TCGA (The Cancer Genome Atlas), 1KGP (1000 Genomes Project), ADSP (Alzheimer's Disease Sequencing Project), and SFARI (Simons Foundation Autism Research Initiative)
- Significantly involved in writing the grant proposal entitled, "Precision Medicine and Clinical Evaluation Technologies for Childhood Rare Cancers", and the application has been approved by the Korea Food & Drug Administration in 2016
- Doubled as a Server Administrator who was responsible for setting up and maintaining large-scale servers (472 cores, 4.25 TB RAM) and storage (756 TB) in Network/Lustre/Fraunhofer File Systems

PUBLICATIONS

- **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 2022;CIR-22-0260*

- **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*

- **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*

- **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*

- **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

- **Gene-wise Variant Burden and Genomic Characterization of Nearly Every Gene.** Yoomi Park†, Heewon Seo, Brian Ryu, and Ju Han Kim* *Pharmacogenomics* 2020 (*in press*)

PUBLICATIONS (CONT'D)

- **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

- **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

- **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

- **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

- **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

- **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

- **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

- **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

- **Disease Markers of Pediatric Idiopathic Nephrotic Syndrome and Markers of Steroid-responsiveness: Whole-transcriptome Sequencing of Peripheral 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

- **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

- **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

- ***cPCDH*** (https://MorrissyLab.ucalgary.ca/public/cPCDH/) is a ShinyApp that incorporates the signature of the clustered protocadherin (cPCDH) genes as cell-of-origin lineage markers in adult and childhood brain tumors.

- ***SYNERGxDB*** (https://SYNERGxDB.ca/) is a web-based database that provides the largest integrated database in drug combination screenings and molecular profiles to discover effective combinations and predictive biomarkers.

- ***CaReAl*** [kæriːəl] (Capturing Read Alignments, https://github.com/lootpiz/CaReAl) is a high-performance alignment capturing tool for visualizing the read-alignment status of nucleotide sequences and obtaining read-level data for evaluating variant calls and detecting technical biases.

- ***VVA*** (Variant Visualization and Annotation, https://github.com/lootpiz/VVA) is a gene- and variant-centred visualization tool to be used for exome sequencing data analysis and is optimized for displaying the overall distribution of variants in a gene at a glance.

- ***KRDK*** (Korean Rare Disease Knowledge base, http://www.snubi.org/software/raredisease/) is a web-based, research oriented data repository that provides comprehensive information for rare disease research: disease review, clinics, directory, mutation database, patient registry, and biobank.

PROGRAMMING LANGUAGE SKILLS

- High-level languages: C/C++
- Scripting languages: R/ShinyApps, Python, HTML, PHP, and javascript (NodeJS)
- Structured query languages (SQL): MySQL, MariaDB, and MongoDB (NoSQL)

REFERENCES

- **A. Sorana Morrissy, PhD,** Sorana.Morrissy@ucalgary.ca
  Assistant Professor,
  Department of Biochemistry and Molecular Biology, University of Calgary, Calgary, AB, Canada

- **Benjamin Haibe-Kains, PhD,** Benjamin.Haibe.Kains@utoronto.ca
  Senior Scientist,
  Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada
  Associate Professor,
  Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada

- **Ju Han Kim, MD, PhD, MS,** juhan@snu.ac.kr
  Chair and Professor, Div. of Biomedical Informatics,
  Seoul National University College of Medicine, Seoul, Korea