**Profile**

* Recent Ph.D. graduate looking for a career in Quantitative Analytics, Data Science, or Software Development

**Education**

**Ph.D. Bioinformatics**  
*University of British Columbia 2011–2015*

**B.Sc. Computer Science (Minor Statistics)**   
*University of British Columbia 2006–2011*

**Coursework**

* Machine Learning, Data Visualization, Numerical Methods, Algorithms, and Software Development

**Technical Skills**

* Proficient in Java, C++, R, Python, JavaScript, PHP, Go, and SQL
* Object-oriented software design
* Large-scale data processing using Hadoop, Apache Spark, and AWS
* HPC programming with Intel Phi-cards
* Data Analysis with ggplot2, dplyr, RMarkdown, NumPy, pandas, scikit-learn, and D3.js

**Experience**

**Bioinformatics Ph.D. Candidate** *2011-2015*   
*Hallam Lab, University of British Columbia*

* Developed MetaPathways, a Python and   
  C++ pipeline for biological sequencing data
* Implemented a Master-worker distributed algorithm to manage a large collection of HPC worker grids
* Enabled sequence-homology search in Hadoop through critical improvements to the YARN scheduler
* Analytically involved in all areas research; author on 10+ peer-reviewed publications

**Teaching Assistant** 2014  
*Marine Biological Laboratory, Woods Hole, MA*

* Conducted tutorials on biological sequence analysis for 60+ post-docs, graduate students, and professors
* Data-literacy topics: data cleaning,
* Statistical topics: Predictive Models, SVMs, Hierarchical Clustering, Dimensionality Reduction,
* General data analysis, visualization, and report generation using R, ggplot2, dplyr, and RMarkdown

**Software Developer** 2011  
*Microbiology and Immunology, UBC*

* Developed a LAMP stack and JavaScript based web application cataloging thousands of bacterial strains
* Ensured accurate requirements via user observation and interviews
* Enhanced the user experience by implementing advanced Boolean search, user histories, and search autocomplete

**Projects**

* **De-Confusion Tables**. An interactive R shiny app. Final project. Developing Data Products. John Hopkins University Coursera Class. July 2014 (<http://goo.gl/Fl78Qh>)
* **D3.js Visualizations**. Variety of visualizations in the d3.framework for analysis of genomic data (<http://goo.gl/fLntUc>)
* **R-Package: fastLSA**. Implementation of the fastLSA correlation method in C and Open-MP (<http://goo.gl/5QPM82>)

**Selected Publications**

For full list see Google Scholar (<http://goo.gl/aYo3Wa>).

* Kim, Konwar, **Hanson**, et al. Koonkie: An Automated Software Tool for Processing Environmental Sequence Information using Hadoop. ASE BigData. Harvard (2014)
* **Hanson**, et al. MetaPathways v2.0: A master-worker model for environmental Pathway/Genome Database construction on grids and clouds. IEEE CIBCB, Hawaii (2014)
* Konwar, **Hanson**, et al. MetaPathways: a modular pipeline for constructing pathway/genome databases from environmental sequence information. BMC Bioinformatics (2013)

**Journal & Conference Reviewing**

* IEEE CIBCB Conference 2014 and 2015
* Symposium on Bioinformatics Research and Applications ISBRA 2015

**Interests**

* Data Analysis, Information Visualization, Machine Learning, Statistics, Software Design
* Photography, Cooking, Badminton, Running, Triathlons