Segun C. Jung, PhD, MBA candidate

August 19, 2022

Associate Director of R&D and Clinical Bioinformatics
NeoGenomics Laboratories
1720 Mission St., South Pasadena, CA 91030

http://dauss75.github.io/
segun.jung@neogenomics.com

Education

- **University of Arizona Global Campus** — Arizona, AZ
  MBA, (GPA: 4.0/4.0) — *2021-expected to graduate in Nov. 2022*
- **New York University** — New York, NY
  PhD, Computational Biology (GPA: 3.8/4.0) — *2013*
  - Advisor: Dr. Tamar Schlick
  - Thesis: Modeling RNA Junctions with Applications to Structure Predictions of Regulatory Regions of Viral RNAs
- **Drexel University** — Philadelphia, PA
  M.Sc. Biomedical Engineering (GPA: 3.82/4.0) — *2007*
  - Advisors: Dr. Louise C. Showe
  - Thesis: One-Class Machine Learning Approach for MicroRNA Gene Predictions
- **Korea University of Technology and Education** — South Korea
  B.E. Information Technology (GPA: 3.86/4.0) — *2003*
  - Graduated with a Major in Electronics Engineering and a Minor in Computater Engineering
  - Thesis: Real-time Digital Video Recording Using Wavelet Compression Algorithm

Research Experience

- Research & Development, NeoGenomics Laboratories, Aliso Viejo, CA — 2018 – present

  1) Hire, train, and manage a group of talented bioinformatician with PhD and MS in biology, bioinformatics, computer science, and computational biology.
  2) Responsible for building/maintaining infrastructure with a high performance cluster used in daily operations (clinical lab) (software – PBS, Torque – and 40+ compute nodes with over 4000 processors).
  3) Responsible for developing/maintaining/operationalizing NGS based DNA and RNA pipelines in collaboration with bench scientists for laboratory developed test (LDT).
  4) Established/operationalized MSI (microsatellite instability) and CNV pipelines.
  5) Presentated research work at conferences (i.e., ASH, AACR)

- Globus Genomics Team, The University of Chicago & Argonne National Laboratories, Chicago, IL 2016 – 2018

  1) A core member of Globus Genomics www.globusgenomics.org, a galaxy-based platform in Amazon EC2 linux instance running Ubuntu 16.04, which provides a cloud-based solution for next generation sequencing data. I am responsible for front- and back-end bioinformatic developments and our source code is maintained in git repo. (programming language: **shell scripting, XML, python**, and **R**).

2) Co-led several projects for building best practice workflows for NGS data analyses via industry and academia collaborations that include Institute for Systems Biology, Colgate-Palmolive, UPenn, Boston Univ., and Fred Hutchinson Research Center. Most recently, we installed and wrapped (>100 tools) the Mothur and QIIME suite and built Microbiome workflows for Colgate-Palmolive.

- Biomedical Informatics Consultant, The University of Illinois at Chicago, Chicago, IL 2017 – 2018

1) Performed genomic data analysis to identify candidate genes in diabetic retinopathy, the leading cause of irreversible vision loss, which show a differential response in expression to glucose between diabetic subjects with and without retinopathy in EBV-transformed lymphoblastoid cell lines (LCLs). The analysis involved controlling confounding factors, gene set enrichment analysis (GSEA), ingenuity pathway analysis (IPA), utilizing eTQL data from GTEx, and imputing gene expression using SNP data from 2829 diabetic subjects using Predixcan. (programming language:**R**).
2) Working on a generic screening test project for child blindness that is caused by automosal recessive manner from both parents. I have been implementing a program that classifies pathogenic rare variants using public data set (gnomAD) based on the American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP) guideline. Then, we perform a hardy-weinberg equilibrium test and estimate the number of affected population from allele frequency for each predicted pathogenic variant (programming language:**python** and **R**).

- Postdoctoral Fellow, Davuluri Lab, Northwestern University Feinberg School of Medicine, Chicago, IL 2013 – 2015

- Worked on next generation sequencing (NGS) data analysis that includes RNA-seq, exon-array, miRNA-seq, SNP array, ChIP-seq, and Whole Exom Sequencing (WES) with various statistical and machine learning techniques (programming language: **R**).

1) Identification of casual generic variants for Prostate Cancer using TCGA RNA-seq and SNP array, and ChIP-seq data (tools - overlapSelect, awk, bedtools, and R packages (DESeq2, foreach and doParallel))
2) Copy number variation analysis using WES and SNP array for native African prostate cancer patients (tools - bwa, fastQC, samtools, bamtools, picard, and GATK)
3) MicroRNA-seq data analysis related to wound/healing in mouse (tools - fastQC, featureCounts, htseq-count, trim_galore(cutadapt wrapper), and R package (DESeq))
4) Illumina beadAarray data analysis for polycythemia vera patients with and without interferon treatment (tools - R packages (limma, illuminaio, lumi))

- Graudate Student, Schlick Lab, NYU, New York 2008 – 2013

- Worked on finding recurrent structural elements in RNA junctions
- Worked on 3D structure modeling of viral RNAs (e.g., fmdv ires) based on knowledge-based tertiary RNA motifs, and conventional molecular dynamics simulations (using NAMD + Amber force field).
- Worked on developing a coarse-grained Monte Carlo program coupled with statistical potentials to predict RNA 3D Structures (programming language: **C++**):

- R&D Intern, Schrodinger Inc, New York 2011

- Worked on developing a new protocol for a ligand placement in PrimeX including an extended sampling and also improved the procedure of structure refinements (programming language: **python**).

- Computational Scientist, Showe Lab, Systems Biology Devision, The Wistar Institute, PA 2006 – 2007
  - Worked on several projects relate to microRNA gene and target prediction utilizing machine learning techniques (programming languages: **shell scripting, perl, cgi, php, awk, matlab**). The webserver I developed are as follows (Note that these web servers are no longer maintained):
  - http://wotan.wistar.upenn.edu/OneClassmiRNA/
  - http://wotan.wistar.upenn.edu/NBmiRTar/
  - http://wotan.wistar.upenn.edu/BayesSVMmiRNAfind/

Professional Experience

- Associate Director of R&D and Clinical Bioinformatics at NeoGenomics Laboratories, CA 2018 – present
- Biomedical Informatics Consultant at the University of Chicago, IL 2016 – 2018
- Postdoctoral Fellow at Northwestern University, IL 2014 – 2015
- Postdoctoral Fellow at the Wistar Institute, PA 2013 – 2014
- R&D Intern at Schrodinger Inc, NY 2011
- Adjunct Instructor at NYU, NY 2010
- Computational Scientist at the Wistar Institute, PA 2006 – 2007
- R&D Intern at Coriell Institute for Medical Research, NJ 2005
- System Engineer at Samsung Electronics Co., Ltd, South Korea 1994 – 1998

Grant

- NSF-STTR-IIP-1648937 (Co-PI: Segun Jung) - www.sbir.gov/sbirsearch/detail/1216233

  STTR Phase I: Development of a High-Performance Clinical Analysis Platform to Support Precision Medicine.

Selected Publications: Full List in Google Scholar

1) Skol AD, **Jung S**, Sokovic AM, Chen S, Fazal S, Sosina O, Borkar PP, Lin A, Sverdlov M, Cao D, Swaroop A, Bebu I, DCCT/ EDIC Study group, Stranger BE, Grassi MA. *Integration of genomics and transcriptomics predicts diabetic retinopathy susceptibility genes.* **Elife.** 9:e59980 (2020). doi:10.7554/eLife.59980..

2) Funk CC, **Jung S**, Richards MA, Rodriguez A, Shannon P, Donovan R, Heavner B, Chard K, Xiao Y, Glusman G, Erteskin-Taner N, Golde T, Toga A, Hood L, Van Horn JD, Kesselman C, Foster I, Ament S, Madduri R, Price ND. *Atlas of Transcription Factor Binding Sites from ENCODE DNase Hypersensitivity Data Across 27 Tissue Types.* **Cell Reports.** 32(7):108029 (2020). doi:10.1016/j.celrep.2020.108029..

3) Jin H, **Jung S**, DebRoy AR, and Davuluri RV., *Identification and validation of regulatory SNPs that modulate transcription factor chromatin binding and gene expression in prostate cancer.* **Oncotarget.** 7(34): 54616–54626 (2016). doi:10.18632/oncotarget.10520.

4) Kim N, Laing C, Elmetwaly S, **Jung S**, Curuksu J, and Schlick T, *Monte Carlo Sampling of 3D RNA Graph Topologies.* **PNAS.** 111(11):4079-4084 (2014).doi:10.1073/pnas.1318893111. PMID:24591615.

5) **Jung S** and Schlick T, *Interconversion between Parallel and Antiparallel Conformations of a 4H RNA junction in Domain 3 of Foot-and-Mouth-Disease Virus IRES Captured by Dynamics Simulations.* **Biophysical J.** 106(2): 447-458 (2014).doi:10.1016/j.bpj.2013.12.008. PMID:24461020.

6) Laing C*, **Jung S***, Kim N, Elmetwaly S, Zahran M, and Schlick T, *Predicting Helical Topologies in RNA Junctions as Tree Graphs.* (*equal contribution) **PLoS One.** 8(8):e71947 (2013).doi:10.1371/journal.pone.0071947. PMID:23991010.

7) **Jung S** and Schlick T, *Candidate RNA Structures for Domain 3 of the Foot-and-Mouth-Disease Virus Internal Ribosome Entry Site.* **Nucleic Acids Res.** 41(3):1483-95 (2012).doi:10.1093/nar/gks1302. PMID:23275533.

8) Yousef M, **Jung S**, Kossenkov AV, Showe LC and Showe MK, *Naive Bayes classifier for microRNA target gene identification.* **Bioinformatics.** 23(22):2987-92 (2007). doi:10.1093/bioinformatics/btm484. PMID:17925304.

Honours

Postdoctoral Professional Development Travel Award 2015
Henry M. MacCracken Fellowship . . . . . 2009 – 2013
Sackler Institute Biomedical Sciences Training Fellowship . . . . . 2007 – 2008
Greater Philadelphia Bioinformatics Alliance (GPBA) Internship Award ($12,000) 2005
Graduate Study Abroad Scholarship by the KOSEF ($60,000) . . . . . 2004 – 2006
Dean's Fellowship . . . . . 2004 – 2006
President Award with Highest Honor 2003
Excellent Undergrad Research Award 2003
Merited-Based Full Scholarship . . . . . 1999 – 2002

Teaching Experience

- General Chemistry Laboratory I — Fall, 2010
  Responsible for teaching, demonstrating, and grading each lab experiment for 32 NYU Undergraduates
- General Chemistry Laboratory II — Spring, 2009
  Responsible for grading, demonstrating, and guiding each lab experiment for 78 NYU Undergraduates