

**Research statement** I'm a computer person who studies the biodata generation process to develop both general and specific tools for interpretable biology.

**Education**

**Ph. D. Computer Science**, 2016  
**M. A. Computer Science**, 2014  
Princeton University, Princeton, NJ  
Thesis: "Targeted analysis of very large genome-wide data collections"  
Advisor: Olga G. Troyanskaya

**B. S. Computer Science**, 2010  
**B. S. Mathematics**, 2010  
The University of Texas at Austin, Austin, TX  
Dean's Scholars Honors  
Turing Scholars Honors  
Thesis: "Developing Scalable Quartet Tree Encodings"  
Advisor: Tandy Warnow

**Experience**

**Postdoctoral fellow**, 2016 - present  
Seoul National University, Seoul, South Korea  
Advisor: V. Narry Kim

**Assistant Instructor**, Fall 2011, Spring 2012  
Princeton University, Princeton, NJ  
An integrated, Quantitative Introduction to the Natural Sciences

**Selected Publications**

1. Kim D\*, Yeo J\*, **Lee YS\***, Jung SJ\*, Seo JJ, Lee YY, Lim J, Chang H, Song J, Yang J, Ahn K, Jung G and Kim VN (under review) "Viral hijacking of the TENT4-ZCCHC14 complex protects viral RNAs via mixed tailing." \*equal contributions
2. Kim H\*, **Lee YS\***, Kim SM, Choi H, Lee JW, Kim TD and Kim VN (submitted) "N6-methyladenosine demethylase FTO orchestrates motile ciliogenesis by stabilizing FOXJ1 mRNA." \*equal contributions
3. **Lee YS**, Krishnan A, Oughtred R, Rust R, Chang CS, Ryu J, Kristensen VN, Dolinski K, Theesfeld CL and Troyanskaya OG (2019) "A Computational Framework

for Genome-wide Characterization of the Human Disease Landscape." *Cell systems* 8 (2), 152-162. e6

4. **Lee YS**, Wong AK, Tadych A, Hartmann BM, Park CY, DeJesus VA, Ramos I, Zaslavsky E, Sealfon SC and Troyanskaya OG (2018) "Interpretation of an individual functional genomics experiment guided by massive public data." *Nature methods* 15 (12), 1049

5. Lim J\*, Kim D\*, **Lee YS\***, Ha M, Lee M, Yeo J, Chang H, Song J, Ahn K and Kim VN (2018) "Mixed tailing by TENT4A and TENT4B shields mRNA from rapid deadenylation." *Science* 361 (6403), 701-704, \*equal contributions

6. **Lee YS**, Krishnan A, Zhu Q and Troyanskaya OG (2013) "Ontology-aware classification of tissue and cell-type signals in gene expression profiles across platforms and technologies." *Bioinformatics* 29 (23), 3036-3044

#### Other Publications

7. Choi Y, Jeong K, Shin S, Lee JW, **Lee YS**, Kim S, Kim SA, Jung J, Kim KP, Kim VN and Kim JS (submitted) "Proteome quantification platform with maximally increased multiplexing for SILAC and di-alkylation labeling."

8. Kwon SC, Baek SC, Choi YG, Yang J, **Lee YS**, Woo JS and Kim VN (2019) "Molecular basis for the single-nucleotide precision of primary microRNA processing." *Molecular cell* 73 (3), 505-518. e5

9. Nirschl C, Suárez-Fariñas M, Izar B, Prakadan S, Dannenfelser R, Tirosh I, Liu Y, Zhu Q, Devi KSP, Carroll SL, Chau D, Rezaee M, Kim TG, Huang R, Fuentes-Duculan J, Song-Zhao GX, Gulati N, Lowes MA, King SL, Quintana FJ, **Lee YS**, Krueger JG, Sarin KY, Yoon CH, Garraway L, Regev A, Shalek AK, Troyanskaya OG and Anandasabapathy N (2017) "IFN $\gamma$ -dependent tissue-immune homeostasis is co-opted in the tumor microenvironment." *Cell* 170 (1), 127-141. e15

10. Bansal M, Yang J, Karan C, Menden MP, Costello JC, Tang H, Xiao G, Li Y, Allen J, Zhong R, Chen B, Kim M, Wang T, Heiser LM, Realubit R, Mattioli M, Alvarez MJ, Shen Y, **NCI-DREAM Community**, Gallahan D, Singer D, Saez-Rodriguez J, Xie Y, Stolovitzky G and Califano A (2014) "A community computational challenge to predict the activity of pairs of compounds." *Nature biotechnology* 32(12), 1213

11. Park CY, Krishnan A, Zhu Q, Wong AK, Lee YS and Troyanskaya OG (2014) "Tissue-aware data integration approach for the inference of pathway interactions in metazoan organisms." *Bioinformatics* 31(7), 1093-1101.

12. Ju W, Greene CS, Eichinger F, Nair V, Hodgins JB, Bitzer M, **Lee YS**, Zhu Q, Kehata M, Li M, Jiang S, Rastaldi MP, Cohen CD, Troyanskaya OG and Kretzler M (2013) "Defining cell-type specificity at the transcriptional level in human disease." *Genome research* 23 (11), 1862-1873

## **Presentations**

"Reconstructing the social network of a cell from big data." Seminar in Bioinformatics, Seoul National University, Seoul, South Korea (2019).

"Guanylation at the 3'-end of mRNA." Mini-symposium on RNA Biology, Seoul National University, Seoul, South Korea (2017).

"Ontology-aware classification of tissue and cell-type signals in gene expression profiles across platforms and technologies." Late Breaking Research, Intelligent Systems for Molecular Biology Conference, Berlin, Germany (2013).

## **Referees**

### **V. Narry Kim, Ph. D.**

Director, RNA Research Center, Institute for Basic Science  
Professor, School of Biological Sciences, Seoul National University  
Seoul, South Korea  
+82-2-880-9120  
narrykim@snu.ac.kr

### **Olga G. Troyanskaya, Ph. D.**

Professor, Department of Computer Science and  
Lewis-Sigler Institute for Integrative Genomics,  
Princeton University  
Princeton, NJ, USA  
(609) 258-1749  
ogt@genomics.princeton.edu

### **Stuart C. Sealfon, M. D.**

Professor, Department of Neurology and  
Center for Advanced Research on Diagnostic Assays,  
Icahn School of Medicine at Mount Sinai  
New York, NY, USA  
(212) 241-7327  
stuart.sealfon@mssm.edu

### **Arjun Krishnan, Ph. D.**

Assistant Professor, Department of Computational Mathematics, Science and  
Engineering and Department of Biochemistry and Molecular Biology,  
Michigan State University  
East Lansing, MI, USA  
(517) 432-0372  
arjun@msu.edu