Kunal Bhutani, PhD

90 W. Cedar St, 5F
Boston, MA, 02114
kbhutani@ucsd.edu | kunalbhutani@gmail.com
Webpage : kunalbhutani.com
Github : @kunalbhutani
832-512-6517

Education

University of California, San Diego, La Jolla, California
Doctor of Philosophy, Bioinformatics & Systems Biology — *February 2017*

University of Texas at Austin, Austin, Texas
Bachelor of Science, Biomedical Engineering — *May 2011*

Research Interests

Applied Bayesian Statistics, Measurement Error, Variant Calling, Transcriptome-wide Association Studies, N-of-1 Studies, Genomics, Transcriptomics, Metabolomics

Research Experience

Transcriptional regulation and imputed gene expression-phenotype associations
Prof. Manolis Kellis and Prof. Nicholas J Schork, MIT — *December 2015 - Present*
Created a framework for estimating uncertainty in trained models of transcriptional regulation and developed a novel hierarchical Bayesian regression model for propagating uncertainty in transcriptome-wide association studies. Performed simulation studies to assess power differences across competing methods, and in application to real data, identified 45 gene expression-phenotype relationships for seven complex diseases.

Molecular surveillance of a germline TP53 mutation carrier
Prof. Nicholas J Schork, JCVI — *April 2015 - Present*
Analyzed metabolomics, microbiome, and RNA-Seq data, and developed statistical methods for identifying deviations from baseline for cancer biomarkers in an n-of-1 study design. Created a latent Bayesian model to separate instrumentation variation from biological variation, and performed power calculations for future longitudinal metabolomic profiling studies.

Mutational burden analysis of three pluripotency induction methods
Prof. Jeanne Loring and Prof. Nicholas J Schork, TSRI — *September 2013 - December 2015*
Identified and characterized mutations induced by reprogramming of iPSCs by analyzing whole-genome sequencing data. Applied statistical tests and performed power calculations to assess the differences in types and annotated features of mutations.

Quality control for single nuclei RNA-Seq
Prof. Roger Lasken and Prof. Nicholas J Schork, JCVI — *January 2015 - May 2015*
Created a quality control pipeline for single nuclei RNA-Seq with emphasis on identifying 3' bias, library complexity, and relative exonic coverage.

Detection of somatic mutations with sample impurity inference
Prof. Vineet Bafna, UCSD — *March 2012 - July 2013*
Developed a Bayesian model that incorporates an estimate of normal contamination during somatic variant calling. Performed simulation studies and compared precision and recall against competing methods.

Publications

Kim S*, Jeong K*, **Bhutani K**, Lee JH, Patel A, Scott E, Nam H, Lee H, Gleeson JG, Bafna V. "Virmid: accurate detection of somatic mutations with sample impurity inference". *Genome biology.* 14(8). 2013

Bhutani K, Nazor KL, Williams R, Tran H, Dai H, Dzakula Z, Cho EH, Pang AW, Rao M, Cao H, Schork NJ, Loring JF. "Whole-genome mutational burden analysis of three pluripotency induction methods". *Nature Communications.* 7. 2016.

Krishnaswami SR*, Grindberg RV*, Novotny M, Venepally P, Lacar B, **Bhutani K**, et al. "Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons". *Nature protocols*. 11(3). 2016.

Buckley AR, Standish KA, **Bhutani K**, Ideker T, Harismendy O, Carter H, Schork NJ. "Pan-Cancer Analysis Reveals Technical Artifacts in The Cancer Genome Atlas (TCGA) Germline Variant Calls". *bioRxiv*. 2016.

Bhutani K*, Sarkar A*, Park Y, Kellis M, Schork NJ. "Modeling prediction error improves power of transcriptome-wide association studies". *bioRxiv*. 2017.

Park Y*, Sarkar A*, **Bhutani K**, Kellis M "Multi-tissue polygenic models for transcriptome-wide association studies". *bioRxiv*. 2017.

Bhutani K*, Magnuson V*, Buckley A, Quarless D, Goetz L, Schork NJ. "Longitudinal metabolome, microbiome, and transcriptome profiling of a germline TP53 mutation carrier". *In Preparation*. 2017.

Presentations

Modeling prediction error improves power of transcriptome-wide association studies
Talk — *Leena Peltonen School of Human Genomics - August 2016*

Propagating uncertainty of predicted expression in transcriptome-wide association studies
Poster — *American Society of Human Genetics - Oct 2016, The Biology of Genomes - May 2016*

Quantifying DNA heterogeneity in a population of cells
Talk — *J. Craig Venter Institute Informatics Seminar - September 2014*

Whole genome sequencing of Li-Fraumeni families reveals heterogeneous de-novo mutation signatures
Poster — *American Society of Human Genetics - October 2013*

Work Experience

Illumina, Inc., La Jolla, CA
Bioinformatics Intern — *July 2012 - September 2012*

Flok, Austin, TX
A Social Geo-location Startup, Co-Founder — *January 2010 - September 2011*

Awards & Achievements

Invited to Leena Peltonen School of Human Genomics, National Science Foundation Graduate Research Fellowship, Texas Exes Award for Scholarship and Leadership

Technical Skills

Programming Languages: Python, R, Bash
Statistics: Bayesian Inference, Generalized Linear Models, Mixed Models, Mixture Models, Measurement Error, Dimensionality Reduction, Outlier Detection, Changepoint Detection
Bioinformatics:
Whole Genome Sequencing: processing, quality control, variant calling: germline and somatic, and annotation
Population Genetics: imputation, genome-wide association studies, transcriptome-wide association studies
RNA-Seq: processing, quality control, and differential expression for single cell and population transcriptomics
Familiarity with metabolomics and microbiome data.