

Vinh Pham

Austin, Texas 78703. 512-915-0719. ✉ phamngocvinh@utexas.edu

github.com/donalbonny

SUMMARY

- Solid experience in molecular genetics, next-generation sequencing (NGS) genomics data processing and data analysis, and statistical analysis.
- Proficient in Unix operating systems, command line interfaces, and high-performance computing
- Skilled in statistical programming R (tidyverse, dplyr, ggplot2, reshape2, caret) and Python (NumPy, Pandas, Matplotlib, scikit-learn)
- Experienced as a field application scientist working with next-generation sequencing.
- A strong passion for solving healthcare challenges with quantitative methodologies, passionate about improving human health by leveraging big data opportunities.

EDUCATION

Doctor of Philosophy in Molecular Genetics/Genomics 2014- present

University of Texas, Austin (UT Austin)

Thesis: Transcriptional regulatory networks in light signaling in Arabidopsis

Master of Science (Integrated Biosciences and Biotechnology) 2011-2013

Pohang University of Science and Technology (POSTECH), South Korea

Bachelor of Science (Honors program in Biochemistry) 2006-2010

Vietnam National University, Hanoi (VNU)

PROFESSIONAL DEVELOPMENT

The University of Texas at Austin: Introduction to programming for researchers January – May 2017

- Topics: Unix/Linux, bash, algorithms, Python, Pandas, NumPy, Matplotlib, parallel programming

The University of Texas at Austin: Bioinformatics January – May 2018

- Topics: R (tidyverse, ggplot2, clustering, logistic regression, classification, cross-validation) and Python

Instructor at Data Skills Boot Camp, Rutgers University May 2018

- Topics: SQL, Introduction to R/OpenRefine

Instructor at American Society of Plant Biologists Conference July 2018

- Topics: Genomics Data analysis using Unix, R and Python

The University of Texas at Austin: Machine Learning Methods in Gene Expression Analysis June 2018

- Topics: clustering, classification (random forests, SVM, LDA, kNN) and regression in R and Python

Udemy: Python for Data Science and Machine Learning September 2018-present

- Topics: Pandas, Scikit-learn, NumPy, Matplotlib, Spark, Seaborn, Machine Learning

Kaggle: Machine Learning October 2018-present

- Topics: Machine learning models, Underfitting and Overfitting, Random forests, Cross-validation

Volunteer, Data Day Texas January 2019

- Topics: Data Science, Analytics, Machine Learning, Deep Learning and Neural Networks

WORK EXPERIENCE

Graduate Research Assistant, UT Austin 2014- present

- **Project 1:** Applied negative binomial generalized linear models for different expression comparison and clustering analysis (k-means and hierarchical clustering) in R using gene expression data to reveal intricate network of developmental pathways regulated by key regulators in Arabidopsis.
- **Project 2:** Develop computational pipelines using Unix/R and statistical based methods using R to analyze co-expression data from 60 different transcriptome sequencing samples to understand transcriptional regulatory network in light signaling and identify the novel key regulators in light response.
- **Project 3:** Apply machine learning based gene regulatory networks analysis methods (mDNA/random forest) using R and Python combined with large sequencing datasets in Arabidopsis to identify the light signaling novel related genes based on expression characteristics patterns.

Field Applications Scientist, Biomedic JSC-Illumina's Exclusive Distributor, Vietnam 2013-2014

- **Technical support:** Prepared library for RNAseq, DNAseq, whole genome sequencing metagenomics, analyzed large-scale data (including genetics screening in lung cancer, HLA typing, prenatal genetic diagnostic, metagenomics analysis).
- **Customer support:** Worked with research scientists and doctors to optimize protocols for HBV, HCV detection and genetic screening in lung cancer mutations by sequencing data analysis.
- Provided training and seminars for internal marketing, sale teams and Vietnamese customers on next-generation sequencing analysis

Instructor, Software/Data Carpentry

2017- present

- Organize and teach workshops on Genomics data analysis using Shell, Git, R and cloud computing.

Teaching Assistant, UT Austin

2017-2018

- Designed and led discussions, instructed students to perform molecular biology, cell culture, microscopy, biochemistry and sequencing data analysis.

SKILLS

Programming languages: •Python (NumPy, Pandas, Scikit-learn) •R (tidyverse, dplyr, ggplot2, reshape2, caret)
•SQLite •Linux/Unix •High-performance computing

Technical Skills: •Data manipulation •Data visualization •Statistical Analysis •Project management • Experimental design •Critical thinking •Technical documentation

High throughput data analysis: •Next-generation sequencing data analysis (whole-genome sequencing, RNA-seq, ChIP-Seq, ATAC-Seq) •Comparative genomics •Cancer Mutations •Epigenomics •Metagenomics analysis
•Differential Gene expression analysis •Regulatory network analysis

Molecular Biology: •DNA/RNA extraction •DNA/RNA library preparation •Primer design •Protein purification and expression •qRT-PCR

AWARDS

- Graduate Dissertation Writing Fellowship, UT Austin (Spring, 2019)
- Teaching Preparation Certificate, Faculty Innovation Center, UT Austin (2018)
- Professional Development Award and Travel award, The University of Texas at Austin (2018)
- US National Science Foundation Travel Award, International Conference on Arabidopsis Research (2017)
- Professional Development Award, The University of Texas at Austin (2017)
- Dean's Prestigious Fellowship, UT Austin (2014, 2015)
- Vietnam Education Foundation (VEF) Ph.D. Fellowship (2014-2016)
- POSTECH Graduate Fellowship (2011-2012)
- Life Science (POSTECH) Undergraduate Fellowship (2010)
- POSCO Asia Fellowships-POSCO TJ Park Foundation (2008)
- Dean's Scholar, Vietnam National University, Hanoi (VNU) (2008, 2010)
- Scholarship for excellent student, Ministry of Education and Training, Vietnam (2009, 2010)

PUBLICATIONS

- Pham VN**, Huq E. Light signaling regulatory network in Arabidopsis (In Preparation)
- Pham VN**, Hoecker U, Huq E. Genomic evidence reveals SPA-regulated developmental and metabolic pathways in dark-grown Arabidopsis seedling. **BMC Genomics** (Under Review)
- Paik I, Chen F, **Pham VN**, Zhu L, Kim JI and Huq E. A phyB-PIF1-SPA1 kinase regulatory module promotes photomorphogenesis. **Nature Communication** (Under Review)
- Pham VN**, Xu X, Huq E. Molecular bases for the constitutive photomorphogenic phenotypes in Arabidopsis. **Development 2018. Selected for the research highlight and interview for "The people behind the papers"**.
- Pham VN**, Kathare PK, Huq E. Dynamic regulation of PIF5 by COP1-SPA complex to optimize photomorphogenesis in Arabidopsis **Plant Journal**. (2018); 96: 260-273 **Selected for the research highlight and cover photo for the issue.**
- Pham VN**, Kathare PK, Huq E. Phytochromes and Phytochrome Interacting Factors. **Plant Physiology**. 2018;176 :1025-1038.
- Xu X, Kathare PK, **Pham VN**, Nguyen A, Bu Q, Huq E. Reciprocal proteasome-mediated degradation of PIFs and HFR1 underlying photomorphogenic development in Arabidopsis. **Development**. 2017;144 (10) :1831-1840

TALKS AND POSTER PRESENTATIONS

- Pham VN**, Xu X, Huq E. PIFs and COP1/SPA: the key regulators of photomorphogenesis in light signaling in Arabidopsis. **Lighting talk, ASPB Plant Biology 2018**, Montreal, Canada, July 2018.
- Pham VN**, Xu X, Huq E. Molecular bases for the constitutive photomorphogenic phenotypes in Arabidopsis. **Graduate Student Competition talk, ASPB, Southern Section Annual Meeting**, New Orleans, USA, March 2018.
- Pham VN**, Kathare PK, Huq E. Dynamic regulation of PIF5 by COP1-SPA complex in Arabidopsis. **28th International Conference on Arabidopsis research**. St Louis, USA. June 2017.
- Pham VN**, Jeong SY, Choi HM, Soh MS, Nam HG. Novel function of nuclear-encoded chloroplast protein in nucleus. **23rd International Conference on Arabidopsis research**. Austria. July 2012.