Phu T. Van, PhD

email pvan@alumni.cmu.edu

website ptvan.github.io

SUMMARY

Seeking a scientific or technical position on collaborative teams working to solve data-intensive problems. I am particularly interested in projects that utilize statistical modeling or machine learning.

WORK EXPERIENCE

2015–Present **Systems Analyst, Analyst Group Leader**

Fred Hutchinson Cancer Research Center

Led the lab's PhD-level analysts in collaboration meeting and journal club. Performed analyses on flow cytometry and transcriptomic datasets. Authored Statistical Analysis Plans & peer-reviewed manuscripts. Created bioinformatic analysis pipelines & interactive data visualizations. Mentored PhD students and bench scientists on statistics, performed reviews of junior analysts' code.
Reference: Evan Henrich · ehenrich@fredhutch.org

2014–2015 **Postdoctoral Research Fellow**

Fred Hutchinson Cancer Research Center

Developed R code to normalize mass cytometry data across experiments using multi-mixture models. Worked with engineers from FlowJo Inc. (Ashland, OR) to enable XML parsing in FlowJo v10. Performed dimension reduction in mass cytometry data and identified correlates of disease progression in early-stage vaccine candidates.
Reference: Evan Greene, PhD · palmergreene@gmail.com

2009-2014 **Doctoral Student**

Carnegie Mellon University

Designed and built a patented high-dynamic-range protein gel imager with robotic gel cutting arm for capturing rare proteins. Developed **SIGILab**, a C++ GUI application controlling gel imager's acquisition of high-dynamic-range images. Developed workflows to quantify protein abundances in 2DE gel images using bash.
Reference: Jonathan Minden, PhD · minden@cmu.edu

2006-2009 **Research Associate**

Institute for Systems Biology

Identified factors that affect peptide detectability in mass spectrometry experiments using R. Constructed a SQLServer database with web frontend for exploring mass spectrometry data and planning proteomic experiments. Developed a regression algorithm in R to model regulators of stress response in microarray data.
Reference: Serdar Turkarslan, PhD · serdar.turkarslan@isbscience.org

TECHNICAL SKILLS

Programming	R, Python, Matlab, C/C++, bash, JavaScript, Perl
Data Analysis	dimension reduction, clustering, regression/classification
Statistics	linear models, variable selection, hypothesis testing
Databases	MySQL/MariaDB, PostgreSQL, SQLServer
Workflows	Nextflow, WDL, Slurm, Docker, GitHub

Bioinformatics STAR, RSEM, BLAST, DEG, GSEA, KEGG, STRING

EDUCATION

2009-2014 **Carnegie Mellon University**

PhD, Biological Sciences

2001-2007 **University of Washington**

B.S., Biology (Physiology specialization), B.S., Wildlife Sciences

SELECTED PUBLICATIONS & PATENTS

Mar. 2020 **Monocyte metabolic programs are associated with resistance to TST/IGRA conversion**

in review

This manuscript reports a potential link between oleic metabolism and *Tuberculosis* resistance in humans. I was the primary bioinformatician on this project, responsible for analyzing data, creating visualizations and making statistical/machine learning recommendations. I also assisted in writing the manuscript.
Authors: Jason Simmons, Phu T. Van, *et al.*

Nov. 2018 **ggCyto: next generation open-source visualization software for cytometry**

Bioinformatics

ggCyto enables grammar-of-graphics plotting of high-dimensional flow cytometry and mass cytometry data in a freely accessible R package. I contributed to the R codebase, performed testing and co-wrote the manuscript.
Authors: Phu T. Van*, Wenxing Jiang*, Raphael Gottardo, Greg Finak (**co-first authors*)

Nov. 2010 **Coordination of frontline defense mechanisms under severe oxidative stress**

Molecular Systems Biology

This manuscript reports a gene network that regulates microbial response to oxidative stress. I wrote the R code responsible for predicting the gene network, processed microarray data, and co-wrote the manuscript.
Authors: Amardeep Kaur, Phu T. Van, *et al.*

Jul. 2019 **US10362237: Structured illumination system for increased dynamic range in quantitative imaging**

United States Patent

"The systems disclosed herein employ an iterative image collection strategy that utilizes structured illumination to achieve greater than 1,000,000-fold dynamic range measurements, representing a dramatic improvement over the prior art."
Inventors: Jonathan Minden, Frederick Lanni, Phu T. Van