

Phu T. Van, PhD

email pvan@alumni.cmu.edu
website ptvan.github.io

WORK EXPERIENCE

TwinStrand Biosciences	2021–	Bioinformatics Scientist II	Performed analyses on duplex sequencing data. Reference: Evan Henrich · evan.henrich@gmail.com
	2014–2020	Postdoc, Bioinformatics Analyst	Coordinated projects among PhD-level analysts, led brainstorming sessions and journal clubs. Performed analyses on flow cytometry and transcriptomic datasets. Created data analysis pipelines & interactive data visualizations. Authored Statistical Analysis Plans & peer-reviewed manuscripts. Reviewed junior analysts' code, mentored PhD students and bench scientists on statistics. Developed R code to normalize mass cytometry data across experiments using multi-mixture models. Performed dimension reduction on mass cytometry data and identified correlates of disease using regression. Reference: Amy Paguirigan, PhD · apagui@fredhutch.org
Fred Hutchinson Cancer Research Center	2009–2014	Doctoral Student	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 bash scripts to quantify protein abundances in 2DE gel images. Lectured in Modern Biology course and mentored junior students. Reference: Jonathan Minden, PhD · minden@cmu.edu
Carnegie Mellon University	2006–2009	Research Associate	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 weighted regression algorithm in R to predict regulators of stress response in microarray data. Reference: Serdar Turkarslan, PhD · serdar.turkarslan@isbscience.org
Institute for Systems Biology			

TECHNICAL SKILLS

Programming	R, Python, Matlab, C/C++, bash, JavaScript, Perl
Data Analysis / Statistics	dimension reduction, clustering, regression/classification, linear models, variable selection, hypothesis testing
Databases	MySQL/MariaDB, PostgreSQL, SQLServer
Workflows	Nextflow, WDL, Slurm, Docker, GitHub
Bioinformatics	STAR, RSEM, BLAST, SAMtools/BCFtools, DEG, GSEA, KEGG, STRING

EDUCATION

2009–2014 Carnegie Mellon University

PhD, Biological Sciences

2001-2007 University of Washington

BS, Biology (Physiology specialization); BS, Wildlife Sciences

SELECTED PUBLICATIONS & PATENTS

complete list: <https://scholar.google.com/citations?user=IDadFEkAAAAJ>

- | | |
|---|---|
| <i>in revision at J. of
Clin. Invest.</i> | <p><i>Feb. 2021</i> Monocyte metabolic programs are associated with resistance to TST/IGRA conversion</p> <p>This manuscript reports a potential link between oleic metabolism and <i>Tuberculosis</i> resistance in humans. As the primary analyst on the project, I aligned the raw transcriptomic data to the genome, quantified gene expression, created data visualizations and made recommendations on statistical/machine learning methods. I also assisted in writing the manuscript.
Authors: Jason Simmons, PHU T. VAN, <i>et al.</i></p> |
| <i>Bioinformatics</i> | <p><i>Nov. 2018</i> ggCyto: next generation open-source visualization software for cytometry</p> <p>This R package enables plotting of high-dimensional flow cytometry and mass cytometry data in the grammar-of-graphics style. I contributed to the R codebase, performed testing and co-wrote the manuscript.
Authors: PHU T. VAN*, Wenxing Jiang*, Raphael Gottardo, Greg Finak (*<i>co-first authors</i>)</p> |
| <i>Electrophoresis</i> | <p><i>Sep. 2014</i> High dynamic range proteome imaging with the structured illumination gel imager</p> <p>I built the Structured Illumination Gel Imager (SIGI), which enabled a 1000-fold increase in dynamic range over then-current methods while imaging 2D electrophoretic gels. I wrote the prototype imager software in Matlab, the production software in C/C++ and this manuscript, which informed the subsequent patent application for the imager.
Authors: PHU T. VAN, Victor Bass, Dan Shiwerski, Frederick Lanni, Jonathan Minden</p> |
| <i>Molecular Systems
Biology</i> | <p><i>Nov. 2010</i> Coordination of frontline defense mechanisms under severe oxidative stress</p> <p>This manuscript reports a previously-undiscovered gene network that regulates cellular response to oxidative stress in microbes. I analyzed microarray data, wrote the R code for predicting the gene network, and co-wrote the manuscript.
Authors: Amardeep Kaur, PHU T. VAN, <i>et al.</i></p> |
| <i>United States
Patent</i> | <p><i>Jul. 2019</i> US10362237: Structured illumination system for increased dynamic range in quantitative imaging</p> <p>"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</p> |