

## Phu T. Van, PhD

[pvan@alumni.cmu.edu](mailto:pvan@alumni.cmu.edu), <https://ptvan.github.io>

### Research Interests

Bioinformatics, analytics, statistical inference, data science

### Education

**Ph.D. in Biological Sciences**, Carnegie Mellon University, 2014

**B.S. in Biology** (Physiology specialization), University of Washington, 2007

**B.S. in Wildlife Sciences**, University of Washington, 2007

### Research Experience

#### Bioinformatics Analyst, Analyst Group Leader

**FredHutch**

**Oct. 2015 - present**

**Seattle, WA**

Led the lab's 3 PhD-level analysts in bi-weekly working meeting and journal club. Coordinated analysis between FredHutch and University of Cape Town (South Africa) on multi-site multi-omics data integration project. Performed variable selection and sensitivity analyses for HIV vaccine trial data. Created interactive dashboards to help collaborators identify candidate genes for follow-up experiments. Developed analysis pipelines for multiple data types. Authored Statistical Analysis Plans and advised collaborators on research projects. Co-authored peer-reviewed manuscripts, talks and posters. Mentored PhD students and non-scientific staff on statistics and experimental design. Performed reviews of junior analysts' code.

#### Postdoctoral Fellow

**FredHutch**

**Aug. 2014 - Oct. 2015**

**Seattle, WA**

Developed R code to normalize mass cytometry experiments using multi-mixture models and identifying outliers for quality control. Developed R code for writing XML representation of cell populations to enable data sharing of flow cytometry experiments. Worked with engineers from FlowJo Inc. (Ashland, OR) to incorporate XML parsing into FlowJo X. Identified correlates of disease in flow cytometry data of early-stage vaccines.

#### Research Associate

**Institute for Systems Biology**

**May 2006 - Aug. 2009**

**Seattle, WA**

Identified biophysical factors that differentially affect peptide detectability in tandem mass spectra experiments using R. Constructed a database and corresponding web frontend for archaeal proteomic data. Identified genetic regulators of oxidative stress response across different environmental stresses by developing a weighted regression algorithm in R for microarray data.

#### PhD Student

**Carnegie Mellon University**

**Aug 2009 - May 2014**

**Pittsburgh, PA**

Designed and built a patent-pending 2-dimensional protein gel imager with integrated gel cutting robot for identifying rare proteins. Improved reduction & alkylation of complex protein samples for mass spectrometric sequencing by developing agarose stacking gels to reduce sample loss between isoelectric focusing and SDS-PAGE. Developed automated workflows to quantify protein abundances and detection biases in 2DE gels using bash shell scripts and C/C++ to process imager output.

### **Computational Skills**

**Data Analysis:** clustering, regression, variable selection, hypothesis testing

**Statistical Computing:** R, Python, Matlab

**Programming and Scripting:** bash, Perl, C

**Computing environments:** Slurm, Docker, Git/GitHub

**Image processing:** image registration, segmentation, PSF deconvolution, HDR mapping

**Databases:** mySQL, postgresSQL, SQLServer

**User Interfaces :** HTML, JavaScript, PHP, Shiny

### **Patents & Selected Publications (\*co-authors)**

Minden, J., Lanni, F. and **Van, P.** “Structured Illumination System for Increased Dynamic Range in Quantitative Imaging”, USPTO 10,362,237 granted July 23, 2019

Simmons, JD, **Van PT**, Stein CM, Velen K, Peterson GJ, Graustein AD, Nguyen FK, Gottardo R, Mayanja-Kizza H, Fielding K, Grant A, Ntshiq T, Wallis RS, Churchyard G, Boom WH, Hawn TR. Resistance to TST/IGRA conversion associates with distinct transcriptional programs related to free fatty acid stimulation. In preparation.

**Van PT\***, Jiang W\*, Gottardo R, Finak G. ggCyto: Next Generation Open-Source Visualization Software for Cytometry. *Bioinformatics* 34(22):3951-3953. 2018

**Van PT**, Ganesan V, Bass V, Parthasarathy A, Schlesinger D, Minden J. In-gel equilibration for 2DE-based proteomic workflows. *Electrophoresis*. 35(20):3012-7. 2014

**Van PT**, Bass V, Shiwarski D, Lanni F, Minden J. High dynamic range proteome imaging with the Structured Illumination Gel Imager. *Electrophoresis*. 35(18):2642-55 2014

**Van PT**, Schmid AK, King NL, Kaur A, Pan M, Whitehead K, Koide T, Facciotti M, Deutsch E, Reiss D, Mallick P and Baliga N. Halobacterium salinarum NRC-1 PeptideAtlas: strategies for targeted proteomics. *Journal of Proteome Research*. 7(9):3755-3764. 2008

**Van PT\***, Tkaczyk FA\*, West SD. Ecology of Myocastor Coypus in the Union Bay area. Bachelor Senior Thesis, College of Forest Resources, University of Washington. 2007