

# Phu T. Van, PhD

*Bioinformatics Solutions Manager*

pvan@alumni.cmu.edu , <https://ptvan.github.io>

## SUMMARY

---

I'm interested in scientific or technical positions on collaborative teams working to solve complex problems in computational biology.

## TECHNICAL STRENGTHS

---

<b>Programming</b>	R, Python, Matlab, C/C++, bash, JavaScript, Perl
<b>Data Analysis</b>	dimension reduction, clustering, regression/classification
<b>Statistics</b>	linear models, variable selection, hypothesis testing
<b>Databases</b>	MySQL/MariaDB, PostgreSQL, SQLServer
<b>Workflows</b>	Nextflow, WDL, Slurm, Docker, GitHub
<b>Markup/Interfaces</b>	LaTeX, Markdown, Shiny, HTML
<b>Bioinformatics</b>	BWA, STAR, RSEM, BLAST, SAMtools, BCFtools VardictJava, varscan, delly, limma, edgeR

## EXPERIENCE

---

<b>TwinStrand Biosciences</b> <i>Bioinformatics Solutions Manager</i>	November 2021 - Present <i>Seattle, WA</i>
--	---

- Served as bioinformatics expert for the company in approving new projects.
- Spearheaded "à la carte" bioinformatics offering to customers to increase revenue.
- Provided bioinformatics guidance to external clients with Field&Technical Support department.
- Assisted Custom Applications department in creating new genomic panels.
- Supervised R&D collaborations between bioinformatic and wetlab technicians.
- Oversaw development of new R data analysis packages.
- Initiated and later supervised data analyses that expanded TwinStrand DuplexSeq(TM) support to two new sequencing platforms.
- Spearheaded data mining collaborations with the Data Engineering department and co-led company's Machine Learning Interest Group.
- Produced whitepapers and reviewed marketing materials for technical accuracy.
- Co-authored customer communication SOPs.

<b>TwinStrand Biosciences</b> <i>Bioinformatics Scientist II</i>	January 2021 - October 2021 <i>Seattle, WA</i>
---	---

- Performed analyses on duplex sequencing data for diverse projects including mutagenesis and cancer MRD.
- Contributed R code to internal bioinformatic pipelines and customer-facing reports.
- Presented findings to commercial and academic clients.
- Performed code reviews of R packages.

<b>Fred Hutchinson Cancer Research Center</b> <i>Bioinformatics Analyst, Analyst Group Leader</i>	October 2015 - December 2020 <i>Seattle, WA</i>
--	--

- Performed differential gene expression and geneset enrichment analysis on transcriptomic datasets.
- Performed variable selection and sensitivity analyses on flow cytometry datasets.
- Led the lab's PhD-level analysts in brainstorming sessions and journal club.

- Coordinated analyses on projects with external collaborators.
- Authored Statistical Analysis Plans and peer-reviewed manuscripts.
- Created bioinformatic analysis pipelines using R, Nextflow and WDL.
- Created interactive data visualizations using RMarkdown, igraph and leaflet.
- Mentored PhD students and bench scientists on statistics, advised hiring and performed reviews of junior analysts' code.

### **Fred Hutchinson Cancer Research Center**

October 2014 - October 2015

*Postdoctoral Fellow*

*Seattle, WA*

- Developed R code to normalize mass cytometry data across experiments using multi-mixture models.
- Worked with engineers from FlowJo Inc. (Ashland, OR) to develop parsing in of XML-encoded flow cytometry data in FlowJo v10.
- Performed dimension reduction on mass cytometry data and identified correlates of disease.

### **Carnegie Mellon University**

August 2009 - May 2014

*PhD Student*

*Pittsburgh, PA*

- 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 processes that improved reduction and alkylation of complex protein samples in preparation for mass spectrometric sequencing.
- Developed workflows to quantify protein abundances in 2DE gel images using bash.

### **Institute for Systems Biology**

May 2006 - August 2009

*Research Associate*

*Seattle, WA*

- 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.

## **EDUCATION**

### **Carnegie Mellon University**

*2014*

PhD, Biological Sciences

### **University of Washington**

*2007*

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

## **CONFERENCES, PATENTS AND REPRESENTATIVE PUBLICATIONS**

1. Annette E Dodge, Danielle LeBlanc, Andrew Williams, Phu Van, Jake Higgins, Fang Yin Lo, Jeffry Yaplee, Clint C Valentine, Jesse J Salk, Carole L Yauk, et al. Characterisation of procarbazine-induced mutation spectrum in the bone marrow of mutamouse males using duplex sequencing. In *ENVIRONMENTAL AND MOLECULAR MUTAGENESIS*, volume 63, pages 92–93. WILEY 111 RIVER ST, HOBOKEN 07030-5774, NJ USA, 2022
2. E Cho, F Marchetti, P Van, J Higgins, FY Lo, J Yaplee, CC Valentine III, J Salk, and CL Yauk. Optimal methodological design for duplex sequencing (tm) in tk6 cells determined through a time

- and concentration response analysis following enu treatment. In *Environmental and Molecular Mutagenesis*, volume 62, page 36, 2021
3. D LeBlanc, M Meier, A Williams, J Buick, J Higgins, J Yaplee, P Van, FY Lo, CC Valentine, JJ Salk, et al. Duplex sequencing reveals an attenuated mutation frequency increase in the germ cells of mutamouse males exposed to n-ethyl-n-nitrosourea and benzo [a] pyrene relative to the tgr lacz assay. In *Environmental and Molecular Mutagenesis*, volume 62, page 87, 2021
  4. AE Dodge, D LeBlanc, A Williams, P Van, J Higgins, FY Lo, J Yaplee, CC Valentine, JJ Salk, and CL Yauk. Duplex sequencing (tm) reveals increases in mutation frequencies and c<sub>i</sub> t transitions in the bone marrow of mutamouse males exposed to procarbazine. In *Environmental and Molecular Mutagenesis*, volume 62, page 87, 2021
  5. J. Simmons, K. Dill-McFarland, C. Stein, P. Van, Chihota V., Ntshiq T., Maenetje P., Peterson G., Benchek P., Velen K., Fielding K., Grant A., Gottard R., Mayanja-Kizza H., Wallis R., Churchyard G., Boom W.H., and Hawn T.R. Monocyte transcriptional responses to Mycobacterium tuberculosis associate with clinical resistance to TST/IGRA conversion in Uganda and South Africa. *mSphere*, 2022. in press
  6. G. Moncunill, J. Carnes, C. Young, L. Carpp, S. de Rosa, J. Campo, A. Nhabomba, M. Mpina, C. Jairoce, G. Finak, P. Haas, C. Murie, P. Van, H. Sanz, S. Dutta, Mordmuller B., S. Agnandji, N. Diez-Padriza, N. Williams, J. Aponte, C. Valim, D. Neafsey, C. Daudenberger, J. McElrath, C. Dobano, K. Stuart, and R. Gottardo. Transcriptional correlates of malaria in rts,s/as01-vaccinated african children: A matched case-control study. *eLife*, 2021. in press
  7. J. Simmons, P. Van, C. Stein, Chihota V., Velen K., Fielding K., Grant A., Ntshiq T., Peterson G., Graustein A., Nguyen F., Gottard R., Mayanja-Kizza H., Wallis R., Churchyard G., Boom W.H., and Hawn T.R. Monocyte metabolic transcriptional programs associate with resistance to tuberculin skin test/interferon- release assay conversion. *Journal of Clinical Investigations*, June 2021
  8. J. S. Minden, F. Lanni, and P. T. Van. USPTO US10362237: Structured illumination system for increased dynamic range in quantitative imaging, July 2019
  9. P. Van, W. Jiang, R. Gottardo, and G. Finak. ggCyto: next generation open-source visualization software for cytometry. *Bioinformatics*, 34(22):3951–3953, 11 2018
  10. P. T. Van, V. Ganesan, V. Bass, A. Parthasarathy, D. Schlesinger, and J. S. Minden. In-gel equilibration for improved protein retention in 2DE-based proteomic workflows. *Electrophoresis*, 35(20):3012–3017, Oct 2014
  11. P. T. Van, V. Bass, D. Shiwarski, F. Lanni, and J. Minden. High dynamic range proteome imaging with the structured illumination gel imager. *Electrophoresis*, 35(18):2642–2655, Sep 2014
  12. U. K. Iheagwara, P. L. Beatty, P. T. Van, T. M. Ross, J. S. Minden, and O. J. Finn. Influenza virus infection elicits protective antibodies and T cells specific for host cell antigens also expressed as tumor-associated antigens: a new view of cancer immunosurveillance. *Cancer Immunology Research*, 2(3):263–273, Mar 2014
  13. A. Kaur, P. T. Van, C. R. Busch, C. K. Robinson, M. Pan, W. L. Pang, D. J. Reiss, J. DiRuggiero, and N. S. Baliga. Coordination of frontline defense mechanisms under severe oxidative stress. *Molecular Systems Biology*, 6:393, Jul 2010
  14. T. Koide, D. J. Reiss, J. C. Bare, W. L. Pang, M. T. Facciotti, A. K. Schmid, M. Pan, B. Marzolf, P. T. Van, F. Y. Lo, A. Pratap, E. W. Deutsch, A. Peterson, D. Martin, and N. S. Baliga. Prevalence of transcription promoters within archaeal operons and coding sequences. *Molecular Systems Biology*, 5:285, 2009

15. P. T. Van, A. K. Schmid, N. L. King, A. Kaur, M. Pan, K. Whitehead, T. Koide, M. T. Facciotti, Y. A. Goo, E. W. Deutsch, D. J. Reiss, P. Mallick, and N. S. Baliga. Halobacterium salinarum NRC-1 PeptideAtlas: toward strategies for targeted proteomics and improved proteome coverage. *Journal of Proteome Research*, 7(9):3755–3764, Sep 2008
16. A. K. Schmid, D. J. Reiss, A. Kaur, M. Pan, N. King, P. T. Van, L. Hohmann, D. B. Martin, and N. S. Baliga. The anatomy of microbial cell state transitions in response to oxygen. *Genome Research*, 17(10):1399–1413, Oct 2007
17. M. Blundon, V. Ganesan, B. Redler, P. T. Van, and J. S. Minden. Two-Dimensional Difference Gel Electrophoresis. *Methods Mol. Biol.*, 1855:229–247, 2019