

Phu T. Van, PhD

Bioinformatics Scientist

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

SUMMARY

I'm seeking a scientific or technical position on collaborative teams working to solve data-intensive problems. I'm particularly interested in projects that utilize bioinformatics and/or machine learning.

TECHNICAL STRENGTHS

| | |
|--------------------------|---|
| 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 |
| Markup/Interfaces | LaTeX, Markdown, Shiny, HTML |
| Bioinformatics | STAR, RSEM, BLAST, SAMtools/BCFtools, DEG, GSEA, KEGG, STRING |

EXPERIENCE

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

- Served as Subject Matter Expert on Duplex Sequencing to internal departments and external customers. Advised Field Application Scientists and Tech Support on computational tools and procedures.

| | |
|---|---|
| TwinStrand Biosciences <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.

| | |
|--|--|
| Fred Hutchinson Cancer Research Center <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 <i>Postdoctoral Fellow</i> | October 2014 - October 2015 <i>Seattle, WA</i> |
|---|---|

- 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

PhD Student

August 2009 - May 2014

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

Research Associate

May 2006 - August 2009

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

PhD, Biological Sciences

2014

University of Washington

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

2007

CONFERENCES, PATENTS AND REPRESENTATIVE PUBLICATIONS

1. 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
2. J. Simmons, K. Dill-McFarland, C. Stein, P. Van, Chihota V., Ntshiq T., Maenetje P., Peterson G., Benckek 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. 2021. in review
3. 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
4. J. S. Minden, F. Lanni, and P. T. Van. USPTO US10362237: Structured illumination system for increased dynamic range in quantitative imaging, July 2019
5. 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

6. 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
7. 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
8. 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
9. 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
10. 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
11. 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
12. 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
13. 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