

Vikas R. Pejaver | Postdoctoral scholar

1959 NE Pacific St, E317, Seattle, WA 98195

✉ vpejaver@uw.edu • 🐦 @vikaspejaver

Education

Ph.D. Informatics (Bioinformatics) 2016

Indiana University, Bloomington, IN

Advisor: Predrag Radivojac

Dissertation: Computational methods for understanding the impact of amino acid substitutions on protein function

M.S. Bioinformatics 2010

Indiana University, Bloomington, IN

B.E. Biotechnology 2008

P.E.S. Institute of Technology, Bangalore, India

Research Interests

Computational biology: Impact of variants on protein structure and function as molecular mechanisms of genetic disease.

Biomedical data science: Variant and genome interpretation. Rare and undiagnosed disorder screening. Deep phenotyping from electronic health records and clinical notes.

Machine learning: Positive-unlabeled learning. Model evaluation. Interpretable and explainable learning.

Awards and Honors

2019: Travel Fellowship from CAGI* Workshop: Assessing the Future of Genome Interpretation

2019: NIH Pathway to Independence Award (K99/R00)

2016: Moore/Sloan and Washington Research Foundation Innovation in Data Science Postdoctoral Fellowship, eScience Institute, University of Washington

2014: ISCB Student Council Travel Fellowship

2004–2008: Certificate of distinction, P.E.S. Institute of Technology

Research Experience

Postdoctoral Scholar, Laboratory of Sean Mooney 2017–Present

Department of Biomedical Informatics and Medical Education, and the eScience Institute, University of Washington

Developing data science approaches towards phenotyping patients in electronic health records.

Research Assistant, Laboratory of Predrag Radivojac 2015–2017, 2012–2013

School of Informatics and Computing, Indiana University

Developed methods for the prediction of pathogenicity of missense variants and their effects on protein structure and function.

Research Intern, Bioinformatics & Computational Biology 2011 (Summer)

Genentech

Integrated ChIP-Seq and microarray data for the identification of transcription factor targets relevant to autoimmune diseases.

Research Intern, Bioinformatics Group 2010 (Summer)
Center for Genomics and Bioinformatics, Indiana University
Refined gene models in the *Nasonia vitripennis* genome through the integration of RNA-Seq, EST and tiling array data.

Research Assistant, Laboratories of Sun Kim and Yves Brun 2008–2010
School of Informatics and Computing, Indiana University
Designed and developed methods and software for the detection, visualization and analysis of conserved and functionally related gene clusters across multiple microbial genomes.

Teaching Experience

Instructor, Software/Data Carpentry Workshop 2018
The eScience Institute, University of Washington
Taught session on the Unix shell as a certified Carpentries Instructor

Teaching Assistant, Department of Computer Science and Informatics 2013–2014, 2010–2011
School of Informatics and Computing, Indiana University
Taught lab sessions, led class discussions, designed and graded assignments and mentored undergraduate & graduate students on course material

- INFO-I502: Human-centered Research Methods (Spring 2014).
- INFO-I400/H400/I590: Data Visualization (Fall 2013, Fall 2014).
- CSCI-B555: Machine Learning (Fall 2011).
- INFO-I308: Information Representation (Spring 2011).
- INFO-I519: Introduction to Bioinformatics (Fall 2010).

Undergraduate Teaching Assistant, Department of Biotechnology 2005
P.E.S. Institute of Technology
Prepared class notes and mentored peers on material for the course titled *Momentum Transfer*

Mentoring Experience

Co-mentor, Department of Biomedical Informatics and Medical Education 2017–Present
School of Medicine, University of Washington
Advised students on developing computational methods for biomedical data

- *Yiliang Ma*: Masters Dissertation Project (2018–Present).
- *Chethan Jujjavarapu*: Doctoral Dissertation Project (2017–Present).
- *Robert Minneker*: NLM Biomedical Health Informatics Undergraduate Internship (2019).
- *Samir Yhann*: NLM Biomedical Health Informatics Undergraduate Internship (2018–2019).

Received UW Undergraduate Research Symposium Population Health Recognition Award.

- *Christian Bock*: Masters Dissertation Project (2017).

Student Co-mentor, Department of Computer Science and Informatics 2015–2016
School of Informatics and Computing, Indiana University

Provided guidance to undergraduate & graduate students in conceptualizing and implementing machine learning algorithms for different biological problems

- *Abraham DaSilvio, Bethune-Cookman University: IU Summer Research Opportunities in Computing* (2016).
- *Tyrese Taylor, Bethune-Cookman University: IU Summer Research Opportunities in Computing* (2016).
- *Kranthi Kalva: Masters Capstone Project* (2015–2016).

Publications

Preprints/In Review.....

1. Jujjavarapu, C., **Pejaver, V.**, Cohen, T. A., Mooney, S. D., Heagerty, P. J., and Jarvik, J. G. A comparison of natural language processing methods for the classification of lumbar spine imaging findings related to lower back pain. *In review*.

2. **Pejaver, V.**, Urresti, J., Lugo-Martinez, J., Pagel, K. A., Lin, G. N., Nam, H., Mort, M., Cooper, D. N., Sebat, J., Iakoucheva, L. M., Mooney, S. D., and Radivojac, P. MutPred2: inferring the molecular and phenotypic impact of amino acid variants. *bioRxiv*, 134981, 2017.

- Accepted, *Nature Communications*.

3. **Pejaver, V.**, Jain, R., Cai, B., Bock, C., Bergquist, T., Mort, M., Cooper, D. N., Wang, Y., Adhikari, A., Brenner, S. E., Currier, R. J., Rinaldo, P., Tang, H., Kwok, P., Koenig, B. A., Puck, J. M., Gallagher, R. C., and Mooney, S. D. Evaluation of exome sequencing to improve newborn screening for very long-chain acyl-CoA dehydrogenase deficiency. *In review*.

Reviews/Book Chapters.....

1. Mooney, S. J. and **Pejaver, V.** Big data in public health: terminology, machine learning, and privacy. *Annu. Rev. Public Health*, 39:95–112, 2018.

2. **Pejaver, V. R.**, Lee, H. and Kim, S. Gene cluster prediction and its application to genome annotation. In *Protein function prediction for omics era*, pages 35–54. Springer, 2011.

Published (*Equal contribution; ^Co-corresponding author).....

1. *Bergquist, T., *Yan, Y., Schaffter, T., Yu, T., **Pejaver, V.**, Hammarlund, N., Prosser, J., Guinney, J., and Mooney, S. D. Piloting a model-to-data approach to enable predictive analytics in healthcare through patient mortality prediction. *J. Am. Med. Inform. Assoc.*, *In press*, 2020.

– Covered by **Healthcare IT News** and **Becker’s Hospital Review**.

2. Joshi, R., **Pejaver, V.**, Hammarlund, N., Sung, H., Lee, S. K., Lee, H., Scott, G., Gombar, S., Shah, N., Shen, S., Mooney, S. D., and Pinsky, B. A predictive tool for identification of SARS-CoV-2 PCR-negative emergency department patients using routine test results. *J. Clin. Virol.*, 104502, 2020.

3. Fernandes, J. D., Sarabipour, S., Smith, C. T., Niemi, N. M., Jadavji, N. M., Kozik, A. J., Holehouse, A. S., **Pejaver, V.**, Symmons, O., Filho, A. W. B., and Haage, A. A survey-based analysis of the academic job market. *eLife*, 9(e54097), 2020.

– Preprint version is fourth most downloaded paper of all time in bioRxiv on scientific communication and education

– Covered by Nature in the *Career News* section.

- 4.** Bergquist, T., **Pejaver, V.**, Hammarlund, N., Mooney, S. D., and Mooney, S. J. Evaluation of the secondary use of electronic health records to detect seasonal, holiday-related, and rare events related to traumatic injury and poisoning. *BMC Public Health*, 20(1):46, 2020.
- 5.** Clark, W. T., Kasak, L., Bakolitsa, C., Hu, Z., Andreoletti, G., Babbi, G., ... **Pejaver, V.**, Wang, M., Wei, L., Moulton, J., Yu, G. K., Brenner, S. E., and LeBowitz, J. H. Assessment of predicted enzymatic activity of alpha-N-acetylglucosaminidase (NAGLU) variants of unknown significance for CAGI 2016. *Hum. Mutat.*, 40(9):1519–1529, 2019.
- 6.** Cline, M., Babbi, G., Bonache, S., Cao, Y., Casadio, R., de la Cruz, X., ... **Pejaver, V.**, ... Sun, Y., Topper, S., Parsons, M. T., Spurdle, A. B., Goldgar, D. E., and ENIGMA Consortium. Assessment of blind predictions of the clinical significance of *BRCA1* and *BRCA2* variants. *Hum. Mutat.*, 40(9):1546–1556, 2019.
- 7.** Voskanyan, A., Katsonis, P., Lichtarge, O., **Pejaver, V.**, Radivojac, P., Mooney, S. D., ... Neuhausen, S. L., Ziv, E., Pal, L. R., Andreoletti, G., Brenner, S., and Kann, M. G. Assessing the performance of in-silico methods for predicting the pathogenicity of variants in the gene *CHEK2*, among hispanic females with breast cancer. *Hum. Mutat.*, 40(9):1612–1622, 2019.
- 8.** **Pejaver, V.**, Babbi, G., Casadio, R., Folkman, L., Katsonis, P., Kundu, K., Lichtarge, O., Martelli, P. L., Miller, M., Moulton, J., Pal, L. R., Savojardo, C., Yin, Y., Zhou, Y., Radivojac, P., and Bromberg, Y. Assessment of methods for predicting the effects of PTEN and TPMT protein variants. *Hum. Mutat.*, 40(9):1495–1506, 2019.
- 9.** Rogers, J., Raveendran, M., Harris, R. A., Mailund, T., Leppälä, K., Athanasiadis, G., ... Zinner, D., Roos, C., Jolly, C. J., Gibbs, R. A., Worley, K. C., and **Baboon Genome Analysis Consortium**. The comparative genomics and complex population history of *Papio* baboons. *Sci. Adv.*, 5(1):eaau6947, 2019.
- 10.** *Ismail, W. A., *Pagel, K. A., **Pejaver, V.**, Zhang, S. V., Casasa, S., Mort, M., Cooper, D. N., Hahn, M. W., and Radivojac, P. The sequencing and interpretation of the genome obtained from a Serbian individual. *PLoS One*, 13(12):e0208901, 2018.
- 11.** *Wang, J., ***Pejaver, V. R.**, Dann, G. P., Wolf, M. Y., Kellis, M., Huang, Y., Garcia, B. A., Radivojac, P., and Kashina, A. Target site specificity and *in vivo* complexity of the mammalian arginylome. *Sci. Rep.*, 8(1):16177, 2018.
- 12.** Glusman, G., Rose, P. W., Prlic, A., Dougherty, J., Duarte, J. M., Hoffman, A. S., ... **Pejaver, V.**, ... Reynolds, S. M., Rokem, A., Schwede, T., Song, S., Tilgner, H., Valasatava, Y., Zhang, Y., and Deutsch, E. W. Mapping genetic variations to three-dimensional protein structures to enhance variant interpretation: a proposed framework. *Genome Med.*, 9(113), 2017.
- 13.** Daneshjou, R., Wang, Y., Bromberg, Y., Bovo, S., Martelli, P. L., Babbi, G., ... **Pejaver, V.**, ... Altman, R. B., Klein, T. E., Hoskins, R. A., Repo, S., Brenner, S. E., and Morgan, A. A. Working toward precision medicine: predicting phenotypes from exomes in

the Critical Assessment of Genome Interpretation (CAGI) challenges. *Hum. Mutat.*, 38(9):1182–1192, 2017.

14. Pejaver, V., Mooney, S. D. and Radivojac, P. Missense variant pathogenicity predictors generalize well across a range of function-specific prediction challenges. *Hum. Mutat.*, 38(9):1092–1108, 2017.

15. Reddy, K. D., Malipeddi, J., DeForte, S., Pejaver, V., Radivojac, P., Uversky, V. N., and Deschenes, R. J. Physicochemical sequence characteristics that influence S-palmitoylation propensity. *J. Biomol. Struct. Dyn.*, 35(11):2337–2350, 2017.

16. Pagel, K. A., Pejaver, V., Lin, G. N., Nam, H., Mort, M., Cooper, D. N., Sebat, J., Iakoucheva, L. M., Mooney, S. D., and Radivojac, P. When loss-of-function is loss of function: assessing mutational signatures and impact of loss-of-function genetic variants. *Bioinformatics*, 33(14):i389–i398, 2017.

– **Won Ian Lawson Van Toch Memorial Award for Outstanding Student Paper at ISMB/ECCB 2017.**

17. *Ioannidis, N. M., *Rothstein, J. H., Pejaver, V., Middha, S., McDonnell, S. K., Baheti, S., ... Ostrander, E. A., Bailey-Wilson, J. E., Radivojac, P., Thibodeau, S. N., Whittemore, A. S., and Sieh, W. REVEL: an ensemble score for predicting the pathogenicity of rare non synonymous variants. *Am. J. Hum. Genet.*, 99(4):877–885, 2016.

– **Highlighted in Dr. Russ Altman’s annual AMIA Translational Bioinformatics 2017 Year-in-Review.**

18. Lugo-Martinez, J., Pejaver, V., Pagel, K. A., Jain, S., Mort, M., Cooper, D. N., Mooney, S. D., and Radivojac, P. The loss and gain of functional amino acid residues is a frequent mechanism causing human inherited disease. *PLoS Comput. Biol.*, 12(8):e10005091, 2016.

19. Khanal, N., Pejaver, V., Li, Z., Radivojac, P., Clemmer, D. E., and Mukhopadhyay, S. Position of proline mediates the reactivity of S-palmitoylation. *ACS Chem. Biol.*, 10(11):2529–2536, 2015.

20. Suzuki, H., Dapper, A. L., Jackson, C. E., Lee, H., Pejaver, V., Doak, T. G., Lynch, M., and Preer Jr., J. R. Draft genome sequence of *Caedibacter varicaedens*, a kappa killer endosymbiont bacterium of the ciliate *Paramecium biaurelia*. *Genome Announc.*, 3(6):e01310–15, 2015.

21. Pejaver, V., Hsu, W., Xin, F., Dunker, A. K., Uversky, V. N., and Radivojac, P. The structural and functional signatures of proteins that undergo multiple events of post-translational modification. *Protein Sci.*, 23(8):1077–1093, 2014.

– **Featured article at the front of the issue.**

– **In the journal’s top-10 list of most cited papers in 2014-2015.**

22. Li, Z., Dilger, J. M., Pejaver, V., Smiley, D., Arnold, R. J., Mooney, S. D., Mukhopadhyay, S., Radivojac, P., and Clemmer, D. E. Intrinsic size parameters for palmitoylated and carboxyamidomethylated peptides. *Int. J. Mass Spectrom.*, 368:6–14, 2014.

23. Pejaver, V. R., An, J., Rhee, S., Bhan, A., Choi, J., Liu, B., Lee, H., Brown, P. J., Kysela, D., Brun, Y. V., and Kim, S. GeneclusterViz: a tool for conserved gene cluster visualization, exploration and analysis. *Bioinformatics*, 28(11):1527–1529, 2012.

24. Pejaver, V. R., and Kim, S. Gene Cluster Profile Vectors: a method to infer functionally related gene sets by grouping proximity-based gene clusters. *BMC Genomics*, 12(Suppl 2 - IEEE International Conference on Bioinformatics and Biomedicine 2010):S2, 2011.

25. Werren, J. H., Richards, S., Desjardins, C. A., Niehuis, O., Gadau, J., Colbourne, J. K., ... Pejaver, V. R., ... Wyder, S., Yamada, T., Yi, S. V., Zecher, C. N., Zhang, L., and Gibbs, R. A. Functional and evolutionary insights from the genomes of three parasitoid *Nasonia* species. *Science*, 327(5963):343–348, 2010.

Invited Talks and Guest Lectures

ITHS KL2-TL1 Seminar Series, University of Washington

Seattle, Washington, 2018 and 2020

Department of Biomedical Informatics, Stony Brook University

Stony Brook, New York, 2020

Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mt. Sinai

New York, New York, 2020

Department of Biostatistics and Informatics, University of Colorado

Aurora, Colorado, 2020

Center for Genomics and Bioinformatics Workshop on Microbial Analysis, Indiana University

Bloomington, Indiana, 2011

Conference Presentations

A performance-based approach to establish standards for missense variant impact prediction tools

CAGI Workshop: Assessing the Future of Genome Interpretation, San Francisco, California, 2019*

Predicting the molecular mechanisms of disease-associated amino acid substitutions

Moore-Sloan Data Science Environment Summit, New Orleans, Louisiana, 2017

Probabilistic prediction of the different notions of missense variant impact

Human Genome Variation Society Meeting, Orlando, Florida, 2017

Predicting the in vitro functional effects of natural and synthetic missense mutations

QBI/Convergence Zone Symposium, San Francisco, California, 2017

MutPred2: predicting the pathogenicity and molecular consequences of missense variants

ISCB Student Council Symposium, Boston, Massachusetts, 2014

Professional Activities

Invited grant reviewer: Advanced Scientific Computing Research Program, Department of Energy, 2020

Chair: Gordon Research Seminar on Human Genetic Variation and Disease, 2018

Organizer:

- CAGI Trainee Salon, CAGI* Workshop: Assessing the Future of Genome Interpretation, 2019
- Breakout session - "To what extent can biomedical and health data be made FAIR?", Moore-Sloan Data Science Environment Summit, 2018

Panelist: Future of genome interpretation and vision for CAGI panel, CAGI* Workshop: Assessing the Future of Genome Interpretation, 2019

Program committee member: International Conference on Intelligent Systems for Molecular Biology, 2018

Reviewer:

- AMIA Annual Symposium, 2020
- BioData Mining
- Bioinformatics
- BMC Bioinformatics
- Genetics and Molecular Biology
- Human Mutation
- IEEE/ACM Transactions on Computational Biology and Bioinformatics
- IEEE Journal of Biomedical and Health Informatics
- International Journal of Epidemiology
- Intrinsically Disordered Proteins
- Nature Computational Science
- Nucleic Acids Research
- Pacific Symposium on Biocomputing (specific sessions)
 - 2019: Precision medicine: improving health through high-resolution analysis of personal data
 - 2017: Computational approaches to understanding the evolution of molecular function
- Patterns
- PeerJ
- PLOS Computational Biology
- PLOS Genetics
- PLOS ONE
- Proceedings of the National Academy of Sciences
- SSM - Population Health
- Translational Research

Member:

- Academic Data Science Alliance (ADSA)
- American Medical Informatics Association (AMIA)
- American Society of Human Genetics (ASHG)
- International Society for Computational Biology (ISCB)