

J. Harry Caufield

Los Angeles, CA, USA

✉ jcaufield@mednet.ucla.edu

☎ 1-484-467-6685

🏠 harrycaufield.net

EMPLOYMENT

David Geffen School of Medicine at UCLA

Assistant Project Scientist and Program Coordinator, Department of Physiology

Los Angeles, CA, USA

2020–Present

EDUCATION AND TRAINING

David Geffen School of Medicine at UCLA

Postdoctoral Fellow, Department of Physiology

Los Angeles, CA, USA

2017–2020

Virginia Commonwealth University

PhD in Integrative Life Sciences

Richmond, VA, USA

2012–2016

Virginia Commonwealth University

MS in Bioinformatics

Richmond, VA, USA

2009–2012

University of Delaware

BS in Biological Sciences

Newark, DE, USA

2004–2008

RESEARCH EXPERIENCE

Information extraction from unstructured biomedical data

University of California, Los Angeles

Los Angeles, CA, USA

Jan. 2017–Present

Developing data science, machine learning, information extraction, and natural language processing approaches for working with structured and unstructured biomedical data. Building ontologies and annotation systems for concepts and relationships in clinical case reports. Designing applications concerning cardiovascular disease, cardiac proteins, and knowledge graph approaches. Supporting operation of the Integrated Data Science Training in CardioVascular Medicine (iDISCOVER) program. Supervised by Profs. Peipei Ping (Depts. of Physiology, Medicine/Cardiology, and Bioinformatics) and Wei Wang (Depts. of Computer Science and Computational Medicine).

Microbial interactomics

Virginia Commonwealth University

Richmond, VA, USA

Dec. 2011–Dec. 2016

Identified, compared, and interpreted protein-protein interactions among microbial and viral proteomes using bioinformatics (e.g., sequence alignments, phylogenetics, and network analysis) and high-throughput molecular biology methods (e.g., yeast two-hybrid). Also served as lab manager. Supervised by Dr. Peter Uetz (VCU Center for the Study of Biological Complexity).

Dissertation: Interactomics-Based Functional Analysis: Using Interaction Conservation To Probe Bacterial Protein Functions.

Ribosomal proteins and bacteriophage genetics in *Staphylococcus aureus*

Virginia Commonwealth University

Richmond, VA, USA

Oct. 2009–May 2012

Investigated the potential for post-translational modification of a ribosomal protein sequence broadly conserved bacterial species and its relation to bacteriophage infection. Employed bioinformatics through genome and protein sequence analysis and phylogenetics. Work supervised by Dr. Gail Christie (VCU Department of Microbiology and Immunology).

PUBLICATIONS

- Mehla, J., Liechti, G., Morgenstein, R., **Caufield, J.H.**, Hosseinnia, A., Gagarinova, A., Phanse, S., Goodacre, N., Brockett, M., Sakhawalkar, N., Babu, M., Xia, R., Montelione, G., Vorobiev, S., den Blaauwen, T., Hunt, J., & Uetz, P. (2021) YhcB (DUF1043), a novel cell division protein conserved across gamma-proteobacteria. **Under revision**.
- **Caufield, J.H.**, Sigdel, D., Choi, H., Fu, J., & Ping, P. (2021) Cardiovascular Informatics: bridging the gap. **Cardiovascular Research**. **Under revision**.
- Zhou, Y., Yan, Y., Han, R., **Caufield, J. H.**, Chang, K.-W., Sun, Y., Ping, P., & Wang, W. (2021) Clinical Temporal Relation Extraction with Probabilistic Soft Logic Regularization and Global Inference. **AAAI 2021 (accepted)**.
- Wang, X., Guan, Y., Liu, W., Chauhan, A., Jiang, E., Li, Q., Liem, D., Sigdel, D., **Caufield, J.H.**, Ping, P., & Han, J. (2020) EVIDENCEMINER: Textual Evidence Discovery for Life Sciences. **ACL 2020 (Demonstration Track)**, 56-62. doi: 10.18653/v1/2020.acl-demos.8
- **Caufield, J. H.**, Zhou, Y., Bai, Y., Liem, D. A., Garlid, A. O., Chang, K.-W., Sun, Y., Ping, P., & Wang, W. (2019) A Comprehensive Typing System for Information Extraction from Clinical Narratives. **medRxiv [Preprint]**. doi: 10.1101/19009118
- **Caufield, J. H.** & Ping, P. (2019) New advances in extracting and learning from protein–protein interactions within unstructured biomedical text data. **Emerging Topics in Life Sciences**, 3(4), 357–369. doi: 10.1042/ETLS20190003
- **Caufield, J. H.**, Zhou, Y., Garlid, A. O., Liem, D. A., Cao, Q., Lee, J. M., Murali, S., Spendlove, S., Wang, W., Zhang, L., Sun, Y., Han, J., Watson, K. A., & Ping, P. (2018). A reference set of curated biomedical data and metadata from clinical case reports. **Scientific Data**, 5, 180258. doi: 10.1038/sdata.2018.258
- **Caufield, J. H.**, Liem, D. A., Garlid, A. O., Watson, K., Bui, A. T., Wang, W., & Ping, P. (2018). A Metadata Extraction Approach for Clinical Case Reports to Enable Advanced Understanding of Biomedical Concepts. **Journal of Visualized Experiments**, (139), e58392. doi:10.3791/58392
- Liem, D., Murali, S., Sigdel, D., Shi, Y., Wang, X., Shen, J., Choi, H., **Caufield, J. H.**, Wang, W., Ping, P., & Han, J. (2018). Phrase Mining of Textual Data to Analyze Extracellular Matrix Protein Patterns Across Cardiovascular Disease. **AJP - Heart and Circulatory Physiology**, 315(4), H910–H924. doi: 10.1152/ajpheart.00175.2018
- Mehla, J., **Caufield, J. H.**, & Uetz, P. (2018). Making the Right Choice: Critical Parameters of the Y2H Systems. In: Oñate-Sánchez (editor), Two-Hybrid Systems: Methods and Protocols, **Methods in Molecular Biology**, vol. 1794, pp. 17-28. doi:10.1007/978-1-4939-7871-7_2
- Wuchty, S., Müller, S. A., **Caufield, J. H.**, Häuser, R., Aloy, P., Kalkhof, S., & Uetz, P. (2018). Proteome Data Improves Protein Function Prediction in the Interactome of *Helicobacter pylori*. **Molecular & Cellular Proteomics**, 17(5), 961–973. doi:10.1074/mcp.RA117.000474
- Babu, M., Bundalovic-Torma, C., Calmettes, C., ... **Caufield, J. H.**, et al. (2018). Global Landscape of Cell Envelope Protein Complexes in *Escherichia coli*. **Nature Biotechnology**, 36(1), 103–112. doi:10.1038/nbt.4024
- Mehla, J., Dedrick, R. M., **Caufield, J. H.**, Wagemans, J., Sakhawalkar, N., Johnson, A., Hatfull, G. F., & Uetz, P. (2017). Virus-host protein-protein interactions of mycobacteriophage Giles. **Scientific Reports**, 7(1), 16514. doi:10.1038/s41598-017-16303-7
- Sakhawalkar, N., **Caufield, J. H.**, & Uetz, P. (2017). Two-hybrid and Related Systems. In: **eLS**. 1–9. John Wiley & Sons, Ltd. doi:10.1002/9780470015902.a0000981.pub2
- **Caufield, J. H.**, Wimble, C., Shary, S., Wuchty, S., & Uetz, P. (2017). Bacterial Protein Meta-Interactomes Predict Cross-Species Interactions and Protein Function. **BMC Bioinformatics**, 18(1), 171.

doi:10.1186/s12859-017-1585-0

- Mehla, J., **Caufield, J. H.**, Sakhawalkar, N., & Uetz, P. (2017). A Comparison of Two-Hybrid Approaches for Detecting Protein–Protein Interactions. In: Shukla, A. K. (editor), **Methods in Enzymology** 586: Proteomics in Biology, Part B. Elsevier. doi:10.1016/bs.mie.2016.10.020
- Mehla, J., Dedrick, R., **Caufield, J. H.**, Siefiring, R., Mair, M., Johnson, A., Hatfull, G., & Uetz, P. (2015). The protein interactome of mycobacteriophage Giles predicts functions for unknown proteins. **Journal of Bacteriology**, 197(15), 2508-16. doi:10.1128/JB.00164-15
- **Caufield, J. H.**, Abreu, M., Wimble, C., & Uetz, P. (2015). Protein complexes in Bacteria. **PLoS Computational Biology**, 11(2), e1004107. doi:10.1371/journal.pcbi.1004107
- Mehla, J., **Caufield, J. H.**, & Uetz, P. (2015). The Yeast Two Hybrid system: A tool for mapping protein-protein interactions. In: Boone, C. (editor), **Budding Yeast: A Laboratory Manual**. CSHL Press.
- Mehla, J., **Caufield, J. H.**, & Uetz, P. (2015). Mapping Protein–Protein Interactions Using Yeast Two-Hybrid Assays. In: Boone, C. (editor), **Budding Yeast: A Laboratory Manual**. CSHL Press.
- Wall, E., **Caufield, J.**, Manning, K., Dokland, T., Christie, G. (2015). Specific N-terminal cleavage of ribosomal protein L27 in *Staphylococcus aureus* and related bacteria. **Molecular Microbiology**, 95(2), 258-69. doi:10.1111/mmi.12862
- Rajagopala, S. V., Sikorski, P., **Caufield, J. H.**, Tovchigrechko, A., & Uetz, P. (2012). Studying protein complexes by the yeast two-hybrid system. **Methods**, 58(4), 392-9. doi:10.1016/j.ymeth.2012.07.015
- **Caufield, J. H.**, Sakhawalkar, N., & Uetz, P. (2012). A comparison and optimization of yeast two-hybrid systems. **Methods**, 58(4), 317-24. doi:10.1016/j.ymeth.2012.12.001

See Google Scholar profile at: [goo.gl/9E2ZTB](https://scholar.google.com/citations?user=goo.gl/9E2ZTB)

CONFERENCE PRESENTATIONS

Oral Presentations

- Exploring informatics approaches to better understand pathogenic pathways underlying COVID-19. **AHA BCVS Scientific Sessions 2020**. Online.
- Building Metadata to Render Clinical Case Reports FAIR. **ISMB/ECCB 2019**. Basel, Switzerland.
- Computational Approaches for Learning from Massive, Unstructured Biomedical Text Data. **Los Angeles Postdoctoral Research Symposium 2019**. Los Angeles, CA, USA.
- Connecting the dots – digital objects in cardiovascular diseases: Meta-Interactomics. **Proteomic Forum 2017**. Potsdam, Germany.
- Meta-Interactomics and the Evolution of Bacterial Protein Complexes. **US HUPO Annual Conference 2017**. San Diego, CA, USA.
- Predicted protein complexomes and their use in studying bacteriophage-mediated protein interactions. **VCU CSBC Research Review 2014**. Richmond, VA, USA.
- Predicted protein complexomes and their use in studying bacteriophage-mediated protein interactions. **Virginia Branch American Society for Microbiology meeting 2014**. Harrisonburg, VA, USA. *Won 2nd prize for Best Student Presentation out of 19 presenters.*
- Protein Complex Conservation across Bacterial Species and Genome Characteristics. **VCU Forbes Day Research Symposium 2014**. Richmond, VA, USA.
- Let's Do Some Networking: Interactomics across domains of life. **UVA German Graduate Conference 2014**. Charlottesville, VA, USA.

- The Impact of Genome Reduction on Conservation of Microbial Protein Complexes and Their Components. **Virginia Branch American Society for Microbiology meeting 2013**. Charlottesville, VA, USA. *Won 3rd prize for Best Student Presentation out of 17 presenters.*
- Protein-protein interaction networks of bacteriophage using the yeast two-hybrid system. **Virginia Branch American Society for Microbiology meeting 2012**. Virginia Beach, VA, USA.
- Potential N-terminal processing of ribosomal protein L27 in *Staphylococcus aureus*. **Virginia Branch American Society for Microbiology meeting 2010**. Lynchburg, VA, USA.

Poster Presentations

- **Caufield, J. H.**, Zhou, Y., Bai, Y., Liem, D.A., Garlid, A., Chang, K-W., Sun, Y., Wang, W., & Ping, P. Resources for comprehensive information extraction and data integration in biomedicine. **ISMB 2020**. Online.
- **Caufield, J. H.**, Zhou, Y., Zinsser, M., Zheng, H., Garlid, A. O., Bai, Y., Zhou, Y., Cao, Q., Lee, J. M., Murali, S., Spendlove, S., Liem, D. A., Chang, K-W., Sun, Y., Wang, W., & Ping, P. Resources for comprehensive information extraction and data integration in biomedicine. **IDRE Early Career Researcher Day 2019**. Los Angeles, CA, USA.
- **Caufield, J. H.**, Zhou, Y., Garlid, A. O., Liem, D. A., Cao, Q., Lee, J. M., Murali, S., Spendlove, S., & Ping, P. Learning Structured Knowledge from Clinical Case Reports. **ISMB/ECCB 2019**. Basel, Switzerland.
- Zhou, Y., **Caufield, J. H.**, Sigdel, D., Liem, D. A., Sun, Y., Chang, K-W., Ping, P., & Wang, W. Metadata Acquired from Clinical Case Reports: a resource for extracting information from clinical narratives. **ISMB/ECCB 2019**. Basel, Switzerland.
- **Caufield, J. H.**, Zhou, Y., Garlid, A. O., Liem, D. A., Cao, Q., Lee, J. M., Murali, S., Spendlove, S., Wu, J., Komya, K., Wang, W., & Ping, P. Transforming clinical case reports into structured data with coordinated annotation and analysis. **International Conference on Systems Biology of Human Diseases 2018**. Los Angeles, CA, USA.
- **Caufield, J. H.**, Zhou, Y., Garlid, A. O., Liem, D. A., Cao, Q., Lee, J. M., Murali, S., Spendlove, S., & Ping, P. Automated and manual biomedical literature analysis unlocks new resources for integrating biomedical data with clinical observations. **UCLA Cardiovascular Theme Symposium 2018**. Los Angeles, CA, USA.
- **Caufield, J. H.** & Ping, P. Implications for cardiac remodeling among the interactions of 14-3-3 proteins. **NHLBI Systems Biology Symposium 2017**. Bethesda, MD, USA.
- **Caufield, J. H.**, Komya, K., Wu, J., Spendlove, S., Zhou, Y., Cao, Q., Liem, D., Wang, W., & Ping, P. Clinical case report analysis provides the foundation for new approaches to biomedical data integration. **NHLBI Systems Biology Symposium 2017**. Bethesda, MD, USA.
- **Caufield, J. H.**, Komya, K., Wu, J., Nanayakkara, P., Zhou, Y., Cao, Q., Liem, D., & Ping, P. HeartCases - a medical case report analysis system for understanding cardiovascular disease. **ISMB/ECCB 2017**. Prague, Czech Republic.
- **Caufield, J. H.** & Uetz, P. Bacterial Proteins and Complexes: Evolution, Function, and Perturbations. **American Society for Microbiology Microbe 2016**. Boston, MA, USA.
- **Caufield, J. H.** & Uetz, P. Meta-interactomics provides functional context for microbial proteins and their interactions. **VCU Institute for Structural Biology, Drug Discovery and Development Summit 2016**. Richmond, VA, USA. *Won 2nd prize for Best Student Presentation out of 8 student presenters.*
- **Caufield, J. H.** & Uetz, P. Meta-interactomics provides functional context for microbial proteins and their interactions. **VCU Integrative Life Sciences Student Research Showcase 2015**. Richmond, VA,

USA.

- **Caufield, J. H., Abreu, M., Wimble, C., & Uetz, P.** Predicted protein complexomes and their use in studying bacteriophage-mediated protein interactions. **VCU Integrative Life Sciences Student Research Showcase 2014.** Richmond, VA, USA.
- **Caufield, J. H., Abreu, M., Wimble, C., Goodacre, N., & Uetz, P.** From protein interaction domains to complexes to networks: computational approaches. **VCU Graduate Student Association Research Symposium 2014.** Richmond, VA, USA.
- **Caufield, J. H., Abreu, M., Wimble, C., Goodacre, N., & Uetz, P.** From protein interaction domains to complexes to networks: computational approaches. **VCU Integrative Life Sciences Student Research Showcase 2013.** Richmond, VA, USA.
- **Caufield, J. H., Abreu, M., Wimble, C., Goodacre, N., & Uetz, P.** The Impact of Genome Reduction on Conservation of Microbial Protein Complexes and Their Components. **Virginia Branch American Society for Microbiology meeting 2013.** Charlottesville, VA, USA.
- **Caufield, J. H., Abreu, M., Wimble, C., Goodacre, N., & Uetz, P.** From protein interaction domains to complexes to networks: computational approaches. **VCU Watts Day Symposium 2013.** Richmond, VA, USA.
- **Caufield, J. H., Dedrick, R. M., del Rosario, N. G., Wagemans, J., Hatfull, G., & Uetz, P.** Phage Interactomics: Identification of Phage-Phage and Phage-Host Protein Interactions by Systematic Yeast Two-Hybrid Screens. **VCU Watts Day Symposium 2012.** Richmond, VA, USA.
- **Caufield, J. H., Dedrick, R. M., del Rosario, N. G., Wagemans, J., Hatfull, G., & Uetz, P.** Phage Interactomics: Identification of Phage-Phage and Phage-Host Protein Interactions by Systematic Yeast Two-Hybrid Screens. **SEA-PHAGES meeting 2012.** Ashburn, VA, USA.

GRANTS, AWARDS, & HONORS

- NIH/NHLBI T32 HL139450 (PI: Ping) *Sep. 2019–Aug. 2020*
iDISCOVER: Integrated Data Science Training in Cardiovascular Medicine
Role: Trainee
- UCLA Institute for Digital Research and Education (IDRE) Postdoctoral Fellowship *2019*
Amount: \$5,000
- Virginia Commonwealth University Graduate School Dissertation Assistantship *Fall 2016*
Amount: \$17,250 + tuition support

TEACHING

- **Teaching Assistant and Instructor - Bioinformatics 201** **Los Angeles, CA, USA**
University of California, Los Angeles *Jan. 2019–Present*
Currently assisting with organization and teaching of a discussion course for 20 to 30 Bioinformatics graduate students quarterly. I guide student interactions as they describe and evaluate recently published findings covering a wide variety of computational methods in the biomedical sciences. I also present lectures and workshops on informatics topics.
- **Teaching Assistant - Life Science 301** **Richmond, VA, USA**
Virginia Commonwealth University *Aug. 2015–Dec. 2015*
Assisted with Integrative Life Sciences Research course for undergraduate students in the life sciences. Led weekly discussion session involving topics relevant to research conduct, including ethics, effective

communication, and experimental design.

- **Teaching Assistant - Bioinformatics 541**

Virginia Commonwealth University

Richmond, VA, USA

Jan. 2015–May 2015

Assisted with Molecular Genetics Lab course for undergraduate and graduate students. Facilitated wet lab exercises and prepared lab materials. Aided class protocol design. Graded tests and quizzes. Lectured on 16S rRNA gene-specific primer design. Advised students on scientific writing.

- **Teaching Assistant - Bioinformatics 251**

Virginia Commonwealth University

Richmond, VA, USA

Aug. 2013–Dec. 2013

Assisted with Phage Discovery Lab course for undergraduates. Facilitated wet lab exercises and prepared lab materials. Prior to assisting with course, attended Phage Phest student symposia in Spring 2012 and Spring 2013.

- **Teaching Assistant - Bioinformatics 301**

Virginia Commonwealth University

Richmond, VA, USA

Jan. 2010–May 2010

Assisted with introductory bioinformatics course for undergraduates. Advised students on introductory bioinformatics and programming concepts using BioBike, a visual programming language.

MENTORING

Completed 8 hours of evidence-based mentor training based on the *Entering Mentoring* series (Pfund, Branchaw, Handelsman 2014).

- **Mentor in Ping Lab**

University of California, Los Angeles

Los Angeles, CA, USA

Jan. 2017–Present

Have directly supervised and served as mentor to seventeen undergraduate students and interns on research and software development projects involving text analysis and biomedical data mining.

- **Mentor in Uetz Lab**

Virginia Commonwealth University

Richmond, VA, USA

May 2012–Dec. 2016

Served as mentor to three graduate and nine undergraduate students.

SERVICE & OUTREACH

Reviewing

I have provided reviews for the following scholarly journals and academic conferences:

- **Bioinformatics Community Conference (BCC) 2020** (Open Bioinformatics Foundation)
- **iScience** (Cell Press)
- **Scientific Reports** (Nature)
- **International Journal of Molecular Sciences (IJMS)** (MDPI)
- **Materials** (MDPI)

Events, Committees, & Organizations

- UCLA Medical Informatics Curriculum Committee

July 2020–Present

Assist with development and application of guidelines for the Medical Informatics graduate program curriculum and exams. Mentor program students and provide guidance regarding requirements.

- Data Science Events at ISMB 2020 *Jan. 2020–July 2020*
Led planning of biomedical data science outreach efforts, including a day-long track of talks and panel discussions, to be hosted as part of the 2020 Intelligent Systems for Molecular Biology conference. Worked closely with the NIH Office of Data Science Strategy to plan events.
- UCLA Physiology Trainee Seminar Series Planning Committee *Dec. 2019–March 2020*
Coordinated an interdepartmental seminar series highlighting the research of UCLA students, post-docs, and project scientists in physiology.
- BD2K Activities at ISMB/ECCB 2019 *July 2019*
Assisted with management of outreach efforts for the NIH Big Data to Knowledge initiative (though the BD2K Centers – Coordination Center, BD2KCCC) during the 5 days of the 2019 Intelligent Systems for Molecular Biology and European Conference on Computational Biology joint conference in Basel, Switzerland. Activities included an outreach table, a Special Track with 11 speakers providing talks on data sharing and open science, and several discussion panels.
- BD2K Activities at ISMB/ECCB 2017 *July 2017*
Coordinated and managed outreach efforts for the NIH Big Data to Knowledge initiative (though the BD2K Centers – Coordination Center, BD2KCCC) during the 5 days of the 2017 Intelligent Systems for Molecular Biology and European Conference on Computational Biology joint conference in Prague, Czech Republic. Activities included an outreach table, a Special Track with 18 speakers providing talks on machine learning and data science, and a discussion session.
- VCU Integrative Life Science Student Organization (ILSSO) *Sep. 2013–Dec. 2016*
Founding member of student organization. Served on executive board. Primary organizer for the 2014 ILS Research Showcase, a research meeting with 21 presenters and more than 70 attendees. Assisted with organization of the 2015 Research Showcase. Vice President for the 2015 - 2016 academic year.

PROFESSIONAL MEMBERSHIPS & EXPERIENCE

- American Heart Association (AHA) – Member *Fall 2018 – Present*
- International Society for Computational Biology (ISCB) – Member *Spring 2017 – Present*
- American Society for Microbiology (ASM), Virginia Branch – Member *Fall 2010 – Fall 2016*
- American Society for Microbiology (ASM) – Member, Division M *Fall 2010 – Fall 2016*

WORKSHOPS, SHORT COURSES, & OTHER MEETINGS ATTENDED

- Academic Job Search Intensive Workshop *Los Angeles, CA, USA*
UCLA Career Center *Sep. 2019*
- CollaborationFest (CoFest) 2019 *Basel, Switzerland*
Open Bioinformatics Foundation *July 2019*
- K Grant Writing Workshop *Los Angeles, CA, USA*
UCLA Clinical and Translational Science Institute *May 2018*
- HUPO-PSI Spring 2018 meeting *Heidelberg, Germany*
HUPO Proteomics Standards Initiative *Apr. 2018*
- CIAPM – third annual meeting *San Diego, CA, USA*

California Initiative to Advance Precision Medicine

- **Stable and Transient Protein-Protein Interactions**
USHUPO 2017 workshop
- **Scientific Writing and Publishing Institute (SWPI) series**
American Society for Microbiology

Oct. 2017

San Diego, CA, USA

Mar. 2017

Online

Jan. 2015–Mar. 2015