

CURRICULUM VITAE

CATHY HUEY-HWA WU, Ph.D., Fellow of ACM and IEEE

Unidel Edward G. Jefferson Chair in Engineering and Computer Science

Founding Director, Data Science Institute (DSI)

Founding Director, Center for Bioinformatics & Computational Biology (CBCB)

University of Delaware

EXECUTIVE SUMMARY

Scholarship & Leadership

- **Contributions to Field of Study:** Conducted Bioinformatics Data Science research for 30 years, contributing to the foundational knowledge of bioinformatics, computational biology, and data science with broad impact to biomedicine and society
- **Research Citations:** Published >290 peer-reviewed publications, with >48,000 citation and h-index of 71 on Google Scholar; Recognized as a “Highly Cited Researcher” (top 1%) for 7 consecutive years (2014-2020)
- **Principal Awards:** IEEE (Institute of Electrical and Electronics Engineers) Fellow, 2022; ACM (Association for Computing Machinery) Fellow, 2021; AAIA (Asia-Pacific Artificial Intelligence Association) Fellow, 2021; Purdue Distinguished Agriculture Alumni Award, 2020
- **Grants and Team Science:** (1) Continuously funded since 1987 from federal and other agencies, totaling over 70 grants and \$60 million funding support for her lab, including \$6M new grants in 2021-2026; (2) Led UD data science initiatives since 2018 to accelerate cross-campus collaborations and multi-institutional team science grant applications, resulting in grant awards totaling \$8.6M and \$20M pending review
- **International Research Resource:** Co-founded the international UniProt Consortium in 2002, with continuous NIH funding through 2026 as MPI to provide a central hub of protein knowledge (with over 8 million pageviews monthly from a million unique users worldwide) for foundational understanding of biological systems, drug discovery, and mapping of personalized genome
- **National Leadership and Boards:** Council of Councils Working Group for the Common Fund Data Ecosystem, NIH (2021-); NIH Advisory Council, National Institute of General Medical Sciences (2017-2021); IEEE/ACM TCBB Steering Committee (2018-) and Associate Editor (2014-), ACM SIGBio Board of Directors (2010-), Board on Research Data and Information (BRDI), National Research Council (NRC) (2008-2011)

Training & Mentoring

- **Graduate Programs:** (1) Founded the Center for Bioinformatics and Computational Biology (CBCB) in 2010 and serves as the Founding Director of Bioinformatics Data Science graduate programs, including MS/PSM programs in Bioinformatics and Computational Biology (2010), PhD program in Bioinformatics Data Science (2012), Graduate Certificate in Bioinformatics (2012), Online Graduate Certificate in Applied Bioinformatics (2017), and Online Graduate Certificate in Biomedical Informatics and Data Science (2019), to train the next-generation of researchers and professionals at the interface of computational sciences and life sciences; (2) Brings together 11 Data Science-related degree and certificate programs and over 70 data science-related courses across campus as the DSI Director
- **Mentoring:** (1) Mentored more than 200 mentees throughout academic career - undergraduate and graduate students, post-graduate trainees, junior scientists, and young investigators who have received many awards and fellowships and have successful careers in academic, industry, and government organizations; (2) mentored 13 resident faculty recruited through campus-wide data science clustering hiring as the DSI Director
- **Diversity, Equity, and Inclusion (DEI):** (1) Women in Computing: helped launching the ACM-W (Women in Computing) chapter at UD, and co-hosted the Women in Bioinformatics panels at several annual ACM conferences; (2) A strong advocate for DEI, mentor in the National Research Mentoring Network (NRMN), and has received training in Culturally Aware Mentoring

OFFICE ADDRESS

Center for Bioinformatics & Computational Biology
University of Delaware
Ammon Pinizzotto Biopharmaceutical (APBio) Innovation Center
590 Avenue 1743, Suite 147
Newark, DE 19713
Phone: (302) 831-8869
Email: wuc@udel.edu
DSI home : <https://dsi.udel.edu/>
CBCB home: <https://bioinformatics.udel.edu/>

TABLE OF CONTENT

EDUCATION	3
POSITIONS AND EMPLOYMENT.....	3
PROFESSIONAL APPOINTMENTS / HONORS / SOCIETIES.....	3
PROFESSIONAL SERVICES	6
Conference Organizing Committees (>70)	6
Study Sections / Grant Review Panels / Site Visit Teams / Academic Reviews (>80)	9
INVITED LECTURES / PANELS (>200).....	12
UNIVERSITY SERVICE.....	19
TEACHING.....	20
Degree Programs Developed.....	21
Courses Taught	21
MENTORING.....	22
RESEARCH GRANTS (>70).....	28
PUBLICATIONS AND PRODUCTS.....	37
Products / Research Resources.....	37
Books / Conference Proceedings / Journal Special Issues (15).....	37
Refereed Papers in Journals and Conference Proceedings (>290)	38
1) Original Papers in Refereed Journals.....	38
2) Reviews or Editorials in Refereed Journals.....	49
3) Publications in Refereed Conference Proceedings and Book Chapters	50

EDUCATION

Institution	Degree	Date	Field
National Taiwan University, Taiwan	B.S.	1974-1978	Plant Pathology
Purdue University, W. Lafayette, IN	M.S.	1980-1982	Plant Pathology
Purdue University, W. Lafayette, IN	Ph.D.	1982-1984	Plant Pathology/Molecular Biology
Michigan State University, E. Lansing, MI	Postdoc.	1985-1986	Molecular Biology
University of Texas at Tyler, Tyler, TX	M.S.	1988-1989	Computer Science

POSITIONS AND EMPLOYMENT

2018-Present	Founding Director, Data Science Institute (DSI), University of Delaware (UD), Newark, Delaware
2017-Present	Unidel Edward G. Jefferson Chair in Engineering and Computer Science; Founding Director, Online Graduate Certificate in Applied Bioinformatics (2017-Present) & Online Graduate Certificate in Biomedical Informatics and Data Science (2019-Present), UD
2013-Present	Affiliate Faculty, Institute of Financial Service Analytics (IFSA), UD
2012-Present	Founding Director, PhD Program in Bioinformatics and Systems Biology (renamed as PhD Program in Bioinformatics Data Science in 2021), UD
2010-Present	Founding Director, Master of Science (MS) Program in Bioinformatics and Computational Biology, Professional Science Master's (PSM) Program in Bioinformatics, and Graduate Certificate in Bioinformatics, UD
2010-Present	Professor, Biomedical Engineering Program, UD
2009-Present	Founding Director of Center for Bioinformatics & Computational Biology, UD
2009-Present	Edward G. Jefferson Chair of Bioinformatics and Computational Biology; Professor of Department of Computer and Information Sciences; Professor of Department of Biological Sciences; Fellow, Delaware Biotechnology Institute (DBI), UD
2008-2010	Founding Co-Director, Bioinformatics Track, M.S. Degree in Biochemistry and Molecular Biology, Georgetown University Medical Center (GUMC), Washington, DC
2002-Present	Adjunct Professor (2009-Present), Professor (2001-2008), Department of Oncology; Member, Lombardi Comprehensive Cancer Center, GUMC
2001-Present	Adjunct Professor (2009-Present), Professor (2001-2008), Department of Biochemistry and Molecular & Cellular Biology, GUMC
2001-Present	Director, Protein Information Resource (PIR), Georgetown University (2001-Present) and University of Delaware (2009-Present)
1999-2002	Director of Bioinformatics, Protein Information Resource (1999-2001), Board of Directors (1999-2002), Vice President (2000-2002), National Biomedical Research Foundation, Washington, DC
1990-1999	Professor (1998-1999), Associate Professor (1994-1998), Assistant Professor (1990-1994), Department of Epidemiology and Biomathematics, University of Texas Health Center at Tyler, Texas
1989-1994	Associate Professor (1994), Assistant Professor (1989-1994), Department of Computer Science, University of Texas at Tyler, Texas
1986-1987	Research Scientist, Department of Plant Pathology & Microbiology, Texas A&M University, Texas
1985-1986	Postdoctoral Fellow, MSU-DOE Plant Research Laboratory, Michigan State University (Advisor: Christopher R. Somerville, Member, National Academy of Sciences), Michigan
1980-1984	Graduate Research Assistant, Department of Botany & Plant Pathology, Purdue University, IN

PROFESSIONAL APPOINTMENTS / HONORS / SOCIETIES

Advisory Boards/Committees

2021-Present Council of Councils (CoC) Working Group for the Common Fund Data Ecosystem, National Institutes of Health (NIH)

2021-Present Board of Regents Comparative Genomics Resource Working Group, National Library of Medicine (NLM), NIH

2020 Board of Scientific Counselors (Ad hoc member), National Library of Medicine (NLM), NIH

2019-Present External Advisory Committee, Center of Quantitative Biology COBRE (Centers of Biomedical Research Excellence), Geisel School of Medicine at Dartmouth, NH

2018-Present Steering Committee, IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)

2017-2021 Advisory Council, NIGMS (National Institute of General Medical Sciences) (2017-2021); Search Committee, Division Director of BBCB (Biophysics, Biomedical Technology, and Computational Biosciences) (2020), NIGMS, NIH

2015-Present Advisory Board, NIAID (National Institute of Allergy and Infectious Diseases) Bioinformatics Integration Support Contract (BISC), NIH

2014 External Scientific Panel, Library of Integrated Network-Based Cellular Signatures (LINCS), Common Funds program, NIH

2014 Ad hoc Member, Advisory Council, NIGMS (National Institute of General Medical Sciences), NIH

2013-2015 Informatics Advisory Committee, Joint Genome Institute (JGI), U.S. Department of Energy (DOE)

2012 External Advisory Panel, NHLBI Proteomics Program, National Heart, Lung, and Blood Institute (NHLBI), NIH

2010-Present Board of Directors, SIGBio (2010-Present), Distinguished Member (2017-2021), Fellow (2021-present), Association for Computing Machinery (ACM)

2010-2013 Scientific Advisory Board, Data-Intensive Academic Grid (DIAG)

2009-2010 NSF Advisory Committee for Cyberinfrastructure Task Force on Grand Challenges, National Science Foundation (NSF)

2008-2011 Board on Research Data and Information (BRDI), National Research Council (NRC)

2008-2011 U.S. National Committee for CODATA (Committee on Data for Science and Technology)

2008-2018 Board of Directors; Chair, Bioinformatics and Biostatistics Subcommittee of US HUPO Initiative and Executive Committee (2010-2015), US Human Proteome Organization (US HUPO)

2006-2010 TeraGrid Scientific Advisory Board, National Science Foundation (NSF)

2005-2014 Council (2012-2014; 2005-2008), Human Proteome Project (HPP) Working Group (2009-2014), Human Proteome Organization (HUPO)

2005-2015 Advisory Committee, Protein Data Bank (PDB)

2002-2013 Advisory Committee, Protein Structure Initiative, National Institute of General Medical Sciences (NIGMS), National Institutes of Health (NIH)

2002-2009 Advisory Board (Chair, 2005-2006), Emerging Information Technology Conference

2000-Present Board of Directors (2000-2004), Education Committee (2003-Present), Senior Member (2016-Present), International Society for Computational Biology (ISCB)

2000-2009 Advisory Board, Association of Chinese Bioinformaticians

2001-2005 Scientific Advisory Board, Silico Insights, Inc., Woburn, MA

2000-2002 Advisory Board of Directors, Data Unlimited International, Gaithersburg, MD

1998-2002 Scientific Advisory Board, D'Trends, Inc., San Ramon, CA

Awards / Honors

- Fellow, Institute of Electrical and Electronics Engineers (IEEE), 2022
- Fellow, Association for Computing Machinery (ACM), 2021
- Fellow, Asia-Pacific Artificial Intelligence Association (AAIA), 2021
- Distinguished Agriculture Alumni Award, Purdue University, 2020
- Recognition of Service Award, Association for Computing Machinery (ACM), 2019

- National Advisory General Medical Sciences Council, NIH, 2017 (appointed by NIH Director)
- Distinguished Member, Association for Computing Machinery (ACM), 2017
- Senior Member, The Institute of Electrical and Electronics Engineers (IEEE), 2017
- Senior Member, International Society for Computational Biology (ISCB), 2016
- Recognized as a “Highly Cited Researcher” (top 1%) by Thomson Reuters/Clarivate Analytics, 2014, 2015, 2016, 2017, 2018, 2019, 2020
- Recognition of Service Award, Association for Computing Machinery (ACM), 2014
- Board of Directors, SIGBio, Association for Computing Machinery (ACM), 2010 (elected), 2015 (re-elected)
- Board of Directors, US Human Proteome Organization (USHUPO), 2011 (elected), 2014 (re-elected)
- Council, Human Proteome Organization (HUPO), 2012 (elected to a 3- year term)
- FIRST Award, R29 LM05524, National Library of Medicine, NIH, 1993-1999
- President's Academic Scholarship, University of Texas at Tyler, 1988
- Du Pont Graduate Student Award, Purdue University, 1983
- Book Coupon Award (top 5% of class), National Taiwan University, Taiwan 1975, 1977, 1978

Editorial Boards

- Editorial Board, Current Opinion in Systems Biology, 2016-Present
- Associate Editor (2014-Present) and Steering Committee (2018-Present), IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)
- Editorial Board, International Journal of Genomic Medicine, 2013-Present
- Editorial Board, Microbiome, 2012-Present
- Executive Editor, Journal of Proteomics and Bioinformatics, 2008-2013
- Editorial Board, “Science, Engineering, and Biology Informatics” SEBI book series (ISSN: 1793-3692), World Scientific, 2004-Present
- Editorial Board, Transactions of the Integrated Bio-medical Informatics and Enabling Technologies Symposium (TIBETS), 2003-2012

Professional Societies

- American Association for the Advancement of Science (AAAS)
- American Medical Informatics Association (AMIA)
- Asia-Pacific Artificial Intelligence Association (AAIA)
- Association for Computing Machinery (ACM) [Fellow; SIGBio Board of Directors]
- Human Proteome Organization (HUPO) [former Council]
- Institute of Electrical and Electronics Engineers (IEEE) [Fellow]
- International Society for Biocuration (ISB)
- International Society for Computational Biology (ISCB) [Senior Member; former Board of Directors]
- Sigma Xi, The Scientific Research Honor Society
- US Human Proteome Organization (US HUPO) [former Board of Directors]

In the News /Profiles

- *Communications of the ACM* ([February 2022](#)): ACM Member News: At the Intersection of Computational Biology and Data Science
- *UDaily* ([9/1/2021](#)): Blue Hen stars: UD Magazine features stories of alumni excellence and expertise
- *UDaily* ([8/30/2021](#)): Surfing the Tidal Wave of Data: At the dawn of a new information age, Cathy Wu’s star keeps rising
- *UDaily* ([8/30/2021](#)): Helping Our Planet and Its People
- *UDMagazine* ([August 2021](#)): Surfing the Tidal Wave of Data
- *UDaily* ([2/18/2021](#)): Data science leader: Cathy Wu, director of UD’s Data Science Institute, discusses the future of this fast-growing field

- *UDaily* ([11/18/2020](#)): Highly cited researchers: Seven UD professors named to 2020 global list
- *UDaily* ([11/20/2019](#)): Very influential researchers: Six UD professors rank in the top 1% cited by peers
- *UD Engineering News* ([10/2/2019](#)): A Major Computational Resource for Delaware
- *UDaily* ([8/16/2019](#)): \$23.5 million to expand Delaware’s biomedical research pipeline: UD-led program funded by NIH and State of Delaware for another five years
- *UDaily* ([7/29/2019](#)): Help at your fingertips: New app designed for opioid users, loved ones, providers
- *UD Engineering News* ([3/1/2019](#)): Celebrating Women’s History Month
- *UDaily* ([12/19/2018](#)): Highly Cited Researchers: Wu, Green, Brown and Janotti make list of top influencers
- *UDMagazine* ([December 2018](#)): A conversation with Cathy Wu
- *UDaily* ([9/7/2018](#)): A ‘momentous day’ for University of Delaware research: Launch of Data Science Institute adds major collaborative muscle
- *UDaily* ([4/3/2018](#)): UD announces new Data Science Institute and founding director
- *UDaily* ([3/27/2018](#)): Searching for the Big Picture: \$1.9 million NIH grant helps UD computer scientists make biomedical research easier to peruse
- *UDaily* ([3/8/2018](#)): Female engineering pioneers: Meet 21 women who made history in UD’s College of Engineering
- *UDaily* ([2/1/2018](#)): Matchmaking for cancer care: New system matches genetic anomalies with precision medicine treatments
- *UDaily* ([12/1/2017](#)): Wu snags multiple honors: UD bioinformatics director is recognized among top researchers in her field
- *UDaily* ([8/2/2017](#)): UD teams get millions in NSF funding
- *UDaily* ([5/23/2017](#)): Biosciences big data: New online graduate certificate in applied bioinformatics launched
- *UDaily* ([2/24/2017](#)): For the Record: Cathy Wu is co-editor of a book
- *UDaily* ([1/10/2017](#)): Wu named Highly Cited Researcher
- *UDaily* ([8/29/2016](#)): Big data for better medicine: Researchers receive funding to facilitate translation of biomedical data
- *UDaily* ([5/28/2016](#)): Doctoral Hooding Convocation: Erin Crowgey Research Prize for her dissertation
- *UDaily* ([10/1/2015](#)): Bioinformatics research underpins development of precision medicine
- *UDaily* ([11/12/2015](#)): Women of research: New issue of University of Delaware Research magazine
- *UDaily* ([11/2/2015](#)): UD partnering on regional big data innovation hub
- *UDaily* ([10/28/2015](#)): Research provides insights into genetic basis of obesity
- *Youtube* ([10/27/2015](#)): Women of Research: Cathy Wu
- *UDaily* ([10/21/2014](#)): UD's Targett, Wu to present at Inspiring Women in STEM conference
- *UDaily* ([7/8/2014](#)): Three UD scientists among Thomson Reuters' 2014 Highly Cited Researchers
- *UDaily* ([1/6/2014](#)): Leadership of statewide biomedical research program announced
- *UDaily* ([12/12/2011](#)): UD-led research team awarded \$4.8 million in NSF, NIH funding for bioinformatics research
- *UDaily* ([4/13/2010](#)): UD launches Professional Science Master's programs in biotechnology, bioinformatics to meet workforce needs
- *UDaily* ([4/22/2009](#)): Cathy Wu delivers inaugural Jefferson Chair lecture at UD
- *UDaily* ([3/3/2009](#)): University welcomes Wu as new Jefferson Chair
- *The Scientist* ([10/15/2001](#)): Cathy Wu at the Crossroads: She saved the Protein Information Resource database and now aims to restore it to the world's best

PROFESSIONAL SERVICES

Conference Organizing Committees (>70)

1. Organizing Committee, NLM Curation at Scale Workshop 2022, National Library of Medicine, NIH, Bethesda, MD, March 28-30, 2022

2. Organizing Committee, Delaware Data Science Symposium, University of Delaware, Newark, DE, November 17, 2021
3. Session Co-Chair, Computational Biology, Genomic Data Science, and Epidemiology, Virtual North East Regional IDEa Conference, v-NERIC 2021, August 16-18, 2021
4. Organizing Committee, Delaware Data Science DARWIN Computing Symposium, University of Delaware, Newark, DE, February 12, 2021
5. Steering Committee, Text Mining COSI track at ISMB 2020, Text Mining Community of Special Interest (COSI), July 13, 2020
6. Organizing Committee & Session Chair, Delaware Data Science DARWIN Computing Symposium, University of Delaware, Newark, DE, February 12, 2020
7. Planning Committee, Delaware Data Science Symposium, University of Delaware, Newark, DE, November 15, 2019
8. Organizing Committee & Session Chair, Delaware IDEAs Symposium, University of Delaware, Newark, DE, November 7, 2019
9. Scientific Advisory Committee, Biocuration-2019, The 12th Conference of the International Society for Biocuration, Cambridge, UK, April 7-10, 2019
10. Conference Co-Chair, 2018 ACM Conference on Bioinformatics, Computational Biology and Health Informatics (ACM BCB), Washington, DC, August 29-September 1, 2018
11. Scientific Committee and Session Chair, Biocuration-2018, The 11th Conference of the International Society for Biocuration, Shanghai, China, April 8-10, 2018
12. Steering Committee, The 6th Critical Assessment for Information Extraction in Biology (BioCreative VI) Workshop, Washington, DC, November, 2017
13. Organizer, Protein Ontology Workshop, Bar Harbor, ME, October 11-12, 2017
14. Program Committee, The 13th Annual US HUPO Conference, San Diego, CA, March, 2017
15. Organizer, Protein Ontology Workshop, Washington, DC, November 14-15, 2016
16. Steering Committee, The 5th Critical Assessment for Information Extraction in Biology (BioCreative V) Workshop, Sevilla, Spain, September 9-11, 2015
17. Scientific Committee, Biocuration-2015, The 8th Conference of the International Society for Biocuration, Beijing, China, April 23-26, 2015
18. Organizer, Protein Ontology Meeting, Washington, DC, June 18-19, 2014
19. Conference Co-Chair, Plant Protein Phosphorylation Symposium, University of Missouri, Columbia, MO, May 28-30, 2014.
20. Steering Committee, The 4th Critical Assessment for Information Extraction in Biology (BioCreative IV) Workshop, Bethesda, Maryland, October 7-9, 2013
21. Conference Co-Chair, 2013 ACM Conference on Bioinformatics, Computational Biology and Biomedicine (ACM BCB), Washington, DC, September 22-25, 2013
22. International Program Advisory Board, 2013 HUPO World Congress, Yokohama, Japan, September 14-18, 2013
23. Session Chair, Biocuration-2013, The Conference of the International Society for Biocuration, Cambridge, UK, April 7-10, 2013
24. Workshop Organizer, PIR Knowledge Mining Tools for Proteins, Complexes and PTMs, International Plant and Animal Genome (PAG-XXI) Conference, San Diego, CA, January 12-16, 2013
25. Program Committee/Session Chair, 2012 HUPO World Congress, Boston, MA, September 9-13, 2012
26. Conference Co-Chair, 2012 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Philadelphia, October 4-7, 2012
27. Program Committee, The 3rd International Conference on Biomedical Ontology (ICBO-2012), Graz, Austria, July 22-25, 2012
28. Workshop Organizer, Immunology Ontologies and Their Applications in Processing Clinical Data, Buffalo, NY, June 11-13, 2012

29. Organizing Committee, BioCreative-2012 Workshop on Interactive Text Mining in the Biocuration Workflow, Washington DC, April 4-5, 2012
30. Organizing Committee, Biocuration-2012, The Conference of the International Society for Biocuration, Washington DC, April 2-4, 2012
31. Workshop Organizer, PIR Knowledge Mining Tools for Proteins, Complexes and PTMs, International Plant and Animal Genome (PAG-XX) Conference, San Diego, CA, January 14-18, 2012
32. Organizer, Alzforum/Protein Ontology Meeting, Buffalo, NY on October 4-5, 2011
33. Panel Chair, ACM Conference on Bioinformatics, Computational Biology and Biomedicine (ACM-BCB 2011), Chicago, August 1-3, 2011
34. Program Committee, International Conference on Biomedical Ontology (ICBO 2011), Buffalo, New York, July 28-30, 2011
35. Organizer, NECC (North East Cyberinfrastructure Consortium) Genome Annotation Workshop, Newark, DE, May 23-27, 2011
36. Session Organizer, ABRF-2011: An International Symposium of the Association of Biomolecular Resource Facilities, San Antonio, TX, February 19-22, 2011
37. Chair, Steering Committee, The 3rd Critical Assessment for Information Extraction in Biology (BioCreative III) Workshop, Bethesda, Maryland, September 13-15, 2010
38. Area Chair, 2010 IEEE International Conference on Bioinformatics and Bioengineering (BIBE), Philadelphia, PA, May 31-June 3, 2010
39. Organizer, NECC (North East Cyberinfrastructure Consortium) Genome Annotation Workshop, Newark, DE, May 10-14, 2010
40. Organizer, The 3rd Annual Protein Ontology Meeting, Newark, DE, April 26-28, 2010
41. Workshop Co-Organizer, PIR (Protein Information Resource) Workshop on Text Mining for Database Curation, Plant and Animal Genome XVIII Conference, San Diego, CA, January 9-13, 2010
42. Program Committee Co-Chair, 2009 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Washington, DC, November 1-4, 2009
43. Chair, Steering Committee, International Symposium on Designing Microbial Research Commons, National Academy of Sciences, Washington, DC, October 8-9, 2009
44. Workshop Co-Organizer, Workshop on Literature Collection and Curation, The 3rd International Biocuration Conference, Berlin, Germany, April 16-19, 2009
45. Session Chair, Session on Databases and Systems Integration, The 5th Annual US HUPO Conference, San Diego, CA, February 22-25, 2009
46. Workshop Organizer, PIR (Protein Information Resource) Workshop on Text Mining for Database Curation, Plant and Animal Genome XVII Conference, San Diego, CA, January 10-14, 2009
47. Program Committee Vice Chair, 2008 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Philadelphia, PA, November 7-9, 2008
48. Organizer, The 2nd Annual Protein Ontology Meeting, Washington, DC, November 18-20, 2008
49. Program Committee, The 16th Annual International Conference on Intelligent Systems for Molecular Biology (ISMB), Toronto, Canada, July 19-23, 2008
50. Session Organizer, Bioinformatics and Systems Biology, the 1st Annual EITC-Bio Workshop – Synergy of Bioinformatics and Biomedical Research, Princeton University, NJ, June 7, 2008
51. Co-Chair, Organizing Committee, The 4th Annual US HUPO Conference, Bethesda, MD, March 16-19, 2008
52. Workshop Organizer, Biodefense Functional and Structural Proteomics Workshop, The 4th Annual US HUPO Conference, Bethesda, MD, March 16-19, 2008
53. Organizer, The 1st Annual Protein Ontology Meeting, Washington, DC, December 3-4, 2007
54. Program Committee, The 15th Annual International Conference on Intelligent Systems for Molecular Biology (ISMB), Vienna, Austria, July 21-25, 2007
55. Session Organizer, Bioinformatics and Systems Biology, The 7th Annual Emerging Information And Technology Conference (EITC), Princeton University, NJ, August 9-10, 2007

56. Conference Scientific Committee, BioCreative (Critical Assessment of Information Extraction systems in Biology), Madrid, Spain, April 23-25, 2007
57. Workshop Organizer, NIAID Proteomics Workshop, The 3rd Annual US HUPO Conference, Seattle, WA, March 4-8, 2007
58. Program Committee, HUPO 5th Annual World Congress, Long Beach, CA, October 28-November 1, 2006
59. Meeting Organizer, The 3rd Annual Programmatic Meeting of the NIAID Biodefense Proteomic Research Program, Georgetown University, Washington, DC, May 16-17, 2006
60. Program Committee, the Pacific Symposium on Biocomputing (PSB), Maui, Hawaii, January 3-7, 2006
61. Conference Organizer, the Emerging Information Technology Conference (EITC), Princeton, NJ, October 29-30, 2004
62. Program Committee, The International Conference on Intelligent Systems for Molecular Biology (ISMB), Glasgow, Scotland, July 31-August 4, 2004
63. Organizer, Full-day workshop on Proteomic Bioinformatics, Georgetown University Medical Center, Washington, DC, June 17, 2003
64. Session Co-Chair, the Pacific Symposium on Biocomputing (PSB), Lihue, Hawaii, January 3-7, 2003
65. Conference Organizer, the Emerging Information Technology Conference (EITC), Princeton, NJ, November 1-2, 2002
66. Program Committee, the International Conference on Intelligent Systems for Molecular Biology (ISMB), Edmonton, Alberta, Canada, August 3-7, 2002
67. Program Committee, the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (BIOKDD), Edmonton, Alberta, Canada, July 23-26, 2002
68. Program Committee, the International Conference on Computational Biology and Genome Informatics (CBGI), Research Triangle Park, NC, March 8-10, 2002
69. Session Co-Chair, the Pacific Symposium on Biocomputing (PSB), Kauai Marriott Resort and Beach Club, Hawaii, January 3-7, 2002
70. Program Committee, the 2nd IEEE International Symposium on Bio-Informatics & Biomedical Engineering (BIBE), Bethesda, MD, November 4-5, 2001
71. Program Committee, the International Workshop on Bioinformatics in Data Mining (BIOKDD), San Francisco, CA, August 26-29, 2001
72. Program Co-Chair, the Atlantic Symposium on Computational Biology and Genome Information Systems & Technology (CBGI), Durham, NC, March 13-15, 2001
73. Organizer, invited session on Bioinformatics, the World Multiconference on Systemics, Cybernetics and Informatics (SCI), Orlando, FL, July 12-16, 1998
74. Organizer, Bioinformatics panel discussion (plenary session) and paper session, the Centennial Meeting of the Texas Academy of Science, Huntsville, TX, March 6-8, 1997
75. Organizing committee, the Third International Conference on Bioinformatics and Genome Research, Tallahassee, FL, June 1-4, 1994

Study Sections / Grant Review Panels / Site Visit Teams / Academic Reviews (>80)

1. Member, Advisory Council, National Institute of General Medical Sciences (NIGMS), May 20, 2021
2. Member, Academic Program Review Panel, Institute for Genomics and Bioinformatics, University of California, Irvine, March 8, 2021
3. Member, Bioinformatics Academic Program Review, Georgia Institute of Technology, February 8-9, 2021
4. Member, Advisory Council, National Institute of General Medical Sciences (NIGMS), February 3, 2021
5. Member, Advisory Council, National Institute of General Medical Sciences (NIGMS), September 17, 2020
6. Member, Advisory Council, National Institute of General Medical Sciences (NIGMS), May 21, 2020
7. Member, Advisory Council, National Institute of General Medical Sciences (NIGMS), NIH, Bethesda, MD, January 16, 2020
8. Member, Master in Data Science Program Academic Program Review, University of San Francisco, San Francisco, CA, 2019

Francisco, CA, November 18-19, 2019

9. Member, Advisory Council, National Institute of General Medical Sciences (NIGMS), NIH, Bethesda, MD, September 19, 2019
10. Member, Review Panel for Harnessing the Data Revolution (HDR) - Frameworks, Directorate for Computer and Information Science and Engineering (CISE), NSF, June 26-27, 2019
11. Member, Advisory Council, National Institute of General Medical Sciences (NIGMS), NIH, Bethesda, MD, May 16-17, 2019
12. Member, Review Panel for Harnessing the Data Revolution (HDR): Data Science Corps (DSC), Directorate for Computer and Information Science and Engineering (CISE), NSF, May 9-10, 2019
13. Member, Advisory Council, National Institute of General Medical Sciences (NIGMS), NIH, Bethesda, MD, January 24-25, 2019
14. Member, Advisory Council, National Institute of General Medical Sciences (NIGMS), NIH, Bethesda, MD, September 13-14, 2018
15. Member, Advisory Council, National Institute of General Medical Sciences (NIGMS), NIH, Bethesda, MD, May 11, 2018
16. Member, Advisory Council, National Institute of General Medical Sciences (NIGMS), NIH, Bethesda, MD, January 18-19, 2018
17. Member, Triennial Reverse Site Review Panel for the Joint Genome Institute (JGI), Office of Biological and Environmental Research, U.S. Department of Energy (DOE), Bethesda, MD, December 5-8, 2017
18. Member, Advisory Council, National Institute of General Medical Sciences (NIGMS), NIH, Bethesda, MD, September 14-15, 2017
19. Member, Advisory Council, National Institute of General Medical Sciences (NIGMS), NIH, Bethesda, MD, January 26-27, 2017
20. Member, Special Emphasis Panel/Scientific Review Group 2016/10 ZLM1 AP-T (01) 1 meeting, T15 Informatics Training Grant Program, National Library of Medicine (NLM), NIH, Bethesda, MD, August 2016
21. Member, External Evaluation Site Visit Team, West Virginia-INBRE (IDeA Network of Biomedical Research Excellence), Marshall University, June 2016
22. Member, Special Emphasis Panel/Scientific Review Group 2016/05 ZRG1 ETTN-D (02) M Emerging Technologies in Neuroscience, Center for Scientific Review (CSR), NIH, Bethesda, MD, March 2016
23. Member, Bioinformatics Academic Program Review, Georgia Institute of Technology, February 2016
24. Member, Year Eight Site Review Panel for the BioEnergy Science Center (BESC), Office of Biological and Environmental Research, U.S. Department of Energy (DOE), November 2015
25. Member, Special Emphasis Panel/Scientific Review Group 2015/08 ZRG1 HDM-Z (52) R - Pediatric Research using Integrated Sensor Monitoring Systems (PRISMS): Informatics Platform Technologies for Asthma (U54), Center for Scientific Review (CSR), Bethesda, NIH, July 2015
26. Member, Special Emphasis Panel/Scientific Review Group 2015/08 ZRG1 IMST-B (81), T32-T15 (supplement): Big Data to Knowledge (BD2K) Training Grant, Center for Scientific Review (CSR), Bethesda, NIH, June 2015
27. Member, Special Emphasis Panel/Scientific Review Group 2015/10 ZRG1 VH-J (90) S - BD2K: Biomedical Data Science Training Coordination Center, Center for Scientific Review (CSR), Bethesda, NIH, June 2015
28. Member, Special Emphasis Panel, ZGM1 PPBC-5 (KB), PharmGKB R24 Review, National Institute of General Medical Sciences (NIGMS), NIH, February 2015
29. Member, Advisory Council, National Institute of General Medical Sciences (NIGMS), NIH, September 2014
30. Member, Special Emphasis Panel, ZGM1 TWD-6 (C1), COBRE (P20) Review, National Institute of General Medical Sciences (NIGMS), NIH, July 2014
31. Member, Special Emphasis Panel/Scientific Review Group 2014/05 ZRG1 BST-N (51) R - Big Data to Knowledge (BD2K) Centers of Excellence for Big Data Computing in the Biomedical Sciences, Center for Scientific Review (CSR), Bethesda, NIH, April 2014
32. Member, Special Emphasis Panel/Scientific Review Group 2013/08 ZCA1 SRLB-V (O1) S - Early Stage and

- Advanced Development of Informatics Technology, National Cancer Institute (NCI), NIH, July 2013
33. Member, Site Visit Panel, NCI Intramural Research Programs, Radiation Oncology Branch, National Cancer Institute (NCI), Bethesda, NIH, June 2013
 34. Member, Special Emphasis Panel, ZGM1 PPBC-5 LG, R24 Legacy Resources Review, National Institute of General Medical Sciences (NIGMS), NIH, March 2013
 35. Member, Year Five Site Review panel for the BioEnergy Science Center (BESC), Office of Biological and Environmental Research, U.S. Department of Energy (DOE), October 2012
 36. Member, External Advisory Panel for NHLBI (National Heart, Lung, and Blood Institute) Proteomics Program, Bethesda, NIH, August 2012
 37. Member, Special Emphasis Panel/Scientific Review Group 2012/01 BDMA, Biodata Management and Analysis Study Section, Center for Scientific Review (CSR), NIH, October 2011
 38. Member, Grant Review Panel, IMST 15 Small Business: Cell, Molecular, and Computational Biology, Center for Scientific Review (CSR), NIH, June 2011
 39. Member, Special Emphasis Panel/Scientific Review Group 2011/05 ZGM1 PPBC-4 (LR), NIGMS Legacy Community-Wide Scientific Resources, National Institute of General Medical Sciences (NIGMS), NIH, April 2011
 40. Member, Site Visit Panel, Biotechnology High Performance Computing Software Applications Institute (BHSAI), US Army Medical Research and Materiel Command (USAMRMC), Department of Defense (DoD), October 2010
 41. Members, Review Committee, RFA-DA-10-014 Systems Biology, HIV-AIDS, and Substance Abuse (R01) applications, National Institute of Drug Abuse (NIDA), NIH, July 2010
 42. *Ad hoc* reviewer, CISE Computing Research Infrastructure (CRI), NSF, April 2010
 43. *Ad hoc* reviewer, Postdoctoral Fellow of the Research Foundation - Flanders (Belgium) (FWO), February 2010
 44. Member, Bio Medium panel, Division of Information and Intelligent Systems (IIS), Directorate for Computer and Information Science and Engineering (CISE), NSF, November 2009
 45. Member, Special Emphasis Panel/Scientific Review Group 2010/01 ZLM1 ZH-C (J2), National Library of Medicine (NLM), NIH, October 2009
 46. Member, Special Emphasis Panel/Scientific Review Group 2009/10 ZGM1 CBCB-3 (BI), National Institute of General Medical Sciences (NIGMS), NIH, July 2009
 47. *Ad hoc* reviewer, 2009/10 ZRG1 BST-M (58) R - RFA OD-09-003 Challenge Grants Panel 4, NIH, June 2009
 48. Member, Special Emphasis Panel, Drug Docking and Screening Data Resource (U01), NIGMS, NIH, August 2008
 49. Member, Special Emphasis Panel, Phase III clinical trial on Multicenter Trial to Study Efficacy of Propranolol in Severely Burned Patients (P01), NIGMS, NIH, August 2008
 50. Member, Site Visit Team, New York Consortium on Membrane Protein Structure (NYCOMPS), NIGMS, NIH, New York, NY, May 2008
 51. Member, Site Visit Team, Center for Eukaryotic Structural Genomics (CESG), NIGMS, NIH, Madison, WI, April 2008
 52. Member, Site Visit Team, Protein Structure Initiative (PSI), National Institute of General Medical Sciences (NIGMS), NIH (3 PSI centers at New York, NY; Buffalo, NY; Piscataway, NJ), May 2007
 53. Member, Site Visit Panel, ZRG1 BST-D (P41 grant), National Center for Research Resources (NCRR), NIH, San Francisco, CA, November 2006
 54. Member, Special Emphasis Panel, Innovative Molecular Analysis Technology (IMAT), National Cancer Institute (NCI), NIH, October 2006
 55. Member, Special Emphasis Panel, Clinical Proteomic Technology Assessment for Cancer (U24), National Cancer Institute (NCI), NIH, July 2006
 56. Member, Special Emphasis Panel, T15 Informatics Training Grant Program, National Library of Medicine (NLM), NIH, May 2006

57. Frequent *Ad hoc* reviewer for Biological Databases and Informatics Program, Directorate for Biological Sciences, NSF, 2002-2006
58. Member, Special Emphasis Panel, E. coli Model Organism Resource Review, ZGM1 GDB-5 EC, National Institute of General Medical Sciences (NIGMS), NIH, December 2005
59. Member, CAREER Review Panel, Information & Intelligent Systems (IIS), Directorate for Computer and Information Science and Engineering (CISE), NSF, November 2005.
60. Member, Special Emphasis Panel, KinaseNET Program Project, Center for Scientific Review (CSR), NIH, March 2005
61. *Ad hoc* reviewer, Wellcome Trust Sanger Institute (WTSI), London, UK, March 2005
62. Member, Special Emphasis Panel, ZRG1 BST-D, Technology Centers for Networks and Pathways, CSR, NIH, July 2004
63. *Ad hoc* reviewer, IRG: ZDA1 SRC(99), NIDA Neuroproteomics Research Centers, National Institute on Drug Abuse (NIDA), NIH, July 2004
64. Member, Site Visit Team, Tomato Genome Sequencing Project, Plant Genome Research Program, Directorate for Biological Sciences, NSF, May 2004
65. *Ad hoc* reviewer, IRG: ZDA1 RXL-E(21), NIDA Neuroproteomics Research Centers, NIDA, NIH, March 2004
66. *Ad hoc* reviewer, Discovery Grant Program, Natural Sciences and Engineering Research Council (NSERC), Canada, January 2004
67. Member, Comparative Medicine Review Committee (CMRC), National Center for Research Resources (NCRR), NIH, June 2003
68. Member, Site Visit Team, Protein Data Bank (PDB), Biological Databases and Informatics (BDI) Program, Directorate for Biological Sciences, NSF, May 2003
69. Member, SSS-H (90), Computational Biology Study Section, CSR, NIH, February 2003
70. Frequent *Ad hoc* reviewer for Plant Genome Research Program, Directorate for Biological Sciences, NSF, 2002-2004
71. Member, SSS-E(01), Computational Biology Study Section, CSR, NIH, 2002
72. Member, Special Emphasis Panel, the National Research Program for Genomic Medicine, National Science Council, Taiwan, 2002
73. Member, Special Emphasis Panel, Biomedical Information Science and Technology Initiative (BISTI) Program, CSR, NIH, 2001
74. Member, Special Emphasis Panel, T15 Informatics Training Grant Program, National Library of Medicine (NLM), NIH, 2001
75. Member, Grant Review Panel, Microbial Genome Program, Department of Energy, 2001
76. Member, Grant Review Panel, Bioinformatics Post-doctoral Program, NSF, 2001
77. Member, Grant Review Panel, Biological Database and Informatics (BDI) Program, NSF, 1999-2000
78. Member, Site Visit Team, the *Arabidopsis* Functional Genomics Consortium, Division of Integrative Biology and Neurosciences, NSF, 2000
79. Reverse Site Visit Team, Plant Genome Research Program, NSF, 2000
80. Member, Grant Review Panel, Computational Biology Activities, Division of Biological Instrumentation and Resources, NSF, 1997
81. *Ad hoc* reviewer for Puerto Rico EPSCoR, 1995
82. Member, Grant Review Panel, Small Business Innovation Research (SBIR), the Division of Information, Robotics and Intelligent Systems, NSF, 1992
83. Member, Grant Review Panel, Special Project Program of Computational Biology Activities, Division of Biological Instrumentation and Resources, NSF, 1992

INVITED LECTURES / PANELS (>200)

1. Speaker, Research Town Hall, University of Delaware, Newark, DE, November 15, 2021
2. Keynote, The Tenth International Conference on Bioinformatics and Biomedical Science (ICBBS 2021),

October 29-31, 2021

3. Speaker, NIH STRIDES AWS Tech Talk - AWS Cloud Credits for Research, October 27, 2021
4. Speaker & Panel, MTBA (Midwest Taiwanese Biotechnology Association) Annual Symposium, Chicago, IL, September 18-19, 2021
5. Speaker, My Analytical Chemistry Story, ACS Fall 2021, American Chemical Society, August 23, 2021
6. Keynote, The Fifth International Conference on Biological Information and Biomedical Engineering (BIBE 2021), July 20-22, 2021
7. Speaker, Research Center Symposium, College of Engineering, University of Delaware, Newark, DE, April 30, 2021
8. Speaker, Distinguished Alum Seminar, Purdue University, West Lafayette, IN, March 19, 2021
9. Speaker & Panel, Biotech & Chemtech Symposium, New Castle County Chamber of Commerce Chamber of Commerce, DE, February 25, 2021
10. Speaker & Panel, BioBreakfast, Delaware BioScience Association, Wilmington, DE, January 21, 2021
11. Speaker & Panel, Workshop on Knowledge Guided Machine Learning (KGML), NSF Harnessing the Data Revolution (HDR) program, August 20, 2020
12. Keynote, Text Mining Community of Special Interest (COSI), The Annual International Conference on Intelligent Systems for Molecular Biology (ISMB), a virtual conference, July 13, 2020
13. Panel, COVID-19, University of Delaware, Newark, DE, June 23, 2020
14. Speaker, Nemours Cardiac Center, Nemours/Alfred I. duPont Hospital for Children, Wilmington, DE, May 19, 2020
15. Speaker, Brookhaven National Laboratory (BNL) Computational Science Initiative Seminar Series, Upton, NY, January 2020
16. Speaker, NIST (National Institute of Standards and Technology) Biomolecular Measurement Division Seminar Series, Gaithersburg, MD, January 21, 2020
17. Speaker, Data Science & Digital Agriculture Seminar Series, Purdue University, West Lafayette, IN, December 6, 2019
18. Speaker/Panelist, NERLSCD /MAD SSCi 19 (Northeast Regional Life Science Core Directors/Mid Atlantic Directors & Staff of Scientific Cores), Philadelphia, PA, November 8, 2019
19. Speaker, NIGMS Joint IDeA (Institutional Development Award) PI Meeting, NIH, Bethesda, MD, September 2019
20. Speaker, Neuroscience and Mathematics Education Workshop, University of Delaware, Newark, DE, September 13, 2019
21. Speaker, 2019 NorthEast Regional IDeA Conference (NERIC), Bretton Woods, NH, August 2019
22. Speaker, Junior Investigators Network (JIN), Delaware IDeA Programs, Newark, DE, June 20, 2019
23. Speaker, National Institute for Mathematical and Biological Synthesis (NIMBioS) Investigative Workshop Scientific Collaboration Enabled by High Performance Computing, Knoxville, TN, May 2019
24. Speaker, The 2019 Annual US HUPO (Human Proteome Organization) Conference, Bethesda, MD, March 2019
25. Speaker, The 6th Catholic University International Symposium on Medical Informatics, Korea Big Data Webinar, November 2018
26. Speaker, The 26th Annual International Conference on Intelligent Systems for Molecular Biology (ISMB), Chicago, IL, July 2018
27. Speaker, National Center for Biotechnology Information, National Library of Medicine (NLM), NIH, Bethesda, MD, May 2018
28. Speaker, Quill and Grill Wilmington Club, Unidel Foundation, Wilmington, DE, May 2018
29. Speaker, 2018 EITA (Emerging Information and Technology Association) Healthcare Innovation and Entrepreneurship Workshop "Smart, Precision, and Preventive Medicine: Challenges, Opportunities, and Future Directions," New York, NY, April 2018
30. Keynote Speaker, The 11th Conference of the International Society for Biocuration (Biocuration-2018),

Shanghai, China, April 2018

31. Speaker, EMBL Cancer Genomics Conference 2017, EMBL, Heidelberg, Germany, November 2017
32. Speaker, Nemours/Alfred I. duPont Hospital for Children, Wilmington, DE, October 2017
33. Speaker/Panelist, Health and Big Data Forum, University of Delaware, Newark, DE, September 2017
34. Speaker, NERIC (North East Regional IDeA Conference), Burlington, VT, August 2017
35. Speaker, Innovative Discoveries Series seminar, Christiana Care Health System, Wilmington, DE, June 2017
36. Speaker, BD2K-CCC (Big Data to Knowledge Centers Coordination Center) Webinar, June 2017
37. Speaker, LINCS-BD2K Systems Biology Data Science (SBDS) Symposium, Cincinnati, OH, May 2017
38. Speaker, The 2017 Annual US HUPO (Human Proteome Organization) Conference, San Diego, CA, March 2017
39. Speaker, International Darwin Day, University of Delaware, February 2017
40. Speaker, LINCS Data Science Research (DSR) Webinar, November 2016
41. Speaker/Panelist, Sustainability Session, NIH BD2K (Big Data to Knowledge) All Hands Meeting, Bethesda, MD, November 2016
42. Plenary Keynote Speaker, International Conference on Computational Advances in Bio and medical Sciences (ICCABS-2016), Georgia Institute of Technology, Atlanta, GA, October 2016
43. Speaker, Delaware-CTR (Clinical Translational Research Award) ACCEL Precision Medicine Retreat, Wilmington, DE, September 2016
44. Speaker, Frederick National Laboratory for Cancer Research, Frederick, MD, July 2016
45. Speaker, NHGRI Large Genomic Resources Meeting, National Human Genome Research Institute, NIH, Rockville, MD, May 2016
46. Speaker/Panelist, Innovations in Curation workshop, FORCE2016 Conference, Portland, OR, April 2016
47. Keynote Speaker, Biomedical Linked Annotation Hackathon (BLAH) Symposium, Mishima, Japan, November 2015
48. Speaker, The Role of Ontology in Big Cancer Data Workshop, National Cancer Institute, Bethesda, MD, May 2015
49. Panelist, NSF Workshop on Mass Spectrometry and Big Data Analysis, Arlington, VA, May 2015
50. Speaker, NIH Proteomics Interest Group Lecture, Bethesda, MD, April 2015
51. Presenter, UD Day on the Hill, Washington DC, March 2015
52. Speaker, Symposium on Data Enabled Research and Education, University of Rhode Island, Kingston, RI, January 2015
53. Panelist, Inspiring Women in Science & Technology (IWST) conference, Delaware BioScience Association, Wilmington, DE, October 2014
54. Speaker, RI-INBRE Seminar Series, University of Rhode Island, Kingston, RI, September 2014
55. Speaker, NIEHS-EPA Workshop for the Development of a Framework for Environmental Health Science Language, Raleigh, NC, September 2014
56. Speaker, INBRE PI-PC Meeting, National Institute of General Medical Sciences (NIGMS), Bethesda, MD, September 2014
57. Panelist, Clinical Study Data Sharing and Reuse Workshop, National Institute of Allergy and Infectious Diseases (NIAID), Bethesda, MD, September 2014
58. Speaker, Fifth Biennial National IDeA Symposium of Biomedical Research Excellence (NISBRE), Washington, DC, June 2014
59. Keynote Speaker, Advancing Computational Biology @ Howard University Symposium, Washington, DC, March 2014
60. Speaker, The 3rd International Conference on Proteomics & Bioinformatics (Proteomics-2013), Philadelphia, PA, July 2013
61. Speaker, The 6th Annual World protein and Peptide Conference (PepCon-2013), Suzhou, China, March 2013
62. Speaker and Mentor, Project Prometheus, Shantou University, China, February 2013

63. Speaker, The Second EITA-Bio Conference (EITA-Bio 2012), Princeton University, Princeton, NJ, October 2012
64. Speaker, Bioinformatics Colloquium, George Mason University, October 2012
65. Speaker, Georgetown Biomedical Informatics Symposium, Georgetown University, October 2012
66. Speaker, Nemours Center for Cancer and Blood Disorders, Research Strategy Retreat Nemours AI DuPont Children Hospital, Wilmington, DE, March 2012
67. Speaker, Workshop on Use of Ontologies for Organizing Plant and Animal Genomics Data, Plant and Animal Genome (PAG-XX) Conference, San Diego, CA, January 2012
68. Speaker, Workshop on PIR Knowledge Mining Tools for Proteins, Complexes and PTMs, PAG-XX Conference, San Diego, CA, January 2012
69. Speaker, AstraZeneca Pharmaceuticals, Wilmington, DE, January 2012
70. Speaker, Biochemistry Department Seminar, UMDNJ-New Jersey Medical School, Newark, NJ, January 2012
71. Speaker, MU Informatics Institute (MUII), University of Missouri, Columbia, MO, November 2011
72. Speaker, AlzForum/Protein Ontology Kick-Off Meeting, Buffalo, New York, October 2011
73. Panelist, Women in Bioinformatics Panel, 2011 ACM Conference on Bioinformatics, Computational Biology and Biomedicine (ACM-BCB), Chicago, IL, August 2011
74. Speaker, 4th Northeast Regional IDeA Meeting, Newport, RI, August 2011
75. Speaker, International Conference on Biomedical Ontology: ICBO 2011, Buffalo, New York, July 2011
76. Speaker, Digital Systems Seminar series, Department of Electrical and Computer Engineering, University of Delaware, April 2011
77. Speaker, The 7th Annual US HUPO (Human Proteome Organization) Conference, Raleigh, NC, March 2011
78. Speaker, ABRF-2011 (International Symposium of the Association of Biomolecular Resource Facilities), San Antonio, TX, February 2011
79. Panelist, Grand Challenges in Proteomics Workshop, National Science and Technology Council (NSTC), Gaithersburg, MD, February 2011
80. Speaker, Fraunhofer-CMB (Center for Molecular Biotechnology), Newark, DE, December 2010
81. Speaker, Department of Biological Sciences, University of Delaware, November 2010
82. Speaker, NIH Proteomics Interest Group (ProtiG), Bethesda, MD, October 2010
83. Plenary Speaker, NISBRE (National IDeA Symposium of Biomedical Research Excellence) Meeting, Washington, DC, June 2010
84. Panelist, International Critical Assessment of Functional Annotation Experiment (CAFAE), Crystal City, VA, May 2010
85. Speaker, DuPont, Wilmington, DE, April 2010
86. Speaker, Delaware INBRE Research Forum, Newark, DE, April 2010
87. Plenary Speaker, DOE Genomic Sciences Contractor-Grantee Workshop and Genomic Sciences Knowledgebase Workshop, Crystal City, VA, February 2010
88. Speaker, Protein Information Resource (PIR) Workshop on Text Mining for Database Curation, Plant and Animal Genome XVIII Conference, San Diego, CA, January 2010
89. Speaker, Protein Ontology workshop, Pacific Symposium of Biocomputing, The Big Island of Hawaii, January 2010
90. Speaker, Chemical Engineering Seminar Series, University of Delaware, December 2009
91. Speaker, American Statistical Association, Delaware Chapter, Newark, DE, November 2009
92. Speaker, Greater Philadelphia Bioinformatics Alliance (GPBA) Fall Retreat, Philadelphia, PA, November 2009
93. Speaker, Chemistry Biology Interface (CBI) Seminar Series, University of Delaware, November 2009
94. Speaker, Nemours Biomedical Research, Alfred I. duPont Hospital for Children, Wilmington, DE, May 2009
95. Speaker, American Society for Quality - ASQ509 Biomed/Biotech Special Interest Group Meeting, Rockville, MD, May 2009

96. Speaker, inaugural lecture as the Edward G. Jefferson Chair of Bioinformatics and Computational Biology, University of Delaware, April 2009
97. Speaker, Centers of Computational Biology, SRI International, Menlo Park, CA, April 2009
98. Speaker, The 5th Annual US HUPO Conference, San Diego, CA, February 2009
99. Keynote Speaker, The MidSouth Computational Biology and Bioinformatics Society (MCBIOS) Annual Meeting, Starkville, MS, February 2009
100. Speaker, Penn Bioinformatics Forum, University of Pennsylvania, Philadelphia, PA, February 2009
101. Speaker and Panelist, Computer Science Research Day, University of Delaware, Newark, DE, December 2008
102. Keynote Speaker, BioLINK SIG, The 16th Annual International Conference on Intelligent Systems for Molecular Biology (ISMB), Toronto, Canada, July 2008
103. Speaker, Taiwanese American Science and Technology Association (TASTA), Washington, DC, July 2008
104. Plenary Speaker, The 1st Annual EITC-Bio Workshop "Synergy of Bioinformatics and Biomedical Research," Princeton, NJ, June 2008
105. Panelist, DOE Genomics: GTL (Genome-To-Life) Program, Systems Biology Network/ Knowledgebase Workshop, Bethesda, MD, May 2008
106. Panelist, Workshop on the Biological Annotation of Novel Proteins, University of California, San Diego, La Jolla, CA, March 2008
107. Speaker, Biodefense Functional and Structural Proteomics Workshop, The 4th Annual US HUPO Conference, Bethesda, MD, March 2008
108. Speaker, Bioinformatics and Computational Biology Seminar Series, University of Delaware, February 2008
109. Speaker, Imaging Science and Information Systems (ISIS) Center Seminar Series, Georgetown University, Washington, DC, February 2008
110. Speaker, The 1st Annual Protein Ontology Meeting, Washington, DC, December 2007
111. Speaker, GenMed Seminar Series, Research Center for Genetic Medicine, Children's National Medical Center, Washington, DC, October 2007
112. Panelist, Workshop on Information Extraction for Biomedical Informatics (IEBI), University of Pennsylvania, Philadelphia, PA, October 2007
113. Speaker, International Workshop on Genomic Databases (IWGD'07), Angra dos Reis, Rio de Janeiro, Brazil, August 2007
114. Keynote Speaker, The 7th Annual Emerging Information & Technology Conference (EITC-2007), Princeton, NJ, August
115. Speaker, Towards Systems Proteomics: A Symposium Sponsored by the NHLBI Proteomics Initiative, Boston, MA, April 2007
116. Speaker, NIAID Proteomics Workshop, 3rd Annual US Conference, Seattle, WA, March 2007
117. Keynote Speaker, IEEE 6th Symposium on Bioinformatics and Bioengineering (BIBE), Washington, DC, October 2006
118. State-of-the-Art Speaker, HUPO World Congress, Long Beach, CA, October 2006
119. Speaker, Swiss-Prot20 Scientific Symposium, Fortaleza, Brazil, July 2006
120. Speaker, HUPO Scientific Meeting, Montreal, Canada, July 2006
121. Panelist, Workshop for a Plant Science Synthesis Center, NSF, Arlington, VA, October 2005
122. Speaker, Taiwanese American Science and Technology Association, Gaithersburg, MD, October 2005
123. Speaker, Intercampus Workshop Series, Georgetown University, Washington, DC, September 2005
124. Speaker, Keystone Symposium: Proteomics and Bioinformatics, Keystone, CO, April 2005
125. Speaker, School of Computational Sciences Colloquium, George Mason University, Fairfax, VA, March 2005
126. Lecturer, Course - BIOC 237 Fundamentals of Proteomics, George Washington University, Washington, DC, March 2005
127. Speaker, Advanced Research Computing (ARC) Seminar Series, Georgetown University, Washington, DC, December 2004

128. Speaker, GMU Bioinformatics Colloquium, George Mason University, Manassas, VA, November 2004
129. Speaker and Panelist, The Society for Neuroscience 34th Annual Meeting, Human Brain Project Satellite Symposium: Brain Maps and Systems Biology, San Diego, CA, October 2004
130. Speaker, Biology Department Seminar Series, Georgetown University, Washington, DC, October 2004
131. Speaker, PDB Symposium: Database Challenges in Biology, Rockville, MD, September 2004
132. Keynote Speaker, The Third Annual Great Lakes Bioinformatics Retreat, Hartland, MI, August 2004
133. Keynote Speaker, The First Annual Indiana Bioinformatics Conference, Indiana University School of Medicine, Indiana, May 2004
134. Panelist, Community Workshop on Long-Lived Data Collections, National Science Board, Arlington, VA, March 2004
135. Lecturer, DIMACS workshop in Data Mining Techniques in Bioinformatics, Rutgers University, Piscataway, New Jersey, October 2003
136. Lecturer, Distinguished Lecture Series in Bioinformatics and Computational Biology, University of Maryland, College Park, MD, October 2003
137. Keynote Speaker, Integrated Biomedical Informatics and Enabling Technologies (IBET), Pasquerilla Conference Center, Johnstown, PA, August
138. Speaker and Panelist, DOE Genomes to Life (GTL) Data Infrastructure workshop in Biological Data Integration, Gaithersburg, MD, July 2003
139. Speaker, Computational Biology Center, IBM Thomas J Watson Research Center, Yorktown Heights, NY, January 2003
140. Speaker and Panelist, NINDS workshop in Proteomics in the Neurosciences, Washington, DC, December 2002
141. Speaker, Feist-Weiller Cancer Center, Louisiana State University Health Sciences Center, Shreveport, LA, November 2002
142. Keynote Speaker, The Second Annual Emerging Information Technology Conference (EITC), Princeton, NJ, November 2002
143. Speaker and Panelist, The 2002 Virginia Biotechnology Summit and Governor's Conference on Technology Transfer and University Research (VABIO-2002), McLean, VA, October 2002
144. Lecturer, LRRRI Symposium on Molecular Approaches for Early Diagnosis and Treatment of Respiratory Disease, Santa Fe, NM, October 2002
145. Director and Lecturer, Full-day workshop on Proteomic Bioinformatics, Tri-Annual NIH proteomics course, October 2002
146. Speaker, Symposium on Biomedical Technology Development, Rockville, MD, September 2002
147. Speaker, The 19th Annual meeting of the Mid-Atlantic Plant Molecular Biology Society (MAPMBS), Beltsville, MD, August 2002
148. Speaker and International Delegate, The Second Chinese Conference on Bioinformatics (CCB02), Beijing, China, June 2002
149. Lecturer, Association for Research in Vision and Ophthalmology (ARVO) Meeting, Ft. Lauderdale, FL, May 2002
150. Speaker, Conference on Proteomics - The New Frontiers, University of Delaware, Newark, DE, March 2002
151. Director and Lecturer, Full-day workshop on Proteomic Bioinformatics, Tri-Annual NIH proteomics course, March 2002
152. Speaker, Molecular and Cellular Biology, Medical College of Ohio, Toledo, OH, November 2001
153. Keynote Speaker, IEEE International Symposium on Bio-Informatics & Biomedical Engineering, Bethesda, MD, November 2001
154. Speaker, School of Computational Sciences, George Mason University, Manassas, VA, October 2001
155. Lecturer, Three one-hour lectures, European Science Foundation Training Course in Functional Genomics: Curation of Databases in Molecular Biology, France, October 2001
156. Keynote Speaker, SIAC 2001-Bioinformatics and Data Mining, Wright State University, Dayton, OH, August

2001

157. Speaker, The Information Technology Research Institute, Wright State University, Dayton, OH, August 2001
158. Speaker, TSMC Conference in Bioinformatics, National Chiao Tung University, Taiwan, August 2001
159. Lecturer, Three one-hour lectures, the Summer Institute on Bioinformatics, Academia Sinica, Taiwan, August 2001
160. Lecturer, Half-day bioinformatics tutorial, the International Conference on Intelligent Systems for Molecular Biology (ISMB-2001), Copenhagen, Denmark, July 2001
161. Speaker, CHI Beyond Genomics 2001 - Proteomics Conference, San Francisco, CA, June 2001
162. Speaker, Cambridge Interpro Workshop, Sanger Center, Hinxton, United Kingdom, June 2001
163. Speaker, GCG/MSI Bioinformatics Seminar at the Walter Reed Army Institute of Research, Washington, DC, May 2001
164. Speaker and Panelist, Protein Sequence Database Workshop, NIH (NLM, NHGRI, NIGMS), Bethesda, MD, May 2001
165. Speaker, Protein Data Bank, Rutgers University, Piscataway, NJ, April 2001
166. Speaker, John Wiley & Sons, Inc., New York, NY, April
167. Speaker, Center for Biophysical Sciences & Engineering, University of Alabama at Birmingham, Birmingham, AL, March 2001
168. Plenary Speaker, Atlantic Symposium on Computational Biology and Genome Information Systems & Technology (CBGI 2001), Durham, NC, March 2001
169. Speaker, Department of Physiology and Biophysics, Georgetown University Medical Center, Washington, DC, March 2001
170. Speaker, Delaware Biotechnology Institute, University of Delaware, Newark, DE, December 2000
171. Speaker, NCI-FCRDC Advanced Biomedical Computing Center, Fredrick, MD, November 2000
172. Lecturer, Grand Rounds lectures, University of Texas Health Center at Tyler, Tyler, TX, November 2000
173. Plenary Speaker and Panelist, International Medical Informatics Symposium, Taiwan, November 2000
174. Speaker, Academia Sinica, Taiwan, November 2000
175. Speaker, National Taiwan University, Taiwan, November 2000
176. Speaker, National Health Research Institute, Taiwan, November 2000
177. Speaker, National Chiao Tung University, Taiwan, November 2000
178. Lecturer, Two-day Bioinformatics Workshop, National Yang Ming University, Taiwan, October 2000
179. Best Poster Presentation, MPSA 2000 (Methods of Protein Structure and Analysis), Charlottesville, VA, September 2000
180. Speaker, The Institute for Genome Research (TIGR), Rockville, MD, July 2000
181. Speaker, Boehringer Ingelheim Pharmaceuticals, Inc., Ridgefield, CT, June 2000
182. Speaker, Celera Genomics, Rockville, MD, July 1999
183. Speaker, International Bioinformatics and Genome Research Conference, San Francisco, CA, June 1999
184. Speaker and Panelist, NIGMS Structural Genomics Targets Workshop, National Institutes of Health, Bethesda, MD, February 1999
185. Speaker, Protein Information Resource, National Biomedical Research Foundation, Washington, DC, May 1998
186. Speaker, Lilly Research Lab., Eli Lilly and Company, Indianapolis, IN, January 1998
187. Speaker and Panelist, Challenges in Annotating Genomes Workshop at Pacific Symposium on Biocomputing, Maui, Hawaii, January 1998
188. Speaker, University of Texas Health Science Center at Houston, June 1997
189. Speaker, Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, June
190. Speaker, The Centennial Meeting of the Texas Academy of Science, Huntsville, TX, March 1997
191. Speaker, Protein Information Resource, National Biomedical Research Foundation, Washington, DC, May 1996

192. Speaker, Department of Chemistry, Stephen A. Austin University, Nacogdoches, TX, October 1995
193. Speaker, Molecular Informatics Group, Baylor College of Medicine, Houston, TX, April 1995
194. Speaker, Department of Computer Science, University of Tennessee, Knoxville, TN, October 1994
195. Speaker, Department of Computer Science, University of North Texas, Denton, TX, July 1994
196. Speaker, The Third International Conference on Bioinformatics and Genome Research, Tallahassee, Florida, June 1994
197. Speaker, Baylor College of Medicine, Houston, TX, May 1994
198. Speaker, Cray Symposium, Austin, TX, May 1994
199. Speaker, Workshop on Computational Neuroscience, Austin, TX, May 1993
200. Speaker and Panelist, The 2nd International Workshop on Open Problems in Computational Molecular Biology, Telluride, CO, July 1992
201. Speaker, The Second International Conference on Bioinformatics, Supercomputing and Complex Genome Analysis, St. Petersburg, FL, June 1992

UNIVERSITY SERVICE

University of Delaware

1. Co-Chair, Provost Search Advisory Committee, University of Delaware, 2021-Present
2. Co-Chair, Search Committee, Computer & Information Sciences Department Chair, College of Engineering, 2021-Present
3. Search Committee, DE-INBRE (IDeA Network of Biomedical Research Excellence) Principal Investigator, University of Delaware, 2021
4. Chair, Search Committee, Delaware Environmental Institute (DENIN) Director, University of Delaware, 2021
5. Advisory Committee to Dean, Named Professors Committee, College of Engineering, 2020-Present
6. Co-Chair, Search Committee, Computer & Information Sciences Department Chair, College of Engineering, 2020-2021
7. Data Coordination Committee, Clinical and Translational Research Task Force, Delaware CTR ACCEL Program, 2020-2021
8. Reopening Steering Committee, University of Delaware, 2020
9. Guiding Coalition Group, College of Engineering, 2020
10. Commission, Engineering-Driven Healthcare, College of Engineering, 2020
11. Research Council, University of Delaware, 2019-Present
12. Search Committee, Dean of Graduate College, 2019-2020
13. Nomination Review Committee, JPMC Faculty Fellows, 2019
14. Co-Chair, College of Engineering Dean Search Committee, 2017-2018
15. Chair of "Applications" Subcommittee, Data Science Faculty Search Committee, 2017-2018
16. Search Committee, Vice President for Information Technologies, University of Delaware, 2017
17. Faculty Search Committee, Department of Biological Sciences, 2016-2017
18. Executive Council for Diversity, University of Delaware, 2014-2019
19. Research Misconduct Inquiry Committee, University of Delaware, 2014-2015
20. Task Force on Research Data Security, University of Delaware, 2014-2015
21. Faculty Search Committee, Department of Biological Sciences, 2014
22. Scientific Advisory Board, Virome program, University of Delaware, 2014-Present
23. Internal Advisory Committee, Nemours COBRE - The Delaware Comprehensive Sickle Cell Research Center, 2014-Present
24. Nomination Review Committee, JPMC Faculty Fellows and Ph.D. Program Launch Leaders, Institute for Financial Services Analytics (IFSA), 2013
25. Search Committee, Unidel Howard E. Cosgrove Chair in the Environment, 2013

26. Nomination Review Committee of Named Professors in Mathematical Sciences, College of Arts & Sciences, 2013
27. Review Committee, Irma Ayers Chair in Human Services, University of Delaware, 2013
28. IGERT-SBE2 Executive Committee and Academic Program Committee, 2012-2019
29. Outstanding Junior Faculty Award Committee, College of Engineering, 2012-2021
30. Chair, Program Committee, PhD Program in Bioinformatics and Systems Biology (renamed Bioinformatics Data Science in 2020), 2012-Present
31. Search Committee, Dean of the College of Engineering, 2012-2013
32. Nomination Review Committee of Named Professors, College of Arts & Sciences, 2011-2012
33. Chair, Steering Committee, Bioinformatics Network of Delaware (BiND), Delaware INBRE, 2011-2016
34. Chair, Program Committee, Master's Program in Bioinformatics and Computational Biology, 2010-Present
35. Chair, Steering Committee, Center for Bioinformatics and Computational Biology, 2009-Present
36. Management Team (Member 2009-2013, Chair 2014-2021), Delaware INBRE (IDeA Networks of Biomedical Research Excellence)
37. Director, Bioinformatics Core Facility, University of Delaware, 2009-2016
38. DENIN Council of Fellows, Delaware Environmental Institute (DENIN), 2009-2016
39. Search Committee, Howard E. Cosgrove Chair in Environment, University of Delaware, 2010-2012
40. Statistics Task Force, University of Delaware, 2010-2011
41. Entrepreneurship Program Steering Committee, College of Engineering, 2010-2011
42. Search Committee, Dean of the Lerner College of Business and Economics, 2010-2011
43. Strategic Planning Committee, College of Arts & Sciences, 2010-2011
44. Chair, Faculty Search Committee, Bioinformatics and Computation Biology, University of Delaware, 2009-2011
45. Research Computing Task Force, University of Delaware, 2009-2011
46. Graduate Special Recruitment Committee, Department of Computer & Information Sciences, University of Delaware, 2009-2010

Georgetown University

1. Georgetown Database Of Cancer (G-DOC) Project Advisory Committee, Georgetown University Medical Center (GUMC), 2008-2009
2. Member, Medical School Clinical Informatics Committee, GUMC, 2007-2009
3. Member, Systems Medicine Curriculum Committee, GUMC, 2007-2009
4. Member, Task Force on Computational and Systems Biology, Georgetown University (GU), 2007-2009
5. Member, Committee on Appointments and Promotions, GUMC, 2005-2009
6. Advisor, Advance Research Computing, University Information System, GU, 2002-2009
7. Member, Research Committee (Chaired by Dean, Research and Translational Science) and Shared Resources and Infrastructure Subcommittee, GUMC, 2002-2005
8. Member, Task force, Virginia Tech Drug Development Program, GUMC, 2004
9. Member, Committee for MS Computational Biosciences degree program, GU, 2002-2004
10. Co-Chair, Task Force for Bioinformatics, Statistics, and Computing (Commissioned by Dean, Graduate School of Arts and Sciences), GU, 2002 [Delivered "Task Force Report on Bioinformatics, Statistics, and Computing"]
11. Member, Science Committee (Chaired by Provost), GU, 2002
12. Member, GU delegate to Virginia Tech for exploring strategic alliance, 2001
13. Member, Bioinformatics and Biostatistics Faculty Search Committee, Lombardi Comprehensive Cancer Center, GUMC, 2003-2004
14. Member, Bioinformatics Working Group, Lombardi Comprehensive Cancer Center, GUMC, 2001-2009

TEACHING

Degree Programs Developed

University of Delaware: <http://bioinformatics.udel.edu/Education>

Online Graduate Certificate program launched in Fall 2019
Online Graduate Certificate in Biomedical Informatics & Data Science

Online Graduate Certificate program launched in Fall 2017
Online Graduate Certificate in Applied Bioinformatics

PhD program launched in Fall 2012, Achieved Permanent Status in 2019
PhD program in Bioinformatics and Systems Biology
(renamed as PhD program in Bioinformatics Data Science in 2020)

Master's programs launched in Fall 2010, Achieved Permanent Status in 2017
Master of Science (MS) Program in Bioinformatics and Computational Biology
Computational Sciences Concentration
Life Sciences Concentration

Professional Science Master's (PSM) Program in Bioinformatics
Computational Sciences Concentration
Life Sciences Concentration

Graduate Certificate in Bioinformatics
Computational Sciences Concentration
Life Sciences Concentration

Georgetown University: <http://bioinformatics.georgetown.edu/>

Launched in Fall 2008
Master of Science (MS) Program in Biochemistry & Molecular Biology - Bioinformatics Track

Courses Taught

University of Delaware

Year	Course No	Course Title	Role
2017 Fall	BINF 644	Bioinformatics	Course Director
2017 Spring	BINF 694	Systems Biology I	Course Director
2016 Fall	BINF 644	Bioinformatics	Course Director
	BINF 865	Bioinformatics Seminar	Course Director
2015 Spring	BINF 694	Systems Biology I	Course Director
	BINF 865	Bioinformatics Seminar	Course Director
2014 Fall	BINF 644	Bioinformatics	Course Director
	BINF 865	Bioinformatics Seminar	Course Director
2014 Spring	BINF 694	Systems Biology I	Course Director
	BINF 865	Bioinformatics Seminar	Course Director
2013 Fall	BINF 644	Bioinformatics	Course Director
	BINF 865	Bioinformatics Seminar	Course Director
2013 Spring	BINF 694	Systems Biology I	Course Director
	BINF 865	Bioinformatics Seminar	Course Director
2012 Fall	BINF 667	Bioinformatics	Course Director
	BINF 865	Bioinformatics Seminar	Course Director
2012 Spring	BINF 667	Systems Biology I	Course Director
	BINF 865	Bioinformatics Seminar	Course Director
2011 Fall	BINF 865	Bioinformatics Seminar	Course Director

2011 Spring	BINF 865	Bioinformatics Seminar	Course Director
-------------	----------	------------------------	-----------------

Georgetown University

Year	Course No	Course Title	Role
2006-2007	BCHB 521	Bioinformatics	Course Director
2005-2006	BCHB 521	Bioinformatics	Course Director
2004-2005	BCHB 621	Advanced Bioinformatics	Course Director
	BCHB 531/2	Medical Biochemistry	Lectures on bioinformatics
2003-2004	BCHB 521	Bioinformatics	Course Co-Director
2002-2003	BCHB 521	Bioinformatics	Course Co-Director
2001-2002	BCHB 521	Bioinformatics	Course Co-Director
2000-2001	BCHB 521	Bioinformatics	Course Co-Director
1999-2000	BCHB 520	Molecular Biology & Bioinformatics	Lectures on bioinformatics

University of Texas at Tyler

Year	Course No	Course Title	Role
1993-1994	COSC 5330-01	Operating Systems	Course Director
	COSC 3311-01	Fortran Programming	Course Director
1992-1993	COSC 5330-81	Operating Systems	Course Director
	COSC 5360-01	Information Structures	Course Director
1991-1992	COSC 5360-81	Information Structures	Course Director
	COSC 5330-81	Operating Systems	Course Director
	COSC 4330-01	Programming Languages	Course Director
1990-1991	COSC 5360-81	Information Structures	Course Director
	COSC 5330-01	Operating Systems	Course Director
	COSC 5308-81	Computing Software	Course Director
	COSC 3308-02	Personal Computing	Course Director
1989-1990	COSC 5360-81	Information Structures	Course Director
	COSC 3309-01	Information System Software	Course Director
	COSC 3308-01	Personal Computing	Course Director
	COSC 3308-02	Personal Computing	Course Director
	COSC 3309-02	Information System Software	Course Director
	COSC 5330-81	Operating Systems	Course Director
	COSC 3308-02	Personal Computing	Course Director
	COSC 3350-01	RPG Programming	Course Director
	COSC 5360-81	Information Structures	Course Director
	COSC 3309-01	Information System Software	Course Director

MENTORING

Students, Trainees, Research Scientists and Faculty Mentored (>200)

University of Delaware (90), Georgetown University (62), University of Texas/Health Center at Tyler (57)

University of Delaware (90 Mentees)

- Research Faculty (8)
- Postdoctoral/Postmasters Research Mentees (5)
- PhD Dissertation Advisees (15: 9 CIS, 6 CBCB)
- MS Thesis/Internship Advisees (18: 6 CIS, 12 CBCB)
- PhD Dissertation/MS Thesis Committee (29)

- Independent Studies/Rotations/Summer Scholars (8)
 - Visiting Scholars/Interns (7)
1. Research Faculty Mentored in the Department of Computer and Information Sciences (CIS) and Center for Bioinformatics and Computational Biology (CBCB)
 - 1) Dr. Hongzhan Huang: Research Associate Professor (2009-present)
 - 2) Dr. Cecilia Arighi: Research Assistant Professor (2009-2012), Research Associate Professor (2012-present)
 - 3) Dr. Chuming Chen: Research Assistant Professor (2009-2018) , Research Associate Professor (2018-present)
 - 4) Dr. Natalia Roberts: Research Assistant Professor (2009-2011)
 - 5) Dr. Shawn Polson: Research Assistant Professor (2010-2017), Research Associate Professor (2017-present) [Director of CBCB Bioinformatics Core and Director of BiND, 2016]
 - 6) Dr. Manabu Torii: Research Assistant Professor (2011-2014) [Scientist, Medical Informatics at Kaiser Permanente]
 - 7) Dr. John Garavelli: Adjunct Professor (2011-2020)
 - 8) Dr. Huseyin Serdar Kuyuk, Adjunct Faculty (2020-2021)
 2. Research Trainee Mentored in the Department of Computer and Information Sciences (CIS) and Center for Bioinformatics and Computational Biology (CBCB)
 - 1) Dr. Jules Nchoutmboube, NIH Post-Master Trainee (2009-2010) [PhD, University of Maryland, 2016; MRL Postdoctoral Scientist, Merck 2016-2019; Scientist/Study Manager, BioReliance]
 - 2) Dr. Qinghua Wang: Postdoctoral Research Associate (2010-2013); Bioinformatics Scientist (2013-Present)
 - 3) Dr. Catalina Oana Tudor: Postdoctoral Research Associate (2011-2013) [Bioinformatics Scientist, UD; Chief Software Engineer & Founder, Urban Mining Company]
 - 4) Dr. Jennifer Wyffels: NIH Re-Entry Postdoctoral Research Associate (2012-2014) [Bioinformatics Scientist, UD]
 - 5) Dr. Karen Ross: NIH Re-Entry Postdoctoral Research Associate (2012-2015) [Assistant Professor, Medical Education, Georgetown University]
 3. Graduate Students Mentored in the Department of Computer and Information Sciences
 - Ph.D. Student (Dissertation Advisor/Co-advisor)
 - 1) Dr. Alvaro Gonzalez (2009-2012) [Bioinformatician at Epinomics]
 - 2) Dr. Yifan Peng (2010-2016) [Assistant Professor, Weill Cornell Medicine]
 - 3) Dr. Luis “David” Lopez (2010-2012) [Principal scientist, Plex-VR]
 - 4) Dr. Ruoyao Ding (2011-2017) [Research Assistant Professor, Jinan University, China]
 - 5) Dr. TianChuan “Michael” Du (2011-2015) [Data Scientist, Microsoft]
 - 6) Dr. Gang Li (2011-2017) [Software Engineer, Google]
 - 7) Dr. Lei Huang (2012-2017) [UD Graduate Fellow, 2016; Research Staff, IBM Almaden Research Center]
 - 8) Dr. Ashique Mahmood (2016-2018) [Research Scientist, Facebook]
 - 9) Dr. Samir Gupta (2016-2019) [Postdoc Fellow, Georgetown University]
 - M.S. Student
 - 1) Ms. Nisha Subramanian (2009-2010)
 - 2) Ms. Amy Siu (2009-2010) [PhD, International Max Planck Research School for Computer Science, Germany]
 - 3) Mr. Kanik Sem (2009-2010)
 - 4) Mr. Miao Zhang (2009-2010)
 - 5) Mr. Tim Armstrong (2010-2011)
 - 6) Mr. Zhiwen Li (2012-2014)

4. Graduate Students Mentored in the Bioinformatics and Computational Biology Program
 - PhD Student (Dissertation Advisor/Co-Advisor)
 - 1) Dr. Erin Crowgey (2013-2016) [UD Graduate Fellow, 2014, UD Dissertation Fellow, 2015, UD Dissertation Interdisciplinary Research Prize, 2016; Director Medical Bioinformatics, Center for Cancer and Blood Disorders, Nemours Children's Health]
 - 2) Mr. Imam Cartealy (2016-2017) [Fulbright Scholar]
 - 3) Ms. Riza Buatista (2016-present) [UD Graduate Scholar Award, 2016; NSF IGERT Scholar; ACM SIGHPC-Intel Fellowship]
 - 4) Dr. Jia Ren (2017-2019) [Software Engineer, Google]
 - 5) Ms. Rita Hayford (2017-present) [UD Graduate Scholar Award, 2017, 2018; UD Dissertation Fellowship Award, 2020]
 - 6) Mr. Sachin Gavali (2017-present) [UD Doctoral Fellowship Award, 2020; Sigma Xi GiAR Award, 2020]
 - MS and PSM Student (Thesis/Internship Advisor)
 - 1) Dr. Xia "Katie" Bi (2010-2012) [MD, Jefferson Medical College]
 - 2) Dr. Sari Khaleel (2010-2012) [MD, Geisel School of Medicine at Dartmouth, 2016; Urologic Oncology Fellow, Memorial Sloan Kettering Cancer Center]
 - 3) Dr. Jia Ren (2012-2013) [Software Engineer, Google]
 - 4) Dr. Mengxi Lv (2012-2014) [UD Doctoral Fellow, 2017] [Senior Data Scientist, IBM]
 - 5) Dr. Irem Celen (2012-2014) [Republic of Turkey Scholarship Scholar] [UD Graduate Fellow, 2016] [Bioinformatics Scientist, Sema4]
 - 6) Ms. Pan Tang (2013-2014)
 - 7) Ms. Julie Cowart (2012-2016) [Bioinformatics Software Engineer, UD]
 - 8) Ms. Melanie Salinas (2015) [PhD student at National University of Singapore]
 - 9) Dr. M. Joseph Tomlinson IV (2015-2016) [Computational Biologist, GSK]
 - 10) Ms. Xu Zhu (2016-2017)
 - 11) Mr. Juan Paolo Sicat (2016-2017) [PhD student at University of Greenwich]
 - 12) Ms. Xihan Qin (2018-2021)
5. Thesis/Dissertation Committee
 - 1) Ms. Abel Licon (2010) [MS, Computer and Information Sciences]
 - 2) Dr. Catalina Oana Tudor (2010-2011) [PhD, Computer and Information Sciences]
 - 3) Dr. Kevin McCormick (2011-2013) [PhD, Computer and Information Sciences]
 - 4) Mr. Brian S. Ladman (2011-2015) [PhD, Animal and Food Sciences]
 - 5) Mr. Jaysheel Bhavsar (2012-2014) [MS, Computer and Information Sciences]
 - 6) Mr. Sanjeev Patra (2012-2013) [MS, Bioinformatics and Computational Biology]
 - 7) Dr. Reza Hammond (2012-2019) [MS, Bioinformatics and Computational Biology, 2012-2014; PhD, Bioinformatics and Systems Biology, 2015-2019] (Senior Scientist, AbbVie)
 - 8) Mr. Abhijit Kapare (2012-2013) [MS, Bioinformatics and Computational Biology]
 - 9) Mr. Christopher Resnyk (2012-2013) [MS, Animal and Food Sciences]
 - 10) Dr. Dongqing Zhu (2013-2014) [PhD, Computer and Information Sciences]
 - 11) Mr. Alparslan Sari (2014) [MS, Bioinformatics and Computational Biology]
 - 12) Mr. Prasad Gajare (2014) [MS, Bioinformatics and Computational Biology]
 - 13) Dr. Boyu Zhang (2014-2015) [PhD, Computer and Information Sciences]
 - 14) Dr. Colin Kern (2014-2015) [PhD, Computer and Information Sciences]
 - 15) Mr. Matthew Ralston (2014-2015) [MS, Bioinformatics and Computational Biology]
 - 16) Ms. Barbra Ferrell (2014-2015) [MS, Bioinformatics and Computational Biology]
 - 17) Ms. Rebecca Sahaoui (2014-2016) [MS, Biological Sciences]
 - 18) Dr. Ramanuja Simha (2014-2016) [PhD, Computer and Information Sciences]
 - 19) Dr. Sangjuncta Barkataki (2015-2018) [PhD, Medical Sciences] (Senior Scientist, AstraZeneca)
 - 20) Dr. Atul Kakrana (2015-2018) [PhD, Bioinformatics and Systems Biology]

- 21) Dr. Moumita Bhattacharya (2016-2019) [PhD, Computer and Information Sciences](Senior Research Scientist, Netflix)
 - 22) Dr. Mengxi Lv (2016-2019) [PhD, Bioinformatics and Systems Biology]
 - 23) Dr. Irem Celen (2016-2019) [PhD, Bioinformatics and Systems Biology]
 - 24) Mr. Imam Cartealy (2019-present) [PhD, Bioinformatics Data Science]
 - 25) Mr. Gongbo Zhang (2019-present) [PhD, Computer and Information Sciences]
 - 26) Mr. Peng Su (2019-present) [PhD, Computer and Information Sciences]
 - 27) Ms. Debarati Roychowdhury (2020-present) [PhD, Computer and Information Sciences]
 - 28) Mr. Jiefu Li (2020-present) [PhD, Computer and Information Sciences]
6. Independent Studies/Rotations/Summer Scholars
 - 1) Mr. Yuanlong “Tyler” Tian (Fall 2012) [PSM, Bioinformatics]
 - 2) Mr. Yubo Xu (Spring 2014) [MS, Bioinformatics and Computational Biology]
 - 3) Mr. David Blickwedel (Spring 2014) [BS, Chemical and Biomolecular Engineering]
 - 4) Ms. Jennifer Niemczyk (Spring 2013) [NSF IGERT Scholar]
 - 5) Ms. Stephanie Luff (Fall 2013) [NSF IGERT Scholar]
 - 6) Ms. Rachel Lawrence (Summer 2018) [NSF EPScoR Undergraduate Summer Scholar]
 - 7) Mr. Connor Herron (Summer 2019) [NSF EPScoR Undergraduate Summer Scholar]
 - 8) Mr. Jonathan Olshin (Summer-Fall 2021) [NSF EPScoR Undergraduate Summer Scholar]
 7. Visiting Scholar [China Scholarship Council Program]
 - Dr. Dongmei Zhang, Associate Professor, Department of Computer Science & Technology, Shandong Jianzhu University, China (2014-2015)
 8. Visiting Interns
 - Undergraduate Student
 - Mr. John Stewart (2009) [Delaware State University]: Independent Study in Bioinformatics
 - Ms. Melissa Early (2011) [Wesley College]: Summer Intern
 - Mr. Gabriel Fernandez (2012) [Wesley College]: Summer Intern
 - Mr. Alexander Gorowara (2012) [Worcester Polytechnic Institute]: Summer Intern
 - Ms. Amber Wang (Summer 2020) [University of Maryland, College Park]: Summer Intern
 - Ms. Dana Niemeyer (Summer 2020) [Delaware State University]: NSF EPScoR Undergraduate Summer Scholar]

Georgetown University (62 Mentees)

- Research Faculty (16)
 - Postdoctoral Research Scientists and Research Associates Trainees (6)
 - Research Assistant Trainees (8)
 - PhD Dissertation Advisees (3)
 - Graduate Internship Advisees (19)
 - Visiting Scholars/Interns (6)
 - High School Mentees (4)
1. Research Faculty Mentored in the Department of Biochemistry and Molecular Biology, Georgetown University Medical Center (GUMC)
 - 1) Dr. Darren Natale: Research Assistant Professor (2002-present)
 - 2) Dr. Leslie Arminski: Research Assistant Professor (2002-2021)
 - 3) Dr. Hongzhan Huang: PIR Senior Bioinformatics Scientist (1999-2002); Research Assistant Professor (2002-2008)
 - 4) Dr. Anastasia Nikolskaya: Research Assistant Professor (2002-2008)
 - 5) Dr. Raja Mazumder: Research Assistant Professor (2002-2012)

- 6) Dr. Sandeep Kumar: Research Assistant Professor (2002-2003)
 - 7) Mr. Baris Suzek: Research Associate (2004-2008); Research Instructor (2008-2012); Research Assistant Professor (2012-2013)
 - 8) Dr. Peter McGarvey: Research Associate Professor (2005-2021); Research Professor (2021-present)
 - 9) Dr. Cecilia Arighi: Research Assistant Professor (2005-2008)
 - 10) Dr. Xin "Jean" Yuan: Research Instructor (2005-2006) [Investigator, GlaxoSmithKline (2007-present)]
 - 11) Dr. Robel Kahsay: Research Instructor (2005-2007)
 - 12) Dr. Zhang-Zhi Hu: PIR Senior Bioinformatics Scientist (2000-2006); Research Assistant Professor (2006-2007); Research Associate Professor (2007-2008)
 - 13) Dr. C.R. Vinayaka: PIR Senior Bioinformatics Scientist (2001-2006); Research Assistant Professor (2006-present)
 - 14) Dr. Sona Vasudevan: PIR Senior Bioinformatics Scientist (2004-2006); Research Assistant Professor (2006-2012)
 - 15) Dr. Natalia Roberts: Research Assistant Professor (2011-2013)
 - 16) Dr. Karen Ross: Assistant Professor on Medical Educator Track (2015-2022); Associate Professor on Medical Educator Track (Sep 2022-)
2. Postdoctoral Research Scientists and Research Associates Mentored at the National Biomedical Research Foundation (NBRF) and Department of Biochemistry and Molecular Biology, GUMC
 - 1) Dr. Chittari V. Shivakumar: Bioinformatics Scientist, NBRF (1999-2000)
 - 2) Dr. Chunlin Xiao: Bioinformatics Scientist, NBRF (1999-2001)
 - 3) Dr. Zhenglin Hou: Bioinformatics Scientist, NBRF (2000-2001)
 - 4) Dr. Peter Kourtesis: Postdoctoral Research Associate, GUMC (2003) [Clinical Development Coordinator at ALAPIS, Athens, Greece, since 2010]
 - 5) Dr. Xianying Wei: Postdoctoral Research Associate, GUMC (2004-2005)
 - 6) Dr. Thanemozhi G. Natarajan: Postdoctoral Research Associate, GUMC (2010-2012) [Genomics Scientist at Human Longevity, Inc.]
 3. Research Assistant Trainees Mentored at NBRF and Department of Biochemistry and Molecular Biology, GUMC
 - 1) Mr. Joseph Janda (B.S.): Technical Coordinator (1999-2000)
 - 2) Mr. Jiang Zhang (M.S.): Graduate Research Assistant (1999-2002)
 - 3) Ms. Kali Lewis (B.S.): Technical Coordinator (2000-2002)
 - 4) Mr. Devin Hendricks: Research Assistant (2003)
 - 5) Ms. Christina Fang (B.S.): Graduate Research Assistant (2004-2006)
 - 6) Mr. Narendra Kambam (Ph.D. student in Bioinformatics, George Mason U): Graduate Research Assistant (2003-2004)
 - 7) Mr. Vincent Hermoso (M.S. student in Biochemistry & Molecular Biology, GUMC): Graduate Research Assistant (2003-2005)
 - 8) Mr. Amar Kalelkar (B.S.): Graduate Research Assistant and Technical Coordinator (2003-2006)
 4. Graduate Students Mentored in the Department of Biochemistry and Molecular Biology, GUMC
 - Ph.D. Students in the Biomedical Program, GUMC
 - 1) Ms. Natalia Petrova, Biochemistry & Molecular Biology PhD program (2002-2008) [Dissertation topics: Structure bioinformatics for protein catalytic site prediction] [Research Assistant Professor at University of Delaware, since Spring 2009]
 - 2) Mr. Daniel Crooks, Biomedical Sciences GU/NIH Partnership Program (2007-2012) [NIH]
 - 3) Mr. Norman Goodacre, Biochemistry and Molecular & Cellular Biology PhD program (2009-2014) [ORISE Postdoctoral Fellow at FDA]
 - Graduate Interns (BCHB910 Biochemistry Internship) and Volunteer Interns in M.S. or Ph.D. program on Biochemistry & Molecular Biology, GUMC
 - 1) Ms. Zhenya Li (Ph.D. student) (2000)

- 2) Ms. Fang Wang (M.S. student) (2001)
 - 3) Mr. Bo Li (M.S. student) (2002)
 - 4) Ms. Maggie Yu (M.S. student) (2002)
 - 5) Ms. Abby Ngau (M.S. student) (2002-2003)
 - 6) Mr. Zachary Khalique (M.S. student) (2004)
 - 7) Mr. Vincent Hermoso (Spring 2005)
 - 8) Ms. Kawther Abdilleh (Spring 2005)
 - 9) Mr. Xiangbin Wang (Fall 2005)
 - 10) Mr. Alireza Amoozmand (Spring 2006)
 - 11) Mr. Ti-Cheng Chang (Spring 2006) [PhD, Penn State University, 2012]
 - 12) Mr. Paul Ramos (Spring 2006)
 - 13) Mr. Michael Brennan (Fall 2006)
 - 14) Ms. Elvena Fong (Spring 2007)
 - 15) Ms. Aparna Pallav (Spring 2007) [Bioinformatician, Virginia Commonwealth University]
 - 16) Ms. Pratibha Dubey (2008-2011)
 - 17) Mr. Lin An (2012-2013)
 - 18) Mr. Kam Ho To (Spring 2008)
 - 19) Lanlan "Iris" Yin (Spring 2008)
5. Visiting Graduate Interns
 - 1) Mr. Jung-Wei Fan, M.S. in Bioinformatics, National Yang Ming University, Taiwan (Summer 2002) [Ph.D. Biomedical Informatics, Columbia University; presently Kaiser Permanente's informatics team]
 - 2) Mr. Hsin-Ta Wu, M.S. in Bioinformatics, Yang Ming U, Taiwan (Summer 2005)
 - 3) Mr. Weihsiang Hsieh, M.S. in Bioinformatics, Yang Ming U, Taiwan (Summer 2005)
 - 4) Ms. Yen-Ting Hsieh, M.S. in Bioinformatics, Yang Ming U, Taiwan (Summer 2007)
 - 5) Chia-Ju "Cheryl" Lee, Northwestern University, Chicago, IL (Summer 2008)
 6. Visiting Scholar [China Scholarship Council-Georgetown Postdoctoral Fellowship Program]
 - 1) Dr. Kai Tang, Assistant Professor, State Key lab of Ocean biology, Xiamen University, Xiamen, China (2008-2009)
 7. High School Students Mentored
 - 1) Mr. Sherwin Yu, Thomas Jefferson High School for Science and Technology, Virginia (Summer 2007) [attending Yale University since Fall 2008]
 - 2) Mr. Winston Dong, Montgomery Blair High School, Maryland (Summer 2007)
 - 3) Mr. Theodore Toth, Winston Churchill High School, Maryland (Summer 2007) [attending Georgetown University since Fall 2008]
 - 4) Ms. Sheron Yin, Thomas Jefferson High School for Science and Technology, Virginia (Summer 2008) [attending Yale University since Fall 2009]

University of Texas/Health Center at Tyler (57 Mentees)

1. 53 Graduate (M.S. in Computer Science) research projects, University of Texas at Tyler (UTT) and University of Texas Health Center at Tyler (UTHCT), 1989-1998
 - Mentored M.S. students on research projects in Research Methods (COSC 5380) or Independent Study (COSC 5199)
2. 4 Postdoctoral Trainees, University of Texas Health Center at Tyler (UTHCT), 1995-1999
 - Dr. Sheng Zhao (1995-1997)
 - Dr. Chittari V. Shivakumar (1996-1998)
 - Dr. Jing Zhao (1998)
 - Dr. Hongzhang Huang (1998-1999)

RESEARCH GRANTS (>70)

[Continuously funded since 1987, totaling \$60M to lab]

1) Current Grants (8)

1. 09/17/2021 – 05/31/2026
2U24HG007822-08, NIH/NHGRI
UniProt: A Protein Sequence and Function Resource for Biomedical Science
\$24,633,173 total; PIR Component: \$4,105,528
Role: MPI and PI of PIR Component, 15.8% effort
[MPI: Alex Bateman, European Bioinformatics Institute (EBI), UK; MPI: Alan Bridge, Swiss Bioinformatics Institute, Switzerland]
2. 08/25/2021 – 07/31/2026
1R35GM141873-01, NIH/NIGMS
Protein Knowledge Networks and Semantic Computing for Disease Discovery
\$2,166,895 (total costs)
Role: PI, 28.3% effort
3. 09/01/2021 – 08/31/2022
358432114088, NIH STRIDES & AWS (Amazon Web Services)
NIGMS AWS Cloud Credit for Research (CCR): Cloud Computing for UniProt Protein Sequence Clustering as a Community Resource
\$54,650 (total costs)
Role: PI, 0% effort
4. 09/01/2021 – 08/31/2026
2125703, NSF/DGE
NRT- HDR: Computing and Data Science Training for Materials Innovation, Discovery, Analytics
\$2,999,011 total
Role: Co-PI, 0.8% effort
[PI: Arthi Jayaraman, University of Delaware]
5. 09/01/2020 – 08/31/2023
20-2020EPSCoR-0024, NASA
NASA EPSCoR Research Project: Building a Competitive and Sustainable Delaware Remote Sensing Big Data Center for Cutting-Edge Coastal and Environmental Change Research and Workforce Development
\$749,807 (total costs)
Role: Co-Science PI, 2.5% effort
[PI: William Matthaues, University of Delaware]
6. 10/01/2019 – 09/30/2022
1919839, NSF/OAC (Office of Advanced Cyberinfrastructure)
MRI: Acquisition of a Big Data and High Performance Computing System to Catalyze Delaware Research and Education
\$1,399,992 (total costs)
Role: Co-PI, 0% effort
[PI: Rudolf Eigenmann, University of Delaware]
7. 09/05/2018 – 06/30/2023
U54GM104941-07, NIH/NIGMS
Delaware Clinical and Translational Research ACCEL Program
\$21.0 million (total costs)
Role: Lead, Biostatistics, Epidemiology & Research Design (BERD) Core, 5% effort
[PI: Stuart Binder-Macleod, University of Delaware]
8. 08/01/2017 – 07/31/2022
NSF/OIA (Office of Integrative Activities)

RII Track-2 FEC: Advanced Biomanufacturing: Catalyzing Improved Host Development and High Quality Medicines through Genome to Phenome Predictions

\$6,000,000 (total costs)

Role: Senior Investigator, 5% effort

[PI: Sarah Harcum, Clemson University]

2) Past Grants (67)

9. 06/01/2018 – 09/30/2021
2U24HG007822-05, NIH/NHGRI
UniProt: A Centralized Protein Sequence and Function Resource
\$18 million total; PIR Component: \$3 million
Role: Consortium PI (PI of PIR Component), 20% effort
[PI: Alex Bateman, European Bioinformatics Institute (EBI), Hinxton, UK]
10. 05/01/2019 – 08/31/2021
2P20GM103446-19, NIH/NIGMS
Delaware INBRE
\$10 million (total costs)
Role: Program Coordinator, 30% effort
[PI: Steven Stanhope, University of Delaware]
11. 07/01/2020 – 05/31/2021
3U24HG007822-07S1, NIH/NHGRI
UniProt - Enhancing functional genomics data access for the Alzheimer's Disease (AD) and dementia-related protein research communities
\$318,651 total; PIR Component: \$191,622 [Administrative Supplement]
Role: Consortium PI (PI of PIR Component), 0% effort (effort in parent grant 2U24HG007822-07)
[PI: Alex Bateman, European Bioinformatics Institute (EBI), Hinxton, UK]
12. 09/01/2019 – 08/31/2020
University of Delaware, Graduate College Innovation Grant in Program Development & Innovation: An Online Biomedical Informatics & Data Science Graduate Certificate Program
\$40,913 (total costs)
Role: MPI, 0% effort
13. 09/01/2016 – 08/31/2020
3R01GM080646-11S1, NIH/NIGMS [Administrative Supplement]
PRO: Protein Ontology in OBO Foundry for Scalable Integration of Biomedical Knowledge
\$40,000 (total costs)
Role: PI, 0% effort (effort in parent grant 2R01GM080646-10)
14. 09/01/2015 – 08/31/2020
2R01GM080646-10, NIH/NIGMS
PRO: Protein Ontology in OBO Foundry for Scalable Integration of Biomedical Knowledge
\$2,810,198 (total costs)
Role: PI, 10% effort
15. 08/05/2016 – 07/31/2020
1U01GM120953-01, NIH/NIGMS
Semantic Literature Annotation and Integrative Panomics Analysis for PTM-Disease Knowledge Network Discovery
\$1,110,161 (total costs)
Role: Contact PI, 10% effort
16. 04/01/2018 – 06/30/2020
UNIDEL Foundation

- Data Science Institute
 \$650,000 (total costs)
 Role: Director, 24% effort
17. 09/20/2019 – 05/31/2020
 3U24HG007822-05S2, NIH/NHGRI
 UniProt Supporting Functional Genomics and Research Community Engagement for Alzheimer's Disease
 \$453,902 total; PIR Component: \$304,812 [Administrative Supplement]
 Role: Consortium PI (PI of PIR Component), 0% effort (effort in parent grant 2U24HG007822-06)
 [PI: Alex Bateman, European Bioinformatics Institute (EBI), Hinxton, UK]
 18. 12/10/2018 – 06/30/2019
 University of Delaware, Unidel Grand Challenge Grants
 Development of a UD wide training grant to support diversity in biomedical research graduate programs
 \$40,000 (total costs)
 Role: Co-PI, 0% effort [PI: Babatunde A. Ogunnaike, UD]
 19. 07/01/2018 – 06/30/2019
 WXWH18R0032, USAMRMC (US Army Medical Research and Materiel Command)
 ISfinder Prokaryotic Insertion Sequence (IS) Database
 \$301,172 total; PIR Component \$142,000
 Role: Co-PI, 2% effort
 20. 09/06/2018 – 05/31/2019
 3U24HG007822-05S1, NIH/NHGRI
 UniProt supporting an integrated multidisciplinary research agenda for Alzheimer's Disease
 \$512,695 total; PIR Component: \$114,319 [Administrative Supplement]
 Role: Consortium PI (PI of PIR Component), 0% effort (effort in parent grant 2U24HG007822-05)
 [PI: Alex Bateman, European Bioinformatics Institute (EBI), Hinxton, UK]
 21. 05/01/2016 – 04/30/2019
 3P20GM103446-16S1, NIH/NIGMS [Administrative Supplement]
 Delaware INBRE
 \$522,433 (total costs)
 Role: Program Coordinator and Co-PI, 0% effort (effort in parent grant 2P20GM103446-14)
 [PI: Steven Stanhope, University of Delaware]
 22. 08/01/2014 – 04/30/2019
 2P20GM103446-14, NIH/NIGMS
 Delaware INBRE
 \$18.2 million (total costs)
 Role: Program Coordinator, 30% effort
 [PI: Steven Stanhope, University of Delaware]
 23. 07/01/2012 – 06/30/2019
 NSF/DGE (Division of Graduate Education)
 IGERT: Systems Biology of Cells in Engineered Environments (SBE2)
 \$3,000,000 (total costs)
 Role: Co-PI, 2.5% effort
 24. 09/22/2015 – 05/31/2018
 1U01HG008390-01, NIH/Common Fund, BD2K
 MACE2K - Molecular and Clinical Extraction: A Natural Language Processing Tool for Personalized Medicine
 \$1,404,531 total; UD Subcontract: \$468,000
 Role: PI of UD Subcontract, 8.3% effort

- [PI: Subha Madhavan, Georgetown University, Washington, DC]
25. 09/01/2016 – 08/31/2018
University of Delaware, E-Learning Grant Program
Excellence and Innovation in E-Learning: An Online Bioinformatics Graduate Certificate Program
\$75,000 (total costs)
Role: PI, 2% effort
 26. 09/18/2014 – 05/31/2018
1U41HG007822-01, NIH/NHGRI & NIGMS
UniProt: A Centralized Protein Sequence and Function Resource
\$21.8 million total; PIR Component: \$3.8 million
Role: Consortium PI (PI of PIR Component), 20% effort
[PI: Alex Bateman, European Bioinformatics Institute (EBI), Hinxton, UK]
 27. 05/01/2016 – 04/30/2018
5U54HL127624-03, NIH/Common Fund, BD2K
Data Coordination and Integration Center for LINCS-BD2K: eDSR-Collaborative Resource for LINCS
Panomics PTM Knowledge Network
\$8 million total; UD subcontract: \$260,000
Role: PI of UD Subcontract, 4.7% effort
[PI: Avi Ma'ayan, Icahn School of Medicine at Mount Sinai, New York, NY]
 28. 05/01/2016 – 04/30/2018
1642535, NSF/MCB (Division of Molecular and Cellular Bioscience
Youth Bioinformatics Symposium
\$3,780 (total costs)
Role: Co-PI, 0% effort
[PI: Kamaljeet Sanghera, George Mason University, Fairfax, VA;
Sponsor: International Society for Computational Biology (ISCB)]
 29. 09/01/2013 – 08/31/2016
DE-SC0010838, DOE
BioCreative Workshops for DOE Genome Sciences: Text Mining for Metagenomics
\$40,000 (total costs)
Role: PI, 2% effort
 30. 09/01/2014 – 08/31/2016
1R561LM011354-01A1, NIH/NLM
Utilizing Imaged-based Features in Biomedical Literature Classification
Role: Co-Investigator, 2.5% effort
\$560,000 (total costs)
[PI: Hagit Shatkay, University of Delaware]
 31. 07/01/2011 – 06/30/2016
1062520, NSF/DBI (Division of Biological Infrastructure)
ABI Development: Integrative Bioinformatics for Knowledge Discovery of PTM Networks
\$1,592,561 (total costs)
Role: PI, 10% effort
 32. 09/21/2011 – 08/31/2015
2 R01GM080646-06, NIH/NIGMS
PRO: A Protein Ontology in OBO Foundry for Integration of Biomedical Knowledge
\$3,285,049 (total costs)
Role: PI, 15% effort
 33. 09/21/2012 – 08//2015
3R01GM080646-07S1, NIH/NIGMS

- PRO: A Protein Ontology in OBO Foundry for Integration of Biomedical Knowledge Supplements to Promote Re-entry into Biomedical and Behavioral Research Careers
\$276,103 (total costs)
Role: PI/Mentor, 0% effort (effort included in parent grant)
34. 09/27/2010 – 07/31/2014
1U41HG006104-01, NIH/NHGRI and NIGMS
UniProt: A Centralized Protein Sequence and Function Resource
\$21.8 million total; PIR Component: \$3,632,500 (total costs), \$2,387,000 (total direct costs)
Role: Consortium PI (PI of the PIR Component), 15% effort
[PI: Rolf Apweiler, European Bioinformatics Institute (EBI), Hinxton, UK]
35. 05/01/2009 – 02/28/2014
8P20GM103446, NIH/NIGMS
Delaware INBRE
\$17.4 million (total costs)
Role: Program Coordinator and Bioinformatics Core Director, 30% effort
[PI: Steven Stanhope, University of Delaware]
36. 09/01/2011 – 08/31/2014
DE-SC0007092, DOE
Experimental Systems-Biology Approaches for Clostridia-Based Bioenergy Production
\$2,232,919 (total costs)
Role: Co-PI, 5% effort
[PI: Terry Papoutsakis, University of Delaware]
37. 08/19/2010 – 08/31/2014
1G08LM010720, NIH/NLM
Linking Text Mining and Data Mining for Biomedical Knowledge Discovery
\$450,000 (total costs)
Role: PI, 0% effort
38. 10/01/2009 – 08/31/2014
0850319, NSF/DBI
Linking Text Mining with Ontology and Systems Biology
\$150,040 (total costs)
Role: PI, 3% effort
39. 11/01/2012 – 02/28/2014
3P20GM103446-12S1, NIH/NIGMS
Delaware INBRE
Supplements to Promote Re-entry into Biomedical and Behavioral Research Careers
\$154,530 (total costs)
Role: Co-PI/Mentor, 0% effort (effort included in parent grant)
[PI: Karl Steiner, University of Delaware]
40. 09/01/2012 – 08/31/2013
1242809, NSF/DBI
BIBM-2012 Travel Awards: Broadening Interdisciplinary Research and Education in Bioinformatics and Biomedicine
\$16,000 (total costs)
Role: PI, 0% effort
41. 09/01/2009 – 04/30/2012
3R01GM080646-04S2, NIH/NIGMS
PRO: A Protein Ontology in Open Biomedical Ontologies (Administrative Supplements)
\$473,350 (total costs), \$270,000 (total direct costs)

- Role: PI, 5% effort
42. 01/01/2010 – 06/30/2012
UNIDEL Foundation
Delaware Center for Bioinformatics and Computational Biology
\$500,000 (total costs)
Role: PI, 5% effort
43. 06/01/2011 – 05/31/2012
1137427, NSF/IIS (Division of Information and Intelligent Systems)
III: Small: Women in Bioinformatics Initiative at ACM BCB 2011
\$22,500 (total costs)
Role: Co-PI, 0% effort
[PI: Wei Wang, University of North Carolina at Chapel Hill]
44. 09/01/2009-08/31/2012
081425, NSF/EPSCoR
Delaware EPSCoR
Role: Bioinformatics Core Director, 2.5% effort
[PI: Don Sparks, University of Delaware]
45. 09/01/2010-08/31/2012
1039979, NSF/DBI
NSF MRI: Acquisition of Long Read DNA Sequencing
\$744,538 (total costs)
Role: Bioinformatics Core Director, 0% effort
[PI: Blake Meyers, University of Delaware]
46. 09/01/2006 – 08/31/2010
2 U01 HG02712-04, NIH/NHGRI and NIGMS
The UniProt Protein Sequence and Function Knowledgebase
\$24 million total; PIR Component: \$6,000,251 (total costs), \$3,991,406 (total direct costs)
Role: Co-PI (PI of the PIR Component), 20% effort
[PI: Rolf Apweiler, European Bioinformatics Institute (EBI), Hinxton, UK]
47. 05/01/2007 – 09/20/2011
1 R01 GM080646-01, NIH/NIGMS
PRO: A Protein Ontology in Open Biomedical Ontologies
\$2,179,589 (total costs), \$1,587,425 (total direct costs)
Role: PI, 15% effort
48. 08/29/2009 – 08/28/2011
3P20RR016472-09S2, NIH/NCRR
Delaware INBRE (Administrative Supplement only)
\$979,810 (total costs)
Role: Co-PI, 5% effort
[PI: Karl Steiner, University of Delaware]
49. 12/01/2009 – 06/30/2011
Strategic Initiative Award, UDRF (University of Delaware Research Foundation)
Intelligent Joint Search in Health and Biomedical Science Databases
\$45,000 (total costs)
Role: Co-PI and Research Mentor, 2.5% effort
[PI: Ben Carterette, University of Delaware]
50. 05/01/2010 – 04/30/2011
Delaware Health Sciences Alliance (DHSA)

- Linking Genotype to Phenotype: A pilot project to create a research data warehouse of biospecimen and omic information
 \$75,000 (total costs)
 Role: Co-PI, 5% effort
 [PI: Jack London, Thomas Jefferson University]
51. 09/01/2006 – 08/31/2010
 DOD Center of Excellence
 A New Paradigm for Breast Cancer Exploiting Low Dose Estrogen-induced Apoptosis
 \$10.7 million total; PIR component: \$646,231 (total), \$416,386 (direct)
 Role: Director of Georgetown Bioinformatics Core, 5% effort
 [PI: Craig Jordan, Fox Chase Comprehensive Cancer Center]
52. 09/01/2010 – 06/30/2011
 U54RR026076-A1; NIH/NCRR
 Georgetown-Howard Universities Center for Clinical and Translational Science (GHUCCTS)
 \$11 million total costs
 Role: Key Personnel (Task Leader of Biomedical Informatics)
 [PI: Joe Verbalis, Georgetown University Medical Center]
53. 09/13/10 – 12/31/2010
 GS-10F-0202K/10RT0098, USAMC/RDECOM
 Bacterial Pathogen Diagnosis
 \$100,000 (total costs)
 Role: PI, 5% effort
54. 10/01/2009 – 09/30/2010
 0960601, NSF/IIS (Division of Information and Intelligent Systems)
 BIBM Conference: Fostering Interdisciplinary Research and Education in Bioinformatics and Biomedicine
 \$20,000 (total costs)
 Role: PI, 2.5% effort
55. 7/01/2004 – 12/31/2009
 HHSN266200400061C, NIH/NIAID
 Administrative Resource for Biodefense Proteomic Research Centers
 \$8.8 million total; PIR component: \$1,094,046 (total), \$705,252 (direct)
 Role: Co-PI and Scientific Director (PI of the PIR Component), 20% effort
 [PI: Margaret Moore, Social and Scientific Systems, Inc. (SSS), Silver Spring, MD]
56. 09/20/2007 – 09/19/2009
 W81XWH0720112, USAMRMC (US Army Medical Research and Materiel Command)
 Literature Mining of Pathogenesis-Related Proteins in Human Pathogens
 \$200,000 (total costs), \$130,103 (total direct costs)
 Role: PI, 5% effort
57. 01/01/2005 – 03/30/2009
 caBIG (cancer Biomedical Informatics Grid), NIH/NCI
 Prime Contract: 1435-04-04-CT-73980; Subcontract: 79524CBS10
 Developer Project (2005-2006): Grid enablement of Protein Information Resource (PIR)
 Adopter Projects (2005-2007): SEED Genome Annotation; GenomeConnect ID Mapping
 ICR (Integrated Cancer Research), VCDE (Vocabularies and Common Data Elements) and Architecture
 Workspace Participant Projects (2005-2009)
 \$981,820 (total), \$634,163 (direct)
 Role: PI, 5% effort (2005-2007); Director (2008-2009)
58. 09/30/2002 – 08/31/2006
 1 U01 HG02712-01, NIH/NHGRI

The UniProt Protein Sequence and Function Knowledgebase
\$21 million total; PIR Component: \$7,089,327 (total costs), \$ 4,420,227 (total direct costs)
Role: Co-PI (PI of the PIR Component), 50% effort

[PI: Rolf Apweiler, European Bioinformatics Institute (EBI), Hinxton, UK]

59. 09/01/2004 – 08/31/2007
IIS-0430743, NSF/SEIII (Science and Engineering Information Integration and Informatics)
BioTagger - Biological Entity Tagging Using Online Resources and Machine Learning
\$823,109 total; PIR component: \$230,748 (total), \$148,678 (direct)
Role: Co-PI (PI of the PIR Component), 5% effort
[PI: Hongfang Liu, Department of Biostatistics, Bioinformatics and Biomathematics, Georgetown University Medical Center]
60. 06/01/2005 – 05/31/2006
Army Medical Command
Identification of biological pathogens
\$171,535 (total), 110,525 (direct)
Role: Co-PI (PI of the PIR Component), 5% effort
[PI: Spiros Dimolitsas, Senior Vice President, Georgetown University]
61. 8/18/2003 – 8/17/2006
W0130230, IBM, Shared University Research (SUR) Grant
Universal Protein Knowledgebase (UniProt) on IBM Server
\$1,149,704 [a gift towards GU's Third Century Campaign in reaching its \$1 billion goal]
Role: PI, 0% effort (an equipment grant with no allocated percent effort)
62. 01/01/2005 – 12/31/2005
GS-35F-4356D, US Air Force (USAF), Epidemic Outbreak Surveillance (EOS) Program
USAF Information Technology Support for the EOS Program
\$1 million total; PIR component: \$175,000 (total), \$113,000 (direct)
Role: Co-PI (PI of the PIR Component), 5% effort
[PI: Doreen Melling, Anteon Corporation, Fairfax, VA]
63. 07/01/2002 – 06/30/2005
DBI-0138188, NSF/BIO/BDI (Biological Database and Informatics)
Integrated Protein Classification for Genomic/Proteomic Research
\$499,928 (total), \$322,119 (direct)
Role: PI, 10% effort
64. 07/01/2002 – 06/30/2005
ITR-0205470, NSF/ITR (Information Technology Research)
Constructing Protein Ontologies Using Text Mining
\$704,430 total; PIR component: \$303,478 (total), \$195,540 (direct)
Role: Co-PI (PI of the PIR Component), 10% effort
[PI: Inderjeet Mani, Department of Linguistics, Georgetown University]
65. 10/01/2002-09/30/2004
SUN Microsystems, Academic Excellence Grant (AEG) Program
Integrated Bioinformatics Resource on SUN Server
\$75,000
Role: PI, 0% effort (an equipment grant with no allocated percent effort)
66. 03/01/2001-09/30/2002
P41 LM05798, NIH/NLM
Protein Information Resource for the Next Millennium
\$ 2,093,805 (total), \$1,311,812 (direct)
Role: PI, 80% effort

67. 03/01/1999-02/28/2001
P41 LM05798, NIH/NLM
Protein Information Resource for the Next Millennium
\$ 2,607,798 (total), \$1,665,955 (direct)
Role: Co-PI, 80% effort
[PI: Robert Ledley, National Biomedical Research Foundation, Washington, DC]
68. 10/01/1999-06/30/2002
DBI-9974855, NSF/BIO/BDI
PIR Classification Database for Genomic Research
\$283,147 (total), \$176,064 (direct)
Role: PI, 10% effort
69. 07/01/2003-12/31/2003
DCWASA (Water and Sewer Authority)
Computational Gene Target Identification for Pathogens
\$60,000 (total), \$38,638 (direct)
Role: Co-PI, 5% effort
[PI: Raja Mazumder, PIR, Department of Biochemistry and Molecular Biology, GUMC]
70. 07/01/2000-12/31/2000
R43 HG02176, NIH/NHGRI, SBIR (Small Business Innovation Research)
Protein Sequence and Annotation Search System
\$100,000 (total), \$58,954 (direct)
Role: Consultant, 5% effort
[PI: Hongzhan Huang, PIR, National Biomedical Research Foundation, Washington, DC]
71. 07/1999-07/2001
Monsanto Company
GeneFIND/ProClass and Neural Network Technology
\$50,000 total
Role: PI, 5% effort
72. 12/1998-05/2000
US1998055, Compaq-Digital, Inc., External Technology Program (ETG)
Digital Alpha Server for Protein Information Resource
\$220,000
Role: PI, 0% effort (an equipment grant with no allocated percent effort)
73. 07/1993-06/1999
R29 LM05524, NIH/NLM
Classification Neural Networks for Genome Research
\$460,641 (total), \$349,922 (direct)
Role: PI, 50% effort
74. 04/1991-12/1995
Cray Research, Inc.
Artificial Neural Systems on Cray for Gene Identification
\$151,000 total
Role: PI, 10% effort
75. 02/1989-12/1992
Cray Research, Inc.
Gene Identification and Classification
\$64,000 total
Role: Co-PI, 10% effort
[PI: George Whitson, Department of Computer Science, University of Texas at Tyler, TX]

76. 01/1987-12/1988
Texas A&M University
Genetic variations of *Phymatotrichum omnivorum* measured by RFLP of mitochondrial DNA
\$45,000 total
Role: PI, 50% effort

PUBLICATIONS AND PRODUCTS

Products / Research Resources

1) Patents

1. United States Patent No. 7943148, issued on 5/17/2011, "Amino acid sites in Flavivirus E proteins useful for development of diagnostics and vaccines," Inventors: Jose-Luis Sagripanti, Raja Mazumder, Cathy Wu
2. United States Patent No. 5845049, issued on 12/1/1998, "A neural network system with n-gram term weighting method for molecular sequence classification and motif identification," Sole Inventor: Cathy Wu

2) Research Resources

3. Protein Information Resource (PIR), 1999-Present
Led the development of PIR (<https://proteininformationresource.org/>) to become a major bioinformatics resource, as profiled in The Scientist (10/15/2001) "Cathy Wu at the Crossroads: She saved the Protein Information Resource database and now aims to restore it to the world's best"
4. UniProt Consortium, 2002-Present
Co-founded the UniProt Consortium (<https://www.uniprot.org/>) with Swiss Institute of Bioinformatics (SIB) and European Bioinformatics Resource (EBI). Many PIR resources have been integrated with the UniProt, including the PIRSF protein families into InterPro. UniProt receives >100 million webhits (8 million pageviews) from nearly 1 million unique sites worldwide on a monthly average, and a daily download of >2 TB of data.
5. Protein Ontology Consortium, 2007-Present
Founded the Protein Ontology Consortium (<https://proconsortium.org>) and led the development of Protein Ontology (PRO) within the Open Biomedical Ontologies (OBO) Foundry. PRO SPARQL endpoint is the top ranked Linked Open Data among 100 life science datasets according to YummyData. PRO website received >20 million hits from 1 million unique sites in the past year.
6. Natural Language Processing (NLP) and BioCreative Text Mining Challenge Evaluations, 2010-Present
Co-led the BioCreative (Critical Assessment of Information Extraction in Biology) Challenge Evaluations and Workshops (2010-2017); developed iPTMnet (launched in 2017) to capture protein post-translational modification (PTM) information from text mining and data mining, and iTextMine (launched in 2018) to integrate multiple literature mining tools resulting from NLP research

Books / Conference Proceedings / Journal Special Issues (15)

1) Books

1. **Wu CH**, Arighi CN and Ross KE (Editors) (2017). *Protein Bioinformatics: From Protein Modifications and Networks to Proteomics*. Series *Methods in Molecular Biology*, Volume 1558, Series Editor: Walker, John M. 472p. Humana Press. ISBN: 978-1-4939-6781-0; DOI: 10.1007/978-1-4939-6783-4
2. **Wu CH** and Chen C. (Editors) (2011). *Bioinformatics for Comparative Proteomics*. Series *Methods in Molecular Biology*, Volume 694, Series Editor: Walker, John M. 387p. Humana Press. ISBN: 978-1-60761-976-5
3. Wang J, **Wu CH** and Wang P. (Editors) (2003) *Computational Biology and Genome Informatics*. World Scientific. ISBN 981-238-257-7

4. **Wu CH** and McLarty J. (2000) *Neural Networks and Genome Informatics*. Methods in Computational Biology and Biochemistry, Volume 1, Series Editor: Konopka, AK. 205 p. Elsevier Science. ISBN 0 08 042800 2

2) Conference Proceedings / Journal Special Issues

5. Arighi CN, Wang Q, **Wu CH**. (Editors) (2017). Proceedings of the BioCreative VI Challenge Evaluation Workshop, October 18-20, 2017
6. Arighi CN, **Wu CH**, Cohen KB, Hirschman L, Krallinger M, Valencia A, Lu Z, Wilbur J, Wiegiers T (2014) BioCreative-IV Virtual Issue. *Database* 2014.
7. **Wu CH**, Aluru S, Catalyurek U, Hannenhalli S, Slonim D, Zhang A. (2014) ACM BCB-2013 (ACM Conference on Bioinformatics, Computational Biology and Biomedical Informatics) Virtual Issue, *Database* 2014
8. Arighi CN, Cohen KB, Hirschman L, Lu Z, Tudor OC, Wiegiers T, Wilbur J, **Wu CH**. (Editors) (2013). Proceedings of the Fourth BioCreative Challenge Evaluation Workshop, Vol 1. ISBN 978-0-615-89815-5
9. **Wu CH**, Arighi CN, Cohen KB, Hirschman L, Krallinger M, Lu Z, Mattingly C, Valencia A, Wiegiers TC, Wilbur WJ. (2012) BioCreative-2012 Virtual Issue, *Database* 2012.
10. Arighi CN, Cohen KB, Hirschman L, Krallinger M, Lu Z, Mattingly C, Valencia A, Wiegiers T, Wilbur J, **Wu CH**. (Editors) (2012). Proceedings of 2012 BioCreative Workshop. University of Delaware Press.
11. Gao J, Alhajj R, Dubitzky W, Ungar L, Wu CH, Christianson A, Liebman M, Hu X. (Editors) (2012). *Proceedings of 2012 IEEE International Conference on Bioinformatics and Biomedicine (BIBM-2012)*, IEEE Computer Society. ISBN 978-1-4673-2558-5
12. Arighi CN, Cohen KB, Hirschman L, Krallinger M, Lu Z, Valencia A, Wilbur J, **Wu CH**. (Editors) (2011). *The Third BioCreative - Critical Assessment of Information Extraction in Biology Challenge*. *BMC Bioinformatics*, Volume 12, Supplement 8.
13. Arighi CN, Cohen KB, Hirschman L, Krallinger M, Lu Z, Valencia A, Wilbur J, **Wu CH**. (Editors) (2010). Proceedings of the BioCreative III (Critical Assessment of Information Extraction in Biology) Workshop. University of Delaware Press. ISBN 978-1-4507-3685-5.
14. Chen XW, Kim S, **Wu CH**, Xu D. (Editors) (2009). *Proceedings of 2009 IEEE International Conference on Bioinformatics and Biomedicine (BIBM-2009)*, IEEE Computer Society. ISBN 978-0-7695-3885-3
15. **Wu CH**, Wang J. and Wang P. (Editors) (2001) *Proceedings of the Atlantic Symposium on Computational Biology and Genome Information Systems & Technology*. March 15-17, 2001, Durham, North Carolina. Association for Intelligent Machinery, Inc. ISBN 0-9707890-0-9

Refereed Papers in Journals and Conference Proceedings (>290)

["Highly Cited Researcher" (top 1%); [Google Scholar](#): >48,000 citations, h-index: 71, i10-index: 206]

1) Original Papers in Refereed Journals

1. Gavali S, Ross KE, Chen C, Cowart JE, **Wu CH**. (2022) PredKinKG: A knowledge graph representation learning approach to predicting novel kinase-substrate interactions. *Molecular Omics* (in press)
2. Hayford R, Serba D, Xie S, Ayyappan V, Thimmapuram J, Saha MC, **Wu CH**, Kalavacharla V. (2022) Global Analysis of Switchgrass (*Panicum virgatum* L.) Transcriptomes in Response to Interactive Effects of Drought and Heat Stresses. *BMC Plant Biology* (in press)
3. Wang Y, Wang Q, Huang H, Huang W, Chen Y, McGarvey PB, **Wu CH**, Arighi CN, on behalf of the UniProt Consortium. (2021) A crowdsourcing tale: an open platform for literature curation in UniProt. *PLoS Biol.* 2021 Dec 6;19(12):e3001464. doi: 10.1371/journal.pbio.3001464
4. Chen C, Ross KE, Gavali S, Cowart JE, **Wu CH**. (2021) COVID-19 knowledge graph from semantic integration of biomedical literature and databases. *Bioinformatics* doi: 10.1093/bioinformatics/btab694
5. Zhang X, Maity TK, Ross KE, Qi Y, Cultraro CM, Bahta M, Pitts S, Keswani M, Gao S, Nguyen KDP, Cowart J, Kirkali F, **Wu CH**, Guha U (2021) Alterations in the global proteome and phosphoproteome in third-

- generation EGFR TKI resistance reveal drug targets to circumvent resistance. *Cancer Research* 2021 Mar 16;canres.2435.2020. doi: 10.1158/0008-5472.CAN-20-2435. PMID: 33727228 [IF: 9.727]
6. Ross K, Varani AM, Snesrud E, Huang H, Alvarenga DO, Zhang J, **Wu CH**, McGann P, Chandler M. TnCentral: A Prokaryotic Transposable Element Database and Web Portal for Transposon Analysis. *mBio*. 2021 Sep 14:e0206021. doi: 10.1128/mBio.02060-21. PMID: 34517763.
 7. Chitwood DG, Wang Q, Elliott K, Bullock A, Jordana D, Li Z, **Wu CH**, Harcum SW, Saski CA. (2021) Characterization of Metabolic Responses, Genetic Variations, and Microsatellite Instability in Ammonia-stressed CHO cells grown in Fed-batch Cultures. *BMC Biotechnology* 2021 Jan 8;21(1):4. doi: 10.1186/s12896-020-00667-2.
 8. Gene Ontology Consortium. (2021) The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res*. 2021 Jan 8;49(D1): D325-D334. doi: 10.1093/nar/gkaa1113. PMID: 33290552; PMCID: PMC7779012.
 9. The UniProt Consortium (2021) UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res* 2021 Jan 8;49(D1):D480-D489. doi: 10.1093/nar/gkaa1100. PMID: 33237286
 10. Chen C, Huang H, Ross KE, Cowart JE, Arighi CN, **Wu CH**, Natale NA. (2020) Protein ontology on the semantic web for knowledge discovery. *Scientific Data* 2020 Oct 12;7(1):337. doi: 10.1038/s41597-020-00679-9.
 11. Blum M, Chang HY, Chuguransky S, Grego T, Kandasaamy S, Mitchell A, Nuka G, Paysan-Lafosse T, Qureshi M, Raj S, Richardson L, Salazar GA, Williams L, Bork P, Bridge A, Gough J, Haft DH, Letunic I, Marchler-Bauer A, Mi H, Natale DA, Necci M, Orengo CA, Pandurangan AP, Rivoire C, Sigrist CJA, Sillitoe I, Thanki N, Thomas PD, Tosatto SCE, **Wu CH**, Bateman A, Finn RD. (2020) The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res*. 2020 Nov 6:gkaa977. doi: 10.1093/nar/gkaa977. Epub ahead of print. PMID: 33156333.
 12. Ovadia EM, Pradhan L, Sawicki LA, Cowart J, R. E. Huber, Polson SW, Chen C, van Golen KL, Ross KE, **Wu CH**, Kloxin AM. (2020) Understanding ER+ breast cancer dormancy using bioinspired synthetic matrices for long-term 3D culture and insights into late recurrence. *Advanced Biosystems* 2020 Jun 30;e2000119. doi: 10.1002/adbi.202000119.
 13. Shamsaei B, Chojnacki S, Pilarczyk M, Najafabadi M, Niu W, Chen C, Ross K, Matlock A, Muhlich J, Chutipongtanate S, Jaffe J, MacCoss M, **Wu CH**, Pillai A, Ma'ayan A, Zheng J, Turner JP, Vidovic D, Schurer S, Kouril M, Medvedovic M, Meller J. (2020) piNET: a versatile web platform for downstream analysis and visualization of proteomics data. *Nucleic Acids Research*, gkaa436, doi: 10.1093/nar/gkaa436
 14. Gavali S, Cowart J, Chen C, Ross KE, Arighi CN, **Wu CH**. (2020) API for iPTMnet: A resource for protein post-translational modification network discovery. *Database (Oxford)* 2020. doi: 10.1093/database/baz157
 15. MacDougall A, Volynkin V, Saidi R, Poggioli D, Zellner H, Hatton-Ellis E, Joshi V, O'Donovan C, Orchard S, Auchincloss AH, Baratin D, Bolleman J, Coudert E, de Castro E, Hulo C, Masson P, Pedruzzi I, Rivoire C, Arighi C, Wang Q, Chen C, Huang H, Garavelli J, Vinayaka CR, Yeh LS, Natale DA, Laiho K, Martin MJ, Renaux A, Pichler K; UniProt Consortium. (2020) UniRule: a unified rule resource for automatic annotation in the UniProt Knowledgebase. *Bioinformatics* 2020 Nov 1;36(17):4643-4648. doi: 10.1093/bioinformatics/btaa485. PMID: 32399560
 16. Breuza L, Arighi CN, Argoud-Puy G, Casals-Casas C, Estreicher A, Famiglietti ML, Georghiou G, Gos A, Gruaz-Gumowski N, Hinz U, Hyka-Nouspikel N, Kramarz B, Lovering RC, Lussi Y, Magrane M, Masson P, Perfetto L, Poux S, Rodriguez-Lopez M, Stoeckert C, Sundaram S, Wang LS, Wu E, Orchard S; IMEx Consortium, UniProt Consortium. (2020) A Coordinated Approach by Public Domain Bioinformatics Resources to Aid the Fight Against Alzheimer's Disease Through Expert Curation of Key Protein Targets. *J Alzheimers Dis*. 2020;77(1):257-273. doi: 10.3233/JAD-200206. PMID: 32716361; PMCID: PMC7592670.
 17. Morgat A, Lombardot T, Coudert E, Axelsen K, Neto TB, Gehant S, Bansal P, Bolleman J, Gasteiger E, de Castro E, Baratin D, Pozzato M, Xenarios I, Poux S, Redaschi N, Bridge A; UniProt Consortium. (2020) Enzyme annotation in UniProtKB using Rhea. *Bioinformatics* 2020 Mar 1;36(6):1896-1901. doi: 10.1093/bioinformatics/btz817. PMID: 31688925; PMCID: PMC7162351.
 18. Bye-A-Jee H, Zaru R, Magrane M, Orchard S; UniProt Consortium. (2020) *Caenorhabditis elegans*

- phosphatase complexes in UniProtKB and Complex Portal. *FEBS J.* 2020 Jan 16. doi: 10.1111/febs.15213.
19. Su P, Li G, **Wu CH**, Vijay-Shanker K. (2019) Using distant supervision to augment manually annotated data for relation extraction. *PLoS One* 14(7): e0216913. doi: 10.1371/journal.pone.0216913 [PMC6667146]
 20. McGarvey PB, Nightingale A, Luo J, Huang H, Martin MJ, **Wu CH**, UniProt Consortium. (2019) UniProt Genomic Mapping for Deciphering Functional Effects of Missense Variants. *Human Mutation* 2019: 1-12. doi: 10.1002/humu.23738 [PMC6563471]
 21. Chitwood DG, Wang Q, Elliott K, Bullock A, Jordana D, Li Z, **Wu CH**, Harcum SW, Saski CA (2019) Microsatellite instability as a tool to diagnose genome instability in CHO cell culture. bioRxiv 823252; doi: <https://doi.org/10.1101/823252>
 22. Chen C, Wang Q, Huang H, Vinayaka CR, Garavelli JS, Arighi CN, Natale DA, **Wu CH**. (2019) PIRSitePredict for Protein Functional Site Prediction Using Position Specific Rules. *Database (Oxford)* 2019. doi: 10.1093/database/baz026
 23. Lu C, Sidoli S, Kulej K, Ross K, **Wu CH**, Garcia BA. (2019) Coordination between TGF- β cellular signaling and epigenetic regulation during epithelial to mesenchymal transition. *Epigenetics & Chromatin* 12:11. doi: 10.1186/s13072-019-0256-y [PMC6368739]
 24. Sawicki LA, Ovadia EM, Pradhan L, Cowart JE, Ross KE, **Wu CH**, Kloxin AM. (2019) Tunable Synthetic Extracellular Matrices to Investigate Breast Cancer Response to Biophysical and Biochemical Cues. *APL Bioengineering* 3, 016101 (2019). doi: 10.1063/1.5064596
 25. The UniProt Consortium (2019) UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* 47(D1): D506-D515 [PMC6323992]
 26. Ding R, Qu Y, **Wu CH**, Vijay-Shanker K. (2018) Automatic gene annotation using GO terms from cellular component domain. *BMC Med Inform Decis Mak.* 2018 Dec 7;18(Suppl 5):119. doi: 10.1186/s12911-018-0694-7. PMID: 30526566; PMCID: PMC6284271.
 27. Ren J, Li G, Ross KE, Arighi CA, McGarvey PB, Rao S, Cowart JC, Madhavan S, Vijay-Shanker K, **Wu CH**. (2018) iTextMine: integrated text-mining system for large-scale knowledge extraction from literature. *Database (Oxford)* 2018. doi: 10.1093/database/bay128 [PMC6301332]
 28. Gupta S, Dingerdissen H, Ross KE, Hu Y, **Wu CH**, Mazumder R, Vijay-Shanker K. DEXTER: Disease-Expression Relation Extraction from Text. *Database (Oxford)* (2018): bay045; doi: 10.1093/database/bay045. PubMed PMID: 29860481; PubMed Central PMCID: PMC6007211.
 29. Cogburn LA, Trakooljul N, Chen C, Huang H, **Wu CH**, Carré W, Wang X, White HB III. (2018) Transcriptional profiling of liver during the critical embryo-to-hatchling transition period in the chicken (*Gallus gallus*). *BMC Genomics* 19(1):695. doi: 10.1186/s12864-018-5080-4
 30. Huang LC, Ross KE, Baffi TR, Drabkin H, Kochut KJ, Ruan Z, D'Eustachio P, McSkimming D, Arighi CN, Chen C, Natale DA, Smith C, Gaudet P, Newton AC, **Wu CH**, Kannan N. (2018) Integrative annotation and knowledge discovery of kinase post-translational modifications and cancer-associated mutations through federated protein ontologies and resources. *Scientific Reports* 8(1):6518. [PMC5916945] doi: 10.1038/s41598-018-24457-1
 31. Huang L, Liao L, **Wu CH**. (2018) Completing sparse and disconnected protein-protein networks by deep learning. *BMC Bioinformatics* 19(1):103. doi: 10.1186/s12859-018-2112-7.
 32. Huang H, Arighi CN, Ross KE, Ren J, Li G, Chen SC, Wang Q, Cowart J, Vijay-Shanker K, **Wu CH**. (2018) iPTMnet: an integrated resource for protein post-translational modification network discovery. *Nucleic Acids Research*, 46(Database issue), D542–D550. doi:10.1093/nar/gkx1104 [PMC5753337]
 33. Huang L, Liao L, **Wu CH**. (2017) Evolutionary analysis and interaction prediction for protein-protein interaction network in geometric space. *PLOS ONE* 12(9):e0183495. [PMC5590856]
 34. Mahmood AS, Rao S, McGarvey PB, **Wu CH**, Madhavan S, Vijay-Shanker K (2017) eGARD: Extracting associations between genomic anomalies and drug responses from text. *PLOS ONE* 12(12):e0189663. [PMC5738129]
 35. Ding R, Boutet E, Lieberherr D, Schneider M, Tognolli M, **Wu CH**, Vijay-Shanker K, Arighi CN. (2017) eGenPub, a text mining system for extending computationally mapped bibliography for UniProt

Knowledgebase by capturing centrality. *Database (Oxford)*. doi: 10.1093/database/bax081. [PMC5691349]

36. The UniProt Consortium (2017) UniProt: the universal protein knowledgebase. *Nucleic Acids Res* 45(D1): D158-D169 [PMC5210571]
37. Natale DA, Arighi CN, Blake JA, Bona J, Chen C, Chen SC, Christie KR, Cowart J, D'Eustachio P, Diehl AD, Drabkin HJ, Duncan WD, Huang H, Ren J, Ross K, Ruttenberg A, Shamovsky V, Smith B, Wang Q, Zhang J, El-Sayed A, **Wu CH**. (2017) Protein Ontology (PRO): enhancing and scaling up the representation of protein entities. *Nucleic Acids Res.* 45(D1): D339-D346 [PMC5210558]
38. Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, Chang HY, Dosztányi Z, El-Gebali S, Fraser M, Gough J, Haft D, Holliday GL, Huang H, Huang X, Letunic I, Lopez R, Lu S, Marchler-Bauer A, Mi H, Mistry J, Natale DA, Necci M, Nuka G, Orengo CA, Park Y, Pesseat S, Piovesan D, Potter SC, Rawlings ND, Redaschi N, Richardson L, Rivoire C, Sangrador-Vegas A, Sigrist C, Sillitoe I, Smithers B, Squizzato S, Sutton G, Thanki N, Thomas PD, Tosatto SC, **Wu CH**, Xenarios I, Yeh LS, Young SY, Mitchell AL. (2017) InterPro in 2017- beyond protein family and domain annotations. *Nucleic Acids Res.* 45(D1): D190-D199 [PMC5210578]
39. The Gene Ontology Consortium (2017). Expansion of the Gene Ontology knowledgebase and resources. *Nucleic Acids Res.* 45(D1): D331-D338 [PMC5210579]
40. Kim S, Dogan RI, Chatr-Aryamontri A, Chang CS, Oughtred R, Rust J, Batista-Navarro RT, Carter J, Ananiadou S, Matos S, Santos A, Campos D, Oliveira J L, Singh O, Jonnalagadda J, Dai H J, Su EC, Chang Y-C, Su Y-C, Chu C-H, Chen CC, Hsu W-L, Peng Y, Arighi C, **Wu CH**, Vijay-Shanker K, Aydin F, Hüsünbeyi Z M, Özgür A, Shin S-Y, Kwon D, Tyers M, Dolinski K, Wilbur WJ, Comeau DC. (2016) BioCreative V BioC track overview: collaborative biocurator assistant task for BioGRID. *Database (Oxford)*. 2016. pii: baw121. doi: 10.1093/database/baw121 [PMC5009341]
41. Wang Q, Abdul SS, Almeida L, Ananiadou S, Balderas-Martínez YI, Batista-Navarro RT, Campos D, Chilton L, Chou H-J, Contreras G, Cooper L, Dai HJ, Ferrell B, Fluck J, Gama-Castro S, Chatr-Aryamontri A, Laulederkind S, Matis-Mitchell S, McEntyre J, Orchard S, Pundir S, Rodriguez-Esteban R, Van Auken K, Lu Z, Schaefer M, **Wu CH**, Hirschman L, Arighi CN (2016) Overview of the interactive task in BioCreative V. *Database (Oxford)*, 2016. pii: baw119. doi:10.1093/database/baw119. [PMC5009325]
42. Du T, Liao L, **Wu CH** (2016) Enhancing interacting residue prediction with integrated contact matrix prediction in protein-protein interaction. *EURASIP Journal on Bioinformatics and Systems Biology* 2016(1):17. [PMC5075339]
43. Du T, Liao L, **Wu CH**, Sun B. Prediction of residue-residue contact matrix for protein-protein interaction with Fisher score features and deep learning. *Methods* 2016 Nov 1; 110:97-105. doi: 10.1016/j.ymeth.2016.06.001. [PMID: 27282356]
44. Huang L, Liao L, **Wu CH** (2016) Protein-protein interaction prediction based on multiple kernels and partial network with linear programming. *BMC Systems Biology* 10(Suppl 2): 45. doi: 10.1186/s12918-016-0296-x [PMC4977483]
45. Gupta S, Ross KE, Tudor CO, **Wu CH**, Schmidt CJ, Vijay-Shanker K (2016) Myriad Ways of Associating microRNAs with Diseases *Journal of Biomedical Semantics* 7, 9 [doi: 101186/s13326-015-0044-y]
46. Peng Y, Arighi CN, **Wu CH**, Vijay-Shanker K (2016) BioC-compatible full-text passage detection for protein-protein interactions using extended dependency graph. *Database (Oxford)* 2016, pii: baw072 [doi: 101093/database/baw072] [PMC4915133]
47. Chen C, Huang H, Mazumder R, Natale DA, McGarvey PB, Zhang J, Polson SW, Wang Y, **Wu CH**, UniProt Consortium (2016) Computational clustering for viral reference proteomes. *Bioinformatics* 32(13):2041-2043. [doi: 101093/bioinformatics/btw110] [PMC4920120]
48. Huang L, Liao L, **Wu CH** (2016) Inference of protein-protein interaction networks from multiple heterogeneous data. *EURASIP Journal on Bioinformatics and Systems Biology* 2016 (1), 8 [PMC4761017]
49. Goodacre N, Edwards N, Danielsen M, Uetz P, **Wu CH** (2017) Predicting nsSNPs that disrupt protein-protein interactions using docking. *IEEE Transactions on Computational Biology and Bioinformatics (TCBB)* 14 (5): 1082-1093 doi: 10.1109/TCBB.2016.2520931 [PMID: 26812731]

50. Celen I, Ross KE, Arighi CN, **Wu CH** (2015) Bioinformatics knowledge map for analysis of beta-catenin function in cancer *PLOS ONE* 10(10): e0141773 [PMC4624812]
51. Resnyk CW, Chen C, Huang H, **Wu CH**, Simon J, Le Bihan-Duval E, Duclos MJ, Cogburn LA (2015) RNA-seq analysis of abdominal fat in genetically fat and lean chickens highlights a divergence in expression of genes controlling adiposity, hemostasis, and lipid metabolism *PLOS ONE* 10, e0139549 [PMC4596860]
52. Li G, Ross KE, Arighi CN, Peng Y, **Wu CH**, Vijay-Shanker K (2015) miRTex: A text mining system for miRNA-gene relation extraction *PLOS Computational Biology* 11(9): e1004391 DOI: 10.1371/journal.pcbi1004391 [PMC4583433]
53. Ding R, Arighi CN, Lee JY, **Wu CH**, Vijay-Shanker K (2015) pGenN, a Gene normalization tool for plant genes and proteins in scientific literature *PLOS ONE* 10(8): e0135305 doi: 10.1371/journal.pone0135305 [PMC4530884]
54. Arighi CN, Shamovsky V, Masci AM, Ruttenberg A, Smith B, Natale DA, **Wu CH**, D'Eustachio P (2015) Toll-Like Receptor signaling in vertebrates: testing the integration of proteins, complexes, and pathway data in the Protein Ontology framework *PLOS One* 10(3), e0122978 doi: 10.1371/journal.pone0122978 [PMC4404318]
55. Tudor CO, Ross KE, Li G, Vijay-Shanker K, **Wu CH**, Arighi CN (2015) Construction of phosphorylation interaction networks by text mining of full-length articles using the eFIP system *Database pii: bav020* doi: 10.1093/database/bav020 [PMC4381107]
56. Torii M, Arighi CN, Li G, Wang Q, **Wu CH**, Vijay-Shanker K (2015) RLIMS-P 2.0: A generalizable rule-based information extraction system for literature mining of protein phosphorylation information *IEEE Transactions on Computational Biology and Bioinformatics (TCBB)* 12 (1): 17-29 doi: 10.1109/TCBB.2014.2372765 [PMC4568560]
57. Selvanathan SP, Graham GT, Erkizan HV, Dirksen U, Natarajan TG, Dakic A, Yu S, Liu X, Paulsen MT, Ljungman ME, **Wu CH**, Lawlor ER, Üren A, Toretsky JA (2015) Oncogenic fusion protein EWS-FLI1 is a network hub that regulates alternative splicing *Proc Natl Acad Sci USA* 112(11), E1307-1316 [PMC4371969]
58. Crowgey E, Stably D, Chen C, Huang H, Polson SW, Sol-Church K, **Wu CH** An Integrated Approach for Analyzing Clinical Genomic Variant Data from Next Generation Sequencing *J Biomol Tech* jbt15-2601-002 [PMC4310222]
59. Gene Ontology Consortium (2015) Gene Ontology Consortium: going forward *Nucleic Acids Res* (Database issue):D1049-1056 [PMC4383973]
60. Crowgey EL, Kolb A, **Wu CH** (2015) Development of bioinformatics pipeline for analyzing clinical pediatric NGS Data *AMIA Joint Summits on Translational Science Proceedings 2015*, 207-211 [PMC4525226]
61. Mitchell A, Chang HY, Daugherty L, Fraser M, Hunter S, Lopez R, McAnulla C, McMenamin C, Nuka G, Pesseat S, Sangrador-Vegas A, Scheremetjew M, Rato C, Yong SY, Bateman A, Punta M, Attwood TK, Sigrist CJ, Redaschi N, Rivoire C, Xenarios I, Kahn D, Guyot D, Bork P, Letunic I, Gough J, Oates M, Haft D, Huang H, Natale DA, **Wu CH**, Orengo C, Sillitoe I, Mi H, Thomas PD, Finn RD (2015) The InterPro protein families database: the classification resource after 15 years *Nucleic Acids Res* 43(Database issue): D213-221 [PMC4383996]
62. The UniProt Consortium (2015) UniProt: a hub for protein information *Nucleic Acids Res* 43(Database issue): D204-212 [PMC4384041]
63. Suzek BE, Wang Y, Huang H, McGarvey P, **Wu CH**, UniProt Consortium (2015) UniRef Clusters: A comprehensive and scalable alternative for improving sequence similarity searches *Bioinformatics* 31(6), 926-932 [PMC4375400]
64. Huang L, Liao L, **Wu CH**. (2015) Evolutionary model selection and parameter estimation for protein-protein interaction network based on differential evolution algorithm. *IEEE Transactions on Computational Biology and Bioinformatics* (TCBB) 12(3):622-631. doi: 10.1109/TCBB.2014.2366748. [PMC4719153]
65. Famiglietti ML, Estreicher A, Gos A, Bolleman J, Géhant S, Breuza L, Bridge A, Poux S, Redaschi N, Bougueleret L, Xenarios I, UniProt Consortium (2014) Genetic variations and diseases in UniProtKB/Swiss-

Prot: the ins and outs of expert manual curation *Hum Mutat* 2014 Aug;35(8):927-35 doi: 101002/humu22594 [PMC4107114]

66. Wyffels J, King BL, Vincent J, Chen C, **Wu CH**, Polson SW (2014) SkateBase, an elasmobranch genome project and collection of molecular resources for chondrichthyan fishes *F1000Research* 3, 191 [PMC4184313]
67. Peng Y, Torii M, **Wu CH**, Vijay-Shanker K (2014) A generalizable NLP framework for fast development of pattern-based biomedical relation extraction systems *BMC Bioinformatics* 15, 285 [Highly Accessed][PMC4262219]
68. Torii M, Li G, Li Z, Oughtred R, Diella F, Çelen I, Arighi CN, Huang H, Vijay-Shanker K, **Wu CH** (2014) RLIMS-P: An online text mining tool for literature-based extraction of protein phosphorylation information *Database* 2014, bau081 [PMC4131691]
69. Peng Y, Tudor CO, Torii M, **Wu CH**, Vijay-Shanker K (2014) iSimp in BioC standard format: Enhancing the interoperability of a sentence simplification system *Database* 2014, bau038 [PMC4028706]
70. Comeau DC, Batista-Navarro RT, Dai HJ, Islamaj Dogan R, Jimeno A, Khare R, Lu Z, Marques H, Mattingly CJ, Neves M, Peng Y, Rak R, Rinaldi F, Tsai RT, Verspoor K, Wiegers TC, **Wu CH**, Wilbur WJ (2014) BioC Interoperability Track Overview *Database* 2014, bau053 [PMC4074764]
71. Chen C, Khaleel SS, Huang H, **Wu CH** (2014) Software for pre-processing Illumina next-generation sequencing short read sequences *Source Code for Biology and Medicine* 9, 8 [PMC4064128] [Highly Accessed]
72. Nguyen NT, Zhang X, **Wu CH**, Lange RA, Chilton RJ, Lindsey ML, Jin YF (2014) Integrative computational and experimental approaches to establish a post-myocardial infarction knowledge map *PLOS Computational Biology* 10(3), e1003472 [PMC3961365]
73. Natale DA, Arighi CN, Blake JA, Bult CJ, Christie K, Cowart J, D'Eustachio P, Diehl A, Drabkin HJ, Helfer O, Huang H, Masci A, Ren J, Roberts N, Ross K, Ruttenberg A, Shamovsky V, Smith B, Yerramalla M, Zhang J, AlJanahi A, Gan C, Çelen I, Lv M, Schuster-Lezell E, **Wu CH** (2014) Protein Ontology: a controlled structured network of protein entities *Nucleic Acids Res* 42(1), D415-421 [PMC3964965] doi: 10.1093/nar/gkt1173
74. UniProt Consortium (2014) Activities at the Universal Protein Resource (UniProt) *Nucleic Acids Res* 42(1), D191-198 [PMC3965022]
75. Wang Q, Venkataramanan KP, Huang H, Papoutsakis ET, **Wu CH** (2013) Transcription factors and genetic circuits orchestrating the complex, multilayered response of *Clostridium acetobutylicum* to butanol and butyrate stress *BMC Systems Biology* 7, 120 [Highly Accessed] [PMC3828012] doi: 10.1186/1752-0509-7-120
76. Crooks DR, Natarajan TG, Jeong SY, Chen C, Park SY, Huang H, Ghosh MC, Tong WH, Haller RG, **Wu CH**, Rouault TA (2013) Elevated FGF21 secretion, PGC-1 α and ketogenic enzyme expression are hallmarks of iron-sulfur cluster depletion in human skeletal muscle *Human Molecular Genetics* 2013, ddt393 [PMC3857942]
77. Chen C, Li Z, Huang H, Suzek BE, **Wu CH**, UniProt Consortium (2013) A Fast Peptide Match Service for UniProt Knowledgebase *Bioinformatics* 2013, btt484 [PMC3799477]
78. Comeau DC, Islamaj Dogan R, Ciccarese P, Cohen KB, Krallinger M, Leitner F, Lu Z, Peng Y, Rinaldi F, Torii M, Valencia A, Verspoor K, Wiegers TC, **Wu CH**, Wilbur WJ (2013) BioC: A Minimalist Approach to Interoperability for Biomedical Text Processing *Database* 2013, bat064 [PMC3889917]
79. Ross KE, Arighi CN, Ren J, Huang H, **Wu CH** (2013) Construction of protein phosphorylation networks by data mining, text mining, and ontology integration: Analysis of the spindle checkpoint *Database* 2013, bat038 [PMC3675891]
80. Lopez LD, Yu J, Arighi C, Tudor CO, Torii M, Huang H, Vijay-Shanker K, **Wu CH** (2013) A framework for biomedical figure segmentation towards image-based document retrieval *BMC Systems Biology* 7(Suppl 4), S8 [PMC3856606]
81. Gana R, Rao S, Huang H, **Wu CH**, Vasudevan S (2013) Structural and functional studies of S-adenosyl-L-methionine binding proteins: a ligand-centric approach *BMC Struct Biol* 13, 6 [PMC3662625] [Highly

Accessed]

82. Gonzalez AJ, Liao L, **Wu CH** (2013) Prediction of contact matrix for protein-protein interaction *Bioinformatics* 29, 1018-1025 [PMC3624801]
83. Ross KE, Arighi CN, Ren J, Natale DA, Huang H, **Wu CH** (2013) Use of the protein ontology for multi-faceted analysis of biological processes: a case study of the spindle checkpoint *Front Genet* 4, 62 [PMC3636526]
84. Arighi CN, Carterette B, Cohen KB, Krallinger M, Wilbur WJ, Fey P, Dodson R, Cooper L, Van Slyke CE, Dahdul W, Mabee P, Li D, Harris B, Gillespie M, Jimenez S, Roberts P, Matthews L, Becker K, Drabkin H, Bello S, Licata L, Chatr-Aryamontri A, Schaeffer ML, Park J, Haendel M, Van Auken K, Li Y, Chan J, Muller HM, Cui H, Balhoff JP, Chi-Yang Wu J, Lu Z, Wei CH, Tudor CO, Raja K, Subramani S, Natarajan J, Cejuela JM, Dubey P, **Wu CH** (2013) An overview of the BioCreative 2012 Workshop Track III: interactive text mining task *Database* 2013, bas056 [PMC3625048]
85. UniProt Consortium (2013) Update on activities at the Universal Protein Resource (UniProt) in 2013 *Nucleic Acids Res* 41(Database issue), D43-47 [PMC3531094]
86. Pedruzzi I, Rivoire C, Auchincloss AH, Coudert E, Keller G, de Castro E, Baratin D, Cuhe BA, Bougueleret L, Poux S, Redaschi N, Xenarios I, Bridge A; UniProt Consortium (2013) HAMAP in 2013, new developments in the protein family classification and annotation system *Nucleic Acids Res* 41(Database issue), D584-589 [PMC3531088]
87. Tudor CO, Arighi CN, Wang Q, **Wu CH**, Vijay-Shanker K (2012) The eFIP system for text mining of protein interaction networks of phosphorylated proteins *Database* 2012, bas044 [PMC3514748]
88. Hirschman L, Burns G, Krallinger M, Arighi C, Cohen KB, Valencia A, **Wu CH**, Chatr-Aryamontri A, Dowell KG, Huala E, Lourenço A, Nash R, Veuthey A-L, Wiegers T, Winter AG (2012) Text mining for the biocuration workflow *Database* 2012, bas020 [PMC3328793]
89. Wang Q, Arighi CN, King BL, Polson SW, Vincent J, Chen C, Huang H, Kingham B, Page ST, Rendino MF, Thomas WK, Udworthy DW, **Wu CH**, North East Bioinformatics Collaborative Curation Team (2012) Community annotation and bioinformatics workforce development in concert – Little skate genome annotation workshops and jamborees *Database* 2012, bar064 [PMC3308154]
90. Gonzalez AJ, Liao L, **Wu CH** (2012) Predicting ligand binding residues and functional sites using multi-positional correlations with graph theoretic clustering and kernel CAA. *IEEE/ACM Trans Comput Biol Bioinform* 9(4), 992-1001, doi: 10.1109/TCBB.2011.136 [PMID: 22025754]
91. UniProt Consortium (2012) Reorganizing the protein space at the Universal Protein Resource (UniProt) *Nucleic Acids Research* 40 (Database issue), D71-75 [PMC3245120]
92. Magrane M, UniProt Consortium (2012) UniProt Knowledgebase: a hub of integrated protein data *Database* 2011, bar009 [PMC3070428]
93. Hunter S, Jones P, Mitchell A, Apweiler R, Attwood TK, Bateman A, Bernard T, Binns D, Bork P, Burge S, de Castro E, Coggill P, Corbett M, Das U, Daugherty L, Duquenne L, Finn RD, Fraser M, Gough J, Haft D, Hulo N, Kahn D, Kelly E, Letunic I, Lonsdale D, Lopez R, Madera M, Maslen J, McAnulla C, McDowall J, McMenamin C, Mi H, Mutowo-Muellenet P, Mulder N, Natale D, Orengo C, Pesseat S, Punta M, Quinn AF, Rivoire C, Sangrador-Vegas A, Selengut JD, Sigrist CJ, Scheremetjew M, Tate J, Thimmajananathan M, Thomas PD, **Wu CH**, Yeats C, Yong SY (2012) InterPro in 2011: new developments in the family and domain prediction database *Nucleic Acids Res* 40(Database issue), D306-312 [PMC3245097]
94. Bult CJ, Drabkin HJ, Evsikov A, Natale D, Arighi C, Roberts N, Ruttenberg A, D'Eustachio P, Smith B, Blake JA, **Wu CH** (2011) The Representation of Protein Complexes in the Protein Ontology (PRO) *BMC Bioinformatics* 12, 371 [Highly Accessed] [PMC3189193]
95. Hu ZZ, Kagan BL, Ariazic EA, Rosenthala DS, Zhanga L, Li JV, Huang H, **Wu CH**, Jordan VC, Riegela AT, Wellsteina A (2011) Proteomic analysis of pathways involved in estrogen-induced growth and apoptosis in breast cancer cells *PLOS One* 6(6), e20410 [PMC3124472]
96. Arighi CN, Lu Z, Krallinger M, Cohen KB, Wilbur J, Valencia A, Hirschman L, **Wu CH** (2011) Overview of the BioCreative III Workshop *BMC Bioinformatics* 12, Suppl 8: p S1 [PMC3269932]
97. Arighi CN, Roberts P, Agarwal S, Bhattacharya S, Cesareni G, Chatr-aryamontri A, Clematide S, Gaudet P,

- Giglio MG, Harrow I, Huala E, Krallinger M, Leser U, Li D, Liu F, Lu Z, Maltais L, Okazaki N, Perfetto L, Rinaldi F, Sætre R, Salgado D, Srinivasan P, Thomas PE, Toldo L, Hirschman L, **Wu CH** (2011) BioCreative III Interactive Task: an Overview *BMC Bioinformatics* 12, Suppl 8: p S4 [PMC3269939]
98. Chen C, Natale DA, Finn RD, Huang H, Zhang J, **Wu CH**, Mazumder R (2011) Representative Proteomes: a stable, scalable and unbiased proteome set for sequence analysis and functional annotation *PLOS One* 6(4), e18910 [PMC3083393]
 99. Huang H, McGarvey PB, Suzek BE, Mazumder R, Zhang J, Chen Y, and **Wu CH** (2011) A comprehensive protein-centric ID mapping service for molecular data integration *Bioinformatics* 27, 1190-1191 [PMC3072559]
 100. Natale DA, Arighi CN, Barker WC, Blake JA, Bult CJ, Caudy M, Drabkin HJ, D'Eustachio P, Evsikov AV, Huang H, Nchoutmboube J, Roberts NV, Smith B, **Wu CH** (2011) The Protein Ontology: A structured representation of protein forms and complexes *Nucleic Acids Research* 39(Database issue), D539-545 [PMC3013777]
 101. UniProt Consortium (2011) Ongoing and future developments at the Universal Protein Resource *Nucleic Acids Research* 39(Database issue):D214-219 [PMC3013648]
 102. Yin L, Xu G, Torii M, Niu Z, **Wu CH**, Hu Z, Liu H (2010) Document classification for mining host pathogen protein-protein interactions *Artificial Intelligence in Medicine* 49, 155-160 [PMC2902599]
 103. Fan J, Traore K, Li W, Amri H, Huang H, **Wu CH**, Chen H, Zirkin B, Papadopoulos V (2010) Molecular mechanisms mediating the effect of mono-(2-ethylhexyl) phthalate (MEHP) on hormone-stimulated steroidogenesis in MA-10 mouse tumor Leydig cells *Endocrinology* 151(7):3348-3362 [PMC2903930]
 104. Tang K, Huang H, Jiao N, **Wu CH** (2010) Phylogenomic analysis of marine Roseobacters *PLOS One* 5(7):e11604 [PMC2904699]
 105. Hinz U; UniProt Consortium (2010) From protein sequences to 3D-structures and beyond: the example of the UniProt Knowledgebase *Cellular and Molecular Life Sciences* 67(7), 1049-1064 [PMC2835715]
 106. UniProt Consortium (2010) The Universal Protein Resource (UniProt) in 2010 *Nucleic Acids Res* 38(Database issue):D142-8 [PMC2808944]
 107. McGarvey PB, Huang H, Mazumder R, Zhang J, Chen Y, Zhang C, Cammer S, Will R, Odle M, Sobral B, Moore M, **Wu CH** (2009) Systems integration of biodefense omics data for analysis of pathogen-host interactions and identification of potential targets *PLOS One* 4: e7162 [PMC2745575]
 108. Arighi CN, Liu H, Natale D, Barker WC, Drabkin HJ, Blake J, Smith B, **Wu CH** (2009) TGF-beta Signaling Proteins and the Protein Ontology *BMC Bioinformatics* 10 Suppl 5:S3 [PMC2679403]
 109. Barker WC, Mazumder R, Vasuvedan S, Sagripanti J-L, **Wu CH** (2009) Sequence signatures in envelope protein may determine whether flaviviruses produce hemorrhagic or encephalitic syndromes *Virus Genes* 39, 1-9 [PMID: 19283462]
 110. Torii M, Hu Z, **Wu CH**, Liu H (2009) BioTagger-GM: A gene/protein name recognition system *J Am Med Inform Assoc* 16, 247-255 [PMC2649315]
 111. Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bork P, Das U, Daugherty L, Duquenne L, Finn RD, Gough J, Haft D, Hulo N, Kahn D, Kelly E, Laugraud A, Letunic I, Lonsdale D, Lopez R, Madera M, Maslen J, McAnulla C, McDowall J, Mistry J, Mitchell A, Mulder N, Natale D, Orengo C, Quinn AF, Selengut JD, Sigrist CJ, Thimma M, Thomas PD, Valentin F, Wilson D, **Wu CH**, Yeats C (2009) InterPro: the integrative protein signature database *Nucleic Acids Res* 37(Database issue):D211-5 [PMC2686546]
 112. UniProt Consortium – Bairoch A, et al (SIB), Apweiler R, et al (EBI), **Wu CH**, et al (PIR) (2009) The Universal Protein Resource (UniProt) 2009 *Nucleic Acids Res* 37(Database issue):D169-74
 113. Hu ZZ, Huang H, Cheema A, Jung M, Dritschilo A, **Wu CH** (2008) Integrated bioinformatics for radiation-induced pathway analysis from proteomics and microarray data *J Proteomics & Bioinformatics* 1, 47-60 [PMC2603135]
 114. UniProt Consortium – Bairoch A, et al (SIB), Apweiler R, et al (EBI), **Wu CH**, et al (PIR) (2008) The Universal Protein Resource (UniProt) *Nucleic Acids Research* 36(Database issue), D190–D195
 115. Zhang C, Crasta O, Cammer S, Will R, Kenyon R, Sullivan D, Yu Q, Wei S, Jha R, Liu D, Xue T, Zhang Y, Moore

- M, McGarvey P, Huang H, Chen Y, Zhang J, Mazumder R, **Wu CH**, Sobral B (2008) An emerging cyberinfrastructure for biodefense pathogen and pathogen-host Data *Nucleic Acids Research* 36(Database issue), D884–D891 [PMC2239001]
116. Natale DA, Arighi CN, Hu ZZ, Liu HF, Blake J, Smith B, **Wu CH** (2007) Framework for a protein ontology *BMC Bioinformatics* 8 (Suppl 9), S1 [Faculty of 1000 Biology] [PMC2217659]
 117. Torii M, Hu ZZ, Song M, **Wu CH**, Liu HF (2007) A comparison study on algorithms of detecting long forms for short forms in biomedical text *BMC Bioinformatics* 8 (Suppl 9), S5 [PMC2217663]
 118. Mazumder R, Hu ZZ, Vinayaka CR, Sagripanti J-L, Frost SDW, Pond SLK, **Wu CH** (2007) Computational analysis and identification of amino acid sites in dengue E proteins relevant to development of diagnostics and vaccines *Virus Genes* 35, 175-186 [PMID: 17508277]
 119. Huang H, Hu ZZ, Arighi C, **Wu CH** (2007) Integration of bioinformatics resources for functional analysis of gene expression and proteomic data *Frontiers in Bioscience* 12, 5071-5088 [PMID: 17569631]
 120. Suzek BE, Huang H, McGarvey P, Mazumder R, **Wu CH** (2007) UniRef: comprehensive and non-redundant UniProt reference clusters *Bioinformatics* 23, 1282-1288 [PMID: 17379688]
 121. Qiu P, Wang ZJ, Liu KJ, Hu ZZ, **Wu CH** (2007) Dependence network modeling for biomarker identification *Bioinformatics* 23, 198-206 [PMID: 17077095]
 122. Mulder NJ, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bork P, et al, **Wu CH**, Yeats C (2007) New developments in the InterPro database *Nucleic Acids Res* 35(Database issue), D224-228 [PMID: 17202162]
 123. UniProt Consortium – Bairoch A, et al (SIB), Apweiler R, et al (EBI), **Wu CH**, et al (PIR) (2007) The Universal Protein Resource (UniProt) *Nucleic Acids Res* 35(Database issue), D193-197
 124. Huang H, Shukla HD, **Wu CH**, Saxena S (2007) Challenges and solutions in proteomics *Current Genomics* 8, 21-28
 125. Hu ZZ, Valencia JC, Huang H, Chi A, Shabanowitz J, Hearing VJ, Appella E, **Wu CH** (2007) Comparative bioinformatics analyses and profiling of lysosome-related organelle proteomes *International Journal of Mass Spectrometry* 259, 147-160 [PMC1828028]
 126. Chi A, Valencia JC, Hu ZZ, Watabe H, Yamaguchi H, Mangini NJ, Huang H, Canfield VA, Cheng KC, Yang F, Abe R, Yamagishi S, Shabanowitz J, Hearing VJ, **Wu, CH**, Appella E and Hunt DF (2006) Proteomic and bioinformatic characterization of the biogenesis and function of melanosomes *Journal of Proteome Research* 5, 3135-3144 [PMID: 17081065]
 127. Han B, Obradovic Z, Hu ZZ, **Wu, CH**, Vucetic S (2006) Substring selection for biomedical document classification *Bioinformatics* 22, 2136-2142 [PMID: 16837530]
 128. Liu H, Hu ZZ, Manabu T, **Wu, CH**, Friedman, C (2006) Quantitative assessment of dictionary-based protein named entity tagging *J Am Med Inform Assoc* 13, 497-507 [PMC1561801]
 129. Petrova NV, **Wu CH** (2006) Prediction of catalytic residues using Support Vector Machine with selected protein sequence and structural properties *BMC Bioinformatics* 7, 312 [Faculty of 1000 Biology] [PMC1534064]
 130. Nikolskaya AN, Arighi C, Huang H, Barker WC, **Wu, CH** (2006) PIRSF family classification system for protein functional and evolutionary analysis *Evolutionary Bioinformatics Online* 2, 209-221 [PMC2674652]
 131. Yuan X, Hu ZZ, Wu HT, Torii M, Narayanaswamy M, Ravikumar KE, Vijay-Shanker K, **Wu, CH** (2006) An online literature mining tool for protein phosphorylation *Bioinformatics* 22, 1668-1669 [PMID: 16644790]
 132. **Wu, CH**, Apweiler, R, Bairoch, A, Natale, DA, Barker, WC, Boeckmann, B, Ferro, S, Gasteiger, E, Huang, H, Lopez, R, Magrane, M, Martin, M J, Mazumder, R, O'Donovan, C, Redaschi, N and Suzek, B (2006) The Universal Protein Resource (UniProt): an expanding universe of protein information *Nucleic Acids Research*, 34, D187-191
 133. Liu, HF, Hu, ZZ, Zhang, J, **Wu, CH** (2006) BioThesaurus: a web-based thesaurus of protein and gene names *Bioinformatics* 22, 103-105 [PMID: 16267085]
 134. Mazumder, R, Natale, D, Murthy, S, Thiagarajan, R, **Wu, CH** (2005) Computational identification of strain-, species- and genus-specific proteins *BMC Bioinformatics* 6, 279 [PMC1310627]

135. Hu, ZZ, Narayanaswamy, M, Ravikumar, KE, Vijay-Shanker, K and **Wu, CH** (2005) Literature mining and database annotation of protein phosphorylation using a rule-based system *Bioinformatics* 21, 2759–2765 [PMID: 15814565]
136. Liu HF, Hu ZZ, **Wu, CH** (2005) DynGO: a tool for visualizing and mining of Gene Ontology and its associations *BMC Bioinformatics* 6, 201 [PMC1199584]
137. Mani I, Hu ZZ, Jang SE, Samuel K, Krause M, Phillips J, **Wu, CH** (2005) Protein name tagging guidelines: lessons learned *Comparative & Functional Genomics*, 6, 72-76 [PMC2448601]
138. Schneider, M, Bairoch, A, **Wu, CH** and Apweiler, R (2005) Plant protein annotation in the UniProt Knowledgebase *Plant Physiology*, 138, 59–66 [PMC1104161]
139. Bairoch, A, Apweiler, R, **Wu, CH**, Barker, WC, Boeckmann, B, Ferro, S, Gasteiger, E, Huang, H, Lopez, R, Magrane, M, Martin, M J, Natale, DA, O Donovan, C, Redaschi, N and Yeh, LS (2005) The Universal Protein Resource (UniProt) *Nucleic Acids Research*, 33, D154-159
140. Mulder NJ, Apweiler, R, TKAttwood, ABairoch, ABateman, DBinns, PBradley, PBork, PBucher, LCerutti, RCopley, ECourcelle, UDas, RDurbin, WFlerschmann, Julian Gough 9, DHaft, NHarte, NHulo, DKahn, AKanapin, MKrestyaninova, DLonsdale, RLopez, lLetunic, MMadera, JMaslen, JMcDowall, ANNikolskaya, et al CSigrist, VSilventoinen, DJStudholme, RVAughan and **Wu, CH** (2005) InterPro, progress and status in 2005 *Nucleic Acids Research*, 33, D201-205 [PMC540060]
141. **Wu, CH** and Nebert, DW (2004) Update on human genome completion and annotations: Protein Information Resource *Human Genomics*, 1 (3), 229-233 [PMC3525084]
142. **Wu, CH**, Huang H, Nikolskaya A, Hu Z, Barker WC (2004) The iProClass integrated database for protein functional analysis *Computational Biology and Chemistry*, 28, 87-96
143. **Wu, CH**, Nikolskaya A, Huang H, Yeh L-S, Natale D, Vinayaka CR, Hu Z, Mazumder R, Kumar S, Kourtesis P, Ledley RS, Suzek BE, Arminski L, Chen Y, Zhang J, Cardenas JL, Chung S, Castro-Alvear J, Dinkov G and Barker WC (2004) PIRSF family classification system at the Protein Information Resource *Nucleic Acids Research*, 32, D112-114
144. Apweiler R, Bairoch A, **Wu, CH**, Barker, WC, Boeckmann, B, Ferro1, S, Gasteiger, E, Huang, H, Lopez, R, Magrane, M, Martin, M J, Natale, DA, O Donovan, C, Redaschi, N, Yeh, LS (2004) UniProt: Universal Protein Knowledgebase *Nucleic Acids Research*, 32, D115-119
145. Hu, Z, Mani, I, Hermoso, V, Liu, H and **Wu, CH** (2004) iProLINK: An integrated protein resource for literature mining *Computational Biology and Chemistry*, 28, 409-416
146. Huang H, Hu, ZZ, Suzek, BE and **Wu, CH** (2004) The PIR integrated protein databases and data retrieval system *Data Science* 3, 163-174
147. Li, W, Amri, H, Huang, H, **Wu, CH** and Papadopoulos, V (2004) Gene and protein profiling of the acute response of MA-10 Leydig tumor cells to human chorionic gonadotropin *Journal of Andrology*, 25, 900-913 [PMID: 15477362]
148. Herbert, KG, Gehan, NH, Piel, WH, Wang, JTL and **Wu, CH** (2004) BIO-AJAX: An extensible framework for biological data cleaning *SIGMOD Record*, 33, 51-57
149. **Wu, CH**, Yeh, L-S, Huang, H, Arminski, L, Castro-Alvear, J, Chen, Y, Hu, Z, Kourtesis, P, Ledley, R S, Suzek, BE, Vinayaka, CR, Zhang, J and Barker, WC (2003) The Protein Information Resource *Nucleic Acids Research*, 31, 345-347
150. Huang, H, Barker, WC, Chen, Y and **Wu, CH** (2003) iProClass: An integrated database of protein family, function, and structure information *Nucleic Acids Research*, 31, 390-392
151. **Wu, CH** (2003) PIR protein knowledge base for functional proteomics *Experimental Lung Research*, 29, Supplement 1, 27-31
152. **Wu, CH**, Huang, H, Arminski, L, Castro-Alvear, J, Chen, Y, Hu, Z, Ledley, R S, Lewis, K C, Mewes, H-W, Orcutt, BC, Suzek, BE, Tsugita, A, Vinayaka, CR, Yeh, L-S, Zhang, J and Barker, WC (2002) The Protein Information Resource: an integrated public resource of functional annotation of proteins *Nucleic Acids Research*, 30, 35-37
153. Ma, Q, Wang, J T L, Shasha, D and **Wu, CH** (2002) DNA sequence classification via an expectation

maximisation algorithm and neural networks: A case study *IEEE Transactions on Systems, Man, and Cybernetics* - Part C: Special Issue on Knowledge Management

154. **Wu, CH**, Xiao, C, Hou, Z, Huang, H, and Barker, W C (2001) *iProClass: An integrated and comprehensive protein classification database* *Nucleic Acids Research*, 29, 52-54
155. Barker, WC, Garavelli, JS, Hou, Z, Huang, H, Ledley, RS, McGarvey, PB, Mewes, H-W, Orcutt, BC, Pfeiffer, F, Tsugita, A, Vinayaka, CR, Xiao, C, Yeh, L-S and **Wu, CH** (2001) Protein Information Resource: a community resource for expert annotation of protein data *Nucleic Acids Research*, 29, 29-32
156. Wang, J T L, Ma, Q, Shasha, D and **Wu, CH** (2001) An Empirical Study of Bioinformatics Tools for Protein Sequence Classification *IBM Systems Journal*, Special Issue on Deep Computing for Life Sciences, 40 (2), 426-441
157. Huang, H, Xiao, C and **Wu, CH** (2000) ProClass protein family database *Nucleic Acids Research*, 28, 273-276
158. McGarvey, P, Huang, H, Barker, W C, Orcutt, B C and **Wu, CH** (2000) The PIR Web site: New resource for bioinformatics *Bioinformatics*, 16, 290-291 [PMID: 10869023]
159. Barker, W C, Garavelli, J S, Huang, H, McGarvey, P B, Marzec, C R, Orcutt, B C, Srinivasarao, G Y, Xiao, C, Yeh, L-S, Ledley, R S, Mewes, H-W, Pfeiffer, F, Tsugita, A and **Wu, CH** (2000) The Protein Information Resource *Nucleic Acids Research*, 28, 41-44
160. **Wu, CH**, H Huang and J McLarty (1999) Gene family identification network design for protein sequence analysis *International Journal of Artificial Intelligence Tools*, Special Issue on Biocomputing, 8 (4), 419-432
161. **Wu, CH**, S Shivakumar and H Huang (1999) ProClass protein family database *Nucleic Acids Research*, 27, 272-274
162. Barker, W C, Garavelli, J S, McGarvey, P B, Marzec, C R, Orcutt, B C, Srinivasarao, GY, Yeh, L-S, Ledley, R S, Mewes, HW, Pfeiffer, F, Tsugita, A and **Wu, CH** (1999) The PIR-International Protein Sequence Database *Nucleic Acids Research*, 27, 39-43
163. **Wu, CH**, S Shivakumar, C V Shivakumar and S Chen (1998) GeneFIND web server for protein family identification and information retrieval *Bioinformatics*, 14, 223-224 [PMID: 9545458]
164. **Wu, CH**, S Shivakumar and W C Barker (1997) Protein family identification and information retrieval using a ProClass database and a motif neural design *Mathematical Modelling and Scientific Computing*, 8
165. **Wu, CH**, S Zhao and W C Barker (1997) Family identification system for electron transfer proteins using a motif neural network design *Protein and Peptide Letters*, 4, 9-16
166. **Wu, CH**, H L Chen and S Chen (1997) Counter-propagation neural networks for molecular sequence classification: Supervised LVQ and dynamic node allocation *Applied Intelligence*, 7, 27-38
167. **Wu, CH**, S Zhao and H L Chen (1996) A protein class database organized with ProSite protein groups and PIR superfamilies *Journal of Computational Biology*, 3 (4), 547-562
168. **Wu, CH**, S Zhao, H L Chen, C J Lo and J McLarty (1996) Motif identification neural design for rapid and sensitive protein family search *CABIOS (Computer Applications in the BioSciences)*, 12 (2), 109-118 [PMID: 8744773]
169. **Wu, CH** (1996) Gene Classification Artificial Neural System *Methods In Enzymology*, 266, 71-88
170. **Wu, CH** and C Y Tsai (1996) Destiny of a 57 Kilodalton protein in maize genotypes responding to heat shock and pathogenic infections *Taiwania*, 41 (4), 329-338
171. **Wu, CH**, H L Chen and S C Chen (1995) Gene classification artificial neural system *International Journal on Artificial Intelligence Tools*, 4 (4), 501-510
172. **Wu, CH**, M Berry, S Shivakumar and J McLarty (1995) Neural networks for full-scale protein sequence classification: Sequence encoding with singular value decomposition *Machine Learning*, 21, 177-193
173. **Wu, CH**, S Shivakumar, H Lin, S Veldurti, and Y Bhatikar (1995) Neural networks for molecular sequence classification *Mathematics & Computers in Simulation*, 40, 23-33
174. **Wu, CH** and S Shivakumar (1994) Back-propagation and counter-propagation neural networks for phylogenetic classification of ribosomal RNA sequences *Nucleic Acids Research*, 22(20), 4291-4299 [PMC331947]
175. **Wu, CH** (1993) Classification neural networks for rapid sequence annotation and automated database

- organization *Computers & Chemistry*, 17(2), 219-227
176. **Wu, CH** and G Whitson (1993) Neural network database systems for genetic sequence classification *Mathematical Modeling and Scientific Computing*, Vol 2, 474-479
 177. **Wu, CH**, G Whitson, J McLarty, A Ermongkonchai and T Chang (1992) Protein classification artificial neural system *Protein Science*, 1, 667-677
 178. **Wu, CH**, H Warren, L Lyznik and C Tsai (1991) Nuclear RNA Polymerase II in maize leaves infected with *Bipolaris maydis* *Maydica*, 36(3): 205-211
 179. **Wu, CH**, H Warren and C Tsai (1991) Accumulation and binding of a polysome-associated-protein in maize responding to infection by *Bipolaris maydis* *Taiwania*, 36: 111-116
 180. **Wu, CH**, T Caspar, J Browse, S Lindquist and C Somerville (1988) Characterization of an HSP70 cognate gene family in *Arabidopsis* *Plant Physiology*, 88: 731-740 [PMC1055652]
 181. **Wu, CH**, HL Warren, K Sitaraman and CY Tsai (1988) Translational alterations in maize leaves responding to pathogen infection, paraquat treatment or heat shock: polysome dissociation and accumulation of a 57 kilodalton protein *Plant Physiology*, 86: 1323-1329 [PMC1054672]
 182. **Wu, CH** and H L Warren (1984) Natural autofluorescence in fungi and correlation with viability *Mycologia*, 76: 1049-1058
 183. **Wu, CH** and H L Warren (1984) Induced autofluorescence in fungi and correlation with viability: potential application of fluorescence microscopy *Phytopathology*, 74: 1353-1358
 184. Wu, W S, **Wu, CH** and K C Wu (1979) *Alternaria brassicicola*, a destructive pathogen in seed production of cauliflower *Plant Protection Bulletin (Taiwan)*, 21: 294-304

2) Reviews or Editorials in Refereed Journals

185. Garcia L, Bolleman J, Gehant S, Redaschi N, Martin M, UniProt Consortium. (2019) FAIR adoption, assessment and challenges at UniProt. *Sci Data* 2019 6(1):175. doi: 10.1038/s41597-019-0180-9. [PMC6754384]
186. Arighi CN, **Wu CH**, Cohen KB, Hirschman L, Krallinger M, Valencia A, Lu Z, Wilbur J, Wieggers T (2014) BioCreative-IV Virtual Issue *Database* 2014, bau039 [PMC4030502]
187. **Wu CH**, Arighi CN, Cohen KB, Hirschman L, Krallinger M, Lu Z, Mattingly C, Valencia A, Wieggers TC, Wilbur WJ (2012) BioCreative-2012 Virtual Issue, *Database* 2012, bas049 [PMC3514749]
188. Hood LE, Omenn GS, Moritz RL, Aebersold R, Yamamoto KR, Amos M, Hunter-Cevera J, Locascio L; Anderson L, Baer T, Banfield J, Beanan M, Beretta L, Boehm J, Boggess M, Britz J, Caprioli R, Carr S, Carter K, *et al*, Thompson D, Tsien R, Wang L, Winick W, Wispelwey J, **Wu CH**, Zangmeister R (2012) New and improved proteomics technologies for understanding complex biological systems: addressing a grand challenge in the life sciences *Proteomics* 12(18), 2773-2783 [PMC4005326]
189. Gaudet P, Arighi C, Bastian F, Bateman A, Blake JA, Cherry MJ, D'Eustachio P, Finn R, Giglio M, Hirschman L, Kania R, Klimke W, Martin MJ, Karsch-Mizrachi I, Munoz-Torres M, Natale D, O'Donovan C, Ouellette F, Pruitt KD, Robinson-Rechavi M, Sansone SA, Schofield P, Sutton G, Van Auken K, Vasudevan S, **Wu CH**, Young J, Mazumder R (2012) Recent advances in biocuration: Meeting Report from the 5th International Biocuration Conference *Database* 2012, bas036 [PMC3483532]
190. Uhler PF (Editor) (2011) *Designing the Microbial Research Commons: Proceedings of an International Workshop* Preface, Introduction, Closing Observations by **Wu CH**, Steering Committee Chair The National Academies Press ISBN: 978-0-309-21979-2 [PMID: 22593950]
191. Legrain P, Aebersold R, Archakov A, Bairoch A, Bala K, Beretta L, Bergeron J, Borchers C, Corthals GL, Costello CE, Deutsch EW, Domon B, Hancock W, He F, Hochstrasser D, Marko-Varga G, Salekdeh GH, Sechi S, Snyder M, Srivastava S, Uhlen M, **Wu CH**, Yamamoto T, Paik YK, Omenn GS (2011) The human proteome project: Current state and future direction *Mol Cell Proteomics* 10 (7), M111009993 [PMC3134076]
192. Chen C, Huang H, **Wu CH** (2011) Protein bioinformatics databases and resources *Methods in Molecular Biology* 694, 3-24 [review] [PMID: 21082424]
193. Chen C, McGarvey PB, Huang H, **Wu CH** (2010) Protein bioinformatics infrastructure for the integration

and analysis of multiple high-throughput omics data *Advances in Bioinformatics* 2010, Article ID 423589, 19 pages, doi:10.1155/2010/423589 [review] [PMID: 20369061]

194. Mazumder R, Natale DA, Julio JA, Yeh LS, **Wu CH** (2010) Community annotation in biology *Biology Direct* 5, 12 [opinion] [Highly Accessed] [PMC2834641]
195. Toronto International Data Release Workshop Authors (2009) Prepublication data sharing *Nature* 461, 168-170 [opinion] [PMC3073843]
196. Apweiler, R, Bairoch, A and **Wu, CH** (2004) Protein sequence databases *Current Opinion in Chemical Biology*, 8, 76-80 [review][PMID: 15036160]
197. **Wu, CH**, Huang, H, Yeh, L-S and Barker, WC (2003) Protein family classification and functional annotation *Computational Biology and Chemistry*, 27, 37-47 [review]
198. Hirschman, L, Park, JC, Tsujii, J, Wong, L and **Wu, CH** (2002) Accomplishments and challenges in literature data mining for biology *Bioinformatics*, 18, 1553-1561 [review] [PMID: 12490438]
199. Peck, J, Douglas, G, IV, **Wu, CH**, and Burbelo, PD (2002) Human RhoGAP domain-containing proteins: Structure, function and evolutionary relationships *FEBS Letters*, 528, 27-34 [review][PMID: 12297274]
200. **Wu, CH** (1997) Artificial neural networks for molecular sequence analysis *Computers & Chemistry*, 21(4), 237-256 [review][PMID: 9415987]

3) Publications in Refereed Conference Proceedings and Book Chapters

201. Gavali S, Ross KE, Cowart J, Chen C, **Wu CH**. (2022) iPTMnet RESTful API for post-translational modification network analysis. *Methods in Molecular Biology – Computational Methods for Predicting Post-Translational Modification Sites* (in press) <https://link.springer.com/book/9781071623169>
202. Gavali S, Chen C, Cowart J, Peng X, Ding S, **Wu CH**, Anderson T. (2021) Understanding the Factors Driving the Opioid Epidemic Using Machine Learning. *Proceedings of IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 1309-1314. doi: 10.1109/BIBM52615.2021.9669486.
203. Eigenmann R, Bagozzi BE, Jayaraman A, Totten B, **Wu CH**. (2021). DARWIN - A Resource for Computational and Data-intensive Research at the University of Delaware and in the Delaware Region, University of Delaware, 2021, URL: <https://udspace.udel.edu/handle/19716/29071>
204. Ma M, Zhao L, Ren J, Tulyakov S, **Wu CH**, Peng X. (2021). SMIL: Multimodal Learning with Severely Missing Modality. *Proceedings of the AAAI Conference on Artificial Intelligence* 35 (3), 2302-2310. <https://www.aaai.org/AAAI21Papers/AAAI-437.MaM.pdf>
205. Li Q, Wang X, Zhang Y, Ling F, **Wu CH**, Han J. (2018) Pattern Discovery for Wide-Window Open Information Extraction in Biomedical Literature. In *Proceedings of IEEE International Conference on Bioinformatics and Biomedicine (BIBM-2018)*
206. Wang X, Zhang Y, Li Q, **Wu CH**, Han J. (2018) PENNER: Pattern-enhanced Nested Named Entity Recognition in Biomedical Literature. In *Proceedings of IEEE International Conference on Bioinformatics and Biomedicine (BIBM-2018)*
207. Ding R, Pan X, Hao T, **Wu CH**, Vijay-Shanker K. (2018) Protein Complex Mention Recognition with Web-based Knowledge learning. *The 2nd International Symposium on User Modeling and Language Learning (UMLL)*, August 22-24, 2018, Thailand.
208. Arighi C, Hirschman L, Lemberger T, Bayer S, Liechti R, Comeau D, **Wu CH**. (2017) Bio-ID Track Overview. In *Proceedings of the BioCreative VI Challenge Evaluation Workshop*: 14-19.
209. Gupta S, Mahmood AS, Ross K, **Wu CH**, Vijay-Shanker K (2017) Identifying comparative structures in biomedical text. Conference of the Association for Computational Linguistics, Vancouver, Canada; In *Proceedings of Workshop on Biomedical Natural Language Processing (ACL BioNLP-2017)*: 206-215. DOI: 10.18653/v1/W17-2326
210. Li G, **Wu CH**, Vijay-Shanker K (2017) Noise reduction methods for distantly supervised biomedical relation extraction. Conference of the Association for Computational Linguistics, Vancouver, Canada; In *Proceedings of Workshop on Biomedical Natural Language Processing (ACL BioNLP-2017)*: 184-193. DOI: 10.18653/v1/W17-2323

211. Chen C, Huang H, **Wu CH**. (2017) Protein bioinformatics databases and resources. *Methods in Molecular Biology* 1558: 3-39. doi: 10.1007/978-1-4939-6783-4_1. [PMID: 28150231]
212. Wang Q, Ross KE, Huang H, Ren J, Li G, Vijay-Shanker K, **Wu CH**, Arighi CN. (2017) Analysis of protein phosphorylation and its functional impact on protein–protein interactions via text mining of the scientific literature. *Methods in Molecular Biology* 1558: 213-232. doi: 10.1007/978-1-4939-6783-4_10. [PMID: 28150240]
213. Ross KE, Huang H, Ren J, Arighi CN, Li G, Tudor CO, Lv M, Lee JY, Chen SC, Vijay-Shanker K, **Wu CH**. (2017) iPTMnet: integrative bioinformatics for studying PTM networks. *Methods in Molecular Biology* 1558: 333-353. doi: 10.1007/978-1-4939-6783-4_16. [PMC5436273]
214. Ross KE, Natale DA, Arighi C, Chen SC, Huang H, Li G, Ren J, Wang M, Vijay-Shanker K, **Wu CH** (2016) Scalable Text Mining Assisted Curation of Post-Translationally Modified Proteoforms in the Protein Ontology *Proceedings of the 7th International Conference on Biomedical Ontology (ICBO 2016)*, http://ceur-ws.org/Vol-1747/BIT103_ICBO2016.pdf
215. Peng Y, Arighi C, **Wu CH**, Vijay-Shanker K (2015) Extended dependency graph for BioC-compatible PPI passage detection in full-text articles In *Proceedings of the BioCreative V Workshop*
216. Huang L, Liao L, **Wu CH** (2015) Protein-protein interaction network inference from multiple kernels with optimization based on random walk by linear programming In *Proceedings of IEEE International Conference on Bioinformatics and Biomedicine (BIBM-2015)*
217. Peng Y, Gupta S, **Wu CH**, Vijay-Shanker K (2015) An extended dependency graph for relation extraction in biomedical texts. Conference of the Association for Computational Linguistics, Beijing, China; In *Proceedings of Workshop on Biomedical Natural Language Processing (ACL BioNLP-2015)*: 21-30
218. Goodacre N, Edwards N, Danielsen M, Uetz P, **Wu CH** (2014) Docking features for predicting binding loss due to protein mutation. *Proceedings of the 5th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (ACM-BCB 2014)*, Association of Computing Machinery, 2014
219. Du T, González A, Wang Q, Huang H, Liao L, **Wu CH** (2014) DDI2PPI: an integrated web server for protein-protein interaction and residue contact matrix predictions. *Proceedings of the 5th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (ACM-BCB 2014)*, 641-642
220. Gupta S, Tudor CO, **Wu CH**, Schmidt CJ, Vijay-Shanker K (2014) Automatically Identifying Biological Functions of microRNAs from the Literature *Proceedings of the 6th International Symposium on Semantic Mining in Biomedicine (SMBM 2014)*
221. Ross K, Tudor CO, Li G, Ding R, Celen I, Cowart J, Arighi CN, Natale D, **Wu CH** (2014) Knowledge Representation of Protein PTMs and Complexes in the Protein Ontology: Application to Multi-Faceted Disease Analysis *Proceedings of the 5th International Conference on Biomedical Ontology (ICBO 2014)*, CEUR-Workshop Proceedings, 1327, 43-46 Houston, Texas, USA, Oct 6-9, 2014
222. Peng Y, Tudor CO, Torii M, **Wu CH**, Vijay-Shanker K (2013) Enhancing the interoperability of iSimp by using the BioC format *Proceedings of the Fourth BioCreative Challenge Evaluation Workshop, Vol 1*, 5-9
223. Torii M, Li G, Li Z, Çelen I, Diella F, Oughtred R, Arighi CN, Huang H, Vijay-Shanker K, **Wu CH** (2013) RLIMS-P: Literature-based curation of protein phosphorylation information *Proceedings of the Fourth BioCreative Challenge Evaluation Workshop, Vol 1*, 247-253
224. Torii M, Arighi CN, Wang Q, **Wu CH**, Vijay-Shanker K (2013) Text mining of protein phosphorylation information using a generalizable rule-based approach In *Proceedings of (ACMBCB-2013)*
225. Lopez L, Yu J, Arighi CN, Torii M, Vijay-Shanker K, Huang H, **Wu CH** (2013) An Image-Text Approach for Extracting Experimental Evidence of Protein-Protein Interactions in the Biomedical Literature In *Proceedings of (ACMBCB-2013)*
226. Du T, Liao L, **Wu CH** (2013) Prediction of Protein-Protein Interaction Sites at Interface Topology Level *BIOCOMP'13: The 14th International Conference on Bioinformatics and Computational Biology*, July 22-25, 2013, Las Vegas, USA
227. Schmidt CJ, Sun L, Arighi CN, Decker K, Vijay-Shanker K, Torii M, Tudor CO, **Wu CH**, D'Eustachio P (2012) Pathway Curation: Application of Text Mining Tools eGIFT and RLIMS-P, In *IEEE International Conference*

- on *Bioinformatics and Biomedicine Workshops*, 523-528 doi: 101109/BIBMW20126470377
228. Peng Y, Tudor CO, Torii M, **Wu CH**, K Vijay-Shanker (2012) iSimp: A Sentence Simplification System for Biomedical Text In *Proceedings of IEEE International Conference on Bioinformatics and Biomedicine* (BIBM-2012), 211-216 [Acceptance rate: 19%] doi: 101109/BIBM20126392671
 229. Lopez L, Yu J, Tudor CO, Arighi C, Huang H, Vijay-Shanker K, **Wu CH** (2012) Robust segmentation of biomedical figures toward an image-based document retrieval In *Proceedings of IEEE International Conference on Bioinformatics and Biomedicine* (BIBM-2012) [Acceptance rate: 19%]
 230. Bi X, Huang H, Matis-Mitchell S, McGarvey PB, Shatkay H, Torii M, **Wu CH** (2012) Building a classifier for identifying sentences pertaining to disease-drug relationships in tardive dyskinesia In *Proceedings of IEEE International Conference on Bioinformatics and Biomedicine* (BIBM-2012)
 231. Lopez L, Arighi C, Shatkay H, Huang H, **Wu CH**, Yu J (2011) A system for extracting figures and captions in biomedical PDF documents In *Proceedings of IEEE International Conference on Bioinformatics and Biomedicine* (BIBM-2011), 578-581
 232. Hu ZZ, Huang H, **Wu CH**, Jung M, Dritschilo A, Riegel AT, Wellstein A (2011) Omics-based molecular target and biomarker identification *Methods in Molecular Biology* 719, 547-571 [PMID: 21370102]
 233. Arighi C, Siu A, Tudor CO, Nchoutmboube J, **Wu CH**, Shanker VK (2011) eFIP: A tool for mining functional impact of phosphorylation from literature *Methods in Molecular Biology* 694, 63-75 [PMID: 21082428]
 234. Vasudevan S, Vinayaka CR, Natale DA, Huang H, Kaysay RY, **Wu CH** (2011) Structure-guided rule-based annotation of protein functional sites in UniProt knowledgebase *Methods in Molecular Biology* 694, 91-105 [PMID: 21082430]
 235. McGarvey PB, Zhang J, Natale DA, **Wu CH**, Huang H (2011) Protein-centric data integration for functional analysis of comparative proteomics data *Methods in Molecular Biology* 694, 323-339 [PMID: 21082443]
 236. Gonzalez AJ, Liao L, **Wu CH** (2010) Predicting ligand binding residues using multi-positional correlations and kernel canonical correlation analysis In *Proceedings of IEEE International Conference on Bioinformatics and Biomedicine* (BIBM-2010) [Acceptance rate: 17%]
 237. **Wu CH**, McLarty JW, Liao L (2010) Neural Networks In *Encyclopedia of Life Sciences 2010*, John Wiley & Sons, [DOI: 101002/9780470015902a0005268pub2]
 238. Gonzalez AJ, Liao L, **Wu CH** (2010) Predicting functional sites in biological sequences using canonical correlation analysis *Proceedings of International Conference on Bioinformatics and Computational Biology* (BIOCOMP'10) [Acceptance rate: 27%]
 239. Petrova NV, **Wu CH** (2010) *In silico* prediction of catalytic residues in proteins using a consensus of prediction (CoP) approach *Proceedings of the 10th International Conference on Bioinformatics and Bioengineering* (BIBE-2010)
 240. Hu ZZ, Cohen KB, Hirschman L, Valencia A, Liu H, Giglio MG, **Wu CH** (2008) iProLINK: A Framework for Linking Text Mining with Ontology and Systems Biology In *Proceedings of IEEE International Conference on Bioinformatics and Biomedicine* (BIBM 2008), pp 467-472
 241. Xu G, Yin L, Torii M, Niu Z, **Wu CH**, Hu Z, Liu H (2008) Document classification for mining host pathogen protein-protein interactions In *Proceedings of IEEE International Conference on Bioinformatics and Biomedicine* (BIBM 2008), pp 461-466
 242. Arighi CN, Liu H, Natale D, Barker WC, Drabkin HJ, Hu ZZ, Blake J, Smith B, **Wu CH** (2008) TGF-beta Signaling Proteins and the Protein Ontology Bio-Ontologies 2008: Knowledge in Biology at ISMB 2008
 243. Narayanaswamy M, Ravikumar1, KE, Hu ZZ, Vijay-Shanker K, **Wu CH** (2008) Information extraction of protein phosphorylation from biomedical literature In "*Information Retrieval in Biomedicine: Natural Language Processing for Knowledge Integration*," publisher IGI-Global Ed Pr Violaine Prince
 244. Drabkin HJ, Arighi CN, **Wu CH**, Blake JA (2008) Functional annotation of protein isoforms and modified forms *Proceedings of the International Conference on Bioinformatics and Computational Biology* (BIOCOMP'08)
 245. Liu HF, Torii M, Hu ZZ, **Wu CH** (2007) Mapping gene/protein names in free text to biomedical databases *Proceedings of the IEEE ICDM Workshop on Mining and Management of Biological Data*

246. Herbert KG, Spirollari J, Wang JTL, Piel WH, Westbrook J, Barker WC, Hu ZZ and **Wu, CH** (2007) Bioinformatic databases In: *Encyclopedia of Computer Science and Engineering* (Cassie Craig Assistant Editor), John Wiley & Sons
247. McGarvey PB, Huang H and **Wu, CH** (2007) Protein bioinformatics In: *Medical Applications of Mass Spectrometry* Eds Akos Vertes and Karoly Vekey Elsevier Science
248. Liu HF, Torii M, Hu ZZ, **Wu CH** (2007) Gene mention and gene normalization based on machine learning and online resources *Proceedings of the Second BioCreative Challenge Evaluation Workshop*
249. Natale DA, Arighi C, Barker WC, Blake J, Chang TC, Hu ZZ, Liu HF, Smith B and **Wu, CH** (2006) Framework for a protein ontology *Proceedings of ACM First International Workshop on Text Mining in Bioinformatics, TMBIO 2006*
250. Torii M, Liu HF, Hu ZZ and **Wu, CH** (2006) A comparison study of biomedical short form definition detection algorithms *Proceedings of ACM First International Workshop on Text Mining in Bioinformatics, TMBIO 2006*
251. **Wu, CH** and Barker, WC (2005) Information flow and data integration of databanks *Database Annotation in Genomics* A M Lesk (Ed), John Wiley & Sons, Ltd pp187-201
252. Barker, WC and **Wu, CH** (2005) Annotation of protein sequences *Database Annotation in Genomics* A M Lesk (Ed), John Wiley & Sons, Ltd pp131-147
253. Huang, H, Nikolskaya AN, Vinayaka CR, Chung S, Zhang J and **Wu, CH** (2005) Family classification and integrative analysis for protein functional annotation *Trends in Bioinformatics Research* Peter V Yan (Ed), Nova Science Publishers, Inc pp 33-57
254. Natale, D, Vinayaka, C R and **Wu, CH** (2005) Large-scale, classification-driven, rule-based functional annotation of proteins *Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics* Bioinformatics Volume, Subramaniam, S (Ed) John Wiley & Sons, Ltd
255. Barker, WC, Mazumder, R, Nikolskaya, A and **Wu, CH** (2005) The PIR superfamily (PIRSF) classification system *Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics* Proteomics Volume, Dunn, M J (Ed) John Wiley & Sons, Ltd
256. Apweiler, R, TKAttwood, ABairoch, ABateman, DBinns, PBradley, LBordoli PBucher, ECourcelle, UDas, RDurbin, V Flegel, WFleischmann, JGouzy, SGriffiths-Jones DHaft, NHarte, NHulo, DKahn, AKanapin, MKrestyaninova, DLonsdale, RLopez, JMcDowall, NMulder, AN Nikolskaya, et al CSigrist, VSilventoinen, DJStudholme and **Wu, CH** (2005) InterPro - prediction of protein families, domains and functional sites *Progress in Genome Research*, Nova Science Publishers
257. Liu, H and **Wu, CH** (2004) A study of text categorization for model organism databases *Proceedings of BioLINK 2004: Linking Biological Literature, Ontologies and Databases*, pp 25-32
258. Mani, I, Hu, Z, Jang, S, Samuel, K, Krause, M, Phillips, J and **Wu, CH** (2004) Protein name tagging guidelines: Lessons learned" *Proceedings of BioLINK SIG, Intelligent Systems for Molecular Biology*, Glasgow, 29 July, 2004
259. Liu, H, **Wu, CH** and Friedman, C (2004) BioTagger: a biological entity tagging system *Proceedings of BioCreative Workshop - A critical assessment of text mining methods in molecular biology*, Granada, Spain, March 28-31, 2004
260. **Wu, CH** and Barker, W C (2004) A family classification approach to functional annotation of proteins *The Practical Bioinformatician* L Wong (Ed), World Scientific, NJ, pp 417-434
261. **Wu, CH** (2004) The Protein Information Resource for proteomic bioinformatics *Functional Proteomics: Principles and Practical Applications* P Ping (Ed), The Humana Press Inc [ISBN: 1-55582-210-X]
262. **Wu, CH** (2003) The Protein Information Resource for functional genomics and proteomics *Computational Biology and Genome Informatics* J Wang, C H Wu and P Wang (Eds), World Scientific pp 117-137
263. **Wu, CH** (2003) The PIR for functional genomics and proteomics *Introduction to Bioinformatics: A Theoretical and Practical Approach* S A Krawetz and D D Womble (Eds), Humana Press, Totowa, NJ, pp 431-442
264. **Wu, CH** and McLarty, J (2002) Neural Networks *The Encyclopedia of the Human Genome* D N Cooper,

Editor-in-Chief, Nature Publishing Group and Macmillan Publishers Ltd

265. Ma, Q, Wang, J T L, and **Wu, CH** (2000) Application of neural networks to biological data mining: A case study in DNA sequence classification *Proceedings of 12th International Conference on Software Engineering & Knowledge Engineering, SEKE-2000*, pp 23-30
266. Wang, J T L, Ma, Q, Shasha, D and **Wu, CH** (2000) Application of neural networks to biological data mining: A case study in protein sequence classification *Proceedings of Sixth ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD-2000*
267. **Wu, CH** and S Shivakumar (1998) ProClass protein family database: New version with motif alignments *Proceedings of the Pacific Symposium on Biocomputing '98*, eds, R B Altman, A K Dunker, L Hunter and T E Klein, World Scientific Publishing, pp 719-730 [PMID: 9697225]
268. **Wu, CH** and S Shivakumar (1998) Gene family identification network design *Proceedings of IEEE International Joint Symposia on Intelligence and Systems*, 103-110
269. **Wu, CH** and H Huang (1998) Gene family identification using neural networks *Proceedings of the World Multiconference on Systemics, Cybernetics and Informatics (SCI'98)*, Vol 4, pp 267-273
270. Ma, Q, J Wang and **Wu, CH** (1998) Detection of ALU sequences in DNA: A neural network approach *Proceedings of the Second International Workshop on Intelligent Control*, pp 392-395
271. Wang, Z, P Johnson, J Wang and **Wu, CH** (1998) Biological software development on the World Wide Web *Proceedings of the Second International Workshop on Intelligent Control*, pp 423-426
272. **Wu, CH**, S Zhao, K Simmons and S Shivakumar (1997) Motif neural network design for large-scale protein family identification *Proceedings of the 1997 International Conference on Neural Networks (ICNN '97)*, 1, 86-89, IEEE, Piscataway, NJ
273. **Wu, CH**, H L Chen, C J Lo and J McLarty (1996) Motif identification neural design for rapid and sensitive protein family search *Proceedings of the Pacific Symposium on Biocomputing '96*, eds, L Hunter and T Klein, World Scientific Publishing, pp 674-685 [PMID: 9390267]
274. **Wu, CH**, H Chen and S Shivakumar (1995) Neural networks for molecular sequence classification and database organization In: *Proceedings of the Third International Conference on Bioinformatics and Genome Research*, eds, H A Lim and C R Cantor, World Scientific Publishing, pp 25-36
275. **Wu, CH** and G Whitson (1995) Neural network database system for genetic sequence analysis In: *Progress in Neural Networks*, Vol 3, ed O M Omidvar, Ablex Publishing Co, pp 235-256
276. **Wu, CH** and H L Chen (1995) Gene classification artificial neural system In: *Proceedings of the First International IEEE Symposium on Intelligence in Neural and Biological Systems*, IEEE Computer Society Press, pp 102-107
277. **Wu, CH** (1994) Neural networks for molecular sequence classification In: *The Protein Folding Problem and Tertiary Structure Prediction*, eds K Merz and S LeGrand, Birkhauser, pp 279-305
278. **Wu, CH**, S Shivakumar, H Lin, S Veldurti, and Y Bhatikar (1994) Neural networks for molecular sequence classification *Proceedings for Workshop on Computational Issues in the Neurosciences*, The University of Texas System, Center for High Performance Computing, Austin, TX
279. **Wu, CH**, C Wang and I Yazdanpanahi (1993) Protein classification artificial neural network: a filter program for database search In: *Bioinformatics, Supercomputing and Complex Genome Analysis*, eds H Lim, J Fickett, C Cantor and R Robbins, World Scientific pp 349-358
280. **Wu, CH**, M Berry, Y Fung and J McLarty (1993) Neural networks for molecular sequence classification In: *Proceedings of First International Conference on Intelligent Systems for Molecular Biology*, eds L Hunter, D Searls & J Shavlik, AAAI Press, pp 429-437
281. **Wu, CH**, G Whitson, C Hsiao and C Huang (1992) Classification artificial neural systems for genome research *Proceedings of 1992 Supercomputing Conference*, 797-803
282. Whitson, G, **Wu, CH**, J Taylor and A Ermongkonchai (1992) CANS: an interactive neural network system for Cray supercomputers *Proceedings of the Symposium on Applied Computing*, 665-668
283. **Wu, CH**, J McLarty and G Whitson (1991) Neural networks for molecular sequence database management *Proceedings of ACM 19th Annual Computer Science Conference*, 588-594

284. **Wu, CH**, A Ermongkonchai, and T Chang (1991) Protein classification using a neural network protein database system *Proceedings of Analysis of Neural Net Applications Conference*, 29-41
285. Whitson, G, **Wu, CH** and P Taylor (1991) An artificial neural expert system shell for hypercubes *Symposium on Applied Computing*, 104-109
286. Whitson, G, **Wu, CH**, P Taylor and A Ermongkonchai (1991) HANS: an interactive neural network system for Intel supercomputers *Proceedings of Intel 1991 International User's Conference*, 231-236
287. He, C and **Wu, CH** (1991) Analysis and implementation of Kohonen algorithm for handwritten character image recognition *Proceedings of 29th ACM Southeast Regional Conference*, 163-168
288. **Wu, CH**, G Whitson and J McLarty (1990) Artificial neural system for gene classification using a domain database *Proceedings of ACM 18th Annual Computer Science Conference*, 288-292
289. **Wu, CH**, GM Whitson and G J Montllor (1990) PROCANS: A protein classification system using a neural network *Proceedings of International Joint Conference of Neural Network*, Vol II, 91-95
290. **Wu, CH**, PA Taylor, GM Whitson and C Smith (1990) CLIPS: a tool for corn disease diagnosis and aid to neural network for knowledge acquisition *First CLIPS Conference Proceedings*, 550-553
291. Whitson, G, **Wu, CH**, A Ermongkonchai and J Weber (1990) A back-propagation system for hypercubes *Symposium on applied Computing*, 71-77
292. Whitson, G, **Wu, CH** and P Taylor (1990) Using artificial neural system to determine the knowledge base of an expert system *Proceedings of ACM Symposium on Small Systems*, 268-270
293. Kulkarni, AD, G Whitson, J Bolin and **Wu, CH** (1989) Some applications of the parallel distributed processing models *Proceedings of Workshop on Applied Computing*, 185-192
294. **Wu, CH** and S Lyda (1988) Application of restriction fragment length polymorphisms in studying genetic variation of *Phymatotrichum omnivorum* *Proceedings of Beltwide Cotton Production Research Conference*, 39-41