Zuo-Fei YUAN

Gender: Male
Birthdate: 1982-09-08
Email: feifei82@gmail.com
Phone: (+1) 901-595-7872
Address: 262 Danny Thomas Place, Memphis, TN 38105

EDUCATION

Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, July 2012
Ph.D., Computer Science, Bioinformatics

Central South University, Changsha, China, July 2004
B.S., Computer Science

RESEARCH INTERESTS

Computational proteomics, protein identification and quantification, tandem mass tag (TMT), histones, data-independent acquisition (DIA), preprocessing for MS and MS/MS, peptide scoring and validation, data analysis and pipeline designing, modification discovery, spectral library search, de novo, proteogenomics search, transcriptome analysis.

"Computational Proteomics" is a computational technique that uses high-performance algorithms and practical software tools to process large-scale protein experiments or simulation data and solve the problems in proteomics research, e.g. protein identification, post-translational modification analysis, protein quantification, protein-protein interaction, protein localization, protein structure or protein dynamics. (translated from cncp.ict.ac.cn/2018/about.html)

RESEARCH EXPERIENCE

1. Sr Bioinformatics Research Scientist
2020/2 – now
St. Jude Children's Research Hospital, Supervisor: Dr. Junmin Peng and Dr. Yuxin Li
The St. Jude Center for Proteomics and Metabolomics provides state-of-the-art mass spectrometry (MS)-based protein analysis services to researchers at St. Jude Children's Research Hospital. The services include protein intact mass characterization, small- and large-scale protein identification (e.g., affinity purification-mass spectrometry), post-translational modification analysis (e.g., phosphorylation, ubiquitination, methylation and acetylation), and comprehensive quantitative profiling of proteome and phosphoproteome directly from cells, tissues and even clinical specimens. (https://www.stjude.org/research/why-st-jude/shared-resources/proteomics-mass-spectrometry.html)

My work in St. Jude Center for Proteomics and Metabolomics focuses on TMT MS-based proteomics analysis:
(1) Developing novel algorithms and software for high-throughput proteomics analysis.
(2) Maintaining and improving the JUMP proteomics suite pipeline (a proteomics software integrating protein database creation, database search for spectrum-peptide matches (PSM),

PSM filtering, protein posttranslational modification (PTM) site localization, and protein quantification).
(3) Delivering high-quality results of quantitative proteomics and bioinformatics analysis to St. Jude researchers.

2. Research Associate
2017/11 – 2019/9
Perelman School of Medicine, University of Pennsylvania (UPENN), PI: Dr. Benjamin Garcia
Dr. Garcia has pioneered high-throughput "bottom-up" proteomic methods for detection of histone PTMs and quantitative comparison of multiple cellular states, and "middle down" proteomic approaches that facilitate computation of specific combinatorial histone proteoforms. These methods have made unique impact in chromatin biology and epigenetics research, and have been fully embraced by a growing number of research groups from all over the world. (https://www.asms.org/about-asms-awards/biemann-medal)

My work in Garcia Lab focused on MS-based proteomics and histone data analysis:
(1) Developing the computational platform for processing Epi-Proteomics mass spectrometry data - EpiProfile.
(2) Developing the pipeline to discriminate and quantify isobaric phosphopeptides - ISOMER.
Note: EpiProfile is an essential tool for histone data analysis and is widely used by researchers in hospitals and institutes for disease and cancer studies.

2017/11 – 2019/9
Children's Hospital of Philadelphia (CHOP), PI: Dr. Kathrin Bernt, Dr. Richard Aplenc
Dr. Bernt is an Attending Physician at CHOP, and Assistant Professor at UPENN. Her research program is devoted to developing new therapies for children with leukemia. Specifically, her laboratory focuses on epigenetic therapies involving a group of proteins called histones.

My work in Bernt Lab focused on MS-based histone data analysis and RNA-seq data analysis. For the first time over, we can comprehensively analyze histone modifications in childhood leukemia. Novel therapeutic approaches to leukemia therapy are urgently needed. Identifying novel disease mechanisms and targets will be critical to this effort.

Dr. Aplenc is a Professor of Pediatrics in the Department of Pediatrics as well an Assistant Vice President and Chief Clinical Research Officer at CHOP. He is directing a large proteomic study aimed at identifying novel receptors on surface of acute myeloid leukemia (AML) cells that can serve as targets for CAR-T cell therapy for AML.

My work in Aplenc Lab focused on proteomics data analysis and RNA-seq data integration. Specifically, I am responsible for the entire proteomics workflow and use it to identify previously undiscovered cell surface proteins. In addition, I play a crucial role in the integration of proteomic data with RNA-seq data which are obtained on these patient samples. Strong computational background is absolutely essential to the success of the project.

3. Postdoc

2012/11 – 2017/10
Perelman School of Medicine, University of Pennsylvania, PI: Dr. Benjamin Garcia
My work in Garcia Lab focused on characterizing and quantifying histone and protein modifications:
(1) Developed an accurate computational approach to quantify histone modifications, obtained the layouts (i.e. the time relationship) of histone modifications, and discriminated isobaric peptides with modifications due to their time overlapping.
(2) Developed lots of data analysis workflows: histone peptide identification, modifications site localization, open database searching, and SILAC and TMT labeling and label-free quantification.

4. Ph.D. Candidate
2004/9 – 2012/7
Institute of Computing Technology, Chinese Academy of Sciences, Advisor: Dr. Si-Min He
Dr. He is an expert in Computational Proteomics. Since 2002 He Lab has focused on developing pFind Studio, a computational solution for mass spectrometry-based proteomics. The search engine Open-pFind was published in Nature Biotechnology in 2018 and was selected for the top ten breakthroughs of Chinese bioinformatics in 2018.

My work in He Lab focused on MS-based computational proteomics:
(1) Accurate Determination of Monoisotopic Peaks in High-resolution Mass Spectra: developed an effective approach to judge overlapping isotopic clusters, designed a new scoring function to identify co-eluted peptides, and proposed a universal approach to validate methods for monoisotopic peak determination.
(2) In-depth Analysis of Massive Spectral Data: initiated an analysis pipeline that includes preprocessing, modification discovery, parameter determination by pre-search, the second database search and filtering, spectral library search, proteogenomics search, sample comparison, and cloud computing.
(3) Initiating the Application of pFind at the Institute of Zoology, Chinese Academy of Sciences: made pFind as a fast, convenient, and functional identification studio outside pFind team for the first time.
Additional professional experience:
(1) The Third MaxQuant Summer School in Max Planck Institute of Biochemistry (2011)
(2) The First China Workshop on Computational Proteomics (2010)

REVIEWER

1. Bioinformatics, 2. Genomics Proteomics and Bioinformatics, 3. Molecular & Cellular Proteomics, 4. Analytical Chemistry, 5. Journal of Proteome Research, 6. Journal of Proteomics, 7. Proteomics Clinical Applications, 8. Chemical Research in Toxicology, 9. Computational and Structural Biotechnology Journal, 10. Journal of The American Society for Mass Spectrometry, 11. BMC Systems Biology, 12. Asia Pacific Bioinformatics Conference.

TALKS

2. Yuan, Z. F.; Sidoli, S.; Kulej, K.; Garcia, B. A., ISObaric Modification Extraction and Resolvement (ISOMER). 67th American Society for Mass Spectrometry (ASMS) 2019.
ASMS was formed in 1969 to promote and disseminate knowledge of mass spectrometry and allied topics, which is attended by more than 6,500 scientists. Dr. Garcia is the recipient of

the 2018 Biemann Medal for contributions to elucidation of the "histone code", the set of posttranslational modifications (PTMs) to histone proteins that are thought to regulate gene expression.

My talk focused on discriminating and quantifying phosphorylated peptide isomers. Phosphorylation plays an important role in the regulation of protein function. It is very common that there are multiple possible phosphorylation sites on a peptide. How to discriminate and quantify these different sites is the critical point, because they are expected to have different functions.

1. Yuan, Z. F., Demonstrating the relationship of epi-proteomics, whole-proteomics, and phospho-proteomics. 5th China Workshop on Computational Proteomics (CNCP) 2018.
CNCP has been held every two years since 2010 by the pFind team in the Institute of Computing Technology of the Chinese Academy of Sciences, which is attended by more than 150 scientists. The main topic is about high-throughput protein identification, quantification, interaction, and structure based on mass spectrometry, including both the dry labs (computational) and the wet labs (experimental), and other omics labs (e.g. genomics, transcriptomics).

My talk focused on demonstrating the relationship of epi-proteomics, whole-proteomics, and phospho-proteomics. Generally, we do the research in each individual proteomics. What if we build the networks between these multiomics data? It will be very clear to find out the relationship between histone modifications and regulated genes. Then it can be used as a direction for drugs' treatment. Therefore, the workflow will build up a bridge to connect foundation research and clinical application.

POSTERS

7. Yuan, Z. F.; Sidoli, S.; Marchione, D. M.; Simithy, J.; Janssen, K. A.; Szurgot, M. R.; Garcia, B. A., EpiProfile 2.0: A Computational Platform for Processing Epi-Proteomics Mass Spectrometry Data. 15th United States Human Proteome Organization (US HUPO) 2019 (Lightning Talk & Poster).

6. Yuan, Z. F.; Sidoli, S.; Garcia, B. A., Evaluation of database searching engines for accurate identification of histone post-translational modifications. 66th American Society for Mass Spectrometry (ASMS) 2018 (Poster).

5. Yuan, Z. F.; Sidoli, S.; Fujiwara, R.; Kulej, K.; Garcia, B. A., Discriminating isobaric phosphopeptides using data-independent acquisition mass spectrometry. 65th American Society for Mass Spectrometry (ASMS) 2017 (Poster).

4. Yuan, Z. F.; Sidoli, S.; Lin, S.; Wang, X. S.; Bhanu, N. V.; Garcia, B. A., Decoding histone post-translational modifications by bottom-up mass spectrometry. 64th American Society for Mass Spectrometry (ASMS) 2016 (Poster).

3. Yuan, Z. F.; Sidoli, S.; Lin, S.; Wang, X. S.; Bhanu, N. V.; Garcia, B. A., Decoding histone post-translational modifications by bottom-up mass spectrometry. 12th United States Human Proteome Organization (US HUPO) 2016 (Poster).

2. Yuan, Z. F.; Lin, S.; Molden, R. C.; Cao, X. J.; Bhanu, N. V.; Wang, X. S.; Sidoli, S.; Liu, S. C.; Garcia, B. A., Quantification of histone post-translational modifications by mass spectrometry. 63th American Society for Mass Spectrometry (ASMS) 2015 (Poster).

1. Yuan, Z. F.; Lin, S.; Garcia, B. A., Evaluation of accessible database searching engines for accurate identification of histone post-translational modifications. 62th American Society for Mass Spectrometry (ASMS) 2014 (Poster).

PUBLICATIONS

38 papers and 1900 citations (updating every day, as shown in the google scholar link below)
https://scholar.google.com/citations?user=spNv0V8AAAAJ&hl=en