

# TSUNG-HENG TSAI

College of Computer and Information Science, Northeastern University  
<https://tsunghengtsai.github.io> | [t.tsai@northeastern.edu](mailto:t.tsai@northeastern.edu) | 703-338-5562

## EDUCATION

---

<b>Ph.D., Electrical Engineering</b> , Virginia Tech	May 2014
Dissertation: Bayesian Alignment Model for Analysis of LC-MS-based Omic Data	
Advisors: Yue Wang (Virginia Tech) and Habtom W. Ressom (Georgetown University)	
<b>M.S., Electrical and Control Engineering</b> , National Chiao Tung University	June 2005
<b>B.S., Power Mechanical Engineering</b> , National Tsing Hua University	June 2003

## RESEARCH EXPERIENCE

---

<b>Northeastern University</b> , Boston, MA	August 2015–present
Postdoctoral Research Associate; Supervisor: Olga Vitek	
<ul style="list-style-type: none"><li>• Development and maintenance of MSstats, an R package for relative protein quantification.</li><li>• Developed a statistical approach to select informative features for protein quantification (with Roche).</li><li>• Developed the R package MSstatsPTM, for quantitative characterization of post-translational modifications (with Genentech).</li><li>• Developed statistical methods for therapeutic protein characterization (with Genentech).</li></ul>	
<b>University of Texas MD Anderson Cancer Center</b> , Houston, TX	July 2014–July 2015
Postdoctoral Research Fellow; Supervisor: W. Fraser Symmans	
<ul style="list-style-type: none"><li>• Developed genomic predictor for the response and prognosis of breast cancer patients to their treatment.</li><li>• Contributed to molecular triaging protocol for the clinical trial NCT02276443 (ARTEMIS: A robust triple-negative breast cancer (TNBC) evaluation framework to improve survival).</li><li>• Grant proposal development for projects on theranostic genomic assay development (NIH R01) and molecular diagnostic and translational omics for triple negative breast cancer (CPRIT MIRA).</li></ul>	
<b>Georgetown Lombardi Comprehensive Cancer Center</b> , Washington, DC	September 2009–May 2014
Graduate Research Assistant; Supervisor: Habtom W. Ressom	
<ul style="list-style-type: none"><li>• Developed a computational workflow for liquid chromatography-mass spectrometry (LC-MS) data.</li><li>• Developed a Bayesian alignment model for analysis of LC-MS data.</li><li>• Led large-scale proteomic and glycomic studies for biomarker discovery (R01CA143420, R01GM086746).</li><li>• Grant proposal development for projects on biomarker discovery and systems biology (NIH R01, U01).</li></ul>	
<b>Academia Sinica</b> , Taiwan	March 2007–June 2009
Research Assistant; Supervisor: Da-Wei Wang	
<ul style="list-style-type: none"><li>• Developed a computational pipeline for analysis of surface-enhanced Raman spectroscopy (SERS) data.</li><li>• Developed a multiscale peak detection method for SERS data.</li></ul>	

## PUBLICATIONS

### Journal Papers

1. **T-H Tsai**, Z Hao, Q Hong, B Moore, C Stella, JH Zhang, Y Chen, M Kim, T Koulis, GA Ryslik, E Verschueren, F Jacobson, WE Haskins, and O Vitek (2017). Statistical characterization of therapeutic protein modifications. *Scientific Reports*, 7(1):7896.
2. M Wang, **T-H Tsai**, C Di Poto, A Ferrarini, G Yu, and HW Ressom (2016). Topic model-based mass spectrometric data analysis in cancer biomarker discovery studies. *BMC Genomics*, 17(Suppl 4):545.
3. **T-H Tsai**, E Song, Z Rui, C Di Poto, M Wang, Y Luo, RS Varghese, MG Tadesse, DH Ziada, CS Desai, K Shetty, Y Mechref, and HW Ressom (2015). LC-MS/MS based serum proteomics for identification of candidate biomarkers for hepatocellular carcinoma. *Proteomics*, 15(13):2369–2381.

4. **T-H Tsai**, M Wang, C Di Poto, Y Hu, S Zhou, Y Zhao, RS Varghese, Y Luo, MG Tadesse, DH Ziada, CS Desai, K Shetty, Y Mechref, and HW Ressom (2014). LC-MS profiling of N-glycans derived from human serum samples for biomarker discovery in hepatocellular carcinoma. *Journal of Proteome Research*, 13(11):4859–4868.
5. **T-H Tsai**, MG Tadesse, C Di Poto, LK Pannell, Y Mechref, Y Wang, and HW Ressom (2013). Multi-profile Bayesian alignment model for LC-MS data analysis with integration of internal standards. *Bioinformatics*, 29(21):2774–2780.
6. **T-H Tsai**, MG Tadesse, Y Wang, and HW Ressom (2013). Profile-based LC-MS data alignment — A Bayesian approach. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 10(2):494–503.
7. JF Xiao, RS Varghese, B Zhou, MR Ranjbar, Y Zhao, **T-H Tsai**, C Di Poto, J Wang, D Goerlitz, Y Luo, AK Cheema, N Sarhan, H Soliman, MG Tadesse, DH Ziada, and HW Ressom (2012). LC-MS based serum metabolomics for identification of hepatocellular carcinoma biomarkers in Egyptian cohort. *Journal of Proteome Research*, 11(12):5914–5923.
8. HW Ressom, JF Xiao, L Tuli, RS Varghese, B Zhou, **T-H Tsai**, MR Ranjbar, Y Zhao, J Wang, C Di Poto, AK Cheema, MG Tadesse, R Goldman, and K Shetty (2012). Utilization of metabolomics to identify serum biomarkers for hepatocellular carcinoma in patients with liver cirrhosis. *Analytica Chimica Acta*, 743:90–100.
9. L Tuli, **T-H Tsai**, RS Varghese, JF Xiao, AK Cheema, and HW Ressom (2012). Using a spike-in experiment to evaluate analysis of LC-MS data. *Proteome Science*, 10(13).
10. GK Befekadu, MG Tadesse, **T-H Tsai**, and HW Ressom (2011). Probabilistic mixture regression models for alignment of LC-MS data. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 8(5):1417–1424.
11. T-T Liu, Y-H Lin, C-S Hung, T-J Liu, Y Chen, Y-C Huang, **T-H Tsai**, H-H Wang, D-W Wang, J-K Wang, Y-L Wang, and C-H Lin (2009). A high speed detection platform based on surface-enhanced Raman scattering for monitoring antibiotic-induced chemical changes in bacteria cell wall. *PLoS One*, 4(5):e5470.
12. Y-H Yi, P-Y Ho, T-W Chen, W-J Lin, V Gukassyan, **T-H Tsai**, D-W Wang, T-S Liu, C-Y Tang, SJ Lo, T-Y Chen, F-J Kao, and C-H Lin (2009). Membrane targeting and coupling of NHE1-integrin $\alpha_{IIB}\beta_3$ -NCX1 by lipid rafts following integrin-ligand interactions trigger  $Ca^{2+}$  oscillations. *Journal of Biological Chemistry*, 284(6):3855–3864.

## Book Chapters

1. Y Zhao, **T-H Tsai**, C Di Poto, LK Pannell, MG Tadesse, HW Ressom (2017). Variability Assessment of Label-free LC-MS Experiments for Difference Detection. In *Statistical Analysis of Proteomics, Metabolomics and Lipidomics Data using Mass Spectrometry* (Eds. S. Datta and B.J.A. Mertens), Springer, pp. 157–176.
2. **T-H Tsai**, M Wang, and HW Ressom (2016). Preprocessing and Analysis of LC-MS-based Proteomic Data. In *Statistical Analysis in Proteomics* (Ed. K. Jung), Springer, pp. 63–76.

## PREPRINTS

1. **T-H Tsai**, M Choi, B Banfai, Y Liu, T Dunkley, and O Vitek (2018). Selection of features with consistent profiles improves relative protein quantification in mass spectrometry experiments.
2. BV Sinn, C Fu, R Lau, J Litton, **T-H Tsai**, R Murthy, A Tam, E Andreopoulou, Y Gong, R Murthy, R Gould, Y Zhang, TA King, A Viale, V Andrade, D Giri, R Salgado, I Laios, C Sotiriou, C Marginean, DN Kwiatkowski, R Layman, D Booser, C Hatzis, V Valero, and WF Symmans (2018). SET<sub>ER/PR</sub> - A robust 18-gene predictor for sensitivity to endocrine therapy for metastatic breast cancer.

## PRESENTATIONS

---

### Invited Talks

- CIBI/OpenMS User Meeting, Berlin, Germany September 2018  
*Statistical Methods for Quantitative Proteomics*
- Aebersold Lab, ETH Zurich, Switzerland July 2018  
*Feature Selection for Relative Protein Quantification in Mass Spectrometry-based Proteomics*
- Biognosys AG, Zurich, Switzerland July 2018  
*Extensions of MSstats Framework for Mass Spectrometry-based Proteomics*
- Genentech Inc., South San Francisco, CA September 2016  
*Statistical Characterization of Therapeutic Protein Modifications*
- Targeted Proteomics International Symposium, Mumbai, India December 2015  
*Statistical Methods in Quantitative Proteomics*
- Department of Bioinformatics and Computational Biology, MD Anderson Cancer Center March 2014  
*Bayesian Alignment Model for LC-MS-based Omics*
- Neurological Research Institute, Baylor College of Medicine February 2014  
*Bayesian Alignment Model for LC-MS-based Omics*
- Center for Biomarker Discovery and Translation, Johns Hopkins University March 2012  
*Bayesian Alignment Model for LC-MS Data*
- Virginia Tech Research Center — Arlington November 2011  
*Bayesian Alignment Model for LC-MS Data*

### Contributed Talks

- Annual World Congress of the Human Proteome Organization (HUPO), Orlando, FL September 2018  
*A Statistical Framework for Relative Quantification of Post-translational Modifications in Global Proteomics Experiments*
- Annual Conference of the US Human Proteome Organization (US HUPO), Boston, MA March 2016  
*gMODs: an Open Source Data Analysis Tool for Quantifying the Differential Site Occupancy of Therapeutic Protein Modifications*
- IEEE International Conference on Bioinformatics and Biomedicine, Atlanta, GA November 2011  
*Bayesian Alignment Model for LC-MS Data*

## SOFTWARE

---

### MSstats (Co-developer)

- R package for relative quantification of proteins and peptides in mass spectrometry-based proteomics.
- Available at <http://msstats.org/>

### MSstatsPTM

- R package for quantitative characterization of post-translational modifications.
- Available at <https://github.com/tsunghengtsai/MSstatsPTM>

## TEACHING

---

### Tutorials and Short Courses

- ETH Zurich, Switzerland July 2018  
*DIA/SWATH Course (1 week)*  
Co-instructor
- Northeastern University, Boston, MA May 2018  
*Advanced R (3 days) of the 2-week May Institute (R25EB023929)*  
Co-designer and co-instructor

- Cold Spring Harbor Laboratory, Cold Spring Harbor, NY August 2017  
*Proteomics* (2 weeks)  
Invited speaker
- Northeastern University, Boston, MA May 2017  
*Advanced R* (3 days) of the 2-week May Institute (R25EB023929)  
Co-designer and co-instructor
- Buck Institute, Novato, CA February 2017  
*Targeted Proteomics* (1 week)  
Co-instructor
- Cold Spring Harbor Laboratory, Cold Spring Harbor, NY June 2016  
*Proteomics* (2 weeks)  
Co-instructor
- Northeastern University, Boston, MA May 2016  
*Computation & Statistics for Targeted Proteomics* (1 week)  
Co-instructor
- Buck Institute, Novato, CA March 2016  
*Targeted Proteomics* (1 week)  
Co-instructor
- Institute of Technology Bombay, India December 2015  
*Targeted Proteomics* (1 week)  
Co-instructor

#### Guest Lectures

- DS 5110 Introduction to Data Management and Processing, Northeastern University September 2017  
*Introduction to Tidy Data*
- CS 7280 Topics in statistics and data analysis, Northeastern University November 2016  
*Data Wrangling and Visualization with R*
- CS 7280 Topics in statistics and data analysis, Northeastern University February 2016  
*Data Wrangling with R*

---

#### PROFESSIONAL SERVICE

**Program Committee Member:** *BIBM* (2014, 2015)

**Referee for Journals:** *Bioinformatics* (1), *BMC Bioinformatics* (1), *Computational Statistics and Data Analysis* (1), *EURASIP Journal on Bioinformatics and Systems Biology* (1), *IEEE/ACM Transactions on Computational Biology and Bioinformatics* (3), *Journal of Applied Statistics* (1), *Journal of Chromatography A* (1), *Nature Communications* (1)

**Referee for Conferences:** *BIBM* (2013, 2014), *BMEI* (2008), *ICASSP* (2011, 2012), *iCBBE* (2009), *ISBI* (2011, 2012), *ISCAS* (2011), *ISMB* (2018), *MLSP* (2011)

---

#### AWARDS AND HONORS

William A. Blackwell Award, ECE Department, Virginia Tech	2014
Best Poster Award, International Conference on Laser Applications in Life Sciences (LALS)	2008
Phi-Tau-Phi Honorary Member, National Chiao Tung University	2005
Mr. Feng-Chang Lu Memorial Scholarship, National Tsing Hua University	2002
Chun-Tsung Scholarship, Chun-Tsung Endowment	2002
Yung Tuan Memorial Scholarship, National Tsing Hua University	2001