CURRICULUM VITAE

Hyungwon Choi, Ph.D.

Associate Professor
Yong Loo Lin School of Medicine
National University of Singapore
Centre for Life Sciences
28 Medical Drive, Singapore 117456

Email: hyung_won_choi@nus.edu.sg
Website: https://www.cssblab.org/
SLING: https://sling.sg

Research Interests

Computational Systems Biology and Bioinformatics
High-dimensional Data Analysis and Statistical Computing
Clinical Application of -Omics Technologies

Current Appointment

Aug 2023 – Present	Programme Leader at Singapore Lipidomics Incubator Life Sciences Institute National University of Singapore
Oct 2018 – Present	Associate Professor Department of Medicine, Yong Loo Lin School of Medicine, NUS Cardiovascular Research Institute, NUHS
Jul 2017 – Sep 2018	Associate Professor Saw Swee Hock School of Public Health, NUS
Mar 2016 – Jul 2019	Joint Principal Investigator Institute of Molecular and Cell Biology, A*STAR
Jan 2011 – Jun 2017	Assistant Professor Saw Swee Hock School of Public Health, NUS

Education and Training

Jan 2009 – Nov 2010	Post-doctoral Research Fellow Department of Pathology, University of Michigan
Aug 2004 – Apr 2009	Ph.D. in Biostatistics University of Michigan at Ann Arbor, MI, USA
Jul 2003 – Jul 2004	Life Science Research Assistant Stanford Human Genome Center, CA, USA
Sep 2001 – Jun 2003	M.S. in Statistics Stanford University, Stanford, CA, USA
Mar 1995 – Feb 2001	B.A. in Spanish Language and Literature Seoul National University, Seoul, Korea

Awards

Visiting Scholar Award, Seoul National University, South Korea (2016)
Moore-Sloan Visiting Scholar Award, New York University, USA (2015)
HUPO Young Investigator Award (2009)

Professional Membership

- Fellow of Royal Society of Chemistry
- Associate Editor, Molecular Omics, Royal Society of Chemistry

NUS Confidential

Teaching Experience

Term 2, 2023/24	PHM5002 Proteomics and metabolomics in precision health and medicine
Term 2, 2023/24	MDG5241 Advanced statistical methods for bioinformatics
Term 2, 2022/23	CDM5104 Computational Genomics (co-teaching)
Term 2, 2022/23	MDG5241 Advanced statistical methods for bioinformatics
Term 2, 2021/22	MDG5241 Advanced statistical methods for bioinformatics
Term 2, 2019/20	SPH6004 Advanced statistical learning
Term 2, 2018/19	SPH6004 Advanced statistical learning
Term 2, 2017/18	CO5232 Data management and survey data analysis
Term 2, 2017/18	SPH6004 Advanced biostatistics
Term 1, 2016/17	CO5232 Data management and survey data analysis
Term 1, 2015/16	CO5232 Data management and survey data analysis
Term 2, 2015/16	SPH6004 Advanced biostatistics
Term 1, 2014/15	CO5232 Data management and survey data analysis
Term 1, 2014/15	CO5103 Quantitative epidemiology (statistics tutorial)
Term 2, 2013/14	SPH6004 Advanced biostatistics
Term 2, 2013/14	MD4150 Statistics for community health project
Term 2, 2012/13	SPH6004 Advanced biostatistics
Term 2, 2012/13	MD4150 Statistics for community health project
Term 1, 2011/12	ST1232 Statistics for life sciences and tutorial

Current Lab Members

Oct 2023 -	Sungdong Lee, Research Fellow
Dec 2022 -	Yiyun Pang, Senior Research Fellow
Sep 2022 -	Youngrae Kim, Research Fellow
Jun 2022 -	Hui Yi Tay, Undergraduate Intern
Oct 2021 -	Zhao Qing (Lexie), Ph.D student
Jan 2021 -	Rachel Lim, Ph.D. student (with F. Torta)
Jun 2018 -	Guoshou Teo, Senior Research Fellow

Past Lab / Systems Biology Core Members

Jun 2019 – May 2022	Joseph Lee, Student Research Assistant
Jan 2016 – Feb 2022	Hiromi Koh, Research Fellow
Jan 2015 – Jun 2021	Soumita Ghosh, Research Fellow
Aug 2014 – Mar 2021	Ginny Xiaohe Li, Research Fellow
Apr 2018 – Jun 2020	Joses Ho, Research Fellow at IMCB
Aug 2018 – Aug 2019	Shuang Gao, Research Assistant
Jan 2017 – Jun 2018	Anusha Chimmiri, Research Associate, Systems Biology Core
Aug 2012 – Jul 2016	Gengbo Chen, Ph.D. in Public Health (Biostatistics)
Jan 2012 – Nov 2015	Guoci Teo, Ph.D. in Statistics
Sep 2011 – Oct 2015	Guoshou Teo, Ph.D. in Statistics

Active Research Grants – Main PI

- Scalable estimation of multivariate hurdle model for network inference in single cell and multi-tissue atlas studies

  Role: PI (SGD 786,655) for computational algorithm development

- Probability models to infer cell type contribution and site of origin for total molecular changes in omics data, Singapore Ministry of Education Tier 2, 2022-2025

  Role: PI (SGD 729,560) for combined analysis of bulk sequencing and single-cell data

Active Research Grants – Co-PI

- MOMMENTUM-CVD Singapore NMRC Centre Grant, 2022-2025 (SGD 13M)

NUS Confidential

Role: Theme PI (SGD 748,595) for bioinformatics and biostatistics core

- IAF-PP NUS-Agilent Hub, 2019-2024 (SGD 10.5M)

  Role: co-PI for the Data Integration Node

Completed Research Grants

- V2P: multi-stage big data approach for proteogenomic analysis, Singapore Ministry of Education Tier 2, 2019-2022

  Role: PI (SGD 553,720) for the re-analysis of large-scale exome DNA sequencing and mRNA splice variants with direct impact on functional units on protein sequences

- Diabetes, Tuberculosis, Neuroscience, NUHS Centre Grant, 2017-2021

  Role: co-I (SGD 396,720) for Systems Biology Core

- DYNAMO: Diabetes study in nephropathy and other microvascular complications, Singapore NMRC Open Fund Large Collaborative Grant, 2017-2021
  Role: co-I (SGD 200,000) for bioinformatics analysis in urinary and plasma biomarker theme

- Proteogenomic analysis of biological networks in cancer multi-omics data, Singapore Ministry of Education Tier 2, 2017-2019

  Role: PI (SGD 466,328) for developing proteogenomic data analysis method for multi-omics data in breast, colorectal, and blood cancers (the cancer genome atlas)

- R01 GM113237 (PI Vogel), Robust statistical approaches for decoding protein and mRNA expression regulation, NSF-NIH Joint Program in Mathematical Biology (USD1.5M), 2014-2018

  Role: co-PI (USD335,000) for statistical modeling of time-course multi-omics data analysis

- Development of computational pipeline for mass spectrometry data analysis, Singapore Ministry of Education Tier 2, 2014-2017

  Role: PI (SGD 506,498) for developing proteomics and metabolomics data analysis platform for clinical studies

- Integrative analysis of multiple –omics datasets, Singapore Ministry of Education Tier 1, 2012-2015

  Role: PI (SGD 138,000) for developing statistical methods for integrative –omics data analysis

- Statistical methods for mass spectrometry-based proteomics. NUS Internal Grant, SGD 176,000 (2011-2014)

  Role: PI (SGD 176,000) for developing computational proteomics methods

Invited Presentations

- EMBL Workshop, Heidelberg, Germany, “Integrative multi-omics data analysis for understanding cellular dynamics” (2024)
- EMBL Workshop, Heidelberg, Germany, “Integrative multi-omics data analysis for understanding cellular dynamics” (2023)
- HUST-NUS lecture series, “Multi-omics data integration for biomarker prioritization in cardiovascular diseases” (2022)
- EMBL Workshop, Heidelberg, Germany, “Integrative multi-omics data analysis for understanding cellular dynamics” (2020)
- Human Proteome Organization Meeting, Adelaide, Australia, “Predictive network of circulating lipids and proteins associated with elevated risk of coronary artery disease using iOmicsPASS” (2019)
- EMBL Workshop, Heidelberg, Germany, “Integrative multi-omics data analysis for understanding cellular dynamics” (2019)

NUS Confidential

- iSLS7, “Data analysis challenges and solutions for mass spectrometry-based metabolomics” (2018)
- 7th Asia Oceania Mass Spectrometry Conference, Singapore, “Customized spectral libraries: the bridge between untargeted and targeted metabolomics” (2017)
- Frontiers in Cancer Science, Singapore, “Pan-cancer analysis of exome sequence variants and other multi-omics data in the Cancer Genome Atlas” (2017)
- Seventh Singapore Statistical Conference, “Guiding integration of high-throughput data sets with inter-molecular relationship” (2017)
- American Society of Mass Spectrometry Meeting, San Antonio, Texas, “MetaboDIA: a novel pipeline for untargeted metabolomics using DIA-MS” (2016)
- Institute for Mathematical Sciences, Singapore, “Taking network biology to the new heights: dynamic systems” (2015)
- Simons Foundation for Data Analysis, New York, “Extraction, preprocessing, and statistical analysis of quantitative proteomics data from next generation mass spectrometry” (2015)
- Department of Mathematics and Statistics, LaTrobe University, Melbourne, Australia, “mapDIA: preprocessing and statistical analysis of quantitative proteomics data in data independent acquisition mass spectrometry” (2015)
- US Human Proteome Organization Meeting, Tempe, Arizona, “MAP-DIA: processing and analysis of DIA-MS data” (2015).
- Center for Synthetic and Systems Biology, University of Texas at Austin, “DIA-MS: Tools to cut a diamond from a lump of coal” (2015).
- Institute of Molecular and Cell Biology, ASTAR, Singapore, “Bioinformatics analysis of quantitative proteomics data with application to interaction mapping studies” (2014).
- International Forum on Novel Techniques for Quantitative Proteome Analysis, Dailan, China, “Computational tools for robust filtering of affinity purification-mass spectrometry data” (2013).
- American Society of Mass Spectrometry Annual Meeting, Minneapolis, “Luciphor – phosphorylation site localization algorithm” (2013).
- US Human Proteome Organization Meeting, Baltimore, “Luciphor – phosphorylation site localization algorithm” (2013)
- RECOMB Computational Proteomics, UC San Diego, “QPROT: Bayesian analysis of differential protein expression” (2012)
- Korean Cancer Association Annual Meeting, “Annotating cancer biomarkers with functional proteomics data” (2011)
- Department of Statistics and Applied Probability, NUS, “Bayesian model for analysis of copy number and gene expression data” (2011)
- Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, “Analysis of quantitative proteomics data: protein expression to protein-protein interaction” (2010)
- Department of Epidemiology and Public Health, National University of Singapore, “Statistical methods for computational proteomics data” (2010)
- Epidemiology and Biostatistics, Sloan Kettering Cancer Research Center (2010)
- Department of Bioinformatics and Computational Biology, MD Anderson (2010)
- International Biometric Society/Eastern North American Region, “Significance analysis of spectral count data for differential protein expression” (2009)
- Workshop on PPI Networks, McGill University Bellairs Institute, “Significance analysis of AP-MS data” (2008)
- Annual International Chinese Statistical Association Meeting, “Combined genomic profiles of copy number and gene expression” (2008)

Invited Teaching

- EMBL Course “Analysis and Integration of Transcriptome and Proteome Data”, EMBL Heidelberg, Germany (2024).
- EMBL Course “Analysis and Integration of Transcriptome and Proteome Data”, EMBL Heidelberg, Germany (2023).
- EMBL Course “Analysis and Integration of Transcriptome and Proteome Data”, EMBL Heidelberg, Germany (2020).
- EMBL Course “Analysis and Integration of Transcriptome and Proteome Data”, EMBL Heidelberg, Germany (2019).
- SLING lipidomics workshop, Life Science Institute, NUS (2017).

NUS Confidential

- Undergraduate Student Seminar, Department of Statistics, Seoul National University, Korea, “Statistical computing in computational biology” (2013).
- Student Seminar, School of Pharmacy, Seoul National University, Korea, “Computational proteomics” (2011).

Peer-reviewed Articles (Main papers)

- H. Koh *et al.*, An integrated signature of extracellular matrix proteins and a diastolic function imaging parameter predicts post-MI long-term outcomes. *Frontiers in Cardiovascular Medicine* 2023, 10, 1123682.
- M.R. Wenk, H. Choi. Abundant circulating lipids – a new opportunity for NSCLC detection? *Nat. Rev. Clin. Oncol.* 2022, News and Views, 1-2.
- S. Ko, G.X.L. Li, **H. Choi***, J-.H. Won*. Computationally scalable regression modelling for ultrahigh-dimensional omics data with ParProx. *Brief. Bioinf*o. 2021, bbab256.
- S. Ghosh, A. Datta, **H. Choi**. multiSLIDE: a web server for exploring connected elements of biological pathways in multi-omics data. *Nat. Comms*. 2021; 12(1):2279.
- Y.M. Tan, Y. Gao, G. Teo., […], Z. Lei, **H. Choi**. Plasma metabolome and lipidome associations with T2D and diabetic nephropathy. *Metabolites* 2021; 11(4):228.
- G. Teo, W. Chew, B. Burla, D. Herr, ES Tai, M. Wenk, F. Torta, **H. Choi**. MRMkit: automated data processing for large-scale targeted metabolomics analysis. *Anal. Chem*. 2020, 92 (20) 13677-13682.
- P. Narayanaswamy, G. Teo, J. Ow, A. Lau, P. Kaldis, S. Tate, **H. Choi**. MetaboKit: a comprehensive data extraction tool for untargeted metabolomics. *Mol. Omics* 2020; 16, 436-447.
- G.X.L. Li, D. Munro, D. Fermin, C. Vogel, **H. Choi**. A protein-centric approach for exome variant aggregation enables sensitive association analysis with clinical outcomes. *Hum. Mutation*. 2020; 41(5): 934-945.
- B. Vitrinel, H. Koh, FM. Kar, S. Maity, J. Rendleman, **H. Choi**, C. Vogel. Exploiting interdata relationships in next-generation proteomics analysis. *Mol. Cell. Proteomics* 2019:18, S5-S14.
- H. Koh, D. Fermin, C. Vogel, KP Choi, RM Ewing, **H. Choi**. iOmicsPASS: network-based integration of multiomics data for predictive subnetwork discovery. *npj Syst. Biol. Appl*. 2019:5(1), 22.
- H. Koh, Y.B. Zhang, C. Vogel, **H. Choi**. EBprotV2: a Perseus plugin for differential protein expression analysis of labelling-based quantitative proteomics data. *J. Proteome Res*. 2019; 18(2): 748-752.
- **H. Choi** *et al*., Plasma protein and microRNA biomarkers of insulin resistance: a network-based integrative -omics analysis. *Front. Physiol*. 2019:10, 379.
- J. Rendleman, **H. Choi**, C. Vogel. Integration of large-scale multi-omics datasets: a protein-centric view. *Current Opinion in Systems Biology* 2018;11:74-81.
- S. Ghosh, A. Datta, K. Tan, **H. Choi**. SLIDE – a web-based tool for interactive visualization of large-scale -omics data. *Bioinformatics* 2018:35(2), 346-348.
- G. Li, C. Vogel, **H. Choi.** PTMscape: an open source tool to predict generic post-translational modifications and map modification crosstalk in protein domains and biological processes. *Mol. Omics* 2018; 14:197-209.
- G. Teo, Y. Zhang, C. Vogel, **H. Choi**. PECAplus: statistical analysis of time-dependent regulatory changes in dynamic single-omics and dual-omics experiments, *npj Systems Biology and Application* 2017; 4(1):3.
- G. Chen[#], S. Walmsley[#], G. Cheung, L. Chen, C. Cheng, R. Beuerman, T. Wong, L. Zhou*, **H. Choi***. Customized consensus spectral library building for untargeted quantitative metabolomics analysis using data independent acquisition mass spectrometry and MetaboDIA workflow, *Analytical Chemistry* 2017; 89(9):4897-4906.
- **H. Choi**, D. Ghosh, Z. Qin, Computationally tractable multivariate HMM in genome-wide mapping studies. *Methods Mol. Biol*. 2017; 1552:135-148.
- G.C. Teo, H. Koh, D. Fermin, J.P. Lambert, J.D. Knight, A.C. Gingras, **H. Choi**. SAINTq: scoring protein-protein interactions in affinity purification-mass spectrometry experiments with fragment or peptide intensity data. *Proteomics*. 2016; 16(15-16):2238-45.
- Z. Cheng[#], G.S. Teo[#], S. Krueger, T.M. Rock, H.W.L. Koh, **H. Choi***, C. Vogel*. Differential dynamics of the mammalian mRNA and protein expression response to misfolding stress. *Mol. Syst. Biol.* 2016; 12(1):855.

NUS Confidential

- G.S. Teo, S. Kim, Ben Collins, A.C. Gingras, A.I. Nesvizhskii, **H. Choi**. mapDIA: a new platform for data normalization and analysis for label-free quantitative proteomics data. *J. Proteomics*. 2015; 129:108-20.
- G. Chen, C. Liang, D. Fermin, C.N. Ong, C.S. Tan, **H. Choi**, MetTailor: Dynamic block summary and data normalization for robust analysis of mass spectrometry data in metabolomics. *Bioinformatics*. 2015; 31(22):3645-52.
- **H. Choi**, S. Kim, D. Fermin, C.C. Tsou, A.I. Nesvizhskii, QPROT: statistical method for testing differential expression using protein-level intensity data in label-free quantitative proteomics. *J. Proteomics*. 2015; 129:121-6.
- H.W.L. Koh, H. Swa, D. Fermin, S.G. Ler, J. Gunaratne, **H. Choi**, EBprot: differential expression analysis in labeling-based quantitative proteomics experiments. *Proteomics*. 2015; 15:2580-91.
- D. Fermin, D. Avtonomov, **H. Choi***, A. Nesvizhskii*. LuciPHOr2: site localization of generic post-translational modifications from tandem mass spectrometry data. *Bioinformatics*. 2014; 31(7):1141-43.
- G.C. Teo, G. Liu, J. Zhang, A. Nesvizhskii, A-.C. Gingras, **H. Choi**, SAINTexpress: improvements and additional features in SAINT software. *J. Proteomics.* 2014; 100:37-43.
- G.S. Teo, C. Vogel, D. Ghosh, S. Kim, **H. Choi**, PECA: a novel statistical tool for deconvoluting time-dependent gene expression regulation. *J. Proteome Res.* 2014; 13(1):29-37.
- D. Fermin, S. Walmsley, A-.C. Gingras, **H. Choi***, A. Nesvizhskii*, Luciphor: algorithm for phosphorylation site localization with false localization rate estimation using modified target-decoy approach, *Mol. Cell. Proteomics* 2013; 12: 3409-3419.
- **H. Choi**, D. Fermin, A. Nesvizhskii, D. Ghosh, Z.S. Qin, Sparsely correlated hidden Markov models with application to genome-wide location studies. *Bioinformatics* 2013; 29(5):533-41.
- **H. Choi**, G. Liu, M. Tyers, A.-C. Gingras, A. Nesvizhskii, Analyzing protein-protein interactions from AP-MS data with SAINT. *Curr. Protoc. Bioinformatics* 2012, Chapter 8:Unit8.15.
- **H. Choi**, Computational assembly of protein complexes in AP-MS studies. *Proteomics* 2012; 12(10):1663-68.
- **H. Choi**, T. Glatter, M. Gstaiger, A. Nesvizhskii, SAINT-MS1: protein-protein interaction scoring using label-free intensity data in AP-MS experiments, *J. Proteome Res*. 2012; 11(4):2619-24.
- **H. Choi**, N. Pavelka, When one and one gives more than two: challenges and opportunities of integrative omics, *Front. Gene*. 2011; 2:105.
- T. Kwon*, **H. Choi***, C. Vogel, A. Nesvizhskii, E.M. Marcotte, MSblender: a probabilistic approach for integrating peptide identifications from multiple database search engines. *J. Proteome Res*. 2011; 10(7):2949-58.
- **H. Choi**, B. Larsen, Z.-Y. Lin, A. Breitkreutz, D. Mellacheruvu, D.Fermin, Z.S. Qin, M. Tyers, A.-C. Gingras, A. Nesvizhskii, SAINT: probabilistic scoring of AP-MS data. *Nat. Methods* 2011; 8:70-73.
- **H. Choi**, S. Kim, A.-C. Gingras, A. Nesvizhskii, Analysis of local protein complexes via model-based bi-clustering of label-free quantitative AP-MS data. *Mol. Syst. Biol*. 2010; 6:385.
- A. Breitkreutz*, **H. Choi***, J. Sharom*, L. Boucher*, V. Neduva* et al., A global kinase and phosphatase interaction network in yeast. *Science* 2010; 328:1043-46.
- **H. Choi**, Z.S. Qin, D. Ghosh, A double-layered mixture model for joint analysis of copy number and gene expression datasets. *J. Comput. Biol*. 2010; 17(2):121-37.
- D. Ghosh and **H. Choi**, Discussion of “A Bayesian model for cross-study differential gene expression" by Scharpf et al., *J. Am. Statist. Assoc.*, 2009; 104(488):1310-14.
- **H. Choi**, A. Nesvizhskii, D. Ghosh, Z.S. Qin, Hierarchical hidden Markov model with application to joint analysis of ChIP-seq and ChIP-chip data. *Bioinformatics*, 2009,; 25:1715-21.
- **H. Choi**, D. Fermin, A. Nesvizhskii, Significance analysis of spectral count data in label-free shotgun proteomics, *Mol. Cell. Proteomics*, 2008; 7(12):2373-85.
- **H. Choi**, A. Nesvizhskii, Semi-supervised model-based validation of peptide identifications in mass spectrometry-based proteomics, *J. Proteome Res.*, 2008; 7(1):254-65.
- **H. Choi**, D. Ghosh, A. Nesvizhskii, Statistical validation of peptide identification in large-scale proteomics using target-decoy database search strategy and flexible mixture modelling, *J. Proteome Res*. 2008; 7(1):286-92.
- **H. Choi**, A. Nesvizhskii, False discovery rates and related statistical concepts in mass spectrometry-based proteomics, *J. Proteome Res.*, 2008; 7(1):47-50.
- **H. Choi**, R. Shen, A. Chinnaiyan, D. Ghosh, A latent variable approach for meta-analysis of gene expression data from multiple microarray experiments, *BMC Bioinformatics*, 2007; 8:364.

NUS Confidential

* indicates co-first or co-corresponding authorship.

Peer-reviewed Articles (Collaborative papers)

- S. Pushalkar *et al*., Complex changes in serum protein levels in COVID-19 convalescents. *Sci. Rep*. 2024, 14(1), 4479.
- Teo *et al*., Oral peroxisome proliferator-activated receptor-a agonist enhances corneal nerve regeneration in patients with type 2 diabetes. *Diabetes*. 2023, 72(7), 932-946.
- J.K. Chang *et al*., Q-RAI data-independent acquisition for lipidomic quantitative profiling. *Sci. Rep*. 2023, 13(1), 19281.
- M.M. Hoppe *et al*., Patterns of oncogene coexpression at single-cell resolution influence survival in lymphoma. *Cancer Discovery* 2023, 13(5): 1144-1163.
- K. Allgoewer *et al*., Re-mining serum proteomic data reveals extensive post-translational modifications upon Zika and dengue infection. *Mol. Omics* 2023, ePub.
- S. Wu *et al*., Proteomic signatures of the serological response to influenza vaccination in a large human cohort study. *Viruses* 2022, 14(11):2479.
- G. Li *et al*., KOPI: Kinase inhibitor Proteome Impact analysis. *Sci. Rep*. 2022, 12(1):13015.
- S. Wu *et al*., Evaluation of determinants of the serological response to the quadrivalent split-inactivated influenza vaccine. *Mol. Syst. Biol*. 2022, 18:e10724.
- T. Tumkaya *et al*., Most primary olfactory neurons have individually neutral effects on behavior. *Elife* 2022, 11, e71238.
- X. Liu *et al*., Alternations in the gut microbiota and metabolome with newly diagnosed unstable angina. *J. Genet. Genomics* 2022, 49(3):240-248.
- S. Tan *et al*., Variability of the plasma lipidome and subclinical coronary atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* 2022, 42(1):100-112.
- A. Chan *et al*., Novel autoantibodies in idiopathic small fiber neuropathy. *Ann. Neurol*. 2022, 91(1): 66-77.
- S. Tan *et al*., Variability of the plasma lipidome and subclinical coronary atherosclerosis. *ATVB* 2022, 42(1): 100-112.
- A. Targa *et al*., Non-genetic and genetic rewiring underlie adaptation to hypomorphic alleles of an essential gene. *EMBO J*. 2021, e107839.
- M. Hoppe *et al*., Quantitative imaging of RAD51 expression as a marker of platinum resistance in ovarian cancer. *EMBO Mol. Med.* 2021, 13(5):e13366.
- J.W. Dear *et al*., A metabolomic analysis of thiol response for standard and modified N-acetyl cysteine treatment regimens in patients with acetaminophen overdose. *Clin. Transl. Sci.* 2021, 4(4):1476-1489.
- K. Allgoewer *et al*., New proteomic signatures to distinguish between Zika and dengue infections. *Mol. Cell. Proteomics* 2021, 20, 100052.
- J.R. Ow *et al*., Remodeling of whole-body lipid metabolism and a diabetic-like phenotype caused by loss of CDK1 and hepatocyte division. *Elife* 2020, 9, e63835.
- M.R. Dewhurst *et al*. Loss of hepatocyte cell division leads to liver inflammation and fibrosis. *PLoS Genetics* 2020, 16 (11), e1009084.
- L.S.W. Loo *et al*., BCL-xL/BCL2L1 is a critical anti-apoptotic protein that promotes the survival of differentiating pancreatic cells from human pluripotent stem cells. *Cell Death Dis*. 2020, 11 (5): 1-18.
- J.Y.H. Seah *et al*. Plasma sphingolipids and risk of cardiovascular diseases: a large-scale lipidomic analysis. *Metabolomics* 2020, 16(9), 1-12.
- S. Ahn *et al*., Self-semi-supervised clustering of large-scale data with a massive null group. *JKSS* 2020, 1-16.
- J. Niska-Blakie *et al*., Knockout of the non-essential gene SUGCT creates diet-linked, age-related microbiome disbalance with a diabetes-like metabolic syndrome phenotype. *Cellular and Molecular Life Sciences*, 2020, 77(17):3423-3439.
- S. Manohar *et al*., Polyubiquitin chains linked by lysine residue 48 selectively target oxidized proteins in vivo. *Antioxidants & Redox Signaling* 2019, 31(15):1133-1149.
- J. Ong *et al*., Insights into early recovery from influenza pneumonia by spatial and temporal quantification of putative lung regenerating cells and by lung proteomics. Cells 2019;8(9):975.
- J. Ho *et al*., Moving beyond P values: data analysis with estimation graphics. *Nat. Methods* 2019;16(7):565-566.

NUS Confidential

- Lim *et al*., Truncated rank correlation as a robust measure of test-retest reliability in mass spectrometry data. *Statist. App. Genet. Mol. Biol*. 2019;18(4):/j/sagmb.2019.18.issue-4.
- L. Li *el al*., Antibody treatment against angiopoietin-like 4 reduces pulmonary edema and injury in secondary pneumococcal pneumonia. *mBio* 2019;10(3):e02469-18.
- J. Rendleman *et al*., New insights into the cellular temporal response to proteostasis stress. *eLife* 2018, 7, e39054.
- Chew *et al*., Large-scale lipidomics identifies associations between plasma sphingolipids and T2DM incidence. *JCI Insight* 2019;5(13):e126925.
- Caldez *et al.*, Metabolic remodelling during liver regeneration. *Dev. Cell* 2019:47(4):425-438.
- Lee *et al*., ABRF Proteome Informatics Research Group (iPRG) 2016 Study: inferring proteoforms from bottom-up proteomics data, *J. Biomol. Tech*. 2018, 29(2): 39-45.
- L. Baltazar *et al*., Concentration-dependent protein loading of extracellular vesicles released by Histoplasma capsulatum after antibody treatment and its modulatory action upon macrophages, *Sci. Rep*. 2018, 8(1), 8065.
- K.S. Hahn *et al*., A least squares method for detecting multiple change points in a univariate time series, *Proc. Pacific Rim Statist. Conf. for Production Eng.* 2018.
- K. Tan *et al.*, In vitro model of fully differentiated human nasal epithelial cells infected with Rhinovirus reveals epithelium-initiated immune responses, *J. Inf. Diseases*, 2018, jix640.
- C. Windpassinger *et al*., CDK10 mutations in humans mice cause severe growth retardation, spine malformations, and developmental delays. *Am. J. Hum. Genet*. 2017, 101(3):391-403.
- H. Shen *et al*., IKK2 regulates cytokinesis during vertebrate development, *Sci. Rep*. 2017, 7(1):8094.
- J Liu *et al*., Profiling of plasma metabolites suggests altered mitochondrial fuel usage and remodeling of sphingolipid metabolism in individuals with type 2 diabetes and kidney disease. *Kidney Int. Rep.* 2017; 2(3):470-480.
- J. Knight *et al.*, ProHits-viz: a suite of web tools for visualizing interaction proteomics data, *Nat. Methods*, 2017; 14(7):645-646.
- D. Zheng *et al*., Differentiation of club cells to alveolar epithelial cells in vitro. *Sci. Rep*. 2017; 7:41661.
- R.L. Sontag *et al*., Identification of novel host interactors of effectors secreted by *Salmonella* and *Citrobacter*. mSystems. 2016; 1(4):pii:e00032-15.
- G. Liu *et al*., Data independent acquisition analysis in ProHits 4.0. *J. Proteomics*. 2016; 149:64-68.
- Y. Sun *et al*., Plasma fatty acids, oxylipins, and risk of myocardial infarction: the Singapore Chinese Health Study. *J. Lipid Res*. 2016; 57(7):1300-1307.
- L. Baltazar *et al.*, Antibody binding alters the characteristics and contents of extracellular vesicles release by Histoplasma capsulatum. *mSphere* 2016; 1(2):e00085-15.
- A. Alli Shaik *et al*., Phosphoprotein network analysis of white adipose tissues unveils deregulated pathways in response to high-fat diet. *Sci. Rep*. 2016; 6:25844
- G. Liu *et al*., Data independent acquisition analysis in ProHits 4.0. *J. Proteomics* 2016; S1874-3919(16):30174-9.
- D.Y. Oh *et al*., Adjuvant-induced human monocyte secretome profiles reveal adjuvant- and age-specific protein signatures. *Mol. Cell. Proteomics*, 2016; 15(6):1877-1894.
- L. Chen *et al*., Plasma metabonomic profiling of diabetic retinopathy, *Diabetes*, 2016; 65(4):1099-1108.
- G. Liu *et al*., Gene essentiality is a quantitative property linked to cellular evolvability. *Cell* 2015, 163(6):1388-99.
- Y. Sun *et al*., Plasma α-linolenic and long-chain ω-3 fatty acids are associated with a lower risk of acute myocardial infarction in Singapore Chinese adults. *J. Nutr*. 2015; 146(2):275-282.
- S. Majumdar *et al*., Environmental effects of nanoceria on seed production of common bean (Phaseolus vulgaris): a proteomic analysis. *Environ. Sci. Technol*. 2015; 49(22):13283-93.
- C.P. Lee *et al*., Mammographic breast density and common genetic variants in breast cancer risk prediction. *PLoS One*, 2015; 10(9):e0136650.
- C. Tsou *et al*., DIA-Umpire: comprehensive computational framework for data-independent acquisition proteomics, *Nat. Methods* 2015, 12(3):258-64.
- J. Knight *et al*., A web-tool for visualizing quantitative protein-protein interaction data. Proteomics, *Proteomics* 2014; 15(8):1432-6.

NUS Confidential

- M. Taipale *et al*., A quantitative chaperone interaction network reveals the architecture of cellular protein homeostasis pathways. *Cell*, 2014; 158(2):434-448.
- K. Tan *et al.*, microRNAs in a regenerating lung: an integrative systems biology analysis in a mouse model of influenza pneumonia. *BMC Genomics*, 2014; 15:587.
- T. Narasaraju *et al.*, Combination therapy with hepatocyte growth factor and oseltamivir confers enhanced protection against influenza viral pneumonia. *Curr. Mol. Med.* 2014; 14(5):690-702.
- K. Kim *et al*., Reinvestigation of aminoacyl-TRNA synthetase core complex by affinity purification-mass spectrometry reveals TARSL2 as a potential member of the complex. *PLoS One*, 2013; 8(12):e81734.
- A. Couzens *et al*., Protein interaction network of the mammalian HIPPO pathway reveals mechanisms of kinase-phosphatase interactions, *Sci. Signal*. 2013; 19:6(302).
- D. Mellacheruvu *et al*., The CRAPome: a contaminant repository for affinity purification-mass spectrometry data, *Nat. Methods*, 2013; 10:730-736.
- E. Bayer-Santos *et al*., Proteomic analysis of Trypanosoma cruzi secretome: characterization of two populations of extracellular vesicles and soluble protein, *J. Proteome Res*., 2013; 12(2):883-97.
- G. Liu *et al*., Using ProHits to store, annotate, and analyze affinity purification-mass spectrometry data, *Curr. Protocols in Bioinformatics*, Chapter 8, Unit 8-16, 2012.
- S. Rodrigues *et al*., Label-free quantitative proteomics reveals differentially regulated proteins in the latex of sticky disease Caprica papaya L. plants, *J. Proteomics*, 2012; 75(11):3191-98.
- J. Bayona *et al*., SUMOylation pathway in Trypanosoma cruzi: functional characterization and proteomic analysis of target proteins. *Mol. Cell. Proteomics*, 2011; 10(12):M110 007369.
- A. Bergerat *et al*., Prestroke proteomic changes in cerebral microvessels in stroke-prone, transgenic[hCETP]-hyperlipidemi, Dahl salt-sensitive hypertensive rats, *Mol. Med.*, 2011; 17(7-8):588-598.
- B. Emmer *et al*., Global analysis of protein palmitoylation in African trypanosomes, *Eukaryotic Cell*, 2011; 10(3): 455-463.
- T. Glatter *et al*., Modularity and hormone sensitivity of the Drosophila melanogaster insulin receptor/target of rapamycin interaction proteome, *Mol. Syst. Biol*. 2011; 7:547.
- D. Skarra *et al*., Label-free quantitative proteomics and SAINT analysis enable interactome mapping for the human Ser/Thr protein phosphatase 5, *Proteomics*, 2011; 11(8):1508-1516.
- S. Vareed *et al*., Metabolites of purine nucleoside phosphorylase (NP) in serum have the potential to delineate pancreatic adenocarcinoma. *PLOS One*, 2011; 6(3):e17177.
- Y. Ding, **H. Choi**, A. Nesvizhskii, Adaptive discriminant function analysis and reranking of MS/MS database search results for improved peptide identification in shotgun proteomics, *J. Proteome Res.*, 2008; 7(11):4878-4889.
- W. Gu, **H. Choi**, D. Ghosh, Global associations between copy number and transcript mRNA microarray data: an empirical study, *Cancer Informatics*, 2008; 6:17-23.
- J. Lim, S. Lee, **H. Choi**. Information loss from censoring in rank-based procedures, *Stat. Prob. Letters*, 2006; 76: 1705-1713.

Abstracts/Conference Proceedings

- PTMscape: a computational tool for global post-translational modification prediction with in-protein false discovery rate estimation. Human Proteome Organization World Congress, Taipei, Taiwan (2016).
- EBprotV2: Statistical analysis of labelling-based quantitative proteomics data with applications to clinical data. Human Proteome Organization World Congress, Taipei, Taiwan (2016).
- PECA2: new developments in the statistical modeling of protein concentration regulation in dynamic systems. Human Proteome Organization World Congress, Taipei, Taiwan (2016).
- MetaboDIA. American Society of Mass Spectrometry Meeting, San Antonio, TX, USA (2016).
- SAINTq. American Society of Mass Spectrometry Meeting, San Antonio, TX, USA (2016).
- Advances in statistical analysis of quantitative proteomics data in data independent acquisition mass spectrometry. American Society of Mass Spectrometry Meeting, St. Louis, MO, USA (2015).
- EBprot: differential expression analysis in labeling-based quantitative proteomics experiments, US HUPO, Tempe (2015).
- PECA: a novel statistical tool for deconvoluting time-dependent gene expression regulation. US HUPO, Seattle (2014).

NUS Confidential

- SAINTexpress: improvements and additional features in SAINT software. US HUPO, Seattle (2014).
- LuciPHOr: an algorithm for phosphorylation site localization with false localization rate estimation. Korean Human Proteome Organization Meeting, Seoul, Korea (2013).
- QPROT: a comprehensive software suite for analyzing label-free quantitative proteomics data. American Society of Mass Spectrometry Meeting, Vancouver, Canada (2012).
- LuciPHOr: an algorithm for phosphorylation site localization. American Society of Mass Spectrometry Meeting, Vancouver, Canada (2012).
- MSblender: a probabilistic approach to integrate multiple database search results in peptide identification. RECOMB satellite proteomics, San Diego, USA (2011).
- Dynamic changes in protein-protein interaction data. American Society of Mass Spectrometry Meeting, Salt Lake City, USA (2010).
- Significance analysis of interactome: filtering high-confidence protein-protein interactions from AP-MS experiments, RECOMB Satellite Computational Proteomics, San Diego, USA (2010).
- Partition clustering of AP-MS data for the reconstruction of local protein complexes. Human Proteome Organization 8th Annual World Congress, Toronto, Canada (2009).
- Significance analysis of spectral counts for differential expression in label-free proteomics. American Society of Mass Spectrometry Meeting, Denver, USA (2008).

Software Packages

Multi-omics data integration

- DLMM: Integrative analysis of DNA copy number and mRNA expression data.
  http://sourceforge.net/projects/dlmm/
- PECA: Protein expression control analysis for gene expression regulation.
  Latest development at https://github.com/PECAplus
  Old version of PECA at http://sourceforge.net/projects/pecanalysis/
- iOmicsPASS: Integrative -omics approach for predictive analysis of cancer subtypes with sub-network signatures. https://github.com/cssblab/iOmicsPASS
- iOmicsPASS+: Extension of iOmicsPASS with Gaussian graphical lasso for partial correlation network estimation and supervised classification analyses (https://github.com/cssblab/iOmicsPASSplus)
- HP-ACCORD: Ultrahigh-dimensional sparse precision matrix estimation algorithm (to be released in 2024).

Data visualization

- SLIDE: Systems-level Interactive Data Exploration.
  Click here for the web version of SLIDE
  GitHub page: https://github.com/cssblab/SLIDE
- multiSLIDE: web-based multi-omics data visualization
  https://github.com/soumitag/multiSLIDE/
- DABEST: Data Analysis with Bootstrapped ESTimation
  https://github.com/ACCLAB/DABEST-python
  https://github.com/ACCLAB/dabestr

Differential expression analysis for proteomics data

- EBprot: Statistical analysis of labeling-based proteomics data.
  http://sourceforge.net/projects/ebprot/
  http://github.com/cssblab/EBprot
- mapDIA: Data processing and statistical analysis of data independent acquisition MS data.
  https://github.com/mapDIA/mapDIA
- QSPEC/QPROT: Statistical protein differential expression analysis for spectral count data.
  http://sourceforge.net/projects/qprot/
- KSA2D: Kinase-substrate joint differential analysis.
  https://github.com/ginnyintifa/KSA2D/

Analysis of interaction proteomics data

NUS Confidential

- nestedCluster: Biclustering algorithm for affinity purification – quantitative proteomics data. http://sourceforge.net/projects/nestedcluster/
- SAINT suite: Significance analysis of affinity purification-MS data. http://sourceforge.net/projects/saint-apms/

Computational proteomics

- LuciPHOr: Probabilistic algorithm to control false localization rates of generic PTM from mass spectrometry-based proteomics data. http://luciphor.sourceforge.net/ http://luciphor2.sourceforge.net/
- MSblender: Integration of database search results in peptide identification (led by TJ Kwon, UNIST, Korea). http://sourceforge.net/projects/msblender/
- PTMscape: an open source tool to predict generic post-translational modifications and map hotspots of modification crosstalk. https://github.com/cssblab/PTMscape/
- PRMkit: fast and accurate data extraction and normalization of parallel reaction monitoring (PRM) data for targeted proteomics experiments (to be released)

Computational metabolomics

- MetMatch: Database search engine for untargeted metabolomics (led by Scott Walmsley, U Colorado Denver). http://github.com/IsoMS/metmatch
- MetaboDIA: spectral library and data extraction for DIA-MS in metabolomics. https://sourceforge.net/projects/metabodia
- MetaboKit: robust workflow for data dependent and data independent acquisition mass spectrometry for untargeted metabolomics. https://github.com/MetaboKit/MetaboKit
- MRMkit: automated data processing for large-scale targeted mass spectrometry. https://github.com/MRMkit/MRMkit
- DIMSkit: mass fingerprinting and quantification in chromatography-free direct infusion MS experiments (to be released in 2023)
- MetaboScan: compound identification and quantification with scanning SWATH (to be released in 2023)
- Lipid-EAD: structural elucidation of lipid molecules using electron activated dissociation spectra (to be released in 2024)

Genome-wide location studies

- ChIPmeta: Integration of multiple ChIP experiments. http://sourceforge.net/projects/chipmeta
- scHMM: Sparsely correlated HMM algorithm for ChIP data analysis. http://sourceforge.net/projects/schmm/

NUS Confidential