Wu Yonghui, PhD

Cell: (+65) 8666-0876
Email: hedy_wu@hotmail.com

Summary:

I'm a scientist working on Data Mining and Bioinformatics for around 24 years. My work mainly focuses on the Genomics/Epigenetic research based on Statistic Data Analytics / Machine Learning / Deep Learning / Foundation Model. Next-Generation Sequencing (NGS) / Microarray / Wearables are important research data sources in the work. With extensive knowledge and experience, I'm confident in handling all the challenges in the domain.

Skills (Technologies, Tools, Platforms):

Statistic Analytics:

Clustering, Classification, PCA, Survival Analysis, etc.

AI:

Foundation Model (LLM, BERT, Transformer)
Deep Learning (ANN, CNN, RNN),
Machine Learning (Logistic Regression, SVM, KNN, Decision Tree),
Feature Selection (tsfresh)
Platform (H2O),
Library (pandas, mlbox, Keras, TensorFlow)

Computer Language:

R, Python, MATLAB, C, Jupyter Notebooks

Bioinformatics:

Microarray and next generation sequencing data processing and analysis, including Affymatrix GeneChip/Illumina BeadArray, Chip-seq, RNA-seq, Exome-seq, etc.
NGS data quality assessment, alignment, SNP variant identification, CHIP-Seq peak identification and motif discovery, RNA-Seq differential expression analysis, alternative splicing events discovery, transcriptome assembly.
NGS software packages (Fastqc, Bowtie aligner, BWA aligner, Tophat, Samtools, MACS, Galaxy, MEME, etc.)

Data Visualization:

Cytoscape, Treeview, OpenGL

Education:

1996-2001 Ph.D. & M.S., Xi'an Jiaotong University, PRC
1991-1995 B.S., Xi'an Jiaotong University, PRC

Professional Experiences:

2023- 2025 **Principal Scientist,** Institute of High Performance Computing (IHPC), A*STAR, Singapore

Projects:

1) Develop a genomic foundation model.

2) AI for exploring genomic data in eye disease.

3) An AI-based TCR-pHLA treatment approach for cancer patients.

2021- 2023 **Senior Scientist**, Institute of High Performance Computing (IHPC), A*STAR, Singapore

Projects:

1) Build an analytics AI platform for genomics data in Modstore (An Instructional HPC-Based Platform for National Science Experiment Big Data Challenge).

2) Integrate genomic data with CT images to assist doctors in prescribing drugs for cancer patients.

3) Generate synthetic data for genetic datasets by GAN.

2021- 2021 **Senior Data Scientist,** Standard Chartered Bank, Singapore

Projects:

Use Machine Learning and NLP technique to predict the Client experiences from Client onboarding Emails

2019- 2020 **Research Scientist,** Diagnostics Development Hub (DXD), A*STAR, Singapore

Projects:

1) Use Machine Learning and Deep Learning to detect reinjure risk for return to sport after Anterior Cruciate Ligament reconstruction

2) Use Machine Learning to detect microRNA Biomarkers for Ovarian Cancer

3) Use Machine Learning to detect microRNA Biomarkers for Sepsis

2018- 2019 **Senior Research Fellow**, National University of Singapore, Singapore

Projects:

Immune gene genetics moderate the association between antenatal maternal depressive symptoms and brain morphology from birth to childhood

2013- 2018 **Senior Research Fellow**, Singapore Institute for Clinical Sciences, A*STAR, Singapore

Projects:

1) Integrative epigenome-wide analysis across umbilical cord and cord blood reveal that common DNA methylation pathway change may

mediate genetic risk associated with preterm birth

2) Next-Generation Sequencing (NGS) data analysis and integration with methylation data for insulin resistant man samples collected by the Singapore Adult Metabolism Study (SAMS)

3) Cell Type Proportion Prediction for samples based on Illumina methylation data

4) Age Study based on methylation data

2007- 2013 **Senior Bioinformatics Specialist**, National Cancer Center of Singapore, Singapore

Projects:

1) Next-Generation Sequencing (NGS) data analysis

* Whole-genome sequencing
* Exome sequencing
* Transcriptomic analysis (RNA-Seq)

2) Gene Expression Microarray data analysis

* Integrated analysis of cross-platform microarray data.
* Co-expression Network
* Module analysis
* Pathway analysis

3) Providing bioinformatics support and analysis for other projects in NCCS and DUKE-NUS as a member of these bioinformatics analysis teams

2004-2007 **Senior Bioinformatics Specialist**, Agenica Research Pte Ltd, Singapore

Projects:

1) **DNA microarray data analysis**

* SNP
* Copy Number variation
* LOH

2) **Cancer clinic data analysis**

* Survival analysis
* Patient Clustering analysis

3) **Bioinformatics support and analysis for other projects**

2003-2004 **Research Fellow**, National University of Singapore, Singapore

Research on Phylogenetic Footprinting of Genes

2001-2002 **Research Assistant**, City University of Hong Kong, HK

Research on Bioinformatics, including:

* The algorithms to recognize protein coding of Gene sequences
* Mining information to predict a protein coding
* Creation of software for Gene visualization
* Microarray normalization

Publications:

Yonghui Wu, Heike Grabsch, Tatiana Ivanova, Iain Beehuat Tan, Jacinta Murray, Chia Huey Ooi, Alexander Ian Wright, Nicholas P West, Gordon G A Hutchins, Jeanie Wu, Minghui Lee, Julian Lee, Jun Hao Koo, Khay Guan Yeoh, Nicole van Grieken, Bauke Ylstra, Sun Young Rha, Jaffer A Ajani, Jae Ho Cheong, Sung Hoon Noh, Kiat Hon Lim, Alex Boussioutas, Ju-Seog Lee, Patrick Tan. Comprehensive genomic meta-analysis identifies intra-tumoural stroma as a predictor of survival in patients with gastric cancer. Gut (2012) Jun 26. [PMID: 22735568] **This paper has been selected as an Editor's Choice by Gut and recommended as being of special significance in its field by F1000.**

Anqi Qiu, Han Zhang, Changqing Wang, Yap-Seng Chong, Lynette P. Shek, Peter D. Gluckman, Michael J. Meaney, Marielle V. Fortier, and **Yonghui Wu.** Canonical TGF-β Signaling Regulates the Relationship between Prenatal Maternal Depression and Amygdala Development in Early Life. Transl Psychiatry. (2021) Mar 15;11(1):170. doi: 10.1038/s41398-021-01292-z. [PMID: 33723212; PMCID: PMC7961018].

Yonghui Wu, Han Zhang, Changqing Wang, Birit F. P. Broekman, Yap-Seng Chong, Lynette P. Shek, Peter D. Gluckman, Michael J. Meaney, Marielle V. Fortier and Anqi Qiu. Inflammatory modulation of the associations between prenatal maternal depression and neonatal brain. Neuropsychopharmacology. (2020) Jul 20. doi: 10.1038/s41386-020-0774-0. [PMID: 32688365]

Yonghui Wu, Xinyi Lin, Ives Yubin Lim, Li Chen, Ai Ling Teh, Julia L. MacIsaac, Kok Hian Tan, Michael S. Kobor, Yap Seng Chong, Peter D. Gluckman and Neerja Karnani. Analysis of two birth tissues provides new insights into the epigenetic landscape of neonates born preterm. Clinical Epigenetics (2019) Feb 11;11(1):26. doi: 10.1186/s13148-018-0599-4. [PMID: 30744680]

Eugenia Migliavacca, Stacey K H Tay, Harnish P Patel, Tanja Sonntag, Gabriele Civiletto, Craig McFarlane, Terence Forrester, Sheila J Barton, Melvin K Leow, Elie Antoun, Aline Charpagne, Yap Seng Chong, Patrick Descombes, Lei Feng, Patrice Francis-Emmanuel, Emma S Garratt, Maria Pilar Giner, Curtis O Green, Sonia Karaz, Narasimhan Kothandaraman, Julien Marquis, Sylviane Metairon, Sofia Moco, Gail

Nelson, Sherry Ngo, Tony Pleasants, Frederic Raymond, Avan A Sayer, Chu Ming Sim, Jo Slater-Jefferies, Holly E Syddall, Pei Fang Tan, Philip Titcombe, Candida Vaz, Leo D Westbury, Gerard Wong, **Wu Yonghui**, Cyrus Cooper, Allan Sheppard, Keith M Godfrey, Karen A Lillycrop, Neerja Karnani and Jerome N Feige. Mitochondrial oxidative capacity and NAD + biosynthesis are reduced in human sarcopenia across ethnicities. Nat Commun. (2019) Dec 20;10(1):5808. doi: 10.1038/s41467-019-13694-1. [PMID: 31862890]

Xinyi Lin, Ives Yubin Lim, **Yonghui Wu**, Ai Ling Teh, Li Chen, Izzuddin M. Aris, Shu E. Soh, Mya Thway Tint, Julia L. MacIsaac, Alexander M. Morin, Fabian Yap, Kok Hian Tan, Seang MeiSaw, Michael S. Kobor, Michael J. Meaney, Keith M. Godfrey, Yap Seng Chong, Joanna D. Holbrook, Yung Seng Lee, Peter D. Gluckman, Neerja Karnani, and on behalf of the GUSTO study group. Developmental pathways to adiposity begin before birth and are influenced by genotype, prenatal environment and epigenome. BMC Med. 2017 Mar 7;15(1):50. doi: 10.1186/s12916-017-0800-1. [PMID: 28264723]

Karen Lillycrop, Robert Murray, Clara Cheong, Ai Ling Teh, Rebecca Clarke-Harris, Sheila Barton, Paula Costello, Emma Garratt, Eloise Cook, Philip Titcombe, Bhuvaneshwari Shunmuganathan, Samantha J. Liew, Yong-Cai Chua, Xinyi Lin, **Yonghui Wu**, Graham C. Burdge, Cyrus Cooper, Hazel M. Inskip, Neerja Karnani, James C. Hopkins, Caroline E. Childs, Carolina Paras Chavez, Philip C. Calder, Fabian Yap, Yung Seng Lee, Yap Seng Chong, Philip E. Melton, Lawrie Beilin, Rae-Chi Huang, Peter D. Gluckman, Nick Harvey, Mark A. Hanson, Joanna D. Holbrook, The EpiGen Consortium, and Keith M. Godfrey. ANRIL Promoter DNA Methylation: A Perinatal Marker for Later Adiposity. EBioMedicine. 2017 May; 19:60-72. doi: 10.1016/j.ebiom.2017.03.037. Epub 2017 Apr 26. [PMID: 28473239]

R C Huang, E S Garratt, H Pan, **Y Wu**, E A Davis, S J Barton, G C Burdge, K M Godfrey, J D Holbrook, and K A Lillycrop. Genome-wide methylation analysis identifies differentially methylated CpG loci associated with severe obesity in childhood. Epigenetics. 2015 Nov; 10(11): 995–1005. [PMID: 26646899]

Hong Pan, Xinyi Lin, **Yonghui Wu**, Li Chen, Ai Ling The, Shu E Soh, Yung Seng Lee, Mya Thway Tint, Julia L MacIsaac, Alexander M Morin, Kok Hian Tan, Fabian Yap, Seang Mei Saw, Michael S Kobor, Michael J Meaney, Keith M Godfrey, Yap-Seng Chong, Peter D Gluckman, Neerja Karnani, Joanna D Holbrook and GUSTO Study Group. HIF3A association with adiposity: the story begins before birth. Epigenomics. 2015; 7(6):937-50. doi: 10.2217/epi.15.45. Epub 2015 May 26. [PMID:26011824]

Zahra Kabiri, Gediminas Greicius, Babita Madan, Steffen Biechele, Zhendong Zhong, Hamed Zaribafzadeh, Edison, Jamal Aliyev, **Yonghui**

Wu, Ralph Bunte, Bart O. Williams, Janet Rossant, and David M. Virshup. Stroma provides an intestinal stem cell niche in the absence of epithelial Wnts. Development (2014) Jun; 141(11): 2206-15. [PMID: 24821987]

Tatiana Ivanova, Hermioni Zouridis, **Yonghui Wu**, Lai Ling Cheng, Iain Beehuat Tan, Veena Gopalakrishnan, Chia Huey Ooi, Julian Lee, Luo Qin, Jeanie Wu, Minghui Lee, Sun Young Rha, Dan Huang, Natalia Liem, Khay Guan Yeoh, Wei Peng Yong, Bin Tean Teh, Patrick Tan. Integrated epigenomics identifies BMP4 as a modulator of cisplatin sensitivity in gastric cancer. Gut. 2013 Jan;62(1):22-33. doi: 10.1136/gutjnl-2011-301113. Epub 2012 Apr 25. [PMID: 22535375]

Hermioni Zouridis, Niantao Deng, Tatiana Ivanova, Yansong Zhu, Bernice Wong, Dan Huang, **Yonghui Wu**, Yingting Wu, Iain Beehuat Tan, Natalia Liem, Veena Gopalakrishnan, Qin Luo, Jeanie Wu, Minghui Lee, Wei Peng Yong, Liang Kee Goh, Bin Tean Teh,Steve Rozen and Patrick Tan. Methylation subtypes and large-scale epigenetic alterations in gastric cancer. Sci Transl Med. 2012 Oct 17; 4(156):156ra140. doi: 10.1126/scitranslmed.3004504. [PMID: 23076357]

Lai Ling Cheng, Yoko Itahana, Zheng Deng Lei, Na-Yu Chia, **Yonghui Wu**, Yingnan Yu, Shen Li Zhang, Aye Aye Thike, Anuradha Pandey, Steve Rozen, Pieter Mathijs Voorhoeve, Qiang Yu, Puay Hoon Tan, Boon Huat Bay, Koji Itahana, and Patrick Tan. TP53 Genomic Status Regulates Sensitivity of Gastric Cancer Cells to theHistone Methylation Inhibitor 3-Deazaneplanocin A(DZNep). Clin Cancer Res August 1, 2012 18: 4201-4212; Published Online First June 6, 2012; doi:10.1158/1078-0432.CCR-12-0036. [PMID: 22675170]

Tracy M. Covey, Simran Kaur, Tina Tan Ong, Kyle D. Proffitt, **Yonghui Wu**, Patrick Tan, David M. Virshup. PORCN moonlights in a Wnt-independent pathway that regulates cancer cell proliferation. PLoS One. 2012; 7(4): e34532. doi: 10.1371/journal.pone.0034532. Epub 2012 Apr 11. [PMID: 22509316]

Jin Zhou, Geng Bo Chen, Yew Chung Tang, **Yonghui Wu**, Chui Sun Yap, Guihua Wang, Junbo Hu, Xianmin Xia, Patrick Tan, Liang Kee Goh, Paul Michael Yen and Rohit AnthonySinha. Genetic and bioinformatic analyses of the expression and function of PI3K regulatory subunit PIK3R3 in an Asian patient gastric cancer library. BMC Med Genomics. 2012 Aug 9;5:34. doi: 10.1186/1755-8794-5-34. [PMID: 22876838]

Philippe Broët, Sophie Camilleri-Broët, Shenli Zhang, Marco Alifano, Dhinoth Bangarusamy, Maxime Battistella, **Yonghui Wu**, Marianne Tuefferd, Jean-François Régnard, Elaine Lim, Patrick Tan, and Lance D. Miller. Prediction of clinical outcome in multiple lung cancer cohorts by integrative genomics: implications for chemotherapy selection. Cancer Res. 2009 Feb 1;69(3):1055-62. doi: 10.1158/0008-5472.CAN-08-1116. Epub

2009 Jan 27. [PMID: 19176396]

Han C. Toh, Li Sun, Yatanar Soe, **Yonghui Wu**, Yee P. Phoon, Whay K. Chia, Jeanie Wu, Kee Y. Wong, Patrick Tan. G-CSF induces a potentially tolerant gene and immunophenotype profile in T cells in vivo. Clin Immunol. 2009 Jul;132(1):83-92. doi: 10.1016/j.clim.2009.03.509. Epub 2009 Apr 2. [PMID: 19345152]

Qingsong Hou, **Yonghui Wu**, Heike Grabsch, Yansong Zhu, Siew Hong Leong, Kumaresan Ganesan, Debra Cross, Lay Keng Tan, Jiong Tao, Veena Gopalakrishnan, Bor Luen Tang, Oi Lian Kon, and Patrick Tan. Integrative Genomics Identifies RAB23 as an Invasive Mediator Gene in Diffuse Gastric Cancer. Cancer Res. 2008; 68: 4623-4630. [PMID: 18559507]

Kumaresan Ganesan, Tatiana Ivanova, **Yonghui Wu**, Vikneswari Rajasegaran, Jeanie Wu, Ming Hui Lee, Kun Yu, Sun Young Rha, Hyun Cheol Chung, Bauke Ylstra, Gerrit Meijer, Kon Oi Lian, Heike Grabsch, and Patrick Tan. Inhibition of Gastric Cancer Invasion and Metastas is by PLA2G2A, a Novel β-catenin/TCF Target Gene. Cancer Res. (2008) 68(11): 4277-428. [PMID: 18519687]

Kaia Davis Tan, Yansong Zhu & **Yonghui Wu**. Amplification and overexpression of PPFIA1, a putative 11q13 invasion suppressor gene, in head and neck squamous cell carcinoma. Genes, Chromosomes and Cancer, Volume 47, Issue 4, pages 353–362, April 2008. [PMID: 18196592]

Alan Wee-Chung Liew, **Yonghui Wu**, Hong Yan, Mengsu Yang: Effective statistical features for coding and non-coding DNA sequence classification for yeast, C. elegans and human. IJBRA 1(2): 181-201 (2005). [PMID: 18048129]

Yonghui Wu, Alan Wee-Chung Liew, Hong Yan and Mengsu Yang. Statistical Features for Classification of Short Human Exons and Introns. The First Asia-Pacific Bioinformatics Conference. Phys. Rev. E 67, 061916 (2003). [PMID: 16241270]

Yonghui Wu, Alan Wee-Chung Liew, Hong Yan and Mengsu Yang. DB-Curve: A Novel 2D Method of DNA Sequence Visualization and Representation. Chemical Physics Letters. Chemical Physics Letters, Volume 367, Issues 1-2, 2 January 2003, Pages 170-176.

Alan Wee-Chung Liew, **Yonghui Wu**, Hong Yan: Selection of Statistical Features Based on Mutual Information for Classification of Human Coding and Non-coding DNA Sequences. ICPR (3) 2004: 766-769.

Yonghui Wu, Dichen Li and Bingheng Lu. Research on artificial Bone Fabrication Based on RP China Mechanical Engineering. Vol.12, No.4,

392-394, 2001.

Yonghui Wu, Dichen Li and Bingheng Lu. The research on ECM fabrication of tissue engineering based on RP. Foreign Medical Sciences (Biomedical Engineering Fascicle). Vol.24, No.3, 102-106, 2001.

Minlin Sun, **Yonghui Wu**. Physical and chemical properties and extended release action of calcium phosphate cement/BMP composite bone graft material. Journal of The Third Military Medical University. 2001.7. 800~802.

Yonghui Wu, Dichen Li and Bingheng Lu. Research for Biological Manufacturing Engineering Based on Rapid Prototyping. Chinese Journal of Mechanical Engineering, Vol.36, No.9, 61-64, 2000.

Yonghui Wu, Dichen Li and Bingheng Lu. Research on Bioactive Bionics Fabrication Based on RP. China Mechanical Engineering. Vol.11, No.10, 1090-1092, 2000.

Yonghui Wu, Dichen Li and Bingheng Lu. Development of RP System for the Fabrication of Artificial Bioactive Bone, Rapid Development Technology of New Products in 21 Century, Shaanxi Science & Technology Publishing House, 152-158, 2000.5.

Bingheng Lu, **Yonghui Wu** and Dichen Li. Rising Bio-Manufacturing Engineering in 21st Century. China Mechanical Engineering. Vol.11, No.1-2, 149-153, 2000.

Dichen Li, **Yonghui Wu**. The Research on Rapid Manufacturing of Artificial Bioactive Bone, China Mechanical Engineering. Vol.11, supplement, 102-104, 2000.

Yonghui Wu, Dichen Li, Xuhong Liu and Bingheng Lu. Research on artificial bone fabrication method based on RP, International Conference on Advanced manufacturing Technology, Xi'an, China, 1999. 6.

Zhilin Qiao, **Yonghui Wu**. The Realization and Application of Image Layer Removing Method in Reverse Engineering, Rapid Prototyping & Rapid Mold Manufacturing, Shaanxi Science & Technology Publishing House, 428-430, 1998.5.

Chengming Luo, **Yonghui Wu**. The fabrication of Title Prototype Based on RP, Rapid Prototyping & Rapid Mold Manufacturing, Shaanxi Science & Technology Publishing House, 59-61,1998.5.

Honors and Awards:

Period for Ph.D.:

- ✧ First-class Excellent Student Honor and Scholarship of Xi'an Jiaotong University (2001)
- ✧ The Honor of National Competition Winner of Xi'an Jiaotong University (2001)
- ✧ The Silver Award of "Challenge Cup" Chinese Business Plan Competition (2000, The First Attendee)
- ✧ Second-class Award of the 2nd National Computer Competition (1999, The Third Attendee)
- ✧ **Xinxing** Scholarship (1999)
- ✧ First-class Excellent Student Honor and Scholarship of Xi'an Jiaotong University (1999)
- ✧ **Xinxing** Scholarship (1997)
- ✧ First-class Excellent Student Honor and Scholarship of Xi'an Jiaotong University (1997)

Period for B.S.:

- ✧ Excellent Student Honor and Scholarships Every Year

Other Technical Skills:

Computer skills:

Proficient in Matlab, R. Lots of experiences in Linux, Java, SPSS, SQL, BC++, C, VC++, C#, Delphi and various other software programs.