William KM Lai

Cornell University
458 Biotechnology Building
Ithaca, NY 14850
(716) – 435 – 9187
wkl29@cornell.edu

SCIENTIFIC EDUCATION & EMPLOYMENT

2020 – Present	**Assistant Research Professor** in the Department of Molecular Biology and Genetics and the Department of Computational Biology, Cornell University.
2018 – 2020	**Assistant Research Professor** in the Department of Biochemistry and Molecular Biology, Pennsylvania State University.
2013 – 2018	**Postdoctoral Researcher** in the Department of Biochemistry and Molecular Biology, Pennsylvania State University in the laboratory of B. Franklin Pugh.
2008 – 2013	**PhD in Biochemistry** from SUNY Buffalo in the laboratory of Michael Buck.
2005 – 2008	**BS in Neuroscience,** University of Pittsburgh, Pennsylvania Minor in Chemistry 2008, Latin Honors: Cum Laude
2003 – 2008	**Laboratory Research Assistant** in the Department of Biological Sciences, SUNY Buffalo in the laboratory of Richard R. Almon, PhD.

JOURNAL PUBLICATIONS

28. Nwosu ZC; Giza H; Nassif M; Charlestin V; Menjivar RE; Kim D; Kemp SB; **Lai WKM**; Loveless I; Steele NG; Hu J; Hu B; Wang S; Magliano MP; Lyssiotis CA. Multi-dimensional analyses identify genes of high priority for pancreatic cancer research. JCI Insight. 2024, (*accepted*).

27. Lang O; Srivastava D; Pugh BF; **Lai WKM***. GenoPipe: identifying the genotype of origin within (epi)genomic datasets. *Nucleic Acids Research.* 2023, Nucleic Acids Research 51 (22), 12054-12068. (*corresponding author)

26. Mittal C; Lang O; **Lai WKM**; Pugh BF. An integrated SAGA and TFIID PIC assembly pathway selective for poised and induced promoters. *Genes & Development*. 2022, 36(17-18):985-1001.

25. John J; Jabbar J; Badjatia N; Rossi MJ; **Lai WKM**; Pugh BF. Genome-wide promoter assembly in E. coli measured at single-base resolution. *Genome Research*. 2022, 32(5):878-892.

24. Shao, D; Kellogg G; Nematbakhsh A; Kuntala PK; Mahony S; Pugh BF; **Lai WKM***. PEGR: a flexible management platform for reproducible epigenomic and genomic research. *Genome Biology.* 2022, 23(1):99. (*corresponding author). https://github.com/seqcode/pegr

23. Sun Q, Nematbakhsh A, Kuntala PK, Kellogg G, Pugh BF, **Lai WKM***. STENCIL: A web templating engine for visualizing and sharing life science datasets. *PLoS Computational Biol*ogy. 2022, 18(2):e1009859. (*corresponding author) https://github.com/CEGRcode/stencil

22. Zhao, T; Vvedenskaya, IO; **Lai WKM**; Basu, S; Pugh, BF; Nickels, BE; Kaplan CD. Ssl2/TFIIH function in transcription start site scanning by RNA polymerase II in Saccharomyces cerevisiae. *eLife*. 2021, (10):71013

21. **Lai, WKM**; Mariani, L; Rothschild, G; Smith, ER; Venters, BJ; Blanda, TR; Kuntala, PK; Bocklund, K; Mairose, J; Dweikat, SN; Mistretta, K; Rossi, MJ; James, D; Anderson, JT; Phanor, SK; Zhang, W; Zhao, Z; Shah, AP; Novitzky, K; McAnarney, E; Keogh MC; Shilatifard A; Basu U; Bulyk M; Pugh, BF. A ChIP-exo screen of 887 Protein Capture Reagents Program transcription factor antibodies in human cells. *Genome Research*. 2021, *31* (9):1663–1679.

20. Rossi MJ; Kuntala PK; **Lai WKM**; Yamada N; Badjatia N; Mittal C; Kuzu G; Bocklund K; Farrell NP; Blanda TR; Mairose JD; Basting AV; Mistretta KS; Rocco DJ; Perkinson ES; Kellogg GD; Mahony S; Pugh BF. A high-resolution protein architecture of the budding yeast genome. *Nature.* 2021, 592, 309-314.

19. Badjatia, N; Rossi MJ; Bataille AR; Mittal C; **Lai WKM**; Pugh BF. Acute stress drives global repression through two independent RNA polymerase II stalling events in Saccharomyces. *Cell Reports.* 2021, *34* (3):108640

18. Qiu C; Jin H; Vvedenskaya I; Llenas JA; Zhao T; Malik I; Visbisky AM; Schwarts SL; Cui P; Čabart P; Han KH; **Lai WKM**; Metz RP; Johnson C; Sze SH; Pugh BF; Nickels BE; Kaplan CD. Universal promoter scanning by Pol II during transcription initiation in Saccharomyces *Genome Biology.* 2020, *21* (1-31)

17. Yamada N; **Lai WKM**; Farrell, N; Pugh BF; Mahony S. Characterizing protein-DNA binding event subtypes in ChIP-exo data. *Bioinformatics.* 2019, *35* (6):903-913

16. Rossi MJ; **Lai WKM**; Pugh BF. Simplified ChIP-exo and ChIP-seq assays. *Nature Communications*. 2018, *9* (2842).

15. Rossi, MJ; **Lai WKM**; Pugh, BF. Genome-wide determinants of sequence-specific DNA binding of general regulatory factors. *Genome Research.* 2018, *28* (4):497-508

14. Rossi, MJ; **Lai WKM**; Pugh BF. Correspondence: DNA shape is insufficient to explain binding. *Nature Communications.* 2017, 8:15643

13. **Lai WKM**; Pugh BF. Understanding nucleosome dynamics and their links to gene expression and DNA replication. *Nature Reviews Molecular Cell Biology*. 2017, *18* (9):548

12. **Lai WKM**; Pugh BF. Genome-wide uniformity of human 'open' pre-initiation complexes. *Genome Research.* 2017, 15-26

11. Paul E; Tirosh I; **Lai W**; Buck MJ; Palumbo MJ; Morse RH. Chromatin Mediation of a Transcriptional Memory Effect in Yeast. *G3 Genes|Genomes|Genetics*. 2015, *5* (5):829-838

10. Puri S*; **Lai WKM***; Rizzo JM*; Edgerton M; Buck MJ. Iron-responsive chromatin remodeling and MAPK signaling enhance adhesion in Candida albicans. *Molecular Microbiology*. 2014, *93* (2):291-305 (*co-first authors)

9. **Lai WKM**; Buck MJ. An Integrative Approach to Understanding the Combinatorial Histone Code at Functional Elements. *Bioinformatics.* 2013, *29* (18):2231-7

8. Givens, R; **Lai, W**; Rizzo, J; Bard, J; Mieczkowski, P; Leatherwood, J; Huberman, J; Buck, M. Chromatin architectures at fission yeast transcriptional promoters and replication origins. *Nucleic Acids Research*. 2012, *40* (15):7176-7189

7. **Lai WKM**; Bard JE; Buck MJ. ArchTEx: accurate extraction and visualization of next-generation sequence data. *Bioinformatics*. 2012, *28* (7):1021-3 https://github.com/WilliamKMLai/ArchTEx

6. **Lai WKM**; Buck MJ. ArchAlign: coordinate-free chromatin alignment reveals novel architectures. *Genome Biology*. 2010, *11* (R126) **Highly Accessed** https://github.com/WilliamKMLai/ArchAlign

5. Escamilla-Hernandez R; Chakrabarti R; Romano RA; Smalley K; Zhu QQ; **Lai W**; Halfon MS; Buck MJ; Sinha S. Genome-wide search identifies *Ccnd2* as a direct transcriptional target of Elf5 in mouse mammary gland. *BMC Molecular Biology.* 2010, *11* (68)

4. Almon RR; DuBois DC; **Lai W**; Xue B; Nie J; Jusko WJ. Gene expression analysis of hepatic roles in cause and development of diabetes in Goto-Kakizaki rats. *Journal of Endocrinology*. 2009, 200 (3):331-46

3. Almon RR; Yang E; **Lai W**; Androulakis IP; Ghimbovschi S; Hoffman EP; Jusko WJ; DuBois DC. Relationships between Circadian Rhythms and Modulation of Gene Expression by Glucocorticoids in Skeletal Muscle. *American Journal of Physiology. Regulatory, Integrative, and Comparative Physiology*. 2008, *295* (4):R1031-47

2. Almon RR; Yang E; **Lai W**; Androulakis IP; DuBois DC; Jusko WJ. Circadian variations in rat liver gene expression: relationships to drug actions. *Journal of Pharmacology and Experimental Therapeutics*. 2008, *326* (3):700-16

1. Almon RR; **Lai W**; DuBois DC; Jusko WJ. Corticosteroid-regulated Genes in Rat Kidney: Mining Time Series Data. *American Journal of Physiology. Endocrinology and Metabolism.* 2005, *289* (5):E870-82

CONFERENCE PUBLICATIONS (ACM-indexed)

3. Gafur J; Goddard S; **Lai WKM**. Adversarial Robustness and Explainability of Machine Learning Models. *Practice and Experience in Advanced Research Computing.* 2024. https://github.com/EpiGenomicsCode/Adversarial_Observation

2. Lang O; Pugh BF; **Lai WKM**. ScriptManager: an interactive platform for reducing barriers to genomics analysis. *Practice and Experience in Advanced Research Computing.* 2022. **Best Short Paper Award.** https://github.com/CEGRcode/scriptmanager

1. Shao D; Kellogg G; Mahony S; **Lai W**; Pugh BF. PEGR: a management platform for ChIP-based next generation sequencing pipelines. *Practice and Experience in Advanced Research Computing.* 2020, 285-292

MANUSCRIPTS UNDER CONSIDERATION

Gafur, J; Lang O; **Lai WKM***. Sequence-free identification of enhancers identifies conserved patterns of chromatin. (*corresponding author). *Nature Communications - submitted*

Krebs JE; Chen H; Lang OW; **Lai WKM**; Pugh BG. Translational and rotational setting of nucleosomes across a human genome. *Nature Genetics - submitted*

Chen H; Krebs JE; Lang OW; **Lai WKM**; Pugh BF. Translational and rotational relationships among transcription factors and nucleosomes. *Nature - submitted*

Arora S; Yang J; Akiyama T; James DQ; Morrissey A; Blanda TR; Badjatia N; **Lai WKM**, Ko MSH; Pugh BF; Mahony S. Joint sequence & chromatin neural networks characterize the differential abilities of Forkhead transcription factors to engage inaccessible chromatin. *bioArxiv*. https://doi.org/10.1101/2023.10.06.561228

MANUSCRIPTS IN PREPARATION

Lyndaker, A; Smugereski S; Mellini D; Wang S; **Lai WKM***. High-resolution chromatin mapping and functional characterization of Pol II Regulatory complexes. (*corresponding author).

RESEARCH SUPPORT

2024-2028 **NIH R35GM155380**

Title: “*Deconvolving the language of protein binding*”

Direct Costs: $1,250,000

2022-2025 **ACCESS**
BIO230120: “Prototyping a science gateway for educational outreach”
BIO230120: “Exploring Alphafold2 as a service”
BIO230120: “Rigor and Reproducibility: making FAIR-compliant software accessible in genomics research”
NSF Reported Value: $338,600

2022-2024 **NIH RM1GM139738 – Early Stage Investigator Pilot Project**
Title: “*High Resolution Chromatin Mapping and Functional Characterization of Pol II Regulatory Complexes*”
Direct Costs: $354,167

2022-2023 **Cornell MISG Track 2**
Title: “*Epigenetic determinants of patient outcome in pancreatic cancer*”
co-PIs: William KM Lai & Rohit Chandwani
Direct Costs: $75,000

2018-2022 **XSEDE Startup Allocation**, supported by National Science Foundation grant number ACI-1548562.
Allocation ID: TG-MCB180094
NSF Reported Value: $50,000

2019 **ICDS Computational and Data Science Seed Grant**, Penn State University.
Direct Costs: $35,000

2018 **NVIDIA Academic GPU Grant**, NVIDIA Corporation, Santa Clara, CA
Title: “*Deconvolution of the gene and epigenetic regulatory code*”

SERVICE

Directorship
2024 – Present Director, Cornell Center for Vertebrate Genomics

2022 – Present Director, Epigenomics Core Facility

Peer Review
2016 – Present Reviewer: BMC Genomics, Gigascience, Nature Methods, Cell Systems, Briefings in Bioinformatics, Genome Research, Nucleic Acids Research
2015 – 2017 Judge: Penn State Undergraduate Exhibit Poster Session

Academic Committees

2021	MBG Faculty Search Process Task Force

External Advisory Boards

2023 – Present	External Advisory Board Member for NSF ACCESS

Internal Advisory Boards

2023 – Present	Faculty Senator
2020 – 2022	EpiGenomics Core Faculty Advisory Board

Society Membership

2012 – 2024	Member, International Society for Computational Biology
2010 – 2012	Member, American Statistical Association

MENTORSHIP

2023 – 2024	Assoc. Director of Training, Cornell Center for Vertebrate Genomics
	- Oversee graduate and post-doc executive committee
2022 – Present	Freshman Undergraduate Advisor
	- Freshman advisor to >20 students
2022 – Present	Computational Biology Undergraduate Honors Thesis Co-Chair
	- Advised and reviewed honors theses for computational biology undergraduates

PhD Thesis Committee Member

Olivia Lang, Computational Biology Field, Cornell University
Justin Cha, Computational Biology Field, Cornell University (Thesis Co-Chair)
Kaiyu Mu, Nutritional Science Field, Cornell University
Brent Basso, Computational Biology Field, Cornell University

TEACHING

2023 – Present	Course Instructor, Cornell University, CB2010
2022 – Present	Course Instructor, Cornell University, BIOG 2290 (Independent Research)
2022	Guest Lecturer, Cornell University, BIOMG1320
2020	Workshop Instructor, Cornell University, Intro to EpiGenomic Analysis
2014	Course Instructor, Penn State University, Molecular and Cellular Biology II
2010	Graduate Teaching Assistant, University at Buffalo, Medical School Foundation
2006 – 2008	Undergraduate Teaching Assistant, University of Pittsburgh, Chemistry I and II

AWARDS

2015	**BBA Gene Regulatory Mechanisms Best Poster**, 34th Penn State Summer Symposium in Molecular Biology, State College, PA
2010	**Elizabeth Olmsted Ross Award for Outstanding Graduate Poster**, SUNY Buffalo, Buffalo, NY
2008-2009	**University at Buffalo Presidential Fellowship**, SUNY Buffalo, Buffalo, NY
2005-2008	**University of Pittsburgh Honors Full Tuition Scholarship**, University of Pittsburgh, Pittsburgh, PA

PRESENTATIONS

2024	Speaker – Element Biosciences, University of Pennsylvania
2024	Speaker – Element Biosciences, New York City, NY
2024	Speaker – Element Biosciences, Houston, TX
2023	Speaker – Northeast Regional Yeast Meeting (NERY), Cornell University
2023	Speaker – Great Lakes Bioinformatics Conference, McGill University
2022	Speaker – Northeast Regional Laboratory Scientists and Core Directors, Rochester, NY
2022	Speaker – FASEB DataWorks
2021	Speaker – UK Dementia Research Institute, Imperial College of London
2021	Speaker – Cornell Day of Data
2019	Poster – 35th Penn State Summer Symposium in Molecular Biology
2018	Speaker/Poster – RECOMB/ISCB Conference on Regulatory and Systems Genomics with DREAM Challenges, New York City, NY
2018	Poster – Transcription regulation: Chromatin and Polymerase II ASBMB Special Symposia
2017	Poster – Mechanism of Eukaryotic Transcription CSHL
2017	Poster – Penn State Cancer Institute Annual Retreat
2015	Poster – Mechanism of Eukaryotic Transcription CSHL
2015	Speaker/Poster– 34th Penn State Summer Symposium in Molecular Biology
2012	Poster – National Graduate Student Research Conference, NIH
2011	Speaker – NorthEast Regional Yeast Meeting (NERY)
2010	Poster – NorthEast Regional Yeast Meeting (NERY)