William KM Lai

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

SCIENTIFIC EDUCATION & EMPLOYMENT

May 2022 – Present **Director of the Epigenomics Core Facility**, Cornell University, Cornell Institute of Biotechnology

June 2020 – Present **Assistant Research Professor** in the Department of Molecular Biology and Genetics, 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, PhD.

2008 – 2013 **PhD in Biochemistry** from State University of New York at Buffalo, Buffalo, New York in the laboratory of Michael Buck, PhD.
> Thesis: "Computational Tools for Investigating the Role of Chromatin in Regulating Genomic Functional Elements"

2005 – 2008 **BS in Neuroscience,** University of Pittsburgh, Pittsburgh, Pennsylvania
> Minor in Chemistry 2008, Latin Honors: Cum Laude

2003 – 2008 **Laboratory Research Assistant** in the Department of Biological Sciences, State University of New York at Buffalo, Buffalo, New York in the laboratory of Richard R. Almon, PhD.

JOURNAL PUBLICATIONS

28. 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)

27. 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*. 2023.

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 PROCEEDINGS

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

2.	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

1.	**Lai WKM**; Buck MJ. Identifying genomic features by BLASTing through chromatin. *Epigenetics & Chromatin* 6, P43 (2013).

PAPERS IN PREPARATION

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

FUNDING

2023	**NIH R21**

Title: "An in silico approach to designing novel immunoreagents"

Direct Costs: $275,000

Impact score: 28, 7th percentile; *Decision pending*

2022	**Cornell MISG Track 2**

Title: "*Epigenetic determinants of patient outcome in pancreatic cancer*"

co-PIs: William KM Lai & Rohit Chandwani

Direct Costs: $75,000

2022   **NIH RM1 – Early Stage Investigator Pilot Project**

Title: "*High Resolution Chromatin Mapping and Functional Characterization of Pol II Regulatory Complexes*"

Direct Costs: $354,167

2022   **XSEDE Startup Allocation**

Title: "*Rigor and Reproducibility: making FAIR-compliant software accessible in genomics research*"

NSF Reported Value: $50,000

2019   **ICDS Computational and Data Science Seed Grant**, Penn State University.

Direct Costs: $35,000

2018   **XSEDE Startup Allocation**, Extreme Science and Engineering Discovery Environment (XSEDE) supported by National Science Foundation grant number ACI-1548562. Allocation ID: **TG-MCB180094**

2018   **NVIDIA Academic GPU Grant**, NVIDIA Corporation, Santa Clara, CA

Title: "*Deconvolution of the gene and epigenetic regulatory code*"

�▬ SERVICE

2023 – Present   External advisory board member for ACCESS
- *ex officio* representative from the Research Advisory Committee

2023 – Present   Faculty Advisor for the Center for Vertebrate Genomics
- Oversee graduate and post-doc executive committee

2022 – Present   Freshman Undergraduate Advisor
- Office of Undergraduate Biology Advising, 8 students

2022 – Present   Computational Biology Undergraduate Honors Thesis Co-Chair
- Advised and reviewed honors theses for computational biology undergraduates

▆ TEACHING EXPERIENCE

Course Instructor, Fall 2023
Cornell University, Ithaca, New York
        BIO2010

Guest Lecturer, Spring 2022
Cornell University, Ithaca, New York
        BIOMG1320

Course Instructor, Fall 2020
Cornell University, Ithaca, New York
> Genome feature and epigenomics data analysis for ChIP-Seq and ATAC-Seq experiments

Course Instructor, Spring 2014
Pennsylvania State University, State College, Pennsylvania
> BMB 252 Honors – Molecular and Cellular Biology II

Graduate Teaching Assistant, Fall 2010
University at Buffalo, Buffalo, New York

Undergraduate Teaching Assistant, Fall 2006 – Spring 2008
University of Pittsburgh, Pittsburgh, Pennsylvania
> General Chemistry Laboratory 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

2023 GLBIO – Speaker
> *"Adversarial attack identifies conserved features of enhancer chromatin architecture"*

2022 NERLSCD Cloud Computing Panel – Speaker
> *"An end-to-end methodology for data analysis, reproducibility, and dissemination"*

FASEB DataWorks 2022 – Speaker
> *"An end-to-end methodology for data analysis, reproducibility, and dissemination"*

Cornell Day of Data 2021 – Speaker
> *"Mapping the yeast epigenome: developing a methodology for data reproducibility and dissemination"*

35th Penn State Summer Symposium in Molecular Biology 2019 – Poster Presentation

"High-throughput validation of renewable monoclonal antibodies in ChIP-seq/exo experiments"

RECOMB/ISCB Conference on Regulatory and Systems Genomics with DREAM Challenges
2018 – Speaker and Poster Presentation
"Methods of defining "success" in ChIP-seq/exo experiments"

Transcription regulation: Chromatin and Polymerase II ASBMB Special Symposia 2018 – Poster
Presentation
"Methods of defining "success" in ChIP-seq/exo experiments"

Mechanism of Eukaryotic Transcription CSHL 2017 – Poster Presentation
"Genome-wide determinants of sequence-specific DNA binding of general regulatory factors"

Penn State Cancer Institute Annual Retreat 2017 – Poster Presentation
"Application of ChIP-exo to tumor tissue reveals differences in the epigenetic profile between cancer types"

Mechanism of Eukaryotic Transcription CSHL 2015 – Poster Presentation
"High-resolution assays reveal details of mammalian initiation complex organization and function"

34th Penn State Summer Symposium in Molecular Biology 2015 – Speaker and Poster
"High-resolution assays reveal details of mammalian enhanceosome organization and function"

National Graduate Student Research Conference 2012 – Poster Presentation
"*Role of Chromatin in Regulating Genomic Functional Elements*"

NorthEast Regional Yeast Meeting (NERY) 2011 – Speaker
"*ArchAlign and ArchBLAST: Next-Generation Tools to Detect and Understand Chromatin Architecture*"

NorthEast Regional Yeast Meeting (NERY) 2010 – Poster Presentation
"ArchAlign: A Next-Generation Alignment Algorithm to Detect Chromatin Architecture"

PROFESSIONAL ACTIVITIES

2021   **Member**, MBG Faculty Search Process Task Force
2020 – Present, **Member**, EpiGenomics Core Faculty Advisory Board
2016 – Present, **Journal Reviewer**, BMC Genomics, Gigascience, Nature Methods, Cell
            Systems, Briefings in Bioinformatics
2015 – 2017, **Judge**, Penn State Undergraduate Exhibit Poster Session
2012 – 2014, **Member**, Interaction Society for Computational Biology
2010 – 2012, **Member**, American Statistical Association

OTHER RELEVANT SKILLS / COURSES

Molecular Biology Courses:
2013   Human Embryonic Stem Cell Culturing Training Course - Hands on training for the culturing and maintenance of H1 and H9 human embryonic cells lines

Statistics Courses:
2012   Statistics for Bioinformatics
2011   Applied Multivariate Statistics, Statistical Genetics, Statistical Comparison and Association, Introduction to Theoretical Statistics II
2010   Regression Analysis, Math Analysis for Biostatistics, Introduction to Theoretical Statistics I

Experimental Design Courses:
2007   Pharmacokinetic – Pharmacodynamic Modeling Concepts and Applications Summer Course – Course on experimental design with a focus on multiple dosing strategies used to analyze cellular and organismal response from a pharmacological perspective.

Programming Languages:
Java, C++, Perl, Python, R