Uri Keich

Associate Professor

School of Mathematics and Statistics
University of Sydney, NSW 2006
Australia
uri.keich@sydney.edu.au
www.maths.usyd.edu.au/u/uri/
0000-0002-3209-5011
dKdP_kcAAAAJ

Research Interests

Statistical methods for computational biology, with emphasis on false discovery analysis in multiple hypothesis testing. Applications to proteomics (tandem mass spectrometry; primary focus), genomics (sequence alignment, motif finding, DNA replication origins). Development of computationally efficient algorithms for statistical significance estimation.

Education

1996 **Ph.D., Mathematics**, *Courant Institute, New York University*, New York, NY

1991 **M.Sc., Mathematics**, *Technion – Israel Institute of Technology*, Haifa, Israel

1987 **B.Sc., Computer Science & Mathematics**, *Hebrew University*, Jerusalem, Israel, *Summa Cum Laude*

Professional Experience

2016–present **Associate Professor**, *University of Sydney*, Sydney, Australia
School of Mathematics and Statistics

2009–2015 **Senior Lecturer**, *University of Sydney*, Sydney, Australia
School of Mathematics and Statistics

2003–2008 **Assistant Professor**, *Cornell University*, Ithaca, NY
Department of Computer Science

2001–2003 **Project Scientist**, *University of California, San Diego*, La Jolla, CA
Department of Computer Science and Engineering

1999–2000 **Assistant Professor**, *University of California, Riverside*, Riverside, CA
Department of Mathematics

1996–1999 **Von Karman Instructor**, *California Institute of Technology*, Pasadena, CA
Applied Mathematics Department

Leadership Roles

2025–present **Faculty of Science Academic Lead for AI and Assessment (0.4 FTE)**, *University of Sydney*

- Working collaboratively with Associate Heads Education across eight schools, Honours coordinators, and faculty staff
- Developing policies and guidelines for assessments of Honours projects in the age of generative AI
- Devised a relational database approach for curriculum assessment mapping, enabling mapping of assessments from units to learning outcomes for majors, programs, and streams
- This approach will be the basis for program-level assessment redesign in Science and was highly commended by Pro-Vice-Chancellor (Learning & Teaching) Prof. Adam Bridgeman

Jan–May 2025 **Deputy Head of School**, *University of Sydney*
School of Mathematics and Statistics. Role concluded early upon appointment to the faculty-wide Academic Lead position.

- Continued leading the school's engagement with generative AI through strategic presentations and organized meetings on assessment policy implications
- Advised colleagues on adapting teaching approaches to the AI landscape
- Drafted school strategic vision on AI integration into the teaching curriculum

2013–2025 **Statistics Honours Coordinator**, *University of Sydney*

Honors and Awards

2024 Faculty of Science Learning & Teaching Award for Teaching and Learning Excellence (Individual)

2024 Student-initiated Faculty of Science Teacher/Unit of Study Commendation: "Taught exceptionally challenging content in an intuitive and engaging manner. Learnt more under his instruction than I have in entire semesters in the past."

2024 Sydney University Postgraduate Representative Association Supervisor of the Year Award

2015 Best Paper Award, RECOMB 2015

2007–2009 NSF CAREER Award

1997 Wilhelm T. Magnus Memorial Prize for Significant Contributions to the Mathematical Sciences, Courant Institute, NYU

1994–1995 Alfred P. Sloan Doctoral Dissertation Fellowship

1990 Wolf Foundation Prize for M.Sc.

1987 Graduated B.Sc. Summa Cum Laude, Hebrew University of Jerusalem

Publications

2025 Freestone J., Noble WS., Keich U. A semi-supervised framework for diverse multiple hypothesis testing scenarios. *Under revision, Journal of the Royal Statistical Society Series B.*

2025 Solivais AJ., Boekweg H., Smith LM., Noble WS., Shortreed MR., Payne SH., Keich U. Improved detection of differentially abundant proteins through FDR-control of peptide-identity-propagation. *Journal of Proteome Research*, 24(9), 4437–4449.

2025 Freestone J., Käll L., Noble WS., Keich U. How to train a post-processor for tandem mass spectrometry proteomics database search while maintaining control of the false discovery rate. *Journal of Proteome Research*, 24(5), 2266–2279.

2025 Wen B., Freestone J., Riffle M., MacCoss MJ., Noble WS., Keich U. Assessment of false discovery rate control in tandem mass spectrometry analysis using entrapment. *Nature Methods*, 22(7), 1454–1463.

2024 Lu Y., Noble WS., Keich U. A BLAST from the past: revisiting blastp's E-value. *Bioinformatics*, 40(12), btae729.

2024 Freestone J., Käll L., Noble WS., Keich U. Semi-supervised Learning While Controlling the FDR with an Application to Tandem Mass Spectrometry Analysis. *LNCS (RECOMB 2024)*, 14758, 448–453.

2024 Freestone J., Noble WS., Keich U. Analysis of tandem mass spectrometry data with CONGA: Combining Open and Narrow searches with Group-wise Analysis. *Journal of Proteome Research*, 23(6), 1894–1906.

2024 Freestone J., Noble WS., Keich U. Re-investigating the correctness of decoy-based false discovery rate control in proteomics tandem mass spectrometry. *Journal of Proteome Research*, 23(6), 1907–1914.

2024 Lin A., See D., Fondrie WE., Keich U., Noble WS. Target-decoy false discovery rate estimation using Crema. *Proteomics*, 24(8).

2023 Ebadi A., Luo D., Freestone J., Noble WS., Keich U. Bounding the FDP in competition-based control of the FDR. *arXiv*, 2302.11837. [Preprint]

2023 Ebadi A., Freestone J., Noble WS., Keich U. Bridging the False Discovery Gap. *Journal of Proteome Research*, 22(7), 2172–2178.

2023 Rajchert A. and Keich U. Controlling the False Discovery Rate via Competition: is the +1 needed? *Statistics and Probability Letters*, 197, 109819.

2023 Luo D., Ebadi A., Emery K., He Y., Noble WS., Keich U. Competition-based control of the false discovery proportion. *Biometrics*.

2023 Hasam S., Emery K., Noble WS., Keich U. A Pipeline for Peptide Detection Using Multiple Decoys. (Invited Chapter) *Methods in Molecular Biology*, 2426:25–34.

2022 Freestone J., Short T., Noble WS., Keich U. Group-walk: a rigorous approach to group-wise false discovery rate analysis by target-decoy competition. *Bioinformatics*, 38(Supplement 2):ii82–ii88.

2022 Lin A., Short T., Noble WS., Keich U. Improving peptide-level mass spectrometry analysis via double competition. *Journal of Proteome Research*, 21(10): 2412–2420.

2022 Heil LR., Fondrie WE., McGann CD., Federation AJ., Noble WS., MacCoss MJ., Keich U. Building Spectral Libraries from Narrow-Window Data-Independent Acquisition Mass Spectrometry Data. *Journal of Proteome Research*, 21(6): 1382–1391.

2021 Lin A., Plubell DL., Keich U., Noble WS. Accurately Assigning Peptides to Spectra When Only a Subset of Peptides Are Relevant. *Journal of Proteome Research*, 20(8): 4153–4164.

2021 Peres N., Lee AR., Keich U. Exactly Computing the Tail of the Poisson-Binomial Distribution. *ACM Transactions on Mathematical Software*, 47(4): 1–19.

2020 Emery K., Hasam S., Noble WS., Keich U. Multiple competition-based FDR control and its application to peptide detection. *RECOMB 2020*, 54–71.

2019 Emery K., Keich U. Controlling the FDR in variable selection via multiple knockoffs. *arXiv*, 1911.09442V2. [Preprint]

2019 Keich U., Tamura K., Noble WS. Averaging Strategy To Reduce Variability in Target-Decoy Estimates of False Discovery Rate. *Journal of Proteome Research*, 18(2): 585–593.

2018 Keich U., Noble WS. Controlling the FDR in imperfect matches to an incomplete database. *Journal of the American Statistical Association*, 113(523): 973–982.

2017 Noble WS., Keich U. Response to "Mass spectrometrists should search for all peptides, but assess only the ones they care about". *Nature Methods*, 14(7): 644.

2017 Keich U., Noble WS. Progressive calibration and averaging for tandem mass spectrometry statistical confidence estimation: Why settle for a single decoy? *LNCS (RECOMB 2017)*, 10229: 99–116.

2017 Wilson H., Keich U. Accurate small tail probabilities of sums of iid lattice-valued random variables via FFT. *Journal of Computational and Graphical Statistics*, 26(1): 223–229.

2016 Wilson H., Keich U. Accurate pairwise convolutions of non-negative vectors via FFT. *Computational Statistics & Data Analysis*, 101: 300–315.

2016 Manescu D., Keich U. A symmetric length-aware enrichment test. *Journal of Computational Biology*, 23(6): 508–525.

2015 Keich U., Kertesz-Farkas A., Noble WS. Improved False Discovery Rate Estimation Procedure for Shotgun Proteomics. *Journal of Proteome Research*, 14(8): 3148–3161.

2015 Manescu D., Keich U. A symmetric length-aware enrichment test. *RECOMB 2015, LNBI* 9029: 224–242. **(Best Paper Award)**

2015 Kertesz-Farkas A., Keich U., Noble WS. Tandem Mass Spectrum Identification via Cascaded Search. *Journal of Proteome Research*, 14(8): 3027–3038.

2015 Keich U., Noble WS. On the Importance of Well-Calibrated Scores for Identifying Shotgun Proteomics Spectra. *Journal of Proteome Research*, 14(2): 1147–1160.

2014 Tanaka E., Bailey TL., Keich U. Improving MEME via a two-tiered significance analysis. *Bioinformatics*, 30(14): 1965–1973.

2013 Liachko I., Youngblood RA., Keich U.*, Dunham MJ. High-resolution mapping, characterization, and optimization of autonomously replicating sequences in yeast. *Genome Research*, 23(4): 698–704. (*co-corresponding author)

2011 Liachko I., Tanaka E., Cox K., Chung SC., Yang L., Seher A., Hallas L., Cha E., Kang G., Pace H., Barrow J., Inada M., Tye BK., Keich U. Novel Features of ARS Selection in Budding Yeast Lachancea kluyveri. *BMC Genomics*, 12: 633.

2011 Tanaka E., Bailey TL., Grant CE., Noble WS., Keich U. Improved similarity scores for comparing motifs. *Bioinformatics*, 27(12): 1603–1609.

2011 Gupta N., Bandeira N., Keich U., Pevzner PA. Target-Decoy Approach and False Discovery Rate: When Things May Go Wrong. *Journal of The American Society for Mass Spectrometry*, 22(7): 1111–1120.

2011 Ng P., Keich U. Alignment Constrained Sampling. *Journal of Computational Biology*, 18(2).

2010 Bhaskar A., Keich U. Confidently estimating the number of DNA replication origins. *Statistical Applications in Genetics and Molecular Biology*, 9(1): Article 28.

2010 Liachko I., Bhaskar A., Li C., Chung SCC., Tye BK., Keich U. A Comprehensive Genome-Wide Map of Autonomously Replicating Sequences in a Naive Genome. *PLoS Genetics*, 6(5).

2009 Oliver HF., Orsi RH., Ponnala L., Keich U., Wang W., Sun Q., Cartinhour SW., Filiatrault MJ., Wiedmann M., Boor KJ. Deep RNA sequencing of *L. monocytogenes* reveals overlapping and extensive stationary phase and sigma B-dependent transcriptomes. *BMC Genomics*, 10: 641.

2009 Nagarajan N., Keich U. Reliability and efficiency of algorithms for computing the significance of the Mann-Whitney test. *Computational Statistics*, 24(4): 605–622.

2008 Ng P., Keich U. Factoring local sequence composition in motif significance analysis. *Genome Informatics*, 21: 15–26.

2008 Ng P., Keich U. GIMSAN: a Gibbs motif finder with significance analysis. *Bioinformatics*, 24(19): 2256–2257.

2008 Keich U., Gao H., Garretson JS., Bhaskar A., Liachko I., Donato J., Tye BK. Computational detection of significant variation in binding affinity across two sets of sequences with application to the analysis of replication origins in yeast. *BMC Bioinformatics*, 9: 372.

2008 Nagarajan N., Keich U. FAST: Fourier transform based Algorithms for Significance Testing of ungapped multiple alignments. *Bioinformatics*, 24(4): 577–578.

2007 Keich U., Ng P. A conservative parametric approach to motif significance analysis. *Genome Informatics*, 19: 61–72.

2007 Zhi D., Keich U., Pevzner P., Heber S., Tang H. Correcting base-assignment errors in repeat regions of shotgun assembly. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 4(1): 54–64.

2006 Nagarajan N., Ng P., Keich U. Refining motif finders with E-value calculations. *RECOMB Satellite Workshop on Regulatory Genomics*, 73–84.

2006 Ng P., Nagarajan N., Jones N., Keich U. Apples to apples: improving the performance of motif finders and their significance analysis in the Twilight Zone. *Bioinformatics*, 22(14): e393–e401.

2006 Keich U., Nagarajan N. A fast and numerically robust method for exact multinomial goodness-of-fit test. *Journal of Computational and Graphical Statistics*, 15(4): 779–802.

2005 Nagarajan N., Jones N., Keich U. Computing the p-value of the information content from an alignment of multiple sequences. *Bioinformatics*, 21(Suppl 1, ISMB): i311–i318.

2005 Buhler J., Keich U., Sun Y. Designing Seeds for Similarity Search in Genomic DNA. *Journal of Computer and System Sciences*, 70(3): 342–363.

2005 Keich U. sFFT: a faster accurate computation of the p-value of the entropy score. *Journal of Computational Biology*, 12(4): 416–430.

2004 Keich U., Nagarajan N. A faster reliable algorithm to estimate the p-value of the multinomial llr statistic. *WABI 2004*.

2004 Keich U., Li M., Ma B., Tromp J. On Spaced Seeds for Similarity Search. *Discrete Applied Mathematics*, 138(3): 253–263.

2003 Buhler J., Keich U., Sun Y. Designing Seeds for Similarity Search in Genomic DNA. *RECOMB 2003*.

2003 Eskin E., Keich U., Gelfand MS., Pevzner PA. Genome-Wide Analysis of Bacterial Promoter Regions. *Pacific Symposium on Biocomputing*.

2003 Keich U. Stationary Tangent – the Discrete and Non-smooth Cases. *Journal of Time Series Analysis*, 24(2): 173–192.

2002 Keich U., Pevzner PA. Finding motifs in the twilight zone. *RECOMB 2002*.

2002 Keich U., Pevzner PA. Subtle motifs: defining the limits of motif finding algorithms. *Bioinformatics*, 18(10): 1382–1390.

2002 Keich U., Pevzner PA. Finding motifs in the twilight zone. *Bioinformatics*, 18(10): 1374–1381.

2001 Cwikel M., Keich U. Optimal decompositions for the K-functional for a couple of Banach lattices. *Arkiv för Matematik*, 39(1): 27–64.

2000 Keich U. A Possible Definition of A Stationary Tangent. *Stochastic Processes and Their Applications*, 88(1): 1–36.

1999 Keich U. Krein's Strings, the Symmetric Moment Problem, and Extending a Real Positive Definite Function. *Communications on Pure and Applied Mathematics*, 52(10): 1315–1334.

1999 Keich U. On L^p Bounds for Kakeya Maximal Functions and the Minkowski Dimension in R^2. *Bulletin of the London Mathematical Society*, 31: 213–221.

1999 Keich U. Absolute Continuity Between the Wiener and Stationary Gaussian Measures. *Pacific Journal of Mathematics*, 188(1): 95–108.

1999 Keich U. The Entropy Distance Between the Wiener and Stationary Gaussian Measures. *Pacific Journal of Mathematics*, 188(1): 109–128.

1996 Aharoni R., Keich U. A Generalization of the Ahlswede Daykin Inequality. *Discrete Mathematics*, 152: 1–12.

Software Contributions

FDRBench	Tool for conducting entrapment experiments including database generation and FDP estimation methods
PIP-ECHO	FDR control for peptide-identity-propagation (implemented in FlashLFQ)
RESET	Flexible procedures for controlling FDR with side-information in p-value and competition settings
Percolator-RESET	MS/MS database search post-processor addressing FDR control issues (to be implemented in Percolator v4.0)
SGPvalue	Tool assigning studentized-Gumbel based p-values to blastp search results
Crema	Open-source Python tool implementing TDC-based FDR estimation methods
bandsfdp	R package providing upper prediction bands on FDP in competition-based multiple hypothesis testing
stepdownfdp	R package implementing step-down procedure for controlling false discovery proportion
CONGA	Post-processor combining narrow window and open modification MS/MS searches with rigorous FDR control
groupwalk	R package for FDR control with intrinsic group structure

aTDC	Variance reduction in FDR estimates (implemented in Crux toolkit)
multicomp	R package for FDR control using multiple competition
dbSearchFDR	R package for FDR control in imperfect matches to incomplete databases
ShiftConvolvePoibin	R package for Poisson-Binomial distribution computation
Tomtom	Motif comparison tool (contributed to score function redesign)
SADMAMA	Tool for motif scanning and binding affinity variation detection
GIMSAN	Gibbs motif finder with significance analysis
ALICO	Alignment constrained sampling
FAST	Fourier transform based algorithms for significance testing of ungapped multiple alignments
aFFT	Accurate convolution of non-negative vectors
sisFFT	Accurate tail probabilities of sums of iid lattice-valued random variables
bagFFT	Computing exact p-value of the llr statistic for multinomial goodness-of-fit test

Teaching Experience

University of Sydney

2009–2025	STAT2911 – Probability and Statistical Models (Advanced)
2019, 2021–2022	MATH1905 – Statistical Thinking with Data (Advanced)
2017–2019	MSH2 – Probability (Honours level)
2018	MATH1005 – Statistical Thinking about Data
2009–2017	MSH8 – Statistical Methods in Bioinformatics
2010–2011, 2013–2015	STAT3914/STAT3014 – Applied Statistics (Advanced)
2010, 2022, 2024	MATH1907/1933/1972 – Mathematics (Special Studies Program)

Cornell University (2003–2008)

CS 426 – Introduction to Bioinformatics; CS 726 – Computational Molecular Biology; CS 628 – Biological Sequence Analysis; CS 280 – Discrete Structures

Earlier Positions

Courses in real analysis, calculus, probability, statistics, and stochastic processes at UC Riverside, Caltech, and NYU (1991–2000)

Selected Invited/Refereed Talks

2026	Lorentz Center Workshop, Leiden, Netherlands
2026	Technical University of Munich
2025	Faculty of Science Education Showcase, University of Sydney [video]
2025	ASMS 2025 Annual Conference (invited workshop presentation)
2024	Tertiary Education Seminar Series, University of Sydney
2023	Combi Seminar, University of Washington Genome Sciences
2023	Statistics Seminar, Melbourne University
2020	RECOMB 2020
2017	RECOMB 2017
2016	Computer Science and Engineering Colloquium, UCSD

2016	Computer Science Department Colloquium, Princeton University
2015	RECOMB 2015 (Best Paper Award)
2015	AMSI BioInfoSummer, Sydney
2013	Bayes by the Bay Workshop, India (fully funded)
2013	Edmond J. Safra Center for Bioinformatics Distinguished Speaker Series, Tel-Aviv University
2012	RECOMB Conference on Bioinformatics Education (keynote) [video]

Professional Service

Editorial

2025	Invited participant, Molecular & Cellular Proteomics (MCP) workshop on updating reporting guidelines
2010–2022	Editorial Board, BMC Bioinformatics

Program Committees

RECOMB Satellite Workshop on Regulatory Genomics (2004–6, 2008–10, 2012–14); RECOMB Satellite Conference on Open Problems in Algorithmic Biology (2012); Asia-Pacific Bioinformatics Conference (2007, 2009–10); IEEE BIBM (2019); GIW (2008–9)

Reviewing

Journals	Nature Methods (2x in 2025), PNAS, Bioinformatics, Journal of Proteome Research, PLoS Computational Biology, Journal of Computational and Graphical Statistics, and others
Conferences	RECOMB, ISMB, ICML, SODA, and others
Grants	Deutsche Forschungsgemeinschaft (DFG)

Funding

2024	\$8,711 USD – Subcontractor on NSF award 2245300 (PI Noble)
2023–2024	\$27,500 USD – Chan-Zuckerberg Initiative: Single cell proteomics (PI Noble)
2020	\$30,639 USD – University of Washington contractor for R01 GM121818 (PI Noble)
2016	\$11,000 AUD – Faculty of Science Seedfunding Scheme
2014	\$38,892 AUD – Faculties of Sciences Research Equipment Scheme (joint)
2007–2009	\$644,880 USD – NSF CAREER Award (terminated early due to relocation)

Supervision

PhD Students

2021–2024	Jack Freestone (9 joint papers; now Lecturer at Macquarie University)
2016–2019	Kristen Emery (3 joint papers; now in industry)
2010–2014	Emi Tanaka (3 joint papers; now Deputy Director, Biological Data Science Institute, ANU)
2005–2011	Patrick Ng, Cornell University (6 joint papers; now in finance)
2004–2006	Niranjan Nagarajan, Cornell University (7 joint papers; now Associate Director, Genome Institute of Singapore)

MSc Students

2021–2023	Arya Ebadi (2 joint papers)
2015–2016	Huon Wilson (2 joint papers; now at Data61)

Honours Students

Temana Short (2023), Jack Freestone (2020), Arya Ebadi (2020), Dong Luo (2018–19), Yilun He (2018), Guo Xiangyuanchai (2017–18), Kristen Emery (2015), and others

Research Project Students

Supervised 28 undergraduate and Masters of Engineering students through research projects and lab rotations at University of Sydney and Cornell University; 11 joint papers with these students