# Ming Li

## List of Publications by Year
### in descending order

| | | | |
|---|---|---|---|
| 53 papers | 13,473 citations | 279487<br>23 h-index | 205818<br>48 g-index |
| 57 all docs | 57 docs citations | 57 times ranked | 14641 citing authors |

| # | Article | IF | Citations |
|---|---------|-----|-----------|
| 1 | Initial sequencing and comparative analysis of the mouse genome. Nature, 2002, 420, 520-562. | 13.7 | 6,319 |
| 2 | PEAKS: powerful software for peptidede novo sequencing by tandem mass spectrometry. Rapid Communications in Mass Spectrometry, 2003, 17, 2337-2342. | 0.7 | 1,156 |
| 3 | An Introduction to Kolmogorov Complexity and Its Applications. , 1997, , . | | 1,072 |
| 4 | The Similarity Metric. IEEE Transactions on Information Theory, 2004, 50, 3250-3264. | 1.5 | 766 |
| 5 | PatternHunter: faster and more sensitive homology search. Bioinformatics, 2002, 18, 440-445. | 1.8 | 713 |
| 6 | An Introduction to Kolmogorov Complexity and Its Applications. Texts in Computer Science, 2008, , . | 0.5 | 682 |
| 7 | Information distance. IEEE Transactions on Information Theory, 1998, 44, 1407-1423. | 1.5 | 377 |
| 8 | De novo peptide sequencing by deep learning. Proceedings of the National Academy of Sciences of the United States of America, 2017, 114, 8247-8252. | 3.3 | 278 |
| 9 | RAPTOR: OPTIMAL PROTEIN THREADING BY LINEAR PROGRAMMING. Journal of Bioinformatics and Computational Biology, 2003, 01, 95-117. | 0.3 | 245 |
| 10 | Deep learning enables de novo peptide sequencing from data-independent-acquisition mass spectrometry. Nature Methods, 2019, 16, 63-66. | 9.0 | 235 |
| 11 | On the closest string and substring problems. Journal of the ACM, 2002, 49, 157-171. | 1.8 | 205 |
| 12 | ZOOM! Zillions of oligos mapped. Bioinformatics, 2008, 24, 2431-2437. | 1.8 | 187 |
| 13 | Shared Information and Program Plagiarism Detection. IEEE Transactions on Information Theory, 2004, 50, 1545-1551. | 1.5 | 180 |
| 14 | PATTERNHUNTER II: HIGHLY SENSITIVE AND FAST HOMOLOGY SEARCH. Journal of Bioinformatics and Computational Biology, 2004, 02, 417-439. | 0.3 | 170 |
| 15 | On spaced seeds for similarity search. Discrete Applied Mathematics, 2004, 138, 253-263. | 0.5 | 110 |
| 16 | Chain Letters and Evolutionary Histories. Scientific American, 2003, 288, 76-81. | 1.0 | 88 |
| 17 | Complete De Novo Assembly of Monoclonal Antibody Sequences. Scientific Reports, 2016, 6, 31730. | 1.6 | 85 |
| 18 | PICKY: a novel SVD-based NMR spectra peak picking method. Bioinformatics, 2009, 25, i268-i275. | 1.8 | 61 |

| # | ARTICLE | IF | CITATIONS |
|---|---------|-----|-----------|
| 19 | Fragment–HMM: A new approach to protein structure prediction. Protein Science, 2008, 17, 1925-1934. | 3.1 | 57 |
| 20 | DeepIso: A Deep Learning Model for Peptide Feature Detection from LC-MS map. Scientific Reports, 2019, 9, 17168. | 1.6 | 53 |
| 21 | Computing exact P-values for DNA motifs. Bioinformatics, 2007, 23, 531-537. | 1.8 | 38 |
| 22 | On the complexity of the spaced seeds. Journal of Computer and System Sciences, 2007, 73, 1024-1034. | 0.9 | 38 |
| 23 | Computationally instrument-resolution-independent de novo peptide sequencing for high-resolution devices. Nature Machine Intelligence, 2021, 3, 420-425. | 8.3 | 35 |
| 24 | A streamlined platform for analyzing tera-scale DDA and DIA mass spectrometry data enables highly sensitive immunopeptidomics. Nature Communications, 2022, 13, . | 5.8 | 34 |
| 25 | Information distance from a question to an answer. , 2007, , . | | 23 |
| 26 | Personalized deep learning of individual immunopeptidomes to identify neoantigens for cancer vaccines. Nature Machine Intelligence, 2020, 2, 764-771. | 8.3 | 22 |
| 27 | PatternHunter II: highly sensitive and fast homology search. Genome Informatics, 2003, 14, 164-75. | 0.4 | 22 |
| 28 | Optimizing Multiple Spaced Seeds for Homology Search. Lecture Notes in Computer Science, 2004, , 47-58. | 1.0 | 21 |
| 29 | New Information Distance Measure and Its Application in Question Answering System. Journal of Computer Science and Technology, 2008, 23, 557-572. | 0.9 | 19 |
| 30 | PROTEIN SECONDARY STRUCTURE PREDICTION USING NMR CHEMICAL SHIFT DATA. Journal of Bioinformatics and Computational Biology, 2010, 08, 867-884. | 0.3 | 19 |
| 31 | ERROR TOLERANT NMR BACKBONE RESONANCE ASSIGNMENT AND AUTOMATED STRUCTURE GENERATION. Journal of Bioinformatics and Computational Biology, 2011, 09, 15-41. | 0.3 | 18 |
| 32 | Optimizing Multiple Spaced Seeds for Homology Search. Journal of Computational Biology, 2006, 13, 1355-1368. | 0.8 | 17 |
| 33 | Knowledge map construction for question and answer archives. Expert Systems With Applications, 2020, 141, 112923. | 4.4 | 16 |
| 34 | INFORMATION DISTANCE AND ITS APPLICATIONS. International Journal of Foundations of Computer Science, 2007, 18, 669-681. | 0.8 | 15 |
| 35 | Designing succinct structural alphabets. Bioinformatics, 2008, 24, i182-i189. | 1.8 | 11 |
| 36 | Enhanced question understanding with dynamic memory networks for textual question answering. Expert Systems With Applications, 2017, 80, 39-45. | 4.4 | 11 |

| # | Article | IF | Citations |
|---|---------|-----|-----------|
| 37 | Question Classification by Approximating Semantics. , 2015, , . | | 8 |
| 38 | A New Multiword Expression Metric and Its Applications. Journal of Computer Science and Technology, 2011, 26, 3-13. | 0.9 | 7 |
| 39 | Merge-Weighted Dynamic Time Warping for Speech Recognition. Journal of Computer Science and Technology, 2014, 29, 1072-1082. | 0.9 | 7 |
| 40 | A tale of solving two computational challenges in protein science: neoantigen prediction and protein structure prediction. Briefings in Bioinformatics, 2022, 23, . | 3.2 | 7 |
| 41 | Can We Determine a Protein Structure Quickly?. Journal of Computer Science and Technology, 2010, 25, 95-106. | 0.9 | 6 |
| 42 | Protein Threading by Linear Programming: Theoretical Analysis and Computational Results. Journal of Combinatorial Optimization, 2004, 8, 403-418. | 0.8 | 5 |
| 43 | Towards Automated Structure-Based NMR Resonance Assignment. Lecture Notes in Computer Science, 2010, , 189-207. | 1.0 | 5 |
| 44 | A note on the single genotype resolution problem. Journal of Computer Science and Technology, 2004, 19, 254-257. | 0.9 | 4 |
| 45 | Deep Omics. Proteomics, 2018, 18, 1700319. | 1.3 | 4 |
| 46 | Deep neural network for detecting arbitrary precision peptide features through attention based segmentation. Scientific Reports, 2021, 11, 18249. | 1.6 | 4 |
| 47 | Information distance between what I said and what it heard. Communications of the ACM, 2013, 56, 70-77. | 3.3 | 3 |
| 48 | JBCB, the first decade. Journal of Bioinformatics and Computational Biology, 2014, 12, 1401001. | 0.3 | 1 |
| 49 | Deep Learning Deciphers Protein–RNA Interaction. Genomics, Proteomics and Bioinformatics, 2019, 17, 475-477. | 3.0 | 1 |
| 50 | Partial Match Distance. Lecture Notes in Computer Science, 2013, , 55-64. | 1.0 | 1 |
| 51 | Information Distance and Its Extensions. Lecture Notes in Computer Science, 2011, , 18-28. | 1.0 | 1 |
| 52 | ChimST: An Efficient Spectral Library Search Tool for Peptide Identification from Chimeric Spectra in Data-Dependent Acquisition. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2021, 18, 1416-1425. | 1.9 | 0 |
| 53 | Invited Talk: Modern Homology Search. , 2007, , 145-145. | | 0 |