

# Gregory Kucherov

## List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/430550/publications.pdf>

Version: 2024-02-01

84  
papers

2,270  
citations

304368

22  
h-index

243296

44  
g-index

95  
all docs

95  
docs citations

95  
times ranked

2790  
citing authors

#	ARTICLE	IF	CITATIONS
1	mreps: efficient and flexible detection of tandem repeats in DNA. Nucleic Acids Research, 2003, 31, 3672-3678.	6.5	333
2	YASS: enhancing the sensitivity of DNA similarity search. Nucleic Acids Research, 2005, 33, W540-W543.	6.5	296
3	NORINE: a database of nonribosomal peptides. Nucleic Acids Research, 2007, 36, D326-D331.	6.5	226
4	Diversity of Monomers in Nonribosomal Peptides: towards the Prediction of Origin and Biological Activity. Journal of Bacteriology, 2010, 192, 5143-5150.	1.0	102
5	Spaced seeds improve $k$ -mer-based metagenomic classification. Bioinformatics, 2015, 31, 3584-3592.	1.8	84
6	Rapid inference of antibiotic resistance and susceptibility by genomic neighbour typing. Nature Microbiology, 2020, 5, 455-464.	5.9	74
7	Using cascading Bloom filters to improve the memory usage for de Bruijn graphs. Algorithms for Molecular Biology, 2014, 9, 2.	0.3	62
8	A UNIFYING FRAMEWORK FOR SEED SENSITIVITY AND ITS APPLICATION TO SUBSET SEEDS. Journal of Bioinformatics and Computational Biology, 2006, 04, 553-569.	0.3	58
9	Optimal Reconstruction of Graphs under the Additive Model. Algorithmica, 2000, 28, 104-124.	1.0	55
10	Reconstructing a Hamiltonian cycle by querying the graph: Application to DNA physical mapping. Discrete Applied Mathematics, 1998, 88, 147-165.	0.5	54
11	Finding approximate repetitions under Hamming distance. Theoretical Computer Science, 2003, 303, 135-156.	0.5	54
12	Multiseed Lossless Filtration. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2005, 2, 51-61.	1.9	50
13	Improved hit criteria for DNA local alignment. BMC Bioinformatics, 2004, 5, 149.	1.2	49
14	Searching for gapped palindromes. Theoretical Computer Science, 2009, 410, 5365-5373.	0.5	48
15	Matching a set of strings with variable length don't cares. Theoretical Computer Science, 1997, 178, 129-154.	0.5	42
16	On maximal repetitions in words. Lecture Notes in Computer Science, 1999, , 374-385.	1.0	42
17	Reconsidering the significance of genomic word frequencies. Trends in Genetics, 2007, 23, 543-546.	2.9	31
18	Using Cascading Bloom Filters to Improve the Memory Usage for de Bruijn Graphs. Lecture Notes in Computer Science, 2013, , 364-376.	1.0	30

#	ARTICLE	IF	CITATIONS
19	Combinatorial Search on Graphs Motivated by Bioinformatics Applications: A Brief Survey. Lecture Notes in Computer Science, 2005, , 16-27.	1.0	29
20	Evolution of biosequence search algorithms: a brief survey. Bioinformatics, 2019, 35, 3547-3552.	1.8	27
21	On repetition-free binary words of minimal density. Theoretical Computer Science, 1999, 218, 161-175.	0.5	25
22	Linear-time computation of local periods. Theoretical Computer Science, 2004, 326, 229-240.	0.5	25
23	Diversity and structure of PIF/Harbinger-like elements in the genome of <i>Medicago truncatula</i> . BMC Genomics, 2007, 8, 409.	1.2	25
24	Simplitigs as an efficient and scalable representation of de Bruijn graphs. Genome Biology, 2021, 22, 96.	3.8	22
25	On Subset Seeds for Protein Alignment. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2009, 6, 483-494.	1.9	20
26	Optimal Linear Arrangement of Interval Graphs. Lecture Notes in Computer Science, 2006, , 267-279.	1.0	20
27	SIGffRid: A tool to search for sigma factor binding sites in bacterial genomes using comparative approach and biologically driven statistics. BMC Bioinformatics, 2008, 9, 73.	1.2	19
28	Minimally overlapping words for sequence similarity search. Bioinformatics, 2021, 36, 5344-5350.	1.8	19
29	Structural pattern matching of nonribosomal peptides. BMC Structural Biology, 2009, 9, 15.	2.3	18
30	Finding repeats with fixed gap. , 0, , .		16
31	Approximate string matching using a bidirectional index. Theoretical Computer Science, 2016, 638, 145-158.	0.5	16
32	Estimating seed sensitivity on homogeneous alignments. , 0, , .		14
33	Absent words in a sliding window with applications. Information and Computation, 2020, 270, 104461.	0.5	14
34	Decidability of Regularity and Related Properties of Ground Normal Form Languages. Information and Computation, 1995, 118, 91-100.	0.5	13
35	Back-translation for discovering distant protein homologies in the presence of frameshift mutations. Algorithms for Molecular Biology, 2010, 5, 6.	0.3	13
36	On relationship between term rewriting systems and regular tree languages. Lecture Notes in Computer Science, 1991, , 299-311.	1.0	13

#	ARTICLE	IF	CITATIONS
37	On maximal repetitions of arbitrary exponent. Information Processing Letters, 2010, 110, 252-256.	0.4	12
38	On the combinatorics of suffix arrays. Information Processing Letters, 2013, 113, 915-920.	0.4	12
39	Cross-document pattern matching. Journal of Discrete Algorithms, 2014, 24, 40-47.	0.7	11
40	Multi-seed Lossless Filtration. Lecture Notes in Computer Science, 2004, , 297-310.	1.0	10
41	Optimal Bounds for Computing $\alpha$ -gapped Repeats. Lecture Notes in Computer Science, 2016, , 245-255.	1.0	9
42	On-line construction of position heaps. Journal of Discrete Algorithms, 2013, 20, 3-11.	0.7	8
43	Prefix Table Construction and Conversion. Lecture Notes in Computer Science, 2013, , 41-53.	1.0	8
44	Searching for Gapped Palindromes. Lecture Notes in Computer Science, 2008, , 18-30.	1.0	8
45	RNF: a general framework to evaluate NGS read mappers. Bioinformatics, 2016, 32, 136-139.	1.8	7
46	Subset Seed Automaton. , 2007, , 180-191.		7
47	Undecidability of ground reducibility for word rewriting systems with variables. Information Processing Letters, 1995, 53, 209-215.	0.4	6
48	Pattern Matching on Sparse Suffix Trees. , 2011, , .		6
49	A new quasi-reducibility testing algorithm and its application to proofs by induction. Lecture Notes in Computer Science, 1988, , 204-213.	1.0	6
50	Minimal Absent Words in a Sliding Window and Applications to On-Line Pattern Matching. Lecture Notes in Computer Science, 2017, , 164-176.	1.0	6
51	Designing Efficient Spaced Seeds for SOLiD Read Mapping. Advances in Bioinformatics, 2010, 2010, 1-12.	5.7	5
52	Regular Language Constrained Sequence Alignment Revisited. Journal of Computational Biology, 2011, 18, 771-781.	0.8	5
53	Decidability of regularity and related properties of ground normal form languages. Lecture Notes in Computer Science, 1993, , 272-286.	1.0	5
54	Cross-Document Pattern Matching. Lecture Notes in Computer Science, 2012, , 196-207.	1.0	5

#	ARTICLE	IF	CITATIONS
55	Space-efficient representation of genomic k-mer count tables. Algorithms for Molecular Biology, 2022, 17, 5.	0.3	5
56	A Unifying Framework for Seed Sensitivity and Its Application to Subset Seeds. Lecture Notes in Computer Science, 2005, , 251-263.	1.0	4
57	Set-Min Sketch: A Probabilistic Map for Power-Law Distributions with Application to $k$ -Mer Annotation. Journal of Computational Biology, 2022, 29, 140-154.	0.8	4
58	Some results on top-context-free tree languages. , 1994, , 157-171.		3
59	Optimal neighborhood indexing for protein similarity search. BMC Bioinformatics, 2008, 9, 534.	1.2	3
60	Improved Filters for the Approximate Suffix-Prefix Overlap Problem. Lecture Notes in Computer Science, 2014, , 139-148.	1.0	3
61	Computing Discriminating and Generic Words. Lecture Notes in Computer Science, 2012, , 307-317.	1.0	3
62	Optimal bounds for computing $\hat{1}$ -gapped repeats. Information and Computation, 2019, 268, 104434.	0.5	2
63	Patterns in Words versus Patterns in Trees: A Brief Survey and New Results. Lecture Notes in Computer Science, 2000, , 283-296.	1.0	2
64	Minimal Discriminating Words Problem Revisited. Lecture Notes in Computer Science, 2013, , 129-140.	1.0	2
65	The complexity of testing ground reducibility for linear word rewriting systems with variables. Lecture Notes in Computer Science, 1995, , 262-275.	1.0	2
66	On Maximal Unbordered Factors. Lecture Notes in Computer Science, 2015, , 343-354.	1.0	2
67	Algorithms for path-constrained sequence alignment. Journal of Discrete Algorithms, 2014, 24, 48-58.	0.7	1
68	Efficient Seeding Techniques for Protein Similarity Search. Communications in Computer and Information Science, 2008, , 466-478.	0.4	1
69	Seed Design Framework for Mapping SOLiD Reads. Lecture Notes in Computer Science, 2010, , 384-396.	1.0	1
70	On-Line Construction of Position Heaps. Lecture Notes in Computer Science, 2011, , 326-337.	1.0	1
71	Full-Fledged Real-Time Indexing for Constant Size Alphabets. Lecture Notes in Computer Science, 2013, , 650-660.	1.0	1
72	Approximate Tandem Repeats. , 2016, , 106-109.		1

#	ARTICLE	IF	CITATIONS
73	Valentin Antimirov (1961–1995). Theoretical Computer Science, 1996, 155, 289-290.	0.5	0
74	Visualization of dynamic automata using Padnon. Lecture Notes in Computer Science, 1998, , 25-28.	1.0	0
75	The complexity of some complementation problems. Information Processing Letters, 1999, 71, 159-165.	0.4	0
76	AN OVERVIEW OF THE ALGORITHMIC SESSION OF MCCMB'05. Journal of Bioinformatics and Computational Biology, 2006, 04, xv-xvi.	0.3	0
77	Efficient alternatives to PSI-BLAST. Bulletin of the Polish Academy of Sciences: Technical Sciences, 2012, 60, 495-505.	0.8	0
78	Full-Fledged Real-Time Indexing for Constant Size Alphabets. Algorithmica, 2017, 79, 387-400.	1.0	0
79	Special Issue on Computer Science Symposium in Russia (2019). Theory of Computing Systems, 2021, 65, 441-443.	0.7	0
80	Information-theoretic problems of DNA-based storage systems. Informatsionno-Upravliaiushchie Sistemy, 2021, , 39-52.	0.3	0
81	Approximate Tandem Repeats. , 2008, , 48-51.		0
82	Back-Translation for Discovering Distant Protein Homologies. Lecture Notes in Computer Science, 2009, , 108-120.	1.0	0
83	CPM's 20th Anniversary: A Statistical Retrospective. Lecture Notes in Computer Science, 2009, , 1-11.	1.0	0
84	Regular Language Constrained Sequence Alignment Revisited. Lecture Notes in Computer Science, 2011, , 404-415.	1.0	0