Heng Li

List of Publications by Year
in descending order

110	192,251	17776 65	29333 108
papers	citations	h-index	g-index

138	138	138	190857
all docs	docs citations	times ranked	citing authors

#	Article	IF	Citations
1	The Sequence Alignment/Map format and SAMtools. Bioinformatics, 2009, 25, 2078-2079.	1.8	49,124
2	Fast and accurate short read alignment with Burrowsâ€“Wheeler transform. Bioinformatics, 2009, 25, 1754-1760.	1.8	43,062
3	A global reference for human genetic variation. Nature, 2015, 526, 68-74.	13.7	13,998
4	Fast and accurate long-read alignment with Burrowsâ€“Wheeler transform. Bioinformatics, 2010, 26, 589-595.	1.8	10,002
5	Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics, 2018, 34, 3094-3100.	1.8	7,764
6	A map of human genome variation from population-scale sequencing. Nature, 2010, 467, 1061-1073.	13.7	7,209
7	An integrated map of genetic variation from 1,092 human genomes. Nature, 2012, 491, 56-65.	13.7	7,199
8	A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics, 2011, 27, 2987-2993.	1.8	5,467
9	Twelve years of SAMtools and BCFtools. GigaScience, 2021, 10, .	3.3	4,546
10	A Draft Sequence of the Neandertal Genome. Science, 2010, 328, 710-722.	6.0	3,588
11	Accurate whole human genome sequencing using reversible terminator chemistry. Nature, 2008, 456, 53-59.	13.7	3,118
12	Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Research, 2008, 18, 1851-1858.	2.4	2,275
13	Inference of human population history from individual whole-genome sequences. Nature, 2011, 475, 493-496.	13.7	2,053
14	The complete genome sequence of a Neanderthal from the Altai Mountains. Nature, 2014, 505, 43-49.	13.7	1,830
15	A High-Coverage Genome Sequence from an Archaic Denisovan Individual. Science, 2012, 338, 222-226.	6.0	1,695
16	Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. Nature Methods, 2021, 18, 170-175.	9.0	1,675
17	Genetic history of an archaic hominin group from Denisova Cave in Siberia. Nature, 2010, 468, 1053-1060.	13.7	1,537
18	The complete sequence of a human genome. Science, 2022, 376, 44-53.	6.0	1,222

#	Article	IF	Citations
19	The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. Nature, 2016, 538, 201-206.	13.7	1,216
20	Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature, 2014, 513, 409-413.	13.7	1,179
21	Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. Bioinformatics, 2016, 32, 2103-2110.	1.8	1,082
22	EnsemblCompara GeneTrees: Complete, duplication-aware phylogenetic trees in vertebrates. Genome Research, 2009, 19, 327-335.	2.4	1,058
23	The sequence and de novo assembly of the giant panda genome. Nature, 2010, 463, 311-317.	13.7	1,058
24	Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. Nature Biotechnology, 2019, 37, 1155-1162.	9.4	1,010
25	A Draft Sequence for the Genome of the Domesticated Silkworm (Bombyx mori). Science, 2004, 306, 1937-1940.	6.0	994
26	Fast and accurate long-read assembly with wtdbg2. Nature Methods, 2020, 17, 155-158.	9.0	917
27	Genome sequence of a 45,000-year-old modern human from western Siberia. Nature, 2014, 514, 445-449.	13.7	856
28	The diploid genome sequence of an Asian individual. Nature, 2008, 456, 60-65.	13.7	834
29	The Genomes of Oryza sativa: A History of Duplications. PLoS Biology, 2005, 3, e38.	2.6	808
30	Toward better understanding of artifacts in variant calling from high-coverage samples. Bioinformatics, 2014, 30, 2843-2851.	1.8	790
31	Great ape genetic diversity and population history. Nature, 2013, 499, 471-475.	13.7	768
32	A survey of sequence alignment algorithms for next-generation sequencing. Briefings in Bioinformatics, 2010, 11, 473-483.	3.2	765
33	Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. Nature Genetics, 2008, 40, 722-729.	9.4	736
34	Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly. Genome Research, 2017, 27, 849-864.	2.4	728
35	Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems. , 2019, , .		671
36	A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. Nature Biotechnology, 2008, 26, 779-785.	9.4	619

#	Article	IF	Citations
37	Tabix: fast retrieval of sequence features from generic TAB-delimited files. Bioinformatics, 2011, 27, 718-719.	1.8	494
38	TreeFam: a curated database of phylogenetic trees of animal gene families. Nucleic Acids Research, 2006, 34, D572-D580.	6.5	465
39	The Date of Interbreeding between Neandertals and Modern Humans. PLoS Genetics, 2012, 8, e1002947.	1.5	402
40	A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. Nature, 2004, 432, 717-722.	13.7	391
41	Three-dimensional genome structures of single diploid human cells. Science, 2018, 361, 924-928.	6.0	347
42	Exploring single-sample SNP and INDEL calling with whole-genome de novo assembly. Bioinformatics, 2012, 28, 1838-1844.	1.8	330
43	Single-cell whole-genome analyses by Linear Amplification via Transposon Insertion (LIANTI). Science, 2017, 356, 189-194.	6.0	303
44	A direct characterization of human mutation based on microsatellites. Nature Genetics, 2012, 44, 1161-1165.	9.4	302
45	New strategies to improve minimap2 alignment accuracy. Bioinformatics, 2021, 37, 4572-4574.	1.8	296
46	TreeFam: 2008 Update. Nucleic Acids Research, 2007, 36, D735-D740.	6.5	294
47	Improving SNP discovery by base alignment quality. Bioinformatics, 2011, 27, 1157-1158.	1.8	275
48	Complete Genomes Reveal Signatures of Demographic and Genetic Declines in the Woolly Mammoth. Current Biology, 2015, 25, 1395-1400.	1.8	263
49	Extremely low-coverage sequencing and imputation increases power for genome-wide association studies. Nature Genetics, 2012, 44, 631-635.	9.4	239
50	The design and construction of reference pangenome graphs with minigraph. Genome Biology, 2020, 21, 265.	3.8	195
51	The Human Pangenome Project: a global resource to map genomic diversity. Nature, 2022, 604, 437-446.	13.7	192
52	HTSlib: C library for reading/writing high-throughput sequencing data. GigaScience, 2021, 10, .	3.3	191
53	No evidence that selection has been less effective at removing deleterious mutations in Europeans than in Africans. Nature Genetics, 2015, 47, 126-131.	9.4	182
54	BFC: correcting Illumina sequencing errors. Bioinformatics, 2015, 31, 2885-2887.	1.8	173

#	Article	IF	Citations
55	Functional equivalence of genome sequencing analysis pipelines enables harmonized variant calling across human genetics projects. Nature Communications, 2018, 9, 4038.	5.8	166
56	pIRS: Profile-based Illumina pair-end reads simulator. Bioinformatics, 2012, 28, 1533-1535.	1.8	163
57	Genes controlling seed dormancy and pre-harvest sprouting in a rice-wheat-barley comparison. Functional and Integrative Genomics, 2004, 4, 84-93.	1.4	157
58	A synthetic-diploid benchmark for accurate variant-calling evaluation. Nature Methods, 2018, 15, 595-597.	9.0	154
59	Targeting a cytokine checkpoint enhances the fitness of armored cord blood CAR-NK cells. Blood, 2021, 137, 624-636.	0.6	147
60	The contribution of rare variation to prostate cancer heritability. Nature Genetics, 2016, 48, 30-35.	9.4	139
61	Haplotype-resolved assembly of diploid genomes without parental data. Nature Biotechnology, 2022, 40, 1332-1335.	9.4	139
62	Neutral evolution of â€˜non-codingâ€™ complementary DNAs. Nature, 2004, 431, 1-2.	13.7	127
63	SOAPindel: Efficient identification of indels from short paired reads. Genome Research, 2013, 23, 195-200.	2.4	115
64	Chromosome-scale, haplotype-resolved assembly of human genomes. Nature Biotechnology, 2021, 39, 309-312.	9.4	109
65	A cohort autopsy study defines COVID-19 systemic pathogenesis. Cell Research, 2021, 31, 836-846.	5.7	93
66	FermiKit: assembly-based variant calling for Illumina resequencing data. Bioinformatics, 2015, 31, 3694-3696.	1.8	92
67	Curated variation benchmarks for challenging medically relevant autosomal genes. Nature Biotechnology, 2022, 40, 672-680.	9.4	90
68	Extreme selective sweeps independently targeted the X chromosomes of the great apes. Proceedings of the National Academy of Sciences of the United States of America, 2015, 112, 6413-6418.	3.3	75
69	Pyroptotic macrophages stimulate the SARS-CoV-2-associated cytokine storm. Cellular and Molecular Immunology, 2021, 18, 1305-1307.	4.8	74
70	Differential DNA methylation of vocal and facial anatomy genes in modern humans. Nature Communications, 2020, 11, 1189.	5.8	69
71	Porcine transcriptome analysis based on 97 non-normalized cDNA libraries and assembly of 1,021,891 expressed sequence tags. Genome Biology, 2007, 8, R45.	13.9	67
72	An Accurate and Comprehensive Clinical Sequencing Assay for Cancer Targeted and Immunotherapies. Oncologist, 2019, 24, e1294-e1302.	1.9	67

#	Article	IF	Citations
73	Using population admixture to help complete maps of the human genome. Nature Genetics, 2013, 45, 406-414.	9.4	61
74	Fast construction of FM-index for long sequence reads. Bioinformatics, 2014, 30, 3274-3275.	1.8	56
75	Metagenome assembly of high-fidelity long reads with hifiasm-meta. Nature Methods, 2022, 19, 671-674.	9.0	56
76	Pathological changes in the lungs and lymphatic organs of 12 COVID-19 autopsy cases. National Science Review, 2020, 7, 1868-1878.	4.6	52
77	BGT: efficient and flexible genotype query across many samples. Bioinformatics, 2016, 32, 590-592.	1.8	46
78	Comprehensive identification of transposable element insertions using multiple sequencing technologies. Nature Communications, 2021, 12, 3836.	5.8	44
79	Accurate SNV detection in single cells by transposon-based whole-genome amplification of complementary strands. Proceedings of the National Academy of Sciences of the United States of America, 2021, 118, .	3.3	41
80	Real-time mapping of nanopore raw signals. Bioinformatics, 2021, 37, i477-i483.	1.8	41
81	Detecting SNPs and estimating allele frequencies in clonal bacterial populations by sequencing pooled DNA. Bioinformatics, 2009, 25, 2074-2075.	1.8	40
82	Fast alignment and preprocessing of chromatin profiles with Chromap. Nature Communications, 2021, 12, 6566.	5.8	39
83	Snap: an integrated SNP annotation platform. Nucleic Acids Research, 2007, 35, D707-D710.	6.5	36
84	Deep short-read sequencing of chromosome 17 from the mouse strains A/J and CAST/Ei identifies significant germline variation and candidate genes that regulate liver triglyceride levels. Genome Biology, 2009, 10, R112.	13.9	36
85	Mapping the Human Reference Genomeâ€™s Missing Sequence by Three-Way Admixture in Latino Genomes. American Journal of Human Genetics, 2013, 93, 411-421.	2.6	36
86	Worldwide genetic variation of the IGHV and TRBV immune receptor gene families in humans. Life Science Alliance, 2019, 2, e201800221.	1.3	33
87	Evaluating the contribution of rare variants to type 2 diabetes and related traits using pedigrees. Proceedings of the National Academy of Sciences of the United States of America, 2018, 115, 379-384.	3.3	28
88	Omicsâ€based profiling of carcinoma of the breast and matched regional lymph node metastasis. Proteomics, 2008, 8, 5038-5052.	1.3	26
89	The anatomy of successful computational biology software. Nature Biotechnology, 2013, 31, 894-897.	9.4	25
90	htsget: a protocol for securely streaming genomic data. Bioinformatics, 2019, 35, 119-121.	1.8	23

#	Article	IF	Citations
91	Identifying centromeric satellites with dna-brnn. Bioinformatics, 2019, 35, 4408-4410.	1.8	22
92	A haplotype-awarede novoassembly of related individuals using pedigree sequence graph. Bioinformatics, 2020, 36, 2385-2392.	1.8	22
93	CRISPAltRations: A validated cloud-based approach for interrogation of double-strand break repair mediated by CRISPR genome editing. Molecular Therapy - Methods and Clinical Development, 2021, 21, 478-491.	1.8	18
94	Haplotype frequencies in a sub-region of chromosome 19q13.3, related to risk and prognosis of cancer, differ dramatically between ethnic groups. BMC Medical Genetics, 2009, 10, 20.	2.1	16
95	Postmortem high-dimensional immune profiling of severe COVID-19 patients reveals distinct patterns of immunosuppression and immunoactivation. Nature Communications, 2022, 13, 269.	5.8	16
96	Test Data Sets and Evaluation of Gene Prediction Programs on the Rice Genome. Journal of Computer Science and Technology, 2005, 20, 446-453.	0.9	14
97	Higher Rates of Processed Pseudogene Acquisition in Humans and Three Great Apes Revealed by Long-Read Assemblies. Molecular Biology and Evolution, 2021, 38, 2958-2966.	3.5	13
98	Pseudo-Sanger sequencing: massively parallel production of long and near error-free reads using NGS technology. BMC Genomics, 2013, 14, 711.	1.2	12
99	The distinct clinicopathological and prognostic implications of PIK3CA mutations in breast cancer patients from Central China . Cancer Management and Research, 2019, Volume 11, 1473-1492.	0.9	10
100	Comprehensive Characterizations of Immune Receptor Repertoire in Tumors and Cancer Immunotherapy Studies. Cancer Immunology Research, 2022, 10, 788-799.	1.6	10
101	PigGIS: Pig Genomic Informatics System. Nucleic Acids Research, 2007, 35, D654-D657.	6.5	9
102	Evolutionary Transients in the Rice Transcriptome. Genomics, Proteomics and Bioinformatics, 2010, 8, 211-228.	3.0	9
103	Bedtk: finding interval overlap with implicit interval tree. Bioinformatics, 2021, 37, 1315-1316.	1.8	5
104	Haplotype-resolved germline and somatic alterations in renal medullary carcinomas. Genome Medicine, 2021, 13, 114.	3.6	5
105	CoLoRd: compressing long reads. Nature Methods, 2022, 19, 441-444.	9.0	5
106	A cross-species alignment tool (CAT). BMC Bioinformatics, 2007, 8, 349.	1.2	4
107	Vasculogenic Mimicry Formation Predicts Tumor Progression in Oligodendroglioma. Pathology and Oncology Research, 2021, 27, 1609844.	0.9	2
108	Aberrant expression of thyroid transcription factor-1 in meningeal solitary fibrous tumor/hemangiopericytoma. Brain Tumor Pathology, 2021, 38, 122-131.	1.1	0

#	Article	IF	Citations
109	The Prognostic Value of Retraction Clefts in Chinese Invasive Breast Cancer Patients. Pathology and Oncology Research, 2021, 27, 1609743.	0.9	0
110	BCOR-CCNB3 sarcoma with concurrent RNF213-SLC26A11 gene fusion: a rare sarcoma with altered histopathological features after chemotherapy. World Journal of Surgical Oncology, 2022, 20, 156.	0.8	0