Tim Hubbard

List of Publications by Year
in descending order

170 papers	120,089 citations	5430 85 h-index	5739 167 g-index
180 all docs	180 docs citations	180 times ranked	122157 citing authors

#	Article	IF	Citations
1	Pharmacogenomic testing in paediatrics: Clinical implementation strategies. British Journal of Clinical Pharmacology, 2022, 88, 4297-4310.	1.1	12
2	Whole genome sequencing for the diagnosis of neurological repeat expansion disorders in the UK: a retrospective diagnostic accuracy and prospective clinical validation study. Lancet Neurology, The, 2022, 21, 234-245.	4.9	74
3	Transcriptional activity and strain-specific history of mouse pseudogenes. Nature Communications, 2020, 11, 3695.	5.8	17
4	Perspectives on ENCODE. Nature, 2020, 583, 693-698.	13.7	123
5	Bi-allelic Loss-of-Function CACNA1B Mutations in Progressive Epilepsy-Dyskinesia. American Journal of Human Genetics, 2019, 104, 948-956.	2.6	45
6	GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Research, 2019, 47, D766-D773.	6.5	2,350
7	The 100â€‰000 Genomes Project: bringing whole genome sequencing to the NHS. BMJ: British Medical Journal, 2018, 361, k1687.	2.4	312
8	Scientists on the Spot: Sequencing the human genome to influence patient healthcare. Cardiovascular Research, 2018, 114, e66-e67.	1.8	0
9	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
10	Automated PDF highlighting to support faster curation of literature for Parkinsonâ€™s and Alzheimerâ€™s disease. Database: the Journal of Biological Databases and Curation, 2017, 2017, .	1.4	1
11	Analysis of diagnoses extracted from electronic health records in a large mental health case register. PLoS ONE, 2017, 12, e0171526.	1.1	9
12	Making sense of big data in health research: Towards an EU action plan. Genome Medicine, 2016, 8, 71.	3.6	190
13	The language of mental health problems in social media. , 2016, , .		60
14	Global implementation of genomic medicine: We are not alone. Science Translational Medicine, 2015, 7, 290ps13.	5.8	146
15	An interactive genome browser of association results from the UK10K cohorts project. Bioinformatics, 2015, 31, 4029-4031.	1.8	12
16	The UK10K project identifies rare variants in health and disease. Nature, 2015, 526, 82-90.	13.7	1,014
17	Ensembl 2014. Nucleic Acids Research, 2014, 42, D749-D755.	6.5	1,211
18	Current status and new features of the Consensus Coding Sequence database. Nucleic Acids Research, 2014, 42, D865-D872.	6.5	140

#	Article	IF	Citations
19	Defining functional DNA elements in the human genome. Proceedings of the National Academy of Sciences of the United States of America, 2014, 111, 6131-6138.	3.3	635
20	Comparative analysis of pseudogenes across three phyla. Proceedings of the National Academy of Sciences of the United States of America, 2014, 111, 13361-13366.	3.3	72
21	Comparative analysis of the transcriptome across distant species. Nature, 2014, 512, 445-448.	13.7	289
22	Reply to Brunet and Doolittle: Both selected effect and causal role elements can influence human biology and disease. Proceedings of the National Academy of Sciences of the United States of America, 2014, 111, E3366.	3.3	25
23	Characterizing genetic variants for clinical action. American Journal of Medical Genetics, Part C: Seminars in Medical Genetics, 2014, 166, 93-104.	0.7	50
24	A Comparison of Peak Callers Used for DNase-Seq Data. PLoS ONE, 2014, 9, e96303.	1.1	71
25	Assessment of transcript reconstruction methods for RNA-seq. Nature Methods, 2013, 10, 1177-1184.	9.0	679
26	Systematic evaluation of spliced alignment programs for RNA-seq data. Nature Methods, 2013, 10, 1185-1191.	9.0	467
27	The zebrafish reference genome sequence and its relationship to the human genome. Nature, 2013, 496, 498-503.	13.7	3,708
28	Chromatin Accessibility Data Sets Show Bias Due to Sequence Specificity of the DNase I Enzyme. PLoS ONE, 2013, 8, e69853.	1.1	61
29	Combining RT-PCR-seq and RNA-seq to catalog all genic elements encoded in the human genome. Genome Research, 2012, 22, 1698-1710.	2.4	50
30	The GENCODE pseudogene resource. Genome Biology, 2012, 13, R51.	13.9	273
31	An integrated encyclopedia of DNA elements in the human genome. Nature, 2012, 489, 57-74.	13.7	15,516
32	GENCODE: The reference human genome annotation for The ENCODE Project. Genome Research, 2012, 22, 1760-1774.	2.4	4,217
33	Ensembl 2012. Nucleic Acids Research, 2012, 40, D84-D90.	6.5	840
34	The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. Genome Research, 2012, 22, 1775-1789.	2.4	4,428
35	Landscape of transcription in human cells. Nature, 2012, 489, 101-108.	13.7	4,484
36	Evidence for Transcript Networks Composed of Chimeric RNAs in Human Cells. PLoS ONE, 2012, 7, e28213.	1.1	61

#	Article	IF	Citations
37	Developing and implementing an institute-wide data sharing policy. Genome Medicine, 2011, 3, 60.	3.6	29
38	A User's Guide to the Encyclopedia of DNA Elements (ENCODE). PLoS Biology, 2011, 9, e1001046.	2.6	1,257
39	The GENCODE exome: sequencing the complete human exome. European Journal of Human Genetics, 2011, 19, 827-831.	1.4	58
40	ITFoM â€“ The IT Future of Medicine. Procedia Computer Science, 2011, 7, 26-29.	1.2	17
41	Dalliance: interactive genome viewing on the web. Bioinformatics, 2011, 27, 889-890.	1.8	99
42	Ensembl 2011. Nucleic Acids Research, 2011, 39, D800-D806.	6.5	630
43	The Origins, Evolution, and Functional Potential of Alternative Splicing in Vertebrates. Molecular Biology and Evolution, 2011, 28, 2949-2959.	3.5	74
44	Modernizing Reference Genome Assemblies. PLoS Biology, 2011, 9, e1001091.	2.6	458
45	Shotgun proteomics aids discovery of novel protein-coding genes, alternative splicing, and â€œresurrectedâ€•pseudogenes in the mouse genome. Genome Research, 2011, 21, 756-767.	2.4	113
46	Metamotifs - a generative model for building families of nucleotide position weight matrices. BMC Bioinformatics, 2010, 11, 348.	1.2	6
47	AnnoTrack - a tracking system for genome annotation. BMC Genomics, 2010, 11, 538.	1.2	9
48	International network of cancer genome projects. Nature, 2010, 464, 993-998.	13.7	2,114
49	iMotifs: an integrated sequence motif visualization and analysis environment. Bioinformatics, 2010, 26, 843-844.	1.8	11
50	Ensembl's 10th year. Nucleic Acids Research, 2010, 38, D557-D562.	6.5	251
51	Novel Candidate Cancer Genes Identified by a Large-Scale Cross-Species Comparative Oncogenomics Approach. Cancer Research, 2010, 70, 883-895.	0.4	40
52	Genome-wide end-sequenced BAC resources for the NOD/MrkTacâ˜† and NOD/ShiLtJâ˜†â˜† mouse genomes. Genomics, 2010, 95, 105-110.	1.3	14
53	From identification to validation to gene count. Genome Biology, 2010, 11, .	3.8	1
54	The consensus coding sequence (CCDS) project: Identifying a common protein-coding gene set for the human and mouse genomes. Genome Research, 2009, 19, 1316-1323.	2.4	476

#	Article	IF	Citations
55	Ensembl 2009. Nucleic Acids Research, 2009, 37, D690-D697.	6.5	721
56	Discovery of Candidate Disease Genes in ENUâ€“Induced Mouse Mutants by Large-Scale Sequencing, Including a Splice-Site Mutation in Nucleoredoxin. PLoS Genetics, 2009, 5, e1000759.	1.5	39
57	Petabyte-scale innovations at the European Nucleotide Archive. Nucleic Acids Research, 2009, 37, D19-D25.	6.5	82
58	Cancer gene discovery in mouse and man. Biochimica Et Biophysica Acta: Reviews on Cancer, 2009, 1796, 140-161.	3.3	13
59	Prepublication data sharing. Nature, 2009, 461, 168-170.	13.7	243
60	Post-publication sharing of data and tools. Nature, 2009, 461, 171-173.	13.7	142
61	Accurate and Sensitive Peptide Identification with Mascot Percolator. Journal of Proteome Research, 2009, 8, 3176-3181.	1.8	399
62	A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. Nature Biotechnology, 2008, 26, 779-785.	9.4	619
63	NestedMICA as an ab initio protein motif discovery tool. BMC Bioinformatics, 2008, 9, 19.	1.2	32
64	Integrating biological data â€“ the Distributed Annotation System. BMC Bioinformatics, 2008, 9, S3.	1.2	87
65	Large-Scale Mutagenesis in p19ARF- and p53-Deficient Mice Identifies Cancer Genes and Their Collaborative Networks. Cell, 2008, 133, 727-741.	13.5	167
66	The Protein Feature Ontology: a tool for the unification of protein feature annotations. Bioinformatics, 2008, 24, 2767-2772.	1.8	19
67	Comparison of Mascot and X!Tandem Performance for Low and High Accuracy Mass Spectrometry and the Development of an Adjusted Mascot Threshold. Molecular and Cellular Proteomics, 2008, 7, 962-970.	2.5	58
68	An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs). Genome Research, 2008, 18, 1518-1529.	2.4	350
69	SISYPHUSâ€”structural alignments for proteins with non-trivial relationships. Nucleic Acids Research, 2007, 35, D253-D259.	6.5	74
70	Large-Scale Discovery of Promoter Motifs in Drosophila melanogaster. PLoS Computational Biology, 2007, 3, e7.	1.5	64
71	The vertebrate genome annotation (Vega) database. Nucleic Acids Research, 2007, 36, D753-D760.	6.5	260
72	Prominent use of distal 5' transcription start sites and discovery of a large number of additional exons in ENCODE regions. Genome Research, 2007, 17, 746-759.	2.4	173

#	Article	IF	Citations
73	Data growth and its impact on the SCOP database: new developments. Nucleic Acids Research, 2007, 36, D419-D425.	6.5	854
74	Ensembl 2007. Nucleic Acids Research, 2007, 35, D610-D617.	6.5	699
75	Priorities for nucleotide trace, sequence and annotation data capture at the Ensembl Trace Archive and the EMBL Nucleotide Sequence Database. Nucleic Acids Research, 2007, 36, D5-D12.	6.5	46
76	Ensembl 2008. Nucleic Acids Research, 2007, 36, D707-D714.	6.5	440
77	Lessons learned from the initial sequencing of the pig genome: comparative analysis of an 8 Mb region of pig chromosome 17. Genome Biology, 2007, 8, R168.	13.9	44
78	New tools and expanded data analysis capabilities at the protein structure prediction center. Proteins: Structure, Function and Bioinformatics, 2007, 69, 19-26.	1.5	31
79	Critical assessment of methods of protein structure predictionâ€”Round VII. Proteins: Structure, Function and Bioinformatics, 2007, 69, 3-9.	1.5	199
80	Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature, 2007, 447, 799-816.	13.7	4,709
81	Integrating sequence and structural biology with DAS. BMC Bioinformatics, 2007, 8, 333.	1.2	68
82	GENCODE: producing a reference annotation for ENCODE. Genome Biology, 2006, 7, S4.	13.9	533
83	EGASP: the human ENCODE Genome Annotation Assessment Project. Genome Biology, 2006, 7, S2.	13.9	228
84	DNA sequence of human chromosome 17 and analysis of rearrangement in the human lineage. Nature, 2006, 440, 1045-1049.	13.7	130
85	The DNA sequence and biological annotation of human chromosomeâ€‰1. Nature, 2006, 441, 315-321.	13.7	211
86	A machine learning strategy to identify candidate binding sites in human protein-coding sequence. BMC Bioinformatics, 2006, 7, 419.	1.2	12
87	Ensembl 2006. Nucleic Acids Research, 2006, 34, D556-D561.	6.5	331
88	Genomic anatomy of the Tyrp1 (brown) deletion complex. Proceedings of the National Academy of Sciences of the United States of America, 2006, 103, 3704-3709.	3.3	30
89	The DNA sequence of the human X chromosome. Nature, 2005, 434, 325-337.	13.7	985
90	Critical assessment of methods of protein structure prediction (CASP)â€”Round 6. Proteins: Structure, Function and Bioinformatics, 2005, 61, 3-7.	1.5	162

#	Article	IF	Citations
91	Adding Some SPICE to DAS. Bioinformatics, 2005, 21, ii40-ii41.	1.8	42
92	NestedMICA: sensitive inference of over-represented motifs in nucleic acid sequence. Nucleic Acids Research, 2005, 33, 1445-1453.	6.5	100
93	The Vertebrate Genome Annotation (Vega) database. Nucleic Acids Research, 2004, 33, D459-D465.	6.5	125
94	Ensembl 2005. Nucleic Acids Research, 2004, 33, D447-D453.	6.5	368
95	An Overview of Ensembl. Genome Research, 2004, 14, 925-928.	2.4	391
96	Ensembl 2004. Nucleic Acids Research, 2004, 32, 468D-470.	6.5	146
97	A New Trade Framework for Global Healthcare R&D. PLoS Biology, 2004, 2, e52.	2.6	87
98	A census of human cancer genes. Nature Reviews Cancer, 2004, 4, 177-183.	12.8	2,868
99	The DNA sequence and analysis of human chromosome 13. Nature, 2004, 428, 522-528.	13.7	91
100	The DNA sequence and comparative analysis of human chromosome 10. Nature, 2004, 429, 375-381.	13.7	74
101	DNA sequence and analysis of human chromosome 9. Nature, 2004, 429, 369-374.	13.7	314
102	Finishing the euchromatic sequence of the human genome. Nature, 2004, 431, 931-945.	13.7	4,232
103	What can we learn from noncoding regions of similarity between genomes?. BMC Bioinformatics, 2004, 5, 131.	1.2	8
104	SCOP database in 2004: refinements integrate structure and sequence family data. Nucleic Acids Research, 2004, 32, 226D-229.	6.5	815
105	Domain Insertions in Protein Structures. Journal of Molecular Biology, 2004, 338, 633-641.	2.0	72
106	The ENCODE (ENCyclopedia Of DNA Elements) Project. Science, 2004, 306, 636-640.	6.0	2,121
107	CASP5 target classification. Proteins: Structure, Function and Bioinformatics, 2003, 53, 340-351.	1.5	20
108	Critical assessment of methods of protein structure prediction (CASP)-round V. Proteins: Structure, Function and Bioinformatics, 2003, 53, 334-339.	1.5	221

#	Article	IF	Citations
109	The DNA sequence and analysis of human chromosome 6. Nature, 2003, 425, 805-811.	13.7	300
110	ddbRNA: detection of conserved secondary structures in multiple alignments. Bioinformatics, 2003, 19, 1606-1611.	1.8	53
111	Ensembl 2002: accommodating comparative genomics. Nucleic Acids Research, 2003, 31, 38-42.	6.5	216
112	Computational Detection and Location of Transcription Start Sites in Mammalian Genomic DNA. Genome Research, 2002, 12, 458-461.	2.4	238
113	DATABASES ANDTOOLS FORBROWSINGGENOMES. Annual Review of Genomics and Human Genetics, 2002, 3, 293-310.	2.5	23
114	SCOP database in 2002: refinements accommodate structural genomics. Nucleic Acids Research, 2002, 30, 264-267.	6.5	403
115	MaxBench: evaluation of sequence and structure comparison methods. Bioinformatics, 2002, 18, 494-495.	1.8	12
116	Biological information: making it accessible and integrated (and trying to make sense of it). Bioinformatics, 2002, 18, S140-S140.	1.8	16
117	The Ensembl genome database project. Nucleic Acids Research, 2002, 30, 38-41.	6.5	1,411
118	The Significance of Performance Ranking in CASPâ€"Response to Marti-Renom et al.. Structure, 2002, 10, 291-292.	1.6	5
119	A physical map of the mouse genome. Nature, 2002, 418, 743-750.	13.7	316
120	Initial sequencing and comparative analysis of the mouse genome. Nature, 2002, 420, 520-562.	13.7	6,319
121	Critical assessment of methods of protein structure prediction (CASP): Round IV. Proteins: Structure, Function and Bioinformatics, 2001, 45, 2-7.	1.5	146
122	Prediction targets of CASP4. Proteins: Structure, Function and Bioinformatics, 2001, 45, 8-12.	1.5	9
123	Assessment of novel fold targets in CASP4: Predictions of three-dimensional structures, secondary structures, and interresidue contacts. Proteins: Structure, Function and Bioinformatics, 2001, 45, 98-118.	1.5	76
124	Peter Andrew Kollman. Proteins: Structure, Function and Bioinformatics, 2001, 45, 2-3.	1.5	42
125	The DNA sequence and comparative analysis of human chromosome 20. Nature, 2001, 414, 865-871.	13.7	287
126	Mining the draft human genome. Nature, 2001, 409, 827-828.	13.7	58

#	Article	IF	Citations
127	Initial sequencing and analysis of the human genome. Nature, 2001, 409, 860-921.	13.7	21,074
128	Quality control in databanks for molecular biology. BioEssays, 2000, 22, 1024-1034.	1.2	16
129	Open annotation offers a democratic solution to genome sequencing. Nature, 2000, 403, 825-825.	13.7	31
130	A browser for expression data. Bioinformatics, 2000, 16, 402-403.	1.8	2
131	SCOP: a Structural Classification of Proteins database. Nucleic Acids Research, 2000, 28, 257-259.	6.5	533
132	BioJava. ACM SIGBIO Newsletter, 2000, 20, 10-12.	0.1	33
133	SCOP: a Structural Classification of Proteins database. Nucleic Acids Research, 1999, 27, 254-256.	6.5	208
134	The DNA sequence of human chromosome 22. Nature, 1999, 402, 489-495.	13.7	1,086
135	RMS/Coverage graphs: A qualitative method for comparing three-dimensional protein structure predictions. Proteins: Structure, Function and Bioinformatics, 1999, 37, 15-21.	1.5	46
136	Critical assessment of methods of protein structure prediction (CASP): Round III. Proteins: Structure, Function and Bioinformatics, 1999, 37, 2-6.	1.5	137
137	Critical assessment of methods of protein structure prediction (CASP): Round III. Proteins: Structure, Function and Bioinformatics, 1999, 37, 2-6.	1.5	27
138	Analysis and assessment of ab initio three-dimensional prediction, secondary structure, and contacts prediction. Proteins: Structure, Function and Bioinformatics, 1999, Suppl 3, 149-70.	1.5	50
139	Critical assessment of methods of protein structure prediction (CASP): Round III. , 1999, 37, 2.		103
140	RMS/coverage graphs: a qualitative method for comparing three-dimensional protein structure predictions. Proteins: Structure, Function and Bioinformatics, 1999, Suppl 3, 15-21.	1.5	11
141	GLASS: A tool to visualize protein structure prediction data in three dimensions and evaluate their consistency. , 1998, 30, 339-351.		3
142	SCOP, Structural Classification of Proteins Database: Applications to Evaluation of the Effectiveness of Sequence Alignment Methods and Statistics of Protein Structural Data. Acta Crystallographica Section D: Biological Crystallography, 1998, 54, 1147-1154.	2.5	31
143	Sequence comparisons using multiple sequences detect three times as many remote homologues as pairwise methods. Journal of Molecular Biology, 1998, 284, 1201-1210.	2.0	498
144	SPEM: a parser for EMBL style flat file database entries. Bioinformatics, 1998, 14, 823-824.	1.8	2

#	Article	IF	Citations
145	Using neural networks for prediction of the subcellular location of proteins. Nucleic Acids Research, 1998, 26, 2230-2236.	6.5	517
146	Assessing sequence comparison methods with reliable structurally identified distant evolutionary relationships. Proceedings of the National Academy of Sciences of the United States of America, 1998, 95, 6073-6078.	3.3	493
147	SCOP: a Structural Classification of Proteins database. Nucleic Acids Research, 1997, 25, 236-239.	6.5	481
148	PROTEIN FOLDS IN THE ALL-Î² AND ALL-Î± CLASSES. Annual Review of Biophysics and Biomolecular Structure, 1997, 26, 597-627.	18.3	104
149	Intermediate sequences increase the detection of homology between sequences. Journal of Molecular Biology, 1997, 273, 349-354.	2.0	217
150	New horizons in sequence analysis. Current Opinion in Structural Biology, 1997, 7, 190-193.	2.6	6
151	Population statistics of protein structures: lessons from structural classifications. Current Opinion in Structural Biology, 1997, 7, 369-376.	2.6	173
152	The Solution Structure of the S1 RNA Binding Domain: A Member of an Ancient Nucleic Acidâ€“Binding Fold. Cell, 1997, 88, 235-242.	13.5	391
153	Critical assessment of methods of protein structure prediction (CASP): Round II. , 1997, 29, 2-6.		114
154	Numerical criteria for the evaluation of ab initio predictions of protein structure. Proteins: Structure, Function and Bioinformatics, 1997, 29, 140-150.	1.5	2
155	Critical assessment of methods of protein structure prediction (CASP): Round II. Proteins: Structure, Function and Bioinformatics, 1997, 29, 2-6.	1.5	22
156	Critical assessment of methods of protein structure prediction (CASP): Round II. , 1997, 29, 2.		82
157	Numerical criteria for the evaluation of ab initio predictions of protein structure. Proteins: Structure, Function and Bioinformatics, 1997, Suppl 1, 140-50.	1.5	12
158	[37] Understanding protein structure: Using scop for fold interpretation. Methods in Enzymology, 1996, 266, 635-643.	0.4	80
159	Protein structure prediction:playing the fold. Trends in Biochemical Sciences, 1996, 21, 279-281.	3.7	1
160	Protein structure prediction: playing the fold. Trends in Biochemical Sciences, 1996, 21, 279-281.	3.7	5
161	Prediction of the structure of GroES and its interaction with GroEL. Proteins: Structure, Function and Bioinformatics, 1995, 22, 199-209.	1.5	14
162	Fold recognition and ab initio structure predictions using hidden markov models and Î²-strand pair potentials. Proteins: Structure, Function and Bioinformatics, 1995, 23, 398-402.	1.5	73

#	Article	IF	Citations
163	Gene duplications in H. influenzae. Nature, 1995, 378, 140-140.	13.7	82
164	SCOP: A structural classification of proteins database for the investigation of sequences and structures. Journal of Molecular Biology, 1995, 247, 536-540.	2.0	3,952
165	SCOP: a structural classification of proteins database for the investigation of sequences and structures. Journal of Molecular Biology, 1995, 247, 536-540.	2.0	5,400
166	Protein design on computers. Five new proteins: Shpilka, grendel, fingerclasp, leather, and aida. Proteins: Structure, Function and Bioinformatics, 1992, 12, 105-110.	1.5	26
167	The role of heat-shock and chaperone proteins in protein folding: possible molecular mechanisms. Protein Engineering, Design and Selection, 1991, 4, 711-717.	1.0	78
168	Knowledge-based protein modelling and design. FEBS Journal, 1988, 172, 513-520.	0.2	236
169	Comparison of solvent-inaccessible cores of homologous proteins: definitions useful for protein modelling. Protein Engineering, Design and Selection, 1987, 1, 159-171.	1.0	185
170	Heat-shock proteins during growth and sporulation ofBacillus subtilis. FEBS Letters, 1985, 188, 209-214.	1.3	34