Kim-Anh Le Cao

List of Publications by Year
in descending order

112 papers	9,472 citations	94433 37 h-index	46799 89 g-index
135 all docs	135 docs citations	135 times ranked	15846 citing authors

#	Article	IF	Citations
1	mixOmics: An R package for â€˜omics feature selection and multiple data integration. PLoS Computational Biology, 2017, 13, e1005752.	3.2	2,279
2	Sparse PLS discriminant analysis: biologically relevant feature selection and graphical displays for multiclass problems. BMC Bioinformatics, 2011, 12, 253.	2.6	699
3	DIABLO: an integrative approach for identifying key molecular drivers from multi-omics assays. Bioinformatics, 2019, 35, 3055-3062.	4.1	496
4	integrOmics: an R package to unravel relationships between two omics datasets. Bioinformatics, 2009, 25, 2855-2856.	4.1	401
5	A Sparse PLS for Variable Selection when Integrating Omics Data. Statistical Applications in Genetics and Molecular Biology, 2008, 7, Article 35.	0.6	398
6	Conservation and divergence in Toll-like receptor 4-regulated gene expression in primary human versus mouse macrophages. Proceedings of the National Academy of Sciences of the United States of America, 2012, 109, E944-53.	7.1	332
7	Citrullinated peptide dendritic cell immunotherapy in HLA risk genotypeâ€"positive rheumatoid arthritis patients. Science Translational Medicine, 2015, 7, 290ra87.	12.4	302
8	Visualising associations between paired â€˜omicsâ€™ data sets. BioData Mining, 2012, 5, 19.	4.0	261
9	Benchmarking single cell RNA-sequencing analysis pipelines using mixture control experiments. Nature Methods, 2019, 16, 479-487.	19.0	259
10	Sparse canonical methods for biological data integration: application to a cross-platform study. BMC Bioinformatics, 2009, 10, 34.	2.6	219
11	Genome-wide characterization of the routes to pluripotency. Nature, 2014, 516, 198-206.	27.8	187
12	Variable selection for generalized canonical correlation analysis. Biostatistics, 2014, 15, 569-583.	1.5	168
13	Quantitative gene profiling of long noncoding RNAs with targeted RNA sequencing. Nature Methods, 2015, 12, 339-342.	19.0	155
14	Dynamic molecular changes during the first week of human life follow a robust developmental trajectory. Nature Communications, 2019, 10, 1092.	12.8	151
15	Malignant transformation of oral epithelial dysplasia: a real-world evaluation of histopathologic grading. Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology, 2014, 117, 343-352.	0.4	148
16	MixMC: A Multivariate Statistical Framework to Gain Insight into Microbial Communities. PLoS ONE, 2016, 11, e0160169.	2.5	139
17	A novel approach for biomarker selection and the integration of repeated measures experiments from two assays. BMC Bioinformatics, 2012, 13, 325.	2.6	129
18	Microbiome profiling reveals gut dysbiosis in a transgenic mouse model of Huntington's disease. Neurobiology of Disease, 2020, 135, 104268.	4.4	118

#	Article	IF	Citations
19	HLA Alleles Associated With Risk of Ankylosing Spondylitis and Rheumatoid Arthritis Influence the Gut Microbiome. Arthritis and Rheumatology, 2019, 71, 1642-1650.	5.6	116
20	Independent Principal Component Analysis for biologically meaningful dimension reduction of large biological data sets. BMC Bioinformatics, 2012, 13, 24.	2.6	111
21	Intestinal Metaproteomics Reveals Host-Microbiota Interactions in Subjects at Risk for Type 1 Diabetes. Diabetes Care, 2018, 41, 2178-2186.	8.6	105
22	Human hepatocellular carcinomas with a periportal phenotype have the lowest potential for early recurrence after curative resection. Hepatology, 2017, 66, 1502-1518.	7.3	87
23	Multivariate Analysis of Multiple Datasets: a Practical Guide for Chemical Ecology. Journal of Chemical Ecology, 2018, 44, 215-234.	1.8	86
24	MINT: a multivariate integrative method to identify reproducible molecular signatures across independent experiments and platforms. BMC Bioinformatics, 2017, 18, 128.	2.6	83
25	Ratios of T-cell immune effectors and checkpoint molecules as prognostic biomarkers in diffuse large B-cell lymphoma: a population-based study. Lancet Haematology,the, 2015, 2, e445-e455.	4.6	74
26	PTRF/cavin-1 neutralizes non-caveolar caveolin-1 microdomains in prostate cancer. Oncogene, 2014, 33, 3561-3570.	5.9	72
27	A fine-scale dissection of the DNA double-strand break repair machinery and its implications for breast cancer therapy. Nucleic Acids Research, 2014, 42, 6106-6127.	14.5	72
28	Temporal development of the oral microbiome and prediction of early childhood caries. Scientific Reports, 2019, 9, 19732.	3.3	65
29	YuGene: A simple approach to scale gene expression data derived from different platforms for integrated analyses. Genomics, 2014, 103, 239-251.	2.9	63
30	A retrospective analysis of clinical features of oral malignant and potentially malignant disorders with and without oral epithelial dysplasia. Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology, 2013, 116, 725-733.	0.4	62
31	In vivo gene expression in granulosa cells during pig terminal follicular development. Reproduction, 2008, 136, 211-224.	2.6	54
32	Genetic Variants in ERAP1 and ERAP2 Associated With Immune‐Mediated Diseases Influence Protein Expression and the Isoform Profile. Arthritis and Rheumatology, 2018, 70, 255-265.	5.6	52
33	An integrated metagenomics and metabolomics approach implicates the microbiota-gut-brain axis in the pathogenesis of Huntington's disease. Neurobiology of Disease, 2021, 148, 105199.	4.4	52
34	Variable selection in microbiome compositional data analysis. NAR Genomics and Bioinformatics, 2020, 2, lqaa029.	3.2	49
35	Ecological consequences of abrupt temperature changes in anaerobic digesters. Chemical Engineering Journal, 2019, 361, 266-277.	12.7	47
36	A Linear Mixed Model Spline Framework for Analysing Time Course 'Omics' Data. PLoS ONE, 2015, 10, e0134540.	2.5	46

#	ARTICLE	IF	CITATIONS
37	Novel Multivariate Methods for Integration of Genomics and Proteomics Data: Applications in a Kidney Transplant Rejection Study. OMICS A Journal of Integrative Biology, 2014, 18, 682-695.	2.0	41
38	A molecular classification of human mesenchymal stromal cells. PeerJ, 2016, 4, e1845.	2.0	41
39	Determinants of Body Fat in Infants of Women With Gestational Diabetes Mellitus Differ With Fetal Sex. Diabetes Care, 2011, 34, 2581-2585.	8.6	40
40	An mRNA atlas of G protein-coupled receptor expression during primary human monocyte/macrophage differentiation and lipopolysaccharide-mediated activation identifies targetable candidate regulators of inflammation. Immunobiology, 2013, 218, 1345-1353.	1.9	40
41	Diffusion-weighted Imaging Is a Sensitive and Specific Magnetic Resonance Sequence in the Diagnosis of Ankylosing Spondylitis. Journal of Rheumatology, 2018, 45, 771-778.	2.0	40
42	PhosphoPICK: modelling cellular context to map kinase-substrate phosphorylation events. Bioinformatics, 2015, 31, 382-389.	4.1	39
43	A Generic Multivariate Framework for the Integration of Microbiome Longitudinal Studies With Other Data Types. Frontiers in Genetics, 2019, 10, 963.	2.3	39
44	Altered Repertoire Diversity and Disease–Associated Clonal Expansions Revealed by T Cell Receptor Immunosequencing in Ankylosing Spondylitis Patients. Arthritis and Rheumatology, 2020, 72, 1289-1302.	5.6	39
45	A Natural History of Actinic Keratosis and Cutaneous Squamous Cell Carcinoma Microbiomes. MBio, 2018, 9, .	4.1	37
46	Personalised pathway analysis reveals association between DNA repair pathway dysregulation and chromosomal instability in sporadic breast cancer. Molecular Oncology, 2016, 10, 179-193.	4.6	36
47	IL-23 favours outgrowth of spondyloarthritis-associated pathobionts and suppresses host support for homeostatic microbiota. Annals of the Rheumatic Diseases, 2019, 78, 494-503.	0.9	36
48	Biomarker development for axial spondyloarthritis. Nature Reviews Rheumatology, 2020, 16, 448-463.	8.0	34
49	A small set of extra-embryonic genes defines a new landmark for bovine embryo staging. Reproduction, 2011, 141, 79-89.	2.6	33
50	Serum Glycoprotein Biomarker Discovery and Qualification Pipeline Reveals Novel Diagnostic Biomarker Candidates for Esophageal Adenocarcinoma. Molecular and Cellular Proteomics, 2015, 14, 3023-3039.	3.8	33
51	Uncoupled Embryonic and Extra-Embryonic Tissues Compromise Blastocyst Development after Somatic Cell Nuclear Transfer. PLoS ONE, 2012, 7, e38309.	2.5	29
52	Dietary intake influences gut microbiota development of healthy Australian children from the age of one to two years. Scientific Reports, 2019, 9, 12476.	3.3	28
53	Interpretation of network-based integration from multi-omics longitudinal data. Nucleic Acids Research, 2022, 50, e27-e27.	14.5	28
54	Integrative mixture of experts to combine clinical factors and gene markers. Bioinformatics, 2010, 26, 1192-1198.	4.1	27

#	Article	IF	Citations
55	Evaluation of Serum Glycoprotein Biomarker Candidates for Detection of Esophageal Adenocarcinoma and Surveillance of Barrett's Esophagus. Molecular and Cellular Proteomics, 2018, 17, 2324-2334.	3.8	25
56	Multiclass classification and gene selection with a stochastic algorithm. Computational Statistics and Data Analysis, 2009, 53, 3601-3615.	1.2	24
57	*Streptococcus* species enriched in the oral cavity of patients with RA are a source of peptidoglycan-polysaccharide polymers that can induce arthritis in mice. Annals of the Rheumatic Diseases, 2021, 80, 573-581.	0.9	24
58	Effect of Perioperative Opioids on Cancer-Relevant Circulating Parameters: Mu Opioid Receptor and Toll-Like Receptor 4 Activation Potential, and Proteolytic Profile. Clinical Cancer Research, 2018, 24, 2319-2327.	7.0	22
59	*In Situ* Stimulation of Thiocyanate Biodegradation through Phosphate Amendment in Gold Mine Tailings Water. Environmental Science & Technology, 2017, 51, 13353-13362.	10.0	20
60	Multi-Omic Data Integration Allows Baseline Immune Signatures to Predict Hepatitis B Vaccine Response in a Small Cohort. Frontiers in Immunology, 2020, 11, 578801.	4.8	20
61	Gene-environment-gut interactions in Huntington's disease mice are associated with environmental modulation of the gut microbiome. IScience, 2022, 25, 103687.	4.1	20
62	Sorting the nuclear proteome. Bioinformatics, 2011, 27, i7-i14.	4.1	18
63	A peripheral blood transcriptomic signature predicts autoantibody development in infants at risk of type 1 diabetes. JCI Insight, 2018, 3, .	5.0	18
64	Integrating independent microbial studies to build predictive models of anaerobic digestion inhibition by ammonia and phenol. Bioresource Technology, 2020, 316, 123952.	9.6	17
65	Prediction of kinase-specific phosphorylation sites through an integrative model of protein context and sequence. Biochimica Et Biophysica Acta - Proteins and Proteomics, 2016, 1864, 1599-1608.	2.3	16
66	Statistical challenges in longitudinal microbiome data analysis. Briefings in Bioinformatics, 2022, 23, .	6.5	16
67	DynOmics to identify delays and co-expression patterns across time course experiments. Scientific Reports, 2017, 7, 40131.	3.3	15
68	*Ski* Overexpression in Skeletal Muscle Modulates Genetic Programs That Control Susceptibility to Diet€Induced Obesity and Insulin Signaling. Obesity, 2012, 20, 2157-2167.	3.0	14
69	An integrated analysis of human myeloid cells identifies gaps in inÂvitro models of inÂvivo biology. Stem Cell Reports, 2021, 16, 1629-1643.	4.8	14
70	Selection of Biologically Relevant Genes with a Wrapper Stochastic Algorithm. Statistical Applications in Genetics and Molecular Biology, 2007, 6, Article29.	0.6	13
71	A multi-modal data harmonisation approach for discovery of COVID-19 drug targets. Briefings in Bioinformatics, 2021, 22, .	6.5	13
72	Epistatic interactions between killer immunoglobulin-like receptors and human leukocyte antigen ligands are associated with ankylosing spondylitis. PLoS Genetics, 2020, 16, e1008906.	3.5	12

#	Article	IF	Citations
73	timeOmics: an R package for longitudinal multi-omics data integration. Bioinformatics, 2022, 38, 577-579.	4.1	12
74	A simple, scalable approach to building a cross-platform transcriptome atlas. PLoS Computational Biology, 2020, 16, e1008219.	3.2	12
75	Genome-resolved metagenomics of an autotrophic thiocyanate-remediating microbial bioreactor consortium. Water Research, 2019, 158, 106-117.	11.3	11
76	Alterations in the Gut Fungal Community in a Mouse Model of Huntington's Disease. Microbiology Spectrum, 2022, 10, e0219221.	3.0	11
77	PhosphoPICK-SNP: quantifying the effect of amino acid variants on protein phosphorylation. Bioinformatics, 2017, 33, 1773-1781.	4.1	10
78	Taxonomic and functional assessment using metatranscriptomics reveals the effect of Angus cattle on rumen microbial signatures. Animal, 2020, 14, 731-744.	3.3	9
79	Community-wide hackathons to identify central themes in single-cell multi-omics. Genome Biology, 2021, 22, 220.	8.8	9
80	Integrating Multi-omics Data to Dissect Mechanisms of DNA repair Dysregulation in Breast Cancer. Scientific Reports, 2016, 6, 34000.	3.3	8
81	How does the early life environment influence the oral microbiome and determine oral health outcomes in childhood?. BioEssays, 2021, 43, e2000314.	2.5	8
82	Glyco-centric lectin magnetic bead array (LeMBA) – proteomics dataset of human serum samples from healthy, Barrett×³s esophagus and esophageal adenocarcinoma individuals. Data in Brief, 2016, 7, 1058-1062.	1.0	6
83	Exploring the association between BMI and mortality in Australian women and men with and without diabetes: the AusDiab study. Diabetologia, 2019, 62, 754-758.	6.3	6
84	Time-course analysis of metabolomic and microbial responses in anaerobic digesters exposed to ammonia. Chemosphere, 2021, 283, 131309.	8.2	6
85	A field guide to cultivating computational biology. PLoS Biology, 2021, 19, e3001419.	5.6	6
86	Identifying Molecular Mechanisms of the Late-Phase Asthmatic Response by Integrating Cellular, Gene, and Metabolite Levels in Blood. Annals of the American Thoracic Society, 2016, 13, S98-S98.	3.2	6
87	Analysis of a simulated microarray dataset: Comparison of methods for data normalisation and detection of differential expression (Open Access publication). Genetics Selection Evolution, 2007, 39, 669.	3.0	5
88	A pilot study on transcriptome data analysis of folliculogenesis in pigs. Animal, 2009, 3, 393-401.	3.3	5
89	Mapping the stabilome: a novel computational method for classifying metabolic protein stability. BMC Systems Biology, 2012, 6, 60.	3.0	5
90	Multiple interaction nodes define the postreplication repair response to UV-induced DNA damage that is defective in melanomas and correlated with UV signature mutation load. Molecular Oncology, 2020, 14, 22-41.	4.6	5

#	Article	IF	Citations
91	Model-based joint visualization of multiple compositional omics datasets. NAR Genomics and Bioinformatics, 2020, 2, lqaa050.	3.2	4
92	Sincast: a computational framework to predict cell identities in single-cell transcriptomes using bulk atlases as references. Briefings in Bioinformatics, 2022, 23, .	6.5	4
93	Predicting qualitative phenotypes from microarray data â€" the Eadgene pig data set. BMC Proceedings, 2009, 3, S13.	1.6	3
94	multiomics: A user-friendly multi-omics data harmonisation R pipeline. F1000Research, 0, 10, 538.	1.6	2
95	Statistical Analysis on Microarray Data: Selection of Gene Prognosis Signatures. , 2009, , 55-76.		2
96	The EADGENE Microarray Data Analysis Workshop (<i>Open Access publication</i>). Genetics Selection Evolution, 2007, 39, 621-631.	3.0	2
97	A new method for integrated analysis applied to gene expression and cytokines secretion in response to LIPO-5 vaccine in HIV-negative volunteers. Retrovirology, 2012, 9, .	2.0	1
98	The surprising benefit of passiveâ€"aggressive behaviour at Christmas parties: being crowned king of the crackers. Medical Journal of Australia, 2014, 201, 694-696.	1.7	1
99	Pathway dysregulation analysis of the nucleotide excision repair mechanisms reveals it is not a common feature of melanomas. Pigment Cell and Melanoma Research, 2019, 32, 336-338.	3.3	1
100	ofw: An<i>R</i>Package to Select Continuous Variables for Multiclass Classification with a Stochastic Wrapper Method. Journal of Statistical Software, 2008, 28, .	3.7	1
101	Analysis of the real EADGENE data set: Multivariate approaches and post analysis (<i>Open Access) Tj ETQq1 1 0.784314 rgBT /Overl	3.0	1
102	Analysis of a simulated microarray dataset: Comparison of methods for data normalisation and detection of differential expression (<i>Open Access publication</i>). Genetics Selection Evolution, 2007, 39, 669-683.	3.0	1
103	A statistician's Christmas party. Significance, 2014, 11, 44-47.	0.4	0
104	Abstract 2499: Diagnostic serum biomarkers for canine hemangiosarcoma: a potential model of human angiosarcoma.. , 2013, , .		0
105	A simple, scalable approach to building a cross-platform transcriptome atlas. , 2020, 16, e1008219.		0
106	A simple, scalable approach to building a cross-platform transcriptome atlas. , 2020, 16, e1008219.		0
107	A simple, scalable approach to building a cross-platform transcriptome atlas. , 2020, 16, e1008219.		0
108	A simple, scalable approach to building a cross-platform transcriptome atlas. , 2020, 16, e1008219.		0

#	Article	IF	Citations
109	Title is missing!. , 2020, 16, e1008906.		0
110	Title is missing!. , 2020, 16, e1008906.		0
111	Title is missing!. , 2020, 16, e1008906.		0
112	Title is missing!. , 2020, 16, e1008906.		0