Supplementary Table S1. Top 15 scientific sources and their metrics

Source	NP	LC	H-index	G-index	M-index	TC	PY start
Bioinformatics	187	3,435	35	64	3.889	4,943	2015
Nature Communications	134	5,244	41	71	5.125	5,439	2016
Nucleic Acids Research	103	2,683	32	76	2.909	5,835	2013
Genome Biology	84	3,761	39	84	3.9	7,377	2014
Methods in Molecular Biology	83	30	12	21	1.091	588	2013
Briefings in Bioinformatics	70	50	14	25	1.556	749	2015
BMC Bioinformatics	68	1,015	19	39	2.375	1,618	2016
Frontiers in Immunology	57	459	11	24	1.833	633	2018
Frontiers in Genetics	54	287	13	26	1.625	757	2016
Nature Methods	54	318	37	54	3.364	7,859	2013
BMC Genomics	50	455	13	34	1.182	1,203	2013
PLOS Computational biology	45	425	14	28	1.556	813	2015
Scientific Reports	41	67	14	24	1.75	657	2016
Nature Biotechnology	36	246	22	36	2.2	10,931	2014
Cell Reports	32	192	18	32	2.25	1,771	2016

This legend examines key bibliometric indicators for the top 15 sources that have published literature on single-cell RNA sequencing (scRNA-seq) analysis. These indicators include the number of publications (NP), local citations (LC), h-index (a metric that evaluates both the productivity and citation impact of the publications), g-index (a measure focusedon highly cited publications), m-index (an evaluation of the impact based on an author's most cited works), total citations (TC), and average publications per year (PY). The analysis of these metrics offers valuable insights into the productivity, citation impact, and overall influence of these leadingsources in the field of scRNA-Seq analysis.

Supplementary Table S2. A complete list of the most common pre-processing analysis features for scRNA-seq analysis

Name	STAR	Seurat	Monocle	kallisto	salmon	CellRanger	Scanpy	velocyto	scran	Harmony	MAST	RaceID	scvi-tools
Quality control						Python/R	Python				R		Python
Normalization & transformation		R				Python/R	Python		R		R		Python
Alignment	C/C++			C/C++	C++	Python/R							
UMIs	C/C++					Python/R							
Dimensionality reduction		R	R			Python/R	Python			R/C++		R/C++	Python
Integration		R							R	R/C++		R/C++	Python
Dimensionality reduction		R	R			Python/R	Python			R/C++		R/C++	Python
Variable gene/alternative splicing		R					Python	Python/R	R				
Marker genes		R	R			Python/R						R/C++	
Imputation		R					Python					R/C++	Python

Systematically categorize crucial steps in preprocessing single-cell RNA sequencing (scRNA-seq) analysis. Each table is divided into two distinct categories for immediate recognition: essential and advanced features. Essential steps, pivotal for the foundational analysis process, are marked in light orange. These steps are critical to ensure the integrity and quality of the data for subsequent analysis. Advanced features, marked in green, denote supplementary analysis techniques that, while not mandatory for all studies, can significantly enrich the analysis. These advanced options offer opportunities for deeper data exploration and can be tailored to address specific research inquiries. This dual-colour coding effectively guides users through the critical workflow of scRNA-seq analysis, distinguishing between indispensable steps and those that offer further depth and refinement to the research. UMIs: unique molecular identifiers.

Supplementary Table S3. A complete list of the most common downstream analysis features for scRNA-seq analysis

Name	STAR	Seurat	Monocle	kallisto	salmon	CellRanger	Scanpy	infer-CNV	CellPhoneDB	BackSPIN	SCENIC	AUCell	velocyto	scran	MAST	RaceID	scvi-tools	SCDE
Clustering		R	R			Python/R	Python			Python	Python/R					R/C++		
Quality control						Python/R	Python								R		Python	
Normalization & transformation		R					Python							R	R		Python	
Gene filtering		R					Python			Python								
Alignment	C/C++					Python/R												
Visualization	C/C++	R	R			Python/R	Python	R	Python		Python/R	R	Python/R	R		R/C++		R
Gene networks									Python		Python/R				R			
Classification																	Python	
Ordering		R	R				Python					R	Python/R			R/C++		
Differential expression		R	R			Python/R	Python								R	R/C++	Python	
Gene sets															R			R
Interactive						Python/R						R		R				
Quantification	C/C++			C/C++	C++	Python/R												
Simulation							Python											
Variants								R										
Expression patterns			R													R/C++		
Cell cycle														R				
Rare cells																R/C++		

The table systematically organizes key steps in single-cell RNA sequencing (scRNA-seq) analysis, highlighting features critical for downstream analysis. It is structured into two main categories for easy identification: essential and advanced features. Essential steps, crucial for the core analysis process, are highlighted in orange. These steps are to maintain data integrity and quality for further analysis. In contrast, advanced features, highlighted in green, represent additional analysis methods that, although not essential for every study, can greatly enhance the analysis. These advanced techniques allow for more in-depth exploration of the data and can be customized to meet specific research needs. The use of dual-colour coding serves as an effective navigation tool through the scRNA-seq analysis workflow, clearly differentiating between fundamental steps and those that provide additional insight and sophistication to the research.