Table S1 The comparison of BP4RNAseq with other tools for processing raw RNA-seq data.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Comparison | Analysis type | Quantification method | Quantification type | Tool type |
| BP4RNAseq | Retrospective and newly generated RNA-seq data analyses | Alignment-based (HISAT2 + StringTie) and alignment-free (Salmon) | Gene and transcript read counts | Software |
| Recount2 (Collado-Torres, et al., 2017) | Retrospective RNA-seq data analysis | Alignment-based (Rail-RNA) | Gene and transcript read counts | Compiled datasets |
| GREIN (Al Mahi, et al., 2019) | Retrospective RNA-seq data analysis | Alignment-free (Salmon) | Gene and transcript read counts | Compiled datasets |
| ARCHS4 (Lachmann, et al., 2018) | Retrospective RNA-seq data analysis | Alignment-free (Kallisto) | Gene and transcript read counts | Compiled datasets |
| Toil (Vivian, et al., 2017) | Retrospective RNA-seq data analysis | Alignment-free (Kallisto) | Transcript read counts | Compiled datasets |
| Skymap (Tsui, et al., 2019) | Retrospective RNA-seq data analysis | Alignment-based (Bowtie2) | Allelic and transcript read counts | Compiled datasets |
| Expression Atlas (Papatheodorou, et al., 2018) | Retrospective RNA-seq data analysis | Alignment-based (Tophat) | Gene read counts | Compiled datasets |
| GREP2 (Al Mahi, et al., 2019) | Retrospective RNA-seq data analysis | Alignment-free (Salmon) | Gene and transcript read counts | Software |
| Lair (Pimentel, et al., 2016) | Retrospective RNA-seq data analysis | Alignment-free (Kallisto) | Gene and transcript read counts | Software |

**References**

Al Mahi, N.*, et al.* GREIN: An Interactive Web Platform for Re-analyzing GEO RNA-seq Data. *Sci Rep-Uk* 2019;9:1-9.

Collado-Torres, L.*, et al.* Reproducible RNA-seq analysis using recount2. *Nat Biotechnol* 2017;35(4):319-321.

Lachmann, A.*, et al.* Massive mining of publicly available RNA-seq data from human and mouse. *Nature Communications* 2018;9:1-10.

Papatheodorou, I.*, et al.* Expression Atlas: gene and protein expression across multiple studies and organisms. *Nucleic Acids Res* 2018;46(D1):D246-D251.

Pimentel, H.*, et al.* The Lair: a resource for exploratory analysis of published RNA-Seq data. *Bmc Bioinformatics* 2016;17:1-6.

Tsui, B.*, et al.* Extracting allelic read counts from 250,000 human sequencing runs in Sequence Read Archive. In: Altman, R.B.*, et al.*, editors, *Pacific Symposium on Biocomputing 2019* 2019;196-207. Singapore: World Scientific Publ Co Pte Ltd.

Vivian, J.*, et al.* Toil enables reproducible, open source, big biomedical data analyses. *Nat Biotechnol* 2017;35(4):314-316.