

HTSinfer: Supplementary materials

- README.pdf: This file describing Zenodo repository contents.
- sample_tables_htsinfer.zip: Zipped directory containing sample tables used for testing HTSinfer.
 - htsinfer_self_generated_samples_table.tsv: TSV file describing 72 RNA-Seq samples generated by the laboratory, then made public on SRA.
 - htsinfer_sra_samples_table.tsv: TSV file describing 720 RNA-Seq samples that were downloaded from SRA.
- results_htsinfer.zip: Zipped directory containing tables that describe the metadata inference results. In case of the 720 RNA-Seq samples, HTSinfer was tested with the “--records” argument set to various values (10 thousand, 100 thousand, 1 million, 10 million, 0 – all records) to assess performance and runtimes. The TSV files contain information regarding metadata from SRA (columns 1-8), the aggregated results from HTSinfer (columns 9-22), the count metrics describing the results of the library source inference (columns 23-34), the count metrics describing the results of the 3’ adapter inference (columns 35-42), the runtimes, in seconds, for the various submodules (columns 43-47), the alignment statistics used for the library type inference (columns 48-51), and the comparison results, describing the concordance between the metadata from SRA and the inference from HTSinfer (columns 51-55).
 - htsinfer_self_generated_samples_results.tsv
 - htsinfer_sra_samples_10k_records_results.tsv
 - htsinfer_sra_samples_100k_records_results.tsv
 - htsinfer_sra_samples_1m_records_results.tsv
 - htsinfer_sra_samples_10m_records_results.tsv
 - htsinfer_sra_samples_all_records_results.tsv
 - htsinfer_sra_samples_1m_records_70median_tin_results.tsv: TSV file describing the results from the 720 RNA-Seq samples downloaded from SRA that had a median TIN score ≥ 70 .
- Supplementary_table_1.pdf: Supplementary Table 1, comparing the functionality of available metadata inference tools.