

Directories

- README.pdf: This file describing Zenodo repository contents.
- APAEVAL.tar.gz: tarball containing the data used in APAeval for genome annotations, ground truth, etc. as well as outputs to method workflows and summary workflows for plots made in the paper. Contains the following directories:
 - benchmarking_results: Contains the outputs and plots for various benchmarking events (identification, absolute quantification, relative quantification).
 - genome_resources: Contains genome annotation files used for data preprocessing, method workflows, and benchmarking workflows. These include general transcriptome annotation files (GTFs) as well as algorithm specific BED files used for annotations of 3'UTRs or PAS sites. See README within for more details.
 - inputs_ground_truth: Contains BED6 files for orthogonal / ground-truth cleavage and polyadenylation sites (PAS) based on simulation or targeting sequencing experiments. Subdirectories contain subsets (i.e. those in terminal exons) and/or transformations (i.e. relative expression of major PAS sites within a terminal exons) used for benchmarking challenges. See README within for more details.
 - inputs_rnaseq: Contains files used to run and summary files resulting from nfcornaseq runs on mapping of all RNA-seq samples for APAeval.
 - method_workflow_outputs: Contains the outputs of all method workflow runs used in APAeval for the three benchmarking events (Identification, Absolute Quantification, and Relative Quantification). Each method workflow was also run on a standard / shared annotation (those ending with _gencode suffix) or the algorithm's preferred annotation (typically RefSeq) (those ending with _preferred suffix). Outputs are BED6 files describing the PAS with the score field indicating absolute or relative quantification, if applicable.
 - Challenge_data_summary.pdf: PDF containing table describing RNA-seq data SRA accession numbers and matching orthogonal data (3'end sequencing or simulation) accession numbers.
- bams: contains bam files resulting from nfcornaseq runs described within APAEVAL/inputs_rnaseq/mapping_samples_nfcornaseq.pdf