Template of the Method Section Using OneStopRNAseq v1.0.0

The RNA-seq analysis was performed using OneStopRNAseq (Li et al., 2020). Specifically, FastQC (Andrews, 2010) and MultiQC (Ewels et al., 2016) were used for raw reads quality control and QoRTs (Hartley & Mullikin, 2015) for post-alignment quality control.

Reads were aligned to the reference genome assembly XXX (e.g. hg38, mm10) with star_2.7.5a (Dobin et al., 2013) and annotated with XXX (e.g. gencode.v34.primary_assembly, gencode.vM25.primary_assembly) (Harrow et al., 2012). Aligned exon reads were counted toward gene expression with featureCounts_2.0.0 (Liao et al., 2014) with default settings except the following parameters XXX (e.g. '`-Q 20 --minOverlap 1 --fracOverlap 0 -p -B -C`' for paired-end strict-mode analysis). Differential expression (DE) analysis was performed with DESeq2_1.28.1 (Love et al., 2014). Within DE analysis, 'ashr' was used to create log2 Fold Change (LFC) shrinkage for each comparison (Stephens, 2016). Significantly differentially expressed genes (DEGs) were filtered with the criteria FDR < XXX (please replace XXX with the MAX_FDR you entered) and absolute log2 fold change (|LFC|) > XXX (please replace XXX with the MIN_LFC you entered). Heatmaps were created with pheatmap (Kolde, n.d.). Gene set enrichment analysis were performed with GSEA (Subramanian et al., 2005).

Alternative splicing analysis was performed with rMATS_4.1.0 (Shen et al., 2014). Differential exon usage analysis was performed with DEXSeq_1.34.0 (Anders et al., 2012).

If you performed transposable element (TE) analysis, please include the following section. Transposable element (TE) expression quantification was performed with SalmonTE (Jeong et al., 2018). The TE expression quantification was combined with regular gene expression profile and analyzed with DESeq2 together for improved normalization.

Allele-specific gene expression quantification was performed with ASEReadCounter from GATK package (Van der Auwera et al., 2013).

Note:
XXX: a placeholder to be filled with information from the software and parameter setting table in the 'About' page.

List of reference genome assemblies and annotations in OneStopRNAseq v1.0.0

Species	Genome	Annotation
Human	hg38	gencode.v34.primary_assembly
Mouse	mm10	gencode.vM25.primary_assembly
Worm (*C. elegans*)	WBcel235	WBcel235.90
Yeast (*S. cerevisiae*)	R64-1-1	R64-1-1.90
Fruit fly (*D. melanogaster*)	BDGP6	BDGP6.22.96

Zebra fish (*D. rerio*)	danRer11	V4.3.2 (Lawson et al., 2020)

References:

Anders, S., Reyes, A., & Huber, W. (2012). Detecting differential usage of exons from RNA-seq data. *Genome Research*, *22*(10), 2008–2017. https://doi.org/10.1101/gr.133744.111

Andrews, S. (2010). FastQC. *Babraham Bioinformatics*. https://doi.org/citeulike-article-id:11583827

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/bts635

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/btw354

Harrow, J., Frankish, A., Gonzalez, J. M., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B. L., Barrell, D., Zadissa, A., Searle, S., Barnes, I., Bignell, A., Boychenko, V., Hunt, T., Kay, M., Mukherjee, G., Rajan, J., Despacio-reyes, G., Saunders, G., … Guigo, R. (2012). GENCODE : The reference human genome annotation for The ENCODE Project. *Genome Research*, 1760–1774. https://doi.org/10.1101/gr.135350.111.

Hartley, S. W., & Mullikin, J. C. (2015). QoRTs: A comprehensive toolset for quality control and data processing of RNA-Seq experiments. *BMC Bioinformatics*. https://doi.org/10.1186/s12859-015-0670-5

Jeong, H. H., Yalamanchili, H. K., Guo, C., Shulman, J. M., & Liu, Z. (2018). An ultra-fast and scalable quantification pipeline for transposable elements from next generation sequencing data. *Pacific Symposium on Biocomputing*, *0*(212669), 168–179. https://doi.org/10.1142/9789813235533_0016

Kolde, R. (n.d.). *pheatmap*. https://github.com/raivokolde/pheatmap

Lawson, N. D., Li, R., Shin, M., Grosse, A., Yukselen, O., Stone, O. A., Kucukural, A., & Zhu, L. (2020). An improved zebrafish transcriptome annotation for sensitive and comprehensive detection of cell type-specific genes. *ELife*. https://doi.org/10.7554/eLife.55792

Li, R., Hu, K., Liu, H., Green, M. R., & Zhu, L. J. (2020). Onestoprnaseq: A web application for comprehensive and efficient analyses of rna-seq data. *Genes*, *11*(10), 1–14. https://doi.org/10.3390/genes11101165

Liao, Y., Smyth, G. K., & Shi, W. (2014). FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/btt656

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*. https://doi.org/10.1186/s13059-014-0550-8

Shen, S., Park, J. W., Lu, Z. X., Lin, L., Henry, M. D., Wu, Y. N., Zhou, Q., & Xing, Y. (2014). rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1419161111

Stephens, M. (2016). False discovery rates: a new deal. *Biostatistics*, kxw041. https://doi.org/10.1093/biostatistics/kxw041

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., & Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.0506580102

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K. V., Altshuler, D., Gabriel, S., & DePristo, M. A. (2013). From fastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Current Protocols in Bioinformatics*. https://doi.org/10.1002/0471250953.bi1110s43