

The battle for reads: evaluating strategies to tackle multi-mapping in RNA-seq quantification in highly repetitive genomes

ADDITIONAL FILE 1: SUPPLEMENTARY TABLES

Aldana A. Cepeda Dean^{1,2}, Carlos A. Buscaglia^{1,2}, Virginia Balouz^{1,2}, Natalia Rego^{3,5*},
Luisa Berná^{3,4,5*}

¹Instituto de Investigaciones Biotecnológicas (IIBio), Universidad Nacional de San Martín (UNSAM), and Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina.

²Escuela de Bio y Nanotecnologías (EByN), UNSAM, Buenos Aires, Argentina.

³Bioinformatic Unit, Institut Pasteur de Montevideo, Uruguay.

⁴Laboratory of Apicomplexan Biology. Institut Pasteur de Montevideo, Uruguay

⁵Laboratorio de Genómica Evolutiva, Facultad de Ciencias, Universidad de la República, Montevideo, Uruguay.

*Both authors contributed equally to this work and should therefore be considered co-senior authors.

Address correspondence to: Luisa Berná (lberna@pasteur.edu.uy) or Natalia Rego (nrego@pasteur.edu.uy), Institut Pasteur de Montevideo, Mataojo 2020, 11400 Montevideo, Departamento de Montevideo, Uruguay

Supplementary Table 1. RNA-Seq methodologies in *T. cruzi*

Reference	Alignment	Parameters	Quantification	Multi-mapping ¹
RNAseq				
Sabalette et al., 2023 (<i>JBC</i>)	Bowtie2	--very-sensitive-local	featureCounts ('-p -t "CDS" -g "ID" -T 40')	Not special strategy indicated; using Bowtie2+featureCounts by default discards multi-mapping reads
Lima et al., 2023 (<i>Parasites & Vectors</i>)	STAR (v2.7.7a)	up to 3 mismatches	HTSeq-count (v0.5.4p3, union model)	Not special strategy indicated; by default HTSeq-count discards multi-mapping
Sabalette et al., 2023 (<i>PLOS NTDs</i>)	Bowtie2	--very-sensitive-local	featureCounts ('-p -t "CDS" -g "ID" -T 40')	Not special strategy indicated; using Bowtie2+featureCounts by default discards multi-mapping reads
Sabalette et al., 2024 (<i>Genomics Data</i>)	Bowtie2	--very-sensitive-local	featureCounts ('-p -t "CDS" -g "ID" -T 40')	Not special strategy indicated; using Bowtie2+featureCount by default discards multi-mapping reads
Cruz-Saavedra et al., 2020 (<i>Virulence</i>)	TopHat (v2.1.0)	default	Not declared	Review unallocated reads and align them using the information on splicing junctions
Callejas-Hernandez et al., 2019 (<i>Sci Rep</i>)	Bowtie2	up to 3 mismatches	Stringtie	Reports one alignment while indicating multiple possible

				mappings
Chavez et al., 2017 (<i>PLOS One</i>)	Bowtie2	--very-sensitive-local	HTSeq (v0.6.0)	Not special strategy indicated
Cruz-Saavedra et al., 2020 (<i>PeerJ</i>)	TopHat (v2.1.0)	default	Cufflink	Review unallocated reads and align them using the information on splicing junctions
Paes et al., 2020 (<i>PLOS NTDs</i>)	Bowtie2 (v2.3.5)	--very-sensitive-local	featureCounts	Only used the uniquely and correctly mapped reads.
Trey Belew et al., 2018 (<i>PLOS Pathogens</i>)	TopHat2 (v2.0.14)	-g 1, -G gff_file	HTSeq	Allowing a single randomly placed mapping per read
García-Huertas et al., 2022 (<i>Parasitology Research</i>)	Bowtie2	default	featureCounts	Not special strategy indicated; using Bowtie2+featureCounts by default discards multi-mapping reads
Single-cell				
Inchausti et al., 2025 (<i>eLife</i>)	kallisto bustools	--filter -em	kallisto	Take into account the multi-mapping reads
Laidlaw et al. 2024 (<i>preprint</i>)	STAR	default parameters	featureCounts	Multi-mapping and multi-overlapping genes were not counted

¹This refers to the strategy used by authors to the handling of multi-mapping reads.

Supplementary Table S2. Mapped reads among selected strategies.

Strategy	Total Reads after trimming	Mapped Reads	Properly Paired	Mapping Rate (%)
Bowtie2	47,994,598	46,173,873	44,145,550	91,98%
STAR	47,994,598	102,603,394	37,391,462	77,90%
Salmon	47,994,598	85,026,076	33,523,814	69,85%

Supplementary Table 3. Median of multi-mapping across *T. cruzi* multigenic families.

Multigenic Family	Gene dosage in YC6	Multi-mapping		
		Bowtie2	STAR	Salmon
MASP	703	74,03%	38,82%	41,82%
Mucin	526	73,58%	37,10%	41,18%
GP63	143	78,24%	51,59%	59,27%
TS	617	73,50%	41,62%	56,90%
DGF-1	102	76,91%	29,86%	51,27%