

Table 1: Selected metrics based on the output of **rnaQUAST**, **HISAT2**, **DETONATE**, **TransRate** and **BUSCO** for the transcripts assembled by all ten tools on the *Escherichia coli* RNA-Seq data set. Details and much more statistics, complementing this evaluation, can be found in the electronic supplement, content S4–S10. The RSEM-EVAL score is divided by 10^9 . Number of ambiguous bases is given in thousand. F1 score – an F1 score of 1 would mean that all nucleotides/contigs in the estimated true assembly were recovered with at least 90% identity. KC score – k -mer compression score reflecting the similarity of each assembly to **DETONATE**s estimated “true” assembly. The Ex90N50 statistic is computed as usual N50 but limited to the top most highly expressed transcripts that represent 90% of the total normalized expression data. Normalized scores in the range between 0 and 1 are displayed. Raw values for each metric and assembly tool are given in subscript next to the normalized values. In the last row, the summarized *metric score (MS)* of (0,1)-normalized scores is given for each assembly tool.

	Trinity	Oases	Trans-ABYSS	SOAP-Trans	Bridger	BinPacker	IDBA-Tran	Shannon	SPAdes-sc	SPAdes-rna
<i>k</i> -mer size	default	25,35,45,55,65	25,35,45,55,65	default	default	default	25,35,45,55,65	default	default	default
Evaluation metrics 1–20										
HISAT2										
1 Overall mapping rate	0.70 _{77.01}	0.24 _{49.16}	1.00 _{95.67}	0.36 _{56.62}	0.86 _{87.35}	0.60 _{71.09}	0.00 _{34.31}	0.69 _{76.69}	0.88 _{88.04}	0.89 _{88.9}
rnaQUAST										
2 Transcripts ≥ 1000 nt	0.42 ₃₄₇	1.00 ₇₄₃	0.56 ₄₄₆	0.55 ₄₃₉	0.44 ₃₆₃	0.00 ₆₆	0.51 ₄₁₄	0.45 ₃₇₂	0.62 ₄₈₉	0.40 ₃₄₀
3 Misassemblies	0.87 ₁₃	0.00 ₁₀₃	0.83 ₁₈	0.99 ₁	0.50 ₅₂	0.92 ₈	1.00 ₀	0.48 ₅₄	0.88 ₁₂	0.90 ₁₀
4 Mismatches per transcript	0.88 _{0.37}	0.70 _{0.71}	0.90 _{0.32}	1.00 _{0.13}	0.93 _{0.27}	0.00 _{2.05}	0.98 _{0.17}	0.97 _{0.19}	0.94 _{0.25}	0.96 _{0.21}
5 Average alignment length	0.20 _{467.04}	0.19 _{457.72}	0.00 _{310.44}	0.07 _{364.88}	0.21 _{472.81}	1.00 _{1100.77}	0.29 _{540.83}	0.22 _{482.66}	0.31 _{558.9}	0.06 _{359.2}
6 95%-assembled isoforms	0.95 ₃₉₆	1.00 ₄₁₄	0.00 ₃₂	0.60 ₂₆₁	0.92 ₃₈₂	0.01 ₃₄	0.57 ₂₅₀	0.95 ₃₉₃	0.47 ₂₁₀	0.41 ₁₈₇
7 Duplication ratio	0.98 _{1.012}	0.00 _{1.443}	0.90 _{1.044}	1.00 _{1.002}	0.97 _{1.014}	0.40 _{1.267}	1.00 _{1.001}	0.94 _{1.027}	0.99 _{1.006}	0.98 _{1.009}
8 Ex90N50	0.46 ₃₆₉	0.21 ₂₂₉	0.12 ₁₇₇	0.11 ₁₇₃	0.50 ₃₉₃	1.00 ₆₇₅	0.34 ₃₀₄	0.45 ₃₆₃	0.00 ₁₁₁	0.24 ₂₄₅
9 # full-length transcripts	0.92 ₄₀₇	0.94 ₄₁₈	0.99 ₄₃₆	0.96 ₄₂₃	0.90 ₄₀₂	0.00 ₃₇	0.89 ₃₉₆	0.90 ₃₉₉	1.00 ₄₄₁	0.75 ₃₃₈
TransRate										
10 Reference coverage	1.00 _{0.32}	0.97 _{0.31}	0.07 _{0.03}	0.61 _{0.19}	0.98 _{0.31}	0.00 _{0.01}	0.50 _{0.16}	0.96 _{0.3}	0.44 _{0.14}	0.55 _{0.18}
11 Mean ORF percentage	1.00 _{75.04}	0.69 _{65.72}	0.92 _{72.73}	0.90 _{71.99}	0.99 _{74.59}	0.00 _{45.13}	0.90 _{71.99}	0.94 _{73.34}	0.81 _{69.47}	0.96 _{73.78}
12 Optimal score ^a	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
13 Percentage bases uncovered ^a	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
14 Number of ambiguous bases	0.43 ₂₁₄₃	0.00 ₃₆₀₂	0.15 ₃₀₈₁	0.30 ₂₅₇₉	0.44 ₂₁₁₇	1.00 ₁₉₁	0.45 ₂₀₆₉	0.43 ₂₁₄₅	0.36 ₂₃₇₀	0.31 ₂₅₅₆
DETONATE										
15 Nucleotide F1	0.86 _{0.64}	0.63 _{0.49}	0.82 _{0.61}	1.00 _{0.74}	0.85 _{0.64}	0.00 _{0.06}	0.89 _{0.66}	0.83 _{0.62}	0.97 _{0.71}	0.97 _{0.72}
16 Contig F1	0.46 _{0.03}	0.38 _{0.03}	0.47 _{0.03}	0.89 _{0.06}	0.43 _{0.03}	0.00 ₀	0.41 _{0.03}	0.38 _{0.03}	0.83 _{0.06}	1.00 _{0.07}
17 KC score	0.92 _{0.82}	0.89 _{0.79}	0.93 _{0.82}	0.98 _{0.86}	0.91 _{0.81}	0.00 _{0.16}	0.94 _{0.83}	0.91 _{0.81}	1.00 _{0.87}	0.96 _{0.84}
18 RSEM EVAL	0.79 _{-0.21}	0.66 _{-0.25}	1.00 _{-0.15}	0.34 _{-0.34}	0.93 _{-0.17}	0.62 _{-0.26}	0.00 _{-0.44}	0.83 _{-0.2}	0.86 _{-0.19}	0.83 _{-0.2}
BUSCO										
19 Complete BUSCOs	0.83 ₂₈₄	0.88 ₂₉₉	0.88 ₂₉₇	0.94 ₃₁₆	0.83 ₂₈₅	0.00 ₅₀	0.87 ₂₉₆	0.82 ₂₈₀	1.00 ₃₃₂	0.73 ₂₅₅
20 Missing BUSCOs	0.96 ₁₉₀	1.00 ₁₇₂	1.00 ₁₇₀	0.99 ₁₇₈	0.96 ₁₉₀	0.00 ₇₁₁	0.95 ₁₉₆	0.95 ₁₉₈	1.00 ₁₇₂	0.96 ₁₈₉
Summed up metric (0,1)-score	13.61	10.36	11.53	12.59	13.55	5.55	11.49	13.08	13.37	12.85

^aNot available for the *E. coli* and *A. thaliana* data set because this metric is only calculated by **TransRate** in the case of paired-end data.