





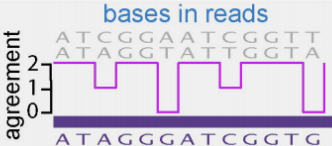




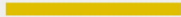

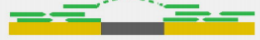









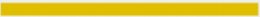


Transcriptome Assembly

Lecture 26
Nov 2, 2016

Announcements

TRANSCRIPTOME ASSEMBLY – END PRODUCT

TRANSCRIPTOME ASSEMBLY – EVALUATION

Error type	Transcripts	Assembly	Read evidence
Family collapse	geneAA  geneAB  geneAC  n=3	 n=1	
Chimerism	 geneC geneB  n=2	 n=1	
Unsupported insertion	 n=1	 n=1	no reads align to insertion 
Incompleteness	 n=1	 n=1	read pairs align off end of contig 
Fragmentation	 n=1	 n=4	bridging read pairs 
Local misassembly	 n=1	 n=1	read pairs in wrong orientation 
Redundancy	 n=1	 n=3	all reads assign to best contig 

TRANSCRIPTOME ASSEMBLY – EVALUATION – TRANSRATE



TransRate

① input data

assembled contigs paired-end reads



② align reads to contigs

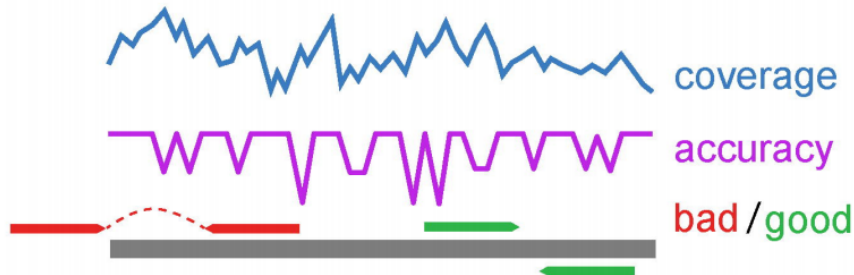


TRANSCRIPTOME ASSEMBLY – EVALUATION – TRANSRATE

③ assign multimapping reads

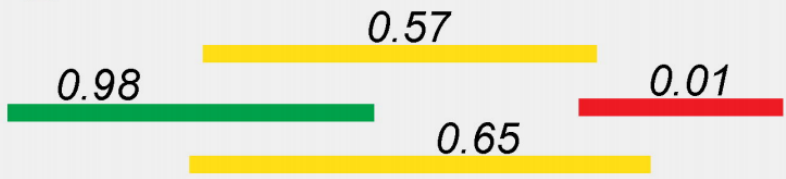


④ collect contig score components

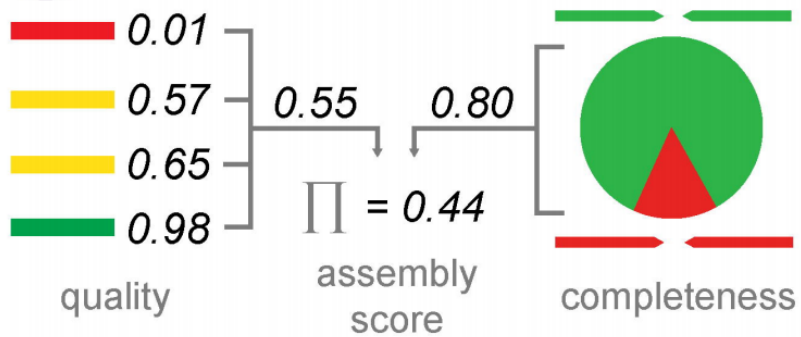


TRANSCRIPTOME ASSEMBLY – EVALUATION – TRANSRate

⑤ calculate contig scores

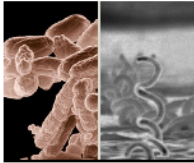


⑥ calculate assembly score



TRANSCRIPTOME ASSEMBLY – EVALUATION – BUSCO

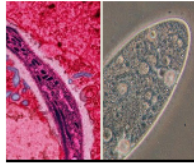
Datasets *(Beta versions, updated sets and additional lineages coming soon)*



Bacteria sets



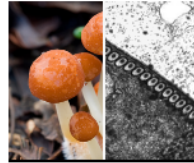
Eukaryota sets



Protists sets



Metazoa sets



Fungi sets



Plants set

[Download all datasets](#)

Image credits

TRANSCRIPTOME ASSEMBLY – ANNOTATION