

Table S2: Assembly statistics for EvidentialGene tr2aacds pipeline merged assembly compared to average statistics for each assembler.

Assembly	Merged Assembly		Trinity	Velvet/Oases	SOAPdenovo-Trans
	Loci	Transcripts			
Total Assembled Transcripts	66 377	247 035	134 766	185 574	271 378
Longest Transcript (Kb)	10 839	10 839	12 167	11 490	27 448
Assembly Length (Mb)	63.45	230.62	123.11	216.75	121.16
Mean Sequence Length (nt)	956	934	910	1 195	392
Predicted ORFs in assembly	38 386	134 827	41 425	86 984	33 707
% ORF ¹	85.46%	78.97%	52.70%	55.73%	12.42%
N90	462	480	449	589	356
N70	798	786	728	1 049	983
N50	1 194	1 113	1 122	1 530	1 802
N30	1 692	1 539	1 697	2 158	2 979
N10	2 853	2 514	2 745	3 441	27 448
GC	46.87%	46.00%	42.55%	42.98%	42.91%
CPG ratio ²	1.71	1.69	1.65	1.65	1.69
% N ³	0.13%	0.13%	0.00%	0.05%	1.08%

¹ mean percentage of a contig covered by an open reading frame (ORF).

² counts of CpG sites relative to the expected number.

³ proportion of bases in the assembly that are N.