

Genome Analysis Report

Taxon Info	
TxID	4530
Species	Oryza sativa
Class	Magnoliopsida
Order	Poales

Warnings

1. Number of genes

- Number monoexon transcript > 25.0% (35.00%)

2. Predicted CDS

- PSAURON score < 90 (84.5)

3. BUSCO/OMArk completeness

No warnings

4. Intron length distribution

No warnings

5. RNA-Seq support

- Percent gene supported < 75.0% (74.05%)

Summary

num_genes	36391
num_transcripts	41048
num_transcripts_without_introns	14365 (35.00%)
mean_exons_per_transcript	4.26

mean_exons_per_multiexon_transcript	6.02
mean_cds_length	1109.73
median_cds_length	894.0
mean_intron_length	408.83
median_intron_length	147.0
BUSCO_annotation	C:97.6[S:84.6,D:13.1],F:0.4,M:2.0
BUSCO_assembly	C:97.9[S:97.0,D:0.9],F:1.2,M:0.9,E:2.4
PSAURON_score	84.5
num_intron_supported	107213 (80.02%)

General Statistics	
num_genes	36391
mean_gene_length	2359.73
median_gene_length	1563.0
num_transcripts	41048
mean_transcripts_per_gene	1.13
median_transcripts_per_gene	1.0
num_exons	175037
mean_exons_per_transcript	4.26
median_exons_per_transcript	2.0
mean_exons_per_multiexon_transcript	6.02
median_exons_per_multiexon_transcript	4.0
num_transcripts_without_introns	14365 (35.00%)
num_exon_3n	89250 (50.99%)
num_exon_3n1	43323 (24.75%)

num_exon_3n2	42464 (24.26%)
mean_cds_length	1109.73
median_cds_length	894.0
total_cds_length	45552283
percentage_cds_coverage	12.20%
num_introns	133989
mean_intron_length	408.83
median_intron_length	147.0
mean_introns_per_transcript	3.26
median_introns_per_transcript	1.0
short_intron_3n/short_intron	19194 (33.71%)
long_intron_3n/long_intron	25586 (33.21%)
stopless_short_intron_3n/stopless_short_intron	3281 (27.01%)
stopless_short_intron_3n1/stopless_short_intron	4427 (36.44%)
stopless_short_intron_3n2/stopless_short_intron	4441 (36.55%)
stop-containing_short_intron_3n/short-containing_intron	15913 (35.52%)
stop-containing_short_intron_3n1/short-containing_intron	14449 (32.25%)
stop-containing_short_intron_3n2/short-containing_intron	14435 (32.22%)

BUSCO	
lineage_dataset	poales_odb12
n_markers	6282
BUSCO Assem	C:97.9[S:97.0,D:0.9],F:1.2,M:0.9,E:2.4
BUSCO Annot	C:97.6[S:84.6,D:13.1],F:0.4,M:2.0
Delta BUSCO	-0.3%

OMArk

OMA_clade	Oryza
num_conserved_hogs	15087
single	11686 (77.46%)
duplicated	2986 (19.79%)
duplicated_unexpected	2749 (18.22%)
duplicated_expected	237 (1.57%)
missing	415 (2.75%)
num_proteins_in_proteome	41044
total_consistent	34598 (84.29%)
consistent_partial_hits	1973 (4.81%)
consistent_fragmented	1977 (4.82%)
total_inconsistent	2456 (5.98%)
inconsistent_partial_hits	567 (1.38%)
inconsistent_fragmented	1541 (3.75%)
total_contaminants	0 (0.00%)
contaminants_partial_hits	0 (0.00%)
contaminants_fragmented	0 (0.00%)
total_unknown	3990 (9.72%)

PSAURON

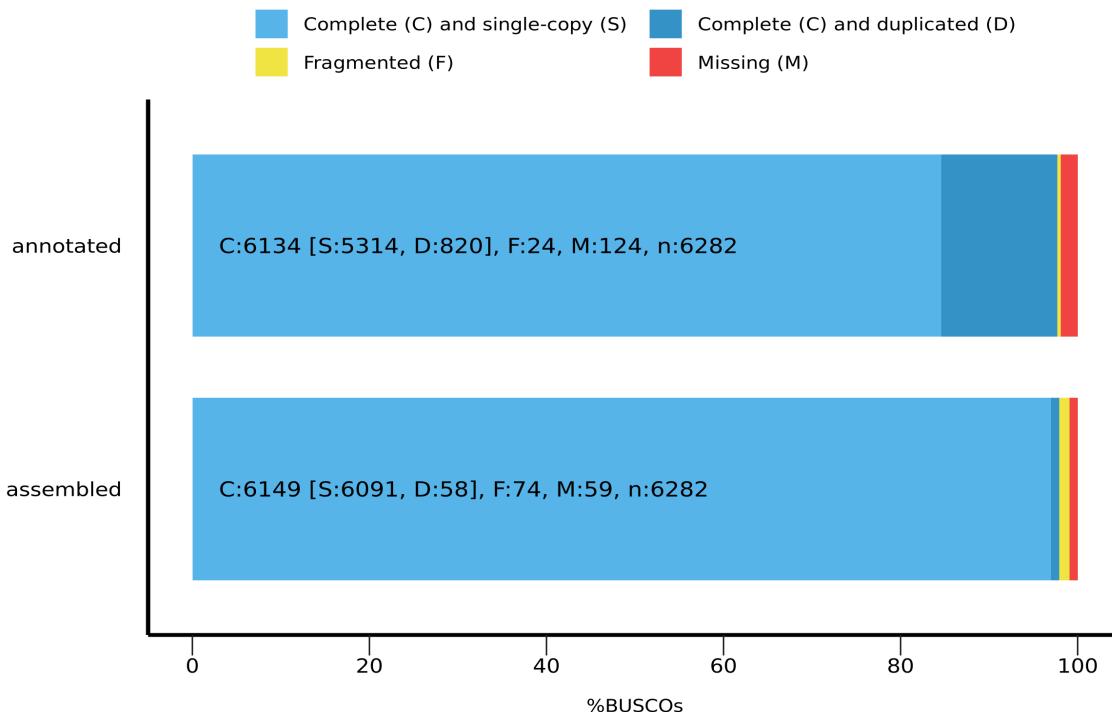
psauron_score	84.5
true_count	34676
false_count	6368

median_score	0.983815
max_score	1.0
min_score	0.00022

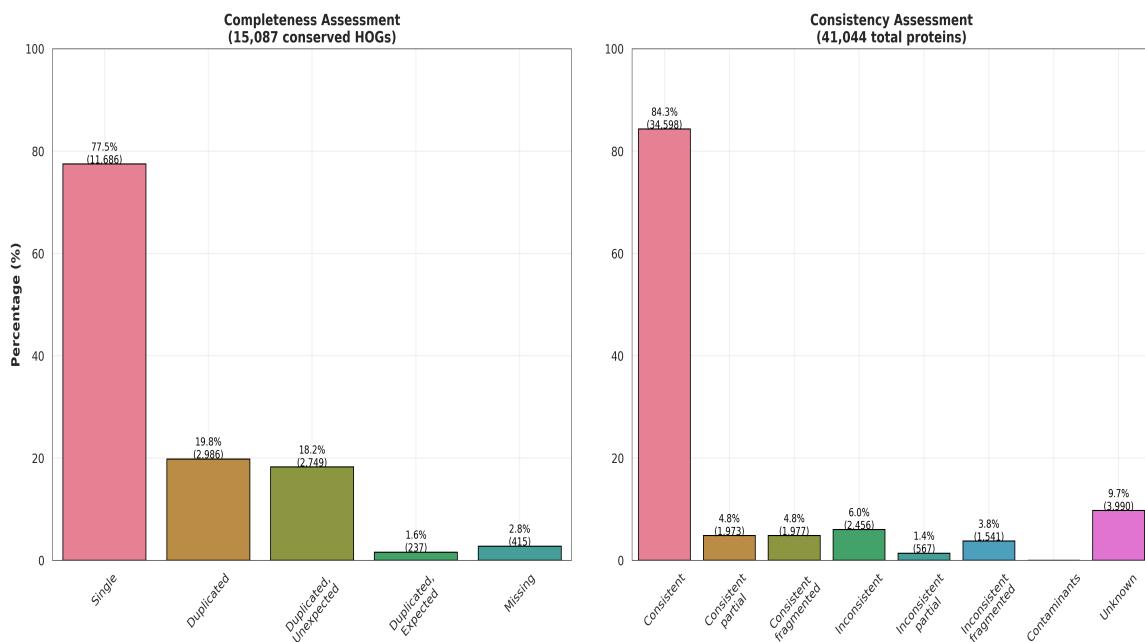
Best Reciprocal Hits	
num_best_reciprocal_hits	28630
num_split_gene_length_<0.75x	695 (2.43%)
num_fusion_gene_length_>1.25x	682 (2.38%)
KL_divergence_normalized	0.0098
JS_divergence_normalized	0.0021
Wasserstein_distance	2.771691

RNASeq	
mapping_rate	96.27%
primary_mapping_rate	95.83%
properly_paired	92.47%
num_gene_supported	26947 (74.05%)
num_exon_supported	152035 (86.86%)
num_exact_intron_boundary	107213 (80.02%)
num_intron_supported_canonical	107133 (99.93%)
num_intron_supported_non_canonical	80 (0.07%)

BUSCO Assessment Results



OMARK Assessment Result



Protein Distribution and Length Relationship by In-Frame Score

