Table 1 Summary of genome assemblies and gene annotations

Assemblies and annotations	Statistics
Assembly	
Contigs	1 112
Total bases	209 067 775
Average length	188 010.59
Minimum length	19 987
Maximum length	2 325 138
N_{50}	581 650
N (%)	0.00
GC (%)	44.25
Gene	
Genes	12 834
Average gene length	2 418.79
Average exon length	425.23
Repeat elements	132 699 478 (63.41%)
Gene coverage	14.83%
BUSco	98.3%; C:745 [S:367, D:378], F:0, M:13, n:758
Annotations	
Blast-hits	10 461 (81.51%)
Kyoto Encyclopedia of Genes and Genomes (KEGG)	7 375 (57.46%)
Gene Ontology (GO)	8 529 (66.46%)
No-hits	2 373 (18.49%)
Candidate secrete effector proteins (CSEP)	455
Secreted proteins (SP)	690
Nonclassical secreted proteins (NCSP)	169
Cytochrome (CYP)	37
Carbohydrate-active enzymes (CAZY)	519

Table caption: This table presents the comprehensive results of the genome assembly and annotation process. It shows the total assembled genome size at 209.07 MB, with 63.41% comprised of repetitive elements, primarily long terminal repeats. A total of 12 834 genes were predicted using a combination of evidence-based, ab initio, and consensus gene modeling approaches. The pie chart illustrates the proportion of genes that match known sequences in the GenBank NR database, highlighting the genetic connectivity with other organisms (Adapted from Pérez-García et al., 2009). The genomic assembly and annotation described by Pérez-García et al. (2009) offer profound insights into the genetic composition and complexity of the organism. The use of long-read sequencing technologies, combined with a haplotype-aware assembler, has allowed for a detailed and comprehensive view of the genome. The high percentage of repetitive sequences and the identification of a substantial number of genes are indicative of a complex genome, which could involve significant regulatory mechanisms and adaptability features. Such genomic data are invaluable for further genetic studies, potentially leading to discoveries related to gene function, evolutionary biology, and biotechnological applications. The alignment of a significant portion of the genes with those in the GenBank NR database also provides a basis for comparative genomics, enhancing our understanding of gene conservation and divergence among related species