Figure S1 Annotation statistics for genome assembly using SP as an example. A. Database source that were used for genome annotation. **B.** Distribution of all calculated sequence similarities by percentages. **C.** Distribution of the amount of Gene Ontology (GO) candidate terms assigned to each sequence. **D.** Distribution of the second level GO terms assigned to each sequence: biological process (P), molecular function (F), and cellular component (C). **E.** Statistics summary of the number of predicted genes that have been blasted and annotated.