

(B)

Seq1 : ATTTATATCAATTTGGAGGAAAAATATATAATATCCCGTTCTTATTTGTACATCTACCATGGACTTT  
 Seq2 : ATTGTTGTCAATTTGGACAAAAAAGGTGTAATATCCCATTCACGGACTTA

Seq1 : A--AAATA--CCGTCAGAGCTCAGTTTTTGTATTATTTAATTTATATTTTAACGGAATTTCATACACG  
 Seq2 : ATAAACAGTETGTCAGAGCTCAGTTTTTCGTATCAAAATTTATATTATAACGGA--TTTCATACACA

Seq1 : AATACACAGACAGAGTGACACTGGTATAATAATAATAAATGCTGATGATAATAATAATAAACTTCAA  
 Seq2 : AATACACAGGCAGAGTGACACTGATATAATAAATAAATAATGTTGATAATGAAAATAAT--CTTCAA

-----

Seq3 : TATTTGAGCCGCGTCACGACAAAACCAACATAATGGGTTTGGCACCAGCATGGATCCAGACCAGCCTG  
 Seq4 : TATTTGAGCCGCGTCACGACAAAACCAACATAATGGGTTTGGCACCAGCATGGATCCAGACCAGCCTG

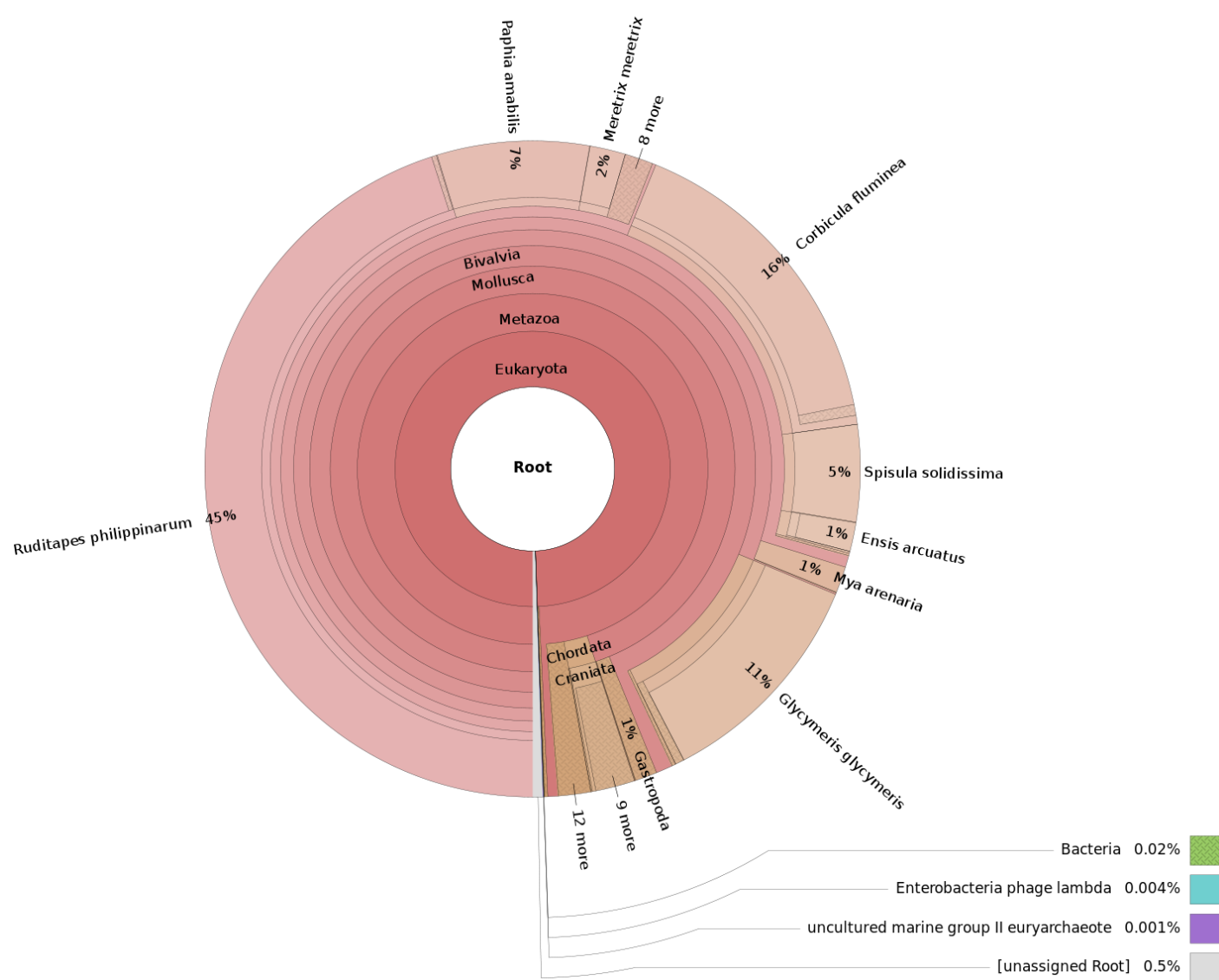
Seq3 : CGCCTCCGCGCAGTCTGATCAGGATCCATGCTGTTTCGCTATCAGTTTCTCTACTTC-----TA  
 Seq4 : CGCATCCGCGCAGTCTGATCAGGATCCATGTTGTTTCGCTTACAAACTCTATAACAAGTAGAGAAACTG

Seq3 : ATAC-----GGTTTGTA-----AG  
 Seq4 : ATACCGAACAGCATGGATCCTGATCAGTCTGCGCGGATGCGCAGGCTGGTCTGGATCCATGCTGGTGG

Seq3 : CAAACCCATTATGTTGGTTTTGTCGTGACGCGGCTCA  
 Seq4 : CAAACCCATTATGTTGGTTTTGTCGTGACGCGGCTCA

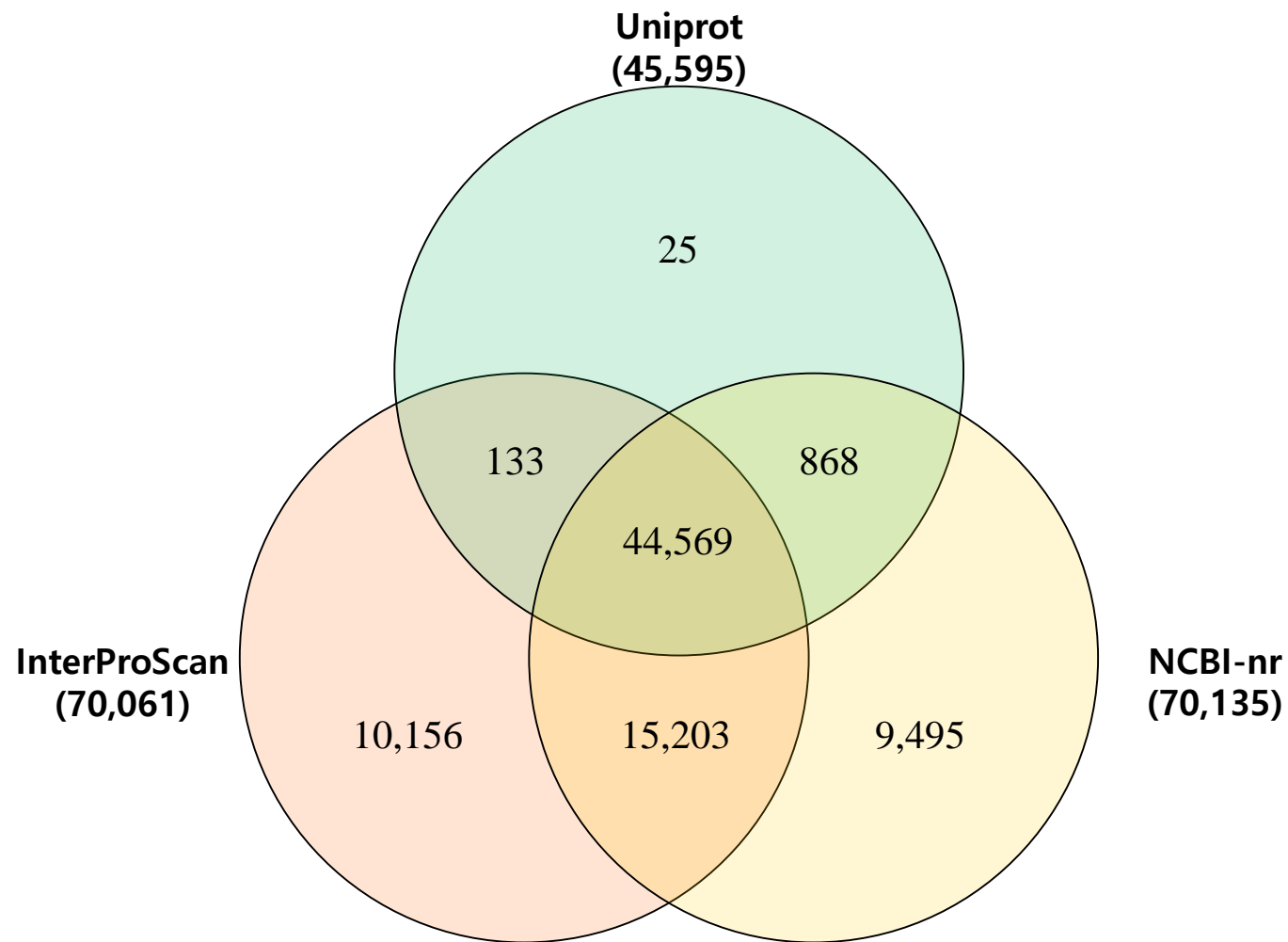
**Figure S1. Estimation of the *R. philippinarum* genome size based on 17-mer analysis.**

(A) The  $x$ -axis and the  $y$ -axis denotes the depth ( $X$ ) and the proportion, respectively, i.e., the frequency at that depth divided by the total frequency at all depths. (B) Example of Heterozygous features in the *R. philippinarum* genome. By performing sequence comparison using large-scale genome alignment tool (LASTZ), high-rate of heterozygosity between allele were defined. Red and blue letters indicate sequence difference and red boxes indicate micro-insertion or deletions between two allele.



**Figure S2. Krona chart for taxonomy profiling of *R. philippinarum*.**

All aligned contigs and scaffolds to NCBI non-redundant nucleotide database (nt database) were assigned to Krona. The abundance of the taxonomic groups corresponds to the percentage of spectra based on sequence identity.



**Figure S3. Summary of *R. philippinarum* gene prediction with three identifier (Uniprot, InterProScan, and NCBI-nr database).**  
The number of unique and shared gene models from three databases represent in the Venn diagram.

