Liu et al. – roles of 6mA in encystment

Supplementary Materials:

Supplemental Figures S1-S8

Figure S1

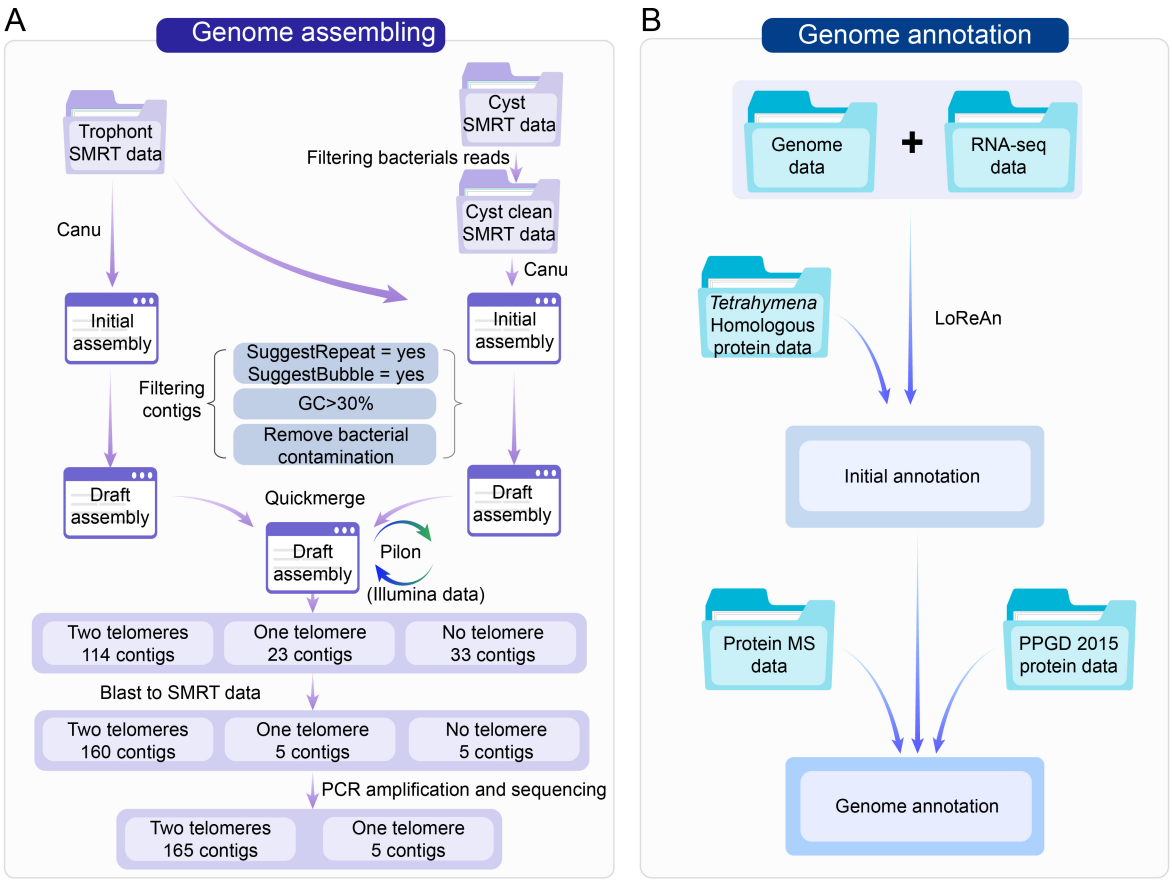

Figure S1. Assembly and annotation workflow for high-quality *P. persalinus* genome assembly

(A) The assembly workflow of the *P. persalinus* genome.

(B) The annotation workflow of the *P. persalinus* genome.

Figure S2

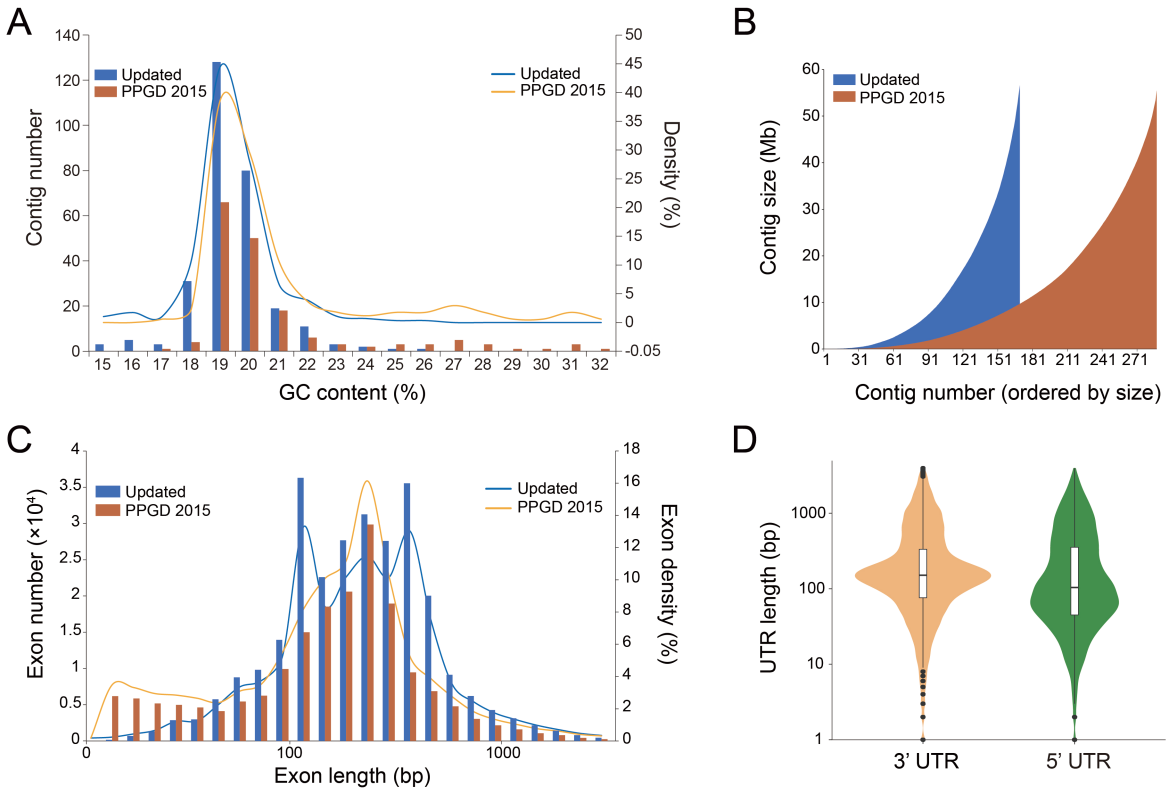

Figure S2. Comparative analysis of two *P. persalinus* genome assemblies

(A) Comparison of the GC content between the updated genome assembly and PPGD 2015. Lines represent contig density and columns contig number.

(B) Size distribution of contigs in the updated genome assembly and PPGD 2015. The x-axis represents the contigs number ordered by size and the y-axis the sum of contigs.

(C) Length distribution of predicted exons in the updated genome assembly and PPGD 2015. The x-axis represents the length of exons and the y-axis the number/density of exons. Lines represent the gene density and columns the gene number.

(D) Length distribution of 3' and 5' UTRs in the updated genome assembly.

Figure S3

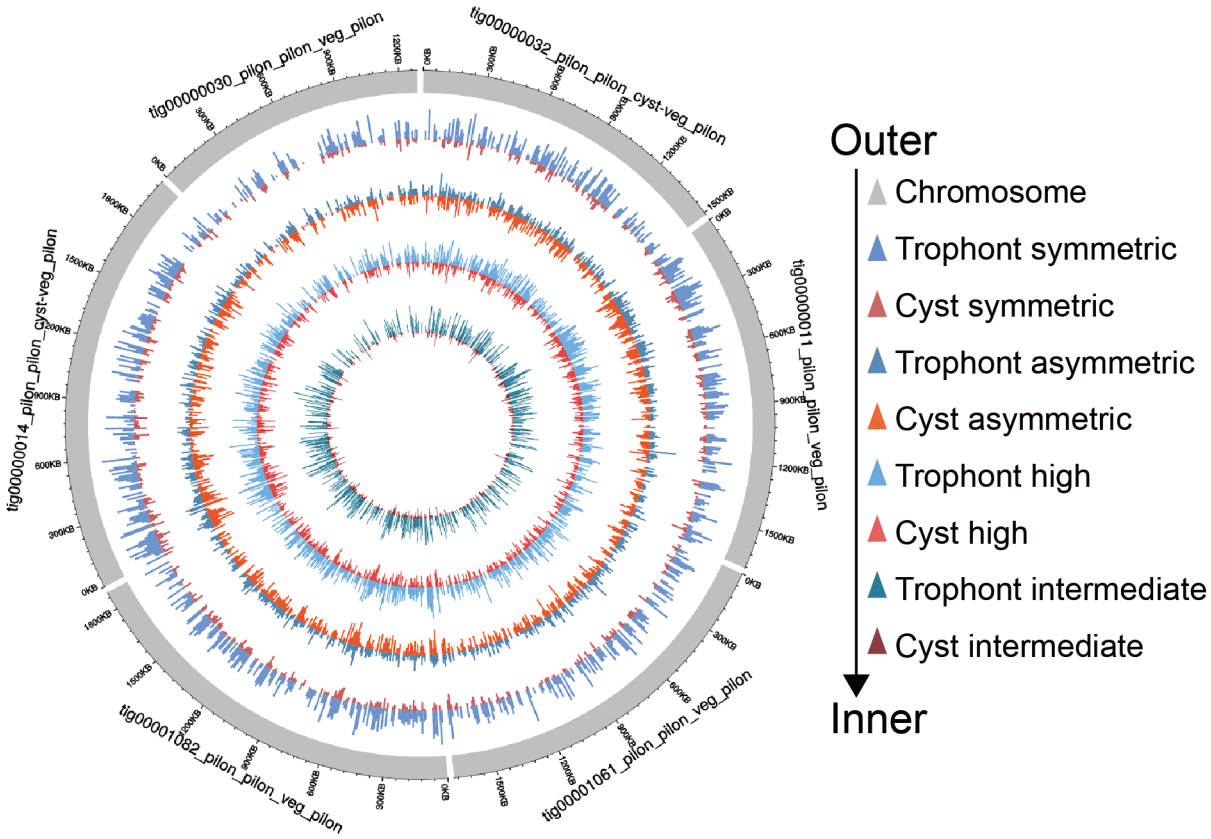

Figure S3. 6mA Circos plot

Circos plots of 6mA distribution across the 6 largest contigs. Nine rings from outer to inner represent chromosomes, distribution of 6mA in different motifs (symmetric, asymmetric) and different methylation levels (high 80–100%, intermediate 20–80%), as denoted on the right side.

Figure S4

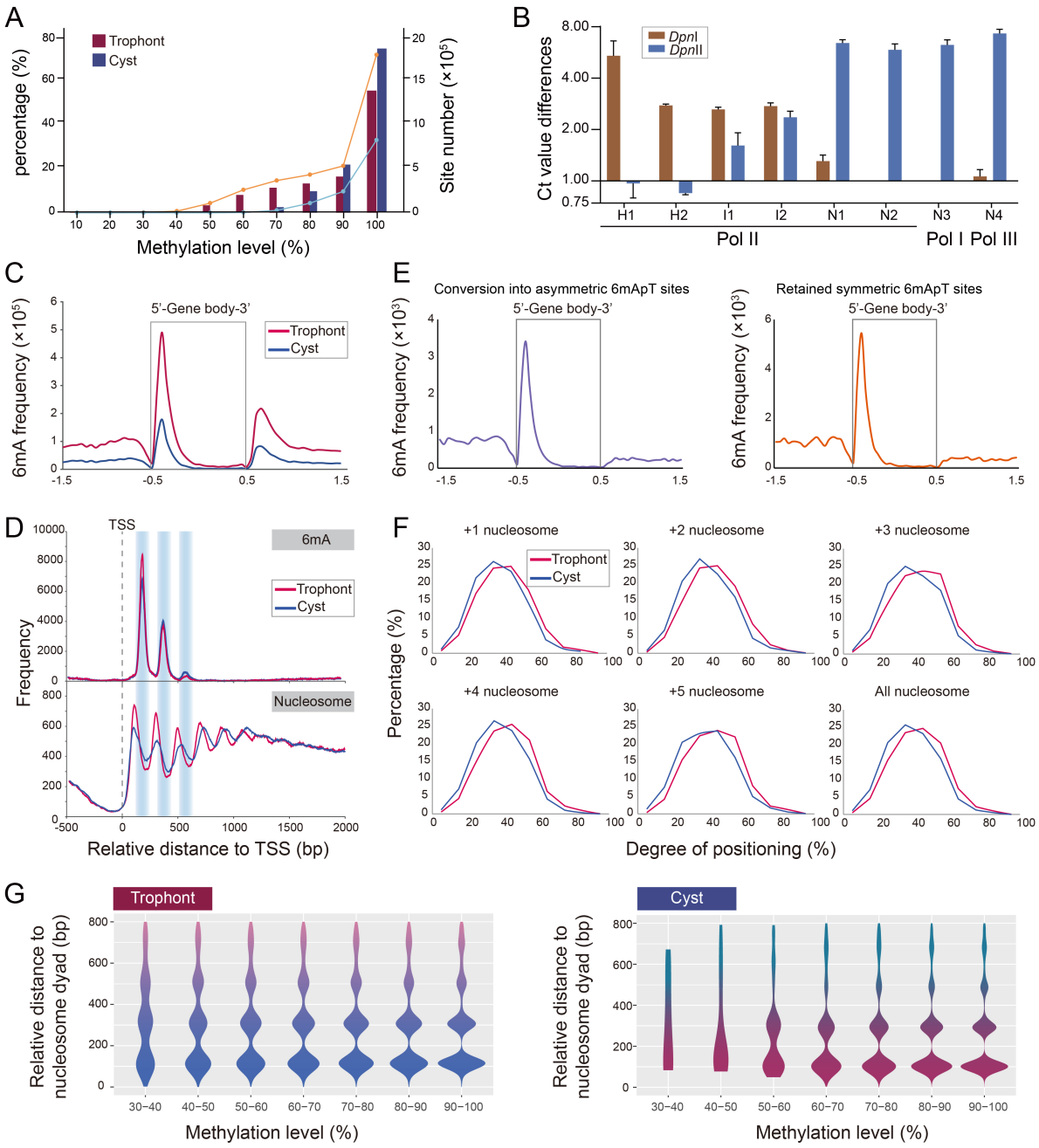

Figure S4. 6mA and nucleosome characteristics in the *P. persalinus* genome

(A) Comparison of 6mA methylation levels in trophonts and cysts. Methylation levels were ranked from low to high and divided into 10 quantiles. Lines represent the percentage of 6mA and columns the number of 6mA sites. The percentage of 6mA was calculated as the number of 6mA sites in each quantile divided by the total number of 6mA sites.

(B) Validation of the methylation status of eight GATC sites by *Dpn*I/*Dpn*II digestion coupled with qPCR analysis. The sites were selected based on their presence on genes transcribed by different RNA

polymerases (Pol I, II and III), and their methylation levels were calculated using SMRT sequencing data (H1-H2: high methylation; I1-I2: intermediate methylation; N1-N2: no methylation; all these are transcribed by Pol II. N3: no methylation transcribed by Pol I; N4: no methylation transcribed by Pol III.). See Materials and Methods for details.

(C) Composite analysis of 6mA distribution on the gene body of trophonts and cysts. Genes were scaled to unit length and were extended to each side by one unit. Distribution frequency was calculated as "6mA amount at a certain position/total 6mA amount".

(D) Distribution profiles of 6mA (top panel) and nucleosomes (bottom panel) around TSS (transcription start site of the 5' UTR) in trophonts and cysts.

(E) Composite analysis of symmetric and asymmetric 6mApT sites on the gene body of cysts. Genes were scaled to unit length and extended to each side by one unit. The distribution frequency was calculated as "6mA amount at a certain position/total 6mA amount".

(F) Nucleosome positioning degree (+1, +2, +3, +4, +5, and all nucleosomes) in trophonts (red) and cysts (blue). Note that the proportion of nucleosomes with high positioning degree in trophonts is higher than that in cysts. The x-axis represents the degree of nucleosome positioning, defined as "the number of fragment centers within ±20 bp of the called nucleosome dyad, relative to the number of all fragment centers within the 147 bp called nucleosome footprint" (Xiong et al. 2016).

(G) 6mA distribution relative to the nucleosome dyad in trophonts (left panel) and cysts (right panel). The violin plots show the density of 6mA between neighboring nucleosome dyads, grouped by methylation levels. Lowly methylated 6mA (0–30%) was excluded due to their scarcity.

Figure S5

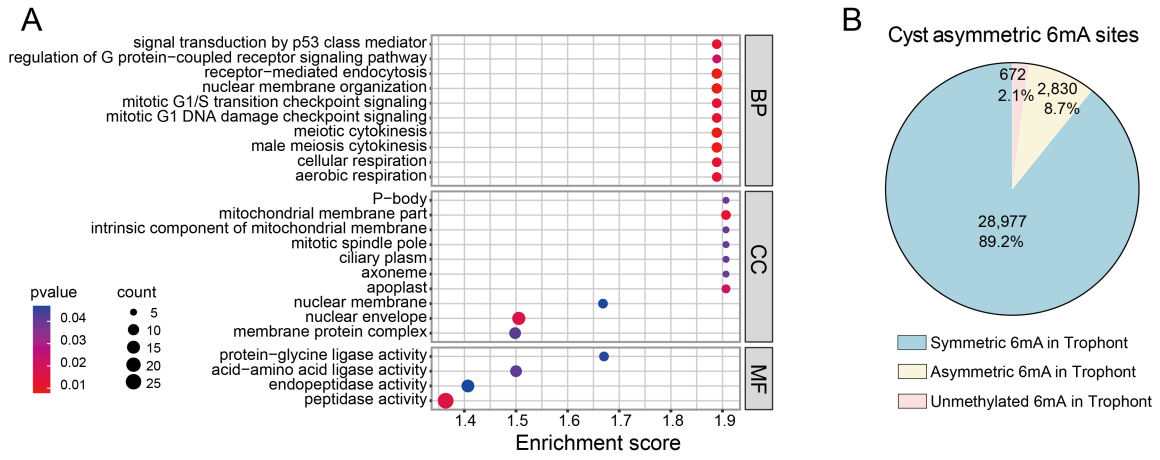

Figure S5. Analysis of differentially expressed genes associated with 6mA

(A) Gene Ontology (GO) analysis of 8,526 DEGs in Figure 4D. All GO terms in molecular function (MF) and cellular component (CC) are shown. The top 10 most significantly enriched GO terms in biological process (BP) are shown.

(B) 6mA pattern conversion of cyst asymmetric 6mApT in 8,526 DEGs in Figure 4D. Three sources of asymmetric 6mApT are shown.

Figure S6

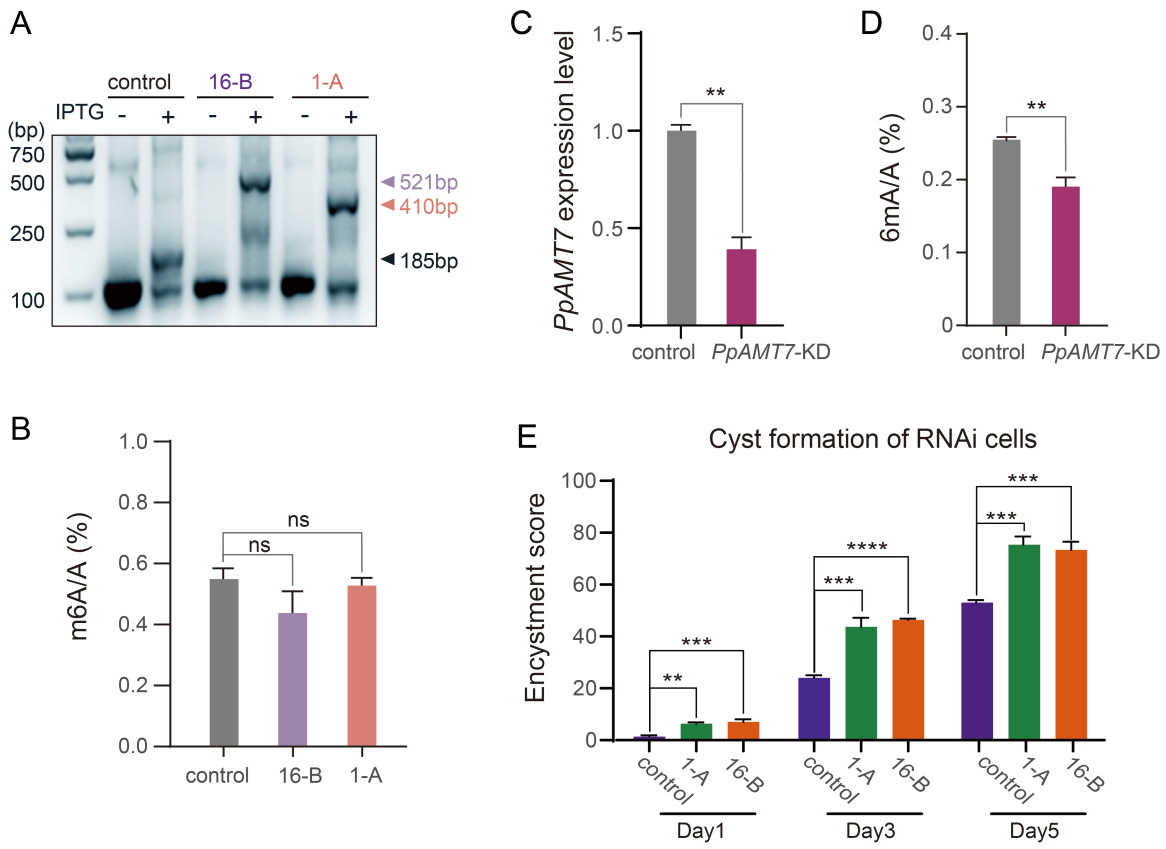

Figure S6. Statistical analysis of *PpAMT7* and *PpAMT1* RNAi

(A) Production of dsRNAs from silencing constructs visualized on a 2% agarose gel, with and without the addition of IPTG. Arrows indicate target dsRNAs generated from the plasmids. The expected sizes are the following: *PpAMT1*-KD 1-A dsRNA is 423bp, 16-B dsRNA is 521bp, and the 185 bp dsRNA is generated from the multiple cloning site (MCS) of the L4440 empty vector.

(B) Mass spectrometry analysis of RNA m6A in control and *PpAMT1*-KD cells after 9 days of feeding. m6A ratios (m6A/A, %) was defined as the number of methylated adenine sites divided by the total number of adenine sites. ns $P > 0.05$ (Student's *t*-test).

(C) Expression levels of *PpAMT7* in control and *PpAMT7*-KD cells after 9 days of feeding. Expression levels were assessed by normalized quantitative RT-PCR data. ** $P < 0.01$ (Student's *t*-test).

(D) Mass spectrometry analysis of 6mA in control and *PpAMT7*-KD cells after 9 days of feeding. 6mA ratio (6mA/A, %) was defined as the number of methylated adenine sites divided by the total number of adenine sites. ** $P < 0.01$ (Student's *t*-test).

(E) Statistical analysis of the encystment scores in Figure 6H (The details of the statistical data are shown in Supplementary Table S3). The proportion of cysts in individual wells was assigned a score for

statistical analysis (0 for trophonts, 1 for 0% < cyst < 50%, and 2 for 50% < cyst < 100%). ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$ (Student's *t*-test).

Figure S7

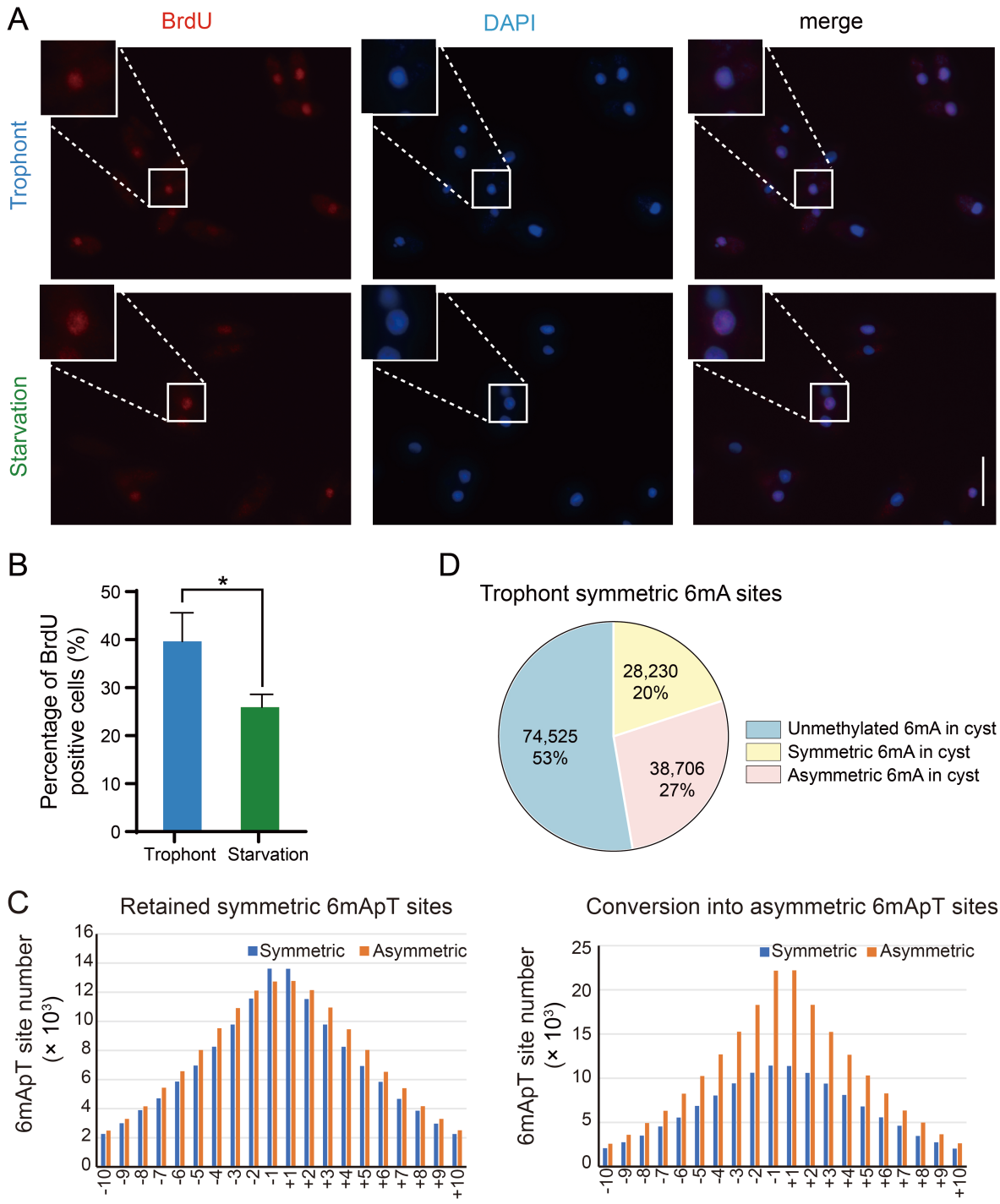

Figure S7. Pattern conversion of symmetric 6mApT sites and BrdU labeling of starved cells

(**A**) Representative immunofluorescence staining images of trophonts and starved cells labeled by BrdU. DAPI, DNA stain; BrdU, signals of cell replication (S phase). Scale bar = 50 μm.

(**B**) Statistical analysis of BrdU-positive cells as presented in (A). Cell images were randomly selected (n = 100). * $P < 0.05$ (Student's *t*-test).

(**C**) Frequencies of symmetric and asymmetric 6mApT sites near the site where 6mApT was either retained in a symmetric methylation state or converted into asymmetric methylation in cysts. The x-axis represents ten 6mApT sites found upstream and downstream (1kb) of the methylation site. The y-axis indicates the number of 6mApT sites.

(**D**) Three outcomes of symmetric 6mApT sites in trophonts after encystment.

Figure S8

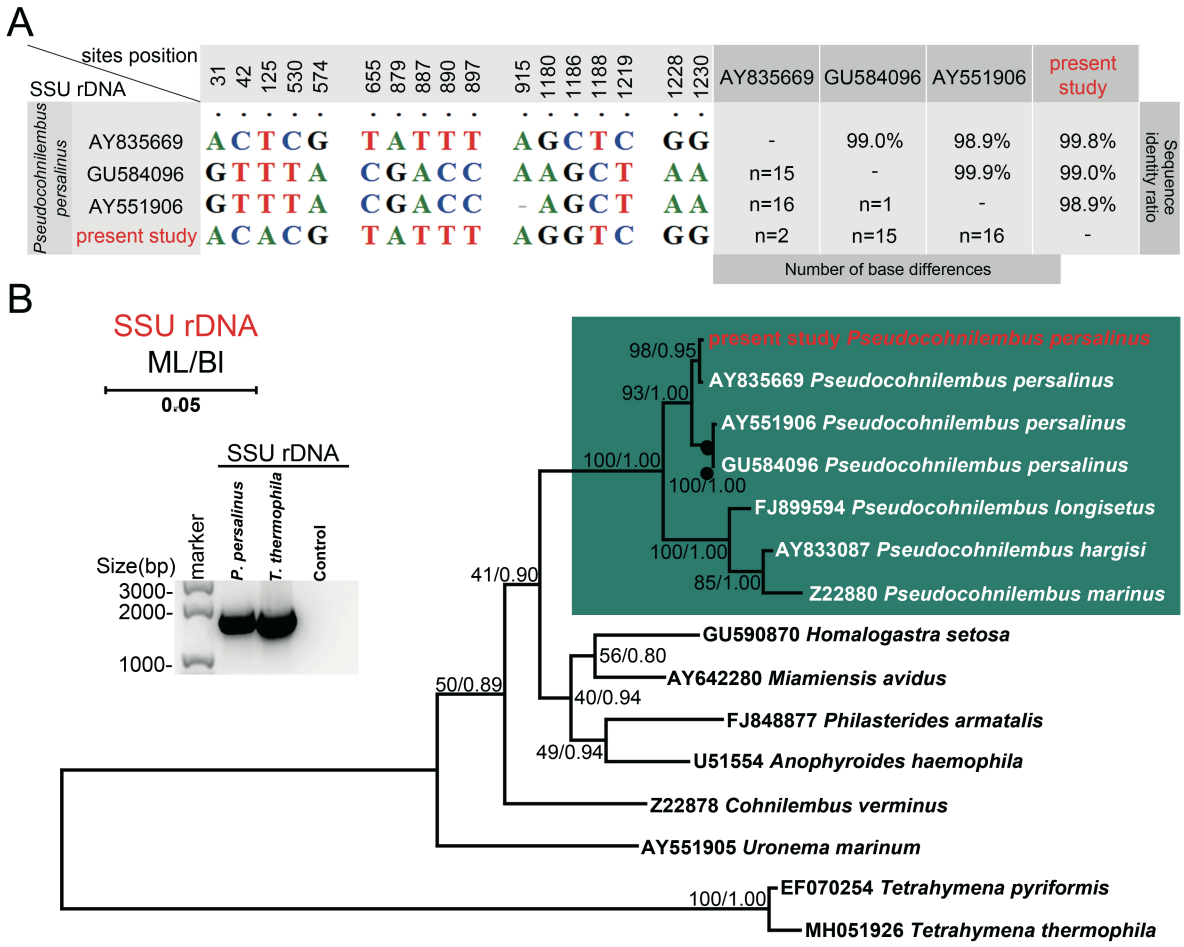

Figure S8. SSU rDNA sequence and phylogenetic identification of *P. persalinus*

(A) SSU rDNA sequence comparison of several *P. persalinus* species and closely related species. Nucleotide positions are given at the top of each column. Insertions and deletions are compensated by introducing alignment gaps (–). The number of nucleotide differences and sequence similarity among these four species are shown. n, number of nucleotide differences.

(B) Maximum likelihood (ML) tree inferred from SSU rDNA sequences of 15 taxa. Numbers near nodes are bootstrap values for ML and posterior probability values for Bayesian inference (BI) trees. PCR amplification products of *P. persalinus* and *T. thermophila* SSU rDNA was evaluated on a 1% agarose gel.

References

Xiong J, Gao S, Dui W, Yang W, Chen X, Taverna SD, Pearlman RE, Ashlock W, Miao W, Liu Y. 2016. Dissecting relative contributions of *cis*- and *trans*-determinants to nucleosome distribution by comparing *Tetrahymena* macronuclear and micronuclear chromatin. *Nucleic Acids Res* **44**: 10091–10105.