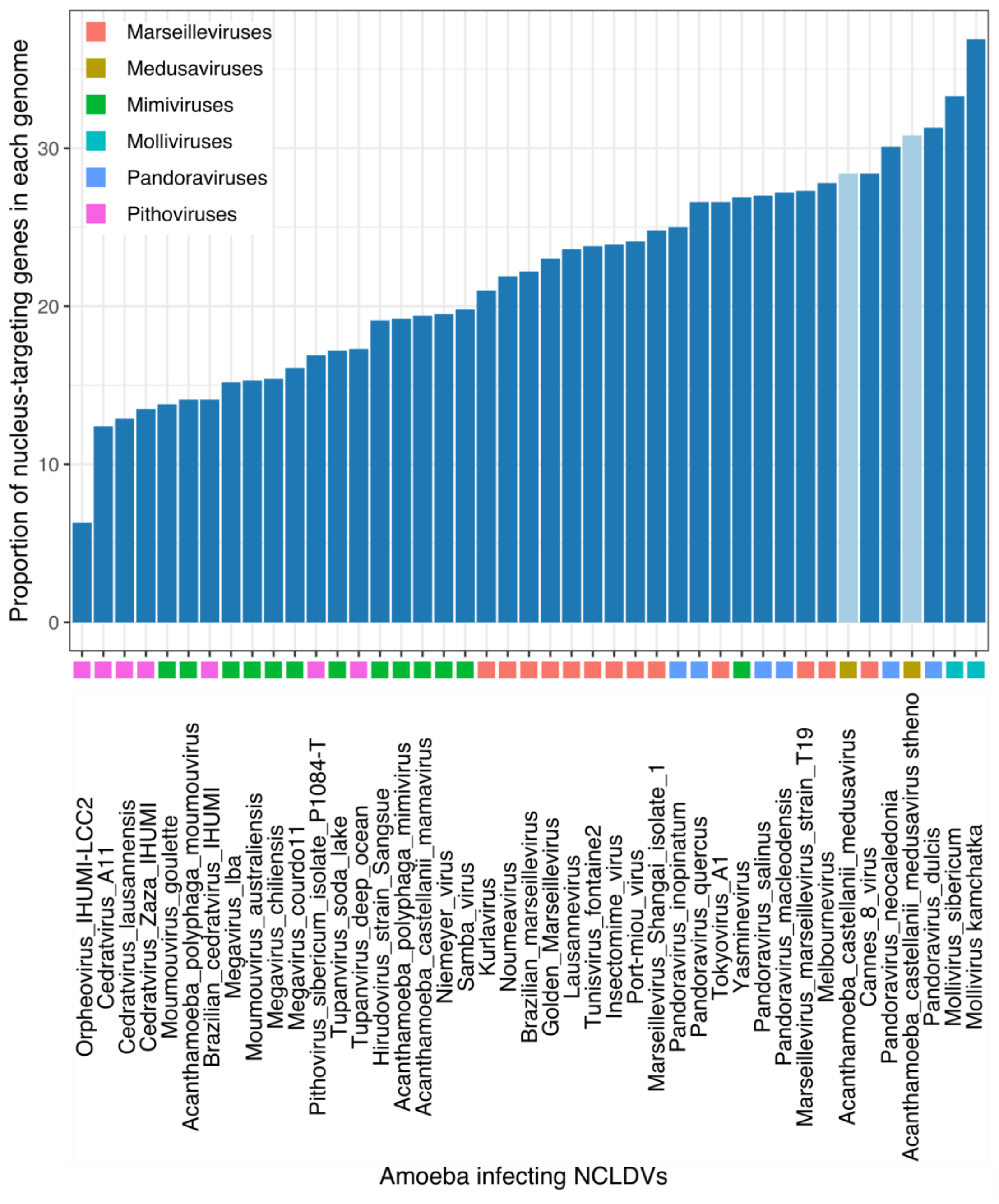
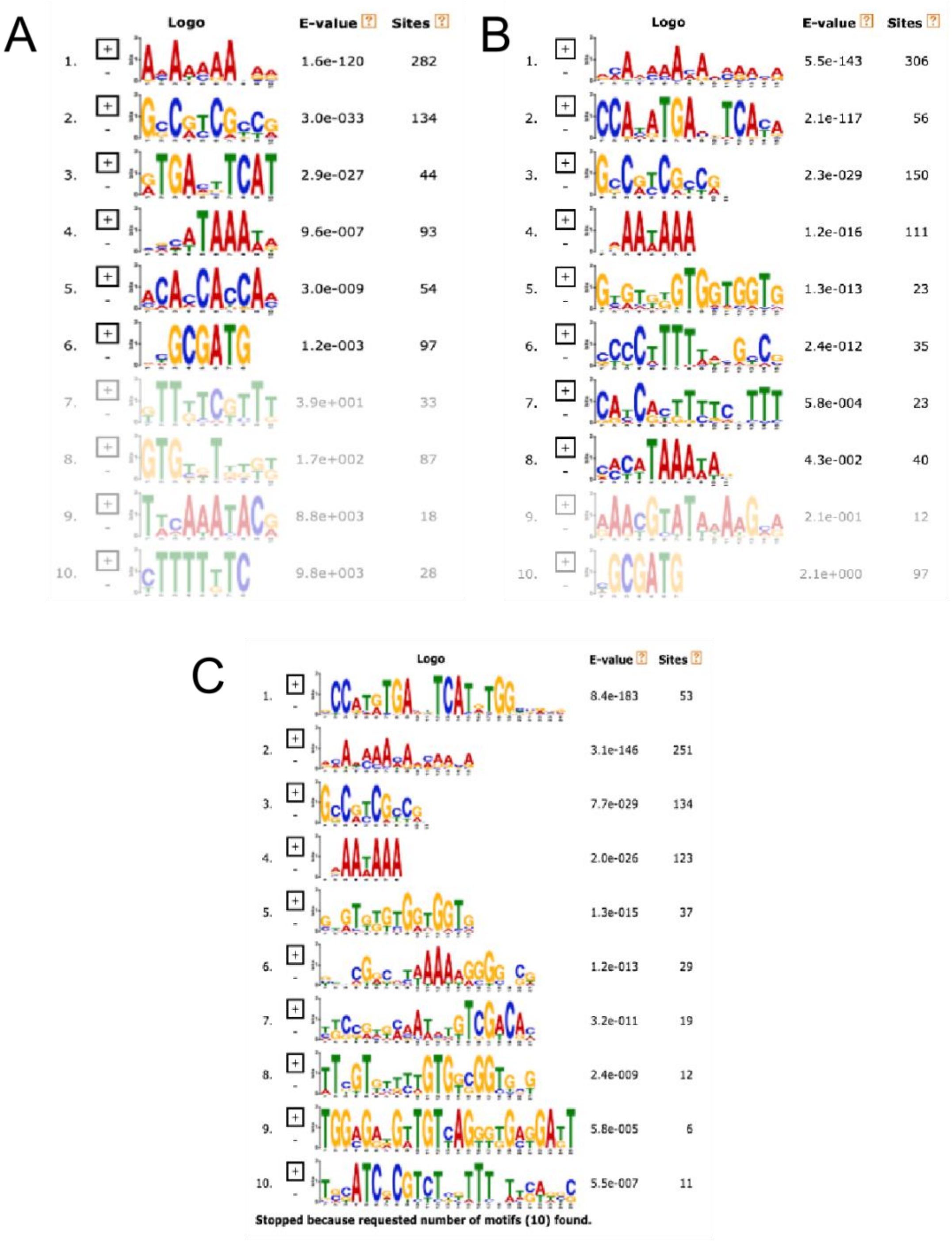


Fig. S1 Deeploc prediction results of other amoeba infecting NCLDVs.



A

	Logo	E-value	Sites
1.		1.6e-120	282
2.		3.0e-033	134
3.		2.9e-027	44
4.		9.6e-007	93
5.		3.0e-009	54
6.		1.2e-003	97
7.		3.9e+001	33
8.		1.7e+002	87
9.		8.8e+003	18
10.		9.8e+003	28

B

	Logo	E-value	Sites
1.		5.5e-143	306
2.		2.1e-117	56
3.		2.3e-029	150
4.		1.2e-016	111
5.		1.3e-013	23
6.		2.4e-012	35
7.		5.8e-004	23
8.		4.3e-002	40
9.		2.1e-001	12
10.		2.1e+000	97

C

	Logo	E-value	Sites
1.		8.4e-183	53
2.		3.1e-146	251
3.		7.7e-029	134
4.		2.0e-026	123
5.		1.3e-015	37
6.		1.2e-013	29
7.		3.2e-011	19
8.		2.4e-009	12
9.		5.8e-005	6
10.		5.5e-007	11

Stopped because requested number of motifs (10) found.

Fig. S2 Promoter motif detected by MEME with motif width ranging from (A) 8-10, (B) 6-15, and (C) 8-25 bp. Motif sequences are ordered according to their e-value.

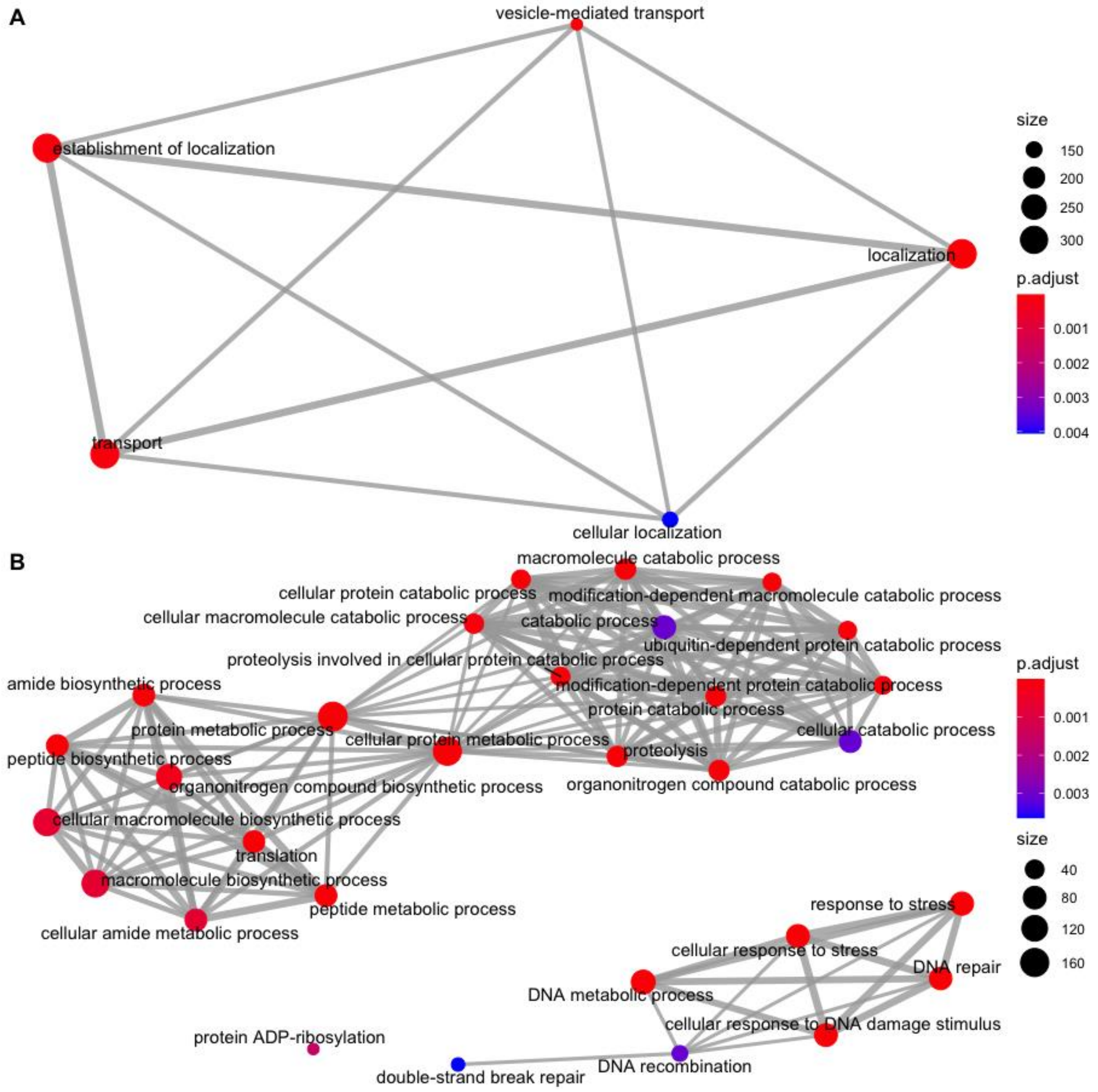


Fig. S3 Enriched map of enriched GO terms. A. GO terms in cluster 1; B. GO terms in cluster 2

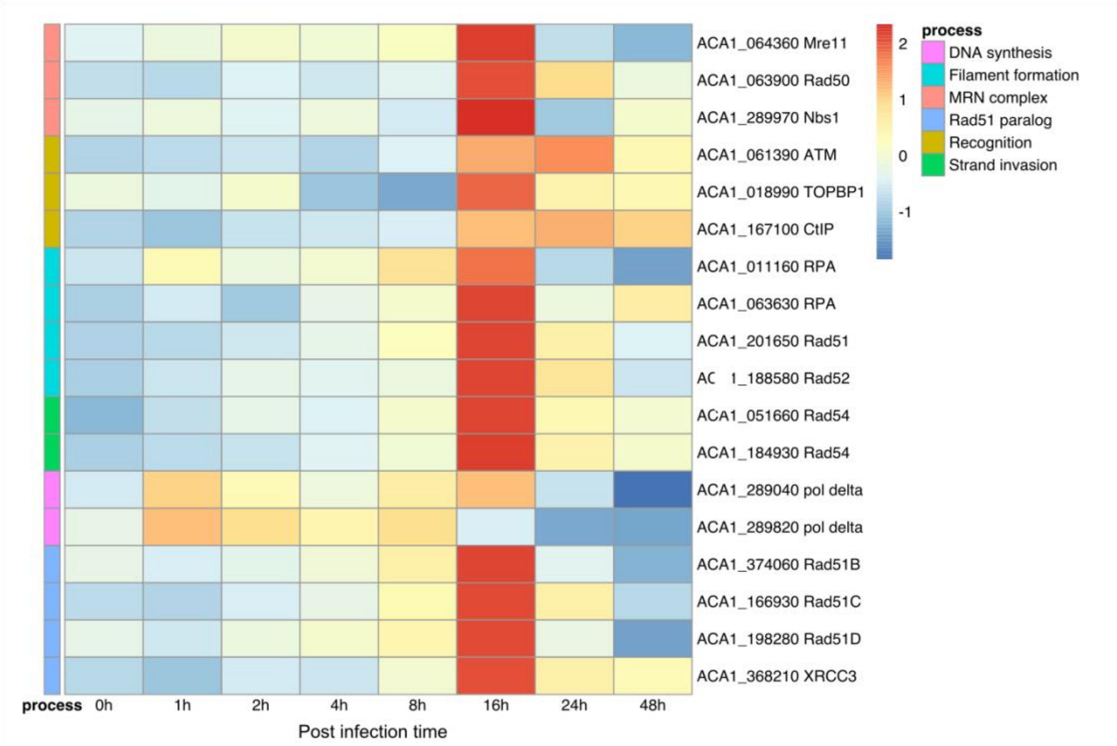


Fig. S4 Expression profile of genes related to homologous recombination repair (KEGG pathway accession number: acan03440). The color scale indicates z-score scaled RPKM values. The process bar on the left indicates the process associated with each gene. Gene names and locus tags are on the right side.

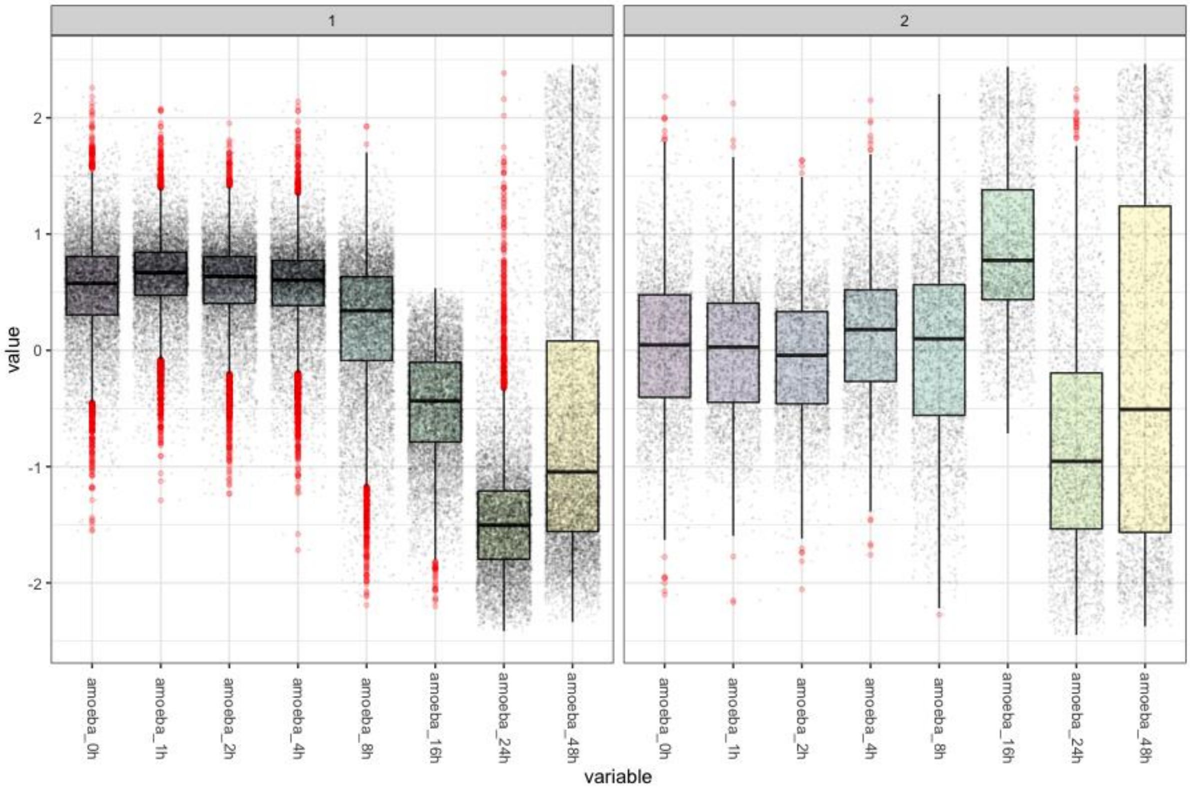


Fig. S5 Expression profile of 2 identified clusters at the end of the experiment.

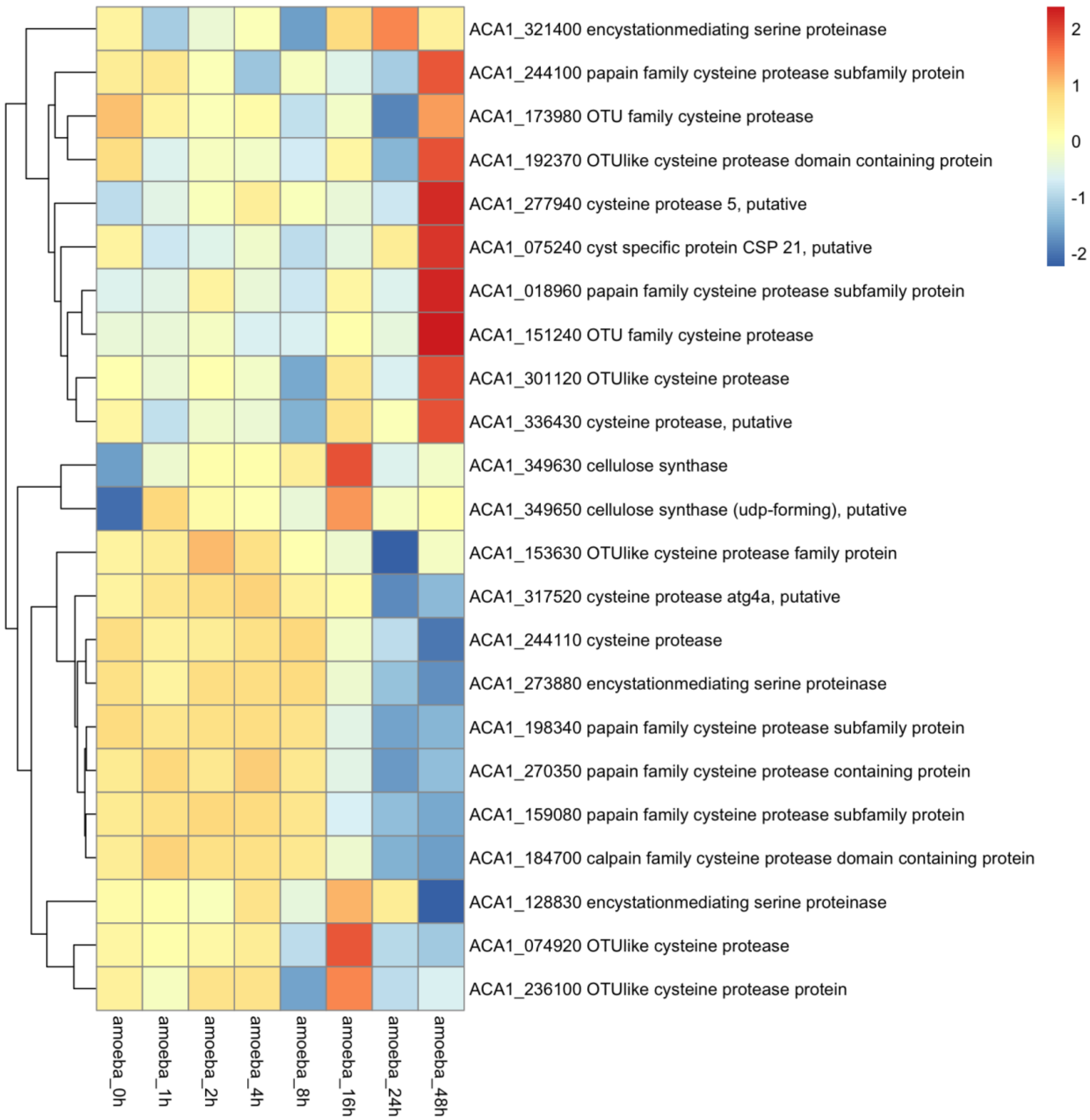


Fig. S6 Expression profile of genes related to encystment (gene name containing cysteine protease, CSP, encystation, or cellulose synthase). The color scale indicates z-score scaled RPKM values. The process bar on the left indicates the process associated with each gene. Gene names and locus tags are on the right side.