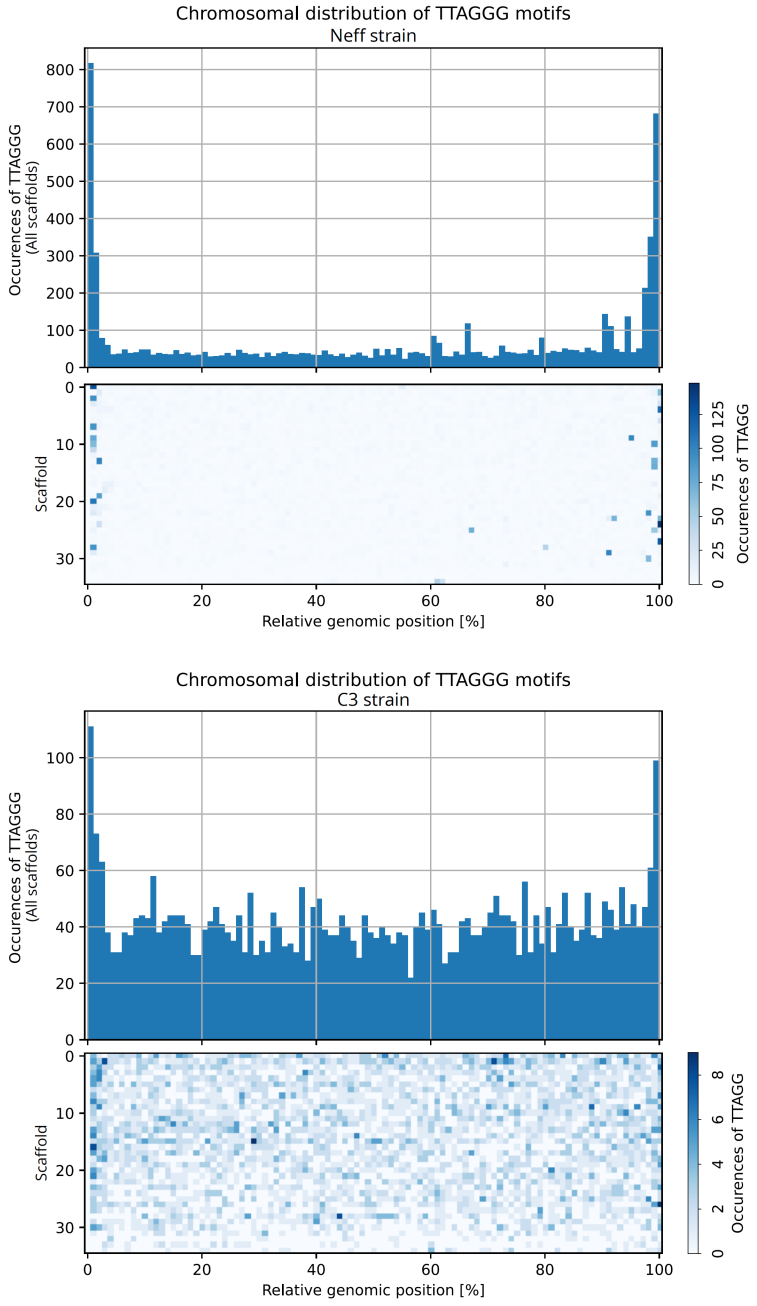
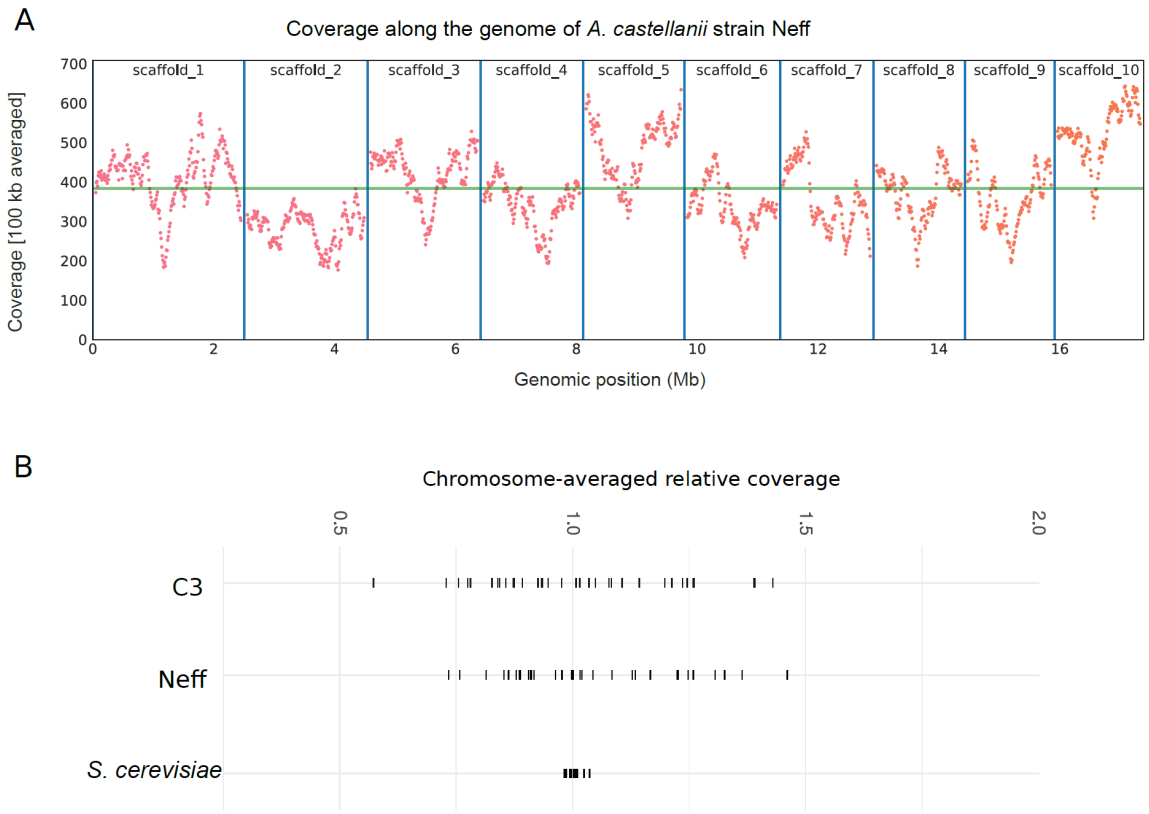
Supplemental Material



Supplementary Figure S1. Genomic distribution of subtelomeric repeats along *Acanthamoeba castellanii* scaffolds.

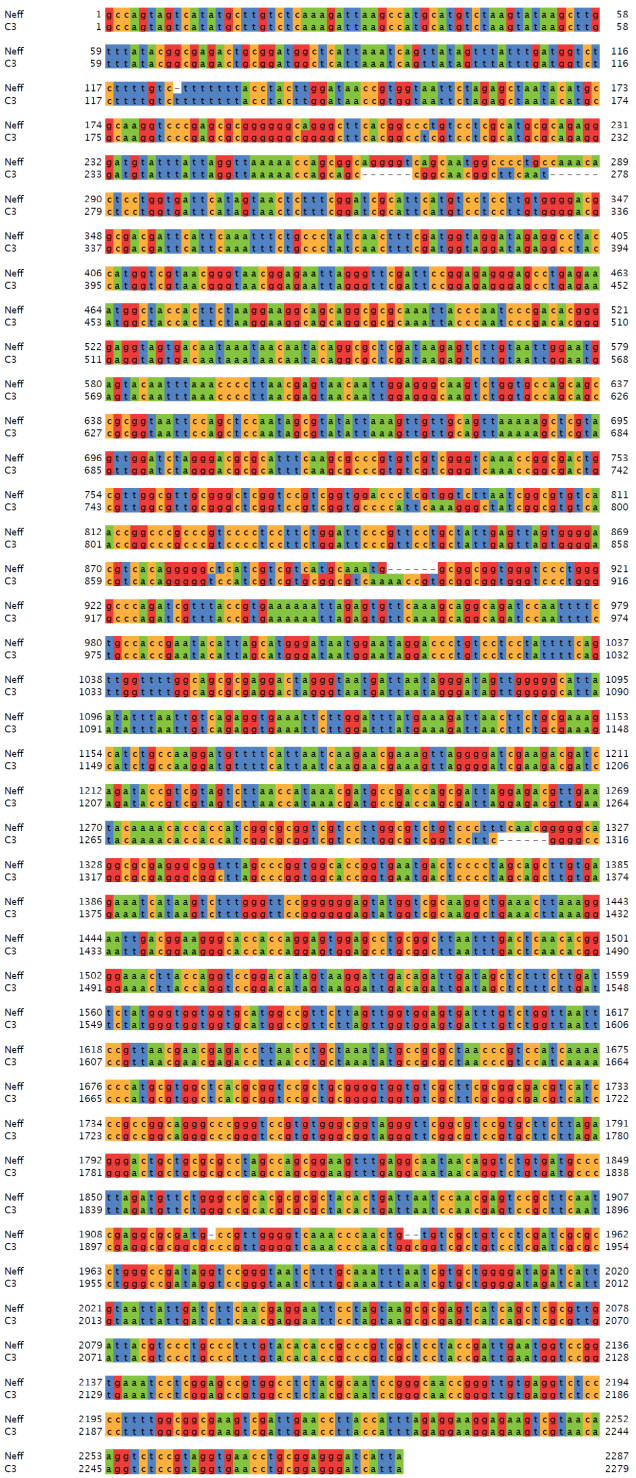
Relative position of "TTAGGG " subtelomeric repeats, from the beginning (0%) to the end (100%) of scaffolds. In *A. castellanii* strains Neff (top) and C3 (bottom). The histograms show the total repeat density summed across the 35 largest scaffolds, while the heatmaps show the distribution for each scaffold on separate rows. Both the heatmap and histogram are binned at 1% of relative scaffold length.



Supplementary Figure S2. Read coverage across scaffolds of *A. castellanii*.

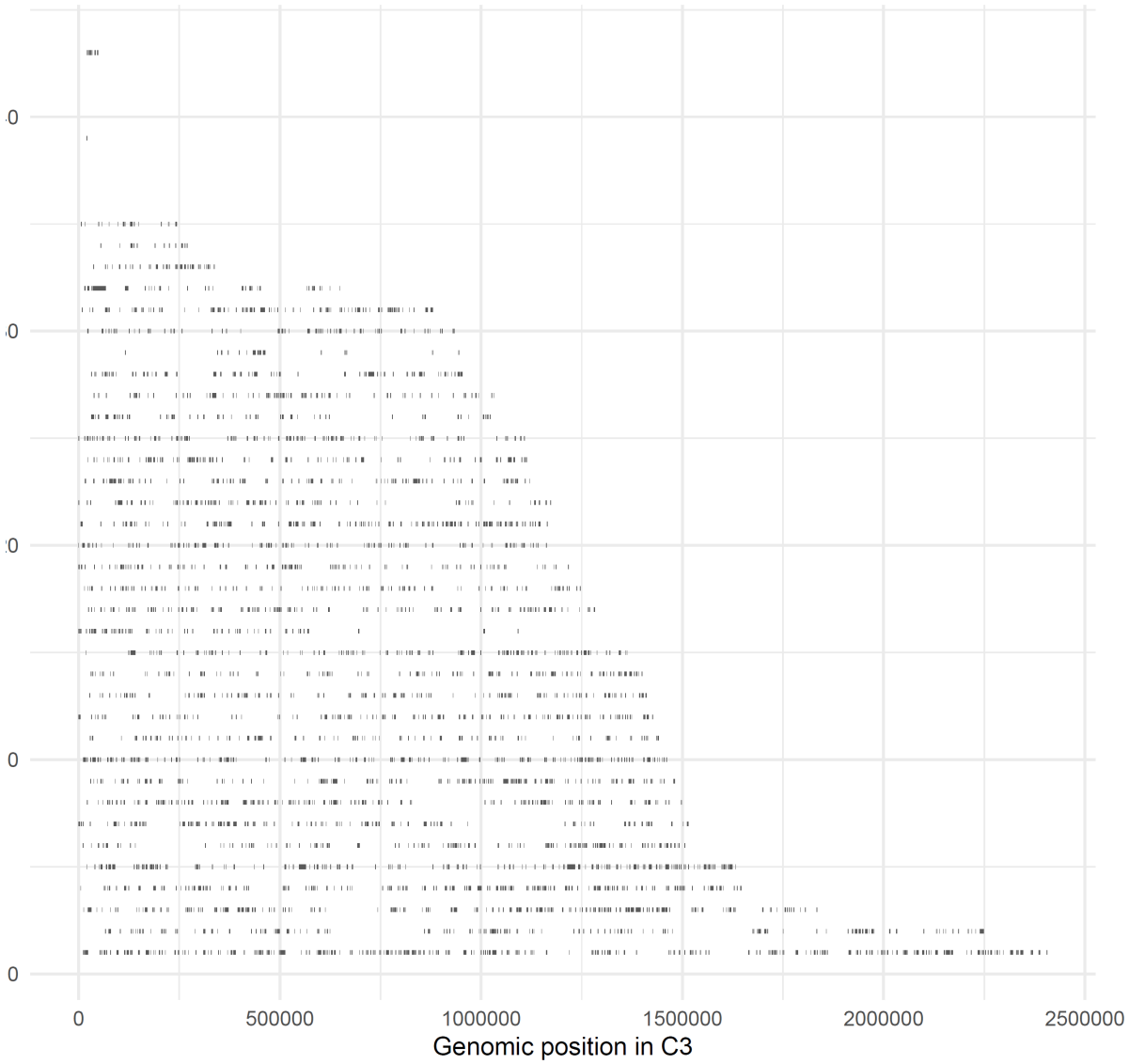
A, Illumina short-reads coverage along the 10 largest scaffolds of *A. castellanii* Neff in a 100 kb sliding window, with the horizontal green line showing genome median coverage.

B, Variability of median coverage per chromosome (relative to genome median) for *A. castellanii* strains C3 and Neff, and asynchronous *Saccharomyces cerevisiae* haploid strain BY4741. For *S. cerevisiae,* library SRR1569870 was used.



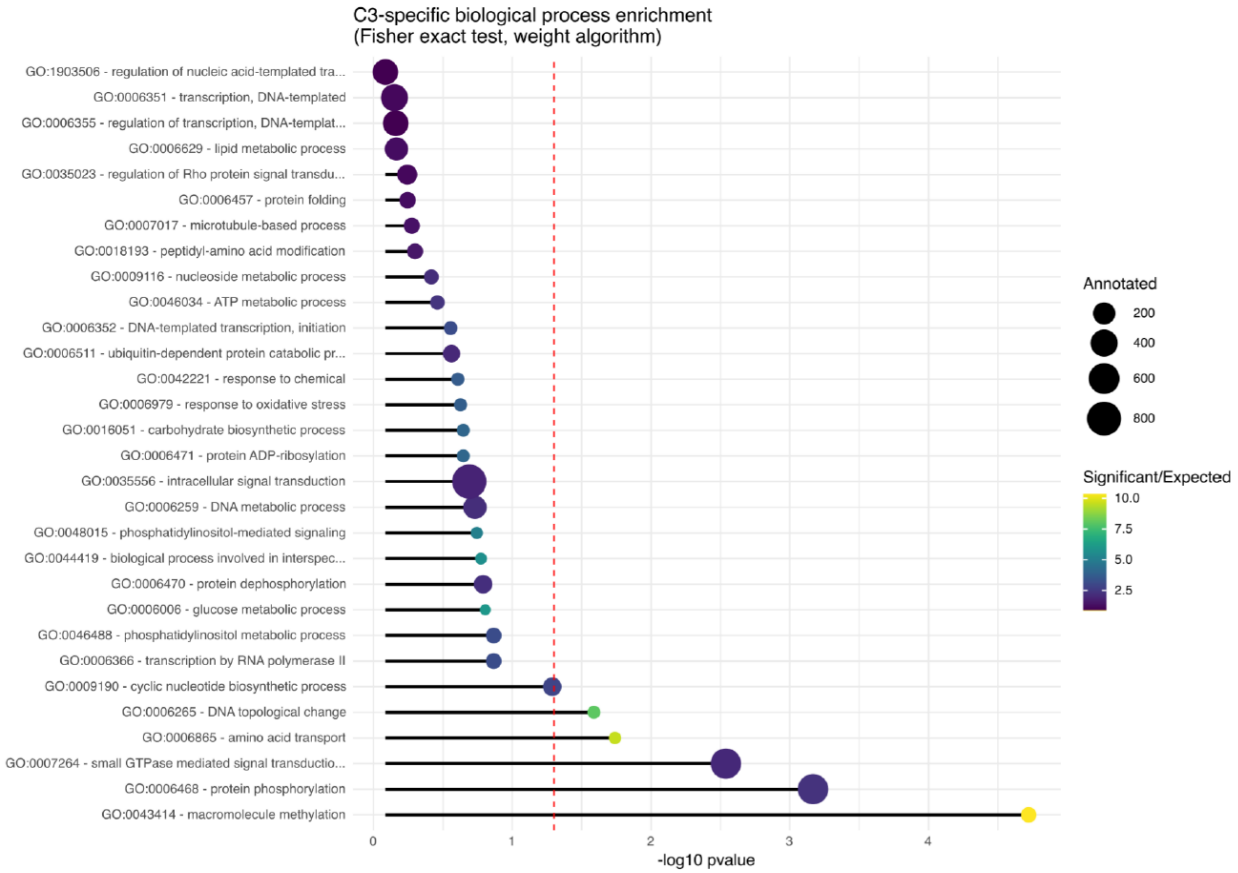
Supplementary Figure S3. Comparison of 18S rDNA sequences from *Acanthamoeba castellanii* C3 and Nef strains.

The *A. castellanii* 18S sequence from NCBI was used to retrieve the 18S rDNA sequences from the C3 and Neff assemblies. These retrieved sequences were aligned using MAFFT with auto setting (v7.4, (Katoh and Standley 2013) and visualized in Jalview (v2.1.1.3, (Waterhouse et al. 2009)).



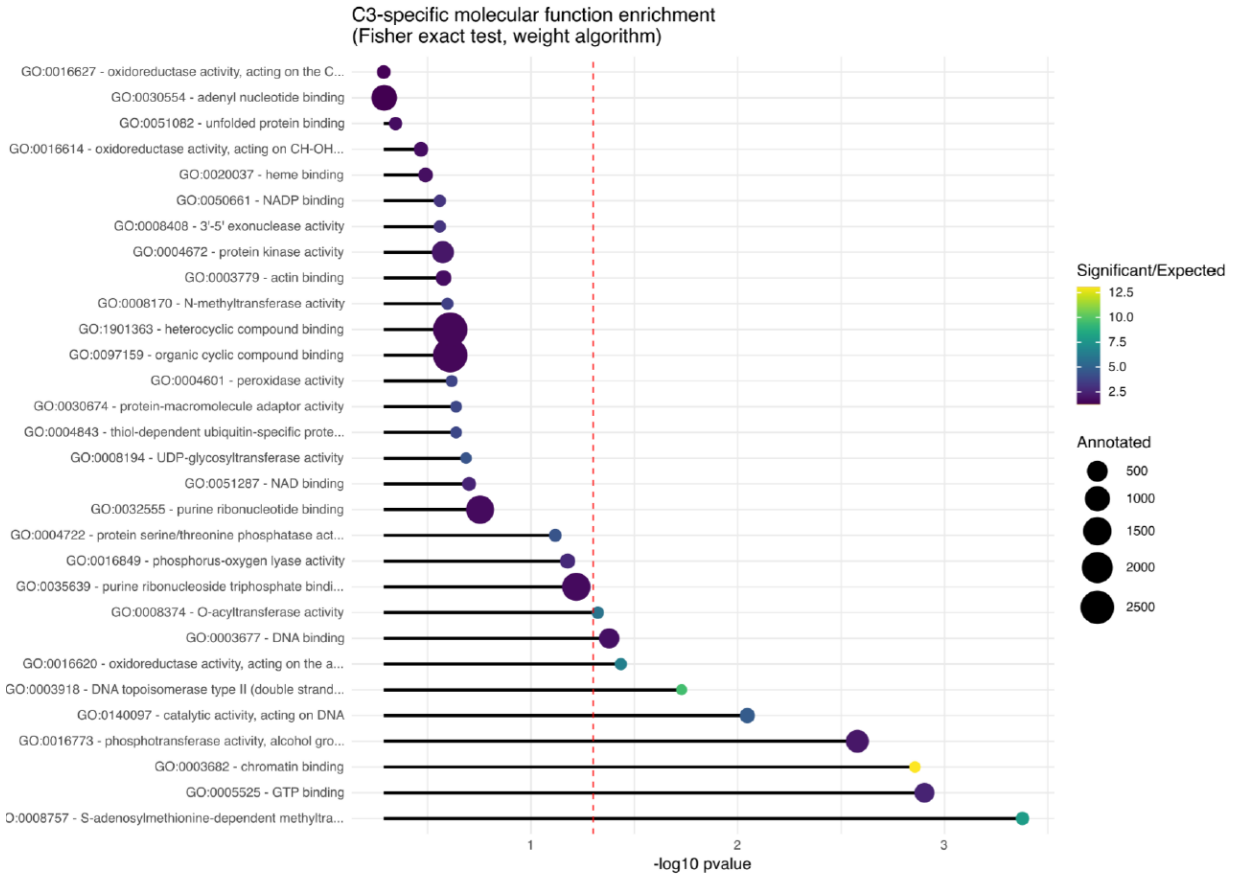
Supplementary Figure S4. Genomic distribution of *Acanthamoeba castellanii* strain C3 regions with no mapping in Neff.

The C3 genome was aligned to the Neff assembly using Minimap2 (v. 2.1) (Li 2018), and the coordinates of C3 regions with no mapping in Neff were retrieved. The figure shows C3-specific regions in black along the 50 largest C3 scaffolds.



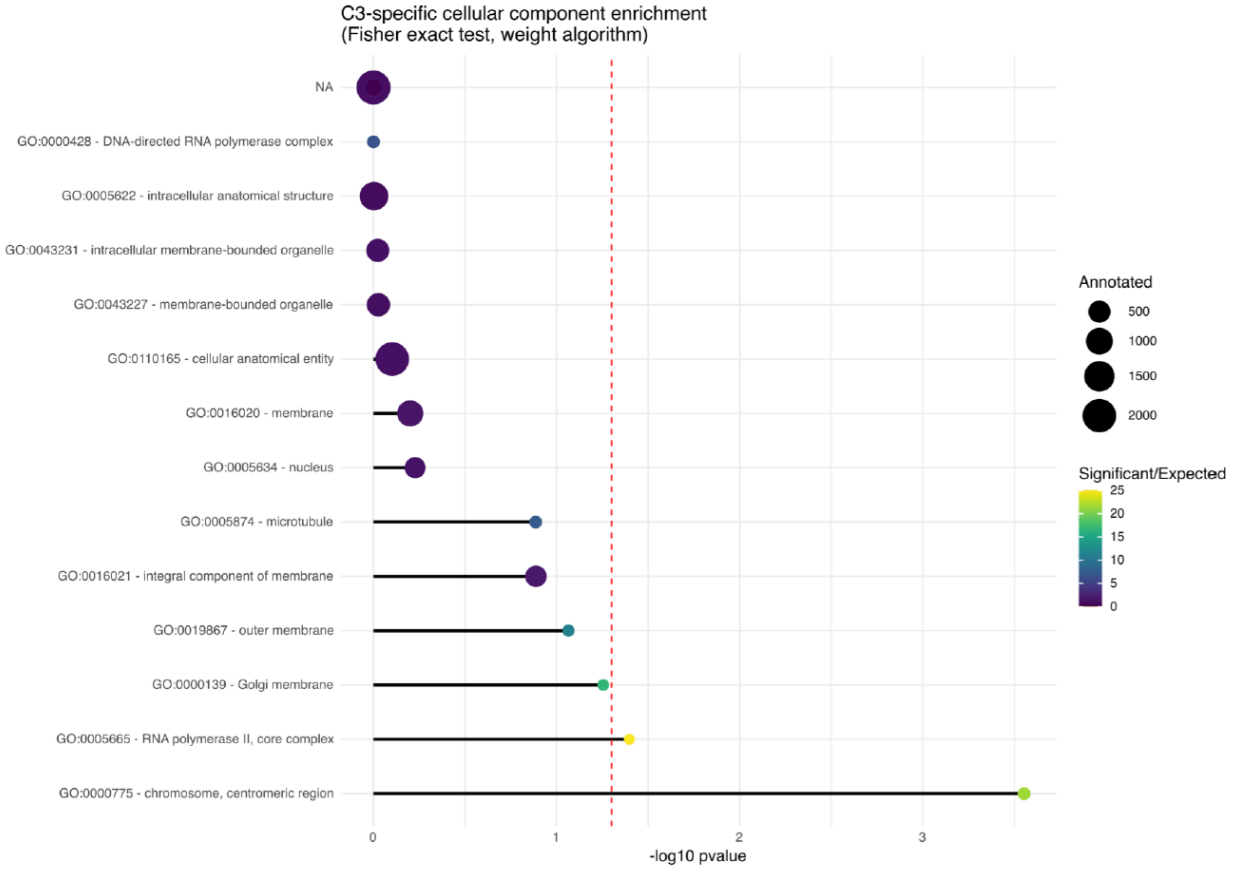
Supplementary Figure S5. Most significant biological process GO term enrichments in genes specific to *Acanthamoeba castellanii* strain C3.

Enrichment was determined using topGO , with nodeSize set to 10 when building the GOdata object. The size of the circles at the end of the bars represents the number of genes annotated under that GO term in the genome, and the colour scale of the circles represents the ratio of how many genes were found in the strain-specific set for that term compared to how many were expected.



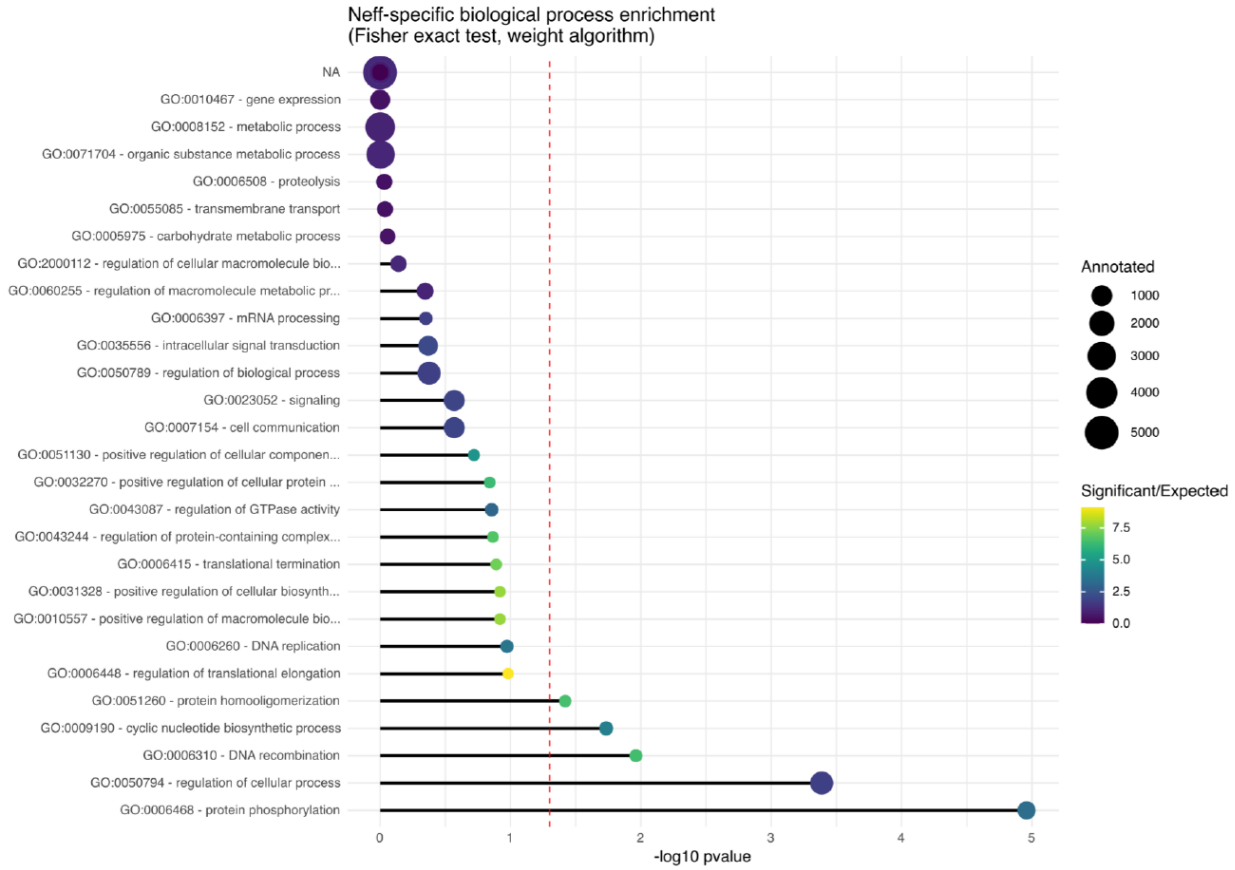
Supplementary Figure S6. Most significant molecular function GO term enrichments in genes specific to *Acanthamoeba castellanii* strain C3.

Enrichment was determined using topGO, with nodeSize set to 10 when building the GOdata object. The size of the circles at the end of the bars represents the number of genes annotated under that GO term in the genome, and the colour scale of the circles represents the ratio of how many genes were found in the strain-specific set for that term compared to how many were expected.

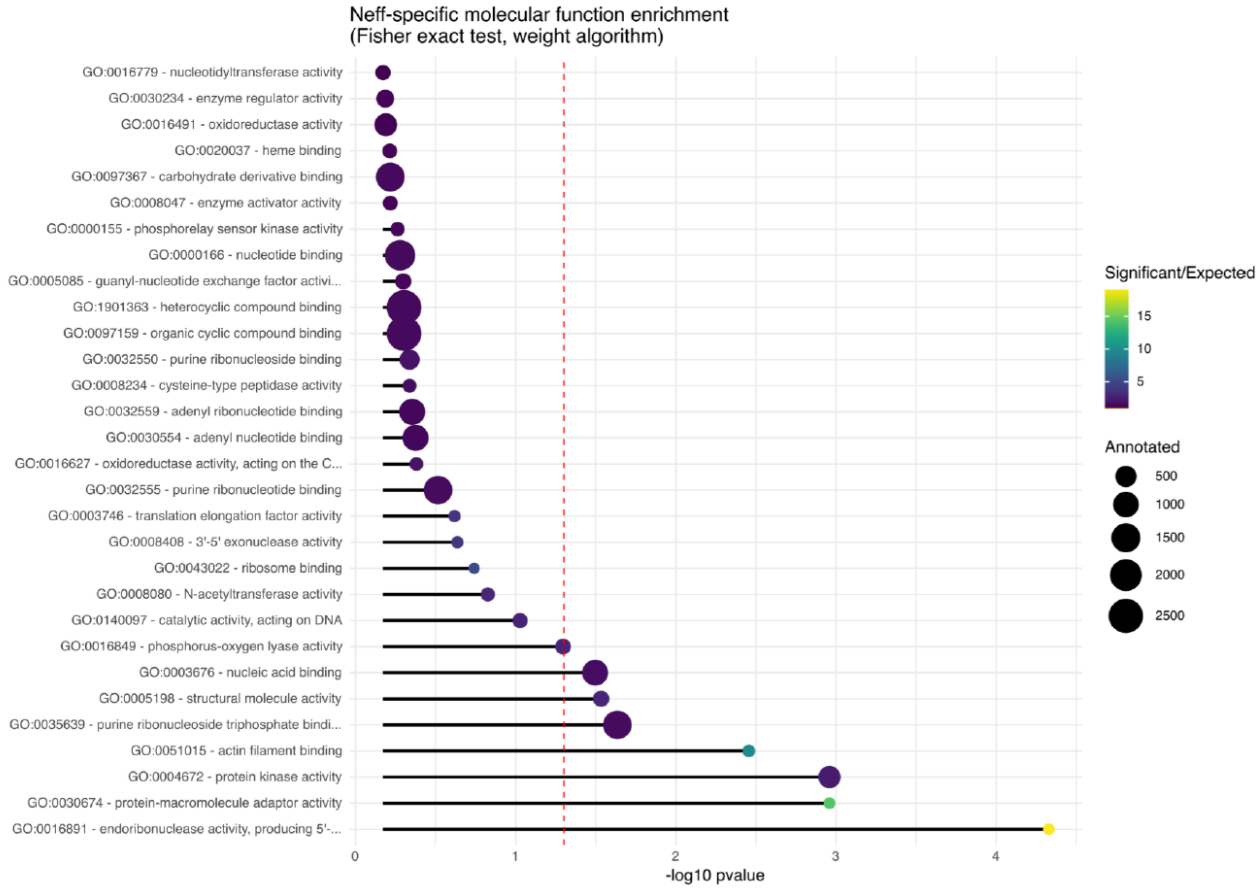


Supplementary Figure S7. Most significant cellular component GO term enrichments in genes specific to *Acanthamoeba castellanii* strain C3.

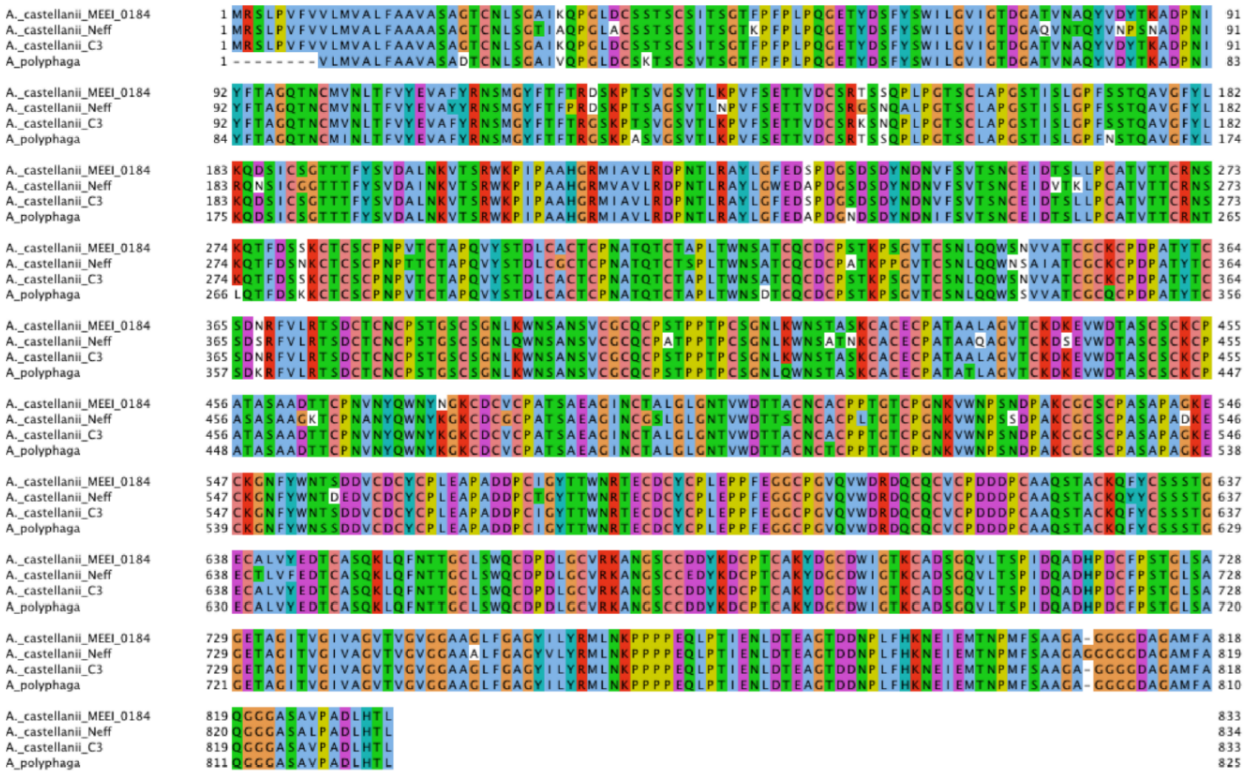
Enrichment was determined using topGO , with nodeSize set to 5 when building the GOdata object. The size of the circles at the end of the bars represents the number of genes annotated under that GO term in the genome, and the colour scale of the circles represents the ratio of how many genes were found in the strain-specific set for that term compared to how many were expected.



Supplementary Figure S8. Most significant biological process GO term enrichments in genes specific to *Acanthamoeba castel lanii* strain Neff. Enrichment was determined using topGO, with nodeSize set to 10 when building the GOdata object. The size of the circles at the end of the bars represents the number of genes annotated under that GO term in the genome, and the colour scale of the circles represents the ratio of how many genes were found in the strain-specific set for that term compared to how many were expected.

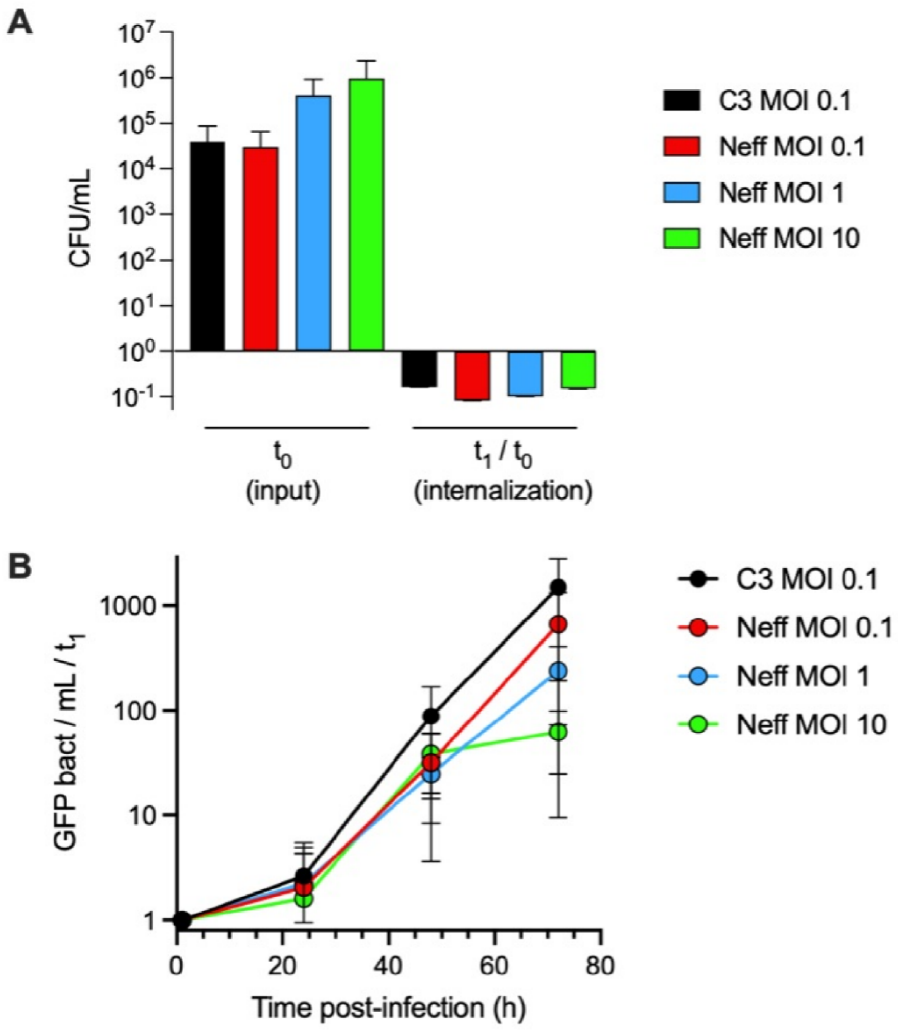


Supplementary Figure S9. Most significant molecular function GO term enrichments in genes specific to *Acanthamoeba castel lanii* strain Neff. Enrichment was determined using topGO, with nodeSize set to 10 when building the GOdata object. The size of the circles at the end of the bars represents the number of genes annotated under that GO term in the genome, and the colour scale of the circles represents the ratio of how many genes were found in the strain-specific set for that term compared to how many were expected.



Supplementary Figure S10. Multiple sequence alignment of mannose binding protein orthologs across three strains of *Acanthamoeba castellanii* and one strain of *Acanthamoeba polyphaga.*

Sites are coloured according to the Clustalx colour scheme and residues differing from the consensus at any given site are not coloured. The alignment was generated with MAFFT- linsi, and was viewed and coloured in Jalview.

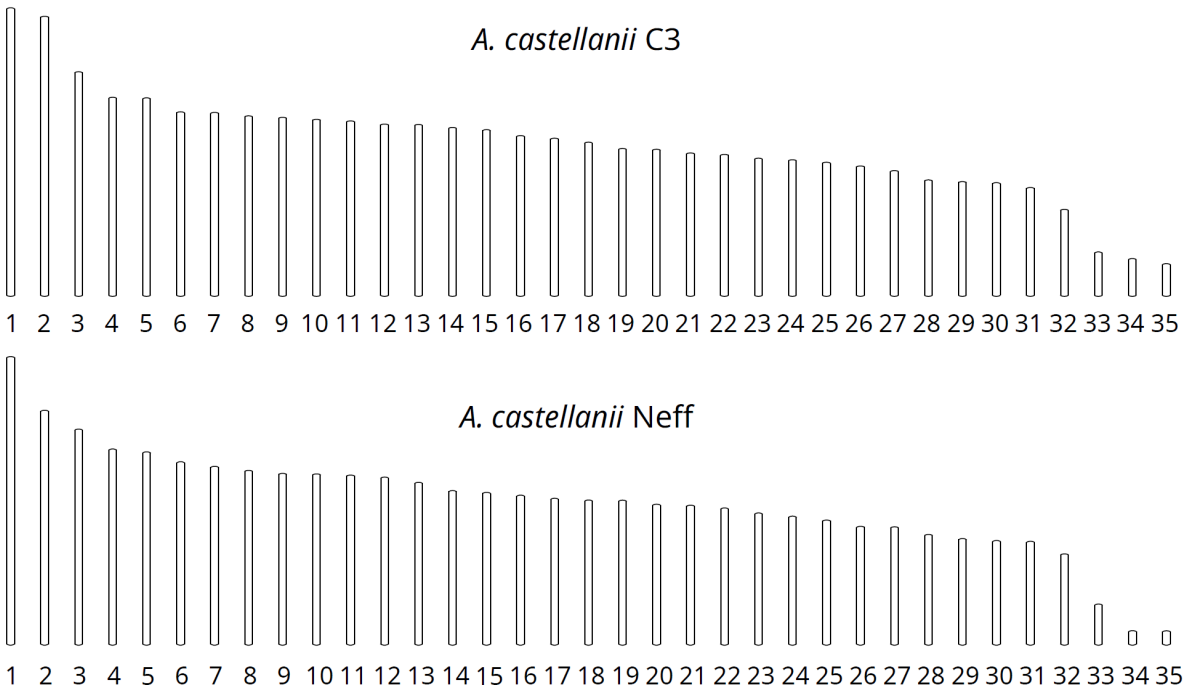


Supplementary Figure S11. Comparative entry and replication of *L. pneumophila* in C3 and Neff strains.

Acanthamoeba castellanii strains C3 and Neff were infected with *Legionella pneumophila* strain Paris constitutively expressing GFP. Strain C3 was infected at an MOI=0.1 and strain Neff was infected at MOI=0.1, MOI=l or MOI=lO.

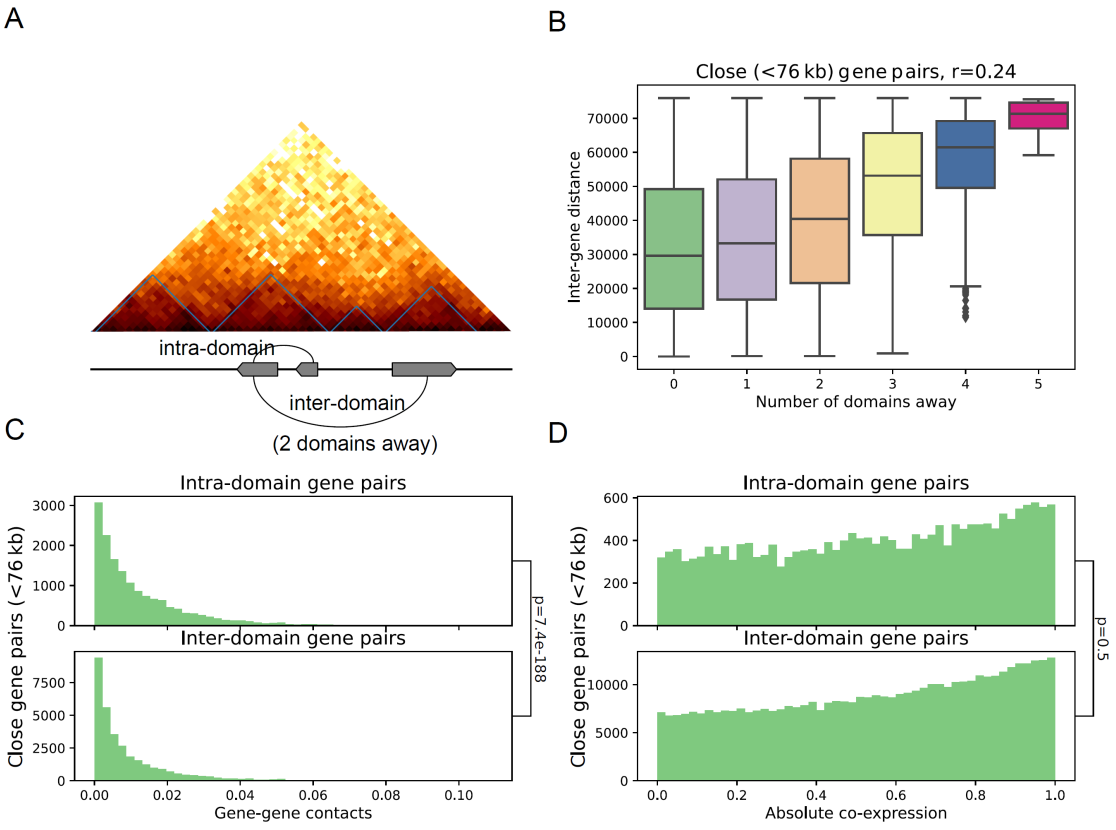
A, At t=O hours post infection (hpi) and at t=l hpi , 300 µL of infected amoeba were removed from the infection cultur e, amoebae were lysed and 100 µL were plated on BCYE plates. CFU were counted to determine bacterial numbers used for infection (input , *t0)* and numbers of internalized bacteria (internalization, ti /t0).

B, At t=O, 1, 24 , 48 and 72 hpi, 300 µL of infected amoeba were removed from the infection culture, amoeba were lysed and 100 µL were placed in 96 well plates in duplicates and analyzed by Flow Cytometry. Graph shows absolute numbers of GFP bacteria per mL. Data were normalized to ti (n=3).



Supplementary Figure S12. Predicted karyotypes of *A. castellanii* strains C3 and Neff

For each strain, 35 scaffolds likely to be chromosomes based on the presence of inter-telomeric contact patterns on the contact maps are ordered by size.



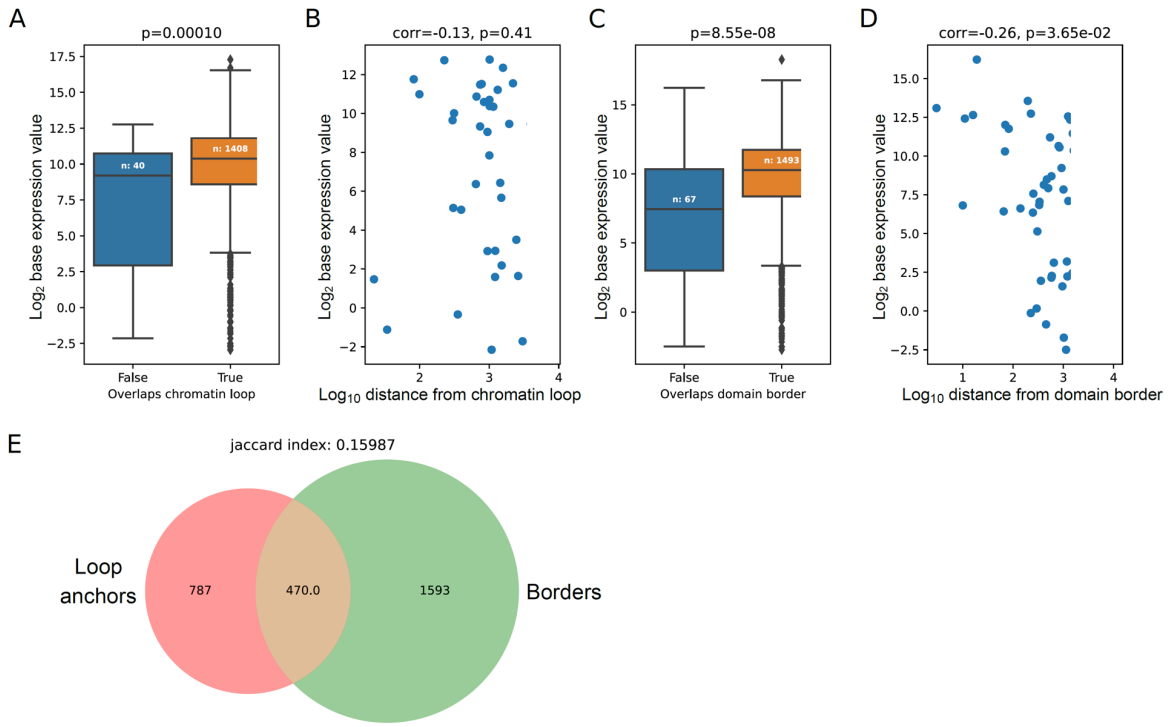
Supplementary Figure S13. Relationship between genes and self-interacting domains.

A, Example of domains detected using Chromosight in the C3 strain, with hypothetical genes indicated for the purpose of illustration.

B, Relationship between inter-gene distance and the number of domains separating them.

C, Distribution of mean inter-gene contacts according to domain separation status.

D, Distribution of gene-pairs co-expression according to domain separation status. For all panels, only gene pairs separated by less than the median domain size are selected.



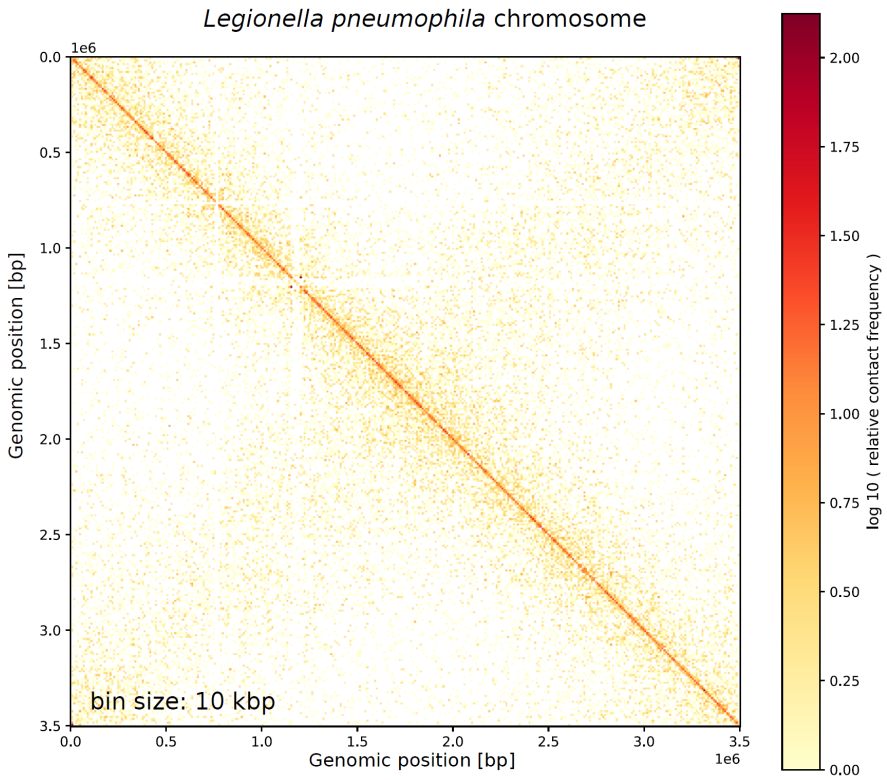
Supplementary Figure S14. Gene expression according to position relative to chromatin loop.

Expression of the closest gene to each loop anchors versus **A**, overlap status with chromatin loops and **B**, distance to closest loop.

Expression of the closest gene to domain borders versus **C**, overlap status with domain borders and **D**, distance to closest border.

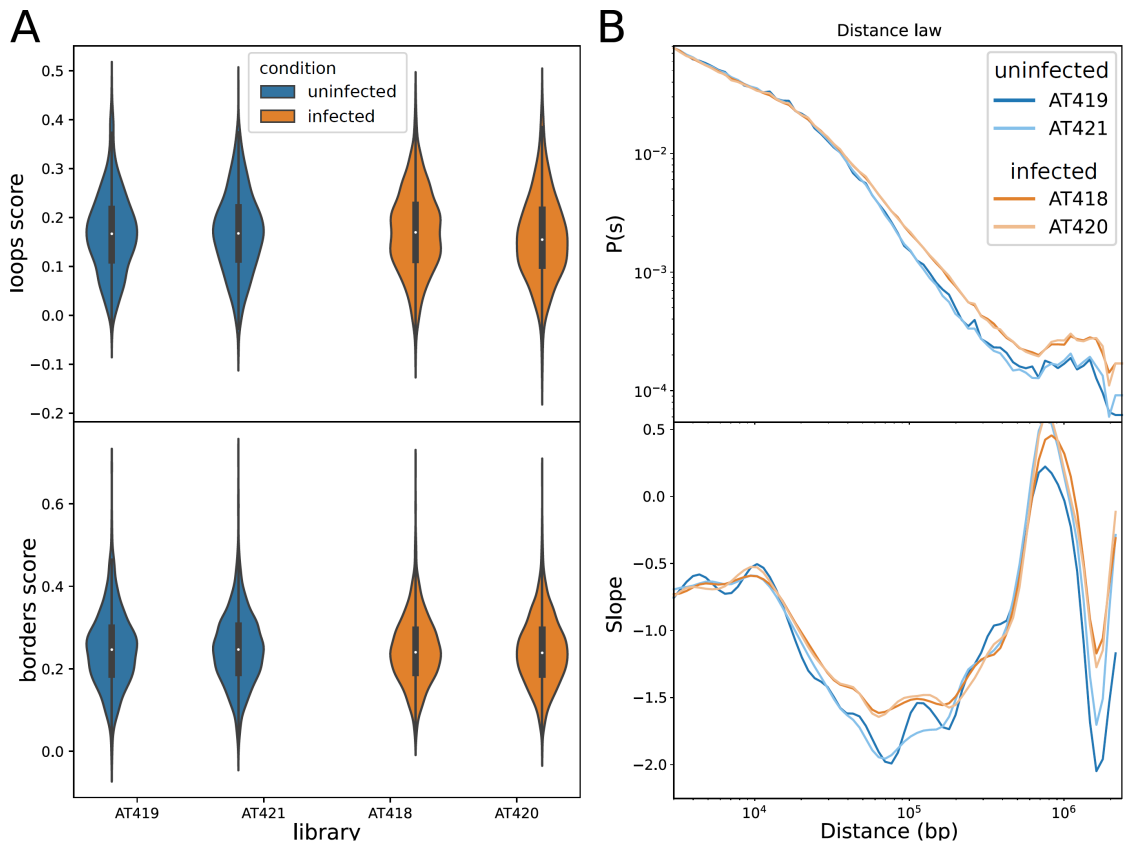
P-values reported for overlap comparisons are obtained using Mann-Whitney U test, correlation coefficients and associated p-values are computed using Spearman 's correlation test.

E, Overlap between chromatin loop anchors and domain borders represented as a Venn diagram.



Supplementary Figure S15. Hi-C contact map of the *Legionella pneumophila* chromosome during infection.

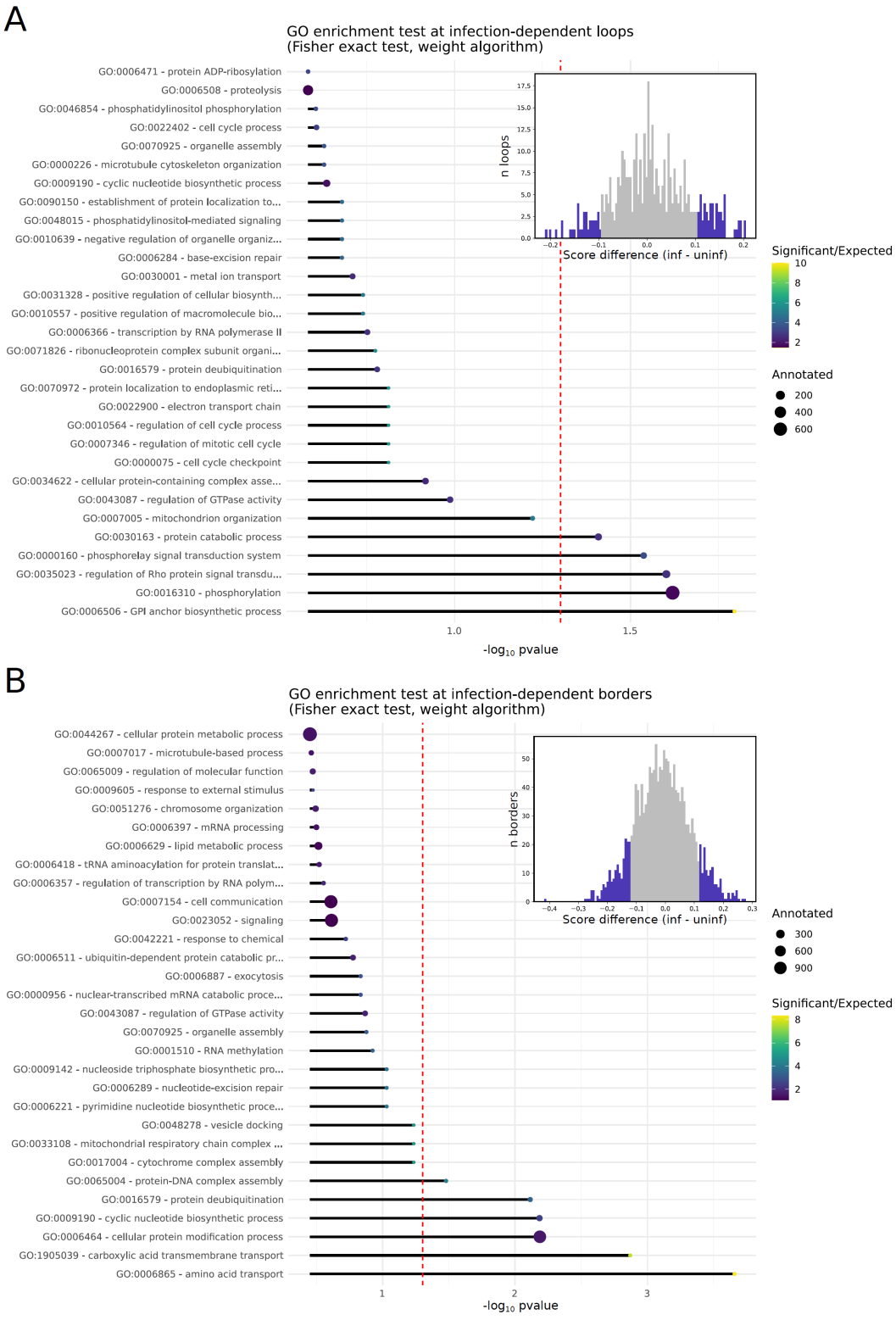
This contact map was generated by aligning all reads from the Hi-C libraries generated on *A. castellanii* C3 strain infected by *L. pneumophila* strain Paris along the reference genome of *L. pneumphila* (NCBI accession NC_006368.1). Bin: 10 kb.



Supplementary Figure S16. Global comparisons of infection Hi-C results between replicates.

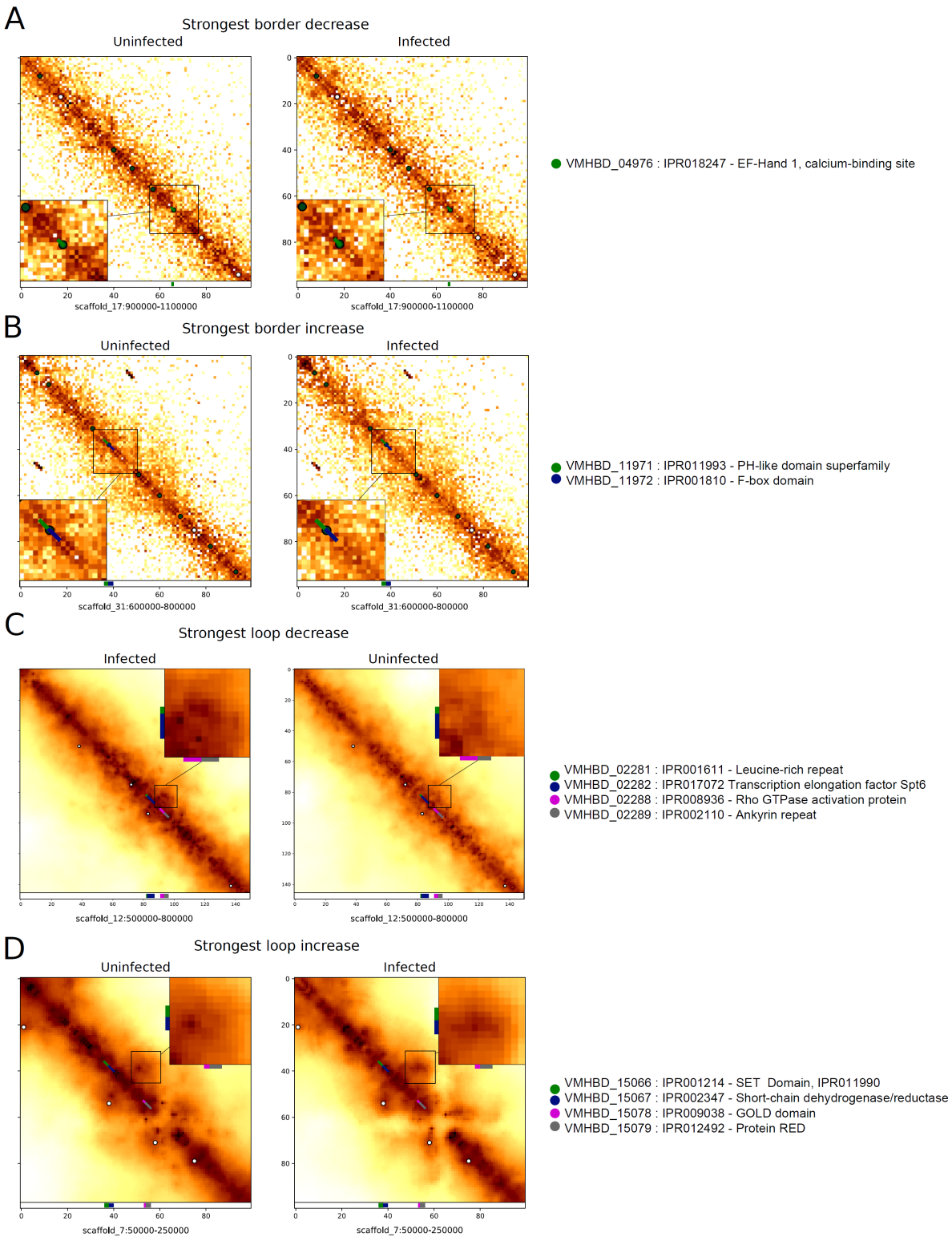
A, Distribution of Chromosight loops and borders scores for all 4 samples.

B, Distance-contact decay function (denoted P(s)) and its slope.

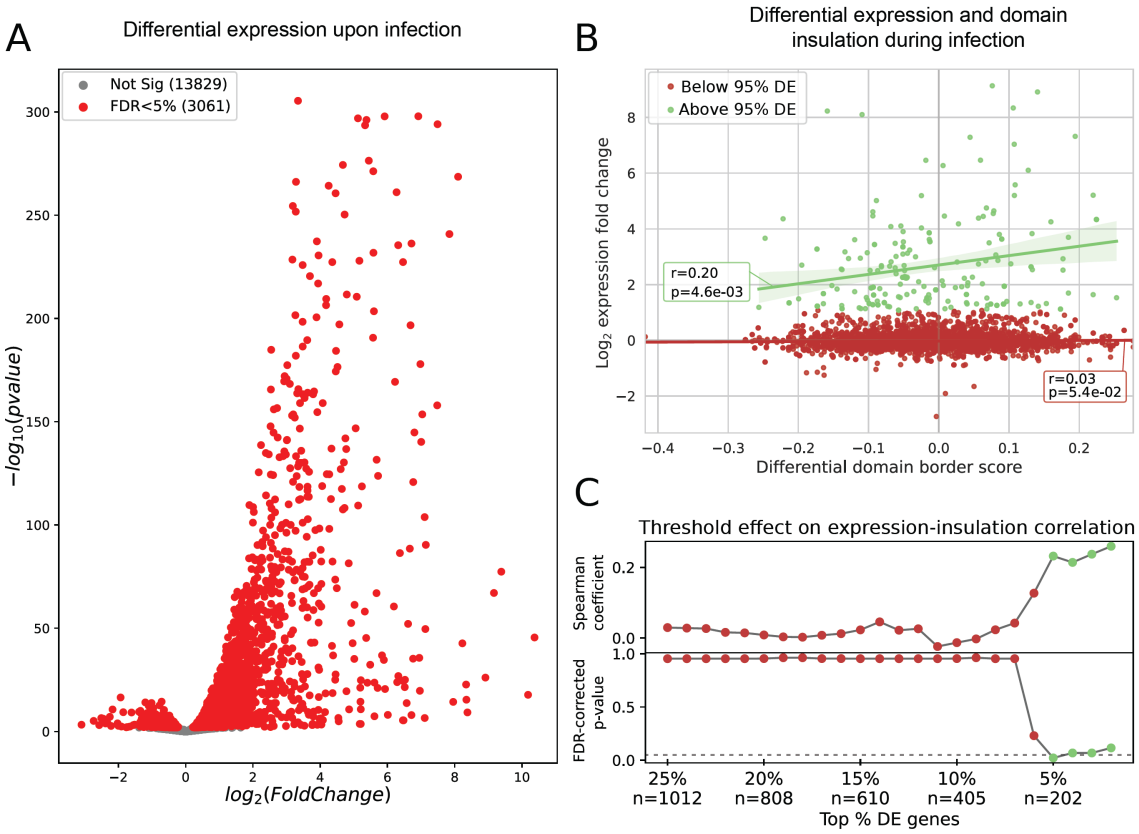


Supplementary Figure S17. GO term enrichment test results for genes overlapping infection-dependent

A, chromatin loops and **B**, domain borders. Histograms show the distribution of loop and border score changes during infection, with highlighted portions showing the 80% percentile threshold used to include genes in the GO enrichment test.



Supplementary Figure S18. Hi-C zooms on strongest pattern changes during infection.
Description of the closest genes are shown below each magnification.
Magnifications of balanced contact map showing **A**, strongest border decrease and **B**, increase.
Serpentine-binned contact maps showing **C**, strongest loop decrease and **D**, increase.



Supplementary Figure S19. Relationship between differential expression and domain insulation during infection.

A, Volcano plot showing differential gene expression (DE) of infected (5h p.i.) versus uninfected amoeba. Genes with significant corrected p-values (FDR<5%) are shown in red.

B, Changes in gene expression and insulation strength of closest domain bord er during infection. Linear regression lines, Spearman correlation coefficients and associated p-values are shown separately for genes with extreme fold change values (95% quantile) and the rest.

C, Spearman correlation coefficient between expression fold change and domain insulation change, and associated FDR-corrected p-values (FDR<5%) for different subsets of genes according to the threshold of extreme fold change. Values are colored according to the 95% threshold selected in b.

Samples					Alignment				Event types					
Used for	Strain	Condition	Library	Mn. pairs	No	Multi	Single	MQ30	Discard	Intra	Inter	Dups.	Mn. used	% in matrix
Infection	C3	infected	AT418	102.5	23.4%	14.2%	62.3%	63%	91.1%	8.8%	2.8%	23%	3.8	3.7%
		uninfected	AT419	94.0	13.9%	19.7%	66.4%	68%	97.9%	2.1%	0.6%	32%	0.8	0.8%
		infected	AT420	94.4	19.1%	15.9%	64.9%	66%	94.9%	5.0%	1.6%	26%	2.0	2.1%
		uninfected	AT421	125.4	30.1%	16.1%	53.7%	55%	94.0%	5.9%	1.6%	44%	2.0	1.5%
Assembly		uninfected	AT337	50.8	6.8%	65.2%	27.8%	30%	23.9%	76.1%	23.7%	5%	6.5	12.9%
		infected	AT407	115.5	19.6%	14.5%	65.8%	64%	88.1%	11.8%	5.4%	7%	6.8	5.9%
		uninfected	AT408	112.2	27.1%	14.2%	58.6%	58%	85.8%	14.1%	6.8%	8%	6.9	6.2%
		infected	PM106	87.3	8.5%	17.0%	74.4%	73%	80.3%	19.6%	8.8%	18%	8.7	10.0%
	Neff	uninfected	AT338	40.0	60.5%	5.5%	33.8%	26%	10.0%	89.9%	37.7%	4%	2.5	6.3%

Supplementary Table S1 Read statistics for *Acanthamoeba castellanii* Hi-C libraries.

The first columns describe each library's sample: For what type of analysis the library was used (infection or genome assembly), what *A. castellanii* strain it contains, its ID and the number of read pairs sequenced in millions. The next columns describe alignment statistics: The percentage of reads which did not align to the reference, aligned more than once or a single time, as well as the total percentage of reads with a mapping quality above 30 (MQ30). The "Event types" columns describe the proportion of different Hi-C events relative to single-aligned reads that passed the MQ30 threshold: discarded events represent undigested restriction fragments or religation on the same fragment, while intra and inter represent valid Hi-C contact within- or between-scaffolds. The remaining columns show general statistics of the libraries, such as the proportion of PCR duplicates, millions and percentage of read pairs used in the final Hi-C contact maps. Despite showing higher percentages of retained reads, libraries AT337, AT407, AT408 and PM106 were not used for infection analysis because they were prepared in separate batches and presented technical variations.

	C3	
	GO term	p-value
Biological process	Macromolecule methylation	1.9 x 10-5
	Protein phosphorylation	0.00068
	Small GTPase mediated signal transduction	0.00289
	Amino acid transport	0.01822
	DNA topological change	0.02584
Molecular function	S-adenosylmethionine-dependent methyltransferase activity	0.0042
	GTP binding	0.00125
	Chromatin binding	0.00139
	Phosphotransferase activity, alcohol group as acceptor	0.00264
	Catalytic activity, acting on DNA	0.0089
	DNA topoisomerase type II (double strand cut, ATP-hydrolyzing) activity	0.01869
	Oxidoreductase activity, acting on the aldehyde or oxo group of donors, NAD or NADP as acceptor	0.03677
	DNA binding	0.0419
	O-acyltransferase activity	0.04753
Cellular component	Chromosome, centromeric region	0.00028
	RNA polymerase II, core complex	0.04003
	Neff	
Biological process	Protein phosphorylation	1.1 x 10-5
	Regulation of cellular process	0.00041
	DNA recombination	0.01092
	Cyclic nucleotide biosynthetic process	0.01844
	Protein homooligomerization	0.038
Molecular function	Endoribonuclease activity, producing 5‘-phosphomonoesters	4.7 x 10-5
	Protein-macromolecule adaptor activity	0.0011
	Protein kinase activity	0.0011
	Actin filament binding	0.0035
	Purine ribonucleoside triphosphate binding	0.0232
	Structural molecule activity	0.0293
	Nucleic acid binding	0.0319
Cellular component	Virion part	1.5 x 10-9

Supplementary Table S2. Functions enriched in C3- or Neff-specific sets of genes.

Strain-specific gene sets were determined based on the consensus of the orthologous clustering programs Broccoli and OrthoFind er, and functions were represented as Gene Ontology terms, separated into the three ontologies: 'biological process', 'molecular function', and 'cellular component'. The R package topGO was used to implement Fisher's exact test with the weight algorithm for the calculation of enriched functions.

Supplementary References

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**: 772–780.

Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**: 3094–3100.

Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. 2009. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**: 1189–1191.