Supporting Information

Brooks et al. 10.1073/pnas.1006741108

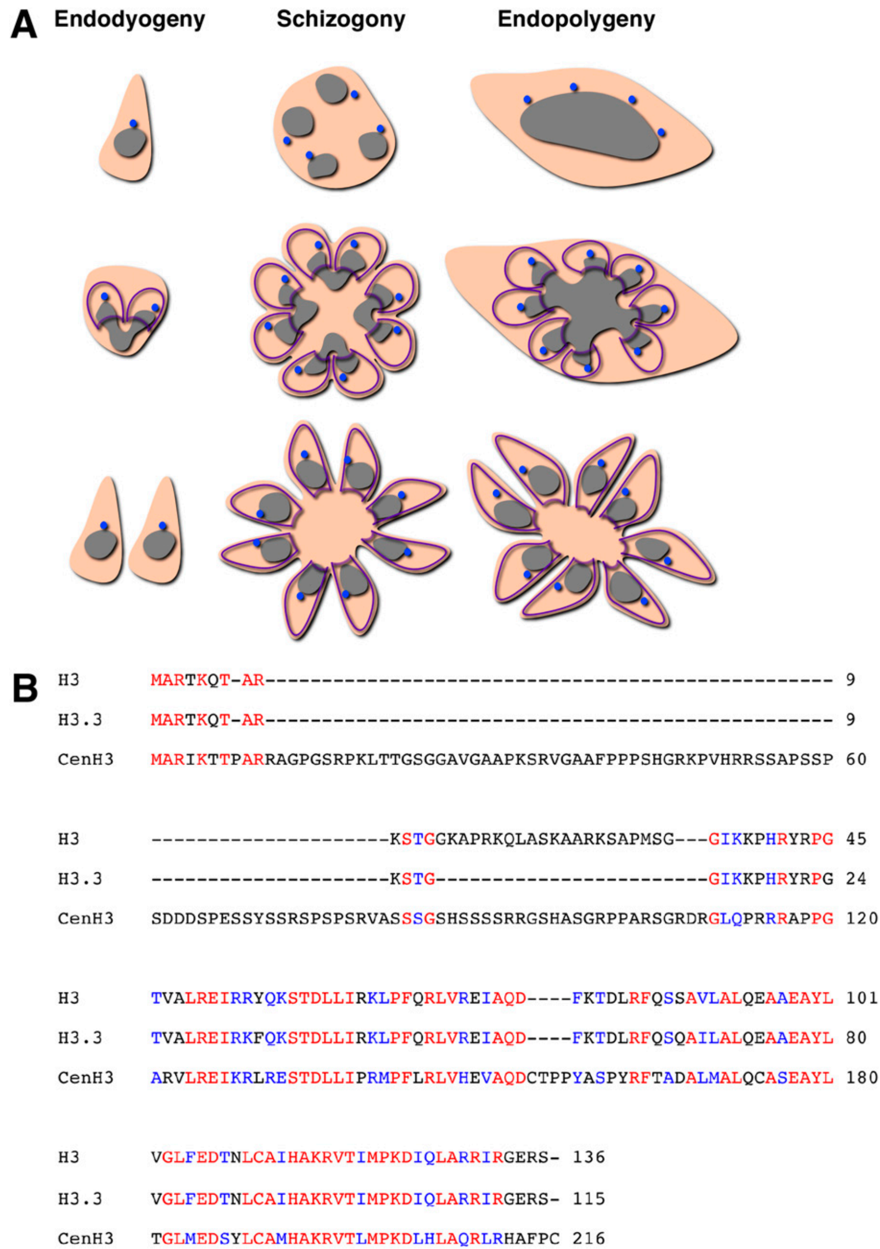

Fig. S1. (*A*) Schematic outline of the three apicomplexan cell division and budding types. Nuclei are shown in gray, centrosomes in blue, and the budding daughter pellicle in purple (modified from ref. 1). (*B*) Multiple protein sequence alignment of the three histone H3 proteins encoded in the *Toxoplasma gondii* genome. Note N-terminal insertion in centromeric H3 variant (CenH3).

1. Striepen B, Jordan CN, Reiff S, van Dooren GG (2007) Building the perfect parasite: Cell division in Apicomplexa. *PLoS Pathog* 3:e78.

ChrIA ChrV ChrIX

ChrIB ChrVI ChrX

ChrII ChrVIIA ChrXI

ChrIII ChrVIII ChrXII

Fig. S2. CenH3-associated DNA is largely free of genes. Detailed view of ChIP-on-chip analysis for centromeric regions of the *T. gondii* genome. Twelve of 14 chromosomes are shown (chromosome IV and VIIb did not produce significant hybridization; Fig. 2). Each panel shows (from top to bottom): histone H3 (red), CenH3-HA (black), and predicted protein-coding genes (blue). Note that these data also can be viewed using the genome browse feature at www.toxodb.org by activating the centromere feature from the ChIP dataset. Chr, chromosome.

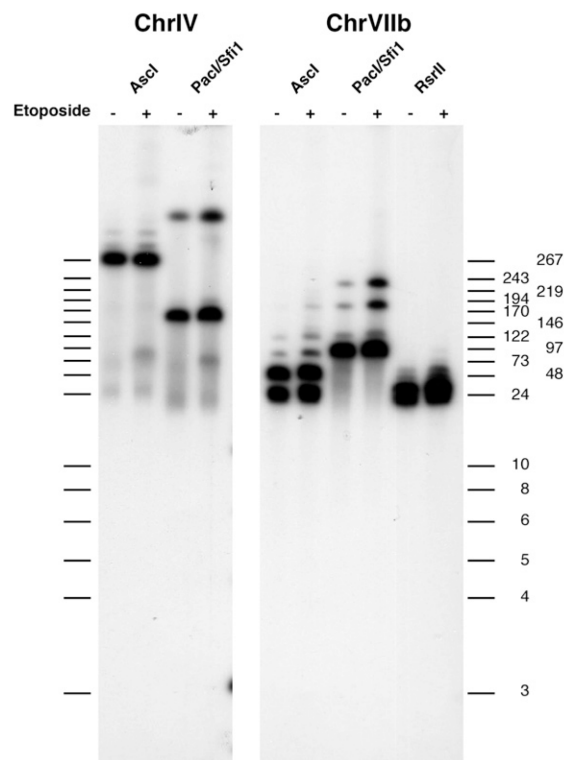

Fig. S3. Etoposide mapping of candidate centromeric regions on chromosomes IV and VIIb. RH strain parasites were cultured in the presence (+) or absence (–) of etoposide as detailed in *Results and Discussion*. The large fragments to be analyzed in this set of experiments (up to 250,000 bp) required precautions to avoid the shearing of genomic DNA. Parasites (10^9/mL) were embedded in 0.7% low-melting-point agarose, digested with proteinase K at 60 °C for 48 h, followed by eight subsequent 2-h Tris-EDTA (pH 8.0) washes. Plugs then were equilibrated twice in restriction buffer for 30 min before overnight digestion with AscI, PacI/SfiI, or RsrII as indicated. Plugs were loaded into a 0.7% agarose/0.5× Tris-borate-EDTA gel for separation by pulse-field electrophoresis (16 h at 5 V/cm with a 120° included angle and 1- to 25-s switch times) before blotting onto nylon membrane. Note that neither of the candidate regions shows fragmentation as a result of etoposide treatment. Further note that they are a poor match for the version 6 chromosome assembly for any strain. For the candidate gene-poor region on chromosome IV (Me49, base pairs 2,136,355–2,205,659, ToxoDB version 6.2), we used a 675-bp probe internal to the candidate centromere region (Me49, base pairs 2,124,104–2,124,778). The GT1 assembly predicts AscI (98,658 bp) and PacI/Sfi: (87,291 bp), and the Me49 assembly predicts AscI (180,453 bp) and PacI/SfiI (71,521 bp). The restriction fragments actually measured for the RH strain are AscI: ∼260,000 bp and PacI/SfiI: ∼160,000 bp. For the candidate gene-poor region on chromosome VIIb (Me49, base pairs 3,720,000–3,745,000), we used a 600-bp probe in the putative centromeric region in the newest assembly (ToxoDB version 6.2). This sequence actually is present twice in the genome. In the Me49 assembly it is found at base pairs 3,715,554–3,716,148, and in the telomere it is found at base pairs 5,003,349–5,003940. In contrast, the GT1 assembly places both sequences close together within the candidate sequence. Our restriction mapping did not produce any match with the Me49 assembly (putative centromere: AscI 54,000 bp; PacI/SfiI: 65,800 bp; RsrII 50,200 bp; telomere: AscI: 20,584 bp; PacI/SfiI: 102,341 bp; RsrII: 25,007 bp) but produced good matches with the GT1 assembly for two digests (predicted: SfiI/PacI: 86,000 bp, and AscI: 45,500 and 28,657 bp). Note that the GT1 assembly would have predicted one slightly larger product for RsrII restriction than actually observed (43,550 and 21,883 bp, respectively).

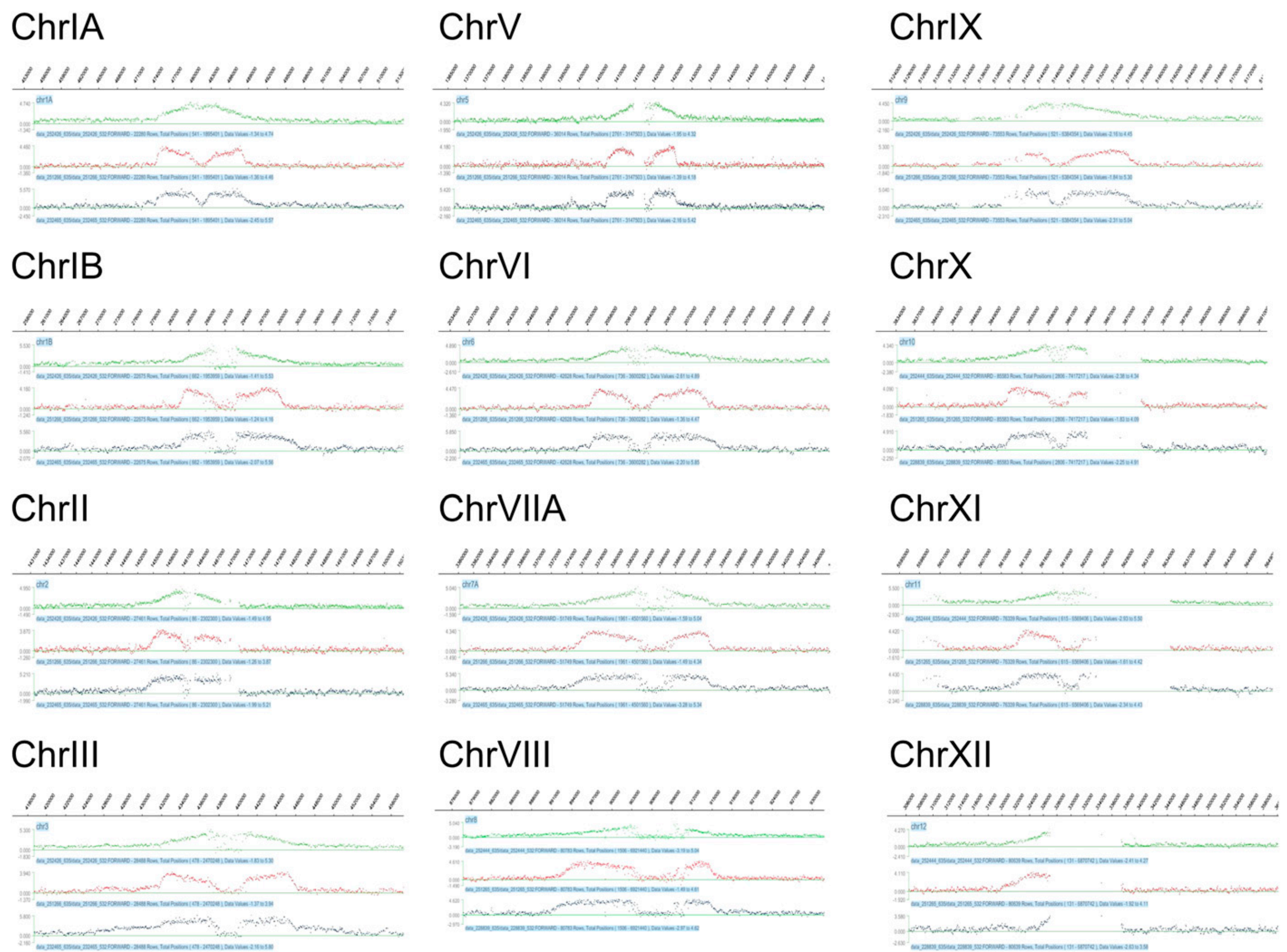

Fig. S4. The *T. gondii* centromeres are marked by flanking histone H3K9 di- and trimethylation. Detailed views of ChIP-on-chip analysis for centromeric regions of 12 chromosomes are shown. Each panel shows (from top to bottom): CenH3-HA (green), H3 lysine 9 dimethylation (H3K9me2) (red), and H3 lysine 9 trimethylation (H3K9me3) (black). Note that these data also can be viewed using the genome browse feature at www.toxodb.org along with additional ChIP data sets [in particular histone 3 acetylated at lysine 9 (H3K9ac) and H3K4me3, which mark active promoter regions] from this and other studies.

Fig. S5. A single significant cluster of CenH3 hybridization per chromosome. Two biological replicates of CenH3 ChIP-chip were analyzed using NimbleScan software to determine consensus peaks of hybridization using a 500-bp sliding window. Positions of peaks with a false-discovery rate of <0.05 within each chromosome are shown here. Significant clustering (more than five consecutive high-confidence peaks) was readily identified for all chromosomes with the exception of IV, VIIb, and XII.

Table S1. Localization of CenH3, H3K9me2, and H3K9me3 hybridization peaks with >40-bp overlap

Chromosome	Start	Stop	Size
Ia	476,538	476,787	249
Ia	478,023	478,132	109
Ia	479,128	481,277	**2,149**
Ia	481,623	481,667	44
Ia	485,893	486,317	424
Ia	486,688	487,172	484
Ia	487,448	489,052	**1,604**
Ib	284,979	285,278	299
Ib	288,299	288,804	505
Ib	292,774	296,404	**3,630**
Ib	296,489	297,056	567
II	1,456,668	1,457,232	564
II	1,457,388	1,457,942	554
II	1,458,083	1,460,234	**2,151**
II	1,463,093	1,463,277	184
II	1,464,088	1,464,505	417
II	1,465,043	1,465,485	442
II	14,66,228	1,466,392	164
II	1,469,958	1,470,208	250
V	1,419,130	1,419,759	629
V	1,420,220	1,423,444	**3,224**
VI	2,057,323	2,060,417	**3,094**
VI	2,070,773	2,071,207	434
VIIa	3,387,880	3,390,334	**2,454**
VIIa	3,391,260	3,391,819	559
VIII	894,798	895,362	564
VIII	897,418	897,467	49
VIII	898,043	899,462	**1,419**
VIII	900,018	901,209	**1,191**
IX	5,143,323	5,143,500	177
IX	5,144,808	5,146,942	**2,134**
IX	5,151,203	5,153,827	**2,624**
IX	5,154,123	5,155,062	939
IX	5,157,098	5,157,148	50
X	816,980	817,134	154
X	3,853,110	3,854,421	**1,311**
X	3,855,085	3,,857859	**2,774**
X	3,861,695	3,864,017	**2,322**
XI	5,600,269	5,600,563	294
XI	5,613,764	5,614,083	319
XI	5,614,819	5,617,933	**3,114**

Peaks falling within the assigned centromeres are shown in bold (www.toxodb.org release 6 coordinates).