PNAS
www.pnas.org

Supplementary Information for

Comparative 3D Genome Organization in Apicomplexan Parasites

Evelien M. Bunnik, Aarthi Venkat, Jianlin Shao, Kathryn E. McGovern, Gayani Batugedara, Danielle Worth, Jacques Prudhomme, Stacey A. Lapp, Chiara Andolina, Leila S. Ross, Lauren Lawres, Declan Brady, Photini Sinnis, Francois Nosten, David A. Fidock, Emma H. Wilson, Rita Tewari, Mary R. Galinski, Choukri Ben Mamoun, Ferhat Ay, and Karine G. Le Roch

Ferhat Ay
Email: ferhatay@lji.org

Karine Le Roch
Email: karinel@ucr.edu

This PDF file includes:

www.pnas.org/cgi/doi/10.1073/pnas.1810815116

Supplementary Materials and Methods

Parasites

An ex vivo *P. knowlesi* blood-stage culture was initiated using iRBCs from a rhesus macaque infected with the Pk1(A+) clone of *P. knowlesi*. Briefly, a cryopreserved vial of ring-stage parasites was thawed, processed and inoculated into a rhesus monkey using standardized procedures. Parasitemia was monitored daily and whole blood was collected into heparin-treated tubes when it reached about 9% ring-stage infected RBCs. This particular strain of *P. knowlesi* is naturally synchronous in rhesus macaques. Collected whole blood was treated with a final concentration of 2 mg/ml ADP (adenosine diphosphate bis salt) for 5 min before passage through a glass bead column equilibrated with Hank's balanced salt solution to remove platelets, followed by passage through a Plasmodipur filter. The filtered blood was centrifuged for 5 min at 800 g and the pellet resuspended in RPMI with L-glutamine, supplemented with 0.25% sodium bicarbonate, 50 ug/ml hypoxanthine, 7.2 mg/ml HEPES, 2 mg/ml glucose, 10% human AB+ serum. The iRBCs were cultured at 2×10^7 iRBCs/ml in 175 cm2 flasks and gassed with 5%: 5%: 90%, O2:CO2:N2. For Hi-C, 1.2×10^9 trophozoite-stage parasites were used.

Female Balb/c mice (n=6 per parasite strain) were injected intraperitoneally with *P. yoelii* XNL PanK1Δ. PanK1 is not essential for parasite replication in the blood stage, and these transgenic parasites undergo normal asexual development and sexual differentiation (1). As a control, we also performed Hi-C on the wild-type *P. yoelii* XNL strain. This sample was discarded from this study due to low sequencing resolution, but showed a good correlation with the PanK1 mutant. We therefore concluded that deletion of PanK1 does not have an effect on genome organization during the blood stages. At high parasitemia (42-71%), cardiac puncture was performed and 400-500 µl blood was collected in EDTA tubes. The blood from each group was combined. The blood was washed twice in 35 ml of warm sterile PBS and depleted of the buffy coat by removal of the upper layer after centrifugation.

Six-to-eight week old female Tuck-Ordinary (TO) or NIHS (Harlan) outbred mice were injected intraperitoneally with wildtype *P. berghei* ANKA and two transgenic strains (Smc2 and Smc4). These transgenic parasites undergo normal asexual development and sexual differentiation. Purification of gametocytes was achieved using a modified protocol from ref (2). Briefly, mice were treated by intraperitoneal injection of 0.2 ml of phenylhydrazine (6 mg/ml) (Sigma) in PBS to encourage reticulocyte formation four days prior to infection with parasites. Day four post infection (p.i.) mice were treated with sulfadiazine (Sigma) at 20 mg/L in their drinking water for two days to eliminate asexual blood stage parasites. On day six p.i. mice were bled by cardiac puncture into heparin and gametocytes separated from uninfected erythrocytes on a NycoDenz gradient made up from 48% NycoDenz (27.6% w/v NycoDenz in 5 mM Tris-HCl, pH 7.20, 3 mM KCl, 0.3 mM EDTA) and coelenterazine loading buffer (CLB), containing PBS, 20 mM HEPES, 20 mM Glucose, 4 mM sodium bicarbonate, 1 mM EGTA, 0.1% w/v bovine serum albumin, pH 7.25. Gametocytes were harvested from the interface, washed and crosslinked as described earlier.

Female CB17/Icr-*Prkdc*^{scid}/IcrIcoCrl mice (n=4 per parasite strain) were injected intraperitoneally with *B. microti* clinical isolate Bm1438 or tick isolate LabS1. When parasitemia reached 15-20% as determined using a Giemsa-stained smear from a drop of blood from the tail vein, 400-600 µl of blood was collected in EDTA tubes by cardiac puncture. The blood was washed twice in 35 ml of warm sterile PBS and depleted of the buffy coat by removal of the upper layer after centrifugation. Parasites directly harvested from infected mice are relatively asynchronous. To also obtain synchronous parasites, 500 µl of blood from infected mice (n=1 per strain) was

washed twice with 1 ml of complete medium (RPMI medium 1640, 30 mg/l hypoxanthine, 25 mM Hepes, 0.225% NaHCO3, 10% FBS, Pen/Strep, 100 µg/ml Kanamycin, 10 µg/ml Gentamycin). After resuspension in 500 µl of complete medium, 250 µl of iRBCs was added to 10 ml of complete media and incubated at 37°C under a gas mixture of 3.5% CO_2, 4% O_2, and 92.5% N_2. Synchronous parasites were harvested after 24 hours (for replicate 1) or 48 hours (for replicate 2).

Toxoplasma gondii strain ME49 tachyzoites were serially passaged and maintained in human foreskin fibroblasts (HFF-1 cell line). Parasites were isolated from the cells by passing through a blunt 23-gauge needle, and a 3 mm filter. For generation of synchronized tachyzoite cultures, temperature shift was utilized (3): purified parasites were counted and inoculated into a new flask of HFF-1 cells at an MOI of 3:1 for 20 min at 4°C. Cultures were then returned to 37°C for one hour prior to removal of uninvaded parasites. For generation of bradyzoite cultures, purified parasites were counted and inoculated into a new flask of HFF-1 cells at an MOI 0.5:1 for 3 hours in cDMEM (5% FBS, 1% Penicillin/Streptomycin, 1X non-essential amino acids, 10 mM HEPES, and 0.05 mg/ml gentamycin) at 37°C, 5% CO2. To induce the transition to bradyzoites, culture media was replaced with pH 8.2 media (DMEM without sodium bicarbonate, 5% FBS, 25 mM HEPES, and 1% Penicillin/Streptomycin) and incubated at 37°C in the absence of CO_2. Media was replaced with fresh pH 8.2 media every 24 hours and parasites were re-isolated after 3 days. Bradyzoite gene expression was confirmed using RT-PCR for BAG1 (TGME49_259020).

Ethics statements

Animal experimental protocols for *B. microti* and *P. yoelii* isolations followed Yale University institutional guidelines for care and use of laboratory animals and were approved by the Institutional Animal Care and Use Committees (IACUC) at Yale University (Protocol #: 2017-11619 and 2016-11301). Yale University is accredited by the American Association for Accreditation of Laboratory Animal Care (AAALAC Number 101), and has an approved Animal Welfare Assurance (#A3230-01, effective until 5/31/2011) on file with the NIH Office for Protection from Research Risks. Rules for ending experiments in animals were to be enacted if animals showed any signs of distress or appeared moribund. This, however, was not the case for any animals in the study.

Animal experiments for *P. knowlesi* were conducted in AAALAC-accredited facilities at the Yerkes National Primate Research Center in accordance with the Animal Welfare Act and the Guide for the Care and Use of Laboratory Animals. Emory University's Animal Care and Use Committee approved all experimental protocols.

Animal experiments for *P. berghei* at Nottingham have passed an ethical review process and were approved by the United Kingdom Home Office. Work was carried out in accordance with the United Kingdom 'Animals (Scientific Procedures) Act 1986' and in compliance with 'European Directive 86/609/EEC' for the protection of animals used for experimental purposes under UK Home Office Project Licenses.

Hi-C Procedure

Day 1: Parasite pellets were thawed on ice in 300 µl ofHi-C lysis buffer (10 mM Tris-HCl at pH 8.0, 10 mM NaCl, 2 mM AEBSF, Roche Complete Mini EDTA-free protease inhibitor cocktail, 0.25% Igepal CA-630). Parasite membranes were disrupted by passing the lysate through a 26.5-gauge needle 15 times with a syringe. Samples were spun at 2,500 × g for 5 min at room temperature (RT). Pellets were washed once with 500 µl ice-cold Hi-C lysis buffer, resuspended in 50 µl of 0.5% SDS and incubated at 62°C for 10 min to solubilize the chromatin. The SDS was neutralized by the addition of 145 µl of nuclease-free water and 25 µl of 10% Triton X-100,

followed by a 15-minute incubation at 37°C. Then, 25 µl of 10× NEBuffer 2, and 4 µl of 25 U/µl MboI restriction enzyme (100 U) was added to digest the DNA overnight at 37°C in a total volume of 254 µl.

Day 2: After digestion, samples were incubated at 62°C for 20 min to inactivate MboI, and then cooled to RT. MboI generated 5' overhangs were filled in by adding 50 µl of fill-in master mix (consisting of 25.4 µl of water, 0.7 µl of 10 mM dATP, 0.7 µl of 10 mM dTTP, 0.7 µl of 10 mM dGTP, 17.5 µl of 0.4 mM Biotin-14-dCTP (Invitrogen), and 5 µl of 5U/µl DNA Polymerase I, Large (Klenow) Fragment (NEB). Samples were mixed by pipetting and incubated for 45 min at 37°C while rotating. Next, blunt-end ligation was performed by adding 900 µl of ligation master mix (consisting of 666 µl of water, 120 µl of 10× T4 DNA Ligase Buffer (NEB), 100 µl of 10% Triton X-100, 12 µl of 100× BSA and 2 µl of 2,000 U/µl T4 DNA Ligase (NEB), and incubating the samples at 20°C for 4 hours with slow rotation. The nuclei were spun down at 2,500 × g for 5 min at RT to remove random ligation products and to reduce the overall reaction volume. The pellet was resuspended in 100 µl of decrosslinking buffer (50 mM Tris-HCl at pH 8.0, 1% SDS, 1 mM EDTA, and 500 mM NaCl). RNA was degraded by adding 5 µl of 20 mg/ml RNaseA (Invitrogen) and incubating for 45 min at 37°C in a water bath. The tubes were mixed by inversion after 30 min. Proteins were degraded by adding 20 µl of 20 mg/ml proteinase K (NEB) and incubating overnight at 45°C in a water bath.

Day 3: The tubes were first allowed to cool to RT. All subsequent steps were performed using low-bind tubes. DNA was extracted using Agencourt AMPure XP beads (Beckman Coulter) and eluted in 130 µl of 10 mM Tris-HCl at pH 8.0. The purified DNA was transferred into a 130 µl Covaris microTube with snap cap and sheared on a Covaris S220 for 65 sec using 10% duty factor, 200 cycles per burst and a peak incident power of 140. The sheared DNA was transferred to a 1.5 µl tube. The empty Covaris microTube was rinsed with 70 µl of water and this was added to the sample to obtain a total volume of 200 µl.

To prepare for pull-down of biotinylated DNA fragments, 150 µl of MyOne Streptavidin T1 magnetic beads (Invitrogen) were washed twice with 400 µl 1× Bind & Wash (B&W) buffer (5 mM Tris-HCl at pH 7.4, 0.5 mM EDTA, 1 M NaCl) containing 0.05% Tween-20 and resuspended in 200 µl 2× B&W buffer. The sonicated DNA sample and the T1 beads were mixed and incubated at RT for 30 min with rotation. The beads were washed twice by adding 600 µl of 1× B&W buffer (pre-warmed to 55°C) and incubating the tubes at 55°C for 2 min with occasional mixing. The beads were then resuspended in 100 µl of 1× T4 DNA Ligase Buffer and transferred to a new tube.

To perform end-repair on DNA fragments bound to the beads, the beads were resuspended in 100 µl of end repair master mix (81 µl of nuclease-free water, 10 µl of 10× T4 DNA Ligase Buffer with 10 mM ATP (NEB), 4 µl of 10 mM dNTPs, 1 µl of 5 U/µl DNA Polymerase I, Large (Klenow) Fragment (NEB), 2 µl of 3 U/µl T4 DNA Polymerase (NEB), 2 µl of 10 U/µl T4 Polynucleotide Kinase (NEB)) and incubated for 30 min at 20°C in a thermomixer (Eppendorf). The beads were washed twice by adding 600 µl of 1× B&W buffer (pre-warmed to 55°C) and incubating the tubes at 55°C for 2 min with occasional mixing. The beads were then resuspended in 100 µl of 1× NEBuffer2 and transferred to a new tube.

A-tailing was performed by resuspending the beads in 50 µl of A-tailing master mix (41 µl of nuclease-free water, 5 µl of 10× NEBuffer2, 1 µl of 10 mM dATP, 3 µl of 5 U/µl Klenow Fragment (3′→5′ exo–) (NEB)) and incubating the sample for 30 min at 37°C in a thermomixer. The beads were washed twice by adding 600 µl of 1× B&W buffer (pre-warmed to 55°C) and

incubating the tubes at 55°C for 2 min with occasional mixing. The beads were then resuspended in 100 µl of 1× T4 DNA Ligase Buffer and transferred to a new tube.

To perform adapter ligation, the beads were then resuspended in a mix of 43 µl of nuclease-free water and 5 µl of 10× T4 DNA Ligase Buffer. Subsequently, 1 µl of 2,000 U/µl T4 DNA ligase and 1 µl of 15 mM NEBNext Illumina Adapter were added. The reaction mixture was incubated at 20°C for 15 min in a thermomixer. After the addition of 3 µl of USER enzyme, the mixture was incubated for an additional 15 min at 37°C in a thermomixer. The beads were washed twice by adding 600 µl of 1× B&W buffer (pre-warmed to 55°C) and incubating the tubes at 55°C for 2 min with occasional mixing. The beads were then resuspended in 100 µl of 10 mM Tris-HCl at pH 8.0 and transferred to a new tube.

To amplify the library, the beads were resuspended in 50 µl of amplification master mix (25 µl of HiFi HotStart ReadyMix (KAPA Biosystems), 23 µl of nuclease-free water, 1 µl of universal forward primer, and 1 µl of barcoded reverse primer) and incubated with the following PCR program: 45 sec at 98°C, 12 cycles of 15 sec at 98°C, 30 sec at 55°C, 30 sec at 62°C and a final extension of 5 min at 62°C. The library was then purified using double-SPRI size selection, with 0.5× right-side selection (25 µl AMPure XP beads) and 1.0× left-side selection (25 µl AMPure XP beads). Libraries were quantified by NanoDrop (Thermo Scientific) and Bioanalyzer (Agilent), prior to multiplexing and sequencing in a 75-bp paired-end run on a NextSeq500 (Illumina).

Genome sequences, gene annotations and specific gene sets

Genome sequences (.fasta) and gene annotations (.gff) were downloaded from PlasmoDB (*P. falciparum* 3D7 version 28, *P. vivax* Sal1 version 29, *P. berghei* ANKA version 35, and *P. yoelii* 17X version 29), ToxoDB (*T. gondii* ME49 version 29), and PiroplasmaDB (*B. microti* RI version 29). For *P. knowlesi*, we used a recently published genome assembly that was based on our Hi-C data in combination with new sequence data generated using Illumina and Pacbio platforms (4). *P. knowlesi* H gene annotations were obtained from PlasmoDB (version 29) and converted to new genome coordinates. The genome of *P. yoelii* strain 17X was recently resequenced (5), resulting in a more complete genome assembly that is now available at PlasmoDB (version 36 and higher). The genome sequence and gene annotations have particularly improved in the subtelomeric regions that contain the *yir* gene family. However, many of these subtelomeric bins are not mappable, which complicates the generation of a 3D model for this genome. Inference of the *P. yoelii* 3D model and all downstream analyses were therefore performed using PlasmoDB version 29 of the genome. For all other parasites, newer DB versions are available, but have minimal, if any, improvements of genome sequence and gene annotations.

Centromere sequences were downloaded directly from PlasmoDB, ToxoDB and PiroplasmaDB. *Plasmodium* virulence genes were identified as "PfEMP1" for *P. falciparum*, "SICAvar" for *P. knowlesi*, and "PIR" for *P. knowlesi*, *P. berghei*, and *P. yoelii*. For *B. microti*, virulence genes included genes of the BMN1 and BMN2 families, as well as vesa-like and tpr related proteins. For *T. gondii*, genes of the TgFAM-A, Tg-Fam-B, Tg-Fam-C, TgFAM-D, Tg-FAM-E and SRS families were included as virulence genes.

Gene expression datasets

The following publicly available gene expression data sets were used for *P. falciparum* trophozoites and gametocytes: (6), *P. vivax* sporozoites: (7), *P. berghei* gametocytes: (5), *P. yoelii* asexual blood stage: time point 5 from (8), *P. knowlesi* trophozoites: time point 4 from (8), *T. gondii* tachyzoites: 8h time point from "Tachyzoite Transcriptome Time Series (ME49)" generated by Brian Gregory, Pennsylvania University (downloaded from ToxoDB version 35), *T.*

gondii bradyzoites: “Bradyzoite in vitro Transcriptome (ME49)” generated by David Sibley, Washington University School of Medicine (downloaded from ToxoDB version 35), and *B. microti*: (9).

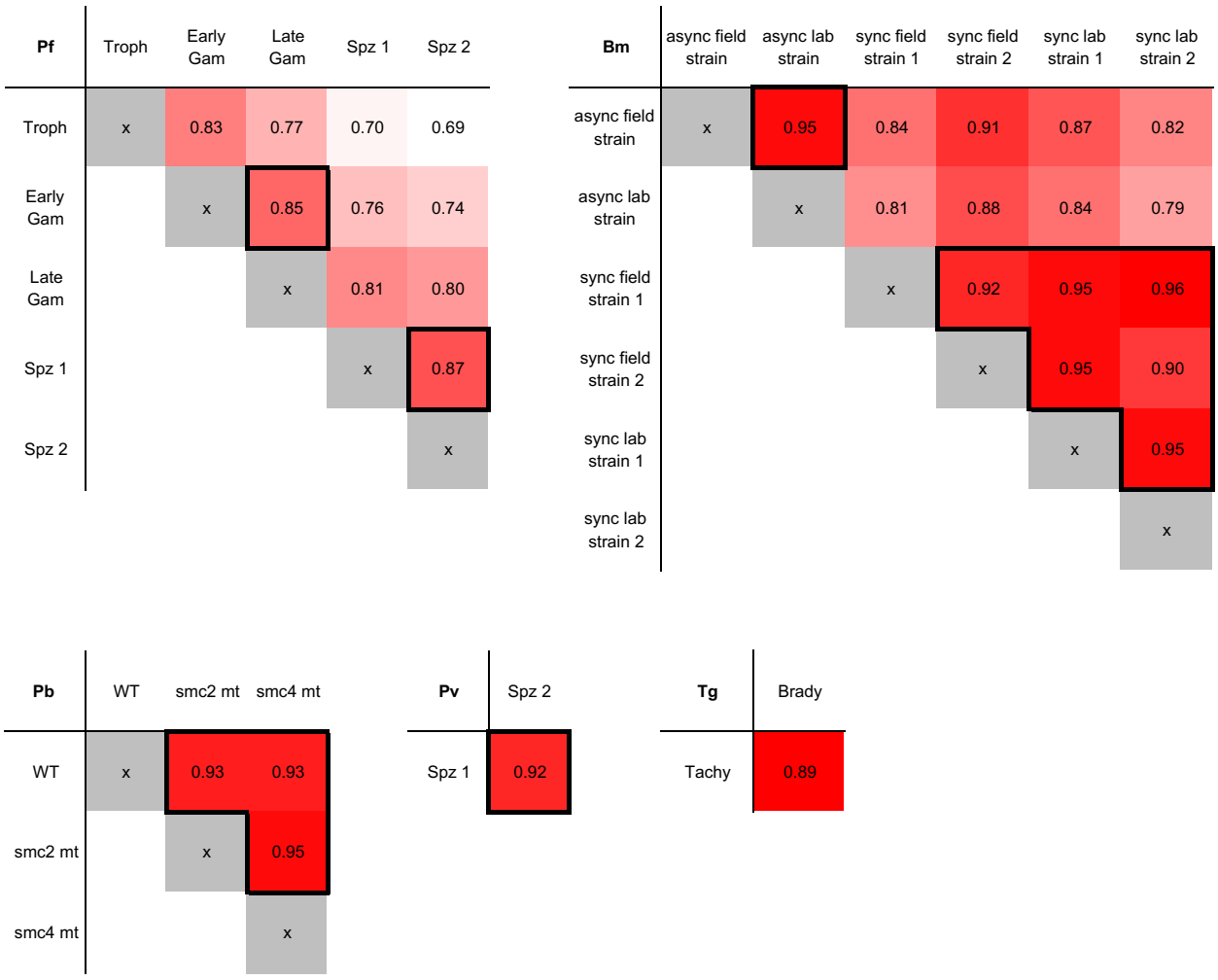


Fig. S1: Correlation between samples of the same organism. Concordance scores calculated using the package GenomeDISCO are shown for samples of the same organism from different stages, different strains, or biological replicates. Samples with boxed values were combined for downstream analysis.

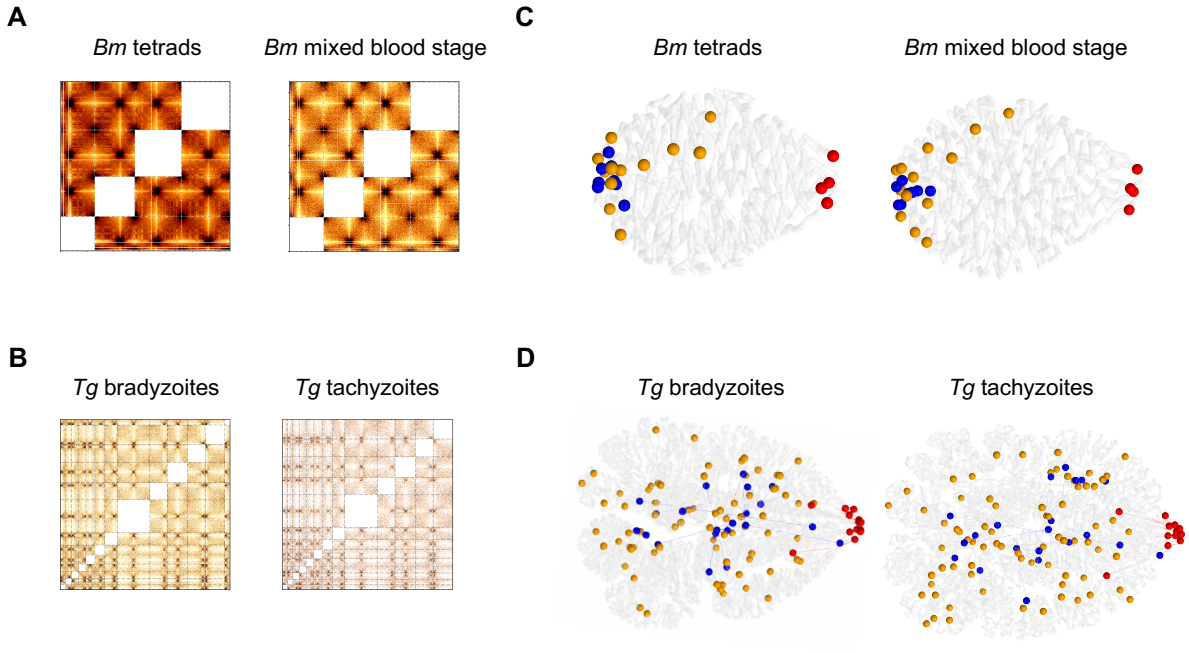


Fig. S2: Hi-C data and 3D genome modeling. A-B) Normalized interchromosomal contact count heatmaps at 10-kb resolution for *B. microti* (A) and *T. gondii* (B) samples. Chromosomes are lined up in numerical order starting with chr1 and chrIa, respectively, in the bottom left corner. Individual chromosomes are delineated by dashed lines. Intrachromosomal contacts are not displayed, hence the white squares along the diagonal of each heatmap. **C-D)** Representative 3D models for genomes of for *B. microti* (C) and *T. gondii* (D) samples. Chromosomes are shown as transparent white ribbons. Centromeres are indicated with red spheres, telomeres with blue spheres and virulence genes with orange spheres. Contact count patterns and genome organization are very similar for samples from the same organism.

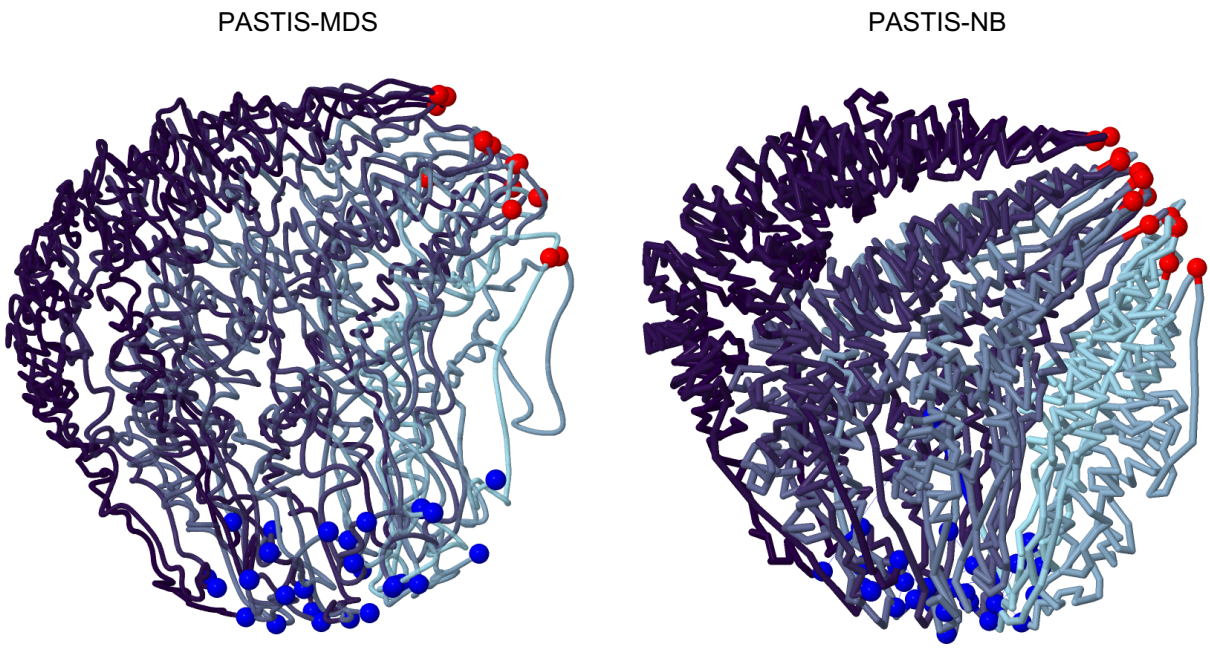


Fig. S3: The effect of different modeling approaches on inference of 3D models of genome organization. The *P. falciparum* trophozoite genome was modeled with PASTIS using the multidimensional scaling approach (left) or the binomial distribution model (right). Major hallmarks of genome organization are conserved between the two approaches. The negative binomial distribution model generates more accurate and robust structures when compared to the multidimensional scaling model, and was therefore used for inferring all models in this study.

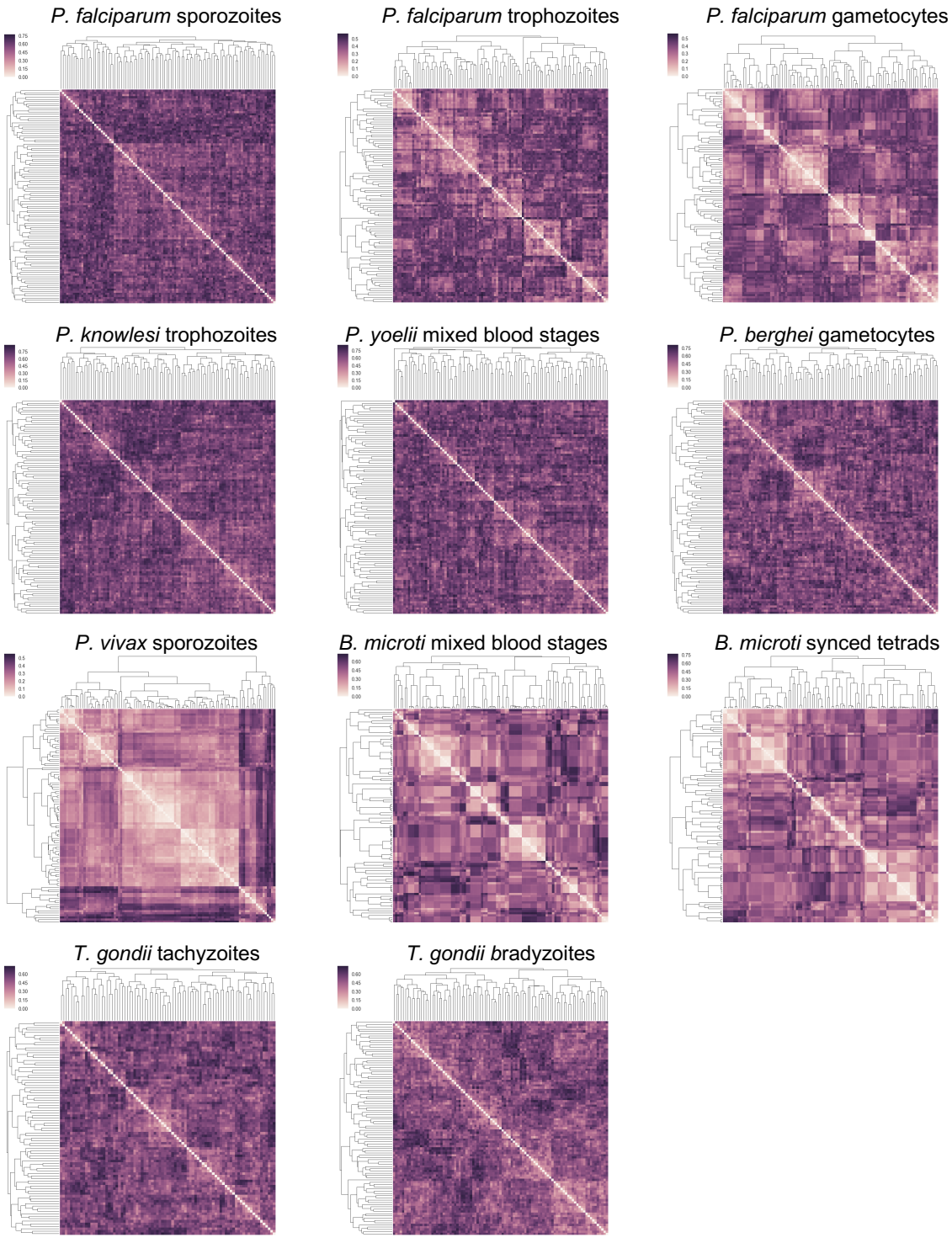


Fig. S4: Robustness of 3D models to random initializations of PASTIS. For each sample, 100 random initializations of the 3D coordinates were used to generate 3D models. The resulting 100 structures for each sample were transformed and clustered, and pairwise disparity scores were visualized in hierarchically clustered heatmaps.

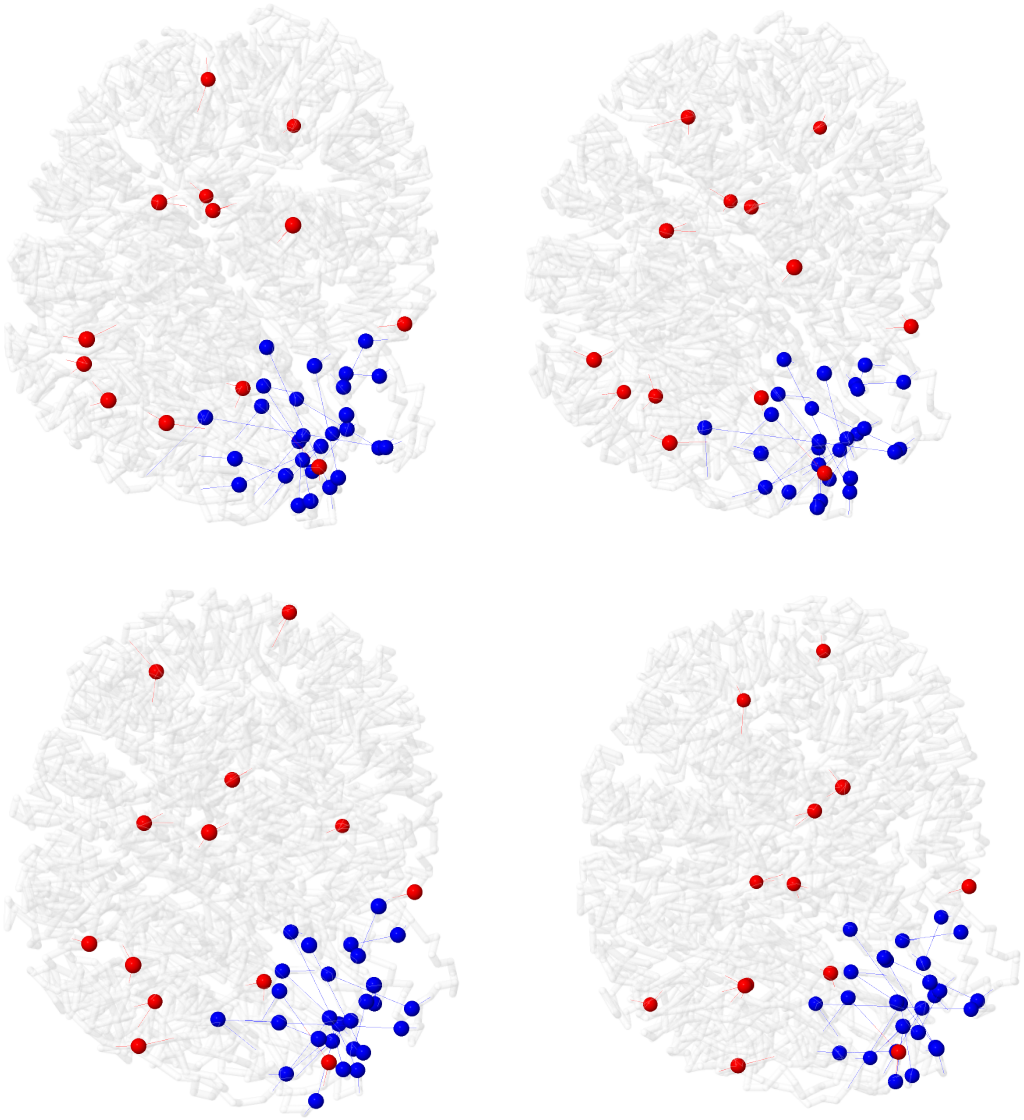


Fig. S5: Representative *P. vivax* 3D models from different initializations. For *P.* vivax, the 100 random initialization of 3D genome modeling resulted in four major clusters of genome structures. Representative models from each of these four clusters are shown here. The main hallmarks of genome organization are preserved in all four 3D models. Centromeres are indicated in red while telomeres are colored in blue.

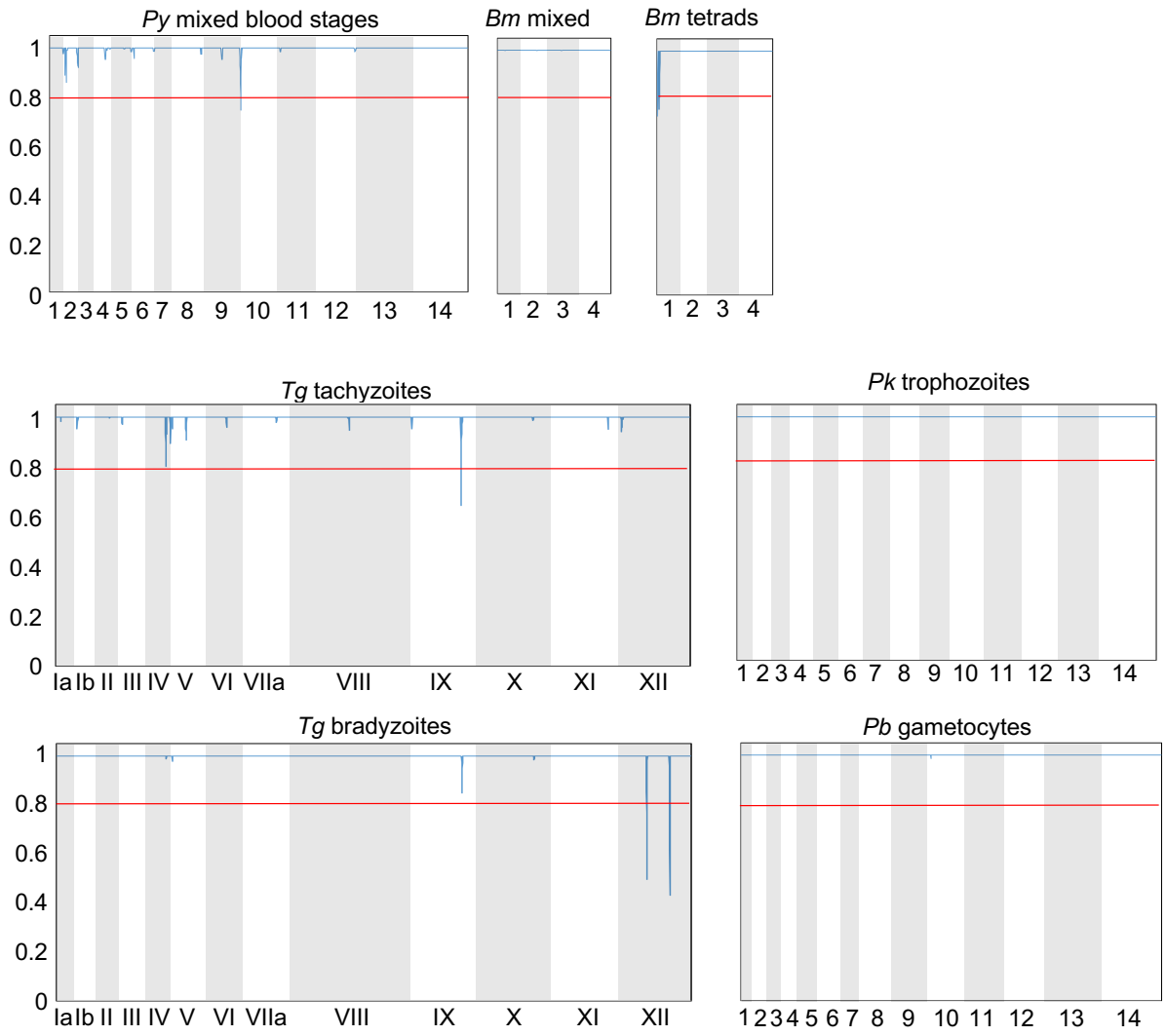


Fig. S6: Misassembly metrics for all samples used in this study. The misassembly metric $S_{observed}/S_{expected}$ was plotted for all chromosomes. To avoid artifacts induced by mappability issues, the score for bins with <50% mappability was set at 1. The threshold for misassembly issues was set at 0.8 and is highlighted with a red line. Chromosomes are indicated with grey and white shading. Results for *P. falciparum* and *P. vivax* samples are included in a recent publication (10).

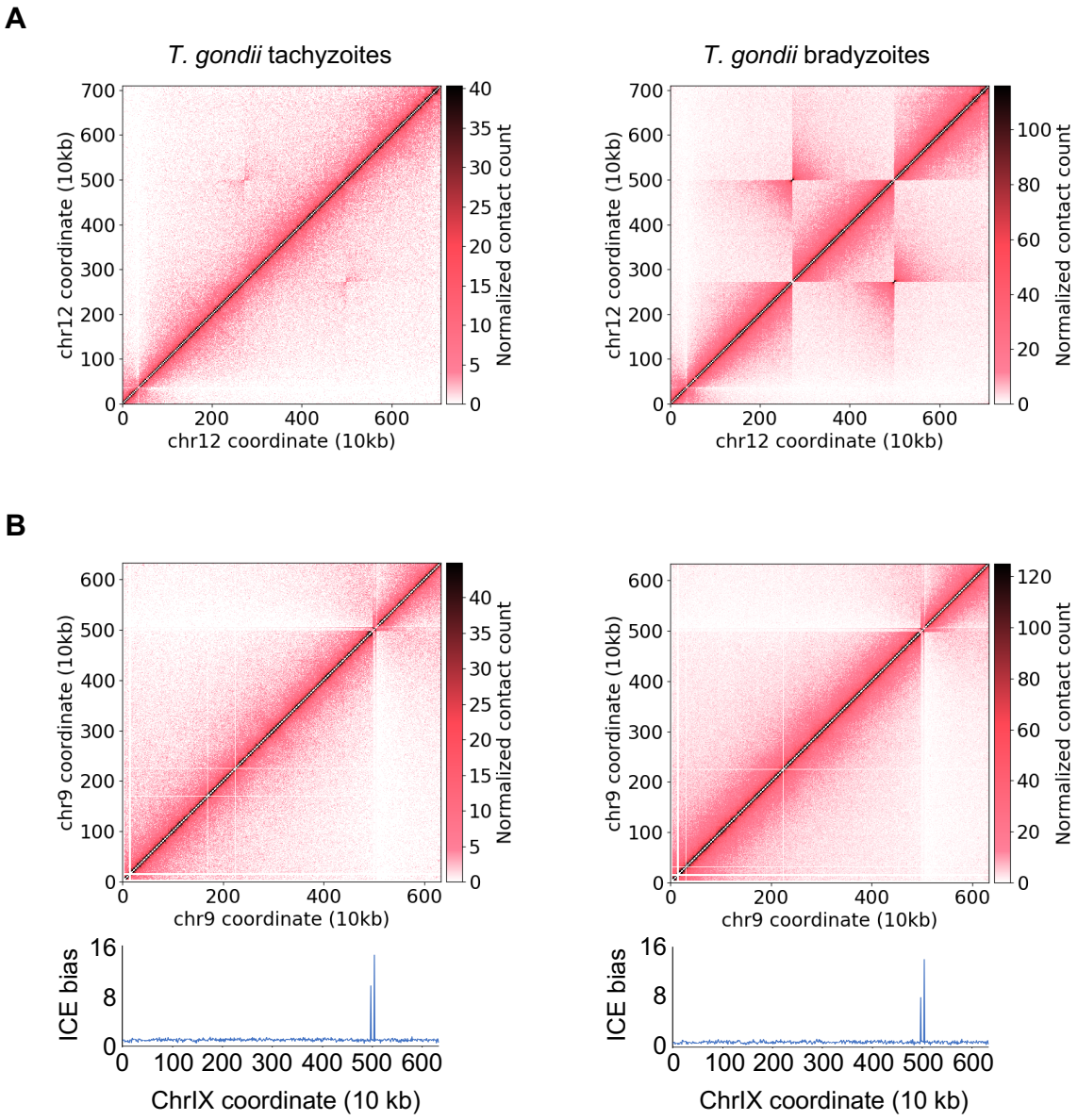


Fig. S7: Detection of inversion and duplication events in the *T. gondii* genome. A) Inversion in chrXII observed in ~10% of the population of tachyzoites (wild-type ME49 strain) and ~100% of the population of bradyzoites (transgenic ME49 strain) used in this study. **B)** Possible amplification of genome sequences in bins 498 and 505 of chrIX. The normalized contact 10-kb resolution heatmaps show an aberrant signal around bin 500 (top). The ICE bias tracks show a higher than expected number of interactions for bins 498 and 505 (bottom).

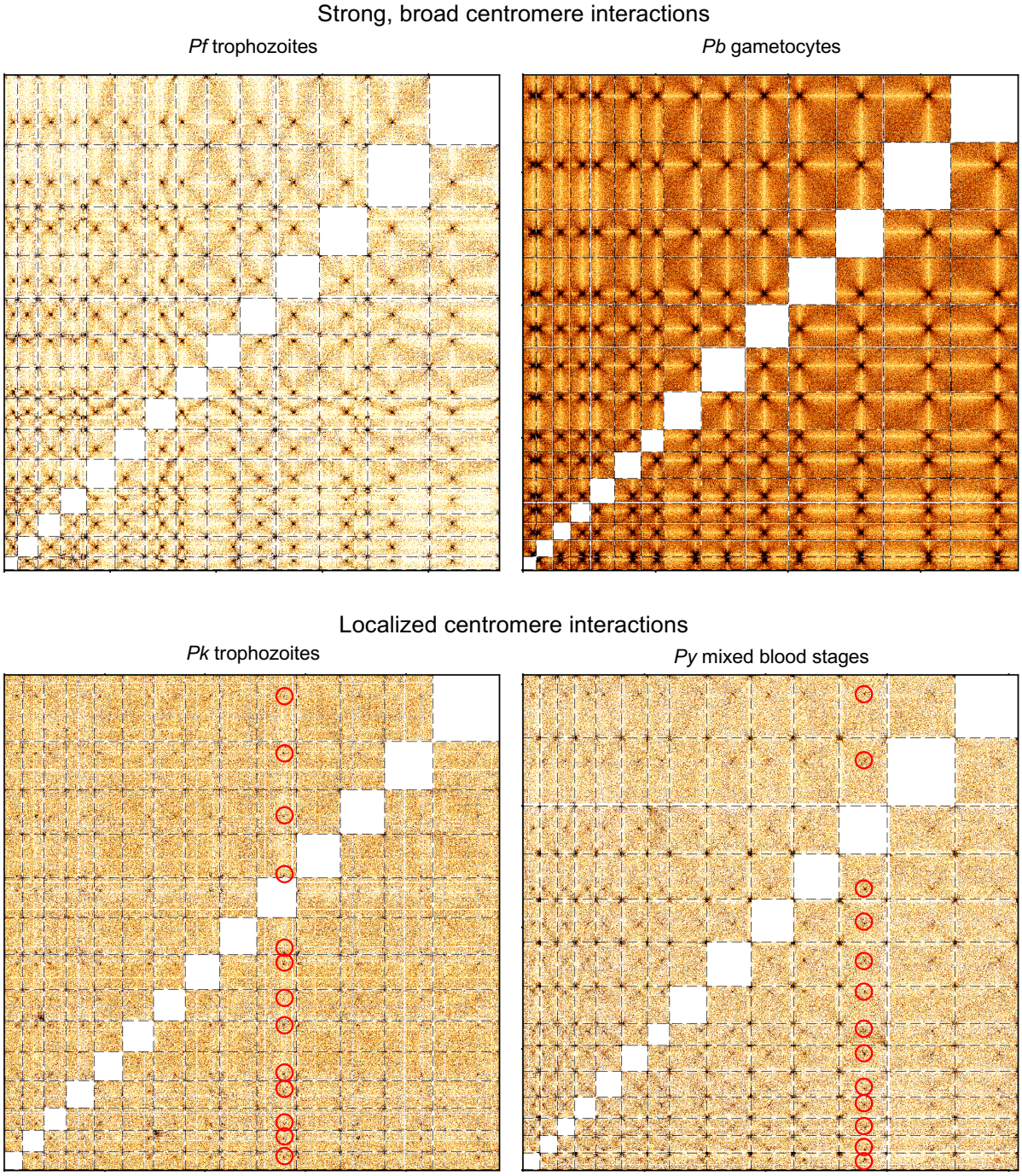


Fig. S8: Differences in centromere interaction patterns. Interchromosomal contact heatmaps of *P. falciparum* and *P. berghei* show strong centromere interactions that do not only involve the centromere itself, but also the surrounding regions (top row). In *P. knowlesi* and *P. yoelii*, centromere interactions are more localized (bottom row). Examples of centromeric interactions are indicated by red circles for one chromosome in each of these two organisms.

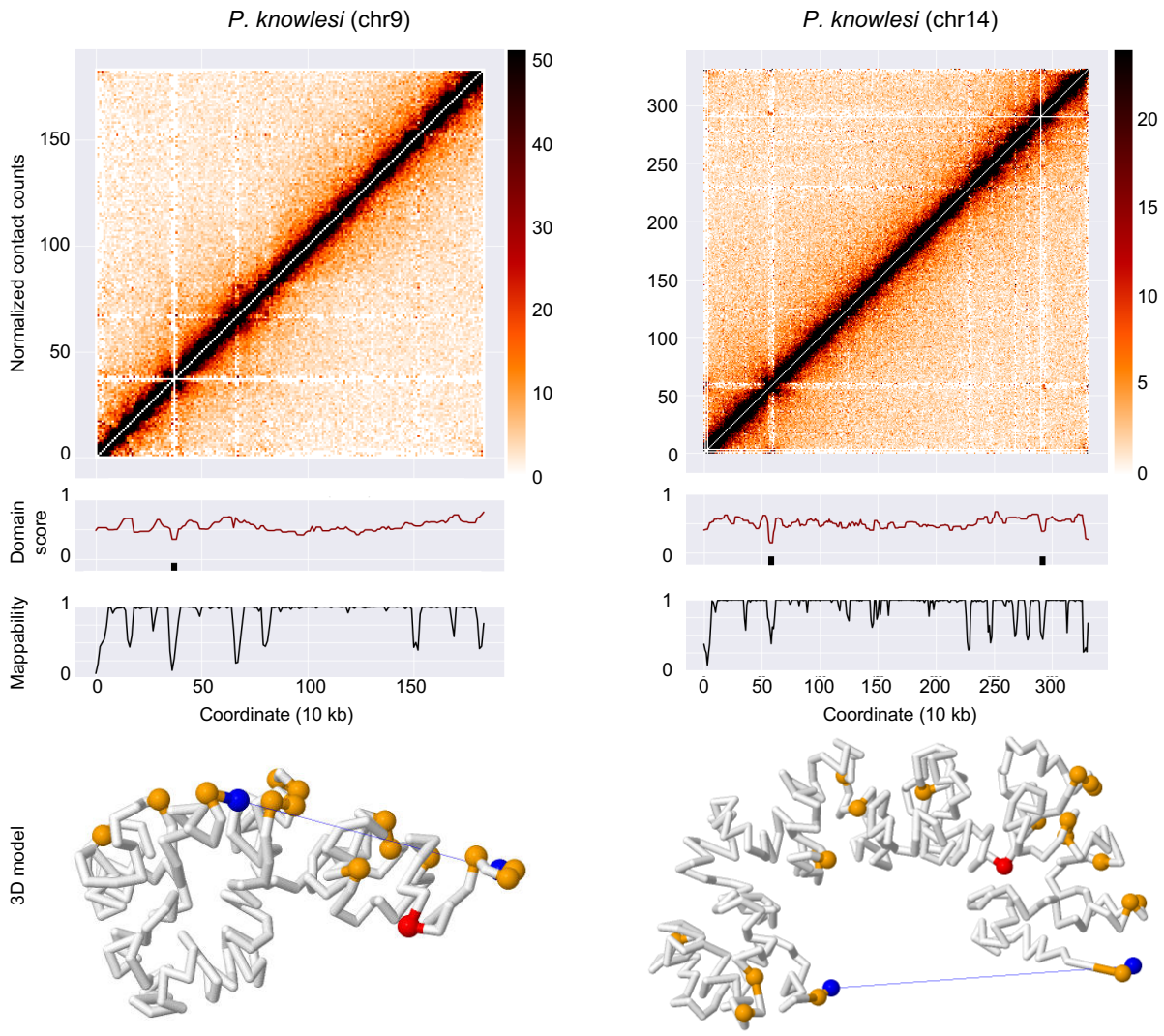


Fig. S9: Formation of domain-like structures and chromosome loops by *SICAvar* genes. Top row: normalized intrachromosomal contact count heatmaps at 10-kb resolution for *P. knowlesi* chr9 and chr14. Second row: domain score tracks. Dips in the tracks that reached the threshold of a domain-like structure are marked with a black box. Third row: mappability tracks. Bottom row: individual chromosome conformation extracted from the 3D model of the full genome.

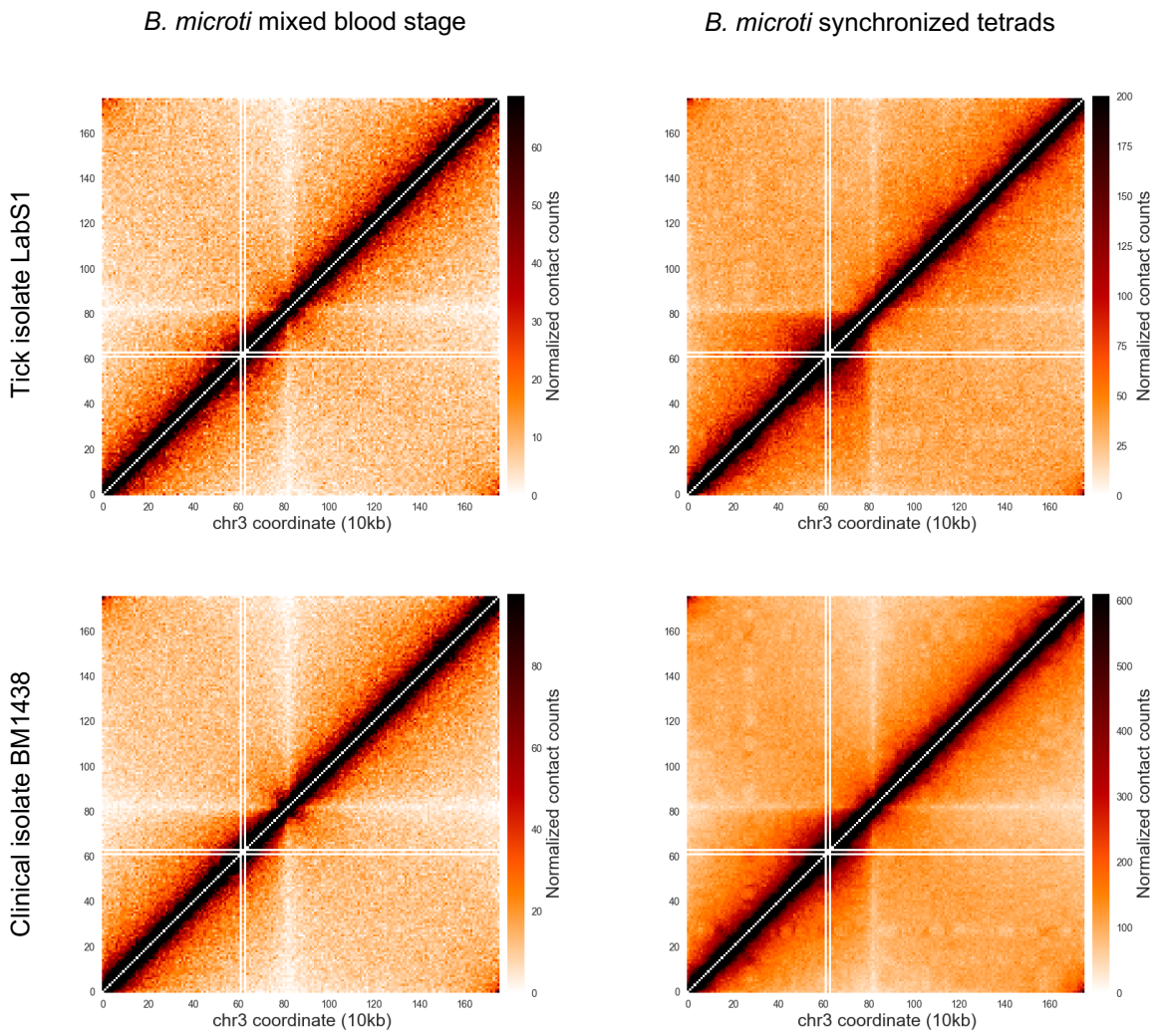
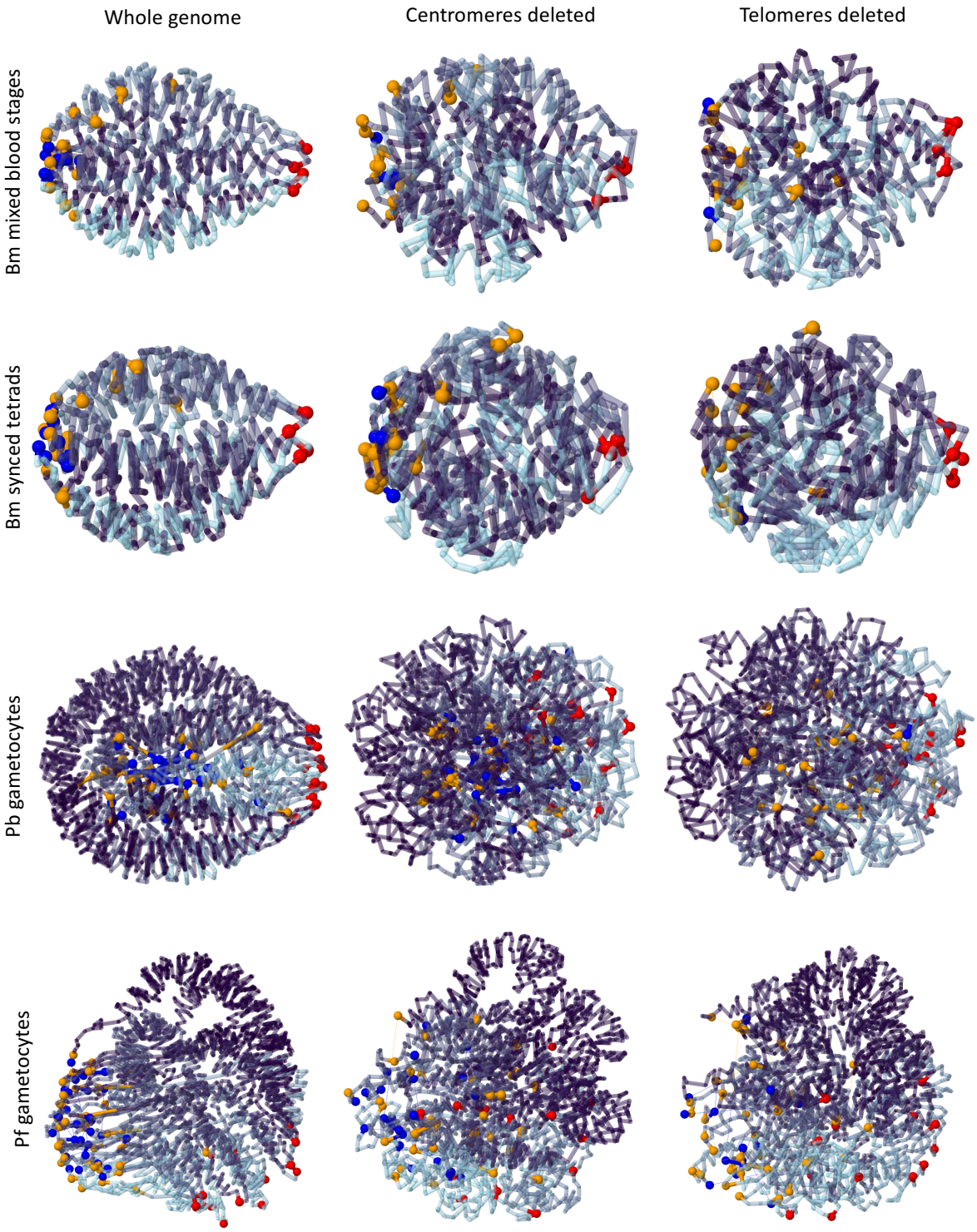
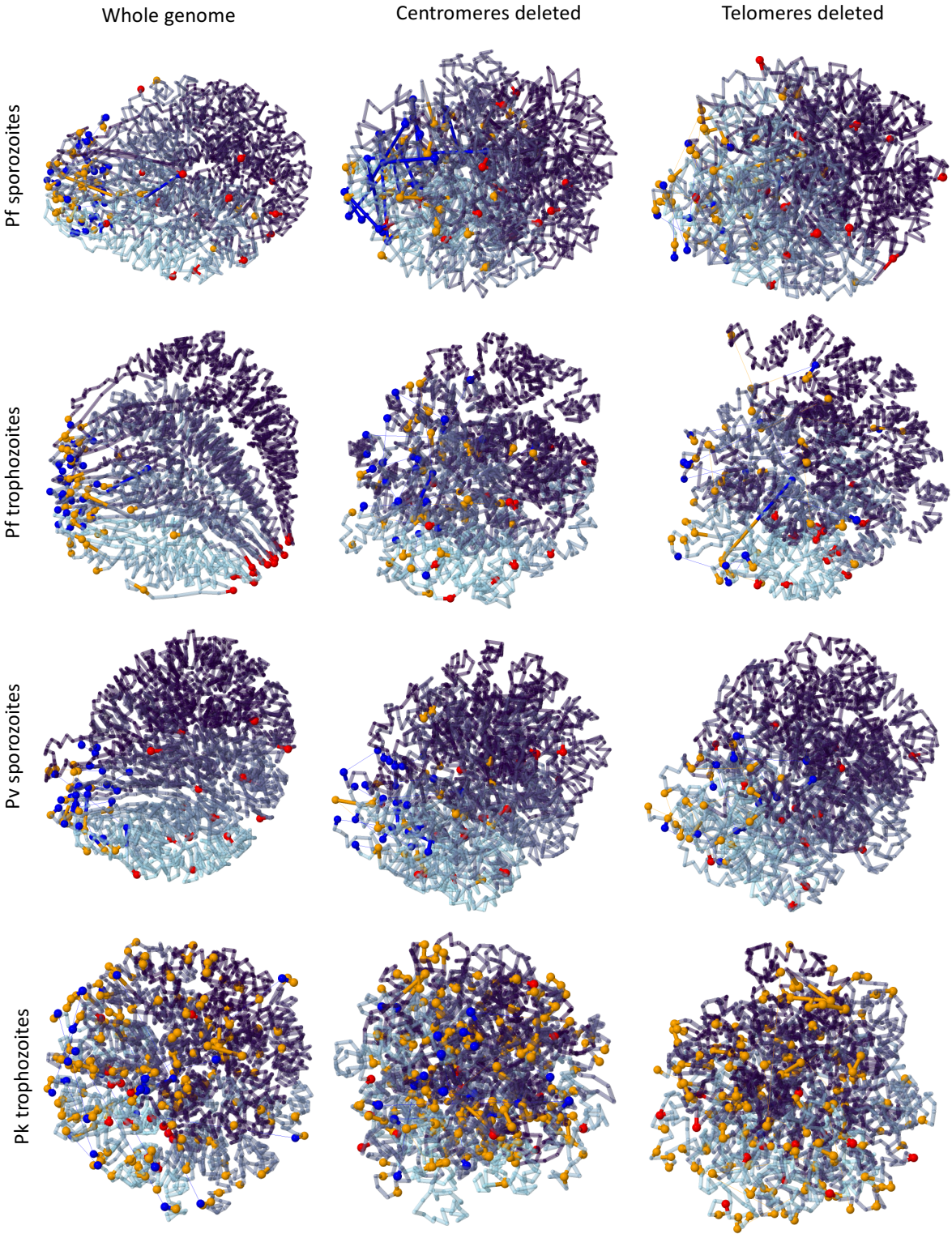


Fig. S10: Comparison between Hi-C data obtained from nonsynchronous and synchronous *B. microti* cultures. Normalized contact heatmaps of chromosome 3 at 10-kb resolution obtained from mixed blood stage culture (left) and from synchronized tetrads (right). The synchronized sample shows contact count patterns that are not visible in the nonsynchronous sample, emphasizing the requirement for the use of tightly synchronized samples to detect stage-specific, cell-cycle dependent or otherwise transient interactions.


Whole genome
Centromeres deleted
Telomeres deleted
Bm mixed blood stages
Bm synced tetrads
Pb gametocytes
Pf gametocytes


Whole genome
Centromeres deleted
Telomeres deleted
Pf sporozoites
Pf trophozoites
Pv sporozoites
Pk trophozoites

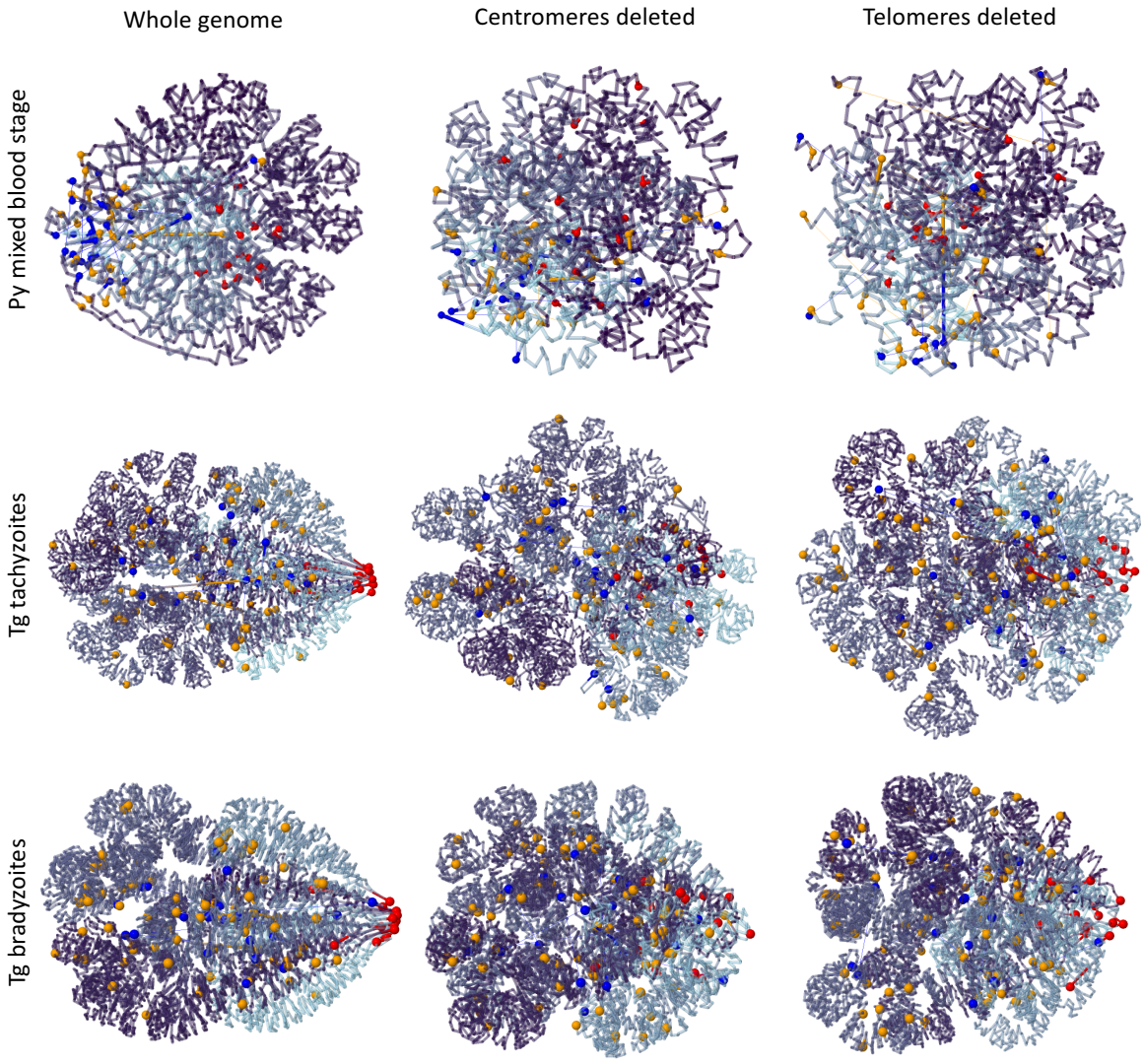


Fig. S11: Effect of deleting hallmarks of genome organization on the 3D model. For each sample, 3D models are shown for the full-length genome (left), the genome in which all bins containing centromeric sequence were deleted (center), and the genome in which 4 bins on all chromosome ends were deleted (right).

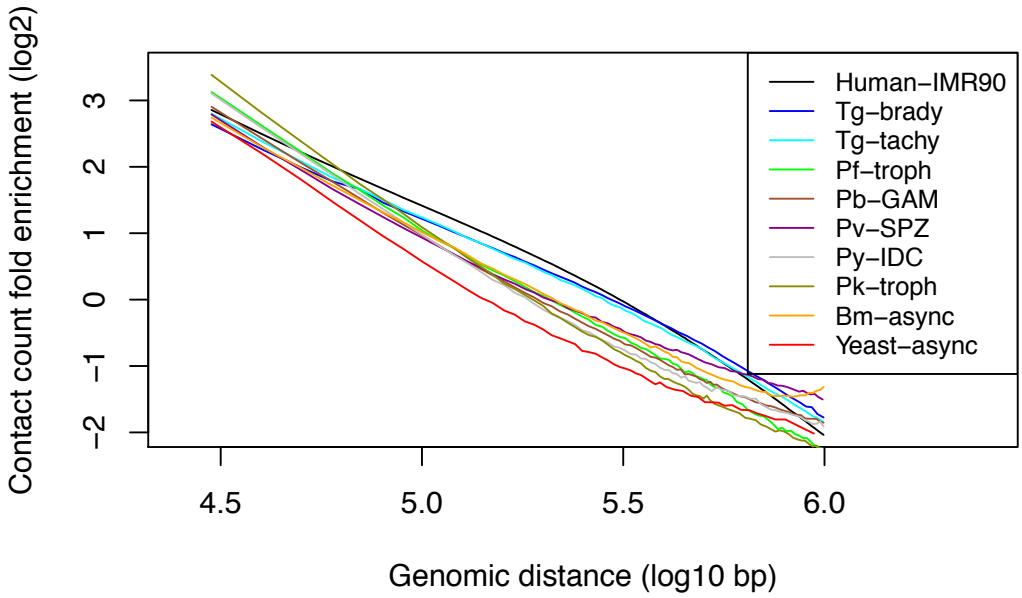


Fig. S12: Contact count fold enrichment as a function of genomic distance for all organisms included in this study. As references, data from yeast and human Hi-C experiments have also been included. The relationship between contact count probability and genomic distance in *T. gondii* is more similar to the human genome than to other apicomplexan parasites.

Table S1: Description of organisms and source material included in this study.

Species	Strain	Natural hosts	Source	Stage	Replicates/strains
Plasmodium falciparum	Lab strain 3D7	Human/ Mosquito	In vitro culture	Trophozoites	1 sample
Plasmodium falciparum	Lab strain NF54	Human/ Mosquito	In vitro culture	Gametocytes	2 samples
Plasmodium falciparum	Lab strain NF54	Human/ Mosquito	*A. stephensi* mosquitoes	Sporozoites	2 biological replicates
Plasmodium vivax	Field strain	Human/ Mosquito	*A. cracens* mosquitoes	Sporozoites	2 biological replicates
Plasmodium knowlesi	Lab strain Pk1(A+)	Long-tailed macaque/ Mosquito	Rhesus macaques	Trophozoites	1 sample
Plasmodium berghei	Lab strain ANKA	Rodents/ Mosquito	Mouse	Gametocytes	3 strain variants
Plasmodium yoelii	Lab strain XNL	Rodents/ Mosquito	Mouse	Mixed blood stage	1 sample
Babesia microti	Tick isolate LabS1 + clinical isolate Bm1438	Warm-blooded animals/ ticks	Mouse	Mixed blood stage	2 strains, 1 sample each
Babesia microti	LabS1 + Bm1438	Warm-blooded animals/ ticks	Ex vivo mouse blood cultures	Synchronized tetrads	2 strains, 2 biological replicates each
Toxoplasma gondii	Lab strain ME49	Warm-blooded animals	In vitro culture	Tachyzoites	1 sample
Toxoplasma gondii	Lab strain ME49	Warm-blooded animals	In vitro culture	Bradyzoites	1 sample

Table S2: Numbers of sequence reads generated in Hi-C experiments and valid interaction pairs retained for downstream analyses.

Merged Sample	Strains/Conditions/Replicates	Raw paired-end reads	Aligned Pairs	Unique Valid Pairs	Percentage of valid pairs		Valid Pairs Within 1kb	Intrachr 1kb-10kb / Total Valid Pairs > 1kb	Intrachr >10kb / Total Valid Pairs > 1kb	Interchr / Total Valid Pairs > 1kb	Publication
Babesia microti mixed blood stage	Mixed strain Bm1438	49,853,688	11,307,312	2,875,817	5.77%	*	385,553	0.167	0.386	0.447	This work
	Mixed strain LabS1	70,606,554	22,106,352	4,427,842	6.27%	*	743,888	0.178	0.371	0.451	
	TOTAL	**120,460,242**	**33,413,664**	**7,303,659**	6.06%	*	1,119,708	0.174	0.377	0.449	
Babesia microti synchronized tetrads	Synced strain Bm1438 rep 1	24,031,449	13,160,646	9,443,749	39.30%		477,626	0.127	0.358	0.515	This work
	Synced strain Bm1438 rep 2	29,384,425	17,111,462	12,977,882	44.17%		620,922	0.145	0.368	0.487	
	Synced strain LabS1 rep 1	68,824,017	40,499,550	29,764,580	43.25%		1,397,860	0.130	0.344	0.525	
	Synced strain LabS1 rep 2	60,108,185	33,887,411	21,601,639	35.94%		1,152,732	0.119	0.335	0.546	
	TOTAL	**182,348,076**	**104,659,069**	**73,787,850**	40.47%		3,599,286	0.129	0.348	0.524	
Plasmodium berghei gametocytes	Smc2 transgenic strain	82,330,220	45,142,471	25,598,041	31.09%		3,395,677	0.186	0.241	0.573	This work
	Smc4transgenic strain	37,336,346	16,429,012	9,565,745	25.62%		1,245,809	0.168	0.243	0.590	
	WT	48,635,834	27,470,015	10,110,737	20.79%		2,367,279	0.198	0.300	0.502	
	TOTAL	**168,302,400**	**89,041,498**	**45,274,523**	26.90%		6,981,165	0.184	0.253	0.562	
Plasmodium falciparum gametocytes	Early gametocytes	150,070,973	100,049,648	61,713,026	41.12%		4,902,051	0.227	0.461	0.312	Bunnik et al *Nature Communications* 2018
	Late gametocytes	75,338,560	46,265,208	17,594,303	23.35%		3,371,750	0.204	0.340	0.456	
	TOTAL	**225,409,533**	**146,314,856**	**79,307,329**	35.18%		8,155,336	0.222	0.437	0.341	
Plasmodium falciparum sporozoites	Sporozoites rep 1	130,539,746	15,697,495	1,740,915	1.33%	*	546,300	0.132	0.275	0.593	Bunnik et al *Nature Communications* 2018
	Sporozoites rep 2	51,856,178	7,652,584	925,159	1.78%	*	281,810	0.131	0.277	0.592	
	TOTAL	**182,395,924**	**23,350,079**	**2,666,074**	1.46%	*	734,277	0.131	0.275	0.593	
P. falciparum trophozoites	**TOTAL**	**60,567,770**	**25,123,926**	**6,968,959**	11.51%		756,940	0.211	0.392	0.397	Ay et al *Genome Research* 2014
P. knowlesi trophozoites	**TOTAL**	**33,857,803**	**15,882,302**	**11,243,440**	33.21%		2,610,076	0.272	0.281	0.447	Lapp et al *Parasitology* 2017
Plasmodium vivax sporozoites	Sporozoites rep 1	138,095,734	25,859,628	9,976,858	7.22%	*	1,770,721	0.154	0.277	0.569	Bunnik et al *Nature Communications* 2018
	Sporozoites rep 2	160,681,970	15,860,169	6,145,121	3.82%	*	1,018,276	0.153	0.281	0.566	
	TOTAL	**298,777,704**	**41,719,797**	**16,121,979**	5.40%	*	2,787,874	0.154	0.278	0.568	
P. yoelii mixed blood stage	PanK1 transgenic strain	58,044,078	18,912,344	5,533,378	9.53%	*	2,047,279	0.231	0.312	0.457	This work
T. gondii bradyzoites	**TOTAL**	**117,710,987**	**54,058,885**	**40,176,192**	34.13%		2,115,709	0.172	0.426	0.402	This work
T. gondii tachyzoites	**TOTAL**	**33,116,669**	**19,060,530**	**13,075,539**	39.48%		1,290,368	0.183	0.436	0.382	This work

* low percentages in these samples are due to host DNA contamination

References for SI

1. Hart RJ, *et al.* (2016) Genetic Characterization of Plasmodium Putative Pantothenate Kinase Genes Reveals Their Essential Role in Malaria Parasite Transmission to the Mosquito. *Sci Rep* 6:33518.
2. Beetsma AL, van de Wiel TJ, Sauerwein RW, & Eling WM (1998) Plasmodium berghei ANKA: purification of large numbers of infectious gametocytes. *Exp Parasitol* 88(1):69-72.
3. Kafsack BF, Beckers C, & Carruthers VB (2004) Synchronous invasion of host cells by Toxoplasma gondii. *Mol Biochem Parasitol* 136(2):309-311.
4. Lapp SA, *et al.* (2017) PacBio assembly of a Plasmodium knowlesi genome sequence with Hi-C correction and manual annotation of the SICAvar gene family. *Parasitology* 145(1):1-14.
5. Otto TD, *et al.* (2014) A comprehensive evaluation of rodent malaria parasite genomes and gene expression. *BMC Biol* 12:86.
6. Bunnik EM, *et al.* (2013) Polysome profiling reveals translational control of gene expression in the human malaria parasite Plasmodium falciparum. *Genome Biol* 14(11):R128.
7. Westenberger SJ, *et al.* (2010) A systems-based analysis of Plasmodium vivax lifecycle transcription from human to mosquito. *PLoS Negl Trop Dis* 4(4):e653.
8. Hoo R, *et al.* (2016) Integrated analysis of the Plasmodium species transcriptome. *EBioMedicine* 7:255-266.
9. Silva JC, *et al.* (2016) Genome-wide diversity and gene expression profiling of Babesia microti isolates identify polymorphic genes that mediate host-pathogen interactions. *Sci Rep* 6:35284.
10. Bunnik EM, *et al.* (2018) Changes in genome organization of parasite-specific gene families during the Plasmodium transmission stages. *Nat Commun* 9(1):1910.