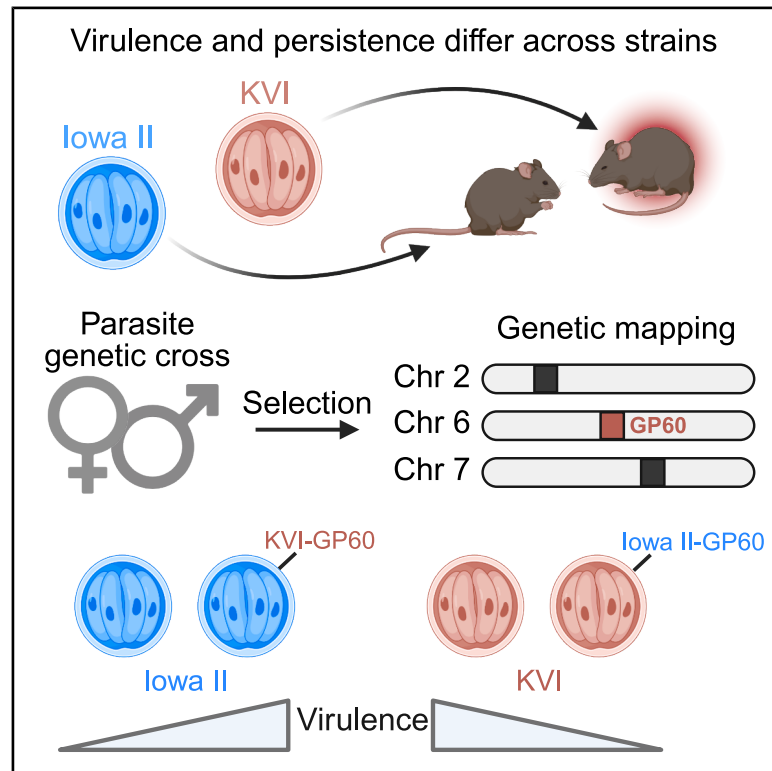# Genetic crosses reveal genomic loci responsible for virulence in *Cryptosporidium parvum* infection

Graphical abstract

Authors

Sebastian Shaw, Xue Li,
Gracyn Y. Buenconsejo, ...,
Daniel P. Beiting, Timothy J.C. Anderson,
Boris Striepen

Correspondence

striepen@upenn.edu

In brief

Shaw et al. used genetic crosses and bulk segregant analysis to identify parasite loci that control virulence in *Cryptosporidium parvum*. Their work establishes a forward genetics framework in this important parasite and reveals *gp60* as a key determinant of virulence.

Highlights

- *Cryptosporidium parvum* strains produce infections with different outcomes

- Genetic crosses identify loci associated with virulence and persistence

- Glycoprotein GP60 determines virulence in a strain-specific manner

- Bulk segregant analysis enables powerful forward genetics for *Cryptosporidium*

CellPress

Article

Genetic crosses reveal genomic loci responsible for virulence in *Cryptosporidium parvum* infection

Sebastian Shaw,[1] Xue Li,[2] Gracyn Y. Buenconsejo,[1] Tiffany H. Zhou,[1] Allison Cohen,[1] Daniel Yasur-Landau,[3] Rui Xiao,[1] Daniel P. Beiting,[1] Timothy J.C. Anderson,[2] and Boris Striepen[1,4,*]

[1]Department of Pathobiology, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA, USA
[2]Texas Biomedical Research Institute, San Antonio, TX, USA
[3]Division of Parasitology, Kimron Veterinary Institute, Bet Dagan, Israel
[4]Lead contact
*Correspondence: striepen@upenn.edu
https://doi.org/10.1016/j.celrep.2025.116315

SUMMARY

The relationship between parasite genotype and pathogenesis is largely unknown for *Cryptosporidium*, a leading cause of diarrheal disease in children. An array of parasites with similar genomes produces varied disease outcomes in different hosts. Here, we isolate and characterize *Cryptosporidium parvum* strains that show marked differences in virulence and persistence in mice. Taking advantage of the sexual life cycle of this eukaryotic pathogen, we use genetic crosses to discover the underlying chromosomal loci. Whole-genome sequencing and bulk segregant analysis of infection-selected progeny mapped three loci on chromosomes 2, 6, and 7 associated with the ability to colonize and persist in mice and the positions of drug resistance genes. The chromosome 6 locus encodes the hyper-polymorphic surface glycoprotein GP60. Reverse genetic studies in both parental strains demonstrate that GP60 controls parasite burden and virulence, but not persistence, and reveal the dominance of the less virulent allele, suggesting it restricts virulence.
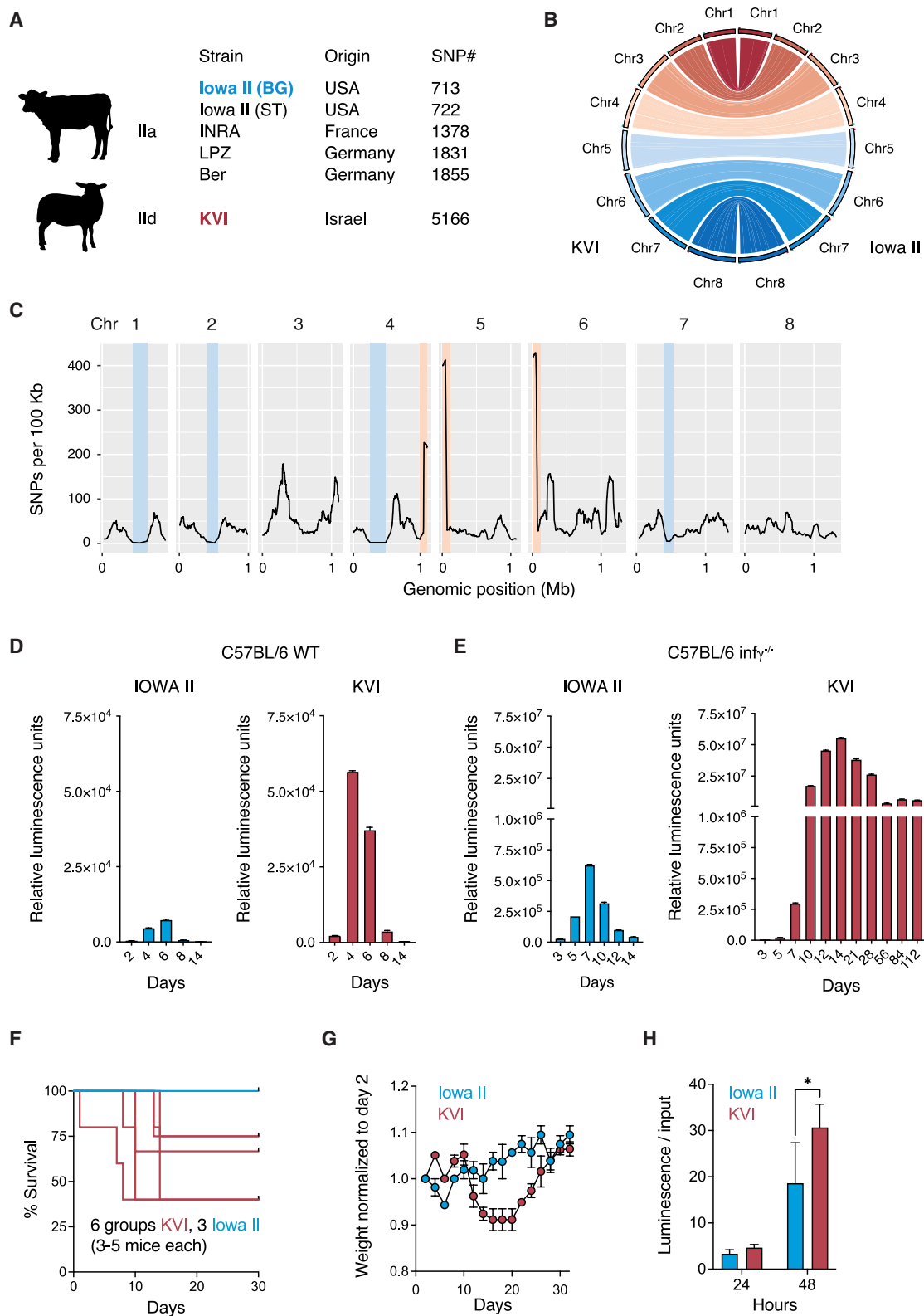
INTRODUCTION

The individual outcomes of infectious diseases are highly variable and can range from fatal illness to the absence of notable signs and symptoms. While the basis of this variability is not always understood, multiple, and at times interacting, contributions of host, pathogen, and environment have been reported. This is also the case for infection with the single-celled eukaryotic parasite *Cryptosporidium*. *Cryptosporidium* is a leading cause of diarrheal disease and early childhood morbidity and mortality.[1,2] Infection occurs around the world and under a variety of epidemiological scenarios. Transmission can be linked to drinking or recreational water contact, direct human-to-human spread, or zoonotic spillover from a variety of domesticated animals. Environmental factors control the endurance and transmissive capacity of the infectious oocyst stage,[3] and abundance or lack of nutrition dramatically impacts the course of infection in children.[4–6] Host genetics,[7] immune status,[8] and co-infection and microbiome[9] status are similarly important. Individuals who lack T cell function, either from birth or due to HIV-AIDS or immunosuppressive treatment, are highly susceptible to severe cryptosporidiosis,[10] while humans and animals exposed to repeated infection develop non-sterile immunity that protects them from disease.[8,11,12] Many species and strains of *Cryptosporidium* have been described,[13] and numerous recent reports highlight strain-specific (and thus likely pathogen-mediated) dif-

ferences in host range, transmissibility, and virulence[12,14–16]; however, the underlying molecules and mechanisms remain largely unknown.

As is typical for apicomplexan parasites, the *Cryptosporidium* life cycle features phases of asexual and sexual replication; however, in this parasite, both phases occur in the same single host. Recent studies have demonstrated that progression from the initial asexual phase to sexual reproduction is obligate, occurs roughly every 48 h, and that this progression and the resulting sex are essential to continuous infection of the host.[17–19] Remarkably, *Cryptosporidium* will thus experience significant genome recombination and meiotic segregation numerous times over the course of a single infection.[20] We take advantage of this unique sexual life cycle to associate specific parasite genes with the traits of virulence and persistence through genetic crosses and linkage analysis. Genetic crosses have been transformational to understanding drug resistance and virulence in the apicomplexan parasites: *Plasmodium* and *Toxoplasma*.[21,22] We recently developed a crossing model for *Cryptosporidium* in mice, where selection for dual drug resistance is used to isolate recombinant progeny to the exclusion of selfing parental types.[23] Crosses were feasible within and between species of diverging host specificity. However, while pairing *C. parvum* and *C. tyzzeri* offered significant phenotypic differences and numerous SNPs for mapping, the overall efficiency of mating was low, resulting in modest genomic recombination.[23] Here,

(legend on next page)

we establish a system within the species of *C. parvum* that balances significant genomic and phenotypic diversity with high mating efficiency to build a powerful forward genetic model to dissect host-parasite interaction.

RESULTS

Collection and genome sequencing of *C. parvum* strains

To enable efficient genetic mapping of virulence within the species of *C. parvum*, we collected *C. parvum* strains previously isolated from cattle in the United States and Europe (generously shared with us by Drs. Laurent, Olias, and Krücken), and we returned to a study site in Israel where ovine *C. parvum* had been found and characterized previously.[24] Feces from lambs in Israel were collected and tested for the presence of parasites by PCR. Oocysts were isolated from positive samples, and parasites were amplified in *Ifnγ*[−/−] mice (see methods section for further details; all mice used in this study were C57BL/6 background). All parasite strains were subjected to whole-genome sequencing, followed by alignment to the most recent telomere-to-telomere assembly of the *C. parvum* Iowa II reference genome[25] and SNP calling (Figure 1A). Analysis of the diagnostic glycoprotein 60 (GP60) locus showed strain KVI (Kimron Veterinary Institute) to be IId genotype and all others to be IIa[26] (Figure S1). The KVI genome showed the largest number of SNPs (5,166) when compared to the *C. parvum* Iowa II reference genome (Figures 1A and 1C show the SNP distribution, highlighting areas of high [pink] and low [blue] density). We generated both Illumina short reads and Nanopore long reads, which were then used for *de novo* genome assembly. We found this assembly to be highly syntenic to the recent *C. parvum* telomere to telomere genome[25] (Figures 1B and S3; Table S1; note that for KVI, we were unable to assemble five of the 16 telomeres). Based on the number of distinguishing SNPs and previous reports of phenotypic difference between parasites of IIa and IId background,[14] we chose to pursue Iowa II (Bunchgrass Farms) and KVI as a mating pair to map loci underlying host-pathogen interaction.

Iowa II and KVI show marked differences in virulence and persistence during mouse infection

To facilitate rigorous side-by-side comparison, we engineered the KVI strain to express nanoluciferase (Nluc) (Figure S2A) to match Iowa II reporter parasites[23] already in hand. We integrated a cassette encoding Nluc as well as the tdNeonGreen fluorescent protein into the phenylalanyl-tRNA synthase locus, selecting for resistance to BRD7929.[23] Age- and sex-matched

C57BL/6 wild-type mice were infected with 20,000 oocysts per mouse of either transgenic Iowa II or KVI, and parasite shedding was measured following luciferase activity in the feces. Both parasite strains produced brief infection of modest burden, typical for *C. parvum* in immunocompetent adult mice,[12] with KVI exceeding the burden of Iowa II (Figure 1D). Next, we infected *Ifnγ*[−/−] mice with 10,000 oocysts per mouse, which revealed pronounced strain differences. Iowa II resulted in a self-limiting 2-week course of infection peaking at the level previously reported.[12] By contrast, KVI infection peaked 100-fold higher, and it remained continuously and robustly patent to the latest time point we measured, day 112 of infection (Figure 1E). 25%–50% of *Ifnγ*[−/−] mice infected with KVI died by day 14, the peak of parasite burden, while all Iowa II-infected mice survived (Figure 1F). This difference in severity was also apparent when following the weight of infected mice (Figure 1G). While Iowa II-infected mice showed only a brief and moderate weight loss, those mice infected with KVI lost weight to a nadir at day 14. Mice that survived regained weight afterward despite continued high infection. Lastly, we measured the growth of both strains *in vitro* using a luciferase assay in human HCT-8 cell cultures.[27] Over 48 h of growth, KVI was significantly, yet moderately elevated when compared to Iowa II (Figure 1H).

Genetic cross of Iowa II and KVI parasites

To discover the genes that underlie the difference in mouse infection between the two parasite strains, we conducted genetic crosses. *Ifnγ*[−/−] mice were each infected with 8,000 oocysts per mouse of the less virulent Iowa II and 2,000 oocysts per mouse of the KVI strain and subjected to dual-drug selection with BRD7929 and paromomycin from days 4–10 (Figures 2A and S4A–S4C, cross 1–3). We also conducted crosses in which mice were left untreated (Figures S4D and S4E, cross 4–5). Fecal luciferase activity was recorded to track parasite burden throughout the infections and Figure 2B (drug-treated) and 2C (untreated) show measurements for both types of experiments. Parasite burden was slightly higher in experiments that used drug treatment. This likely reflects the effect of the antibiotic paromomycin on the bacterial microbiome, which is known to enhance parasite infection.[9,27] We observed sustained infection in all experiments, suggesting selection for the superior ability of the KVI parental strain to colonize and persist in mice. We collected oocysts for DNA sequencing across the infection and into the persistence phase and specific time windows for each collection pool, which are indicated by gray boxes in Figures 2 and S4. Genomic DNA was extracted from 1 to 5 million purified

---

Figure 1. Identification of *C. parvum* strains that differ in virulence and persistence

(A) Type IIa and a newly collected IId strain of *C. parvum* were subjected to whole-genome sequencing, with KVI showing the largest number of SNPs when compared to the reference genome.

(B–E) (B) The KVI genome was highly syntenic with the reference, and (C) shows SNP distribution with areas of especially low (blue) or high SNP density (pink) highlighted. WT (D) or *Ifnγ*[−/−] mice (E) were infected with Iowa II (blue) or KVI (red) parasites expressing Nluc, and parasite burden was measured by following fecal luciferase activity. KVI yields much higher infection and persists (in the absence of IFNγ). One of two replicates is shown; *n* = 3 mice per group per replicate.

(F) Survival of mice infected with the indicated strain, each line shows a cage of 3–5 mice, and 6 cages for KVI and 3 for Iowa II.

(G) Mice infected with the strain indicated were weighed every 2 days, and the weight is plotted normalized to day 2 of the infection. Data are mean ± SD of *n* = 2 biological replicates.

(H) HCT-8 cultures were infected with the indicated strain, and luminescence was measured after 24 and 48 h and plotted normalized to the input used for infection (individual cultures are shown as dots and their mean as bars). Data are mean ± SEM of *n* = 3 biological replicates. Significance was evaluated using two-way ANOVA, *$p < 0.05$.

oocysts of each pool. We conducted whole-genome Illumina sequencing, and variants were called (SNP loci with coverage below 30-fold in either of the compared pools were excluded from analysis).

Drug resistance loci are readily identified by bulk segregant analysis

Genetic crosses in *Plasmodium* and *Toxoplasma* traditionally relied on the isolation of clonal lines of recombinant progeny and assessed genotype and phenotype individually for each clone to establish linkage.[21,22] Here, we turned to bulk segregant analysis as a method that uses progeny populations instead of clones to conduct genetic mapping (for a broader conceptual review of this approach[28]). The key element of bulk segregant analysis is selection; by comparing allele frequencies across the entire genome among progeny exposed to different selective pressures, we can identify loci for which inheritance from one parent appears enriched or depleted in response to pressure. This is well-suited for *Cryptosporidium*, where we can collect parasites across the course of *in vivo* infections with ease but lack a continuous tissue culture system required for cloning.

We first asked whether our crosses might identify the positions of the marker transgenes used. We compared three paromomycin/BRD7929-treated and two untreated crosses (Figure S4), using the genome sequences of the first oocyst pool collected in each experiment. Visual inspection of the aligned Illumina reads showed a loss of the wild-type sequence and fixation of the resistance genes in the *pheRS* and *TK* loci in all drug-selected pools (Figure S5). Next, we counted reads corresponding to the genotypes of each parent and calculated allele frequencies for all variable SNPs. These frequencies were then used for bulk segregant analysis. We identified two quantitative trait loci (QTLs) (Figure 3A) on chromosomes 3 and 5, respectively, each characterized by very high G values (G values represent the level of statistical support at each SNP, and G′ curves average across multiple neighboring SNPs). Zoomed-in views of the allele distribution of these chromosomes show them to be inherited in equal measure from each parent in the absence of drugs (Figures 3B and 3C, blue shades), but to be highly selected upon drug treatment for the parent that carried the respective marker (Figures 3B and 3C, red shades). We found the insertion sites at or in the close vicinity of the highest confidence SNP. We noted a difference in the shape of the allele frequency and G′ value curve for the markers. While the *TK* locus was found in the middle of the overall peak and the 95th percentile confidence interval, *PheRS* was located outside that interval, and right at the transition between plateau and slope of a broader peak (Figures 3D and 3E). We conclude that *Cryptosporidium* crosses combined with bulk segregant analysis readily identify genes under drug selection with high confidence. The results of the crosses are highly reproducible upon multiple repeats (Figures S6 and S7), and SNP-by-SNP analysis may be superior to reliance on broad, weighted moving averages to identify linked genes.

Loci on chromosomes 2, 6, and 7 are linked to virulence and/or persistence

We hypothesized that, analogous to drug treatment, we might be able to use the host to impose selection suitable for mapping by

bulk segregant analysis. We considered that such selection may act immediately, based on differences in the ability to colonize a specific host species or only unfold over time, reflecting differences in parasite persistence or interactions with the host's developing adaptive immunity. To account for these scenarios, we followed the progeny through infection and sampled and whole-genome sequenced longitudinally at multiple points of infection for 40–49 days (Figure S4). We observed three QTL on chromosomes 2, 6, and 7, with marked changes in allele frequency (Figures 4A, 4B, and S7; Table S1 for an overview of G- and G′ values of all five crosses). These changes occurred regardless of whether parasites were subjected to drug treatment or not, and all three reflected preferential inheritance of the KVI genome. Bulk segregant analysis identified highly significant QTL for these positions; however, there were kinetic differences. While the locus on chromosome 6 showed almost immediate impact and reached significance at 11–14 days, the two other loci produced shifts in allele frequencies and significant QTLs only after the infections proceeded to later time points (Figures 4C and 4D).

The locus on chromosome 7 contains dense granule proteins, one of which is required for high parasite burden

The QTL peak on chromosome 7 encompasses multiple genes, including three presumptive secretory proteins and two ABC transporters. The measurements of mRNA abundance comparing the two parental strains found no significant differences for any of these genes (Figure 5A). Three genes *cgd7_4490*, *cgd7_4500*, and *cgd7_4530* were predicted to likely encode dense granule proteins in a recent spatial proteomic study,[29] with *cgd7_4500* showing the highest degree of polymorphism and a dN/dS ratio (non-synonymous to synonymous substitutions) of 2.93 across the sequences currently available on the CryptoDB database. We began analysis by introducing sequences, encoding a C-terminal influenza hemagglutinin (HA) epitope into *cgd7_4490* and *cgd7_4500* in the KVI strain background (Figure S2B). Immunofluorescence assays using anti-HA antibodies using expansion microscopy confirmed that both proteins localize to the dense granule compartment in sporozoites, and we will thus refer to the proteins as dense granule proteins, DG5 and DG6, respectively (Figures 5B and 5C). Following secretion during invasion, DG5 is found in the parasitophorous vacuole surrounding the intracellular parasite, while DG6 is localized to the host parasite interface. Next, we targeted each locus in the KVI strain by introducing an Nluc/*Neo* marker that disrupted the coding sequence (Figure S2B). Drug-resistant parasites were recovered in both cases, albeit with pronounced delay for DG6. Uniform disruption was confirmed by PCR mapping of the loci (Figure S2B). In side-by-side infection experiments in $Inf\gamma^{-/-}$ mice, we noted an 85-fold reduction of parasite burden for the ΔDG6 mutant when calculating the area under the curve, while deletion of DG5 had a more modest impact on infection; here, the area under the curve was half when compared to wild type (Figures 5D and 5E). Consistent with reduced parasite burden, mice infected with ΔDG6 mutants showed less weight loss when compared to those infected with KVI wild type (Figure 5F), while mice infected with ΔDG5 mutants had weights
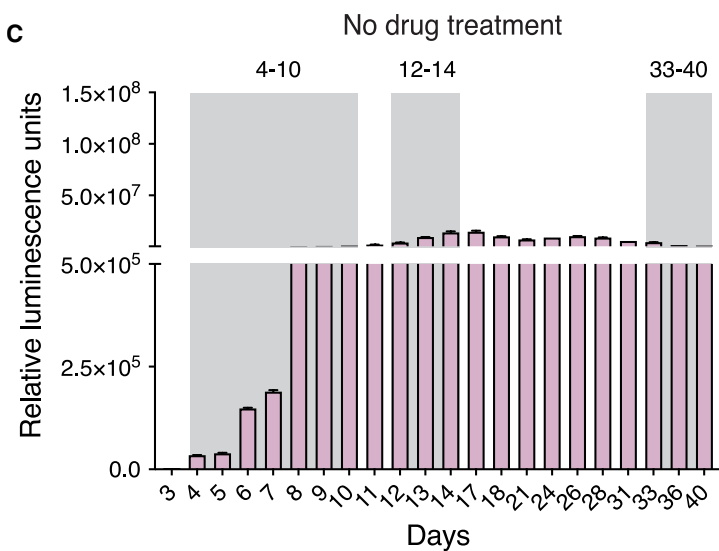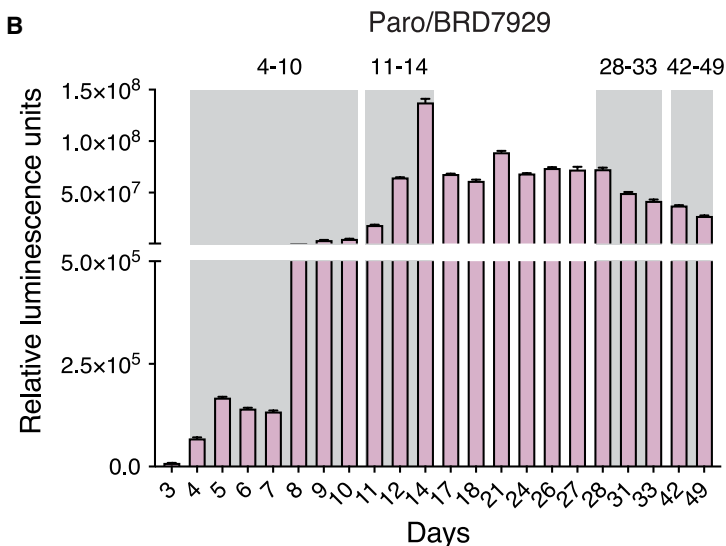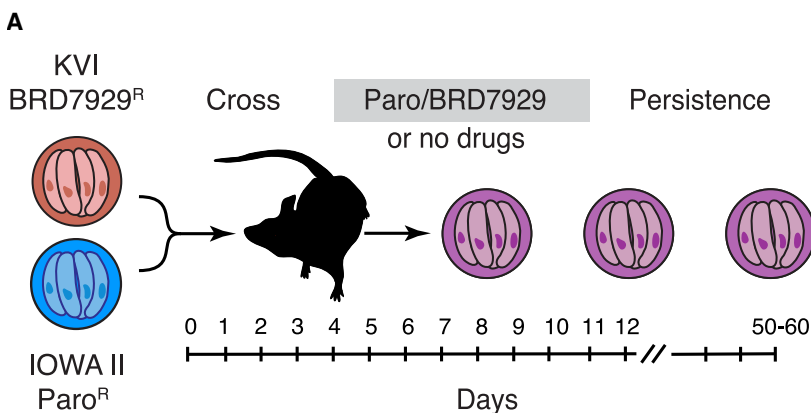
A

Figure 2. Genetic crosses of Iowa II and KVI parasites

(A) KVI and Iowa II parasites were engineered to be resistant to BRD7929 or paromomycin (paro), respectively. *Ifnγ*$^{-/-}$ mice were infected with both strains and treated with both drugs from day 4 to 10, or not. Five crosses were conducted in total (all are shown in Figure S4).

(B and C) Parasite burden was measured by following fecal luciferase activity for two exemplary crosses, one with drug treatment (B) and one without (C); $n = 3$ mice per cross. Gray boxes indicate pools of collected feces that were used for oocyst purification and whole genome sequencing.
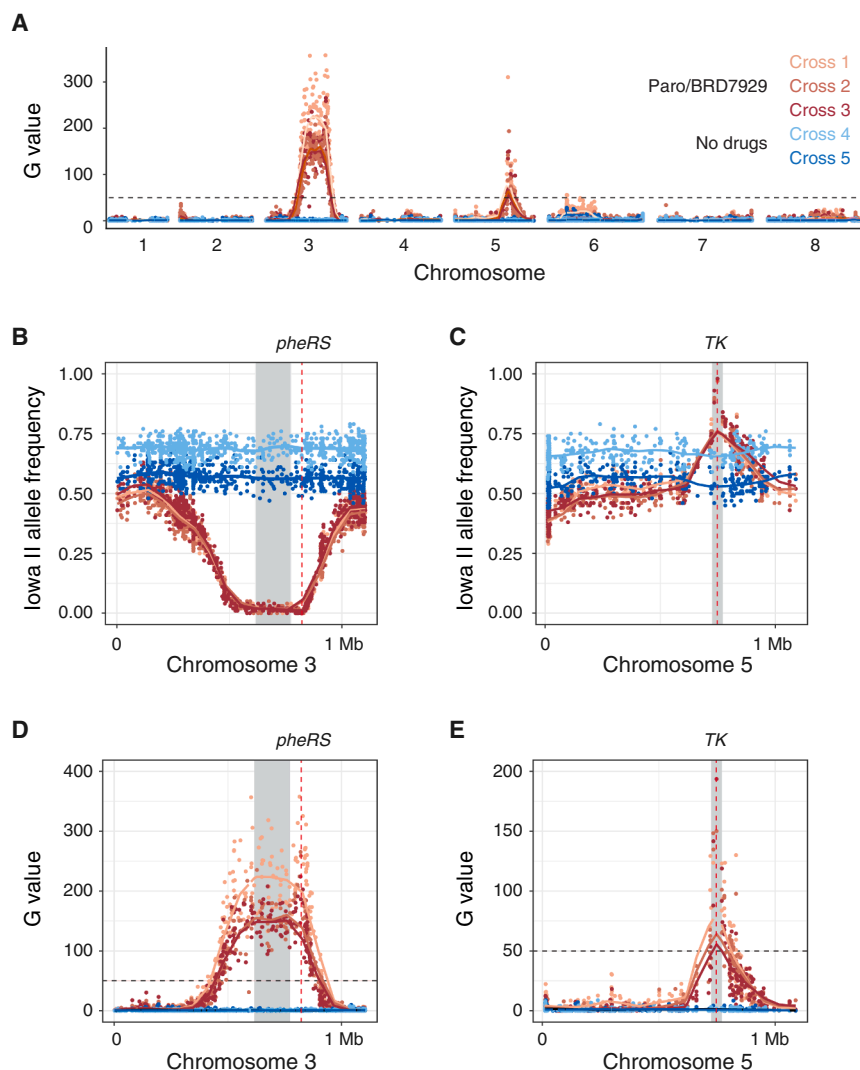
B

C

Figure 3. Mapping of drug resistance loci by bulk segregant analysis

(A) Bulk segregant analysis comparing three treated (red) with two untreated (blue) genetic crosses between KVI and Iowa II parasites (the first pool of oocysts collected in each experiment was used for this analysis). The G value for each individual SNP is shown as a dot, the lines are the weighted moving averages for the G′-values, and the significance threshold is shown as a black dashed line. Note highly significant QTL on chromosomes 3 and 5.

(B and C) Allele frequencies of SNPs using the Iowa II genome as a reference for chromosomes 3 and 5. While untreated parasites show balanced inheritance from both parents, treatment shifts regions of chromosome 3 to KVI and chromosome 5 to Iowa II, respectively.

(D and E) Zoomed-in chromosomal views, G values for both loci. Note that SNPs with peak G values fall into close proximity of the PheRS and TK loci (vertical red dashed lines), which carry the engineered drug resistance determinants. The 95th percentile of each QTL is shown in gray.

GP60 controls parasite burden and virulence but not persistence

Figure 6A shows a closer inspection of the QTL on chromosome 6, plotting average SNP G-values for bins of 1,000 bp, along with annotated gene models (all individual G-values are found in Table S1). The peak falls into gene *cgd6_1080*, which encodes GP60, a polymorphic gene that is frequently used to type *Cryptosporidium*.[26,30] GP60 has been suspected to be linked to virulence,[31] and a recent study has explored its contributions to host infectivity.[32] When comparing the GP60 gene between the parental Iowa II and KVI strains, we noted 149 SNPs and three indels (Figure S8A) and found that the selected progeny largely carry the KVI allele (Figure 4C). Both alleles encode an N-terminal signal peptide, a C-terminal glycosyl-phosphatidylinositol (GPI) anchor signal, numerous potential O-glycosylation sites, and a single potential N-linked glycosylation site, albeit not at the same position in the protein (Figure S8B shows an annotated sequence alignment). The QTL includes 5′ upstream regions with the presumptive promoter, and we thus considered that KVI and Iowa II might express GP60 at different levels as previously noted for ROP18 alleles in *T. gondii* strains that differ in virulence.[33,34] However, qPCR analysis of both parents showed similar GP60 mRNA abundance (Figure 6B). Next, we used reverse genetic engineering to test the impact of GP60 in isolation. We considered that GP60 from the more virulent KVI strain might be a dominant factor conferring virulence. We introduced KVI GP60 along with its 5′ and 3′ regulatory sequences as a transgene into the dispensable thymidine kinase locus of Iowa II parasites (this also introduced a Nluc reporter, Figures 6C and S2D). When we infected *Infγ*$^{-/-}$

similar to controls (Figure 5G). In HCT-8 tissue culture experiments, *in vitro* growth of the ΔDG6 mutant was found to be indistinguishable from KVI wild type (Figure 5H). To rigorously control whether the observed burden reduction was due to modification of the target locus alone, we next complemented the mutant. We reintroduced the gene amplified from either the Iowa II or the KVI strain as a transgene 3′ to the PheRS locus using BRD7929 selection, including the respective 5′ and 3′ sequences from each allele type to ensure native levels of expression (Figure S2C). Parasite burden in both complemented strains rebounded to levels comparable to KVI wild type, while the ΔDG6 mutant again showed lower infection (Figure 5I).

We conclude that while DG5 is dispensable under the conditions tested, loss of DG6 results in reduced infection *in vivo* but not *in vitro*, and this is rescued by complementation. We observed that sequences from either parent provided similar levels of complementation, suggesting that the strain-specific difference responsible for the QTL resides within a different gene.
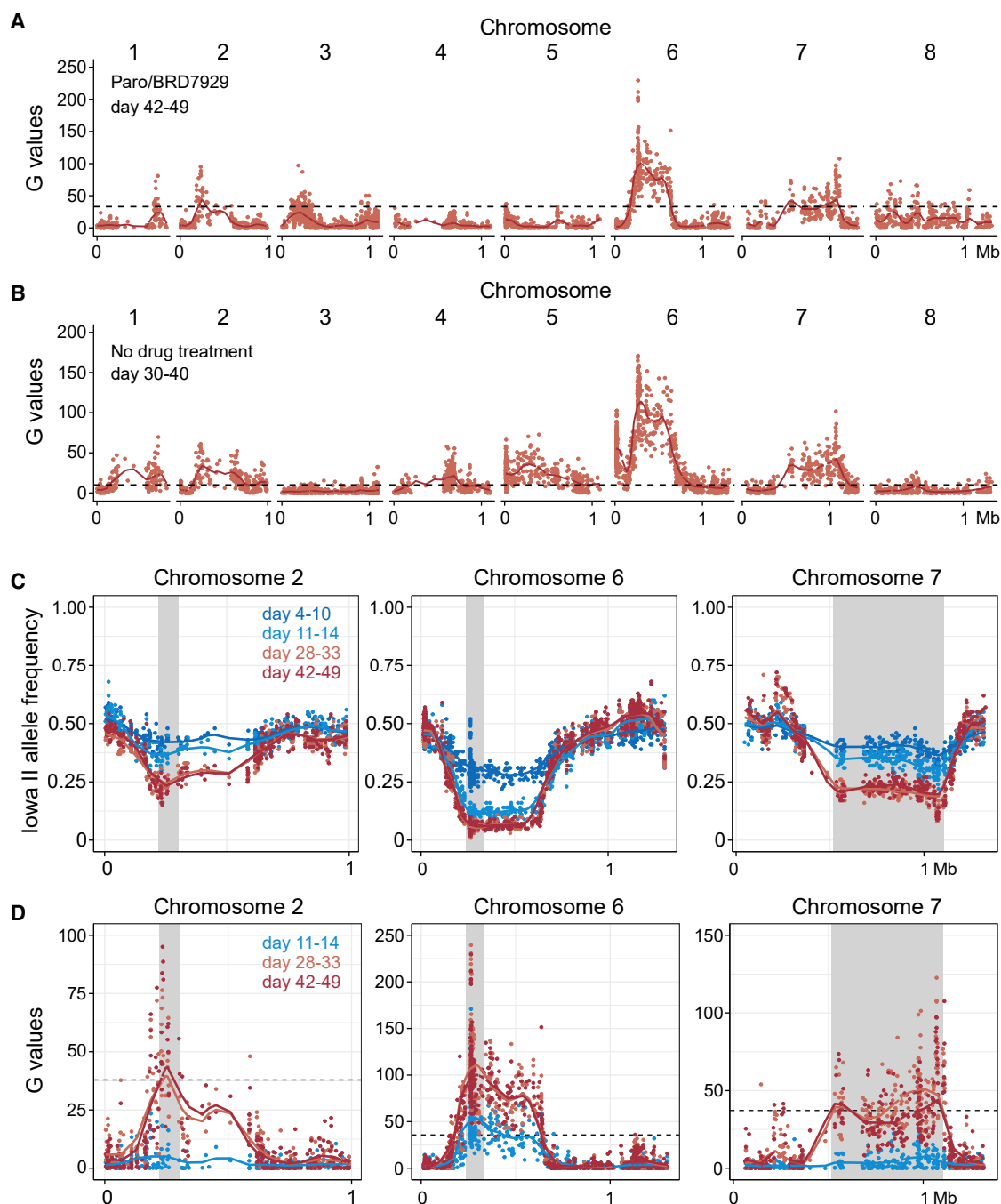
Figure 4. Longitudinal analysis of genetic crosses identifies loci associated with virulence and persistence

(A and B) Genetic crosses were conducted between Iowa II and KVI parasites with (A) or without (B) treatment with BRD7929 and paromomycin. Parasites were sampled as shown in Figures 2D and 2C, and whole genome was sequenced. G values of bulk segregant analysis are shown, comparing the first and last sampling windows for each experiment. There are three significant QTL in both treated and untreated populations.

(C and D) (C) Allele frequencies and (D) G values for these three loci are shown in more detail for the cross that used drug selection (all experiments are found in Figure S6). While the QTL on chromosome 6 emerges early, the loci on chromosomes 2 and 7 only show selection upon prolonged infection (samples from day 26 onward). In all plots, the horizontal dashed line represents the threshold for QTL detection. The 95th percentile of each QTL is shown in gray.
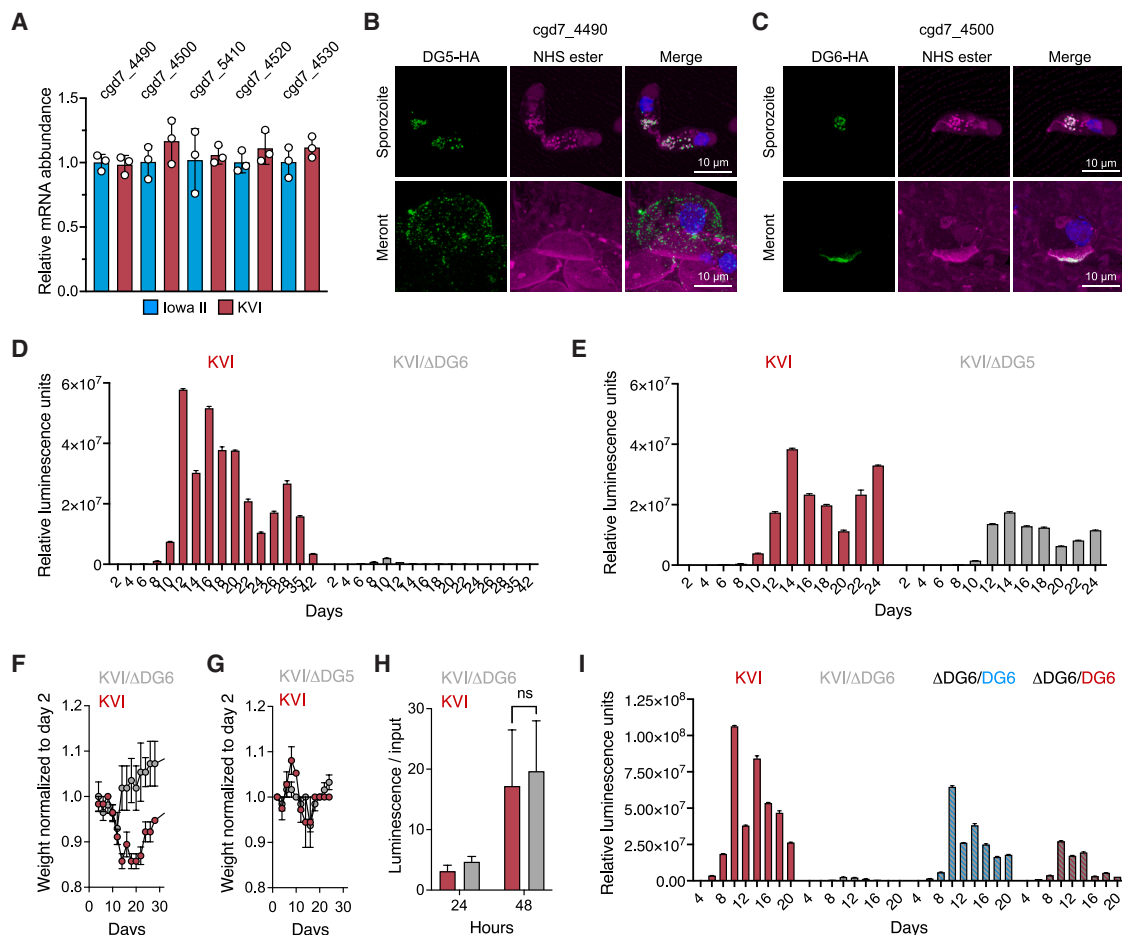
Figure 5. Loss of dense granule protein DG6 but not DG5 results in attenuation *in vivo*

(A) The core of the QTL on chromosome 7 includes three presumptive dense granule proteins and two ABC transporters. Comparison of their relative expression level by qPCR found no difference between KVI and Iowa II (a representative of two experiments with three biological and three technical replicates each is shown). The $\Delta\Delta$Ct method was used to determine the relative expression with GAPDH as the control gene. The fold change of the KVI strain is relative to the GAPDH control and Iowa II samples. Data represent three biological repeats and is plotted as the mean. A list of all primers is provided in Table S2.

(B and C) Expansion microscopy fluorescence micrographs of HCT-8 cells either infected with HA-tagged DG5 (B) or DG6 (C) in green, NH-ester in magenta, and Hoechst in blue. Scale bars, 10 μm.

(D and E) $Ifn\gamma^{-/-}$ mice were infected with ΔDG5 (D) or ΔDG6 (E) mutant parasites in direct comparison with KVI parental wild type, and parasite burden was measured by fecal luciferase assay.

(F and G) Body weight of mice infected with ΔDG6 (F) or ΔDG5 (G), each compared to the KVI parent, normalized to the weight at day 2 of infection. Data are mean ± SD of n = 2 biological replicates.

(H) HCT-8 cell culture assays comparing the growth of ΔDG5 and ΔDG6 to KVI parent. Data are mean ± SEM of n = 3 biological replicates. Significance was evaluated using two-way ANOVA, *p < 0.05.

(I) The ΔDG6 was complemented in *trans* with the DG6 gene of Iowa II (blue) or KVI (maroon). These parasites were used to infect $Ifn\gamma^{-/-}$ mice in comparison to the parental ΔDG6 and KVI wild type. Both genes confer enhanced infectivity to the mutant. For all infections in (D) and (I), one out of two replicates is shown; n = 3 mice per replicate.

mice with the resulting stable transgenic parasites, we found, to our surprise, the resulting infection to be moderate and similar to Iowa II (Figure 6E). Hypothesizing that alternatively, the low virulence Iowa II allele might be dominant, we next disrupted the coding sequence of the endogenous GP60 gene by insertion of a PheRS marker cassette in the bi-allelic strain (Figures 6D and S2D). This resulted in a marked increase in parasite burden in side-by-side infections (Figure 6E). We also engineered reciprocal modifications in the highly virulent KVI parent (Figure S2E). Ex-

pressing Iowa II GP60 as a transgene decreased KVI parasite burden, again suggesting the dominance of this allele in reducing virulence (Figure 6F). We also considered that expression of two distinct alleles of GP60 might have a dominant negative effect, limiting parasite growth in a way unrelated to virulence. To control for this, we ablated the original KVI gene in the pseudo diploid strain, leaving only the single Iowa II allele behind. This did not change parasite burden, which remained low, arguing that parasite growth *in vivo* is not impeded by the expression of two alleles
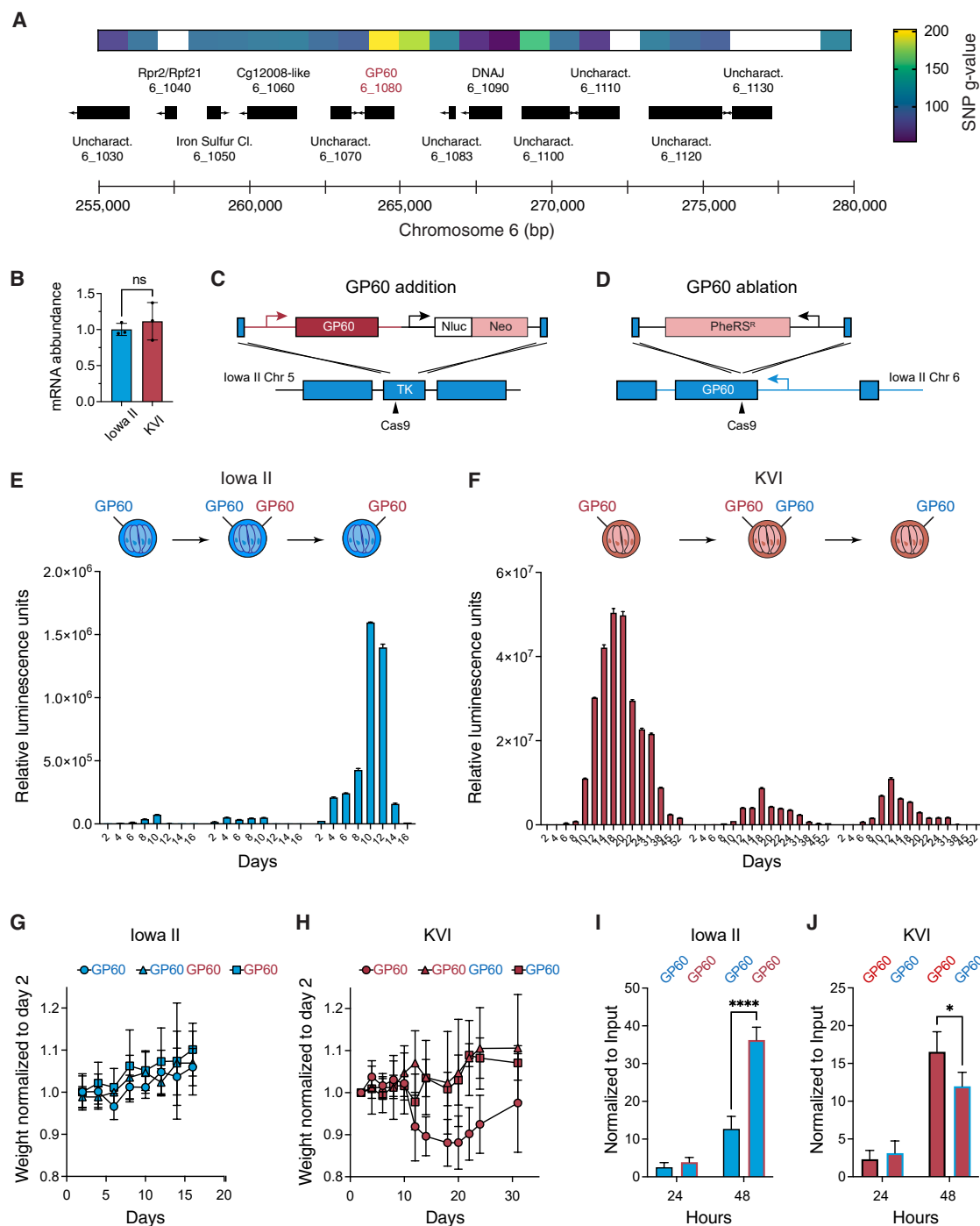
Figure 6. GP60 controls parasite burden and virulence

(A) Section of chromosome 6 that includes the QTL showing the coding regions of genes as black boxes with arrows indicating direction of transcription, and mean G values binned in 1000-bp intervals (white indicates areas devoid of SNPs distinguishing the parents).

(B) qPCR analysis of KVI GP60 mRNA abundance relative to Iowa II. The ΔΔCt method was used to determine the relative expression with GAPDH as the control gene. The fold change of the KVI strain is relative to the GAPDH control and Iowa II samples. Data represent three biological repeats and are plotted as the mean. A list of all the primers is provided in Table S2.

(C and D) (C) The GP60 gene, including its regulatory regions of the KVI strain, was introduced as a transgene in the thymidine kinase (TK) locus of the Iowa II strain using *Neo*, the native GP60 locus was then disrupted (D) using the PheRS marker, generating the strains used in (E). Reciprocal modifications were engineered into KVI (F). See Figure S2B and S2C for details on all allelic replacement mutants.

(legend continued on next page)

of GP60 but by the presence of the Iowa II allele (Figure 6F). We monitored the weight of all infected mice as a measure of virulence. Removing Iowa II GP60 in Iowa II alone was not sufficient to cause enhanced weight loss, and this might be linked to the fact that while parasite burden increased dramatically, the duration of the infection remained short (Figure 6G). Expression of Iowa II GP60 prevented the substantial weight loss associated with KVI infection (Figure 6H). When these replacement mutants were used to infect HCT-8 cultures, we observed that replacing GP60 in the Iowa II strain with the KVI allele enhanced growth, while exchange of GP60 in the KVI strain reduced growth when compared to the respective parent (Figures 6I and 6J). We were unable to ablate GP60 completely in the strains we studied.

We conclude that bulk segregant analysis correctly mapped and identified GP60 as a major strain-specific contributor to virulence. Interestingly, reverse genetic analysis in both parental backgrounds suggests that Iowa II GP60 moderates virulence in a dominant fashion. While changes in GP60 changed virulence, they had no impact on persistence, Iowa II did not gain persistence, and KVI did not lose it when the gene was replaced.

DISCUSSION

Sexual recombination has been hypothesized to be critical for the ability of *Cryptosporidium* to adapt to opportunity and challenge through rapid evolution of complex traits. Recent studies have highlighted the likely role of sex in the emergence of enhanced transmissibility, resulting in an epidemiological sweep of *C. hominis* strain IfA12G1R5 across the United States.[15,35] Recombination was similarly documented in studies of *C. hominis* across the African continent.[36] Increased host specialization resulting in narrowing adaptation to humans has been noted in the typically more broadly zoonotic *C. parvum*, and this has been discussed as the potential result of sex.[37] Interestingly, the overall rate of such genomic exchange appears higher in recent human history, likely reflecting increased long-distance travel of both humans and domestic animals, enhancing mating opportunities of geographically separated parasite strains.[38] In this study, we show, through controlled laboratory experimentation, that sex indeed allows parasites to recombine and rapidly select beneficial traits, including drug resistance and superior colonization ability. We also show that this can be exploited to discover parasite gene function in forward genetic mapping experiments.

While *Cryptosporidium* is not amenable to mapping based on the isolation and study of individual progeny clones, it proves well suited for bulk segregant analysis, which identified the loci of the drug resistance markers used with high statistical support. We noted relatively broad G′ linkage peaks that contain one or multiple much more narrow peaks when considering the G values of individual SNPs. This could indicate that multiple loci contribute to the overall QTL; however, this was also observed for the PheRS drug resistance locus, where a single gene drives selection. For most organisms, crosses involve a single round of recombination, and the progeny thus represent the F1 generation. However, *Cryptosporidium* experiences continuous mating among the progeny as sex occurs every third day of an ongoing infection.[17,18] We note that mapping was successful even in the absence of dual drug selection to enforce mating. This suggests an overall high rate of recombination and that crosses with unmarked strains, e.g., for species not amenable to transfection, are likely feasible. Using marker genes biases selection for their loci, which may fix chromosomes, limiting resolution. By the same token, careful placement of markers could enhance segregation of loci difficult to resolve.

Several additional factors likely influence the resolution of genetic mapping, including the specific biology of the trait under selection, the number of recombination events between the strains used, and the chromosomal location of the genes to be distinguished. One obvious factor is the number of SNPs that discern the parents. Higher SNP density improves the ability to pinpoint the position of selected genes, as seen in the highly polymorphic GP60 locus (Figure 6), while low density diminishes precision. The remarkable similarity of *C. parvum* IIa strains, even when isolated from different continents (Figure 1A), was an early impediment to our efforts. IIa strains are highly prevalent in cattle, and parasite uniformity may reflect host uniformity due to the global distribution and dominance of certain highly productive cattle breeds.[39] Isolation of IId parasites from lambs was important to success, and other studies have noted more pronounced differences between IIa and IId strains as well.[14,40] Future sampling of diverse hosts could yield diverse parasites, further improve mapping resolution, and provide insight into the mechanistic basis of host specificity. Another factor to consider is the strength of the selection applied.[41] The highly potent BRD7929 kills parasites fast[42] and the rigorous selection imposed may have resulted in rapid fixation of a large chromosomal segment, while the less potent drug paromomycin provided slower selection, thus adding opportunity for recombination and resolution. Lastly, time itself is important, and the chronic nature of KVI infection in the mouse model chosen here offered the opportunity for longitudinal evaluation. Two of the three QTLs only rose to significance over time, suggesting that they are required for chronic persistence but not initial colonization.

More work is needed to understand the specific mechanisms by which these loci impact persistence. However, we note the presence of polymorphic secretory proteins, in particular dense granule proteins, a subset of proteins that in *T. gondii* was shown

(E and F) WT parasites, parasite expressing GP60 of both parents, and parasite expressing only the GP60 of the other parent were used to infect *Ifnγ*$^{-/-}$ mice, and parasite burden was measured by fecal luciferase assay (note that all strains carried a single Nluc gene). Expression of Iowa II GP60 is associated with lowered burden in both parental backgrounds, while adding or removing KVI GP60 had little effect; one out of two replicates is shown; *n* = 3 mice per replicate.
(G and H) Mean weight of infected mice from (D) and (E) normalized to day 2 of infection (combined data of two experiments using four mice each, error bars show standard deviation).
(I and J) HCT-8 cultures were infected with the indicated strains, and luminescence was measured after 24 and 48 h and plotted normalized to the input used for infection (individual cultures are shown as dots and their mean as bars). Data are mean ± SEM of *n* = 3 biological replicates. Significance was evaluated using two-way ANOVA; *$p < 0.05$ and ****$p < 0.0001$.

to contribute to immune evasion by manipulating host defense pathways and their transcriptional and regulatory circuits.[43–45] The late emergence of these loci might suggest their dependence on selection imposed by the development of adaptive immunity. Crosses in hosts with varied immune competence may provide differentiating selection to link loci to the evasion of specific restriction mechanisms or may yield additional loci only revealed by the change of selection applied. The advanced tools accumulated for the study of immunity in mice offer a multitude of genetic backgrounds and reagents to remove or add specific elements of the immune system. We and others have shown *C. parvum* to be highly sensitive to such manipulations (reviewed in[10,11,46]). This breadth of selection systems, combined with the relative ease of the population-based crossing model developed in this study, holds significant promise for the future of forward genetic analysis of host-pathogen interaction in *Cryptosporidium*. We note that there are drawbacks to using population instead of clone-based strategies in the analysis of genetic crosses, most importantly that it is difficult to recognize and resolve epistatic interactions of multiple loci. This and other potential limitations have been reviewed in detail before.[28]

GP60 stands out as the most significantly selected locus in our experiments, leading to rapid fixation of the KVI allele, indicating that it exerts a profound influence on the ability of *C. parvum* to colonize mice, reach high parasite burdens, and cause disease as measured by weight loss and mortality. Reverse genetic analysis rigorously validated this finding, replacing the GP60 locus alone changes parasite load and virulence in both parental genomes as background. The well-acknowledged high degree of polymorphism of this locus[13,26] suggests that it is under significant diversifying selection,[47] a hallmark of rapid host-pathogen coevolution. GP60 is proteolytically processed into GP40 and GP15, which remain associated following cleavage and membrane-bound due to a GPI anchor on GP15.[48] GP40 is particularly variable and heavily glycosylated[30,49] (Figure S8). The protein has been localized to the surface of the extra- and intracellular parasite stages, and this localization depends on the presence of signal peptides for ER insertion and GPI anchor addition.[32] GP60 has been suggested to play a role in host cell invasion;[49] however, its specific molecular function remains to be revealed. Li and colleagues found that replacing GP60 in a IId strain with a IIa allele resulted in less virulent parasites.[32] We engineered replacement mutants in IIa and IId parental backgrounds along with intermediate pseudo-diploid strains expressing both alleles. The phenotypic analysis of these mutants demonstrated a dominance of the less virulent Iowa IIa allele in the genomic background of both parents; virulence was restricted by the presence of the Iowa II allele, rather than the KVI allele, enhancing virulence. This is intriguing, and the mechanism underlying this dominance is still unknown. A different GP60 allele might impact the parasite's ability to invade and grow. While the strain difference was noticeable in tissue culture, the effect was much more pronounced in animals, particularly in those lacking IFNγ. Overall, this may suggest that GP60 is recognized by the host as a trigger for immune restriction and that the IIa allele produces a stronger response in mice. Such a scenario underlies the relative burden of *T. gondii* in different rat hosts. Dense granule proteins secreted by the parasite trigger a potent inflammasome response, restricting the path-

ogen.[50] *C. parvum* activates a range of innate mechanisms through the NLRP6 inflammasome,[51,52] a subset of Toll-like receptors,[53,54] and likely additional pathways.[55] GP60 is also a target of adaptive immunity and a prominent antigen recognized by antibody and cell-mediated responses.[49,56,57] Genetic mapping studies in *Plasmodium* have identified strain-specific antigens.[58] Exploring the impact of GP60 alleles in the context of mouse mutants lacking specific restriction pathways will be helpful to further distinguish these models.

The crosses reveal the amplitude of parasite burden and the persistence of parasite presence as distinct strain-specific traits governing *Cryptosporidium* infection. GP60 controls parasite burden in the mouse model, but it does not impact persistence. Switching the GP60 allele to KVI in the Iowa parent, parasites did not gain persistence, nor did the KVI parent lose persistence in the reverse swap. Additional loci on chromosomes 2 and 7 were significantly selected in chronic infection, suggesting that they may harbor persistence-conferring genes. Initial studies on two genes encoding dense granule proteins demonstrated one of them, DG6, to be important for high parasite burden *in vivo* but not *in vitro*. However, complementation studies argue against a strain-specific function of this factor, making additional genes in this locus high-value candidates to be pursued in future studies. Importantly, both high burden and longer persistence were required for the KVI strain to cause severe life-threatening disease.

Our experimentation shows that forward genetic studies in *Cryptosporidium* are feasible, and with the help of bulk segregant analysis, highly capable of discovering the genomic location of genes associated with drug resistance and virulence. Continued recombination and continued selection yield linkage of different parasite genes to different phases of the infection. Varying the specific circumstance under which host and parasite encounter each other will tune the selection pressure and is likely to reveal different aspects of interaction and the numerous parasite factors that modulate such interaction. The small genome size of *Cryptosporidium* and the ability to use mice to conduct crosses make this a straightforward discovery model to understand how host restriction and parasite evasion govern parasite infection and the evolution of host specificity.

Limitations of the study

This study used population-based bulk segregant analysis to map genetic association. Potential limitations of the approach are that epistatic multi-gene effects are more difficult to detect and to resolve than when analyzing individual progeny clones. While mapping studies identified loci on chromosomes 2, 6, and 7, reverse genetic studies have thus far only pinpointed a specific gene in one of the loci. Further tests of additional genes on chromosomes 2 and 7 are required to enhance resolution. Linkage obtained for the specific parasite and host strains used in this study may not always extrapolate to other host-pathogen pairings.

Materials availability

This study did not generate new, unique reagents.

Data and code availability

- Whole-genome sequencing data (Illumina and Nanopore) have been deposited at NCBI as BioProject: PRJNA1263880 and are publicly available as of the date of publication.
- All original code has been deposited at GitHub and is publicly available at https://github.com/ruicatxiao/Automated_Bulk_Segregant_Analysis as of the date of publication.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

AUTHOR CONTRIBUTIONS

S.S. and B.S. conceived the study. D.Y.-L. isolated and genotyped the KVI strain in Israel; S.S. performed genetic crosses with support from G.Y.B.; reverse genetic characterization experiments with contributions from T.H.Z. to tagging and gene ablation; and A.C. to expansion microscopy. X.L. performed bulk segregant analysis with support from T.J.C.A.; R.X. generated the automated bulk segregant analysis and the *de novo* assembly of the KVI strain genome; and S.S. and B.S. wrote the manuscript with contributions from all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  ○ Mouse models of infection
  ○ Parasite strains
  ○ Cell culture
- METHOD DETAILS
  ○ Generation of transgenic parasites
  ○ Cell culture and *in vitro* growth assay
  ○ Microscopy
  ○ qPCR
  ○ Genetic crosses and sample collection
  ○ Oocyst purification and genomic DNA extraction
  ○ Library preparation and sequencing of cross progeny
- QUANTIFICATION AND STATISTICAL ANALYSIS
  ○ *De novo* genome assembly
  ○ Genotype calling
  ○ Bulk segregant analysis
- ADDITIONAL RESOURCES
  ○ Automated bulk segregant analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2025.116315.

REFERENCES

1. Checkley, W., White, A.C., Jr., Jaganath, D., Arrowood, M.J., Chalmers, R. M., Chen, X.M., Fayer, R., Griffiths, J.K., Guerrant, R.L., Hedstrom, L., et al. (2015). A review of the global burden, novel diagnostics, therapeutics, and vaccine targets for cryptosporidium. Lancet Infect. Dis. *15*, 85–94. https://doi.org/10.1016/S1473-3099(14)70772-8.

2. Striepen, B. (2013). Parasitic infections: Time to tackle cryptosporidiosis. Nature *503*, 189–191.

3. Robertson, L.J., Campbell, A.T., and Smith, H.V. (1992). Survival of Cryptosporidium parvum oocysts under various environmental pressures. Appl. Environ. Microbiol. *58*, 3494–3500. https://doi.org/10.1128/aem.58.11.3494-3500.1992.

4. Kotloff, K.L., Nataro, J.P., Blackwelder, W.C., Nasrin, D., Farag, T.H., Panchalingam, S., Wu, Y., Sow, S.O., Sur, D., Breiman, R.F., et al. (2013). Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. Lancet *382*, 209–222. https://doi.org/10.1016/S0140-6736(13)60844-2.

5. Korpe, P.S., and Petri, W.A., Jr. (2012). Environmental enteropathy: critical implications of a poorly understood condition. Trends Mol. Med. *18*, 328–336. https://doi.org/10.1016/j.molmed.2012.04.007.

6. Mondal, D., Haque, R., Sack, R.B., Kirkpatrick, B.D., and Petri, W.A., Jr. (2009). Attribution of malnutrition to cause-specific diarrheal illness: evidence from a prospective study of preschool children in Mirpur, Dhaka, Bangladesh. Am. J. Trop. Med. Hyg. *80*, 824–826.

7. Wojcik, G.L., Korpe, P., Marie, C., Mentzer, A.J., Carstensen, T., Mychaleckyj, J., Kirkpatrick, B.D., Rich, S.S., Concannon, P., Faruque, A.S.G., et al. (2020). Genome-Wide Association Study of Cryptosporidiosis in Infants Implicates PRKCA. mBio *11*, e03343-19. https://doi.org/10.1128/mBio.03343-19.

8. Kabir, M., Alam, M., Nayak, U., Arju, T., Hossain, B., Tarannum, R., Khatun, A., White, J.A., Ma, J.Z., Haque, R., et al. (2021). Nonsterile immunity to cryptosporidiosis in infants is associated with mucosal IgA against the sporozoite and protection from malnutrition. PLoS Pathog. *17*, e1009445. https://doi.org/10.1371/journal.ppat.1009445.

9. Charania, R., Wade, B.E., McNair, N.N., and Mead, J.R. (2020). Changes in the Microbiome of Cryptosporidium-Infected Mice Correlate to Differences in Susceptibility and Infection Levels. Microorganisms *8*, 879. https://doi.org/10.3390/microorganisms8060879.

10. Cohn, I.S., Henrickson, S.E., Striepen, B., and Hunter, C.A. (2022). Immunity to Cryptosporidium: Lessons from Acquired and Primary Immunodeficiencies. J. Immunol. *209*, 2261–2268. https://doi.org/10.4049/jimmunol.2200512.

11. Pardy, R.D., Wallbank, B.A., Striepen, B., and Hunter, C.A. (2024). Immunity to Cryptosporidium: insights into principles of enteric responses to infection. Nat. Rev. Immunol. *24*, 142–155. https://doi.org/10.1038/s41577-023-00932-3.

12. Sateriale, A., Slapeta, J., Baptista, R., Engiles, J.B., Gullicksrud, J.A., Herbert, G.T., Brooks, C.F., Kugler, E.M., Kissinger, J.C., Hunter, C.A., and Striepen, B. (2019). A Genetically Tractable, Natural Mouse Model of Cryptosporidiosis Offers Insights into Host Protective Immunity. Cell Host Microbe *26*, 135–146.e135. https://doi.org/10.1016/j.chom.2019.05.006.

13. Feng, Y., Ryan, U.M., and Xiao, L. (2018). Genetic Diversity and Population Structure of Cryptosporidium. Trends Parasitol. *34*, 997–1011. https://doi.org/10.1016/j.pt.2018.07.009.

14. Jia, R., Huang, W., Huang, N., Yu, Z., Li, N., Xiao, L., Feng, Y., and Guo, Y. (2022). High infectivity and unique genomic sequence characteristics of Cryptosporidium parvum in China. PLoS Negl. Trop. Dis. *16*, e0010714. https://doi.org/10.1371/journal.pntd.0010714.

15. Huang, W., Guo, Y., Lysen, C., Wang, Y., Tang, K., Seabolt, M.H., Yang, F., Cebelinski, E., Gonzalez-Moreno, O., Hou, T., et al. (2023). Multiple introductions and recombination events underlie the emergence of a hyper-transmissible Cryptosporidium hominis subtype in the USA. Cell Host Microbe *31*, 112–123.e4. https://doi.org/10.1016/j.chom.2022.11.013.

16. Slapeta, J. (2013). Cryptosporidiosis and Cryptosporidium species in animals and humans: a thirty colour rainbow? Int. J. Parasitol. *43*, 957–970. https://doi.org/10.1016/j.ijpara.2013.07.005.

17. English, E.D., Guérin, A., Tandel, J., and Striepen, B. (2022). Live imaging of the Cryptosporidium parvum life cycle reveals direct development of male and female gametes from type I meronts. PLoS Biol. *20*, e3001604. https://doi.org/10.1371/journal.pbio.3001604.

18. Tandel, J., English, E.D., Sateriale, A., Gullicksrud, J.A., Beiting, D.P., Sullivan, M.C., Pinkston, B., and Striepen, B. (2019). Life cycle progression and sexual development of the apicomplexan parasite Cryptosporidium parvum. Nat. Microbiol. *4*, 2226–2236. https://doi.org/10.1038/s41564-019-0539-x.

19. Walzer, K.A., Tandel, J., Byerly, J.H., Daniels, A.M., Gullicksrud, J.A., Whelan, E.C., Carro, S.D., Krespan, E., Beiting, D.P., and Striepen, B. (2024). Transcriptional control of the Cryptosporidium life cycle. Nature *630*, 174–180. https://doi.org/10.1038/s41586-024-07466-1.

20. Kimball, A., Funkhouser-Jones, L., Huang, W., Xu, R., Witola, W.H., and Sibley, L.D. (2024). Mendelian segregation and high recombination rates facilitate genetic analyses in Cryptosporidium parvum. PLoS Genet. *20*, e1011162. https://doi.org/10.1371/journal.pgen.1011162.

21. Behnke, M.S., Dubey, J.P., and Sibley, L.D. (2016). Genetic Mapping of Pathogenesis Determinants in Toxoplasma gondii. Annu. Rev. Microbiol. *70*, 63–81. https://doi.org/10.1146/annurev-micro-091014-104353.

22. Su, X., Hayton, K., and Wellems, T.E. (2007). Genetic linkage and association analyses for trait mapping in Plasmodium falciparum. Nat. Rev. Genet. *8*, 497–506. https://doi.org/10.1038/nrg2126.

23. Shaw, S., Cohn, I.S., Baptista, R.P., Xia, G., Melillo, B., Agyabeng-Dadzie, F., Kissinger, J.C., and Striepen, B. (2024). Genetic crosses within and between species of Cryptosporidium. Proc. Natl. Acad. Sci. USA *121*, e2313210120. https://doi.org/10.1073/pnas.2313210120.

24. Tako, S., Fleiderovitz, L., Markovich, M.P., Mazuz, M.L., Behar, A., and Yasur-Landau, D. (2023). Cryptosporidium parvum gp60 subtypes in diarrheic lambs and goat kids from Israel. Parasitol. Res. *122*, 2237–2241. https://doi.org/10.1007/s00436-023-07925-0.

25. Baptista, R.d.P., Xiao, R., Li, Y., Glenn, T.C., and Kissinger, J.C. (2023). New T2T assembly of Cryptosporidium parvum IOWA annotated with reference genome gene identifiers. Preprint at bioRxiv, 2023.06.13.544219. https://doi.org/10.1101/2023.06.13.544219.

26. Robinson, G., Chalmers, R.M., Elwin, K., Guy, R.A., Bessonov, K., Troell, K., and Xiao, L. (2025). Deciphering a cryptic minefield: A guide to Cryptosporidium gp60 subtyping. Curr. Res. Parasitol. Vector. Borne. Dis. *7*, 100257. https://doi.org/10.1016/j.crpvbd.2025.100257.

27. Vinayak, S., Pawlowic, M.C., Sateriale, A., Brooks, C.F., Studstill, C.J., Bar-Peled, Y., Cipriano, M.J., and Striepen, B. (2015). Genetic modification of the diarrhoeal pathogen Cryptosporidium parvum. Nature *523*, 477–480. https://doi.org/10.1038/nature14651.

28. Li, X., Kumar, S., Brenneman, K.V., and Anderson, T.J.C. (2022). Bulk segregant linkage mapping for rodent and human malaria parasites. Parasitol. Int. *91*, 102653. https://doi.org/10.1016/j.parint.2022.102653.

29. Guerin, A., Strelau, K.M., Barylyuk, K., Wallbank, B.A., Berry, L., Crook, O.M., Lilley, K.S., Waller, R.F., and Striepen, B. (2023). Cryptosporidium uses multiple distinct secretory organelles to interact with and modify its host cell. Cell Host Microbe *31*, 650–664.e6. https://doi.org/10.1016/j.chom.2023.03.001.

30. Strong, W.B., Gut, J., and Nelson, R.G. (2000). Cloning and sequence analysis of a highly polymorphic Cryptosporidium parvum gene encoding a 60-kilodalton glycoprotein and characterization of its 15- and 45-kilodalton zoite surface antigen products. Infect. Immun. *68*, 4117–4134.

31. Bouzid, M., Hunter, P.R., Chalmers, R.M., and Tyler, K.M. (2013). Cryptosporidium pathogenicity and virulence. Clin. Microbiol. Rev. *26*, 115–134. https://doi.org/10.1128/CMR.00076-12.

32. Li, M., Yang, F., Hou, T., Gong, X., Li, N., Sibley, L.D., Feng, Y., Xiao, L., and Guo, Y. (2024). Variant surface protein GP60 contributes to host infectivity of Cryptosporidium parvum. Commun. Biol. *7*, 1175. https://doi.org/10.1038/s42003-024-06885-0.

33. Taylor, S., Barragan, A., Su, C., Fux, B., Fentress, S.J., Tang, K., Beatty, W.L., Hajj, H.E., Jerome, M., Behnke, M.S., et al. (2006). A secreted serine-threonine kinase determines virulence in the eukaryotic pathogen Toxoplasma gondii. Science *314*, 1776–1780. https://doi.org/10.1126/science.1133643.

34. Saeij, J.P.J., Boyle, J.P., Coller, S., Taylor, S., Sibley, L.D., Brooke-Powell, E.T., Ajioka, J.W., and Boothroyd, J.C. (2006). Polymorphic secreted kinases are key virulence factors in toxoplasmosis. Science *314*, 1780–1783. https://doi.org/10.1126/science.1133690.

35. Jex, A.R., and Tyler, K.M. (2023). American variants of Cryptosporidium hominis: Over-sexed and over here? Cell Host Microbe *31*, 5–7. https://doi.org/10.1016/j.chom.2022.12.012.

36. Tichkule, S., Jex, A.R., van Oosterhout, C., Sannella, A.R., Krumkamp, R., Aldrich, C., Maiga-Ascofare, O., Dekker, D., Lamshöft, M., Mbwana, J., et al. (2021). Comparative genomics revealed adaptive admixture in Cryptosporidium hominis in Africa. Microb. Genom. *7*, mgen000493. https://doi.org/10.1099/mgen.0.000493.

37. Nader, J.L., Mathers, T.C., Ward, B.J., Pachebat, J.A., Swain, M.T., Robinson, G., Chalmers, R.M., Hunter, P.R., van Oosterhout, C., and Tyler, K.M. (2019). Evolutionary genomics of anthroponosis in Cryptosporidium. Nat. Microbiol. *4*, 826–836. https://doi.org/10.1038/s41564-019-0377-x.

38. Corsi, G.I., Tichkule, S., Sannella, A.R., Vatta, P., Asnicar, F., Segata, N., Jex, A.R., van Oosterhout, C., and Cacciò, S.M. (2023). Recent genetic exchanges and admixture shape the genome and population structure of the zoonotic pathogen Cryptosporidium parvum. Mol. Ecol. *32*, 2633–2645. https://doi.org/10.1111/mec.16556.

39. Talenti, A., Powell, J., Hemmink, J.D., Cook, E.A.J., Wragg, D., Jayaraman, S., Paxton, E., Ezeasor, C., Obishakin, E.T., Agusi, E.R., et al. (2022). A cattle graph genome incorporating global breed diversity. Nat. Commun. *13*, 910. https://doi.org/10.1038/s41467-022-28605-0.

40. Feng, Y., Li, N., Roellig, D.M., Kelley, A., Liu, G., Amer, S., Tang, K., Zhang, L., and Xiao, L. (2017). Comparative genomic analysis of the IId subtype family of Cryptosporidium parvum. Int. J. Parasitol. *47*, 281–290. https://doi.org/10.1016/j.ijpara.2016.12.002.

41. Chantepie, S., and Chevin, L.M. (2020). How does the strength of selection influence genetic correlations? Evol. Lett. *4*, 468–478. https://doi.org/10.1002/evl3.201.

42. Vinayak, S., Jumani, R.S., Miller, P., Hasan, M.M., McLeod, B.I., Tandel, J., Stebbins, E.E., Teixeira, J.E., Borrel, J., Gonse, A., et al. (2020). Bicyclic azetidines kill the diarrheal pathogen Cryptosporidium in mice by inhibiting parasite phenylalanyl-tRNA synthetase. Sci. Transl. Med. *12*, eaba8412. https://doi.org/10.1126/scitranslmed.aba8412.

43. Braun, L., Brenier-Pinchart, M.P., Hammoudi, P.M., Cannella, D., Kieffer-Jaquinod, S., Vollaire, J., Josserand, V., Touquet, B., Couté, Y., Tardieux, I., et al. (2019). The Toxoplasma effector TEEGR promotes parasite persistence by modulating NF-kappaB signalling via EZH2. Nat. Microbiol. *4*, 1208–1220. https://doi.org/10.1038/s41564-019-0431-8.

44. Hakimi, M.A., Olias, P., and Sibley, L.D. (2017). Toxoplasma Effectors Targeting Host Signaling and Transcription. Clin. Microbiol. Rev. *30*, 615–645. https://doi.org/10.1128/Cmr.00005-17.

45. Olias, P., Etheridge, R.D., Zhang, Y., Holtzman, M.J., and Sibley, L.D. (2016). Toxoplasma Effector Recruits the Mi-2/NuRD Complex to Repress STAT1 Transcription and Block IFN-gamma-Dependent Gene Expression. Cell Host Microbe *20*, 72–82. https://doi.org/10.1016/j.chom.2016.06.006.

46. Laurent, F., and Lacroix-Lamandé, S. (2017). Innate immune responses play a key role in controlling infection of the intestinal epithelium by Cryptosporidium. Int. J. Parasitol. *47*, 711–721. https://doi.org/10.1016/j.ijpara.2017.08.001.

47. Brockhurst, M.A., Chapman, T., King, K.C., Mank, J.E., Paterson, S., and Hurst, G.D.D. (2014). Running with the Red Queen: the role of biotic conflicts in evolution. Proc. Biol. Sci. *281*, 20141382. https://doi.org/10.1098/rspb.2014.1382.

48. O'Connor, R.M., Wanyiri, J.W., Cevallos, A.M., Priest, J.W., and Ward, H.D. (2007). Cryptosporidium parvum glycoprotein gp40 localizes to the sporozoite surface by association with gp15. Mol Biochem Parasit *156*, 80–83. https://doi.org/10.1016/j.molbiopara.2007.07.010.

49. Cevallos, A.M., Zhang, X., Waldor, M.K., Jaison, S., Zhou, X., Tzipori, S., Neutra, M.R., and Ward, H.D. (2000). Molecular cloning and expression of a gene encoding Cryptosporidium parvum glycoproteins gp40 and gp15. Infect. Immun. *68*, 4108–4116.

50. Cavailles, P., Flori, P., Papapietro, O., Bisanz, C., Lagrange, D., Pilloux, L., Massera, C., Cristinelli, S., Jublot, D., Bastien, O., et al. (2014). A highly conserved Toxo1 haplotype directs resistance to toxoplasmosis and its associated caspase-1 dependent killing of parasite and host macrophage. PLoS Pathog. *10*, e1004005. https://doi.org/10.1371/journal.ppat.1004005.

51. McNair, N.N., Bedi, C., Shayakhmetov, D.M., Arrowood, M.J., and Mead, J.R. (2018). Inflammasome components caspase-1 and adaptor protein apoptosis-associated speck-like proteins are important in resistance to Cryptosporidium parvum. Microbes Infect. *20*, 369–375. https://doi.org/10.1016/j.micinf.2018.04.006.

52. Sateriale, A., Gullicksrud, J.A., Engiles, J.B., McLeod, B.I., Kugler, E.M., Henao-Mejia, J., Zhou, T., Ring, A.M., Brodsky, I.E., Hunter, C.A., and Striepen, B. (2021). The intestinal parasite Cryptosporidium is controlled by an enterocyte intrinsic inflammasome that depends on NLRP6. Proc. Natl. Acad. Sci. USA *118*, e2007807118. https://doi.org/10.1073/pnas.2007807118.

53. Gibson, A.R., Sateriale, A., Dumaine, J.E., Engiles, J.B., Pardy, R.D., Gullicksrud, J.A., O'Dea, K.M., Doench, J.G., Beiting, D.P., Hunter, C.A., and Striepen, B. (2022). A genetic screen identifies a protective type III interferon response to Cryptosporidium that requires TLR3 dependent recognition. PLoS Pathog. *18*, e1010003. https://doi.org/10.1371/journal.ppat.1010003.

54. O'Hara, S.P., Bogert, P.S.T., Trussoni, C.E., Chen, X., and LaRusso, N.F. (2011). TLR4 promotes Cryptosporidium parvum clearance in a mouse model of biliary cryptosporidiosis. J. Parasitol. *97*, 813–821. https://doi.org/10.1645/GE-2703.1.

55. Mead, J.R. (2023). Early immune and host cell responses to Cryptosporidium infection. Front. Parasitol. *2*, 1113950. https://doi.org/10.3389/fpara.2023.1113950.

56. Wang, Y., Gao, M., Li, X., Zhu, W., Zhao, M., Li, J., Liu, X., Cao, L., Li, S., Zhang, S., et al. (2023). Structural Analyses of a Dominant Cryptosporidium parvum Epitope Presented by H-2K(b) Offer New Options To Combat Cryptosporidiosis. mBio *14*, e0266622. https://doi.org/10.1128/mbio.02666-22.

57. Bhalchandra, S., Gevers, K., Heimburg-Molinaro, J., van Roosmalen, M., Coppens, I., Cummings, R.D., and Ward, H.D. (2023). Identification of the glycopeptide epitope recognized by a protective Cryptosporidium monoclonal antibody. Infect. Immun. *91*, e0027523. https://doi.org/10.1128/iai.00275-23.

58. Martinelli, A., Cheesman, S., Hunt, P., Culleton, R., Raza, A., Mackinnon, M., and Carter, R. (2005). A genetic approach to the de novo identification of targets of strain-specific immunity in malaria parasites. Proc. Natl. Acad. Sci. USA *102*, 814–819. https://doi.org/10.1073/pnas.0405097102.

59. Stanke, M., Diekhans, M., Baertsch, R., and Haussler, D. (2008). Using native and syntenically mapped cDNA alignments to improve de novo gene finding. Bioinformatics *24*, 637–644. https://doi.org/10.1093/bioinformatics/btn013.

60. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

61. Agyabeng-Dadzie, F., Beaudry, M.S., Deyanov, A., Slanis, H., Duong, M.Q., Turner, R., Khan, A., Arias, C.A., Kissinger, J.C., Glenn, T.C., and de Paula Baptista, R. (2025). Evaluating the Benefits and Limits of Multiple Displacement Amplification With Whole-Genome Oxford Nanopore Sequencing. Mol. Ecol. Resour. *25*, e14094. https://doi.org/10.1111/1755-0998.14094.

62. Astashyn, A., Tvedte, E.S., Sweeney, D., Sapojnikov, V., Bouk, N., Joukov, V., Mozes, E., Strope, P.K., Sylla, P.M., Wagner, L., et al. (2024). Rapid and sensitive detection of genome contamination at scale with FCS-GX. Genome Biol. *25*, 60. https://doi.org/10.1186/s13059-024-03198-7.

63. Kolmogorov, M., Bickhart, D.M., Behsaz, B., Gurevich, A., Rayko, M., Shin, S.B., Kuhn, K., Yuan, J., Polevikov, E., Smith, T.P.L., and Pevzner, P.A. (2020). metaFlye: scalable long-read metagenome assembly using repeat graphs. Nat. Methods *17*, 1103–1110. https://doi.org/10.1038/s41592-020-00971-x.

64. Shumate, A., and Salzberg, S.L. (2021). Liftoff: accurate mapping of gene annotations. Bioinformatics *37*, 1639–1643. https://doi.org/10.1093/bioinformatics/btaa1016.

65. Marcais, G., Delcher, A.L., Phillippy, A.M., Coston, R., Salzberg, S.L., and Zimin, A. (2018). MUMmer4: A fast and versatile genome alignment system. PLoS Comput. Biol. *14*, e1005944. https://doi.org/10.1371/journal.pcbi.1005944.

66. Vaser, R., Sović, I., Nagarajan, N., and Šikić, M. (2017). Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res. *27*, 737–746. https://doi.org/10.1101/gr.214270.116.

67. Alonge, M., Lebeigle, L., Kirsche, M., Jenike, K., Ou, S., Aganezov, S., Wang, X., Lippman, Z.B., Schatz, M.C., and Soyk, S. (2022). Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. Genome Biol. *23*, 258. https://doi.org/10.1186/s13059-022-02823-7.

68. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., and Li, H. (2021). Twelve years of SAMtools and BCFtools. GigaScience *10*, giab008. https://doi.org/10.1093/gigascience/giab008.

69. Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A., and Korobeynikov, A. (2020). Using SPAdes De Novo Assembler. Curr. Protoc. Bioinformatics *70*, e102. https://doi.org/10.1002/cpbi.102.

70. Garrison, E., Kronenberg, Z.N., Dawson, E.T., Pedersen, B.S., and Prins, P. (2022). A spectrum of free software tools for processing the VCF variant call format: vcflib, bio-vcf, cyvcf2, hts-nim and slivar. PLoS Comput. Biol. *18*, e1009123. https://doi.org/10.1371/journal.pcbi.1009123.

71. Lacroix-Lamande, S., Mancassola, R., Naciri, M., and Laurent, F. (2002). Role of gamma interferon in chemokine expression in the ileum of mice and in a murine intestinal epithelial cell line after Cryptosporidium parvum infection. Infect. Immun. *70*, 2090–2099. https://doi.org/10.1128/IAI.70.4.2090-2099.2002.

72. Xiao, L., Morgan, U.M., Limor, J., Escalante, A., Arrowood, M., Shulaw, W., Thompson, R.C., Fayer, R., and Lal, A.A. (1999). Genetic diversity within Cryptosporidium parvum and related Cryptosporidium species. Appl Environ Microb *65*, 3386–3391.

73. Xiao, L., Alderisio, K., Limor, J., Royer, M., and Lal, A.A. (2000). Identification of species and sources of Cryptosporidium oocysts in storm waters

with a small-subunit rRNA-based diagnostic and genotyping tool. Appl. Environ. Microbiol. *66*, 5492–5498. https://doi.org/10.1128/AEM.66.12.5492-5498.2000.

74. Feng, Y., Li, N., Duan, L., and Xiao, L. (2009). Cryptosporidium genotype and subtype distribution in raw wastewater in Shanghai, China: evidence for possible unique Cryptosporidium hominis transmission. J. Clin. Microbiol. *47*, 153–157. https://doi.org/10.1128/JCM.01777-08.

75. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. *25*, 3389–3402.

76. Alves, L.M., Heisler, E.G., Kissinger, J.C., Patterson, J.M., and Kalan, E.B. (1979). Effects of Controlled Atmospheres on Production of Sesquiterpenoid Stress Metabolites by White Potato Tuber: Possible Involvement of Cyanide-resistant Respiration. Plant Physiol. *63*, 359–362. https://doi.org/10.1104/pp.63.2.359.

77. (2022). Cryptosporidiosis (WOAH Terrestrial Manual).

78. Kato, S., Jenkins, M.B., Ghiorse, W.C., and Bowman, D.D. (2001). Chemical and physical factors affecting the excystation of Cryptosporidium parvum oocysts. J. Parasitol. *87*, 575–581. https://doi.org/10.1645/0022-3395(2001)087[0575:CAPFAT]2.0.CO;2.

79. Pawlowic, M.C., Vinayak, S., Sateriale, A., Brooks, C.F., and Striepen, B. (2017). Generating and Maintaining Transgenic Cryptosporidium parvum Parasites. Curr. Protoc. Microbiol. *46*, 20B.2.1–20B.2.32. https://doi.org/10.1002/cpmc.33.

80. Dos Santos Pacheco, N., and Soldati-Favre, D. (2021). Coupling Auxin-Inducible Degron System with Ultrastructure Expansion Microscopy to Accelerate the Discovery of Gene Function in Toxoplasma gondii. Methods Mol. Biol. *2369*, 121–137. https://doi.org/10.1007/978-1-0716-1681-9_8.

81. Liffner, B., Cepeda Diaz, A.K., Blauwkamp, J., Anaguano, D., Frolich, S., Muralidharan, V., Wilson, D.W., Dvorin, J.D., and Absalon, S. (2023). Atlas of Plasmodium falciparum intraerythrocytic development using expansion microscopy. eLife *12*. https://doi.org/10.7554/eLife.88088.

82. Dadzie, F.A., Beaudry, M.S., Deyanov, A., Slanis, H., Duong, M.Q., Turner, R., Khan, A., Arias, C.A., Kissinger, J.C., Glenn, T.C., and de Paula Baptista, R. (2024). Evaluating the Benefits and Limits of Multiple Displacement Amplification with Whole-Genome Oxford Nanopore Sequencing. Preprint at bioRxiv, 2024.02.09.579537. https://doi.org/10.1101/2024.02.09.579537.

83. Bellinzona, G., Nardi, T., Castelli, M., Batisti Biffignandi, G., Adjou, K., Betson, M., Blanchard, Y., Bujila, I., Chalmers, R., Davidson, R., et al. (2024). Comparative genomics of Cryptosporidium parvum reveals the emergence of an outbreak-associated population in Europe and its spread to the United States. Genome Res. *34*, 877–887. https://doi.org/10.1101/gr.278830.123.

84. Mansfield, B.E., Oltean, H.N., Oliver, B.G., Hoot, S.J., Leyde, S.E., Hedstrom, L., and White, T.C. (2010). Azole Drugs Are Imported By Facilitated Diffusion in Candida albicans and Other Pathogenic Fungi. PLoS Pathog. *6*, e1001126. https://doi.org/10.1371/journal.ppat.1001126.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
No antibodies were used in this study		N/A
Bacterial and virus strains		
Escherichia coli GC5	Genesee Scientific	42–653
Biological samples		
No biological samples were used in this study		N/A
Chemicals, peptides, and recombinant proteins		
Ultrapure Phenol:Chloroform:Isoamyl Alcohol	Thermo Scientific	15593031
chloroform	Sigma-Aldrich	25666-100ML
RNAse	New England Biolabs	T3018L
ProtK	New England Biolabs	P8107S
Paraformaldehyde	Electron Microscopy Sciences	15710
Poly-D-Lysine	EMD Millipore	A-003-E
Acrylamide	Sigma	A4058
Formaldehyde	Fisher	AA33314AP
Tetramethylethylendiamine	ThermoFisher	17919
Ammonium persulfate	ThermoFisher	17874
N,N′-methylenbisacrylamide	SIGMA	M1533
Sodium Acrylate	AKscientific	R624
PBS	Sigma-Aldrich	11666789001
Bovine Serum Albumin	Sigma-Aldrich	A2153-500g
Critical commercial assays		
Quick-DNA Fecal/Soil Microbe Kit	Zymo Research	D6010
REPLI-G MINI kit WGA	Qiagen	150023
RNeasy minikit	Qiagen	73404
SuperScript IV reverse transcriptase	Thermo Fisher Scientific	18090050
SsoAdvanced Universal SYBR Green Supermix	Bio-Rad	1725271
Illumina Sequencing Kit	Illumina	20050264
Illumina Nextera Whole-genome DNA library prep kit	Illumina	20060059
Nanopore Library prep kit	Oxford Nanopore	SQK-NBD114.24
Nanopore FlowCell	Oxford Nanopore	FLO-MIN114
Deposited data		
Raw reads Illumina sequencing	NCBI	BioProject: PRJNA1263880
Raw reads Nanopore sequencing	NCBI	BioProject: PRJNA1263880
Experimental models: Cell lines		
HCT8	ATCC	CCL-244
Experimental models: Organisms/strains		
Cryptosporidium parvum	Bunchgrass Farms, ID	IOWA strain
Cryptosporidium parvum	Kimron Veterinary Institute, Israel	KVI strain
Cryptosporidium parvum	University of Arizona	Sterling strain
Cryptosporidium parvum	INRA Tours, France	INRA strain
Cryptosporidium parvum	University of Giessen	Berlin strain
Cryptosporidium parvum	University of Giessen	Leipzig strain

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
ifnγ [-/-] mice	Jackson Laboratory	Strain 002287
C57BL/6J mice	Jackson Laboratory	Strain 000664
Oligonucleotides		
Primers for cloning and genotyping see Table S2	This paper	N/A
Software and algorithms		
PRISM	GraphPad	https://www.graphpad.com/scientific-software/prism/
ImageJ	ImageJ	https://imagej.net/Welcome
ABSA (v0.1)	This paper	https://github.com/ruicatxiao/Automated_Bulk_Segregant_Analysis
AGAT (v1.4.0)	NA	https://github.com/NBISweden/AGAT
BRAKER (v3.0.8)	Stanke et al.[59]	https://github.com/Gaius-Augustus/BRAKER
BSA.optimization (v2022.01.26	Brennemann et al.[28]	https://github.com/emilyli0325/BSA.optimization
BWA (0.7.17-r1188)	Li et at.[60]	https://github.com/lh3/bwa
CADECT (v1.0.2)	Agyabeng-Dadzie et al.[61]	https://github.com/rpbap/CADECT
Dorado (v0.9.1)	NA	https://github.com/nanoporetech/dorado
fcs (v0.5.0)	Astashyn et al.[62]	https://github.com/ncbi/fcs
FindTelomeres (v2020.08.09)	NA	https://github.com/JanaSperschneider/FindTelomeres
Flye (v2.9.5)	Kolmogorov et al.[63]	https://github.com/mikolmogorov/Flye
GATK (v4.6.1.0)	978–1491975190	https://github.com/broadinstitute/gatk
Liftoff (v1.6.3)	Shumate et al.[64]	https://github.com/agshumate/Liftoff
MUMmer (v4.0.0)	Marçais et al.[65]	https://github.com/mummer4/mummer
Python (v3.0.11)	NA	https://www.python.org/
Racon (v1.4.3)	Vaser e al.[66]	https://github.com/lbcb-sci/racon
RagTag (v2.1.0)	Alonge et al.[67]	https://github.com/malonge/RagTag
samtools (v1.19.2)	Danecek et al.[68]	https://github.com/samtools/samtools
Singularity (v4.2.0)	10.5281/zenodo.5570766	https://github.com/sylabs/singularity
SPAdes (v4.0.0)	Prjibelski et al.[69]	https://github.com/ablab/spades
TrimGalore (v0.6.10)	10.5281/zenodo.7598955	https://github.com/FelixKrueger/TrimGalore
vcffilter (1.0.10)	Garrison et al.[70]	https://github.com/vcflib/vcflib
R (v4.3.0)	R Core Team, 2020	https://www.R-project.org/
Rstudio (v2024.04.1 + 748)	R Core Team, 2020	http://www.rstudio.com/

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Mouse models of infection

All parasites were maintained in 6- to 8-week-old male and female C57BL/6 *Ifnγ* [-/-] (Jackson Laboratory stock no. 002287) mice bred in-house. C57BL/6 wildtype mice were purchased from Jackson. Protocols for animal maintenance, care and experimentation were approved by the Institutional Animal Care and Use Committee of the University of Pennsylvania (protocol #806292).

Parasite strains

Oocysts of *C. parvum* strains were obtained originally from the following sources: Iowa II (BG), Bunch Grass Farms, Deary, Idaho; Iowa II (ST), Dr. Reed, University of Arizona; LPZ and Ber, Dr. Olias, University of Giessen, Germany, and Dr. Kücken, Free University Berlin, Germany; INRAE,[71] Dr. Fabrice Laurent, INRAE and University of Tours, Nouzilly, France. KVI (Kimron Veterinary Institute) strain oocysts were obtained in March 2023 from a clinical case of cryptosporidiosis in a 2.5-week-old lamb on a farm in Beit Dagan, Israel. Fecal samples found to harbor oocysts detected using modfied Ziehl-Neelsen staining were subjected to DNA extraction using the Presto Stool DNA Extraction Kit (Geneaid, Taiwan). The species of *Cryptosporidium* was determined by sequencing PCR ampli-fied SSU rRNA[72,73] and subtyped using GP60 amplicons.[74] Amplicons were matched to sequences deposited in GenBank using BLAST,[75] and GP60 subtype was established by multiple sequence alignment with IIa and IId subtypes and inspection.[76] For isola-tion, oocysts were subjected to sedimentation as described in the WOAH Terrestrial Manual[77] with the omission of ethyl acetate, and

all centrifugation steps were performed for 10 min at 2000 RPM and 4°C. Recovered oocysts were washed once in 0.5% household bleach prior to wash, storage and shipment in PBS at 4°C.

Cell culture

HCT8 obtained from ATCC were cultured in DMEM supplemented with 10% FBS at 37°C in presence of 5% CO_2.

METHOD DETAILS

Generation of transgenic parasites

To generate transgenic parasites, 1.25×10^7 oocysts were incubated at 37°C for 1 h in 10 mM HCl followed by two washes with phosphate buffered saline (PBS) and an incubation at 37°C for 1 h in 0.2 mM sodium taurocholate and 22 mM sodium bicarbonate to induce excystation.[78] Excysted sporozoites were electroporated and used to infect four-to eight-week-old mice pretreated with antibiotics by oral gavage as previously described.[27,79] Transgenic parasites were selected with paromomycin in the drinking water (16g/L) or BRD7929 by daily gavage (5mg/kg/day formulated in in 70% PEG400 and 30% aqueous glucose (5% in H_2O). Feces were collected from each mouse in the cage, mixed, and 20 mg were used to measure parasite burden by nLuc luciferase assay.[27,79] All transgenic parasites generated in this study are documented in Figure S2 and in STAR Methods Table S2.

KVI tdNeon

The repair encodes the last 113 bp of the pheRS gene (*cgd3_3320*, recodonized) including the mutation that confers resistance (L482V) to BRD7929. This short sequence is followed by a nanoluciferase reporter fused to tdNeonGreen under the enolase promoter.

Iowa with ectopic GP60-KVI

The repair encodes the GP60-KVI version under its own promoter and 3′UTR followed by a nanoluciferase reporter fused to the *Neo* resistance gene under the enolase promoter. It is integrated into the thymidine kinase locus.

Iowa with ectopic GP60-KVI and endogenous GP60 knockout

The repair encodes the pheRS gene including the mutation that confers resistance (L482V) to BRD7929 followed by a nanoluciferase reporter under the enolase promoter. The parental strain for this transgenic is the Iowa with ectopic GP60-KVI strain.

KVI with ectopic GP60-Iowa

The repair encodes the GP60-Iowa version under its own promoter and 3′UTR followed by a nanoluciferase reporter fused to the *Neo* resistance gene under the enolase promoter. It is integrated into the thymidine kinase locus.

KVI with ectopic GP60-Iowa and endogenous GP60KO

The repair encodes the pheRS gene including the mutation that confers resistance (L482V) to BRD7929 followed by a nanoluciferase reporter under the enolase promoter. The parental strain for this transgenic is the KVI with ectopic GP60-Iowa strain.

KVI cgd7_4490 knockout

The repair encodes a nanoluciferase reporter fused to the *Neo* resistance gene under the enolase promoter.

KVI cgd7_4490 Ha-tag

The repair encodes the HA-tag followed by a nanoluciferase reporter fused to the *Neo* resistance gene under the enolase promoter.

KVI cgd7_4500 knockout

The repair encodes a nanoluciferase reporter fused to the *Neo* resistance gene under the enolase promoter.

KVI cgd7_4500 Ha-tag

The repair encodes the HA-tag followed by a nanoluciferase reporter fused to the *Neo* resistance gene under the enolase promoter.

KVI cgd7_4500 knockout and endogenous Iowa II cgd7_4500

The repair encodes the last 113 bp of the pheRS gene (*cgd3_3320*, recodonized) including the mutation that confers resistance (L482V) to BRD7929. This short sequence is followed by the Iowa II *cgd7_4500* version under its own promoter and 3′UTR. The parental strain for this transgenic is KVI *cgd7_4500* knockout strain.

KVI cgd7_4500 knockout and endogenous Iowa II cgd7_4500

The repair encodes the last 113 bp of the pheRS gene (cgd3_3320, recodonized) including the mutation that confers resistance (L482V) to BRD7929. This short sequence is followed by the *KVI cgd7_4500* version under its own promoter and 3′UTR. The parental strain for this transgenic is *KVI cgd7_4500* knockout strain.

Cell culture and *in vitro* growth assay

HCT-8 cells were purchased from ATCC (CCL-224TM) and maintained in RPMI 1640 medium (Sigma-Aldrich) supplemented with 10% Cosmic calf serum (HyClone). Cells were infected with 40000 excysted sporozoites and serum was reduced to 1%. Medium was aspirated after 24h and 48 h, cells were lysed and mixed with NanoGlo substrate (Promega) and luminescence was measured using a Glomax reader (Promega). For normalization, luminescence was measured after 4h post infection (input).

Microscopy

Infected cells were fixed with 4% paraformaldehyde (Electron Microscopy Science) in PBS and stained with Hoechst (1 μg/mL). Coverslips were then mounted on glass slides with fluorogel (Electron Microscopy Science) mounting medium. Ultrastructure expansion

microscopy was performed as previously described.[29,80,81] Sporozoites were allowed to excyst in culture media for 2 h before allowing them to settle onto poly-D-lysine-coated coverslips prior to fixation. For intracellular stages cells were fixed 24 h post-infection. Crosslinking prevention and anchoring were performed in a mixture of 2% acrylamide and 1.4% formaldehyde overnight at 37°C. Samples were embedded in a hydrogel, which was subjected to denaturation, expansion, and staining with antibodies to detect HA epitope-tagged proteins, Atto647-conjugated NHS ester dyes to detect protein density, and Hoechst to detect DNA. Stained gels were imaged on a Leica Stellaris 8 inverted confocal microscope with a 63× water immersion objective.

qPCR

Infections were set up in a 24-well plate, with 200,000 oocysts per well. RNA was extracted at 48 h by direct lysis in the well with 350 μL of RLT lysis buffer from the RNeasy minikit (Qiagen). The manufacturer's instructions were followed, and RNA was eluted in 40 μL of RNase-free water. cDNA was prepared per sample using SuperScript IV reverse transcriptase (Thermo Fisher Scientific). The cDNA was diluted 1:10 before setting up the quantitative PCR (qPCR) in a 10 μL reaction, which included SsoAdvanced Universal SYBR Green Supermix (Bio-Rad). The reaction was loaded into the ViiA 7 Real-Time PCR system (Thermo Fisher Scientific) with the following conditions: initial incubation at 95°C for 3 min, 95°C for 15 s and 60°C for 30 s for 40 cycles, and a single melt curve. The $\Delta\Delta C_t$ method was used to determine the relative expression with GAPDH as the control gene. A list of all of the primers is provided in Table S2.

Genetic crosses and sample collection

In total, five crosses between the *C. parvum* Iowa II (BG) and KVI strains were conducted. 3 crosses with initial dual drug selection (paromomycin and BRD7929). Mice were infected with 2000/8000 (KVI/BG; cross 1 and cross 2) and 1000/4000 (KVI/BG; cross 3) oocysts per mouse, respectively. Recombinant progeny was selected by dual drug selection from day 4 to day 10. 2 crosses were conducted in the absence of drug selection with 100/9900 (KVI/BG) oocysts per mouse for each cross. Fecal material was collected daily and stored at 4°C. 14 Segregant pools were collected as indicated in Figure S4 (gray boxes).

Oocyst purification and genomic DNA extraction

Fecal material of infected mice was collected and oocyst were purified using a sucrose gradient and CsCl flotation.[79] Genomic DNA was extracted using phenol/chloroform as described.[23]

Library preparation and sequencing of cross progeny

We prepared Illumina libraries from extracted genomic DNA and sequenced both parents and 14 segregant pools. The library preparation was carried out using Illumina DNA Prep (former Nextera DNA Flex kit, Illumina Inc.). Subsequently, sequencing was performed on the Illumina NextSeq 2000 sequencer, utilizing P2 300 cycle flowcell kit. We prepared Oxford Nanopore library from whole genome amplified genomic DNA using REPLI-G MINI kit WGA (CAT #150023, Qiagen, Hilden, Germany) followed by a T7 Endonuclease I debranching (CAT #M0302, New England BioLabs, Ipswich, MA) as described.[82] We used native barcoding kit SQK-NBD114.24 to construct ligation sequencing library. Total of 400ng of final library was loaded onto an R10.4.1 flowcell, and sequenced on an Oxford Nanopore MinION machine for 72 h.

QUANTIFICATION AND STATISTICAL ANALYSIS

De novo genome assembly

Illumina short read (Nextseq 2000, 2 × 150bp) and Nanopore long read (MinION R10.4.1) sequences were generated and used for genome assembly. Illumina short reads were adapter and quality (>Q30) trimmed with TrimGalore (v0.6.10). Raw nanopore reads are base called and demultiplexed with ONT Dorado (v0.9.1) with Super Accuracy (SUP) plus duplex mode. CADECT (v1.0.2) was used on base called long reads to remove concatemers resulted from whole genome amplification. Long reads were assembled with Flye (v2.9.5) into 69 contigs. Hybrid genome assembly was performed with SPAdes (v4.0.0) using both illumina short-read and Flye long-read assembled contigs. NCBI fcs-adapter (v0.5.0) and fcs-gx(v0.5.0) were used to remove adapter contamination and non-*Cryptosporidium* sequences from the assembled genome. The genome was re-orientated and scaffolded based on the T2TCpBGF genome[25] with RagTag(v2.1.0) into 8 chromosomes. The genome was polished with 3 rounds of illumina short-reads and 3 rounds of Nanopore long-reads, both using Racon (v1.4.3). The final genome size is 9,083,795bp and we identified 11 out of 16 telomeres from the assembly using FindTelomeres. The consistency between the KVI *de novo* assembly and the T2TCpBGF genome was assessed using MUMmer (Figure S3). The Iowa II and KVI genomes aligned well, showing high similarity with only a few small inversions across the genome. A single translocation (∼500bp) was observed at the end of chr 7, caused by rearrangement of the Hemogen gene. The inversions ranged from 100 bp to 53 kb, which are unlikely to impact recombination. We annotated the KVI genome with Liftoff (v1.6.3) using the T2TCpBGF annotation as well as *de novo* BRAKER (v3.0.8). Annotations from Liftoff and BRAKER3 were merged and cleaned with AGAT (v1.4.0). The final annotation of KVI identified 3928 protein coding, 45 tRNA, and 15 rRNA genes.

Reference KVI genome can be accessed from BioProject: PRJNA1263880. Alternatively, KVI genome is fasta format and annotation in gff format can also be accessed here: https://github.com/ruicatxiao/Automated_Bulk_Segregant_Analysis/tree/main/KVI_genome_annotation.

Genotype calling

The Iowa II *de novo* assembly (T2T reference genome[25]) was then used as the reference genome to identify single-nucleotide polymorphism (SNP) distinguishing the two parents, which were then used for bulk segregant analysis. Whole-genome sequencing reads for each library were individually mapped to the Iowa II *de novo* assembly using the BWA-MEM alignment algorithm with default parameters. The resulting alignments were converted to SAM format, sorted into BAM format, and deduplicated using Picard tools. Variants for each sample were called using HaplotypeCaller and were subsequently aggregated across all samples using GenotypeGVCFs (see https://github.com/ruicatxiao/Automated_Bulk_Segregant_Analysis for detailed parameters).

We compared the genome-wide distribution of SNPs between the parental parasites and identified three regions with extremely high SNP density (>30 SNPs per kb) (Figure 1C). The regions with high SNP density include: chr 4:1,099kb–1,106kb (7kb), chr 5: 0-12kb (12kb), and chr 6: 4kb–14kb (10kb). We excluded these regions from further bulk segregant analysis, leaving a total of 3,570 SNPs for subsequent analyses. We also identified four genomic regions with notably low SNP density (<5 SNPs per 100 kb). The regions with low SNP density include: chr 1: 390kb–700kb (310kb), chr 2: 345kb–559kb (214kb), chr 4: 199kb–530kb (331kb), and chr 7: 387kb–530kb (143kb). Similar high SNP density distribution patterns were observed in population data at the same locations[83] (Figure 4), suggesting that these are common hypervariable regions in *C. parvum*. However, no regions with extremely low SNP density were identified in the currently published population data.

Bulk segregant analysis

SNP loci with coverage below 30× in either of the compared pools were excluded from each bulk segregant analysis. At each variable locus, we counted reads corresponding to the genotypes of each parent and calculated allele frequencies. Iowa II allele frequencies were plotted across the genome (Figure S6), and outliers were removed using Hampel's rule with a window size of 100 loci. Bulk segregant analysis was performed using the R package *QTLseqr*. Extreme QTLs were defined as loci with false discovery rates (FDRs, Benjamini-Hochberg adjusted *p*-values) below 0.01. Once a QTL was detected, a 95% confidence interval was estimated using Li's method to localize causative genes.[84]

ADDITIONAL RESOURCES

Automated bulk segregant analysis

We also developed an Automated Bulk Segregant Analysis (ABSA) pipeline implemented in Python (v3.10.12) to streamline the workflow. This links the software needed: raw reads are quality-trimmed using Trim Galore, aligned to a reference genome with BWA, processed using SAMTools, and variants are called with GATK4 (v4.6.1.0) and filtered via vcffilter. Downstream analyses and visualization are performed in R. ABSA supports both manual installation and containerized deployment via Singularity, with a provided Singularity definition file based on Ubuntu 24.04 that bundles all required dependencies into a single, reproducible image. The code can be accessed at (https://github.com/ruicatxiao/Automated_Bulk_Segregant_Analysis).