

RESEARCH ARTICLE

# Sexual reproduction in a natural *Trypanosoma cruzi* population

Alexander S. F. Berry<sup>1</sup>, Renzo Salazar-Sánchez<sup>2</sup>, Ricardo Castillo-Neyra<sup>2,3</sup>, Katty Borrini-Mayorí<sup>2</sup>, Claudia Chipana-Ramos<sup>2</sup>, Melina Vargas-Maquera<sup>2</sup>, Jenny Ancca-Juarez<sup>2</sup>, César Náquira-Velarde<sup>2</sup>, Michael Z. Levy<sup>2,3</sup>, Dustin Brisson<sup>4\*</sup>, the Chagas Disease Working Group in Arequipa<sup>†</sup>

**1** Division of Gastroenterology, Hepatology, and Nutrition, Children’s Hospital of Philadelphia, Philadelphia, Pennsylvania, United States of America, **2** Universidad Peruana Cayetano Heredia/University of Pennsylvania Chagas Disease Field Laboratory, Arequipa, Peru, **3** Department of Biostatistics, Epidemiology and Informatics, The Perelman School of Medicine of the University of Pennsylvania, Philadelphia, Pennsylvania, United States of America, **4** Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania, United States of America

<sup>†</sup> Membership of the Chagas Disease Working Group in Arequipa is listed in the Acknowledgments.

\* [dbrisson@sas.upenn.edu](mailto:dbrisson@sas.upenn.edu)



**OPEN ACCESS**

**Citation:** Berry ASF, Salazar-Sánchez R, Castillo-Neyra R, Borrini-Mayorí K, Chipana-Ramos C, Vargas-Maquera M, et al. (2019) Sexual reproduction in a natural *Trypanosoma cruzi* population. PLoS Negl Trop Dis 13(5): e0007392. <https://doi.org/10.1371/journal.pntd.0007392>

**Editor:** Daniella Castanheira Bartholomeu, Universidade Federal de Minas Gerais, BRAZIL

**Received:** November 7, 2018

**Accepted:** April 15, 2019

**Published:** May 20, 2019

**Copyright:** © 2019 Berry et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All genomic sequencing data are available on the Sequence Read Archive (SRA) accession number PRJNA517719.

**Funding:** This work was funded by grants from the National Institute of Health (<https://www.nih.gov/>) (T32HG000046[AB], T32AI007532[AB], R01AI142572[DB], R01AI097137[DB], and NIAID 5R01AI101229[MZL], NIAID 5P50AI074285[CN]), the National Science Foundation (<https://nsf.gov/>) (DEB-1354184[DB]), and Burroughs Wellcome

## Abstract

### Background

Sexual reproduction provides an evolutionary advantageous mechanism that combines favorable mutations that have arisen in separate lineages into the same individual. This advantage is especially pronounced in microparasites as allelic reassortment among individuals caused by sexual reproduction promotes allelic diversity at immune evasion genes within individuals which is often essential to evade host immune systems. Despite these advantages, many eukaryotic microparasites exhibit highly-clonal population structures suggesting that genetic exchange through sexual reproduction is rare. Evidence supporting clonality is particularly convincing in the causative agent of Chagas disease, *Trypanosoma cruzi*, despite equally convincing evidence of the capacity to engage in sexual reproduction.

### Methodology/ Principle Findings

In the present study, we investigated two hypotheses that can reconcile the apparent contradiction between the observed clonal population structure and the capacity to engage in sexual reproduction by analyzing the genome sequences of 123 *T. cruzi* isolates from a natural population in Arequipa, Peru. The distribution of polymorphic markers within and among isolates provides clear evidence of the occurrence of sexual reproduction. Large genetic segments are rearranged among chromosomes due to crossing over during meiosis leading to a decay in the genetic linkage among polymorphic markers compared to the expectations from a purely asexually-reproducing population. Nevertheless, the population structure appears clonal due to a high level of inbreeding during sexual reproduction which increases homozygosity, and thus reduces diversity, within each inbreeding lineage.

Fund (<https://www.bwfund.org/>) (1012376[DB]). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## Conclusions/ Significance

These results effectively reconcile the apparent contradiction by demonstrating that the clonal population structure is derived not from infrequent sex in natural populations but from high levels of inbreeding. We discuss epidemiological consequences of this reproductive strategy on genome evolution, population structure, and phenotypic diversity of this medically important parasite.

### Author summary

The rearrangement of alleles among individuals in a population during sexual reproduction maintains high allelic diversity within individuals in a population at polymorphic genes. Allelic diversity within individuals can be particularly important for parasites as it enhances their ability to evade host immune systems. Despite the potential benefits of sexual reproduction for parasites, natural populations of the protozoan parasite—and causative agent of human Chagas disease—*Trypanosoma cruzi*, exhibit clonal population structures indicative of asexual reproduction. This is particularly surprising as *T. cruzi* has the capacity for sexual reproduction. Here, we resolve this apparent contradiction by sequencing whole genomes of 123 *T. cruzi* isolates from a natural population in Arequipa, Peru. Evidence of past sexual reproduction and allelic rearrangements are common in this *T. cruzi* population. However, the majority of sexual reproduction events occur between close relatives resulting in an apparent clonal population structure. Sexual reproduction with distant relatives in areas with greater strain diversity has the potential to affect public health by increasing diversity in immune evasion genes within individuals and enhancing within-host survival, rapidly diversifying antigens that could affect the sensitivity of serological diagnostics, and by generating diversity in pathogenicity or drug resistance.

## Introduction

An increasing body of evidence suggests that many eukaryotic microparasites exhibit highly-clonal population structures [1–9]. That is, isolates sampled in nature tend to represent independently evolving, asexual lineages with limited genetic exchange among lineages. Despite increasingly detailed studies supporting the apparent rarity of sexual reproduction in many eukaryotic microparasites, this diverse group of organisms is capable of sexual reproduction and evidence of sex has been observed using population genetic methods and experimentally demonstrated in several species [10–16]. While sexual reproduction can generate untested combinations of alleles that are less successful than the parental combinations on average (recombinational load [17]), sex is also expected to provide several evolutionary advantages to parasites including increasing within-individual diversity in immune evasion genes where diversity is paramount for immune evasion and thus survival. We investigated two hypotheses that could explain the apparent contradiction between the observed clonal population structure and the capacity to engage in sexual reproduction in the protozoan parasite and causative agent of Chagas disease, *Trypanosoma cruzi*. Like many microparasites, *T. cruzi* tends to exhibit highly clonal population structures in nature [18–20] with rare exceptions [21,22], despite experimental evidence of the capacity for sexual reproduction [23]. In the first hypothesis, meiosis and fertilization are exceedingly rare such that mutational processes occur much

more frequently than meiosis (Fig 1) and in the second, sexual reproduction is common but occurs between closely-related individuals resulting in the high homozygosity across genomes observed in natural populations.

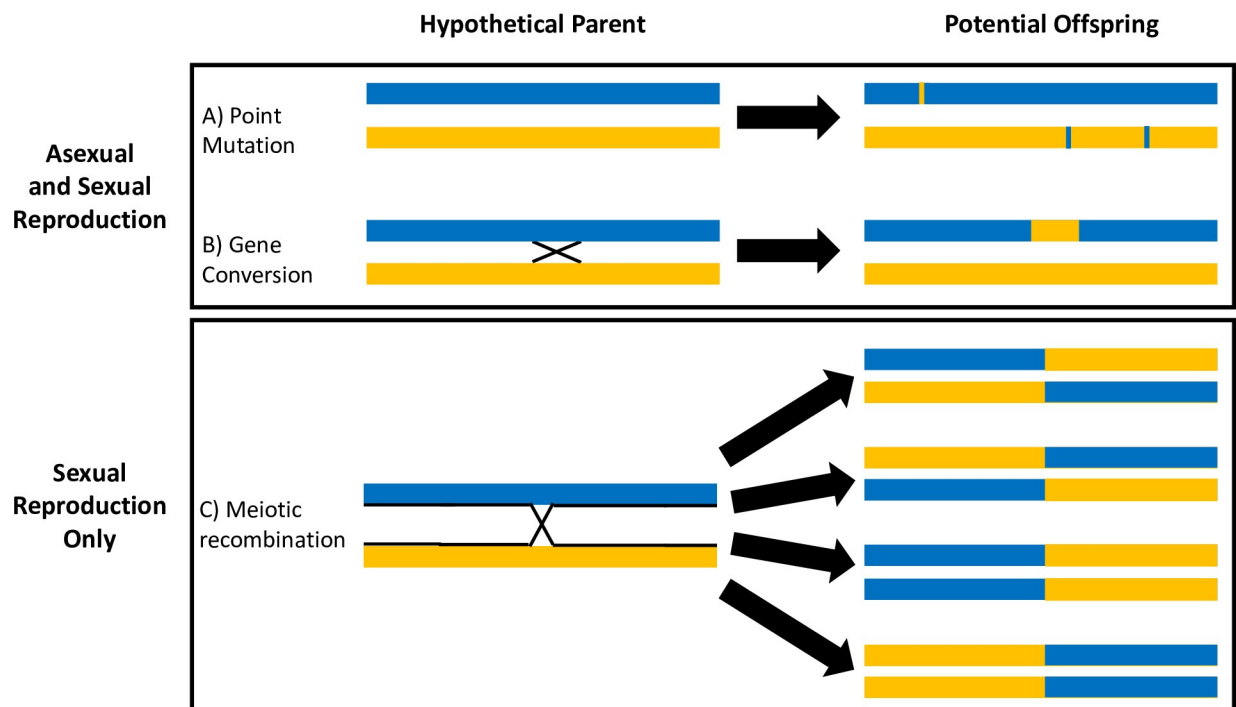
The clonal population structure of *T. cruzi* is well-documented leading many to accept that this parasite reproduces asexually despite its capacity for sexual reproduction [5]. Asexual reproduction results in an excess of homozygosity compared to sexual populations as well as non-random associations of alleles across loci (linkage disequilibrium), both of which have been observed in natural *T. cruzi* populations [18–20]. For example, Oliveira *et al.* 1998 [18] found that homozygosity at, and linkage disequilibrium between, eight microsatellite loci among 24 *T. cruzi* strains was significantly greater than expected assuming random mating, suggesting a limited role for sexual reproduction [5]. Nevertheless, the capacity to engage in sexual reproduction in *T. cruzi* has been demonstrated in controlled experiments where clones transfected with different markers were mixed *in vitro* and progeny harboring both markers were observed [23]. Additionally, hybrid clones have been isolated from natural populations that formed through sexual reproduction between distantly-related lineages [24–28].

The demonstrated capacity of *T. cruzi* to reproduce sexually along with the potential advantages of sex is remarkable given that the population genetic patterns associated with sex (Fig 1) are not observed in natural populations [18–20]. That is, regular sexual reproduction randomly combines alleles at each locus resulting in the proportions of heterozygotes and homozygotes that conform to Hardy-Weinberg expectations, a result that has not been observed in any *T. cruzi* population investigated [18–20]. Similarly, sexual reproduction is expected to break down allelic associations among loci due to the independent assortment of chromosomes as well as cross-over events that occur during meiosis. By contrast, empirical studies demonstrate high levels of linkage disequilibrium among alleles in natural populations [18,29,30]. In the present study, we investigated the pattern of genomic signatures associated with the combinations of allelic markers expected in sexual and asexual populations to investigate the potential causes of the apparent contradiction between the observed clonal population structure and the capacity to engage in sexual reproduction.

## Results

The genome sequences from 123 *T. cruzi* isolates collected throughout Arequipa, Peru (Fig 2) were highly similar, suggesting that all isolates descended from a recent common ancestor. The 15,357bp maxicircle was nearly identical among isolates with an average pairwise difference of 1.04 [31] and only 12,256 sites in the ~56Mbp diploid genome were polymorphic. There is no evidence of introgression of chromosomes or segments of chromosomes into this population from outside populations as all runs of polymorphic markers are distributed throughout the phylogeny of these 123 isolates. Additionally, there are no runs of polymorphic markers that are similar to any of the outgroup genome sequences. These data suggest that the common ancestor of all sampled individuals is relatively recent and contained much of the variation observed in the isolates sequenced.

Evidence of sexual reproduction was inferred from the distribution of polymorphic markers within and among sequenced isolates. The exchange of genetic material between non-sister chromatids of homologous chromosomes during meiosis results in gametes that contain chromosomes comprised of a set of polymorphic markers derived from one parental chromosome and a set derived from the other parental chromosome. When these gametes fuse and produce diploid offspring, some offspring will have regions of their chromosomes that are uniformly homozygous at the polymorphic markers while the same region will be heterozygous in other offspring (Fig 1C). Evidence of sexual reproduction was detected by identifying large segments



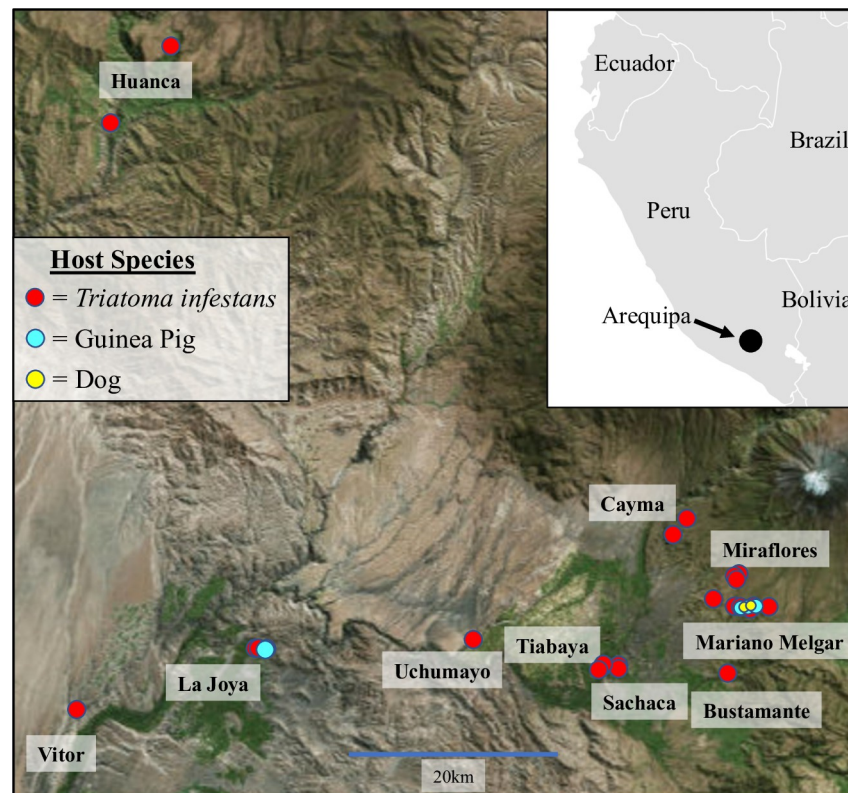
**Fig 1. Reproductive strategies can be inferred from the distribution of polymorphic markers within and among strains in a population.** Both asexual and sexual populations experience both A) point mutations and B) gene conversion events while only populations reproducing sexually experience C) meiotic recombination that reassort large genomic regions among chromosomes as well as chromosomes among individuals. A) Point mutations alter random bases individually such that spatial clustering of mutations is not expected in most cases. Whole chromosomes carrying novel point mutations, along with all linked markers, will be passed to offspring during asexual reproduction. These markers can disassociate due to cross-over events and independent assortment of chromosomes during meiosis and sexual reproduction. B) Gene conversion events homogenize small regions of homologous chromosomes (<10kb), effectively reducing diversity at multiple markers in one chromosomal region. C) Sexual reproduction can result in the disassociation of polymorphic markers due to the independent assortment of chromosomes as well as crossing over between homologous chromosomes during meiosis. Crossing over results in the exchange of large chromosomal segments (up to Mbs) among homologous chromosomes, disrupting associations among polymorphic markers on each chromosome. Crossing over thus results in gametes that contain chromosomes comprised of a set of polymorphic sites derived from the one parental chromosome and a set derived from the other parental chromosome. As only a small number of cross-over events are expected to occur on each chromosome during meiosis (~1), a small number of large genomic regions reassort on each resulting chromosome, a genetic signature that is distinct from point mutations and gene conversion. Random fusion of gametes with these recombined chromosomes will produce offspring with some chromosomal regions that are uniformly homozygous at polymorphic markers which are heterozygous in their siblings.

<https://doi.org/10.1371/journal.pntd.0007392.g001>

(>10kb) of chromosomes where all of the markers ( $\geq 20$ ) that are heterozygous in some isolates are homozygous in other isolates within the same population. Such linkage blocks containing loci in high linkage disequilibrium are common in the *T. cruzi* genomes in Arequipa (Fig 3; S1 Fig). Using a conservative algorithm to detect chromosomal segments derived from meiotic recombination, we identified 474 recombined segments ranging in size from 10kb to 468kb (S2 Fig). Evidence of meiotic recombination events were detected on 69 of the 80 (86.25%) contigs longer than 100kb, suggesting that this phenomenon occurs throughout the *T. cruzi* genome.

The conservative criteria used to identify these meiotic recombination events are unlikely to be met on contigs shorter than 100kb (S3 Fig). Reassortment of chromosomal segments through sexual reproduction (meiotic recombination or independent assortment of chromosomes) accounts for 95% of the average pairwise differences among *T. cruzi* isolates, with the remaining differences explained by point mutations.

Isolates derived from geographically-proximal locations commonly have the same recombination events, suggesting identity-by-descent (S4 and S5 Figs). Samples collected in the same



**Fig 2. Geographic distribution of isolates.** The location of all 123 sampled *T. cruzi* are represented by dots and colored by the host species. Inset shows the location of Arequipa. Map of Arequipa was obtained from <https://landlook.usgs.gov> and was modified using QGIS v. 2.18 [69]. Inset was modified from [https://commons.wikimedia.org/wiki/Atlas\\_of\\_the\\_world](https://commons.wikimedia.org/wiki/Atlas_of_the_world).

<https://doi.org/10.1371/journal.pntd.0007392.g002>

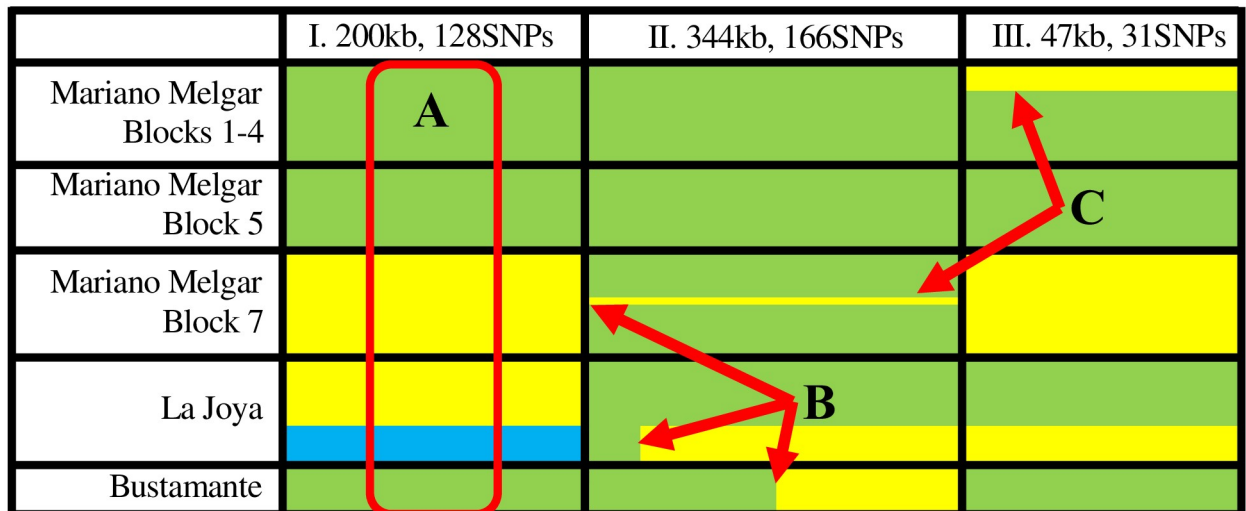
year or from the same host, as opposed to collections from geographically similar areas, do not form monophyletic clades (S5 Fig). These data suggest that geography is the most important predictor of phylogenetic relationship.

The frequency of heterozygotes in this population is higher than expected from a Hardy-Weinberg population at 95% of polymorphic markers while only 0.10% of markers have more homozygotes than expected. Of those markers in Hardy-Weinberg equilibrium, only one (0.2%) site had a minor allele frequency greater than 0.1. Of the ten markers with more homozygotes than expected, only 1 (10%) had a minor allele frequency greater than 0.1. All polymorphic markers contain only two character states.

## Discussion

The evolutionary advantages of sex for eukaryotic microparasites are numerous and include the potential to increase the diversity of immune evasion genes within individuals where diversity is paramount for survival within hosts. Yet many populations of eukaryotic microparasites that have the capacity for sexual reproduction exhibit a clonal population structure suggesting that sexual reproduction either does not occur with sufficient regularity to disrupt clonal population structures in nature or that sexual reproduction occurs primarily among closely related individuals. The genome sequence data analyzed here supports the hypothesis that *T. cruzi*, a eukaryotic microparasite in which clonal population structures are commonly observed despite the demonstrated capacity for sexual reproduction, engages in sexual reproduction





**Fig 3. The distribution of polymorphic markers within and among individuals provides clear evidence of meiotic recombination.** This representative subset of the data portrays the distributions of 325 polymorphic markers throughout three regions of the genome in 54 isolates (grouped by geographic location: Mariano Melgar Blocks 1–4 (N = 14), Block 5 (N = 11), Block 7 (N = 14), La Joya (N = 14), and Bustamante (N = 1)) that indicate the presence of sexual reproduction as A) the independent assortment of chromosomes among individuals in a population; B) meiotic recombination events that reassort polymorphic markers along a chromosome; and C) evidence of very recent reassortment of polymorphic markers via sexual reproduction. Markers that are polymorphic in the population can be heterozygous (green) or homozygous in each individual (yellow representing a site that is homozygous for one base, blue representing individuals homozygous for the alternative base). For clarity, only polymorphic sites with a minor allele frequency >16% in the population are shown here. (A) Independent assortment of chromosomes and subsequent fusion of gametes has resulted in a region (I) that is homozygous in some individuals and heterozygous in others across a 200kb region containing 128 polymorphic markers on an 829kb contig. In this region, individuals can be homozygous for either set of linked polymorphic markers. Interestingly, genotypes are geographically clustered, indicating identity by descent. Runs of identical color indicate identical sequence. (B) Meiotic recombination has resulted in a region (II) of a 357kb contig that partially retain the linkage patterns of both parental chromosomes while the rest retains the linkage pattern of only one parental chromosome, showing how crossing over can affect parts of chromosomes. In this region, there is evidence of at least three independent meiotic recombination events. (C) Sexual reproduction after a recent colonization of a city block has resulted in the genetic divergence of closely related samples over a run of polymorphic markers. For example, a meiotic recombination event is apparent in region II in one sampled individual on Block 7. The remaining contigs (S1 Fig) contain near-complete sequence similarity among all clones isolated from Block 7 suggesting that all individuals on this block share a recent common ancestor and the meiotic event occurring in region II occurred after the Block 7 was colonized. Similarly, a meiotic recombination event in a region (III) of a 130kb contig occurred in one lineage inhabiting Blocks 1–3 in Mariano Melgar. The regions shown here contain only some of the 474 meiotic recombination events distributed across 151 contigs, with each detected meiotic recombination event spanning from 10kb to nearly 500kb. The relative positions of each contig within the genome are unknown. For data spanning the whole genome, see S1 Fig.

<https://doi.org/10.1371/journal.pntd.0007392.g003>

with high levels of inbreeding. The clonal population structure observed in this and other *T. cruzi* populations is mostly likely accounted for by periodic meiosis and the fusion of gametes with those of closely related individuals.

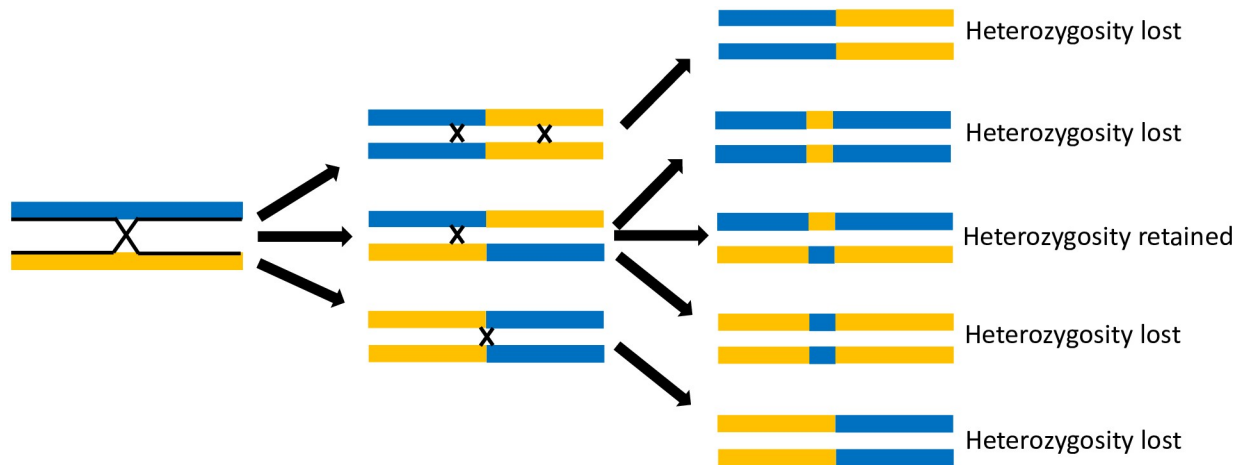
The meiotic processes necessary for gamete formation and sexual reproduction create novel associations of polymorphic markers through recombination between homologous chromosomes. Fusion of a gamete containing a recombined chromosome with a gamete from the same individual containing one of the parental chromosomes during sexual reproduction will result in homozygosity throughout one large region of the chromosome while the sites that were heterozygous in the parent will remain heterozygous in the offspring in the other chromosomal region. Here, we looked for evidence of recombined chromosomes derived from meiotic processes to detect evidence of sexual recombination. Among the genome sequences from 123 *T. cruzi* isolates collected in Arequipa, Peru, 474 independent meiotic recombination events were identified where all of the markers ( $\geq 20$ ) that are heterozygous in some isolates are homozygous in other isolates across a large chromosomal segment ( $>10$ kb). Sexual reproduction occurs with sufficient regularity to observe meiotic recombination events within subpopulations that have only recently established in blocks within the city (Fig 3C).

These types of meiotic recombination events are present throughout the *T. cruzi* genome and range in size from 10kb to 468kb. The distributions of polymorphic markers observed across large segments of DNA is substantially more likely to result from meiotic recombination than identical mutations recurring in multiple lineages or through gene conversion events which impact shorter chromosomal regions (<10kb).

The population structure of *T. cruzi* appears clonal despite the occurrence of sexual reproduction in this natural population due to both the irregularity of sexual reproduction and the high level of inbreeding. Despite evidence of multiple meiosis and fertilization events in this relatively young population, asexual reproduction remains the dominant form of *T. cruzi* reproduction. The dominance of asexual reproduction has maintained both the higher than expected levels of heterozygosity observed at the majority of polymorphic markers as well as the high levels of linkage disequilibrium among the polymorphic markers within the observed linkage blocks on many contigs. In agreement with disproportionate asexual reproduction, heterozygote frequencies in this population are higher than expected from a randomly-mating sexual population in 95% of polymorphic markers. Although heterozygosity is expected to decay with repeated inbreeding, the limited number of sexual reproduction events experienced in this population since the heterozygous common ancestor have resulted in only a moderate reduction in heterozygosity among the extant descendants. The geographic structure of this population, where subpopulations cluster within city blocks (S4 and S5 Figs), may prevent gene flow, and thus recombination, between subpopulations in different geographic regions [22]. This phenomenon, known as the Wahlund effect [32], is expected to cause deviations from the expected Hardy-Weinberg equilibrium and level of linkage disequilibrium, even when sexual recombination is occurring within geographically-isolated subpopulations. These data also suggest that the common ancestor of all sampled *T. cruzi* genomes contained nearly all of the polymorphisms analyzed, which is also supported by data showing that most (>95%) of the genetic diversity among isolates results from the reassortment of a set of ancestral polymorphic markers. While the data suggest that the common ancestor was likely heterozygous at most polymorphic sites identified here, heterozygous sites likely comprised a very small percentage of the ancestral genome because less than 0.1% of sites are polymorphic in this population.

Variation in the number of chromosomes within each *T. cruzi* strain, or aneuploidy [23,33], can hinder population genomic analyses. Although the proportion of base counts derived from sequencing depth across heterozygous sites suggests this population is primarily diploid (S6 Fig), the conclusions presented here regarding the number and location of sexual recombination events observed are robust to variable ploidy levels. That is, many contigs contain internal recombination breakpoints that are detected in multiple isolates (Fig 3B), suggesting that these chromosomes have experienced a meiotic recombination event, regardless of the number of other chromosomes in the strains where these contigs were identified.

High levels of inbreeding also cause the population structure of this *T. cruzi* population to appear clonal. Inbreeding limits the exchange of diverse alleles among individuals within the population resulting in multiple independently evolving lineages. Further, inbreeding reduces the benefits of sexual reproduction as it results in the continuous decay of diversity within each lineage until the absorbing boundary of complete homozygosity is attained (Fig 4). That is, only half of the offspring that result from self-mating, where gametes fuse only with gametes of the same parent, in a heterozygote lineage will be heterozygous while all offspring of homozygotes would be homozygous, thus reducing the proportion of heterozygotes in the population by half with each generation. No evidence of outcrossing was detected in the current dataset (Fig 3A). It is unclear, however, if the apparent inbreeding results from obligatory selfing or if outbreeding occurs but cannot be observed due to the limited genome-wide diversity



**Fig 4. Inbreeding results in a decay of heterozygosity.** After one generation of inbreeding, half of the offspring retain the heterozygosity found in the parent while half become homozygous. As diversity cannot be restored in homozygous regions in the absence of outcrossing and heterozygosity decays by half each generation, population-level heterozygosity continually declines in inbreeding populations until all lineages are homozygous.

<https://doi.org/10.1371/journal.pntd.0007392.g004>

in Arequipa. The obligatory selfing hypothesis is supported by published investigations in *T. cruzi* populations with greater diversity in which multiple *T. cruzi* lineages are frequently found in the same host [34] and vector [18,35] without evidence of outcrossing. Future studies are necessary to assess the population or molecular mechanisms that result in high levels of selfing in areas with opportunities for outcrossing.

The biological mechanism of sexual reproduction in *T. cruzi* may restrict outcrossing. The location (in what host or host tissue), frequency, and mechanism affecting the probability of outcrossing in *T. cruzi* have not been elucidated, although it is known that *Trypanosoma brucei* produces gametes in its vector [12]. By contrast, experimental evidence suggests that *T. cruzi* undergoes recombination in the mammalian host [23]. Haploid gametes have not yet been detected in *T. cruzi*, potentially due to the practical challenges of isolating an ephemeral stage of a relatively rare organism, or that *T. cruzi* recombination may occur via parasexual recombination. Parasexual reproduction, which occurs in several eukaryotic microbes, is the creation of a tetraploid cell through the fusion of diploid cells, followed by chromosomal crossing over during mitosis and stochastic chromosome loss to return to a diploid state [23,36,37]. The data presented here are broadly consistent with both parasexual recombination and meiosis as both mechanisms result in the exchange of large genomic segments [36]. However, the stochastic chromosome loss that occurs during parasexual reproduction is expected to generate considerable aneuploidy which was not detected in the current dataset (S6 Fig).

Sexual reproduction in parasites like *T. cruzi* may be particularly important for maintaining a diverse assortment of immune evasion genes within individuals [38,39]. The strong selection pressure generated by the host immune system has resulted in the extensive and diverse immune evasion gene family in *T. cruzi*, the trans-sialidase genes [27,40,41]. Sexual reproduction can maintain immune evasion gene diversity in the face of homogenization events that commonly occur among paralogs within genomes by redistributing variation among trans-sialidase genes within a population through chromosome sorting or cross-over events [42–44]. While we did not specifically investigate trans-sialidase genes, it is possible that even in populations where inbreeding is common, low frequencies of sex can increase trans-sialidase diversity within individuals, thus increasing the probability of survival within a host. Future studies



could assess whether recombination hotspots are conserved across *T. cruzi* populations and if recombination events impact the diversity of trans-sialidase genes.

Since the *T. cruzi* isolates used in this analysis were maintained in culture for variable periods of time, meiosis could have occurred during the isolation and culturing process. However, 65% (310) of all crossing over events were present in multiple samples. Thus, a substantial proportion of the crossing over detected here must have occurred prior to isolation.

Although the ubiquitous observations of clonal population structures have resulted in the hypothesis that the frequency of sexual reproduction is insufficient to disrupt the clonal population structure in natural *T. cruzi* populations [25,45], the data presented here suggest that sexual recombination is not uncommon in natural populations. Sexual reproduction occurred repeatedly in the recent evolutionary history of this population and has occurred within sub-populations that have only recently established on city blocks (Fig 3C). The number of meiotic recombination events reported here is conservative, as events occurring between homozygous regions cannot be detected and our criteria for identifying recombined regions was strict. Nevertheless, the extant linkage disequilibrium across large linkage blocks suggests that asexual reproduction has been the most common mode of reproduction in this population. This mixed life-history strategy has important medical and evolutionary implications. Sexual reproduction may allow for rapid diversification of antigens which may contribute to the variability in serological diagnostics [46–48] and has the potential to generate genetic and phenotypic diversity [49,50] in pathogenicity [51,52], host and vector propensity [53,54], and vulnerability to drugs [55,56] if outcrossing is common in populations with greater diversity. Primarily asexual reproduction coupled with periodic sexual inbreeding, on the other hand, will result in clonal population structures that maximize the diversity among lineages and minimize the possibility that virulence factors or drug resistance will introgress into other lineages. Primarily asexual reproduction coupled with periodic sexual outcrossing may be significant for the adaptive evolution in novel environments and may be particularly important for invading urbanizing environments where rapid adaptation may be at a premium.

## Materials and methods

### Ethics Statement

The Institutional Animal Care and Use Committee (IACUC) of Universidad Peruana Cayetano Heredia reviewed and approved the animal-handling protocol used for this study (identification number 59605). The IACUC of Universidad Peruana Cayetano Heredia is registered in the National Institutes of Health at the United States of America with PHS Approved Animal Welfare Assurance Number A5146-01 and adheres to the Animal Welfare Act of 1990 [57].

### Sample collection and study site

Genomic DNA from 123 *T. cruzi* isolates was sequenced to uncover evidence of sexual reproduction (Fig 2). Samples were isolated from *Triatoma infestans* bugs collected in Arequipa, Peru (N = 114) and an additional nine samples were isolated from the blood of guinea pigs (N = 7) and dogs (N = 2) in Arequipa (S1 Table). Uninfected *T. infestans* were allowed to feed on six of the nine blood samples in order to transfer the parasite to a natural vector. For all samples (lab- and naturally-infected vectors), *T. cruzi* was isolated by injecting feces from infected vectors into guinea pigs or mice and re-isolating *T. cruzi* from the blood of each infected mammal as previously described [58]. *T. cruzi* was isolated directly from the blood samples of three guinea pigs collected in Arequipa without passage through *T. infestans*.

## Sequencing

DNA from all laboratory cultures was extracted using Qiagen DNEasy DNA Purification Kit. 150bp single-end read libraries were prepared using TruSeq Nano kit and sequenced to an average depth of ~50X using Illumina's NextSeq500 (reads available at Sequence Read Archive (SRA) under the BioProject PRJNA517719). Six *T. cruzi* isolates were prepared in duplicate, and one in triplicate, to allow estimation of sequencing error. Low quality bases were trimmed from raw reads using trimmomatic-0.32 [59].

## Whole genome assembly

Phylogenetic analyses demonstrate that the reference genome most closely related to the Arequipan isolates is the discrete typing unit 1 (DTU1) genome TcJR clone 4 (S7 Fig). Arequipan genomes were assembled using the TcJR clone 4 genome, obtained from TriTrypDB (<http://tritrypdb.org/tritrypdb/>), as a reference using bowtie2 [60]. Only the 333 contigs longer than 10kb were used for the assembly to avoid spurious alignments to short contigs, for a total genome assembly that includes nearly 28Mbp (S2 Table). This assembly largely excluded the extensive repeat regions found throughout the *T. cruzi* genome. Duplicate reads were removed using Picard MarkDuplicates [61].

## SNP calling

Individual gVCF files containing SNP data for each sample were generated using GATK HaplotypeCaller [61,62] following GATK's Best Practices procedure [63,64]. A joint genotype file containing all polymorphic sites from all samples was created using GATK GenotypeGVCF. Indels were excluded. Polymorphic loci were hard-filtered by quality using GATK VariantFiltration, requiring Fisher strand bias (FS) <40, mapping quality (MQ) >30, and quality by depth (QD) >10. Only loci for which all samples achieved a minimum depth of 20 and a Genotype Quality score (GQ) greater than 40 were included. These filters maximized the number of polymorphic sites identified while ensuring that duplicate and triplicate sequences resulted in identical SNP datasets. The final consensus SNP panel included 9271 polymorphic sites. To assess ploidy in these isolates, the proportion of reads containing each nucleotide was calculated for all heterozygous sites in all 123 isolates.

## Detection of meiotic recombination events

Cross-over events during meiosis create a novel combination of polymorphic markers by combining large genetic regions of homologous chromosomes (Fig 1C). During meiosis, non-sister chromatids of homologous chromosomes can exchange genetic material resulting in gametes containing chromosomes that differ from either parental chromosome. The chromosomes that result from a meiotic cross-over event in gametes are comprised of two sections, one retaining the linkage association among markers found on one of the parental chromosomes joined with another chromosomal segment with the linkage associations found on the other parental chromosome. Fusion of gametes containing chromosomes that have experienced different cross-over events will result in regions of the chromosome that are heterozygous in some diploid offspring and homozygous in others. Thus, detection of large segments (>10kb) of chromosomes where all of the polymorphic markers (>20) that are heterozygous in some isolates are homozygous in other isolates within the same population indicates that a meiotic cross-over event created a novel combination of markers. Long chromosomal regions with many (>20) polymorphic markers that are heterozygous in some isolates but homozygous in others are unlikely to result from independent point mutations due to the clustering of all of

the mutational events in one region of a chromosome. Additionally, the point mutation hypothesis would necessitate that the same combination of mutations occurred independently in multiple isolates. Gene conversion could also produce the pattern expected from meiotic crossing over although gene conversion events are typically restricted to much shorter chromosomal segments (<10kb).

We examined the genomic sequence data for the distributions of polymorphic markers within and among individuals that are expected to result from meiosis. To this end, we first aligned all polymorphic sites by contig and position using VCF tools v0.1.13 [65] and PLINK [66]. The data were converted to a VCF file containing SNP data for all 123 samples and subsequently to PLINK format using VCF tools. PLINK's runs of homozygosity function was used to generate a list of regions that were entirely homozygous in some isolates and heterozygous in at least one isolate, a pattern suggestive of a meiotic recombination event. A conservative estimate of the number and size of recombinant regions was obtained by defining a recombinant region as a run of homozygosity containing at least 20 consecutive, homozygous polymorphic markers that span at least 10kb with a density of at least 1 SNP per 50kb. Runs of homozygosity were determined by sliding a window of 25 SNPs across the genome and determining if each window position contained a run of homozygosity as defined. This method produces a conservative estimate of the number and size of recombination events because the minimum size of 20 consecutive homozygous SNPs will not identify runs of homozygosity on short contigs nor regions with low densities of polymorphic markers. Additionally, regions that have acquired point mutations will not be identified as the algorithm requires that a run of homozygosity cannot contain any heterozygous sites. Further, recombination between identical homozygous regions were not detected because runs of homozygosity contained by all samples were ignored. Runs of homozygosity shared by multiple samples were considered identical by descent based on sequence similarity (S8 Fig). Potential recombinant regions were verified using Recombination Analysis Tool (RAT) [67] (S1 Text).

### Hardy-Weinberg equilibrium

The frequency of sexual reproduction affects the proportions of heterozygotes and homozygotes in a population. While regular sexual reproduction results in proportions of heterozygotes and homozygotes that conform to Hardy-Weinberg expectations, infrequent sexual reproduction or inbreeding will result in allele frequencies that deviate from Hardy-Weinberg expectations. To determine the frequency of sex and the randomness of mating in this *T. cruzi* population, the  $\chi^2$  test for Hardy-Weinberg equilibrium was performed for each polymorphic locus [68]. Briefly, for each locus, the difference between the expected genotype frequencies ( $E$ ) and observed genotype frequencies ( $O$ ) was squared and normalized by the expected genotype frequencies  $[(O-E)^2/E]$ . Loci are assumed to be in Hardy-Weinberg equilibrium if the sum of the  $\chi^2$  values across genotypes is less than 3.841.

### Supporting information

**S1 Fig. Evidence of recombination is evident throughout the *T. cruzi* genome.** Character state for each of 9271 polymorphic markers across all 123 *T. cruzi* isolates is shown. Each row represents a *T. cruzi* isolate which are grouped geographically by district and block within districts in the same order displayed in S1 Table. Columns represent polymorphic markers where heterozygous loci are represented in green, markers that are homozygous for one set of markers are represented in yellow, and markers that are homozygous for the other set of markers are represented in blue. All markers contain only two alleles. Black columns separate contigs.

For visualization, the genome is divided into four sections. Evidence of recombination is exhibited throughout the genome, with telltale runs of homozygosity present in 151 contigs. (TIF)

**S2 Fig. Length of meiotic recombination events.** 474 runs of homozygosity that range in size from 10kb to 468kb were identified. Short recombination events are more common than long runs, however because of the short length of contigs, long runs are less likely to be found intact than short runs. (TIF)

**S3 Fig. Number of recombination events per contig.** Recombination events occur throughout the genome. Larger contigs generally have evidence of more recombination because the recombination event criteria are more likely to be met. Further, if crossover events occur randomly, more events are expected to occur in larger regions. (TIF)

**S4 Fig. Isolates on city blocks commonly have the same recombination events indicating that some recombination events occurred after colonization of a city block.** Geographic clustering of the 56 isolates collected in Mariano Melgar was assessed using 100 independent iterations of ADMIXTURE [70] for each number of genetic clusters (K, ranging from 2 to 8) assuming linkage disequilibrium until the log-likelihood increased by less than  $\epsilon = 10^{-4}$  between iterations. The optimal alignment of the 100 iterations was calculated using CLUMPP [71]. **A.** Cross-validation scores are shown for each genetic cluster (K) averaged across 100 iterations. Standard error bars are shown for each value. K = 4 was determined to be the optimal number of genetic clusters. **B.** The proportion of each color within a bar represents the likelihood that a sample belongs to each of four genetic clusters, where each color represents a unique genetic cluster. The clustering of identical genotypes within blocks 1–4, within block 5, and within block 7 recapitulates the phylogenetic relationship among isolates (S5 Fig). A single isolate collected from Block 7 (yellow, red, and brown) is genetically distinct based on both the ADMIXTURE and phylogenetic analyses. (TIF)

**S5 Fig. *T. cruzi* isolates in Arequipa tend to cluster geographically.** The maximum likelihood phylogeny includes all 123 *T. cruzi* samples collected from Arequipa and was calculated using presence/absence for each of the 474 recombination events. Branches and collapsed clades are colored blue or yellow when they include at least one sample collected from a guinea pig or dog, respectively. While samples collected in the same year or from the same host do not form monophyletic clades, samples from the same city block tend to form monophyletic clades (e.g. Mariano Melgar blocks 1–4, 5, and 7), suggesting that geography is the most important predictor of phylogenetic relationship. Bootstrap support (1000 replicates) is labeled on all nodes. Terminal nodes with bootstrap support >0.70 were combined to highlight the effect of geography on phylogenetic relationship. The clade colored black contains samples from 10 districts and has poor bootstrap support (<0.70). Bold numbers within collapsed clades are the total number of samples included. The number of samples from guinea pigs (GP) and dogs are shown where applicable. Sample collection date ranges are also shown within collapsed clades. When samples representing multiple locations are contained within a single collapsed clade, the number of samples from each location is shown in parentheses. (TIF)

**S6 Fig. Frequency of heterozygous sites in which reads contain the reference nucleotide suggests the isolates in this study are primarily diploid.** The single mode of base calls around

50% at all heterozygous sites in all 123 genomes is indicative of diploidy. Bimodal distributions around 33% and 66% would indicate triploidy while trimodal distributions around 25%, 50%, and 75% would indicate tetraploidy. Sites where >80% of reads contained the same base were considered homozygous.

(TIF)

**S7 Fig. Arequipan isolates share a recent common ancestor with TCJRcl4.** 41 distinct Arequipan genomes were aligned to both the TCJRcl4 reference genome and the maxicircle sequence of the Silvio reference strain. A neighbor-joining phylogenetic tree shows that all 41 genomes from both assemblies share a more recent common ancestor with TCJRcl4 than Silvio, and that the isolates are definitively discrete typing unit I. All nodes have >99% bootstrap support.

(TIF)

**S8 Fig. Criteria used to distinguish runs of homozygosity.** Two runs of homozygosity were considered identical by descent (IBD) if they were 99% identical and had ten or fewer non-overlapping homozygous SNPs. That is, a run of homozygosity in one sample was considered identical to a run of homozygosity in another if it shared 99% sequence similarity and were of similar lengths, defined as ten or fewer non-overlapping homozygous SNPs on either end. Relaxing these assumptions did not qualitatively alter the results nor the conclusions.

(TIF)

**S1 Table. Collection locations and years for each *T. cruzi* sample.**

(PDF)

**S2 Table. Reference genome and genome assembly summary statistics.**

(TIF)

**S1 Text. Supplemental methods.**

(DOCX)

## Acknowledgments

The authors would like to thank Paul Sniegowski for helpful comments and suggestions regarding the manuscript. The authors would also like to acknowledge Stephanie Seifert and Jill Devine for their assistance in the laboratory. The authors gratefully acknowledge the members of the Chagas Disease Working Group in Arequipa including those at the Universidad Peruana Cayetano Heredia and the University of Pennsylvania Zoonotic Disease Research Lab in Arequipa, Peru, for their contributions, especially Carlos Condori and Luis Zamudio. The authors also thank Chagas Disease Working Group members Danitza Pamo, Jose Ylla, Jose Qusipe, Paul Picardo and Gabriela Bustamante for their contribution during the isolation and maintenance of the *T. cruzi* strains. In addition, the authors wish to acknowledge the advisory support provided by the following institutions: Ministerio de Salud del Perú (MINSa), the Dirección General de Salud de las Personas (DGSP), the Estrategia Sanitaria Nacional de Prevención y Control de Enfermedades Metaxénicas y Otras Transmitidas por Vectores (ESNP-CEMOTVS), the Dirección General de Salud Ambiental (DIGESA), the Gobierno Regional de Arequipa, the Gerencia Regional de Salud de Arequipa (GRSA), the PanAmerican Health Organization (PAHO/OPS) and the Canadian International Development Agency (CIDA).

## Author Contributions

**Conceptualization:** Alexander S. F. Berry, Dustin Brisson.



**Data curation:** Alexander S. F. Berry.

**Formal analysis:** Alexander S. F. Berry.

**Funding acquisition:** Michael Z. Levy, Dustin Brisson.

**Investigation:** Alexander S. F. Berry.

**Methodology:** Alexander S. F. Berry, Renzo Salazar-Sánchez, Ricardo Castillo-Neyra, Katty Borrini-Mayorí, Claudia Chipana-Ramos, Melina Vargas-Maquera, Jenny Ancca-Juarez, César Náquira-Velarde.

**Project administration:** Dustin Brisson.

**Resources:** Michael Z. Levy, Dustin Brisson.

**Supervision:** Michael Z. Levy, Dustin Brisson.

**Writing – original draft:** Alexander S. F. Berry.

**Writing – review & editing:** Alexander S. F. Berry, Renzo Salazar-Sánchez, Ricardo Castillo-Neyra, Michael Z. Levy, Dustin Brisson.

## References

1. Mihok S, Otieno LH, Darji N. Population genetics of *Trypanosoma brucei* and the epidemiology of human sleeping sickness in the Lambwe Valley, Kenya. *Parasitology*. Cambridge University Press; 1990; 100: 219–233.
2. Tibayrenc M, Kjellberg F, Ayala FJ. A clonal theory of parasitic protozoa: the population structures of *Entamoeba*, *Giardia*, *Leishmania*, *Naegleria*, *Plasmodium*, *Trichomonas*, and *Trypanosoma* and their medical and taxonomical consequences. *Proceedings of the National Academy of Sciences of the United States of America*. 1990; 87: 2414–2418. <https://doi.org/10.1073/pnas.87.7.2414> PMID: 2320563
3. Tibayrenc M, Kjellberg F, Arnaud J, Oury B, Brenière SF, Dardé ML, et al. Are eukaryotic microorganisms clonal or sexual? A population genetics vantage. *Proceedings of the National Academy of Sciences of the United States of America*. 1991; 88: 5129–5133. <https://doi.org/10.1073/pnas.88.12.5129> PMID: 1675793
4. Pujol C, Reynes J, Renaud F, Raymond M, Tibayrenc M, Ayala FJ, et al. The yeast *Candida albicans* has a clonal mode of reproduction in a population of infected human immunodeficiency virus-positive patients. *Proceedings of the National Academy of Sciences*. *National Acad Sciences*; 1993; 90: 9456–9459.
5. Ayala FJ. Is sex better? Parasites say “no.” *Proceedings of the National Academy of Sciences of the United States of America*. *The National Academy of Sciences*; 1998; 95: 3346–3348.
6. Tibayrenc M, Ayala FJ. How clonal are *Trypanosoma* and *Leishmania*? *Trends in Parasitology*. Elsevier; 2013; 29: 264–269.
7. Tibayrenc M, Ayala FJ. *Cryptosporidium*, *Giardia*, *Cryptococcus*, *Pneumocystis* genetic variability: cryptic biological species or clonal near-clades? *PLOS Pathogens*. *Public Library of Science*; 2014; 10: e1003908.
8. Tibayrenc M, Ayala FJ. Reproductive clonality of pathogens: a perspective on pathogenic viruses, bacteria, fungi, and parasitic protozoa. *Proceedings of the National Academy of Sciences*. *National Acad Sciences*; 2012;: 201212452.
9. Tibayrenc M, Ayala FJ. *Trypanosoma cruzi* and the model of predominant clonal evolution. Elsevier; 2017;: 475–495.
10. Babiker HA, Walliker D. Current views on the population structure of *Plasmodium falciparum*: implications for control. *Parasitology Today*. Elsevier; 1997; 13: 262–267.
11. Guttery DS, Holder AA, Tewari R. Sexual Development in *Plasmodium*: Lessons from Functional Analyses. Laura J K, editor. *PLOS Pathogens*. 2012; 8: e1002404–3. <https://doi.org/10.1371/journal.ppat.1002404> PMID: 22275863
12. Peacock L, Bailey M, Carrington M, Gibson W. Meiosis and Haploid Gametes in the Pathogen *Trypanosoma brucei*. *Current Biology*. *The Authors*; 2014; 24: 181–186.

13. Ramesh MA, Malik S-B, Logsdon JM Jr. A phylogenomic inventory of meiotic genes: evidence for sex in *Giardia* and an early eukaryotic origin of meiosis. *Current Biology*. Elsevier; 2005; 15: 185–191.
14. Gibson W. Liaisons dangereuses: sexual recombination among pathogenic trypanosomes. *Research in microbiology*. Elsevier; 2015; 166: 459–466.
15. Segatto M, Ribeiro LS, Costa DL, Costa CHN, de Oliveira MR, Carvalho SFG, et al. Genetic diversity of *Leishmania infantum* field populations from Brazil. *Memorias do Instituto Oswaldo Cruz*. SciELO Brasil; 2012; 107: 39–47.
16. Akopyants NS, Kimblin N, Secundino N, Patrick R, Peters N, Lawyer P, et al. Demonstration of genetic exchange during cyclical development of *Leishmania* in the sand fly vector. *Science*. American Association for the Advancement of Science; 2009; 324: 265–268.
17. Michod RE, Bernstein H, Nedelcu AM. Adaptive value of sex in microbial pathogens. *Infection, Genetics and Evolution*. Elsevier; 2008; 8: 267–285.
18. Oliveira RP, Broude NE, Macedo AM, Cantor CR, Smith CL, Pena SDJ. Probing the genetic population structure of *Trypanosoma cruzi* with polymorphic microsatellites. *Proceedings of the National Academy of Sciences of the United States of America*. The National Academy of Sciences; 1998; 95: 3776–3780.
19. Diosque P, Barnabé C, Padilla AM, Marco JD, Cardozo RM, Cimino RO, et al. Multilocus enzyme electrophoresis analysis of *Trypanosoma cruzi* isolates from a geographically restricted endemic area for Chagas' disease in Argentina. *International Journal for Parasitology*. 2003; 33: 997–1003.
20. Tibayrenc M. Kinetoplastid Biology and Disease. *Kinetoplastid Biol Dis*. BioMed Central; 2003; 2: 12–6.
21. Ramírez JD, Guhl F, Messenger LA, Lewis MD, Montilla M, Cucunuba Z, et al. Contemporary cryptic sexuality in *Trypanosoma cruzi*. *Molecular Ecology*. Wiley/Blackwell (10.1111); 2012; 21: 4216–4226.
22. de Paula Baptista R, D'Ávila DA, Segatto M, do Valle ÍF, Franco GR, Valadares HMS, et al. Evidence of substantial recombination among *Trypanosoma cruzi* II strains from Minas Gerais. *Infection, Genetics and Evolution*. Elsevier; 2014; 22: 183–191.
23. Gaunt MW, Yeo M, Frame IA, Stothard JR, Carrasco HJ, Taylor MC, et al. Mechanism of genetic exchange in American trypanosomes. *Nature*. Macmillan Magazines Ltd. SN; 2003; 421: 936–939.
24. Brisse S, Barnabé C, Banuls A-L, Sidibé I, Noël S, Tibayrenc M. A phylogenetic analysis of the *Trypanosoma cruzi* genome project CL Brener reference strain by multilocus enzyme electrophoresis and multiplex random amplified polymorphic DNA fingerprinting. *Molecular and biochemical parasitology*. Elsevier; 1998; 92: 253–263.
25. Brisse S, Henriksson J, Barnabé C, Douzery EJP, Berkvens D, Serrano M, et al. Evidence for genetic exchange and hybridization in *Trypanosoma cruzi* based on nucleotide sequences and molecular karyotype. *Infection, Genetics and Evolution*. 2003; 2: 173–183. PMID: [12797979](https://pubmed.ncbi.nlm.nih.gov/12797979/)
26. Westenberger SJ, Barnabé C, Campbell DA, Sturm NR. Two hybridization events define the population structure of *Trypanosoma cruzi*. *Genetics*. Genetics Soc America; 2005; 171: 527–543.
27. El-Sayed NM, Myler PJ, Bartholomeu DC, Nilsson D, Aggarwal G, Tran A-N, et al. The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science*. American Association for the Advancement of Science; 2005; 309: 409–415.
28. Machado CA, Ayala FJ. Nucleotide Sequences Provide Evidence of Genetic Exchange among Distantly Related Lineages of *Trypanosoma cruzi*. *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences; 2001; 98: 7396–7401.
29. Tibayrenc M, Ward P, Moya A, Ayala FJ. Natural populations of *Trypanosoma cruzi*, the agent of Chagas disease, have a complex multiclonal structure. *Proc Natl Acad Sci USA*. 1986; 83: 115. <https://doi.org/10.1073/pnas.83.1.115> PMID: [3510428](https://pubmed.ncbi.nlm.nih.gov/3510428/)
30. Tibayrenc M, Ayala FJ. Isozyme variability in *Trypanosoma cruzi*, the agent of Chagas' disease: genetic, taxonomical, and epidemiological significance. *Evolution*. Wiley Online Library; 1988; 42: 277–292.
31. Berry ASF, Salazar Sánchez R, Castillo Neyra R, Borrini-Mayorí K, Chipana-Ramos C, Vargas-Maquera M, et al. Immigration and establishment of urban *Trypanosoma cruzi* populations. *bioRxiv*. Cold Spring Harbor Laboratory; 2019;: 515049.
32. Wahlund S. Zusammensetzung von Populationen und Korrelationserscheinungen vom Standpunkt der Vererbungslehre aus betrachtet. *Hereditas*. Wiley Online Library; 1928; 11: 65–106.
33. Minning TA, Weatherly DB, Flibotte S, Tarleton RL. Widespread, focal copy number variations (CNV) and whole chromosome aneuploidies in *Trypanosoma cruzi* strains revealed by array comparative genomic hybridization. *BMC genomics*. BioMed Central; 2011; 12: 139.
34. López-Cancino SA, Tun-Ku E, la Cruz-Felix De HK, Ibarra-Cerdeña CN, Izeta-Alberdi A, Pech-May A, et al. Landscape ecology of *Trypanosoma cruzi* in the southern Yucatan Peninsula. *Acta Tropica*. 2015; 151: 58–72. <https://doi.org/10.1016/j.actatropica.2015.07.021> PMID: [26219998](https://pubmed.ncbi.nlm.nih.gov/26219998/)

35. Dumonteil E, Ramirez-Sierra M-J, Pérez-Carrillo S, Teh-Poot C, Herrera C, Gourbière S, et al. Detailed ecological associations of triatomines revealed by metabarcoding and next-generation sequencing: implications for triatomine behavior and *Trypanosoma cruzi* transmission cycles. *Scientific Reports*. Springer US; 2018; 8: 1–13.
36. Lewis MD, Llewellyn MS, Gaunt MW, Yeo M, Carrasco HJ, Miles MA. Flow cytometric analysis and microsatellite genotyping reveal extensive DNA content variation in *Trypanosoma cruzi* populations and expose contrasts between natural and experimental hybrids. *International Journal for Parasitology*. Elsevier; 2009; 39: 1305–1317.
37. Feretzaki M, Heitman J. Unisexual reproduction drives evolution of eukaryotic microbial pathogens. *PLOS Pathogens*. Public Library of Science; 2013; 9: e1003674.
38. Gjini E, Haydon DT, Barry JD, Cobbold CA. Critical interplay between parasite differentiation, host immunity, and antigenic variation in trypanosome infections. *The American Naturalist*. The University of Chicago Press; 2010; 176: 424–439.
39. McCulloch R, Morrison LJ, Hall JP. DNA recombination strategies during antigenic variation in the African trypanosome. *American Society of Microbiology*; 2015;: 409–435.
40. Schenkman S. *Trypanosoma cruzi* trans-sialidase as a multifunctional enzyme in Chagas' disease. *Cellular microbiology*. Wiley Online Library; 2012; 14: 1522–1530.
41. Weatherly DB, Peng D, Tarleton RL. Recombination-driven generation of the largest pathogen repository of antigen variants in the protozoan *Trypanosoma cruzi*. *BMC genomics*. BioMed Central; 2016; 17: 729.
42. Freitas-Junior LH, Bottius E, Pirrit LA, Deitsch KW, Scheidig C, Guinet F, et al. Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of *P. falciparum*. *Nature Publishing Group*. Nature Publishing Group; 2000; 407: 1018.
43. Barry AE, Leliwa-Sytek A, Tavul L, Imrie H, Migot-Nabias F, Brown SM, et al. Population genomics of the immune evasion (var) genes of *Plasmodium falciparum*. *PLOS Pathogens*. Public Library of Science; 2007; 3: e34.
44. Graves CJ, Ros VI, Stevenson B, Sniegowski PD, Brisson D. Natural selection promotes antigenic evolvability. *PLOS Pathogens*. Public Library of Science; 2013; 9: e1003766.
45. de Freitas JM, Augusto-Pinto L, Pimenta JR, Bastos-Rodrigues L, Gonçalves VF, Teixeira SMR, et al. Ancestral Genomes, Sex, and the Population Structure of *Trypanosoma cruzi*. *PLOS Pathogens*. 2006; 2: e24–10. <https://doi.org/10.1371/journal.ppat.0020024> PMID: 16609729
46. Longhi SA, Brandariz SB, Lafon SO, Niborski LL, Luquetti AO, Schijman AG, et al. Evaluation of in-house ELISA using *Trypanosoma cruzi* lysate and recombinant antigens for diagnosis of Chagas disease and discrimination of its clinical forms. *The American Journal of Tropical Medicine and Hygiene*. ASTMH; 2012; 87: 267–271.
47. Verani JR, Seitz A, Gilman RH, LaFuente C, Galdos-Cardenas G, Kawai V, et al. Geographic variation in the sensitivity of recombinant antigen-based rapid tests for chronic *Trypanosoma cruzi* infection. *The American Journal of Tropical Medicine and Hygiene*. 2009; 80: 410–415. PMID: 19270291
48. Abras A, Gállego M, Llovet T, Tebar S, Herrero M, Berenguer P, et al. Serological diagnosis of chronic Chagas disease: is it time for a change? Gilligan PH, editor. *J Clin Microbiol*. American Society for Microbiology; 2016; 54: 1566–1572.
49. Laurent JP, Barnabe C, Quesney V, Noel S, Tibayrenc M. Impact of clonal evolution on the biological diversity of *Trypanosoma cruzi*. *Parasitology*. 1997; 114: 213–218. PMID: 9075341
50. Zingales B, Miles MA, Campbell DA, Tibayrenc M, Macedo AM, Teixeira MM, et al. The revised *Trypanosoma cruzi* subspecific nomenclature: rationale, epidemiological relevance and research applications. *Infection, Genetics and Evolution*. Elsevier; 2012; 12: 240–253.
51. Andrade SG. Caracterização de cepas do *Trypanosoma cruzi* isoladas no Recôncavo Baiano. *Revista de Patologia Tropical*. 1974; 3: 65–121.
52. Andrade V, Barral-Netto M, Andrade SG. Patterns of resistance of inbred mice to *Trypanosoma cruzi* are determined by parasite strain. *Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e biologicas*. 1985; 18: 499–506. PMID: 3939191
53. de Lana M, da Silveira Pinto A, Barnabé C, Quesney V, Noël S, Tibayrenc M. *Trypanosoma cruzi*: compared vectorial transmissibility of three major clonal genotypes by *Triatoma infestans*. *Experimental Parasitology*. Elsevier; 1998; 90: 20–25.
54. Roellig DM, McMillan K, Ellis AE, Vandeberg JL, Champagne DE, Yabsley MJ. Experimental infection of two South American reservoirs with four distinct strains of *Trypanosoma cruzi*. *Parasitology*. Cambridge University Press; 2010; 137: 959–966.

55. Revollo S, Oury B, Laurent J-P, Barnabé C, Quesney V, Carrière V, et al. *Trypanosoma cruzi*: impact of clonal evolution of the parasite on its biological and medical properties. *Experimental Parasitology*. Elsevier; 1998; 89: 30–39.
56. Veloso VM, Carneiro CM, Toledo M, Lana M, Chiari E, Tafuri WL, et al. Variation in susceptibility to benznidazole in isolates derived from *Trypanosoma cruzi* parental strains. *Memorias do Instituto Oswaldo Cruz*. *SciELO Brasil*; 2001; 96: 1005–1011.
57. Castillo Neyra R, Barbu CM, Salazar R, Borrini K, Naquira C, Levy MZ. Host-seeking behavior and dispersal of *Triatoma infestans*, a vector of Chagas disease, under semi-field conditions. *PLOS Neglected Tropical Diseases*. *Public Library of Science*; 2015; 9: e3433.
58. Castillo Neyra R, Borrini-Mayori K, Salazar Sánchez R, Ancca Suarez J, Xie S, Náquira Velarde C, et al. Heterogeneous infectiousness in guinea pigs experimentally infected with *Trypanosoma cruzi*. *Parasitology International*. 2016; 65: 50–54. <https://doi.org/10.1016/j.parint.2015.09.009> PMID: 26432777
59. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. *Oxford University Press*; 2014; 30: 2114–2120.
60. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012; 9: 357–359. <https://doi.org/10.1038/nmeth.1923> PMID: 22388286
61. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*. 2010; 20: 1297–1303. <https://doi.org/10.1101/gr.107524.110> PMID: 20644199
62. Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Van der Auwera GA, et al. Scaling accurate genetic variant discovery to tens of thousands of samples. 2017;: 1–22.
63. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*. *Nature Publishing Group, a division of Macmillan Publishers Limited*. All Rights Reserved. SN; 2011; 43: 491–498.
64. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, del Angel G, Levy-Moonshine A, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics*. *Hoboken, NJ, USA: John Wiley & Sons, Inc*; 2013; 43: 11.10.1–33.
65. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. *Bioinformatics*. 2011; 27: 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330> PMID: 21653522
66. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*. 2007; 81: 559–575. <https://doi.org/10.1086/519795> PMID: 17701901
67. Etherington GJ, Dicks J, Roberts IN. Recombination Analysis Tool (RAT): a program for the high-throughput detection of recombination. *Bioinformatics*. *Oxford University Press*; 2004; 21: 278–281.
68. Nielsen R, Slatkin M. An introduction to population genetics: theory and applications. *Sinauer Associates Sunderland*; 2013.
69. QGIS Development Team. QGIS Geographic Information System [Internet]. 2nd ed. *Open Source Geospatial Foundation Project*. <http://qgis.osgeo.org>
70. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*. 2009; 19: 1655–1664. <https://doi.org/10.1101/gr.094052.109> PMID: 19648217
71. Jakobsson M, Rosenberg NA. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*. 2007; 23: 1801–1806. <https://doi.org/10.1093/bioinformatics/btm233> PMID: 17485429