2018

Dispersal Patterns And Reproductive Strategies Of Trypanosoma Cruzi In An Urban Environment

Alexander Scott Berry
University of Pennsylvania, berrya@sas.upenn.edu

Dispersal Patterns And Reproductive Strategies Of Trypanosoma Cruzi In An Urban Environment

Abstract

Interactions with novel environments result in both population extinctions and explosions, depending on life history strategies, ecological circumstances and historical contingencies. Understanding how populations interact with novel environments is more important now than ever because of the rapid increase in anthropogenic environmental alterations. While many native populations perish during urbanization, other populations successfully invade and thrive in these novel environments including many pests and pathogens that are detrimental to human health and economy. Pathogens contain a diversity of life history strategies and are thus good models for understanding how life history strategies interact with novel environments to promote or hinder the dynamic processes of invasion. In order to determine how life history strategies affect the evolution of a population during invasion, we studied a population of Trypanosoma cruzi – a protozoan parasite and causative agent of Chagas disease in humans – in the city of Arequipa in southern Peru. To estimate population structure and evolutionary history, we sequenced and assembled 133 T. cruzi genomes collected throughout Arequipa (N=123) and South America, and performed phylogenetic and population genetic analyses. We found that (1) the extant T. cruzi population in Arequipa was founded by a single introduction; (2) T. cruzi readily disperses between houses in a city block, but rarely disperses between blocks; (3) We resolve an apparent contradiction between the perceived clonal population structure of T. cruzi and its capacity for sexual reproduction by showing that this population regularly underwent meiosis and fertilization as the it expanded throughout the city, but that exclusive inbreeding resulted in a clonal population structure. As urbanization increases worldwide, it is important to understand how life history strategies affect the invasion of urban environments.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Biology

First Advisor
Dustin Brisson

Keywords
Chagas disease, Invasion biology, Population genetics, Trypanosoma cruzi, Urban ecology

Subject Categories
Biology | Parasitology

DISPERSAL PATTERNS AND REPRODUCTIVE STRATEGIES OF *TRYPANOSOMA CRUZI* IN AN

URBAN ENVIRONMENT

Alexander Scott Berry

A DISSERTATION

in

Biology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2018

Supervisor of Dissertation

_____

Dustin Brisson

Associate Professor of Biology

Graduate Group Chairperson

_____

Michael Lampson, Associate Professor of Biology

Dissertation Committee:

Paul S. Schmidt (Chair), Professor of Biology; Michael Z. Levy, Associate Professor of Epidemiology;

Paul Sniegowski, Professor of Biology; David S. Roos, E. Otis Kendall Professor of Biology

ACKNOWLEDGMENTS

I am grateful for the financial support of two NIH training grants – Computational Biology and Parasitology – in addition to the support of the Burroughs Wellcome Fund and the National Institute of Allergy and Infectious Diseases, which allowed me to focus on the research I was most interested in conducting.

I would like to thank my advisor, Dustin Brisson, for his immense support throughout this arduous process. I always enjoyed standing in front of a white board while bouncing science ideas back and forth. I'm certain that he was tired of my terrible ideas, but I think it turned into something that resembles 'good science'. Dustin has shaped me into a significantly better scientist, writer, and coffee connoisseur. The lessons I have learned will follow me forever.

I would also like to thank Mike Levy who served as a committee member, mentor, and Arequipa expert. Mike had a huge impact on what my dissertation became. He always gave me the correct advice, despite my occasional insistence to the contrary, which helped to mold my thesis project into the best version of itself. I vow to never again make hypotheses about the effects of rivers until I confirm that those rivers actually contain water.

I want to thank my other committee members: Paul Schmidt, Paul Sniegowski, and David Roos. The meaningful conversations I had with all of them throughout the years have made their marks on the following pages and have contributed to my development as a scientist. Thank you for your constant scientific guidance.

I wholeheartedly appreciate the tireless efforts of the group in Arequipa who collected all of the samples used in these projects. Without their contributions none of this would

Most of all, I wouldn't have survived without the eternal support of my beautiful wife, Hannah (and our beautiful dog, Elsa). One finishing medical school while the other finishes graduate school makes for a lot of stressful days and nights, but we persevered together. I doubt that we'll look back on this time and think 'that wasn't so hard', but it would have been impossible without you.

<u>ABSTRACT</u>

DISPERSAL PATTERNS AND REPRODUCTIVE STRATEGIES OF *TRYPANOSOMA*

CRUZI IN AN URBAN ENVIRONMENT

Alexander Scott Berry

Dustin Brisson

Interactions with novel environments result in both population extinctions and explosions, depending on life history strategies, ecological circumstances and historical contingencies. Understanding how populations interact with novel environments is more important now than ever because of the rapid increase in anthropogenic environmental alterations. While many native populations perish during urbanization, other populations successfully invade and thrive in these novel environments including many pests and pathogens that are detrimental to human health and economy. Pathogens contain a diversity of life history strategies and are thus good models for understanding how life history strategies interact with novel environments to promote or hinder the dynamic processes of invasion. In order to determine how life history strategies affect the evolution of a population during invasion, we studied a population of *Trypanosoma cruzi* – a protozoan parasite and causative agent of Chagas disease in humans – in the city of Arequipa in southern Peru. To estimate population structure and evolutionary history, we sequenced and assembled 133 *T. cruzi* genomes collected throughout Arequipa (N=123) and South America, and performed phylogenetic and population genetic analyses. We found that *(1)* the extant *T. cruzi* population in Arequipa was founded by a single introduction; *(2) T. cruzi* readily disperses between houses in a city block, but rarely disperses between blocks; *(3)* We resolve an apparent contradiction between the

perceived clonal population structure of *T. cruzi* and its capacity for sexual reproduction by showing that this population regularly underwent meiosis and fertilization as the it expanded throughout the city, but that exclusive inbreeding resulted in a clonal population structure. As urbanization increases worldwide, it is important to understand how life history strategies affect the invasion of urban environments.

TABLE OF CONTENTS

LIST OF TABLES

Table 1.1. Population Genetic Statistics.

Table 1.2. Distance matrix showing average pairwise SNP distance between samples (bottom triangle) and Euclidean distance between sample collection locations (top triangle).

Table 2.1. Population Genetic Statistics

Table 2.2. AMOVA across Arequipan *T. cruzi* isolates shows no significant population structure between districts but significant structure within Mariano Melgar district.

Supplemental Table 2.1. Using only presence/absence for each of 474 recombination events in place of genotype data, AMOVA performed across Arequipan *T. cruzi* isolates shows significant structure within blocks.

Supplemental Table 2.2. Sample collection locations and years.

Supplemental Table 3.1. Collection locations and years for each *T. cruzi* sample.

LIST OF FIGURES

<u>INTRODUCTION</u>

Interactions with novel environments results in both population extinctions and explosions, depending on ecological circumstances and historical contingencies. While there is no one set of rules that governs the dynamics of populations that enter a novel environment, several common themes reappear. Among these: ecological release (whereby a population rapidly expands in a new environment, unrestrained by previous growth restrictions) (Keane & Crawley 2002; Colautti *et al.* 2004; Bolnick *et al.* 2010), extinction due to inhospitable abiotic (e.g. altitude and temperature) or biotic (e.g. new competitors, predators, or pathogens) factors (Tilman *et al.* 2004; Hilker *et al.* 2005; Herborg *et al.* 2007), and establishment in areas that resemble a prior niche (Williamson 1996; Sax & Brown 2000; Mata *et al.* 2013)

The urban environment is becoming one of the most common environments in the world (Seto *et al.* 2012). Many populations are forced to interact with these novel habitats (Seto *et al.* 2011); some thrive while many perish (Hahs *et al.* 2009; Beninde *et al.* 2015). The ascendance of urban ecology will lead to an understanding of how populations interact with novel urban environments. One set of species that are important in these environments are pathogens of humans or human-domesticated species (Alirol *et al.* 2011; Hassell *et al.* 2017). The structure of human-constructed environments is unique and can provide both opportunities and hindrances for the migration, establishment, and dispersal of pathogens. For example, pathogens can rapidly disperse through cities and among cities due to the densities of hosts and the ease of human movement within and among these environments (Patz *et al.* 2004; Bradley & Altizer 2007; Rosenthal *et al.* 2008; Nelson *et al.* 2011) . Examples of such dispersal includes the rapid spread of Zika

virus among humans in Rio de Janerio, Brazil, followed by its spread across the Americas likely facilitated by human travel via aircraft (Faria *et al.* 2016; Metsky *et al.* 2017).

Pathogens contain a diversity of life history strategies and are thus good models for understanding how life history strategies interact with novel environments to promote or hinder the dynamic processes of invasion (Barrett *et al.* 2008). Some hypotheses predicting how life history strategies augment the success of migration, colonization, establishment, and dispersal in novel environments include: (1) Invasion success may depend on the timing of urbanization if recently-built habitats are more hospitable or if transmission is more probable. (2) An increase in density of susceptible hosts within a local habitat may reduce dispersal by disincentivizing migration. (3) An increase in density of pathogens or vectors may increase dispersal if the carrying capacity of a local habitat is reached. In the following three chapters, analyses of how life history strategies affect patterns of invasion in urban environments will focus primarily on a population of the protozoan parasite *Trypanosoma cruzi*—the causative agent of Chagas disease-- in the city of Arequipa, Peru.

Arequipa, Peru is an ideal model for which to study patterns of invasion in urban environments because it is a recently-urbanized, well-characterized environment that underwent a recently-controlled epidemic of *T. cruzi* (Bowman *et al.* 2008; Bayer *et al.* 2009; Hunter *et al.* 2012; Levy *et al.* 2014). The Chagas disease system in Arequipa provides a good model for studying the processes of invasion, establishment, and dispersal because the low rate of *T. cruzi* transmission between insect vector and mammalian host causes invasion processes to occur on timescales that can be measured using realistic sampling strategies. The following three chapters will investigate how life

history strategies affect the dynamic process of invasion. Chapter One will describe how life history strategies interact with ecological and historical factors when migrating to and colonizing a new environment. Chapter Two will explore how life history strategies interact with local habitats to facilitate or hinder dispersal between and among recently-colonized urban habitats. Chapter Three will describe how reproductive strategy affects the generation of diversity as populations expand.

Literature Cited

Alirol, E., Getaz, L., Stoll, B., Chappuis, F., & Loutan, L. (2011). Urbanisation and infectious diseases in a globalised world. *The Lancet Infectious Diseases*, *11*(2), 131–141.

Barrett, L. G., Thrall, P. H., Burdon, J. J., & Linde, C. C. (2008). Life history determines genetic structure and evolutionary potential of host–parasite interactions. *Trends in Ecology & Evolution*, *23*(12), 678–685.

Bayer, A. M., Hunter, G. C., Gilman, R. H., Cornejo del Carpio, J. G., Naquira, C., Bern, C., & Levy, M. Z. (2009). Chagas Disease, Migration and Community Settlement Patterns in Arequipa, Peru. *PLOS Neglected Tropical Diseases*, *3*(12), e567. http://doi.org/10.1371/journal.pntd.0000567

Beninde, J., Veith, M., & Hochkirch, A. (2015). Biodiversity in cities needs space: a meta-analysis of factors determining intra-urban biodiversity variation. *Ecology Letters*, *18*(6), 581–592. http://doi.org/10.1111/ele.12427

Bolnick, D. I., Ingram, T., Stutz, W. E., Snowberg, L. K., Lau, O. L., & Paull, J. S. (2010). Ecological release from interspecific competition leads to decoupled changes in population and individual niche width. *Proc. R. Soc. B*, *277*(1689), 1789–1797.

Bowman, N. M., Kawai, V., Levy, M. Z., Cornejo del Carpio, J. G., Cabrera, L., Delgado, F., et al. (2008). Chagas Disease Transmission in Periurban Communities of Arequipa, Peru. *Clinical Infectious Diseases*, *46*(12), 1822–1828. http://doi.org/10.1086/588299

Bradley, C. A., & Altizer, S. (2007). Urbanization and the ecology of wildlife diseases. *Trends in Ecology & Evolution*, *22*(2), 95–102. http://doi.org/10.1016/j.tree.2006.11.001

Colautti, R. I., Ricciardi, A., Grigorovich, I. A., & MacIsaac, H. J. (2004). Is invasion success explained by the enemy release hypothesis? *Ecology Letters*, *7*(8), 721–733.

Faria, N. R., Azevedo, R. D. S. D. S., Kraemer, M. U. G., Souza, R., Cunha, M. S., Hill, S. C., et al. (2016). Zika virus in the Americas: Early epidemiological and genetic findings. *Science*, *352*(6283), 345–349. http://doi.org/10.1126/science.aaf5036

Hahs, A. K., McDonnell, M. J., McCarthy, M. A., Vesk, P. A., Corlett, R. T., Norton, B. A., et al. (2009). A global synthesis of plant extinction rates in urban areas. *Ecology Letters*, *12*(11), 1165–1173.

Hassell, J. M., Begon, M., Ward, M. J., & Fèvre, E. M. (2017). Urbanization and disease emergence: Dynamics at the wildlife–livestock–human interface. *Trends in Ecology & Evolution*, *32*(1), 55–67.

Hilker, F. M., Lewis, M. A., Seno, H., Langlais, M., & Malchow, H. (2005). Pathogens can slow down or reverse invasion fronts of their hosts. *Biological Invasions*, *7*(5), 817–832.

Hunter, G. C., Borrini-Mayorí, K., Ancca Juárez, J., Castillo Neyra, R., Verastegui, M. R., Malaga Chavez, F. S., et al. (2012). A Field Trial of Alternative Targeted Screening Strategies for Chagas Disease in Arequipa, Peru. *PLOS Neglected Tropical Diseases*, *6*(1), e1468. http://doi.org/10.1371/journal.pntd.0001468

Keane, R. M., & Crawley, M. J. (2002). Exotic plant invasions and the enemy release hypothesis. *Trends in Ecology & Evolution*, *17*(4), 164–170.

Levy, M. Z., Barbu, C. M., Castillo Neyra, R., Quispe-Machaca, V. R., Ancca Juárez, J., Escalante-Mejia, P., et al. (2014). Urbanization, land tenure security and vector-borne Chagas disease. *Proceedings. Biological Sciences*, *281*(1789), 20141003–20141003. http://doi.org/10.1098/rspb.2014.1003

Mata, T. M., Haddad, N. M., & Holyoak, M. (2013). How invader traits interact with resident communities and resource availability to determine invasion success. *Oikos*, *122*(1), 149–160.

Metsky, H. C., Matranga, C. B., Wohl, S., Schaffner, S. F., Freije, C. A., Winnicki, S. M., et al. (2017). Zika virus evolution and spread in the Americas. *Nature Publishing Group*, *546*(7658), 411–415. http://doi.org/10.1038/nature22402

Nelson, M. I., Lemey, P., Tan, Y., Vincent, A., Lam, T. T.-Y., Detmer, S., et al. (2011). Spatial Dynamics of Human-Origin H1 Influenza A Virus in North American Swine. *PLOS Pathogens*, *7*(6), e1002077. http://doi.org/10.1371/journal.ppat.1002077

Patz, J. A., Daszak, P., Tabor, G. M., Aguirre, A. A., Pearl, M., Epstein, J., et al. (2004). Unhealthy landscapes: policy recommendations on land use change and infectious disease emergence. *Environmental Health Perspectives*, *112*(10), 1092.

Rosenthal, B. M., LaRosa, G., Zarlenga, D., Dunams, D., Chunyu, Y., Mingyuan, L., & Pozio, E. (2008). Human dispersal of Trichinella spiralis in domesticated pigs. *Infection, Genetics and Evolution*, *8*(6), 799–805.

Sax, D. F., & Brown, J. H. (2000). The paradox of invasion. *Global Ecology and Biogeography*, *9*(5), 363–371.

Seto, K. C., Fragkias, M., Güneralp, B., & Reilly, M. K. (2011). A meta-analysis of global urban land expansion. *PLoS ONE*, *6*(8), e23777.

Seto, K. C., Güneralp, B., & Hutyra, L. R. (2012). Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proceedings of the National Academy of Sciences*, *109*(40), 16083–16088.

Tilman, D. (2004). Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion, and community assembly. *Proceedings of the National Academy of Sciences*, *101*(30), 10854–10861.

Williamson, M. (1996). Biological invasions, *15*.

CHAPTER 1: Invasion of urban environments by the Chagas disease agent, *Trypanosoma cruzi*

Abstract

Human altered habitats cause indelible changes to biological communities. While many native species suffer in the novel habitats, migrant species can immigrate and establish robust populations. The dynamic ecological interactions impacting immigration and initial establishment of foreign species in these human-altered environments are pivotal to both the conservation of native species and for the control of pest species with negative effects on human health and economy. The repeated emergence of disease systems in urban areas worldwide highlights the importance of understanding these dynamic processes in the current and future distribution and abundance of disease-causing microbes in urban environments. In this study, we examine the pattern of invasion of *Trypanosoma cruzi*—the causative agent of human Chagas disease— in Arequipa, Peru. Phylogenetic analyses of 136 *T. cruzi* isolates from Arequipa and other South American locations suggest that only one *T. cruzi* immigrant established a population in vectors in Arequipa as all extant isolates form a recent monophyletic group within the broader South American phylogeny. We discuss several hypotheses that may explain the limited establishment of *T. cruzi* despite multiple introductions of the parasite.

Introduction

Anthropogenic habitat alterations have transformed biological communities worldwide (McKinney 2002; Foley *et al.* 2005; Seto *et al.* 2012). The distribution and abundance of many species facing dynamic environments depends upon their interactions with novel environmental features during immigration and the establishment of a growing population (Theoharides & Dukes 2007). Although many species fail to establish thriving populations in human altered habitats, many others are well-suited to prosper in these novel environments. For example, populations of many plant (Kowarik 1995; Hahs *et al.* 2009; Aronson *et al.* 2014), insect (Denys & Schmidt 1998, McIntyre 2000; Deguines *et al.* 2016; New 2018), mammal (Mackin-Rogalska *et al.* 1988; Pekin & Pijanowski 2012), and bird (Blair 2001; Strohbach *et al.* 2009) species are severely diminished or suffer local extinctions in urban environments (McKinney 2002) while several microbial species benefit from the abundance of humans and human-associated hosts or vectors (Alirol *et al.* 2011). Although conservation efforts have focused on the impacts of local and global extinctions, establishment or growth of disease-causing microbe populations can negatively impact human health and economy as well as the populations of native flora and fauna (Crowl *et al.* 2008). Mitigating the rate or impacts of invasion of disease-causing microbes is best accomplished through empirical investigations of the dynamic processes of immigration and establishment as these processes occur. The regularity at which disease systems are emerging in many urban and urbanizing areas underscores the importance of understanding how disease-causing microbes migrate to, and establish in, urban environments (Alirol *et al.* 2011). In this

study, we examine the patterns of invasion of *Trypanosoma cruzi*—the causative agent of Chagas disease in humans— into the city of Arequipa, Peru.

Invasion of a new environment by a pathogen occurs in three stages: (1) immigration, or the transport of an individual to the new environment; (2) establishment of a population via reproduction and population growth; and (3) local dispersal or spread (Theoharides & Dukes 2007). Many studies focusing on the outbreaks of disease systems in urban areas have resulted in a wealth of knowledge concerning factors affecting population growth (Deplazes *et al.* 2004; Bradley & Altizer 2007) and considerable progress in understanding local dispersal (Harrington *et al.* 2005). For example, prior studies established that human-created containers increase the abundance of standing water that provide breeding habitats for the mosquitos that spread dengue virus (Adalja *et al.* 2012). Relatively few studies, by contrast, have investigated the early stages of invasion including the immigration and colonization processes in urban environments due to the practical difficulties of collecting the necessary data before a pest species is established. With the worldwide increase in urban habitats, it is necessary to understand the colonization process including factors affecting immigration rates and probabilities of establishment in new urban environments. Identifying species characteristics and environmental features that promote or impede immigration or establishment is necessary to estimate the distribution and abundance of populations in an increasingly urban world.

The Chagas disease system in Arequipa, Peru, provides an ideal system in which to study the early invasion processes in urban environments. Recent urbanization in Arequipa has permitted the colonization and population growth of *T. cruzi* (Levy *et al.* 2014). *T. cruzi* is propagated among domestic animals and humans in the city by its

primary insect vector, *Triatoma infestans* (Levy *et al.* 2006, 2007, 2011; Bowman *et al.* 2008; Hunter *et al.* 2012; Foley *et al.* 2013). Here, we performed phylogenetic analyses of maxicircle DNA, a non-recombining circular element analogous to mitochondrial DNA, to estimate the number of independent *T. cruzi* colonization events in Arequipa and to estimate the timing of each colonization event. We assessed whether extant *T. cruzi* in Arequipa form a single monophyletic clade, indicative of the establishment of a single immigration event, or multiple diverse clades, indicative of multiple independent immigration and establishment events. The results suggest that all analyzed *T. cruzi* isolates are descendants of a single common ancestor that established a population in Arequipa in the very recent past.

Methods

Sample collection

DNA from 133 *T. cruzi* isolates collected over 7 years (2008 to 2015) were analyzed to determine the phylogenetic relationships among samples (Fig 1 and 2). The majority of samples were isolated from *T. infestans* bugs collected from houses throughout Arequipa (N=114). Three of these samples were obtained using xenodiagnosis as described in Chiari & Galvão (1997). An additional ten samples from Arequipa were isolated from the blood of guinea pigs (N=7), dogs (N=2), and a human (N=1). Six isolates were derived from *Panstrongylus lignarius* (N=5) - known in Peru as *P. herreri* (Cáceres *et al.* 2002) - and one guinea pig (N=1) collected in the small towns of La Esperanza, Campo Florido, and Naranjal in northern Peru (Alroy *et al.* 2015). Cultures of three previously established strains isolated from humans in Bolivia (Bol-SH001 and Bol-DH29) and São

Paulo, Brazil (TC-y) were provided by the Infectious Diseases Research Laboratory at Universidad Peruana Cayetano Heredia (Fig 2).

T. cruzi was isolated from vectors and vertebrates using an adaptation of an artificial feeding system that was originally described in Harington (1960). Briefly, each blood sample was collected with citrate-phosphate-dextrose, transferred into a small plastic jar, and covered with a latex membrane fitted tightly with a rubber band. The jars were placed into an incubator and gradually heated to 35° C. Once the temperature was reached, the jars were inverted to allow uninfected *T. infestans* to feed through the membrane for 15 minutes. *T. cruzi* from all 114 naturally-infected *T. infestans*, five naturally-infected *P. lignarius*, and the eight laboratory-infected *T. infestans* were passaged through guinea pigs or mice in order to avoid isolating other microbes present in the vector, as described in Castillo-Neyra *et al.* (2016). Feces from infected vectors were injected into guinea pigs or mice and *T. cruzi* was isolated from the blood of each experimentally-infected mammal. *T. cruzi* were directly isolated in LIT culture media from the blood samples of three naturally-infected guinea pigs collected in Arequipa without passage through *T. infestans*.

Reference sequences of three *T. cruzi* isolates obtained from NCBI database were used in subsequent analyses: Silvio, isolated from a human in Para State, Brazil; Esmeraldo, isolated from a human in Bahia State, Brazil; and CL Brener, isolated from a human in Rio Grande, Brazil (Geer *et al.* 2009) (Fig 2).

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. 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 (Bolger *et al.* 2014).

Sequence assembly

Bowtie2 (Langmead & Salzberg, 2012) was used to assemble maxicircle sequences to the

most closely related reference sequence, Silvio (gi|225217165|gb|FJ203996.1), obtained

from NCBI (Geer *et al.* 2009). Duplicate reads were removed from the assembly using

Picard's MarkDuplicates functionality (McKenna *et al.* 2010). The assembly had an

average depth of >600X across all maxicircles. Maxicircle consensus sequences were

determined using VarScan (Koboldt *et al.* 2009), ensuring highly-confident base calls by

requiring a 60% match to call each SNP.

Maxicircle alignment

All assembled maxicircle sequences and the reference were aligned to the Silvio partial

maxicircle sequence (gi|225217165|gb|FJ203996.1|), Esmeraldo strain complete

maxicircle (gi|85718082|gb|DQ343646.1), and the CL Brener complete maxicircle

(gi|85718081|gb|DQ343645.1) downloaded from the NCBI database. The sequences were

aligned using MUSCLE as implemented in MEGA7 (Kumar *et al.* 2016). The ends were

trimmed so that all sequences started and ended on the same nucleotide, resulting in a

final alignment of 15357bp.

Phylogenetic analyses

Phylogenetic analyses of the 15357bp maxicircle sequence from all samples and

reference strains were performed using BEAST 1.8.4 (Drummond *et al.* 2012).

Phylogenetic analyses assumed a model of sequence evolution in which the rates of

A→T, C→G, and G→T are equal (123343) with γ-distributed rate heterogeneity. An Extended Bayesian Skyline tree prior (Heled & Drummond, 2008) with constant evolutionary rates across lineages (strict clock) was chosen based on BEAST Model Test implemented in BEAST2 (Bouckaert *et al.* 2014). Starting with a UPGMA tree and running one Markov chain Monte-Carlo chain for each of five independent runs of 20 million iterations sampling every 2000 iterations ensured sufficient mixing after a 10% burn-in (ESS values >200 in Tracer v1.6.0) (Rambaut *et al.* 2018). Tree files were combined using LogCombiner1.8.4, excluding a 10% burn-in for each. A Maximum Clade Credibility tree was generated from the combined tree file using TreeAnnotator 1.8.4 and FigTreev1.4.2 was used to visualize tree files (available at http://beast.bio.ed.ac.uk). Phylogenetic analyses were performed using the BEAGLE library to increase computation speed (Suchard & Rambaut, 2009; Ayres *et al.* 2011).

Statistical Analyses

The metrics of population genetic variation, π and θ, were calculated using MEGA7 (Kumar *et al.* 2016). Assuming the sample from 123 infected *T. infestans* is representative of the *T. cruzi* population in vectors in Arequipa, the probability that a distinct lineage representing an independent establishment event could be co-circulating, but not detected by chance, can be calculated using a binomial distribution. The minor lineage must constitute less than 2.41% of the total population in order for there to be a statistically significant chance that a distinct lineage was not detected in any of 123 observations.

Results

The maxicircle sequence is a useful tool for population genetics and phylogenetics analyses because it is conserved between diverse *T. cruzi* lineages and it likely non-recombining (Gaunt *et al* 2003). While there is substantial maxicircle sequence diversity among samples across South America, there is little diversity among samples from Arequipa. Over 13% (2055/15367bp) of the maxicircle sites were polymorphic among all 136 samples while only 16 sites were polymorphic (0.1%) among the 123 samples collected within Arequipa (Table 1). Similarly, estimates of diversity among the samples from Arequipa derived from population genetic statistics were substantially lower than the total diversity across all samples ($\pi=6.8*10^{-5}$ vs $8.18*10^{-3}$; $\theta=1.93*10^{-4}$ vs $2.44*10^{-2}$; average pairwise distance 1.04 vs 126). In contrast to the limited genetic diversity within Arequipa, other locales from which multiple isolates were sampled contain considerably greater diversity despite considerably fewer samples (Fig 3).

The monophyletic group containing all 123 samples derived from *T. infestans* and domestic animals collected in Arequipa coalesce in the very recent past, despite collection sites extending throughout Arequipa and surrounding towns (Fig 3). One sample derived from an infected human in Arequipa (Fig 3C) belongs to a distant lineage from the other samples in Arequipa and is closely related to samples collected in Rio Grande, Brazil, suggesting that *T. cruzi* can immigrate to Arequipa but this lineage has yet to establish in the vector population (Fig 3). The population size of such a second, unsampled *T. cruzi* lineage – if present in *T. infestans* – must be at least 42 times smaller than the dominant population in Arequipa to have remained undetected by chance (p<0.05).

In contrast to the monophyletic ancestry found in Arequipa, genetic diversity was apparent in the samples from other regions investigated here, despite limited sampling. Genetic diversity metrics among samples collected La Esperanza, Peru, a town of 57 houses, (π=4.18*10^{-2}; θ=4.18*10^{-2}) are much larger than those in Arequipa despite limited sampling (n=3)(Alroy *et al.* 2015). There is no statistical correlation between genetic relatedness and geographic distance among South American samples, despite sampling across South America (Table 2). For example, while proximal towns La Esperanza and Campo Florido, Peru have closely related *T. cruzi,* isolates from cities around Brazil encompass nearly the total genetic diversity.

Discussion

The invasion of human-altered environments by non-native species impacts the population health of native species as well as the human health and economies. Investigation of the dynamic process of immigration and establishment of non-native species into these rapidly expanding environments has the potential to mitigate the impacts of pests and pathogens that are detrimental to human populations as well as agricultural and native species (Crowl *et al.* 2008).

The analyses presented here suggest that the population of *T. cruzi* responsible for the recent, now controlled, epidemic, descended from a single recent invasion of Arequipa, Peru. This conclusion is supported by the extremely limited genetic diversity among *T. cruzi* isolates sampled within and around the city despite considerable genetic diversity observed regionally (Table 1; Fig 1). Several non-exclusive hypotheses may explain these results, including that the low probability of establishing a population in Arequipa is due to a low immigration rate; that immigrants rarely establish populations as a result of the

low transmission rate between hosts and vectors; and that there is a high replacement rate among lineages of *T. cruzi*.

Successful invasion of a novel geographic area is a function of the rate of immigration, the temporal duration that a habitat has been suitable for establishment, and the probability that an immigrant establishes a population. The influx of humans into Arequipa over the last ~60 years of urbanization and economic growth (Pedersen *et al.* 2008; Levy *et al.* 2014) has provided many opportunities for *T. cruzi* immigration. However, many human immigrants moved to locations in the city without established *T. infestans* populations (Bayer *et al.* 2009; Levy *et al.* 2014), which could have resulted in few opportunities for *T. cruzi* transmission from infected immigrant humans to their insect vector. Further, many people immigrating to Arequipa came from areas, such as the neighboring regions of Puno and Cusco, that are outside of the range of *T. infestans* (Bayer *et al.* 2009). A low probability of population establishment due to limited contact between *T. infestans* and infected immigrant humans is consistent with the observation of only a single established *T. cruzi* lineage in vectors in Arequipa.

Prior studies suggest that the majority of immigrants from most species that reach a novel geographic area fail to establish due to both inhospitable local environmental conditions (McKinney 2002; Thomas 2011) and stochasticity (Pitelka & Plant Migration Workshop Group 1997; Lockwood *et al.* 2005). Environmental factors that can reduce establishment probabilities include unfavorable abiotic conditions, limited food resources or vectors, or an abundance of predators or competitors. The establishment probability of immigrant *T. cruzi* in an urban environment is likely depressed by a low transmission rate from infected humans to vectors (Gürtler *et al.* 1998; Cohen & Gürtler 2001; Gürtler *et*

al. 2007; Llewellyn *et al.* 2009b). However, given a sizeable vector population, a long-living infected human is likely to transmit the parasite to a bug. Subsequent transmission of *T. cruzi* to a new host by a single infected vector is unlikely because of the low transmission rate (Rabinovich *et al.* 1990; Cohen & Gürtler 2001; Nouvellet *et al.* 2013). Thus, the low human-to-vector and subsequent vector-to-host transmission rates reduce the probability that a *T. cruzi* lineage that has immigrated to Arequipa will establish a population.

The data suggest that multiple *T. cruzi* lineages have immigrated to Arequipa in the recent past with all but one failing to transmit sufficiently to establish a population. This hypothesis is supported by the apparent immigration of a genetically distinct *T. cruzi* sample derived from an infected human at a hospital in Arequipa (Fig 3C) that has yet to establish a population in Arequipa. The potential for identifying additional lineages was hindered by our inability to collect samples from multiple humans. The presence of detectable immigration suggests that the low establishment probability is a key factor affecting the limited invasion of Arequipa and other location.

The observed establishment probability is likely independent of competition. Competitive exclusion—where an existing population prevents the invasion of new immigrants—appears unlikely as the majority of city blocks do not contain *T. cruzi* (Levy *et al.* 2014; Levy *et al.* 2015) despite substantial vector populations (Billig 2017). Under the competitive exclusion hypothesis, we might expect different *T. cruzi* lineages establishing in different areas of the city.

The limited genetic diversity within Arequipa could be the result of a recent replacement of a previously dominant lineage through natural population processes. While the

continuous substitution of a dominant strain through natural selection or drift is common in well-mixed populations, geographic structure within populations tends to result in the persistence of genetically diverse subpopulations (Kerr *et al.* 2002). For example, multiple *T. cruzi* lineages can co-circulate within the same locality (Oliveira *et al.* 1998; Higo *et al.* 2004; Llewellyn *et al.* 2009a; Llewellyn *et al. 2009b*; Curtis-Robles *et al.* 2017; Hodo *et al.* 2018), as seen in the samples sequenced here from La Esperanza (Fig 3), and even within the same host (Perez *et al.* 2014; Dumonteil *et al.* 2018). The absence of samples deriving from a previously established *T. cruzi* lineage in the fragmented urban and inter-district landscapes, much of which contains an active vector population but no *T. cruzi,* is suggestive that no previous populations dominated this area. Additionally, no minor sub-population circulating in the bug population was detected despite the temporal range of samples in our dataset (7 years).

In conclusion, all relevant data suggest that the vast majority, if not all, of *T. cruzi* circulating in vector populations prior to the recently-controlled epidemic in the city of Arequipa descended from a single immigrant. The fact that one strain found in a human patient was unrelated to this lineage suggests that *T. cruzi* may regularly immigrate to the city but that immigrants rarely establish growing populations. The limited genetic diversity, coupled with limited gene flow into the city, could limit the phenotypic diversity that results in variable effectiveness of treatment options and diagnoses (Revollo *et al.* 1998; Veloso *et al.* 2001). However, a potential downside of the limited diversity is that many currently available diagnostics may be ineffective against the local lineage (Longhi *et al.* 2012; Abras *et al.* 2016) as seen in Arequipa (Verani *et al.* 2009). It may be necessary to use locally-optimized diagnostics in endemic regions, which

highlights the difficulties in diagnosing *T. cruzi* infection in immigrants to non-endemic

countries such as the United States.

Figures and Tables
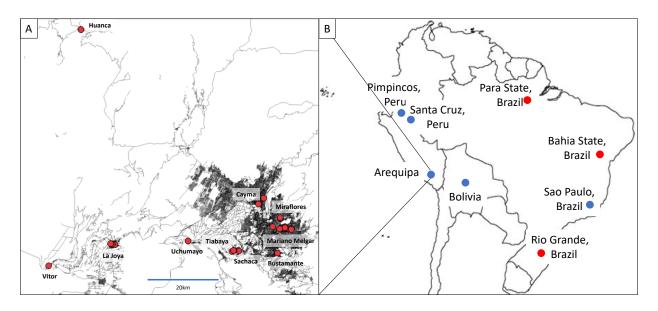
Figure 1. Spatial distribution of samples collected in (A) Arequipa, Peru and (B) South America. (A) The names of ten districts from which *T. cruzi* samples were collected are labeled. Houses from which isolates were collected are represented by red dots. Lines represent major roadways. Densely populated areas appear grey due to the density of roads. **(B)** The sites where isolates were collected are represented by blue dots. Sequences obtained from NCBI database are represented by red dots.
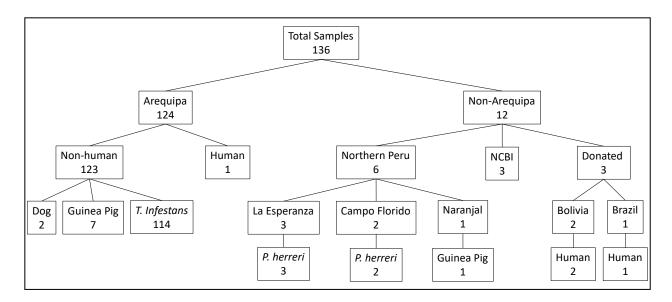
Figure 2. Number of samples collected from each host species per location. Most samples (N=124) were collected in Arequipa and 6 from small towns in northern Peru. 3 isolates were provided by the Infectious Diseases Research Laboratory at Universidad Peruana Cayetano Heredia. Maxicircle sequences for 3 *T. cruzi* lineages were downloaded from the NCBI database.

Table 1. Population Genetic Statistics.

	Average Pairwise Distance	π	θ	Segregating Sites
All samples (n=136)	126	$8.18*10^{-3}$	$2.44*10^{-2}$	2055
Arequipa (n=123)	1.04	$6.80*10^{-5}$	$1.93*10^{-4}$	16
South America (n=13)	728	$4.73*10^{-2}$	$4.24*10^{-2}$	2022
La Esperanza (n=3)	643	$4.18*10^{-2}$	$4.18*10^{-2}$	964

```
┌─ Bahia State, Brazil
└─ São Paulo, Brazil

┌─ (C)Arequipa, Peru-Human
└─ Rio Grande, Brazil
   (B) La Esperanza, Campo
   Florido, Naranjal, Peru
   Bolivia
   Bolivia
   (A)Arequipa, Peru
   La Esperanza, Peru
   La Esperanza, Peru
   Para State, Brazil
```

100 SNPs
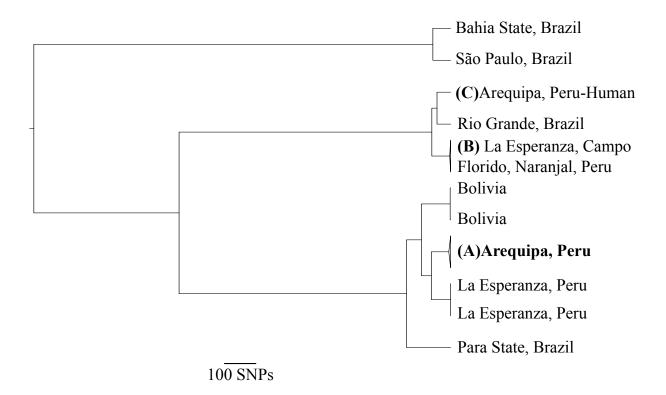
Figure 3. The extant *T. cruzi* population in Arequipa arose from a single, recent

introduction. Maximum clade credibility (MCC) tree shows that **(A)** all 123 *T. cruzi*

isolated from bugs and domestic mammals in Arequipa form a monophyletic clade with a

single, recent common ancestor, indicative of a single immigration event in the recent

past. Despite substantial genetic diversity among *T. cruzi* throughout South America,

those collected in Arequipa show little diversity. **(B)** Three samples collected in Campo

Florido and Naranjal, Peru and one sample from La Esperanza, Peru have nearly identical

maxicircle sequences and form a monophyletic clade. La Esperanza, Peru contains at

least two distinct *T. cruzi* lineages, suggesting multiple independent introductions. **(C)**

The only *T. cruzi* sample isolated from a human in Arequipa is distinct from all other

samples from Arequipa, suggesting that this introduction has not colonized the city. All

tips represent a single sample except (A) (N=123) and (B) (N=4). All nodes have strong

support (posterior probability≥0.99). Nodes are collapsed when the samples contained

have nearly identical maxicircle sequences.

Table 2. Distance matrix showing average pairwise SNP distance between samples (bottom triangle) and Euclidean distance between sample collection locations (top triangle).

	(A) Arequipa, Peru	La Esperanza, Peru	La Esperanza, Peru	(B) La Esperanza, Campo Florido, & Naranjal, Peru	Bolivia	Bolivia	(C) Arequipa, Peru - Human	São Paulo, Brazil	Para State, Brazil	Bahia State, Brazil	Rio Grande, Brazil
(A) Arequipa, Peru		1400km	1400km	1400km	700km	700km	0km	2700km	2500km	3200km	2500km
Pimpincos, Peru	95		0km	60km	2000km	2000km	1400km	4000km	3000km	4100km	3900km
Pimpincos, Peru	95	0		60km	2000km	2000km	1400km	4000km	3000km	4100km	3900km
(B) La Esperanza, Campo Florido, & Naranjal, Peru	970	966	966		2000km	2000km	1400km	4000km	3000km	4100km	3900km
Bolivia	137	137	137	957		0km	700km	2000km	1800km	2500km	2000km
Bolivia	137	137	137	958	2		700km	2000km	1800km	2500km	2000km
(C) Arequipa, Peru - Human	975	971	971	89	967	968		2700km	2500km	3200km	2500km
São Paulo, Brazil	1044	1045	1045	1033	1045	1046	1058		2100km	1300km	1100km
Para State, Brazil	216	216	216	984	205	205	992	1059		1500km	2800km
Bahia State, Brazil	1284	1289	1289	1234	1287	1287	1258	68	1308		2500km
Rio Grande, Brazil	1005	1003	1003	82	1001	1001	77	1052	1026	1284	

(A) 123 samples isolated from bugs, dogs, and guinea pigs in Arequipa are represented here.

(B) 4 samples collected in Campo Florido, Naranjal, and La Esperanza are represented here. Euclidean distances (top triangle) are displayed for Santa Cruz.

(C) The only sample isolated from a human in Arequipa is represented here.

Literature Cited

Abras, A., Gállego, M., Llovet, T., Tebar, S., Herrero, M., Berenguer, P., et al. (2016). Serological Diagnosis of Chronic Chagas Disease: Is It Time for a Change? *Journal of Clinical Microbiology*, *54*(6), 1566–1572. http://doi.org/10.1128/JCM.00142-16

Adalja, A. A., Sell, T. K., Bouri, N., & Franco, C. (2012). Lessons Learned during Dengue Outbreaks in the United States, 2001–2011. *Emerging Infectious Diseases*, *18*(4), 1–7. http://doi.org/10.3201/eid1804.110968

Alirol, E., Getaz, L., Stoll, B., Chappuis, F., & Loutan, L. (2011). Urbanisation and infectious diseases in a globalised world. *The Lancet Infectious Diseases*, *11*(2), 131–141.

Alroy, K. A., Huang, C., Gilman, R. H., Quispe-Machaca, V. R., Marks, M. A., Ancca Juárez, J., et al. (2015). Prevalence and transmission of Trypanosoma cruzi in people of rural communities of the high jungle of Northern Peru. *PLOS Neglected Tropical Diseases*, *9*(5), e0003779.

Aronson, M. F., La Sorte, F. A., Nilon, C. H., Katti, M., Goddard, M. A., Lepczyk, C. A., et al. (2014). A global analysis of the impacts of urbanization on bird and plant diversity reveals key anthropogenic drivers. *Proc. R. Soc. B*, *281*(1780), 20133330.

Ayres, D. L., Darling, A., Zwickl, D. J., Beerli, P., Holder, M. T., Lewis, P. O., et al. (2012). BEAGLE: An Application Programming Interface and High-Performance Computing Library for Statistical Phylogenetics. *Systematic Biology*, *61*(1), 170–173. http://doi.org/10.1093/sysbio/syr100

Bayer, A. M., Hunter, G. C., Gilman, R. H., Cornejo del Carpio, J. G., Naquira, C., Bern, C., & Levy, M. Z. (2009). Chagas Disease, Migration and Community Settlement

Patterns in Arequipa, Peru. *PLOS Neglected Tropical Diseases*, *3*(12), e567. http://doi.org/10.1371/journal.pntd.0000567

Billig, Erica. (2017). Detecting And Controlling Insect Vectors In Urban Environments: Novel Bayesian Methods For Complex Spatial Data. *Publicly Accessible Penn Dissertations*. 2193. https://repository.upenn.edu/edissertations/2193

Blair, R. B. (2001). Birds and butterflies along urban gradients in two ecoregions of the United States: Is urbanization creating a homogeneous fauna? *Biotic Homogenization*, 33–56.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., et al. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology*, *10*(4), e1003537. http://doi.org/10.1371/journal.pcbi.1003537

Bowman, N. M., Kawai, V., Levy, M. Z., Cornejo del Carpio, J. G., Cabrera, L., Delgado, F., et al. (2008). Chagas Disease Transmission in Periurban Communities of Arequipa, Peru. *Clinical Infectious Diseases*, *46*(12), 1822–1828. http://doi.org/10.1086/588299

Bradley, C. A., & Altizer, S. (2007). Urbanization and the ecology of wildlife diseases. *Trends in Ecology & Evolution*, *22*(2), 95–102. http://doi.org/10.1016/j.tree.2006.11.001

Castillo Neyra, R., Borrini-Mayorí, K., Salazar Sánchez, R., Ancca Suarez, J., Xie, S., Náquira Velarde, C., & Levy, M. Z. (2016). Heterogeneous infectiousness in guinea

pigs experimentally infected with Trypanosoma cruzi. *Parasitology International, 65*(1), 50–54.

Cáceres, A. G., Troyes, L., Gonzáles-Pérez, A., Llontop, E., Bonilla, C., Heredia, N., et al. (2002). Enfermedad de Chagas en la región nororiental del Perú. I. Triatominos (Hemiptera, Reduviidae) presentes en Cajamarca y Amazonas. *Revista Peruana De Medicina Experimental Y Salud Publica, 19*(1), 17–23.

Chiari, E., & Galvão, L. (1997). Diagnóstico parasitológico da doença de Chagas. *Clínica E Terapêutica Da Doença De Chagas: Uma Abordagem Prática Para O Clínico Geral. Rio De Janeiro: Fiocruz,* 85–97.

Cohen, J. E., & Gürtler, R. E. (2001). Modeling household transmission of American trypanosomiasis. *Science, 293*(5530), 694–698. http://doi.org/10.1126/science.1060638

Crowl, T. A., Crist, T. O., Parmenter, R. R., Belovsky, G., & Lugo, A. E. (2008). The spread of invasive species and infectious disease as drivers of ecosystem change. *Frontiers in Ecology and the Environment, 6*(5), 238–246. http://doi.org/10.1890/070151

Curtis-Robles, R., Snowden, K. F., Dominguez, B., Dinges, L., Rodgers, S., Mays, G., & Hamer, S. A. (2017). Epidemiology and Molecular Typing of Trypanosoma cruzi in Naturally-Infected Hound Dogs and Associated Triatomine Vectors in Texas, USA. *PLOS Neglected Tropical Diseases, 11*(1), e0005298–18. http://doi.org/10.1371/journal.pntd.0005298

Deguines, N., Julliard, R., De Flores, M., & Fontaine, C. (2016). Functional homogenization of flower visitor communities with urbanization. *Ecology and Evolution*, *6*(7), 1967–1976.

Denys, C., & Schmidt, H. (1998). Insect communities on experimental mugwort (Artemisia vulgaris L.) plots along an urban gradient. *Oecologia*, *113*(2), 269–277.

Deplazes, P., Hegglin, D., Gloor, S., & Romig, T. (2004). Wilderness in the city: the urbanization of Echinococcus multilocularis. *Trends in Parasitology*, *20*(2), 77–84. http://doi.org/10.1016/j.pt.2003.11.011

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, *29*(8), 1969–1973. http://doi.org/10.1093/molbev/mss075

Dumonteil, E., Ramirez-Sierra, M.-J., rez-Carrillo, S. P. X., Teh-Poot, C., Herrera, C., re, S. X. B. G. X., & Waleckx, E. (2018). Detailed ecological associations of triatomines revealed by metabarcoding and next- generation sequencing: implications for triatomine behavior and Trypanosoma cruzi transmission cycles. *Scientific Reports*, 1–13. http://doi.org/10.1038/s41598-018-22455-x

Foley, E. A., Khatchikian, C. E., Hwang, J., Ancca Juárez, J., Borrini-Mayorí, K., Quıspe-Machaca, V. R., et al. (2013). Population structure of the Chagas disease vector, Triatoma infestans, at the urban-rural interface. *Molecular Ecology*, *22*(20), 5162–5171.

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., et al. (2005). Global consequences of land use. *Science*, *309*(5734), 570–574.

Gaunt, M. W., Yeo, M., Frame, I. A., Stothard, J. R., Carrasco, H. J., Taylor, M. C., et al. (2003). Mechanism of genetic exchange in American trypanosomes. *Nature*, *421*(6926), 936–939. http://doi.org/10.1038/nature01438

Geer, L. Y., Marchler-Bauer, A., Geer, R. C., Han, L., He, J., He, S., et al. (2010). The NCBI BioSystems database. *Nucleic Acids Research*, *38*(suppl_1), D492–D496. http://doi.org/10.1093/nar/gkp858

Gürtler, R. E., Cecere, M. C., Lauricella, M. A., Cardinal, M. V., Kitron, U., & Cohen, J. E. (2007). Domestic dogs and cats as sources of Trypanosoma cruzi infection in rural northwestern Argentina. *Parasitology*, *134*(1), 69–82.

Gürtler, R. E., Cohen, J. E., Cecere, M. C., Lauricella, M. A., Chuit, R., & Segura, E. L. (1998). Influence of humans and domestic animals on the household prevalence of Trypanosoma cruzi in Triatoma infestans populations in northwest Argentina. *The American Journal of Tropical Medicine and Hygiene*, *58*(6), 748–758.

Hahs, A. K., McDonnell, M. J., McCarthy, M. A., Vesk, P. A., Corlett, R. T., Norton, B. A., et al. (2009). A global synthesis of plant extinction rates in urban areas. *Ecology Letters*, *12*(11), 1165–1173.

Harington, J. S. (1960). A simple apparatus for the artificial feeding of Rhodnius prolixus (Hemiptera, Reduviidae). *Parasitology*, *50*(1-2), 273–277.

Harrington, L. C., Scott, T. W., Lerdthusnee, K., Coleman, R. C., Costero, A., Clark, G. G., et al. (2005). Dispersal of the dengue vector Aedes aegypti within and between rural communities. *The American Journal of Tropical Medicine and Hygiene*, *72*(2), 209–220. http://doi.org/10.4269/ajtmh.2005.72.209

Heled, J., & Drummond, A. J. (2008). Bayesian inference of population size history from multiple loci. *BMC Evolutionary Biology*, *8*(1), 289–15. http://doi.org/10.1186/1471-2148-8-289

Higo, H., Miura, S., Horio, M., Mimori, T., Hamano, S., Agatsuma, T., et al. (2004). Genotypic variation among lineages of Trypanosoma cruzi and its geographic aspects. *Parasitology International*, *53*(4), 337–344. http://doi.org/10.1016/j.parint.2004.06.001

Hodo, C. L., Wilkerson, G. K., Birkner, E. C., Gray, S. B., & Hamer, S. A. (2018). Trypanosoma cruzi Transmission Among Captive Nonhuman Primates, Wildlife, and Vectors. *EcoHealth*, 1–11.

Hunter, G. C., Borrini-Mayorí, K., Ancca Juárez, J., Castillo Neyra, R., Verastegui, M. R., Malaga Chavez, F. S., et al. (2012). A Field Trial of Alternative Targeted Screening Strategies for Chagas Disease in Arequipa, Peru. *PLOS Neglected Tropical Diseases*, *6*(1), e1468. http://doi.org/10.1371/journal.pntd.0001468

Kerr, B., Riley, M. A., Feldman, M. W., & Bohannan, B. J. (2002). Local dispersal promotes biodiversity in a real-life game of rock–paper–scissors. *Nature Publishing Group*, *418*(6894), 171.

Koboldt, D. C., Chen, K., Wylie, T., Larson, D. E., McLellan, M. D., Mardis, E. R., et al. (2009). VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics*, *25*(17), 2283–2285. http://doi.org/10.1093/bioinformatics/btp373

Kowarik, I. (1995). On the Role of Alien Species in Urban Flora and Vegetation. *Plant Invasions-General Aspects and Special Problems*, 85–103.

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, *33*(7), 1870–1874. http://doi.org/10.1093/molbev/msw054

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359.

Levy, M. Z., Barbu, C. M., Castillo Neyra, R., Quispe-Machaca, V. R., Ancca Juárez, J., Escalante-Mejia, P., et al. (2014). Urbanization, land tenure security and vector-borne Chagas disease. *Proceedings. Biological Sciences*, *281*(1789), 20141003–20141003. http://doi.org/10.1098/rspb.2014.1003

Levy, M. Z., Bowman, N. M., Kawai, V., Waller, L. A., Cornejo del Carpio, J. G., Benzaquen, E. C., et al. (2006). Periurban Trypanosoma cruzi–infected Triatoma infestans, Arequipa, Peru. *Emerging Infectious Diseases*, *12*(9), 1345–1352.

Levy, M. Z., Kawai, V., Bowman, N. M., Waller, L. A., Cabrera, L., Pinedo-Cancino, V. V., et al. (2007). Targeted Screening Strategies to Detect Trypanosoma cruzi Infection in Children. *PLOS Neglected Tropical Diseases*, *1*(3), e103. http://doi.org/10.1371/journal.pntd.0000103

Levy, M. Z., Small, D. S., Vilhena, D. A., Bowman, N. M., Kawai, V., Cornejo del Carpio, J. G., et al. (2011). Retracing Micro-Epidemics of Chagas Disease Using Epicenter Regression. *PLoS Computational Biology*, *7*(9), e1002146. http://doi.org/10.1371/journal.pcbi.1002146

Llewellyn, M. S., Lewis, M. D., Acosta, N., Yeo, M., Carrasco, H. J., Segovia, M., et al. (2009a). Trypanosoma cruzi IIc: Phylogenetic and Phylogeographic Insights from Sequence and Microsatellite Analysis and Potential Impact on Emergent Chagas

Disease. *PLOS Neglected Tropical Diseases*, *3*(9), e510.

http://doi.org/10.1371/journal.pntd.0000510

Llewellyn, M. S., Miles, M. A., Carrasco, H. J., Lewis, M. D., Yeo, M., Vargas, J., et al.

(2009b). Genome-Scale Multilocus Microsatellite Typing of Trypanosoma cruzi

Discrete Typing Unit I Reveals Phylogeographic Structure and Specific Genotypes

Linked to Human Infection. *PLOS Pathogens*, *5*(5), e1000410.

http://doi.org/10.1371/journal.ppat.1000410

Lockwood, J. L., Cassey, P., & Blackburn, T. (2005). The role of propagule pressure in

explaining species invasions. *Trends in Ecology & Evolution*, *20*(5), 223–228.

Longhi, S. A., Brandariz, S. B., Lafon, S. O., Niborski, L. L., Luquetti, A. O., Schijman,

A. G., et al. (2012). 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*, *87*(2), 267–

271.

Mackin-Rogalska, R., Pinowski, J., Solon, J., & Wojcik, Z. (1988). Changes in

vegetation, avifauna, and small mammals in a suburban habitat. *Polish Ecological

Studies*, *14*(1-2), 293–330.

McIntyre, N. E. (2000). Ecology of Urban Arthropods: A Review and a Call to Action.

Annals of the Entomological Society of America, *93*(4), 825–835.

http://doi.org/10.1603/0013-8746(2000)093[0825:EOUAAR]2.0.CO;2

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al.

(2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-

generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303.

http://doi.org/10.1101/gr.107524.110

McKinney, M. L. (2002). Urbanization, Biodiversity, and Conservation. *BioScience*,

52(10), 883–890. http://doi.org/10.1641/0006-3568(2002)052[0883:UBAC]2.0.CO;2

New, T. R. (2018). Promoting and developing insect conservation in Australia's urban

environments. *Austral Entomology*.

Nouvellet, P., Dumonteil, E., & Gourbière, S. (2013). The improbable transmission of

Trypanosoma cruzi to human: the missing link in the dynamics and control of Chagas

disease. *PLOS Neglected Tropical Diseases*, *7*(11), e2505.

Oliveira, R. P., Broude, N. E., Macedo, A. M., Cantor, C. R., Smith, C. L., & Pena, S. D.

J. (1998). Probing the genetic population structure of Trypanosoma cruzi with

polymorphic microsatellites. *Proceedings of the National Academy of Sciences of the

United States of America*, *95*(7), 3776–3780.

Pedersen, D., Tremblay, J., Errázuriz, C., & Gamarra, J. (2008). The sequelae of political

violence: assessing trauma, suffering and dislocation in the Peruvian highlands.

Social Science & Medicine, *67*(2), 205–217.

Pekin, B. K., & Pijanowski, B. C. (2012). Global land use intensity and the endangerment

status of mammal species. *Diversity and Distributions*, *18*(9), 909–918.

Perez, C. J., Lymbery, A. J., & Thompson, R. C. A. (2014). Chagas disease: the challenge

of polyparasitism? *Trends in Parasitology*, *30*(4), 176–182.

http://doi.org/10.1016/j.pt.2014.01.008

Pitelka, L. F., & Group, P. M. W. (1997). Plant migration and climate change: a more

realistic portrait of plant migration is essential to predicting biological responses to

global warming in a world drastically altered by human activity. *American Scientist*, 464–473.

Rabinovich, J. E., Wisnivesky-Colli, C., Solarz, N. D., & Gürtler, R. E. (1990). Probability of transmission of Chagas disease by Triatoma infestans (Hemiptera: Reduviidae) in an endemic area of Santiago del Estero, Argentina. *Bulletin of the World Health Organization*, *68*(6), 737.

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., Suchard, M. A., & Susko, E. (2018). Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic Biology*, *22*, 341–syy032 VL – IS –. http://doi.org/10.1093/sysbio/syy032

Revollo, S., Oury, B., Laurent, J.-P., Barnabé, C., Quesney, V., Carrière, V., et al. (1998). Trypanosoma cruzi: impact of clonal evolution of the parasite on its biological and medical properties. *Experimental Parasitology*, *89*(1), 30–39.

Seto, K. C., Güneralp, B., & Hutyra, L. R. (2012). Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proceedings of the National Academy of Sciences*, *109*(40), 16083–16088.

Strohbach, M. W., Haase, D., & Kabisch, N. (2009). Birds and the city: urban biodiversity, land use, and socioeconomics. *Ecology and Society*, *14*(2).

Suchard, M. A., & Rambaut, A. (2009). Many-core algorithms for statistical phylogenetics. *Bioinformatics*, *25*(11), 1370–1376. http://doi.org/10.1093/bioinformatics/btp244

Theoharides, K. A., & Dukes, J. S. (2007). Plant invasion across space and time: factors affecting nonindigenous species success during four stages of invasion. *New Phytologist*, *176*(2), 256–273. http://doi.org/10.1111/j.1469-8137.2007.02207.x

Thomas, C. D. (2011). Translocation of species, climate change, and the end of trying to recreate past ecological communities. *Trends in Ecology & Evolution*, *26*(5), 216–221. http://doi.org/10.1016/j.tree.2011.02.006

Veloso, V. M., Carneiro, C. M., Toledo, M., Lana, M., Chiari, E., Tafuri, W. L., & Bahia, M. T. (2001). Variation in susceptibility to benznidazole in isolates derived from Trypanosoma cruzi parental strains. *Memorias Do Instituto Oswaldo Cruz*, *96*(7), 1005–1011.

Verani, J. R., Seitz, A., Gilman, R. H., LaFuente, C., Galdos-Cardenas, G., Kawai, V., et al. (2009). Geographic Variation in the Sensitivity of Recombinant Antigen-based Rapid Tests for Chronic Trypanosoma cruzi Infection. *The American Journal of Tropical Medicine and Hygiene*, *80*(3), 410–415. http://doi.org/10.4269/ajtmh.2009.80.410

CHAPTER 2: Dispersal patterns of *Trypanosoma cruzi* through an urban landscape

Abstract

Anthropogenic alteration of natural habitats through the process of urbanization threatens thousands of species with both local and global extinction. Urbanization also creates novel environments that allow some species, including many human pests and pathogens, to thrive. Elucidating patterns of establishment and dispersal in urban environments may guide efforts to control the spread of pests and pathogens. *Trypanosoma cruzi*, a protozoan parasite and the causative agent of Chagas disease in humans, and its insect vector *Triatoma infestans* have spread through the city of Arequipa, Peru. Here, we use *T. cruzi* as a model to elucidate patterns of establishment and dispersal in urban ecosystems. We sequenced whole genomes of 123 *T. cruzi* isolates collected throughout 10 districts in and around the city of Arequipa, Peru. We used population genomic and phylogenomic tools to determine patterns of *T. cruzi* dispersal throughout Arequipa. The data show significant population structure within city blocks such that parasites in the same block tend to be very closely related, but no population structure among blocks within districts, such that parasites in neighboring blocks are no more closely related to one another than to parasites in distant districts. These data suggest that *T. cruzi* dispersal within a block occurs regularly and that occasional long-range dispersal events allow the establishment of new *T. cruzi* populations in distant blocks. Movement of humans and domestic animals may be the primary mechanism of longer-distance *T. cruzi* dispersal in the city. As the number and size of urban environments increase globally, it is more important than ever to understand the dispersal dynamics of pests and pathogens.

.

Introduction

Human activities have resulted in major changes to biological communities worldwide (McKinney 2002; Foley *et al.* 2005; Seta *et al.* 2012). Anthropogenic alteration of natural habitats threatens thousands of species with both local and global extinction (McKinney 2002). Anthropogenic alteration can also create novel environments that can be successfully exploited by numerous species including many human pests and pathogens. Understanding how life history traits interact with features of human altered ecosystems to facilitate or hinder the proliferation of populations is imperative for both conservation of biodiversity and for efforts to control human pests and pathogens that prosper in urban environments. Proliferation in an ecosystem is accomplished by immigration to a new habitat, establishment (reproduction and population growth), and dispersal to new habits within the ecosystem (Theoharides & Dukes 2007). Investigations into the dynamic processes of immigration, establishment, and dispersal of disease-causing agents in urban ecosystems are particularly important for public health risk management because the distribution and abundance of these pest species is correlated with the incidence of disease in humans (Alirol *et al.* 2011). Identifying environmental features that facilitate or hinder the proliferation of populations elucidates the mechanisms underlying their distribution and abundance in the urban ecosystem. In this study, we examine the dispersal pattern of *Trypanosoma cruzi*, the causative agent of Chagas disease, in and around the city of Arequipa, Peru.

The Chagas disease system in Arequipa, Peru is ideal to study the factors associated with dispersal in recently urbanized environments for two reasons. First, a single

introduction of *T. cruzi* recently established in Arequipa and has since dispersed throughout many districts in the city (Chapter 1). A single population allows dispersal patterns through the city to be assessed without being confounded by repeated immigration from outside. Second, the high rate of human immigration into Arequipa has increased drastically over the past ~50 years (Pedersen *et al.* 2008; Levy *et al.* 2014). The current urban expansion has resulted in a mosaic of several types of habitat, each of which provides unique challenges and opportunities for *T. cruzi* to establish and proliferate (Foley *et al.* 2013; Levy *et al.* 2014; Khatchikian *et al.* 2015)]. Here, we provide a model for understanding how populations disperse through urban environments using whole genome sequences of 123 *T. cruzi* isolates collected throughout the city of Arequipa, Peru.

Methods

Sample Collection and Study Site

Sample collection and genome sequencing are described in Chapter 1. Briefly, 123 *T. cruzi* samples were isolated from infected *Triatoma infestans* bugs (N=114), dogs (N=2), and guinea pigs (N=7) from houses throughout Arequipa, Peru (Figure 1A). Nearly half of the samples isolated from *T. infestans* (N=56) were collected during intensive inspections of houses and surrounding areas along an established transect within the Mariano Melgar district in 2010 and 2011 as described in Levy *et al.* 2014 (Figure 1B). Many blocks in Mariano Melgar contained bug populations uninfected with *T. cruzi* (Levy *et al.* 2014), which is a pattern that has been observed in other parts of Arequipa (Levy *et al.* 2015). *T. cruzi* DNA was extracted from each sample and sequenced to an average depth of >50X as described in Chapter 1.

Whole genome assembly

Genomes were assembled using the most closely related reference genome, TcJR clone 4, obtained from TriTrypDB, using bowtie2 (Langmead & Salzberg 2012). 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 28Mbp. This assembly largely excluded the extensive repeat regions found throughout the *T. cruzi* genome. Duplicate reads were removed using Picard MarkDuplicates (McKenna *et al.* 2010).

SNP calling

Individual gVCF files containing SNP data for each sample were generated using GATK HaplotypeCaller (McKenna *et al.* 2010; Poplin *et al.* 2017) following GATK's Best Practices procedure (De Pristo *et al.* 2011; Van der Auwera *et al.* 2013). 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.

Population Genetics Analyses

VCF files were converted to plink format using vcftools (Danecek *et al.* 2011) and plink (Purcell *et al.* 2007). Violin plots were used to visualize the pairwise SNP distance

between samples collected at various spatial scales using ggplot2 (Wickham, 2009) in R (R Core Team, 2017). AMOVA was performed using the poppr package in R (Kamvar *et al.* 2014). PCA was calculated using Tassel 5 (Bradbury *et al.* 2007). A maximum clade credibility (MCC) phylogenetic tree was reconstructed using BEAST 1.8.4 (Drummond *et al.* 2012) implementing a Kimura 3-parameter substitution model, an Extended Bayesian Skyline coalescent tree prior (Heled & Drummond 2008), with equal base frequencies, a strict clock, and a UPGMA starting tree chosen based on BEAST Model Test implemented in BEAST2 (Bouckaert *et al.* 2014). A 10% burn-in was used and log files were examined for convergence using Tracer v1.6.0 (Rambaut *et al.* 2018) ensuring ESS values above 200. A MCC tree was generated using TreeAnnotator 1.8.4. and visualized using FigTree (available at http://beast.bio.ed.ac.uk). Phylogenetic analyses were performed using the BEAGLE library to increase computation speed (Suchard & Rambaut 2009; Ayres *et al.* 2011). 100 independent iterations of ADMIXTURE (Alexander *et al.* 2009) were run for each number of genetic clusters (K, ranging from 2 to 10) assuming linkage disequilibrium until the log-likelihood increased by less than $\varepsilon=10^{-4}$ between iterations. The optimal number of clusters was estimated to be 4 by the cross-validation score averaged across 100 iterations (Supp Fig 2). The optimal alignment of the 100 iterations was calculated using CLUMPP (Jakobsson *et al.* 2007).

Results

A total of 9271 of the >28 million sites in the genome sequences from 123 *T. cruzi* parasites collected from ten spatially separated districts throughout the city of Arequipa, Peru were polymorphic. The average pairwise distance between genomes was 1726 SNPs (Table 1). Between samples from different houses within each city block, however, the

average pairwise distance is 620 (Fig 2). There is significantly less variation among houses within each block than expected by chance (p<0.001) (Fig 2; Table 2). The limited genetic diversity observed within each block using population genetic analyses can also be observed on the phylogeny as most city blocks contain only a single *T. cruzi* lineage with all subsequent diversity generated by mutations *in situ* (Fig 3). This result is also supported by ADMIXTURE analyses of the 56 samples from Mariano Melgar—the most densely-sampled district—which demonstrate that most city blocks are dominated by a single *T. cruzi* genetic cluster (Supp Fig 1). Generally, if one house is infected with *T. cruzi,* multiple houses on that block are infected. 8 out of 12 blocks in Mariano Melgar with an active *T. cruzi* infection have more than one infected house and the average number of infected houses per infected block is 2.83.

T. cruzi collected from different blocks tend to be unrelated regardless of spatial proximity. There is strong evidence that *T. cruzi* gene flow is significantly restricted (p<0.001) among blocks in analyses of the whole dataset and in analyses focusing only on Mariano Melgar (Table 2). This is consistent with the result that *T. cruzi* in different blocks within Mariano Melgar tend to belong to unique genetic clusters (Supp Fig 1). Further, the lineages occupying neighboring city blocks are often phylogenetically divergent suggesting that geographic distance is not correlated with evolutionary distance (Fig 3).

There is no evidence supporting restricted gene flow among the districts of Arequipa (Table 1). The majority of the total genomic diversity is present in the two most densely-sampled districts—Mariano Melgar and La Joya (Fig 3). Further, the lineages occupying

city blocks within Mariano Melgar or La Joya are commonly most closely related to lineages in other districts, supporting common dispersal among districts in the city.

Accounting for Recombination

In addition to the genome-wide SNP data, a dataset containing presence/absence of 474 unique recombination events was generated (described in detail in Chapter 3). Violin plots were used to visualize the pairwise differences in recombination events (Supp Fig 4). An AMOVA was also performed using this dataset to determine if analyses performed using genome-wide SNP data were consistent with those performed using presence of recombination events (Supp Table 1). Analyses using the presence or absence of each recombination event have limited power because these analyses use only the 474 events, each with two character states (presence or absence) as opposed to the nearly 10,000 polymorphic loci in the SNP dataset. However, the genome-wide SNP data may be over-weighing some regions of the genome. For example, one recombination event may homogenize 500 formerly-heterozygous loci in one lineage while another equally likely recombination event only affects 50 loci in another lineage. Here, analyses of genome-wide SNPs may weigh the former event 10X more than the latter. While this may affect divergence time estimation, the similar results between the violin plots (Fig 2 vs. Supp Fig 4) and AMOVAs (Table 2 vs. Supp Table 1) show that it had no effect on estimations of relatedness likely because all recombination events are identical by descent. The results show that both genome-wide SNPs and presence/absence of recombination events provide similar representations of genetic relatedness.

Discussion

Investigations into the impact of human altered environments have focused primarily on habitat destruction and extinction of natural fauna (Mackin-Rogalska *et al.* 1988; Kowarik 1995; Denys & Schmidt 1998; McIntyre 2000; Blair 2001; McKinney 2002). Equally impactful is the creation of novel habitats that are exploited by human-associated pests (Alirol *et al.* 2011). Understanding how wildlife exploit human altered habitats, especially in human-dense urban habitats, is important for understanding the distribution and abundance of human diseases. Here, we investigate the dispersal patterns of *Trypanosoma cruzi*, the cause of human Chagas disease, through the urban ecosystem of Arequipa, Peru. The genomic data support a pattern of dispersal in which dispersal proceeds with little impediment among houses within a city block after a population is established. However, establishment of a novel *T. cruzi* population in a city block is a rare event as *T. cruzi* was undetected in a majority of blocks and blocks supporting a population tend to contain only a single genetic lineage. Geographic proximity to blocks with an active population increases the establishment probability of that block only slightly. In fact, long distance immigration from distant blocks, even blocks in other districts, is as common as immigration from proximal blocks. The dispersal pattern observed in Arequipa likely results from a combination of relatively rare immigration to blocks without an established *T. cruzi* population and a low probability of establishing a population after immigration. These patterns support a model of gene flow in which city blocks are relatively homogenous, high-quality patches that allow within-block gene flow, separated by inhospitable barriers (i.e. roads and streets) that greatly reduce gene flow between blocks.

The population genetic and phylogenetic analyses suggest that each city block contains limited barriers to *T. cruzi* gene flow. After *T. cruzi* establishes in a new block, the population readily disperses to neighboring houses within the block, likely carried by their vectors which move between houses through shared walls (Levy et al. 2008; Khatchikian *et al.* 2015). In fact, most (67%) blocks with any *T. cruzi*-infected bugs had two or more houses containing infected bugs (Fig 3). *T. cruzi* population expansion within each block is likely the result of the abundance of resources for both the parasite and its vector. For example, many houses accommodate guinea pigs – which are particularly important reservoirs of *T. cruzi* (Herrer 1955; Cohen & Gürtler 2001; Levy *et al.* 2006; Levy *et al.* 2008; Coffield *et al.* 2013; Levy *et al.* 2014; Levy *et al.* 2015; Castillo-Neyra *et al.* 2016), dogs, and humans, which offer an abundance of blood meals for the vector and competent hosts for the parasite, providing stable reservoirs for *T. cruzi* infection (Levy *et al.* 2014). These data suggest that blocks remain free of *T. cruzi* due to either low rates of dispersal or low probability of establishment as populations tend to thrive once established.

The distribution of *T. cruzi* among city blocks suggests that dispersal between blocks is rare and likely limited by *T. infestans* dispersal between blocks, similar to reports showing rare between-habitat urban dispersal in other species (Angold *et al.* 2006; Cheptou *et al.* 2008; Niemelä & Kotze 2009). *T. cruzi* was not detected in a majority of city blocks surveyed between 2008-2012 (Levy *et al.* 2014). Blocks containing *T. cruzi* generally supported only a single lineage suggesting a single establishment event. These data suggest either strong barriers to dispersal among city blocks or barriers that hinder the establishment of new populations. Prior investigations demonstrated that streets are a

barrier to dispersal for *T. infestans*, the primary vector of *T. cruzi* in the area, among

blocks (Barbu *et al.* 2013; Khatchikian *et al.* 2015). Nevertheless, there is evidence of

dispersal for both *T. cruzi* and *T. infestans* between neighboring blocks, especially in

areas that urbanized in the more distant past. For example, the older *T. cruzi* populations

likely dispersed between Blocks 1-4 and from Block 7 to 6 as opposed to the recently

established population in Block 5 which has not progressed to neighboring blocks (Levy

et al. 2014)(Fig 3). These data suggest that dispersal to proximal blocks is rare but can

occur given sufficient time, supporting the hypothesis of a low rate of dispersal.

Rare establishment of new city blocks observed in the data may also result from

barriers that hinder the establishment of new populations. These barriers could result

from both life history characteristics of *T. cruzi* – such as inefficient transmission from

vectors to hosts (Rabinovich *et al.* 1990; Cohen & Gürtler 2001) – and specific features

of a block that reduce the probability of establishing a population after introduction –

such as the presence of a vector population, competitive interactions with an existing *T.

cruzi* population, or human-specific activity patterns. Low transmission probability from

bug to host reduces the establishment probability of *T. cruzi* because the parasite is not

vertically (Kribs-Zaleta & Mubayi 2012) or horizontally (Levy *et al.* 2015) transmitted

between bugs. Therefore, *T. cruzi* may be unable to establish a population in a new block

even if its infected vector successfully establishes a population. The discordant

occupancy patterns between *T. infestans* (Khatchikian *et al.* 2015) and *T. cruzi* is

suggestive of the inefficiency of *T. infestans* as a mode of between-block *T. cruzi*

dispersal.

The patchwork of city blocks in urban ecosystems can result in a mosaic of environments that can differentially affect dispersal patterns and the probability of establishment (Kowarik 2011). For example, one block within the heavily sampled Mariano Melgar district (Fig 3, Block 6) appears to have a distinctly higher rate of establishment than all other sampled blocks. This block supports *T. cruzi* representing six unique lineages (Fig 3) including one that likely immigrated from a proximal block (Block 7) and five from locations around Arequipa. The establishment of multiple lineages from adjacent and distant blocks into this block may be indicative of a specific environmental features or human activities such as the breeding or trading of guinea pigs (Kim 2013) that may have increased the immigration rate or establishment probabilities of *T. cruzi* in this block.

The source of the immigrants that establish in each block are no more likely to be from proximal blocks than from distant blocks in other districts. That is, the closest relatives of the lineages inhabiting each block are often found across Arequipa and not necessarily the closest block with an active infestation. The long-range *T. cruzi* gene flow seen here is likely facilitated by human movement as opposed to non-human mediated dispersal within a bug or host. *T. infestans* actively fly long distances (~1km) only under stressful conditions such as starvation, which may be unlikely in human- and animal-occupied houses (Ceballos *et al.* 2005; Richer *et al.* 2007). Individual *T. infestans* could not actively travel the large distances between some districts. Active dispersal by dogs and other small mammals may facilitate some *T. cruzi* dispersal, however, humans likely provide the primary mechanism of *T. cruzi* dispersal. Human movement facilitates

passive transportation of *T. infestans* and reservoir species of *T. cruzi* (Pinto Dias 2013) among blocks within and between districts.

The data presented here show that urban areas are a patchwork of city blocks where within-patch gene flow occurs regularly but where city streets hinder non-human mediated gene flow between patches. City streets do not prevent human movement between patches (Noulas *et al.* 2012), facilitating gene flow between distant patches independent of distance. Such long-range dispersal across the urban ecosystem may be possible for any species capable of exploiting human activity as a mechanism for migration. Without human-mediated migration, a T. cruzi population may be confined to a single patch with infrequent dispersal across the barriers between patches. These results highlight the difficulty of controlling the dispersal of parasites like *T. cruzi*. Eliminating the vector from a city block containing an established *T. cruzi* population would not eliminate the risk of *T. cruzi* dispersal to a distant block. Long-term, city-wide vector control campaigns and continuous vector surveillance activities may be required to prevent establishment of *T. cruzi* in new blocks.

Figures and Tables
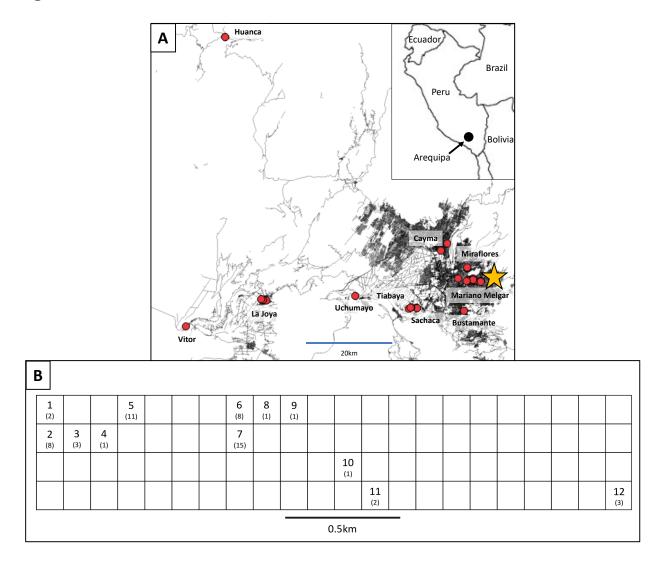
Figure 1. Map of sample collection locations in A) Arequipa, Peru. All 123 *T. cruzi*

samples used in this analysis were collected within 100km of Arequipa, Peru from 2008-

2015. Sampled *T. cruzi* are represented by red dots. Lines represent major roadways.

Densely populated areas appear grey due to the density of roads. Mariano Melgar—the

city center and most densely-sampled district— is indicated with a star. Inset shows the

location of Arequipa in southern Peru. **B) Mariano Melgar district.** 56 samples were

collected from the Mariano Melgar district in Arequipa (star in A). Blocks are displayed

in a grid to maintain privacy. Blocks from which samples were collected are numbered 1-

12. Number of samples collected per block are displayed in parentheses. No parasites were collected in unlabeled blocks. Black lines represent streets separating city blocks.
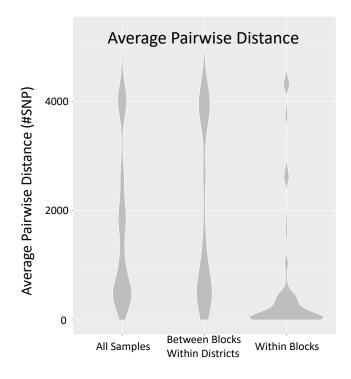
Figure 2. *T. cruzi* **isolates collected from the same block are genetically similar.**

Violin plots show the number of pairwise SNP difference between all pairs of 123

samples, among pairs of samples collected in different blocks from the same district, and

among pairs of samples collected in the same block. The distribution of genetic distances

among samples between blocks among districts is significantly larger than the

distribution among all samples and the genetic distances among samples within blocks

are significantly smaller than expected given the genetic diversity in the dataset

(p<0.001). These data show that genetically-similar samples cluster in blocks while the

overall diversity is distributed among districts.

0.47

Block 6

0.78

Block 5

0.38

Blocks 1-4

0.23

0.62

Block 6

Block 6

Block 7

500 SNPs
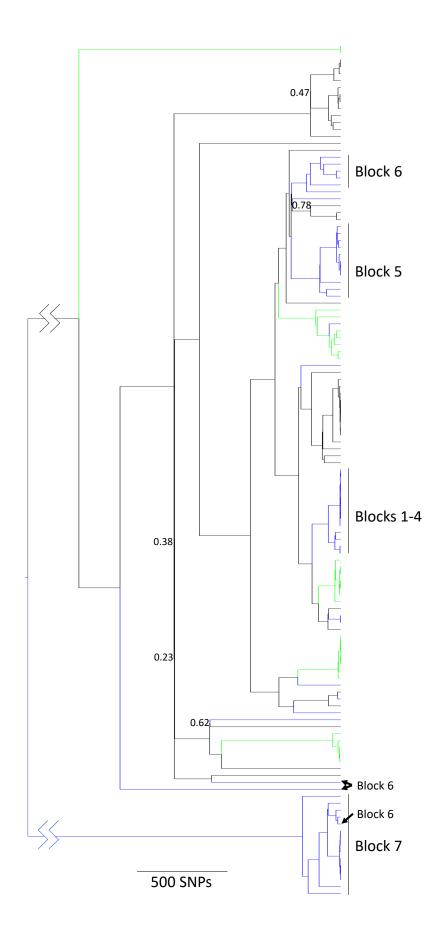
Figure 3. Maximum clade credibility phylogeny shows that the *T. cruzi* collected from many blocks form a monophyly. The phylogeny includes all 123 samples collected from Arequipa. Branches are colored for samples collected in the two most densely-sampled districts: La Joya (green) and Mariano Melgar (blue). Block numbers are labeled for the 48 samples from Blocks 1-7 in Mariano Melgar. Black branches represent samples collected from eight other locations across Arequipa: Bustamante, Cayma, Huanca, Miraflores, Sachaca, Tiabaya, Uchumayo, and Vitor. *T. cruzi* from both Mariano Melgar and La Joya span the phylogeny and confirm that most of the genetic variation in all of Arequipa is contained within a single district. The *T. cruzi* in many blocks within Mariano Melgar were likely introduced from outside of the district, since the MRCA for each block is typically shared with other districts and not from other blocks within Mariano Melgar. Jagged lines in most basal branch indicate that the branch was shortened for visualization. Posterior probabilities less than 0.80 are labeled for basal nodes. An unabridged phylogeny containing all tip labels and posterior probabilities can be found in Supplemental Figure 5.
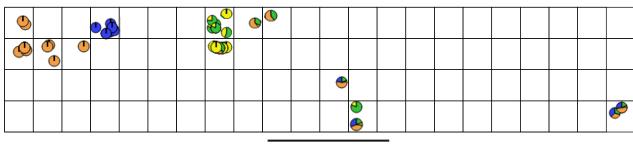
Table 1. Population Genetic Statistics

	Avg. Pairwise distance	No. Segregating Sites	π	θ
123 *T. cruzi* isolates	1761.96	12256	0.14374	0.185659

Table 2. AMOVA across Arequipan *T. cruzi* isolates shows no significant population structure between districts but significant structure within Mariano Melgar district.
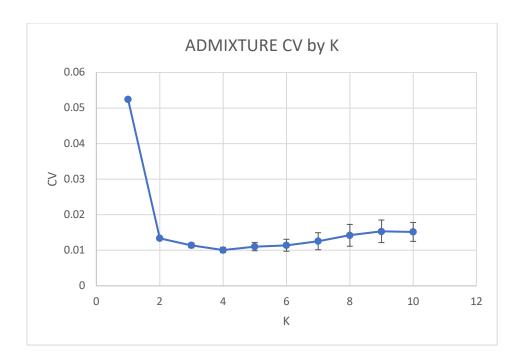
Test	P-value	% total variation
Variation within blocks	0.001	27.828949
Variation among blocks within districts	0.001	67.370526
Variations between districts	0.41	4.800524
Variation within blocks in Mariano Melgar	N/A	27.44506
Variation between blocks in Mariano Melgar	0.001	72.55494

Supplemental Materials
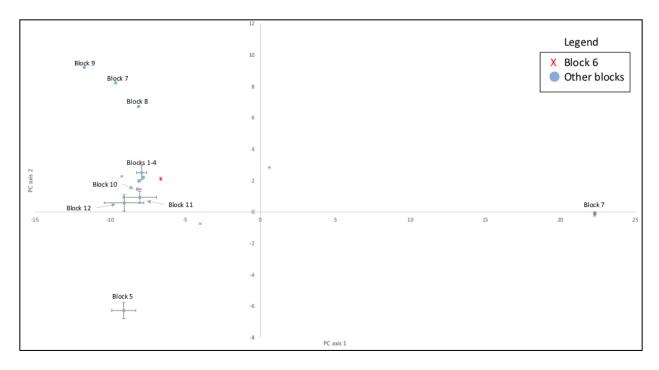
Supplemental Figure 1. ADMIXTURE analysis for K=4 genetic clusters shows that similar genotypes tend to cluster within a block. The colors of each pie chart represent the likelihood that a sample belongs to each of four genetic clusters. Each color (blue, yellow, orange, grey) represent a unique genetic cluster. There is significant clustering of identical genotypes within blocks 1-4, 5, and 7. Block 6 contains an exceptionally diverse population of *T. cruzi*. Block 8-12 have few samples per block, but the samples collected in each block are distinct from those collected elsewhere in the district. Blocks are displayed in a grid to maintain privacy.

Supplemental Figure 2 Cross-validation scores 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.
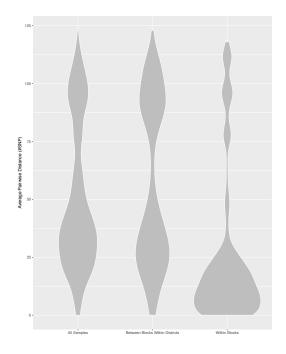
Supplemental Figure 3. Principal component analysis shows genetic similarity of

***T. cruzi* collected from the same city block within the Mariano Melgar district.**

PCA was calculated using 56 samples. The centroids of all samples from each city

block with standard error bars are shown here. Each of the seven samples collected in

Block 6 are represented by a red X because they occupy disparate portions of the

PCA space. One sample collected in Block 7 is unique and is thus represented by its

own dot.

Supplemental Table 1. Using only presence/absence for each of 474 recombination events in place of genotype data, AMOVA performed across Arequipan *T. cruzi* isolates shows significant structure within blocks.

Test	P-value	% total variation
Variation within blocks	0.001	37.131922
Variation among blocks within districts	0.001	62.5097975
Variations between districts	0.477	0.3582805
Variation within blocks in Mariano Melgar	N/A	34.94804
Variation between blocks in Mariano Melgar	0.001	65.05196

Supplemental Figure 4. Using only presence/absence for each of 474 recombination events in place of genotype data, the *T. cruzi* isolates collected from the same block are genetically highly similar. Violin plots show the number of pairwise differences in unique recombination events. Difference between all pairs of 123 samples, among pairs of samples collected in different blocks from the same district, and among pairs of samples collected in the same block, are shown. These results recapitulate the results found using genome-wide SNP data: The distribution of differences among samples between blocks among districts is larger than the distribution among all samples, the differences among samples within blocks are significantly smaller than expected given the diversity in the dataset (p<0.001). The results that genetically-similar samples cluster in blocks while the overall diversity is distributed among districts do not vary regardless of the data set analyzed (recombination events or genome-wide SNP data).

500 SNPs

Supplemental Figure 5. Maximum clade credibility phylogenetic reconstruction with sample labels and all posterior probabilities labeled. Collection locations (district and block number within district) and dates for each sample can be found in Supplemental Table 2.

Supplemental Table 2. Sample collection locations and years.

Sample ID	District	Block	Year Collected
TC001	La Joya	4	2008
TC002	La Joya	4	2008
TC003	La Joya	5	2008
TC004	La Joya	6	2008
TC010	Mariano Melgar	5	2008
TC014	La Joya	3	2008
TC015	La Joya	7	2008
TC016	La Joya	4	2008
TC019	La Joya	6	2008
TC020	La Joya	4	2008
TC022	La Joya	2	2008
TC023	La Joya	9	2008
TC026	La Joya	4	2008
TC027	La Joya	6	2008
TC029	La Joya	6	2008
TC031	La Joya	9	2008
TC033	La Joya	6	2008
TC034	La Joya	8	2008
TC035	La Joya	9	2008
TC036	La Joya	9	2008

TC037	La Joya	6	2008
TC038	La Joya	9	2008
TC039	La Joya	9	2008
TC040	La Joya	1	2008
TC041	Mariano Melgar	5	2010
TC042	Mariano Melgar	2	2010
TC043	Mariano Melgar	2	2010
TC044	Mariano Melgar	6	2010
TC045	Mariano Melgar	2	2010
TC046	Mariano Melgar	7	2010
TC047	Mariano Melgar	7	2010
TC048	Mariano Melgar	7	2010
TC049	Mariano Melgar	5	2010
TC051	Mira	1	2010
TC053	Tiabaya	1	2010
TC055	Mariano Melgar	6	2010
TC057	Tiabaya	1	2010
TC058	Tiabaya	1	2010
TC061	Mariano Melgar	2	2010
TC063	Uchumayo	1	2010
TC064	Mariano Melgar	4	2010
TC065	Mariano Melgar	7	2011
TC067	Sachaca	1	2011

TC068	Mariano Melgar	12	2011
TC069	Mariano Melgar	12	2011
TC070	Mariano Melgar	2	2011
TC071	Mariano Melgar	1	2011
TC072	Mariano Melgar	3	2011
TC073	Mariano Melgar	2	2011
TC074	Mariano Melgar	3	2011
TC075	Mariano Melgar	2	2011
TC076	Mariano Melgar	5	2011
TC077	Mariano Melgar	5	2011
TC078	Mariano Melgar	5	2011
TC079	Mariano Melgar	2	2011
TC080	Mariano Melgar	5	2011
TC081	Mariano Melgar	7	2011
TC082	Mariano Melgar	7	2011
TC083	Mariano Melgar	7	2011
TC084	Mariano Melgar	7	2011
TC085	Mariano Melgar	7	2011
TC086	Mariano Melgar	7	2011
TC088	Mariano Melgar	6	2011
TC089	Mariano Melgar	7	2011
TC090	Mariano Melgar	6	2011
TC091	Mariano Melgar	7	2011

TC092	Mariano Melgar	7	2011
TC095	Mariano Melgar	12	2011
TC097	Mariano Melgar	6	2011
TC098	Mariano Melgar	8	2011
TC099	Mariano Melgar	6	2011
TC100	Mariano Melgar	9	2011
TC101	Mariano Melgar	11	2011
TC102	Mariano Melgar	6	2011
TC103	Mariano Melgar	7	2011
TC104	Mariano Melgar	5	2011
TC105	Mariano Melgar	3	2011
TC106	Sachaca	1	2011
TC107	Mariano Melgar	6	2011
TC108	Mariano Melgar	11	2011
TC110	Mariano Melgar	10	2011
TC111	Mariano Melgar	5	2011
TC112	Mariano Melgar	5	2011
TC113	Sachaca	2	2011
TC114	Sachaca	2	2011
TC115	Mariano Melgar	1	2011
TC116	Mariano Melgar	5	2011
TC119	Cayma	1	2011
TC120	Uchumayo	1	2011

TC122	Cayma	2	2011
TC123	Sachaca	1	2012
TC124	Mariano Melgar	7	2012
TC125	Bustamante	1	2012
TC126	Bustamante	1	2012
TC127	Tiabaya	5	2012
TC129	La Joya	1	2013
TC130	La Joya	1	2013
TC131	La Joya	1	2013
TC132	Mira	3	2013
TC133	Mira	2	2013
TC134	Mira	2	2013
TC135	Murco	N/A	2013
TC136	Mira	4	2014
TC137	Mira	4	2014
TC138	Vitor	1	2013
TC139	Tiabaya	4	2015
TC140	Tiabaya	2	2015
TC141	Huanca	N/A	2015
TC142	Huanca	N/A	2015
TC143	Huanca	N/A	2015
TC144	Huanca	N/A	2015
TC145	Huanca	N/A	2015

TC146	Huanca	N/A	2015
TC147	La Joya	1	2015
TC148	Huanca	N/A	2015
TC149	Huanca	N/A	2015
TC150	Huanca	N/A	2015
TC151	Huanca	N/A	2015
TC152	Huanca	N/A	2015
TC153	Tiabaya	3	2015
TC154	Tiabaya	3	2015
TC155	La Joya	1	2015
TC156	Tiabaya	3	2015

Literature Cited

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664. http://doi.org/10.1101/gr.094052.109

Angold, P. G., Sadler, J. P., Hill, M. O., Pullin, A., Rushton, S., Austin, K., et al. (2006). Biodiversity in urban habitat patches. *Science of the Total Environment*, *360*(1-3), 196–204.

Ayres, D. L., Darling, A., Zwickl, D. J., Beerli, P., Holder, M. T., Lewis, P. O., et al. (2012). BEAGLE: An Application Programming Interface and High-Performance Computing Library for Statistical Phylogenetics. *Systematic Biology*, *61*(1), 170–173. http://doi.org/10.1093/sysbio/syr100

Barbu, C. M., Hong, A., Manne, J. M., Small, D. S., Calderón, J. E. Q., Sethuraman, K., et al. (2013). The effects of city streets on an urban disease vector. *PLoS Computational Biology*, *9*(1), e1002801.

Beninde, J., Veith, M., & Hochkirch, A. (2015). Biodiversity in cities needs space: a meta-analysis of factors determining intra-urban biodiversity variation. *Ecology Letters*, *18*(6), 581–592. http://doi.org/10.1111/ele.12427

Blackburn, T. M., Cassey, P., Duncan, R. P., Evans, K. L., & Gaston, K. J. (2004). Avian extinction and mammalian introductions on oceanic islands. *Science*, *305*(5692), 1955–1958.

Blair, R. B. (2001). Birds and butterflies along urban gradients in two ecoregions of the United States: Is urbanization creating a homogeneous fauna? *Biotic Homogenization*, 33–56.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., et al. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology*, *10*(4), e1003537. http://doi.org/10.1371/journal.pcbi.1003537

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*(19), 2633–2635. http://doi.org/10.1093/bioinformatics/btm308

Castillo Neyra, R., Borrini-Mayorí, K., Salazar Sánchez, R., Ancca Suarez, J., Xie, S., Náquira Velarde, C., & Levy, M. Z. (2016). Heterogeneous infectiousness in guinea pigs experimentally infected with Trypanosoma cruzi. *Parasitology International*, *65*(1), 50–54.

Ceballos, L. A., Vazquez-Prokopec, G. M., Cecere, M. C., Marcet, P. L., & Gürtler, R. E. (2005). Feeding rates, nutritional status and flight dispersal potential of peridomestic populations of Triatomainfestans in rural northwestern Argentina. *Acta Tropica*, *95*(2), 149–159.

Cheptou, P. O., Carrue, O., Rouifed, S., & Cantarel, A. (2008). Rapid evolution of seed dispersal in an urban environment in the weed Crepis sancta. *Proceedings of the National Academy of Sciences*, *105*(10), 3796–3799.

Coffield, D. J., Jr, Spagnuolo, A. M., Shillor, M., Mema, E., Pell, B., Pruzinsky, A., & Zetye, A. (2013). A model for Chagas disease with oral and congenital transmission. *PLoS ONE*, *8*(6), e67267.

Cohen, J. E., & Gürtler, R. E. (2001). Modeling household transmission of American trypanosomiasis. *Science*, *293*(5530), 694–698. http://doi.org/10.1126/science.1060638

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. http://doi.org/10.1093/bioinformatics/btr330

Denys, C., & Schmidt, H. (1998). Insect communities on experimental mugwort (Artemisia vulgaris L.) plots along an urban gradient. *Oecologia*, *113*(2), 269–277.

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, *43*(5), 491–498. http://doi.org/10.1038/ng.806

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, *29*(8), 1969–1973. http://doi.org/10.1093/molbev/mss075

Hahs, A. K., McDonnell, M. J., McCarthy, M. A., Vesk, P. A., Corlett, R. T., Norton, B. A., et al. (2009). A global synthesis of plant extinction rates in urban areas. *Ecology Letters*, *12*(11), 1165–1173.

Heled, J., & Drummond, A. J. (2008). Bayesian inference of population size history from multiple loci. *BMC Evolutionary Biology*, *8*(1), 289–15. http://doi.org/10.1186/1471-2148-8-289

Herrer, A. (1955). Trypanosomiasis americana en el Perú: III. Importancia del cobayo como reservorio de la enfermedad de Chagas en la región sudoccidental. *Revista Peruana De Medicina Experimental Y Salud Publica*, *9*(1-2), 45–55.

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23*(14), 1801–1806. http://doi.org/10.1093/bioinformatics/btm233

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, *2*, e281.

Khatchikian, C. E., Foley, E. A., Barbu, C. M., Hwang, J., Ancca Juárez, J., Borrini-Mayorí, K., et al. (2015). Population Structure of the Chagas Disease Vector Triatoma infestans in an Urban Environment. *PLOS Neglected Tropical Diseases*, *9*(2), e0003425–12. http://doi.org/10.1371/journal.pntd.0003425

Kim, E. (2013, July 26). *Modeling Domestic Guinea Pigs and Trypanosoma cruzi Persistence in Arequipa, Peru*. (S. Levin, Ed.). Princeton, NJ. Retrieved from http://arks.princeton.edu/ark:/88435/dsp01gb19f592r

Kowarik, I. (1995). On the Role of Alien Species in Urban Flora and Vegetation. *Plant Invasions-General Aspects and Special Problems*, 85–103.

Kowarik, I. (2011). Novel urban ecosystems, biodiversity, and conservation. *Environmental Pollution*, *159*(8-9), 1974–1983.

Kribs-Zaleta, C. M., & Mubayi, A. (2012). The role of adaptations in two-strain competition for sylvatic Trypanosoma cruzi transmission. *Journal of Biological Dynamics*, *6*(2), 813–835.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359.

LaPoint, S., Balkenhol, N., Hale, J., Sadler, J., & van der Ree, R. (2015). Ecological connectivity research in urban areas. *Functional Ecology*, *29*(7), 868–878. http://doi.org/10.1111/1365-2435.12489

Levy, M. Z., Barbu, C. M., Castillo Neyra, R., Quispe-Machaca, V. R., Ancca Juárez, J., Escalante-Mejia, P., et al. (2014). Urbanization, land tenure security and vector-borne Chagas disease. *Proceedings. Biological Sciences*, *281*(1789), 20141003–20141003. http://doi.org/10.1098/rspb.2014.1003

Levy, M. Z., Bowman, N. M., Kawai, V., Waller, L. A., Del Carpio, J. G. C., Benzaquen, E. C., et al. (2006). Periurban Trypanosoma cruzi–infected Triatoma infestans, Arequipa, Peru. *Emerging Infectious Diseases*, *12*(9), 1345.

Levy, M. Z., Quispe-Machaca, V. R., Ylla-Velasquez, J. L., Waller, L. A., Richards, J. M., Rath, B., et al. (2008). Impregnated netting slows infestation by Triatoma infestans. *The American Journal of Tropical Medicine and Hygiene*, *79*(4), 528–534.

Levy, M. Z., Tustin, A., Castillo Neyra, R., Mabud, T. S., Levy, K., Barbu, C. M., et al. (2015). Bottlenecks in domestic animal populations can facilitate the emergence of Trypanosoma cruzi, the aetiological agent of Chagas disease. *Proc. R. Soc. B*, *282*(1810), 20142807.

MacArthur, R. H., & Wilson, E. O. (2001). The theory of island biogeography, *1*.

Mackin-Rogalska, R., Pinowski, J., Solon, J., & Wojcik, Z. (1988). Changes in vegetation, avifauna, and small mammals in a suburban habitat. *Polish Ecological Studies*, *14*(1-2), 293–330.

McIntyre, N. E. (2000). Ecology of Urban Arthropods: A Review and a Call to Action. *Annals of the Entomological Society of America*, *93*(4), 825–835. http://doi.org/10.1603/0013-8746(2000)093[0825:EOUAAR]2.0.CO;2

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303. http://doi.org/10.1101/gr.107524.110

McKinney, M. L. (2002). Urbanization, Biodiversity, and Conservation. *BioScience*, *52*(10), 883–890. http://doi.org/10.1641/0006-3568(2002)052[0883:UBAC]2.0.CO;2

Niemelä, J., & Kotze, D. J. (2009). Carabid beetle assemblages along urban to rural gradients: a review. *Landscape and Urban Planning*, *92*(2), 65–71.

Noulas, A., Scellato, S., Lambiotte, R., Pontil, M., & Mascolo, C. (2012). A tale of many cities: universal patterns in human urban mobility. *PLoS ONE*, *7*(5), e37027.

Pinto Dias, J. C. (2013). Human chagas disease and migration in the context of globalization: some particular aspects. *Journal of Tropical Medicine*, *2013*.

Poplin, R., Ruano-Rubio, V., DePristo, M. A., Fennell, T. J., Carneiro, M. O., Van der Auwera, G. A., et al. (2017). Scaling accurate genetic variant discovery to tens of thousands of samples, 1–22. http://doi.org/10.1101/201178

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based

Linkage Analyses. *The American Journal of Human Genetics*, *81*(3), 559–575. http://doi.org/10.1086/519795

R Core Team. (2017). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*. Retrieved from https://www.R-project.org/

Rabinovich, J. E., Wisnivesky-Colli, C., Solarz, N. D., & Gürtler, R. E. (1990). Probability of transmission of Chagas disease by Triatoma infestans (Hemiptera: Reduviidae) in an endemic area of Santiago del Estero, Argentina. *Bulletin of the World Health Organization*, *68*(6), 737.

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., Suchard, M. A., & Susko, E. (2018). Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic Biology*, *22*, 341–syy032 VL – IS –. http://doi.org/10.1093/sysbio/syy032

Richer, W., Kengne, P., Cortez, M. R., Perrineau, M. M., Cohuet, A., Fontenille, D., & Noireau, F. (2007). Active dispersal by wild Triatoma infestans in the Bolivian Andes. *Tropical Medicine & International Health*, *12*(6), 759–764. http://doi.org/10.1111/j.1365-3156.2007.01846.x

Suchard, M. A., & Rambaut, A. (2009). Many-core algorithms for statistical phylogenetics. *Bioinformatics*, *25*(11), 1370–1376. http://doi.org/10.1093/bioinformatics/btp244

Theoharides, K. A., & Dukes, J. S. (2007). Plant invasion across space and time: factors affecting nonindigenous species success during four stages of invasion. *New Phytologist*, *176*(2), 256–273. http://doi.org/10.1111/j.1469-8137.2007.02207.x

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., et al. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Current Protocols in Bioinformatics*, *43*, 11.10.1–33. http://doi.org/10.1002/0471250953.bi1110s43

Wickham, H. (2016). ggplot2: elegant graphics for data analysis. http://doi.org/10.1007/978-3-319-24277-4

Wood, J. R., Alcover, J. A., Blackburn, T. M., Bover, P., Duncan, R. P., Hume, J. P., et al. (2017). Island extinctions: processes, patterns, and potential for ecosystem restoration. *Environmental Conservation*, *44*(4), 348–358.

Chapter 3: Evidence of sexual reproduction in a natural *Trypanosoma cruzi* population

Abstract

Sexual reproduction confers several evolutionary advantages for microparasites including an increased diversity of immune evasion genes within an individual. Reassortment of alleles among individuals is often essential to exploit novel hosts or populations. However, many eukaryotic microparasites exhibit highly-clonal population structures suggesting that these species rarely engage in genetic exchange through sexual reproduction. Evidence of clonality is particularly convincing in the causative agent of Chagas disease, *Trypanosoma cruzi*, despite equally convincing evidence of the capability for sex. 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 genome sequences of 123 isolates from the *T. cruzi* population in Arequipa, Peru. The distribution of polymorphic markers within and among isolates provides clear evidence of the occurrence of sexual reproduction, albeit with high levels of inbreeding. Sexual reproduction has resulted in the reassortment of large genetic segments among chromosomes due to crossing over during meiosis. These cross over events, as well as independent assortment of chromosomes, during sexual reproduction have resulted in a decay in the genetic linkage among polymorphic markers. Despite clear evidence of sexual reproduction, the population structure remains clonal due to a high level of inbreeding that increases homozygosity, and thus reduces diversity, within each inbreeding lineage. These results effectively reconcile the apparent contradiction by demonstrating that the clonal

population structure is derived not from the absence of sex but from high levels of inbreeding. We discuss epidemiological consequences of reproductive strategy on the genome evolution, population structure, and phenotypic diversity of this medically important parasite.

Introduction

An increasing body of evidence suggests that many eukaryotic microparasites exhibit highly-clonal population structures (Kilgour *et al.* 1975; Gibson *et al.* 1980; Mehlitz *et al.* 1982; Sargeaunt *et al.* 1982; Gibson & Gashumba 1983; Allsopp & Newton 1985; Gibson & Wellde 1985; Nash *et al.* 1985; Lanotte *et al.* 1986; Maazoun *et al.* 1986; Moreno *et al.* 1986; Pratlong *et al.* 1986; Safrin *et al.* 1986; Cariou & Pernin 1987; Proctor *et al.* 1987; Dardé *et al.* 1988; Gashumba *et al.* 1988; Meloni *et al.* 1988; Nadler & Honigberg 1988; Desjeux & Dedet 1989; Lehmann *et al.* 1989; Meloni *et al.* 1989; Dardé *et al.* 1990; Mihok *et al.* 1990; Tibayrenc *et al.* 1990; Tibayrenc *et al.* 1991; Brandt *et al.* 1993; Pujol *et al.* 1993; Ayala 1998). That is, isolates sampled in nature tend to represent independently evolving asexual lineages with limited genetic exchange between individuals. 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 or experimentally demonstrated in several species (Babiker & Walliker 1997; Guttery *et al.* 2012; Peacock *et al.* 2014). Paradoxically, sexual reproduction would be expected to provide several advantages to parasites including increasing within-individual diversity in immune evasion genes where diversity is paramount for 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,* which exhibits highly clonal population structures in nature (Oliveira *et al.* 1998; Diosque *et al.* 2003; Tibayrenc 2003) despite experimental evidence of the capacity for sexual reproduction

(Gaunt *et al.* 2003). First, meiosis and fertilization may be exceedingly rare such that mutational processes occur much more frequently than meiosis (Fig 1) and second, sexual reproduction may be 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 (Ayala 1998). 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 (Oliveira *et al.* 1998; Diosque *et al.* 2003; Tibayrenc 2003). For example, Oliveira *et al.* (1998) found that homozygosity at eight microsatellite loci among 24 *T. cruzi* strains was significantly greater than expected assuming random mating, and linkage disequilibrium was significantly higher than expected, suggesting a limited role for sexual reproduction (Ayala 1998). Nevertheless, the capacity for sexual reproduction in *T. cruzi* has been demonstrated in controlled experiments as well as by observations of hybrids that have formed through sexual outcrossing. Meiosis and fertilization in *T. cruzi* was detected *in vitro* by transfecting clones with two different markers and observing progeny with both markers (Gaunt *et al.* 2003). Further, hybridization between distantly-related lineages resulted in the CL Brener reference strain (Brisse *et al.* 1998; Brisse *et al.* 2003; Westenberger *et al.* 2005; El-Sayed *et al.* 2005). Subsequent meiosis in the evolutionary history of this strain has resulted in the transfer of large genomic regions among the divergent chromosomes (Machado & Ayala 2001).

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

are not observed in natural populations (Fig 1). 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. Similarly, sexual reproduction is expected to break down the non-random 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 (Tibayrenc *et al.* 1986; Oliveira *et al.* 1998; Tibayrenc & Ayala 1988). In the present study, we investigated the pattern of genomic signatures associated with the random assortment 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.

Methods

Sample Collection and Study Site

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. Uninfected *T. infestans* were allowed to feed on 6 of the 9 blood samples to 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 (Castillo-Neyra *et al.* 2016). *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. 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 (Bolger *et al.* 2014).

Whole genome assembly

Genomes were assembled using the most closely related reference genome, TcJR clone 4, obtained from TriTrypDB (http://tritrypdb.org/tritrypdb/), using bowtie2 (Langmead & Salzberg 2012). 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 28Mbp. This assembly largely excluded the extensive repeat regions found throughout the *T. cruzi* genome. Duplicate reads were removed using Picard MarkDuplicates (McKenna *et al.* 2010).

SNP calling

Individual gVCF files containing SNP data for each sample were generated using GATK HaplotypeCaller (McKenna *et al.* 2010; Poplin *et al.* 2017) following GATK's Best Practices procedure (De Pristo *et al.* 2011; Van der Auwera *et al.* 2013). 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.

Evidence of Sexual reproduction

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 (Danecek *et al.* 2011) and plink (Purcell *et al.* 2007). 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 a low density 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.

Two runs of homozygosity were considered identical by descent (IBD) if they were 99% identical and had ten or fewer non-overlapping homozygous SNPs. Relaxing these assumptions did not qualitatively alter the results nor the conclusions.

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, Hardy-Weinberg equilibrium calculations were performed for each polymorphic locus.

Results

The genome sequences from 123 *T. cruzi* isolates collected throughout Arequipa, Peru 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 (Chapter 1) 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 by gene flow 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 genomes sequenced. 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 the paternal chromosome and a set derived from the maternal 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). Thus, evidence of sexual reproduction was detected by identifying large segments (>10kb) of chromosomes where all of the markers (≥20) that are heterozygous in some isolates are homozygous in other isolates within the same population. These runs of homozygous chromosomal segments are common in the *T. cruzi* genomes in Arequipa (Fig 3). Using a conservative algorithm to detect chromosomal segments derived from meiotic recombination, we identified 474 recombined segments ranging in size from 10kb to 468kb (Supp Fig 1). Evidence of meiotic recombination events were detected on 69 of the 80 (86.25%) contigs longer than 100kb, suggesting that this phenomenon occurs on all *T. cruzi* chromosomes. The conservative criteria used to identify cross-over events are unlikely to be met on contigs shorter than 100kb (Supp Fig 2). Meiotic recombination and fertilization of gametes with random chromosomes accounts for 95% of the average pairwise differences among *T. cruzi* isolates, with the remaining differences explained by point mutations.

The frequency of heterozygotes is higher than expected in Hardy-Weinberg populations 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. Yet many populations of eukaryotic microparasites that have the capacity for sexual reproduction exhibit a clonal population structure suggesting that sexual reproduction is nearly absent in nature or possibly that sexual reproduction occurs but primarily with 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 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 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 (≥20) that are heterozygous in some isolates are homozygous in other isolates across a large chromosomal segment (>10kb). Sexual reproduction occurs with sufficient regularity to observe meiotic recombination events within subpopulations that have only recently established in blocks within the recently established city (Fig 3C). These 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. 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 clonal

reproduction, heterozygote frequencies in this population are higher than expected from a randomly-mating sexual population in 95% of polymorphic markers. 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.

High levels of inbreeding also increased the apparent clonality of the population. Inbreeding limits the exchange of diversity 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). For example, 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 always 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 in Arequipa; that is, all individuals in the population are closely related (Chapter 1). 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 (Lopez-Cancino *et al.* 2015) and vector (Oliveira *et al.* 1998; Dumonteil *et al.* 2018) 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 and host tissue), frequency, and mechanism affecting the probability of outcrossing in *T. cruzi* have not been elucidated, however *Trypanosoma brucei* produces gametes via meiosis in its vector (Peacock *et al.* 2014). By contrast, experimental evidence suggests that *T. cruzi* undergoes meiosis and fertilization in the mammalian host (Gaunt *et al.* 2003). 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.

Sexual reproduction in parasites like *T. cruzi* may be particularly important for maintaining a diverse assortment of immune evasion genes within individuals (Gjini *et al.* 2010; McCulloch *et al.* 2015). 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 (El-Sayed *et al.* 2005; Schenkman 2012; Weatherly *et al.* 2016). 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 (Freitas-Junior *et al.* 2000; Barry *et al.* 2007; Graves 2013). 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.

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.

The ubiquitous observations of clonal population structures have resulted in the hypothesis that sexual reproduction is nearly absent in natural *T. cruzi* populations (Brisse *et al.* 2003; de Freitas *et al.* 2006). The data presented here suggest that sexual recombination is not uncommon in natural populations. Sexual reproduction occurred repeatedly in the recent evolutionary history of the population and has occurred within subpopulations 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 not be detected and our criteria for identifying recombined regions was strict. Nevertheless, the extant linkage disequilibrium within large linkage blocks suggests that clonal 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 (Longhi *et al.* 2002; Verani *et al.* 2009; Abras *et al.* 2016) and has the potential to generate genetic and phenotypic diversity (Laurent *et al.* 1997; Zingales *et al.* 2012) in pathogenicity (Andrade *et al.* 1974; Andrade *et al.* 1985), host and vector propensity (de Lana *et al.* 1998; Roellig *et al.* 2010), and vulnerability to drugs (Revollo *et al.* 1998) 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.

Figures
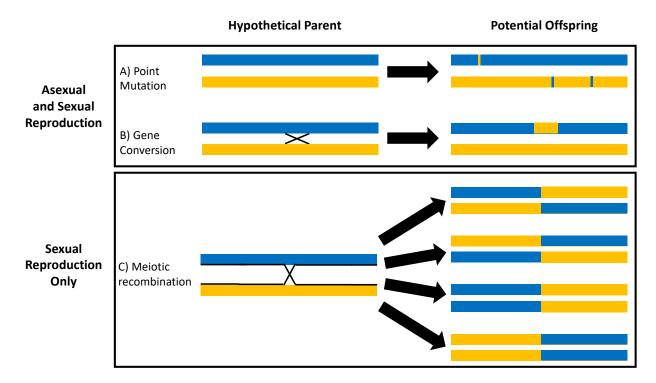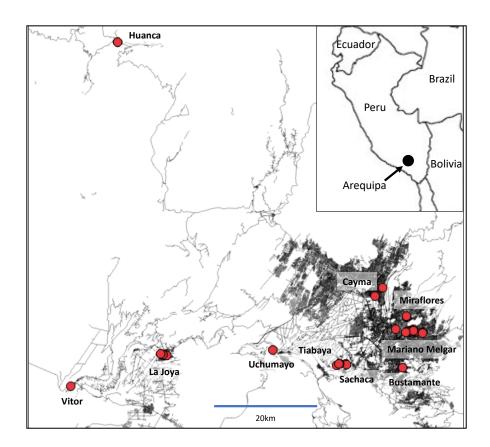
Figure 1. The reproductive strategies employed can be inferred from the distribution of polymorphic markers within and among strains of a population. Both asexual and sexual populations experience both A) point mutations and B) gene conversion events while only populations reproducing sexually experience 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 while these markers could become disassociated due to cross over events and the independent assortment of chromosomes during meiosis. **B)** Gene conversion events homogenize small regions of homologous

chromosomes (10s of kb), 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 paternal chromosome and a set derived from the maternal 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 that are heterozygous in their siblings.

Figure 2. Map of sample collection locations. The location of all 123 sampled *T. cruzi* are represented by red dots. Roads are represented by black lines. Inset shows the location of Arequipa.
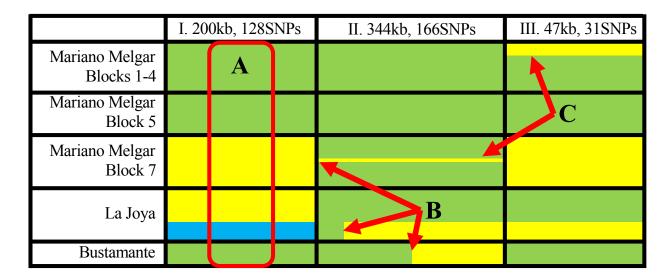
Figure 3. The distribution of polymorphic markers within and among individuals provides clear evidence of meiotic recombination. This representative subset of the data portrays the distributions of 325 polymorphic markers throughout three regions of the genome in 55 isolates (grouped by geographic location: Mariano Melgar Blocks 1-4 (N=14), Block 5 (N=11), Block 7 (N=14), La Joya (N=14), and Bustamante (N=2)) that indicate the presence and frequency 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). Note that only polymorphic sites with a minor allele frequency >16% in the population are represented in order to remove non-polymorphic sites and interspersed SNPs for clarity. **(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 decent. Note that 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 that crossing over can affect parts of chromosomes or whole chromosomes. In this region there is evidence of at least three independent meiotic recombination events. **(C)** Sexual reproduction after the colonization of a city block in the recent past 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 (Supplemental Fig 3) contain significant sequence similarity suggesting that this individual is closely related to other individuals in Block 7 with the exception of the meiotic event occurring in region II. Similarly, a meiotic recombination event in a region (III) of a 130kb contig occurred in one lineage that migrated from Block 2 to Blocks 1 and 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 whole genome data see Supplemental Figure 3.
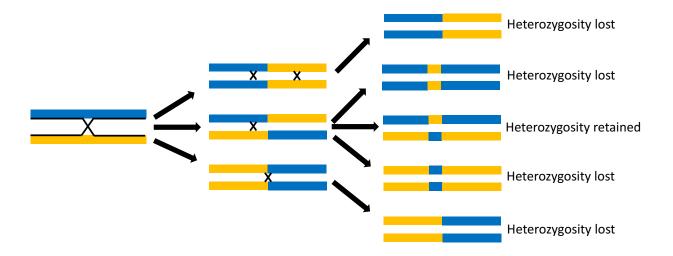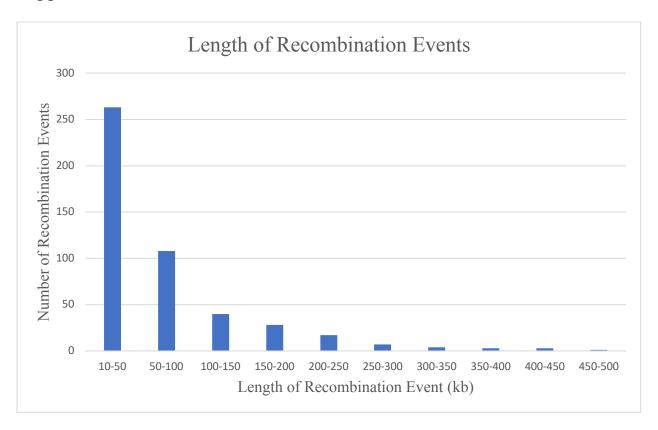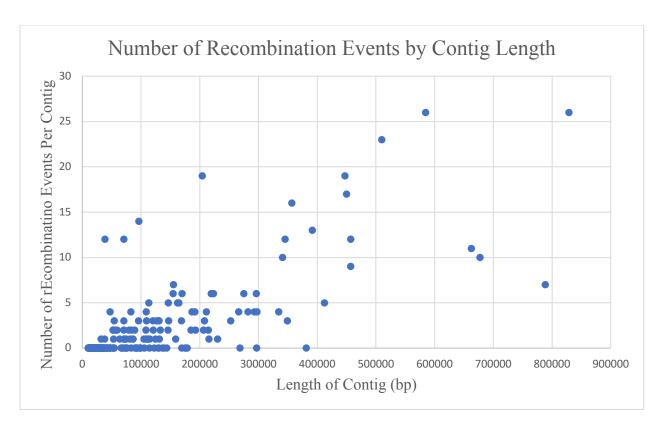
Figure 4. Inbreeding results in a decay of heterozygosity. After one generation of inbreeding, half of the potential offspring retain the heterozygosity found in the parent while half become homozygous. As diversity cannot be restored in homozygous regions in the absence of outcrossing and heterozygosity decays by half each generation, population-level heterozygosity continually declines in inbreeding populations until all lineages are homozygous.

Supplemental Materials
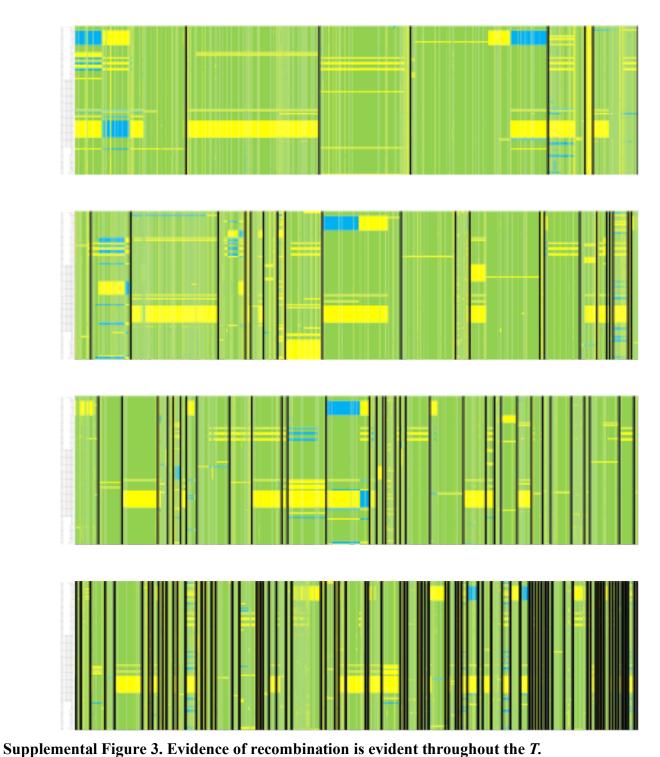
Supplemental Figure 1. Length of meiotic recombination events. 474 runs of

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.

Supplemental Figure 2. 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.

Supplemental Figure 3. 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

Supplemental Table 1. 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.

Supplemental Table 1. Collection locations and years for each *T. cruzi* sample.

Sample ID	District	Block	Year Collected
TC126	Bustamante	1	2012
TC125	Bustamante	1	2012
TC119	Cayma	1	2011
TC122	Cayma	2	2011
TC135	Huanca	N/A	2013
TC152	Huanca	N/A	2015
TC151	Huanca	N/A	2015
TC150	Huanca	N/A	2015
TC149	Huanca	N/A	2015
TC148	Huanca	N/A	2015
TC146	Huanca	N/A	2015
TC145	Huanca	N/A	2015
TC144	Huanca	N/A	2015
TC143	Huanca	N/A	2015
TC142	Huanca	N/A	2015
TC141	Huanca	N/A	2015
TC155	La Joya	1	2015
TC147	La Joya	1	2015
TC131	La Joya	1	2013
TC130	La Joya	1	2013

TC129	La Joya	1	2013
TC040	La Joya	1	2008
TC022	La Joya	2	2008
TC014	La Joya	3	2008
TC026	La Joya	4	2008
TC020	La Joya	4	2008
TC016	La Joya	4	2008
TC002	La Joya	4	2008
TC001	La Joya	4	2008
TC003	La Joya	5	2008
TC037	La Joya	6	2008
TC033	La Joya	6	2008
TC029	La Joya	6	2008
TC027	La Joya	6	2008
TC019	La Joya	6	2008
TC004	La Joya	6	2008
TC015	La Joya	7	2008
TC034	La Joya	8	2008
TC039	La Joya	9	2008
TC038	La Joya	9	2008
TC036	La Joya	9	2008
TC035	La Joya	9	2008
TC031	La Joya	9	2008

TC023	La Joya	9	2008
TC115	Mariano Melgar	1	2011
TC071	Mariano Melgar	1	2011
TC079	Mariano Melgar	2	2011
TC075	Mariano Melgar	2	2011
TC073	Mariano Melgar	2	2011
TC070	Mariano Melgar	2	2011
TC061	Mariano Melgar	2	2010
TC045	Mariano Melgar	2	2010
TC043	Mariano Melgar	2	2010
TC042	Mariano Melgar	2	2010
TC105	Mariano Melgar	3	2011
TC074	Mariano Melgar	3	2011
TC072	Mariano Melgar	3	2011
TC064	Mariano Melgar	4	2010
TC116	Mariano Melgar	5	2011
TC112	Mariano Melgar	5	2011
TC111	Mariano Melgar	5	2011
TC104	Mariano Melgar	5	2011
TC080	Mariano Melgar	5	2011
TC078	Mariano Melgar	5	2011
TC077	Mariano Melgar	5	2011
TC076	Mariano Melgar	5	2011

TC049	Mariano Melgar	5	2010
TC041	Mariano Melgar	5	2010
TC010	Mariano Melgar	5	2008
TC107	Mariano Melgar	6	2011
TC102	Mariano Melgar	6	2011
TC099	Mariano Melgar	6	2011
TC097	Mariano Melgar	6	2011
TC090	Mariano Melgar	6	2011
TC088	Mariano Melgar	6	2011
TC055	Mariano Melgar	6	2010
TC044	Mariano Melgar	6	2010
TC124	Mariano Melgar	7	2012
TC103	Mariano Melgar	7	2011
TC092	Mariano Melgar	7	2011
TC091	Mariano Melgar	7	2011
TC089	Mariano Melgar	7	2011
TC086	Mariano Melgar	7	2011
TC085	Mariano Melgar	7	2011
TC084	Mariano Melgar	7	2011
TC083	Mariano Melgar	7	2011
TC082	Mariano Melgar	7	2011
TC081	Mariano Melgar	7	2011
TC065	Mariano Melgar	7	2011

TC048	Mariano Melgar	7	2010
TC047	Mariano Melgar	7	2010
TC046	Mariano Melgar	7	2010
TC098	Mariano Melgar	8	2011
TC100	Mariano Melgar	9	2011
TC110	Mariano Melgar	10	2011
TC108	Mariano Melgar	11	2011
TC101	Mariano Melgar	11	2011
TC095	Mariano Melgar	12	2011
TC069	Mariano Melgar	12	2011
TC068	Mariano Melgar	12	2011
TC051	Miraflores	1	2010
TC134	Miraflores	2	2013
TC133	Miraflores	2	2013
TC132	Miraflores	3	2013
TC137	Miraflores	4	2014
TC136	Miraflores	4	2014
TC123	Sachaca	1	2012
TC106	Sachaca	1	2011
TC067	Sachaca	1	2011
TC114	Sachaca	2	2011
TC113	Sachaca	2	2011
TC058	Tiabaya	1	2010

TC057	Tiabaya	1	2010
TC053	Tiabaya	1	2010
TC140	Tiabaya	2	2015
TC156	Tiabaya	3	2015
TC154	Tiabaya	3	2015
TC153	Tiabaya	3	2015
TC139	Tiabaya	4	2015
TC127	Tiabaya	5	2012
TC120	Uchumayo	1	2011
TC063	Uchumayo	1	2010
TC138	Vitor	1	2013

Literature Cited

Abras, A., Gállego, M., Llovet, T., Tebar, S., Herrero, M., Berenguer, P., et al. (2016). Serological Diagnosis of Chronic Chagas Disease: Is It Time for a Change? *Journal of Clinical Microbiology*, *54*(6), 1566–1572. http://doi.org/10.1128/JCM.00142-16

Allsopp, B. A., & Newton, S. D. (1985). Characterization of Trypanosoma (Duttonella) vivax by isoenzyme analysis. *International Journal for Parasitology*, *15*(3), 265–270.

Andrade, S. G. (1974). Caracterização de cepas do Trypanosoma cruzi isoladas no Recôncavo Baiano. *Revista De Patologia Tropical*, *3*, 65–121.

Andrade, V., Barral-Netto, M., & Andrade, S. G. (1985). 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*, *18*(4), 499–506.

Ayala, F. J. (1998). Is sex better? Parasites say "no." *Proceedings of the National Academy of Sciences of the United States of America*, *95*(7), 3346–3348.

Babiker, H. A., & Walliker, D. (1997). Current views on the population structure of Plasmodium falciparum: implications for control. *Parasitology Today*, *13*(7), 262–267.

Barry, A. E., Leliwa-Sytek, A., Tavul, L., Imrie, H., Migot-Nabias, F., Brown, S. M., et al. (2007). Population genomics of the immune evasion (var) genes of Plasmodium falciparum. *PLOS Pathogens*, *3*(3), e34.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120.

Brandt, M. E., Bragg, S. L., & Pinner, R. W. (1993). Multilocus enzyme typing of Cryptococcus neoformans. *Journal of Clinical Microbiology*, *31*(10), 2819–2823.

Brisse, S., Barnabé, C., Banuls, A.-L., Sidibé, I., Noël, S., & Tibayrenc, M. (1998). A phylogenetic analysis of the Trypanosoma cruzi genome project CL Brener reference strain by multilocus enzyme electrophoresis and multiprimer random amplified polymorphic DNA fingerprinting. *Molecular and Biochemical Parasitology*, *92*(2), 253–263.

Brisse, S., Henriksson, J., Barnabé, C., Douzery, E. J. P., Berkvens, D., Serrano, M., De Carvalho, M. R. C., Buck, G. A., Dujardin, J.-C., & Tibayrenc, M. (2003a). Evidence for genetic exchange and hybridization in Trypanosoma cruzi based on nucleotide sequences and molecular karyotype. *Infection, Genetics and Evolution*, *2*(3), 173–183. http://doi.org/10.1016/S1567-1348(02)00097-7

Brisse, S., Henriksson, J., Barnabé, C., Douzery, E. J., Berkvens, D., Serrano, M., De Carvalho, M. R. C., Buck, G. A., Dujardin, J.-C., & Tibayrenc, M. (2003b). Evidence for genetic exchange and hybridization in Trypanosoma cruzi based on nucleotide sequences and molecular karyotype. *Infection, Genetics and Evolution*, *2*(3), 173–183.

Cariou, M. L., & Pernin, P. (1987). First evidence for diploidy and genetic recombination in free-living amoebae of the genus Naegleria on the basis of electrophoretic variation. *Genetics*, *115*(2), 265–270.

Castillo Neyra, R., Borrini-Mayorí, K., Salazar Sánchez, R., Ancca Suarez, J., Xie, S., Náquira Velarde, C., & Levy, M. Z. (2016). Heterogeneous infectiousness in guinea

pigs experimentally infected with Trypanosoma cruzi. *Parasitology International*, *65*(1), 50–54.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. http://doi.org/10.1093/bioinformatics/btr330

Dardé, M. L., Bouteille, B., & Pestre-Alexandre, M. (1988). Isoenzymic characterization of seven strains of Toxoplasma gondii by isoelectrofocusing in polyacrylamide gels. *The American Journal of Tropical Medicine and Hygiene*, *39*(6), 551–558.

Dardé, M. L., Bouteille, B., & Pestre-Alexandre, M. (1990). Comparison of isoenzyme profiles ofToxoplasma gondii tachyzoites produced under different culture conditions. *Parasitology Research*, *76*(4), 367–371.

de Freitas, J. M., Augusto-Pinto, L., Pimenta, J. R., Bastos-Rodrigues, L., Gonçalves, V. F., Teixeira, S. M. R., et al. (2006). Ancestral Genomes, Sex, and the Population Structure of Trypanosoma cruzi. *PLOS Pathogens*, *2*(3), e24–10. http://doi.org/10.1371/journal.ppat.0020024

de Lana, M., da Silveira Pinto, A., Barnabé, C., Quesney, V., Noël, S., & Tibayrenc, M. (1998). Trypanosoma cruzi: compared vectorial transmissibility of three major clonal genotypes byTriatoma infestans. *Experimental Parasitology*, *90*(1), 20–25.

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, *43*(5), 491–498. http://doi.org/10.1038/ng.806

Desjeux, P., & Dedet, J. P. (1989). Isoenzyme characterization of 112 Leishmania isolates from French Guiana. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *83*(5), 610–612.

Diosque, P., Barnabé, C., Padilla, A. M., Marco, J. D., Cardozo, R. M., Cimino, R. O., et al. (2003). Multilocus enzyme electrophoresis analysis of Trypanosoma cruzi isolates from a geographically restricted endemic area for Chagas' disease in Argentina. *International Journal for Parasitology*, *33*(10), 997–1003. http://doi.org/10.1016/S0020-7519(03)00139-5

Dumonteil, E., Ramirez-Sierra, M.-J., rez-Carrillo, S. P. X., Teh-Poot, C., Herrera, C., re, S. X. B. G. X., & Waleckx, E. (2018). Detailed ecological associations of triatomines revealed by metabarcoding and next- generation sequencing: implications for triatomine behavior and Trypanosoma cruzi transmission cycles. *Scientific Reports*, 1–13. http://doi.org/10.1038/s41598-018-22455-x

El-Sayed, N. M., Myler, P. J., Bartholomeu, D. C., Nilsson, D., Aggarwal, G., Tran, A.-N., et al. (2005). The genome sequence of Trypanosoma cruzi, etiologic agent of Chagas disease. *Science*, *309*(5733), 409–415.

Freitas-Junior, L. H., Bottius, E., Pirrit, L. A., Deitsch, K. W., Scheidig, C., Guinet, F., et al. (2000). Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of P. falciparum. *Nature Publishing Group*, *407*(6807), 1018.

Gashumba, J. K., Baker, R. D., & Godfrey, D. G. (1988). Trypanosoma congolense: the distribution of enzymic variants in East and West Africa. *Parasitology*, *96*(3), 475–486.

Gaunt, M. W., Yeo, M., Frame, I. A., Stothard, J. R., Carrasco, H. J., Taylor, M. C., et al. (2003). Mechanism of genetic exchange in American trypanosomes. *Nature*, *421*(6926), 936–939. http://doi.org/10.1038/nature01438

Gibson, W. C., & Gashumba, J. K. (1983). Isoenzyme characterization of some Trypanozoon stocks from a recent trypanosomiasis epidemic in Uganda. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *77*(1), 114–118.

Gibson, W. C., & Wellde, B. T. (1985). Characterization of Trypanozoon stocks from the South Nyanza sleeping sickness focus in Western Kenya. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *79*(5), 671–676.

Gibson, W. C., Marshall, T. de C., & Godfrey, D. G. (1980). Numerical analysis of enzyme polymorphism: a new approach to the epidemiology and taxonomy of trypanosomes of the subgenus Trypanozoon, *18*, 175–246.

Gjini, E., Haydon, D. T., Barry, J. D., & Cobbold, C. A. (2010). Critical interplay between parasite differentiation, host immunity, and antigenic variation in trypanosome infections. *The American Naturalist*, *176*(4), 424–439.

Graves, C. J., Ros, V. I., Stevenson, B., Sniegowski, P. D., & Brisson, D. (2013). Natural selection promotes antigenic evolvability. *PLOS Pathogens*, *9*(11), e1003766.

Guttery, D. S., Holder, A. A., & Tewari, R. (2012). Sexual Development in Plasmodium: Lessons from Functional Analyses. *PLOS Pathogens*, *8*(1), e1002404–3. http://doi.org/10.1371/journal.ppat.1002404

Kilgour, V., Godfrey, D. G., & Na'isa, B. K. (1975). Isoenzymes of two aminotransferases among Trypanosoma vivax in Nigerian cattle. *Annals of Tropical Medicine & Parasitology*, *69*(3), 329–335.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359.

Lanotte, G., Rioux, J. A., & Serres, E. (1986). Leishmania: Taxonomie et Phylogénèse: Applications Eco-Epidémiologiques, Colloques Internationaux de Centre National de la Recherche Scientifique/Institut National de la Santé et de la Recherche Médicale. *Paris: CNRS/INSERM*, 27–40.

Laurent, J. P., Barnabe, C., Quesney, V., Noel, S., & Tibayrenc, M. (1997). Impact of clonal evolution on the biological diversity of Trypanosoma cruzi. *Parasitology*, *114*(3), 213–218.

Lehmann, P. F., Kemker, B. J., Hsiao, C. B., & Dev, S. (1989). Isoenzyme biotypes of Candida species. *Journal of Clinical Microbiology*, *27*(11), 2514–2521.

Longhi, S. A., Brandariz, S. B., Lafon, S. O., Niborski, L. L., Luquetti, A. O., Schijman, A. G., et al. (2012). 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*, *87*(2), 267–271.

López-Cancino, S. A., Tun-Ku, E., la Cruz-Felix, De, H. K., Ibarra-Cerdeña, C. N., Izeta-Alberdi, A., Pech-May, A., et al. (2015). Landscape ecology of Trypanosoma cruzi in the southern Yucatan Peninsula. *Acta Tropica*, *151*, 58–72. http://doi.org/10.1016/j.actatropica.2015.07.021

Maazoun, R., Pratlong, F., Lanotte, G., & Rioux, J. A. (1986). Leishmania: Taxonomie et Phylogénèse: Applications Eco-Epidémiologiques, Colloques Internationaux de Centre National de la Recherche Scientifique/Institut National de la Santé et de la Recherche Médicale. *Paris: CNRS/INSERM*, 119–129.

Machado, C. A., & Ayala, F. J. (2001). 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*, *98*(13), 7396–7401.

McCulloch, R., Morrison, L. J., & Hall, J. P. (2015). DNA recombination strategies during antigenic variation in the African trypanosome, 409–435.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303. http://doi.org/10.1101/gr.107524.110

Mehlitz, D., Zillmann, U., Scott, C. M., & Godfrey, D. G. (1982). Epidemiological studies on the animal reservoir of Gambiense sleeping sickness. Part III. Characterization of trypanozoon stocks by isoenzymes and sensitivity to human serum. *Tropenmedizin Und Parasitologie*, *33*(2), 113–118.

Meloni, B. P., Lymbery, A. J., & Andrew, T. R. (1988). Isoenzyme electrophoresis of 30 isolates of Giardia from humans and felines. *The American Journal of Tropical Medicine and Hygiene*, *38*(1), 65–73.

Meloni, B. P., Lymbery, A. J., & Thompson, R. A. (1989). Characterization of Giardia isolates using a non-radiolabeled DNA probe, and correlation with the results of

isoenzyme analysis. *The American Journal of Tropical Medicine and Hygiene*, *40*(6), 629–637.

Mihok, S., Otieno, L. H., & Darji, N. (1990). Population genetics of Trypanosoma brucei and the epidemiology of human sleeping sickness in the Lambwe Valley, Kenya. *Parasitology*, *100*(2), 219–233.

Moreno, G., Rioux, J. A., Lanotte, G., Pratlong, F., & Serres, E. (1986). Leishmania: Taxonomie et Phylogenese: Applications Eco-Epidemiologiques. *Colloques Intemationaux De Centre National De La Recherche Scientifique/Institut National De La Sante Et De La Recherche Mddicale (CNRS/INSERM, Paris)*, 105–119.

Nadler, S. A., & Honigberg, B. M. (1988). Genetic differentiation and biochemical polymorphism among trichomonads. *The Journal of Parasitology*, 797–804.

Nash, T. E., McCutchan, T., Keister, D., Dame, J. B., Conrad, J. D., & Gillin, F. D. (1985). Restriction-endonuclease analysis of DNA from 15 Giardia isolates obtained from humans and animals. *Journal of Infectious Diseases*, *152*(1), 64–73.

Oliveira, R. P., Broude, N. E., Macedo, A. M., Cantor, C. R., Smith, C. L., & Pena, S. D. J. (1998). Probing the genetic population structure of Trypanosoma cruzi with polymorphic microsatellites. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(7), 3776–3780.

Peacock, L., Bailey, M., Carrington, M., & Gibson, W. (2014). Meiosis and Haploid Gametes in the Pathogen Trypanosoma brucei. *Current Biology, 24*(2), 181–186. http://doi.org/10.1016/j.cub.2013.11.044

Poplin, R., Ruano-Rubio, V., DePristo, M. A., Fennell, T. J., Carneiro, M. O., Van der Auwera, G. A., et al. (2017). Scaling accurate genetic variant discovery to tens of thousands of samples, 1–22. http://doi.org/10.1101/201178

Pratlong, F., Lanotte, G., Ashford, R. W., & Rioux, J. A. (1986). Leishmania: Taxonomie et Phylogénèse: Applications Eco-Epidémiologiques, Colloques Internationaux de Centre National de la Recherche Scientifique/Institut National de la Santé et de la Recherche Médicale. *Paris: CNRS/INSERM*, 129–139.

Proctor, E. M., Wong, Q., Yang, J., & Keystone, J. S. (1987). The electrophoretic isoenzyme patterns of strains of Entamoeba histolytica isolated in two major cities in Canada. *The American Journal of Tropical Medicine and Hygiene*, *37*(2), 296–301.

Pujol, C., Reynes, J., Renaud, F., Raymond, M., Tibayrenc, M., Ayala, F. J., et al. (1993). 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*, *90*(20), 9456–9459.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, *81*(3), 559–575. http://doi.org/10.1086/519795

Revollo, S., Oury, B., Laurent, J.-P., Barnabé, C., Quesney, V., Carrière, V., et al. (1998). Trypanosoma cruzi: impact of clonal evolution of the parasite on its biological and medical properties. *Experimental Parasitology*, *89*(1), 30–39.

Roellig, D. M., McMillan, K., Ellis, A. E., Vandeberg, J. L., Champagne, D. E., & Yabsley, M. J. (2010). Experimental infection of two South American reservoirs with four distinct strains of Trypanosoma cruzi. *Parasitology*, *137*(6), 959–966.

Safrin, R. E., Lancaster, L. A., Davis, C. E., & Braude, A. I. (1986). Differentiation of Cryptococcus neoformans serotypes by isoenzyme electrophoresis. *American Journal of Clinical Pathology*, *86*(2), 204–208.

Sargeaunt, P. G., Williams, J. E., Jackson, T., & Simjee, A. E. (1982). A zymodeme study of Entamoeba histolytica in a group of South African schoolchildren. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *76*(3), 401–402.

Schenkman, S. (2012). T rypanosoma cruzi trans-sialidase as a multifunctional enzyme in Chagas' disease. *Cellular Microbiology*, *14*(10), 1522–1530.

Tibayrenc, M. (2003). Kinetoplastid Biology and Disease. *Kinetoplastid Biology and Disease*, *2*(1), 12–6. http://doi.org/10.1186/1475-9292-2-12

Tibayrenc, M., & Ayala, F. J. (1988). Isozyme variability in Trypanosoma cruzi, the agent of Chagas' disease: genetical, taxonomical, and epidemiological significance. *Evolution*, *42*(2), 277–292.

Tibayrenc, M., Kjellberg, F., & Ayala, F. J. (1990). 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*, *87*(7), 2414–2418.

Tibayrenc, M., Kjellberg, F., Arnaud, J., Oury, B., Brenière, S. F., Dardé, M. L., & Ayala, F. J. (1991). Are eukaryotic microorganisms clonal or sexual? A population genetics vantage. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(12), 5129–5133.

Tibayrenc, M., Ward, P., Moya, A., & Ayala, F. J. (1986). Natural populations of Trypanosoma cruzi, the agent of Chagas disease, have a complex multiclonal structure. *Proc Natl Acad Sci USA*, *83*(1), 115.

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., et al. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Current Protocols in Bioinformatics*, *43*, 11.10.1–33. http://doi.org/10.1002/0471250953.bi1110s43

Veloso, V. M., Carneiro, C. M., Toledo, M., Lana, M., Chiari, E., Tafuri, W. L., & Bahia, M. T. (2001). Variation in susceptibility to benznidazole in isolates derived from Trypanosoma cruzi parental strains. *Memorias Do Instituto Oswaldo Cruz*, *96*(7), 1005–1011.

Verani, J. R., Seitz, A., Gilman, R. H., LaFuente, C., Galdos-Cardenas, G., Kawai, V., et al. (2009). Geographic Variation in the Sensitivity of Recombinant Antigen-based Rapid Tests for Chronic Trypanosoma cruzi Infection. *The American Journal of Tropical Medicine and Hygiene*, *80*(3), 410–415. http://doi.org/10.4269/ajtmh.2009.80.410

Weatherly, D. B., Peng, D., & Tarleton, R. L. (2016). Recombination-driven generation of the largest pathogen repository of antigen variants in the protozoan Trypanosoma cruzi. *BMC Genomics*, *17*(1), 729.

Westenberger, S. J., Barnabé, C., Campbell, D. A., & Sturm, N. R. (2005). Two

hybridization events define the population structure of Trypanosoma cruzi. *Genetics*,

171(2), 527–543.

Zingales, B., Miles, M. A., Campbell, D. A., Tibayrenc, M., Macedo, A. M., Teixeira, M.

M., et al. (2012). The revised Trypanosoma cruzi subspecific nomenclature: rationale,

epidemiological relevance and research applications. *Infection, Genetics and Evolution*,

12(2), 240–253.

<u>Conclusions</u>

Summary

In Chapter One we sought to determine how life history strategies interact with environmental and historical features to affect the ability of a population to immigrate to and establish in a new environment. We showed that only one *T. cruzi* lineage established a population in vectors in Arequipa, Peru, despite evidence of multiple immigration events. The low probability of establishment likely resulted from *(a)* historical factors including the immigration of humans to regions of the city without vectors; combined with, *(b)* life history strategies such as stercorarian transmission that cause a low transmission rate from bug to host..

In Chapter Two we aimed to determine how ecological factors affect the patterns of population expansion and dispersal after immigration to a new environment. We showed that *T. cruzi* dispersal between houses within a city block is relatively unimpeded and is likely facilitated by its vector. Dispersal between houses on different blocks occurs independent of the distance between blocks and is likely mediated by human activities, possibly including the trading of infected guinea pigs.

In Chapter Three we assessed the mechanisms that generate and maintain diversity as the population of the parasite expanded. Despite the benefits conferred by sexual reproduction and the capacity for sexual reproduction in *T. cruzi*, natural populations appear clonal. We resolved this apparent contradiction by showing that meiosis and fertilization occurred repeatedly as the population dispersed through the city, but that exclusive inbreeding and frequent mitosis caused a decay in heterozygosity. While most of the genetic differences between isolates in this population were due to recombination

and fertilization events since the common ancestor, cell division via mitosis is likely much more frequent than meiosis, which prolonged the maintenance of heterozygosity.

Perspectives and Future Directions

Invasion dynamics across spatial scales

Our analyses show that immigration to and establishment of a novel environment is improbable, however subsequent dispersal within the environment occurs readily. The dynamic processes of immigration, establishment, and dispersal share many themes across multiple spatial scales. Establishment of a new urban environment (Chapter 1) and a new habitat within an urban environment (Chapter 2) is improbable even when sufficiently hospitable. At two spatial scales (city and block) we attributed the low probability of establishment to a low immigration rate to habitats that contain vectors and a low transmission rate of *T. cruzi* from vector to host. Subsequent dispersal within the habitat (both city and block) occurs readily. For example, after establishment in the city (which likely only occurred once over ~60 years), *T. cruzi* readily expanded throughout the urban environment. Similarly, establishment of a city block is improbable (most city blocks remain uninfected), however after immigration to a block the immigrant tends to establish a thriving population within that block (most blocks that contain one parasite contain multiple). At both spatial scales analyzed here, invasion of a new habitat is hindered by the parasite's transmission rate and immigration rate.

Chagas disease risk for uninfected humans in an uninfected city is very low because the low transmission rate (Rabinovich *et al.* 1990; Gürtler *et al.* 1998; Gürtler *et al.* 2007; Llewellyn *et al.* 2009b; Cohen & Gürtler 2001; Nouvellet *et al.* 2013) causes a low probability of colonizing an uninfected city even if it contains competent vectors and

some immigration of *T. cruzi*-infected humans. If the contact rate between humans and vectors can be further reduced, the establishment of a *T. cruzi* population in a city like Arequipa may be postponed by years. The timing of invasion for the Chagas disease system is contrasted by systems with high transmission rates where the contact rate may need to be reduced to near-zero in order to minimize disease risk, likely requiring local elimination of the vector. For example, the high transmission rate from mosquito to host allowed multiple independent introductions of the Zika virus to colonize Rio de Janiero, Brazil in 2014 (Metsky *et al.* 2017).

Future work should aim to determine if the patterns of invasion of cities (Chapter 1) and city blocks (Chapter 2) can be applied to the connected urban network by comparing the genetic diversity of urban lineages of a diverse set of species. Invasion and colonization of any urban habitat from a sylvatic habitat may be very improbable, as seen in urban environments throughout northern Venezuela where the dominant *T. cruzi* lineage is unrelated to local sylvatic populations (Llewellyn *et al.* 2009). However, after colonization of an urban environment, subsequent migration between connected urban habitats may be relatively trivial. For example, after its introduction in Rio de Janiero, Brazil, a single Zika virus lineage rapidly spread between urban environments across the Americas (Faria *et al.* 2016). Invasion of any urban environment by a single sylvatic lineage may propel the colonizing lineage toward a global distribution via human-mediated dispersal within the network of connected urban environments. On the other hand, conservation efforts of non-pest species may benefit from prioritizing species that are now exclusively found in urban environments if the global genetic diversity of the species is low.

I hypothesize that within-block dispersal patterns of species with similar life history strategies to *T. cruzi* will tend to mimic sylvatic within-habitat dispersal patterns, but that between-block dispersal is distance-independent and reliant on associations with humans, even if between-habitat dispersal readily occurs in sylvatic populations. Between-habitat dispersal in sylvatic environments often occurs in a distance-dependent manner (e.g. steady migration of *Ixodes scapularis* from south to north in the northeastern US) (Khatchikian *et al.* 2015), however associations with humans may provide a mechanism for distance-independent migration between habitats. Human associations may be particularly important in urban environments where the barriers between city blocks may prohibit sylvatic dispersal mechanisms for many species. Of course, dispersal of species with certain life history strategies would not be hindered by roads, nor would some require the use of human activity to migrate long distances (e.g. birds) (Fernández-Juricic 2000; Goddard *et al.* 2010). On the other hand, species that cannot disperse between blocks using sylvatic dispersal mechanisms and cannot associate with humans may become isolated within city blocks, often leading to local extinction.

Importance of guinea pigs in T. cruzi transmission cycle

Guinea pigs are important reservoirs for *T. cruzi* in Peru because of the number of bugs that feed on each animal (Herrer 1955; Cohen & Gürtler 2001; Levy *et al.* 2006; Levy *et al.* 2008; Coffield *et al.* 2013; Castillo-Neyra *et al.* 2016). Many residents of Arequipa own and house domestic guinea pigs (Levy *et al.* 2014; Levy *et al.* 2015). The trading and selling of guinea pigs among residents of different localities in Arequipa (Kim 2013) may be a primary mechanism of long-distance *T. cruzi* dispersal. This hypothesis would be supported if the patterns of guinea pig trading mimic

phylogeographic reconstructions of the *T. cruzi* population in Arequipa. Recently-traded infected guinea pigs in blocks containing uninfected bugs may indicate that the local bug population is likely to become infected.

Importance of humans in T. cruzi transmission cycle

While the importance of guinea pigs and dogs for *T. cruzi* transmission has been well-documented (Herrer 1955; Córdova *et al.* 1969; Cohen & Gürtler 2001; Levy *et al.* 2006; Levy *et al.* 2008; Coffield *et al.* 2013; Levy *et al.* 2014; Castillo-Neyra *et al.* 2016), substantial sampling of bugs, humans, and small mammals in the same region are required to determine the importance of humans in the *T. cruzi* transmission cycle. The *T. cruzi* isolates from dogs (N=2) and guinea pigs (N=7) analyzed here provide additional support that these mammals are integral to the maintenance of *T. cruzi* in the urban environment because the parasites isolated from these mammals are indistinguishable from those isolated from bugs. Thus, establishment of *T. cruzi* in a new block likely involves transmission between bugs and small mammals. The absence of multiple *T. cruzi* lineages within the bug and small mammal transmission cycle in Arequipa despite a non-zero immigration rate of infected humans suggests that infected immigrant humans are not effectively contributing the parasite to the transmission cycle. The presence of *T. cruzi* lineages in the human population that are not present in the bug population would suggest that humans are often a dead end for the parasite and possibly that establishment probability—as opposed to rare migration— is the limiting factor in *T. cruzi* colonization. However, if most infected humans harbor the same lineage isolated from vectors and small mammals then it is likely that migration and establishment are both low and that humans are important for establishment of the parasite.

Since there is no entirely effective cure for chronic *T. cruzi* infection, local elimination of *T. cruzi* requires all infected hosts to perish. Prevention of transmission between hosts and vectors must therefore be maintained for the lifetime of the longest-living hosts: decades for humans. However, if humans are not important for *T. cruzi* transmission, then effectively eliminating the parasite may be possible by maintaining the prevention of contact between small mammals and bugs for a much shorter time. If contact between hosts and vectors resumes while the remaining *T. cruzi* infection is confined to human hosts, it may be unlikely that the parasite enters the transmission cycle between bugs and small mammals.

Potential meiotic mechanisms and implications of outcrossing

The compelling evidence of regular sexual reproduction presented here and elsewhere (Weatherly *et al.* 2016; Schwabl *et al.* 2018) should reinvigorate the field to elucidate the mechanism of meiosis in *T. cruzi*. Prior *T. cruzi* research combined with an understanding of the mechanisms of meiosis in related species allows speculation into the mechanism of meiosis in *T. cruzi*. The literature contains several insights into the location, timing, and mechanism of meiosis including: *(a)* Recombination in *T. cruzi* occurs in the mammalian host (Gaunt *et al.* 2003); however *T. brucei* undergoes meiosis in its vector, suggesting the possibility that *T. cruzi* also undergoes meiosis in its vector (Peacock *et al.* 2014). *(b)* The potential for outcrossing in nature is shown by the hybridization of unrelated *T. cruzi* lineages (Brisse *et al.* 1998; Brisse *et al.* 2003; Westenberger *et al.* 2005; El-Sayed *et al.* 2005), however the frequency of outcrossing has not been determined. *(c)* A lack of evidence for outcrossing in populations that contain multiple *T. cruzi* lineages (Oliveira *et al.* 1998; Lopez-Cancino *et al.* 2015; Dumonteil *et al.* 2018) may suggest that outcrossing

is less frequent than inbreeding, however the potential for outcrossing is eliminated in the population surveyed here because there is only a single, genetically similar lineage in Arequipa.

Relatively infrequent outcrossing may be due to *(i)* the timing and location of meiosis or *(ii)* the mechanism of meiosis. *(i)* Depending on the brevity of the presence of gametes before fusion and the specificity of the location in which meiosis occurs, outcrossing may require that multiple strains undergo meiosis simultaneously in the same location. If ephemeral gametes are fertilized in a precise location, outcrossing may be restricted by the timing and location of meiosis as opposed to the molecular mechanism. *(ii)* The mechanism of meiosis and fertilization may prohibit outcrossing. For example, diploid yeast cells undergo meiosis and fertilization within a protective casing which may minimize, but not eliminate outcrossing (Guilliermond & Tanner 1920; Johnson *et al.* 2004; Taxis *et al.* 2005; Ruderfer *et al.* 2006; Tsai *et al.* 2008; Murphy & Zeyl 2010). *T. cruzi* may undergo meiosis via a similar mechanism that restricts but not prohibits outcrossing. However, it is also possible that the mechanism of meiosis in *T. cruzi* causes the frequency of outcrossing to vary across populations as seen in some populations of *Plasmodium falciparum*, the causative agent of malaria in humans, which undergoes frequent meiosis and regular outcrossing (Mu *et al.* 2005; Neafsey *et al.* 2008).

The amount of outcrossing and inbreeding in *T. cruzi* has important medical and evolutionary implications. Large genetic differences between *T. cruzi* lineages that diverged millions of years ago cause variation in traits including host organ system preference, virulence, and host and vector propensity (Zingales *et al.* 2012). Outcrossing may allow rapid diversification of these and other traits, which may allow for adaptations

to new environments. Outcrossing may also allow for the swapping or introgression of drug resistance genes. If outcrossing is frequent, widespread applications of new chemical agents may result in the emergence of drug-resistant parasites. If *T. cruzi* primarily undergoes asexual reproduction coupled with periodic sexual inbreeding, as seen in the population studied here, the introgression of drug-resistant genes will be unlikely. In fact, this reproductive strategy allows linkage disequilibrium to persist across large genomic regions, which may allow genome-wide association studies to be used to map genes involved in drug-resistance and virulence (Mu *et al.* 2005). The population structure resulting from outcrossing or inbreeding should be considered when performing population genetics and evolutionary analyses, especially when testing hypotheses based on population genetic and phylogenetic models such as phylogeographic reconstructions based on inferred phylogenetic relationships. The applications of population structure outlined above highlight the importance of elucidating the mechanism of meiosis and determining the frequency of inbreeding and outcrossing in natural *T. cruzi* populations.

Literature Cited

Brisse, S., Barnabé, C., Banuls, A.-L., Sidibé, I., Noël, S., & Tibayrenc, M. (1998). A phylogenetic analysis of the Trypanosoma cruzi genome project CL Brener reference strain by multilocus enzyme electrophoresis and multiprimer random amplified polymorphic DNA fingerprinting. *Molecular and Biochemical Parasitology*, *92*(2), 253–263.

Brisse, S., Henriksson, J., Barnabé, C., Douzery, E. J., Berkvens, D., Serrano, M., et al. (2003). Evidence for genetic exchange and hybridization in Trypanosoma cruzi based on nucleotide sequences and molecular karyotype. *Infection, Genetics and Evolution*, *2*(3), 173–183.

Castillo Neyra, R., Borrini-Mayorí, K., Salazar Sánchez, R., Ancca Suarez, J., Xie, S., Náquira Velarde, C., & Levy, M. Z. (2016). Heterogeneous infectiousness in guinea pigs experimentally infected with Trypanosoma cruzi. *Parasitology International*, *65*(1), 50–54.

Ceballos, L. A., Vazquez-Prokopec, G. M., Cecere, M. C., Marcet, P. L., & Gürtler, R. E. (2005). Feeding rates, nutritional status and flight dispersal potential of peridomestic populations of Triatomainfestans in rural northwestern Argentina. *Acta Tropica*, *95*(2), 149–159.

Coffield, D. J., Jr, Spagnuolo, A. M., Shillor, M., Mema, E., Pell, B., Pruzinsky, A., & Zetye, A. (2013). A model for Chagas disease with oral and congenital transmission. *PLoS ONE*, *8*(6), e67267.

Cohen, J. E., & Gürtler, R. E. (2001). Modeling household transmission of American trypanosomiasis. *Science*, *293*(5530), 694–698. http://doi.org/10.1126/science.1060638

Córdova, E., Montesinos, J., & Náquira, F. (1969). Estudio epidemiológico sobre la enfermedad de Chagas en el valle de Vitor. *Arch. Per Pat. Clin*, *23*, 257–270.

Dumonteil, E., Ramirez-Sierra, M.-J., rez-Carrillo, S. P. X., Teh-Poot, C., Herrera, C., re, S. X. B. G. X., & Waleckx, E. (2018). Detailed ecological associations of triatomines revealed by metabarcoding and next- generation sequencing: implications for triatomine behavior and Trypanosoma cruzi transmission cycles. *Scientific Reports*, 1–13. http://doi.org/10.1038/s41598-018-22455-x

El-Sayed, N. M., Myler, P. J., Bartholomeu, D. C., Nilsson, D., Aggarwal, G., Tran, A.-N., et al. (2005). The genome sequence of Trypanosoma cruzi, etiologic agent of Chagas disease. *Science*, *309*(5733), 409–415.

Fernández Juricic, E. (2000). Avifaunal use of wooded streets in an urban landscape. *Conservation Biology*, *14*(2), 513–521.

Goddard, M. A., Dougill, A. J., & Benton, T. G. (2010). Scaling up from gardens: biodiversity conservation in urban environments. *Trends in Ecology & Evolution*, *25*(2), 90–98.

Guilliermond, A., & Tanner, F. W. (1920). The Yeasts. *New York John Wiley and Sons*, 424.

Gürtler, R. E., Cecere, M. C., Lauricella, M. A., Cardinal, M. V., Kitron, U., & Cohen, J. E. (2007). Domestic dogs and cats as sources of Trypanosoma cruzi infection in rural northwestern Argentina. *Parasitology*, *134*(1), 69–82.

Gürtler, R. E., Cohen, J. E., Cecere, M. C., Lauricella, M. A., Chuit, R., & Segura, E. L. (1998). Influence of humans and domestic animals on the household prevalence of Trypanosoma cruzi in Triatoma infestans populations in northwest Argentina. *The American Journal of Tropical Medicine and Hygiene*, *58*(6), 748–758.

Herrer, A. (1955). Trypanosomiasis americana en el Perú: III. Importancia del cobayo como reservorio de la enfermedad de Chagas en la región sudoccidental. *Revista Peruana De Medicina Experimental Y Salud Publica*, *9*(1-2), 45–55.

Johnson, L. J., Koufopanou, V., Goddard, M. R., Hetherington, R., Schäfer, S. M., & Burt, A. (2004). Population genetics of the wild yeast Saccharomyces paradoxus. *Genetics*, *166*(1), 43–52.

Khatchikian, C. E., Prusinski, M. A., Stone, M., Backenson, P. B., Wang, I.-N., Foley, E., et al. (2015). Recent and rapid population growth and range expansion of the Lyme disease tick vector, Ixodes scapularis, in North America. *Evolution*, *69*(7), 1678–1689. http://doi.org/10.1111/evo.12690

Kim, E. (2013, July 26). *Modeling Domestic Guinea Pigs and Trypanosoma cruzi Persistence in Arequipa, Peru*. (S. Levin, Ed.). Princeton, NJ. Retrieved from http://arks.princeton.edu/ark:/88435/dsp01gb19f592r

Levy, M. Z., Barbu, C. M., Castillo Neyra, R., Quispe-Machaca, V. R., Ancca Juárez, J., Escalante-Mejia, P., et al. (2014). Urbanization, land tenure security and vector-borne Chagas disease. *Proceedings. Biological Sciences*, *281*(1789), 20141003–20141003. http://doi.org/10.1098/rspb.2014.1003

Levy, M. Z., Bowman, N. M., Kawai, V., Waller, L. A., Del Carpio, J. G. C., Benzaquen, E. C., et al. (2006). Periurban Trypanosoma cruzi–infected Triatoma infestans, Arequipa, Peru. *Emerging Infectious Diseases*, *12*(9), 1345.

Levy, M. Z., Quispe-Machaca, V. R., Ylla-Velasquez, J. L., Waller, L. A., Richards, J. M., Rath, B., et al. (2008). Impregnated netting slows infestation by Triatoma infestans. *The American Journal of Tropical Medicine and Hygiene*, *79*(4), 528–534.

Levy, M. Z., Tustin, A., Castillo Neyra, R., Mabud, T. S., Levy, K., Barbu, C. M., et al. (2015). Bottlenecks in domestic animal populations can facilitate the emergence of Trypanosoma cruzi, the aetiological agent of Chagas disease. *Proc. R. Soc. B*, *282*(1810), 20142807.

Llewellyn, M. S., Miles, M. A., Carrasco, H. J., Lewis, M. D., Yeo, M., Vargas, J., et al. (2009). Genome-Scale Multilocus Microsatellite Typing of Trypanosoma cruzi Discrete Typing Unit I Reveals Phylogeographic Structure and Specific Genotypes Linked to Human Infection. *PLOS Pathogens*, *5*(5), e1000410. http://doi.org/10.1371/journal.ppat.1000410

López-Cancino, S. A., Tun-Ku, E., la Cruz-Felix, De, H. K., Ibarra-Cerdeña, C. N., Izeta-Alberdi, A., Pech-May, A., et al. (2015). Landscape ecology of Trypanosoma cruzi in the southern Yucatan Peninsula. *Acta Tropica*, *151*, 58–72. http://doi.org/10.1016/j.actatropica.2015.07.021

Metsky, H. C., Matranga, C. B., Wohl, S., Schaffner, S. F., Freije, C. A., Winnicki, S. M., et al. (2017). Zika virus evolution and spread in the Americas. *Nature Publishing Group*, *546*(7658), 411–415. http://doi.org/10.1038/nature22402

Mu, J., Awadalla, P., Duan, J., McGee, K. M., Joy, D. A., McVean, G. A., & Su, X.-Z. (2005). Recombination hotspots and population structure in Plasmodium falciparum. *PLoS Biology*, *3*(10), e335.

Murphy, H. A., & Zeyl, C. W. (2010). Yeast Sex: Surprisingly High Rates of Outcrossing between Asci. *PLoS ONE*, *5*(5), e10461–6. http://doi.org/10.1371/journal.pone.0010461

Neafsey, D. E., Schaffner, S. F., Volkman, S. K., Park, D., Montgomery, P., Milner, D. A., et al. (2008). Genome-wide SNP genotyping highlights the role of natural selection in Plasmodium falciparum population divergence. *Genome Biology*, *9*(12), R171.

Nouvellet, P., Dumonteil, E., & Gourbière, S. (2013). The improbable transmission of Trypanosoma cruzi to human: the missing link in the dynamics and control of Chagas disease. *PLOS Neglected Tropical Diseases*, *7*(11), e2505.

Oliveira, R. P., Broude, N. E., Macedo, A. M., Cantor, C. R., Smith, C. L., & Pena, S. D. J. (1998). Probing the genetic population structure of Trypanosoma cruzi with polymorphic microsatellites. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(7), 3776–3780.

Rabinovich, J. E., Wisnivesky-Colli, C., Solarz, N. D., & Gürtler, R. E. (1990). Probability of transmission of Chagas disease by Triatoma infestans (Hemiptera: Reduviidae) in an endemic area of Santiago del Estero, Argentina. *Bulletin of the World Health Organization*, *68*(6), 737.

Richer, W., Kengne, P., Cortez, M. R., Perrineau, M. M., Cohuet, A., Fontenille, D., & Noireau, F. (2007). Active dispersal by wild Triatoma infestans in the Bolivian

Andes. *Tropical Medicine & International Health*, *12*(6), 759–764. http://doi.org/10.1111/j.1365-3156.2007.01846.x

Ruderfer, D. M., Pratt, S. C., Seidel, H. S., & Kruglyak, L. (2006). Population genomic analysis of outcrossing and recombination in yeast. *Nature Genetics*, *38*(9), 1077.

Schwabl, P., Imamura, H., Van den Broeck, F., Costales, J. A., Maiguashca, J., Miles, M. A., et al. (2018). Parallel sexual and parasexual population genomic structure in Trypanosoma cruzi. *bioRxiv*, 338277.

Taxis, C., Keller, P., Kavagiou, Z., Jensen, L. J., Colombelli, J., Bork, P., et al. (2005). Spore number control and breeding in Saccharomyces cerevisiae: a key role for a self-organizing system. *The Journal of Cell Biology*, *171*(4), 627–640.

Tsai, I. J., Bensasson, D., Burt, A., & Koufopanou, V. (2008). Population genomics of the wild yeast Saccharomyces paradoxus: quantifying the life cycle. *Proceedings of the National Academy of Sciences*, *105*(12), 4957–4962.

Weatherly, D. B., Peng, D., & Tarleton, R. L. (2016). Recombination-driven generation of the largest pathogen repository of antigen variants in the protozoan Trypanosoma cruzi. *BMC Genomics*, *17*(1), 729.

Westenberger, S. J., Barnabé, C., Campbell, D. A., & Sturm, N. R. (2005). Two hybridization events define the population structure of Trypanosoma cruzi. *Genetics*, *171*(2), 527–543.

Zingales, B., Miles, M. A., Campbell, D. A., Tibayrenc, M., Macedo, A. M., Teixeira, M. M., et al. (2012). The revised Trypanosoma cruzi subspecific nomenclature: rationale, epidemiological relevance and research applications. *Infection, Genetics and Evolution*, *12*(2), 240–253.