Molecular Epidemiology of Parasitic Diseases: The Chagas Disease Model

Juan David Ramírez and Felipe Guhl
Centro de Investigaciones en Microbiología y Parasitología Tropical (CIMPAT), Universidad de los Andes, Bogotá, Colombia

1. Introduction

Parasitic diseases represent one of the most important issues in public health. More than one billion people worldwide are infected by parasites causing different disease scenarios. Parasitic diseases are closely related to geographic, social and economic factors driving the prevalence and incidence of these pathologies (WHO, 2010). These represent a broad group of eukaryotic organisms that may cause severe diseases in animal and human populations. Parasites are the causative agents of pathologies such as Malaria. In 2008, there were 247 million cases of Malaria and nearly one million deaths from the disease, mostly among children living in Africa. In Africa, a child dies of Malaria every 45 seconds; the disease accounts for 20% of all childhood deaths. Leishmaniasis threatens approximately 350 million men, women and children in 88 countries around the world. As many as 12 million people are believed to be currently infected by this disease, with approximately 1-2 million estimated new cases occurring every year. Additionally, an estimated of 10 million people are infected worldwide by Chagas disease (American trypanosomiasis), mostly in Latin America, where Chagas disease is endemic. More than 25 million people are at risk of acquiring this disease. It is estimated that in 2008, Chagas disease killed more than 10,000 people. Schistosomiasis is a chronic, parasitic disease caused by blood flukes (trematode worms) of the genus Schistosoma. More than 207 million people are infected with these organisms worldwide, with an estimated 700 million people at risk in 74 endemic countries. Lymphatic filariasis affects more than 1.3 million people in 81 countries. Approximately 65% of those infected live in Southeast Asia, 30% in Africa and the remainder in other tropical areas. Lymphatic filariasis afflicts over 25 million men with genital disease and over 15 million people with lymphoedema. Because the prevalence and intensity of infection are linked to poverty, elimination can contribute to achieving the United Nations Millennium Development Goals. Human African Trypanosomiasis (HAT) affects mostly poor populations living in remote rural areas of Africa. If untreated, it is usually fatal. Travellers also risk becoming infected if they venture through regions where the insect vector (tsetse flies) is common. Generally, the disease is not found in urban areas, although some cases have been reported in suburban areas of Kinshasa, the capital of the Democratic Republic of Congo, and Luanda, the capital city of Angola. In 2004, the number of new reported cases fell to 17,616, which the WHO considered to be due to increased control, estimating the cumulative rate to be between 50,000 and 70,000 cases (WHO, 2010). These trends show the
importance of developing strategies to mitigate the prevalence of these parasitic diseases, and molecular epidemiology arises as a potential tool to understand disease dynamics.

Molecular epidemiology is considered a powerful tool for understanding the genetic variation and evolution of pathogens. The use of the technologies based on molecular biology techniques has allowed the scientific community to reveal disease determinants and the genetic structure of parasites that provoke diseases and cause millions of deaths each year. In recent years, new studies have been conducted with the purpose of elucidating the genetic structure of the etiological agents of these pathologies with the aim of designing strategies that could help to mitigate the associated diseases in human and animal populations. Molecular epidemiology and population genetics have shown to be powerful strategies to understand the genetic structure of parasites, with special emphasis on understanding disease and transmission dynamics. The objective of this chapter is to illustrate for the reader the paramount importance of molecular epidemiology in parasitic diseases, showing some clear examples of parasite disease and transmission dynamics. The focus will be on how molecular epidemiology can be a helpful tool to mitigate disease transmission, prevalence and incidence and can be used as a reliable tool for disease surveillance as well as the need to create synergy between molecular epidemiology and public health programmes to reduce the prevalence of parasitic diseases.

2. Importance and relevance of molecular epidemiology

Molecular epidemiology may be defined as a tool focused on the contribution of potential genetic factors identified by molecular techniques to the aetiology, distribution and prevention of disease across populations (Kilbourne, 1973). It is a field of study that has recently emerged from the integration of molecular biology, epidemiology, biochemistry and public health systems (Figure 1). This approach has been useful in attempting to determine the pathogenesis of certain diseases as well as the genetic variation and genetic structure of pathogens.

![Molecular Epidemiology](https://www.intechopen.com)

Fig. 1. Molecular epidemiology is considered an interdisciplinary science that is a composite of different sciences.
Molecular epidemiology has recently gained paramount importance in the fields of human genetics and in molecular virology. One clear example of this is the use of molecular epidemiology to track viruses that generate severe acute respiratory syndrome (SARS). This allowed researchers to develop strategies with the purpose of tracking the transmissibility and dispersal of this virus, observing that the positive selection pressure associated with human hosts resulted in the emergence of lineages of the virus that became readily transmissible between humans, causing the epidemic outbreak of 2002-2003 (Zhao, 2007). The use of molecular markers also permitted the establishment of prevention and control strategies to mitigate the transmission of these genotypes, thus avoiding increased disease prevalence. Thus, this example shows how traceable pathogens can be and clearly demonstrates the great utility of molecular biology on the basis of molecular epidemiology in obtaining a deeper understanding regarding parasitic diseases.

Molecular epidemiology has been established as a promising science for studying the contribution of potential genetic markers and environmental risk factors of parasitic diseases representing a close synergy between molecular biology and epidemiology. Some of the main objectives of molecular epidemiology focused on the study of parasitic diseases are as follows:

- To enhance our understanding of the pathogenesis of parasitic diseases: Some authors have used the *Toxoplasma gondii* model to describe relationships between nucleoside triphosphate hydrolase (NTPase) isoforms and *Toxoplasma* strain virulence in human toxoplasmosis, reporting that different isoforms are involved in clinical forms of this parasitic disease (Johnson et al., 2003).

- To define genetic susceptibility with genetic markers: The use of human pedigrees to observe patterns of susceptibility to visceral leishmaniasis in Brazil has allowed researchers to develop action plans to mitigate this tropical disease in endemic areas (Jamieson et al., 2007).

- To allow evaluation of subclinical or early disease markers: Prognostic markers are one of the milestone deliverables in molecular epidemiology. In the case of *T. gondii*, a quantitative real-time PCR assay for amniotic fluid has been developed to provide a prognostic marker of foetal infection in pregnant women (Romand et al., 2004).

- To provide new standards for descriptive epidemiology: Some of the problems involved in descriptive epidemiology show how difficult it is to track some kinds of diseases. In particular, the parasites that cause gastrointestinal syndromes fall into this category, such as *Entamoeba histolytica*, *Taenia solium*, *Ascaris lumbricoides*, *Giardia intestinalis*, *Enterobius vermicularis* and others. Microscopic identification becomes tedious and, in some case confusing due to similar morphologies among some parasites. Molecular detection based on PCR assays has provided the field of descriptive epidemiology with a more reliable way to analyse data in population descriptive studies (Singh et al., 2009; Pecson et al., 2006).

- To improve precision in analytical epidemiology: While descriptive epidemiology provides the what, who, when and where; analytical epidemiology attempts to provide the why and how. Few examples are listed related to the detection of emergent genotypes in disease surveillance. A good example is that of HAT, which is caused by two sympatric subspecies (*Trypanosoma brucei rhodesiense* and *T. b. gambiense*); each subspecies is involved in disease severity causing a large number of annual deaths in Africa (Morrison et al., 2011).
As has been shown thus far, molecular epidemiology as applied to the study of parasites is considered an important and relevant tool to investigate these organisms. The important point to focus is on the correct use of the information obtained. Molecular methods are currently available and becoming cheaper every day, and the emergence of new accurate and feasible molecular methods shows promising results for the molecular epidemiology field; however, in molecular epidemiology, the clinical question must be always highlighted to obtain the most reliable results. Thus, the income data become the critical point in the study of parasitic diseases and the basis for obtaining a good outcome that can be translated to meet the main aims of molecular epidemiology studies (Figure 2). This chapter will discuss the molecular methods available to develop molecular epidemiology studies focused on parasites, with some clear examples of how useful and necessary molecular epidemiology is for understanding disease outcomes, transmission dynamics and the current genetic structure of parasitic diseases.

Fig. 2. Flow of information based on the accurate and reliable use of molecular epidemiology focused on the study of parasitic diseases.

3. Molecular biology tools applied in the analysis of parasitic diseases

Molecular biology has made important contributions in the last ten years with respect to understanding the genetics of parasites causing human illness. A broad description of the markers used in molecular epidemiology and their features is presented in Table 1. The first techniques used to track pathogens and disease dispersal were based on biochemical markers, with Multilocus Enzyme Electrophoresis (MLEE) being broadly used to study parasites such as *T. cruzi*, *Leishmania* spp, *E. histolytica* and *T. brucei* (Miles et al., 1977; Pinto et al., 2005; Mathews et al., 1983; Nijokou et al., 2004). This technique is based on differences between loci; according to the obtained banding pattern, it is possible to distinguish among lineages. In the case of *Leishmania*, this technique was used to differentiate species of the genus involved in visceral and cutaneous leishmaniasis (Bañuls et al., 2000; Bañuls et al., 2002; Zhang et al., 2006). A drawback of this technique is the need to culture large quantities of parasites, which are quite difficult to obtain in most parasitic diseases. Subsequently,
molecular biology techniques based on the use of DNA were developed, among which RAPDs (Random Amplified Polymorphic DNA), AFLPs (Amplified Fragment Length Polymorphism), PFGE (Pulse Field Gel Electrophoresis) and RFLPs (Restriction Fragment Length Polymorphism) were the most used techniques within the scientific community. In this sense, analysis of ribosomal markers using RAPDs was of paramount importance in attempting to develop assays to distinguish among morphologically similar amoebas. *E. histolytica* and *E. dispar*, which represent pathogenic and non-pathogenic amoeba species, respectively, have been suggested by many authors to have evolved identically up to the point when a mutation generated a cryptic speciation pattern (Clark et al., 2006). The use of RAPD techniques has also been an important aid in understanding genetic variation within *E. histolytica* isolated from different hosts (Gomes et al., 2000; Prakash et al., 2000). AFLPs and RFLPs are the most recent techniques to be employed to track pathogen dispersal and elucidate their genetic diversity. Fingerprinting based on AFLPs has been useful in differentiating subspecies of *T. brucei*, which is the aetiological agent of sleeping sickness, a pathology that affects more than 8 million people in Africa. These techniques have shown great reproducibility in distinguishing *T. b. rhodesiense* and *T. b. gambiense*, two sympatric species that generate different symptomatologies, demonstrating how molecular epidemiology can assist in understanding disease outcomes in certain pathologies, thus aiding in developing proper treatment and management measures (Agbo et al., 2002; Masiga et al., 2006). RFLPs have been applied to parasites such as *T. gondii* and *Trichinella spiralis*, which are two species for which pigs play an important role in transmission dynamics. *T. gondii* is considered to display a clonal population, but its isolates have been divided into three types that are geographically clustered and, in some cases, are involved in disease outcomes (Wang et al., 1995; Su et al., 2002; Fuentes et al., 2001). These molecular markers have been shown to be important in discriminating species as well as evaluating genetic variability among isolates from the same species. The great advantage of these molecular markers is that they can be used to show the pattern of variation across the whole genome of a pathogen, rather than just a specific region, as will be shown later.

<table>
<thead>
<tr>
<th>Feature</th>
<th>MLEE1</th>
<th>RAPD2 and AFLP3</th>
<th>RFLP4 and PFGE5</th>
<th>Microarrays</th>
<th>MLMT6</th>
<th>MLST7</th>
<th>qPCR8</th>
<th>Genome Sequencing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Culturing</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Analysis of distinct loci</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Cost</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>Labor</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>Informative</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Portability</td>
<td>Low</td>
<td>Medium</td>
<td>Medium</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>Medium</td>
<td>Low</td>
</tr>
</tbody>
</table>

1 Multilocus Enzyme Electrophoresis, 2 Random Amplified Polymorphic DNA, 3 Amplified Fragment Length Polymorphisms, 4 Restriction Fragment Length Polymorphisms, 5 Pulse Field Gel Electrophoresis, 6 Multilocus Microsatellite Typing, 7 Multilocus Sequence Typing, 8 Quantitative Real Time PCR

Table 1. Molecular markers used in molecular epidemiology studies.
The development of Polymerase Chain Reaction (PCR) was of great importance and represents an incredible advance in molecular biology. Since 1990, PCR has been broadly used to study parasites. Modifications of PCR, such as PCR-RFLP, Nested PCR, RT-PCR, AP-PCR (Allele Polymorphic Polymerase Chain Reaction), SHELA-PCR (Solution Hybridization Enzyme-Linked Assay Polymerase Chain Reaction) and qPCR (Quantitative Real Time Polymerase Chain Reaction), have been applied to study parasites. In these tests, the only thing that a researcher requires is a few aliquots of DNA, which simplifies the analysis. The use of PCR-RFLP is widely reported in discriminating Leishmania species. In these studies, the use of Heat Shock proteins with different molecular weights allows discrimination based on PCR amplification of genes that are subsequently digested with restriction endonucleases. According to the band patterns obtained in such analyses, it has been possible to distinguish between the subgenera Viannia and Leishmania as well as species from the different subgenera in some cases (Volpini et al., 2004; Montalvo et al., 2006). Discrimination among Leishmania species or their complexes is necessary in conducting studies on treatment resistance and clinical manifestations associated with leishmaniasis, in which some species cause cutaneous forms, and others cause visceral forms. One of the most important advances in the modification of PCR assays has been the development of quantitative Real-Time PCR (qPCR). This assay involves the quantification of DNA copies in each PCR cycle. The first applications of this technique have been for diagnostic purposes in cases where it has been possible to estimate parasitic loads in infected patients. In the case of L. infantum, the species involved in visceral leishmaniasis manifestations, assays have been developed to estimate parasitic loads in biopsies of infected patients and, thus, to estimate the efficacy of treatment using qPCR (Mary et al., 2004; Bretagne et al., 2001; Ranasinghe et al., 2008). Additionally, qPCR using SYBR green and Hybridisation probe chemistry has allowed the development of melting temperature (Tm) analysis according to a dissociation curve. This permits screening for genotypes or species according to the specific temperature of an amplicon and enables observation of single nucleotide polymorphisms, all in the same reaction. Thus, a qPCR Real Time protocol has been proposed to identify Leishmania species focused on spliced leader genes and minicircle kDNA regions; according to distinct temperatures, investigators were able to discriminate Leishmania species (Wortmann et al., 2005). This approach has also been applied to other parasites, such as Giardia, for which qPCR assays were developed to detect G. lamblia and to discriminate its genotypes in stool specimens (Guy et al., 2004). Additionally, it has been used to study schistosomiasis, an helminthic disease that affects populations in Africa, Asia and America, with qPCR assays being developed to discriminate species in water where the infective form (cercariae) lives and is transmitted to humans (Lier et al., 2006). qPCR has been shown to be a reliable, feasible, fast and accurate method in molecular epidemiology to discriminate species as well as to determine genotypes within species. This suggests the need to pursue studies involving this method, though in some cases, validation studies are required, and further research is needed. A problem involved in working on parasites is sensitivity because in some parasite diseases, the parasitic loads are quite low and even undetectable in some cases, such that concentration methods must be applied or it may be necessary to analyse the whole sample.

Other molecular markers include microarrays and Southern blot and northern blot techniques. The drawback of microarrays is that they are time consuming, expensive, and in some cases, they do not provide the desired information. In the last decade, the rise of
sequencing procedures has been an important addition to molecular epidemiology investigations. The ability to obtain DNA sequences has allowed researchers to unravel the genetic structure of parasites and to go further in the analyses that can be applied. Thus, molecular phylogenetics and population genetics have provided molecular epidemiology with certain, reliable tools for understanding the genetic structure of parasites. Molecular phylogenetics is the science focused on understanding the evolutionary relationships among groups of organisms based on molecular sequences, and these techniques have been broadly applied to parasites. Hence, molecular phylogenetics has allowed the reconstruction of phylogenetic trees based on maximum parsimony and/or maximum composite likelihood methods with the aim of understanding the evolutionary history of parasites and, in some cases, developing analysis involving loci. Based on the use of DNA sequences and phylogenetic reconstructions, new methods such as MLST (Multilocus Sequence Typing) have arisen. MLST has been broadly used in bacteria and yeast but has only recently been applied to parasites; the drawback of MLST in addressing protozoan parasites associated with working with clonal diploids instead of clonal haploids, such as bacteria. The genetic structure of clonal diploid pathogenic organisms is important in terms of elucidating the drivers of disease prevalence, installation and outcomes as well as the virulence factors and geographical distribution related to the disease. In recent years, the population structure of microorganisms such as *Plasmodium*, *Giardia*, *Entamoeba*, *Trypanosomes* (*T. brucei, T. congolense* and *T. cruzi*), *Candida*, *Leishmania* and *Toxoplasma* has gained paramount relevance due to the discussion of clonal propagation versus sexual recombination (De Meeus et al., 2006; Benett et al., 2010; Grigg and Suzuki, 2003; Morrison et al., 2009; Rougeron et al., 2010; Mzilahowa et al., 2007). There are three hypotheses that describe the genetic structure observed in clonal diploid organisms. In 1987, Harvey and Keymer suggested a panmictic population structure in which sexual recombination is frequent. In 1991, Tibayrenc et al. proposed the clonal theory of parasitic protozoa, suggesting that these organisms display a clonal propagation mode associated with infrequent sexual recombination events. Finally, in 1993, Maynard-Smith et al. proposed an epidemic population structure with a background level of frequent sexual recombination and with occasional clonal expansion of particular genotypes. These hypotheses have been tested using a large number of parasitic protozoa; however, the debate still continues.

The use of new methods for typing and elucidating the genetic variability of parasites like MLST has gained importance because it can be considered to be an improvement of MLEE. MLST makes use of different loci involved in parasite metabolism. Thus, MLST strategies have been developed in *Leishmania* for species identification using five metabolic enzymes that are able to discriminate species and genotypes within complexes (Zemanova et al., 2007). An MLST approach has also recently been described for discriminating among subtypes of *Blastocystis*, which is a protozoan parasite involved in bowel inflammation and acute diarrhoea in immunodeficient patients (Stensvold and Clark, 2011). MLST analyses are not only used to discriminate genotypes of species but also to detect recombination or likely genetic exchanges among parasite populations (diploid clonals). In the case of sexual parasites such as helminthes, these approaches are employed to discriminate among species or to detect genotypes. Another important development in molecular epidemiology has been the use of microsatellite markers in developing MLMT (Multilocus Microsatellite Typing) strategies. Microsatellite markers are defined as tandem repetitions of 1-6 base pair segments of DNA; they are neutral and co-dominant and are useful in developing
population genetics analyses. The variability of microsatellites is due to their higher rate of
mutation compared to other neutral regions of DNA. The use of microsatellite markers is
widely reported for purposes ranging from species identification to detection of
recombination based on population genetics statistics. In *P. vivax*, polymorphic
microsatellite markers have been amplified and analysed to unravel the genetic structure of
this parasite and to understand its co-evolution with other *Plasmodium* species (Gomes et al.,
2003; Imgwon et al., 2006). Population genetics tools present a limitation when working with
clonal diploids related to the assumption of Hardy-Weinberg equilibrium and other
statistics, such as *Fis* and *Fst*.

In recent years, DNA sequencing has become an important tool in understanding
microorganisms, particularly those involved in human pathologies. Sequencing procedures
have been improved, and pyrosequencing has become an important method to obtain more
feasible and accurate DNA sequences. Pyrosequencing is a method of DNA sequencing
based on the "sequencing by synthesis" principle. It differs from Sanger sequencing in that it
relies on the detection of pyrophosphate release upon nucleotide incorporation, rather than
chain termination with dideoxynucleotides. Genome sequencing methods developed in
bacteria have also been applied to sequence whole genomes in parasites. The first parasite
genome sequenced was that of *P. falciparum* (Gardner et al., 2002), followed by the genomes
of other parasites, such as *Leishmania*, *T. brucei*, *T. cruzi* and *T. gondii* (Ivens et al., 2005;
Elsayed et al., 2005; Bontell et al., 2009). New initiatives have been developed to sequence
larger genomes, such as those from helminthes including *Ascaris*, *Taenia*, *Schistosoma*
and *Echinococcus*. These approaches have allowed scientists to develop projects aimed at
annotating parasite genomes for the purpose of detecting possible pharmaceutical markers
to develop drugs against these microorganisms. Genome sequencing is becoming cheaper
due to the advances made by Illumina, which will allow the scientific community to begin
sequencing genomes instead of single genes. The possibility of obtaining this type of
metadata permits the application of tools in bioinformatics, metabolomics, immunomics,
vaccinomics, proteomics and other field with the purpose of transitioning into the OMICS
era. The OMICS era is considered to be associated with the most advanced techniques for
understanding the molecular epidemiology of parasitic diseases. These tools will be of
paramount importance in developing new drugs against parasites as well as evaluating
surveillance disease markers or prognostic disease markers to understand the relatedness
between disease outcomes and parasite genetic variability, which is one of the main
objectives in molecular epidemiology.

4. Comparative molecular epidemiology: The Chagas disease model

Chagas disease, which is caused by the parasite *T. cruzi*, is a complex zoonosis that is widely
distributed throughout the American continent. The infection can be acquired through
triatomine faeces, blood transfusion, oral and congenital transmission and laboratory
accidents. Chagas disease represents an important public health problem, with estimates by
the Pan American Health Organization in 2005 of at least 7.7 million people being infected
with *T. cruzi* and another 110 million being at risk (WHO, 2007). Additionally, immigration
of infected people from endemic countries is now making Chagas disease a relevant health
issue in other regions, including Europe and the United States (Rassi et al., 2009). Chagas
disease is comprised of two stages, with the acute phase occurring approximately one week
after the initial infection and approximately 30-40% of infected patients developing the 
chronic phase of the disease, in which cardiomypathy is the most frequent and severe 
clinical manifestation (Rassi et al., 2009). Chagas has lately gained more importance due to 
recent reports of imported cases in Europe, the United States and Canada (Schmunis and 
Yadon, 2010).

Obtaining a full understanding of the aetiology and epidemiology of Chagas disease across 
its distribution has proved elusive and complex and remains the subject of intense 
investigation to the present day. The difficulty in completely defining the epidemiology of 
Chagas disease is attributable to several factors. First, Chagas disease is a zoonosis, and a 
variety of widely distributed mammals serve as reservoirs for T. cruzi. Moreover, all 
mammals are susceptible to T. cruzi infection. An additional factor that contributes to the 
complexity of Chagas disease as a zoonosis is the variety of vectors involved, as they are not 
simply represented by a range of related species or genera, as is the case for all other known 
insect vector-associated diseases. Triatomine bugs are a subfamily of insects, and across this 
relatively broad taxonomic range, there are members from all groups that can harbour T. 
cruzi. However, most transmission is attributable to three main genera: Rhodnius, 
Panstrongylus, and Triatoma, but this diversity still represents two different tribes of the 
subfamily (Rhodniini and Triatomini). Furthermore, the insects vary in more than their 
ancestry, being associated with a diverse range of vertebrate hosts and ecological 
associations. The third factor that complicates the epidemiology of Chagas disease and 
accounts for variation in the clinical manifestation of the disease is the subspecific diversity 
of T. cruzi itself. Much work has been conducted over the past 40 years to elucidate the 
variation of T. cruzi across its geographical distribution and associations with hosts and 
vector species.

The T. cruzi parasite comprises a heterogeneous population that displays clonal propagation 
due to its different cycles of transmission and the possibility of recombination exchanges, 
which can be found in nature and have previously been reported in vitro (Gaunt et al., 2003, 
Sturm et al., 2003; Westenberger et al., 2002). T. cruzi is genetically diverse and is classified 
into a series of strains or subtypes. This genetic diversity was initially discovered using a 
panel of isoenzyme markers to investigate differences between parasites involved in 
domestic and sylvatic cycles in Bahia state in Brazil (Miles et al., 1977). This study 
represented a breakthrough, revealing that in Bahia, there were substantial genetic 
differences between the parasites involved in sympatric sylvatic and domestic transmission 
cycles. These variants were designated zymodemes I and II (ZI and ZII). Soon thereafter, it 
was revealed that the widespread strain associated with the sylvatic cycle in Brazil (ZI) was 
the predominant cause of human disease in Venezuela (Miles et al., 1981). These 
groundbreaking findings paved the way for investigating the aetiology of Chagas disease, 
allowing host–vector–parasite associations and comparative geographical distributions to 
be explored, as reviewed by Miles et al. (2009). Subsequently, four additional zymodemes 
were described from Brazil, Paraguay, and Bolivia. In the following two decades, various 
authors proceeded to characterise strains of T. cruzi, applying other molecular methods as 
they became available. As a result, further diversity was discovered within the original 
zymodemes. However, the designations of subtypes in the literature began to become 
confusing. Recently, a new nomenclature for T. cruzi has been adopted that includes six 
Discrete Taxonomic Units (DTUs) designated T. cruzi I (TcI), T. cruzi II (TcII), T. cruzi III
(TcIII), *T. cruzi* IV (TcIV), *T. cruzi* V (TcV) and *T. cruzi* VI (TcVI) based on different molecular markers and biological features (Zingales et al., 2009). These DTUs are broadly distributed in the American continent in diverse ecotopes (Figure 3). Discrimination of the six DTUs has become an important issue in the molecular epidemiology of *T. cruzi*. There are many reports showing algorithms for the molecular characterisation of these DTUs by performing RAPD, PCR-RFLP, qPCR, MLST, MLMT and DNA sequencing analyses, but to date, there is no consensus protocol for strain typing (Lewis et al., 2009; Rozas et al., 2007; Ramírez et al., 2010; Duffy et al., 2009; Yeo et al., 2011; Llewelly et al., 2009). One of the most recent and reliable algorithms for *T. cruzi* typing was reported by Ramírez et al., 2010 and has been applied on biological samples (Figure 4).

![Geographical distribution of *T. cruzi* DTUs in the American continent](image-url)

**Fig. 3.** Geographical distribution of *T. cruzi* DTUs in the American continent based on Patterson and Guhl, 2010.
In a sense, the findings of Miles et al. (1977) were the tip of the iceberg in unravelling the genetic structure of *T. cruzi*, but at the same time, they hit the nail on the head. The observed predominance of Tcl in the human populations in Venezuela and Colombia and TcII, TcV, and TcVI mostly infecting human hosts in Brazil was to prove representative (i.e., it has since been demonstrated that Tcl predominates in countries north of the Amazon and TcII-TcVI in Southern Cone countries, but this distribution is not absolute). This is particularly illuminating given that there are distinct clinical differences between patients presenting with Chagas disease in these two geographical regions. Strains appear to differ in terms of both their pathogenicity and response to treatment. Tcl and TcII-VI are all associated with cardiac lesions in human infections, but it appears that only TcII, TcV, and TcIV are also associated with digestive tract lesions (Prata, 2001), despite the recent report of digestive tract lesions in Colombia caused by Tcl (Mantilla et al., 2010). Tcl is generally considered to be less pathogenic with lower parasitemia (Burgos et al., 2007) and more chronic cases being asymptomatic compared to Chagas caused by TcII, TcV, and TcVI in Argentina, Brazil, Chile, Paraguay, and Uruguay. Tcl is almost the only form found in human infections north of the Amazon region. Moreover, there is an observed general partitioning of former TcII subtypes between sylvatic and domestic transmission cycles; the human disease is associated with TcII, while TcV is rarely associated with sylvatic hosts (Yeo et al., 2005), and TcIII and TcIV are predominantly sylvatic.

Tcl has remained a constant grouping in the nomenclature since it was first described. However, recent studies based on mini-exon gene (SL-IR) sequences have shown...
polymorphism on this region, with four genotypes being reported within Tcl. These genotypes have also been reported in various regions of South America, where five Tcl genotypes have been detected (Figure 5) (Cura et al., 2010; Herrera et al., 2007; Guhl and Ramírez, 2011). Different molecular markers, including a 48 set of microsatellite loci, have also shown the great diversity in Tcl (Guhl and Ramírez, 2011; Llewellyn et al., 2009; Spotorno et al., 2008; Ramírez et al., 2011). Primers designed based on Tcl sequences confirmed the existence of three genotypes (Ia, Ib and Id) and a new genotype found in the Southern cone countries designated Tcl_e (Cura et al., 2010; Falla et al., 2009).

Fig. 5. Geographical distribution of *T. cruzi* I genotypes based on SL-IR region (Guhl and Ramírez, 2011)

Genetic variability has been clearly demonstrated in *T. cruzi*, with reports of homogeneous (TcII) and heterogeneous groups considered to be hybrids due to recombination events (TcIII-TcVI) (Gaunt et al., 2003, Sturm et al., 2003; Westenberger et al., 2002). It has been shown that TcIII and TcIV are likely to be a product of recombination of Tcl and TcII and...
TcIV-TcVI to be a product of recombination of TcII and TcIII/TcIV (Brisse et al., 2003), although this last statement is still controversial. The recent advances in sequencing procedures have allowed three complete *T. cruzi* genomes to be obtained. The first strain fully sequenced was CL Brener (TcVI), which showed a large number of repetitive elements along the core genome (Elsayed et al., 2005). Likewise, the recent sequencing of the Esmeraldo (TcII) and Sylvio X10 (TcI) genomes has shown the relationship between repetitive elements and mucin-like proteins, which are closely associated with parasite cell invasion and survival, showing this area of inquiry to be quite promising with respect to obtaining more information about the genetic structure of *T. cruzi* DTUs (Franzen et al., 2011; Andersson, 2011).

The molecular epidemiology and distribution of *T. cruzi* genotypes may have important implications with respect to characteristics of the disease. However, few correlations have related *T. cruzi* genetic variability and disease outcome, though it has been shown that TcI is more closely related to patients with cardiomyopathy in Colombia and Venezuela, while TcII-TcVI are more associated with patients with digestive syndrome (megaesophagus/megacolon) (Mantilla et al., 2011; Rassi et al., 2009). The distribution of *T. cruzi* genotypes and reservoirs is implicated in the genetic epidemiology of the disease. In the southern part of the American continent, infection of *Canis familiaris* has been found to be related to TcIV, V and VI, whereas infections in the north are associated with genotypes Ia and Ib (Falla et al., 2009; Herrera et al., 2009). Furthermore, a significant number of *D. marsupialis* are infected with the Tcld genotype, which suggests an association with the sylvatic transmission cycle. Similar studies in primates have demonstrated that TcI predominantly infects arboreal reservoirs (Cura et al., 2010; Falla et al., 2009). There are several hypotheses regarding the distribution of different genetic groups of *T. cruzi*, suggesting that reservoirs belonging to arboreal ecotopes are preferentially infected with TcI and that terrestrial ecotopes are infected with TcII-TcVI (Yeo et al., 2005). This hypothesis is controversial in light of recent reports demonstrating that the arboreal ecotope reservoirs *Monodelphis brevicaudata*, *Philander frenata* and *Didelphis aurita* are infected with TcIII, TcIV and TcII, respectively (Marcili et al., 2009; Llewellyn et al., 2009b).

However, the associations are not absolute, and in the case of Tc, there is no apparent clustering of particular TcI genotypes with *Didelphis* in comparison to isolates from other arboreal mammals (Llewellyn et al., 2009b). Additionally, with respect to phylogeographical analyses of TcIII, the results indicate that isolates cluster according to geography rather than host association (Marcili et al., 2009; Llewellyn et al., 2009b). This could also be supported by the recent analysis developed in mammals naturally infected with TcI using microsatellite markers revealing the role of mammalian reservoirs in diversifying selection on *T. cruzi* (Llewellyn et al., 2011). Two interesting studies on host responses to different strains have confirmed, by comparative artificial infection, that in the southern USA, two species of opossum (*Monodelphis domestica* and *Didelphis virginiana*) seem to be resistant to TcIV (Roellig et al., 2009; Roellig et al., 2010). This highlights a mechanism for the association of a vertebrate host with one strain over others. The strong association between TcI and *Rhodnius* species can be explained by a similar mechanism: comparative studies on artificial infection of *R. prolixus* with various strains revealed a tendency for this species to be resistant to infection by TcII (Mello et al., 1995). In triatomines, susceptibility or resistance to trypanosome infections seems to be modulated by intestinal symbionts that are vital for
development. *T. cruzi* is considered to be subpathogenic for triatomines, whereas *Trypanosoma rangeli* is a species that commonly infects *Rhodnius* species and causes pathogenicity based on reduction of the number of symbionts (Vallejo et al., 2009). Studies using different species of triatomines, such as *R. pallescens*, *T. dimidiata*, *R. colombiensis* and *P. geniculatus*, have shown the affinity of Tcl for infecting these species in comparison with TcII (Mejia-Jaramillo et al., 2009). At least half of all species of triatomin bugs have been found to be naturally infected with *T. cruzi* (Lent and Wygodzinsky, 1979; Schofield, 1994). Unfortunately, the vast majority of these records do not include specific strain associations. This is clearly an area of potential research. In the context of dispersal triggered by starvation, there is evidence that starvation decreases *T. cruzi* infection in triatomines (Kollien and Schaub, 1998), and in some species, starvation may clear the infection altogether (Phillips et al., 1967; Vargas et al., 1985). This factor could help to explain paradigms such as that observed in Venezuela, where sylvatic and domestic bugs seem to be in panmixia, but Tcl shows discrete general clustering of sylvatic and domestic cycles (Fitzpatrick et al., 2009; Llewellyn et al., 2009b). Triatomin bugs directly determine the aetiology of the strains of *T. cruzi* involved in human transmission cycles. This is clear because despite Tcl and Didelphis being widespread, it is the northern distribution of *Rhodnius* that corresponds with its occurrence in human cycles. Overall, the aspects of epidemiological relevance are that associations between terrestrial ecology, *T. infestans*, terrestrial mammals, and *T. cruzi* strains TcII/TcIV have led to the prominence of TcII, TcV, and TcIV in human infections in the southern cone countries of South America. In the northern cone countries of South America, human American Trypanosomiasis infections seem to stem from Tcl associated with arboreal *Rhodnius* and arboreal mammals.

The definition of *T. cruzi* nomenclature must be related to the biological, clinical and pathological characteristics associated with specific populations of *T. cruzi* (Campbell et al., 2004; Zafra et al., 2009). To our knowledge, few correlations reported have been demonstrated to date regarding differences of the host humoral response to specific *T. cruzi* genotypes; however, these findings were flawed because of the low reliability of the diagnostic tests used, leading to a high proportion of false negatives due to variability in the *T. cruzi* strain used for the diagnosis. The implication of Tcl in severe forms of myocarditis in cardiac samples from chronic chagasic patients in Argentina and the lack of any specific clinical manifestation related to *T. cruzi* DTUs in Bolivian chagasic patients indicate the pleomorphism of *T. cruzi* (Ramírez et al., 2009; Moncayo and Ortiz, 2006; Burgos et al., 2010; del Puerto et al., 2010). There have been studies reporting detection of *T. cruzi* in blood samples. Direct detection of *T. cruzi* DTUs in the blood of chronic Chagasic patients was carried out by amplification of the 24Sα rDNA divergent domain and the use of mitochondrial house-keeping genes (Zafra et al., 2009). In this study, molecular characterisation of *T. cruzi* DTUs showed that most of the patients were infected with Tcl, while some patients were found to be infected with TcII (9.9%). Recently, a new approach to *T. cruzi* DTU detection in chronic Chagasic patients was developed indicating that Tcl is the predominant DTU, though TcII was also detected, and it was reported that the genetic characteristics of TcII parasites found in Colombia were similar to those of TcII found in Bolivia and Chile (González et al., 2010). Regarding the genetic variability of the parasite, prognostic markers based on mitochondrial genes where the presence of specific mutations can trigger complications of the chronic phase of the disease in asymptomatic patients have also been demonstrated (dos Santos et al., 2009; Carranza et al., 2009). Despite the observed
genetic variability, it is important to consider the presence of *T. cruzi* clones that can be found in different tissues. Several studies have demonstrated a specific histiotropism of *T. cruzi* in mice showing differences in the pathological, immunological and clinical features that the parasite can elicit in the host (Andrade et al., 2002; Ramírez et al., 2010; Manoel-Caetano et al., 2008). Moreover, some authors have shown that the *T. cruzi* population in a patient's bloodstream could be dissimilar to the parasite population that causes tissue damage (Vago et al., 2000). Differences were found in *T. cruzi* populations in the bloodstreams of patients with chronic Chagasic cardiomyopathy and those of Chagasic patients without cardiomyopathy (Venegas et al., 2009). Microsatellite analyses have also shown multiclonality in heart samples and in the bloodstreams of infected patients, demonstrating that specific populations of *T. cruzi* can probably determine disease outcome (Burgos et al., 2007; Valadares et al., 2008).

Molecular epidemiological studies on *T. cruzi* have attempted to establish the effects of different DTUs in the clinical progression of Chagas disease. Several studies have shown the effect of genetic variability on the host immune response (dos Santos et al., 2009; Melquiades-Rodriguez et al., 2010; Ramírez et al., 2009). It was previously known that cardiopathies in southern cone countries were caused by TcII, TcV and TcVI, but it has recently been demonstrated that TcI can play an important role specifically in severe cardiopathies related to Chagas disease. Studies of cardiac biopsies from Argentinian patients revealed that patients with severe myocarditis were infected with TcI, whereas those with moderate or absent myocarditis were infected with TcII, TcV or TcVI (Burgos et al., 2010). At the same time that the TcI genotype was found in severe myocarditis patients, it was demonstrated that in patients with chronic chagasic cardiopathy, the Tcla genotype was most commonly found in the bloodstream, whereas Tcld was most commonly found in cardiac biopsies. These results are consistent with reports from patients in Colombia, where the least and most prevalent Tcl genotypes in adult patients with chronic chagasic cardiopathy were Tcld and Tcla, respectively (Ramírez et al., 2010). This suggests a possible type of histiotropism associated with Tcl genotypes as well as the epidemiological importance of this DTU in southern countries, where cardiopathies were previously thought to be caused primarily by TcII, TcV and TcVI. A model of clonal histiotropism has been previously reported showing how a composite of clones is related to disease outcome. Recently reported results from Colombia support this premise, with cardiac biopsies being observed to be infected with Tcld, while Tcla is found circulating in the bloodstream (Zafra et al., 2011; Ramírez et al., 2010). This suggests the need to pursue studies to correlate the association between *T. cruzi* genotypes and clinical manifestations of Chagas disease. New studies are also necessary to determine the specific *T. cruzi* populations generating tissue damage in infected patients.

Most of the research performed in Chagas disease is related to understanding the molecular epidemiology of this endemic pathology. Many questions are continually emerging every day in this field based on epidemiological circuits with the aim of better estimating the transmission dynamics of *T. cruzi* in endemic areas. The involvement of *T. cruzi* genetic variability in clinical manifestations is of paramount importance and could resolve the question regarding the high pleomorphism displayed by this clinical entity. New initiatives must be created with interdisciplinary groups with the purpose of unravelling the molecular comparative epidemiology of Chagas disease and attempting to mitigate this pathology in endemic countries.
5. Concluding remarks

In this chapter, many examples regarding the usefulness of molecular epidemiology in parasitic diseases were addressed. These examples illustrated different applications of molecular methods to understand the pathogens that cause human parasitic diseases. It is important to consider the need for synergy between descriptive, analytical and molecular epidemiological methods to develop robust and unbiased data. As a relevant example, we presented the case of *T. cruzi* and described how molecular methods have been useful in defining hypotheses about the parasite’s geographical distribution, host associations and the implications of different genotypes for clinical manifestations related to the heart. Despite the studies reported in the literature on molecular epidemiology in parasites, public health systems do not consider the importance of integration between these two areas. We propose the integration of molecular epidemiology and public health systems to mitigate and reduce the prevalence of tropical diseases caused mainly by parasites, and this combination could become a potential tool for disease prevention and control as well as for the development of appropriate programmes for disease surveillance in endemic countries.

6. References


Andersson B. (2011). The *Trypanosoma cruzi* genome; conserved core genes and extremely variable surface molecule families. Research in Microbiology 162:619-625


www.intechopen.com


This special issue resulted from the invitation made to selected authors to contribute with an overview of a specific subject of their choice, and is based on a collection of papers chosen to exemplify some of the interests, uses and views of the epidemiology across different areas of research and practice. Rather than the comprehensiveness and coherence of a conventional textbook, readers will find a set of independent chapters, each of them of a great interest in their own specialized areas within epidemiology. Taken together, they illustrate the contrast between the attempt to extend the limits of applicability of epidemiological research, and the “regular” scientific activity in this field or an applied epidemiology. Epidemiologists with different levels of expertise and interests will be able to find informative and inspiring readings among the chapters of this book.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: