Investigations on the epidemiology and diversity of *Leishmania tropica* and *L. aethiopica* and the differentiation of their sand fly vectors

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium

(Dr. rer. nat.)

im Fach Biologie

eingereicht an der

Lebenswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von

Dipl.-Biologin Lena Krayter

Präsident der Humboldt-Universität zu Berlin
Prof. Dr. Jan-Hendrik Olbertz

Dekanin/Dekan der Lebenswissenschaftlichen Fakultät
Prof. Dr. Richard Lucius

Gutachter/innen: 1. Prof. Dr. Richard Lucius
2. Prof. Dr. Wolfgang Presber
3. Prof. Dr. Charles Jaffe

Tag der mündlichen Prüfung: 08.09.2015

1 Abstract

Leishmania tropica is the causative agent of human cutaneous leishmaniasis in foci ranging from Africa through the Middle East to northern India. It has also been isolated from human cases of visceral leishmaniasis and wild and domestic animals. It was long supposed to be transmitted anthroponotically, but, in certain foci, rock hyraxes have been incriminated as their natural reservoir hosts.

By multilocus microsatellite typing (MLMT), the world-wide population structure of this species and its closely related species *L. aethiopica* has been revealed applying methods based on both genetic distances and allele frequencies. The 195 strains of *L. tropica* and eight strains of *L. aethiopica* largely clustered according to their geographical origins. A large group of Israeli and Palestinian strains was separated from all other strains at the highest hierarchical level. All other strains formed clusters and groups at lower levels. The strains of *L. aethiopica* formed a distinct group, although clustering among other African strains of *L. tropica*. Intrafocal diversity was found among Moroccan strains isolated in two neighbouring foci, which were separated into two clusters independent of their origins. No correlation was found between microsatellite profiles and disease manifestation when Indian strains isolated from visceral and cutaneous cases were analysed. They formed two groups, which were genetically similar but not identical. Preliminary data obtained through a whole genome sequencing approach including strains of *L. tropica*, *L. aethiopica* and *L. major* have largely corroborated the results of the MLMT approach.

To reveal the reasons for the high genetic variability among strains of *L. tropica*, a Functional Cloning approach was conducted using N-methyl-N'-nitro-N-nitrosoguanidin (MNNG) as an indicator for a functioning or impaired mismatch repair (MMR) system. The transfectants retrieved from the transfection of an acceptor strain exhibiting high MNNG tolerance (hypothetically impaired MMR system) with a cosmid library bearing the genomic DNA of a donor strain with a reduced MNNG tolerance (hypothetically functioning MMR system) were screened for the restored phenotype of the donor strain. Unfortunately, the MNNG tolerance level was not stable in neither independent experiments of the same clone or in experiments testing the MNNG tolerance of different sub-clones. The identification of the cosmid inserts of two clones exhibiting low MNNG tolerance was not successful by BLAST searches, probably because a reference sequence for *L. tropica* is not yet available at the NCBI database.

The time- and cost-efficient identification of a large amount of sand flies is important since several thousands are caught during field studies. Here, a multiplex ligation-dependent probe amplification approach (MLPA) for the identification of sand flies endemic to the Middle East is introduced. The

target for this approach is the 18S rRNA. The unambiguous identification of *Phlebotomus syriacus*, *P. arabicus* and *P. papatasi* was possible with this approach using species-specific probes. Furthermore, this technique has the potential to discriminate more species and to be applied to pooled sand fly specimens.

2 Zusammenfassung

Leishmania tropica ist der Auslöser von kutaner Leishmaniose beim Menschen und kommt in Afrika über den Mittleren Osten bis nach Nordindien vor. Der Parasit wurde auch von humanen Fällen von viszeraler Leishmaniose und aus Wild- und Haustieren isoliert. Lange Zeit ging man davon aus, dass die Übertragung anthroponotisch erfolgte, aber, in manchen Foci werden mittlerweile Klippschliefer als natürlicher Wirt angenommen.

Mittels Mikrosatellitentypisierung (MLMT) wurde die weltweite Populationsstruktur dieser Spezies und der nahe verwandten Spezies *L. aethiopica* aufgedeckt, indem sowohl Methoden angewandt wurden, die auf genetischen Distanzen beruhen als auch solche, die auf der Analyse von Allelfrequenzen basieren. Die 195 Stämme von *L. tropica* sowie die acht Stämme von *L. aethiopica* gruppierten hauptsächlich gemäß ihrer geographischen Herkunft. Eine große Gruppe von israelischen und palästinensischen Stämmen wurde von allen anderen Stämmen auf der höchsten hierarchischen Ebene separiert. Alle anderen Stämme bildeten genetische Entitäten erst auf niedrigeren Niveaus. Die Stämme von *L. aethiopica* stellten eine eigene Gruppe dar, die allerdings innerhalb der afrikanischen Stämme von *L. tropica* gruppierte. Intrafokale Diversität wurde zwischen marokkanischen Stämmen aus zwei benachbarten Foci gefunden, die in zwei Cluster unabhängig von ihrer Herkunft gruppiert. Die Analyse von indischen *L. tropica*-Stämmen, die von kutanen und viszeralen Fällen isoliert wurden, ergab keine Hinweise für eine Korrelation zwischen den Mikrosatellitenprofilen und dem jeweiligen Krankheitsbild. Die zwei Gruppen waren genetisch ähnlich, aber nicht identisch. Vorläufige Ergebnisse einer genomweiten SNP-Analyse haben die Ergebnisse der Mikrosatellitenanalyse weitgehend bestätigt.

Um die Gründe für die hohe genetische Variabilität innerhalb der Spezies *L. tropica* herauszufinden, wurde eine Funktionelle Klonierung durchgeführt, in der N-methyl-N'-nitro-N-nitrosoguanidin (MNNG) als Indikator für ein funktionierendes oder eingeschränktes Mismatch Repair (MMR)-System eingesetzt wurde. Dafür wurde ein Akzeptorstamm (hohe MNNG-Toleranz, hypothetisch eingeschränktes MMR-System) mit einer Cosmidbibliothek, die genomische DNA eines Donorstammes (niedrige MNNG-Toleranz, hypothetisch funktionierendes MMR-System)enthielt, transfiziert. Die erhaltenen Transfektanten wurden dann auf ihre MNNG-Toleranz getestet. Leider blieb der MNNG-Toleranzlevel aber nicht stabil, weder in unabhängigen Experimenten desselben Klons noch von einzelnen Subklonen. Die Identifizierung der Cosmidinserts von zwei Klonen, die eine niedrige MNNG-Toleranz zeigten, war mit BLAST-Suchen nicht erfolgreich, möglicherweise aufgrund der fehlenden Referenzsequenz von *L. tropica* in NCBI.

Die zeit- und kosteneffiziente Identifizierung von großen Mengen an Sandmücken ist wichtig für Feldstudien, in denen mehrere Tausend Mücken gefangen werden. Hier wird eine multiplexe Technik zur ligationsabhängigen Sondenamplifizierung vorgestellt, die die Identifizierung von Sandmücken-Spezies im Mittleren Osten ermöglicht. Als Targetsequenz dient dabei die 18S rRNA. Die Spezies *Phlebotomus syriacus*, *P. arabicus* und *P. papatasi* können durch diese Methode mit spezies-spezifischen Sonden eindeutig identifiziert werden. Außerdem kann diese Methode dazu genutzt werden, weitere Spezies zu diskriminieren und auf gepoolte Sandmücken angewandt werden.

List of abbreviations

L.	*Leishmania*
CL	cutaneous leishmaniasis
VL	visceral leishmaniasis
MCL	mucocutaneous leishmaniasis
DCL	diffuse cutaneous leishmaniasis
LCL	local cutaneous leishmaniasis
P.	*Phlebotomus*
P.	*Procavia*
MLMT	multilocus microsatellite typing
FCA	Factorial Correspondence Analysis
NJ tree	Neighbour Joining tree
F_{ST}	mean fixation index
H_o	observed heterozygosity
H_e	expected heterozygosity
A	number of alleles
F_{IS}	inbreeding coefficient
MLEE	multilocus enzyme electrophoresis
MLPA	multiplex ligation-dependent probe amplification
PCR	polymerase chain reaction
fw	forward
rev	reverse
AT	annealing temperature
aqua dest.	destilled water
RFLP	restriction fragment length polymorphism
ITS	internal transcribed spacer
kDNA	kinetoplastid DNA
hsp	gene coding for heat shock protein

3 Introduction

3.1 Leishmaniasis

With about 278,000 counted cases per year and an estimate of 0.9-1.6 million cases each year, the leishmaniases are one of the most neglected tropical diseases [1]. The human leishmaniases are endemic to tropic and sub-tropic regions in 98 countries on five continents and are caused by at least 17 different *Leishmania* species [2]. The disease is divided into different forms; cutaneous leishmaniasis (CL) is the least fatal form of the disease with mostly self-healing lesions leaving the patient with scars, but without severe health problems (Figure 1); visceral leishmaniasis (VL) is characterized by the affection of inner organs and is lethal if left untreated; mucocutaneous leishmaniasis (MCL) is present almost exclusively in the New World and includes destructive lesions in mucocutaneous regions like nose, mouth and throat. Post-kala azar dermal leishmaniasis (PKDL) is a sequel of VL presenting with dermal lesions. Annually, 20,000-40,000 people die of the visceral form of the disease[1].

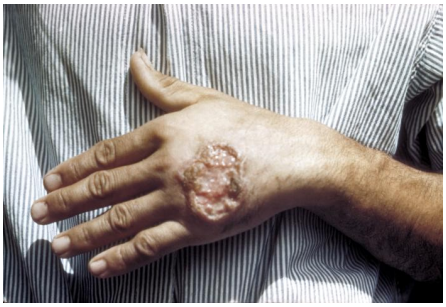

Figure 1. Cutaneous lesion of a patient suffering from CL. Source: Centers for Disease Control and Prevention's Public Health Image Library (PHIL) / Dr. D.S. Martin

The outcome of the disease is largely dependent on the causative species. In the Old World, CL is mainly caused by *L. major*, *L. tropica* and *L. aethiopica*; *L. donovani* and *L. infantum* are the main causative agents of VL.

However, in some cases this correlation between the infecting species and the clinical symptoms is obsolete. There are numerous examples of a usually dermotropic species causing visceral symptoms and vice versa. Strains of *L. tropica* were found to be the causative agent of VL in several foci, often in young children. *L. tropica* has been isolated and identified from two children suffering from VL in Kenya, a 2-year old girl and a 9-year old boy [3]. In Israel, a strain of *L. tropica* was isolated from an infantile case of VL [4]. One strain of *L. tropica* was isolated from the viscera of a rat in Baghdad, Iraq [5]. In Morocco, two cases of canine VL caused by *L. tropica* [6, 7] and several cases of human and canine CL caused by *L. infantum* have been reported [8-10]. In Himachal Pradesh, north-western India, both CL and VL cases were caused by *L. tropica* and *L. donovani* [11-14]. Zymodeme MON-24 is

just one example of *L. infantum*, a species usually causing VL, which has been found to cause CL in humans in countries around the Mediterranean Sea. It was first identified in Algeria [15], and later in Tunisia [16] and Morocco [10]. It was shown to be responsible for 70 % of the Italian cases of CL [17, 18]. An even more unusual case was described in Israel. *L. tropica* was isolated from a dog suffering from MCL in Ma'ale Adumim, a classical CL focus [19].

Leishmania researchers largely agree that the disease manifestation is the result of the interplay of the *Leishmania* species, the host´s immune status and the virulence of the parasite [20-22]. The numerous examples of an unusual species – symptoms correlation indicate the involvement of other factors in this equation. Interestingly, in most cases where *L. tropica* is causing visceral symptoms, the parasites were isolated from either non-human hosts (dog or rat) or young children whose immune system is not yet fully developed. Another hint for the involvement of the immune system is the fact that cases of *L. infantum* in the Mediterranean basin are mostly affecting children. HIV-immunosuppression also influences the disease presentation [23]. In CL cases, the characteristics and distribution of lesions changes and the parasite can also visceralize. Patients initially presented as VL cases can develop cutaneous lesions upon HIV co-infection. Generally, the disease is found to be more severe in immunocompromised patients.

This study focuses on the cutaneous form of leishmaniasis caused by *L. tropica* and *L. aethiopica* and their vectors.

3.2 Distribution and heterogeneity of *L. tropica* and *L. aethiopica*

Parasites of *L. tropica* are endemic to many countries in the Old World. Most of them are found in a stretch through the Middle East, reaching from Greece to Afghanistan (Figure 2). Additional foci are in northern India, Morocco, East Africa and Namibia.

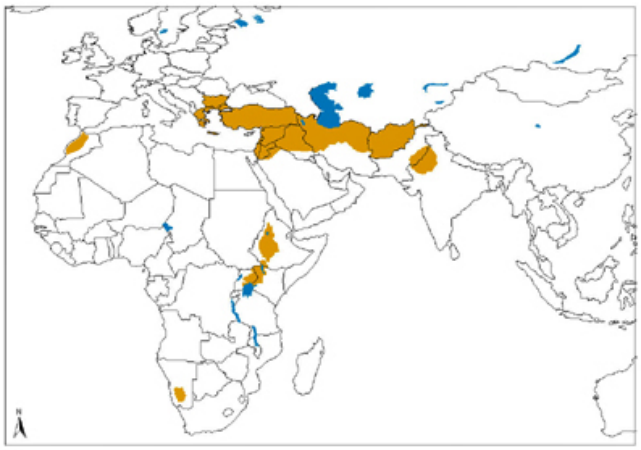

Figure 2. Worldwide distribution of *L. tropica*. Source: http://www.who.int/leishmaniasis/leishmaniasis_maps/en/index1.html

Leishmania tropica is a very heterogeneous species associated with various clinical forms of the disease. In most foci, it is the causative agent of CL, but it has also been isolated from human and canine cases of VL [24-26]. Several studies indicated this species´ great serological, biochemical and molecular heterogeneity. The characterization of the EF (excreted factor) serotypes of strains of *L. tropica* revealed various serological types in several studies [4, 24, 27]. A multilocus enzyme electrophoresis (MLEE) analysis including 1048 strains of the causative agents of Old World cutaneous leishmaniasis found 32 zymodemes within 329 strains of *L. tropica*, while in 638 strains of *L. major* only 12 different zymodemes were identified [28]. Several genetic targets were used to discriminate different strains of *L. tropica*. Schönian *et al.* revealed intra-species variety using different PCR-based techniques, including PCR fingerprinting, ITS-RFLP and ITS-SSCP [29]. Extensive genetic variation was confirmed by Schwenkenbecher *et al.*, who found ten different genetic groups among 117 strains of *L. tropica* from African and Asian foci in a microsatellite typing study [30].

Parasites of the species *L. aethiopica* are geographically restricted to Ethiopia and Kenya. Their neighbouring countries, Djibouti and Eritrea, also report cases of CL, but the causative species has not been identified [1]. In Ethiopia, *L. major*, *L. tropica* and *L. aethiopica* are the causative agents of the estimated annual 50,000 cases of CL. Rarely, *L. donovani* causes PKDL, usually the causative agent of VL [1]. Three clinical forms of the leishmaniases are present in Ethiopia, localised CL, diffuse CL and mucosal leishmaniasis, all mainly caused by *L. aethiopica*. The MLEE study of Pratlong *et al.* revealed a very heterogenic nature of *L. aethiopica* [28]. Among the small sample set of 40 strains, they found 23 different zymodemes.

3.3 Life cycle of the parasites

The life cycle of the *Leishmania* parasites includes mammalian hosts and sand fly vectors (Figure 3). The infective promastigote forms of the leishmanial parasite are transmitted to the mammalian host by the bite of a female sand fly and phagocytised by macrophages. In the macrophages, they transform into amastigotes and multiply. During the next blood meal of the sand fly, the macrophages containing the amastigotes are ingested and those amastigotes are transformed again into promastigotes and amplified in the mid gut of the vector.

It has long been assumed that leishmanial transmission is either anthroponotic or zoonotic, depending on the species. For example, *L. tropica* was considered to be an anthroponotic species, which is transmitted from human to human. In bigger cities this scenario is possible, but it is rather unlikely in rural environments, where the density of humans is low and CL is only distributed focally. In these areas, animal hosts are suspected to be involved as reservoirs. In Israel, *L. tropica* was isolated from rock hyraxes (*Procavia capensis*) [31], domestic dogs [19], golden jackals (*Canis aureus*) and red foxes (*Vulpes vulpes*) [32]. Also in other regions *L. tropica* was isolated from different

mammals, indicating that these may play a role in the leishmanial life cycle. Several strains of *L. tropica* were isolated from canine cases of VL in Greece [33], Iran [34], Morocco [6] and Turkey [35]. Dogs with cutaneous symptoms were found in Morocco [8], one dog presented with concurrent mucosal and visceral involvement in Iran [36] and several dogs with cutaneous symptoms in Morocco. Rats were also found infected by *L. tropica,* e.g. in Iraq [5] and Kenya [37]. The rock hyrax *P. capensis* was incriminated to be the reservoir not only in Israel, but also in Kenya and Namibia [2].

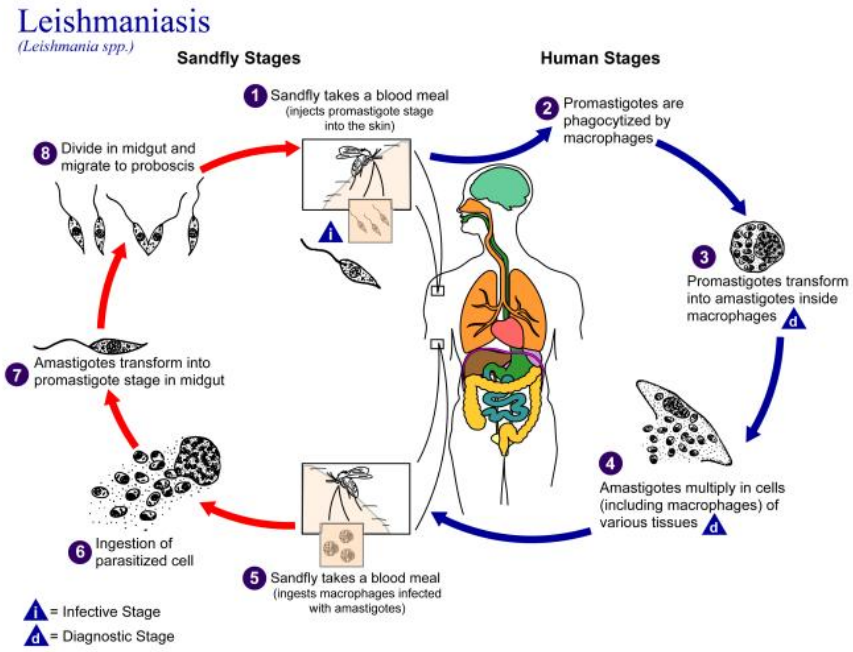

Figure 3. Life cycle of the *Leishmania* parasites. The cycle involves an amastigote form in the macrophages of the mammalian host and the transformation into promastigotes in the sand fly vector. Source: Centers for Disease Control and Prevention's Public Health Image Library (PHIL) / Alexander J. da Silva, PhD, Blaine Mathison

Parasites of *L. aethiopica* have been isolated from the three hyrax species *P. capensis*, *Heterohyrax brucei* and *Dendohyrax arboreus* and the giant pouch rat *Cricetomys sp.*, the former two having been confirmed as their natural reservoir hosts [1, 2, 38].

3.4 Phlebotomine sand flies as vectors

The *Leishmania* parasites are transmitted by sand flies of the genus *Phlebotomus* in the Old World and *Lutzomyia* in the New World (Figure 4). Some sand fly species transmit only parasites of one *Leishmania* species, while others are permissive and able to transmit more than one species. Of the *Leishmania* species causing CL in the Old World, *L. major* is mainly transmitted by *P. papatasi* in the Middle East and northern Africa and by *P. duboscqi* in Africa, but has also been isolated from *P. bergeroti*, *P. salehi*, *P. caucasicus*, *P. ansaril* and *P. mongolensis* [2]. *L. tropica* is mainly transmitted by *P. sergenti* throughout its worldwide distribution, but was also detected in specimens of *P. arabicus*, *P. guggisbergi*, *P chabaudi*, *P. rossi*, *P. saevus* and *P. grovei*. The zoonotic life cycle of

L. aethiopica includes *P. longipes*, *P. pedifer* and *P. sergenti* as the vectors in Ethiopia and *P. pedifer* and *P. aculeatus* as the vectors in Kenya.

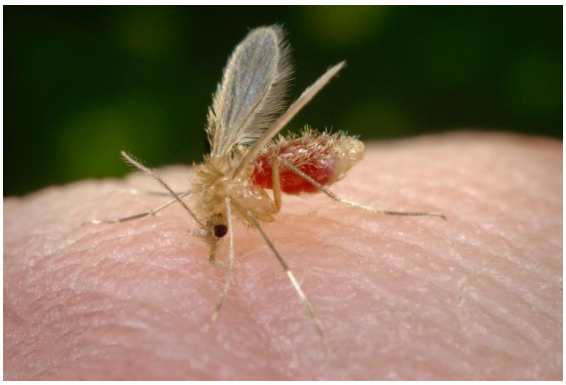

Figure 4. Female sand fly during its blood meal. This figure shows a sand fly of the species *P. papatasi*. Source: Centers for Disease Control and Prevention's Public Health Image Library (PHIL) / Frank Collins

Several sand fly species are endemic to the Middle East, one important focus of this study. Those sand fly species which play a role in the transmission of *L. tropica*, *L. major* and *L. infantum* which are the leishmanial species endemic in the Middle East are listed in Table 1, data summarized from [2].

Table 1. Distribution of the sand fly species in the Middle East and the *Leishmania* species they transmit.

sand fly species	country	*Leishmania* species
P. papatasi	Afghanistan, Egypt, Iran, Iraq, Israel, Jordan, Pakistan, Palestine, Saudi Arabia, Syria, Yemen	*L. major*
P. sergenti	Afghanistan, Iran, Iraq, Israel, Jordan, Palestine, Saudi Arabia, Syria, Turkey, Yemen	*L. tropica*
P. tobbi	Cyprus, Israel, Palestine, Syria, Turkey	*L. infantum*
P. syriacus	Israel, Lebanon, Palestine, Syria, Turkey	*L. infantum*
P. perfiliewi	Israel, Palestine	*L. infantum*
P. alexandri	Iraq, Turkey	*L. donovani*, *L. infantum*
P. caucasicus	Afghanistan, Iran	*L. major*
P. salehi	Iran	*L. major*
P. neglectus	Turkey	*L. infantum*
P. arabicus	Israel	*L. tropica*
P. orientalis	Yemen	*L. donovani*
P. ansaril	Iran	*L. major*
P. bergeroti	Yemen	*L. major*
P. duboscqi	Yemen	*L. major*
P. galilaeus	Syria	*L. infantum*
P. halepensis	Syria	*L. infantum*
P. kandelakii	Iran	*L. infantum*
P. langeroni	Egypt	*L. infantum*

3.5 Cutaneous leishmaniasis in the main foci of this study

3.5.1 Israel and the Palestinian Authority

In Israel and Palestine, CL in humans is caused by three different species, *L. major* and *L. tropica* and, rarely, *L. infantum*. The annual incidences of CL in Israel and Palestine are shown in Figure 5 and Figure 6.

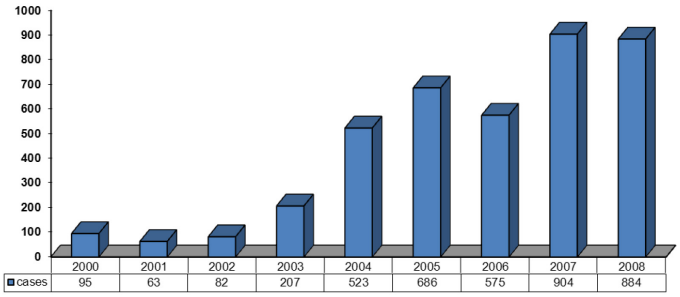

Figure 5. Cutaneous leishmaniasis trend in Israel. The cases of CL, disregarding their causative species, are plotted against the year from 2000 to 2008. Source: [1]

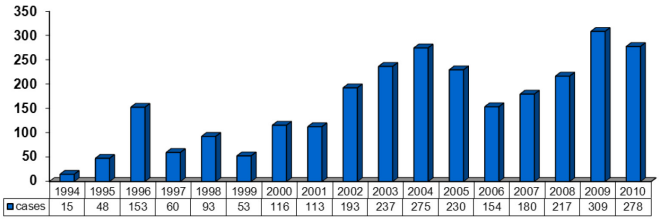

Figure 6. Cutaneous leishmaniasis trend in Palestine. The cases of CL, disregarding their causative species, are plotted against the year from 1994 to 2010. Source: [1]

While CL caused by *L. major* is found predominantly in semi-arid environments, infections caused by *L. tropica* used to be restricted to the more hilly and mountainous areas. During the last decades, studies documented the rapid geographical expansion and drastic increase of cases of CL caused by *L. tropica* in urban and rural areas. New endemic foci emerged in Tiberias, Ma'ale Adumim and Pedu'el in Israel [1, 31, 39] as well as in the Palestinian Jenin District and in Jericho in the West Bank [40, 41].

The main vector of *L. tropica* in Israel and Palestine is *P. sergenti*. In a focus at the northern side of the Sea of Galilee *P. arabicus* was found to be the preferred vector. There, rock hyraxes of the species *P. capensis* are the natural hosts of the parasites [42].

3.5.2 Morocco

In Morocco, CL is caused by three species of *Leishmania*, *L. major*, *L. tropica* and *L. infantum*. Although the disease is still suspected to be underreported, the incidence rates have constantly increased during the last decade ([1], Figure 7).

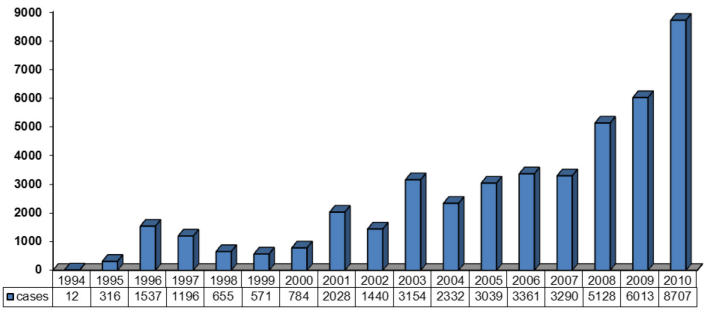

Figure 7. Cutaneous leishmaniasis trend in Morocco. The cases of CL, disregarding their causative species, are plotted against the year from 1994 to 2010. Source: [1]

The majority of the cases of CL are due to *L. major* and *L. tropica*. Those cases caused by *L. infantum* are rare and restricted to the north of the country; *L. infantum* was the infectious species of several human [9] and at least one case of canine CL [8]. *L. infantum* is usually the cause for VL in the Mediterranean area [9, 43]. In the desert region south of the Atlas Mountains, *L. major* is found to be the main causative agent of CL [1]. *L. tropica* is the most widely spread species in Morocco, its distribution stretches from the Atlantic Ocean along the Atlas Mountains almost to the Mediterranean Sea [9]. The infections caused by this species have been considered a major public health threat in Morocco since 1997 [1]. In some cases, the disease outcome is unusual; two cases of canine VL and seven cases of canine CL caused by *L. tropica* have been reported [6-8]. The transmitting vectors of *L. tropica* in Morocco are *P. sergenti* and *P. chabaudi*, which are anthropophilic vectors [1, 44, 45].

3.5.3 India

In India, the vast majority of the *Leishmania* infections, more than 10,000 cases annually, are due to *L. donovani*, resulting in VL [46]. Half of the worldwide burden of VL is found in the state of Bihar, north-eastern India. Three species have been found to cause CL in India: *L. major*, *L. tropica* and *L. donovani* [1]. Human CL caused by *L. tropica* and *L. major* is predominantly found in the north-western states of India, Rajasthan and Punjab [1, 47]. Other CL foci are reported in Himachal Pradesh northern India [11], and Kerala State, south-west India [48].

Bikaner City in Rajasthan State is the main focus of *L. tropica* in India [47, 49]. The first cases were reported in 1971 [50] and the causative species was shown to be *L. tropica* and *P. sergenti* being its vector [51, 52]. One strain of *L. tropica* was isolated from a dog [53, 54]. The number of cases increased from 2001 to 2009, but has decreased now between 2010 and 2011 [49]. In Himachal Pradesh, strains of *L. tropica* and *L. donovani* were both identified as the causative species for 161 cases of CL [11] and one case of VL had been caused by *L. tropica* [12]. Another case of VL due to *L. tropica* has been found in Kolkata [13, 14]. Several cases of CL were diagnosed in Kerala, south-west India [48]. The causing parasite is unknown although there are indications suggesting

L. donovani as the causative species; sand flies collected nearby were identified as *P. argentipes*, the usual vector of *L. donovani* in India. Furthermore, strains isolated from human cases of CL in Sri Lanka were identified as *L. donovani* [55].
No animal reservoir could be identified in India, suggesting the transmission to be anthroponotic [1]. *P. sergenti* is the vector transmitting *L. tropica* in India.

3.6 Microsatellite typing as tool for genetic analyses

Microsatellite markers are oligonucleotide repeats of 1-6 bp in different regions of the DNA of most taxa. Microsatellites evolve co-dominantly, are highly variable and, thus, enable differentiation of strains at an intraspecific level. The mutation rates of microsatellites are considerably higher than in other regions of the genome and range from 10^{-6} to 10^{-2} per generation [56, 57]. Microsatellites have proven to be able to estimate the relatedness of individuals and to determine the population structure of different groups of individuals and given the high mutation rates compared to other genetic markers, to be informative even for populations which have gone through genetic bottlenecks recently [58]. Size variability within microsatellite loci is mainly caused by slippage of the DNA polymerase during DNA replication and, possibly, corroborated by an inefficient mismatch repair system leading to stepwise mutation [57]. Recombination events such as unequal crossing over and gene conversion can cause non-stepwise mutation of microsatellites, although this is estimated to happen to a much lesser extent than the stepwise mutation.

Microsatellite markers are prone to homoplasy. Homoplasy describes the phenomenon when alleles are of different lineages, but have the same size. It is not possible to determine whether a certain change in repeat number is due to deletion or insertion events [59]. Thus, microsatellite typing is not able to reveal the directed evolution of the strains. To minimize the effect of homoplasy a considerable number of markers is required for microsatellite typing in order to gain sufficient data from informative loci. The number of included markers, between 10 and 20, is always a compromise between the number required for a sufficient discriminatory power on the one hand and the economics of labour and costs on the other hand. However, microsatellites have been successfully applied to reveal the population structures and identify genetic groups among strains of several *Leishmania* species [60]. Microsatellite markers have proven more discriminatory than isoenzyme typing by MLEE; e.g. the strains identified as isoenzyme type MON-1 were successfully subdivided by microsatellite typing [61]. The microsatellite marker set is usually species-specific, although single markers are able to differentiate different species. Sets of microsatellite markers have been developed for the *L. donovani* complex [62], *L. major* [63], *L. tropica* [64], and *L. aethiopica* [65] in the Old World as well as for the *Viannia* sub-species in the New World [66]. There have been no inter-species analyses, except for strains of *L. infantum* and *L. donovani*, which are both parts of the

L. donovani species complex [62, 67]. The marker set developed for *L. aethiopica* comprises, however, only five polymorphic markers, thus lacking sufficient statistical support of the results. Previous microsatellite typing studies have shown that population structures and genetic distances calculated based on the microsatellite data are largely in correlation to the strains´ geographical origins [30, 62, 63].

An earlier microsatellite typing study including 117 strains of *L. tropica* from 12 different Asian and African countries has assigned these strains to ten genetic clusters [30]. The strains largely clustered according to their geographical origins in an analysis based on Bayesian statistics. Thus, these clusters were named according to their strains´ geographical origins; Middle East, Asia, Adana, Galilee, MoroccoA, MoroccoB, Tunisia, KenyaA, KenyaB and Namibia. In general, these clusters were confirmed by genetic distance analyses, e.g. by constructing a Neighbour Joining (NJ) tree. The marker set developed for *L. tropica* comprised 21 markers [64], making the typing very cost-intensive and time- and labour-consuming. Some of these markers had been originally developed for *L. major* and *L. donovani* [63, 68, 69].

3.7 An impaired mismatch repair system as possible cause for intraspecific heterogeneity

One possible explanation for the great genetic variability within the species *L. tropica* is an impaired mismatch repair (MMR) system. The components of the MMR machinery are highly conserved in all organisms [70]. In prokaryotic cells, MutS recognizes the mismatch; MutL and MutH bind to the marked site and force the DNA into a conformational change (Figure 8). Upon incision of the erroneous strand by the methylation-sensitive endonuclease MutH, the mismatch is cut out and re-synthesized.

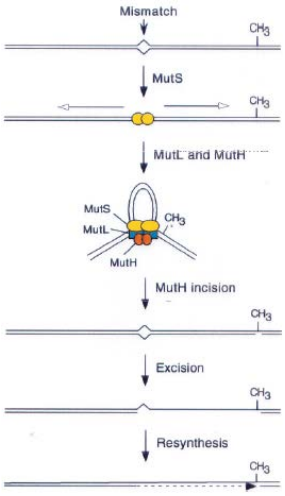

Figure 8. MMR system in bacterial cells. Upon recognition of the mismatch by MutS proteins, the conformational change of the DNA is conducted by MutL. MutH cuts the mismatched strand and the introduced gap is filled by the polymerase. In eukaryotic cells, MSH and MLH proteins take over the tasks of MutS, MutL and MutH as described in [71].

In eukaryotic cells, like *Leishmania sp.*, the process is similar, although more sophisticated and not yet fully elucidated. It involves enzymes, which are homologs to their prokaryotic counterparts, thus called MutS-like homologs, MSH, and MutL-like homologs, MLH. Harpe *et al*. summarize the enzymes involved in the MMR system in several eukaryotic organisms (Table 2) [71]. Although there are slight differences in the composition of the involved proteins, several MSH and MLH proteins are involved in the repair of the DNA with specialised functions in various forms of DNA damage. In the yeast *Saccharomyces cerevisiae* as described in [71], MSH2 is supposed to be essential for the mismatch recognition. In combination with MSH3 and MSH6 it binds to insertion/deletion loops exclusively or to insertion/deletion loops and base-base mismatches, respectively [72, 73]. All yeast MSH proteins possess ATPase binding activity. MSH4 and MSH5 play a critical role in the recombination during meiosis, neither of the two has a detectable role in mismatch repair [74, 75]. MLH1 is the central MutL homolog in yeast forming heterodimers with PMS1, MLH2 and MLH3 for mismatch repair [76]. The proteins involved in MMR have not been described in *Leishmania* sp. So far, *Trypanosoma* are the closest relative of *Leishmania* which have been studied for their MMR genes. Bell *et al*. found three MutS-homologs (MSH2, MSH3 and MSH8) and two MutL-homologs (PMS1 and MLH1) in the genome of *T. brucei* by BLAST search and confirmed their role in mismatch repair [77].

Table 2. Eukaryotic Mut Sand MutL homologs. Source: [71]

E. coli	Yeast	Mammals	Worms	Flies	Plants	Eukaryotic Functions[a]
MutS	MSH1	—[b]	—[b]	—[b]	MSH1	Mutation avoidance in mitochondria
	MSH2	MSH2	MSH2	SPEL1	MSH2	Forms heterodimers with MSH3 and MSH6 to: Repair replication errors Repair mismatches in recombination intermediates Remove nonhomologous tails (MSH2-MSH3 only) Inhibit recombination between nonidentical sequences Respond to DNA damage (mammals)
	MSH3	MSH3	—[b]	—[b]	MSH3	Forms heterodimers with MSH2
	MSH4	MSH4	MSH4	—[b]	MSH4	Forms heterodimers with MSH5 to promote crossing-over in meiosis
	MSH5	MSH5	MSH5	—[b]	—[b]	Forms heterodimers with MSH4 to promote crossing-over in meiosis
	MSH6	MSH6	MSH6	MSH6	MSH6 MSH7	Forms heterodimers with MSH2
MutL	PMS1	PMS2	PMS1	PMS2	PMS2	Forms heterodimers with MLH1 to: Repair replication errors Repair mismatches in recombination intermediates Inhibit recombination between nonidentical sequences Respond to DNA damage (mammals)
	MLH1	MLH1	MLH1	MLH1	MLH1	Forms heterodimers with PMS1, MLH2 and MLH3
	MLH2	PMS1	—[b]	—[b]	—[b]	Forms heterodimers with MLH1 to: Repair replication errors Repair mismatches in recombination intermediates
	MLH3	MLH3	Ced H12	—[b]	MLH3	Forms heterodimers with MLH1 to: Repair replication errors Promote crossing-over in meiosis

The MMR system has multiple functions in the diverse DNA repair mechanisms. It does not only repair nucleotide mismatches and small insertions and deletions, it also plays an important regulatory role in homologous recombination [71]. Mismatches occurring during the recombination process are repaired leading to the genetic phenomenon of gene conversion. The MMR proteins also have a regulatory effect on the recombination between non-identical sequences during mitosis and

meiosis. This prevents the recombination of the DNA of different species, thus having an impact during the speciation process. The recombination efficiency was increased by the inactivation of the *MSH2* gene in a study generating *L. infantum* / *L. major* hybrids [78]. Organisms which lack one or more important components of the MMR machinery exhibit a mutator phenotype, resulting in microsatellite instability in bacteria [79], yeast [80] and mammals [81].

Initial studies have shown that the nucleotide sequences of the *MLH1*, *MSH2*, *MSH4* and *MSH5* genes are more diverse within *L. tropica* than in *L. major* [82, 83]. Figure 9 shows the maximum parsimony trees calculated on the basis of the nucleotide sequences of *MSH2* and *MLH1*. The base substitutions at DNA level were shown not to be silent, but also had impact on the deduced amino acid sequences, thus leading to parsimony trees of the deduced amino acid sequences with greater diversity within *L. tropica* than within *L. major*. The great genetic heterogeneity among the MMR genes within the species *L. tropica* led to the idea that an impaired MMR system might play a crucial role in the diversification of this species.

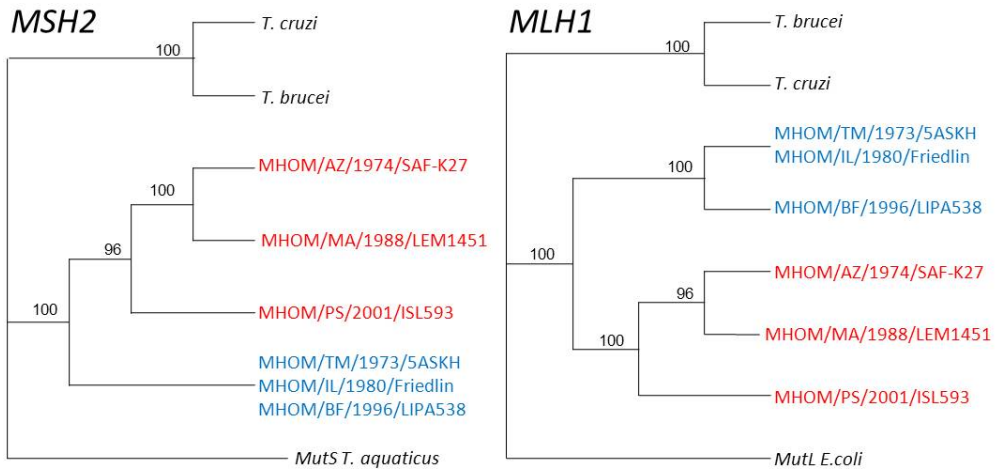

Figure 9. Maximum parsimony tree of nucleotide sequences of the *MSH2* and *MLH1* genes. Strains of *L. tropica* are highlighted in red; strains of *L. major* are blue. *Trypanosoma brucei* and *T. cruzi* are included for comparison and the prokaryotic homologs MutS of *T. aquaticus* and MutL of *E. coli* were used as outgroup. The bootstrap values are indicated at the nodes. Modified after [82]

4 Objectives

Leishmania (*L.*) *tropica* is causing mainly cutaneous leishmaniasis in the Old World and a very heterogeneous species as revealed by several serological, biochemical and molecular methods. The reasons for this considerable heterogeneity are not known. *L. aethiopica* is endemic to a locally very restricted area in East Africa and genetically close to *L. tropica*. Depending on the applied approach, it was found most closely related to either *L. tropica* or *L. major*.

The main aim of this study was to explore further the extensive genetic heterogeneity of strains of *L. tropica* and to elucidate possible mechanisms underlying this immense diversity. In addition, it was attempted to get insights in the relationship between strains of *L. tropica* and of *L. aethiopica,* and to develop a molecular assay for the differentiation of the *Phlebotomus spp*. in pools of sand flies caught during field studies in CL foci.

- Multilocus microsatellite typing (MLMT) has been proven to be a good tool for molecular epidemiological studies of *Leishmania sp*. This work intended to draw a global population study summarizing all previously published microsatellite data and adding new strains of various geographical origins. Particular objectives were
    - To expose the global epidemiology of *L. tropica* and *L. aethiopica* in correlation to the strains´ origins, times of isolation, hosts and disease manifestations.
    - To optimize the previously used microsatellite marker set for *L. tropica* by excluding, adding and modifying several markers. The set of markers was re-evaluated and customized for this study. The number of markers was reduced due to convenience reasons and cost-effectiveness.
    - To reveal the population genetics of *L. tropica* in the Middle East, particularly in Israeli and Palestinian foci.
    - To investigate the inter- and/or intra-focal diversity of *L. tropica* in Morocco by analysing the microsatellite profiles of strains isolated from human cases of CL from Marrakech and Chichaoua Province.
    - To reveal a possible correlation between the genetic make-up of the parasite and the disease outcome by comparing Indian strains of *L. tropica* isolated from CL cases to those isolated from visceral cases.
    - To apply the microsatellite marker set to *L. aethiopica* and reveal their phylogenetic position in comparison to 195 strains of *L. tropica*.
- Different scenarios could possibly explain the high diversity within the species *L. tropica*: enhanced mutation and recombination rates, and an impaired mismatch repair (MMR)

system. The mutation rate is too low ($1.99x10^{-9}$ per bp per generation) for being the exclusive cause of the high heterogeneity; also recombination seems to be an event too rare to sufficiently explain the high genetic diversity within this species. Initial studies have shown that the nucleotide and their deduced amino acid sequences of genes involved in different steps of the MMR (*MSH2*, *MLH1*, *MSH4* and *MSH5*) are very variable in different strains of *L. tropica* compared to those in strains of *L. major*. This study aimed

  - To reveal the genetic mechanisms underlying the large genetic diversity of *L. tropica* by a Complementation Assay. The genomic DNA of one strain of *L. tropica* showing low mutation tolerance, indicating a functioning MMR system, was cloned into a strain of *L. tropica* showing high mutation tolerance and the resulting transfectants were screened for the DNA sequence which restored the donor phenotype. These stretches of DNA were supposed to encompass ORFs responsible for the low mutation tolerance thus indicating a functioning MMR system as the cause for the high level of intra-species genetic variation within the species *L. tropica*.

- In Israel and Palestine, the main focus of this work which was performed in the frame of a trilateral project including Israeli, Palestinian and German collaborators, three *Leishmania* species cause leishmaniasis, *L. major*, *L. tropica* and *L. infantum*. These species are transmitted by different phlebotomine sand fly vectors. Identifying the phlebotomine species within particular foci gives a clue to the *Leishmania* species endemic to these foci and their distribution within certain areas. To date thousands of collected sand flies in the field have to be evaluated morphologically, which is very time-consuming and needs good training of the scientific staff. There are PCR based methods, which target different sequences in the DNA of the sand flies. These PCR techniques are able to detect and identify only one fly at a time. There is a need to make this process more efficient. Multiplex ligation-dependent probe amplification (MLPA) is a molecular technique able to detect and identify or quantify microorganisms. The aim of this study was
  - To develop an MLPA approach targeting species-specific SNPs in the 18S rRNA sequence for the differentiation of phlebotomine sand fly species caught in different Israeli and Palestinian leishmaniasis foci. The 18S rRNA sequence is the most frequently published sequence in the NCBI database and, thus, was available for many phlebotomine species.

5 Materials and Methods

5.1 Materials

5.1.1 Devices

device	***manufacturer***	***model***
centrifuge	Heraeus Sepatech	Biofuge 28RS
rotor	Heraeus Sepatech	#3744 HFA 15.16
		#3751 HFA 5.50
centrifuge	eppendorf	5415D
centrifuge	eppendorf	5415R
photometer	eppendorf	BioPhotometer
waterbath	bio-med	Thermocycler 60
microliter cuvette	Implen	LabelGuard Microliter Cell
counting chamber		Neubauer Improved
microscope	Olympus	BH-2
shaker (for *E. coli* culture)	Sartorius	Certomat U
incubator (for *E. coli* culture)	Sartorius	Certomat HK
PCR cycler	analytikjena	FlexCycler
PCR cycler	Stratagene	Robocycler Gradient 96
incubator (for *Leishmania* culture)	Friocell	Friocell 222 - Comfort
gel documentation system	Syngene	G:BOX
concentrator	eppendorf	5301
power supply	Biotec-Fischer	Phero-stab. 500
		Phero-stab. 300
	Consort	E143
pH meter	WTW	pH 537
FACScan	Becton Dickinson	

5.1.2 Chemicals

Chemicals for buffers and solutions

sodium chloride (NaCl)	Merck
ethylene diamine tetraacetic acid (EDTA)	Roth
Tris(hydroxymethyl)aminomethan (Tris)	Roth
hydrochloric acid (HCl)	Merck
boric acid (H_3BO_3)	Roth
glycerol	Roth
bromophenol blue	Serva
OrangeG	Sigma
sodium hydroxide (NaOH)	AppliChem
acetic acid ($C_2H_4O_2$)	Merck

phosphate buffered saline (Dulbecco´s PBS)	PAA
magnesium chloride hexahydrate ($MgCl_2$)	Merck
potassium acetate	Merck
potassium chloride (KCl)	Roth

DNA extraction

sodium dodecyl sulfate (SDS)	Fluka
RNase from bovine pancreas	Sigma
Proteinase K	Merck
Triton X-100	Fisher Scientific
phenol: chloroform: isoamyl alcohol (25:24:1, v/v)	Roth
chloroform	Merck
isoamyl alcohol	Roth
sodium acetate trihydrate (Na-acetate)	Merck
96% ethanol	J.T.Baker
isopropanol	Merck

PCR, restriction enzymes and fragment analysis

AmpliTaq + 10x PCR buffer	Applied Biosystems
dNTP mix	GE Healthcare
fluorescence-labelled primers (MLMT)	Sigma-Aldrich
primers (unlabelled)	TIB Molbiol
aqua *ad iniectabilia*	Braun
dimethyl sulfoxide (DMSO)	Roth
restriction enzymes (HaeIII, XbaI, EcoRV)	New England Biolabs
X059-A1-sandfly probe mix	MRC-Holland
ligase-65	MRC-Holland
MLPA buffer	MRC-Holland
ligase buffers A and B	MRC-Holland
SALSA PCR primer mix	MRC-Holland
SALSA polymerase	MRC-Holland
GeneScan™ 500 ROX™ Size Standard	Applied Biosystems

Electrophoresis

Agarose NA	GE Healthcare
Metaphor agarose (small fragments)	Biozym
SeaKem LE agarose (large fragments)	FMC Bioproducts
1kb DNA Extension ladder	invitrogen
HyperLadderV	Bioline
Lambda DNA	New England Biolabs
HDGreenPlus	iNTAS
Ethidium Bromide	Roth

Cell culture

Medium M199	invitrogen
fetal bovine serum (FCS)	invitrogen
Hepes	Roth
sodium hydrogen carbonate ($NaHCO_3$)	Merck
Gentamycin	invitrogen
hemin, bovine	Sigma
ethylendiamin-N,N,N',N'-tetra-2-propanol (EDTP)	Merck
Adenin	Sigma
6-Biopterin	Sigma
L-Glutamin / PenStrep	Sigma
Geneticindisulfat (G418)	Roth
paraformaldehyd	Merck
Cryo-SFM	Promocell
N-methyl-N'-nitro-N-nitrosoguanidin (MNNG)	TCI Europe
Caltag Counting Beads	invitrogen
tryptone / peptone from casein	Roth
yeast extract	Roth
agar-agar	Roth
Ampicillin	Roth
glycerol	Roth
calcium chloride dihydrate ($CaCl_2$)	Roth
E. coli XL1-Blue	Stratagene
pUC18	Stratagene

5.1.3 Buffers, solutions and cell culture media

Lysis buffer

50 mM NaCl
10 mM EDTA (pH 8.0)
5 mM Tris-HCl (pH 7.4)
ad 500 ml aqua dest.

Tris-HCl

Molecular weight: 121.14 g/mol
100 mM → 12.114 g/l
pH to 7.4 or 8.0 with HCl

TE buffer

10 mM Tris-HCl (pH 7.4)
1 mM EDTA (pH 8.0)

TBE buffer (10x)

450 mM Tris-HCl
450 mM boric acid
10 mM EDTA (pH 8.0)
ad 2 l aqua dest.

5x DNA loading buffer

30% glycerol
50 mM Tris-HCl (pH7.4)
5 mM EDTA
0.1% bromophenolblue
Storage at 4 °C (for daily use) or -20 °C (long-term)

RNase solution

100 mg RNase
10 mM Tris-HCl (pH7.4)
15 mM NaCl
ad 10 ml aqua dest.
Storage at -20 °C

Modified M199 medium for Leishmania cell culture

component	stock solution	final concentration	preparation of stock solution
M199	10x	1x	
inactivated FCS	100%	20%	30 min. at 56 °C
Hepes, pH 7.4	0.8 mM	40 mM	95.32 g / 500 ml
$NaHCO_3$	2%	0.0352%	4 g / 200 ml
Gentamycin	50 mg/ml	0.02 mg/ml	
Hemin	5 mg/ml	0.01 mg/ml	0.25 g + 25 ml aq. dest. + 25 ml EDTP → 37 °C overnight
Adenine	0.5 mg/ml	0.0135 mg/ml	0.1 g / 200 ml
6-Biopterin	2 mg/ml	0.0012 mg/ml	50 mg + 1 ml 1 M HCl, ad 25 ml
L-Glutamin/PenStrep	200 mM L-Glutamin 10000 U/ml Penicillin 10 mg/ml Streptomycin	2 mM 100 U/ml 0.1 mg/ml	
aqua dest.			

Geneticindisulfat (G418) was used as selection antibiotic for cells containing the cosmid pcosTL (resistance gene: neo). The final concentration was 10 µg/ml.

Fixation solution

4 g paraformaldehyd
50 ml aqua dest.
1 ml 1 M NaOH
dissolve paraformaldehyde at 60 °C
10 ml 10x PBS
set pH to 7.4 at room temperature
ad 100 ml aqua dest.

LB medium for E. coli cultures

20 g tryptone/peptone from casein
10 g Yeast extract
20 g NaCl
ad 2 l aqua dest.

For solid medium, 7.5 g agar-agar was added, autoclaved and poured into petri dishes.
Ampicillin was used as antibiotic for the selection of cells with cosmid pcosTL. The final concentration was 75 µg/ml.

SOC medium

1 g tryptone/peptone from casein
0.25 g yeast extract
8.5 mM NaCl
2.5 mM KCl
10 mM $MgCl_2$
20 mM glucose
ad 50 ml aqua dest.
→ pH 7.0
→ aliquot à 1 ml
Storage at 4 °C

Cosmid preparation

Solution1

50 mM glucose
10 mM EDTA
25 mM Tris-HCl (pH 8.0)
Storage at 4 °C

Solution2

0.2 N NaOH
1% SDS
Storage at room temperature

Solution 3

3 M potassium acetate
2 M acetic acid
Storage at 4 °C

5.2 Multilocus Microsatellite Typing (MLMT)

5.2.1 Ethical statement

For this study, 89 *Leishmania* strains were collected from different health care centres mainly in Israel, Palestine and other countries in the Middle East, in India, Morocco and Ethiopia. They were all collected according to the respective institutional ethical guidelines. Sampling and the use for scientific projects were approved by their ethical committees. Data from additional 114 strains was included in the analysis for comparison which had been published earlier together with the respective ethical clearances.

5.2.2 Parasite strains

All strains of *L. tropica* and *L. aethiopica* typed for this study are listed in Table 3, stating their WHO codes, geographical origins and source of isolation. Most of the strains were isolated from human cases of CL (n=177). The strain set also included eight strains isolated from human cases of VL, three strains isolated from other mammalian hosts (two from hyraxes of the species *P. capensis* and one from a rat) and 15 strains isolated from sand fly vectors, eight from *P. sergenti*, five from *P. arabicus* and two from *P. rossi*. Upon sampling the strains were cultivated either in liquid or semi-solid media and cryopreserved until DNA isolation.

Strain MHOM/PS/2001/ISL590, whose microsatellite loci had been sequenced [64], was used as reference strain in all PCR and fragment analysis runs.

Table 3. Strains of *L. tropica* and *L. aethiopica* used in this study. The WHO code gives information about the strain´s host as well as the country and the year of its isolation. The column "geographic origin" specifies the focus of the strains as precise as it is known. The column "source" indicates the host and, if human, the form of the disease from which the samples were collected. The eight strains of *L. aethiopica* were all collected in Ethiopia.

WHO code	geographic origin	source
MHOM/PS/2002/31JnM17	Tubas, Jenin District, Palestine	human CL
MHOM/PS/2002/35JnF45	El-Yamoon, Jenin District, Palestine	human CL
MHOM/IL/2003/LRC-L999	Jordan Valley, Israel	human CL
MHOM/PS/2003/ISLAH721	Jordan Valley, Palestine	human CL
MHOM/PS/2003/151JnF32	Tayaseer, Jenin District, Palestine	human CL
MHOM/PS/2003/185JnM27	Serees, Jenin District, Palestine	human CL
MHOM/PS/2003/152JnF32	Tayaseer, Jenin District, Palestine	human CL
MHOM/PS/2003/184Jn01	El-Yamoon, Jenin District, Palestine	human CL
MHOM/PS/2003/186JnM12	Beer Al Basha, Jenin District	human CL
MHOM/PS/2003/161JnF80	Al-Shuhada, Jenin District, Palestine	human CL
ISER/IL/2004/LRC-L1167	Tiberias, Israel	*P. sergenti*
MHOM/IL/2009/LRC-L1374	Anatot, Israel	human CL
MHOM/PS/2002/ISL698	Jericho, Palestine	human CL
MHOM/PS/2001/ISL572	Jericho, Palestine	human CL
MHOM/PS/2003/149JnM9	El-Yamoon, Jenin District, Palestine	human CL

WHO code	geographic origin	source
MHOM/PS/2003/163JnM30	El-Yamoon, Jenin District, Palestine	human CL
MHOM/IL/2003/LRC-L1024	Tiberias, Israel	human CL
MHOM/PS/2003/178JnM75	El-Yamoon, Jenin District, Palestine	human CL
MHOM/IL/2003/LRC-L1021	Tiberias, Israel	human CL
MHOM/PS/2002/41JnF12	Silat El-Hartheyia, Jenin District, Palestine	human CL
MHOM/IL/2003/LRC-L1006	Tzarka Amir, Israel	human CL
MHOM/IL/2009/LRC-L1366	Maale Adumim, Israel	human CL
MHOM/PS/2008/335JnM59	Salit El-Hartheyia, Jenin District, Palestine	human CL
MHOM/IL/2003/LRC-L1026	Maale Adumim, Israel	human CL
MHOM/IL/2009/LRC-L1379	Almon, Israel	human CL
MHOM/PS/2000/ISL535	Jericho, Palestine	human CL
MHOM/IL/2003/LRC-L1001	Maale Adumim, Israel	human CL
MHOM/IL/2003/LRC-L1017	Anatot, Israel	human CL
MHOM/IL/2003/LRC-L1016	Israel	human CL
MHOM/IL/2003/LRC-L1018	Maale Adumim, Israel	human CL
MHOM/IL/2004/LRC-L1105	Maale Adumim, Israel	human CL
MHOM/PS/2004/ISLAH802	Jericho, Palestine	human CL
MHOM/IL/2008/LRC-L1352	Herodion, Israel	human CL
MHOM/IL/2009/LRC-L1362	Nokdim, Israel	human CL
MHOM/PS/2010/391Jn	Jenin District, Palestine	human CL
MHOM/PS/2010/394Jn	Jenin District, Palestine	human CL
MHOM/PS/2010/414Jn	El-Yamoon, Jenin District, Palestine	human CL
MHOM/PS/2010/417JnM13	El-Yamoon, Jenin District, Palestine	human CL
MHOM/PS/2010/433JnJM3	Jenin city, Palestine	human CL
MHOM/PS/2011/448Jn	El-Yamoon, Jenin District, Palestine	human CL
MHOM/PS/2010/LRUJ-1702	Jericho, Palestine	human CL
MHOM/IL/2010/LRC-L1486	Maale Adumim, Israel	human CL
MHOM/PS/2011/ LRUJ-1779	Bethlehem, Palestine	human CL
MHOM/PS/2011/ LRUJ-1781	Bethlehem, Palestine	human CL
MHOM/SU/1966/III	Soviet Union	human CL
MHOM/AF/1988/KK26	Afaghanistan	human CL
MHOM/SY/1995/LSL25	Syria	human CL
MHOM/IR/1995/5043	Iran	human CL
MHOM/TR/1999/EP32	Turkey	human CL
MHOM/IR/2000/LEM4036	Iran	human CL
MHOM/SY/2009/LRUJ-1550	Syria	human CL
MHOM/SY/2009/LRUJ-1564	Syria	human CL
MHOM/PS/2002/63JnF21	El-Yamoon, Jenin District, Palestine	human CL
MHOM/PS/2002/18JnMF4	Meselya, Jenin District, Palestine	human CL
MHOM/IL/2003/LRC-L961	Korazim, northern side of Sea of Galilee, Israel	human CL
MPRV/IL/2003/HYRAX107	Amnum, northern side of Sea of Galilee, Israel	hyrax (*Procavia sp.*)
MHOM/MA/1989/LEM1781	Morocco	human CL
MHOM/MA/1995/UERL9	Morocco	human CL
MHOM/MA/1992/LPN79	Morocco	human CL
MHOM/MA/1995/LEM3015	Morocco	human CL
MHOM/YE/1986/LEM1000	Yemen	human CL

WHO code	geographic origin	source
MHOM/KE/1991/EB135	Kenya	human CL
MHOM/PS/2002/20JnYM3	El-Yamoon, Jenin District, Palestine	human CL
MHOM/PS/2002/52JnYM18	El-Yamoon, Jenin District, Palestine	human CL
MHOM/IN/2006/BKC-1	Bikaner, Rajasthan, India	human CL
MHOM/IN/2006/BKC-2	Bikaner, Rajasthan, India	human CL
MHOM/IN/2007/BKC-10	Bikaner, Rajasthan, India	human CL
MHOM/IN/2007/BKC-11	Bikaner, Rajasthan, India	human CL
MHOM/IN/2007/BKC-15	Bikaner, Rajasthan, India	human CL
MHOM/IN/2007/BKC-28	Bikaner, Rajasthan, India	human CL
MHOM/IN/2006/BKC-3	Bikaner, Rajasthan, India	human CL
MHOM/IN/2006/BKC-5	Bikaner, Rajasthan, India	human CL
MHOM/ET/1994/Gere	Arbaminch, southern Ethiopia	human DCL
MHOM/ET/1994/Abauye	Arbaminch, southern Ethiopia	human DCL
MHOM/ET/1994/1610	Wolega, northern Ethiopia	human LCL
MHOM/ET/1994/1470	Shoa, central Ethiopia	human LCL
MHOM/ET/1994/Wandera	Arbaminch, southern Ethiopia	human LCL
MHOM/ET/1987/Kassaye	Gojam, northern Ethiopia	human DCL
MHOM/ET/1972/L100	northern Ethiopia	human DCL
MHOM/ET/1985/LRC-L494	Ethiopia	human CL
MHOM/MA/2000/INH-W02	Chichaoua province, Morocco	human CL
MHOM/MA/2000/INH-W04	Chichaoua province, Morocco	human CL
MHOM/MA/2000/INH-W05	Chichaoua province, Morocco	human CL
MHOM/MA/2000/INH-W13	Chichaoua province, Morocco	human CL
MHOM/MA/2000/INH-W14	Chichaoua province, Morocco	human CL
MHOM/MA/2000/INH-W16	Chichaoua province, Morocco	human CL
MHOM/MA/2000/INH-W17	Chichaoua province, Morocco	human CL
MHOM/MA/2000/INH-W10	Chichaoua province, Morocco	human CL
MHOM/MA/2000/INH-W09	Chichaoua province, Morocco	human CL

5.2.3 DNA isolation from *Leishmania* strains

The DNA of all *Leishmania* strains was isolated by phenol-chloroform extraction. This method includes steps of cell lysis, DNA extraction, ethanol precipitation and washing. Depending on the type of the sample (cultured promastigotes or smears on microscope slides) the protocol differed slightly.

Cultured promastigotes were harvested by centrifugation at 1260 *g* for 10 min. One to two ml lysis buffer were added to the cell pellet, depending on its size. The cells were re-suspended, SDS added to a final concentration of 0.5% and shaken until the solution became viscous. To remove RNA and all DNA damaging components in the solution, first, RNase was added to a final concentration of 100 µg/ml and incubated at 37 °C for 30 minutes. Then Proteinase K was added to a final concentration of 100 µg/ml and incubated at 60 °C for at least 3 h. The DNA was extracted with an equal volume of phenol: chloroform: isoamyl alcohol (25:24:1, v/v) and gentle mixing for 2-

3 minutes. Centrifugation at 16000 *g* for 10 minutes separated the aqueous from the organic phase. The aqueous phase was transferred to a new tube and the two previous steps were repeated until the interphase disappeared (usually two times was sufficient). An equal volume of chloroform: isoamyl alcohol (24:1, v/v) was added to the aqueous phase, the solution gently mixed and centrifuged as before. To precipitate the DNA, 1/10 volume of 3 M Na-acetate, pH 5.2, and 2-2 ½ volumes of ice-cold 96% ethanol were mixed with the aqueous phase and incubated for at least 1 h (better overnight) at -20 °C. The DNA pellet was centrifuged to the bottom of the tube for 30 minutes at 16000 *g* and washed with 70% ethanol. The supernatant was removed and the pellet dried at room temperature. Finally, the DNA was resolved in 100 µl TE buffer, pH 8.0, and stored at 4 °C. A spectrophotometer was used to measure the DNA concentration. The optical density of the solution at 260 nm (OD_{260}) determined the DNA concentration according to the formula

$$DNA\ concentration = OD_{260} \times dilution\ factor \times 50.$$

The purity of the DNA solution was evaluated using the ratio between the optical densities at OD_{260} and OD_{280}. Pure DNA solutions had a ratio between 1.8 and 2.

The DNA isolation from microscopic slides differed from the above protocol with regard to the initial steps. Lysis buffer was added to the sample adhered to the slide, the tissue material was scraped off with a pipette tip and transferred to a microfuge tube. This step was repeated using 100 µl and 150 µl lysis buffer, respectively. TritonX-100 was added to a final concentration of 1% and ProteinaseK to a final concentration of 100-200 µg/ml and incubated overnight at 60 °C. The steps of DNA extraction and precipitation were identical to those mentioned above in the protocol for DNA isolation from leishmanial cultures. The DNA pellet was resolved in a volume of 50-100 µl TE buffer depending on its size.

5.2.4 *Leishmania* species identification

For the detection and species identification of the collected *Leishmania* samples, a PCR-RFLP approach targeting the ribosomal ITS1 region was used as described in [84]. First, the ITS1 region was amplified by PCR. Each PCR reaction contained 1x PCR buffer, 200 µM dNTP mix, 25 pmol forward primer LITSR (5'-CTGGATCATTTTCCGATG-3') and 25 pmol reverse primer L5.8S (5'-TGATACCACTTATCGCACTT-3'), 2.5 % DMSO, 1 unit Taq polymerase and 2-3 µl DNA in a total volume of 50 µl. The thermo-profile of the PCR started with an initial denaturation step for 2 min at 95°C, followed by 33 cycles of amplification (20 sec at 95°C, 30 sec at 53°C, 1 min at 72°C) and ended with a final extension step for 6 min at 72°C. For the different species of *Leishmania*, the resulting PCR products varied in size between 300 and 350 bp. Depending on the intensity of the fragment in the gel electrophoresis, 10 -20 µl of it were digested with 10 units *Hae*III restriction enzyme in a reaction

containing 1x restriction enzyme buffer as provided by the manufacturer of the enzyme. After 1-2 h of incubation at 37°C, the restriction fragments were separated in a 2 % Metaphor agarose gel and visualized in a UV photo documentation system.

The restriction fragments of *L. tropica* (185, 57, 53 and 24 bp) and *L. aethiopica* (200, 57, 54 and 23 bp) clearly identified these two species since no other species prevalent in the respective range of distribution resulted in the same fragment patterns.

5.2.5 Microsatellite typing

5.2.5.1 Optimization of the microsatellite marker set

The 12 microsatellite markers used in this study are listed in Table 4. The markers GA1, GA2, GA6, LIST7010, LIST7011, LIST7027, LIST7033, LIST7039, LIST7040 and 4GTG were adopted as published in [69] and [64]. The primers for marker GA9 were modified resulting in a longer fragment; this marker was therefore named GA9n. One marker was newly added to the set, 27GTGn. All strains included in this study, also the ones already analysed with the old marker set, were typed for this locus.

Table 4. Microsatellite markers used in this study. The sequences of the labelled forward and unlabelled reverse primer for each marker are given as well as the type of fluorescence label (6FAM, blue; HEX, green) which was tagged to the 5'-end of the primer. The annealing temperature (AT) is specified and both the fragment size and repeat array in the reference strain MHOM/PS/2001/ISL590 are shown (*). The column "ref" refers to the original publication of the repsective marker, and, in case of the markers GA9n and 27GTGn, also to the publication in which their modification was described.

m/sat	fw primer (5´→3´)	rev primer (5´→3´)	fluorescence label	AT (°C)	fragment size (bp)*	repeat array*	ref
GA1	TCGGAGTCACCTCGCACCGC	GGTGGGGCAGGTAAAGCGGC	6FAM	56	66	(GA)11	[64]
GA2	GATCACAGCGACGTCTGAAG	CCTGCTGCCACCATCTTAAGC	HEX	56	62	(GA)8	[64]
GA6	GTGTGAGCTAATCGATTGGG	CGCTCTCTCTGTCTCTGTCT	6FAM	42	61	(GA)8	[64]
GA9n	CAAGTCCAAATCAGAAGAGC	CTCTATCCACTGCGTTTCTC	HEX	60	112	(GA)7	[64, 85]
LIST7010	CGGTGAATGCCTAAAGAGAGA	AGGAACGCATACTTGGAAGG	6FAM	42	190	(TA)28	[69]
LIST7011	CGGCGACATGCACACATA	CACACACATTGAAGATGGAGGA	HEX	42	186	(TA)15	[69]
LIST7027	CTCTCTCGTCACCACAGCAC	AGGGGACAAGACACAGATGG	6FAM	50	181	(CA)12	[69]
LIST7033	CATTGCTGAGTGCTGCTAGTG	ATGAGCGTACTGGGCACAC	6FAM	44	180	(GT)8	[69]
LIST7039	CTCGCACTCTTTCGCTCTTT	GAGACGAGAGGAACGGAAAA	6FAM	44	205	(CA)16	[69]
LIST7040	GCAGAGCGAGACACACAGAC	GTGCACGTTGATGTGCTTCT	HEX	50	245	(GT)23	[69]
4GTG	CGGTTTGGCGCTGAAAGCGG	CGTGAGGACGCCACCGAGGC	6FAM	58	62	(GTG)5	[64]
27GTGn	GATAGCGTTGGAGGCAAGC	CTATCCGCACCACGATCC	6FAM	60	106	(GTG)5	[64, 85]

5.2.5.2 Microsatellite amplification

The microsatellite containing sequences were amplified by PCR. Strain MHOM/PS/2001/ISL590 was included in each run as size reference. Each reaction was done in a total volume of 25 µl that contained 1x PCR buffer, 200 µM dNTP mix, 5 pmol fluorescence-labelled forward and unlabelled reverse primer, 1 unit Taq polymerase and 20 ng genomic DNA isolated from cultured *Leishmania*

cells. Two µl of DNA were used when the sample was directly isolated from mammalian or sand fly hosts without prior culturing; they included a vast amount of host DNA which did not interfere in the PCR reaction since the primer pairs were specific for *Leishmania* sequences. All primer sequences, annealing temperatures for amplification, and fragment sizes and repeat arrays in the reference strain are listed in Table 4. The thermo-profile of the PCR started with an initial denaturation step for 5 min at 95°C, followed by 35 cycles (30 sec at 95°C, 30 sec at marker-specific annealing temperature (AT, see Table 4), and 1 min at 72°C) of amplification. A final extension was carried out at 72°C for 6 min. The resulting PCR products were separated by electrophoresis on a 2.5% agarose gel to which 2.5 µl of HDGreenPlus was added as DNA stain. 1 kb DNA ladder was used as size standard. The electrophoresis was carried out using 4-5 V/cm for about 1.5-2 h. The fragments in the gel were visualized and their relative concentrations estimated. The PCR products were diluted for the fragment analysis in an ABI sequencer. The labelled forward primers together with GeneScan™ 500 ROX™ Size Standard, including fragments with the lengths of 35, 50, 75, 100, 139, 150, 160, 200, 250, 300, 340, 350, 400, 450, 490, and 500 bases, were used for detection of the fragments and to determine their sizes.

5.2.5.3 Fragment lengths analysis

The fragment detection in an ABI sequencer was carried out by SMB Services in Molecular Biology GmbH. The received output files were analysed with GeneMapper software version 3.7 (Applied Biosystems, Foster City, USA). Each peak was compared to the one of the reference strain and their sizes rounded accordingly. Stutter peaks as well as peaks resulting from A-overhangs, a result of the amplification using Taq polymerase, were not considered for the analysis. Single peaks represented homozygous alleles; double peaks were considered as the result of heterozygous alleles. Strains with missing data for more than two markers were excluded from the analysis.

5.2.5.4 Normalization of newly and previously typed fragments

Fragment lengths had to be normalized for several reasons. Some of the microsatellite profiles of previously typed strains were retrieved from a Beckman Coulter, and not from an ABI sequencer, using different fluorescence dyes. This altered the resulting fragment lengths by up to 2 bp. Furthermore, the primer sequences for marker GA9n were changed, enlarging the resulting fragment. To combine all data from both sequencing systems and those from all used primer pairs, the results had to be normalized. The repeat number of each fragment was calculated in relation to the reference strain MHOM/PS/2001/ISL590 and multiplied by the size of the di- or trinucleotide microsatellite repeat, and finally the size of the flanking region was added.

5.2.5.5 Creation of the input file

All normalized sizes of clear peaks were collected in an excel file. Values from single peaks were used for both alleles for homozygous loci, and both peaks were used for heterozygous loci, one for each allele. In the case of missing data, when amplification failed for single markers in a strain, those cells in the excel sheet were filled with "nd". A template for the input file with all necessary values is shown in Table 5. The complete file was saved as tab-delimited text file and used as input file for the subsequent analyses.

Table 5. Template for excel input file. The table shows an example for five invented strains. The number in the upper left cell specifies the two column format of the table. d, outbred individual; 1, group number. The microsatellite marker GA1 contains dinucleotide repeats (2), and its flanking region is 44 bp, microsatellite marker 27GTGn contains trinucleotide repeats (3) and its flanking region is 91 bp. The fragment resulting from the PCR amplifying locus GA1 in strain 1 has a length of 66 bp (homozygous); the same fragment amplified in strain 2 is heterozygous and is 58 and 66 bp in length. nd, missing data.

2			2	2	3	3	repeat size
			44	44	91	91	length of flanking region
			GA1	GA1	27GTGn	27GTGn	locus name
strain 1	d	1	66	66	112	118	fragment sizes
strain 2	d	1	58	66	112	112	
strain 3	d	1	66	66	112	115	
strain 4	d	1	58	66	nd	nd	
strain 5	d	1	68	68	115	115	

5.2.6 Population genetic analyses

Several programs were used to deduce the variability and population structure of the strains of *L. tropica* and *L. aethiopica*. The approaches were based on various parameters like allele frequencies and genetic distances. A summary of all population genetic software used for this study is shown in Table 6. The applied programs differ in their implemented algorithms. By comparing the results obtained with the various approaches, one could evaluate whether they did or did not support each other.

For all programs particular input file formats were needed. The conversion between the different files was done by MSA 4.05 [86] and CONVERT 1.31 [87] which both provide a variety of data and format outputs.

Table 6. Software used for the population genetic analyses of this study.

software	purpose	required input file	reference
GeneMapper 3.7	determination of the fragment sizes of the amplified microsatellite markers	*.fsa	Applied Biosystem, Foster City, USA
MS Excel 2003	creation of initial input file as text file	fragment sizes, for details see Table 5	Microsoft
MSA 4.05	conversion of input files and calculation of genetic distances (F_{ST} values)	*.txt (tab-delimited)	[86]
	F statistics	*.txt (tab-delimited)	
populations 1.2.34	creation of NJ tree based on chord distances	*.gen (GENEPOP)	http://bioinformatics.org/~tryphon/populations
MEGA 5.1	visualization of NJ tree	*<NO EXTENSION> (populations-output file)	[88]
STRUCTURE 2.3.4	estimation of the membership coefficients for each strain based on Bayesian statistics	*.struct	[89]
MS Excel 2010	estimation of the most probable number of populations	Ln P(D) values from STRUCTURE run	[90]
CONVERT 1.31	conversion of input files	*.txt (CONVERT) / *.gen (GENEPOP)	[87]
SplitsTree 4.12.8	creation of phylogenetic network	*.txt (NEXUS)	[91]
Genetix 4.05	Factorial Correspondence Analysis (FCA)	*<NO EXTENSION> (GENEPOP)	[92]
GDA 1.1	descriptive statistics	*.nex (NEXUS)	[93]

STRUCTURE is based on Bayesian statistics which uses patterns of allele frequencies for the identification of distinct sub-populations and determines membership coefficients for each strain [89]. This software was used for studying the nature and extent of genetic variation within and between populations and providing information about gene flow, population ancestries and mixed genotypes. The length of burn-in period and the number of Markov chain Monte Carlo (MCMC) repetitions were set to 20,000 and 200,000, respectively. The deltaK calculation estimated the most likely number of populations in the sample set for 10 replicate runs for each K [90]. For this calculation, at least 10 iterations for each K were needed. DeltaK was estimated as the mean of the absolute values of $L''(K)$ averaged over 10 runs divided by the standard deviation of $L(K)$, $\Delta K = m(|L''(K)|)/s[L(K)]$, which expands to $\Delta K = m(|L(K+1) - 2L(K) + L(K-1)|)/s[L(K)]$. The resulting graph indicated the most likely number of populations at its maximum peak.

Neighbour Joining (NJ) trees were inferred by POPULATIONS software (http://bioinformatics.org/~tryphon/populations) which uses microsatellite genetic distances (Chord distances were used here) for the calculation. Bootstrapping was done on the basis of 100 iterations resulting in consensus trees with the bootstrap value on the nodes of the branches. The graphical presentation of the distance trees was done using MEGA 5.1 [88].

A neighbour network was created to analyse the genetic interactions between the strains. This algorithm implemented in SplitsTree 4 accounts for reticulation events such as hybridisation, horizontal gene transfer and recombination [91].

Factorial correspondence analysis (FCA) is an approach to identify genetic populations based on allele frequencies [92]. The graphical output presented each strain in a three-dimensional space allowing delimiting different clusters of strains, although the strains needed to be pre-assigned to different populations in the input file. This pre-assignment was done according to the Bayesian statistics.

F statistics as implemented in MSA software, which was already used for the conversion of the input file, is a method to test for genetic isolation of populations. The pairwise comparisons resulted in different levels of genetic differentiation (<0.05, little; 0.05-0.15, moderate; 0.15-0.25, great; >0.25, very great) where F_{ST} = 0 was no genetic differentiation and F_{ST} =1 was complete differentiation. The *p* values indicated the statistical significance of the F_{ST} values (p<0.05, significant; p>0.05, not significant). The original input file was modified by categorizing the strains into the populations deduced by the Bayesian statistics.

The descriptive analyses of the microsatellite loci and the different populations was done using GDA software [93]. With this program the mean number of alleles (A), the observed (H_o) and expected (H_e) heterozygosity and the inbreeding coefficients (F_{IS}, [94]) were calculated.

5.3 Functional Cloning

In this study, the cause for the high genetic variability among strains of *L. tropica* was aimed to be elucidated. The working hypothesis was that an impaired MMR machinery was responsible for an elevated mutation rate in *L. tropica*. Functional Cloning is based on genetic complementation, restoring the donor´s phenotype in cells of the acceptor strain. The method was described by Clos *et al.* as a means to identify genes involved in drug resistance in *Leishmania* [95].

Here, the tolerance towards N-methyl-N'-nitro-N-nitrosoguanidin (MNNG) was used an indicator for the functionality of the MMR system. Besides the nucleotide excision repair and its role in genetic recombination, an additional function of the MMR system is the stimulation of DNA damage induced cell cycle checkpoints and apoptosis of the cell [96]. MNNG is an alkylating agent which incorporates O^6-methyl-guanine into the replicating DNA which can pair with cytosine or thymine. A functioning MMR system recognizes these mismatches but fails to repair them since they are found in the template strand which is not being repaired (Figure 10). As a consequence to the failure of repair, damage signalling pathways are activated, leading to cell cycle arrest and apoptosis [96, 97]. In cells exhibiting an impaired MMR system, the damage signalling pathways are not activated. The

mismatches are tolerated, leaving a mutation after the next round of replication. Thus, an impaired MMR system causes a mutator phenotype in the cell [80, 98].

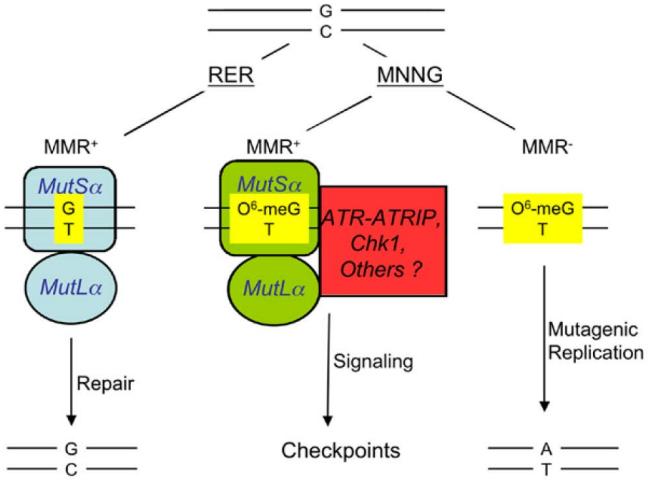

Figure 10. Effect of MNNG on the MMR machinery. The damage caused by MNNG cannot be repaired by the MMR system. This leads to apoptosis in a cell with a functioning MMR system, in a cell exhibiting an impaired MMR system the mismatch is ignored, thus leading to a mutation after the next replication of the DNA. RER, replication error; MMR^+, mismatch repair proficient cell; MMR^-, mismatch repair deficient cell. Source: [97]

In our functional cloning approach, the gDNA of a strain with a functioning MMR system, characterized by low MNNG tolerance was introduced via the shuttle vector pcosTL into a strain of *L. tropica* exhibiting high MNNG tolerance (Figure 11). The inserts of those clones exhibiting low MNNG tolerance were characterized by restriction analysis and alignment to whole genome sequences of different *Leishmania* species.

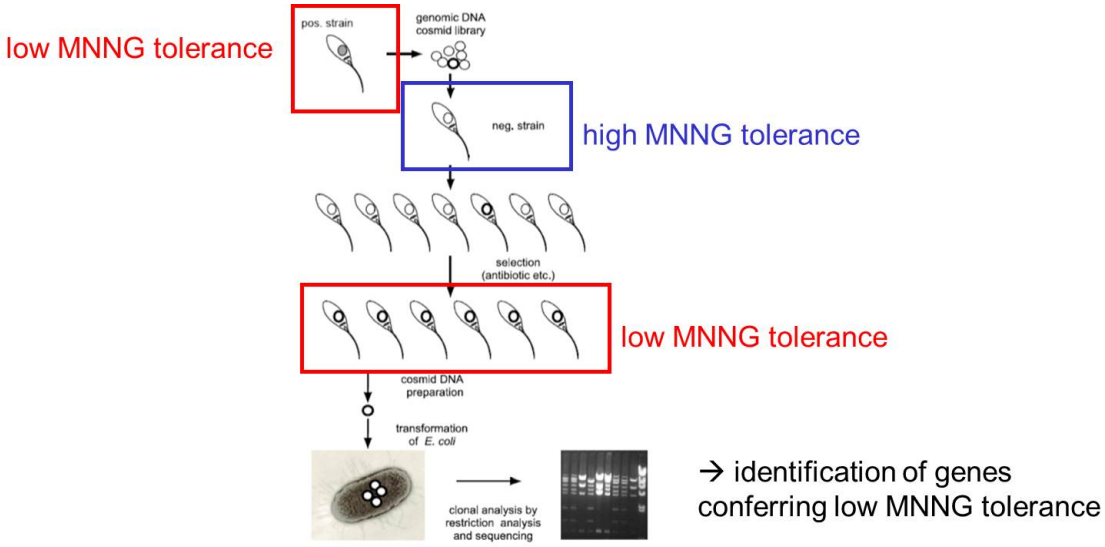

Figure 11. Flow chart of the Functional Cloning approach. A cosmid library is created with the gDNA of the donor strain (here: exhibiting low MNNG tolerance). This library is then used to transfect the acceptor strain (here: exhibiting high MNNG tolerance). The transfectants are screened for the donor phenotype and these cosmids are then analysed for the genes conferring the low MNNG tolerance. Figure modified after [95]

5.3.1 *Leishmania* strains

All twelve strains of *L. tropica*, which were tested for their MNNG susceptibility were isolated between 2001 and 2008 in Israeli and Palestinian foci of CL (Table 3).

Table 7. Strains of *L. tropica* for the selection of donor and acceptor strain for the Functional cloning approach. The strains are called by their lab code throughout the manuscript for convenience reasons.

lab code	WHO code
LRC-L830	MHOM/PS/2001/ISL593
LRC-L881	MHOM/PS/2002/63JnF21
LRC-L885	MHOM/PS/2002/79JnF20
LRC-L886	MHOM/PS/2002/89JnF
LRC-L887	MHOM/PS/2002/87JnM
LRC-L889	MHOM/PS/2002/64JnF4
LRC-L890	MHOM/PS/2002/52JnM18
LRC-L892	MHOM/PS/2002/50Jn20
LRC-L1321	MHOM/IL/2008/LRC-L1321
LRC-L1322	MHOM/PS/2008/335JnM59
LRC-L1323	MHOM/PS/2008/333JnYM27
LRC-L1324	MHOM/PS/2008/336JnM71

5.3.2 Cell culture

5.3.2.1 Determination of cell numbers

The cell numbers in liquid medium were determined using three different methods. For an absolute result of only a few samples the Neubauer counting chamber was used. For the analysis of a greater amount of samples the absolute cell number was determined using a fluorescence-activated cell sorter (FACS). When the comparison between two cultures was addressed, the more time-efficient method of photometric determination of the cell density was employed.

Neubauer counting chamber

The *Leishmania* culture was diluted with fixation solution to an approximate cell density of 10^6/ml. After mixing of the cell solution, it was loaded to the Neubauer counting chamber. Under the microscope the cells of ten of the small squares were counted. The cell number per ml in the original solution was calculated using the formula

$$cells\ per\ ml = \frac{counted\ cells/square}{counted\ surface\ (mm^2) \times\ chamber\ depth\ (mm) \times\ dilution}$$

where the surface of one square was 0.04 mm^2 and the chamber depths was 0.1 mm.

Fluorescence activated cell sorting

The absolute cell number in a culture was assessed by comparing the cell count to the number of Caltag counting beads (Invitrogen) which were of known concentration. Equal amounts of liquid

culture containing the *Leishmania* cells and counting bead solution were mixed and analysed with FACS. Both groups of data points on the dot plot were assessed and the absolute cell number was calculated with the formula

$$cells\ per\ \mu l = \frac{counted\ cells}{counted\ beads} \times beads\ per\ \mu l$$

A minimum of 1000 beads was evaluated for statistical significance.

Photometric determination

If only the relative number of cells was needed, e.g. during the MNNG sensitivity tests, the cell density was evaluated photometrically. The photometer was calibrated against the culture medium at 600 nm and the default values for measuring cell densities were used. The grown cell culture was pipetted to a microcuvette and the optical densities of the cultures grown with and without MNNG determined.

5.3.2.2 Cell culture of promastigote *Leishmania* cells

The *Leishmania* strains were cultured in liquid medium (modified M199) in different volumes. For strain maintenance, culture flasks were used. The initial screening for MNNG sensitive clones was performed in 96well-microtiter plates and positive clones were then re-tested in a 10 ml volume in culture flasks. Independent from the culture volume, all cultures were inoculated with $1x10^6$ cells/ml for strain maintenance, and $5x10^5$ cells/ml for the selection assay, respectively, and incubated at 26 °C. A new flask was inoculated when the cell number reached 2-3 cells/ml to remain a logarithmic growth. For the culturing of the cosmid clones, the medium was completed with 10 µg/ml G418 (geneticindisulfat).

5.3.2.3 Cryopreservation of promastigote *Leishmania* cells

The *Leishmania* culture was grown to a density of $2\text{-}5x10^7$ cells/ml and 1.8 ml of it centrifuged at 4 °C and 300 *g* for 5-10 min. The supernatant was discarded and the cell pellet resuspended in 500 µl CrySFM. This cell solution was transferred to a cryotube and gradually frozen down to -80 °C. After 24 h the cryotube was stored in liquid nitrogen.

For thawing, the tube was transferred from the liquid nitrogen to a warm water bath (30 °C) and incubated until approximately 90% of the cells were thawed. After the rest had thawed at room temperature, the cell suspension was transferred into 10 ml of culture medium and incubated at 26 °C. The parasites were checked microscopically for movement.

5.3.2.4 Single-cell cloning

The transfected *Leishmania* cells, LRC-L1322/6-pcosTL-830 were cloned by limiting dilution [99] to a final concentration of 2.5 cells/ml. Of this diluted sample, 200 µl were filled in each well of the microtiter plate, making it a theoretical 0.5 parasite per well. After incubation at 26 °C about one-third to half of the wells contained grown *Leishmania* cultures. These were then used for the downstream analyses as single clones.

5.3.2.5 MNNG sensitivity tests

The recombinant *Leishmania* clones were tested for their MNNG sensitivity, which, according to the working hypothesis, indicated a functioning or impaired MMR system. In a microtiter plate, one half of the wells were filled with 200 µl of medium containing 2 µM MNNG, the other half with medium lacking this selection drug. Each of these wells was then inoculated with $5x10^5$ cells/ml and grown for 48 h at 26 °C. As positive and negative controls, both the donor LRC-L830 and the acceptor strain LRC-L1322/6 were inoculated along with the other strains in each plate. Positive clones, which were identified in this way, were re-tested in larger culture volumes to minimize pipetting inaccuracies. For this, 10 ml of medium were inoculated as described above in culture flasks. The *p* value for each clone was calculated using the tool at http://graphpad.com/quickcalcs/ttest1/.

5.3.3 Cosmid characterization

5.3.3.1 Cosmid preparation from *Leishmania* culture

The cosmid DNA was isolated from the leishmanial cells by alkaline lysis and subsequent phenol-chloroform-extraction. Fifty ml of *Leishmania* culture were grown to high density (~$2x10^7$ cells/ml) and cells harvested by centrifugation at 4 °C and 1260 *g* for 10 min. The cells were washed twice with PBS (pH 7.0) and were lysed with 1 ml of solution 1 and 2 ml of solution 2. After gentle shaking for 2-5 min 1.5 ml of solution 3 were added and incubated on ice for 5-10 min. After centrifugation of this suspension at 4 °C and 3220 *g* for 30 min, the supernatant was filtered to exclude all cell particles, followed by an RNase digestion (40 µg/ml) at 37 °C for at least 30 min. For the extraction of the DNA, 1 volume of phenol: chloroform: isoamyl alcohol (25:24:1, v/v) was added, mixed and centrifuged at 20 °C and 3220 *g* for 4-6 min. The lower, organic phase was discarded. These steps were repeated twice, with the last repetition being modified by adding an equal volume of chloroform: isoamyl alcohol (24:1, v/v) and transferring the upper, aqueous phase into a fresh tube. The DNA was precipitated by the addition of 0.7 volumes of isopropanol and incubation at room temperature for 10 min. After the DNA was centrifuged to the bottom of the tube at the same conditions as before for 30 min, the supernatant was carefully discarded and the DNA pellet washed with ice-cold 70%

ethanol. The supernatant was removed thoroughly and the pellet dissolved in 20-40 µl of TE buffer (pH 8.0). The cosmid DNA was stored at 4 °C.

5.3.3.2 Transformation of *E. coli* with cosmid pcosTL

The cosmids, probably bearing the inserts with the genes responsible for the low MNNG tolerance, had to be amplified in *E. coli*.

Cells of the *E. coli* strain XL1-Blue were brought into a transformation competent state by serial washing steps of the cells with ice-cold 0.1 M $CaCl_2$. Frozen cells were colonized on LB agar plates and incubated overnight at 37 °C. The next day a pre-culture of 10 ml was inoculated with one colony and again incubated overnight at 130 rpm and 37 °C. A volume of 100 ml LB medium was inoculated with this overnight culture to an optical density of OD_{600}=0.1 in a 500 ml flask. The cell density was monitored photometrically and the growth was stopped at an OD_{600} of 0.6-0.7 by chilling the cells on ice for 20 minutes. After a centrifugation step for 10 min at 4° C and 5000 *g*, the supernatant was discarded and the cell pellet resuspended in 10 ml ice-cold sterile 0.1 M $CaCl_2$. These washing steps were repeated with 4 ml ice-cold sterile 0.1 M $CaCl_2$.The volume of the competent cells was estimated and sterile glycerol was added to a final concentration of 10%. The cells were then aliquoted in 200 µl and either stored at 4 °C for use within the next 48 h or frozen in liquid nitrogen and stored at -80 °C until use. The transformation efficiency was tested for each lot using the plasmid pUC18.

For the transformation, the competent cells were thawed on ice and mixed with 1-5 ng of the cosmid preparation. This mix was placed on ice for 30 min and then incubated in a water bath of 42 °C for exactly 30 sec. After the heat shock, the cells were cooled on ice for 2 min, then suspended in 1 ml warm SOC medium without antibiotics and incubated at 37 °C and 250 rpm for 1 h. For the selection of successfully transformed cells, 200 µl of this culture were plated on an LB agar plate which contained 75 µg/ml ampicillin. The grown colonies were used for the downstream analyses.

For long-time storage, 300 µl of overnight culture was mixed with 300 µl of 10% glycerol. This mix was frozen and kept in a cryotube at -80 °C.

5.3.3.3 Isolation of cosmid DNA from *E. coli*

One colony was used to inoculate an LB agar plate (+ampicillin) and 4 ml of liquid LB medium. Both were incubated overnight at 37 °C. The agar plate was stored at 4 °C as back-up, the cells grown in liquid medium were harvested by centrifugation at 4 °C and 1400 *g* for 15-30 min. They were resuspended in 200 µl of solution 1 (for recipes see chapter 5.1.3) transferred to an Eppendorf tube and mixed with 400 µl of solution 2 by inverting the tube several times and incubating at room

temperature for 3 min. Upon addition of 300 µl of solution 3, the tubes were incubated for 5 min on ice, and then centrifuged at 4 °C and 13000 *g* for 30-60 min. The supernatant was mixed thoroughly with 1 ml of isopropanol by inverting the tube several times to precipitate the DNA. After an additional centrifugation step under the same conditions, the DNA pellet was washed with 700 µl 70% ethanol. The dried DNA pellet was suspended in 20 µl TE/RNase (20 µg/ml) and incubated at 37 °C for up to 120 min. The DNA was stored at 4 °C.

5.3.3.4 Restriction analysis of the recombinant cosmids

The cosmids were subjected to restriction analyses in order to identify differences in the fragment patterns due to the presence of different inserts in the cosmid clones. One reaction consisted of 3-4 µl DNA, 20 U EcoRV and XbaI in a total volume of 15 µl. The buffer and BSA were used according to the manufacturer´s instructions. The reaction was incubated for 15 minutes at 37 °C.

The cosmid fragments were analysed by electrophoresis. Each resulting fragment pattern contained two fragments originating from the cosmid backbone at 7422 and 815 bp. The electrophoresis was carried out at ~100 V for 1.5-2 h in a 0.7% agarose gel. A special agarose (SeaKem LE agarose) was used for the separation of large fragments. The DNA was visualized after soaking the gel into a staining bath containing 100 µl ethidium bromide (10 mg/ml) per 2 l aqua dest. for 20 min. For estimation of the fragments´ sizes, the 1kb DNA Extension ladder was used.

5.3.3.5 Identification of the cosmid inserts in positive clones

The cosmid inserts were partially sequenced to identify their sequences for those clones that exhibited low mutation tolerance in the presence of MNNG. The primers to initiate the sequencing reaction from both ends of the insert were complementary to the cosmid sequence near the ligation sites. Their sequences were 5´-CCGTGTGATATCAGATGCCC-3´ (CH2) and 5´-GGAAACAGCTATGACCATG-3´ (M13rev). The resulting DNA sequences each contained a part of the insert. They were then used for a BLAST search using the NCBI database to determine the start and end points of the insert. All available sequences of different *Leishmania* species were used for the alignments. VectorNTI Advance 11 and BioEdit 7.0.9.1 [100] were used for the handling of the sequencing data.

5.4 Multiplex Ligation-dependent Probe Amplification (MLPA)

5.4.1 Sand fly strains

Forty-five single sand flies and 20 sand fly DNA samples were provided by the entomology unit of the Ministry of Health in Jerusalem, Israel. All sand flies were male and collected at different sites in

Israel and the West Bank (Table 8). The sample set comprised specimens of the species *P. papatasi*, *P. syriacus*, *P. sergenti*, *P. tobbi*, *P. arabicus*, *P. perfiliewi*.

Table 8. Dates and locations of collection of the sand fly specimens used in this study. Each sample was given a serial number.

ID	species	Date of collection	Location	District
***pap*01**	*P. papatasi*	25/10/2011	Karkom	northern hills (Sea of Galilee)
***pap*02**	*P. papatasi*	25/10/2011	Karkom	northern hills (Sea of Galilee)
***pap*03**	*P. papatasi*	13/09/2011	Alrashayda	Judean desert
***pap*04**	*P. papatasi*	13/09/2011	Alrashayda	Judean desert
***pap*05**	*P. papatasi*	06/06/2012	Kfar Adumim	Jerusalem east (hills)
***pap*06**	*P. papatasi*	06/06/2012	Kfar Adumim	Jerusalem east (hills)
***pap*07**	*P. papatasi*	06/06/2012	Kfar Adumim	Jerusalem east (hills)
***pap*08**	*P. papatasi*	06/06/2012	Kfar Adumim	Jerusalem east (hills)
***pap*09**	*P. papatasi*	06/06/2012	Kfar Adumim	Jerusalem east (hills)
***pap*10**	*P. papatasi*	06/06/2012	Kfar Adumim	Jerusalem east (hills)
***pap*11**	*P. papatasi*	06/06/2012	Kfar Adumim	Jerusalem east (hills)
***pap*12**	*P. papatasi*	12/06/2012	Kinneret lake, Susita beach	northern (Sea of Galilee)
***pap*13**	*P. papatasi*	20/06/2012	Ma'ale Adumim	Jerusalem east (hills)
***pap*14**	*P. papatasi*	20/06/2012	Ma'ale Adumim	Jerusalem east (hills)
***ser*01**	*P. sergenti*	31/10/2011	Pedu'el	central (hills)
***ser*02**	*P. sergenti*	20/07/2011	Alrashayda	Judean desert
***ser*03**	*P. sergenti*	04/05/2011	Ma'ale Adumim	Jerusalem east (hills)
***ser*04**	*P. sergenti*	06/06/2012	Ma'ale Adumim	Jerusalem east (hills)
***ser*05**	*P. sergenti*	06/06/2012	Ma'ale Adumim	Jerusalem east (hills)
***ser*06**	*P. sergenti*	06/06/2012	Ma'ale Adumim	Jerusalem east (hills)
***ser*07**	*P. sergenti*	06/06/2012	Ma'ale Adumim	Jerusalem east (hills)
***ser*08**	*P. sergenti*	06/06/2012	Kfar Adumim	Jerusalem east (hills)
***ser*09**	*P. sergenti*	06/06/2012	Kfar Adumim	Jerusalem east (hills)
***ser*10**	*P. sergenti*	26/07/2012	Etri Ruins	Judean foothills
***ser*11**	*P. sergenti*	26/07/2012	Etri Ruins	Judean foothills
***ser*12**	*P. sergenti*	26/07/2012	Etri Ruins	Judean foothills
***ser*13**	*P. sergenti*	26/07/2012	Etri Ruins	Judean foothills
***syr*01**	*P. syriacus*	25/10/2011	Karkom	northern hills (Sea of Galilee)
***syr*02**	*P. syriacus*	25/10/2011	Karkom	northern hills (Sea of Galilee)
***syr*03**	*P. syriacus*	25/10/2011	Karkom	northern hills (Sea of Galilee)
***syr*04**	*P. syriacus*	06/06/2012	Ma'ale Adumim	Jerusalem east (hills)
***syr*05**	*P. syriacus*	06/06/2012	Ma'ale Adumim	Jerusalem east (hills)
***syr*06**	*P. syriacus*	02/07/2012	Kfar Adumim	Jerusalem east (hills)
***syr*07**	*P. syriacus*	02/07/2012	Kfar Adumim	Jerusalem east (hills)
***syr*08**	*P. syriacus*	05/07/2012	Nataf	Jerusalem west
***syr*09**	*P. syriacus*	05/07/2012	Nataf	Jerusalem west
***syr*10**	*P. syriacus*	05/07/2012	Nataf	Jerusalem west
***syr*11**	*P. syriacus*	05/07/2012	Nataf	Jerusalem west
***syr*12**	*P. syriacus*	11/07/2012	Pisgat Zeev	Jerusalem north
***syr*13**	*P. syriacus*	11/07/2012	Pisgat Zeev	Jerusalem north

ID	species	Date of collection	Location	District
***per*01**	*P. perfiliewi*	25/10/2011	Karkom	northern hills (Sea of Galilee)
***per*02**	*P. perfiliewi*	25/10/2011	Karkom	northern hills (Sea of Galilee)
***per*03**	*P. perfiliewi*	25/10/2011	Karkom	northern hills (Sea of Galilee)
***per*04**	*P. perfiliewi*	12/06/2012	Arik bridge pond	northern (Sea of Galilee)
***per*05**	*P. perfiliewi*	12/06/2012	Arik bridge pond	northern (Sea of Galilee)
***per*06**	*P. perfiliewi*	12/06/2012	Arik bridge pond	northern (Sea of Galilee)
***per*07**	*P. perfiliewi*	07/08/2013	Sde Eliyahu	Northern (Jordan valley)
***per*08**	*P. perfiliewi*	17/09/2013	Amnun	northern hills (Sea of Galilee)
***ara*01**	*P. arabicus*	25/10/2011	Karkom	northern hills (Sea of Galilee)
***ara*02**	*P. arabicus*	25/10/2011	Karkom	northern hills (Sea of Galilee)
***ara*03**	*P. arabicus*	24/10/2013	Kahal	northern hills (Sea of Galilee)
***ara*04**	*P. arabicus*	08/10/2013	Kadarim	northern hills (Sea of Galilee)
***ara*05**	*P. arabicus*	17/09/2013	Amnun	northern hills (Sea of Galilee)
***ara*06**	*P. arabicus*	14/10/2013	Karkom	northern hills (Sea of Galilee)
***ara*07**	*P. arabicus*	25/10/2011	Karkom	northern hills (Sea of Galilee)
***ara*08**	*P. arabicus*	08/09/2012	Karkom	northern hills (Sea of Galilee)
***tob*01**	*P. tobbi*	07/09/2011	Pedu'el	central (hills)
***tob*02**	*P. tobbi*	25/10/2011	Karkom	northern hills (Sea of Galilee)
***tob*03**	*P. tobbi*	06/06/2012	Ma'ale Adumim	Jerusalem east (hills)
***tob*04**	*P. tobbi*	06/06/2012	Ma'ale Adumim	Jerusalem east (hills)
***tob*05**	*P. tobbi*	06/06/2012	Ma'ale Adumim	Jerusalem east (hills)
***tob*06**	*P. tobbi*	23/10/2013	Pedu'el	central (hills)
***tob*07**	*P. tobbi*	23/10/2013	Pedu'el	central (hills)
***tob*08**	*P. tobbi*	21/10/2013	Eliav settlement	Judean foothills
***tob*09**	*P. tobbi*	25/10/2011	Karkom	northern hills (Sea of Galilee)

5.4.2 Species identification

All sand flies were examined individually by microscopy and the species were identified morphologically by the staff of the entomology unit of the Ministry of Health in Jerusalem.

5.4.3 DNA isolation

The DNA was isolated from single sand flies. In our laboratory, the flies were homogenized in 250 µl lysis buffer using a sterile glass rod. Triton X-100 and Proteinase K were added to final concentrations of 1 % and 400 µg/ml, respectively, and then incubated overnight at 60°C. The phenol-chloroform-extraction of the DNA was carried out as described for the extraction of *Leishmania* DNA (chapter 5.2.3.) The DNA pellet was dissolved in 50 µl aqua dest. and stored at 4°C.

In the entomology unit of the Ministry of Health in Jerusalem the DNA was extracted using the "DNeasy 96 Blood & Tissue" extraction kit from Qiagen.

5.4.4 The MLPA technique

Multiplex ligation-dependent probe amplification (MLPA) is based on the hybridisation of one specific probe containing a SNP and one unspecific probe complementary to the sequence next to the SNP, followed by the ligation of those probes (Figure 12). The ligation is only successful if the specific probe anneals to the template DNA in full length, thus bringing the two ligation ends next to each other. The probes have to be designed in a way that the targeted SNP is located at the 3'end of the left probe oligo (LPO). The right probe oligo (RPO) contains a stuffer sequence which is variable in length making multiplexing of the reaction possible by including several probes in one tube. The probes are amplified using a universal PCR primer pair attached to the ends of each probe. The resulting fragments can then be detected by electrophoretic separation and identified by their lengths [101, 102].

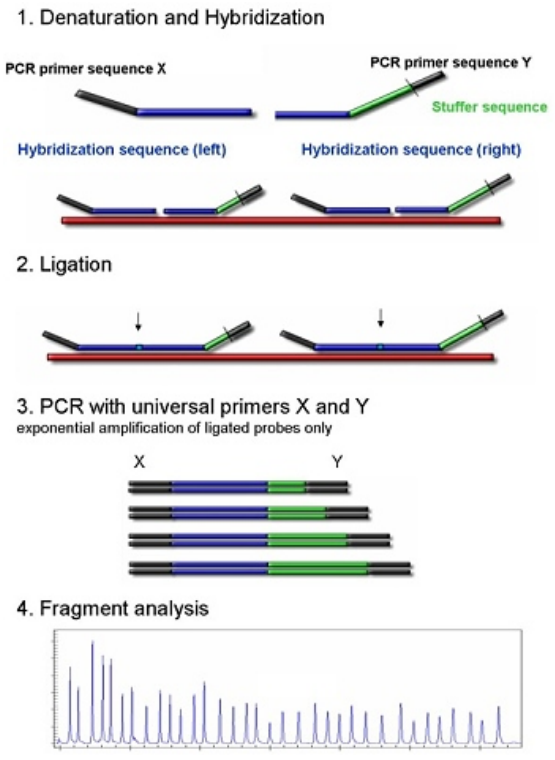

Figure 12. Multiplex ligation-dependent probe amplification. Both probes contain primer and hybridisation sequences; additionally, the LPO contains the species-specific SNP and the RPO contains a stuffer sequence. After the hybridisation of both probes to their complementary strand of the template DNA, the two probes ligate only in case the SNP is present at the ligation site. Thus, only the ligated probes are amplified in a PCR step using universal primers for all different probes included in the reaction tube. Fragment analysis reveals which SNPs were detected by the probes according to their lengths. Source: https://mlpa.com/WebForms/WebFormMain.aspx?Tag=zjCZBtdOUyAt3KF3EwRZhNWLtcfv9pVl/tHJIM%5Cfa9FWO8KMqctOGIoqYwxaGF9Y

5.4.5 Alignment

The 18S rRNA gene sequences of 11 sand fly species available at the NCBI (Table 9) and known to be present in Israel and Palestine were aligned by using MEGA 5.1 software and SNPs differentiating those sand fly species identified.

Table 9. Accession numbers of the 18S rRNA sequences used for the probe design.

sand fly species	accession numbers
P. fallax	AJ244426.1, AJ244427.1
P. simici	AJ244361.1, AJ244363.1, AJ244364.1, AJ244365.1, AJ244366.1
P. perniciosus	AJ244393.1, AJ244394.1, AJ391728.1
P. sergenti	AJ244402.1, AJ244403.1, AJ244404.1, AJ244405.1, AJ391727.1
P. alexandri	AJ244395.1, AJ244396.1, AJ244397.1, AJ244398.1, AJ244399.1, AJ244400.1, AJ244401.1
P. papatasi	AJ244406.1, AJ244407.1, AJ244408.1, AJ244409.1, AJ244410.1, AJ244411.1, AJ244412.1, AJ244413.1, AJ244414.1, AJ391726.1
P. syriacus	AJ244375.1, AJ244376.1
P. neglectus	AJ244367.1, AJ244368.1, AJ244369.1, AJ244370.1, AJ244371.1, AJ244372.1, AJ244373.1, AJ244374.1
P. tobbi	AJ244377.1, AJ244378.1, AJ244379.1, AJ244380.1, AJ244381.1, AJ244382.1, AJ244383.1, AJ244384.1, AJ244385.1
P. perfiliewi	AJ244386.1, AJ244387.1, AJ244388.1, AJ244389.1, AJ244390.1, AJ244391.1, AJ244392.1
P. arabicus	FJ560382.1

5.4.6 Probe design

Based on the alignment of the 18S rRNA sequences, 12 probes were designed by Kiki Tuin at MRC-Holland in Amsterdam, Netherlands. The design aimed to create probes which would be able to distinguish the 11 *Phlebotomus* species investigated in this study. The probes containing SNPs, that were specific to one or few species, were located next to the ligation sites. All probes were conjugated to stuffer sequences of variable lengths (Figure 12). Finally, all probes contained sequences for the annealing of the universal PCR primer pairs. The amplification of all probes was tested on plasmids that contained nucleotide sequences with those SNPs found in the alignments. The sequences of the probes and lengths of the resulting PCR fragments are shown in Table 10. Since species-specific probes could not be identified for each of the species included in the study, a determination table is required for the identification of the sand fly species (Table 23 in the Results section).

Table 10. The probes for species identification including their lengths, specificities and sequences.
The specific SNP is highlighted in yellow, the LPO (left probe oligo) in small letters, the RPO (right probe oligo) in capital letters. The probe SERG consists of three parts and has two ligation sites.

probe	length	specificity	probe sequence
FAL	144	*S. fallax* specific	ttgacggaagggcaccaccaggagtggg**a**- CTGCGGCTTAATTTGACTCAACACGGGAAAACTTACCAGGTCCGAAC
SIM	169	*P. simici* specific	cctgcttgcaggactatggtgctgaaaccatgt**a**- TATTCCTCTGATTGGATGGGCCTCGTGTCTGTTCGATTTGT
PERN1	193	used for *P. perniciosus* determination	ccgacttttacttcctctaaataatcaagttcggtcaacttc**a**- GTGTAACATACCAAATTGCCCAATAAAGGACAATTAGTAGTACACTTCCAGAGACCTCACTAAACAATTCATCAGTAGTAGC
SERG	225	*P. sergenti*, *P. alexandri*	ccgatgaactctggataattatggctgatcgtatggtctt**a**- GCACCGACGATAGATCATTCAAATGTCTGCCCTATCAACTATTGATGGTAGTATAGA-GGACTACCATGGTTGCAACGGGTAACGGG
ALEX	250	*P. alexandri* specific	ccgctgctctgctttgctttcgggttaaagt**a**- GAGTGGTTTGATTTATGTATGGTGGAGTCATACCTGTTTGGTTTGTC
PAP	281	*P. papatasi* specific	ccgcccctatattgactttcgagtttgatataggt**a**- AAGCATATTAATACGCCCCTTATGTTAAAAGACCCATTTTAGTGCTCTTTACCGA
SYR	310	*P. syriacus* specific but low signal also on *P. neglectus* (C>T)	ccataattatccagattcatctaaaacaaacctacaccgt**c**- AAGTGCAGATATTGGTTTTAAAGTCTAATAAAAGCACATGTCTCAATCC
NEG	340	*P. neglectus* specific but low signal also on *P. syriacus* (T>C)	ccatgatactgacatcgacgtaactgtagcataattatccagattcatctaaaacaaacctacaccgt**t**- AAGTGCAGATATTGGTTTTAAAGTCTAATAAAAGCACATGTCTCAATCC
TOB	379	non-specific: *tobbi*, *perfiliewi*, *neglectus*, *syriacus*, *simici*, *perniciosus*, 1x *papatasi*	caaaaggctcagtacaacagctatcatttatttgatcaat**t**- AAAACAGTTACTTGGATAACTGTGGTAATTCCAGAGCTAATACATGCAAACAA
PERF	422	*P. perfiliewi* specific	cctttaagtgtgacggatgttttaccggtgagggaaacgattat**a**- CATATAGTGTTTCATTTTCATAATGGACTCTATTGTTAATACGCCC
ARA	460	*P. arabicus*, *P. simici*	ccacacgggtaatactcccaagtgcgggtt**a**- GTATTACTTGTTGCATGTGTATGTTTACTGTAAAAGGTGGC
PERN2	507	used for *P. perniciosus* determination	cctaacatctgttggtttggtatgccacac**a**- GAAGTTGACCGAACTTGATTATTTAGAGGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGC

Several control fragments (Q) were included in the probe mix. They were 64, 70, 76 and 82 bp in length. These Q fragments contained sequences complementary to the primers that were used for the amplification of the probes. In contrast to the probes, both of the primer annealing sequences were present on the same molecule, thus ligation was not required prior to PCR amplification. In the presence of sufficient amounts of sand fly DNA, amplification of the Q fragments will be outcompeted by the sand fly PCR products. Otherwise, if the amount of template DNA would be too low or the ligation reaction not successful, the Q fragments will be visible in the product mix.

5.4.7 MLPA reaction

The MLPA reaction includes steps of denaturation, hybridisation, ligation and amplification. All necessary components for the MLPA reaction, such as the X059-A1-Sandfly probe mix (containing the 12 sand fly probes and the four Q fragments), ligase-65, PCR ingredients and all buffers were supplied as ready-to-use-solutions by MRC-Holland. Fifteen ng/µl of sand fly DNA were denatured for 5 minutes at 98 °C. As negative control the DNA was replaced by aqua dest. in one reaction tube. For the hybridisation of the probes, 1.5 µl MLPA buffer and 1.5 µl sand fly probe mix were added to each reaction and incubated for 1 h at 95 °C and 16-20 h at 60 °C. The ligation of the probes was done by adding 3 µl each of the ligase buffers A and B, 1 µl ligase-65 and 25 µl aqua dest. and subsequent incubation for 15 minutes at 54 °C. After heat inactivation of the ligase for 5 minutes at 98 °C, the probes were amplified using the supplied SALSA PCR primer mix and SALSA polymerase. The thermo-profile of the PCR consisted of 35 cycles of amplification (30 sec at 95 °C, 30 sec at 60 °C, 60 sec at 72 °C) followed by a final extension for 20 minutes at 72 °C. Ten µl of the PCR products were then separated in a 2.5% agarose gel. The gels were stained using 2.5 µl HDGreenPlus per 100 ml agarose. As size standards, 16 µl of the 1kb DNA ladder and 5 µl of the HyperladderV were used. The fragments were separated at 4-5 V/cm for 2.5-3 h. Depending on their intensities the fragments were diluted and mixed with GeneScan™ 500 ROX™ Size Standard. The detection of the fragments was done in an ABI sequencer by SMB Services in Molecular Biology GmbH, Berlin.

The files obtained during the fragment analysis were evaluated using GeneMapper software version 3.7 (Applied Biosystems, Foster City, USA) using the default settings except for adjusting the size standard to “ROX50-400”

6 Results

6.1 Multilocus Microsatellite Typing (MLMT)

6.1.1 The strains of *L. tropica* and *L. aethiopica* analysed in this study

The geographical origins of the strains of *L. tropica* covered the species´ broad range of distribution throughout the Old world (Table 11) from as far west as Morocco to India in the east (Figure 13 and Figure 14). In total, 203 strains were collected in 18 countries. The strain set included strains of African and Asian origins, with a majority of them originating in the Middle East region. All eight samples of *L. aethiopica* were collected in Ethiopia.

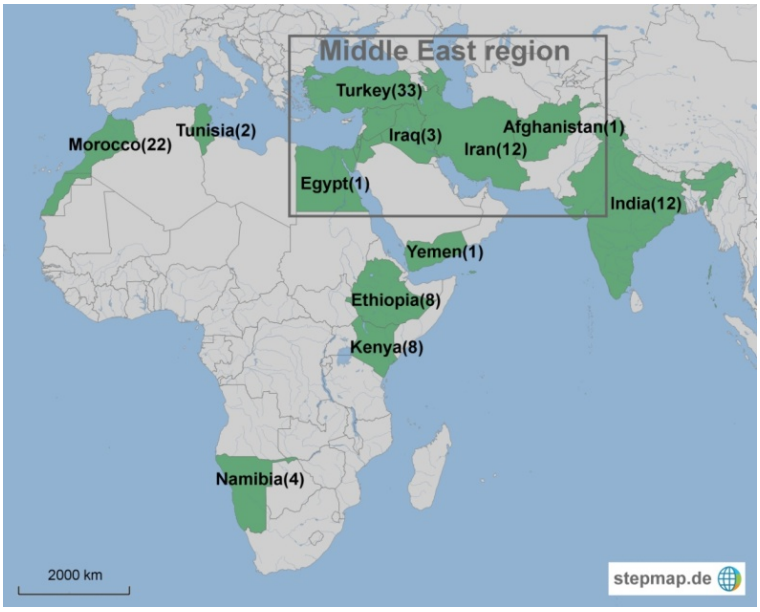

Figure 13. The world-wide distribution of the strains of *L. tropica* and *L. aethiopica* analysed in this study. The number of samples collected in each country is given in parentheses. One strain from the former Soviet Union is not shown in this map since the exact site of its collection is not known. The Middle East region is shown in more detail in Figure 14.

Figure 14. Detailed map of the collection sites in the Middle East region. The numbers of strains of *L. tropica* isolated in the respective countries are specified in parentheses.

Most of the strains (n=177) were isolated from human cases of CL and only eight from human cases of VL. Strains from other mammalian hosts, two from the rock hyrax *P. capensis* and one from a rat (*Rattus sp.*), as well as 15 samples from leishmanial sand fly vectors (eight from *P. sergenti*, five from *P. arabicus* and two from *P. rossi*) were included in the sample set.

6.1.2 Optimization of the microsatellite marker set

The marker set for typing *L. tropica* applied in previous studies comprised 21 markers as published by Schwenkenbecher *et al*. [30]. For this study, the number of markers was decreased to 12 (Table 4). This led to a sufficient amount of data points [60] to deduce the population genetic structure at an intra-species level. Additionally, this reduction minimized the costs for typing tremendously and made it more time-efficient. All markers were informative for the strains of *L. tropica* and *L. aethiopica* studied herein. Only eight markers were amplified successfully in MHOM/IL/1980/Friedlin, a strain of *L. major*.

6.1.3 Normalization of the microsatellite data

The study included data of previously typed strains of *L. tropica* (n=117)[30], for which some of the markers needed to be re-analysed, and newly collected strains (n=86), which were typed with the new marker set. Since some of these data were retrieved from another sequencing system and the primer pair for one marker needed to be re-designed, all typing results had to be normalized. This was done by calculating the repeat number of the microsatellite with regard to the reference strain MHOM/PS/2001/ISL590 and subsequently multiplying this number by the size of the microsatellite (di- or trinucleotide) and adding the lengths of the flanking regions. Normalized fragment lengths for all markers and all strains were used in all subsequent studies and are summarized in Table 11.

Table 11. Microsatellite profiles of all strains of *L. tropica* and *L. aethiopica* analysed in this study. Those strains newly typed for this study are given in bold, all other strains have been typed previously [30] and were completed during this study for marker 27GTGn and missing data. The WHO codes contain information about the host, the country of origin and the year of isolation. The column "geographic origin" gives detailed information about the strains´ origins where available. The column "source" specifies the host and, if human, the manifestation of the disease, CL = cutaneous leishmaniasis, VL = visceral leishmaniasis, DCL = diffuse cutaneous leishmaniasis, LCL = local cutaneous leishmaniasis. The microsatellite profiles are designated as proposed in [85]. The column "population" gives the results of the Bayesian statistics; parentheses indicate discrepancies between the Bayesian and other approaches, which either presented the respective strain as outlier or assigned it to a different cluster or group. The colours of the strains in this column correspond to the colours in Figure 16 to 20. Doc = documentation: *, first published in [30]; **, first published in [85]; *, first published in [103]; ****, first published in [104]; *****, first published in [105]. The fragment lengths are given in bp for all strains and each marker and have been normalized to the reference strain of *L. tropica* MHOM/PS/2001/ISL590.**

						fragment lengths of the microsatellite markers (bp)																							
WHO code	Geographic origin	Source	microsatellite profile	population	Doc	GA1		GA2		GA6		GA9n		LIST7010		LIST7011		LIST7027		LIST7033		LIST7039		LIST7040		4GTG		27GTGn	
MHOM/IL/1996/P837	Anatot, Judean Desert, Israel	human CL	*Ltro* MS 034	Israel/Palestine	*	66	66	62	62	61	61	112	112	190	190	186	186	181	181	180	180	207	207	245	245	62	62	106	106
ISER/IL/1998/LRC-L757	Kfar Adumim, Judean Desert, Israel	*P. sergenti*	*Ltro* MS 034	Israel/Palestine	*	66	66	62	62	61	61	112	112	190	190	186	186	181	181	180	180	207	207	245	245	62	62	106	106
MHOM/IL/1997/P963	Tiberias, western side of Sea of Galilee, Israel	human CL	*Ltro* MS 035	Israel/Palestine	*	66	66	62	62	61	61	112	112	180	180	186	186	181	181	180	180	207	207	245	245	62	62	106	106
ISER/IL/1998/LRC-L758	Kfar Adumim, Judean Desert, Israel	*P. sergenti*	*Ltro* MS 036	Israel/Palestine	*	66	66	62	62	61	61	112	112	190	190	186	186	181	181	180	180	197	197	245	245	62	62	106	106
MHOM/EG/1990/LPN65	Mt. Sinai, Sinai Desert, Egypt	human CL	*Ltro* MS 037	Israel/Palestine	*	66	66	62	62	61	61	112	112	194	194	186	186	183	183	180	180	207	207	245	245	62	62	nd	nd
ISER/IL/1998/LRC-L747	Kfar Adumim, Judean Desert, Israel	*P. sergenti*	*Ltro* MS 038	Israel/Palestine	*	66	66	62	62	61	61	112	112	190	192	186	186	181	181	180	180	205	205	245	245	62	62	106	106
MHOM/PS/2000/GOKS17	Samaria, Palestine	human CL	*Ltro* MS 039	Israel/Palestine	*	66	66	62	62	61	61	112	112	190	192	186	186	181	181	180	180	205	205	245	245	62	62	nd	nd
MHOM/PS/2000/GOKS23	Samaria, Palestine	human CL	*Ltro* MS 039	Israel/Palestine	*	66	66	62	62	61	61	112	112	190	192	186	186	181	181	180	180	205	205	245	245	62	62	nd	nd
MHOM/PS/2001/ISL589	Jericho, Palestine	human CL	*Ltro* MS 040	Israel/Palestine	*	66	66	62	62	61	61	112	112	188	188	186	186	181	181	180	180	205	205	245	245	65	65	106	106
MHOM/PS/2001/ISL588	Jericho, Palestine	human CL	*Ltro* MS 041	Israel/Palestine	*	66	66	62	62	61	61	112	112	190	190	186	186	181	181	180	180	205	205	245	245	62	62	106	106
MHOM/PS/2001/ISL590	Jericho, Palestine	human CL	*Ltro* MS 041	Israel/Palestine	*	66	66	62	62	61	61	112	112	190	190	186	186	181	181	180	180	205	205	245	245	62	62	106	106
MHOM/PS/2001/ISL592	Jericho, Palestine	human CL	*Ltro* MS 041	Israel/Palestine	*	66	66	62	62	61	61	112	112	190	190	186	186	181	181	180	180	205	205	245	245	62	62	106	106
MHOM/PS/2002/31JnM17	**Tubas, Jenin District, Palestine**	**human CL**	***Ltro* MS 041**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2002/35JnF45	**El-Yamoon, Jenin District, Palestine**	**human CL**	***Ltro* MS 041**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/IL/2003/LRC-L999	**Jordan Valley, Israel**	**human CL**	***Ltro* MS 041**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2003/ISLAH721	**Jordan Valley, Palestine**	**human CL**	***Ltro* MS 041**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2003/151JnF32	**Tayaseer, Jenin District, Palestine**	**human CL**	***Ltro* MS 041**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2003/185JnM27	**Serees, Jenin District, Palestine**	**human CL**	***Ltro* MS 041**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2003/152JnF32	**Tayaseer, Jenin District, Palestine**	**human CL**	***Ltro* MS 041**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2003/184Jn01	**El-Yamoon, Jenin District, Palestine**	**human CL**	***Ltro* MS 041**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2003/186JnM12	**Beer Al Basha, Jenin District**	**human CL**	***Ltro* MS 041**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2003/161JnF80	**Al-Shuhada, Jenin District, Palestine**	**human CL**	***Ltro* MS 041**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
ISER/IL/2004/LRC-L1167	**Tiberias, Israel**	***P. sergenti***	***Ltro* MS 041**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/IL/2009/LRC-L1374	**Anatot, Israel**	**human CL**	***Ltro* MS 041**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2002/ISL698	**Jericho, Palestine**	**human CL**	***Ltro* MS 041**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2001/ISL572	**Jericho, Palestine**	**human CL**	***Ltro* MS 041**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2003/149JnM9	**El-Yamoon, Jenin District, Palestine**	**human CL**	***Ltro* MS 041**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2003/163JnM30	**El-Yamoon, Jenin District, Palestine**	**human CL**	***Ltro* MS 041**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/IL/2003/LRC-L1024	**Tiberias, Israel**	**human CL**	***Ltro* MS 041**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2003/178JnM75	**El-Yamoon, Jenin District, Palestine**	**human CL**	***Ltro* MS 041**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/IL/2003/LRC-L1021	**Tiberias, Israel**	**human CL**	***Ltro* MS 041**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2002/41JnF12	**Silat El-Hartheyia, Jenin District, Palestine**	**human CL**	***Ltro* MS 041**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/IL/2003/LRC-L1006	**Tzarka Amir, Israel**	**human CL**	***Ltro* MS 041**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/IL/2009/LRC-L1366	**Maale Adumim, Israel**	**human CL**	***Ltro* MS 041**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2008/335JnM59	**Salit El-Hartheyia, Jenin District, Palestine**	**human CL**	***Ltro* MS 041**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2001/ISL593	Jericho, Palestine	human CL	*Ltro* MS 042	Israel/Palestine	*	66	66	62	62	61	61	112	112	188	190	186	186	181	181	180	180	205	205	245	245	62	62	106	106
MHOM/IL/2003/LRC-L1026	**Maale Adumim, Israel**	**human CL**	***Ltro* MS 042**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**188**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2001/ISL595	Jericho, Palestine	human CL	*Ltro* MS 043	Israel/Palestine	*	66	66	62	62	63	63	112	112	190	190	186	186	181	181	180	180	205	205	245	245	62	62	106	106
MHOM/IL/2009/LRC-L1379	**Almon, Israel**	**human CL**	***Ltro* MS 043**	**Israel/Palestine**	****/*******	**66**	**66**	**62**	**62**	**63**	**63**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/IL/2001/LRC-L836	Tiberias, western side of Sea of Galilee, Israel	human CL	*Ltro* MS 044	Israel/Palestine	*	66	66	62	62	61	61	112	112	190	190	186	186	181	181	180	180	207	225	245	245	62	62	106	106
MHOM/IL/2001/LRC-L837	Tiberias, western side of Sea of Galilee, Israel	human CL	*Ltro* MS 045	Israel/Palestine	*	66	66	62	62	61	61	112	112	190	190	186	186	181	181	180	180	205	205	245	245	62	62	106	112
MHOM/IL/2001/LRC-L838	Tiberias, western side of Sea of Galilee, Israel	human CL	*Ltro* MS 046	Israel/Palestine	*	66	66	60	60	61	61	112	112	190	190	186	186	183	183	180	180	205	205	245	245	62	62	106	106
MHOM/PS/2001/ISL594	Jericho, Palestine	human CL	*Ltro* MS 047	Israel/Palestine	*	66	66	62	62	61	61	112	112	188	188	186	186	183	183	180	180	205	205	245	245	62	62	106	106
MHOM/PS/2002/ISL682	Jericho, Palestine	human CL	*Ltro* MS 048	Israel/Palestine	*	66	66	62	62	61	61	112	112	190	192	186	186	183	183	180	180	205	223	245	245	62	62	106	106
MHOM/PS/2002/ISL681	Jericho, Palestine	human CL	*Ltro* MS 049	Israel/Palestine	*	66	66	62	62	61	61	112	112	190	190	186	186	183	183	180	180	205	223	245	245	62	62	106	106
MHOM/PS/2002/ISL676	Jericho, Palestine	human CL	*Ltro* MS 049	Israel/Palestine	*	66	66	62	62	61	61	112	112	190	190	186	186	183	183	180	180	205	223	245	245	62	62	106	106

						fragment lengths of the microsatellite markers (bp)																							
WHO code	Geographic origin	Source	microsatellite profile	population	Doc	GA1		GA2		GA6		GA9n		LIST7010		LIST7011		LIST7027		LIST7033		LIST7039		LIST7040		4GTG		27GTGn	
MHOM/PS/2002/ISL688	Jericho, Palestine	human CL	*Ltro* MS 050	Israel/Palestine	*	66	66	62	62	61	61	112	112	180	180	186	186	183	183	180	180	205	223	245	245	62	62	106	106
MHOM/PS/2002/ISL692	Jericho, Palestine	human CL	*Ltro* MS 051	Israel/Palestine	*	66	66	62	62	61	61	112	112	180	180	186	186	183	183	180	180	207	225	245	245	62	62	106	106
MHOM/PS/2002/89JnF	Silat El-Hartheyia, Jenin District, Palestine	human CL	*Ltro* MS 051	Israel/Palestine	*	66	66	62	62	61	61	112	112	180	180	186	186	183	183	180	180	207	225	245	245	62	62	106	106
MHOM/PS/2002/ISL700	Jericho, Palestine	human CL	*Ltro* MS 052	Israel/Palestine	*	66	66	62	62	61	61	112	112	180	180	186	186	183	183	180	180	207	207	245	245	62	62	106	106
MHOM/PS/2002/34JnF4	Salit El-Hartheyia, Jenin District, Palestine	human CL	*Ltro* MS 054	Israel/Palestine	*	66	66	62	62	61	61	112	112	190	190	186	186	183	183	180	180	207	225	245	245	62	62	106	106
MHOM/PS/2002/87JnM	El-Yamoon, Jenin District, Palestine	human CL	*Ltro* MS 054	Israel/Palestine	*	66	66	62	62	61	61	112	112	190	190	186	186	183	183	180	180	207	225	245	245	62	62	106	106
MHOM/PS/2002/79JnF20	Salit El-Hartheyia, Jenin District, Palestine	human CL	*Ltro* MS 055	Israel/Palestine	*	66	66	62	62	63	63	112	112	190	190	186	186	181	181	180	180	207	225	245	245	62	62	106	106
MHOM/PS/2002/64JnF4	Salit El-Hartheyia, Jenin District, Palestine	human CL	*Ltro* MS 056	Israel/Palestine	*	66	66	62	62	61	61	112	112	190	190	186	186	183	183	180	180	207	207	245	245	62	62	106	106
MHOM/PS/2002/50Jn20	El-Yamoon, Jenin District, Palestine	human CL	*Ltro* MS 057	Israel/Palestine	*	66	66	62	62	61	61	112	112	192	192	186	186	183	183	180	180	207	207	245	245	62	62	106	106
MHOM/PS/2002/5JnM5	El-Yamoon, Jenin District, Palestine	human CL	*Ltro* MS 058	Israel/Palestine	*	66	66	62	62	61	61	112	112	188	188	186	186	179	179	180	180	197	207	245	245	62	62	106	106
ISER/IL/2002/LRC-L913	Kfar Adumim, Judean Desert, Israel	*P. sergenti*	*Ltro* MS 059	Israel/Palestine	*	66	66	60	62	61	61	112	112	190	190	186	186	185	185	180	180	207	207	245	245	62	62	106	106
MHOM/PS/2000/ISL535	**Jericho, Palestine**	**human CL**	***Ltro* MS 060**	Israel/Palestine	**/*****	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**188**	**188**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/IL/2003/LRC-L1001	**Maale Adumim, Israel**	**human CL**	***Ltro* MS 061**	Israel/Palestine	**/*****	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**192**	**186**	**186**	**179**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/IL/2003/LRC-L1017	**Anatot, Israel**	**human CL**	***Ltro* MS 062**	Israel/Palestine	**/*****	**66**	**66**	**58**	**62**	**61**	**61**	**112**	**112**	**190**	**192**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/IL/2003/LRC-L1016	**Israel**	**human CL**	***Ltro* MS 063**	Israel/Palestine	**/*****	**66**	**66**	**58**	**58**	**61**	**61**	**112**	**112**	**192**	**192**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/IL/2003/LRC-L1018	**Maale Adumim, Israel**	**human CL**	***Ltro* MS 063**	Israel/Palestine	**/*****	**66**	**66**	**58**	**58**	**61**	**61**	**112**	**112**	**192**	**192**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/IL/2004/LRC-L1105	**Maale Adumim, Israel**	**human CL**	***Ltro* MS 064**	Israel/Palestine	**/*****	**66**	**66**	**60**	**60**	**63**	**63**	**112**	**112**	**192**	**192**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2004/ISLAH802	**Jericho, Palestine**	**human CL**	***Ltro* MS 065**	Israel/Palestine	**/*****	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**192**	**192**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/IL/2008/LRC-L1352	**Herodion, Israel**	**human CL**	***Ltro* MS 066**	Israel/Palestine	**/*****	**68**	**68**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/IL/2009/LRC-L1362	**Nokdim, Israel**	**human CL**	***Ltro* MS 067**	Israel/Palestine	**/*****	**66**	**66**	**58**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2010/391Jn	**Jenin District, Palestine**	**human CL**	***Ltro* MS 065**	Israel/Palestine	*****	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**192**	**192**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2010/394Jn	**Jenin District, Palestine**	**human CL**	***Ltro* MS 105**	Israel/Palestine	*****	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**192**	**192**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**109**
MHOM/PS/2010/414Jn	**El-Yamoon, Jenin District, Palestine**	**human CL**	***Ltro* MS 106**	Israel/Palestine	*****	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**109**
MHOM/PS/2010/417JnM13	**El-Yamoon, Jenin District, Palestine**	**human CL**	***Ltro* MS 041**	Israel/Palestine	*****	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2010/433JnJM3	**Jenin city, Palestine**	**human CL**	***Ltro* MS 107**	Israel/Palestine	*****	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**194**	**194**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2011/448Jn	**El-Yamoon, Jenin District, Palestine**	**human CL**	***Ltro* MS 108**	Israel/Palestine	*****	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**109**	**109**
MHOM/PS/2010/LRUJ-1702	**Jericho, Palestine**	**human CL**	***Ltro* MS 042**	Israel/Palestine	*****	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**188**	**190**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/IL/2010/LRC-L1486	**Maale Adumim, Israel**	**human CL**	***Ltro* MS 109**	Israel/Palestine	*****	**64**	**64**	**62**	**62**	**61**	**61**	**112**	**112**	**190**	**192**	**186**	**186**	**181**	**181**	**180**	**180**	**205**	**205**	**245**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2011/ LRUJ-1779	**Bethlehem, Palestine**	**human CL**	***Ltro* MS 124**	Israel/Palestine	*****	**66**	**66**	**58**	**58**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**229**	**180**	**180**	**205**	**205**	**229**	**245**	**62**	**62**	**106**	**106**
MHOM/PS/2011/ LRUJ-1781	**Bethlehem, Palestine**	**human CL**	***Ltro* MS 125**	Israel/Palestine	*****	**66**	**66**	**60**	**60**	**61**	**61**	**112**	**112**	**190**	**190**	**186**	**186**	**181**	**229**	**180**	**180**	**205**	**205**	**229**	**245**	**62**	**62**	**106**	**106**
MHOM/SU/1966/III	**Soviet Union**	**human CL**	***Ltro* MS 110**	Middle East	*****	**66**	**66**	**56**	**58**	**61**	**61**	**112**	**112**	**176**	**180**	**178**	**178**	**169**	**175**	**180**	**180**	**195**	**195**	**229**	**247**	**62**	**62**	**109**	**109**
MHOM/AF/1988/KK26	**Afaghanistan**	**human CL**	***Ltro* MS 112**	Middle East	*****	**66**	**66**	**56**	**58**	**61**	**61**	**112**	**112**	**176**	**180**	**178**	**178**	**169**	**175**	**178**	**178**	**195**	**195**	**229**	**247**	**62**	**62**	**109**	**109**
MHOM/SY/1995/LSL25	**Syria**	**human CL**	***Ltro* MS 117**	Middle East	*****	**66**	**66**	**56**	**58**	**61**	**61**	**112**	**114**	**176**	**180**	**178**	**178**	**169**	**175**	**178**	**178**	**195**	**195**	**229**	**247**	**62**	**62**	**109**	**109**
MHOM/IR/1995/5043	**Iran**	**human CL**	***Ltro* MS 118**	Middle East	*****	**66**	**66**	**56**	**58**	**61**	**61**	**112**	**114**	**176**	**180**	**178**	**178**	**169**	**175**	**178**	**178**	**195**	**195**	**229**	**247**	**62**	**62**	**112**	**112**
MHOM/TR/1999/EP32	**Turkey**	**human CL**	***Ltro* MS 119**	Middle East	*****	**66**	**66**	**56**	**58**	**61**	**61**	**112**	**114**	**180**	**180**	**178**	**178**	**169**	**175**	**178**	**178**	**195**	**195**	**229**	**247**	**62**	**62**	**109**	**109**
MHOM/IR/2000/LEM4036	**Iran**	**human CL**	***Ltro* MS 121**	Middle East	*****	**66**	**66**	**56**	**58**	**61**	**61**	**112**	**112**	**180**	**180**	**178**	**178**	**169**	**175**	**178**	**178**	**195**	**195**	**229**	**247**	**62**	**62**	**109**	**109**
MHOM/SY/2009/LRUJ-1550	**Syria**	**human CL**	***Ltro* MS 122**	Middle East	*****	**66**	**66**	**56**	**58**	**61**	**61**	**112**	**112**	**176**	**180**	**178**	**178**	**169**	**175**	**178**	**178**	**195**	**195**	**229**	**247**	**62**	**62**	**115**	**115**
MHOM/SY/2009/LRUJ-1564	**Syria**	**human CL**	***Ltro* MS 123**	Middle East	*****	**66**	**66**	**58**	**58**	**61**	**61**	**112**	**112**	**176**	**180**	**178**	**178**	**169**	**175**	**178**	**178**	**195**	**195**	**229**	**247**	**62**	**62**	**115**	**115**
MHOM/PS/2002/63JnF21	**El-Yamoon, Jenin District, Palestine**	**human CL**	***Ltro* MS 020**	India/Palestine	**/*****	**66**	**66**	**56**	**56**	**61**	**61**	**112**	**114**	**176**	**180**	**192**	**192**	**169**	**177**	**178**	**178**	**197**	**197**	**229**	**247**	**62**	**62**	**112**	**112**
MHOM/PS/2002/20JnM3	El-Yamoon, Jenin District, Palestine	human CL	*Ltro* MS 021	India/Palestine	*	66	66	58	58	61	61	112	114	176	180	192	192	169	177	178	178	195	195	229	247	62	62	112	112
MHOM/PS/2002/52JnM18	El-Yamoon, Jenin District, Palestine	human CL	*Ltro* MS 022	India/Palestine	*	66	66	56	58	61	61	112	114	176	180	192	192	169	177	178	178	197	197	229	247	62	62	112	112
MHOM/IN/2006/BKC-1	**Bikaner, Rajasthan, India**	**human CL**	***Ltro* MS 030**	India/Palestine	**	**66**	**66**	**56**	**58**	**61**	**61**	**112**	**114**	**176**	**176**	**192**	**192**	**169**	**177**	**178**	**178**	**195**	**195**	**229**	**247**	**62**	**62**	**112**	**118**
MHOM/IN/2006/BKC-2	**Bikaner, Rajasthan, India**	**human CL**	***Ltro* MS 030**	India/Palestine	**	**66**	**66**	**56**	**58**	**61**	**61**	**112**	**114**	**176**	**176**	**192**	**192**	**169**	**177**	**178**	**178**	**195**	**195**	**229**	**247**	**62**	**62**	**112**	**118**
MHOM/IN/2007/BKC-10	**Bikaner, Rajasthan, India**	**human CL**	***Ltro* MS 030**	India/Palestine	**	**66**	**66**	**56**	**58**	**61**	**61**	**112**	**114**	**176**	**176**	**192**	**192**	**169**	**177**	**178**	**178**	**195**	**195**	**229**	**247**	**62**	**62**	**112**	**118**
MHOM/IN/2007/BKC-11	**Bikaner, Rajasthan, India**	**human CL**	***Ltro* MS 030**	India/Palestine	**	**66**	**66**	**56**	**58**	**61**	**61**	**112**	**114**	**176**	**176**	**192**	**192**	**169**	**177**	**178**	**178**	**195**	**195**	**229**	**247**	**62**	**62**	**112**	**118**
MHOM/IN/2007/BKC-15	**Bikaner, Rajasthan, India**	**human CL**	***Ltro* MS 030**	India/Palestine	**	**66**	**66**	**56**	**58**	**61**	**61**	**112**	**114**	**176**	**176**	**192**	**192**	**169**	**177**	**178**	**178**	**195**	**195**	**229**	**247**	**62**	**62**	**112**	**118**
MHOM/IN/2007/BKC-28	**Bikaner, Rajasthan, India**	**human CL**	***Ltro* MS 030**	India/Palestine	**	**66**	**66**	**56**	**58**	**61**	**61**	**112**	**114**	**176**	**176**	**192**	**192**	**169**	**177**	**178**	**178**	**195**	**195**	**229**	**247**	**62**	**62**	**112**	**118**
MHOM/IN/2006/BKC-3	**Bikaner, Rajasthan, India**	**human CL**	***Ltro* MS 031**	India/Palestine	**	**66**	**66**	**56**	**58**	**61**	**61**	**112**	**114**	**176**	**176**	**192**	**192**	**169**	**177**	**178**	**178**	**195**	**195**	**231**	**249**	**62**	**62**	**112**	**118**
MHOM/IN/2006/BKC-5	**Bikaner, Rajasthan, India**	**human CL**	***Ltro* MS 032**	India/Palestine	**	**nd**	**nd**	**56**	**58**	**61**	**61**	**112**	**114**	**176**	**176**	**192**	**192**	**169**	**177**	**178**	**178**	**195**	**195**	**229**	**247**	**62**	**62**	**112**	**118**
MHOM/IL/1980/Singer	Arad, Israel	human CL	*Ltro* MS 077	(IL/PS/JO/IN)	*	66	66	56	56	61	61	114	114	184	184	192	192	175	175	176	176	211	211	229	229	62	62	115	115
MHOM/PS/2002/18JnMF4	**Meselya, Jenin District, Palestine**	**human CL**	***Ltro* MS 102**	(IL/PS/JO/IN)	*****	**66**	**66**	**56**	**58**	**61**	**61**	**112**	**114**	**176**	**180**	**192**	**192**	**169**	**175**	**178**	**178**	**195**	**195**	**226**	**226**	**62**	**62**	**109**	**112**
MHOM/IN/1984/C87	Vaishali, Bihar, India	human VL	*Ltro* MS 018	(IL/PS/JO/IN)	*	66	66	56	56	61	61	112	114	176	180	192	192	169	177	176	176	195	195	247	247	62	62	112	118
MHOM/IN/1997/K26	Vaishali, Bihar, India	human VL	*Ltro* MS 018	(IL/PS/JO/IN)	*	66	66	56	56	61	61	112	114	176	180	192	192	169	177	176	176	195	195	247	247	62	62	112	118
MHOM/JO/1999/K-Z	Jordan	human CL	*Ltro* MS 019	(IL/PS/JO/IN)	*	66	66	56	56	61	61	112	114	176	180	192	192	169	181	176	176	195	195	229	247	62	62	nd	nd
MHOM/AZ/1958/OD	Kirovobad, Azerbaijan	human CL	*Ltro* MS 001	old strains/TR	*	66	66	56	56	61	61	112	114	174	174	192	192	169	177	176	176	197	197	229	247	62	62	112	118
MHOM/IL/1949/LRC-L43	Abu Ghosh, Israel	human VL	*Ltro* MS 001	old strains/TR	*	66	66	56	56	61	61	112	114	174	174	192	192	169	177	176	176	197	197	229	247	62	62	112	118
MHOM/AZ/1974/SAF-K27	Kirovobad, Azerbaijan	human CL	*Ltro* MS 002	old strains/TR	*	66	66	56	56	61	61	112	114	176	176	192	192	169	177	176	176	197	197	229	247	62	62	112	118
MHOM/AZ/1958/NLB305	Kirovobad, Azerbaijan	human CL	*Ltro* MS 002	old strains/TR	*	66	66	56	56	61	61	112	114	176	176	192	192	169	177	176	176	197	197	229	247	62	62	112	118

						fragment lengths of the microsatellite markers (bp)																																			
WHO code	**Geographic origin**	**Source**	**microsatellite profile**	**population**	**Doc**	**GA1**		**GA2**		**GA6**		**GA9n**		**LIST7010**		**LIST7011**		**LIST7027**		**LIST7033**		**LIST7039**		**LIST7040**		**4GTG**		**27GTGn**	
MRAT/IQ/1973/MRCB-IBF	Baghdad, Iraq	rat (*Rattus sp.*)	*Ltro* MS 003	old strains/TR	*	66	66	56	56	61	61	112	114	176	180	192	192	177	177	176	176	197	197	229	247	62	62	nd	nd
MHOM/KE/1981/NLB030B	Kikumini, Masinga, Kenya	human VL	*Ltro* MS 005	old strains/TR	*	66	66	56	56	61	61	112	114	176	176	192	192	169	177	176	176	197	197	247	247	62	62	112	118
MHOM/IQ/1979/Ldj1	Iraq	human VL	*Ltro* MS 005	old strains/TR	*	66	66	56	56	61	61	112	114	176	176	192	192	169	177	176	176	197	197	247	247	62	62	112	118
MHOM/KE/1981/NLB029B	Kikumini, Masinga, Kenya	human VL	*Ltro* MS 006	old strains/TR	*	66	66	56	56	61	61	112	114	176	176	192	192	169	177	176	176	197	197	247	247	62	62	nd	nd
MHOM/IQ/1966/L75	Baghdad, Iraq	human CL	*Ltro* MS 007	old strains/TR	*	66	66	56	56	61	61	112	114	176	180	192	192	171	179	176	176	197	201	229	247	62	62	118	118
MHOM/AZ/1980/K28	Kirovobad, Azerbaijan	human CL	*Ltro* MS 012	old strains/TR	*	66	66	56	56	61	61	112	114	176	180	192	194	169	177	176	176	197	225	229	247	62	62	nd	nd
MHOM/AZ/1973/K25A	Kirovobad, Azerbaijan	human CL	*Ltro* MS 013	old strains/TR	*	66	66	56	56	61	61	112	114	176	180	192	192	169	177	176	176	197	197	229	247	62	62	112	118
MHOM/IN/1991/K112	Vaishali, Bihar, India	human VL	*Ltro* MS 017	old strains/TR	*	66	66	56	56	61	61	112	114	176	176	192	192	169	177	176	176	207	207	247	247	62	62	112	118
MHOM/PS/2002/18JnF4	El-Yamoon, Jenin District, Palestine	human CL	*Ltro* MS 023	old strains/TR	*	66	66	56	56	61	61	112	114	176	180	192	192	169	177	176	176	197	197	229	247	62	62	112	112
MHOM/TR/1995/URFA3	Sanliurfa, Turkey	human CL	*Ltro* MS 025	(old strains/TR)	*	66	66	56	56	61	61	112	114	192	192	194	194	169	177	176	176	197	197	229	247	62	62	nd	nd
MHOM/IN/1979/DD7	India	human VL	*Ltro* MS 004	(old strains/TR)	*	66	66	56	56	61	61	114	114	180	180	192	192	169	175	176	176	197	197	229	229	62	62	112	118
MHOM/IL/1990/P283	Kfar Adumim, Judean Desert, Israel	human CL	*Ltro* MS 008	(old strains/TR)	*	66	66	56	56	61	61	112	114	180	180	194	194	169	177	178	178	197	197	229	247	62	62	118	118
MHOM/TR/1995/EP13	south-eastern Turkey	human CL	*Ltro* MS 009	(old strains/TR)	*	66	66	56	56	61	61	112	114	nd	nd	194	194	171	179	176	176	197	197	229	247	62	62	nd	nd
MHOM/TR/1998/YO007	Aydin region, south of Izmir, Turkey	human CL	*Ltro* MS 014	(old strains/TR)	*	66	66	56	56	61	61	112	114	180	180	194	194	169	177	176	176	197	197	229	247	62	62	nd	nd
MHOM/TR/1998/YO004	Adana, Turkey	human CL	*Ltro* MS 014	(old strains/TR)	*	66	66	56	56	61	61	112	114	180	180	194	194	169	177	176	176	197	197	229	247	62	62	nd	nd
MHOM/TR/1998/YO005	Edessa, south-eastern Turkey	human CL	*Ltro* MS 015	(old strains/TR)	*	66	66	56	56	61	61	112	114	180	180	194	194	169	177	176	176	197	229	229	247	62	62	nd	nd
MHOM/IL/1959/LRC-L22	Negev Desert, Israel	human CL	*Ltro* MS 016	(old strains/TR)	*	66	66	56	56	61	61	112	114	174	178	192	192	169	177	180	180	197	197	229	247	62	62	112	118
MHOM/TR/1995/URFA7	Sanliurfa, Turkey	human CL	*Ltro* MS 028	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	112	118
MHOM/TR/1995/URFA40	Sanliurfa, Turkey	human CL	*Ltro* MS 028	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	112	118
MHOM/TR/1995/URFA1	Sanliurfa, Turkey	human CL	*Ltro* MS 024	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	nd	nd	62	62	nd	nd
MHOM/TR/1995/URFA5	Sanliurfa, Turkey	human CL	*Ltro* MS 026	Sanliurfa	*	66	66	56	56	nd	nd	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	nd	nd
MHOM/TR/1995/URFA10	Sanliurfa, Turkey	human CL	*Ltro* MS 027	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	nd	nd
MHOM/TR/1995/URFA11	Sanliurfa, Turkey	human CL	*Ltro* MS 027	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	nd	nd
MHOM/TR/1995/URFA12	Sanliurfa, Turkey	human CL	*Ltro* MS 027	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	nd	nd
MHOM/TR/1995/URFA14	Sanliurfa, Turkey	human CL	*Ltro* MS 027	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	nd	nd
MHOM/TR/1995/URFA15	Sanliurfa, Turkey	human CL	*Ltro* MS 027	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	nd	nd
MHOM/TR/1995/URFA16	Sanliurfa, Turkey	human CL	*Ltro* MS 027	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	nd	nd
MHOM/TR/1995/URFA19	Sanliurfa, Turkey	human CL	*Ltro* MS 027	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	nd	nd
MHOM/TR/1995/URFA20	Sanliurfa, Turkey	human CL	*Ltro* MS 027	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	nd	nd
MHOM/TR/1995/URFA21	Sanliurfa, Turkey	human CL	*Ltro* MS 027	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	nd	nd
MHOM/TR/1995/URFA23	Sanliurfa, Turkey	human CL	*Ltro* MS 027	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	nd	nd
MHOM/TR/1995/URFA27	Sanliurfa, Turkey	human CL	*Ltro* MS 027	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	nd	nd
MHOM/TR/1995/URFA39	Sanliurfa, Turkey	human CL	*Ltro* MS 027	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	nd	nd
MHOM/TR/1995/URFA29	Sanliurfa, Turkey	human CL	*Ltro* MS 028	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	112	118
MHOM/TR/1995/URFA42	Sanliurfa, Turkey	human CL	*Ltro* MS 028	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	112	118
MHOM/TR/1995/URFA46	Sanliurfa, Turkey	human CL	*Ltro* MS 028	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	112	118
MHOM/TR/1995/URFA57	Sanliurfa, Turkey	human CL	*Ltro* MS 028	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	112	118
MHOM/TR/1995/URFA62	Sanliurfa, Turkey	human CL	*Ltro* MS 028	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	112	118
MHOM/TR/1995/URFA65	Sanliurfa, Turkey	human CL	*Ltro* MS 028	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	112	118
MHOM/TR/1995/URFA26	Sanliurfa, Turkey	human CL	*Ltro* MS 028	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	112	118
MHOM/TR/1995/URFA36	Sanliurfa, Turkey	human CL	*Ltro* MS 029	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	112	112
MHOM/TR/1995/URFA37	Sanliurfa, Turkey	human CL	*Ltro* MS 029	Sanliurfa	*	66	66	56	56	61	61	112	114	178	178	194	194	169	177	176	176	197	197	229	247	62	62	112	112
MHOM/ET/1994/Gere	**Arba Minch, southern Ethiopia**	**human DCL**	***Laet* MS 001**	***L. aethiopica***	****	**68**	**68**	**52**	**66**	**61**	**61**	**112**	**114**	**172**	**172**	**nd**	**nd**	**175**	**187**	**206**	**208**	**193**	**193**	**213**	**221**	**59**	**68**	**94**	**94**
MHOM/ET/1994/Abauye	**Arba Minch, southern Ethiopia**	**human DCL**	***Laet* MS 002**	***L. aethiopica***	****	**68**	**68**	**52**	**66**	**61**	**61**	**110**	**110**	**172**	**172**	**nd**	**nd**	**175**	**187**	**206**	**208**	**193**	**193**	**213**	**221**	**59**	**68**	**94**	**94**
MHOM/ET/1994/1610	**Wolega, northern Ethiopia**	**human LCL**	***Laet* MS 003**	***L. aethiopica***	****	**70**	**70**	**60**	**60**	**61**	**61**	**110**	**110**	**172**	**172**	**184**	**184**	**209**	**215**	**184**	**184**	**nd**	**nd**	**255**	**255**	**68**	**68**	**106**	**106**
MHOM/ET/1994/1470	**Shoa, central Ethiopia**	**human LCL**	***Laet* MS 004**	***L. aethiopica***	****	**68**	**68**	**62**	**66**	**61**	**61**	**110**	**110**	**172**	**172**	**182**	**182**	**175**	**175**	**190**	**196**	**nd**	**nd**	**257**	**257**	**68**	**68**	**94**	**94**
MHOM/ET/1994/Wandera	**Arba Minch, southern Ethiopia**	**human LCL**	***Laet* MS 005**	***L. aethiopica***	****	**68**	**68**	**52**	**66**	**61**	**61**	**nd**	**nd**	**172**	**172**	**nd**	**nd**	**175**	**187**	**206**	**208**	**193**	**193**	**213**	**221**	**59**	**68**	**94**	**94**
MHOM/ET/1987/Kassaye	**Gojam, northern Ethiopia**	**human DCL**	***Laet* MS 006**	***L. aethiopica***	****	**66**	**66**	**60**	**60**	**61**	**61**	**nd**	**nd**	**172**	**172**	**182**	**184**	**177**	**177**	**184**	**184**	**nd**	**nd**	**255**	**255**	**68**	**68**	**94**	**94**
MHOM/ET/1972/L100	**Wolega, northern Ethiopia**	**human DCL**	***Laet* MS 007**	***L. aethiopica***	****	**60**	**66**	**60**	**60**	**61**	**61**	**110**	**110**	**172**	**172**	**174**	**180**	**209**	**209**	**180**	**184**	**203**	**203**	**255**	**255**	**68**	**68**	**106**	**106**
MHOM/ET/1985/LRC-L494	**Ethiopia**	**human LCL**	***Laet* MS 008**	***L. aethiopica***	****	**66**	**70**	**60**	**60**	**61**	**61**	**110**	**110**	**172**	**172**	**160**	**160**	**175**	**175**	**190**	**190**	**nd**	**nd**	**257**	**257**	**68**	**68**	**94**	**94**
MHOM/IL/2000/LRC-L786 (=Omri)	northern side of Sea of Galilee, Israel	human CL	*Ltro* MS 081	Northern Galilee	*	66	66	66	66	61	61	110	110	180	180	190	190	173	173	nd	nd	201	201	245	245	62	62	nd	nd
IARA/IL/2000/Amnumfly	northern side of Sea of Galilee, Israel	*P. arabicus*	*Ltro* MS 082	Northern Galilee	*	66	66	66	66	61	61	110	110	180	180	190	190	173	173	180	180	201	201	245	245	65	65	115	115
MHOM/IL/2002/LRC-L863 (=Ofri)	northern side of Sea of Galilee, Israel	human CL	*Ltro* MS 083	Northern Galilee	*	66	66	66	66	61	61	110	110	182	184	190	190	175	175	180	180	201	219	245	245	65	65	115	115
IARA/IL/2002/LRC-L906	northern side of Sea of Galilee, Israel	*P. arabicus*	*Ltro* MS 084	Northern Galilee	*	66	66	66	66	63	63	110	110	180	180	190	190	173	173	180	180	203	203	245	245	65	65	115	115
IARA/IL/2002/LRC-L907	northern side of Sea of Galilee, Israel	*P. arabicus*	*Ltro* MS 085	Northern Galilee	*	66	66	66	66	61	61	110	110	180	180	190	190	175	175	180	180	203	203	245	245	65	65	97	97
ISER/IL/2002/LRC-L909	northern side of Sea of Galilee, Israel	*P. sergenti*	*Ltro* MS 086	Northern Galilee	*	66	66	64	64	61	61	110	110	182	182	190	190	177	177	180	180	203	203	245	245	65	65	115	115
IARA/IL/2002/LRC-L910	northern side of Sea of Galilee, Israel	*P. arabicus*	*Ltro* MS 087	Northern Galilee	*	66	66	64	64	61	61	110	110	180	180	190	190	177	177	180	180	203	203	245	245	65	65	115	115
IARA/IL/2002/LRC-L947	northern side of Sea of Galilee, Israel	*P. arabicus*	*Ltro* MS 088	Northern Galilee	*	66	66	64	64	61	61	110	110	180	182	190	190	177	177	180	180	203	203	245	245	65	65	115	115

						fragment lengths of the microsatellite markers (bp)																									
WHO code	Geographic origin	Source	microsatellite profile	population	Doc	GA1		GA2		GA6		GA9n		LIST7010		LIST7011		LIST7027		LIST7033		LIST7039		LIST7040		4GTG		27GTGn	
MHOM/IL/2003/LRC-L961	**Korazim, northern side of Sea of Galilee, Israel**	**human CL**	***Ltro* MS 089**	**Northern Galilee**	****/*******	**66**	**66**	**66**	**66**	**61**	**61**	**110**	**110**	**180**	**180**	**190**	**190**	**175**	**175**	**180**	**180**	**201**	**201**	**245**	**245**	**65**	**65**	**115**	**115**
MPRV/IL/2003/HYRAX107	**Amnum, northern side of Sea of Galilee, Israel**	**hyrax (*Procavia sp.*)**	***Ltro* MS 089**	**Northern Galilee**	****/*******	**66**	**66**	**66**	**66**	**61**	**61**	**110**	**110**	**180**	**180**	**190**	**190**	**175**	**175**	**180**	**180**	**201**	**201**	**245**	**245**	**65**	**65**	**115**	**115**
MHOM/MA/1988/LEM1314	Marrakech area, Morocco	human CL	*Ltro* MS 090	MoroccoA/TR	*	66	66	54	54	63	63	112	112	nd	nd	194	194	175	175	176	176	217	217	229	233	62	62	115	115
MHOM/MA/1988/LEM1451	Marrakech area, Morocco	human CL	*Ltro* MS 091	MoroccoA/TR	*	66	66	54	54	61	61	112	112	nd	nd	192	192	175	175	176	176	217	217	229	229	62	62	115	115
MHOM/MA/1988/LEM1528	Marrakech area, Morocco	human CL	*Ltro* MS 093	MoroccoA/TR	*	66	66	56	56	61	61	112	112	194	194	194	194	175	175	176	176	217	217	229	233	62	62	115	115
MHOM/MA/1989/LEM1591	Marrakech area, Morocco	human CL	*Ltro* MS 094	MoroccoA/TR	*	66	66	56	56	61	61	112	112	188	188	188	188	175	175	176	176	217	221	229	229	62	62	115	115
MHOM/MA/1990/LEM1880	Marrakech area, Morocco	human CL	*Ltro* MS 094	MoroccoA/TR	*	66	66	56	56	61	61	112	112	188	188	188	188	175	175	176	176	217	221	229	229	62	62	115	115
MHOM/MA/1989/LEM1879	Marrakech area, Morocco	human CL	*Ltro* MS 097	MoroccoA/TR	*	66	66	54	54	61	61	112	112	194	194	192	192	175	175	176	176	217	217	229	229	62	62	115	115
MHOM/MA/2000/INH-W02	**Chichaoua province, Morocco**	**human CL**	***Ltro* MS 098**	**MoroccoA/TR**	*******	**66**	**66**	**54**	**56**	**61**	**61**	**112**	**112**	**188**	**190**	**188**	**188**	**175**	**175**	**178**	**178**	**221**	**223**	**229**	**229**	**62**	**62**	**115**	**115**
MHOM/MA/2000/INH-W04	**Chichaoua province, Morocco**	**human CL**	***Ltro* MS 098**	**MoroccoA/TR**	*******	**66**	**66**	**54**	**56**	**61**	**61**	**112**	**112**	**188**	**190**	**188**	**188**	**175**	**175**	**178**	**178**	**221**	**223**	**229**	**229**	**62**	**62**	**115**	**115**
MHOM/MA/2000/INH-W05	**Chichaoua province, Morocco**	**human CL**	***Ltro* MS 098**	**MoroccoA/TR**	*******	**66**	**66**	**54**	**56**	**61**	**61**	**112**	**112**	**188**	**190**	**188**	**188**	**175**	**175**	**178**	**178**	**221**	**223**	**229**	**229**	**62**	**62**	**115**	**115**
MHOM/MA/2000/INH-W13	**Chichaoua province, Morocco**	**human CL**	***Ltro* MS 099**	**MoroccoA/TR**	*******	**66**	**66**	**56**	**56**	**61**	**61**	**112**	**112**	**188**	**188**	**188**	**188**	**175**	**175**	**178**	**178**	**211**	**219**	**229**	**229**	**62**	**62**	**115**	**118**
MHOM/MA/2000/INH-W14	**Chichaoua province, Morocco**	**human CL**	***Ltro* MS 099**	**MoroccoA/TR**	*******	**66**	**66**	**56**	**56**	**61**	**61**	**112**	**112**	**188**	**188**	**188**	**188**	**175**	**175**	**178**	**178**	**211**	**219**	**229**	**229**	**62**	**62**	**115**	**118**
MHOM/MA/2000/INH-W16	**Chichaoua province, Morocco**	**human CL**	***Ltro* MS 099**	**MoroccoA/TR**	*******	**66**	**66**	**56**	**56**	**61**	**61**	**112**	**112**	**188**	**188**	**188**	**188**	**175**	**175**	**178**	**178**	**211**	**219**	**229**	**229**	**62**	**62**	**115**	**118**
MHOM/MA/2000/INH-W17	**Chichaoua province, Morocco**	**human CL**	***Ltro* MS 099**	**MoroccoA/TR**	*******	**66**	**66**	**56**	**56**	**61**	**61**	**112**	**112**	**188**	**188**	**188**	**188**	**175**	**175**	**178**	**178**	**211**	**219**	**229**	**229**	**62**	**62**	**115**	**118**
MHOM/MA/1989/LEM1781	**Morocco**	**human CL**	***Ltro* MS 113**	**MoroccoA/TR**	*********	**66**	**66**	**56**	**56**	**61**	**61**	**112**	**112**	**192**	**194**	**178**	**178**	**175**	**175**	**178**	**178**	**197**	**197**	**229**	**229**	**62**	**62**	**115**	**115**
MHOM/MA/1995/UERL9	**Morocco**	**human CL**	***Ltro* MS 120**	**MoroccoA/TR**	*********	**66**	**66**	**54**	**56**	**61**	**61**	**112**	**112**	**192**	**192**	**178**	**178**	**173**	**175**	**178**	**178**	**215**	**215**	**229**	**229**	**62**	**62**	**115**	**115**
MHOM/TR/1997/YO001	south-eastern Turkey	human CL	*Ltro* MS 078	(MoroccoA/TR)	*	66	66	56	56	61	61	nd	nd	180	180	188	188	173	173	176	176	227	227	233	233	62	62	nd	nd
MHOM/TR/1998/YO002	Adana, Turkey	human CL	*Ltro* MS 079	(MoroccoA/TR)	*	66	66	56	56	61	61	112	112	nd	nd	188	188	173	173	176	176	227	227	233	233	62	62	nd	nd
MHOM/MA/1992/LPN79	**Morocco**	**human CL**	***Ltro* MS 115**	**(MoroccoA/TR)**	*********	**66**	**66**	**54**	**54**	**61**	**61**	**112**	**112**	**192**	**194**	**178**	**178**	**175**	**175**	**178**	**178**	**173**	**173**	**233**	**233**	**62**	**62**	**109**	**109**
MHOM/MA/1995/LEM3015	**Morocco**	**human CL**	***Ltro* MS 116**	**(MoroccoA/TR)**	*********	**66**	**66**	**54**	**58**	**61**	**61**	**110**	**112**	**182**	**188**	**178**	**178**	**169**	**175**	**178**	**188**	**217**	**217**	**229**	**243**	**62**	**62**	**103**	**103**
MHOM/MA/1988/LEM1452	Marrakech area, Morocco	human CL	*Ltro* MS 092	MoroccoB	*	62	62	58	58	59	59	110	114	nd	nd	190	190	169	169	168	168	221	221	229	229	62	62	112	112
ISER/MA/1989/LEM1694	Marrakech area, Morocco	*P. sergenti*	*Ltro* MS 095	MoroccoB	*	62	62	58	58	59	59	110	110	188	188	190	190	169	169	188	190	221	221	225	225	62	62	112	112
ISER/MA/1989/LEM1828	Marrakech area, Morocco	*P. sergenti*	*Ltro* MS 096	MoroccoB	*	62	62	58	58	59	59	110	110	182	190	190	190	169	169	186	186	221	225	225	225	62	62	115	115
MHOM/MA/2000/INH-W10	**Chichaoua province, Morocco**	**human CL**	***Ltro* MS 101**	**MoroccoB**	*******	**62**	**62**	**58**	**58**	**59**	**59**	**110**	**110**	**182**	**190**	**190**	**190**	**169**	**169**	**190**	**190**	**223**	**223**	**225**	**225**	**62**	**62**	**112**	**112**
MPRV/NA/1976/HYRAX2	Sandmodder, Namibia	hyrax (*Procavia sp.*)	*Ltro* MS 068	Namibia	*	66	66	62	62	59	59	114	114	178	178	174	174	171	171	176	176	197	197	253	253	68	68	112	112
IROS/NA/1976/ROSSI-II	Keetmanshoop, Namibia	*P. rossi*	*Ltro* MS 069	Namibia	*	66	66	62	62	59	59	114	114	178	178	174	174	171	171	176	176	199	199	253	253	68	68	112	112
IROS/NA/1980/HD3	Hardapdam (near Marienthal), Namibia	*P. rossi*	*Ltro* MS 070	Namibia	*	66	66	62	62	59	59	112	112	178	178	174	174	171	171	178	178	197	197	253	253	68	68	112	115
MHOM/NA/1984/K1	Kamanjab (near Ethosa-NP), Namibia	human CL	*Ltro* MS 071	Namibia	*	66	66	62	62	59	59	112	112	178	178	174	174	171	171	178	178	197	197	nd	nd	68	68	nd	nd
MHOM/KE/1981/NLB162	Lion Camp, Nakuru, Kenya	human CL	*Ltro* MS 073	(Namibia)	*	66	66	62	62	61	61	112	112	180	180	190	190	175	175	178	178	199	199	253	253	68	68	106	106
MHOM/KE/1984/NLB254	Laikipia, Kenya	human CL	*Ltro* MS 072	(KE/TN/YE/MA)	*	58	66	58	68	61	61	110	110	176	176	184	184	171	191	180	180	199	199	245	245	62	62	112	112
MHOM/KE/1985/NLB545	Laikipia, Kenya	human CL	*Ltro* MS 072	(KE/TN/YE/MA)	*	58	66	58	68	61	61	110	110	176	176	184	184	171	191	180	180	199	199	245	245	62	62	112	112
MHOM/KE/1984/NLB297	Laikipia, Kenya	human CL	*Ltro* MS 075	(KE/TN/YE/MA)	*	66	66	70	70	63	63	110	110	176	176	182	182	173	173	180	180	199	199	245	245	68	68	106	106
MHOM/TN/1988/TAT3	southern Tunisia	human CL	*Ltro* MS 076	(KE/TN/YE/MA)	*	66	66	58	58	61	61	110	110	182	182	184	184	171	171	178	178	199	199	229	229	62	62	106	106
MHOM/TN/1980/LEM163	southern Tunisia	human CL	*Ltro* MS 080	(KE/TN/YE/MA)	*	66	66	58	58	61	61	110	110	190	190	184	184	169	169	180	180	209	223	227	227	62	62	nd	nd
MHOM/MA/2000/INH-W09	**Chichaoua province, Morocco**	**human CL**	***Ltro* MS 100**	**(KE/TN/YE/MA)**	*******	**66**	**66**	**54**	**58**	**61**	**61**	**110**	**112**	**178**	**178**	**184**	**184**	**169**	**175**	**178**	**178**	**213**	**213**	**227**	**229**	**62**	**62**	**106**	**115**
MHOM/YE/1986/LEM1000	**Yemen**	**human CL**	***Ltro* MS 111**	**(KE/TN/YE/MA)**	*********	**66**	**66**	**62**	**62**	**61**	**61**	**112**	**112**	**182**	**182**	**178**	**178**	**169**	**169**	**180**	**180**	**203**	**203**	**247**	**247**	**62**	**62**	**109**	**109**
MHOM/KE/1991/EB135	**Kenya**	**human CL**	***Ltro* MS 114**	**(KE/TN/YE/MA)**	*********	**66**	**66**	**70**	**70**	**61**	**61**	**110**	**110**	**176**	**176**	**178**	**178**	**171**	**171**	**180**	**180**	**197**	**197**	**245**	**245**	**68**	**68**	**106**	**106**
MHOM/KE/1984/NLB248	Laikipia, Kenya	human CL	*Ltro* MS 074	shared membership	*	66	66	66	66	61	61	114	114	184	184	182	182	175	175	178	178	195	195	245	245	68	68	112	112
MHOM/PS/2002/20JnYM3	**El-Yamoon, Jenin District, Palestine**	**human CL**	***Ltro* MS 103**	**shared membership**	*********	**66**	**66**	**58**	**58**	**61**	**61**	**112**	**114**	**176**	**180**	**192**	**192**	**169**	**175**	**178**	**178**	**195**	**195**	**229**	**247**	**62**	**62**	**109**	**112**
MHOM/PS/2002/52JnYM18	**El-Yamoon, Jenin District, Palestine**	**human CL**	***Ltro* MS 104**	**shared membership**	*********	**66**	**66**	**56**	**58**	**61**	**61**	**112**	**114**	**176**	**180**	**192**	**192**	**169**	**175**	**178**	**178**	**195**	**195**	**229**	**247**	**62**	**62**	**109**	**112**

6.1.4 Worldwide population structure of *L. tropica* and *L. aethiopica*

All strains of *L. tropica* and *L. aethiopica* were analysed using various programs each using a different algorithm for their calculations. Most of the results were congruent between the different approaches, while some showed differences.

6.1.4.1 Descriptive statistics of the microsatellite loci

In total, 121 different microsatellite profiles were found for the 195 strains of *L. tropica* (Table 11). Of these, 17 profiles were shared by two strains each, two were shared by three strains each and others were shared by four, six, nine, 12 and 27 strains, respectively (Table 12). The 27 strains, which presented the same profile, *Ltro*MS 041, were all collected in Israel and Palestine. The 12 and nine strains sharing the profiles *Ltro*MS 027 and *Ltro*MS 028, respectively, were all isolated in Sanliurfa, Turkey, during a local outbreak in 1995. The latter two varied only in one marker, the first failed to amplify locus 27GTGn, the latter did not (Table 13).

Table 12. Microsatellite profiles which were shared by more than one strain of *L. tropica*.

microsatellite profile	number of strains	geographic origin	source
***Ltro*MS 001**	2	Kirovobad, Azerbaijan	human CL
		Abu Ghosh, Israel	human VL
***Ltro*MS 002**	2	Kirovobad, Azerbaijan	human CL
***Ltro*MS 005**	2	Kikumini, Masinga, Kenya	human VL
		Iraq	human VL
***Ltro*MS 014**	2	Aydin region, south of Izmir, Turkey	human CL
		Adana, Turkey	human CL
***Ltro*MS 018**	2	Vaishali, Bihar, India	human VL
***Ltro*MS 027**	12	Sanliurfa, Turkey	human CL
***Ltro*MS 028**	9	Sanliurfa, Turkey	human CL
***Ltro*MS 029**	2	Sanliurfa, Turkey	human CL
***Ltro*MS 030**	6	Bikaner, Rajasthan, India	human CL
***Ltro*MS 034**	2	Judean Desert, Israel	human CL
		Judean Desert, Israel	*P. sergenti*
***Ltro*MS 039**	2	Samaria, Palestine	human CL
***Ltro*MS 041**	27	Israel/Palestine (14x Jenin District, 5x Jericho)	human CL / 1x *P. sergenti*
***Ltro*MS 042**	3	Israel/Palestine	human CL
***Ltro*MS 043**	2	Israel/Palestine	human CL
***Ltro*MS 049**	2	Jericho, Palestine	human CL
***Ltro*MS 051**	2	Palestine	human CL
***Ltro*MS 054**	2	Jenin District, Palestine	human CL
***Ltro*MS 063**	2	Israel	human CL
***Ltro*MS 065**	2	Palestine	human CL
***Ltro*MS 072**	2	Laikipia, Kenya	human CL
***Ltro*MS 089**	2	northern side of Sea of Galilee, Israel	human CL / 1x hyrax
***Ltro*MS 094**	2	Marrakech area, Morocco	human CL
***Ltro*MS 098**	3	Chichaoua province, Morocco	human CL
***Ltro*MS 099**	4	Chichaoua province, Morocco	human CL

Table 13. Microsatellite profiles of the strains of group "Sanliurfa". n, number of strains sharing this profile. The fragment lengths are given in bp.

microsatellite profile	n	GA1		GA2		GA6		GA9n		LIST7010		LIST7011	
***Ltro*MS 024**	1	66	66	56	56	61	61	112	114	178	178	194	194
***Ltro*MS 026**	1	66	66	56	56	nd	nd	112	114	178	178	194	194
***Ltro*MS 027**	12	66	66	56	56	61	61	112	114	178	178	194	194
***Ltro*MS 028**	9	66	66	56	56	61	61	112	114	178	178	194	194
***Ltro*MS 029**	2	66	66	56	56	61	61	112	114	178	178	194	194

microsatellite profile	LIST7027		LIST7033		LIST7039		LIST7040		4GTG		27GTGn	
***Ltro*MS 024**	169	177	176	176	197	197	nd	nd	62	62	nd	nd
***Ltro*MS 026**	169	177	176	176	197	197	229	247	62	62	nd	nd
***Ltro*MS 027**	169	177	176	176	197	197	229	247	62	62	nd	nd
***Ltro*MS 028**	169	177	176	176	197	197	229	247	62	62	112	118
***Ltro*MS 029**	169	177	176	176	197	197	229	247	62	62	112	112

Each strain of *L. aethiopica* had its unique profile (Table 11). None of the profiles found in *L. aethiopica* was identical to one of those of *L. tropica*.

The descriptive statistics per microsatellite locus is shown in Table 14. The number of different alleles per locus ranged between three (GA6, GA9n) and 20 (LIST7039) with a mean of 9.75. The observed heterozygosity (H_o) was always lower than the expected heterozygosity (H_e). The inbreeding coefficient F_{IS} was positive for all loci, ranging from 0.281 (GA9n) to 1.000 (GA6) with a mean of 0.726.

Table 14. Descriptive statistics per microsatellite locus. A, number of alleles; H_o, observed heterozygosity; H_e, expected heterozygosity; F_{IS}, inbreeding coefficient

locus	A	H_o	H_e	F_{IS}
GA1	7	0.020	0.124	0.840
GA2	10	0.163	0.718	0.773
GA6	3	0.000	0.141	1.000
GA9n	3	0.347	0.482	0.281
LIST7010	11	0.208	0.837	0.752
LIST7011	11	0.015	0.782	0.981
LIST7027	14	0.406	0.817	0.504
LIST7033	11	0.035	0.667	0.948
LIST7039	20	0.131	0.824	0.841
LIST7040	15	0.355	0.703	0.496
4GTG	4	0.015	0.226	0.934
27GTGn	8	0.240	0.727	0.671
all	**9.75**	**0.161**	**0.587**	**0.726**

6.1.4.2 Population structure revealed by Bayesian statistics

All strains were analysed using STRUCTURE software, which uses allele frequencies to analyse the strains and assigns them to different populations. The most probable number of populations was revealed by deltaK calculation. The populations identified by the clustering approach were all re-analysed separately in order to test for hidden sub-structures. This revealed a hierarchical population structure for the strain set studied as shown in Figure 15 and Figure 16. The different hierarchical levels were named “population”, “sub-population”, “cluster” and “group” (Figure 16). The expression “genetic entity” refers to either of the mentioned entities without indicating their hierarchical level.

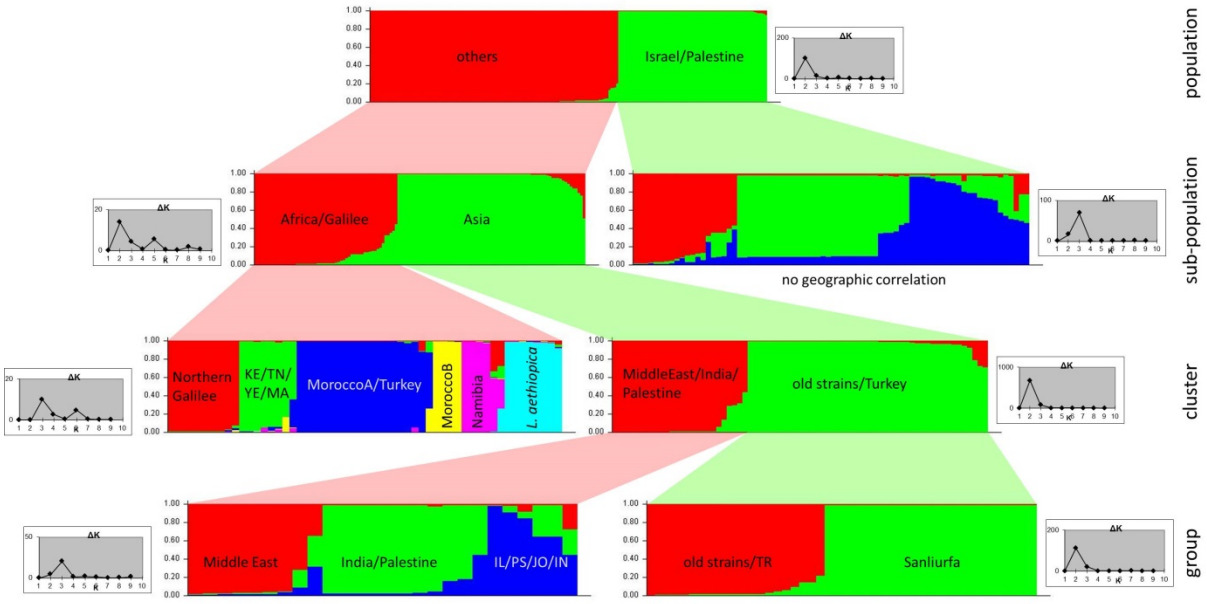

Figure 15. Results of the Bayesian statistics. The upper row shows the structuring of all strains. The rows below show the sub-structures of the found populations. The colours are the default colours of STRUCTURE and do not correspond to the colours in the following figures. The results of the calculation of deltaK for each strain set are presented along with each bar plot.

The bar plots in Figure 15 show the results of the Bayesian statistics. Each bar represents one strain. The clustering approach identified two main populations in the total data set ($\Delta K = 2$) and separated 75 strains from Israel and Palestine and one strain from the Sinai Peninsula, Egypt, from all other strains investigated. When this population, called “Israel/Palestine”, was re-analysed separately, it was split into three sub-populations ($\Delta K = 3$) without any further geographical correlation.

The microsatellite profiles of the strains which formed the population “others” were also re-analysed separately and sub-divided into two sub-populations ($\Delta K = 2$). One of them contained mainly strains from Africa, but also included a group of Israeli strains that were all isolated at the northern side of the Sea of Galilee. The second sub-population comprised mainly strains from Asian foci, but also two Kenyan strains isolated in 1981. Thus, those sub-populations were named “Africa/Galilee” and “Asia”.

Further analysis of the strains of “Africa/Galilee” assigned them to six clusters. Here, the deltaK calculation showed two peaks, one at three and the other at six clusters. The distribution of the strains into three clusters was not consistent throughout the ten iterations. The assignment of the strains to six clusters was statistically better supported, being consistent throughout nine of the ten iterations. Thus, the six genetic entities were accepted as the clusters “*L. aethiopica*”, “Northern Galilee”, “MoroccoA/Turkey”, “MoroccoB”, “Namibia” and “KE/TN/YE/MA” with their names referring to their strains´ origins. The cluster “*L. aethiopica*” contained the eight strains of *L. aethiopica* included in this study. All ten strains of *L. tropica* from the focus at the northern side of the Sea of Galilee in Israel formed the cluster “Northern Galilee”. Twenty-one of the 22 Moroccan strains were assigned to two clusters. The cluster “MoroccoA/Turkey” comprised 17 strains from Morocco and two from Turkey and the remaining four strains from Morocco presented cluster “MoroccoB”. The cluster “Namibia” contained four Namibian strains and one from Kenya. The cluster “KE/TN/YE/MA” contained four strains from Kenya, two from Tunisia and one each from Yemen and Morocco. One strain, MHOM/KE/1984/NLB248, had shared membership of four clusters, “Namibia” (membership coefficient 0.306), “*L. aethiopica*” (0.291), “Northern Galilee” (0.253) and “KE/TN/YE/MA” (0.142). The assignment of the strains into three clusters ($\Delta K = 3$) was largely consistent with the separation into six clusters, with the difference that these three clusters grouped with each other in varying combinations. For example, the cluster “Northern Galilee” grouped together with “MoroccoB” in five iterations, with “*L. aethiopica*” in one iteration and without another cluster in four iterations; cluster “*L. aethiopica*” grouped together with strains from Kenya, Namibia and Tunisia in eight iterations, with “Northern Galilee” in one iteration and alone in another; only the cluster “MoroccoA/Turkey” was consistent throughout all iterations, not closely related to another cluster.

The sub-population “Asia” was divided into two clusters ($\Delta K = 2$), “Middle East/India/Palestine” and “old strains/Turkey”. Re-analysis showed that the first one consisted of three groups, “Middle East”, “India/Palestine” and “IL/PS/JO/IN”. The group “Middle East” comprised nine strains from five countries, Syria, Iran, Afghanistan, Turkey and the former Soviet Union. Eight Indian and three Palestinian strains belonged to the group “India/Palestine”. The group “IL/PS/JO/IN” contained one Israeli, one Palestinian, one Jordanian and two Indian strains. Two strains of the sub-population “Asia”, MHOM/PS/2002/20JnYM3 and MHOM/PS/2002/52JnYM18, had similar membership coefficients for all three groups (Table 15).

Table 15. Membership coefficients of strains MHOM/PS/2002/20JnYM3 and MHOM/PS/2002/52JnYM18. The values are the mean calculated from 10 iterations.

	Middle East	India/Palestine	IL/PS/JO/IN
MHOM/PS/2002/20JnYM3	0.362	0.318	0.321
MHOM/PS/2002/52JnYM18	0.300	0.294	0.406

The cluster "old strains/Turkey" was sub-structured into one group containing 25 Turkish strains collected at a local outbreak of CL near Sanliurfa in 1995, "Sanliurfa", and another one containing two and three Turkish strains isolated in 1995 and 1998, respectively, and 15 strains of various other origins which were isolated in or before 1991, "old strains/TR".

Figure 16 summarizes the results of the Bayesian statistics and introduces a colour code for all genetic entities. These colours will be used in the figures of all following analyses.

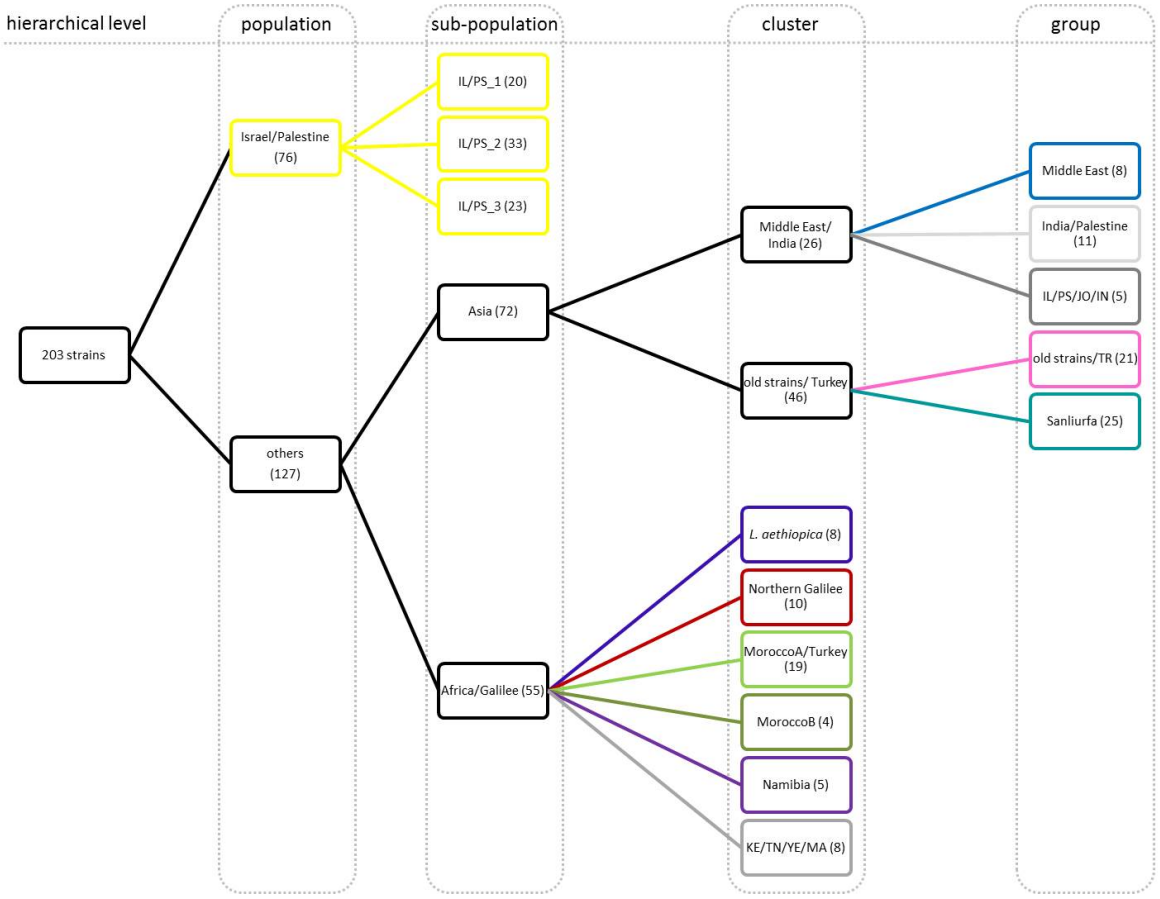

Figure 16. Hierarchical population structure as inferred by Bayesian statistics. All strains were assigned to populations by STRUCTURE software and the resulting populations were re-analysed to infer sub-structures. The names of the various hierarchical levels are indicated in the column headings. The numbers of strains are given in parentheses after the respective names; strains with shared membership between two or more clusters or groups are not counted. The colour code is applied to the figures resulting from all following analyses.

6.1.4.3 Genetic distance analyses

The NJ tree shows the phylogenetic relationships among the strains of *L. tropica* and *L. aethiopica* based on Chord distances (Figure 17). The genetic entities of the Bayesian clustering approach are coloured according to Figure 16.

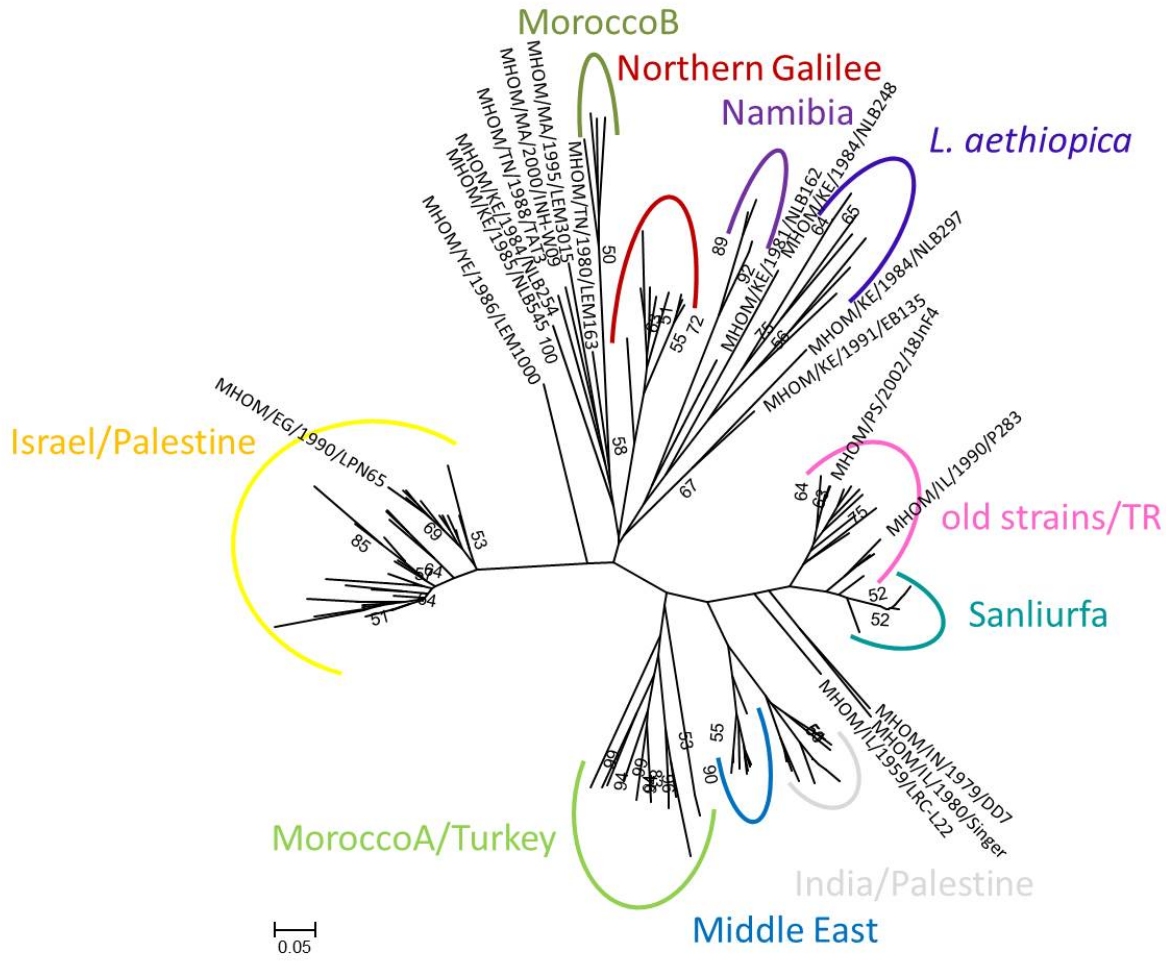

Figure 17. Neighbour Joining tree of all strains. The calculation of the phylogenetic relationship between the strains is based on Chord distance estimates. Bootstrap values >50 are indicated at the nodes. Bayesian results are indicated by colours according to the colour code in Figure 16. Only those strains are labelled with their WHO codes which either deviate from the name of their population, sub-population, cluster or group, respectively, or do not form a monophyletic group in the tree. Detailed information about all strains is summarized in Table 11.

The population "Israel/Palestine" identified by the Bayesian statistics was clearly separated from all other strains. It included the same 26 Israeli, 49 Palestinian and a single strain from Egypt as in the Bayesian statistics.

A second monophyletic group contained the African strains of *L. tropica* and *L. aethiopica*, except for those of the cluster "MoroccoA/Turkey", as well as the strains from the focus north of the Sea of Galilee. However, the branches in the sub-population "Africa/Galilee" were long, indicating a long time span since the separation of these strains. The clusters "Northern Galilee" and "*L. aethiopica*" were confirmed by this analysis. The cluster "Namibia" was also largely confirmed though excluding

the Kenyan strain MHOM/KE/1981/NLB162. Interestingly, strains of the cluster “MoroccoA/Turkey” were not included in this monophyletic group but rather formed a separate small one. One strain, MHOM/MA/1995/LEM3015, previously assigned to the cluster “MoroccoA/Turkey”, was found at the bottom of the branch of “MoroccoB” here. Close relationships were confirmed for five of the eight strains forming the cluster KE/TN/YE/MA in the NJ tree. They were put near the node of the branch of “MoroccoB”. The strain from Yemen, MHOM/YE/1986/LEM1000, was also found in proximity to these strains, although clearly separated from them. Another four Kenyan strains, two from the cluster KE/TN/YE/MA, one from the cluster “Namibia” and the strain with mixed memberships, MHOM/KE/1984/NLB248, intermingled between the African strains.

The sub-population “Asia” identified by the Bayesian clustering could not be confirmed with the genetic distance analysis. The strains of this sub-population were very paraphyletic in the NJ tree. The strains of “Middle East” and “India/Palestine” each formed a monophyletic group with their branches originating from an identical node. The cluster “old strains/Turkey” was largely confirmed by the NJ tree. The old strains formed a very compact monophyletic group which had short branches supporting the close relationship between these strains. The four Turkish strains were not part of this monophyletic group but rather clustered together with the Turkish strains from the local outbreak of CL in Sanliurfa, which in turn formed a clear monophyletic group in the NJ tree.

The Neighbour network presents the phylogenetic relationships among the strains of *L. tropica* and *L. aethiopica* and, additionally, shows reticulation events indicated by cross connections (Figure 18). The results of the Bayesian clustering approach are indicated by the colours as in Figure 16.

The population “Israel/Palestine” is strongly confirmed in the network. These strains form a clear monophyletic group.

All the other strains were largely paraphyletic and showed many cross connections between them. However, some of the entities previously described were present here as well. Like in the Bayesian statistics and the NJ tree, the “Northern Galilee" strains were related to the African strains. The clusters “Northern Galilee”, “Namibia”, “*L. aethiopica*” and “MoroccoB” each formed a monophyletic group in the network. As in the NJ tree, those strains belonging to the cluster “KE/TN/YE/MA” intermingled between the other strains without showing a closer relationship to any of them. The strain from Yemen, MHOM/YE/1986/LEM1000, was found isolated from all other strains supporting its remote phylogenetic position in the NJ tree. Interestingly, the cluster “MoroccoA” was closer to the Asian strains like in the NJ tree and not among the African strains like in the Bayesian results. Some of the strains previously assigned to the cluster “MoroccoA/Turkey” were not members in this

monophyletic group in the network. The two Turkish strains were located on the opposite side of the main branch and three of the Moroccan strains from this cluster were found at separate branches.

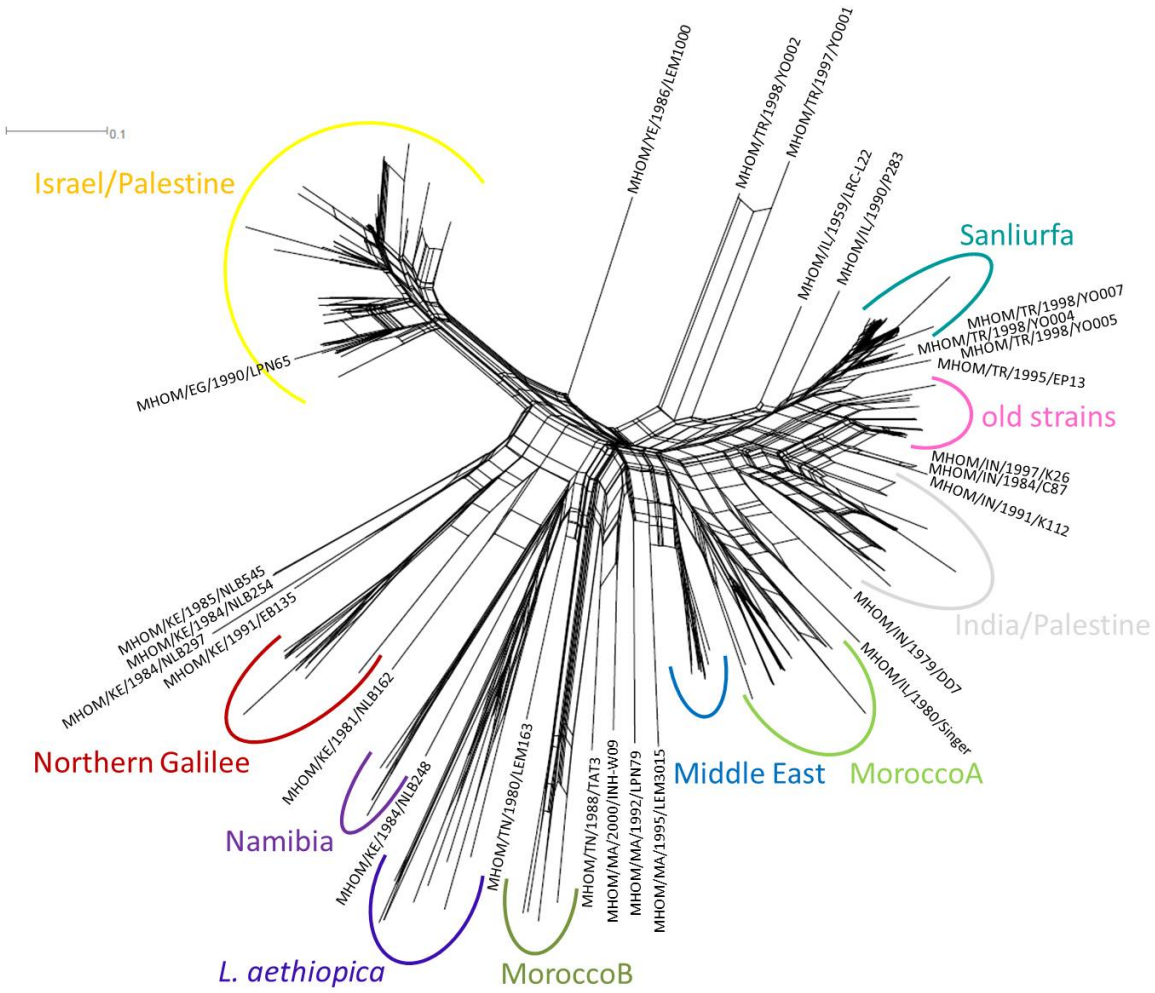

Figure 18. Neighbour network of all strains. Cross connections indicate probable reticulation events like hybridisation, recombination, or horizontal gene transfer between the strains. Bayesian results are indicated by colours according to the colour code in Figure 16. Only those strains are labelled with their WHO codes which either deviate from the respective name of their population, sub-population, cluster or group, respectively, or intermingle between the distinguished bundles of strains in the network. Detailed information about all strains is summarized in Table 11.

The Asian strains of *L. tropica,* previously assigned to the groups “Middle East” and “Sanliurfa”, represented monophyletic groups in the network. The remaining two groups, “old strains” and “India/Palestine” were rather paraphyletic with many cross-connections between their strains´ branches. The strains which were members of the group “IL/PS/JO/IN” according to Bayesian statistics intermingled between the other strains and did not form a distinct genetic group in the network.

6.1.4.4 Genetic relationship between the strains

For the FCA all strains had to be assigned to distinct genetic entities prior to the analysis. This pre-assignment was based on the results of the Bayesian clustering. The three sub-populations of

"Israel/Palestine" were not considered as distinct entities since their division showed no geographical correlation and was not supported by the other analyses. All strains with large proportions of shared membership between two or more genetic entities in the Bayesian statistics were excluded from the FCA since they could not be assigned unambiguously to one of the genetic entities.

Figure 19 and Figure 20 show the results of the FCA. The colours correspond to those of the Bayesian statistics in Figure 16.

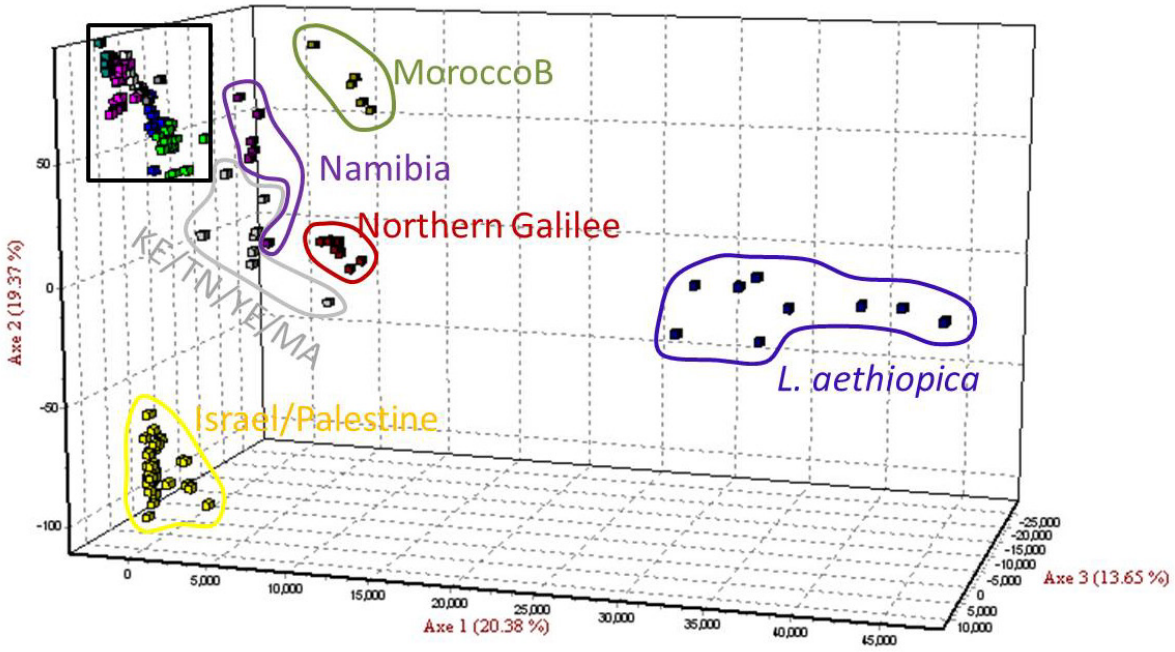

Figure 19. Phylogenetic relationships between all strains of *L. tropica* and *L. aethiopica* included in this study calculated by Factorial Correspondence Analysis (FCA). Bayesian results are indicated by colours according to the colour code in Figure 16. All strains were assigned to the genetic entities proposed by Bayesian clustering before applying FCA. The section indicated by the rectangle is magnified in Figure 20.

The strains of *L. aethiopica* were clearly remote from the strains of *L. tropica* (Figure 19). These strains did not cluster together with the African strains of *L. tropica* as in the previous analyses. The population "Israel/Palestine" presented as distinct entity in this analysis and had the largest genetic distance from all other strains of *L. tropica*. The African strains were quite dispersed, but grouped according to the results of the other approaches. The Kenyan strain MHOM/KE/1981/NLB162 belonging to the cluster "Namibia" according to Bayesian statistics which was not supported by the NJ tree and the network, was closer to the other Kenyan strains in the FCA. The cluster "MoroccoA/Turkey" was located between the two sub-populations "Africa/Galilee" and "Asia".

The Asian strains and those found in the cluster "MoroccoA/Turkey" were genetically closely related. To get a more detailed view on those strains, the rectangle indicated in Figure 19 was magnified and is shown in Figure 20. Although the strains of each group found by the Bayesian statistics tended to cluster together, there was no clear distinction between these groups.

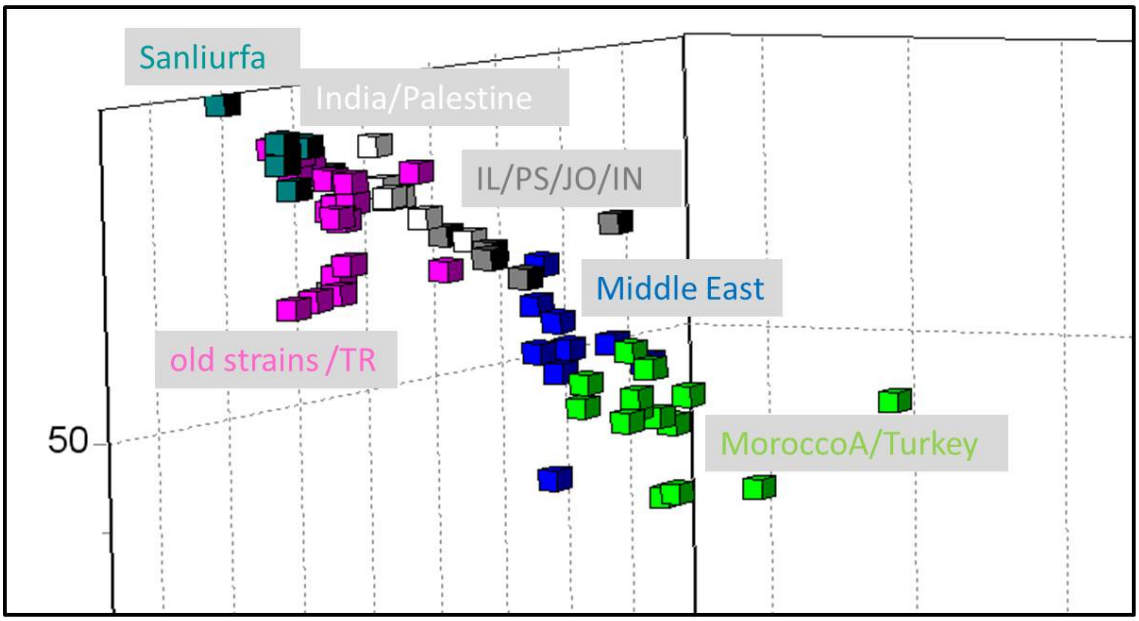

Figure 20. Detailed view on the results of Factorial Correspondence Analysis for the strains in the rectangle shown in Figure 19 which has been magnified for a better resolution of the strains. Bayesian results are indicated by colours according to the colour code in Figure 16.

6.1.4.5 Descriptive statistics by populations

The descriptive statistics for the entities found by the Bayesian statistics are summarized in Table 16. The mean number of different alleles per population (A) ranged between 1.33 ("Sanliurfa") and 3.42 ("MoroccoA/Turkey" and "KE/TN/YE/MA") with a mean of 2.31. In eight of 12 cases the observed heterozygosity (H_o) was lower than the expected heterozygosity (H_e) resulting in a positive F_{IS} value.

Table 16. Descriptive statistics of the genetic entities found by the Bayesian statistics. A, mean number of alleles; H_o, observed heterozygosity; H_e, expected heterozygosity; F_{IS}, inbreeding coefficient

	A	H_o	H_e	F_{IS}
Israel/Palestine	2.75	0.035	0.143	0.755
Middle East	1.67	0.333	0.268	-0.266
India/Palestine	1.75	0.402	0.269	-0.532
IL/PS/JO/IN	2.17	0.343	0.379	0.103
old strains/TR	2.50	0.327	0.306	-0.070
Sanliurfa	1.33	0.318	0.170	-0.921
L. aethiopica	3.25	0.266	0.507	0.492
Northern Galilee	1.83	0.025	0.226	0.895
MoroccoA/Turkey	3.42	0.143	0.370	0.620
MoroccoB	1.83	0.118	0.278	0.624
Namibia	1.75	0.021	0.302	0.942
KE/TN/YE/MA	3.42	0.126	0.544	0.780
mean	**2.31**	**0.205**	**0.314**	**0.354**

6.1.4.6 Genetic distances between the populations

The F_{ST} values indicate the genetic distances by pairwise comparisons between the genetic entities identified by the Bayesian statistics (Table 17). The greater the value, the larger is the genetic distance between the two populations. The largest distance was found between the populations "Israel/Palestine" and "MoroccoB" (F_{ST}=0.820), the smallest one was found between "old strains/TR" and "IL/PS/JO/IN" (F_{ST}=0.180). Noteworthy, the strains of population "Israel/Palestine" showed the largest genetic distances to all other populations, supporting the remote position of those strains in the NJ tree (Figure 17) and the network (Figure 18). Some of the F_{ST} values were not significant which was probably due to the small sample sizes of these populations.

Table 17. Genetic distances between the genetic entities calculated by *F* statistics. Pairwise comparisons give the F_{ST} values (mean fixation indices) that are an indicator for the genetic distances between two genetic entities. The values are categorized into little (<0.05), moderate (0.05-0.15), great (0.15-0.25), and very great (>0.25) differentiation. *, not significant

	Israel/ Palestine	Middle East	India/ Palestine	IL/PS/JO/IN	old strains/ TR	Sanliurfa	*L. aethiopica*	Northern Galilee	MoroccoA/ Turkey	MoroccoB	Namibia	KE/TN/ YE/MA
Israel/ Palestine	0.000											
Middle East	0.750	0.000										
India/ Palestine	0.761	0.385	0.000									
IL/PS/JO/IN	0.748	0.354*	0.227	0.000								
old strains/ TR	0.734	0.491	0.350	0.180	0.000							
Sanliurfa	0.793	0.673	0.619	0.533	0.291	0.000						
L. aethiopica	0.757	0.559	0.579	0.474*	0.573	0.702	0.000					
Northern Galilee	0.759	0.664	0.692	0.621	0.651	0.777	0.532	0.000				
MoroccoA/ Turkey	0.687	0.353	0.449	0.324	0.414	0.548	0.496	0.568	0.000			
MoroccoB	0.820	0.662*	0.660	0.589*	0.643	0.777	0.532*	0.679*	0.580	0.000		
Namibia	0.778	0.628*	0.631	0.541*	0.588	0.705	0.481	0.681	0.538	0.627*	0.000	
KE/TN/ YE/MA	0.610	0.361	0.398	0.320*	0.437	0.590	0.305	0.402	0.381	0.395*	0.386	0.000

6.3 A Functional Cloning approach for the identification of genes conferring high mutation tolerance in *L. tropica*

The Functional Cloning approach as described by Clos *et al*. was customized for the identification of factors leading to a low MNNG tolerance, hypothetically due to an impaired MMR system (Figure 11). This approach included the identification of a donor and an acceptor strain (exhibiting low and high MNNG tolerance, respectively), the creation of a cosmid library bearing the gDNA of the donor strain, the transfection of the acceptor strain with this library, the screening for complemented cells followed by cosmid rescue and, finally, the identification of the genes in the cosmids´ inserts conferring the low mutation tolerance phenotype.

6.3.1 Selection of donor and acceptor strains differing in MNNG tolerance

A total of 12 strains of *L. tropica* isolated in Israel and Palestine (Table 7) between 2001 and 2008 were previously tested for their MNNG sensitivity using an AlamarBlue cell viability test by Boysen *et al*. [106]. The viability of the different strains in the presence of different concentrations of MNNG was monitored after an incubation time of 48 h by measuring the fluorescence intensity.

Based on the results obtained, strain MHOM/PS/2001/ISL593 (or LRC-L830) was selected as donor and MHOM/PS/2008/335JnM59 (or LRC-L1322) as acceptor strain (Figure 21). The acceptor strain LRC-L1322 was cloned to get a culture originating from a single genetic clone. This clone was called LRC-L1322/6 in all further experiments.

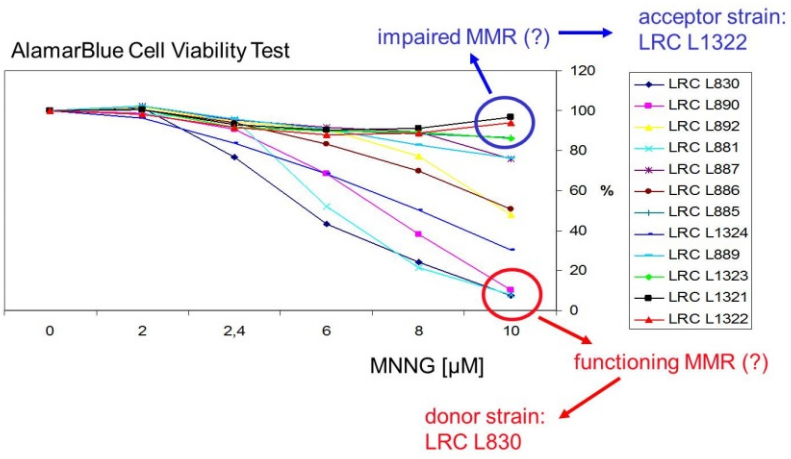

Figure 21. MNNG susceptibility of 12 strains of *L. tropica*. The viability of the strains was monitored as the fluorescence intensity. 106 parasites / ml were incubated in medium supplemented with different concentrations of MNNG in a total volume of 25 µl for 48 h. Twenty-five µl AlamarBlue were added and incubated for an additional 24 h. The fluorescence emission was measured at 590 nm after an excitation pulse at 540 nm. All experiments were conducted in triplicates. The strains with the lowest and highest MNNG sensitivity were selected as donor and acceptor strain, respectively [106].

6.3.2 Creation of a cosmid library

A cosmid library containing the gDNA of the donor strain LRC-L830 was created by Boysen *et al*. as described in [95, 107]. Shortly, high molecular weight DNA was isolated from ~4x10^9 cells of the donor and partially digested using the restriction enzyme Sau3AI. The cosmid pcosTL, which is appropriate for the propagation in both *Leishmania* (G418 selection) and *E. coli* (ampicillin selection) [107], was digested using BamHI and SmaI, resulting in two vector arms with Sau3AI compatible sticky ends at the BamHI restriction site (Figure 22). The isolated genomic DNA was ligated into these vector arms and the resulting cosmid packed into phage particles. The phages selectively took up those linear fragments which had a cos site on each end and were between 38 and 52 kb in size (insert lengths: 29.5-43.5 bp). During the injection into *E. coli* the DNA was circularized, forming circular cosmids containing different sequences of the genomic *Leishmania* DNA. These cosmids were then isolated from *E. coli* and transferred into the *Leishmania* acceptor strain LRC-L1322/6 for establishing a representative cosmid library in this strain. The transfectants were named LRC-L1322/6-pcosTL-830 and cloned by limiting dilution to get pure cultures each originating from a single genetic clone.

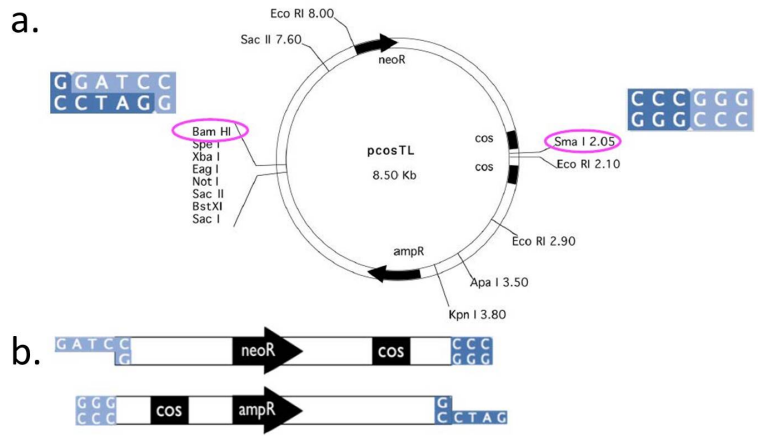

Figure 22. The shuttle vector pcosTL showing the restriction sites (a). Digestion using BamHI and SmaI resulted in two vector arms with Sau3AI compatible sticky ends (b). neoR; neomycin resistance gene, ampR; ampicillin resistance gene

6.3.3 Screening for clones with a complemented donor phenotype by MNNG susceptibility tests

Prior to screening the transfected *Leishmania* parasites for their MNNG tolerance, the optimal amount of inoculum and MNNG concentration had to be determined by testing different cell numbers and different MNNG concentrations in the culture medium, each in four independent experiments. The survival rates of the cultures were determined as the percentage of counted cells in the culture containing MNNG compared to those in the culture without MNNG after 48 h. The optimal conditions were achieved when culture medium containing 2 µM MNNG was inoculated with 5x10^5 cells/ml (Table 18).

Table 18. Survival rates of the donor (LRC-L830) and acceptor strain (LRC-L1322/6) upon addition of different concentrations of MNNG to the culture medium. The data are given as percentage of cell number that were grown with MNNG in comparison to those grown without MNNG (=100%) after an incubation time of 48 h. All data are the mean of four independent experiments.

		MNNG concentration						MNNG concentration				
inoculum		**1 µM**	**2 µM**	**3 µM**	**4 µM**	**5 µM**		**1 µM**	**2 µM**	**3 µM**	**4 µM**	**5 µM**
5x10^5 cells/ml	**LRC-L830**	68.7	43.8	41.3	37.2	45.4	**LRC-L1322/6**	112.2	98.9	82.3	85	78.6
1x10^6 cells/ml		61.8	51.1	41.5	29.7	32.8		119.6	92.9	87.2	75	75.8
2x10^6 cells/ml		86.3	51.4	42.4	45.5	43.6		106.9	92.6	85.9	71.5	63.2
3x10^6 cells/ml		114.9	73.5	61.1	67.7	43.6		99.6	91	82.2	75.8	65.2

The required number of clones to cover the whole sequence of the *Leishmania* genome was calculated using the Poisson distribution with the formula

$$N = \ln(1-P) \Big/ \ln\left[1-\left(\frac{F}{G}\right)\right]$$

where N is the number of independent clones, P is the probability of coverage of the genome, F is the size of the inserts and G is the size of the *Leishmania* genome. With a minimum insert size of 29,500 bp, a probability of 99% to cover the whole genome and a genome size of 3.2×10^7 bp, the required number of clones was 4,993. However, the redundancy is reached with ~2,000 clones (J. Clos, personal communication).

In total, 2,603 clones were obtained and, subsequently, tested for their MNNG sensitivity in cultures with and without MNNG. There was no clear distinction between the two expected phenotypes. Some of the clones showed a reduced tolerance against MNNG, some showed a high tolerance and many others intermingled between the two phenotypes. Since the MNNG tolerance of the acceptor strain itself varied considerably between different experiments (minimum = 53.3%, maximum = 134.6%, Table 19 and Table 20), the threshold of a clone being determined as positive, had to be related to the acceptor´s phenotype in each experiment. The threshold was set to a reduced MNNG tolerance of at least 50% in comparison to that of the acceptor strain. Twenty-four out of the 2603 analysed clones (~0.92%) showed the restored phenotype, i. e. low MNNG tolerance, during the first screening (Table 19). These clones were then re-tested. Noteworthy, the survival rates of the clones varied considerably between the different, independently conducted experiments like it was the case in the acceptor strain (Table 19 and Table 20). The statistical significance of the difference in MNNG susceptibility compared to that of the acceptor strain was confirmed for seven clones (# 285, 260, 261, 382, 797, 2612, 2874) using an unpaired *t* test. The resulting *p* values are also included inTable 19.

Table 19. Survival rates of the clones with a reduced mutation tolerance. Each clone was tested for its MNNG sensitivity in several independent experiments. All values are given in % and were truncated to one position after the decimal point. Those experiments, in which the clone´s survival rate was ≤50% compared to the survival rate of the acceptor strain, are indicated in green. The *p* value was calculated with an unpaired *t* test and indicated the degree of significance of its difference in comparison to the acceptor strain. *, difference statistically not significant.

	survival rates of clone cultures with vs. without MNNG														
	1		2		3		4		5		6		7		*p* value
clone #	clone	acceptor	clone	acceptor	clone	acceptor	clone	acceptor	clone	acceptor	clone	acceptor	clone	acceptor	
20	41.2	96.6	103.3	90.8	80.4	109.2	63.5	128.3	-	-	-	-	-	-	0.075*
125	12.7	106.8	81.7	75.3	119.5	90.8	-	-	-	-	-	-	-	-	0.6705*
169	25.2	106.8	96.4	75.3	117.1	90.8	-	-	-	-	-	-	-	-	0.7811*
285	3.9	99.5	6.0	79.1	19.3	53.3	49.0	74.7	29.0	71.8	47.9	83.3	38.2	109.2	0.0001
	44.0	128.3	53.3	90.4	46.2	85.4	53.4	101.6	-	-	-	-	-	-	
246	26.0	99.5	147.2	79.1	58.3	53.3	55.6	74.7	-	-	-	-	-	-	0.8668*
260	13.6	99.5	46.4	79.1	42.7	53.3	59.9	74.7	-	-	-	-	-	-	0.0381
261	49.1	99.5	35.5	79.1	44.8	53.3	53.3	74.7	-	-	-	-	-	-	0.023
278	45.5	99.5	70.8	79.1	45.3	53.3	0.7	74.7	-	-	-	-	-	-	0.0833*
382	23.7	71.8	44.6	83.3	31.0	109.2	40.2	128.3	60.7	90.4	72.9	85.4	56.0	101.6	0.0008
704	51.5	126.9	95.9	103.8	81.9	98.1	82.2	90.5	114.2	96.5	-	-	-	-	0.1733*
732	54.5	126.9	115.3	103.8	90.3	98.1	99.5	90.5	95.6	96.5	-	-	-	-	0.3365*
749	58.2	126.9	99.6	103.8	83.5	98.1	86.0	90.5	124.5	96.5	-	-	-	-	0.3373*
797	32.7	71.8	55.1	83.3	97.3	109.2	61.4	128.3	44.7	90.4	64.9	85.4	85.0	101.6	0.0225
979	30.6	82.8	90.8	96.4	87.6	101.0	106.4	97.2	104.6	117.6	-	-	-	-	0.3445*
1049	32.5	82.8	98.8	96.4	94.2	101.0	101.8	97.2	90.3	117.6	-	-	-	-	0.3029*
1338	31.4	83.2	121.1	102.4	98.3	99.8	-	-	-	-	-	-	-	-	0.6973*
1439	28.6	134.6	72.5	94.8	52.2	134.5	57.0	108.5	71.9	66.9	-	-	-	-	0.057*
1451	48.5	102.9	84.7	114.0	-	-	-	-	-	-	-	-	-	-	0.1576*
1728	36.3	129.9	67.5	99.4	57.9	77.6	121.8	72.3	-	-	-	-	-	-	0.3267*
1793	43.8	129.9	61.1	99.4	68.8	77.6	80.0	72.3	-	-	-	-	-	-	0.0836*
1932	31.3	67.8	106.2	77.6	140.4	72.3	-	-	-	-	-	-	-	-	0.5686*
2612	30.3	73.0	25.3	83.3	58.4	109.2	23.8	128.3	54.1	71.8	31.3	94.0	48.8	104.5	0.0002
2717	43.8	107.4	164.6	94.0	119.6	104.5	-	-	-	-	-	-	-	-	0.8457*
2874	30.9	75.3	33.3	83.3	52.0	109.2	36.7	128.3	41.0	71.8	28.4	87.3	40.9	101.6	0.0001

Table 20. Descriptive statistics of the survival rates of the positive clones and the acceptor strain. Δsr, difference between the highest and the lowest survival rate of the independently conducted experiments

clone #	LRC-L1322/6	285	260	261	382	797	2612	2874
Δsr	81.3	49.5	46.3	17.9	49.2	64.6	34.7	23.6
standard deviation	18.40	17.43	16.87	6.62	15.98	20.63	13.40	7.36

The difference in MNNG susceptibility between clone and acceptor strain was categorized as "extremely statistically significant" for the clones 285, 382, 2612 and 2874 by the *t* test, while the differences for clones 260, 261 and 797 were considered to be only "statistically significant". Clones with the highest statistical significance were selected for the downstream analyses.

6.3.4 Characterization of the clones with a reduced MNNG tolerance

The cosmids of the clones 285, 382, 797, 2612 and 2874, exhibiting the restored phenotype of low MNNG tolerance, were isolated and amplified in *E. coli* strain XL1blue. The transformation of the *E. coli* cells with the cosmid DNA of clone 2612 was not successful; it was thus excluded from the downstream analyses. The restricted cosmid DNA isolated from several *E. coli* colonies was separated in a gel electrophoresis (Figure 23). The restriction pattern was identical throughout all *E. coli* clones bearing either the cosmid pcosTL-830/285 or pcosTL-830/382. Besides the fragments originating from the cosmid backbone at 815 and 7422 bp, restriction of the cosmids isolated from the clones XL1-pcosTL-830/285 resulted in fragments with the lengths of ~1200, ~3000, ~4400, four fragments >8000 bp and probably uncut cosmids at >40000 bp which is the upper limit of the 1kb DNA Extension ladder. Restriction of the cosmids isolated from the clones XL1-pcosTL-830/382 resulted in fragments of ~1200, ~1500, ~1600, ~4400, three fragments >8000 bp and the uncut cosmid at >40000 bp. Two different fragment patterns were identified for each of the cosmids pcosTL-830/797 and pcosTL-830/2874, named a and b. The fragments of clone XL1-pcosTL-830/797a were ~1500, ~3000, ~3400 and >10000 bp, those of XL1-pcosTL-830/797b were ~700, ~2200, ~3200 and >40000 bp. The restriction of clone XL1-pcosTL-830/2874a yielded fragments at ~3100, ~3600, ~6000 and two >10000 bp; that of clone XL1-pcosTL-830/2874b yielded fragments at ~1000, ~2100, ~4500, ~8000 and >40000 bp. Noteworthy, the lengths of the fragments >8000 bp were hard to determine precisely due to the logarithmic separation of the DNA fragments. These fragments probably contained uncut or only partially cut cosmid DNA.

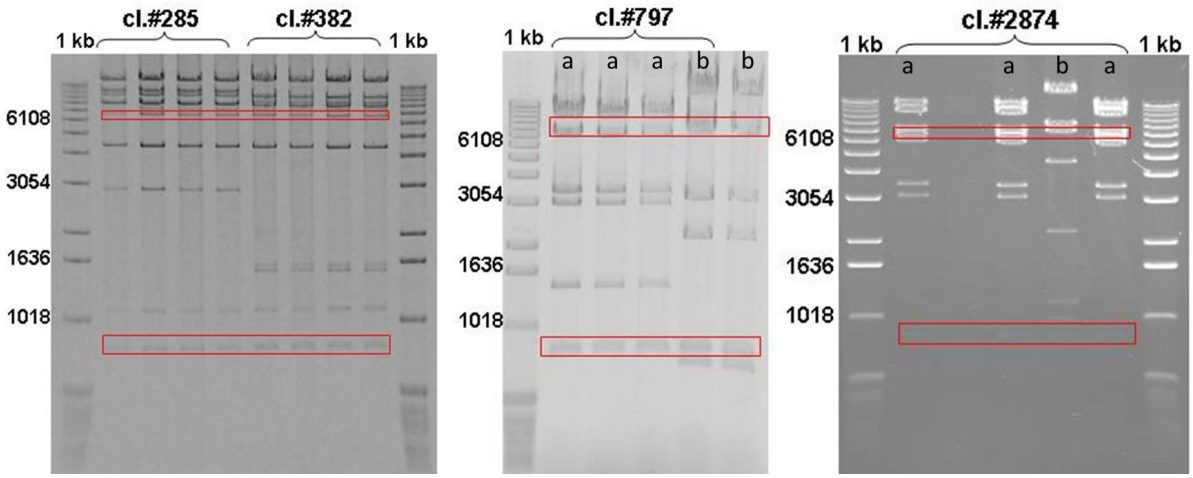

Figure 23. Restriction patterns of cosmid DNA isolated from different *E. coli* colonies. The numbers indicate the *Leishmania* clones with the restored donor phenotype of low MNNG tolerance. The complete restriction reaction of 15 µl was loaded on the gel. The lanes show cosmids of different *E. coli* clones. The fragments highlighted with a red rectangle are those originating from the vector backbone at 815 and 7422 bp.

The cells of LRC-L1322/6-pcosTL-830/797 and 2874 were sub-cloned to select for each of the detected cosmid sub-populations seen in the restriction analysis (Figure 23). In total, 94 sub-clones for the first and 56 sub-clones for the latter one were analysed for their MNNG tolerance. The survival rates of the sub-clones ranged from ~45 to >125% for clone # 797 and from >30 to >125% for clone # 2874 (Figure 24). Two and six sub-clones of clone # 797 and # 2874, respectively, showed a decreased MNNG tolerance compared to the acceptor strain LRC-L1322/6 of more than 50%. When re-testing these sub-clones, the low MNNG tolerance of <50% compared to the acceptor strain could not be confirmed for any of them. The difference in MNNG tolerance compared to that of the acceptor strain was statistically significant in only two sub-clones of clone 2874, 2874/10 and 2874/20 (Table 21).

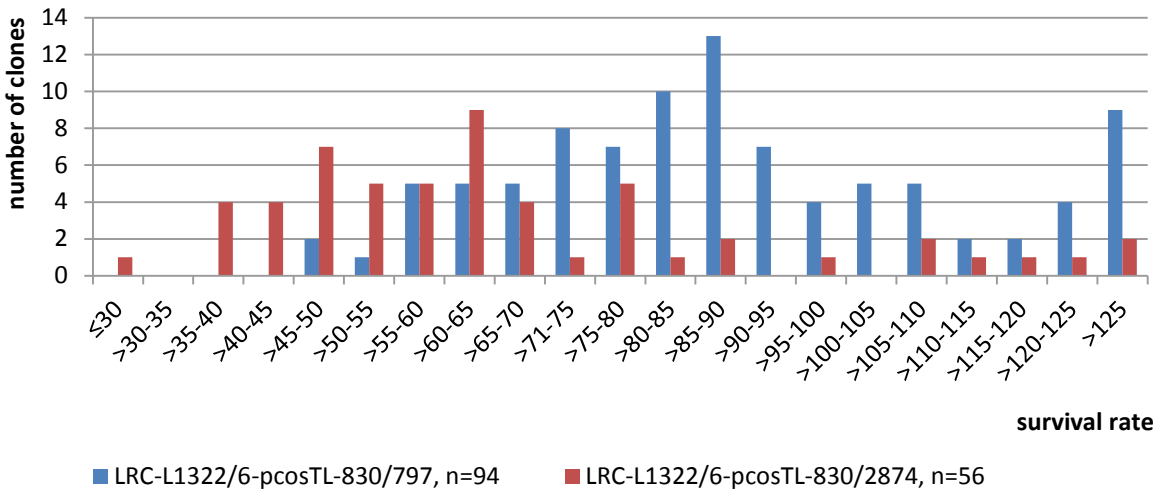

Figure 24. Distribution of the sub-clones of LRC-L1322/6-pcosTL-830/797 and LRC-L1322/6-pcosTL-830/2874 showing different survival rates.

Table 21. MNNG sensitivity test of the sub-clones of LRC-L1322/6-pcosTL-830/797 and LRC-L1322/6-pcosTL-830/2874. All values are given in % and were truncated to one position after the decimal point. Those experiments, in which the clone´s survival rate was ≤50% compared to the survival rate of the acceptor strain, are indicated in green. The *p* value was calculated with an unpaired *t* test and indicated the degree of significance of its difference in comparison to the acceptor strain. *, difference statistically not significant.

	survival rates of clone cultures with vs. without MNNG								
	1		2		3		4		*p* value
clone #	clone	acceptor	clone	acceptor	clone	acceptor	clone	acceptor	
797/77	45.7	95.9	123.7	67.3					0.9473*
797/87	47.5	95.9	102.5	67.3					0.8511*
2874/10	29.6	102.3	62.7	68.4	40.6	77.4	45.5	67.9	0.0176
2874/12	40.8	102.3	85.5	68.4	72.0	77.4			0.3758*
2874/17	39.6	102.3	55.4	68.4	114.0	77.4	62.9	67.9	0.5627*
2874/20	36.7	102.3	52.3	68.4	60.2	77.4	61.8	67.9	0.038
2874/44	38.6	102.3	81.4	68.4	111.1	77.4			0.8202*
2874/45	38.1	102.3	57.8	68.4	63.7	77.4	69.9	67.9	0.0876*

The cosmids pcosTL-830/285 and pcosTL-830/382 were partially sequenced using primers complementary to sequences on each side of the ligation site (CH2 and MH13rev). The resulting sequences of ~500-1000 bp were aligned to the complete nucleotide collection at NCBI; the hits of the BLAST searches are summarized in Table 22. For both cosmids three different alignment sequences were retrieved from the database, one for each of the three *Leishmania* species *L. major*, *L. infantum* and *L. donovani*, where *L. major* showed the highest sequence similarity (89-95% identity). The two insert sequences of the two clones #285 and #382 were in close proximity to each other (<1,300 bp separating them) (Figure 25). Unexpectedly, the insert lengths, calculated as the number of nucleotides from the BLAST hit of the sequence resulting from primer CH2 to the BLAST hit of the sequence resulting from primer MH13rev, ranged between 1125 and 1469 bp. This contradicted not only the theoretical insert length of between 29,500 and 43,500 bp, but also the fragments in the gel electrophoresis (Figure 23), which summed up to ~40,600 bp (pcosTL-830/285) and ~32,700 bp (pcosTL-830/382) when not taking into account those fragments larger than 40,000, which are likely to represent the uncut vector.

Table 22. Results of the BLAST search of the partially sequenced inserts of the cosmids pcosTL-830/285 and pcosTL-830/382. The complete genomes of the indicated strains were used as reference sequences.

clone #	reference sequence	GenBank accession	BLAST match	insert length	identities		gaps	
					CH2	MH13rev	CH2	MH13rev
285	*Leishmania major* Friedlin	FR796431.1	nt 1581291-1582595	1304 bp	940/1043(90%)	630/665(95%)	46/1043(4%)	5/665(0%)
	Leishmania infantum JPCM5	FR796467.1	nt 1568651-1570104	1453 bp	398/463(86%)	584/659(89%)	36/463(7%)	14/659(2%)
	Leishmania donovani BPK282A1	FR799622.2	nt 1593389-1594858	1469 bp	399/462(86%)	582/659(88%)	34/462(7%)	14/659(2%)
382	*Leishmania major* Friedlin	FR796431.1	nt 1567144-1568307	1163 bp	490/548(89%)	517/559(92%)	11/548(2%)	9/559(1%)
	Leishmania donovani BPK282A1	FR799622.2	nt 1579587-1580713	1126 bp	401/469(86%)	533/612(87%)	15/469(3%)	12/612(1%)
	Leishmania infantum JPCM5	FR796467.1	nt 1554867-1555992	1125 bp	401/469(86%)	533/612(87%)	16/469(3%)	12/612(1%)

However, no open reading frames (ORF) or annotated genes could be identified on the sequences detected for the two clones (Figure 25).

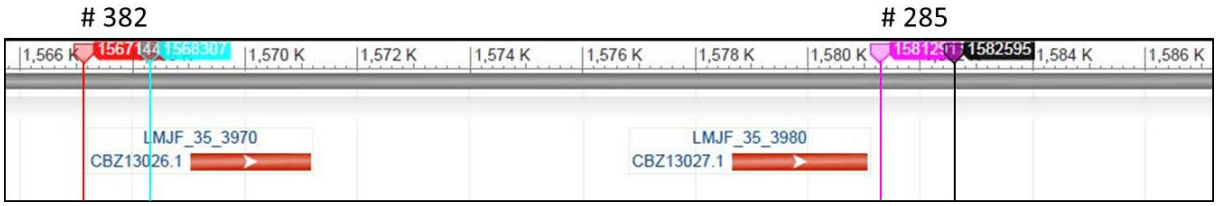

Figure 25. Location of the inserts aligned to the *L. major* genome at NCBI. The inserts of the two cosmids are highlighted and indicated above.

6.4 Multiplex approach for the identification and discrimination of sand fly species endemic to Israel and Palestine

6.4.1 Identification of SNPs specific for the different sand fly species studied

The aim of the MLPA approach was to differentiate 11 sand fly species which are all endemic to the Middle East, the main region of interest in this study. The target selected was the 18S rRNA gene which is a multi-copy gene and, therefore, of advantage for the detection of low amounts of template DNA. By aligning the sequences available at the NCBI using MEGA software, SNPs were identified that would allow to distinguish the different species by the presence/absence of the respective MLPA fragments. The company MRC-Holland in Amsterdam was entrusted with the design of the hybridization probes. Since single SNPs specific for each of the sand fly species included in this study could not be found, a combination of more than one SNP was needed for the unambiguous identification of the sand fly species. The determination scheme is given in Table 23.

Some of the designed probes gave a hybridization signal in more than one sand fly species, namely PERN1, SERG, TOB, ARA, PERN2 (marked with * in Table 23). Probes SYR and NEG were only specific in combination with each other (**). There is more than one probe or probe combination for some of the species since there are several non-identical sequences in the NCBI database. *P. perniciosus* (***) should give signals at 193 and 379 bp, (PERN1, TOB) or one fragment only at 507 bp (PERN2). Fragments at 281 and 507 bp (PAP, PERN2) were expected for eight sequences of *P. papatasi* (****), but only the fragment at 281 bp (PAP) for the sequence AJ244411.1 in the NCBI database. Furthermore, the *P. papatasi* sequence AJ391726.1 indicated the amplification of a 379 bp PCR product only.

Table 23. Determination scheme for the differentiation of the sand fly species of interest (provided by and modified after MRC Holland).
+, signal; -, no signal; ±, probably signal, but not entirely sure (ligation site is the same, surrounding sequence is slightly different); P, probably a low signal; *, probe is not specific; **, 310 and 340 are only specific in combination with each other; *, in case of *P. perniciosus* either probe 193 or probe 507 should give a signal (because there are only three *P. perniciosus* sequences on the NCBI website and these are not identical); **** 8 sequences: AJ244406.1, AJ244407.1, AJ244408.1, AJ244409.1, AJ244410.1, AJ244412.1, AJ244413.1 and AJ244414.1; PAP, This probe is specific for nine *P. papatasi* sequences, however does not detect the last *P. papatasi* sequence AJ391726.1. Therefore absence of this probe does not necessarily mean that the species is not *P. papatasi*; TOB, This probe fails to detect *P. tobbi* sequence AJ244383.1 and *P. neglectus* sequence AJ244373.1. Therefore absence of this probe does not necessarily mean that the species is not *P. tobbi* or *P. neglectus*; PERF, This probe is specific for four *P. perfiliewi* sequences, however might not detect the last three *P. perfiliewi* sequences AJ244386.1, AJ244388.1 and AJ244390.1 due to some mismatches. Therefore absence of this probe does not necessarily mean that the species is not *P. perfiliewi*.**

			SPECIES														
			S. fallax	*P. simici*	*P. perniciosus**** AJ244393.1 and AJ244394.1	*P. perniciosus**** AJ391728.1	*P. sergenti*	*P. alexandri*	*P. papatasi* (8 sequences****)	*P. papatasi* AJ244411.1	*P. papatasi* AJ391726.1	*P. syriacus*	*P. neglectus*	*P. tobbi*	*P. perfiliewi*	*P. arabicus*	
LENGTHS	144	FAL	+	-	-	-	-	-	-	-	-	-	-	-	-	-	Specific
	169	SIM	-	+	-	-	-	-	-	-	-	-	-	-	-	-	Specific
	193*	PERN1	-	+	+	-	-	-	-	-	-	-	-	-	-	-	Use 169 and 379 for determination
	225*	SERG	-	-	-	-	+	+	-	-	-	-	-	-	-	-	Use 250 for determination
	250	ALEX	-	-	-	-	-	+	-	-	-	-	-	-	-	-	Specific
	281	PAP	-	-	-	-	-	-	+	+	-	-	-	-	-	-	Specific
	310**	SYR	-	-	-	-	-	-	-	-	-	+	P	-	-	-	Use 340 for determination
	340**	NEG	-	-	-	-	-	-	-	-	-	P	+	-	-	-	Use 310 for determination
	379*	TOB	-	+	+	-	-	-	-	-	+	+	+	+	+	-	Use all the other probes for determination
	422	PERF	-	-	-	-	-	-	-	-	-	-	-	-	+	-	Specific
	460*	ARA	-	+	-	-	-	-	-	-	-	-	-	-	-	+	Use 169 for determination
	507*	PERN2	-	-	-	+	+	+	+	-	-	-	-	-	-	-	Use 225, 250 and 281 for determination

6.4.2 Test on sand fly samples by electrophoresis

The MLPA technique was tested on 65 sand flies of the species *P. papatasi*, *P. syriacus*, *P. sergenti*, *P. tobbi*, *P. arabicus* and *P. perfiliewi*. All specimens were collected in Israel at different foci (Figure 26 and Table 8 in the M&M section) and identified morphologically. In some of the samples DNA was not detectable by photometric determination, while others had only low amounts of DNA. Since PCR is a very sensitive method, all samples, disregarding of their low DNA contents, were used for MLPA.

Figure 26. Collection sites of the sand flies included in this study. The detailed views show the respective regions in more detail. Tel Aviv and Jerusalem are highlighted in blue as landmarks.

The amplification of MLPA fragments was successful in all samples. Figure 27 a-e shows the MLPA fragments obtained for the different sand fly species on 2.5% agarose gels. The fragment patterns of the specimens of *P. papatasi*, *P. sergenti*, *P. syriacus* and *P. perfiliewi* were consistent within the respective species. All 14 specimens of *P. papatasi* showed one fragment at ~280 bp and another at ~500 bp (Figure 27a), whereas all 13 samples of *P. sergenti* presented fragments of ~220 bp and ~500 bp (Figure 27b). Three different fragment sizes were found throughout all 13 samples of *P. syriacus*, which were ~310 bp, ~340 bp and ~380 bp in length (Figure 27c). All fragments of the eight specimens of *P. perfiliewi* had the same sizes at ~380 bp and ~420 bp (Figure 27d).

The electrophoretic separation of the MLPA products of *P. arabicus* and *P. tobbi* resulted, however, in intra-specific variations of the fragment patterns. All eight samples of *P. arabicus* showed fragments at ~170 bp, ~200 bp, ~380 bp and ~460 bp (Figure 27d), but one sample, *ara*07, had an additional band at ~420 bp. In the sample *ara*08, MLPA products were visible in the agarose gel, but there were also blurry fragments detected between ~60-80 bp. This is the size range of the Q-fragments (64, 70, 76, 82 bp) which were included in the probe mix as amplification controls. Since they were visible in the gel in this case, the amount of template DNA was probably quite low. Figure

27e shows the MPLA products of the nine samples of *P. tobbi*. Six of them presented only one fragment at ~380 bp. The sample *tob*07 had additional bands at ~220 bp and ~500 bp; sample *tob*09 had these two additional bands as well but lacked the ~380 bp band. The sample *tob*08 had additional weak bands at ~220 bp, ~310 bp, ~340 bp and ~420 bp compared to the majority of samples of this species. All samples of *P. tobbi* showed several very weak bands smaller than 220 bp which might be unspecific. The negative controls of the three different MLPA reactions did not show fragments within the size range of the MLPA probes. All three negative controls had one or several unclear bands of ~60-80 bp, which correspond to the size range of the control fragments included in the probe mix.

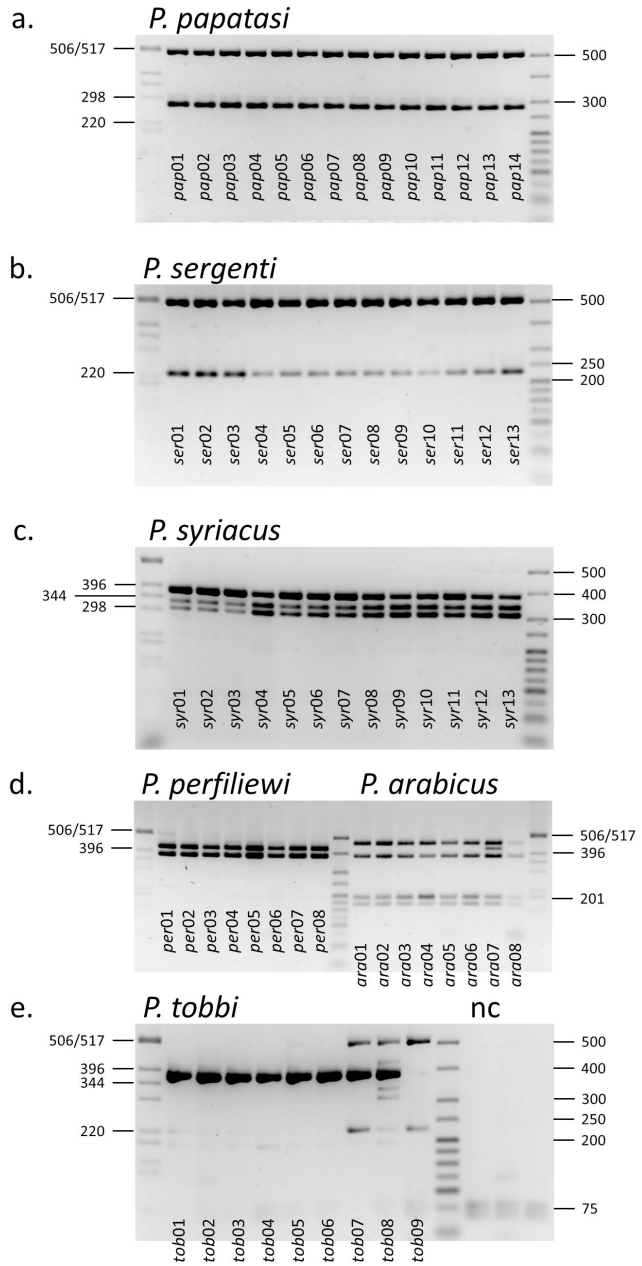

Figure 27. MLPA of 14 samples of *P. papatasi* (a), 13 samples of *P. sergenti* (b), 13 samples of *P. syriacus* (c), eight samples of each *P. perfiliewi* and *P. arabicus* (d), nine samples of *P. tobbi* and the negative controls (e). Ten µl of the MLPA reaction were loaded on the gel. Relevant fragments of the 1 kb DNA ladder and the hyperladder V are indicated.

6.4.3 Test of MLPA probes on sand fly specimens by fragment analysis

All samples were analysed by fragment analysis in an ABI sequencer. The resulting data were evaluated using the GeneMapper software version 3.7.

Figure 28 shows the fragment analyses for each of the different peak patterns successfully identified for the different sand fly species and Table 24 summarizes all fragment sizes retrieved from gel electrophoresis and fragment analysis for all sand fly specimens.

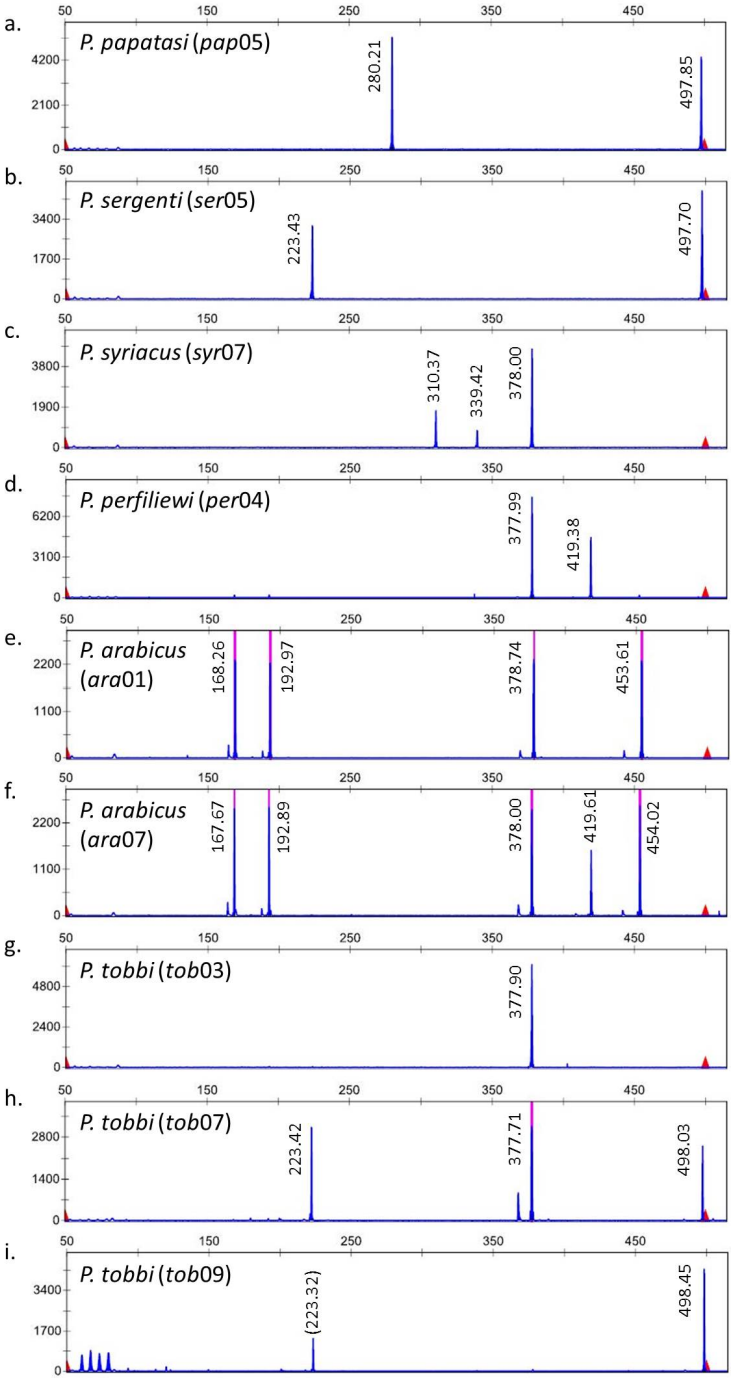

Figure 28. Fragment analysis of each of the identified peak patterns. One example is shown for each peak pattern. The sizes of the peaks are indicated in bp. The sample names are indicated in the respective panels.

A successful MLPA reaction was supposed to contain only very low amount of Q fragments. Thus, the peaks at the respective sizes, 64, 70, 76, and 82 bp, should be either very weak or invisible. Only in the negative control, an MLPA reaction without template DNA, the Q fragments should be clearly visible. In the negative controls of three runs the Q fragments were detected by the fragment analysis at ~61 bp (61.11, 60.86, 60.83), ~67 bp (67.05, 66.8, 66.8), ~73 bp (73.05, 73.0, 73.08), and ~79 bp (79.18, 79.24, 79.5). An example of the peaks of the negative control is shown in Figure 29.

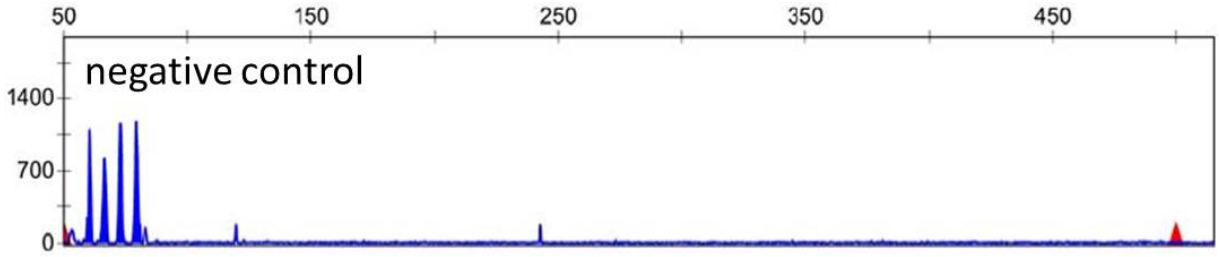

Figure 29. Fragment analysis of the negative control.

The Q fragments were either absent or very weak in the fragment analyses of 63 out of the 65 sand fly samples tested (Figure 28 a-h). Uncertainty occurred only in one sample of *P. arabicus*, *ara*08, and in one sample of *P. tobbi*, *tob*09, which presented quite high peaks for the Q fragments almost as high as the probe fragments (Figure 30). Most probably, in these two cases the DNA concentration was rather low. However, unspecific peaks were detected at ~53 bp and ~83 bp in almost all of the 65 cases. They were low in intensity, and in a size range which did not interfere with the probe signals. The probe signals were clearly seen in all 65 analysed samples.

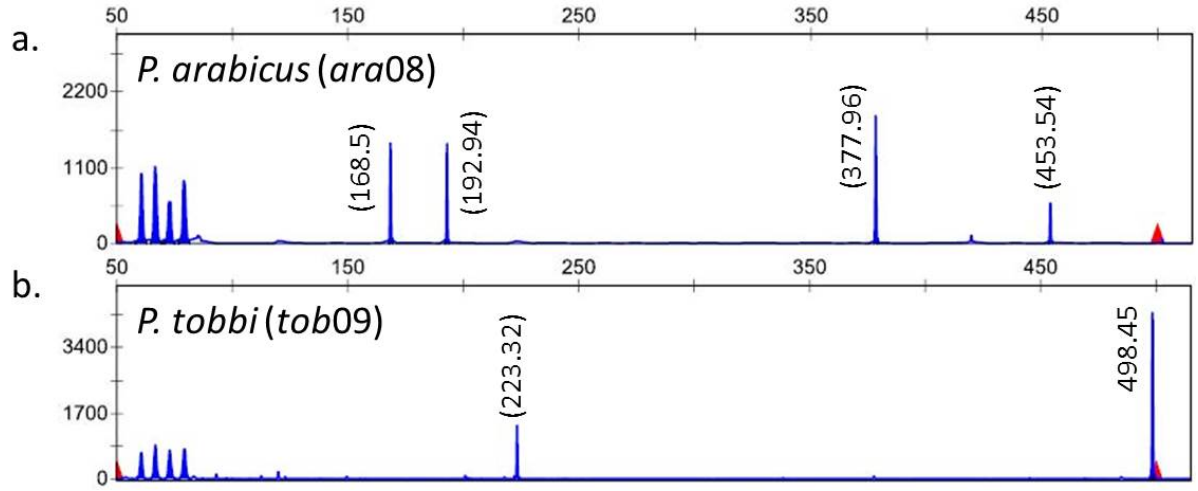

Figure 30. Fragment analysis of the two samples presenting control fragments which peak intensities indicate low DNA template amount or a weak ligation reaction.

Table 24. Summary of MLPA results of all samples. Each sample is presented in one row giving its positive bands in the gel electrophoresis, +, and their sizes as determined in the fragment analysis. (+), weak band in the electrophoresis; *, peak detected in the fragment analysis, but not significantly higher than the control fragments.

	MLPA probes, lengths in bp											
sand fly sample	144	169	193	225	250	281	310	340	379	422	460	507
pap 01						+ 280.19						+ 497.93
pap 02						+ 280.22						+ 497.89
pap 03						+ 280.20						+ 498.17
pap 04						+ 280.18						+ 498.01
pap 05						+ 280.21						+ 497.85
pap 06						+ 280.18						+ 497.85
pap 07						+ 280.20						+ 497.99
pap 08						+ 280.21						+ 497.86
pap 09						+ 280.18						+ 498.01
pap 10						+ 280.16						+ 498.34
pap 11						+ 280.22						+ 498.35
pap 12						+ 280.20						+ 499.04
pap 13						+ 280.28						+ 497.85
pap 14						+ 280.15						+ 497.85
ser 01				+ 223.30								+ 498.45
ser 02				+ 223.45								+ 498.01
ser 03				+ 223.24								+ 498.14
ser 04				+ 223.37								+ 497.57
ser 05				+ 223.43								+ 497.70
ser 06				+ 223.44								+ 498.03
ser 07				+ 223.42								+ 498.91
ser 08				+ 223.39								+ 497.87
ser 09				+ 223.46								+ 497.67
ser 10				+ 223.41								+ 498.01
ser 11				+ 223.43								+ 497.81
ser 12				+ 223.44								+ 497.63
ser 13				+ 223.49								+ 497.76

	MLPA probes, lengths in bp																							
sand fly sample	144		169		193		225		250		281		310		340		379		422		460		507	
syr 01													+	310.54	+	339.52	+	378.71						
syr 02													+	310.52	+	339.52	+	377.94						
syr 03													+	310.49	+	338.45	+	377.80						
syr 04													+	310.24	+	339.35	+	378.00						
syr 05													+	310.32	+	339.35	+	377.97						
syr 06													+	310.31	+	339.35	+	377.99						
syr 07													+	310.37	+	339.42	+	378.00						
syr 08													+	310.36	+	339.42	+	377.87						
syr 09													+	310.44	+	339.49	+	377.90						
syr 10													+	310.27	+	339.28	+	377.90						
syr 11													+	310.05	+	339.17	+	378.11						
syr 12													+	310.32	+	339.36	+	377.96						
syr 13													+	310.44	+	339.43	+	377.96						
per 01																	+	377.64	+	419.47				
per 02																	+	377.67	+	419.49				
per 03																	+	377.73	+	419.31				
per 04																	+	377.99	+	419.38				
per 05																	+	377.87	+	419.46				
per 06																	+	377.95	+	419.39				
per 07																	+	377.91	+	419.35				
per 08																	+	377.93	+	419.36				
ara 01			+	168.26	+	192.97											+	378.74			+	453.61		
ara 02			+	168.37	+	193.21											+	377.68			+	453.95		
ara 03			+	168.53	+	192.96											+	377.93			+	453.59		
ara 04			+	168.54	+	193.11											+	377.97			+	453.31		
ara 05			+	168.53	+	193.04											+	377.83			+	453.57		
ara 06			+	168.52	+	192.98											+	377.97			+	453.48		
ara 07			+	167.67	+	192.89											+	378.00	+	419.61	+	454.02		
ara 08			+	168.5*	+	192.94*											+	377.96*			+	453.54*		

	MLPA probes, lengths in bp																							
sand fly sample	144		169		193		225		250		281		310		340		379		422		460		507	
tob 01																	+	377.54						
tob 02																	+	377.51						
tob 03																	+	377.90						
tob 04																	+	377.88						
tob 05																	+	377.79						
tob 06																	+	377.83						
tob 07							+	223.42									+	377.71					+	498.03
tob 08							(+)						(+)		(+)		+	377.92	(+)				+	
tob 09							+	223.32*															+	498.45

The probes of 144 bp and 250 bp length were not amplified in any of the analysed sand flies and the probes of 169 bp, 193 bp and 460 bp exclusively in the eight specimens of *P. arabicus*. A 225 bp PCR product was amplified in all 13 samples of *P. sergenti* and two samples of *P. tobbi*. The 281 bp amplicon was unique to the 14 samples of *P. papatasi*. In the fragment analysis, the 310 bp and 340 bp PCR products were unique to *P. syriacus*, although there were bands of these sizes seen in the gel electrophoresis for the *P. tobbi* sample *tob*08. The 379 bp MLPA fragment was amplified in *P. syriacus*, *P. perfiliewi*, *P. arabicus* and eight of the ten samples of *P. tobbi* and the 422 bp fragment in all eight *P. perfiliewi* samples as well as one of *P. arabicus*, *ara*07. The 507 bp PCR product was yielded for all 14 specimens of *P. papatasi*, all 13 of *P. sergenti* and two of *P. tobbi*, *tob*07 and *tob*09. In total, six hybridization probes were specific for one species in the fragment analysis leading to the amplification of 169 bp, 193 bp and 460 bp products for *P. arabicus*, 281 bp product for *P. papatasi* and 310 bp and 340 bp products for *P. syriacus*, although the latter two probes showed a positive signal in the gel electrophoresis in *tob*08.

Differences between the fragment analysis and the gel electrophoresis results were detected only in *tob*08 (Table 24 and Figure 31). The positive signals at ~220 bp, ~310 bp, ~340 bp and ~420 bp could not be confirmed by fragment analysis. Only the 379 bp fragment was detected by both methods. Since the fragment analysis is more reliable in terms of reproducibility and objective evaluation of the results, these results should be the decisive ones.

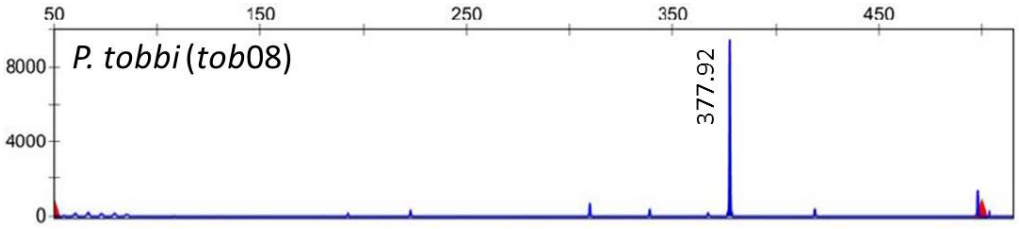

Figure 31. Fragment analysis of *tob*08. This sample showed a special band pattern in the gel electrophoresis. The weak bands detected in the electrophoresis were too small to be considered as correct peaks in the fragment analysis.

The minimum, maximum and mean sizes of all amplified probes are summarized in Table 25. The smaller MLPA products (169 bp to 379 bp) showed an absolute deviation of less than 2 bp to the theoretical value calculated from the probe design. The three longer products, 422 bp, 460 bp and 507 bp, deviated from the calculated value by less than 3, 7 and 10 bp, respectively.

Table 25. Descriptive statistics of all amplified probes. The amount of samples with a positive signal for this fragment (n), the minimum and maximum sizes and the mean of the sizes are given for each MLPA probe.

	MLPA probes, lengths in bp											
	144	**169**	**193**	**225**	**250**	**281**	**310**	**340**	**379**	**422**	**460**	**507**
n	0	8	8	15	0	14	13	13	33	9	8	29
minimum	-	167.67	192.89	223.24	-	280.15	310.05	338.45	377.51	419.31	453.31	497.57
maximum	-	168.54	193.21	223.49	-	280.28	310.54	339.52	378.74	419.61	454.02	499.04
mean	-	168.37	193.01	223.40	-	280.20	310.36	339.32	377.92	419.42	453.63	498.04

7 Discussion

7.1 Microsatellite diversity in *L. tropica* and *L. aethiopica*

7.1.1 Detailed view on certain populations, clusters and groups

Compared to an earlier study by Schwenkenbecher *et al.* [30], a reduced set of microsatellite markers was applied here for characterizing a total of 195 strains of *L. tropica* of different geographical origins and eight of the closely related species *L. aethiopica* isolated from different hosts and representing different clinical manifestations of leishmaniasis. The smaller marker set, 12 vs. 21 microsatellite markers, proved to be sufficiently discriminatory and informative for studying the diversity and population genetic structure in the global strain set used herein.

Four phylogenetic analysis methods, each based on different algorithms, were applied to the microsatellite profiles obtained and their results were compared to each other. The results of these analyses were congruent for most of the strains studied, but were also discrepant in some cases. Those findings which were backed by all of the algorithms were considered as reliable, while others had to be treated with caution. Those strains for which the assignment to certain genetic entities varied, are put in parenthesis in Table 11.

7.1.1.1 Strains of *L. tropica* from Israeli and Palestinian foci are diverse and belong to different genetic groups

The 40 Israeli and 56 Palestinian strains included in the data set were isolated between 1949 and 2011. All Palestinian and 27 of the Israeli strains were obtained from human cases of CL. Eleven strains from Israel were isolated from sand fly vectors, six from *P. sergenti* and five from *P. arabicus*, one strain from a rock hyrax and one from a human case of VL. The collection sites of all strains are summarized in Table 3 and shown in Figure 32. Both territories, Israel and Palestine, should be considered as one area for several reasons. First, they cover only a relatively small area and, despite the political situation and the boarder separating these two countries, people are extensively travelling between them. Furthermore, the WHO code gives the country where human cases were diagnosed, but patients might have gotten infected elsewhere.

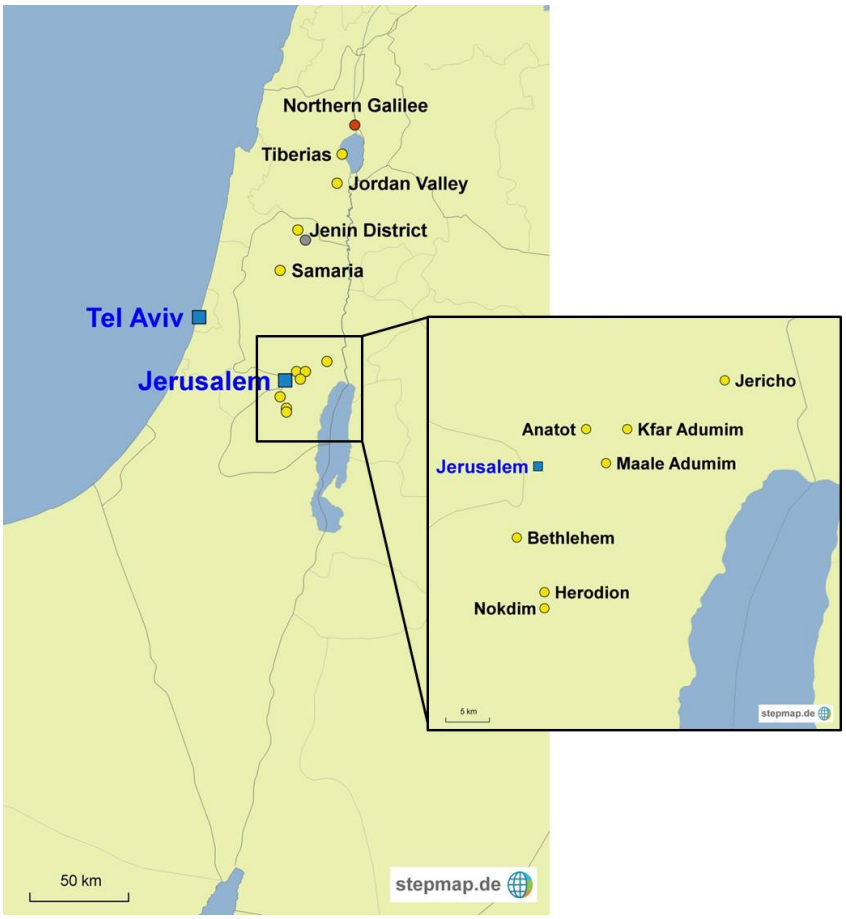

Figure 32. Collection sites of the strains of *L. tropica* in Israel and Palestine. The circles are coloured according to their assignment by the Bayesian statistics; yellow, "Israel/Palestine", red, "Northern Galilee", grey, "India/Palestine". Jerusalem and Tel Aviv are indicated for better orientation.

In total, 59 different microsatellite profiles were determined for the 96 Israeli and Palestinian strains of *L. tropica* studied herein. Twenty-seven Israeli and Palestinian strains collected between 2001 and 2010 shared the same microsatellite profile, *Ltro*MS 041 (Table 11). Another profile, *Ltro*MS 042 was presented by three strains and nine other profiles by two strains each. The remaining 48 strains from this area had unique microsatellite profiles. None of the profiles of Israeli and Palestinian strains were shared with strains of other geographical origins.

The strains of *L. tropica* from Israeli and Palestinian foci were found in four different genetic entities (Figure 16). At the highest hierarchical level, the Bayesian clustering approach assigned 75 of the Israeli and Palestinian strains of *L. tropica* to one population named "Israel/Palestine" together with one strain from Egypt, MHOM/EG/ 1990/LPN65 (Figure 16) and separated them from all other strains under study. The close genetic relationship of these strains as well as their large genetic distance to all other strains tested (Table 17) was confirmed by all subsequent analyses. They formed monophyletic groups in the NJ tree (Figure 17) and the network (Figure 18) and were also remotest to all other groups of *L. tropica* in the FCA (Figure 19). Low heterozygosity was confirmed for this population by the descriptive statistics (H_o=0.035, Table 16).

The remaining 21 Israeli and Palestinian strains could be assigned to distinct genetic entities only at cluster or group levels. Ten Israeli strains isolated from human cases of CL, a hyrax, *P. capensis,* and from different sand fly vector species, *P. sergenti* and *P. arabicus,* formed the cluster "Northern Galilee" identified by Bayesian statistics and confirmed by genetic distance analyses, the NJ tree and the network, and the FCA.

The other 11 strains belonged to different genetic clusters and groups of the sub-population "Asia" (Figure 12). The group "India/Palestine" of the cluster "Middle East/India" contained three Palestinian strains isolated in the Jenin District. This group was, again, confirmed by all approaches used, although the monophyletic nature of this group was not as strong as in the previously mentioned genetic entities of Israeli and Palestinian strains. One strain each from Israel and Palestine was found in the group "IL/PS/JO/IN". Another two strains, both from the Jenin District in Palestine, could not be assigned clearly to any of the three groups of the cluster "Middle East/India". The group "IL/PS/JO/IN" should be, in fact, not considered as a confirmed genetic entity but rather as an artificial group created by the Bayesian statistics approach which has the disadvantage to summarize all strains which do not have sufficient genetic similarity to any of the other groups, into a new one.

The group "old strains/TR" of the cluster "old strains/Turkey" identified by the Bayesian clustering approach included three strains from Israel isolated in 1949, 1959 and 1990 and one strain from Palestine isolated in 2002. Two of these strains, MHOM/IL/1949/LRC-L43 and MHOM/PS/2002/18JnF4, clustered together with the old strains from other Asian foci in all analyses. However, the strains MHOM/IL/1990/P283 and MHOM/IL/1959/LRC-L22 were outliers in the NJ tree and the network. Because of the great differences between the results of the different approaches the genetic relationship between these latter strains could not be confirmed.

The separation of the Israeli and Palestinian strains into several clusters is in agreement with earlier studies [30, 108]. According to Azmi *et al.*, 12 Palestinian and three Israeli strains of *L. tropica* belonged to two clades by using a kDNA-RFLP typing approach, which was in accordance to MLEE typing, EF serotyping and the MLMT results of Schwenkenbecher *et al.* The present study, which included 12 of the strains analysed in the kDNA-RFLP study, corroborated the existence of these two clades. Three strains which had been assigned to clade A were part of the sub-population "Asia", while the nine strains assigned to clade B were part of the sub-population "Israel/Palestine".

The strong monophyletic nature of the population "Israel/Palestine" suggests that these parasites might have evolved from one single founder strain of *L. tropica*. This could have been the oldest strain in this population, MHOM/EG/1990/LPN65, which was isolated on the Sinai Peninsula. It might have been introduced into Israel/Palestine and, subsequently, spread throughout this area.

Transmission of *L. tropica* in Israeli and Palestinian foci is likely to be zoonotic as has been also suspected for several African foci but is in contrast to the anthroponotic transmission reported for other Asian foci [42]. Rock hyraxes infected by *L. tropica* have been found in different Israeli foci and the animals were shown highly abundant in Palestinian areas endemic for *L. tropica*. They are therefore incriminated as the parasites´ natural mammalian reservoirs. Recent house building activities of the Israeli population in close proximity to the natural habitats of the rock hyraxes, possibly, facilitated the transmission of the *L. tropica* parasites from the animals to humans and vice versa [31]. Interestingly, the "Northern Galilee" cluster is related to the African strains of *L. tropica* in all analysis performed and rock hyraxes are the proven animal reservoir for *L. tropica* in Kenya and Namibia [2]. Thus, the Galilean strains might have been introduced into Israel from Africa by infected humans or animals. Hyraxes of the species *P. capensis* are known to live along the Great Rift Valley which stretches from Syria via the Jordan Valley through East Africa and ends in South Africa [109].

Based on the above findings, the establishment of genetically different *L. tropica* parasites in Israeli and Palestinian foci is, most probably, the consequence of their multiple importations into this area rather than due to diversifying evolution from a common ancestor. The multiple introductions of genetically different *L. tropica* parasites into Israel and Palestine is probably the result of extensive travelling and migration of humans or animals, from a great variety of countries into this region, which started several thousand years ago and is still ongoing.

7.1.1.2 Intrafocal diversity of Moroccan strains of *L. tropica*

A total of 22 Moroccan strains of *L. tropica* were analysed in this study. Most of them were isolated in the neighbouring Chichaoua and Marrakech provinces which belong to the region Marrakech-Tensift-Al Haouz in western Morocco, where CL is caused by *L. tropica* [110] and transmitted by *P. sergenti* [45, 111] (Figure 33).

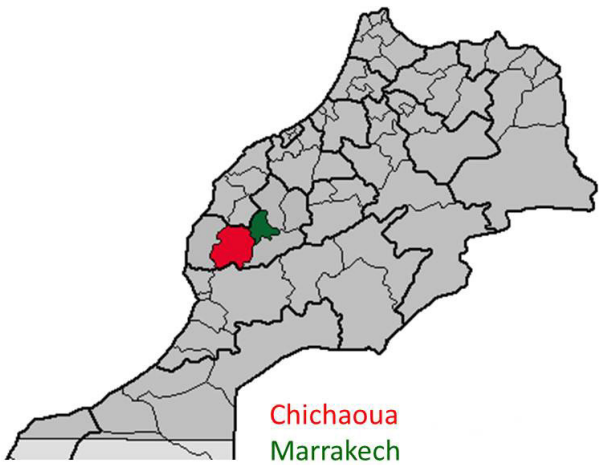

Figure 33. Collection sites of the Moroccan strains of *L. tropica*. The two provinces Chichaoua and Marrakech are highlighted. Modified from *Die Provinz Chichaoua in der Region Marrakech-Tensift-Al Haouz* by Vincent van Zeijst/Wikimedia Commons/CC BY-SA 3.0.

Sixteen microsatellite profiles were found among the Moroccan strains of *L. tropica*, 13 of which were unique to one strain each. Two strains from Marrakech presented the same profile, *Ltro*MS 094; the profiles *Ltro*MS 098 and *Ltro*MS 099 were shared by three and four strains from Chichaoua, respectively.

Bayesian statistics divided the Moroccan strains into two clusters of the sub-population "Africa/Galilee". The first one contained seven strains from Chichaoua, six from Marrakech and four Moroccan strains of unknown origin and, additionally, two strains from Turkey and was named "MoroccoA/Turkey". "MoroccoB" consisted of one strain from Chichaoua and three from Marrakech. The strain MHOM/MA/2000/INH-W09 from Chichaoua was assigned to the cluster "KE/TN/YE/MA", but had a considerable partial membership coefficient for "MoroccoA/Turkey" (0.311). Also strain MHOM/MA/1995/LEM3015 had shared membership coefficients of "MoroccoA/Turkey" (0.620) and "MoroccoB" (0.264). The existence of two separate Moroccan clusters was largely confirmed by the different analyses, with a few exceptions in the strains' placement by the distance-based approaches. Both clusters were confirmed by the FCA (Figure 19 and Figure 20). Discrepant results were observed for three Moroccan strains. The strain MHOM/MA/1992/LPN79, assigned to "MoroccoA/Turkey" by Bayesian statistics had a unique position in the network (Figure 18), and the strains MHOM/MA/1995/LEM3015 and MHOM/MA/2000/INH-W09, which were assigned to "MoroccoA/Turkey" and "KE/TN/YE/MA", respectively, by Bayesian statistics, were outliers in both the NJ tree (Figure 17) and the network (Figure 18) and more closely related to the strains of "MoroccoB". The genetic distance between the two clusters, "MoroccoA" and "MoroccoB", was categorized as very great with an F_{ST} value of 0.580. The cluster "KE/TN/YE/MA" might be an artificially created group, because STRUCTURE was not able to assign those strains to one of the other genetic entities.

The high genetic diversity of Moroccan strains revealed by MLMT and the existence of genetically distinct clusters in the same CL focus is corroborating the results of previous studies that have used MLMT [30] and multilocus enzyme electrophoresis (MLEE) approaches [28]. MLEE revealed a particularly high degree of isoenzyme heterogeneity among strains of *L. tropica* from Morocco, even within the same geographical foci [112]. Ten zymodemes were distinguished among 178 Moroccan strains of *L. tropica*. For comparison, only one zymodeme of *L. major* (MON-25) was found in Morocco in this study and only two zymodemes of *L. infantum* (MON-1 and MON-24) were identified among Moroccan strains in another study [113].

The intra-focal diversity within the Moroccan strains indicates that not only the geographical origin is a factor in correlating the strains to their genetic properties. The possible explanations for genetic diversity within a single focus range from (i) the subsequent introduction of distinct and separate

genetic strains into this area from other geographical regions via (ii) mutation events, possibly in combination with genetic exchange to (iii) recent hybridisation followed by recombination as was described for *L. infantum* [114]. Mutation events alone are not likely to account for the great genetic differences between the two clusters of Moroccan strains of *L. tropica*. The low heterozygosity and the relatively high F_{IS} values of 0.620 ("MoroccoA/Turkey") and 0.624 ("MoroccoB") (Table 16), indicating high inbreeding rates, contradict this explanation of hybridisation / genetic exchange events. Given the infrequency of meiotic cell divisions in *Leishmania* parasites as shown by Rogers *et al.* for *L. infantum* ($7.5x10^4$ mitotic cell divisions per meiosis [114]) and summarizing the above considerations, the most likely hypothesis for the intra-focal genetic diversity of *L. tropica* in Morocco is the consecutive introduction of different strains from other geographical origins.

A further question is why the strains of these two genetic clusters have not interbred and thus mixed their genotypes. The strains from Marrakech were isolated in 1988 and 1989, while those from Chichaoua were isolated in 2000. Genetic exchange is supposed to happen within the female sand fly vector [115].Thus, one sand fly had to get infected by parasites of both variants at about the same time, which is a rather unlikely event. Additionally, the parasites of the two variants could possibly be transmitted by different sand fly vectors or vector populations, which would prevent co-infection of one sand fly with and, subsequently, the possible genetic exchange between individuals of different genetic groups of *L. tropica*. In Morocco, *L. tropica* was isolated from specimens of *P. sergenti* and *P. chabaudi* [1]. Only two strains of *L. tropica* studied here were isolated from sand flies of the species *P. sergenti*, both assigned to cluster "MoroccoB". All other strains were isolated from human cases of CL.

7.1.1.3 Indian strains of *L. tropica* from cutaneous and visceral cases of leishmaniasis belong to different genetic groups

The strain set included 12 strains of *L. tropica* collected from patients in India. Eight of these strains were isolated from human cases of CL in Bikaner, north-west India in 2006 and 2007 and the other four strains were isolated from human cases of VL, three of them in Bihar, north-east India and one of unknown origin between 1979 and 1997 (Table 11 and Figure 34). Additionally, four strains from other cases of human VL were included for comparison. These strains were from Kenya (n=2), Iraq (n=1) and Israel (n=1).

Six of the eight strains isolated from CL cases from Bikaner shared one microsatellite profile, *Ltro*MS 030; the other two strains had unique profiles. Two of the Bihar strains from VL patients had the same profile, *Ltro*MS 018, and one Kenyan and the Iraqi strain presented an identical profile,

*Ltro*MS 005. The Israeli strain shared its profile with a strain from Azerbaijan isolated from a human case of CL. The other strains isolated from human VL cases had unique profiles.

Figure 34. Collection sites of the Indian strains of *L. tropica*.

The Bayesian clustering approach divided the Indian strains of *L. tropica* among three different groups of the sub-population "Asia" (Figure 15 and Figure 16). All eight CL strains from Bikaner were assigned to the group "India/Palestine" and this clustering was confirmed by all subsequent analyses. For the Indian VL strains the results varied in the different analyses. Bayesian statistics assigned two of these strains to "IL/PS/JO/IN" and two to "old strains/TR". In the NJ tree and the network all but one of the VL strains clustered together in the group "old strains/TR" (Figure 17 and Figure 18). Strain MHOM/IN/1979/DD7 which was assigned to "old strains/TR" by the Bayesian statistics was an outlier in the NJ tree and the network. In the FCA, the three groups "India/Palestine", "IL/PS/JO/IN" and "old strains/TR" were very closely related and did not appear as different groups (Figure 20). In conclusion, MLMT revealed that the strains of *L. tropica* from Bikaner isolated from human CL cases fell into the same sub-population as the strains from Bihar isolated from human VL cases, but were not identical to them.

Several reasons can possibly explain the divergence of the Indian strains of *L. tropica*. The distance between the two foci Bikaner in north-western India and Bihar in north-eastern India is about 1300 km which makes it long enough to serve as one explanation. Geographical correlation to the parasites' microsatellite profiles has been shown in several previous MLMT studies on different *Leishmania* species [62, 63, 67] and, also, is the main finding of this study. However, geographical distance seems to be not the only parameter to separate genetic populations of *L. tropica*. Strains from a wide variety of origins were assigned to the same group as the Indian strains from Bihar, including quite remote countries like Israel, Iraq and Kenya. The clustering of many strains which share an early time point of isolation in one monophyletic group ("old strains") indicates the importance of this property. The parasites might have evolved during the last decades and in this

course spread throughout the world. In this study, the time span between the isolation of the Indian strains of *L. tropica* isolated from VL cases and those isolated from CL cases varied between 10 and 28 years. The effect of passing time has been shown in previous studies on strains of the *L. donovani* complex and *L. tropica* [30, 62]. The long time span might have allowed an increase of genetic mutations in the genome. The parasites were derived from different strain collections where, most probably methods of cultivation and numbers of passaging have varied. This might have led to differences in the accumulation of mutations due to varying evolutionary pressure. A third possibility for the genetic separation of the Indian strains isolated from CL and VL cases could be intrinsic genetic differences, causing one type to be dermotropic and the other to be viscerotropic. The microsatellite markers used herein are located in the non-coding regions of the genome, thus a correlation between the genotype and the phenotype is highly speculative and could, so far, not be confirmed by previous studies [67, 116]. On the other hand, it has been demonstrated that the hosts´ nutritional and immune status effects the manifestation of different forms of leishmaniases considerably [20-22]. Interestingly, most of the VL cases of *L. tropica* described in the literature occurred in young children and in non-human hosts (dogs and rats) [3-7]. Also, parasites of the *L. infantum* zymodeme MON-24, which usually cause CL, have been found to cause VL in children and HIV immunodeficient individuals [16]. These examples suggest that natural or acquired immunodeficiency in the hosts might cause parasite invasion of internal organs.

To elucidate the role of the parasite´s genotype in the development of the disease, more strains of *L. tropica* isolated from cases of VL have to be analysed. Two groups of strains which do not differ considerably in terms of geographical origin and collection date should be included in the study to exclude these factors. In Himachal Pradesh, northern India, *L. tropica* was found to cause both human CL and VL [11, 12], which qualifies this focus as an appropriate focus to study this question.

7.1.1.4 Clonal expansion of Turkish strains of *L. tropica*

Thirty-three strains from Turkey were included in this study. All were isolated from human cases of CL between 1995 and 1999, the vast majority of them, 26 strains, in the vicinity of Sanliurfa, south-eastern Turkey. These latter strains were all collected during a local outbreak in 1995. In total, 12 different microsatellite profiles were found among the Turkish strains (Table 11). All strains from Sanliurfa clustered, however, together in one group, “Sanliurfa”. This result was consistent throughout all approaches, except for one strain, MHOM/TR/1995/URFA3, which was assigned to the group “old strains/TR” by Bayesian statistics (Figure 15 to Figure 20). Only five different microsatellite profiles were found among the strains from Sanliurfa, of which *Ltro*MS 027 was shared by 12 strains, *Ltro*MS 028 was shared by nine strains and *Ltro*MS 029 was shared by two strains (Table 11 and Table 13).These five microsatellite profiles differed mainly in the presence or absence of certain markers.

*Ltro*MS 028 contained values for all markers, while one or two markers were not amplified successfully in *Ltro*MS 024 (LIST7040, 27GTGn), *Ltro*MS 026 (GA6, 27GTGn) and *Ltro*MS 027 (27GTGn). Only *Ltro*MS 029 showed a difference in fragment length in marker 27GTGn, which was homozygous (112 bp) and not, like in *Ltro*MS 028, heterozygous (112 and 118 bp). Thus, when considering missing data (nd) not as distinct microsatellite value, the strains from Sanliurfa displayed only two different profiles.

The strain MHOM/TR/1995/URFA3 did not cluster with the other strains from Sanliurfa and instead was assigned to the group "old strains/TR" by Bayesian statistics. In contrast, the genetic distance analyses, both the NJ tree and the network, showed, that this strain was related to the other strains from Sanliurfa. Four other strains from Turkey assigned to "old strains/TR" by Bayesian clustering, were also found near the Sanliurfa strains in the NJ tree and the network.

Interestingly, the descriptive statistics showed that the group of strains from Sanliurfa was close to complete outcrossing (F_{IS} = -0.921) (Table 16). This indicates the strains from Sanliurfa to have evolved clonally, since clonal diploids are expected to accumulate heterozygosity over time at every locus (H_O = 0.318).

Taking all the above findings together the CL epidemic in the vicinity of Sanliurfa was presumably caused by one strain which was introduced into this region and then expanded clonally. *L. tropica* was not present in this area before and was probably introduced by a single infection source. One model to explain the rapid spread of this clone is the anthroponotic transmission within this city with 526,247 inhabitants in its urban area (as at 2012, http://en.wikipedia.org/wiki/sanliurfa). Transmission of CL in this region has always been suggested to be anthroponotic before [117]. The identification of a distinct genetic group consisting of strains from Sanliurfa proves that the microsatellite typing approach is not only suitable for the discrimination of geographically isolated populations of *Leishmania*, but also detects populations which arose from epidemic outbreaks.

7.1.1.5 Strains of *L. aethiopica* formed one genetically isolated cluster

In this microsatellite typing study, eight strains of *L. aethiopica* were included to reveal their phylogenetic relation to the strains of *L. tropica*. The strains were all collected in Ethiopia between 1972 and 1994.

The amplification of at least ten of the microsatellite loci was successful in all strains of *L. aethiopica* (Table 11), which allowed their inclusion in the phylogenetic analyses together with strains of *L. tropica*. Each of the strains of *L. aethiopica* had an individual microsatellite profile. Thus, despite the low sample number of only eight strains, this cluster consisted of the largest mean number of

alleles (A=3.25) indicating large heterogeneity (Table 16). This finding is in accordance to the results of an MLEE analysis, which found *L. aethiopica* to be very heterogeneous [28, 118].

All approaches assigned the strains of *L. aethiopica* to one genetic entity, which was clearly separated from the strains of *L. tropica*. Bayesian statistics assigned them to the sub-population "Africa/Galilee" together with African strains of *L. tropica* and the Israeli strains from the northern side of the Sea of Galilee (Figure 15 and Figure 16). The strains of *L. aethiopica* formed a monophyletic group in the NJ tree (Figure 17) and also in the network (Figure 18). While they were found among the African strains of *L. tropica* in the Bayesian statistics, the NJ tree and the network, their genetic distance to all strains of *L. tropica* was very large in the FCA (Figure 19). The genetic separation was greatest to the cluster "Israel/Palestine" in the *F* statistics (F_{ST}=0.757) and least to the cluster "KE/TN/YE/MA" (F_{ST}=0.305).

Three of eight strains were collected in northern, one in central and three in southern Ethiopia; the exact point of isolation of one strain is not known. Schönian *et al*. applied ITS1-RFLP and PCR fingerprinting to ten strains of *L. aethiopica*, six of which were analysed in this study as well [119]. They found two sub-structures within these ten strains correlating with their point of isolation. This separation could not be confirmed in this study.

The strains were isolated from patients suffering from two different forms of CL, diffuse and local CL, but their microsatellite profiles did not show any correlation with the disease manifestation. This is in accordance to the results of the previously mentioned study [119].

L. tropica and *L. aethiopica* were shown to be closely related by some phylogenetic studies, while others suggested a closer relationship between *L. aethiopica* and *L. major*. An MLEE approach using 15 enzyme markers revealed that strains of *L. aethiopica* clustered in one monophyletic group together with strains of *L. major* [28, 118]. The close relationship between these two species was confirmed by sequencing of the 370 bp fragment of the coding region of the *hsp*20 gene [120]. A study combining ITS1-RFLP using ten different enzymes with DNA fingerprinting also supported a monophyletic group consisting of *L. aethiopica* and *L. major* [119]. On the contrary, RFLP of the *hsp*70 gene using *Hae*III is not sufficient to discriminate between these two species, an additional enzyme is necessary for their distinction [121]. This suggested that *L. aethiopica* might be more closely related to *L. tropica.* In a reverse line dot blot assay for the detection of *L. tropica*, two of the three probes were positive for *L. aethiopica* samples, indicating similar DNA sequences of these two species [122]. All other probes in this assay were species-specific. *L. tropica* and *L. aethiopica* formed a monophyletic group in an MLSA approach analysing 552 polymorphic sites in seven single copy coding DNA sequences located on six different chromosomes, giving a 4,677 bp-long concatenated

sequence [123]. However, despite the revealed close phylogenetic relationship to *L. tropica* and *L. major*, *L. aethiopica* was clearly defined as a distinct species not only by the above methods, but also by many others, like ITS1-RFLP [84], High Resolution Melt Analysis (HRM) [124] and PCR approaches using species-specific primers [65, 125].

The marker set used in this study was also applied to a strain of *L. major* (MHOM/IL/80/Friedlin) to be included as an out-group in the phylogenetic analyses. It was, however, only possible to amplify eight of the *L. tropica* microsatellite markers in this *L. major* strain. This was insufficient, since at least 10 loci should to be typed for getting a satisfying statistical support of the analyses. This result also indicates a closer relationship of *L. aethiopica* to *L. tropica* than *L. major*.

7.1.2 Phylogenetic structure of *L. tropica*, *L. major* and *L. aethiopica* revealed by whole genome sequencing (WGS)

The population genetic analysis of *Leishmania spp*. using microsatellite markers has two major obstacles. The markers are prone to homoplasy and more or less species-specific. The amplification of a sufficient amount of markers was not possible in *L. major*, the closest relative of *L. tropica*/*L. aethiopica*, and thus, the distance-based NJ tree and the network could not be rooted. Microsatellite data are therefore not appropriate to study phylogenetic evolution [60]. They do not allow drawing conclusions about the evolutionary history and the direction of the evolution.

For this reason, a whole genome sequencing (WGS) approach was started by our group (Moser *et al*., unpublished data). The preliminary results of this study shall be discussed here in comparison to the MLMT results. The WGS approach included 153 strains of *L. tropica*, 71 of which were also included in the MLMT study, six strains of *L. aethiopica*, all studied by MLMT, and 13 strains of *L. major*. The strains were collected in a great variety of endemic regions, ranging from Namibia through the Middle East to India between 1949 and 2013. In total, 44,020 SNPs were analysed, which were located in selected housekeeping genes, genes involved in mismatch repair and genes encoding for hypothetical proteins annotated in *L. donovani and L. major* (n=252). The gene sequences retrieved from the whole genome sequences of *L. donovani* were aligned to the draft whole genome sequence of *L. tropica*. Thus, the location of the SNPs in coding regions is only an assumption based on the annotated genes in other *Leishmania* species. Since no reference genome of *L. tropica* has been published to date, it could not be confirmed whether the regions containing the SNPs were coding or non-coding. The concatenated SNPs were used for Bayesian statistics and for creating a NJ tree.

Three populations were found by the Bayesian clustering approach (Figure 35). One contained all 13 strains of *L. major*, the second Israeli and Palestinian strains of *L. tropica* and all six strains of *L. aethiopica,* and the third strains of *L. tropica* of various other origins. Interestingly, the strains of

L. aethiopica showed a great amount of shared membership to those of *L. major*, although closer related to *L. tropica*. Re-analysis of the latter two populations revealed substructures at lower hierarchical levels (Figure 35). When re-analysing only the 151 strains of *L. tropica*, a group of 69 strains from Israeli and Palestinian foci were separated from all other strains. All strains not falling into this sub-population "Israel/Palestine" were then re-analysed by Bayesian statistics. Three groups of strains were separated from all other strains at this hierarchical level, "Northern Galilee", "Kenya/Namibia" and "MoroccoB". These clusters have also been found by the MLMT analysis of this study, corroborating these results.

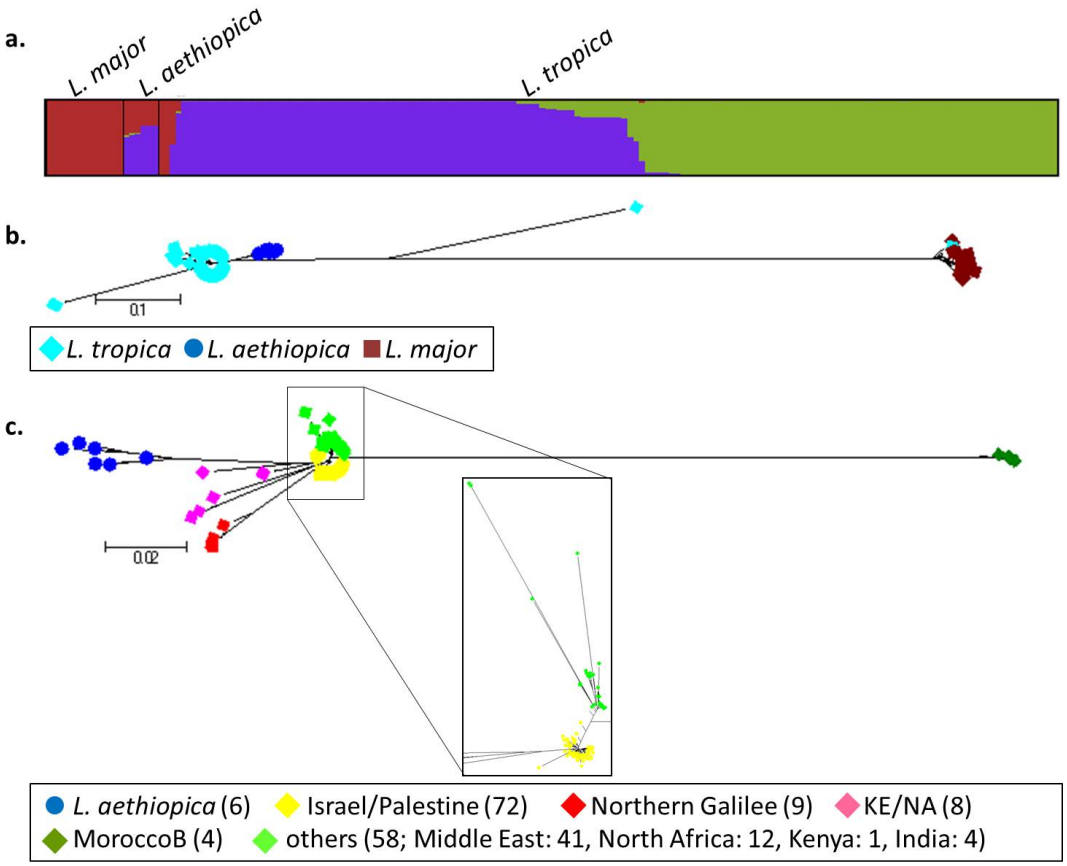

Figure 35. Results of the WGS approach based on the concatenated sequence of 44,020 SNPs. a. Bayesian statistics of 172 strains of *L. tropica*, *L. aethiopica* and *L. major*. b. Midpoint-rooted NJ tree including 172 strains of *L. tropica*, *L. aethiopica* and *L. major*. The colours indicate the previously identified species. c. Midpoint-rooted NJ tree including only the unambiguous strains of *L. tropica* and *L. aethiopica*. The populations indicated by different colours are those found by the Bayesian statistics. Their numbers are indicated in parentheses

The midpoint-rooted NJ trees constructed based on the concatenated SNP data clearly shows that the strains of *L. aethiopica* are closer related to the strains of *L. tropica* than to those of *L. major*, which is in congruence with the MLMT results (Figure 35a). To get a more precise view of the genetic distances among the strains of *L. tropica* and *L. aethiopica*, the strains, which clearly belonged to these two species, were re-analysed (Figure 35b). All clusters revealed by the Bayesian statistics were confirmed here. The strains of population "Israel/Palestine" formed a monophyletic group and, with the least genetic distance, the cluster "others" was found in close proximity. The strains of

L. aethiopica and the cluster “Northern Galilee” also appeared as monophyletic groups. The strains of the cluster “KE/NA” showed close genetic relationships, although not being completely monophyletic. A very long branch located the strains of “MoroccoB” in a position very remote from all other strains.

The WGS analysis revealed that two strains previously typed as *L. tropica* were in fact strains of *L. major* and probably misidentified earlier. More interestingly, it suggested a hybrid structure for strain MHOM/PS/2001/ISL593 which was located between *L. tropica*/*L. aethiopica* (membership coefficient = 0.6306) and *L. major* (0.3694) populations in both the Bayesian statistics and the NJ tree.

7.1.3 Conclusions on the population genetic structure of *L. tropica/L. aethiopica*

The MLMT approach revealed a largely geographical correlation of the population structure for the strains of *L. tropica* analysed in this study. Similar correlations were also found in studies analysing the population genetics of other *Leishmania* species [62, 63]. At the highest hierarchical level of the Bayesian statistics the strains of population “Israel/Palestine” were separated from all other strains of *L. tropica*. This population was also well-defined by the distance-based methods and the FCA. The remaining strains formed genetic entities only at subordinated levels, largely according to the strains´ geographical origins. Exceptions to this geographical correlation were the group of strains collected at the northern side of the Sea of Galilee, which seem to be more closely related to the African strains, and the group containing old strains from a great variety of origins all isolated in or before 1991, which formed a monophyletic group among the Asian strains. Also, the two Moroccan clusters were not in accordance to their geographical origins; although they were found among other African strains of *L. tropica*, both clusters contained strains from each of two foci.

Analysis of 44,020 concatenated SNPs identified by WGS confirmed the genetic entities “Israel/Palestine”, “*L. aethiopica*”, “Northern Galilee”, “Kenya/Namibia” and “MoroccoB” of the MLMT approach revealed in this study, while others, mainly those of the sub-population “Asia” found in the MLMT study were not exposed by the WGS approach.

Interestingly, the strains of *L. aethiopica* were very diverse. They were located on long branches in the NJ tree (Figure 17) and the network (Figure 18), within the also highly paraphyletic group of African strains of *L. tropica* in these two analyses. The FCA presented the strains of *L. aethiopica* clearly as one group with relatively large distances between the single strains, but quite distantly related to the strains of *L. tropica* (Figure 19).

The close relationship between *L. tropica* and *L. aethiopica* was clearly confirmed by the WGS-SNP approach. Based on these data one can speculate whether *L. aethiopica* is a distinct species or just part of an "*L. tropica* species complex". Whereas strains of *L. major* were clearly separated from *L. tropica*/*L. aethiopica,* the strains of *L. aethiopica* were more similar to the strains of *L. tropica* from Israel, Palestine and other regions (Figure 35), even the Moroccan strains of *L. tropica* showed a greater genetic distance. Since this WGS approach included a considerable amount of markers, it is, unlike the methods previously implemented for this species discrimination (chapter 7.1.1.5), statistically very well supported and thus, of great importance for addressing this question. In future studies, the time of divergence between *L. tropica* and *L. aethiopica,* and to *L. major,* should be estimated for elucidating their relationship in detail.

There are multiple explanations for the diversification of *Leishmania* parasites and their distribution and genetic evolution within newly founded foci. In the past and, even more today, tremendous human migration between different countries and even continents has facilitated the co-migration of leishmanial parasites. Travellers, refugees or pilgrims, who got infected with leishmanial parasites, can import their strains to other world-wide regions endemic to competent sand fly vectors. If the female of a competent vector species would take up the promastigotes during a blood meal, the parasite could get transmitted to a competent host, either human or animal. Those newly imported *Leishmania* strains could then propagate either clonally or by recombining with already existing *Leishmania* populations in this area. Rogers *et al*. investigated the reproduction of a population of *L. infantum* in Cukurova, Turkey and proposed an initial hybridisation between two phylogenetically distinct parental strains and subsequent clonal reproduction combined with rare recombination events [114]. Clonal expansion is probable in the case of the CL outbreak in Sanliurfa based on their genetic homogeneity. In the population "Israel/Palestine" the strain MHOM/EG/1990/LPN65 from the Sinai Peninsula might be a putative founder strain for this population, although this is highly speculative. The strains might have been introduced into this area and subsequently interbred with already existing strains. Prerequisites for this scenario of introduction into a new area, which is probably true for at least the strains from the CL outbreak in Sanliurfa, Turkey, and those strains forming the population "Israel/Palestine", are a competent sand fly vector and a mammalian reservoir host.

Since microsatellites are not able to analyse directed evolution and draw conclusions regarding the time of divergence between certain genetic groups, the WGS approach will help to complete the picture revealed by the MLMT analysis. Of course, the choice of markers has a great impact on the outcome of such an analysis. One important difference between microsatellite loci and the SNPs analysed in this WGS approach is that the former are located in intergenic non-coding regions and

the latter presumably in coding regions of the genome. The diversification of the strains in the WGS approach can also indicate functional differences between the clusters. Noteworthy, all clusters located on long branches involve different sand fly vector species and mammalian hosts in their transmission cycles.

The capability for adaptation to a new vector can be of great evolutionary advantage to the *Leishmania* parasites, thus posing a positive evolutionary selection. Such a scenario has been reported for the strains from a recently emerged focus at the northern side of the Sea of Galilee. They can be transmitted not only by *P. sergenti*, the predominant vector of *L. tropica* in Israel and most of its geographical range, but also by a new vector species, *P. arabicus* and the rock hyrax is their natural reservoir host [42]. In South and East Africa, both *L. tropica* and *L. aethiopica* have been isolated from additional sand fly species and animals [37, 38]. One possible explanation for the intra-focal diversity in Morocco as discussed in chapter 7.1.1.2 is the involvement of another vector species. Field studies, addressing the infection of endemic sand fly species with *L. tropica* in different foci and investigating the genetic diversity of the parasites can elucidate the role of adaptation to different vectors for the diversification of the parasites.

7.2 Evaluation of the Functional Cloning approach for the identification of factors contributing to the high variability of *L. tropica*

7.2.1 Functional Cloning as appropriate technique to identify genes responsible for a certain trait

Functional Cloning, also known as genetic complementation, is an approach which has been used in several studies in different species of *Leishmania*. It identified the genes involved in lipophosphoglycan synthesis, *LPG1* and *LPG2* in *L. donovani* [126, 127]. Functional Cloning has also been applied for the identification of genes conferring drug resistance [128]. For example, Coelho *et al.* identified an open-reading frame which encodes a probable ABC transporter protein playing a role in resistance of *L. major* to pentamidin [129] and Choudhury *et al.* found a gene in *L. infantum* mediating resistance to miltefosine and trivalent antimony (Sb^{III}) [130]. Moreover, Functional Cloning has also been applied successfully to cross-species analyses, which identified factors responsible for species-specific organ tropism [131] and virulence [132].

The technique has one convincing advantage over the classical reverse genetic strategies to analyse a gene´s function: It is unbiased. This implies that unknown or unexpected genes can be detected which would probably not have been included as candidate genes in a classical overexpression or gene replacement approach. Moreover, cosmids can be easily isolated by standard alkaline lysis and

the gene(s) responsible for the selected trait can be identified by partial sequence analysis followed by a database search.

Several smaller caveats should be considered when conducting a Functional Cloning assay. First, the use of selection antibiotics in the medium contributes to cellular stress and might interfere with the selection procedure. This can be avoided by not using the antibiotic during the screening procedure. Second, the genes of the gDNA of the acceptor strain are still in place and might interfere with the introduced gene(s) and complementing their function.

7.2.2 Functional Cloning to evaluate the role of the MMR system in the high genetic variability of *L. tropica*

L. tropica has shown high genetic variability in the MLMT analysis of this study, which was corroborated by the WGS approach discussed before. Two possible scenarios might explain the high genetic heterogeneity. High levels of recombination are rather unlikely, since sexual propagation is supposed to be a rare event in *Leishmania* parasites. An elevated mutation rate, possibly corroborated by an impaired MMR system might be responsible for this heterogeneity. An impaired MMR system allows replication errors to be left unrepaired and causes mutations in the daughter cells. To examine this hypothesis, a Functional Cloning approach was conducted using MNNG to cause mismatches in the genomic DNA, which are unrepairable and instead would lead to apoptosis of the respective cell, initiated by a functioning MMR system. An impaired MMR system can, however, be of advantage, when the parasite is in the need to adapting to environmental changes like new sand fly vectors or new hosts.

The Functional Cloning method was applied for the identification of factors in the leishmanial genome, which account for a low mutation tolerance. A cosmid library containing the DNA of a strain showing reduced growth when incubated with MNNG (hypothetically functioning MMR system) was used for the transfection of another strain showing normal growth with MNNG (hypothetically impaired MMR system). The clones were then screened for the restored phenotype.

However, several problems were faced throughout the steps of the screening for transfectants exhibiting the donor phenotype of low MNNG tolerance. The clones showed a wide range of MNNG tolerance levels, not only between different clones, but also between several experiments of the same clone (Table 19 and Table 20). Even the sub-clones originating from LRC-L1322/6-pcosTL-830/797 and 2874, which were determined as positive, showed considerable variance in their MNNG susceptibility levels (Table 21), although they were expected to be either of the two, sensitive or tolerant to MNNG. This made the screening for positive clones very difficult. A clear distinction between the two phenotypes, high and low MNNG tolerance, was almost impossible. Nevertheless,

the cosmids of those clones which showed a statistically significant difference in their MNNG tolerance compared to that of the acceptor strain LRC-L1322/6 were partially sequenced to identify their insert sequences. The BLAST search resulted in sequences much shorter than the expected size of 29,500-43,500 bp. One possible explanation for this phenomenon is that the insert ends could not be aligned to the genome of *L. tropica*, since it is not yet annotated and available. The best alignments of the BLAST search were found when comparing the insert sequences to the genome of the *L. major* reference strain MHOM/IL/1980/Friedlin. This species is closely related to *L. tropica*, but not identical. Possibly, the genome of *L. tropica* bears additional sequences at these sites. This question could be addressed when the whole genome sequence of *L. tropica* is available.

One should also keep in mind, that proteins which are able to transport the MNNG out of the cell could neutralize the effect of MMR genes of the insert. With the MNNG no longer being detrimental to the cell, the growth rate, and, thus, the survival rate of the culture incubated with and that incubated without MNNG would be identical. The clones possibly containing one or more MMR genes would not be detected by the screening in this analysis.

In this study, it was not possible to perform the selection experiments without the selection antibiotic G418, since the phenotype which was screened for in the transfectants was negative; those clones exhibiting a reduction in growth under MNNG selection pressure should be selected. Probably, the cells would have excluded the cosmid bearing a detrimental gene (the hypothetical MMR system genes) from the cell without the selection pressure of the antibiotic in the medium. The addition of an antibiotic reduces the overall growth rate of the cells. In this study, the presence or absence of G418 was a main difference between the culturing of the transfectant clones and the acceptor strain. Possibly, this might have influenced the difference between their growth rates independent from the cosmid´s insert sequence.

7.2.3 Outlook: additional approaches which could address the involvement of the MMR system in the genetic heterogeneity of *L. tropica*

Despite the various difficulties with the method, the hypothesis has not been disproven. The question, whether the high genetic variability of *L. tropica* is due to an impaired MMR system, still remains open. It should be addressed by different approaches.

The applied Complementation assay can be modified in a way that the gDNA is not inserted randomly into the acceptor strain, but certain MMR genes of the donor strain are cloned into the shuttle vector and used to transfect the acceptor strain to overexpress those genes and evaluating their influence on the MNNG tolerance. One can also imagine replacing the acceptor´s MMR genes by those retrieved from the donor´s gDNA. Possible candidate genes can be selected on the basis of

the available whole genome sequences of *L. tropica*. The increase or decrease of MNNG susceptibility can be used for testing the hypothesis of involvement of the MMR system in MNNG tolerance. Noteworthy, this modified complementation approach would no longer be unbiased and many genes possibly involved in mismatch repair would have to be analysed, making it labour-intensive.

With the availability of whole genome sequences it becomes possible to search for mutations in the MMR genes, which could be directly correlated to the MNNG tolerance of strains of *L. tropica*. Also, additional mutations throughout the whole genome can be detected and compared between MNNG-sensitive and MNNG-tolerant strains. One possibility is that the genes do not differ only in their nucleotide sequence but also in their copy number. This copy number variation (CNV) could also be correlated to the phenotype of the leishmanial strains.

Furthermore, the impairment of the MMR system might be not only due to mutations in the MMR genes at the DNA level. The mRNA levels of certain genes differing in MNNG-sensitive and MNNG-tolerant cells can also be important to the functioning of the MMR system. Since the leishmanial gene expression is characterized by the overall lack of transcription regulation and large portions of DNA are transcribed as unidirectional, polycistronic transcription units [133], the stability of the mRNA might play a role here. The mRNA levels could be determined by quantitative Real-Time PCR (RT-PCR) and correlated to the MNNG sensitivity.

7.3 Evaluation of the MLPA approach for the differentiation of *Phlebotomus spp.* caught in Israeli CL foci

In Israel and Palestine CL can be caused by three *Leishmania* species, *L. major*, *L. tropica* and *L. infantum*, which are transmitted by several sand fly species of which some are permissive for different leishmanial species while others are specific for just one [134, 135]. Thus, the spectrum of sand fly species can be indicative for the *Leishmania* species present in a certain focus, which is important for epidemiological studies. Moreover, the biting habits of the different sand fly species differ with regard to the time of the blood meal during the daytime, the preferred mammalian blood source and the preference of living inside or outside houses [136]. Thus, the determination of the sand fly species endemic to a certain area plays an important role in the choice of protective measures against them.

To investigate the epidemiology of sand fly populations, thousands of specimens are being caught each year [137, 138]. The species identification is conducted by microscopic examination by the scientific staff of the public health care centres and research institutions. This examination needs considerable skills and taxonomic expertise, is laborious and time-consuming. In addition, significant morphological similarity between some species may cause difficulties during the identification

process. DNA-based approaches have become increasingly useful for the rapid identification of various sand fly species. Many approaches have been introduced that use PCR-RFLP involving different gene targets and restriction enzymes [139-141]. Sequencing of a 700 bp fragment of the mitochondrial cytochrome oxidase gene subunit I (COI) discriminated Colombian sand fly species [142]. An identification approach based on the sand flies´ protein patterns by matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS) was established for those species present in the Middle East [143]. This technique requires very expensive equipment and is thus not feasible for many researchers and in the field. All of these molecular methods have the advantage that they are more precise and less error-prone since the human element of uncertainty is largely excluded. They have, however, the disadvantage that only single sand flies can be investigated. They do not allow differentiation of different species in pools of sand flies collected in field studies and are, thus, very labour- and cost-intensive. Therefore, there is an urgent need to develop a technique which is able to detect, identify and quantify sand flies of various species contained in pools, using unambiguous molecular markers.

MLPA has been applied successfully for different purposes and to a great variety of organisms ranging from viruses to humans. Approaches have been developed for targeting DNA and RNA, for quantitative and qualitative analyses, and for addressing epidemiologic and diagnostic questions. Only few of the numerous examples are given here. A multiplex MLPA approach has been developed which allows simultaneous detection of various SNPs thus enabling the discrimination of the species within the *M. tuberculosis* complex, the characterization of strains of *Mycobacterium tuberculosis,* the determination of different genetic lineages, the detection of a range of drug resistance markers, and the identification of the clinically most relevant non-tuberculous mycobacterial species [144]. MLPA was also applied for the screening of honeybee viruses, most of which are positive stranded RNA viruses. The assay developed by De Smet *et al.* detects ten different strains of honeybee viruses in one reaction and also determines whether these are actively replicating by selectively screening for either the positive or the negative strand of the RNA virus [145]. Moreover, MLPA approaches are used for molecular diagnosis of human genetic disorders. It allows quantifying genes and became the gold standard for the diagnosis of pathologies that are due to gene copy number variations [146]. It can also be applied to diagnose diseases that are caused by abnormal DNA methylation [146].

Because of the possibility of the MLPA to detect multiple SNPs simultaneously, we attempted to develop an assay based on SNPs in the 18S rRNA sequence of sand flies species. This assay was able to identify and discriminate six different sand fly species, *P. papatasi*, *P. syriacus*, *P. sergenti*, *P. tobbi*, *P. arabicus*, and *P. perfiliewi*. All fragment patterns detected are summarized in Figure 36 and Table 26. It also has the potential to identify five additional species, *P. fallax*, *P. simici*, *P. perniciosus*,

P. alexandri and *P. neglectus*, for which probes were included in the probe mix, but were not tested here. The probes were designed under the assumption that other species, such as *P. mascittii*, *P. notteghemae*, *P. argentipes*, *P. duboscqi*, *P. gigas*, *Lutzomyia sp.*, *Sergentomyia sp.*, *Warileya sp.* and *Brumptomyia sp.* are either not present in the study area or not caught with the traps used.

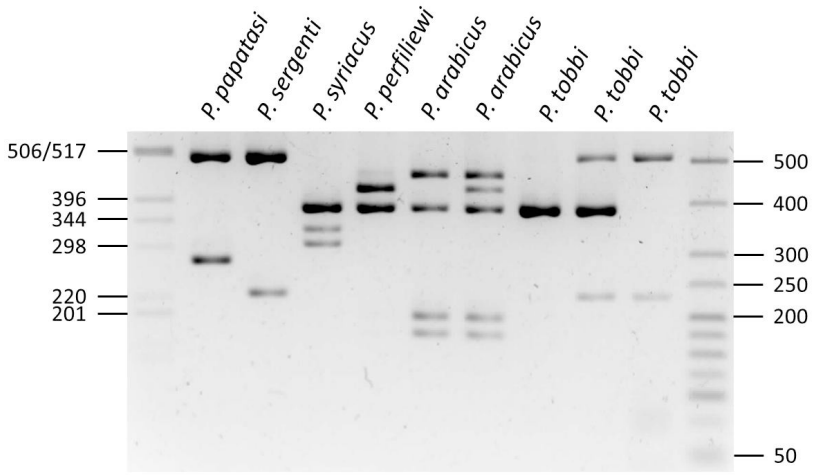

Figure 36. All detected fragment patterns of all analysed sand fly species. Ten µl of the MLPA reaction was loaded on the gel. Relevant fragments of the 1 kb DNA ladder and the hyperladder V are indicated.

Table 26. Comparison between all profiles detected by fragment analysis in all analysed sand flies (green) and the theoretical profiles (blue). The lengths of the fragments are given in bp. n; sample size, x; clear peak, (x); weak peak

		P. papatasi				*P. sergenti*		*P. syriacus*		*P. perfiliewi*		*P. arabicus*			*P. tobbi*			
					n=14		n=13		n=13		n=8		n=7	n=1		n=7	n=1	n=1
LENGTHS	144																	
	169												x	x				
	193												x	x				
	225					x	x										(x)	x
	250																	
	281	x	x		x													
	310							x	x									
	340							(x)	x									
	379			x				x	x	x	x		x	x	x	x		x
	422									x	x			x				
	460											x	x	x				
	507	x			x	x	x										x	x

Some of the probes were only amplified in one of the sand fly species and can be considered as species-specific (Table 26). The ~310 and ~340 bp fragments were both amplified in *P. syriacus* specimens only. A fragment of ~460 bp was unique to the species *P. arabicus* and was also the one fragment predicted for this species from the sequence alignment. Additionally, all specimens of *P. arabicus* had positive signals at ~169 and ~193 bp, which were not predicted for *P. arabicus,* but were only detected in individuals of this species. The fragment of ~281 bp was amplified exclusively for all individuals of *P. papatasi* in this study, while one of the aligned sequences lacked this SNP and

was positive for the probe of ~379 bp instead. This predicted fragment was positive for specimens of *P. syriacus*, *P. perfiliewi*, *P. arabicus* and *P. tobbi*.

Some of the fragment patterns were in conflict with those deduced from the sequenced samples at the NCBI database. The species *P. papatasi, P. sergenti, P. syriacus* and *P. perfiliewi* could be typed unambiguously; they showed consistent fragment patterns for all specimens tested (Table 26) and their profiles were in accordance to those deduced from the database. However, the samples of *P. papatasi* were identical to two of the three possible theoretical profiles, represented by eight sequences in the database. *P. arabicus* and *P. tobbi* showed intra-species variation. Seven of the eight specimens of *P. arabicus* showed an identical fragment profile, of which only the ~460 bp fragment had to be expected from the sequences submitted to NCBI. One individual fly of this species, *ara*07, had an additional peak at ~422 bp. The samples of *P. tobbi* were more diverse; three different fragment patterns were found among them. Seven of the nine included samples presented the profile predicted by the sequences of *P. tobbi* retrieved from the database. One sample had additional peaks at ~225 bp and ~507 bp and another lacked the predicted ~379 bp fragment and, thus, its fragment pattern was identical to that of all samples of *P. sergenti*. It remains to be established whether this can be attributed to an incorrect identification during microscopic examination, mix-up or contamination of DNA samples or to mutation at the probe binding site that would prevent probe hybridization. Specimens which showed fragments not expected based on the BLAST search, should be sequenced and aligned to the sand fly´s 18S rRNA sequence to determine whether they were the result of unspecific binding of the probe inside or outside the 18 S rRNA. Incorrect species determination of the sand fly from which the DNA sequence submitted to NCBI was retrieved, should also be considered.

The initial aim of this study was to develop a one-tube multiplex assay which can be used for differentiating different sand fly species in pools containing 10 to 20 individual flies. This would be of great importance regarding the huge amount of specimens being normally caught in field studies. While the simultaneous use of different probes in one reaction has been applied successfully, the pooling of sand flies seems to be problematic with this probe mix. Specific probes are lacking for some species and some species´ fragment patterns overlap. For example, if *P. perfiliewi* and *P. arabicus* were in the same tube, *P. perfiliewi* would not be detected since both fragments detected in this species, ~379 and ~422 bp, could be found in *P. arabicus* as well. The discrimination between *P. sergenti* and *P. tobbi* could also be problematic, since at least some strains of *P. tobbi* presented the same fragments typically found in *P. sergenti* (~225 and 507 bp). To solve this problem, additional probes should be designed which are species-specific. The SNP detected by the new probe does not necessarily be part of the 18S rRNA gene. It can be in any other sequence of the sand fly genome.

The use of single-copy genes might be, however inappropriate when included in the same assay as the 18S rRNA target because the fragment intensity of the latter could leave single gene fragments undetected.

Also, the quantification of sand flies of a certain species in a pool is not yet possible, since the intensities of the probe signals vary considerably. Only the species´ presence or absence can be detected in a pool of sand flies. This problem could be possibly overcome by including DNA samples isolated from a sand fly pool with known species composition in order to normalization for the varying peak intensities. Such a control pool would have to be included in each PCR and fragment analysis run. But even this normalization has weaknesses, since the fragment yield in the PCR reaction is highly sensitive to the exact amount of template DNA. Therefore, slight differences in template amount can have considerable effects on the quantitative outcome of the reaction.

7.3.1 Optimization of the MLPA approach

An MLPA approach for the species discrimination of sand flies endemic to Israel was developed in this study. Specific probes were found for *P. papatasi*, *P. syriacus* and *P. arabicus*. Since some species were lacking a specific probe, only one specimen at a time could be analysed.

Future studies should aim at the inclusion of probes targeting more sand fly species in the assay. Besides the probes for the six *Phlebotomus* species analysed here, probes for the detection of *P. fallax*, *P. simici*, *P. perniciosus*, *P. alexandri* and *P. neglectus* are already available, but DNA isolated from these species has not been tested yet. An MLPA approach as presented here should always be adjusted to the species composition in the studied area. Thus, for the application of this MLPA approach to other regions, the composition of the species-specific probes would have to be adapted to the species variety in the studied area by either adding new probes or, also, the exclusion of certain probes, which would detect sand flies not endemic to the region of interest. Of course, each region-specific assay has to be evaluated carefully for interference of their probes in order to avoid cross-specificity. A universal, *Phlebotomus* specific, probe should also be designed as positive control. Possible targets would be housekeeping genes, which sequences are highly conserved throughout all sand fly species. Also, in the 18S rRNA sequence there are regions which are conserved in all sand fly species.

For the multiplexing of this approach two basic approaches should be evaluated in the further development. Either additional specific probes for the missing species should be designed and tested or the sand fly DNA should be pooled and used for the detection by one probe only. In the latter scenario the quantification of the sand flies in the pool should be the long-term objective.

8 References

1. **Alvar J, Velez ID, Bern C, Herrero M, Desjeux P, Cano J, Jannin J, den Boer M: Leishmaniasis worldwide and global estimates of its incidence. *PLoS One* 2012, 7(5):e35671.**

2. **Control of the Leishmaniasis, Report of a meeting of the WHO Expert Committee on the Control of Leishmaniases, Geneva, 22-26 March 2010. In: *WHO Technical Report Series.* 2010: 186.**

3. **Mebrahtu Y, Lawyer P, Githure J, Were JB, Muigai R, Hendricks L, Leeuwenburg J, Koech D, Roberts C: Visceral leishmaniasis unresponsive to pentostam caused by *Leishmania tropica* in Kenya. *Am J Trop Med Hyg* 1989, 41(3):289-294.**

4. **Schnur LF, Nasereddin A, Eisenberger CL, Jaffe CL, El Fari M, Azmi K, Anders G, Killick-Kendrick M, Killick-Kendrick R, Dedet JP, Pratlong F, Kanaan M, Grossman T, Jacobson RL, Schonian G, Warburg A: Multifarious characterization of *Leishmania tropica* from a Judean desert focus, exposing intraspecific diversity and incriminating *Phlebotomus sergenti* as its vector. *Am J Trop Med Hyg* 2004, 70(4):364-372.**

5. **Aljeboori TI, Evans DA: *Leishmania spp*. in Iraq. Electrophoretic isoenzyme patterns. I. Visceral leishmaniasis. *Trans R Soc Trop Med Hyg* 1980, 74(2):169-177.**

6. **Guessous-Idrissi N, Berrag B, Riyad M, Sahibi H, Bichichi M, Rhalem A: Short report: *Leishmania tropica*: etiologic agent of a case of canine visceral leishmaniasis in northern Morocco. *Am J Trop Med Hyg* 1997, 57(2):172-173.**

7. **Lemrani M, Nejjar R, Pratlong F: A new *Leishmania tropica* zymodeme--causative agent of canine visceral leishmaniasis in northern Morocco. *Ann Trop Med Parasitol* 2002, 96(6):637-638.**

8. **Dereure J, Rioux JA, Gallego M, Perieres J, Pratlong F, Mahjour J, Saddiki H: *Leishmania tropica* in Morocco: infection in dogs. *Trans R Soc Trop Med Hyg* 1991, 85(5):595.**

9. **Rhajaoui M, Nasereddin A, Fellah H, Azmi K, Amarir F, Al-Jawabreh A, Ereqat S, Planer J, Abdeen Z: New clinico-epidemiologic profile of cutaneous leishmaniasis, Morocco. *Emerg Infect Dis* 2007, 13(9):1358-1360.**

10. **Rioux JA, Mahjour J, Gallego M, Demeure J, Perrieres J, Lahmrani A, Riera C, Saddiki A, Mouki B: Leishmaniose cutanée humaine à *Leishmania infantum* Mon-24 au Maroc. *Bulletin de la Société Française de Parasitologie* 1996, 14:5.**

11. **Sharma NL, Mahajan VK, Kanga A, Sood A, Katoch VM, Mauricio I, Singh CD, Parwan UC, Sharma VK, Sharma RC: Localized cutaneous leishmaniasis due to *Leishmania donovani* and *Leishmania tropica*: preliminary findings of the study of 161 new cases from a new endemic focus in Himachal Pradesh, India. *Am J Trop Med Hyg* 2005, 72(6):819-824.**

12. **Sharma NL, Chang KP, Loomba R, Arora S, Isabel M, Mahajan V, Negi A, Sharma R, Ready P, Mahadev PV, Kaushal K: An Overview of Satluj Valley focus of cutaneous and visceral Leishmaniasis. In: *Abstract book WorldLeish4: 03.-07.February 2009; Lucknow, India*; 2009: 95.**

13. **Khanra S, Bandopadhyay SK, Chakraborty P, Datta S, Mondal D, Chatterjee M, Naskar K, Roy S, Manna M: Characterization of the recent clinical isolates of Indian Kala-azar patients by RAPD-PCR method. *J Parasit Dis* 2011, 35(2):116-122.**

14. Khanra S, Datta S, Mondal D, Saha P, Bandopadhyay SK, Roy S, Manna M: RFLPs of ITS, ITS1 and hsp70 amplicons and sequencing of ITS1 of recent clinical isolates of Kala-azar from India and Bangladesh confirms the association of *L. tropica* with the disease. *Acta Trop* 2012, 124(3):229-234.

15. Belazzoug S, Lanotte G, Maazoun R, Pratlong F, Rioux JA: Un nouveau variant enzymatique de *Leishmania infantum*, Nicolle, 1908, agent de la leishmaniose cutanée du Nord de l'Algérie. *Annales de Parasitologie Humaine et Comparée* 1985, 60:1-3.

16. Gramiccia M, Ben-Ismail R, Gradoni L, Ben Rachid MS, Ben Said M: A *Leishmania infantum* enzymatic variant, causative agent of cutaneous leishmaniasis in north Tunisia. *Trans R Soc Trop Med Hyg* 1991, 85(3):370-371.

17. Gramiccia M, Gradoni L, Angelici MC: Epidemiology of Mediterranean leishmaniasis by *L. infantum*: isoenzyme and kDNA analysis for the identification of parasites from man, vectors and reservoirs. In: *Monograph on Leishmaniasis.* Edited by Hart DJ. New York: Plenum Press; 1988.

18. Gramiccia M, Gradoni L: Successful *in vitro* isolation and cultivation of Italian dermotropic strains of *Leishmania infantum sensu lato*. *Trans R Soc Trop Med Hyg* 1989, 83(1):76.

19. Baneth G, Zivotofsky D, Nachum-Biala Y, Yasur-Landau D, Botero AM: Mucocutaneous *Leishmania tropica* infection in a dog from a human cutaneous leishmaniasis focus. *Parasit Vectors* 2014, 7:118.

20. McMahon-Pratt D, Alexander J: Does the *Leishmania major* paradigm of pathogenesis and protection hold for New World cutaneous leishmaniases or the visceral disease? *Immunol Rev* 2004, 201:206-224.

21. Lipoldova M, Demant P: Genetic susceptibility to infectious disease: lessons from mouse models of leishmaniasis. *Nat Rev Genet* 2006, 7(4):294-305.

22. Sohrabi Y, Havelkova H, Kobets T, Sima M, Volkova V, Grekov I, Jarosikova T, Kurey I, Vojtiskova J, Svobodova M, Demant P, Lipoldova M: Mapping the Genes for Susceptibility and Response to *Leishmania tropica* in Mouse. *PLoS Negl Trop Dis* 2013, 7(7):e2282.

23. Zijlstra EE: PKDL and other dermal lesions in HIV co-infected patients with Leishmaniasis: review of clinical presentation in relation to immune responses. *PLoS Negl Trop Dis* 2014, 8(11):e3258.

24. Schnur LF, Chance ML, Ebert F, Thomas SC, Peters W: The biochemical and serological taxonomy of visceralizing *Leishmania*. *Ann Trop Med Parasitol* 1981, 75:131-144.

25. Magill AJ, Grogl M, Gasser RA, Jr., Sun W, Oster CN: Visceral infection caused by *Leishmania tropica* in veterans of Operation Desert Storm. *N Engl J Med* 1993, 328(19):1383-1387.

26. Sacks DL, Kenney RT, Kreutzer RD, Jaffe CL, Gupta AK, Sharma MC, Sinha SP, Neva FA, Saran R: Indian kala-azar caused by *Leishmania tropica*. *Lancet* 1995, 345(8955):959-961.

27. Jacobson RL, Eisenberger CL, Svobodova M, Baneth G, Sztern J, Carvalho J, Nasereddin A, El Fari M, Shalom U, Volf P, Votypka J, Dedet JP, Pratlong F, Schonian G, Schnur LF, Jaffe CL, Warburg A: Outbreak of cutaneous leishmaniasis in northern Israel. *J Infect Dis* 2003, 188(7):1065-1073.

28. Pratlong F, Dereure J, Ravel C, Lami P, Balard Y, Serres G, Lanotte G, Rioux JA, Dedet JP: Geographical distribution and epidemiological features of Old World cutaneous leishmaniasis foci, based on the isoenzyme analysis of 1048 strains. *Trop Med Int Health* 2009, 14(9):1071-1085.

29. Schonian G, Schnur L, el Fari M, Oskam L, Kolesnikov AA, Sokolowska-Kohler W, Presber W: Genetic heterogeneity in the species *Leishmania tropica* revealed by different PCR-based methods. *Trans R Soc Trop Med Hyg* 2001, 95(2):217-224.

30. Schwenkenbecher JM, Wirth T, Schnur LF, Jaffe CL, Schallig H, Al-Jawabreh A, Hamarsheh O, Azmi K, Pratlong F, Schonian G: Microsatellite analysis reveals genetic structure of *Leishmania tropica*. *Int J Parasitol* 2006, 36(2):237-246.

31. Talmi-Frank D, Jaffe CL, Nasereddin A, Warburg A, King R, Svobodova M, Peleg O, Baneth G: *Leishmania tropica* in rock hyraxes (*Procavia capensis*) in a focus of human cutaneous leishmaniasis. *Am J Trop Med Hyg* 2010, 82(5):814-818.

32. Talmi-Frank D, Kedem-Vaanunu N, King R, Bar-Gal GK, Edery N, Jaffe CL, Baneth G: *Leishmania tropica* infection in golden jackals and red foxes, Israel. *Emerg Infect Dis* 2010, 16(12):1973-1975.

33. Ntais P, Christodoulou V, Tsirigotakis N, Dokianakis E, Dedet JP, Pratlong F, Antoniou M: Will the introduction of *Leishmania tropica* MON-58, in the island of Crete, lead to the settlement and spread of this rare zymodeme? *Acta Trop* 2014, 132:125-130.

34. Hajjaran H, Mohebali M, Mamishi S, Vasigheh F, Oshaghi MA, Naddaf SR, Teimouri A, Edrissian GH, Zarei Z: Molecular identification and polymorphism determination of cutaneous and visceral leishmaniasis agents isolated from human and animal hosts in Iran. *Biomed Res Int* 2013, 2013:789326.

35. Toz SO, Culha G, Zeyrek FY, Ertabaklar H, Alkan MZ, Vardarli AT, Gunduz C, Ozbel Y: A real-time ITS1-PCR based method in the diagnosis and species identification of *Leishmania* parasite from human and dog clinical samples in Turkey. *PLoS Negl Trop Dis* 2013, 7(5):e2205.

36. Mohebali M, Malmasi A, Hajjaran H, Jamshidi S, Akhoundi B, Rezaei M, Janitabar S, Zarei H, Charehdar S: Disseminated Leishmaniasis Caused by *Leishmania tropica* in a Puppy from Karaj, Central Iran. *Iran J Parasitol* 2011, 6(2):69-73.

37. Massamba NN, Mutinga MJ, Kamau CC: Characterisation of *Leishmania* isolates from Laikipia District, Kenya. *Acta Trop* 1998, 71(3):293-303.

38. Lemma W, Erenso G, Gadisa E, Balkew M, Gebre-Michael T, Hailu A: A zoonotic focus of cutaneous leishmaniasis in Addis Ababa, Ethiopia. *Parasit Vectors* 2009, 2(1):60.

39. Vinitsky O, Ore L, Habiballa H, Cohen-Dar M: Geographic and epidemiologic analysis of the cutaneous Leishmaniasis outbreak in northern Israel, 2000-2003. *Isr Med Assoc J* 2010, 12(11):652-656.

40. Al-Jawabreh A, Schnur LF, Nasereddin A, Schwenkenbecher JM, Abdeen Z, Barghuthy F, Khanfar H, Presber W, Schonian G: The recent emergence of *Leishmania tropica* in Jericho (A'riha) and its environs, a classical focus of *L. major*. *Trop Med Int Health* 2004, 9(7):812-816.

41. **Azmi K, Schonian G, Nasereddin A, Schnur LF, Sawalha S, Hamarsheh O, Ereqat S, Amro A, Qaddomi SE, Abdeen Z: Epidemiological and clinical features of cutaneous leishmaniases in Jenin District, Palestine, including characterisation of the causative agents in clinical samples. *Trans R Soc Trop Med Hyg* 2012, 106(9):554-562.**

42. **Svobodova M, Votypka J, Peckova J, Dvorak V, Nasereddin A, Baneth G, Sztern J, Kravchenko V, Orr A, Meir D, Schnur LF, Volf P, Warburg A: Distinct transmission cycles of *Leishmania tropica* in 2 adjacent foci, Northern Israel. *Emerg Infect Dis* 2006, 12(12):1860-1868.**

43. **Amro A, Hamdi S, Lemrani M, Mouna I, Mohammed H, Mostafa S, Rhajaoui M, Hamarsheh O, Schonian G: Moroccan *Leishmania infantum*: Genetic Diversity and Population Structure as Revealed by Multi-Locus Microsatellite Typing. *PLoS One* 2013, 8(10):e77778.**

44. **Guilvard E, Rioux JA, Gallego M, Pratlong F, Mahjour J, Martinez-Ortega E, Dereure J, Saddiki A, Martini A: [*Leishmania tropica* in Morocco. III--The vector of *Phlebotomus sergenti*. Apropos of 89 isolates]. *Ann Parasitol Hum Comp* 1991, 66(3):96-99.**

45. **Guernaoui S, Boumezzough A, Pesson B, Pichon G: Entomological investigations in Chichaoua: an emerging epidemic focus of cutaneous leishmaniasis in Morocco. *J Med Entomol* 2005, 42(4):697-701.**

46. **Ostyn B, Gidwani K, Khanal B, Picado A, Chappuis F, Singh SP, Rijal S, Sundar S, Boelaert M: Incidence of symptomatic and asymptomatic *Leishmania donovani* infections in high-endemic foci in India and Nepal: a prospective study. *PLoS Negl Trop Dis* 2011, 5(10):e1284.**

47. **Sharma MID, Suri JC, Kalra NL, Mohan K, Swami PN: Epidemiological and entomological features of an outbreak of cutaneous leishmaniasis in Bikaner, Rajasthan, during 1971. *J Com Dis* 1973, 5:54-72.**

48. **Simi SM, Anish TS, Jyothi R, Vijayakumar K, Rekha RP, Nimmy R: Searching for Cutaneous Leishmaniasis in Tribals from Kerala, India. *J Glob Infect Dis* 2010, 2(2):95-100.**

49. **Aara N, Khandelwal K, Bumb RA, Mehta RD, Ghiya BC, Jakhar R, Dodd C, Salotra P, Satoskar AR: Clinco-epidemiologic study of cutaneous leishmaniasis in Bikaner, Rajasthan, India. *Am J Trop Med Hyg* 2013, 89(1):111-115.**

50. **Sharma MID, Suri JC, Kalra NL, Mohan K: Studies on cutaneous leishmaniasis in India III. Detection of a zoonotic focus of cutaneous leishmaniasis in Rajasthan. *J Com Dis* 1973, 5:149-153.**

51. **Mohan K, Suri JC: Studies on cutaneous leishmaniasis in India V. Isolation of *Leishmania tropica* from gerbils, sandflies and human. *J Com Dis* 1975, 7:353-357.**

52. **Kumar R, Bumb RA, Ansari NA, Mehta RD, Salotra P: Cutaneous leishmaniasis caused by *Leishmania tropica* in Bikaner, India: parasite identification and characterization using molecular and immunologic tools. *Am J Trop Med Hyg* 2007, 76(5):896-901.**

53. **Peters W, Chance ML, Chowdhury AB, Ghosh D, B., Nandy A, Kalra, J.L., Sanyal RK, Sharma MID, Srivastava L, Schnur LF: The identity of some stocks of *Leishmania* isolated in India. *Ann Trop Med Parasitol* 1981, 75:247-249.**

54. **Le Blancq SM, Peters W: *Leishmania* in the Old World: 2. Heterogeneity among *L. tropica* zymodemes. *Trans R Soc Trop Med Hyg* 1986, 80(1):113-119.**

55. **Karunaweera ND: *Leishmania donovani* causing cutaneous leishmaniasis in Sri Lanka: a wolf in sheep's clothing? *Trends Parasitol* 2009, 25(10):458-463.**

56. **Schlotterer C: Evolutionary dynamics of microsatellite DNA. *Chromosoma* 2000, 109(6):365-371.**

57. **Toth G, Gaspari Z, Jurka J: Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Res* 2000, 10(7):967-981.**

58. **Hedrick PW: Perspective: Highly variable loci and their interpretation in evolution and conservation. *Int J Org Evol* 1999, 53(2):313-318.**

59. **Selkoe KA, Toonen RJ: Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecol Lett* 2006, 9(5):615-629.**

60. **Schonian G, Kuhls K, Mauricio IL: Molecular approaches for a better understanding of the epidemiology and population genetics of *Leishmania*. *Parasitology* 2011, 138(4):405-425.**

61. **Kuhls K, Chicharro C, Canavate C, Cortes S, Campino L, Haralambous C, Soteriadou K, Pratlong F, Dedet JP, Mauricio I, Miles M, Schaar M, Ochsenreither S, Radtke OA, Schonian G: Differentiation and gene flow among European populations of *Leishmania infantum* MON-1. *PLoS Negl Trop Dis* 2008, 2(7):e261.**

62. **Kuhls K, Keilonat L, Ochsenreither S, Schaar M, Schweynoch C, Presber W, Schonian G: Multilocus microsatellite typing (MLMT) reveals genetically isolated populations between and within the main endemic regions of visceral leishmaniasis. *Microbes Infect* 2007, 9(3):334-343.**

63. **Al-Jawabreh A, Diezmann S, Muller M, Wirth T, Schnur LF, Strelkova MV, Kovalenko DA, Razakov SA, Schwenkenbecher J, Kuhls K, Schonian G: Identification of geographically distributed sub-populations of *Leishmania (Leishmania) major* by microsatellite analysis. *BMC Evol Biol* 2008, 8:183.**

64. **Schwenkenbecher JM, Frohlich C, Gehre F, Schnur LF, Schonian G: Evolution and conservation of microsatellite markers for *Leishmania tropica*. *Infect Genet Evol* 2004, 4(2):99-105.**

65. **Kebede N, Oghumu S, Worku A, Hailu A, Varikuti S, Satoskar AR: Multilocus microsatellite signature and identification of specific molecular markers for *Leishmania aethiopica*. *Parasit Vectors* 2013, 6:160.**

66. **Oddone R, Schweynoch C, Schonian G, de Sousa Cdos S, Cupolillo E, Espinosa D, Arevalo J, Noyes H, Mauricio I, Kuhls K: Development of a multilocus microsatellite typing approach for discriminating strains of *Leishmania* (*Viannia*) species. *J Clin Microbiol* 2009, 47(9):2818-2825.**

67. **Downing T, Stark O, Vanaerschot M, Imamura H, Sanders M, Decuypere S, de Doncker S, Maes I, Rijal S, Sundar S, Dujardin JC, Berriman M, Schonian G: Genome-wide SNP and microsatellite variation illuminate population-level epidemiology in the *Leishmania donovani* species complex. *Infect Genet Evol* 2012, 12(1):149-159.**

68. **Jamjoom MB, Ashford RW, Bates PA, Kemp SJ, Noyes HA: Polymorphic microsatellite repeats are not conserved between *Leishmania donovani* and *Leishmania major*. *Molecular Ecology Notes* 2002, 2(2):104-106.**

69. Jamjoom MB, Ashford RW, Bates PA, Kemp SJ, Noyes HA: Towards a standard battery of microsatellite markers for the analysis of the *Leishmania donovani* complex. *Ann Trop Med Parasitol* 2002, 96(3):265-270.

70. Kolodner R: Biochemistry and genetics of eukaryotic mismatch repair. *Genes Dev* 1996, 10(12):1433-1442.

71. Harfe BD, Jinks-Robertson S: DNA mismatch repair and genetic instability. *Annu Rev Genet* 2000, 34:359-399.

72. Habraken Y, Sung P, Prakash L, Prakash S: Binding of insertion/deletion DNA mismatches by the heterodimer of yeast mismatch repair proteins MSH2 and MSH3. *Curr Biol* 1996, 6(9):1185-1187.

73. Alani E: The *Saccharomyces cerevisiae* Msh2 and Msh6 proteins form a complex that specifically binds to duplex oligonucleotides containing mismatched DNA base pairs. *Mol Cell Biol* 1996, 16(10):5604-5615.

74. Ross-Macdonald P, Roeder GS: Mutation of a meiosis-specific *MutS* homolog decreases crossing over but not mismatch correction. *Cell* 1994, 79(6):1069-1080.

75. Hollingsworth NM, Ponte L, Halsey C: *MSH5*, a novel *MutS* homolog, facilitates meiotic reciprocal recombination between homologs in *Saccharomyces cerevisiae* but not mismatch repair. *Genes Dev* 1995, 9(14):1728-1739.

76. Wang TF, Kleckner N, Hunter N: Functional specificity of MutL homologs in yeast: evidence for three Mlh1-based heterocomplexes with distinct roles during meiosis in recombination and mismatch correction. *Proc Natl Acad Sci U S A* 1999, 96(24):13914-13919.

77. Bell JS, Harvey TI, Sims AM, McCulloch R: Characterization of components of the mismatch repair machinery in *Trypanosoma brucei*. *Mol Microbiol* 2004, 51(1):159-173.

78. Coelho AC, Leprohon P, Ouellette M: Generation of *Leishmania* hybrids by whole genomic DNA transformation. *PLoS Negl Trop Dis* 2012, 6(9):e1817.

79. Levinson G, Gutman GA: High frequencies of short frameshifts in poly-CA/TG tandem repeats borne by bacteriophage M13 in *Escherichia coli* K-12. *Nucleic Acids Res* 1987, 15(13):5323-5338.

80. Strand M, Prolla TA, Liskay RM, Petes TD: Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nature* 1993, 365(6443):274-276.

81. de Wind N, Dekker M, Berns A, Radman M, te Riele H: Inactivation of the mouse *Msh2* gene results in mismatch repair deficiency, methylation tolerance, hyperrecombination, and predisposition to cancer. *Cell* 1995, 82(2):321-330.

82. Schulz A: Identifizierung von Komponenten des Mismatch Repair Systems in *Leishmania major* und *Leishmania tropica*. *Diploma thesis.* Würzburg: Julius-Maximilians Unviversität Würzburg; 2005.

83. Karrenbrock M: Charakterisierung der *msh4*- und *msh5*-Gene bei *Leishmania major*, *Leishmania tropica* und *Leishmania turanica*. *Diploma thesis.* Berlin: Humboldt-Universität zu Berlin; 2006.

84. Schonian G, Nasereddin A, Dinse N, Schweynoch C, Schallig HD, Presber W, Jaffe CL: PCR diagnosis and characterization of *Leishmania* in local and imported clinical samples. *Diagn Microbiol Infect Dis* 2003, 47(1):349-358.

85. **Krayter L, Bumb RA, Azmi K, Wuttke J, Malik MD, Schnur LF, Salotra P, Schonian G: Multilocus microsatellite typing reveals a genetic relationship but, also, genetic differences between Indian strains of *Leishmania tropica* causing cutaneous leishmaniasis and those causing visceral leishmaniasis. *Parasit Vectors* 2014, 7:123.**

86. **Dieringer D, Schlötterer C: Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes* 2003, 3:167-169.**

87. **Glaubitz JC: CONVERT: A user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. *Molecular Ecology Notes* 2004, 4(2):309-310.**

88. **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S: MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 2011, 28(10):2731-2739.**

89. **Pritchard JK, Stephens M, Donnelly P: Inference of population structure using multilocus genotype data. *Genetics* 2000, 155(2):945-959.**

90. **Evanno G, Regnaut S, Goudet J: Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 2005, 14(8):2611-2620.**

91. **Huson DH, Bryant D: Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* 2006, 23(2):254-267.**

92. **Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F: GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. In. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France); 1996-2004.**

93. **Lewis PO, Zaykin D: Genetic Data Analysis: Computer program for the analysis of allelic data. In., 1.0 edn. http://hydrodictyon.eeb.uconn.edu/people/plewis/software.php; 2001.**

94. **Wright S: Genetical structure of populations. *Nature* 1950, 166(4215):247-249.**

95. **Clos J, Choudhury K: Functional cloning as a means to identify *Leishmania* genes involved in drug resistance. *Mini Rev Med Chem* 2006, 6(2):123-129.**

96. **Iyer RR, Pluciennik A, Burdett V, Modrich PL: DNA mismatch repair: functions and mechanisms. *Chem Rev* 2006, 106(2):302-323.**

97. **Wang JY, Edelmann W: Mismatch repair proteins as sensors of alkylation DNA damage. *Cancer Cell* 2006, 9(6):417-418.**

98. **Cox EC, Degnen GE, Scheppe ML: Mutator gene studies in *Escherichia coli*: the *mutS* gene. *Genetics* 1972, 72(4):551-567.**

99. **Rosario V: Cloning of naturally occurring mixed infections of malaria parasites. *Science* 1981, 212(4498):1037-1038.**

100. **Hall TA: BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In: *Nucl Acids Symp.* Oxford University Press; 1999: 95-98.**

101. **Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G: Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res* 2002, 30(12):e57.**

102. Ohnesorg T, Turbitt E, White SJ: The many faces of MLPA. *Methods Mol Biol* 2011, 687:193-205.

103. Krayter L, Alam MZ, Rhajaoui M, Schnur LF, Schonian G: Multilocus Microsatellite Typing reveals intra-focal genetic diversity among strains of *Leishmania tropica* in Chichaoua Province, Morocco. *Infect Genet Evol* 2014, 28:233-239.

104. Krayter L, Schnur L, Schonian G: The genetic relationship between *Leishmania aethiopica* and *Leishmania tropica* revealed by comparing microsatellite profiles. *PLOS ONE, under review* 2015.

105. Azmi A, Krayter L, Nasereddin A, Ereqat S, Schnur L, Al-Jawabreh A, Abdeen Z, Schonian G: Genetic Diversity and Population Structure of Palestinian and Israeli strains of Leishmania tropica as Revealed by Multi-Locus Microsatellite Typing *manuscript in preparation* 2015.

106. Boysen K, Schulz A, Schonian G: Identification of putative mismatch repair system components in *Leishmania tropica* and *Leishmania major*. In: *Abstract book WorldLeish4: 03.-07.February 2009; Lucknow, India*; 2009: 240.

107. Kelly JM, Das P, Tomás AM: An approach to functional complementation by introduction of large DNA fragments into *Trypanosoma cruzi* and *Leishmania donovani* using a cosmid shuttle vector. *Mol Biochem Parasitol* 1993, 65:51-62.

108. Azmi K, Schnur L, Schonian G, Nasereddin A, Pratlong F, El Baidouri F, Ravel C, Dedet JP, Ereqat S, Abdeen Z: Genetic, serological and biochemical characterization of *Leishmania tropica* from foci in northern Palestine and discovery of zymodeme MON-307. *Parasit Vectors* 2012, 5:121.

109. Olds N, Shoshani J: *Procavia capensis*. *Mammalian Species* 1982, 171:1-7.

110. Rhajaoui M, Sebti F, Fellah H, Alam MZ, Nasereddin A, Abbasi I, Schonian G: Identification of the causative agent of cutaneous leishmaniasis in Chichaoua province, Morocco. *Parasite* 2012, 19(1):81-84.

111. Ramaoui K, Guernaoui S, Boumezzough A: Entomological and epidemiological study of a new focus of cutaneous leishmaniasis in Morocco. *Parasitol Res* 2008, 103(4):859-863.

112. Pratlong F, Rioux JA, Dereure J, Mahjour J, Gallego M, Guilvard E, Lanotte G, Perieres J, Martini A, Saddiki A: [*Leishmania tropica* in Morocco. IV--Intrafocal enzyme diversity]. *Ann Parasitol Hum Comp* 1991, 66(3):100-104.

113. Ait-Oudhia K, Harrat Z, Benikhlef R, Dedet JP, Pratlong F: Canine *Leishmania infantum* enzymatic polymorphism: a review including 1023 strains of the Mediterranean area, with special reference to Algeria. *Acta Trop* 2011, 118(2):80-86.

114. Rogers MB, Downing T, Smith BA, Imamura H, Sanders M, Svobodova M, Volf P, Berriman M, Cotton JA, Smith DF: Genomic confirmation of hybridisation and recent inbreeding in a vector-isolated *Leishmania* population. *PLoS Genet* 2014, 10(1):e1004092.

115. Akopyants NS, Kimblin N, Secundino N, Patrick R, Peters N, Lawyer P, Dobson DE, Beverley SM, Sacks DL: Demonstration of genetic exchange during cyclical development of *Leishmania* in the sand fly vector. *Science* 2009, 324(5924):265-268.

116. Lukes J, Mauricio IL, Schonian G, Dujardin JC, Soteriadou K, Dedet JP, Kuhls K, Tintaya KW, Jirku M, Chocholova E, Haralambous C, Pratlong F, Obornik M, Horak A, Ayala FJ, Miles MA: Evolutionary and geographical history of the *Leishmania donovani* complex with a revision of current taxonomy. *Proc Natl Acad Sci U S A* 2007, 104(22):9375-9380.

117. Zeyrek FY, Korkmaz M, Ozbel Y: Serodiagnosis of anthroponotic cutaneous leishmaniasis (ACL) caused by Leishmania tropica in Sanliurfa Province, Turkey, where ACL Is highly endemic. *Clin Vaccine Immunol* 2007, 14(11):1409-1415.

118. Rioux JA, Lanotte G, Serres E, Pratlong F, Bastien P, Perieres J: Taxonomy of *Leishmania*. Use of isoenzymes. Suggestions for a new classification. *Ann Parasitol Hum Comp* 1990, 65(3):111-125.

119. Schonian G, Akuffo H, Lewin S, Maasho K, Nylen S, Pratlong F, Eisenberger CL, Schnur LF, Presber W: Genetic variability within the species *Leishmania aethiopica* does not correlate with clinical variations of cutaneous leishmaniasis. *Mol Biochem Parasitol* 2000, 106(2):239-248.

120. Fraga J, Montalvo AM, Van der Auwera G, Maes I, Dujardin JC, Requena JM: Evolution and species discrimination according to the *Leishmania* heat-shock protein 20 gene. *Infect Genet Evol* 2013, 18:229-237.

121. Montalvo AM, Fraga J, Monzote L, Montano I, De Doncker S, Dujardin JC, Van der Auwera G: Heat-shock protein 70 PCR-RFLP: a universal simple tool for *Leishmania* species discrimination in the New and Old World. *Parasitology* 2010, 137(8):1159-1168.

122. Nasereddin A, Bensoussan-Hermano E, Schonian G, Baneth G, Jaffe CL: Molecular diagnosis of Old World cutaneous leishmaniasis and species identification by use of a reverse line blot hybridization assay. *J Clin Microbiol* 2008, 46(9):2848-2855.

123. El Baidouri F, Diancourt L, Berry V, Chevenet F, Pratlong F, Marty P, Ravel C: Genetic structure and evolution of the *Leishmania* genus in Africa and Eurasia: what does MLSA tell us. *PLoS Negl Trop Dis* 2013, 7(6):e2255.

124. Talmi-Frank D, Nasereddin A, Schnur LF, Schonian G, Toz SO, Jaffe CL, Baneth G: Detection and identification of old world *Leishmania* by high resolution melt analysis. *PLoS Negl Trop Dis* 2010, 4(1):e581.

125. Kuru T, Janusz N, Gadisa E, Gedamu L, Aseffa A: *Leishmania aethiopica*: development of specific and sensitive PCR diagnostic test. *Exp Parasitol* 2011, 128(4):391-395.

126. Ryan KA, Garraway LA, Descoteaux A, Turco SJ, Beverley SM: Isolation of virulence genes directing surface glycosyl-phosphatidylinositol synthesis by functional complementation of *Leishmania*. *Proc Natl Acad Sci U S A* 1993, 90(18):8609-8613.

127. Descoteaux A, Luo Y, Turco SJ, Beverley SM: A specialized pathway affecting virulence glycoconjugates of *Leishmania*. *Science* 1995, 269(5232):1869-1872.

128. Cotrim PC, Garrity LK, Beverley SM: Isolation of genes mediating resistance to inhibitors of nucleoside and ergosterol metabolism in *Leishmania* by overexpression/selection. *J Biol Chem* 1999, 274(53):37723-37730.

129. Coelho AC, Beverley SM, Cotrim PC: Functional genetic identification of PRP1, an ABC transporter superfamily member conferring pentamidine resistance in *Leishmania major*. *Mol Biochem Parasitol* 2003, 130(2):83-90.

130. Choudhury K, Zander D, Kube M, Reinhardt R, Clos J: Identification of a *Leishmania infantum* gene mediating resistance to miltefosine and SbIII. *Int J Parasitol* 2008, 38(12):1411-1423.

131. Hoyer C, Mellenthin K, Schilhabel M, Platzer M, Clos J: Use of genetic complementation to identify gene(s) which specify species-specific organ tropism of *Leishmania*. *Med Microbiol Immunol* 2001, 190(1-2):43-46.

132. Zhang WW, Matlashewski G: In vivo selection for *Leishmania donovani* miniexon genes that increase virulence in *Leishmania major*. *Mol Microbiol* 2004, 54(4):1051-1062.

133. Myler PJ, Beverley SM, Cruz AK, Dobson DE, Ivens AC, McDonagh PD, Madhubala R, Martinez-Calvillo S, Ruiz JC, Saxena A, Sisk E, Sunkin SM, Worthey E, Yan S, Stuart KD: The *Leishmania* genome project: new insights into gene organization and function. *Med Microbiol Immunol* 2001, 190(1-2):9-12.

134. Killick-Kendrick R: Some epidemiological consequences of the evolutionary fit between *Leishmaniae* and their phlebotomine vectors. *Bull Soc Pathol Exot Filiales* 1985, 78(5 Pt 2):747-755.

135. Sacks DL: *Leishmania*-sand fly interactions controlling species-specific vector competence. *Cell Microbiol* 2001, 3(4):189-196.

136. Killick-Kendrick R: The biology and control of phlebotomine sand flies. *Clin Dermatol* 1999, 17(3):279-289.

137. Orshan L, Szekely D, Khalfa Z, Bitton S: Distribution and seasonality of *Phlebotomus* sand flies in cutaneous leishmaniasis foci, Judean Desert, Israel. *J Med Entomol* 2010, 47(3):319-328.

138. Arroub H, Hamdi S, Ajaoud M, Habbari K, Lemrani M: Epidemiologic study and molecular detection of *Leishmania* and sand fly species responsible of cutaneous leishmaniasis in Foum Jamaa (Azilal, Atlas of Morocco). *Acta Trop* 2013, 127(1):1-5.

139. Alam MS, Kato H, Fukushige M, Wagatsuma Y, Itoh M: Application of RFLP-PCR-Based Identification for Sand Fly Surveillance in an Area Endemic for Kala-Azar in Mymensingh, Bangladesh. *J Parasitol Res* 2012, 2012:467821.

140. Tiwary P, Kumar D, Rai M, Sundar S: PCR-RFLP based method for molecular differentiation of sand fly species *Phlebotomus argentipes*, *Phlebotomus papatasi*, and *Sergentomyia babu* found in India. *J Med Entomol* 2012, 49(6):1515-1518.

141. Minter LM, Yu T, Florin DA, Nukmal N, Brown GC, Zhou X: Molecular identification of sand flies (*Diptera*: *Psychodidae*) in eastern North America by using PCR-RFLP. *J Med Entomol* 2013, 50(4):920-924.

142. Contreras Gutierrez MA, Vivero RJ, Velez ID, Porter CH, Uribe S: DNA barcoding for the identification of sand fly species (*Diptera*, *Psychodidae*, *Phlebotominae*) in Colombia. *PLoS One* 2014, 9(1):e85496.

143. Dvorak V, Halada P, Hlavackova K, Dokianakis E, Antoniou M, Volf P: Identification of phlebotomine sand flies (*Diptera*: *Psychodidae*) by matrix-assisted laser desorption/ionization time of flight mass spectrometry. *Parasit Vectors* 2014, 7:21.

144. Bergval I, Sengstake S, Brankova N, Levterova V, Abadia E, Tadumaze N, Bablishvili N, Akhalaia M, Tuin K, Schuitema A, Panaiotov S, Bachiyska E, Kantardjiev T, de Zwaan R, Schurch A, van Soolingen D, van 't Hoog A, Cobelens F, Aspindzelashvili R, Sola C, Klatser P, Anthony R: Combined species identification, genotyping, and drug resistance detection of *Mycobacterium tuberculosis* cultures by MLPA on a bead-based array. *PLoS One* 2012, 7(8):e43240.

145. De Smet L, Ravoet J, de Miranda JR, Wenseleers T, Mueller MY, Moritz RF, de Graaf DC: BeeDoctor, a versatile MLPA-based diagnostic tool for screening bee viruses. *PLoS One* 2012, 7(10):e47953.

146. Stuppia L, Antonucci I, Palka G, Gatta V: Use of the MLPA Assay in the Molecular Diagnosis of Gene Copy Number Alterations in Human Genetic Diseases. *Int J Mol Sci* 2012, 13(3):3245-3276.

Selbständigkeitserklärung

Ich erkläre ausdrücklich, dass es sich bei der von mir eingereichten Dissertation mit dem Titel "Investigations on the epidemiology and diversity of *Leishmania tropica* and *L. aethiopica* and the differentiation of their sand fly vectors" um eine von mir selbständig und ohne fremde Hilfe verfasste Arbeit handelt.

Ich erkläre ausdrücklich, dass ich sämtliche in der oben genannten Arbeit verwendeten fremden Quellen, als solche kenntlich gemacht habe.

Datum: _____ Unterschrift: _____

List of publications

Krayter L, Bumb RA, Azmi K, Wuttke J, Malik MD, Schnur LF, Salotra P, Schönian G: **Multilocus microsatellite typing reveals a genetic relationship but, also, genetic differences between Indian strains of *Leishmania tropica* causing cutaneous leishmaniasis and those causing visceral leishmaniasis**. *Parasit Vectors* 2014, **7**:123

Krayter L, Alam MZ, Rhajaoui M, Schnur LF, Schönian G: **Multilocus Microsatellite Typing reveals intra-focal genetic diversity among strains of *Leishmania tropica* in Chichaoua Province, Morocco**. *Infect Genet Evol* 2014, **28**:233-239

Krayter L, Schnur LF, Schönian G: **The genetic relationship between *Leishmania aethiopica* and *Leishmania tropica* revealed by comparing microsatellite profiles.** PLOS ONE 2015, *under review*

Azmi K*, Krayter L*, Nasereddin A, Ereqat S, Schnur LF, Al-Jawabreh A, Abdeen Z, Schönian G: **Genetic Diversity and Population Structure of Palestinian and Israeli strains of *Leishmania tropica* as Revealed by Multi-Locus Microsatellite Typing.** 2015, *manuscript with co-authors* (* contributed equally to this work)

Krayter L, Tuin K, Orshan L, Schönian G: **Development of Multiplex ligation-dependent probe amplification (MLPA) probes discriminating phlebotomine sand fly species from the Middle East.** 2015, *manuscript in preparation*

Datum: _____ Unterschrift: _____

Acknowledgements

I thank Prof. Dr. Richard Lucius for the supervision of my thesis and for giving me the opportunity to apply for my doctoral degree at the Humboldt-University Berlin. Thanks also to all reviewers of this thesis for their effort.

Special thanks go to Dr. Gabriele Schönian for her supervision of my studies during all steps, from the first experiment to the final version of my thesis. She gave me the opportunity to conduct the experimental work in her lab and helped me with critical discussion of my work.

Prof. Dr. Wolfgang Presber not only provided new viewpoints to the work in progress, he also introduced other interesting research fields during our weekly lab meetings.

PD Dr. Joachim Clos and Dr. Andrea Nühs from the Bernhard-Nocht-Institut in Hamburg helped with every problem of the Functional Cloning assay.

The *Leishmania* strains included in this work were provided by Lionel F. Schnur (Hebrew University, Jerusalem, Israel), Poonam Salotra (National Institute of Pathology, New Delhi, India), Mohamed Rhajaoui (Institut National d'Hygiène, Rabat, Morocco) and Kifaya Azmi (Al-Quds University, Palestine).

The sand flies for the MLPA approach were provided by Laor Orshan (Ministry of Health, Jerusalem, Israel).

Tina Moser processed and provided the WGS-SNP data and discussed the results with me.

Anita Masic read the manuscript critically and provided me with helpful comments.

I also thank all lab members, especially Carola Schweynoch, Tina Moser and Olivia Stark as well as all guests from many countries for a nice working atmosphere and vivid discussions.

This work was conducted in the framework of the project "Emergence of Cutaneous Leishmaniasis due to *Leishmania tropica* and *L. major* in The Palestinian Authority and Israel ". I thank the Deutsche Forschungsgemeinschaft for funding this trilateral project.

Thanks to all Israeli and Palestinian collaborators for the introduction into their exciting and beautiful country and good scientific collaboration.

Thanks to all of you!