Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

Pós-Graduação em Biologia Celular e Molecular

MARIANA CÔRTES BOITÉ

Desenvolvimento e aplicação de marcadores moleculares para estudos taxonômicos, genéticos e epidemiológicos em *Leishmania* (*Viannia*).

Tese apresentada ao Instituto Oswaldo Cruz como parte dos requisitos para obtenção do título de Doutor em Ciências

Orientadora: Drª Elisa Cupolillo

RIO DE JANEIRO
ANO 2014

Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

Pós-Graduação em Biologia Celular e Molecular

AUTOR: MARIANA CÔRTES BOITÉ

DESENVOLVIMENTO E APLICAÇÃO DE MARCADORES MOLECULARES PARA ESTUDOS TAXONÔMICOS, GENÉTICOS E EPIDEMIOLÓGICOS EM *LEISHMANIA* (*VIANNIA*)

ORIENTADORA: Prof. Drª Elisa Cupolillo

Aprovada em: _____/_____/_____

EXAMINADORES:

Prof. Dr. Jeffrey Jon Shaw
Prof. Dr. Fernando Araújo Monteiro
Prof. Dr. Paulo Eduardo Martins Ribolla
Prof. Dr. Luiza de Oliveira Ramos Pereira
Prof. Dr. Gonzalo José Bello Bentacor

Rio de Janeiro, 13 de março de 2014

A todos aqueles que contribuíram e contribuem para o estudo das leishmanioses

AGRADECIMENTOS

Ao Steve, meu amor, companheiro e amigo, por todo incentivo;

Aos meus pais Carlos e Ana, de quem recebi tanto carinho e apoio imensuráveis em todas as situações e por toda minha vida, mas especialmente durante o período de execução deste trabalho;

Aos meus irmãos Luciano e Adriana, que me apoiaram de maneira incondicional, com carinho e paciência, abrindo suas casas para me receber sempre que precisei, oferecendo abrigo e aquela comidinha caseira que renova as forças, mesmo nos dias mais cansativos;

Aos meus cunhados (e irmãos!) Agenor e Thaissa, por sempre me receberem com tanto carinho em suas casas, emprestando seus computadores, impressoras, etc....oferecendo todo apoio logístico que precisei, e ainda com bom humor como brinde;

Aos meus tios Severino (tio Si), Jorginho e Maristela, grandes incentivadores desde o mestrado; à tia Miza, tia Marilda e tia Zelinha pela torcida e carinho constantes;

À Elisa Cupolillo, minha orientadora de todas as horas e situações. Paciente, mas com a firmeza que me ajudava a ser objetiva nos momentos de dúvidas. Este agradecimento ultrapassa a orientação acadêmica, passa pela amizade, pelos ensinamentos em diferentes esferas e, principalmente, pela confiança depositada em mim que me fez ser uma profissional mais madura.

Aos pesquisadores Renato Porrozzi e Patricia Cuervo, pelo apoio, orientação e amizade;

Ao pesquisador Alexandre Peixoto, que sempre esteve disposto a contribuir e ensinar oferecendo um bate papo científico estimulante e divertido;

À Barbara Neves dos Santos, amiga "faça chuva ou sol", amiga a -10ºC ou a 50ºC, longe ou perto... amiga sem fronteiras. Sem a sua ajuda durante o doutorado sanduiche tudo teria sido bem mais difícil! Além da amizade, sua colaboração e experiência foram essenciais neste trabalho.

Ao Gabriel Eduardo Melim Ferreira, amigo e indispensável colaborador, sempre disposto a ajudar e a ensinar com calma, paciência e bom humor;

À Mariel Marlow e professor Mario Steindel, pessoas com quem apreciei muito trabalhar;

À Amanda dos Santos Cavalcanti, amiga de longa data, com quem dividi as dúvidas e certezas profissionais;

Ao Marcos Augusto Trannin, por toda a gentileza e por ser esse querido aluno-colaborador-amigo;

À Fernanda Morgado, pela amizade calma e conselhos que me ajudaram tanto nesta fase final tão estressante;

À Lina Carrillo, por toda paciência e apoio;

À Juliana Bragazzi Cunha, amiga e conselheira incansável;

Ao pesquisador José Paulo Leite, pela amizade, conselhos, confiança e indicação que me trouxe até o LPL;

Ao Davi Coe Torres, responsável pelos primeiros treinamentos que recebi no LPL;

Ao colega bioinformata Diogo Tschoeke;

À Taíse Salgado de Oliveira, minha primeira aluna de iniciação científica, por toda ajuda no desenvolvimento deste trabalho;

À Caroline Batista, nossa caçulinha, cabeça-dura inteligente, "menina" que admiro e que sempre me deu uma ajuda especial com as isoenzimas;

Ao Israel Pinto, amigo que oferece aquela conversa boa de quem entende de roça;

Aos colegas do LPL: Andrés, Camila Lemos, Camila Braga, Cintia, Monica, Natalia, Priscila, Grazielle, Virgínia, Joana, Antônio Teva, Rosane, Eduardo, Ricardo, Selma e Carlos Henrique;

À Isabel Mauricio, ao professor Michael Miles e aos colegas da London School of Hygiene and Tropical Medicine, pela oportunidade e ensinamentos;

À Katrin Kuhls, minha *Doktor Mutter*, colaboradora e amiga, ao professor Marcus Fhrome pela confiança, e aos colegas de TH Wildau, Alemanha, por me acolherem;

À equipe da Plataforma de Sequenciamento PDTIS;

À secretaria de pós-graduação, em especial à Daniele Lobato;

À Luiza de Oliveira Ramos Pereira, por aceitar prontamente o convite para ser a revisora desta tese e realizar a tarefa com tanto carinho;

Muito obrigada.

"O ideal é viver como se fossemos imortais. Continuar nossa vida com os mesmos projetos e as mesmas ilusões com que começamos a viver. Isso é possível se fazemos o que gostamos, se nossa vida está dedicada a materializar uma vocação, o que significa que a recompensa obtida é o ato mesmo de exercitar essa vocação."

Mario Vargas Llosa

Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

DESENVOLVIMENTO E APLICAÇÃO DE MARCADORES MOLECULARES PARA ESTUDOS TAXONÔMICOS, GENÉTICOS E EPIDEMIOLÓGICOS EM *LEISHMANIA* (*VIANNIA*)

RESUMO

TESE DE DOUTORADO

Mariana Côrtes Boité

As leishmanioses englobam um largo espectro de doenças causadas pelo parasita do gênero *Leishmania*. É notável o número de espécies descritas, e, apesar de tratar-se de um gênero com alta variabilidade genética, a validade de algumas dessas espécies vêm sendo questionada. A associação entre *Leishmania* spp. e as diferentes formas clínicas sugere a participação do parasita na manifestação e no prognóstico da doença. Métodos moleculares estão sendo cada vez mais empregados para diagnóstico, estudos taxonômicos, filogenéticos e epidemiológicos envolvendo este parasita. Os métodos utilizados são, entretanto, diversos, comprometendo a reunião de dados e a comparação inter-laboratorial. Sendo assim, o desenvolvimento de um marcador único e robusto, que permita a troca de informações, enriquecerá o conhecimento sobre a diversidade fenotípica e molecular das leishmânias, atendendo tanto a demanda por uma revisão taxonômica do gênero, como para a elaboração de um sistema integrado de identificação e tipagem deste organismo, ponto essencial em estudos epidemiológicos. O principal objetivo deste trabalho foi desenvolver marcadores *multilocus* para *Leishmania* (*Viannia*) e avaliar sua aplicabilidade em estudos epidemiológicos. Os objetivos específicos incluíram: i) o desenvolvimento de um painel *multi locus sequence analisys* (MLSA) para cepas de *Leishmania* (*Viannia*) que circulam no Brasil que permita, concomitantemente, a identificação de isolados e a realização de estudos filogenéticos; ii) o estudo da estrutura da população de cepas de *Leishmania* (*Viannia*) quanto à clonalidade e ocorrência de eventos de recombinação pela determinação de perfis de microssatélites; iii) avaliação da aplicabilidade do painel MLSA como ferramenta epidemiológica para as leishmanioses; iv) a descrição de eventos moleculares em leishmânia que podem interferir nos resultados obtidos a partir de marcadores moleculares empregados. Após construção do painel MLSA e determinação de perfis de microssatélites (MLMT) foi possível comparar os dois marcadores. Os resultados a partir de MLMT definiram duas populações principais, uma com cepas de *L.* (*V.*) *guyanensis* da região Amazônia e outra apresentando as cepas de *L.* (*V.*) *braziliensis*, oriundas da costa Atlântica brasileira. Um terceiro grupo, muito heterogêneo, pôde ser definido com cepas de *L.* (*V.*) *braziliensis* da região norte e com as cepas das espécies *L.* (*V.*) *shawi*, *L.* (*V.*) *naiffi*, e *L.* (*V.*) *lainsoni*. Sinais de recombinação foram detectados no grupo formado por cepas identificadas como *L.* (*V.*) *guyanensis* e também naquele composto por cepas de *L.* (*V.*) *braziliensis* da região costeira. A recombinação pode ser uma das justificativas tanto para a grande diversidade encontrada como para a ausência de estruturação clara da população. O MLSA separou as espécies de acordo com a classificação prévia, exceto para *L.* (*V.*) *shawi* e também apontou para ocorrência de recombinação pelo padrão reticulado das redes construídas. Ao aplicar o MLSA em uma situação de surto em Santa Catarina, se detectou associação entre características epidemiológicas dos pacientes e os grupos MLSA, com separação de casos importados e autóctones, sinalizando para o potencial como marcador epidemiológico. A partir dos dados MLSA também foi possível realizar um estudo sobre a ocorrência de variação intra-cepa detectada em sequências de DNA de isolados clonados de *Leishmania*. Foi possível concluir que o painel MLSA construído e validado apresenta-se como boa alternativa para tipagem, estudos taxonômicos e epidemiológicos em *L.* (*Viannia*). Como tal pode representar tanto um substituto molecular para o *multilocus enzyme electroforesis* (MLEE), método-ouro para tipagem de *Leishmania*, quanto um marcador padrão a ser utilizado em revisão taxonômica do gênero. Mesmo com o advento do sequenciamento completo de genomas, a contribuição do MLSA ainda é relevante, e este se apresenta como potencial marcador epidemiológico, apesar da inclusão de novos genes ser necessária. O estudo com MLMT evidenciou que o mesmo é ferramenta de escolha para estudos de genética de populações em *L.* (*Viannia*), mas não para identificação e avaliação taxonômica do subgênero. Demonstrou-se ainda que a plasticidade genômica das leishmânias gera diversidade em tipos de sequencia de DNA e, portanto deve ser sempre considerada em estudos baseados em marcadores moleculares.

Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

DESENVOLVIMENTO E APLICAÇÃO DE MARCADORES MOLECULARES PARA ESTUDOS TAXONÔMICOS, GENÉTICOS E EPIDEMIOLÓGICOS EM *LEISHMANIA* (*VIANNIA*)

ABSTRACT

TESE DE DOUTORADO

Mariana Côrtes Boité

Leishmaniasis represent a broad spectrum of diseases caused by parasite of the genus *Leishmania*. The number of species described is remarkably high and the status of some has been questioned, although there is indeed a notable genetic variability within this genus. The association of *Leishmania* spp. and the different clinical forms suggests the involvement of the parasite in the prognosis and clinical outcome of the disease. Molecular methods have been often applied for diagnosis, studies of taxonomy, phylogeny and epidemiology. However, the approaches used are distinct, hampering comparisons between laboratories and data assembly. The development of a standard marker which allows exchange of information would contribute to the knowledge over the phenotypic and molecular diversity of *Leishmania*. It would also enable a taxonomic review as well as the establishment of a standardized and integrated typing system - essential in epidemiological studies. The main objective of this study was to develop *multi locus* markers for *Leishmania* (*Viannia*) and to evaluate its applicability in epidemiological studies. The specific objectives were: i) to develop a *multi locus sequence analyses* (MLSA) panel with *Leishmania* (*Viannia*) strains from Brazil that allows the species identification and phylogenetic inferences to be performed; ii) to execute a population structure study through microsatellites profiles (MLMT); iii) to evaluate the MLSA panel as an epidemiological tool; iv) to describe molecular events that may interfere in the results obtained by the molecular markers employed. After MLSA and MLMT, the results could be compared. MLMT defined two main populations, one comprising *L.* (*V.*) *guyanensis* and other *L.* (*V.*) *braziliensis* strains from the Atlantic coast; recombination signs were detected for both. A third group, quite heterogeneous, included *L.* (*V.*) *braziliensis* strains from northern Brazil and the other species *L.* (*V.*) *shawi*, *L.* (*V.*) *naiffi*, and *L.* (*V.*) *lainson*. Recombination may justify all the diversity observed and the lack of clear structuration. MLSA was able to differentiate species in agreement with previous classification, except for *L.* (*V.*) *shawi*. The reticulate pattern in the networks also pointed to recombination occurrence. After applying MLSA over strains isolated from an outbreak in Santa Catarina state, association between MLSA groups and epidemiological characteristics from the patients was detected, in a way that autochthonous and imported cases could be differentiated. MLSA data also allowed the description of intra-strain variation among DNA sequences from cloned *Leishmania* cultures. The results lead to the following conclusions: the MLSA panel developed and validated represents a good option for typing, taxonomy and epidemiology studies for *L.* (*Viannia*). It can be chosen as the molecular substitute for *multilocus enzyme electroforesis* (MLEE), the gold standard for *Leishmania* typing, and as the standard marker for a taxonomic review as well. Even considering the whole genome sequence approach, the contribution of MLSA is considered relevant, although more *loci* should be included. MLMT was revealed as the best approach for population genetics in *L.* (*Viannia*), but not for taxonomic inferences for this subgenus. It was also demonstrated that genomic plasticity in *Leishmania* raises intra-strain DNA sequence diversity, and therefore this aspect should be addressed in studies based on molecular markers.

Sumário

1. Introdução

As doenças tropicais negligenciadas (DTN) formam um grupo de mais de 17 doenças parasitárias que têm grande impacto na saúde pública, especialmente em países em desenvolvimento. Dentre as DTN, as leishmanioses têm destaque pelo número de pessoas acometidas (12 milhões) e sob o risco de contrair a doença (350 milhões), assim como pela distribuição geográfica mundial (Alvar et al., 2012). Trata-se de uma protozooze onde pelo menos 21 espécies de *Leishmania*, das mais de 30 já descritas, foram apontadas como infectantes para humanos (WHO, 2010). Os parasitos podem ser transmitidos por diferentes espécies (Coura, 2005) e populações (Ferreira et al., 2012) de vetor, insetos flebotomíneos, apresentando ciclos de transmissão complexos envolvendo hospedeiros / reservatórios de diferentes ordens de mamíferos (Coura, 2005). Ao acometerem humanos observa-se pleomorfismo clínico, podendo ocorrer respostas variadas ao tratamento, inclusive com antimoniais (Ait-Oudhia et al., 2011; Yasinzai et al., 2013), resultando em falha terapêutica, recidivas e surgimento de formas clínicas distintas da inicial. A severidade das manifestações clínicas em pessoas imunocompetentes e a resposta ao tratamento dependem, entre outros fatores, da espécie infectante. Há, por exemplo, espécies e cepas responsáveis por infecções autor resolutivas ou brandas, e que podem apresentar-se em simpatria com variantes mais virulentas (Arevalo et al., 2007; Chakravarty & Sundar, 2010; Maltezou, 2010), limitando o diagnóstico clínico por orientação geográfica a regiões onde diferenças na distribuição de formas clínicas já foram reconhecidas (de Brito et al., 2012; Guerbouj et al., 2001). É provável que a variabilidade existente para este parasita favoreça a sobrevivência em sistemas ecológicos diversos e determine a distribuição das formas clínicas observadas da doença (Calvopina et al., 2004; Lucas et al., 1998).

Ainda não há vacina segura e eficiente contra nenhuma forma de leishmaniose humana e o espectro e a eficiência de drogas são limitados (WHO, 2010). Uma das razões que provavelmente dificultam avanços nessas áreas encontra-se na diversidade do próprio parasito, reflexo de sua plasticidade genética. (Dujardin et al., 2007; Rogers et al., 2011). Sendo assim, o estudo da variabilidade genética do parasita aparece como um elemento importante para compreender os mecanismos que levam à complexidade clínica e epidemiológica. Diante disso, a identificação rápida da espécie ou cepa envolvida na infecção permitiria uma abordagem clínica diferenciada, poupando o paciente dos efeitos colaterais e distúrbios proporcionados pelo tratamento convencional. Desta forma, distinguir entre as espécies de *Leishmania* e compreender sua heterogeneidade podem contribuir para o

diagnóstico e prognóstico da doença, assim como para a tomada de decisões referentes ao tratamento e combate à doença.

Para realizar estudos que têm como objetivo a identificação de cepas e a análise da diversidade de um gênero de parasita são pontos decisivos a determinação dos marcadores para a classificação taxonômica, além da compreensão de aspectos moleculares do parasita.

1.1. Taxonomia de *Leishmania* e marcadores utilizados

A taxonomia reúne organismos em grupos definidos, promove nomenclatura apropriada para os grupos formados e sistematiza a identificação de organismos previamente desconhecidos. O termo tem sua origem na palavra grega *taxis* – que significa ordem, arranjo – e nomos (lei, ciência). Assim, a taxonomia de um organismo é a prática e a ciência da classificação, usando unidades taxonômicas, chamadas *taxa*, para analisar e formular uma estrutura hierárquica. Tipicamente, tal formulação obedece a relações de tipo e subtipo, que por definição devem apresentar as mesmas propriedades e comportamentos. Como se trata de uma abordagem criada pelo homem, e não uma regra natural deve ser constantemente revista e submetida à validação.

Para o estudo da diversidade de um parasito a classificação taxonômica é ferramenta primordial. Classicamente a taxonomia utiliza as origens e características comuns, estabelecendo a identificação, classificação, nomenclatura e as relações filogenéticas de um grupo de organismos. No caso das leishmânias, as variações morfológicas são pouco significativas entre as espécies, há simpatria e diferentes espécies podem causar a mesma forma clínica. Ainda assim, uma vez que estas eram as informações disponíveis, a classificação de *Leishmania* foi inicialmente baseada em critérios clínicos, bioecoepidemiológicos, tais como relação com vetores, distribuição geográfica, tropismo tissular, propriedades antigênicas e as manifestações clínicas. Contudo, análises imunológicas, bioquímicas e moleculares confirmaram que tais critérios eram muitas vezes insuficientes. Sendo assim, para uma classificação apropriada dos isolados deste gênero necessita-se de ferramentas moleculares e/ou bioquímicas.

A taxonomia numérica a partir dos dados obtidos por *multilocus enzyme electroforesis* (MLEE; ver item 1.2) e análise por marcadores moleculares, associadas aos conceitos que foram utilizados para descrição de cada espécie, resultou no sistema atual de classificação das leishmânias (Figura 1). Este grupo de parasitas encontra-se então classificado em três subgêneros, descritos em seguida.

O subgênero *L.* (*Sauroleishmania*) inclui apenas as espécies parasitas de répteis, que embora para alguns autores seja considerado como outro gênero (Lainson et al., 1987), estudos moleculares corroboraram sua classificação como subgênero (Fraga et al., 2010; Noyes et al., 1998). As espécies de *Leishmania* que infectam mamíferos são classificadas em dois subgêneros, *L.* (*Leishmania*) e *L.* (*Viannia*). O subgênero *L.* (*Viannia*) foi determinado em 1987, por Lainson e Shaw (Killick-Kendrick, 1987) para classificar os parasitas representantes da seção peripilaria. Definida com base em estudos de ciclo intravetorial de *Leishmania*, as espécies classificadas no subgênero *L.* (*Viannia*) apresentam desenvolvimento pobre em meio de cultura e lento em *hamsters* experimentalmente infectados; desenvolvem-se no intestino posterior do flebotomíneo aderidas à parede, na região do piloro, por isso chamada de seção peripilaria. Este grupo inclui apenas espécies encontradas nas Américas, e por isso correspondem à maioria daquelas referidas como do Novo Mundo. Já as espécies do subgênero *L.* (*Leishmania*) crescem facilmente em cultura (geralmente), provocam grandes lesões nodulares em *hamsters*, com metástase para as extremidades, e se desenvolvem no intestino médio e anterior de flebotomíneos; estão incluídas neste grupo espécies que circulam nas Américas, África, Ásia e Europa, razão pela qual são chamadas do Novo e do Velho Mundos, dependendo da região geográfica onde são observadas. O presente trabalho tem como foco apenas as espécies do subgênero *L.* (*Viannia*).

É notável o elevado número de espécies descritas (Coura, 2005; Rioux et al., 1990). Uma das razões pode ser a ausência de critérios claros para determinação de espécies em *Leishmania*. O conceito de espécie é motivo de debate em várias áreas. Independente do conceito criado será sempre uma classificação - de certa forma - arbitrária, e por isso dificilmente capaz de classificar a vida de forma perfeita. Ainda assim a investigação científica seria impossível sem essas categorias. O conceito biológico de Dobzhansky (Dobzhansky, 1935) que se refere à impossibilidade de dois indivíduos gerarem prole fértil não pode ser aplicado, considerando a estrutura preponderantemente clonal de muitos microrganismos, incluindo *Leishmania* (Tibayrenc et al., 1990; Tibayrenc & Ayala, 2012, 2013). Já o conceito filogenético (Davis M. Hillis, 1996) parece mais apropriado, onde espécie é considerada um grupo basal, com ancestralidade comum e distinguível de outros grupos. Para *Leishmania*, na prática, a definição continua arbitrária, misturando características extrínsecas e intrínsecas e ainda não há consenso sobre os critérios para uma revisão taxonômica. Adicionalmente, falta um método único de identificação e classificação, que determine os caracteres a serem considerados e que seja aceito entre os pesquisadores e padronizado entre os laboratórios para contribuir com esta questão.

A validade taxonômica de várias espécies de *Leishmania* tem sido constantemente discutida (Banuls et al., 2000; Fraga et al., 2010; Fraga et al., 2013), apesar da constatação de que realmente existe uma variabilidade elevada para este gênero (El Baidouri et al., 2013; Kuhls et al., 2011; Oddone et al., 2009). Por exemplo, uma análise molecular da região codificante do gene da proteína de choque térmico *hsp70* suporta a existência de oito espécies com relevância médica, dentro das 17 definidas por MLEE (Fraga et al., 2010). As principais espécies que têm seu *status* taxonômico questionado estão representadas na Figura 1. Há muitos anos *L*. (*L*.) *chagasi* já era apontada como espécie muito similar à *L*. (*L*.) *infantum* do Velho Mundo (Cupolillo et al., 1994; Mauricio et al., 2001). Recentemente ficou demonstrado que *L*. (*L*.) *chagasi* representa de fato cepas importadas de *L*. (*L*.) *infantum* para o Novo Mundo, e, portanto não é uma espécie válida (Leblois et al., 2011). *L*. (*V*.) *peruviana*, uma espécie próxima à *L*. (*V*.) *braziliensis*, já teve sua validade questionada, mas atualmente é aceita pela comunidade científica. *L*. (*V*.) *panamensis* e *L*. (*V*.) *shawi* foram apontadas como espécies muito próximas de *L*. (*V*.) *guyanensis*, sendo classificadas como parte do mesmo complexo por MLEE (Banuls et al., 2000). Com base nos resultados de ferramentas moleculares mais atuais, *L*. (*V*.) *panamensis* e *L*. (*V*.) *shawi* permanecem como pontos de debate (Cupolillo et al., 1994). Outra discussão iniciada mais recentemente sugere a classificação destes parasitas em subespécies (Fraga et al., 2010), uma proposta apresentada já no início da década de 70 (Lainson & Shaw, 1972).

Os dados reunidos atualmente indicam que uma revisão taxonômica deste gênero se mostra de fundamental importância (Schonian et al., 2010). Para que tal revisão seja possível, entretanto, deve ser estabelecido, primeiro, um critério para se determinar espécies em *Leishmania*, qual forma de classificação e unidade taxonômica será aceita, qual o método para tipagem de cepas, padronização deste método entre laboratórios e qual o poder de resolução ideal para diferenciar as principais espécies descritas. Além disso, um painel de cepas amplo o suficiente para representar a diversidade do gênero deve ser preparado para que uma revisão apropriada seja realizada.

1.2. *Multi locus enzyme electroforesis* - MLEE

Atualmente a eletroforese de isoenzimas, ou MLEE é um dos métodos preconizados para a identificação de isolados de *Leishmania* em nível de espécies. Trata-se de um ensaio bioquímico em que padrões variados de migração de várias proteínas (enzimas) podem ser observados e comparados aos padrões de cepas de referencia (cepas já caracterizadas representando espécies descritas e consideradas válidas). Os diferentes padrões são

consequência da codificação da enzima por alelos distintos do *locus* alvo, portanto refletem indiretamente diferenças na sequencia do DNA. A partir da comparação é possível identificar um isolado de *Leishmania* em nível específico e determinar em qual zimodema está classificado.

Apesar de outras metodologias moleculares terem sido desenvolvidas e frequentemente utilizadas, o MLEE ainda aparece como uma referência comum, como base de comparação para os novos métodos testados. Tal fato se deve ao MLEE ter sido o marcador escolhido para classificação de muitas espécies (Cupolillo et al., 1994; Rioux et al., 1990) e suas variantes (Belazzoug et al., 1988), e também para a realização de revisões taxonômicas do gênero. MLEE aparece ainda como a metodologia mais indicada para identificação de isolados de *Leishmania* pela Organização Mundial de Saúde. Esta metodologia permite a distinção das principais espécies de *Leishmania*, mesmo utilizando um número reduzido de *loci* para análise (Cupolillo et al., 1995). Provavelmente em razão dessas características, diversos estudos epidemiológicos fazem uso do MLEE, associado ou não a outros marcadores (Azeredo-Coutinho et al., 2007; Brito et al., 2009). Nas principais Coleções de *Leishmania* do mundo a classificação ainda se baseia nos resultados de isoenzimas, mesmo quando outros métodos são utilizados.

Mesmo diante de tantos pontos positivos envolvendo o MLEE como método de identificação, há atualmente uma busca por outro marcador, comum e padronizado entre os laboratórios, e que possa substituir o ensaio de isoenzimas (Schonian et al., 2011). A justificativa para esta busca baseia-se nas limitações apresentadas por esta metodologia. São elas: i) o painel de enzimas utilizado e o método de eletroforese empregado diferem entre os países, comprometendo a comparação direta dos resultados; ii) a menor resolução quando comparado aos métodos moleculares, uma vez que zimodemas indistinguíveis podem ser gerados por distintos genótipos; iii) diferenças na mobilidade eletroforética podem ser resultado de ocorrência de heterozigoto em apenas um nucleotídeo (Mauricio et al., 2006), e não de uma grande variação na sequencia do DNA codificante; iv) para realizar o ensaio é necessária uma grande quantidade de enzima, o que significa que o parasita isolado precisa ser cultivado até que uma concentração de células adequada seja obtida, e esse processo demanda tempo e recursos físico, técnico e financeiro.

Buscando superar as restrições da abordagem bioquímica, métodos moleculares vêm sendo desenvolvidos (Jamjoom et al., 2004; Oddone et al., 2009; Odiwuor et al., 2011; Schonian et al., 2011), mas na maioria das vezes usando ainda como base a classificação por

MLEE. Os resultados, no entanto, nem sempre corroboram aqueles obtidos por MLEE, exatamente pelas razões já descritas.

1.3. Marcadores moleculares

Com a popularização da PCR muitas abordagens moleculares vêm sendo utilizadas com o objetivo de identificar espécies e cepas de *Leishmania* (da Silva et al., 2010; Graca et al., 2012). Tais técnicas podem envolver o uso de enzimas de restrição em produtos amplificados de alvos específicos (*restriction fragment polimorphim* - RFLP), produto amplificado randomicamente a partir de todo o genoma (*amplified fragment length polymorphism* - AFLP) (Odiwuor et al., 2011), amplificação randômica a partir de polimorfismos na sequencia de DNA (*randomly amplified polimorphic* DNA – RAPD), ou podem ser baseadas na análise das sequencias de DNA de regiões codificantes e não-codificantes. Cada um desses métodos tem poder discriminatório específico, vantagens e limitações. Também podem ser considerados mais ou menos apropriados dependendo do tipo de estudo a ser desenvolvido e das perguntas a serem respondidas. Isso se deve às diferenças no nível de resolução e outras características do marcador, como por exemplo, se são ou não afetados por seleção natural (neutralidade). Schonian e colaboradores (Schonian et al., 2011) resumiu tais aspectos em uma revisão recente (Figura 2).

De maneira geral, as características desejáveis para marcadores utilizados na identificação de isolados, em estudos filogenéticos e de populações são a estabilidade após passagens *in vitro* e *in vivo* do patógeno, reprodutibilidade, e apresentar dados armazenáveis para comparação e transferência entre laboratórios.

Apesar das variadas contribuições das técnicas moleculares atuais, um método único, padronizado entre laboratórios ainda não existe. Tal fato impede reunião e comparação de dados, limitando correlações. Assim, busca-se, através de um marcador comum, reunir dados de diferentes regiões, gerados por laboratórios distintos, e cobrir assim maior parte da diversidade do gênero *Leishmania* em nível regional e mundial.

Uma vez estabelecido, um sistema molecular de identificação e tipagem de *Leishmania* integrado pode contribuir: i) para o desenvolvimento de técnicas de diagnóstico direto a partir de polimorfismos específicos; ii) para estudos de correlação entre genótipos com fenômenos biológicos e eco-epidemiológicos; iii) para avaliar a estrutura da população de *Leishmania* em relação à clonalidade e recombinação genética pela presença de heterozigotos e de alelos parentais na população estudada; e iv) promover a revisão taxonômica do gênero através de uma abordagem baseada em sequências de DNA.

1.3.1. Marcadores *multi locus* - MLST e MLSA

Análise de populações bacterianas por "tipagem por sequencias *multi locus*" (*multilocus sequence typing* - MLST) foi proposta em 1998 (Maiden et al., 1998) e tem funcionado como uma eficiente ferramenta no estudo de microrganismos diversos. O método se baseia na análise de sequencias de DNA de aproximadamente 700pb, de diferentes *loci*, que apresentam evolução conservada. Alelos são determinados para cada genótipo haplóide e então usados para estudos populacionais, rastreamento de epidemias e detecção de genótipos mosaicos causados por eventos de recombinação (Maiden et al., 1998; Maiden et al., 2013). Como MLEE, o MLST descrito para bactérias utiliza alelos como unidade de comparação, ao invés da sequencia nucleotídica. As sequencias, no entanto, podem ser utilizadas quando for apropriado. Para cada *locus* com uma sequencia de DNA única, um número é arbitrariamente escolhido para representar o alelo. A combinação de alelos encontrada para uma cepa gera um perfil denominado "tipo de sequencia" (*sequence type* - ST).

MLST já foi empregado para organismos não haploides (Debourgogne et al., 2010; Lauthier et al., 2012) inclusive para *Leishmania* (Mauricio et al., 2006; Tsukayama et al., 2009; Zemanova et al., 2007). Entretanto, especificamente em *Leishmania*, deve ser usado com maior cautela, levando em consideração aspectos como, por exemplo, presença de heterozigotos. Estes, apesar de oferecem um terceiro nível de informação ao processo de diferenciação de cepas podem ser também um fator complicador, caso sejam erroneamente detectados e interpretados (Cortes et al., 2012; Cupolillo et al., 1994; Odds & Jacobsen, 2008). Populações policlonais (Laurent et al., 2007) que podem ser selecionadas por pressão do meio (Pacheco et al., 1995) (hospedeiros envolvidos no ciclo de transmissão e meios de culturas), moisaiscismo cromossomal (Sterkers et al., 2011), e eventos de duplicação gênica (Rogers et al., 2011), são pontos relevantes que interferem na determinação de heterozigotos e na interpretação do resultado MLST em *Leishmania*. Outro cuidado envolve a avaliação dos marcadores a serem utilizados em MLST. Esta análise deve ser feita a partir de um painel de cepas coletadas em diferentes ecótopos e regiões, pois as localidades amostradas podem apresentar frequências alélicas distintas para um dado *locus* polimórfico, que podem não ser detectadas caso o número de indivíduos analisados seja reduzido (erro tipo II) (Davis M. Hillis, 1996). Cepas isoladas de formas clínicas e perfis isoenzimáticos variados também precisam ser consideradas. Além disso, é importante a validação os resultados obtidos por MLST pela comparação com diferentes metodologias.

A escolha e padronização dos genes representam uma etapa fundamental do processo de construção de MLST para *Leishmania*, porque uma abordagem genômica para identificar e diferenciar espécies baseia-se na diversidade dos genes alvo. Genes mais polimórficos podem contribuir para distinção de espécies próximas, ou mesmo cepas. Já os genes conservados contribuem para a estruturação em níveis taxonômicos mais elevados. Os genes de escolha podem variar de acordo com a proposta da pesquisa. Genes constitutivos, de seleção neutra, são preferíveis em estudos filogenéticos e de genética de populações. Por outro lado, genes não-neutros, sujeitos à pressão do meio, podem contribuir para identificação de traços epidemiológicos importantes, como virulência e resistência a drogas (Maiden et al., 2013; Schonian et al., 2011).

A combinação apropriada de *loci* em termos de resolução também pode variar dependendo da espécie ou subgênero em estudo. Se o marcador detecta muita variação, ele pode não representar bem as relações entre espécies mais distantes. Por outro lado, se ele detecta poucas variações, não vai explorar de maneira adequada as relações existentes entre espécies mais próximas. Já a combinação adequada pode ser filogeneticamente informativa e alcançar diferentes níveis taxonômicos (Davis M. Hillis, 1996).

O MLST permite que análises sejam realizadas tanto a partir do perfil de alelos encontrados, utilizando estes como unidade de comparação e não considerando o grau de variação nucleotídica entre as sequencias de DNA, ou fazendo uso de cada polimorfismo encontrado entre elas como informação. A possibilidade de gerar os dados das duas formas permite que inferências filogenéticas e epidemiológicas sejam feitas em conjunto (Odds & Jacobsen, 2008). Isso porque a determinação do tipo de sequencia permite o agrupamento de cepas com geração de complexos clonais favorecendo estudos de correlação com origem geográfica, forma clínica, etc. Já as variações nas sequencias nucleotídicas fornecem informações de relação evolutiva entre as cepas e os grupos, ao invés de apenas semelhança.

Até o presente, diversos estudos foram realizados com *Leishmania* seguindo o racional do MLST (El Baidouri et al., 2013; Mauricio et al., 2006; Tsukayama et al., 2009; Zemanova et al., 2007). Trabalho com cepas de *Leishmania* que circulam na China demonstrou que a história evolutiva desse parasita é mais complexa do que previamente descrito para aquele país (Zhang et al., 2013). Ainda, levantou a hipótese que a origem das cepas pode estar mais relacionada aos achados filogenéticos do que à forma clínica descrita. Outro estudo com 222 cepas isoladas no continente africano e eurásia permitiu que inferências taxonômicas e filogenéticas importantes sobre o gênero fossem feitas (El Baidouri et al., 2013). Os resultados corroboraram a classificação por MLEE, exceto para as espécies responsáveis pela

leishmaniose visceral (*L*. (*L*.) *donovani*, *L*. (*L*.) *archibaldi* e *L*. (*L*.) *infantum*). Com cepas de *L*. (*Viannia*) outro estudo foi desenvolvido, porém apenas com objetivo diagnóstico (Tsukayama et al., 2009).

Apesar dos diferentes trabalhos publicados, um sistema de tipagem propriamente ainda não existe para *Leishmania*. Isso se deve ao fato de, na maioria dos estudos, os marcadores utilizados e região sequenciada não serem os mesmos; quando os marcadores coincidem, os tipos de sequencia encontrados muitas vezes não são comparados para estabelecimento de um painel único. Sendo assim, a proposta de um sistema comum entre laboratórios ainda não foi alcançada. Pela falta de consenso, o termo mais apropriado para analise *multi locus* em *Leishmania* seria o MLSA, ou *multi locus sequence analysis* e não MLST, já que o sistema de tipagem propriamente ainda não existe.

Trabalhos apontam para o fato de o MLST corroborar a classificação das principais espécies de *Leishmania* caracterizadas por MLEE, mas com maior resolução, diferenciando cepas de mesmo zimodema (El Baidouri et al., 2013; Tsukayama et al., 2009), indicando que este marcador pode ser uma opção válida tanto para padronização da identificação de isolados quanto para a revisão taxonômica do gênero.

1.3.2. *Multi locus* microssatélite - MLMT

Sequências microssatélites, ou *short tandem repeats* (STRs) são repetições de 1-6 nucleotídoes encontradas em genomas de eucariotos e procariotos. Podem estar presentes em regiões codificantes (mais raramente) ou não-codificantes, e representam fatores que geram e ajudam a manter variabilidade genética (Toth et al., 2000). De forma geral, apresentam elevados níveis de polimorfismo. A taxa de mutação nesses *loci* é alta, assim como pode ser o número de alelos por *locus*. Os elevados níveis de polimorfismo podem ser devido às características moleculares do processo de replicação do DNA nessas regiões. Acredita-se que a DNA polimerase pode passar por um processo de "deslizamento" no qual adiciona unidades de repetição, e que esse processo é muito mais comum em regiões repetitivas menores. Outra possibilidade para explicar os polimorfismos de microssatélites seria a formação de "harpins" que levariam a perda de unidades repetitivas durante a replicação e essa seria a tendência para as regiões com maior número de repetições (Lai et al., 2003).

Os microssatélites estão randomicamente distribuídos, e em *Leishmania* já foram descritos em todos os cromossomos de diferentes cepas (Rossi et al., 1994). A variação no tamanho das sequências dos microssatélites é resultado da perda ou ganho de unidades de repetição, e pode ser detectada após amplificação do segmento com iniciadores específicos,

que flanqueiam a região repetitiva. A informação gerada pelos diferentes tamanhos pode ser usada para inferências diversas, como genética de populações e epidemiologia molecular (Schonian et al., 2011).

Como marcador molecular, a tipagem a partir de regiões microssatélites (*multi locus microssatélite typing* – MLMT) combina as vantagens do alto poder discriminatório e da codominância quando comparado ao MLEE, RAPD e PCR-RFLP (Schonian et al., 2011). Já foi aplicado em diversos organismos, como outros protozoários, fungos e insetos (Barnabe et al., 2011; L'Ollivier et al., 2012; Santos et al., 2013). Por apresentarem alta taxa de mutação, quando comparado a outras regiões do DNA, os microssatélites são particularmente interessantes para estudos de organismos similares (Schonian et al., 2008). Entretanto, exatamente pelos microssatélites apresentarem uma evolução rápida, os estudos utilizando MLMT mostraram que essa abordagem parece não ser indicada como técnica ouro para identificação das espécies de *Leishmania*, sendo que o conjunto de marcadores desenvolvidos para uma espécie comumente funcionam apenas nas espécies mais próximas.

Após a descrição dos microssatélites em *Leishmania* estudos buscaram aplicar o conhecimento para a diversidade do gênero (Rossi et al., 1994). Contudo, as análises dos *loci* escolhidos mostraram um reduzido número de alelos, o que limitava seu poder discriminativo. Posteriormente, foram identificadas duas regiões repetitivas com resolução suficiente para separar alguns complexos de espécies de *Leishmania* e até espécies dentro do complexo. Posteriormente, um painel de microssatélites foi desenvolvido e apresentado como um marcador polimórfico com capacidade de discriminar entre espécies do subgênero *L.* (*Viannia*) (Russell et al., 1999). Desde então, diferentes regiões de microssatélites já foram utilizadas na caracterização de cepas de várias espécies de *Leishmania* com apropriado poder discriminatório. Para as espécies do subgênero *L.* (*Viannia*), um painel de microssatélites desenvolvido permitiu a discriminação em nível intraespecífico, com importantes contribuições para a epidemiologia e genética de populações dos parasitas deste subgênero (Bulle et al., 2002; Jamjoom et al., 2002; Oddone et al., 2009; Rougeron et al., 2008). Trabalhos recentes com microssatélites em cepas de *L.* (*L.*) *infantum* que circulam no Brasil (Ferreira et al., 2012; Motoie et al., 2013) detectaram ocorrência de populações distintas, que podem estar associadas a padrões eco-epidemiológicos da leishmaniose visceral no país.

Desta forma, o MLMT representa uma ferramenta apropriada para o estudo de genética de populações em *Leishmania*, e por isso pode permitir a identificação de pontos importantes na dinâmica de transmissão e dispersão deste parasita. Porém, seu alto

polimorfismo não faz deste marcador um bom método para identificação de isolados ou inferências taxonômicas.

1.4. Leishmaniose tegumentar no Brasil e utilização de marcadores moleculares para estudos epidemiológicos e da diversidade do agente

A leishmaniose tegumentar americana (LTA) apresenta-se em fase de expansão geográfica. No Brasil, é endêmica em todos os estados e uma média anual de 27.250 casos tem sido registrada entre 1990-2010. Devido à variedade de formas clínicas observadas, a LTA é melhor apresentada como um complexo de doenças (Marlow et al., 2013), e não apenas como uma única patologia. Os agentes causadores são principalmente espécies do subgênero *L.* (*Viannia*): *L.* (*V.*) *braziliensis, L.* (*V.*) *guyanensis, L.* (*V.*) *lainsoni, L.* (*V.*) *naiffi, L.* (*V.*) *shawi e L.* (*V.*) *lindenbergi*; e *L.* (*Leishmania*) *amazonensis* como única não pertencente ao mesmo subgênero (Grimaldi & Tesh, 1993). Cada uma dessas espécies tem um ciclo de transmissão distinto, com diferentes espécies de vetor e de animal reservatório, e com variada distribuição geográfica (Grimaldi et al., 1989). A diversidade genética observada neste grupo de parasitas é relatada em diversos estudos (Boite et al., 2012; Cupolillo et al., 1997; Cupolillo et al., 1998) e pode estar relacionada aos diferentes ecótopos onde circulam os parasitas. Tal diversidade pode também explicar as variadas formas clínicas e diferentes respostas ao tratamento observadas na LTA em pacientes imunocompetentes, dependendo da espécie e cepa envolvida na infecção. Mas como ainda não há marcadores para cepas mais virulentas, e muitas vezes a identificação da espécie de *Leishmania* não ocorre, a decisão do tratamento é baseada na experiência prática regional do médico (McGwire & Satoskar, 2014).

Leishmania (*V.*) *braziliensis* é a espécie apontada como o principal agente causador da leishmaniose tegumentar no Brasil, e com a mais ampla distribuição pelo território nacional (Grimaldi et al., 1989). Pode causar as formas clínicas cutânea, mucosa ou disseminada; pacientes imunocompetentes infectados com esta espécie podem apresentar diferentes respostas ao tratamento (Eichner et al., 2013). A complexidade clínica pode variar, entre outros fatores, em função da cepa ou população de parasitos envolvidos na infecção (Baptista et al., 2009; Baptista et al., 2012). Sua diversidade intraespecífica já foi descrita no Brasil (Cupolillo et al., 1997; Cupolillo et al., 2003) - mesmo dentro de um único estado (de Brito et al., 2012; Schriefer et al., 2009). O principal reservatório para esta espécie ainda não foi determinado e diferentes espécies do inseto vetor estão envolvidas na transmissão do parasita (Coura, 2005).

A distribuição geográfica localizada das diferentes formas clínicas causada por *L.* (*V.*) *braziliensis*, como por exemplo, casos de leishmaniose muco-cutânea (LMC) e disseminada na Bahia, é indício de que grupos de genótipos (que podem ser os responsáveis pelos fenótipos) contribuam para a ocorrência das formas clínicas da LTA (Schriefer et al., 2009). Esse indício sustenta o uso de marcadores moleculares na busca de correlação entre genótipos e clínica. É possível que os diferentes ciclos de transmissão existentes para *L.* (*V.*) *braziliensis* colaborem para a geração e manutenção da diversidade intraespecífica (Cupolillo et al., 2003). Portanto, esses ciclos, uma vez caracterizados, também podem ser correlacionados a grupos de genótipos.

Tais dados apontam para a importância em compreender a variabilidade apresentada por *L.* (*V.*) *braziliensis* a fim de contribuir com prognóstico da leishmaniose tegumentar no Brasil. Através de estudos de genética de populações é possível buscar correlação entre populações específicas desta espécie e informações eco-epidemiológicas. Neste contexto, resultados prévios obtidos por MLSA (Alam et al., 2014) e MLMT (Oddone et al., 2009), inclusive com isolados de *Leishmania* (*Viannia*), demonstram a habilidade em discriminar grupos de genótipos desses marcadores. Indicam, portanto, que ambos apresentam potencial para aplicação em tal proposta, podendo contribuir para o prognóstico e controle da doença causada por *L.* (*V.*) *braziliensis*.

1.5. Aspectos moleculares e reprodução das leishmânias

As leishmânias apresentam peculiaridades quanto aos mecanismos de reprodução, regulação da expressão gênica e constituição cariotípica (Sterkers et al., 2011) que são estudados e questionados até os dias de hoje. Esses aspectos são de grande relevância em virtualmente todos os estudos sobre as leishmanioses, pois podem determinar diferenças fenotípicas, como por exemplo, virulência e resistência ao tratamento, observadas entre espécies próximas e cepas de uma mesma espécie. No caso de estudos envolvendo o desenvolvimento de marcadores taxonômicos moleculares, estudos filogenéticos e de genética de populações, tais aspectos precisam ser considerados, pois são determinantes para os resultados encontrados. O impacto nesses estudos será discutido no Capítulo IV.

O genoma de *Leishmania* sp. apresenta elevada sintenia entre as espécies (conservação da ordem de *loci* nos cromossomos) e a comparação dos diferentes genomas mostrou reduzido número de genes espécie-específicos (Peacock et al., 2007). Apesar desta informação, as diferenças fenotípicas entre as espécies são marcantes. O genoma de *Leishmania* é, em sua maior parte, constitutivamente expresso, e parte da regulação da

expressão gênica ocorre em níveis traducionais e pós-traducionais (Kazemi, 2011). Outro aspecto marcante em *Leishmania* e que tem impacto direto sobre os estudos com marcadores moleculares é seu cariótipo. A discussão sobre a ploidia deste parasita vem oscilando entre diploidia e aneuploidia nos últimos 20 anos, tendo como base estudos utilizando *pulsed filed electroforesis* (PFGE) e knock-out (Bastien et al., 1992; Dujardin et al., 2007; Scholler et al., 1986). Mais recentemente, o estado aneuploide foi mais bem descrito através do sequenciamento completo do genoma e de hibridização fluorescente *in vitro* (FISH) (Rogers et al., 2011; Sterkers et al., 2011). Durante a mitose uma distribuição assimétrica dos cromossomos, ou duplicação irregular dos mesmos, leva a formação de células filhas com diferentes cariótipos. Este modelo foi chamado de mosaicismo aneuploide (Sterkers et al., 2012). Os mesmos estudos detectaram variabilidade cariotípica em nível inter e intraespecifico. A aneuploidia é descrita como prejudicial para o *fitnes* da célula de mamíferos, provavelmente pelo desequilíbrio de síntese proteica (Williams et al., 2008). Porém, esse estado também pode significar rápida adaptação de patógenos e conferir, por exemplo, características como resistência a drogas (Polakova et al., 2009). Em *Leishmania* a aneuploidia já foi apontada como importante mecanismo do parasita para responder às modificações em seu ambiente, por exemplo, o número de cópias de cromossomos supranumerários foi relacionada ao nível de resistência do parasita a drogas utilizadas no tratamento das leishmanioses (Ubeda et al., 2008).

Quanto à reprodução, a teoria da clonalidade foi descrita para *Leishmania* e outros tripanossomatídeos e é aceita ainda hoje (Tibayrenc & Ayala, 1999). Eventos de recombinação já foram descritos (Akopyants et al., 2009), mas parecem ser infrequentes o bastante para não descaracterizar uma estrutura da população prevalentemente clonal. Porém, diante das informações referentes ao mosaicismo cromossomal, o conceito de clonalidade volta a ser questionado, uma vez que o mosaicismo gerado após uma duplicação celular (formação de “clones”) gera células filhas não idênticas. Também por essa razão se questiona o uso de modelos para organismos diploides nas análises populacionais, pois a frequência alélica detectada pode não ser totalmente representativa.

1.5.1. Impacto da plasticidade cromossômica sobre estudos de filogenia e genética de populações

À medida que a aneuploidia foi caracterizada em *Leishmania*, os testes baseados na hipótese de diploidia em estudos usando sequências de DNA (MLST) ou análises de

fragmentos (MLMT) passaram a ser questionados, assim como o conceito de clonalidade. As razões são as apresentadas em seguida.

Tradicionalmente, esses marcadores co-dominantes (MLST e MLMT) são escolhidos pela sua habilidade em detectar heterozigostos, permitir inferências sobre recombinação, resolver variabilidade genética e identificar estruturação da população. Os estudos baseados em sequencias de DNA (MLST) utilizam comumente o sistema Big Dye em plataformas de sequenciamento tipo Sanger. Isso significa que é possível observar picos duplos, sobrepostos nos eletroferogramas que caracterizam sítios ambíguos. O sitio ambíguo aponta para uma quantidade equivalente de dois nucleotídeos na amostra, e, portanto, é comumente interpretado como um heterozigoto. No MLMT dois ou mais picos de intensidade parecida e tamanhos distintos também podem ser observados, e também são interpretados como heterozigotos. Entretanto, em nenhum dos casos é possível determinar com precisão se os picos sobrepostos (MLST) ou duplos (MLMT) sinalizam uma mistura de cepas, policlonalidade ou heterozigose. Isso se deve ao fato de o mosaicismo cromossomal em uma cepa cultivada poder dar origem, por exemplo, a uma proporção equivalente de duas populações de células homozigotas, mas com alelos distintos. Como resultados serão observados picos sobrepostos em eletroferogramas e picos de intensidades semelhantes, mas de tamanhos distintos, durante a análise de fragmentos e ainda assim não representar um heterozigoto. Por outro lado, pode ocorrer um heterozigoto verdadeiro, mas formado por um terceiro alelo menos frequente e, portanto não proporcional nas células que formam o isolado sob estudo. Por isso pode ser interpretado apenas como ruído e não ser incluído nas análises. Com isso em mente, o conceito de clonalidade em *Leishmania* parece mais plausível do ponto de vista biológico que genético, uma vez que as células filhas não necessariamente serão idênticas apesar de a reprodução ocorrer por bipartição (Tibayrenc & Ayala, 2012, 2013).

2. Justificativa

As leishmanioses são um complexo de antropozoonoses, amplamente distribuídas pelo território nacional. A despeito das medidas de controle, o número de casos registrados vem aumentando e atingindo novas regiões (Marlow et al., 2013). Mesmo em estados onde há descrição de declínio da incidência, permanece o risco de contrair a doença. As razões para essa expansão podem variar de questões socioeconômicas e precipitação (Goncalves Neto et al., 2013) a migração humana, alterações ambientais com a expansão urbana desordenada, construção de estradas e ferrovias ou usinas hidroelétricas (Karagiannis-Voules et al., 2013). As estratégias de controle não mudaram no Brasil ao longo dos últimos 60 anos para leishmaniose visceral e os programas propostos encontram obstáculos, principalmente em diagnosticar e reportar novos casos (Dantas-Torres & Brandao-Filho, 2006).

A diversidade observada entre as espécies e cepas de *Leishmania* que circulam no Brasil é descrita em vários estudos, assim como a complexidade eco-epidemiológica da doença (Boite et al., 2012; Brito et al., 2009; Cupolillo et al., 2003). Essas características apontam para a importância da identificação de cepas e espécies no tratamento e controle das leishmanioses.

A identificação de isolados de *Leishmania* e classificação taxonômica atual do gênero ainda são baseadas no ensaio de isoenzimas (MLEE). A validade de muitas espécies tem sido, no entanto, questionada tendo como base ensaios moleculares, que não podem ser comparados entre laboratórios por utilizarem marcadores / alvos distintos. Há, entre muitos pesquisadores, a intenção de substituir o MLEE por um marcador molecular comum entre os laboratórios para que os dados possam ser disponibilizados e comparados facilmente. Uma vez estabelecido, um sistema integrado de identificação e tipagem molecular de *Leishmania* pode contribuir: i) para o desenvolvimento de técnicas de diagnóstico direto a partir de polimorfismos específicos; ii) para estudos de correlação entre genótipos com fenômenos biológicos e eco-epidemiológicos; iii) para avaliar a estrutura da população de *Leishmania* em relação à clonalidade e recombinação genética pela presença de heterozigotos e de alelos parentais na população estudada; iv) para suplantar as desvantagens do MLEE e promover a revisão taxonômica do gênero através de uma abordagem baseada em sequências de DNA.

O conhecimento da estrutura genética das leishmânias também é de grande importância para o estudo desses organismos, pois pode permitir a correlação entre a distribuição geográfica de padrões genéticos com a dinâmica de transmissão e epidemiologia dessa doença ainda negligenciada.

Diante do exposto, acredita-se que o desenvolvimento e análise de marcadores moleculares, que permitam tanto a identificação de isolados de *Leishmania* quanto o estudo da diversidade do gênero, e que funcionem também como ferramenta epidemiológica, representam uma contribuição valiosa para o estudo, controle e diagnóstico das leishmanioses.

3. Objetivos

O objetivo geral deste trabalho foi o desenvolvimento e avaliação de marcadores moleculares para o estudo da diversidade observada no subgênero *Leishmania* (*Viannia*) e da epidemiologia molecular das leishmanioses.

Objetivos específicos

Desenvolver um painel MLST para cepas de *Leishmania* (*Viannia*) que circulam no Brasil, que permita a identificação de isolados e a realização de estudos filogenéticos;

Determinar a estrutura da população de cepas de *Leishmania* (*Viannia*) quanto à clonalidade e ocorrência de eventos de recombinação;

Avaliar a aplicabilidade do painel MLST como ferramenta epidemiológica para as leishmanioses;

Descrever eventos moleculares em leishmânia que podem interferir nos resultados obtidos a partir de marcadores moleculares.

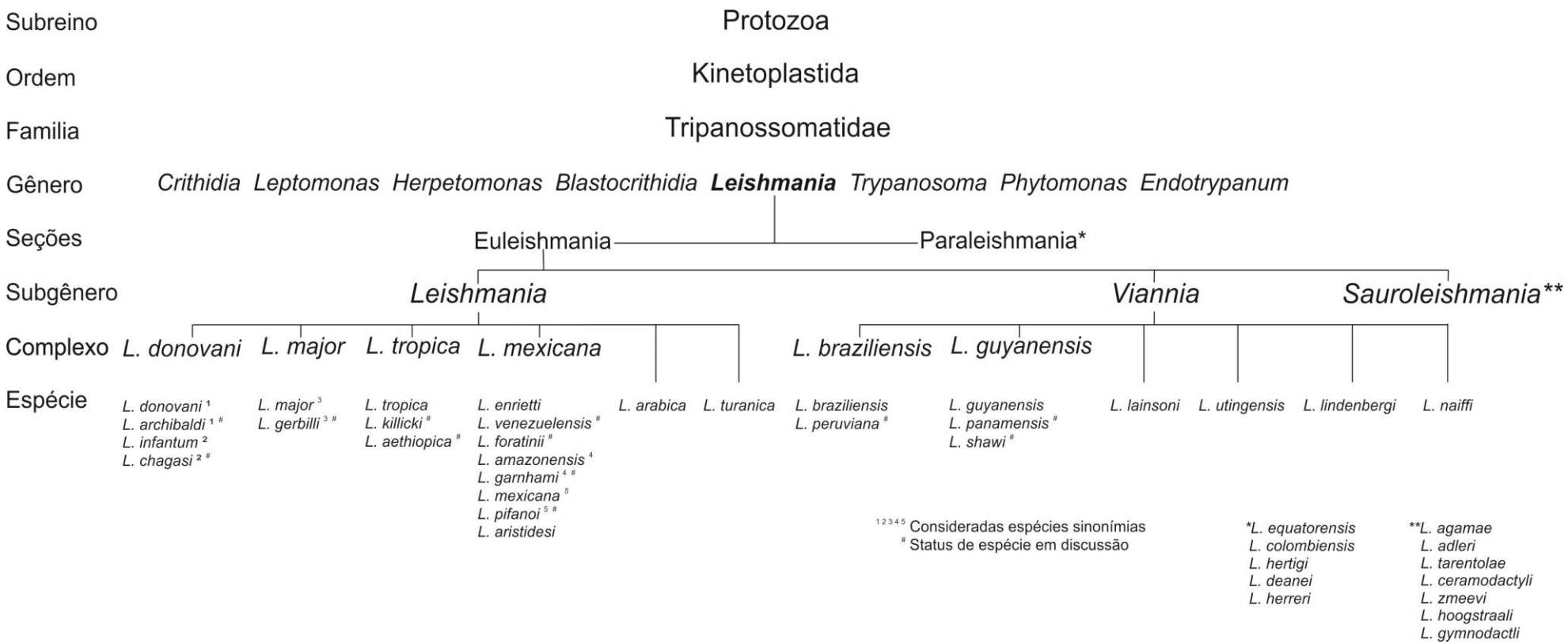

Figura 1. Atual sistema de classificação das leishmânias. Modificado (WHO, 2010).

	MLEE[a]	Species spec. PCR	Genus-spec. PCR (species id by RFLP,[b] hybridization, sequencing, HRM)	PCR RFLP minicircle kDNA	MLST[c]	MLMT[d]
Species typing						
Identification of sympatric and imported species	x	(x)	x		x	
Identification of species associated with treatment failure	x	(x)	x		x	
Identification of new non-human reservoir hosts	x	x	x		x	
Incrimination of new phlebotomine vectors	x	x	x		x	
Detection of new parasite-insect-host-combinations	x	x	x		x	
Detection of hybrids	x				x	
Strain typing						
Comparison of strains isolated from different hosts	x			x		x
Differentiation of zoonotic versus anthroponotic transmission cycles	x			x		x
Identification of new genotypes in mammal and insect hosts	x			x		x
Identification of the origins of new genotypes	x					x
Search for genotypes related to drug resistance				x		x
Differentiation between persistence or re-inoculation in relapses				x		x
Detection of epidemics	x			x		x
Identification of sources of *Leishmania*	x					x
Testing for population structure in *Leishmania* species						x
Testing for gene flow between populations and for hybrids	x				x	x

Abbreviations: [a] MLEE, multilocus enzyme electrophoresis; [b] RFLP, restriction fragment length polymorphism; [c] MLST, multilocus sequence typing; [d] MLMT, multilocus microsatellite typing. (X) limited application.

Figura 2. Questões epidemiológicas que podem ser abordadas por diferentes métodos na identificação de isolados de *Leishmania* (Schonian et al., 2011).

4. Material, métodos e resultados - Artigos gerados.

As seções “Material e métodos” e “Resultados” serão substituídas pela apresentação dos artigos gerados durante o projeto de doutorado, como sugerem as normas para elaborações de teses do programa de pós-graduação do Instituto Oswaldo Cruz.

A presente tese de doutorado gerou quarto artigos, sendo três já publicados e o último aceito para publicação. Os trabalhos serão apresentados em seguida na sua forma original de publicação, separados em capítulos precedidos de uma breve descrição dos resultados do estudo e de sua relação com os objetivos propostos (item 3).

4.1. Capítulo I

New insights on taxonomy, phylogeny and population genetics of *Leishmania* (*Viannia*) parasites based on multi locus sequence analysis.

Boité MC, Mauricio IL, Miles MA, Cupolillo E.
PLoS Neglected Tropical Diseases
2012; 6(11):e1888. doi: 10.1371/journal.pntd.0001888
Artigo publicado

Este estudo evidenciou o valor de uma abordagem *multi locus* na determinação das relações inter e intra-especificas para o subgênero *L*. (*Viannia*). Foi possível demonstrar que a análise de *multi locus* (MLSA) representa uma ferramenta útil para validação de algumas das espécies de *Leishmania* descritas, além de contribuir para o estudo filogenético do subgênero. Estudos epidemiológicos e análises de genética de populações baseados em haplótipos também podem fazer uso da abordagem apresentada. O artigo apresenta a avaliação de um painel de marcadores para cepas do subgênero *L*. (*Viannia*), que podem ser empregados no desenvolvimento de um sistema de tipagem baseado em *multi locus* – MLST, representando a conclusão do primeiro e objetivo desta tese.

OPEN ACCESS Freely available online

PLOS | NEGLECTED TROPICAL DISEASES

New Insights on Taxonomy, Phylogeny and Population Genetics of *Leishmania (Viannia)* Parasites Based on Multilocus Sequence Analysis

Mariana C. Boité[1], Isabel L. Mauricio[2,3], Michael A. Miles[2], Elisa Cupolillo[1]*

1 Laboratório de Pesquisa em Leishmaniose, Fundação Oswaldo Cruz, Instituto Oswaldo Cruz, Rio de Janeiro, Brazil, **2** Department of Pathogen Molecular Biology, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom, **3** Instituto de Higiene e Medicina Tropical/Unidade de Parasitologia e Microbiologia Médicas, Lisboa, Portugal

Abstract

The *Leishmania* genus comprises up to 35 species, some with status still under discussion. The multilocus sequence typing (MLST)—extensively used for bacteria—has been proposed for pathogenic trypanosomatids. For *Leishmania*, however, a detailed analysis and revision on the taxonomy is still required. We have partially sequenced four housekeeping genes—glucose-6-phosphate dehydrogenase (G6PD), 6-phosphogluconate dehydrogenase (6PGD), mannose phosphate isomerase (MPI) and isocitrate dehydrogenase (ICD)—from 96 *Leishmania (Viannia)* strains and assessed their discriminatory typing capacity. The fragments had different degrees of diversity, and are thus suitable to be used in combination for intra- and inter-specific inferences. Species-specific single nucleotide polymorphisms were detected, but not for all species; ambiguous sites indicating heterozygosis were observed, as well as the putative homozygous donor. A large number of haplotypes were detected for each marker; for 6PGD a possible ancestral allele for *L. (Viannia)* was found. Maximum parsimony-based haplotype networks were built. Strains of different species, as identified by multilocus enzyme electrophoresis (MLEE), formed separated clusters in each network, with exceptions. NeighborNet of concatenated sequences confirmed species-specific clusters, suggesting recombination occurring in *L. braziliensis* and *L. guyanensis*. Phylogenetic analysis indicates *L. lainsoni* and *L. naiffi* as the most divergent species and does not support *L. shawi* as a distinct species, placing it in the *L. guyanensis* cluster. BURST analysis resulted in six clonal complexes (CC), corresponding to distinct species. The *L. braziliensis* strains evaluated correspond to one widely geographically distributed CC and another restricted to one endemic area. This study demonstrates the value of systematic multilocus sequence analysis (MLSA) for determining intra- and inter-species relationships and presents an approach to validate the species status of some entities. Furthermore, it contributes to the phylogeny of *L. (Viannia)* and might be helpful for epidemiological and population genetics analysis based on haplotype/diplotype determinations and inferences.

Citation: Boité MC, Mauricio IL, Miles MA, Cupolillo E (2012) New Insights on Taxonomy, Phylogeny and Population Genetics of *Leishmania (Viannia)* Parasites Based on Multilocus Sequence Analysis. PLoS Negl Trop Dis 6(11): e1888. doi:10.1371/journal.pntd.0001888

Editor: Christian Tschudi, Yale School of Public Health, United States of America

Received June 15, 2012; **Accepted** September 17, 2012; **Published** November 1, 2012

Funding: This work was supported by Fundação Oswaldo Cruz (FIOCRUZ) - http://portal.fiocruz.br, Fundação de Amparo a Pesquisa do Estado do Rio de Janeiro (FAPERJ) - http://www.faperj.br, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) - http://www.cnpq.br. This research was also supported by a grant from the European Union (LeishEpinetSA, INCO-CT2005-015407). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

* E-mail: ecupoli@ioc.fiocruz.br

Introduction

Leishmania are the causative agents of leishmaniasis, which can present in different forms, from simple cutaneous to the deadly visceral disease, and are found in most tropical and sub-tropical regions. In spite of their morphological homogeneity, more than 20 species have been described for the *Leishmania* genus. Phenotypic diversity is observed not only between species, but also even in virulence levels among clones [1] which, upon interaction with the host's immunological response, contributes to determining the observed clinical pleiotropy and affects the efficiency of therapy applied [2]. *Leishmania* genetic diversity may also compromise vaccine development, although key antigen genes are highly conserved. Understanding leishmaniasis and the development of measures to counter its spread depend on the ability to identify *Leishmania* species and characterize genetic variants.

Multilocus enzyme electrophoresis (MLEE) is still considered by many to be the gold standard for *Leishmania* identification, but several DNA based methods have proven useful to study *Leishmania* genetic diversity [3,4]. DNA sequencing and PCR-RFLP of *hsp70* genes have been shown to be promising for the identification of *Leishmania* parasites [5,6], although too conserved for intra-specific diversity studies. Highly polymorphic markers, such as microsatellites, perform poorly at taxonomic levels higher than species, whilst most other genotyping methods rely on multicopy genes that are more difficult to analyze.

A standardized, sensitive, practical and reproducible typing method, such as multilocus sequence typing (MLST), must form the basis for a robust classification of *Leishmania* species, which is achievable in most laboratories even in the advent of fast and cheaper genome sequencing.

MLST, as proposed in 1998 for bacterial pathogens [7], provides a portable, reproducible, and quantitative typing system. It has since been applied to diploid organisms [8–10], including

Author Summary

Leishmania is a protozoan genus comprising many species, some associated with a human neglected disease called leishmaniasis. This parasite is found worldwide and is transmitted by sand flies, having numerous domestic and sylvatic animals as reservoirs. *Leishmania* is genetically and ecologically diverse and it has been argued that this has an impact on the epidemiology of the disease. Many typing methods have been proposed for the study of this diversity, although a generally agreed methodology is still required. Also, there is still a lack of consensus on the validity of some species. Multilocus sequence typing (MLST) is a method for studying the population structure and diversity of pathogens, but before an MLST scheme can be proposed it is essential to undertake a detailed analysis and selection of the sequences that are to be included in the system. Here, we sequenced four gene fragments of 96 *L. (Viannia)* strains, representing most species from this subgenus. Our results showed a good agreement between the current species assignment and the multilocus sequence analysis. Evidence of genetic recombination was found and the phylogenetic relationships were determined. Overall the results point to the feasibility of an MLST scheme for *Leishmania* and indicate that the four gene fragments analyzed could form part of this typing system. This will certainly be a valuable approach for taxonomy, population genetics, and epidemiological studies of this pathogen.

Leishmania, with ten markers described for the *L. donovani* complex [11–13] and four targets proposed for *L. (Viannia)* spp. [14]. MLST databases for *Leishmania* have not yet been implemented and the published number of strains and species typed is relatively low. Alongside MLST, multilocus sequence analysis (MLSA) has been shown to be a good tool for strain characterization and epidemiological surveillance, as well as population structure and evolutionary studies [9,10,15,16]. The implementation of an MLST system demands careful evaluation of markers beforehand to study their diversity and phylogenetic consistency [17,18].

We report here the evaluation of four candidate coding regions, located in different chromosomes, to be included in an MLSA system of the subgenus *L. (Viannia)*. These regions are part of the genes for the metabolic enzymes glucose-6-phosphate dehydrogenase (G6PDH, EC 1.1.1.49), phosphogluconate dehydrogenase (6PGD, EC 1.1.1.44), mannose phosphate isomerase (MPI, EC 5.3.1.8) and isocitrate dehydrogenase (ICD, EC 1.1.1.42). Three of these genes had been studied previously, but mainly for Peruvian isolates of the *L. braziliensis* complex. Herein a large number of strains and most species of the subgenus *L. (Viannia)* were included. The sequences were used to conduct haplotype analysis, including the construction of haplotype networks. Diploid multilocus analyses were performed and a concatenated tree was defined, as well as clonal complexes (CC). Using an MLSA approach, we report the genetic diversity of the four gene fragments evaluated in the *L. (Viannia)* group and demonstrate their value for taxonomic, phylogenetic and population genetic studies of *Leishmania* parasites.

Methods

Ethics statement

Research in this study was subject to ethical review by the European Commission and approved as part of contract negotiation for Project LeishEpiNetSA (contract 01547): the work conformed to all relevant European regulations. The research was also reviewed and approved by the ethics committee of the London School of Hygiene and Tropical Medicine (approval 5092). The *Leishmania* strains analyzed were principally strains derived from international cryobanks such as Coleção de Leishmania do Instituto Oswaldo Cruz (CLIOC) and the cryobank of the London School of Hygiene and Tropical Medicine (LSHTM). In all cases *Leishmania* were isolated from patients as part of normal diagnosis and treatment with no unnecessary invasive procedures and with written and/or verbal consent recorded at the time of clinical examination. Data on isolates were coded and anonymised. The *Leishmania* strains are deposited either at the LSHTM (n = 9) or at CLIOC (n = 87) as open data. CLIOC is a Depository Authority by the Ministry of the Environment [Fiel Depositária pelo Ministério do Meio Ambiente, MMA] (D.O.U. 05.04.2005). All samples were used for research purposes only and the data were analyzed anonymously in the scope of resolution 21 (August 31, 2006 – CGEN/MMA), for which authorization is not required.

Samples and GenBank available sequences

Ninety six strains were obtained from frozen stocks from CLIOC (n = 87), and from the London School of Hygiene and Tropical Medicine (LSHTM) (n = 9) (Table S1). Strains were chosen to be representative of the zymodeme and geographical diversity of species of the subgenus *Leishmania (Viannia)* in South America, and in particular Brazil: *L. (V) braziliensis*, *L. (V) guyanensis*, *L. (V) naiffi*, *L. (V) lainsoni*, *L. (V) shawi*, *L. (V) utingensis* and *L. (V) lindenbergi*. Most strains had been previously characterized by MLEE [19]. DNA sequences were retrieved from GenBank of *L. (Viannia)* species: 14 of G6PD; 24 of 6PGD; 20 of MPI and 1 of ICD (Table S2). Only those presenting full coverage with the alignment of sequences obtained in the present study were included in the analysis: for G6PD: 3 *L. braziliensis*, 2 *L. guyanensis*, 1 *L. peruviana*, 1 *L. panamensis*, 1 *L. lainsoni*; for 6PGD: 10 *L. peruviana*, 07 *L. braziliensis*, 1 *L. panamensis*; for MPI: 11 *L. peruviana*, 9 *L. braziliensis*; for ICD: 1 *L. braziliensis*.

DNA extraction and PCR

Promastigotes were grown at 25°C in Schneider's medium supplemented with 20% (v/v) heat-inactivated fetal bovine serum to a density of 1×10^9 cells/mL (late log phase), as estimated by counting in a Neubauer chamber. DNA extraction was performed using the Wizard DNA purification Kit (Promega, Madison, USA) following the manufacturer's instructions.

The chosen loci are distributed on different chromosomes according to *L. (Viannia) braziliensis* genome sequencing data: 6PGD, on chromosome 34; G6PD, on chromosome 20; MPI, on chromosome 32; ICD, on chromosome 33.

Primers (Table 1) were designed in conserved regions of gene sequences from the published *L. major* and the *L. braziliensis* genomes in Genbank (www.ncbi.nlm.nih.gov/Genbank) to amplify sections of the coding regions of the genes that would be amenable to full sequencing using the PCR primers and that include putative species-specific polymorphisms as well as singleton SNPs. Amplification reactions had, for 50 µl total volume, 0,1 mM of each primer, reaction buffer (100 mM Tris–HCl, pH 8.8; 500 mM KCl, 1% Triton X-100; 15 mM $MgCl_2$), 0.25 mM deoxyribonucleotide triphosphate (dNTPs), 0.025 U FideliTaq/GoTaq polymerase and approximately 50 ng DNA. Amplification conditions were: 94°C for 2 minutes, followed by 35 cycles at 94°C for 30 seconds, 55°C (for ICD, 6PGD and G6PD) or 58°C (for MPI) for 30 seconds and 68°C for 1 minute, with a final extension at 68°C for 5 min.

Table 1. Detail of target regions of each locus studied.

Locus	Gene ID*	Gene length	Amplicon size (bp)	Primer positions	Primers sequence 5′-3′
G6PD	LbrM.20.0160	1686	881	173	ATGGAAGCGTGTGATCGAAT
				1015	GGCTCAACACACTTCAGCAA
6PGD	LbrM.34.3250	1440	836	143	CTCAAGGAACATGAGCACGA
				940	TTGTCCTTGACTTGCTCACG
MPI	LbrM.32.1750	1287	681	128	GGCAAGATGTATGCGGAGTT
				770	CTCCCCAGGAACCATCTGTA
ICD	LbrM.33.2820	1278	1022	99	GAATCGGGAAGGAGATCACA
				1082	CATCATAGCCCCAGAGAGGA

Bp- base pairs; Fw - Foward; Rv – Reverse.
*related to *L. braziliensis* genome.
doi:10.1371/journal.pntd.0001888.t001

DNA sequencing

PCR products were purified with the Wizard SV Clean-up System (Promega). The final DNA concentration was estimated by comparison with a DNA Ladder Marker (Promega) in 2% agarose gel. Sequencing was performed with the same primers used for amplification, using the ABI PRISM BigDyeTerminator v3.1 Cycle Sequencing Kit, and products analyzed in an automated DNA sequencer (ABI-3730). Consensus sequences were generated and edited in Phred/Phrap/Consed Version: 0.020425.c [20] from two forward and two reverse strands. Sequences with Phred values below ten over their extent were discarded and only sequence segments with values above twenty were used for contig construction. Ambiguous (heterozygous) sites were coded using the standard IUPAC codes for combinations of two or more bases. Contigs from all samples were manually assembled and aligned in MEGA4 [21]. The homologous sequences available in GeneBank for other *L. (Viannia)* isolates were obtained using the Basic Local Alignment Search Tool (BLAST) algorithm hosted by NCBI, National Institute of Health, USA (http://www.ncbi.nlm.nih.gov) and included in the alignments.

Haplotype analyses

Haplotype reconstruction was done through DNAsp5 [22] using the PHASE algorithm, which automatically assigns a haplotype number (H) for each unique haploid sequence. Haplotype data was then used in DNAsp to calculate the synonymous and non-synonymous substitution rates and haplotype diversity (HD). Haplotype diversity (HD) was calculated as: $HD = N(1-\Sigma xi^2)/(N-1)$ where xi is the haplotype frequency of each haplotype in the sample and N is the sample size. The average haplotype diversity for species was compared through ANOVA.

The discriminatory capacity of each marker employed was calculated through Simpson's diversity index (D), as follows: $D = \Sigma n(n-1)/N(N-1)$, where n = the total number of strains of a particular haplotype and N = total number of strains analyzed. Phylogenetic congruence between markers was assessed by comparison of maximum parsimony (MP) and median-joining (MJ) networks generated by the Network free software [23], using individual haplotype data.

Diploid multilocus analysis

For each sample a sequence type number (ST) was defined with each marker, which in the homozygous strains was identical to the haplotype number (H), and for the heterozygous strains was a combination of the two possible alleles. A diploid sequence type (DST) was defined for the final combination of STs of the four markers. The sequences from the four genes were concatenated using BioEdit *v7.0.9* and unique concatenated sequences were identified with the DST previously assigned. To evaluate the phylogenetic information provided by the four markers and to investigate the presence of recombination signatures, a Neighbor-Net network was built in SplitsTree 4.0 [24] based on genetic distances calculated according to the Kimura-2 parameter model of nucleotide substitutions from concatenated data.

ST data combinations (with the exception of heterozygous strains) were analyzed in e-BURST v.3 (http://eburst.mlst.net/v3/enter_data/single/) to define clonal complexes (CC), which are sets of related strains containing pairs of strains that share at least (L-1) identical alleles at the L loci with at least one other member of the CC.

Results and Discussion

We report here the most comprehensive multilocus sequence analysis study to date of the subgenus *Leishmania (Viannia)*, based on fragments of four different metabolic enzyme coding genes from a wide range of species, zymodemes and geographical origins.

With the aim of developing a working MLST scheme, we selected polymorphic coding regions that could be sequenced with forward and reverse primers. The full coding regions of three of the studied genes (6PGD, G6PD and MPI) had already been sequenced and characterized for some *L. (Viannia)* strains, mainly of Peruvian *L. braziliensis* and *L. peruviana*. However, we also partially analyzed the coding region for ICD and studied a large number of strains (96), to include most known species or complexes of *L. (Viannia)*.

Genetic diversity of the four housekeeping genes used in the MLSA of *Leishmania (Viannia)*: Polymorphic, ambiguous and multi-allelic sites

The sequences obtained for each gene varied between 589 (MPI) and 914 bp (ICD) (Table 2) and covered between 38% (G6PD) and 71% (ICD) of the corresponding gene. Sequences were deposited in GenBank (www.ncbi.nlm.nih.gov/Genbank) with accession numbers JN996517–JN996708 and JQ181608–JQ181801 and are also available at the CLIOC website (htpp://clioc.fiocruz.br).

Gene fragments had 50–78 polymorphic sites (respectively, 7.3–8.7%) for the strains studied here, of which 46 (G6PD) to 76 (ICD) were parsimony informative (PI) sites (Table 2). 6PGD had the

Table 2. Characteristics of each sequenced region for the 96 strains used in the present study.

Gene	Length	PS	PI (bp)	Singletons	Ambiguous sites	
	(bp)	(bp)	(% of length)	(% of length)	Number of sites (% of length)	Number of isolates (% total strains)
G6PD	683	50	46 (6.7%)	4 (0.6%)	6 (0.9%)	8 (8.3%)
6PGD	716	63	60 (8.4%)	1 (0.1%)	5 (0.7%)	9 (9.4%)
MPI	589	50	49 (8.3%)	1 (0.2%)	2 (0.3%)	2 (2.1%)
ICD	914	78	76 (8.3%)	2 (0.2%)	6 (0.7%)	10 (10.4%)

Percentages are relative to the total number of sites (% of length) or strains (% of total strains). PS – Polymorphic sites; PI - parsimony informative sites.
doi:10.1371/journal.pntd.0001888.t002

highest percentage of PI sites (8.4%), with the lowest found in G6PD (6.7%). The studied G6PD region had a much higher number and percentage of singletons, 4 (0.6%), in relation to the respective total template length than the other three gene regions (0.1–0.2%; Table 2).

The four sequenced fragments, representing different loci, had different degrees of diversity between species groups as seen by the haplotype diversity (Figure 1) and are thus suitable to be used in combination for intra- and inter-specific inferences. Although theoretically including more loci improves discriminatory capacity, it has been described that increasing the number of genes above four did not increase discrimination of MLST [18]. Furthermore, the index of diversity [25] for the targets included in the present study implies that the typing results can be interpreted with confidence.

Overall, after including published gene sequences for the same regions as studied here (except for ICD, which had no available sequences) gene fragments provided 51 (MPI) and 64 (6PGD) PI sites (Table 3).

Most polymorphic sites were bi-allelic, although sites with three variants occurred in three genes, G6PD, 6PGD and in ICD (Table S3). Recently, the genome plasticity of some *Leishmania* species was analyzed [26]. The evaluation of the *L. braziliensis* genome showed that 30 of 35 chromosomes are clearly trisomic; three are tetrasomic (chromosomes 4, 5 and 29) and one hexasomic (chromosome 31). Moreover, the same study showed that multicopy genes are found preferentially on non-supranumerary chromosomes. Such peculiarities are of great importance in all molecular analyses proposed for the *Leishmania* genus. In the present work, all markers used are located on different chromosomes, which are described as non-supranumerary [26]. However, the description of a trisomic genome for *L. braziliensis* indicates that multiallelic possibilities might be quite frequent when polymorphic sites are analyzed in DNA sequences of *Leishmania (Viannia)* species.

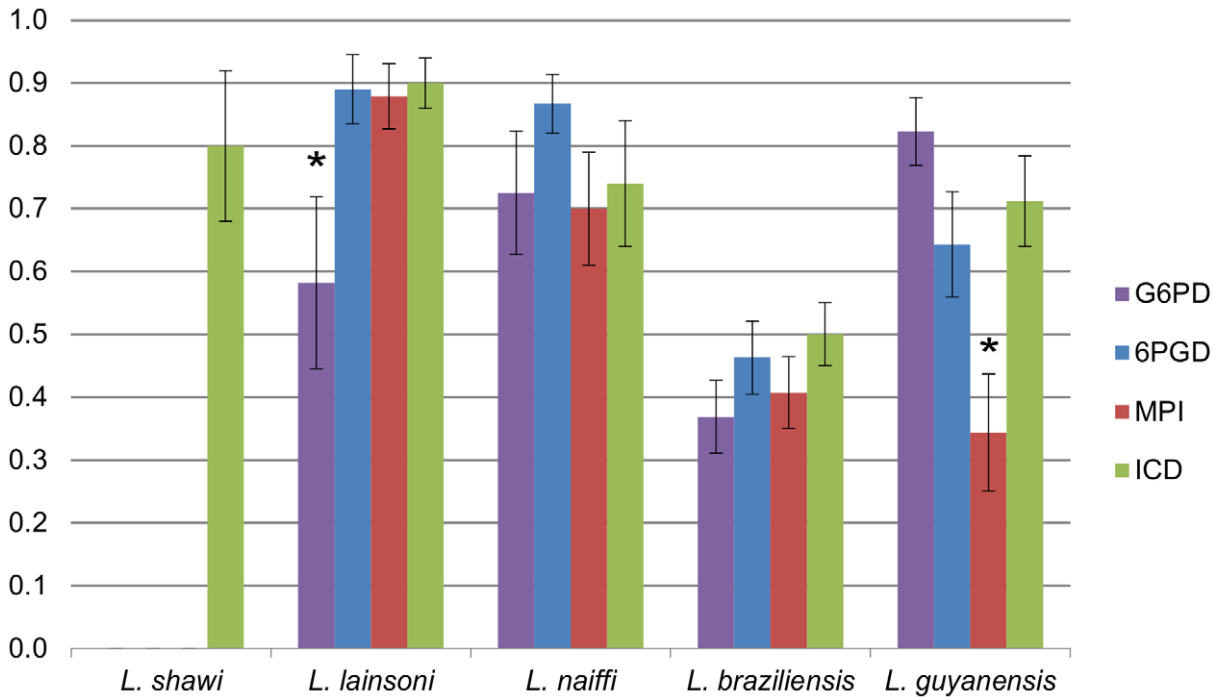

Figure 1. Comparison of haplotype diversity values of each locus within species groups. ICD was the only polymorphic marker for *L. shawi* strains. * G6PD and MPI were the least polymorphic markers ($P<0.001$) within species for *L. lainsoni* and *L. guyanensis* respectively.
doi:10.1371/journal.pntd.0001888.g001

Table 3. Characteristics of each sequenced region for all available *L. (Viannia)* (retrieved from GenBank plus the 96 strains included) and for *L. braziliensis* complex samples.

Gene		Length (bp)	PI*	Singletons	Ambiguous sites	
			(% of length)	(% of lentgh)	Number of sites	Number of isolates
G6PD	*L. (Viannia)*	683	54 (7.9%)	4 (0.6%)	6	8
	L. braziliensis		11 (1.6%)	0	2	5
6PGD	*L. (Viannia)*	716	64 (8.9%)	1 (0.1%)	5	9
	L. braziliensis		27 (3.8%)	0	3	5
MPI	*L. (Viannia)*	589	51 (8.7%)	1 (0.2%)	2	2
	L. braziliensis		17 (2.9%)	1 (0.2%)	3	3

Percentages are relative to the total number of sites (% of length).
*PI - parsimony informative sites.
doi:10.1371/journal.pntd.0001888.t003

Moreover, gene copy numbers might also change between species and generations of the same strain.

Twenty-six isolates had at least one locus with double peaks in the chromatograms (Table 2 and Table S4). The number of sites showing double peaks was similar in three gene regions, but much smaller for MPI, which also had a much smaller number of strains with double peaks in their sequences compared to the other targets (Table 2). Different strains presenting double peaks at the same site were observed in all markers, except MPI (Table S4).

Although most strains were homozygous, almost a third had at least one locus with double peaks, except *L. shawi*. MPI was the most homozygous gene, even though it was not the most conserved. All the other markers presented more than one double peak per strain and more than one strain with double peaks. It is possible that this is a sampling effect of the sequenced region in MPI, or that gene conversion in this gene is higher [27].

The presence of ambiguous sites and possible parental alleles among the samples studied strongly suggests real heterozygosis and also the occurrence of some level of recombination. Ambiguous sites were observed in twenty positions over the four gene fragments analyzed (Table S4), and putative homozygous parental were observed in eight of them. Many strains had two or more heterozygous sites. Heterozygosis can be caused by mutation in one allele but it can also be caused by genetic exchange between strains with different alleles. Mutation is more likely for single heterozygous sites but recombination is a more parsimonious explanation for two or more sites [28]. The majority of heterozygous alleles among our sample presented just one ambiguous site. Exceptions were encountered in one *L. braziliensis* strain showing HP 8/17 in 6PGD and 6/11 in ICD, with two and three ambiguous sites respectively. This observation reinforces the fact that real heterozygosis might be occurring, because these two loci are on different chromosomes and in both sequence alignments at least two ambiguous sites were observed in that sample. Moreover, after determination of the minimum number of recombination events by using DNAsp, recombination between sites was detected for the four markers, with highest frequency for 6PGD and least for G6PD (data not shown).

The assumption that ambiguities in the chromatograms are the result of heterozygosis can only be reliably postulated if the target is a single copy gene. In our case, the four selected genes are single copy genes, as estimated by in silico analysis of the four available *Leishmania* species reference genomes (data not shown). Heterozygous samples would need confirmation through biological cloning of isolates, to exclude polyclonal populations. The strains we have used were not cloned, except for the reference strains. As far as we know, none of the studies aiming to construct an MLST system for *Leishmania* have used biologically and/or genetically cloned the samples. Following an overview of sample profiles, without cloning, if ambiguous sites are detected, the strains containing them should be selected for deeper analysis and cloning.

Species-specific SNPs

Putative species-specific SNPs, as determined through the data presented here and available in GenBank, were detected for *L. guyanensis* in 6PGD (A105); for *L. naiffi* in MPI (15T, 82A, 98A, 270A, 306C, 498C, 543T) and G6PD (320C, 341C, 432A, 467A); for *L. lainsoni* in MPI (33G, 105G, 294T) and G6PD (116A, 239T, 247G, 326A, 464C, 482C, 572A, 577T) and for *L. shawi* in MPI (135G). No species-specific SNPs were detected for *L. braziliensis*. ICD was the only marker that did not present species-specific SNPs.

In a previous study, species-specific SNPs for all species, including *L. braziliensis* were shown [14]. This incongruence between the two studies might be a consequence of differences in the gene fragments analyzed as well as in the strains studied. SNP markers should thus be used for species identification with great care, even when more strains are studied in future. Either full sequences should be obtained, or a large panel of SNP markers should be used for reliable identification and characterization.

Identification of assorted *Leishmania (Viannia)* haplotypes

Here we found a large number of unique haplotypes, which are likely to be rare or recent in the population. In contrast, some haplotypes were shared across species, such as haplotype H1 of 6PGD, which was detected in strains from different species: almost all *L. braziliensis* (n = 40), nine *L. peruviana* as well as two previously published *L. braziliensis* sequences, one *L. lainsoni* (IOCL 2500 from Acre) and one *L. naiffi* (IOCL 855 from Amazonas) strain. This could be due to recombination or convergent evolution, but it could also represent the most likely ancestral haplotype, as depicted from the haplotype network constructed for this gene (Figure 2). Remarkably, those species present different electrophoretic mobility for the 6PGDH enzyme system. Such incongruence might be due to the fact that the entire coding region was not sequenced here, so sequence sections coding for differences in MLEE might not be present in the current analysis, or it might be due to post-transcriptional and post-translational modifications [28].

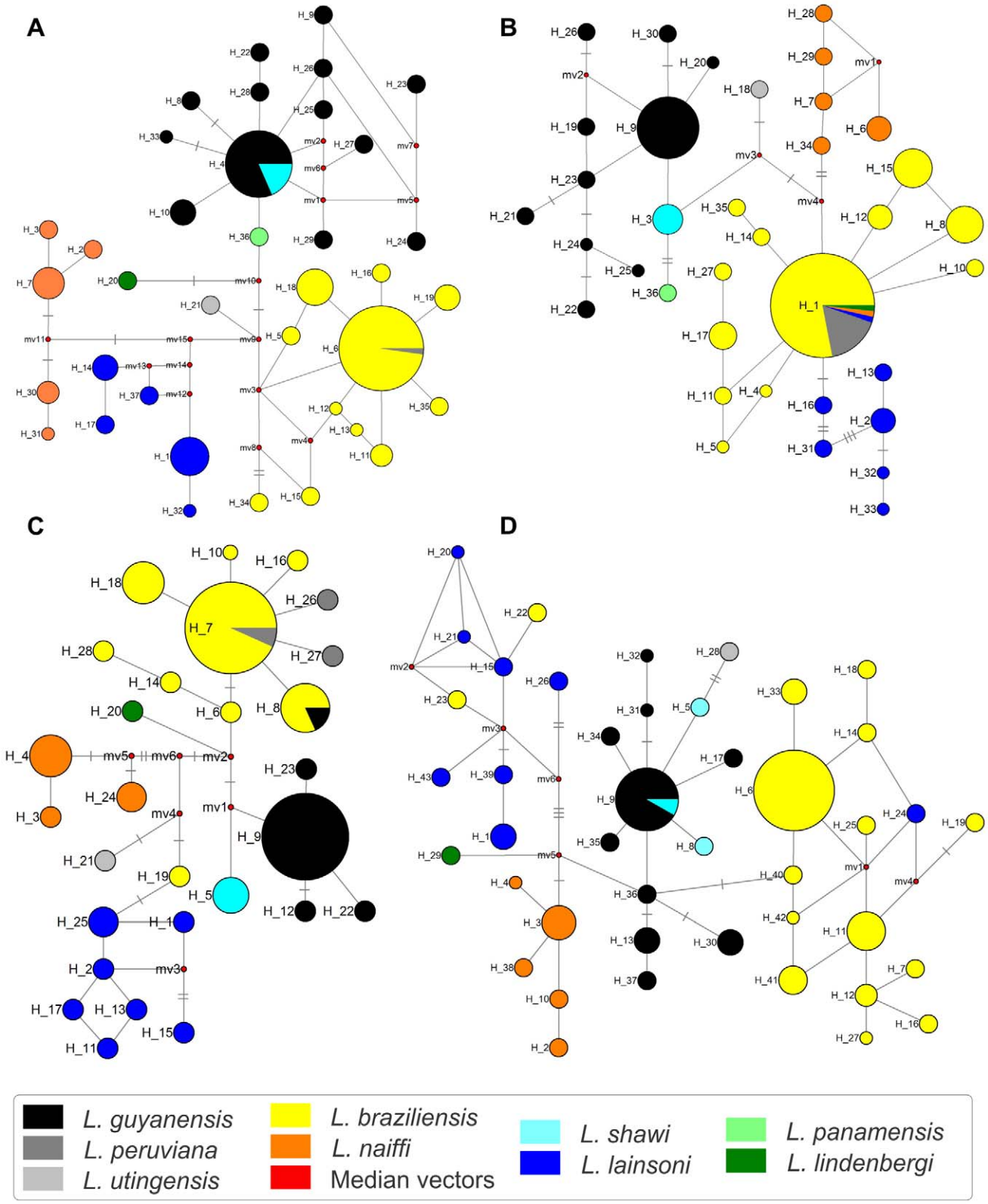

Figure 2. Maximum parsimony-based haplotype networks. A) G6PD; B) 6PGD; C) MPI; D) ICD. The 96 strains selected for the study (Table S1) plus sequences of *Leishmania* (*Viannia*) strains retrieved from GenBank (Table S2) were included. Haplotype frequency is represented by the size of each node and the numbers of polymorphisms are indicated in the branches by dashes: one = 02 to 05 polymorphism; two = 06 to 10; three = more than 10 polymorphisms. Each species, considering the MLEE characterization, were assigned by different colors as coded by the legend.
doi:10.1371/journal.pntd.0001888.g002

Table 4. Haplotype analysis per marker for the 96 strains included in the present study.

Gene	Length (bp)	N° of different alleles	HD*	N° unique alleles
G6PD	683	33	0.78	21 (63.6%)
6PGD	716	34	0.76	22 (64.7%)
MPI	589	25	0.77	16 (64%)
ICD	914	43	0.85	31 (72.1%)

Percentages are relative to the number of different alleles.
*HD - haplotype diversity.
doi:10.1371/journal.pntd.0001888.t004

Upon haplotype assignment it was found that among the 96 studied strains there were 25 (MPI) to 43 (ICD) different haplotypes, of which 16 (MPI) to 31 (ICD) were represented by only one strain. Nevertheless, ICD had the highest percentage of exclusive haplotypes per strain (72.1%) and G6PD had the lowest (63.6%; Table 4).

Strains identified as *L. braziliensis* represented more than 50% of the strains studied. The greatest number of haplotypes was observed within the *L. braziliensis* group for all markers except G6PD for which the greatest number of haplotypes was observed within the *L. guyanensis* group (Table 3). Around 30% of the haplotypes in each gene were detected in *L. braziliensis* (Table 3). However, almost no singletons and no species-specific SNPs were observed among the *L. braziliensis* strains, but some ambiguous sites were seen (Table S4). *L. utingensis* and *L. lindenbergi* had unique haplotypes for all markers, with the exception of 6PDG for *L. lindenbergi*, as mentioned above.

Regarding G6PD, 12 haplotypes were shared by at least two strains. The most common haplotype, H6, was found in 45 strains, characterized as *L. braziliensis* or *L. peruviana*. Four additional haplotypes were found in previously published sequences, one in *L. panamensis* (H36), two *L. braziliensis* (H34, from Peru, and H35, from Brazil) and one *L. lainsoni* (H37, from Peru). One published *L. guyanensis* sequence had haplotype H4, the most common *L. guyanensis*/*L. shawi* haplotype in this analysis (Table S2). Except for *L. shawi* and *L. peruviana*, which shared haplotypes with *L. guyanensis* and *L. braziliensis* respectively, the network depicted from this target presented the best clustering accordingly to species (Figure 2A).

Concerning 6PGD, haplotype H15 was only found in one isolate of *L. braziliensis* from Acre, northern Brazil, but it was identical to four previously published *L. braziliensis* strains from Peru. Similarities among strains from these two adjacent areas were reported previously based on MLMT [29], corroborating the suggestion that a geographic cluster and probably a hierarchical population structure of *L. braziliensis* exist in this area. *L. guyanensis* had 10 haplotypes, of which H9 was the most frequent, and *L. shawi* had one distinct haplotype (H3). Only two 6PGD haplotypes found in previously published sequences (one *L. braziliensis* from Peru and one *L. panamensis*) were not detected among the strains studied here, indicating a good coverage of haplotype diversity by our study.

The two most frequent alleles observed, corresponding to 6PGD H1 and H8 (Figure 2B), were also observed combined in one heterozygous *L. braziliensis* strain (IOCL 2833). These haplotypes were also observed in one (H1; IOCL 2494) and two (H8 IOCL 918 and 2538) other heterozygous strains, combined with other haplotypes (Table S4). Some 6PGD haplotypes were unique to heterozygous strains, as H4/H5 and H30/H31 (Table S1).

MPI *L. braziliensis* haplotype H8 was also found in one *L. guyanensis* strain. Haplotype H7 was the most common in *L. braziliensis* (41 strains) and was also found in eight previously published sequences of *L. peruviana* and eight of *L. braziliensis* (Table S2). MPI was previously reported as a good target to discriminate between *L. braziliensis* and *L. peruviana* [14,30]. Our conflicting results could be related to differences in the fragment regions analyzed, or might reflect a bias in the sample analyzed previously. Two published sequences of *L. peruviana* strains presenting an ambiguous site had the most common H7 and a new haplotype 26 after allele reconstruction. One available *L. peruviana* and one *L. braziliensis* sequence from Peru presented additional haplotypes H27 and H28, respectively (Figure 2C; Table S2).

The greatest number of haplotypes was found in ICD, for almost all species, although there were similar levels for G6PD and 6PGD in *L. naiffi*. ICD H6 was by far the most common *L. braziliensis* haplotype, including one previously published sequence (Table S2). ICD H9, the most common *L. guyanensis* haplotype, was also found in one *L. shawi* (IOCL 3200) (Figure 2D; Table S1).

Overall, and including published gene sequences for the same regions as studied here, there are 28 to 43 different described alleles in *L. (Viannia)*, of which 11 to 16 are in the *L. braziliensis* complex.

Haplotype diversity and haplotype network selection of marker combinations for resolution of inter- and intra-species relationships

Overall, haplotype diversity (HD) was higher for ICD (0.85), similar among the other gene regions (0.76–0.78) and as compared in ANOVA not significantly different ($P>0.01$). However, upon analysis by species, MPI was the least polymorphic marker for *L. guyanensis* ($P<0.001$) and G6PD was the least polymorphic locus for *L. lainsoni* ($P<0.001$). Among the three *L. shawi* isolates, ICD was the only polymorphic marker (Figure 1).

The Simpson index of diversity (D), which provides an indication of the discriminatory capacity of each marker, was similar for MPI, G6PD and 6PGD, (0.77, 0.77 and 0.78 respectively), but higher for ICD ($D = 0.89$). Analyses of the discriminatory capacity indicate a high level of strain discrimination, of almost 90%, using these four loci. However, higher values were observed when applied to *L. braziliensis* and *L. guyanensis* (96% and 100%, respectively). Other species were represented by few strains and were not analyzed. This suggests that, even though the number of markers is smaller than that usually used in MLST (7), these four markers are sufficient for studies in *L. (Viannia)*. However, more detailed population genetics studies may require more markers. Regarding the discriminatory power of each marker, the Simpson index showed that ICD had the highest diversity. Indeed, this is a very polymorphic locus in MLEE, which is able to detect intra-species variation [19].

MPI was here found to be a good marker to distinguish between the species, although HP8 was shared among *L. braziliensis* and one *L. guyanensis* strain. This enzyme is used as a marker to differentiate *L. peruviana* from *L. braziliensis*, which was supported by DNA sequencing in a previous study [14] by the detection of a specific SNP, as well as others useful to differentiate closely related *L. (Viannia)* species. We could not confirm this, given that our sequences were shorter and did not include that SNP locus.

Maximum parsimony-based haplotype networks built for each gene (Figure 2) showed that species, as identified by MLEE, clearly formed separated clusters in each gene network, with a few exceptions confined to some strains that for some markers were not grouped in their species haplogroups. Reticulate patterns were observed in some clusters for all loci studied. Even with markers

for which *L. shawi* presented different haplotypes from *L. guyanensis*, these two species always clustered close together.

Haplogroups were, in general, consistent with the species (color-coded nodes), although exceptions occurred for all markers: G6PD (H4 and H6; Figure 2A); 6PGD (H1; Figure 2B); MPI (H7, H8 and H19; Figure 2C) and in the ICD network (H9, H22, H23, H24; Figure 2D). A single haplotype MPI (H19) comprising an *L. braziliensis* strain (IOCL 2541, *L. braziliensis* from Pernambuco) was at the base of the *L. lainsoni* cluster, but for all the other markers this strain clustered within the *L. braziliensis* species group. One *L. lainsoni* strain (IOCL 2500 from Acre) was part of the *L. braziliensis* cluster for ICD (H24) and 6PGD (H1). Conversely, the *L. lainsoni* cluster included two haplotypes exclusively from *L. braziliensis* strains (haplotypes H22 and H23, IOCL 2498 and IOCL 2499 respectively, from Acre), even though both clusters were located in opposite sides of the network.

The most frequent haplotypes (node size) were often the founding haplotype for a given haplogroup, as clearly observed for *L. braziliensis* (MPI H7, ICD H6, G6PD H6, 6PGD H1) and *L. guyanensis* (MPI H9, ICD H9, G6PD H4, 6PGD H9). Eighteen out of 55 *L. braziliensis* strains were frequently observed in the most common haplotypes, but for *L. guyanensis* the composition of the most common haplotypes was different between the markers. The most frequent haplotypes were also usually those in which inter-specific sharing of sequences was observed (Figure 2).

L. guyanensis formed a diverse cluster, whereas *L. shawi* strains presented a profile coherent with a subpopulation of the *L. guyanensis* group for all markers, commonly sharing the most common *L. guyanensis* haplotype or differing from it in at most two polymorphisms (Figure 2C).

Although we used a relevant number of strains, the data analysis by Network software generated median vectors. The presence of median vectors in the networks might indicate that: i) intermediate haplotypes were present in lost populations; ii) haplotypes from populations that not included in this analysis; iii) the ancestors of these strains suffered rapid adaptive evolution with expansion of these extant strains. Therefore, even without sequencing strains representing all genetic diversity, statistical tools may predict the variability.

Neighbour Joining (NJ) trees were built for each marker to evaluate the phylogenetic relationship between the haplotypes. Almost no incongruence was observed between the markers. Four monophyletic groups were clearly observed for each marker, representing basically *L. lainsoni* and *L. naiffi*, the most divergent groups, and *L. braziliensis* and *L. guyanensis* (in this case, including *L. shawi*), which were very closely related for any marker (data not shown), corroborating the genetic distance MLEE-based tree [19]. *L. lindenbergi* (except in 6PGD) and *L. utingensis* were each in a separate and independent branch, but grouping closer to *L. naiffi* and *L. guyanensis* respectively.

Available sequences of *L. peruviana* for G6PD, 6PGD and MPI, and for *L. panamensis* for G6PD and MPI, were included in the respective network constructions. *L. peruviana* sequences either presented the most frequent haplotype for *L. braziliesnis* strains (in G6PD H6, 6PGD H1 MPI H7; Figure 2) or differed from it in one polymorphic site (in the MPI H26 and H27; Figure 2C; Table S2). A previous study using random amplification of polymorphic DNA (RAPD) and MLEE [31] reported that *L. peruviana* and *L. braziliensis* corresponded to two closely related, but distinct monophyletic lines, which was not corroborated by *hsp70* gene sequence analysis [32]. *L. panamensis* sequences presented new haplotypes (G6PD H36 and 6PGD H36) within the *L. guyanensis* and *L. shawi* haplogroup. Previously [33] MLEE and RAPD analysis questioned the distinction between *L. panamensis* and *L. guyanensis*, since data did not indicate distinct monophyletic lines. In individual NJ trees for the gene fragments studied here (data not shown), the *L. panamensis* sequences clustered within the *L. guyanensis*/*L. shawi* group and *L. peruviana* clustered within the *L. braziliensis* group. It was not possible to include *L. panamensis* and *L. peruviana* in the final MLSA conclusions since there were no available sequences for all genes studied. Furthermore, more strains from both species should be sequenced for the four gene targets to infer properly on the monophyletic origin of them.

Concatenated NeighborNet confirms species-specific clusters and suggests relatively frequently recombination occurring in *L. braziliensis* and *L. guyanensis*

Among the 96 *L. (Viannia)* strains, 75 final diploid sequence types (DSTs) were assigned. Although we detected a high number of DSTs, many DSTs are unique, while others are more prevalent, widely distributed and presenting temporal stability, which might reflect limited genetic recombination involving these DSTs [17]. The only species with strains sharing the same DST was *L. braziliensis*. DST12, for example, was found in 18 strains of *L. braziliensis* from different Brazilian endemic regions related to the Atlantic rain forest, except one, and present zymodeme diversity although most are classified in zymodeme 27 (Table S1). DST12 is not only the most frequent but also shows temporal stability, as the strains included in this DST had been isolated between 1987 and 2001. Strains typed as DST12 were isolated from patients presenting distinct clinical manifestations. This raises the intriguing proposition that the apparent dominance of DST12 in endemic locations associated with urban areas of the Atlantic rain forest region may be a consequence of higher fitness of this DST to the modified environment.

Three other *L. braziliensis* DSTs (16, 27, and 35) comprised two or three strains. All the other *L. braziliensis* were assigned to unique DSTs (n = 30). DST16 comprised two heterozygous *L. braziliensis* strains from Peru, whilst DST27 comprised two strains (zymodemes 27 and 74) from different localities (Pernambuco and Bahia, respectively), and DST35 three strains from the same zymodeme (IOC/Z27) and geographic origin (Bahia, Northern Brazil).

This demonstrates that the MLSA approach allows both detection of different genotypes and the level of separation between strains through the number of polymorphic sites.

A NeighborNet network was obtained with the concatenated sequences represented by the DSTs (Figure 3). The clusters were in agreement with MLEE for species groups, with the exception of *L. shawi*, which clusters together with all *L. guyanensis* strains. The reticulate aspects of the *L. guyanensis* group suggest recombination events occurring among the strains, including *L. shawi* (Figure 3). This same aspect was observed for the *L. braziliensis* cluster, but two strains were more divergent (IOC/L2498 and IOC/L2499, DSTs 32 and 57). Both were isolated from Acre state, a region bordering Peru. Recently we have demonstrated that these two strains clustered together with *L. peruviana* and Peruvian *L. braziliensis* by microsatellite typing [29]. However, upon removal of ICD sequences from this analysis these two strains grouped very closed to the other *L. braziliensis* strains (data not shown).

It is clear that *L. braziliensis*, *L. guyanensis*, *L. naiffi* and *L. lainsoni* all represent distinct species, forming monophyletic groups in the NeighborNet. *L. lindenbergi* and *L. utingensis* were placed close to the monophyletic groups corresponding to *L. naiffi* and *L. guyanensis* respectively, corroborating MLEE and MLMT data (unpublished). More isolates from both species should be studied to infer their taxonomic status. However, it is important to mention that these two species shared no alleles with all the other species/strains, except for *L. lindenbergi*, which shared alleles in 6PGD with

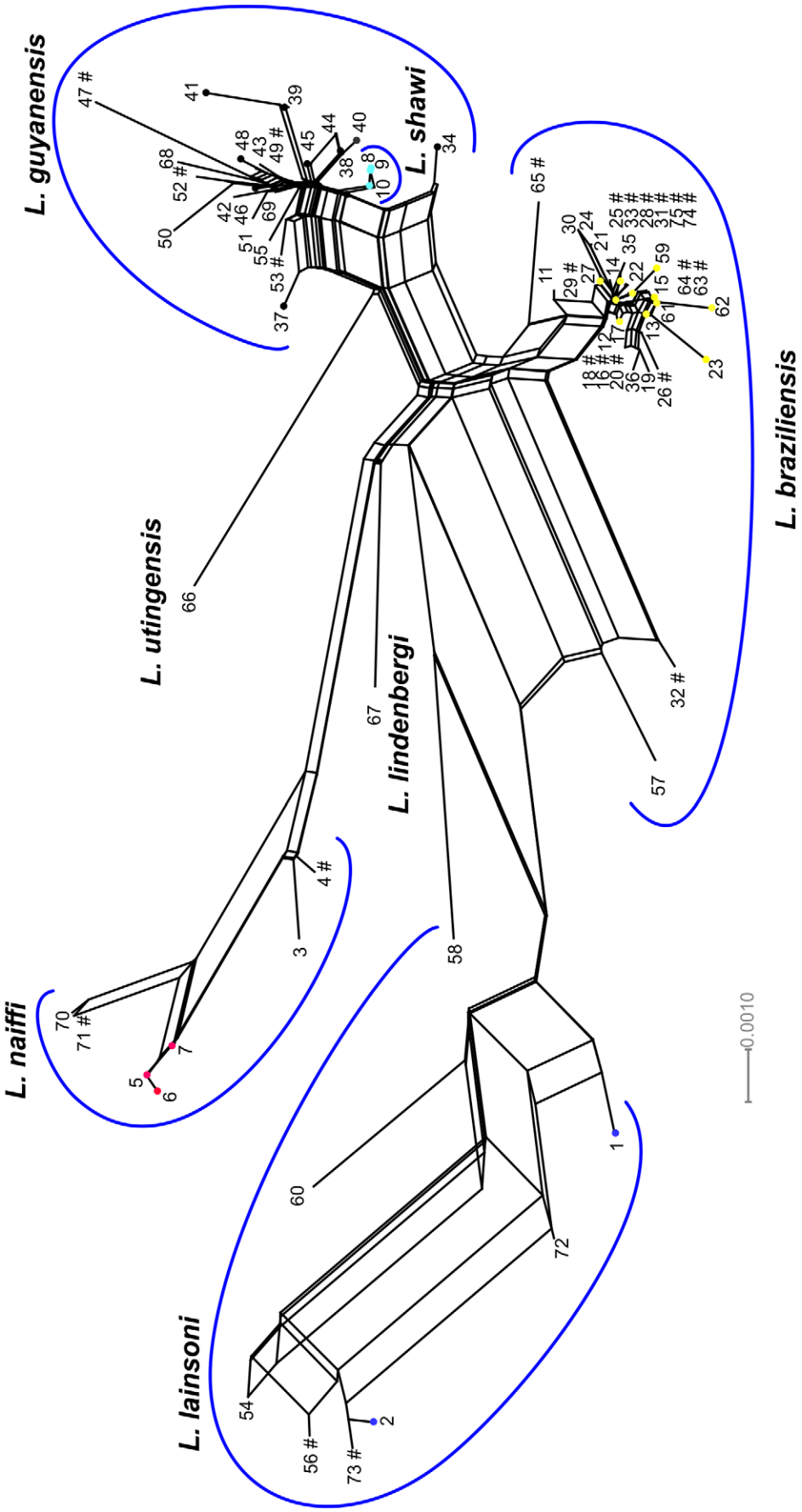
L. guyanensis
L. shawi
L. utingensis
L. lindenbergi
L. braziliensis
L. naiffi
L. lainsoni
0.0010

Figure 3. NeighborNet graph based on concatenated sequences of four gene fragments of *Leishmania (Viannia)* subgenus. A bushy network structure is observed indicating homologous recombination. # = DSTs presenting ambiguous sites (possible heterozygous) in at least one locus; DSTs gathered in clonal complex are represented by different colored circles depending on the species: yellow - *L. braziliensis*; black – *L. guyanensis*; light blue – *L. shawi*; blue – *L. lainsoni*; red – *L. naiffi*. Singletons are represented by DST number only.
doi:10.1371/journal.pntd.0001888.g003

L. braziliensis (Figure 2B). The estimate of the average genetic distance between *L. lindenbergi* or *L. utingensis* and the other *L. (Viannia)* species is comparable to those observed between each of the species analyzed. The highest values were observed between *L. lainsoni* and any other species.

Corroborating the results observed for the haplotype network, *L. naiffi* and *L. lainsoni* were the most divergent and the large splits observed suggesting no influence of genetic exchange. DST58, represented by a single *L. lainsoni* strain was positioned in a split between *L. lainsoni* and *L. naiffi*. The extensive reticulation suggests that recombination has occurred relatively frequently in *L. braziliensis* and *L. guyanensis*, as proposed by other authors [34]. As expected, *L. braziliensis* and *L. naiffi* presented higher within group average genetic distance values than *L. guyanensis*, even including *L. shawi* strains in the *L. guyanensis* group, corroborating previous studies on the genetic diversity of both *L. braziliensis* and *L. naiffi* [35,36,37,38]. However, as far as we know, this is the first report of a high level of genetic diversity within *L. lainsoni*. This was the species giving the highest within group average genetic distance value, although it consists of only two zymodemes. Reports on the occurrence of *L. lainsoni* and *L. naiffi* in distinct endemic areas have constantly increased, indicating that both species are adapted to several environments where different sand fly species participate in parasite transmission. Notably, *L. lainsoni* was the most divergent species, followed by *L. naiffi*, which might contribute to the widespread dispersion of these two species [39–41].

BURST analysis contributes to the understanding of relatedness among DSTs

Considering that a large number of unique DSTs were detected, the BURST algorithm was employed to evaluate the relationships between 52 homozygous DSTs, which corresponded to 70 strains. Here we applied eBURST to analyze the allele profiles as an efficient method to recognize natural discontinuities that would merit taxonomic status as species or subspecies. The BURST algorithm first identifies mutually exclusive groups of related genotypes in the population (typically an MLST database), and attempts to identify the founding genotype (sequence type, ST) of each group. The algorithm then predicts the descent from the predicted founding genotype to the other genotypes in the group, displaying the output as a radial diagram, centered on the predicted founding genotype. The procedure was developed for use with the data produced by MLST (STs and their allelic profiles).

The clonal complex model has proved to be valuable in analyzing many MLST datasets, making the results more amenable for epidemiological analysis. A clonal complex (CC) comprises genetically related, but not identical, organisms. In this study six CCs were observed (Figure 4); each one represented by strains classified as the same *Leishmania* species. Strains typed as *L. braziliensis* were grouped in two clonal complexes. The CCs observed were: CC1, *L. lainsoni* (comprising 2 of 6 strains; 2/6); CC2, *L. naiffi* (3/5); CC3, *L. shawi* (3/3); CC4, *L. braziliensis* (30/37); CC5, *L. braziliensis* (2/37); CC6, *L. guyanensis* (13/17). CC6 included 13 *L. guyanensis* DSTs from 17 analyzed, indicating DST37 as the predict founder genotype. CC4 included 12 out of 19 *L. braziliensis* DSTs analyzed and the predicted founder genotype was DST12, the most common DST observed for this species.

Most of the isolates were not grouped in any clonal complex, showing a high prevalence of singletons (n = 15; 4 *L. lainsoni*, 2 *L. naiffi*, 5 *L. braziliensis*, 4 *L. guyanensis*), differing at two or more loci from all other DSTs (Table S1), typical of populations with a high rate of recombination relative to mutation and for which eBURST does not reliably indicate ancestry [42]. Although eBURST does not contribute to phylogenetic analysis in *Leishmania*, this is a helpful clustering tool to analyze MLST results from this organism, which minimizes the differences observed between each *Leishmania* strain. For example, although all *L. guyanensis* strains were assigned to a distinct DST, almost all of them clustered in the same clonal group (CC5), corroborating the already demonstrated homogeneity of the *L. guyanensis* population in Brazil [19] and contrasting with the results with microsatellites showing diversity within the *L. guyanensis* complex [34]. We cannot infer the dispersion capacity of this group, but temporal stability is certainly a characteristic of this clonal complex.

Another scenario was observed for *L. braziliensis*. Out of 55 *L. braziliensis*, we obtained groups of 18, 3, 2 and 2 strains that were assigned to the same DSTs. However, two clonal complexes were observed for this species, one represented by 57.5% of the strains analyzed (CC4), the other one by only two strains (CC5) and 27.5% representing singletons. This result would fit a hypothesis of *L. braziliensis* having better fitness than *L. guyanensis*, leading to an improved dispersion. However, more *L. guyanensis* isolates from other regions not located in Brazil and its 'sister' species, *L. panamensis*, would benefit this discussion.

Although the main purpose of this study was not to investigate the epidemiological findings from MLSA of *L. (Viannia)*, it is important to mention some aspects. The capacity of dispersion and adaptation of *L. braziliensis* clonal complex CC4 is evident as well as its temporal stability. Although formed by only two *L. braziliensis* strains, CC5 represents an interesting clonal complex. The two strains were isolated from the same locality, a municipality in Pernambuco state, in northeastern Brazil, where a unique zymodeme is found. Furthermore, it is interesting that a high level of genetic diversity was already demonstrated for the *L. braziliensis* population in Pernambuco, depicted by MLEE analysis. It seems that in this region two transmission cycles co-exist; involving or not involving the sylvatic environment [35,43]. This could influence or favor the maintenance of two clonal complexes in the area.

The putative founder for the clonal cluster CC4 (*L. braziliensis*) and CC6 (*L. guyanensis*) were determined. However, it is not possible to infer if they represent the ancestor of each group. In fact, for *L. braziliensis* CC4 the putative founder is the most prevalent DST, but for *L. guyanensis* this is not true as for this species each strain was assigned as one DST.

To evaluate the robustness of each CC and the relationship between them, a NJ tree was built using the concatenated sequences of DSTs included by eBURST in clonal complexes, excluding the singletons (Figure 5). Clusters with high bootstrap supports were observed, corresponding to each CC. DST24 and DST30, from CC5, were the only two DSTs inside the *L. braziliensis* cluster forming a group supported by the bootstrap value. Inside the *L. guyanensis* group a high supported cluster was

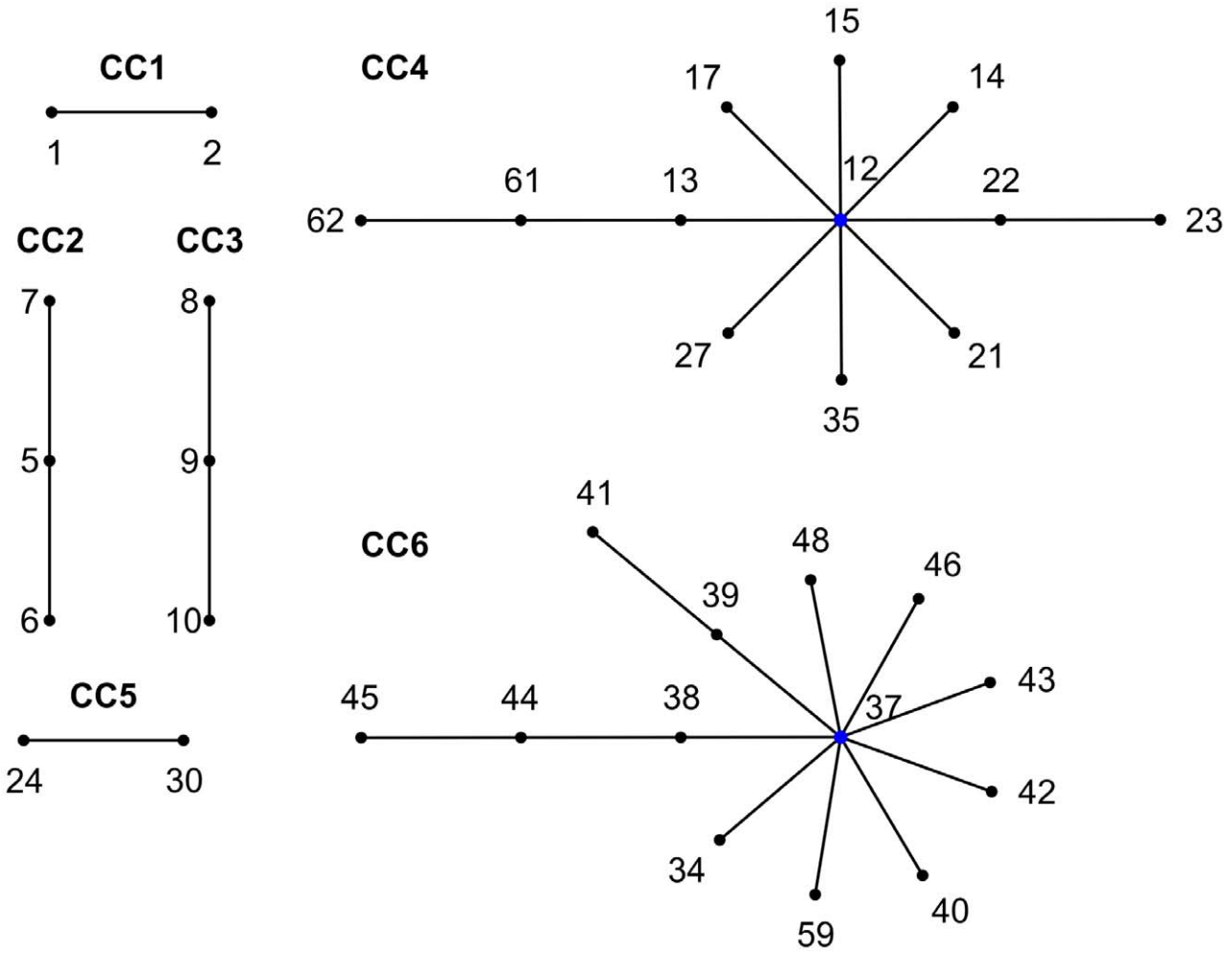

Figure 4. eBURST diagrams of analyzed *Leishmania (Viannia)* strains. Six clonal complexes were found: CC1: *L. lainsoni*; CC2 *L. naiffi*; CC3 *L. shawi*; CC4 *L. braziliensis*; CC5 *L. braziliensis*; CC6 *L. guyanensis*. Predicted founders were identified for CC4 (DST12) and CC6 (DST37; IOC/L565). Only homozygous DSTs were used (Table S1).
doi:10.1371/journal.pntd.0001888.g004

observed composed of all CC3 DSTs, corresponding to *L. shawi*. Low bootstrap values were observed among almost all DSTs from CC4 and CC6, corroborating the fact that they belong to the same complex. Evidence was again found for the divergence of *L. lainsoni* and *L. naiffi* clonal complexes.

Concluding remarks

The results obtained here strongly supported the established taxonomy of *L. (Viannia)*, considering the species that have been found in circulation in Brazil. Specifically, our data support monophyly of all but one Brazilian *L. (Viannia)* species analyzed here and highlight the close relationships between *L. braziliensis* and *L. guyanensis* and the recombination events occurring in both species. Some aspects merit special mention, however. The taxonomic validity of *L. shawi* has been questioned and indeed the markers studied here suggested that *L. shawi* isolates were closely related to or were part of the *L. guyanensis* group. The same was observed in the NeighborNet network with concatenated sequences. However, the eBURST analysis indicated that these isolates form a distinct clonal complex from *L. guyanensis*, although closely related as observed in the NeighborNet network. Recently, *hsp70* gene sequence analysis has indicated that *L. shawi* is a subgroup of *L. guyanensis* [5]. In addition, *L. shawi* and *L. guyanensis* had the lowest average genetic distance of the *L. (Viannia)* studied here. Our results, thus, indicate that *L. guyanensis* has different clonal populations, such as those observed for *L. braziliensis*, with *L. shawi* being one of them.

It is noteworthy that the MLSA clades derived here are in good agreement with MLEE clusters reported previously. Taken together, our data point to the combined four gene scheme used here as a reasonable approach that provides extensive differentiation and offers evolutionarily accurate clustering. The analysis performed herein should be extended to species other than those studied here and should be used as a starting point to develop an MLST scheme for *Leishmania* spp. genetic typing. While providing complete genome sequences is not possible as a routine approach, MLST generates evidence for similarities and differences between *Leishmania* species and/or strains, offering a number of advantages over most typing methods and providing results helpful for taxonomic, population genetics, evolutionary and, in consequence, epidemiological studies [4].

Such data are likely to revolutionize the systematics of *Leishmania*, consolidate our view of what constitutes a *Leishmania* 'species', provide evidence concerning the epidemiology of these

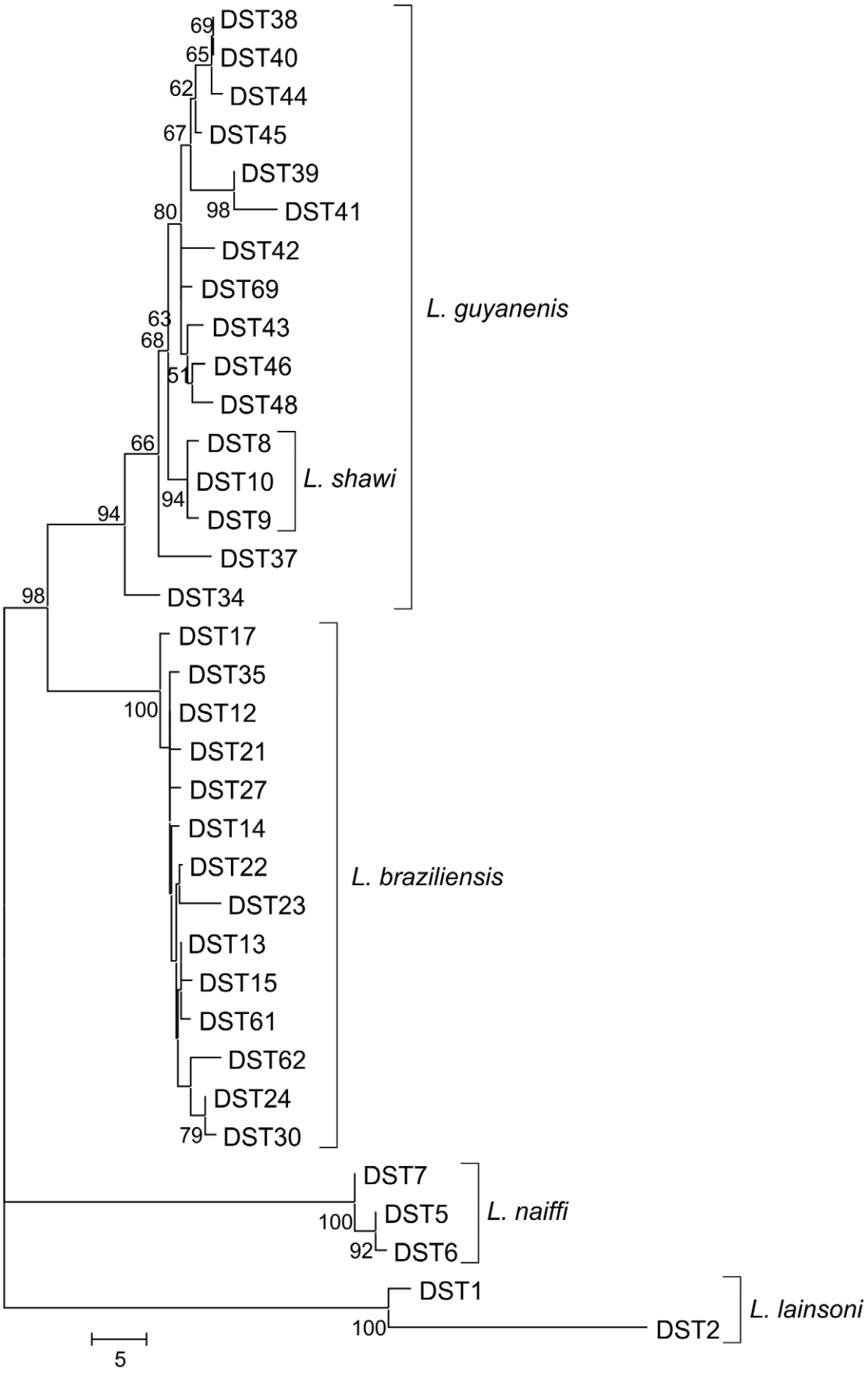
DST38
DST40
DST44
DST45
DST39
DST41
DST42
DST69
DST43
DST46
DST48
DST8
DST10
DST9
DST37
DST34
DST17
DST35
DST12
DST21
DST27
DST14
DST22
DST23
DST13
DST15
DST61
DST62
DST24
DST30
DST7
DST5
DST6
DST1
DST2
69
65
62
67
98
80
63
68
51
66
94
94
98
100
79
100
92
100
5
L. guyanenis
L. shawi
L. braziliensis
L. naiffi
L. lainsoni

Figure 5. Phylogenetic relationship of *Leishmania (Viannia)* inferred from concatenated sequences of G6PD, 6PGD, MPI, ICD fragments. The phylogenetic tree was constructed using number of differences and the NJ method. Only DSTs comprised in any clonal complexes (Figure 4 and Table S1) were included. Bootstrap values (percentages of 500 replicates) above 50% are indicated at the nodes.
doi:10.1371/journal.pntd.0001888.g005

pathogens and might even be applicable directly to clinical samples.

Definition of *Leishmania* species and knowledge of the genetic structure of all *Leishmania* species will provide a useful framework for exploring the evolutionary dynamics and phylogenetic distribution of relevant strain properties. Recognizing the urgent need for a standardized globally acceptable species definition and typing method for *Leishmania*, we are now sequencing other genes and including more species and strains in our analysis aiming to propose that species within the *Leishmania* genus could be defined as a group of strains that share a determined level of similarity in the concatenated nucleotide sequences of the genes selected. To achieve this, establishment of a consensus MLST gene set that provides optimum differentiation for *Leishmania* species and/or strains is required.

Supporting Information

Table S1 Taxonomic and collection data (clinical, geographical, biochemical and molecular Information) for *L. (Viannia)* strains used for MLSA.
(DOCX)

Table S2 Sequences retrieved from GenBank and those included in the analysis performed in the present study, indicating the assigned sequence type (ST) for each marker.
(DOCX)

Table S3 Multi-alleles sites observed for each fragment-gene alignment and the distribution of each allele in the studied strains.
(DOCX)

Table S4 Strains presenting ambiguous sites (IUPAC symbols) for the targets with the respective site position in length alignment and the most common nucleotide.
(DOCX)

Acknowledgments

We thank the Plataforma de Sequenciamento do Instituto Oswaldo Cruz (PDTIS) for performing all the sequencing. We also would like to thank Gabriel Eduardo Melim Ferreira for his contribution on the figures and for his critical review on the text and for Global Science Editing, UK.

Author Contributions

Conceived and designed the experiments: EC MCB MAM ILM. Performed the experiments: MCB. Analyzed the data: MCB EC. Contributed reagents/materials/analysis tools: EC ILM MAM. Wrote the paper: EC MCB MAM ILM.

References

1. Garin YJ, Sulahian A, Pratlong F, Meneceur P, Gangneux JP, et al. (2001) Virulence of *Leishmania infantum* is expressed as a clonal and dominant phenotype in experimental infections. Infect Immun 69: 7365–7373.
2. Arevalo J, Ramirez L, Adaui V, Zimic M, Tulliano G, et al. (2007) Influence of *Leishmania (Viannia)* species on the response to antimonial treatment in patients with American tegumentary leishmaniasis. J Infect Dis 195: 1846–1851.
3. Banuls AL, Hide M, Prugnolle F (2007) *Leishmania* and the leishmaniases: a parasite genetic update and advances in taxonomy, epidemiology and pathogenicity in humans. Adv Parasitol 64: 1–109.
4. Schonian G, Mauricio I, Cupolillo E (2010) Is it time to revise the nomenclature of *Leishmania*? Trends Parasitol.
5. da Silva LA, de Sousa CdS, da Graça GC, Porrozzi R, Cupolillo E (2010) Sequence analysis and PCR-RFLP profiling of the hsp70 gene as a valuable tool for identifying *Leishmania* species associated with human leishmaniasis in Brazil. Infect Genet Evol 10: 77–83.
6. Montalvo AM, Fraga J, Monzote L, Montano I, De Doncker S, et al. (2010) Heat-shock protein 70 PCR-RFLP: a universal simple tool for Leishmania species discrimination in the New and Old World. Parasitology 137: 1159–1168.
7. Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, et al. (1998) Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci USA 95: 3140–3145.
8. Debourgogne A, Gueidan C, Hennequin C, Contet-Audonneau N, de Hoog S, et al. (2010) Development of a new MLST scheme for differentiation of Fusarium solani Species Complex (FSSC) isolates. J Microbiol Methods 82: 319–323.
9. Odds FC (2010) Molecular phylogenetics and epidemiology of *Candida albicans*. Future Microbiol 5: 67–79.
10. Odds FC, Bougnoux ME, Shaw DJ, Bain JM, Davidson AD, et al. (2007) Molecular phylogenetics of Candida albicans. Eukaryot Cell 6: 1041–1052.
11. Mauricio IL, Gaunt MW, Stothard JR, Miles MA (2001) Genetic typing and phylogeny of the Leishmania donovani complex by restriction analysis of PCR amplified gp63 intergenic regions. Parasitology 122: 393–403.
12. Mauricio IL, Yeo M, Baghaei M, Doto D, Pratlong F, et al. (2006) Towards multilocus sequence typing of the *Leishmania donovani* complex: resolving genotypes and haplotypes for five polymorphic metabolic enzymes (ASAT, GPI, NH1, NH2, PGD). Int J Parasitol 36: 757–769.
13. Zemanova E, Jirku M, Mauricio IL, Horak A, Miles MA, et al. (2007) The *Leishmania donovani* complex: genotypes of five metabolic enzymes (ICD, ME, MPI, G6PDH, and FH), new targets for multilocus sequence typing. Int J Parasitol 37: 149–160.
14. Tsukayama P, Lucas C, Bacon DJ (2009) Typing of four genetic loci discriminates among closely related species of New World *Leishmania*. Int J Parasitol 39: 355–362.
15. Jacobsen MD, Boekhout T, Odds FC (2008) Multilocus sequence typing confirms synonymy but highlights differences between *Candida albicans* and *Candida stellatoidea*. FEMS Yeast Res 8: 764–770.
16. Jacobsen MD, Rattray AM, Gow NA, Odds FC, Shaw DJ (2008) Mitochondrial haplotypes and recombination in *Candida albicans*. Med Mycol 46: 647–654.
17. Lauthier JJ, Tomasini N, Barnabe C, Rumi MM, D'Amato AM, et al. (2012) Candidate targets for Multilocus Sequence Typing of *Trypanosoma cruzi*: validation using parasite stocks from the Chaco Region and a set of reference strains. Infect Genet Evol 12: 350–358.
18. Yeo M, Mauricio IL, Messenger LA, Lewis MD, Llewellyn MS, et al. (2011) Multilocus sequence typing (MLST) for lineage assignment and high resolution diversity studies in *Trypanosoma cruzi*. PLoS Negl Trop Dis 5: e1049.
19. Cupolillo E, Grimaldi G, Momen H (1994) A General Classification of New-World Leishmania Using Numerical Zymotaxonomy. Am J Trop Med Hyg 50: 296–311.
20. Gordon D (2002) Viewing and Editing Assembled Sequences Using Consed. Current Protocols in Bioinformatics: John Wiley & Sons, Inc.
21. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. Molecular Biology and Evolution 24: 1596–1599.
22. Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451–1452.
23. Bandelt HJ, Dress AW (1992) Split decomposition: a new and useful approach to phylogenetic analysis of distance data. Mol Phylogenet Evol 1: 242–252.
24. Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 23: 254–267.
25. Hunter PR, Gaston MA (1988) Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. J Clin Microbiol 26: 2465–2466.
26. Rogers MB, Hilley JD, Dickens NJ, Wilkes J, Bates PA, et al. (2011) Chromosome and gene copy number variation allow major structural change between species and strains of Leishmania. Genome Res 21: 2129–2142.
27. Mauricio IL, Gaunt MW, Stothard JR, Miles MA (2007) Glycoprotein 63 (gp63) genes show gene conversion and reveal the evolution of Old World *Leishmania*. Int J Parasitol 37: 565–576.
28. Cuervo P, Domont GB, De Jesus JB (2010) Proteomics of trypanosomatids of human medical importance. J Proteomics 73: 845–867.
29. Oddone R, Schweynoch C, Schonian G, de Sousa CS, Cupolillo E, et al. (2009) Development of a multilocus microsatellite typing approach for discriminating strains of *Leishmania (Viannia)* species. J Clin Microbiol 47: 2818–2825.
30. Zhang WW, Miranda-Verastegui C, Arevalo J, Ndao M, Ward B, et al. (2006) Development of a genetic assay to distinguish between *Leishmania (Viannia)* species on the basis of isoenzyme differences. Clin Infect Dis 42: 801–809.

31. Banuls AL, Dujardin JC, Guerrini F, De Doncker S, Jacquet D, et al. (2000) Is *Leishmania (Viannia) peruviana* a distinct species? A MLEE/RAPD evolutionary genetics answer. J Eukaryot Microbiol 47: 197–207.
32. Fraga J, Montalvo AM, De Doncker S, Dujardin JC, Van der Auwera G (2010) Phylogeny of *Leishmania* species based on the heat-shock protein 70 gene. Infect Genet Evol 10: 238–245.
33. Banuls AL, Jonquieres R, Guerrini F, Le Pont F, Barrera C, et al. (1999) Genetic analysis of leishmania parasites in Ecuador: are *Leishmania (Viannia) panamensis* and *Leishmania (V.) guyanensis* distinct taxa? Am J Trop Med Hyg 61: 838–845.
34. Rougeron V, De Meeus T, Hide M, Waleckx E, Dereure J, et al. (2010) A battery of 12 microsatellite markers for genetic analysis of the *Leishmania (Viannia) guyanensis* complex. Parasitology 137: 1879–1884.
35. Brito ME, Andrade MS, Mendonca MG, Silva CJ, Almeida EL, et al. (2009) Species diversity of *Leishmania (Viannia)* parasites circulating in an endemic area for cutaneous leishmaniasis located in the Atlantic rainforest region of northeastern Brazil. Trop Med Int Health 14: 1278–1286.
36. Rougeron V, De Meeus T, Hide M, Waleckx E, Bermudez H, et al. (2009) Extreme inbreeding in *Leishmania braziliensis*. Proc Natl Acad Sci USA 106: 10224–10229.
37. Cupolillo E, Brahim LR, Toaldo CB, de Oliveira-Neto MP, de Brito ME, et al. (2003) Genetic polymorphism and molecular epidemiology of *Leishmania (Viannia) braziliensis* from different hosts and geographic areas in Brazil. J Clin Microbiol 41: 3126–3132.
38. Cupolillo E, Aguiar Alves F, Brahim LR, Naiff MF, Pereira LO, et al. (2001) Recent advances in the taxonomy of the New World leishmanial parasites. Med Microbiol Immunol 190: 57–60.
39. Kato H, Gomez EA, Yamamoto Y, Calvopina M, Guevara AG, et al. (2008) Natural infection of Lutzomyia tortura with *Leishmania (Viannia) naiffi* in an Amazonian area of Ecuador. Am J Trop Med Hyg 79: 438–440.
40. van der Meide WF, Jensema AJ, Akrum RA, Sabajo LO, Lai AFRF, et al. (2008) Epidemiology of cutaneous leishmaniasis in Suriname: a study performed in 2006. Am J Trop Med Hyg 79: 192–197.
41. Simon S, Veron V, Carme B (2010) *Leishmania* spp. identification by polymerase chain reaction-restriction fragment length polymorphism analysis and its applications in French Guiana. Diagn Microbiol Infect Dis 66: 175–180.
42. Turner KM, Feil EJ (2007) The secret life of the multilocus sequence type. Int J Antimicrob Agents 29: 129–135.
43. Brandao-Filho SP, Brito ME, Carvalho FG, Ishikawa EA, Cupolillo E, et al. (2003) Wild and synanthropic hosts of *Leishmania (Viannia) braziliensis* in the endemic cutaneous leishmaniasis locality of Amaraji, Pernambuco State, Brazil. Trans R Soc Trop Med Hyg 97: 291–296.

Table S1: Taxonomic and collection data (clinical, geographical, biochemical and molecular Information) for *L. (Viannia)* strains used for MLSA.

IOC/L	Species	International code	Geographic origin (Brazilian state)	Clinical form	Zymodeme	MPI	G6PD	6PGD	ICD	DST	CC
1023	*L. lainsoni*	MHOM/BR/1981/M6426	PA	CL	15	1	1	2	1	1	CC1
1266	*L. lainsoni*	MAGO/BR/1983/IM1721	PA	CL	15	2	1	2	1	2	CC1
855	*L. naiffi*	MHOM/BR/1986/IM2773	AM	*	36	3	2	1	2	3	singl
1365	*L. naiffi*	MDAS/BR/1979/M5533	PA	VL	36	4	3	28	3; 4	4	htz
854	*L. naiffi*	ISQU/BR/1985/IM2264	PA	*	38	4	7	6	3	5	CC2
1939	*L. naiffi*	ISQU/BR/1994/IM3936	AM	*	49	4	7	6	10	6	CC2
1953	*L. naiffi*	MHOM/BR/1994/IM4000	AM	CL	49	4	7	7	3	7	CC2
1545	*L. shawi*	MCEB/BR/1984/M8408	PA	CL	26	5	4	3	5	8	CC3
3199	*L. shawi*	MHOM/BR/1999/M17997	PA	*	*	5	4	3	8	9	CC3
3200	*L. shawi*	MHOM/BR/1999/M17998	PA	*	*	5	4	3	9	10	CC3
1731	*L. braziliensis*	MAGO/BR/1992/IM154	RO	CL	53	6	5	1	6	11	singl
849	*L. braziliensis*	MHOM/BR/1987/J.CARLOS	RJ	MCL	27	7	6	1	6	12	CC4
2139	*L. braziliensis*	MHOM/BR/1996/GBS	PB	CL	27	7	6	1	6	12	CC4
2159	*L. braziliensis*	MHOM/BR/1996/JPS	RJ	CL	27	7	6	1	6	12	CC4
2287	*L. braziliensis*	MHOM/BR/1998/AFS	PE	CL	72	7	6	1	6	12	CC4
2288	*L. braziliensis*	MHOM/BR/1996/MAS	PE	CL	45	7	6	1	6	12	CC4
2291	*L. braziliensis*	MHOM/BR/1997/ASB	PE	MCL	73	7	6	1	6	12	CC4
2419	*L. braziliensis*	MHOM/BR/1999/JAS	PE	CL	73	7	6	1	6	12	CC4
2420	*L. braziliensis*	MHOM/BR/1999/PCS	PE	CL	45	7	6	1	6	12	CC4
2427	*L. braziliensis*	MHOM/BR/1999/SJB	PE	CL	74	7	6	1	6	12	CC4
2467	*L. braziliensis*	MHOM/BR/2001/LTCP14182	BA	DL	27	7	6	1	6	12	CC4
2472	*L. braziliensis*	MHOM/BR/2001/LTCP13455	BA	CL	27	7	6	1	6	12	CC4
2475	*L. braziliensis*	MHOM/BR/2001/LTCP14214	BA	CL	27	7	6	1	6	12	CC4
2480	*L. braziliensis*	MHOM/BR/2001/LTCP13980	BA	MCL	27	7	6	1	6	12	CC4
2510	*L. braziliensis*	MHOM/BR/2000/CEA	PE	CL	27	7	6	1	6	12	CC4
2515	*L. braziliensis*	MHOM/BR/2001/TSS	PE	CL	78	7	6	1	6	12	CC4
2660	*L. braziliensis*	MHOM/BR/2004/ARARAQUARA-1	SP	CL	27	7	6	1	6	12	CC4
2836	*L. braziliensis*	MHOM/BR/1994/LTCP9845	BA	CL	27	7	6	1	6	12	CC4
2838	*L. braziliensis*	MHOM/BR/1997/LTCP11245	BA	CL	27	7	6	1	6	12	CC4

IOC/L	Species	International code	Geographic origin (Brazilian state)	Clinical form	Zymodeme	MPI	G6PD	6PGD	ICD	DST	CC
2148	*L. braziliensis*	MHOM/BR/2001/VLNC	PE	MCL	105	7	6	1	11	13	CC4
2481	*L. braziliensis*	MHOM/BR/2000/LTCP13490	BA	CL	27	7	6	1	14	14	CC4
2502	*L. braziliensis*	MHOM/BR/2002/NMT-RBO040	AC	CL	82	7	6	1	25	15	CC4
iz 26	*L. braziliensis*	MORY/PE/84/AO23	-	*	*	7	6	1	40; 41	16	htz
iz 28	*L. braziliensis*	MRAT/PE/84/A1	-	*	*	7	6	1	40; 41	16	htz
iz 31	*L. braziliensis*	MORY/PE/84/AC20	-	*	*	7	6	1	42	17	CC4
iz 27	*L. braziliensis*	MORY/PE/84/AO8	-	*	*	7	6	1	6; 41	18	htz
2491	*L. braziliensis*	MHOM/BR/2002/NMT-RBO 005	AC	CL	78	7	6	8	16	19	singl
2535	*L. braziliensis*	MHOM/BR/2001/HC-JS	ES	CL	27	7	6	8	12; 27	20	htz
2344	*L. braziliensis*	MHOM/BR/1997/NMT-MAO 229P	AM	CL	*	7	6	10	6	21	CC4
2693	*L. braziliensis*	MNEC/BR/2003/NECTOMYS	PE	CL	74	7	6	12	6	22	CC4
2495	*L. braziliensis*	MHOM/BR/2002/NMT-RBO025	AC	CL	78	7	6	12	19	23	CC4
2950	*L. braziliensis*	MHOM/BR/2006/CEN	PE	CL	75	7	6	17	33	24	CC5
2494	*L. braziliensis*	MHOM/BR/2002/NMT-RBO 018	AC	CL	80	7	6	1; 11	18	25	htz
1734	*L. braziliensis*	MHOM/BR/1991/IM3713	AM	CL	35	7	6	4; 5	7	26	htz
2509	*L. braziliensis*	MHOM/BR/2001/CRFN	PE	CL	74	7	18	1	6	27	CC4
2929	*L. braziliensis*	MHOM/BR/2005/NMT-LTCP16011-P	BA	CL	27	7	18	1	6	27	CC4
iz 33	*L. braziliensis*	MORY/PE/84/ABR23	-	*	*	7	18	1	6; 41	28	htz
2833	*L. braziliensis*	MHOM/BR/2001/LTCP14349	BA	MCL	27	7	18	1; 8	6	29	htz
2951	*L. braziliensis*	MHOM/BR/2006/CM	PE	CL	75	7	19	17	33	30	CC5
2511	*L. braziliensis*	MHOM/BR/2001/JS	PE	CL	105	7	11; 16	1	6	31	htz
2499	*L. braziliensis*	MHOM/BR/2002/NMT-RBO 035	AC	CL	81	7	12; 13	1	23	32	htz
2468	*L. braziliensis*	MHOM/BR/2001/LTCP14183	BA	MCL	27	7	6; 11	1	6	33	htz
iz 18	*L. guyanensis*	MCHO/BR/80/M6202	*	*	*	8	4	9	9	34	CC6
2466	*L. braziliensis*	MHOM/BR/2001/LTCP13183	BA	CL	27	8	6	1	6	35	CC4
2476	*L. braziliensis*	MHOM/BR/2001/LTCP14278	BA	DL	27	8	6	1	6	35	CC4
2847	*L. braziliensis*	MHOM/BR/2002/LTCP15476	BA	CL	27	8	6	1	6	35	CC4
2152	*L. braziliensis*	MHOM/BR/1996/AFN	RJ	CL	27	8	6	8	12	36	singl
565	*L. guyanensis*	MHOM/BR/1975/M4147	PA	CL	23	9	4	9	9	37	CC6
2963	*L. guyanensis*	MHOM/BR/2007/033-MECM	AM	CL	23	9	4	9	13	38	CC6
2937	*L. guyanensis*	MHOM/BR/2007/065	AM	CL	23	9	4	9	30	39	CC6

IOC/L	Species	International code	Geographic origin (Brazilian state)	Clinical form	Zymodeme	MPI	G6PD	6PGD	ICD	DST	CC
2971	*L. guyanensis*	MHOM/BR/2007/021-R	AM	CL	23	9	4	9	37	40	CC6
2969	*L. guyanensis*	MHOM/BR/2007/AC	AM	CL	23	9	4	26	30	41	CC6
2334	*L. guyanensis*	MHOM/BR/1997/NMT-MAO 202P	AM	CL	23	9	8	9	9	42	CC6
2341	*L. guyanensis*	MHOM/BR/1997/NMT-MAO 223P	AM	CL	23	9	9	9	9	43	CC6
2350	*L. guyanensis*	MHOM/BR/1997/NMT-MAO 237P	AM	CL	23	9	10	9	13	44	CC6
2970	*L. guyanensis*	MHOM/BR/2007/021-HMB	AM	CL	23	9	10	9	36	45	CC6
2936	*L. guyanensis*	MHOM/BR/2007/069	AM	CL	23	9	22	9	9	46	CC6
2938	*L. guyanensis*	MHOM/BR/2007/063	AM	CL	23	9	23	19	31; 32	47	htz
2956	*L. guyanensis*	MHOM/BR/2007/011	AM	CL	23	9	24	9	9	48	CC6
2957	*L. guyanensis*	MHOM/BR/2007/014-JIS	AM	CL	23	9	25	9; 20	9	49	htz
2961	*L. guyanensis*	MHOM/BR/2007/019-WDSN	AM	CL	23	9	27	22	9	50	singl
2962	*L. guyanensis*	MHOM/BR/2007/031-LOP	AM	CL	23	9	28	23	35	51	singl
2966	*L. guyanensis*	MHOM/BR/2007/039	AM	CL	23	9	29	24; 25	9	52	htz
iz 34	*L. guyanensis*	MCHO/BR/80/M6200	*	*	*	9	4; 33	32	9	53	htz
2490	*L. lainsoni/ L. naiffi*	MHOM/BR/2002/NMT-RBO004	AC	CL	87	11	14	2	15	54	singl
2493	*L. guyanensis*	MHOM/BR/2002/NMT-RBO 013	AC	CL	110	12	4	9	17	55	singl
2497	*L. lainsoni*	MHOM/BR/2002/NMT-RBO 027P	AC	CL	86	13	1	13	20; 21	56	htz
2498	*L. braziliensis*	MHOM/BR/2002/NMT-RBO 029	AC	CL	84	14	15	14	22	57	singl
2500	*L. lainsoni*	MHOM/BR/2002/NMT-RBO036	AC	CL	86	15	17	1	24	58	singl
2501	*L. braziliensis*	MHOM/BR/2002/NMT-RBO037	AC	CL	79	16	6	15	6	59	singl
2503	*L. lainsoni*	MHOM/BR/2002/NMT-RBO044	AC	CL	86	17	14	16	26	60	singl
566	*L. braziliensis*	MHOM/BR/1975/M2903	PA	CL	27	18	6	1	11	61	CC4
921	*L. braziliensis*	MHOM/BR/1987/H-210	CE	CL	27	18	6	27	11	62	CC4
918	*L. braziliensis*	MRAT/BR/1987/C18.454	CE	VL	27	18	6	8; 11	11	63	htz
2538	*L. braziliensis*	MHOM/BR/2002/EMM	RJ	CL	27	18	19	8; 17	6; 11	64	htz
2541	*L. braziliensis*	MHOM/BR/1999/MJAA-II	PE	CL	74	19	6; 16	1	6	65	htz
2689	*L. utingensis*	ITUB/BR/1977/M4964	PA	*	101	20	20	18	28	66	singl
2690	*L. lindenbergi*	MHOM/BR/1966/M15733	PA	CL	102	21	21	1	29	67	singl
2960	*L. guyanensis*	MHOM/BR/2007/029-ZAV	AM	CL	23	22	26	21	34	68	singl
2964	*L. guyanensis*	MHOM/BR/2007/034-MFPS	AM	CL	23	23	4	9	9	69	singl
992	*L. naiffi*	MDAS/BR/1987/IM3280	PA	VL	42	24	30	28	3	70	singl

IOC/L	Species	International code	Geographic origin (Brazilian state)	Clinical form	Zymodeme	MPI	G6PD	6PGD	ICD	DST	CC
995	*L. naiffi*	MDAS/BR/1987/IM3292	PA	VL	36	24	30; 31	29	38	71	htz
iz 35	*L. lainsoni*	IUBI/BR/00/M12025	PA	*	*	25	1	33	43	72	singl
iz 25	*L. lainsoni*	MHOM/BR/82/M6887	PA	*	*	25	1; 32	30; 31	39	73	htz
2463	*L. braziliensis*	MHOM/BR/2001/JOLIVAL	BA	MCL	27	7; 10	6; 11	1	6	74	htz
2571	*L. braziliensis*	MHOM/BR/2003/NJS	MS	CL	7	7; 8	6	1	6	75	htz

From the left: IOC/L (Institute Oswaldo Cruz Leishmania Collection) – culture collection code; *Leishmania* species; International code – see WHO, 2010 for description, BR= Brazil and PE= Peru; Brazilian states – AC: Acre, AM: Amazonas, BA: Bahia, CE: Ceará, PA: Pará, PB: Paraíba, PE: Pernambuco, ES: Espírito Santo, MS: Mato Grosso do Sul, RJ: Rio de Janeiro, RO: Rondônia, SP: São Paulo; clinical form observed in the host: VL – visceral leishmaniasis, CL – cutaneous leishmaniasis, MCL – mucocutaneous leishmaniasis; zymodemes according to the CLIOC system ; sequence types observed for each marker; DST: diploid sequence type; CC: clonal complexes formed after e-burst analysis; singl: singletons after e-burst analysis; htz: heterozygous strains – not included in the e- burst analysis. * Information not available.

Table S2. Sequences retrieved from GenBank and those included in the analysis performed in the present study, indicating the assigned sequence type (ST) for each marker.

Marker	GenBank available sequences included in the ST alignments	Full sequence cover?*	ST
6PGD	*Leishmania braziliensis* MHOM/BR/75/M2904 6-phosphogluconate dehydrogenasedecarboxylating (LbrM34_V2.3250)	yes	1
	Leishmania peruviana strain MHOM/PE/87/PAB2941 6-phosphogluconate dehydrogenase (6pgd) gene complete cds	yes	1
	Leishmania peruviana strain MHOM/PE/92/LEH1563 6-phosphogluconate dehydrogenase (6pgd) gene complete cds	yes	1
	Leishmania peruviana strain MHOM/PE/92/LEH1384 6-phosphogluconate dehydrogenase (6pgd) gene complete cds	yes	1
	Leishmania peruviana strain MHOM/PE/92/LEH1339 6-phosphogluconate dehydrogenase (6pgd) gene complete cds	yes	1
	Leishmania peruviana strain MHOM/PE/91/LEH0521 6-phosphogluconate dehydrogenase (6pgd) gene complete cds	yes	1
	Leishmania peruviana strain MHOM/PE/87/PAB2885 6-phosphogluconate dehydrogenase (6pgd) gene complete cds	yes	1
	Leishmania peruviana strain MHOM/PE/92/LEH1449 6-phosphogluconate dehydrogenase (6pgd) gene complete cds	yes	1
	Leishmania peruviana strain MHOM/PE/92/LEH1347 6-phosphogluconate dehydrogenase (6pgd) gene complete cds	yes	1
	Leishmania peruviana strain MHOM/PE/91/LEH0547 6-phosphogluconate dehydrogenase (6pgd) gene complete cds	yes	1
	Leishmania peruviana strain MHOM/PE/87/PAB2880 6-phosphogluconate dehydrogenase (6PGD) gene complete cds	yes	1
	Leishmania braziliensis strain MHOM/PE/93/LEH1642 6-phosphogluconate dehydrogenase (6pgd) gene complete cds	yes	15
	Leishmania braziliensis strain MHOM/PE/94/LEH1829 6-phosphogluconate dehydrogenase (6pgd) gene complete cds	yes	15
	Leishmania braziliensis strain MHOM/PE/91/LEH0375 6-phosphogluconate dehydrogenase (6pgd) gene complete cds	yes	15
	Leishmania braziliensis strain MHOM/PE/93/LEH0022 6-phosphogluconate dehydrogenase (6pgd) gene complete cds	yes	15
	Leishmania braziliensis strain MHOM/PE/88/BAA2079 6-phosphogluconate dehydrogenase (6PGD) gene complete cds	yes	New-35
	Leishmania braziliensis strain MHOM/BR/84/LTB300 6-phosphogluconate dehydrogenase (6PGD) gene complete cds	yes	1
	Leishmania panamensis strain MHOM/PA/71/LS94 6-phosphogluconate dehydrogenase (6PGD) gene complete cds	yes	New-36
	Leishmania guyanensis strain MHOM/BR/75/M4147 6-phosphogluconate dehydrogenase (6PGD) gene complete cds	no	..
	Leishmania lainsoni strain MHOM/PE/88/BAB1730 6-phosphogluconate dehydrogenase (6PGD) gene complete cds	no	..
	Leishmania braziliensis 6-phosphogluconate dehydrogenase gene partial cds	no	..
	Leishmania guyanensis 6-phosphogluconate dehydrogenase gene partial cds	no	..
	Leishmania panamensis 6-phosphogluconate dehydrogenase gene partial cds	no	..
	Leishmania lainsoni 6-phosphogluconate dehydrogenase gene partial cds	no	..
G6PD	*Leishmania braziliensis* strain MHOM/BR/1975/M2903 glucose-6-phosphate dehydrogenase (G6PD) mRNA partial cds	no	...
	Leishmania peruviana strain MHOM/PE/1984/LC39 glucose-6-phosphate dehydrogenase (G6PD) mRNA partial cds	no	...
	Leishmania guyanensis strain MHOM/BR/1975/M4147 glucose-6-phosphate dehydrogenase (G6PD) mRNA complete cds	no	...

Marker	GenBank available sequences included in the ST alignments	Full sequence cover?*	ST
	Leishmania panamensis strain MHOM/PA/1971/LS94 glucose-6-phosphate dehydrogenase (G6PD) mRNA partial cds	no	...
	Leishmania lainsoni strain MHOM/BR/1981/M6426 glucose-6-phosphate dehydrogenase (G6PD) mRNA partial cds	no	...
	Leishmania braziliensis strain MHOM/BR/84/LTB300 glucose-6-phosphate dehydrogenase (G6PD) gene complete cds	yes	6
	Leishmania braziliensis strain MHOM/PE/88/BAA2079 glucose-6-phosphate dehydrogenase (G6PD) gene complete cds	yes	New-34
	Leishmania braziliensis strain MHOM/PE/88/BAA2072 glucose-6-phosphate dehydrogenase (G6PD) gene complete cds	yes	6
	Leishmania braziliensis strain MDID/BR/1995/IM4159 glucose-6-phosphate dehydrogenase (G6PD) mRNA partial cds	no	...
	Leishmania braziliensis strain MAGO/BR/1992/IM154 glucose-6-phosphate dehydrogenase (G6PD) mRNA partial cds	no	...
	Leishmania peruviana strain MHOM/PE/87/PAB2880 glucose-6-phosphate dehydrogenase (G6PD) gene complete cds	yes	6
	Leishmania guyanensis strain MHOM/BR/75/M4147 glucose-6-phosphate dehydrogenase (G6PD) gene complete cds	yes	4
	Leishmania panamensis strain MHOM/PA/71/LS94 glucose-6-phosphate dehydrogenase (G6PD) gene complete cds	yes	New-36
	Leishmania lainsoni strain MHOM/PE/88/BAB1730 glucose-6-phosphate dehydrogenase (G6PD) gene complete cds	yes	New-37
	Leishmania braziliensis MHOM/BR/75/M2904_glucose-6-phosphate_dehydrogenase_(LbrM20_V2.0160)_partial_mRNA	yes	New-35
ICD	*Leishmania braziliensis* MHOM/BR/75/M2904 isocitrate dehydrogenase putative (LbrM33_V2.2820) partial mRNA	yes	6
MPI	*Leishmania peruviana* strain MHOM/PE/87/PAB2941 mannose phosphate isomerase (mpi) gene complete cds	yes	New-7;26**
	Leishmania peruviana strain MHOM/PE/92/LEH1339 mannose phosphate isomerase (mpi) gene complete cds	yes	7
	Leishmania peruviana strain MHOM/PE/92/LEH1384 mannose phosphate isomerase (mpi) gene complete cds	yes	7
	Leishmania peruviana strain MHOM/PE/92/LEH1449 mannose phosphate isomerase (mpi) gene complete cds	yes	7
	Leishmania peruviana strain MHOM/PE/91/LEH0521 mannose phosphate isomerase (mpi) gene complete cds	yes	New-27
	Leishmania peruviana strain MHOM/PE/92/LEH1563 mannose phosphate isomerase (mpi) gene complete cds	yes	7
	Leishmania peruviana strain MHOM/PE/92/LEH1347 mannose phosphate isomerase (mpi) gene complete cds	yes	7
	Leishmania peruviana strain MHOM/PE/87/PAB2885 mannose phosphate isomerase (mpi) gene complete cds	yes	New-7;26**
	Leishmania peruviana strain MHOM/PE/91/LEH0547 mannose phosphate isomerase (mpi) gene complete cds	yes	7
	Leishmania peruviana strain MHOM/PE/87/PAB2880 mannose phosphate isomerase (MPI) gene complete cds	yes	7
	Leishmania peruviana strain MHOM/PE/90/HB44 mannose phosphate isomerase gene complete cds	yes	7
	Leishmania braziliensis strain MHOM/PE/93/LEH1642 mannose phosphate isomerase (mpi) gene complete cds	yes	7
	Leishmania braziliensis strain MHOM/PE/94/LEH1829 mannose phosphate isomerase (mpi) gene complete cds	yes	7
	Leishmania braziliensis strain MHOM/PE/93/LEH0022 mannose phosphate isomerase (mpi) gene complete cds	yes	7
	Leishmania braziliensis strain MHOM/PE/91/LEH0375 mannose phosphate isomerase (mpi) gene complete cds	yes	7
	Leishmania braziliensis strain MHOM/BR/84/LTB300 mannose phosphate isomerase (MPI) gene complete cds	yes	7
	Leishmania braziliensis strain MHOM/PE/88/BAA2079 mannose phosphate isomerase (MPI) gene complete cds	yes	New-28

Marker	GenBank available sequences included in the ST alignments	Full sequence cover?*	ST
	Leishmania braziliensis strain MHOM/PE/88/BAA2072 mannose phosphate isomerase (MPI) gene complete cds	yes	7
	Leishmania braziliensis MHOM/BR/75/M2904 phosphomannose isomerase putative (LbrM32_V2.1750) partial mRNA	yes	7
	Leishmania braziliensis strain MHOM/BR/91/2195 mannose phosphate isomerase gene complete cds	yes	7

Available sequences for each marker included in the Sequence Types (STs) alignment for comparison regarding coverage* – which allows comparison or not – and ST found for the respective strain. ** A new heterozygous sequence type.

Table S3. Multi-alleles sites observed for each fragment-gene alignment and the distribution of each allele in the studied strains.

Marker	Position	Nucleotide	Number of strains presenting the nucleotide / total number of strains
G6PD	179	A	5/7 *L. naiffi*
		C	21/22 *L. guyanensis*; 3/*3 L. shawi*; 1/1 *L. utingensis*
		G	All other strains
	341	C	7/7 *L. naiffi*
		A	21/22 *L. guyanensis*; 3/3 *L. shawi*
		G	All other strains
	467	T	1/55 *L. braziliensis*
		A	7/7 *L. naiffi*
		C	All other strains
6PGD	9	T	3/7 *L. naiffi*; 2/55 *L. braziliensis*
		G	1/22 *L. guyanensis*
		C	All other strains
	90	A	3/8 *L. lainsoni*; 2/22 *L. guyanensis*
		C	3/7 *L. naiffi*; 1/8 *L. lainsoni*; 2/55 *L. braziliensis*
		T	All other strains
ICD	52	C	1/55 *L. braziliensis*
		T	1/3 *L. shawi*
		G	All other strains
	97	T	1/1 *L. utingensis*; 1/55 *L. braziliensis*
		C	1/22 *L. guyanensis*
		G	All other strains

Table S4. Strains presenting ambiguous sites (IUPAC symbols) for the targets with the respective site position in length alignment and the most common nucleotide.

		G6PD						6PGD					MPI			ICD					
IOC/L	***Species***	**228**	**316**	**416**	**524**	**538**	**557**	**391**	**460**	**466**	**562**	**602**	**259**	**319**	**480**	**147**	**174**	**318**	**384**	**401**	**492**
	Nucleotides*	G	G	G	C	C	G	C/G/T	T/C	C/T	C	A	A	C/T	A	A/C	C/T	C	G/A	T	T/C
918	*L. braziliensis*	...	...	...	...	...	...	...	Y	...	...	...	...	...	...	...	...	...	...	...	...
1734	*L. braziliensis*	...	...	...	...	...	...	S	...	...	...	...	...	...	...	...	...	...	...	...	...
2463	*L. braziliensis*	...	...	...	Y	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2468	*L. braziliensis*	...	...	...	Y	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2494	*L. braziliensis*	...	...	...	...	...	...	S	...	...	...	...	...	...	...	...	...	...	...	...	...
2499	*L. braziliensis*	...	...	...	Y	Y	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2511	*L. braziliensis*	...	...	...	...	...	...	...	...	...	...	...	W	....	....	...	...	...	...	...	...
2535	*L. braziliensis*	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	Y	...	...	...	...
2538	*L. braziliensis*	...	...	...	...	...	...	S	Y	Y	...	...	...	...	...	M	Y	...	...	...	...
2541	*L. braziliensis*	...	...	...	...	Y	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2571	*L. braziliensis*	...	...	...	...	...	...	...	...	...	...	...	...	Y	...	...	...	...	...	...	...
2833	*L. braziliensis*	...	...	...	...	...	...	...	Y	...	...	...	...	...	...	...	...	...	...	...	...
iz 26	*L. braziliensis*	...	...	...	...	...	...	...	...	...	...	...	...	...	...	M	Y	...	...	...	...
iz 27	*L. braziliensis*	...	...	...	Y	...	...	...	...	...	...	...	...	...	...	M	Y	...	...	...	Y
iz 28	*L. braziliensis*	...	...	...	...	...	...	...	...	...	...	...	...	...	...	M	Y	...	...	...	...
iz 31	*L. braziliensis*	...	...	...	...	...	...	...	...	...	...	...	...	...	...	M	...	...	...	...	...
iz 33	*L. braziliensis*	...	...	...	...	...	...	...	...	...	...	...	...	...	...	M	Y	...	...	...	Y
2938	*L. guyanensis*	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	Y	...
iz 34	*L. guyanensis*	R	...	R	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2957	*L. guyanensis*	...	...	...	...	...	...	...	...	...	...	R	...	...	...	...	...	...	...	...	...
2966	*L. guyanensis*	...	...	...	...	...	...	...	...	...	...	R	...	...	...	...	...	...	...	...	...
2497	*L. lainsoni*	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	Y	...	...
iz 25	*L. lainsoni*	...	...	...	...	...	R	...	...	...	Y	...	...	...	...	...	...	...	...	...	...
854	*L. naiffi*	...	...	...	...	...	...	S	Y	...	...	...	...	...	...	...	...	...	...	...	...

995	*L. naiffi*	...	R	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
1365	*L. naiffi*	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	Y	...	...	...	
....	***L. peruviana*	...	...	...	...	...	...	...	...	...	...	...	....	....	Y	...	...	...	...	...	...	

*Nucleotides: possible nucleotides observed in the alignment, presented accordingly to their frequencies, most frequent to less frequent observed. **Sequence retrieved from GenBank for MPI.

4.2. Capítulo II

Multilocus sequence analysis for *Leishmania braziliensis* outbreak investigation.

Marlow MA, **Boité MC**, Ferreira GEM, Steindel M, Cupolillo E
PLoS Neglected Tropical Diseases
2014, PNTD-D-13-01511R2 10.1371
Artigo publicado

Neste artigo apresentamos a avaliação do painel MLSA desenvolvido nesta tese como ferramenta epidemiológica. Para isso, analisamos diversas cepas identificadas como *L.* (*V.*) *braziliensis* e que foram isoladas de diferentes regiões brasileiras, incluindo cepas isoladas após surto de LTA no estado de Santa Catarina. A utilização de cepas isoladas de um surto recente, contendo informações epidemiológicas importantes, permitiu que a abordagem MLSA fosse aplicada e testada como ferramenta epidemiológica em uma situação real. A associação entre os grupos formados após MLSA e as características epidemiológicas dos pacientes apresentou um padrão que permitiu a separação de casos importados e autóctones em Santa Catarina. Os resultados demonstraram que o MLSA apresenta resolução adequada para investigação de surtos e, portanto tem potencial para uso como ferramenta na epidemiologia molecular das leishmanioses. Com este estudo foi alcançado o terceiro objetivo apresentado pela presente tese.

OPEN ACCESS Freely available online

PLOS | NEGLECTED TROPICAL DISEASES

Multilocus Sequence Analysis for *Leishmania braziliensis* Outbreak Investigation

Mariel A. Marlow[1], Mariana C. Boité[2], Gabriel Eduardo M. Ferreira[2], Mario Steindel[1], Elisa Cupolillo[2]*

1 Laboratório de Protozoologia, Departamento de Microbiologia, Imunologia e Parasitologia, Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina, Brazil, **2** Laboratório de Pesquisa em Leishmaniose, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

Abstract

With the emergence of leishmaniasis in new regions around the world, molecular epidemiological methods with adequate discriminatory power, reproducibility, high throughput and inter-laboratory comparability are needed for outbreak investigation of this complex parasitic disease. As multilocus sequence analysis (MLSA) has been projected as the future gold standard technique for *Leishmania* species characterization, we propose a MLSA panel of six housekeeping gene loci (6*pgd*, *mpi*, *icd*, *hsp*70, *mdh*mt, *mdh*nc) for investigating intraspecific genetic variation of *L.* (*Viannia*) *braziliensis* strains and compare the resulting genetic clusters with several epidemiological factors relevant to outbreak investigation. The recent outbreak of cutaneous leishmaniasis caused by *L.* (*V.*) *braziliensis* in the southern Brazilian state of Santa Catarina is used to demonstrate the applicability of this technique. Sequenced fragments from six genetic markers from 86 *L.* (*V.*) *braziliensis* strains from twelve Brazilian states, including 33 strains from Santa Catarina, were used to determine clonal complexes, genetic structure, and phylogenic networks. Associations between genetic clusters and networks with epidemiological characteristics of patients were investigated. MLSA revealed epidemiological patterns among *L.* (*V.*) *braziliensis* strains, even identifying strains from imported cases among the Santa Catarina strains that presented extensive homogeneity. Evidence presented here has demonstrated MLSA possesses adequate discriminatory power for outbreak investigation, as well as other potential uses in the molecular epidemiology of leishmaniasis.

Citation: Marlow MA, Boité MC, Ferreira GEM, Steindel M, Cupolillo E (2014) Multilocus Sequence Analysis for *Leishmania braziliensis* Outbreak Investigation. PLoS Negl Trop Dis 8(2): e2695. doi:10.1371/journal.pntd.0002695

Editor: Gabriele Schönian, Charité University Medicine Berlin, Germany

Received October 1, 2013; **Accepted** December 27, 2013; **Published** February 13, 2014

Funding: MAM and MS were funded by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) and MCB, GEMF, and EC were funded by FIOCRUZ (Fundação Oswaldo Cruz), CNPq and FAPERJ (Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

* E-mail: ecupoli@ioc.fiocruz.br

Introduction

Leishmaniasis, a vector-borne disease caused by protozoan parasites of genus *Leishmania* [1], represents one of the highest disease burdens among the neglected tropical diseases in developing nations [2]. While not often fatal like the visceral form, the cutaneous form of the disease contributes substantially to leishmaniasis disease burden as it requires a lengthy and costly treatment regimen, results in apparent scarring, and can progress to a severely disfiguring mucosal form [1]. In recent years, leishmaniasis outbreaks have been described with increasing frequency [3–5], including those in sub-tropical regions or regions not previously endemic across the global [6–8]. In Brazil, beginning in 2005, an outbreak of human cutaneous leishmaniasis occurred in the southern Brazilian state of Santa Catarina, where the disease had not been observed previously as endemic. Overtime, cutaneous leishmaniasis has emerged in the region with evidence of a continued transmission cycle [9]. The species responsible for this outbreak has been incriminated as *Leishmania* (*Viannia*) *braziliensis* [9], the most widely distributed *Leishmania* species in Brazil to date [10,11]. However, many questions still remain regarding the outbreak, such as: is one main strain or various strains responsible for the outbreak; is the emergence of *L.* (*V.*) *braziliensis* in the region a recent event; and how are Santa Catarina strains related to other strains in Brazil? A wide range of molecular tools are available for the investigation of molecular epidemiology of leishmaniasis, but choosing which method and/or markers to use continues to be a challenge [12]. Particularly for New World species, open access databases based on gold-standard genetic markers have not been developed. Currently, outbreak investigation of leishmaniasis, mainly conducted for visceral leishmaniasis outbreaks caused by *L.* (*Leishmania*) *donovani* species complex [13,14], commonly employs multilocus microsatellite typing (MLMT). This technique has been proven to discriminate at the intra-species level [15] with high discriminatory power and is useful for determining outbreak strain origin when a database of MLMT strains is available for the *Leishmania* species of interest [13]. At the present moment, an open access MLMT database for *L.* (*V.*) *braziliensis*, has not been developed. The high discriminatory power of this technique has its drawbacks depending on the type of epidemiological question or analysis. In some cases, almost 20 "different" genotypes can be identified in one focus [13,16,17]. Dividing the isolates into many different genotypes reduces the statistical power of analyses involving epidemiological variables, such as clinical and demographic characteristics of the patient. Such reductions in statistical power greatly reduce the ability of researcher to conclude the relationship of factors like clinical form and disease virulence with a particular genotype.

Thus, epidemiological tools with appropriate discriminatory power, increased reliability and inter-laboratory reproducibility

Author Summary

Molecular epidemiology of infectious diseases, which uses pathogen genetics to determine risk factors in the human population, is commonly employed to assist in outbreak investigation. While definitive genetic markers and techniques have been developed for several other bacterial, viral, and parasitic pathogens, the scientific community has yet to agree on an international standard for inter- and intra-species differentiation of *Leishmania*, the parasite that causes the disease leishmaniasis. As leishmaniasis represents one of the highest disease burdens among the neglected tropical diseases, development of molecular techniques, which allow for inter-laboratory comparability through international sequence databases, is imperative for moving forward with disease control. Based on the current standard technique employed for bacteria, the authors propose a panel of six genetic markers for multilocus sequence analysis (MLSA) for intraspecific differentiation of *Leishmania braziliensis*, the most widely distributed of the *Leishmania* species in Brazil. Using strains from a recent outbreak in the sub-tropical non-endemic southern Brazil in comparison with strains from eleven other Brazilian states, the authors provide a practical example of how this technique can be applied in a real world outbreak situation.

and comparability urgently are required. With these characteristics in mind, the method of multilocus sequence analysis (MLSA) provides a promising alternative. Projected as the future gold standard species typing method [12], MLSA involves sequencing a panel of house-keeping gene loci based on the panel of enzymes used in MLEE [18]. Several markers of these conserved regions have already been described, including ten markers for *L.* (*L.*) *donovani* [19,20], and six markers for New World species [18,21]. However, for *L.* (*Viannia*) species, these studies have mainly focused on interspecies discrimination and phylogenetic/taxonomic analysis and have employed only up to four markers. Given the challenges described above, we propose a panel of six gene loci, including three new markers described here for the first time, as an epidemiological tool for investigation of *L.* (*V.*) *braziliensis* outbreaks. In the present study, the recent outbreak in Santa Catarina is used to demonstrate the applicability of this technique in outbreak settings. The overarching objective of this work will be to generate interest in the community of leishmaniasis investigators to create an international sequence database based on these gene markers, as well as other markers from the original MLEE panel, for a more comprehensive and unified investigation into the distribution and epidemiological characteristics of *Leishmania* species.

Methods

Ethics statement

Ethical approval for the use of patient data and their respective sample was received from the UFSC Ethics Committee. CLIOC is a Depository Authority of the Ministry of the Environment [Fiel Depositária pelo Ministério do Meio Ambiente, MMA] (D.O.U. 05.04.2005). Following Resolution 21 (August 31, 2006 – CGEN/MMA), authorization was not required for usage of samples previously deposited in CLIOC since the samples were used for research purposes only and data were analyzed anonymously.

Data and sample collection

Leishmania (*Viannia*) *braziliensis* strains from eleven Brazilian states (n = 53) were obtained from the *Leishmania* Collection of the Oswaldo Cruz Institute (Coleção de *Leishmania* do Instituto Oswaldo Cruz- CLIOC) in Rio de Janeiro, Brazil, and strains from Santa Catarina (n = 33) were obtained from the cryobank of the Laboratório de Protozoologia of the Universidade Federal de Santa Catarina (UFSC), Florianópolis, Santa Catarina, Brazil. Patient data from Santa Catarina used in this study were investigated as part of routine reportable disease surveillance and collection procedures have been previously described in [9]. Santa Catarina isolates were deposited in CLIOC and subjected to MLEE characterization, according to routine procedures employed by CLIOC.

PCR amplification and sequencing

Leishmania promastigotes were cultured at 25°C in Schneider's medium supplemented with 20% heat-inactivated fetal bovine serum. DNA extraction was conducted using the Wizard DNA purification Kit (Promega, Madison, USA), according to manufacturer's instructions.

Amplification was performed for a panel of six housekeeping gene loci listed in Table 1. Primers and PCR conditions have been previously described for 6-*phosphogluconate dehydrogenase* (6*pgd*), *manose*-6-*phosphate isomerase* (*mpi*), *isocitrate dehydrogenase* (*icd*) [18] and for the *heat shock protein* 70 (*hsp*70) [22,23]. Primers for mitochondrial *malate dehydrogenase* (*mdh*mt) and nuclear *malate dehydrogenase* (*mdh*nc) are described here for the first time. Both follow the reaction condition: for 50 µl, 0,2 mM of each primer, 100 mM Tris–HCl, pH 8.8; 500 mM KCl, 1% Triton X-100; 15 mM $MgCl_2$, 0.25 mM deoxyribonucleotide triphosphate (dNTPs), 0.025 U FideliTaq/GoTaq polymerase and 50 ng DNA. Amplification conditions were 94°C for 2 min, followed by 34 cycles at 94°C for 30 s, 52°C for 30 s and 72°C for 1 min, with a final extension at 72°C for 5 min. PCR products were purified and subsequently sequenced with the same primers used in the PCR.

Table 1. Amplicon size, analyzed sequence fragment length and primer sequence of the target regions for the six loci studied.

Locus	Amplicon size (bp)	Analyzed sequence length (bp)	Primer sequence (5′-3′)
6-*phosphogluconate dehydrogenase* (6*pgd*)	836	666	CTCAAGGAACATGAGCACGA TTGTCCTTGACTTGCTCACG
Mannose-6-phosphate isomerase (*mpi*)	681	569	GGCAAGATGTATGCGGAGTT CTCCCCAGGAACCATCTGTA
Isocitrate dehydrogenase (*icd*)	1022	755	GAATCGGGAAGGAGATCACA CATCATAGCCCCAGAGAGGA
Heat-shock protein 70 (*hsp*70)	1022	896	GGACGAGATCGAGCGCATGGT TCCTTCGACGCCTCCTGGTTG
Malate dehydrogenase mitochondrial (*mdh*mt)	821	666	TGCCGACCTCTTCCATATTC GAGTGAGGTGCGTCTTCACA
Malate dehydrogenase nuclear (*mdh*nc)	1010	803	TCACAACCGCAACTACGA CTACTCACGATAACGGCAGA

doi:10.1371/journal.pntd.0002695.t001

Consensus sequences were obtained and edited in the software package Phred/Phrap/Consed Version: 0.020425.c (University of Washington, Seattle, WA, USA) and only those with Phred values above 20 were used as contigs. Analyzed sequence fragment lengths for each marker are provided in Table 1. Contigs of all strains were mounted and aligned in MEGA4 (Molecular Evolutionary Genetics Analysis version 4) [24]. Ambiguous sites were divided into two of the possible alleles for all markers using the PHASE algorithm in DnaSP5 [25].

Determination of clonal complexes

Clonal complexes (CC) were defined through BURST analysis in the software eBURSTv3 [26]. The BURST algorithm identified groups of mutually exclusive genotypes associated with a MLSA population and the founding genotype sequence within each group. Then, the algorithm provided the predicted descent from the founding genotype for all other genotypes [26,27]. For this analysis, criterion for CC formation was fixed at the most stringent level with at least five identical alleles for the six loci defining a CC. Sequences which were not able to be grouped into a clonal complex remained in the analysis as unique sequences.

Analysis of population structure

Haploid sequences rebuilt from the PHASE algorithm in DNAsp containing homozygous and heterozygous alleles were imported into STRUCTURE 2.3.4 (University of Chicago, Chicago, IL, USA) to investigate the population structure of the 86 samples of *L.* (*V.*) *braziliensis* based on the six MLSA loci. Using a Bayesian statistical approach, STRUCTURE applies a model-based clustering method to infer population structure and assign individuals to clusters based on multilocus genotype data [28]. Genetically distinct clusters (K) are identified based on the frequency of alleles, attributing the fraction of each genotype for each sample. In STRUCTURE, runs were performed using a burn-in period of 200,000 iterations followed by 600,000 running iterations. Runs were repeated three times to obtain data suitable for estimating the value of ΔK (defined as the rate of variation of the log likelihood of data between successive values of K), which provides the most likely K value for the data to be used in STRUCTURE HARVESTER [29]. STRUCTURE HARVESTER generates graphs for the change in the log of k and calculation of ΔK of STRUCTURE results, which were compared for choosing the K that best fit the data. Next, CLUMPP version 1.1.2 [30] was employed to align the multiple replicate analyzes of the same data set. Hierarchical analysis of two to seven K clusters was performed to define the assignment of borderline strains.

Based on clusters found in STRUCTURE, we used Microsatellite Analyser (MSA) [31] to estimate F_{ST} values and Genetic Data Analysis (GDA) version 1.1 [32] to calculate expected heterozygosity (H_e), observed heterozygosity (H_o), and inbreeding coefficient (F_{IS}). Recombination analysis was performed in Recombination Detection Program (RDP) [33].

Development of median-joining network

To view genetic relationships (phylogenetic network) among strains and differentiation provided by the six markers, the median-joining network was mounted in the program SplitsTree 4.0 [34]. The median-joining network was constructed using concatenated character nucleotide sequences with ambiguous sites for all loci and strains. Nodes of the network, representing individual or groups of strains, were labeled by size, color and/or year/location to reflect epidemiological variables associated with the patient from whom the strain was isolated.

Statistical analysis and mapping

Associations between genetic and epidemiological variables were analyzed in Stata SE 13 (StataCorp LP, College Station, TX, USA). Chi-squared test, or Fisher's exact test when appropriate, was used to assess the relationships between categorical variables. Maps were created in ArcGIS 10 (ESRI, Redlands, CA, USA).

Results

BURST analysis identified three clonal complexes (CC) among the 86 strains of *L.* (*V.*) *braziliensis*, with over half (54.7%, 47/86) of the strains not belonging to any of the three CCs and remaining separate as unique sequence types (Supporting Information S1). A total of 76 distinct sequence types were observed among strains. The analysis was heavily weighted by the homogeneity and large number of strains from Santa Catarina included in the analysis, with the large majority (84.8%, 28/33) of Santa Catarina strains being grouped into one nearly exclusive clonal complex (CC1). Five out of six strains from Santa Catarina that did not group with CC1 were registered as imported cases in the epidemiological investigation. No association was found between CC and clinical form ($p = 0.660$). Figure 1 shows the geographical distribution of the CCs by state, revealing proportionally higher genetic variation in states from the Amazon biome (94.1% (n = 16/17) unique sequence types) (Supporting Information S1).

Through calculation of ΔK in the STRUCTURE analysis, the *L.* (*V.*) *braziliensis* strains included in the present study from 12 Brazilian states were found to best fit into three clusters (POP) (Supporting Information S1). Overall, 41.9% (36/86) of strains belonged to POP1, 40.7% (35/86) to POP2, and 16.3% (14/86) to POP3. As in the BURST analysis, the large majority (87.9%, 29/33) of Santa Catarina strains formed their own cluster (POP2), which also included four strains from Pernambuco, one from Mato Grosso and one from Bahia (Figure 2). The four Santa Catarina strains that did not cluster with POP2 were registered as imported cases in the epidemiological analysis. These four strains were the same strains from imported cases that did not cluster in the BURST analysis. Complete strain information can be found in Supporting Information S2.

As shown in Figure 3, POP1 demonstrated the most extensive geographical distribution, including strains from all states analyzed in this study. A distinction can be made between the genetic variation and genetic structure of coastal states, which contain Atlantic forest, and northern states, which are located in the Amazon basin. States of the Amazon region were predominately comprised of POP1 strains, while strains of POP2 and POP3 were mainly found in coastal states.

A significant association between the genetic cluster designated by STRUCTURE and leishmaniasis clinical form of the patient from which the strain was isolated was observed ($p = 0.030$) (Table 2). Most strains from cases presenting the mucocutaneous clinical form (4/7) belonged to POP3, including one case from Rio de Janeiro State, one from Pernambuco and two from Bahia.

Based on the scale for the interpretation of F_{ST} suggested by Wright (1978), the estimates showed significant genetic differentiation among the STRUCTURE clusters (Table 3). POP1 and POP3 showed moderate genetic differentiation ($F_{ST} = 0.1087$), while POP2 showed great genetic differentiation with POP1 and POP3 ($F_{ST} = 0.1540$ and 0.2028, respectively). POP1 had the highest average number of alleles per locus (23.3), while both clusters POP2 and POP3 were similar in mean number of alleles, being approximately five alleles per locus. Positive values of F_{IS} were found for all clusters. F_{IS} values for POP1 and POP3 were particularly high (Table 4). All loci were polymorphic for POP1

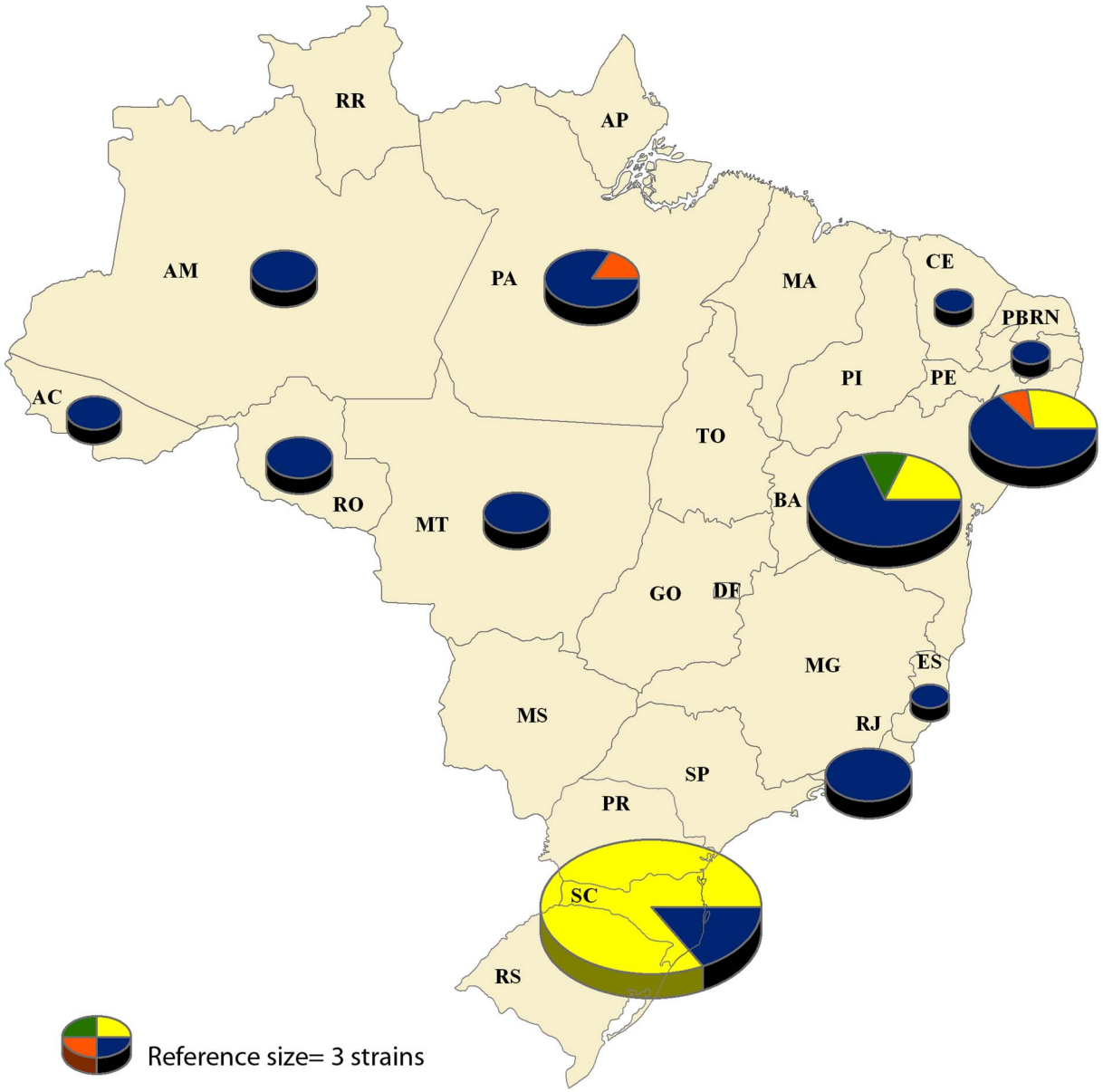

Figure 1. Geographical distribution of clonal complexes by state among Brazilian strains of *L.* (*V.*) *braziliensis* (n = 86). Colors represent clonal complex (CC) as follows: yellow, CC1; green, CC2; orange, CC3; blue, unique sequence types. The size of the individual pie charts was weighted according to number of strains from the state included in the analysis. A chart representing three samples is provided as a reference. State names are the same as described in Supporting Information S1.
doi:10.1371/journal.pntd.0002695.g001

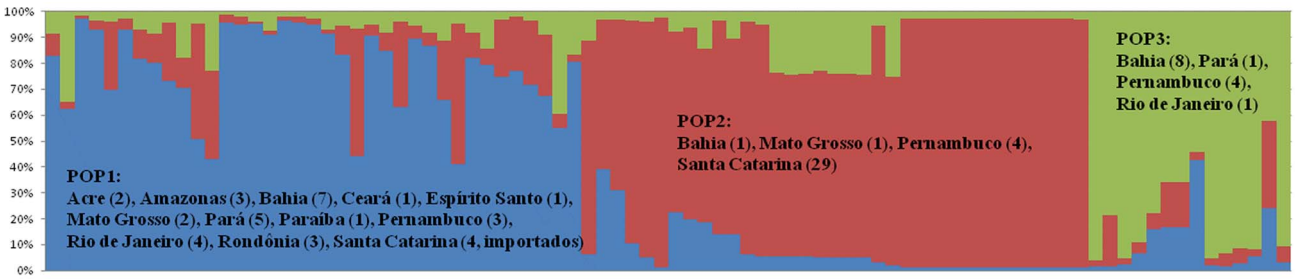

Figure 2. STRUCTURE output estimating genetic structure of the *L.* (*V.*) *braziliensis* strains from twelve Brazilian states (n = 86, *K* = 3). POP1 – Cluster 1; POP2 – Cluster 2; POP3 – Cluster 3; State (number of strains).
doi:10.1371/journal.pntd.0002695.g002

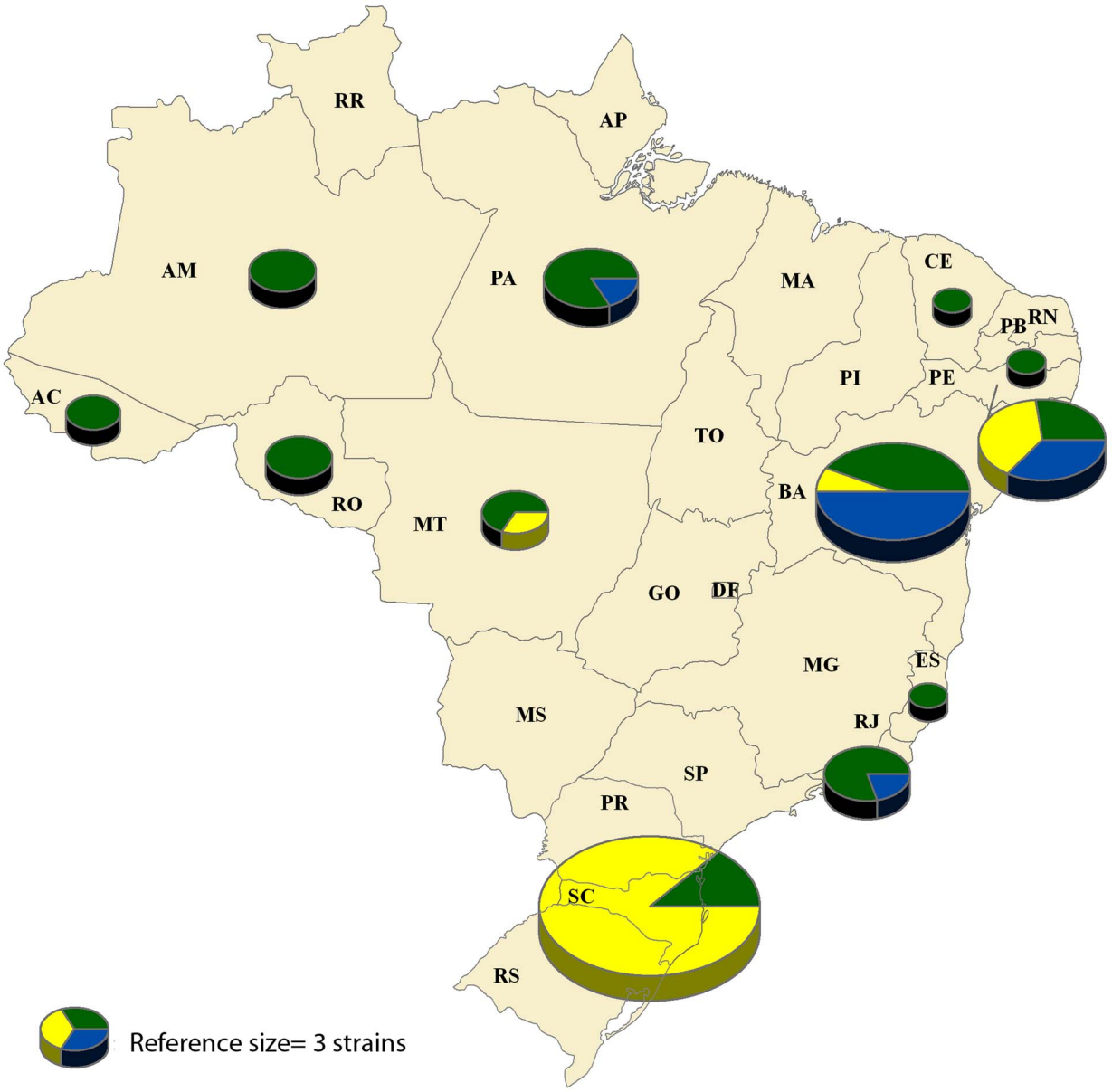

Figure 3. Geographical distribution of genetic clusters of Brazilian *L.* (*V.*) *braziliensis* strains as designated by STRUCTURE by state (*K* = 3, n = 86). Colors represent genetic clusters (POP) as follows: green, POP1; yellow, POP2; blue, POP3. The size of the individual pie charts was weighted according to number of strains from the state included in the analysis. A chart representing three samples is provided as a reference. State names are the same as described in Supporting Information S1.
doi:10.1371/journal.pntd.0002695.g003

Table 2. Leishmaniasis clinical form of patient by genetic cluster as attributed by STRUCTURE analysis of 86 *L.* (*V.*) *braziliensis* strains from Brazil.

	Cluster			
Clinical form	**POP1**	**POP2**	**POP3**	**Total**
Cutaneous	29	32	10	71
Disseminated	2	0	0	2
Mucocutaneous	1	2	4	7
Total	32	34	14	80

doi:10.1371/journal.pntd.0002695.t002

and POP2 and five (83.3%) of the six loci were polymorphic for POP3. The marker 6*pgd* was not polymorphic for POP3. In general, the new markers *hsp*70, *mdhnc* and *mdh*mt showed the highest number of alleles of 35, 40 and 44, respectively, in comparison to 15–30 alleles for the other three markers.

Results of the BURST and STRUCTURE analysis were found to be significantly associated ($p<0.001$) (Table 5). The majority (37/48) of unique sequences in the BURST analysis were forced into their own population (POP1) in the STRUCTURE analysis, representing mainly strains from the Amazon regions.

Recombination events were detected by seven algorithms in RDP software ($p<0.05$). However, neither the beginning nor ending breakpoints could be identified, which may have resulted in recombinant misidentification. Nonetheless, one sample from

Table 3. Matrix of F_{ST} values and corresponding p values for the three clusters identified in STRUCTURE of 86 Brazilian *L.* (*V.*) *braziliensis* strains.

p\F$_{ST}$	POP1	POP2	POP3
POP1	0	0.1540	0.1087
POP2	0.0003	0	0.2028
POP3	0.0003	0.0003	0

F_{ST} values can be found in the upper right triangle and p values with Bonferroni correction are reported in the lower left triangle of the matrix.
doi:10.1371/journal.pntd.0002695.t003

Santa Catarina (185) and one sample from Bahia (IOC/L 2871) were indicated as potentially parental or recombinant. Thirty-one samples from Santa Catarina had sequences with partial evidence of the same recombination event.

The median-joining network was created from concatenated sequences of the six gene loci for the 86 strains of *L.* (*V.*) *braziliensis* from Brazil. Majority of Santa Catarina strains presented as an evident cluster. Other strains close to the Santa Catarina cluster were from Pernambuco (n = 2), Rio de Janeiro (n = 1), and Pará (n = 1). When the nodes of Santa Catarina strains were highlighted by case origin, all cases not clustered with the principal cluster were imported cases, with the exception of strain 605 (Figure 4). This 605 strain also was grouped within the main Santa Catarina CC and POP in both the BURST and STRUCTURE analyses.

When the strains from cases of mucocutaneous and disseminated clinical form were highlighted, those from Bahia were clustered, while mucocutaneous cases from other Brazilian states appeared closer to the main cluster of Santa Catarina (Figure 5). When the median-joining network was reduced to only strains from Santa Catarina, the resulting network presented three principal branches. Marked by year and city of leishmaniasis case diagnosis, a main cluster can be observed in the center of the network, representing the epicenter of the outbreak which occurred in 2006 in the municipality of Blumenau (Figure 6). From this main epicenter, autochthonous cases branched separately, appearing to evolve over time and space to the neighboring municipality of the capital municipality of Florianópolis. The map in Figure 6 shows this main cluster of related Santa Catarina strains was distributed over a distance of 140 km in four years from Blumenau to Florianópolis.

Discussion

Multilocus sequencing analysis (MLSA) was successful in detecting epidemiological patterns among *L.* (*V.*) *braziliensis* strains from twelve Brazilian states. Additionally, the technique was able to detect intra-species variation compatible with epidemiological characteristics within a specific outbreak focus, demonstrating the potential of this technique as a molecular tool for outbreak investigation.

Table 4. Characterization of the three clusters found in the STRUCTURE analysis of *L.* (*V.*) *braziliensis* strains from Brazil.

Population	N	P	A	H_e	H_o	F_{IS}
POP1	37	1.000	23.333	0.825	0.239	0.714
POP2	35	1.000	5.833	0.345	0.276	0.202
POP3	14	0.833	5.167	0.506	0.119	0.772

n – Sample size; P – Proportion of polymorphic loci; A – Mean number of alleles per locus; H_e - Expected heterozygosity; H_o – Observed heterozygosity; F_{IS} – Inbreeding coefficient.
doi:10.1371/journal.pntd.0002695.t004

Table 5. Comparison of the results of the BURST and STRUCTURE analyses of the 86 *L.* (*V.*) *braziliensis* strains from Brazil.

Clonal Complex	Cluster			
	POP1	POP2	POP3	Total
CC1	0	31	3	34
CC2	0	0	2	2
CC3	0	0	2	2
Unique sequence	37	4	7	48
Total	37	35	14	86

doi:10.1371/journal.pntd.0002695.t005

In the BURST analysis, strains were found to group into three clonal complexes. Samples from the Amazon region presented largely as unique sequence types, demonstrating a proportionally higher level of heterogeneity in comparison to coastal states. This distinction is particularly apparent when compared to Santa Catarina. Since the BURST analysis permitted samples not to be grouped into a specific CC and remain as unique sequences, the STRUCTURE analysis was observed to force these unique sequences to form a genetic cluster. This was also evident in the significant association between the two analyses ($p<0.001$). POP1 was comprised of almost entirely unique sequence types. This high heterogeneity is characteristic of strains from the Amazon biome, as previously observed in other studies, and reflects the large variety of vectors and hosts in the region [10,35,36]. Furthermore, the emergence of *L.* (*V.*) *braziliensis* in the state of Santa Catarina, Brazil appears to be a recent event, given the high homogeneity observed among the analyzed strains. This conclusion is based on the assumptions of the Hardy-Weinberg equilibrium model of populations, which states if no evolutionary pressure mechanism, such as migration in or out of the population or mutation over a long period of time, is acting upon a given population, then the genetic frequencies will remain unaltered [37,38]. Therefore, during the period from which the samples were collected during the outbreak, the strains were largely uninfluenced by outside strains, remaining as their own apparently unique population. This could also be caused by a specific transmission cycle in which other *Leishmania* strains or species were not easily incorporated. Specific vector-parasite relationships would remove the possibility of recombination given the major selective force on *Leishmania* populations occurs in the vector hosts during the development of the parasite [39].

MLSA utilizing the panel of six markers was able to distinguish epidemiological characteristics among *L.* (*V.*) *braziliensis* strains. In all three analyses (BURST, STRUCTURE, and median-joining), MLSA results were compatible with case origin evaluated in the epidemiological investigation of Santa Catarina strains. These results demonstrate the potential of this method for use in future outbreak investigations and surveillance. Despite being registered as an imported case in the epidemiological investigation, strain 605 was grouped within the main Santa Catarina cluster in all three analyses, pointedly suggesting this patient was most likely an autochthonous case. In such cases, the molecular characteriza-

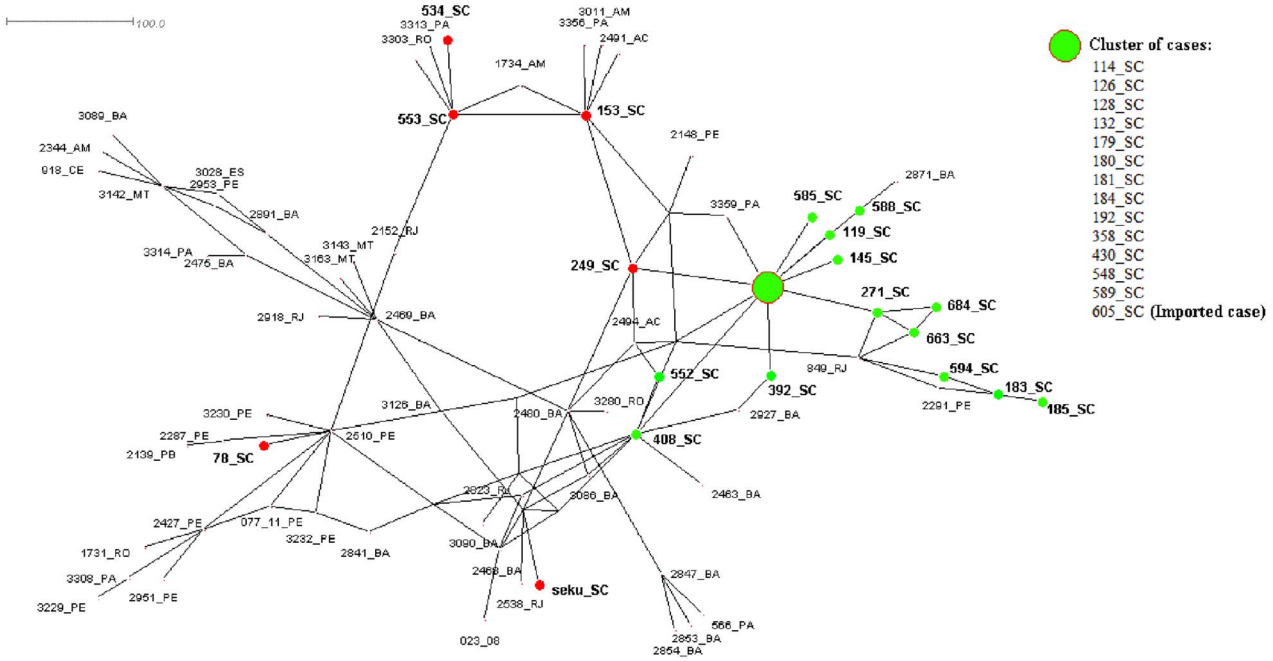

Figure 4. Median-joining network based on concatenated sequences of six gene fragments for 86 strains of *L.* (*V.*) *braziliensis* from Brazil, with case origin designated for strains from Santa Catarina. Green nodes represent autochthonous cases and red nodes represent imported cases.
doi:10.1371/journal.pntd.0002695.g004

tion proved to be a more reliable and precise tool than the epidemiological interview to determine if a case acquired the infection locally or outside of a given region. Results also show the methodology possesses discriminatory power to differentiate imported and autochthonous cases at state macroregion levels. Knowledge on the origin of a case is important for predicting case

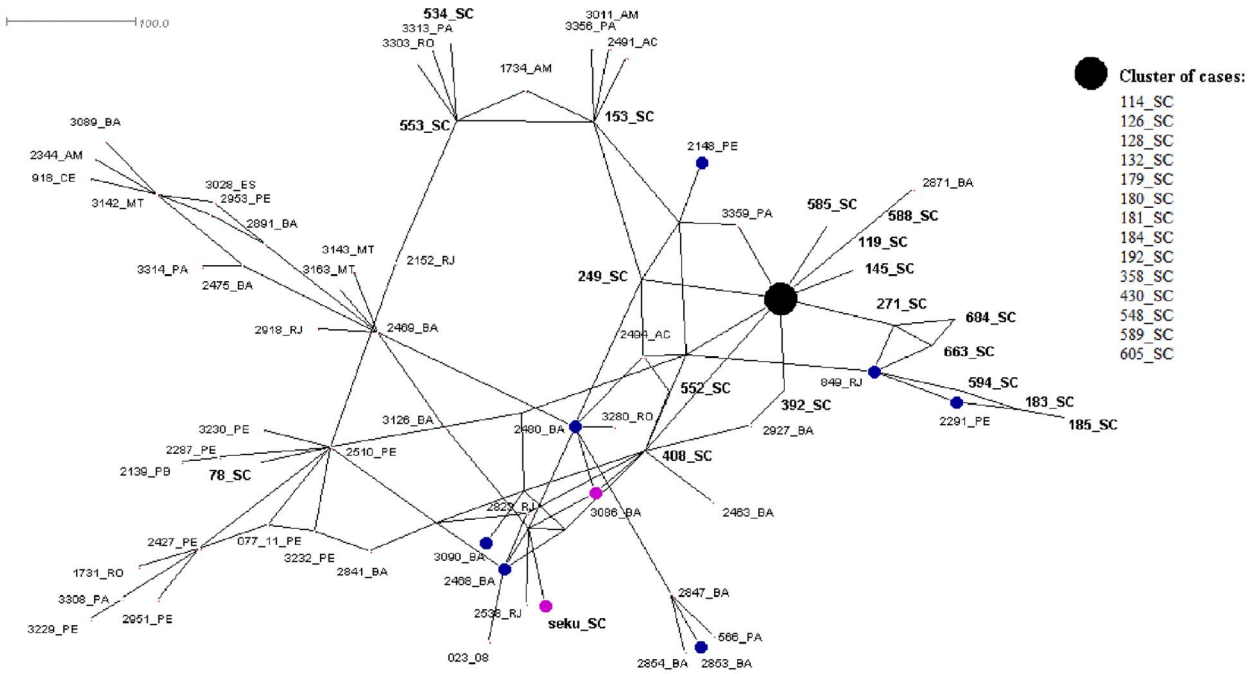

Figure 5. Median-joining network based on concatenated sequences of six gene fragments for 86 strains of *L.* (*V.*) *braziliensis* from Brazil, with clinical form of the case designated. Blue nodes represent mucocutaneous cases and purple nodes represent disseminated cases. Non-highlighted or black nodes signify cutaneous cases.
doi:10.1371/journal.pntd.0002695.g005

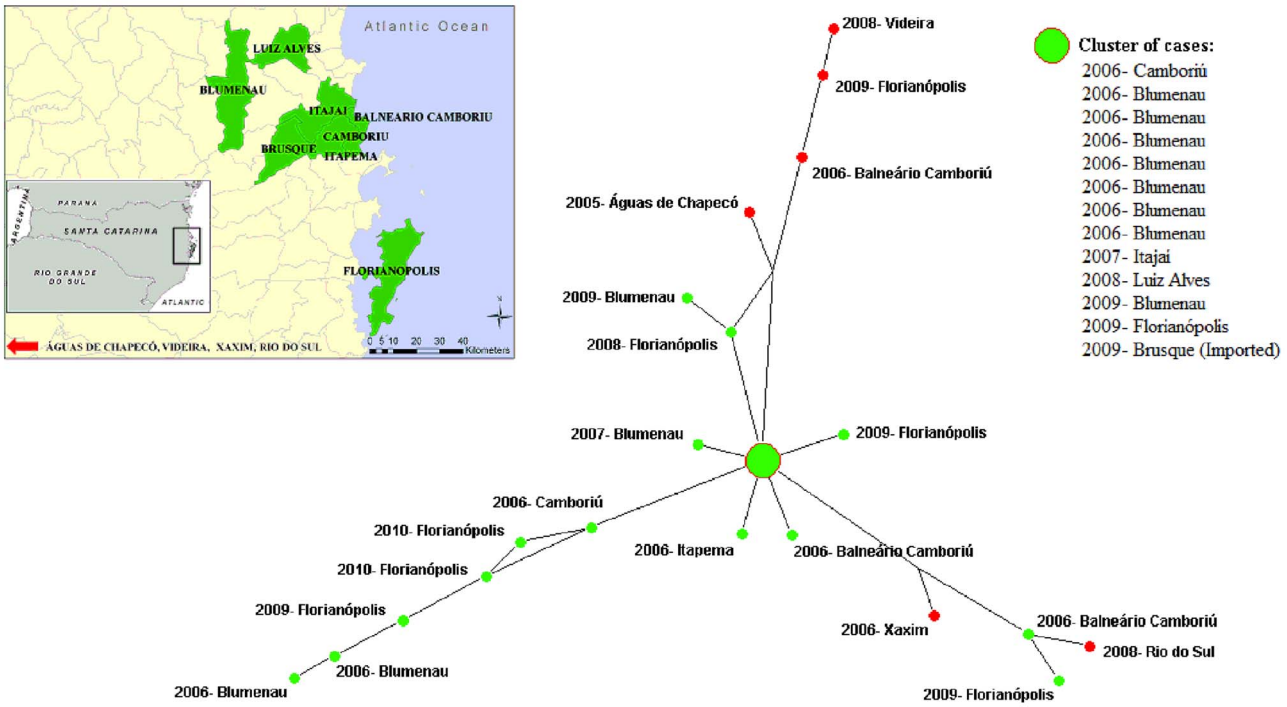

Figure 6. Reduced median-joining network based on concatenated sequences of six gene fragments for 33 strains of *L.* (*V.*) *braziliensis* from Santa Catarina, Brazil, with year, city of residence, and origin of case designated. Green nodes represent autochthonous cases and red nodes represent imported cases.
doi:10.1371/journal.pntd.0002695.g006

outcome and treatment course, since several studies have shown a relationship between specific characteristics of the infecting parasite and geographical location with the outcome of the patient [9,40,41].

Along the same lines, MLSA showed a significant association between clusters in the STRUCTURE analysis and patient clinical form among the samples analyzed in the present study ($p = 0.0296$). However, the current study only included seven cases of the mucocutaneous form. A study involving a large representative sample of these cases with controls is necessary to validate these findings. Identification of a genetic marker of *Leishmania* virulence has not been identified at the present moment [42], and the identification of such a marker would have important clinical and pharmacological significance. Despite the limited number of samples in this study, this methodology could be promising for the identification of a specific *L.* (*V.*) *braziliensis* cluster predisposed to the mucocutaneous form, and therefore, warrants further investigation.

Recombination is often difficult to detect within species because of low inter-strain diversity and/or apparent low diversity due to inappropriate sampling [43]. However, RDP results of the present study were able to reveal recombination occurring between the *L. braziliensis* strains. This suggests the strains from Santa Catarina may be the result of a clonal expansion from a recombinant event, and the resulting strains then encountered proper conditions to propagate in the state. Previous studies on recombination, including a study on population genetics for inbreeding [44] and a previous MLSA phylogenetic study [18] specifically for *L. braziliensis*, were able to detect recombination signals as well. In these situations, homologous recombination may have been the responsible mechanism. This phenomenon also may have produced the well-structured clonal complexes in *Leishmania* in the present study which allowed for the epidemiological inferences to be made.

As no definitive set of markers for MLSA has been defined for the study of populations within a given species of *Leishmania*, the markers evaluated here could be defined as potential candidates in the panel used for this type of study. Interestingly, the three new markers, *hsp*70, *mdh*nc and *mdh*mt, were the most polymorphic of the six markers, suggesting their addition provided the increase in discriminatory power that allowed for intra-species differentiation. Taken together, these six markers provided adequate discriminatory power to answer epidemiological questions surrounding genetic clusters of a single species. An important benefit of MLSA is the ability to create and store sequences in an international database for global comparison of *Leishmania* species and strains [15]. The next step will be to determine the viability and discriminatory power of this six loci panel for other species of *Leishmania* and increase the number of markers and strains sequenced. Four of the markers (6*pgd*, *mpi*, *icd* and *hsp*70) have already proven to be discriminatory among species of the *Leishmania* subgenus *Viannia*, including *L.* (*V.*) *shawi*, *L.* (*V.*) *lainsoni*, *L.* (*V.*) *naiffi* and *L.* (*V.*) *guyanensis* [18,21].

With the recent increase in development of genetic markers and new statistical methods for analyzing them, the choice of which software is most adequate to your specific analysis is becoming increasingly difficult. No definitive guidelines currently exist [45]. For this reason, we opted to evaluate our MLSA results from three different perspectives, using diverse software (BURST, STRUCTURE and Splitstree) to arrive at our inferences regarding the genetic structure among the *L.* (*V.*) *braziliensis* included in the present study. The BURST analysis, which is commonly used for MLSA of haploid organisms, such as bacteria, permitted a better comprehension of the genetic variability among the samples using conservative parameters for differentiating clonal complexes. As

almost all Santa Catarina strains fit into one clonal complex and the remaining strains were mainly unique sequences, we can conclude the cluster in this state is highly homogeneous in comparison to other states. However, with over half of the strains not grouped in a clonal complex, comparison of genotypes with epidemiological factors was not possible. The STRUCTURE analysis forced all strains into a cluster, resulting in the grouping of all of these unique sequences into their own cluster. This phenomenon shows that, despite the high diversity among the samples from the Amazon region, strains from Santa Catarina continue to be genetically distinct from other Brazilian strains analyzed here. In other words, the diverse genetic clusters within POP1 of the Amazon region, as a whole, are still genetically more distinct from Santa Catarina strains than within themselves, as also shown by F_{ST} and F_{IS}. Interestingly, our study found high positive F_{IS} values (high inbreeding coefficients) among the populations of *L.* (*V.*) *braziliensis*, which negates the hypothesis of strictly clonal reproduction among *Leishmania* species. High F_{IS} values have also been observed in various MLMT studies for *Leishmania*, including a study on *L.* (*V.*) *braziliensis* in Bolivia and Peru [46], a study on *L.* (*L.*) *infantum* in Old World and New World strains [47] and a study on *L.* (*L.*) *donovani* in Ethiopia [48]. In these studies, possible explanations of these high F_{IS} values were the presence of considerable inbreeding and/or sub-structuring of the population, reflecting a possible Wahlund effect.

Despite being too complex for comparing all strains among themselves, the median-joining network was the best visual representation for comparing Santa Catarina strains with all other strains from Brazil. This type of analysis is most applicable in an outbreak situation in which strains from a specific area can be compared to other reference strains, allowing for the distinguishing of imported cases and other epidemiological differences. Overall, until software capable of addressing specific genetic *Leishmania* characteristics, such as infrequent recombination, is created, use of all three types of genetic analyses can be used as an alternative to provide a robust MLSA analysis.

This information on the genetic variability of circulating strains is important for public health and control efforts. Considering drug resistance and complications in treatment have not been observed in Santa Catarina cases, control of leishmaniasis in Santa Catarina where the parasite strains are genetically homogeneous would be expected to be much more efficient than in regions where the parasite presents genetic heterogeneity and a more complex transmission cycle. This factor emphasizes the need for more urgent and active control methods to prevent further introduction of *Leishmania* strains and/or species, as well as geographical spread of the disease.

Conclusion

MLSA revealed epidemiological patterns among *L.* (*V.*) *braziliensis* strains from twelve Brazilian states, even within the state of Santa Catarina where the strains presented extensive homogeneity. The addition of three markers, *hsp*70, *mdh*nc and *mdh*mt to the previously described panel of markers increased the discriminatory power of the technique, permitting the identification of three genetic clusters within *L.* (*V.*) *braziliensis* strains. All three analyses (BURST, STRUCTURE and median-joining network) provided a complementary and integral part in the interpretation of the MLSA results. When used in tandem with MLMT, these two methods could provide a more robust approach to the molecular epidemiology of leishmaniasis and increased validity of the population structure model. A prospective study design that seeks to include a representative sample of the patient population and active collection of their *Leishmania* strains is needed to validate this method as a molecular epidemiology tool. However, the present study has provided sufficient evidence of the effectiveness of this method for pursuing further validation of MLSA for leishmaniasis outbreak investigation.

Supporting Information

Supporting Information S1 Summary of clonal complex and genetic cluster of strains by state.
(XLS)

Supporting Information S2 Strains by sequence type, with data on sex, clinical outcome, origin, and state of origin provided for Santa Catarina strains only. Doubled columns present alternative alleles for those sequences presenting ambiguous sites.
(XLS)

Acknowledgments

The authors would like to thank the Plataforma de Sequenciamento do Instituto Oswaldo Cruz (PDTIS) for performing sequencing and CLIOC for providing strains and MLEE typing.

Author Contributions

Conceived and designed the experiments: MAM MCB MS EC. Performed the experiments: MAM MCB. Analyzed the data: MAM MCB GEMF. Contributed reagents/materials/analysis tools: MS EC. Wrote the paper: MAM MCB GEMF MS EC.

References

1. Reithinger R, Dujardin JC, Louzir H, Pirmez C, Alexander B, et al. (2007) Cutaneous leishmaniasis. Lancet Infect Dis 7: 581–596.
2. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, et al. (2012) Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380: 2163–2196.
3. Aguado M, Espinosa P, Romero-Mate A, Tardio JC, Cordoba S, et al. (2013) Outbreak of cutaneous leishmaniasis in Fuenlabrada, Madrid. Actas Dermosifiliogr 104: 334–342.
4. Desjeux P (2001) The increase in risk factors for leishmaniasis worldwide. Trans R Soc Trop Med Hyg 95: 239–243.
5. Varani S, Cagarelli R, Melchionda F, Attard L, Salvadori C, et al. (2013) Ongoing outbreak of visceral leishmaniasis in Bologna Province, Italy, November 2012 to May 2013. Euro Surveill 18: 20530.
6. Villinski JT, Klena JD, Abbassy M, Hoel DF, Puplampu N, et al. (2008) Evidence for a new species of Leishmania associated with a focal disease outbreak in Ghana. Diagn Microbiol Infect Dis 60: 323–327.
7. Salomon OD, Sosa-Estani S, Ramos K, Orellano PW, Sanguesa G, et al. (2006) Tegumentary leishmaniasis outbreak in Bella Vista City, Corrientes, Argentina during 2003. Mem Inst Oswaldo Cruz 101: 767–774.
8. Faucher B, Gaudart J, Faraut F, Pomares C, Mary C, et al. (2012) Heterogeneity of environments associated with transmission of visceral leishmaniasis in South-Eastern France and implication for control strategies. PLoS Negl Trop Dis 6: e1765.
9. Marlow MA, da Silva Mattos M, Makowiecky ME, Eger I, Rossetto AL, et al. (2013) Divergent profile of emerging cutaneous leishmaniasis in subtropical Brazil: new endemic areas in the southern frontier. PLoS One 8: e56177.
10. Grimaldi G, Jr., Tesh RB, McMahon-Pratt D (1989) A review of the geographic distribution and epidemiology of leishmaniasis in the New World. Am J Trop Med Hyg 41: 687–725.
11. Banuls AL, Hide M, Prugnolle F (2007) Leishmania and the leishmaniases: a parasite genetic update and advances in taxonomy, epidemiology and pathogenicity in humans. Adv Parasitol 64: 1–109.
12. Schonian G, Kuhls K, Mauricio IL (2011) Molecular approaches for a better understanding of the epidemiology and population genetics of Leishmania. Parasitology 138: 405–425.
13. Gelanew T, Cruz I, Kuhls K, Alvar J, Canavate C, et al. (2011) Multilocus microsatellite typing revealed high genetic variability of Leishmania donovani strains isolated during and after a Kala-azar epidemic in Libo Kemkem district, northwest Ethiopia. Microbes Infect 13: 595–601.

14. Motoie G, Ferreira GE, Cupolillo E, Canavez F, Pereira-Chioccola VL (2013) Spatial distribution and population genetics of Leishmania infantum genotypes in Sao Paulo State, Brazil, employing multilocus microsatellite typing directly in dog infected tissues. Infect Genet Evol 18: 48–59.
15. Schonian G, Mauricio I, Gramiccia M, Canavate C, Boelaert M, et al. (2008) Leishmaniases in the Mediterranean in the era of molecular epidemiology. Trends Parasitol 24: 135–142.
16. Chicharro C, Llanes-Acevedo I, Garcia E, Nieto J, Moreno J, et al. (2013) Molecular typing of Leishmania infantum isolates from a leishmaniasis outbreak in Madrid, Spain, 2009 to 2012. Euro Surveill 18: 20545.
17. Seridi N, Amro A, Kuhls K, Belkaid M, Zidane C, et al. (2008) Genetic polymorphism of Algerian Leishmania infantum strains revealed by multilocus microsatellite analysis. Microbes Infect 10: 1309–1315.
18. Boite MC, Mauricio IL, Miles MA, Cupolillo E (2012) New insights on taxonomy, phylogeny and population genetics of Leishmania (Viannia) parasites based on multilocus sequence analysis. PLoS Negl Trop Dis 6: e1888.
19. Mauricio IL, Yeo M, Baghaei M, Doto D, Pratlong F, et al. (2006) Towards multilocus sequence typing of the Leishmania donovani complex: resolving genotypes and haplotypes for five polymorphic metabolic enzymes (ASAT, GPI, NH1, NH2, PGD). Int J Parasitol 36: 757–769.
20. Zemanova E, Jirku M, Mauricio IL, Horak A, Miles MA, et al. (2007) The Leishmania donovani complex: genotypes of five metabolic enzymes (ICD, ME, MPI, G6PDH, and FH), new targets for multilocus sequence typing. Int J Parasitol 37: 149–160.
21. Tsukayama P, Lucas C, Bacon DJ (2009) Typing of four genetic loci discriminates among closely related species of New World Leishmania. Int J Parasitol 39: 355–362.
22. da Silva LA, de Sousa Cdos S, da Graca GC, Porrozzi R, Cupolillo E (2010) Sequence analysis and PCR-RFLP profiling of the hsp70 gene as a valuable tool for identifying Leishmania species associated with human leishmaniasis in Brazil. Infect Genet Evol 10: 77–83.
23. Garcia L, Kindt A, Bermudez H, Llanos-Cuentas A, De Doncker S, et al. (2004) Culture-independent species typing of neotropical Leishmania for clinical validation of a PCR-based assay targeting heat shock protein 70 genes. J Clin Microbiol 42: 2294–2297.
24. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24: 1596–1599.
25. Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451–1452.
26. Feil EJ, Li BC, Aanensen DM, Hanage WP, Spratt BG (2004) eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. J Bacteriol 186: 1518–1530.
27. Spratt BG, Hanage WP, Li B, Aanensen DM, Feil EJ (2004) Displaying the relatedness among isolates of bacterial species – the eBURST approach. FEMS Microbiol Lett 241: 129–134.
28. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155: 945–959.
29. Earl D, vonHoldt B (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4: 359–361.
30. Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23: 1801–1806.
31. Dieringer D, Schlötterer C (2003) microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. Molecular Ecology Notes 3: 167–169.
32. Lewis POZ, D. (2002) Genetic Data Analysis: Computer program for the analysis of allelic data, Version 1.1.
33. Martin DP, Lemey P, Lott M, Moulton V, Posada D, et al. (2010) RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26: 2462–2463.
34. Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 23: 254–267.
35. Cupolillo E, Brahim LR, Toaldo CB, de Oliveira-Neto MP, de Brito ME, et al. (2003) Genetic polymorphism and molecular epidemiology of Leishmania (Viannia) braziliensis from different hosts and geographic areas in Brazil. J Clin Microbiol 41: 3126–3132.
36. Gomes RF, Macedo AM, Pena SD, Melo MN (1995) Leishmania (Viannia) braziliensis: genetic relationships between strains isolated from different areas of Brazil as revealed by DNA fingerprinting and RAPD. Exp Parasitol 80: 681–687.
37. Jameson DL (1977) Benchmark Papers in Genetics 8. Evolutionary Genetics. Stroudsburg, PA: Dowden, Hutchinson and Ross.
38. Weinberg W (1908) Über den Nachweis der Vererbung beim Menschen. Jahresh. Ver Vaterl Naturkd Württemb 64: 369–382.
39. Lainson RS, J.J. (1987) Evolution, classification and geographical distribution. The Leishmaniases in Biology and Medicine; Peters WK-K, R., editor. London: Academic Press.
40. Arevalo J, Ramirez L, Adaui V, Zimic M, Tulliano G, et al. (2007) Influence of Leishmania (Viannia) species on the response to antimonial treatment in patients with American tegumentary leishmaniasis. J Infect Dis 195: 1846–1851.
41. Souza AS, Giudice A, Pereira JM, Guimaraes LH, de Jesus AR, et al. (2010) Resistance of Leishmania (Viannia) braziliensis to nitric oxide: correlation with antimony therapy and TNF-alpha production. BMC Infect Dis 10: 209.
42. Hartley MA, Ronet C, Zangger H, Beverley SM, Fasel N (2012) Leishmania RNA virus: when the host pays the toll. Front Cell Infect Microbiol 2: 99.
43. Prugnolle F, De Meeus T (2010) Apparent high recombination rates in clonal parasitic organisms due to inappropriate sampling design. Heredity (Edinb) 104: 135–140.
44. Mannaert A, Downing T, Imamura H, Dujardin JC (2012) Adaptive mechanisms in pathogens: universal aneuploidy in *Leishmania*. Trends Parasitol 28: 370–376.
45. Halkett F, Simon JC, Balloux F (2005) Tackling the population genetics of clonal and partially clonal organisms. Trends Ecol Evol 20: 194–201.
46. Rougeron V, De Meeus T, Hide M, Waleckx E, Bermudez H, et al. (2009) Extreme inbreeding in Leishmania braziliensis. Proc Natl Acad Sci U S A 106: 10224–10229.
47. Kuhls K, Alam MZ, Cupolillo E, Ferreira GE, Mauricio IL, et al. (2011) Comparative microsatellite typing of new world leishmania infantum reveals low heterogeneity among populations and its recent old world origin. PLoS Negl Trop Dis 5: e1155.
48. Gelanew T, Kuhls K, Hurissa Z, Weldegebreal T, Hailu W, et al. (2010) Inference of population structure of Leishmania donovani strains isolated from different Ethiopian visceral leishmaniasis endemic areas. PLoS Negl Trop Dis 4: e889.

4.3. Capítulo III

Population structure and evidence for both clonality and recombination among Brazilian strains of the subgenus *Leishmania* (*Viannia*).

Kuhls K, Cupolillo E, Silva SO, Schweynoch C, **Boité MC**, Mello MN, Mauricio I, Miles M, Wirth T, Schönian G.
PLoS Neglected Tropical Diseases
2013 Oct 31;7(10):e2490. doi: 10.1371/journal.pntd.0002490.
Artigo publicado

Este trabalho aplicou marcadores microssatélites desenvolvidos previamente para o subgênero *L.* (*Viannia*), corroborando estudos prévios onde significativa diversidade para este subgênero já havia sido descrita. Duas populações principais foram definidas por análises utilizando MLMT: uma contendo cepas de *L.* (*V.*) *guyanensis* da região Amazônia e outra apresentando as cepas de *L.* (*V.*) *braziliensis*, oriundas da costa Atlântica brasileira. Um terceiro grupo, muito heterogêneo, pôde ser definido com cepas de *L.* (*V.*) *braziliensis* da região norte e com as cepas das espécies *L.* (*V.*) *shawi*, *L.* (*V.*) *naiffi*, e *L.* (*V.*) *lainsoni*, todas também isoladas na região norte do país. Sinais de recombinação foram detectados no grupo formado por cepas identificadas como *L.* (*V.*) *guyanensis* e também naquele composto por cepas de *L.* (*V.*) *braziliensis* da região costeira. A recombinação pode ser uma das justificativas tanto para a grande diversidade encontrada como para a ausência de estruturação clara da população. Entretanto, a observação de clones aparentemente estáveis entre as cepas estudadas indicam que os eventos de recombinação não são frequentes o suficiente para romper uma estrutura basicamente clonal da população de *Leishmania* (*Viannia*). Com esses resultados pudemos concluir o segundo objetivo desta tese.

OPEN ACCESS Freely available online

PLOS | NEGLECTED TROPICAL DISEASES

Population Structure and Evidence for Both Clonality and Recombination among Brazilian Strains of the Subgenus *Leishmania* (*Viannia*)

Katrin Kuhls[1¤], Elisa Cupolillo[2], Soraia O. Silva[3], Carola Schweynoch[1], Mariana Côrtes Boité[2], Maria N. Mello[3], Isabel Mauricio[4,5], Michael Miles[5], Thierry Wirth[6], Gabriele Schönian[1]*

1 Institut für Mikrobiologie und Hygiene, Charité Universitätsmedizin Berlin, Berlin, Germany, **2** Laboratório de Pesquisa em Leishmaniose, Instituto Oswaldo Cruz - Fiocruz, Rio de Janeiro, Brazil, **3** Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, **4** Instituto de Higiene e Medicina Tropical, Lisboa, Portugal, **5** Department of Pathogen Molecular Biology, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom, **6** Ecole Pratique des Hautes Etudes, Muséum National d'Histoire Naturelle, Département de Systématique et Évolution, UMR-CNRS 7205, Paris, France

Abstract

Background/Objectives: Parasites of the subgenus *Leishmania (Viannia)* cause varying clinical symptoms ranging from cutaneous leishmaniases (CL) with single or few lesions, disseminated CL (DL) with multiple lesions to disfiguring forms of mucocutaneous leishmaniasis (MCL). In this population genetics study, 37 strains of *L. (V.) guyanensis*, 63 of *L. (V.) braziliensis*, four of *L. (V.) shawi*, six of *L. (V.) lainsoni*, seven of *L. (V.) naiffi*, one each of *L. (V.) utingensis* and *L. (V.) lindenbergi*, and one *L. (V.) lainsoni*/*L. naiffi* hybrid from different endemic foci in Brazil were examined for variation at 15 hyper-variable microsatellite markers.

Methodology/Principal findings: The multilocus microsatellite profiles obtained for the 120 strains were analysed using both model- and distance-based methods. Significant genetic diversity was observed for all *L. (Viannia)* strains studied. The two cluster analysis approaches identified two principal genetic groups or populations, one consisting of strains of *L. (V.) guyanensis* from the Amazon region and the other of strains of *L. (V.) braziliensis* isolated along the Atlantic coast of Brazil. A third group comprised a heterogeneous assembly of species, including other strains of *L. braziliensis* isolated from the north of Brazil, which were extremely polymorphic. The latter strains seemed to be more closely related to those of *L. (V.) shawi*, *L. (V.) naiffi*, and *L. (V.) lainsoni*, also isolated in northern Brazilian foci. The MLMT approach identified an epidemic clone consisting of 13 strains of *L. braziliensis* from Minas Gerais, but evidence for recombination was obtained for the populations of *L. (V.) braziliensis* from the Atlantic coast and for *L. (V.) guyanensis*.

Conclusions/Significance: Different levels of recombination versus clonality seem to occur within the subgenus *L.* (*Viannia*). Though clearly departing from panmixia, sporadic, but long-term sustained recombination might explain the tremendous genetic diversity and limited population structure found for such *L.* (*Viannia*) strains.

Citation: Kuhls K, Cupolillo E, Silva SO, Schweynoch C, Côrtes Boité M, et al. (2013) Population Structure and Evidence for Both Clonality and Recombination among Brazilian Strains of the Subgenus *Leishmania* (*Viannia*). PLoS Negl Trop Dis 7(10): e2490. doi:10.1371/journal.pntd.0002490

Editor: Christophe Ravel, University of Montpellier, France

Received November 19, 2012; **Accepted** September 8, 2013; **Published** October 31, 2013

Funding: This work was supported by grants from the EU (INCO-CT2005-015407) and from the Brazilian financial agencies CNPq (304683/2009-4; 476727/2010-3; 400132/2011-7), FAPERJ (CNE-E-26/101.535/2010), FINEP (Sicol IV-01.06.1227.00), and FAPEMIG (CBB - APQ-01123-09). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

* E-mail: gabriele.schoenian@charite.de

¤ Current address: Molecular Biotechnology and Functional Genomics, University of Applied Sciences Wildau, Germany.

Introduction

The species of the subgenus *Leishmania (Viannia)* Lainson and Shaw, 1987, are exclusively endemic in the New World (NW) and infections of humans with these protozoan parasites constitute a significant public health problem in at least 18 countries of Latin America [1]. Subgenus *L. (Viannia)* parasites are capable of causing a variety of clinical symptoms ranging from cutaneous leishmaniasis (CL) with single or few lesions that may heal spontaneously, disseminated CL (DL) with multiple lesions, to disfiguring forms of mucocutaneous leishmaniasis (MCL) that may occur concomitantly or after remission of CL [2]. The outcome of human infections by *Leishmania* parasites is thought to be influenced by the immune status of the host and virulence of the infecting parasite [3]. At present, multilocus enzyme electrophoresis (MLEE) is the reference technique for the identification of *Leishmania* and was employed in most of the classification schemes, although MLEE is likely to be partially superseded by multilocus sequence typing (MLST). The application of numerical taxonomy and cladistic techniques to electrophoretic data has resulted in the identification of two species complexes in the subgenus *L. (Viannia)*, namely the *L. (V.) braziliensis* complex comprising *L. (V.) braziliensis* and *L. (V.) peruviana*, and the *L. (V.) guyanensis* complex comprising *L. (V.) guyanensis*, *L. (V.) panamensis* and *L. (V.) shawi*, and of at least four single species, *L. (V.) lainsoni*, *L. (V.) naiffi*, *L. (V.) lindenbergi* and *L. (V.) utingensis* (for review see [4]). This classification has been largely supported by a recent molecular study comparing *hsp*70 sequences of different *Leishmania* species [5].

Author Summary

Cutaneous leishmaniasis (CL) constitutes a significant public health problem in all federal states of Brazil. Most cases are caused by parasites of the subgenus *Leishmania (Viannia)* which can cause a variety of clinical symptoms ranging from single or few lesions, disseminated CL with multiple lesions, to disfiguring forms of mucocutaneous leishmaniasis. This study has used a multilocus microsatellite typing approach for exploring the genetic diversity and population structure among 120 strains representing different subgenus *L. (Viannia)* species and different Brazilian CL foci. Genetic diversity within the subgenus was much higher than expected, especially within *L. (V.) braziliensis*, *L. (V.) shawi*, *L. (V.) naiffi*, and *L. (V.) lainsoni* which were all from the north of Brazil. These strains could not be assigned to well-defined populations, but presented a rather loosely associated group. Strains of *L. (V.) braziliensis* isolated along the Atlantic coast of Brazil and strains of *L. (V.) guyanensis* formed, however, two clearly separated populations exhibiting remarkable levels of sexual exchange. The latter finding is in contrast to previous studies suggesting clonal modes of propagation or inbreeding for natural populations of *Leishmania* parasites and might explain the genetic heterogeneity and limited population structure for Brazilian strains of subgenus *L.* (*Viannia*) observed in this study.

In Brazil, CL is endemic in all federal states and an annual mean of 27,250 CL cases has been registered from 1990–2010 (http://portal.saude.gov.br/portal/saude/profissional/area.cfm?id_area = 1560). The disease is caused by six species of the subgenus *L.* (*Viannia*), *L. (V.) braziliensis*, *L. (V.) guyanensis*, *L. (V.) shawi*, *L. (V.) lainsoni*, *L. (V.) naiffi* and *L. (V.) lindenbergi* plus one species of the subgenus *L. (Leishmania)*, *L. (L.) amazonensis*. Around 6–7% of the CL patients will develop symptoms of MCL [6] mainly after infection with *L. (V.) braziliensis* and, to a lesser extent, also with *L. (V.) guyanensis* [7,8]. So far, DL has been also mainly associated with *L. (V.) braziliensis*, and only sporadic cases with *L. (V.) guyanensis* [1]. Severe anergic diffuse cutaneous leishmaniasis (DCL) may be a long term sequel in a minority of *L. (L.) amazonensis* infections [1].

Transmission of species of the subgenus *L.* (*Viannia*) involves different species of phlebotomine sand flies and a wide variety of wild and domestic animals have been implicated as reservoir hosts. Some species have a more restricted transmission cycle, whereas others are more complex with several different vectors and hosts in different ecological and geographical regions [1,3]. Sympatry of different subgenus *L. (Viannia)* species has been reported particularly in the Amazon region [9] where separate epidemiological patterns have been described, involving different sand fly species [10]. There is also increasing evidence that pathogenic *Leishmania* strains can be maintained in both sylvatic cycles, involving wild animals and sylvatic sand flies, and urban cycles involving domestic animals and peridomestic sand flies [11]. Subgenus *L.* (*Viannia*) parasites are characterized by tremendous genetic diversity and the described species vary enormously in their degree of such diversity [12,13]. Intra-specific polymorphisms are, for example, very frequent in *L. (V.) braziliensis* and *L. (V) naiffi*, and it has been suggested that the genetic diversity of the parasites is most probably related to the sand fly vector(s) and/or animal reservoir(s) involved in the transmission cycles. On the other hand, *L. (V.) guyanensis*, particularly strains circulating in the Brazilian Amazon region, and *L. (V.) shawi* have been found to be rather homogenous by MLEE and ITS-RFLP typing [12,13].

Multilocus microsatellite typing (MLMT) is currently the method of choice for molecular epidemiological and population genetic studies of different species of *Leishmania* (reviewed in [14]). It combines the advantages of co-dominance and higher discriminatory power when compared to MLEE, RAPD and the PCR-RFLP approaches used in many studies. Different sets of microsatellite markers have been designed and successfully applied for discriminating strains of subgenus *L. (Viannia)* with special emphasis on *L. (V) braziliensis* and *L. (V.) guyanensis* [15–19]. In a preliminary study we have demonstrated that our microsatellite marker set is highly discriminatory at intra-species level. Moreover, this genotyping scheme allows the detection of mild genetic structures at different levels, and is thus, relevant for epidemiological and population genetic studies of strains within the subgenus *L.* (*Viannia*) [17].

In the present study we investigated microsatellite variation in strains of the subgenus *L.* (*Viannia*) from different Brazilian foci endemic for CL. Strains of *L. (V.) braziliensis* isolated along the Atlantic coast of Brazil and strains of *L. (V.) guyanensis* formed two clearly separated populations. Evidence for significant levels of recombination was obtained for both of these populations, and in Minas Gerais the emergence of an epidemic clone of *L. (V.) braziliensis* was identified. A third loosely associated group comprised *L. (V.) braziliensis* strains and several other subgenus *L.* (*Viannia*) species, all from northern Brazil.

Materials and Methods

Ethics statement

Research in this study was subject to ethical review by the European Commission and approved as part of contract negotiation for Project LeishEpiNetSA (contract 01547): the work conformed to all relevant European regulations. The research was also reviewed and approved by the ethics committee of the London School of Hygiene and Tropical Medicine (approval 5092). The *Leishmania* strains isolated from human and animal hosts and analysed in this microsatellite analysis, were received from the "Coleção de Leishmania do Instituto Oswaldo Cruz – CLIOC ([http://clioc.fiocruz.br], registered at the World Data Centre for Microorganisms under the number WDCM731 and recognized as the depository authority by the Brazilian Ministry of the Environment, MMA/CGEN Deliberação CGEN 97 de 22/03/2005, Processo 02000.003672/2004-34), and from the collection hosted by the Universidade Federal de Minas Gerais in Belo Horizonte. Only previously gathered samples from animals have been used in this study. All human strains of *Leishmania* had been isolated from patients as part of normal diagnosis and treatment with no unnecessary invasive procedures and with written and/or verbal consent recorded at the time of clinical examination. Data on human isolates were coded and anonymised.

Parasite and DNA samples

Sources, designation, geographical origins, MLEE identification, if known, and clinical manifestation for the 120 subgenus *L.* (*Viannia*) strains from Brazil that were used in this study are listed in Table S1. These included 37 strains of *L. (V.) guyanensis*, 63 of *L. (V.) braziliensis*, four of *L. (V.) shawi*, six of *L. (V.) lainsoni*, seven of *L. (V.) naiffi*, one each of *L. (V.) utingensis* and *L. (V.) lindenbergi*, and one *L. (V.) lainsoni*/*L. (V.) naiffi* hybrid. Most of the strains were isolated from human CL cases, three from DL cases, and three strains from patients suffering from MCL. The reference strains were cloned, but all other strains represented uncloned material. Seven strains were isolated from sand fly vectors, and 18 from different animals, such as opossums (4), rodents (3), armadillos (3), dogs (2), sloths (2), pacas (2), a capuchin monkey (1) and a

Table 1. Numbers of strains per species, region, clinical picture and host of the 120 strains studied.

Species	Origin	Strains	CL	DL	MCL	Sand fly	Wild animal	Dog	nd	Zymodeme IOC/Z
L. guyanensis (total 37)	Amazonas	36	29			1	6			Z23
	Acre	1	1							Z110
L. braziliensis (total 63)	Pernambuco	14	12				2			Z45, 72–75, 78, 105
	Bahia	9	3	3	3					Z27
	Rio de Janeiro	4	4							Z27
	Espírito Santo	2						2		Z27
	Minas Gerais	15	15							nd
	Paraná	3	3							nd
	Ceará	2	1				1			Z27
	Pará	4	2			1			1	Z27
	Acre	8	8							Z78-84
	Amazonas	1	1							Z27
	Rondonia	1					1			Z53
L. shawi (total 4)	Pará	4				3	1			Z26
L. lainsoni (total 6)	Acre	3	3							Z86
	Pará	2	1				1			Z15
	Rondonia	1					1			Z15
L. naiffi (total 7)	Pará	5	1			1	3			Z36, 37, 38, 41
	Amazonas	2	2							Z36
L. naiffi/L. lainsoni hybrid (total 1)	Acre	1								Z87
L. utingensis (total 1)	Pará	1				1				Z101
L. lindenbergi (total 1)	Pará	1	1							Z102
overall		120	88	3	3	7	16	2	1	

VL – visceral leishmaniasis, CL – cutaneous leishmaniasis, MCL – mucocutaneous leishmaniasis, DL – disseminated cutaneous leishmaniasis, nd – not defined, IOC/Z – zymodemes according to the CLIOC system [11].

doi:10.1371/journal.pntd.0002490.t001

porcupine (1). Table 1 summarises the number of strains per species according to geographical origin, zymodeme and clinical picture.

Most strains were obtained from the FIOCRUZ *Leishmania* collection (Coleção de *Leishmania* do Instituto Oswaldo Cruz, CLIOC, WDCM731, http://clioc.fiocruz.br). Seventeen strains of *L. (V.) braziliensis*, 15 from Minas Gerais and two from Pará, were obtained from the collection of the Universidade Federal de Minas Gerais, Belo Horizonte. Sample preparation and MLEE typing, based on the electrophoretic mobility of 11 enzymes in agarose gel electrophoresis, were performed as previously described [12].

DNA was isolated using proteinase K- phenol/chloroform extraction [20] or the Wizard^{TM} Genomic DNA Purification System (Promega, Madison, WI, USA) according to the manufacturer's protocol, suspended in TE-buffer or distilled water and stored at 4°C until use.

PCR amplification assays and electrophoretic analysis of the microsatellite markers

The standard set of 15 primer pairs (CSg46, CSg47, CSg48, CSg53, CSg55, CSg59, 6F, 7G, 10F, 11H, 11C, B3H, B6F, AC01R, and AC16R), specific for *L. (Viannia)*, was used for amplification of microsatellite containing fragments, as previously described [17]. PCRs were performed with fluorescence-conjugated forward primers. Screening of length variations of the amplified markers was done by automated fragment analysis using the ABI PRISM GeneMapper (Applied Biosystems, Foster City, CA). After manual checking the microsatellite repeat numbers were calculated for all loci by comparing the sizes of the respective fragments to those of the strains MHOM/BR/00/LTB300 (*L. braziliensis*) and MHOM/SR/87/TRUUS1 (*L. guyanensis*), which were included as reference strains in every experiment, and for which the repeat numbers had been determined by sequencing. These repeat numbers were then multiplied by two, since we have used dinucleotide microsatellites throughout, after which the size of the flanking region was added, as determined by sequencing of the reference strains. This rigorous normalization process was applied to correct for small size differences that could occur due to the use of different sequencing machines and/or fluorescent dyes during the analyses. These normalized fragment sizes for all markers were assembled into a multilocus microsatellite profile for every strain under study. In 1.7% of all loci three or four peaks were observed of which only the two most prominent bands were included in the microsatellite profiles. The microsatellite profiles of all strains analysed in this study are given in Table S1.

Data analysis

Population structure was investigated using the STRUCTURE software [21], which applies a Bayesian model-based clustering approach. This algorithm identifies genetically distinct populations on the basis of allele frequencies. Genetic clusters are constructed from the genotypes identified, estimating for each strain the fraction of its genotype that belongs to each cluster. The following parameters were used: "burn-in" period of 20,000 iterations,

probability estimates based on 200,000 Markov Chain Monte Carlo iterations. The most appropriate number of populations was determined by comparing log likelihoods for values of K between 1 and 10, with ten runs performed for each K, and by calculating ΔK, which is based on the rate of change in the log probability of data between successive K values [22].

Factorial correspondence analysis (FCA) implemented in GENETIX v 4.03 software [23] was performed, which places the individuals in a three-dimensional space according to the degree of their allelic state similarities.

Phylogenetic analysis was based on microsatellite genetic distances, calculated with the program POPULATIONS 1.2.28 (http://bioinformatics.org/~tryphon/populations) for the numbers of repeats within each locus using the Chord-distance [24], which follows the infinite allele model (IAM). Neighbor-joining trees (NJ) were constructed with the POPULATIONS software and visualized with MEGA [25]. Additionally, phylogenetic networks were inferred from the distance matrix obtained from the microsatellite dataset by using the Neighbor-Net method in SplitsTree4 [26].

Microsatellite markers as well as populations were analysed with respect to diversity of alleles (*A*), expected (gene diversity) and observed heterozygosity (H_e and H_o, respectively) applying GDA (http://hydrodictyon. eeb.uconn.edu/people/plewis/software. php). F_{IS} is a measure of heterozygosity that assesses the level of identity within individuals compared to that between individuals. It ranges between −1 and 1, where a negative value corresponds to an excess of heterozygotes, and a positive value to heterozygote deficiency. $F_{IS} = 0$ indicates Hardy-Weinberg allele proportions. Mean F_{IS} estimates over loci in each population were calculated with the software FSTAT (version 2.9.3.2) [27] using Weir and Cockerman's (1984) unbiased estimators [28]. Confidence intervals per locus were assessed by randomization and bootsraping procedures over loci and individuals, implemented in GENETIX [23] using 1,000 random permutations. We also analysed the data, by computing estimates and tests of significance for various population genetic parameters. Genetic differentiation and gene flow was assessed by *F*-statistics [28–30] with the corresponding *P*-values (confidence test) using the MSA software [31]. Linkage between all pairs of loci in populations 1 and 2 was tested using the software ARLEQUIN, version 3.5 [32] and FSTAT [27]. *P*-values for multiple tests were corrected using a sequential Bonferroni correction to minimize the likelihood of Type 1 errors [33]. Composite digenic disequilibrium values were estimated and their significance was tested using Chi-square statistics as described by Weir [34]. An exact test for association between alleles across loci based on permutation [35] was also employed.

To assess the level of multilocus linkage disequilibrium, the Index of Association (I_A, multilocus) and the r_d statistic were calculated in MULTILOCUS 1.3b [36,37]. *P* values were derived through comparison to a null distribution of 1,000 randomizations; median values were taken from 1,000 diploid resamplings of the multiallelic dataset.

Results

Genetic diversity of Brazilian strains of subgenus *L.* (*Viannia*)

All 15 microsatellite markers were polymorphic in the 120 strains of the subgenus *L.* (*Viannia*) analysed here (Table 2). Taking together all strains, the number of alleles varied between 7–29, with a mean value of 15, with markers CSg47 and CSg48 being the most variable. The overall observed heterozygosity per marker ranged from 0.117 to 0.706 and was lower than the expected (0.646–0.940) for all markers. Overall inbreeding coefficients varied between 0.245 and 0.855. The discrepancy between expected and observed heterozygosity and the high F_{IS} values most probably reflects population substructuring (Wahlund effect) although the existence of a considerable amount of inbreeding cannot be ruled out. A total of 107 strains had unique MLMT profiles (Table S1). Five profiles, Lgua13, Lbra6, Lbra22, Lbra48 and Lsha3, were each shared by two strains. Nine of the 15 strains of *L. braziliensis* isolated between 1986 and 1992 from human CL cases in Minas Gerais presented indistinguishable MLMT profiles (Lbra26). Only two of the Minas Gerais strains were different from the predominating genotype (Figure 1, Figure S1).

More than two peaks were found for 30 of the 1800 loci (1.7%) analysed in this study. Twenty-three of the 120 strains presented such possibly aneuploid loci. Strain *L. braziliensis* L 2516 had more than two peaks in four loci, and strains *L. braziliensis* L-0018, *L. lainsoni* L-2500 and L-2503, and *L. naiffi* L-991 in two loci each. One putative aneuploid locus was seen in five strains of *L. guyanensis*, eight of *L. braziliensis*, two of *L. lainsoni*, two of *L. naiffi* and in the *L. naiffi*/*L. lainsoni* hybrid L-2490.

Population structure of Brazilian strains of the subgenus *L.* (*Viannia*)

The multilocus microsatellite profiles consisting of the repeat numbers for 15 markers were processed using both model-based and distance-based methods and the results of both analyses are compared in Figure 1 and Figure S1.

STRUCTURE analysis assigned the 120 Brazilian strains of the subgenus *L. (Viannia)* to three main populations (Table S1, Figure S2) as inferred by ΔK calculation. Population 1 consisted exclusively of strains of *L. (V.) guyanensis* from the state of Amazonas (n = 36). Only one strain of *L. (V.) guyanensis*, strain L-2493 from Acre, was not part of this population. This strain was previously identified as a new enzymatic variant, Z110, of *L. (V.) guyanensis* by MLEE (unpublished results) and presented few differences in the *hsp70* sequence and RFLP [38]. Population 2 comprised 43 strains of *L. (V.) braziliensis* isolated in the eastern states of the country, namely Pernambuco, Bahia, Minas Gerais, Espírito Santo, Rio de Janeiro, and Paraná. Population 3 (n = 41) presented a mixture of 20 strains of *L. (V.) braziliensis*, the single strain of *L. (V.) guyanensis* L-2493 from Acre mentioned above, all strains of *L. (V.) shawi*, *L. (V.) lainsoni*, *L. (V.) naiffi*, the *L. (V.) lainsoni*/*L. naiffi* hybrid and the single strains of *L. (V.) utingensis* and *L. (V.) lindenbergi*. Thirty-five of these strains were isolated in the north of Brazil, five in the northeast and one in Paraná. Several strains of *L. (V.) braziliensis* had mixed membership coefficients for Populations 2 and 3, especially those from Pará (Table S1), and strain L-2446 from Pernambuco for all three populations indicating gene flow between the populations.

The existence of Populations 1 and 2 was supported by FCA. However, strains of Population 3 were shown to be highly heterogeneous compared to Populations 1 and 2 (Figure 2). Two strains of Population 2, namely L-2516 and L-2446 both from Pernambuco, grouped within the cloud formed by Population 3. *F*-statistics revealed significant genetic differentiation between the three populations identified by STRUCTURE, especially between Populations 1 and 2 (Table 3). The correlation between population assignment and the geographical origin of the strains is shown in Figure 3.

The strains of *L. (V.) lainsoni* and of *L. (V.) naiffi*, including the *L. (V.) lainsoni*/*L. (V.) naiffi* hybrid, were assigned to distinct genetic groups only when Population 3 was re-analysed by STRUCTURE in order to check for hidden substructures within Population 3 (Table S1, Figure S2). Strains of *L. (V.) shawi*, however, did not present a separate entity in this analysis. When STRUCTURE was performed on Populations 1 and 2 separately,

Table 2. Characterization of the microsatellite markers for all strains and the three populations inferred by STRUCTURE.

Marker	Population *K*3	n	Repeat array	Fragment sizearray [bp]	A	H_e	H_o	F_{IS}
CSg46	Pop1	36	(AC) 6–16	71–91	4	0.651	1.000	−0.548
	Pop2	43	(AC) 11–23	81–105	4	0.113	0.093	0.178
	Pop3	41	(AC) 7–22	73–103	12	0.895	0.441	0.511
	overall	**120**	**(AC) 6–23**	**71–105**	**14**	**0.788**	**0.486**	**0.383**
CSg47	Pop1	36	(TG) 8–18	87–107	11	0.859	0.750	0.129
	Pop2	43	(TG) 6–29	83–129	15	0.845	0.884	−0.046
	Pop3	41	(TG) 2–34	75–139	21	0.881	0.475	0.464
	overall	**120**	**(TG) 2–34**	**75–139**	**29**	**0.940**	**0.706**	**0.250**
CSg48	Pop1	36	(TG) 7–8	76–78	2	0.081	0.028	0.660
	Pop2	43	(TG) 4–14	70–90	4	0.524	0.070	0.868
	Pop3	41	(TG) 3–27 (34)[1]	68–117 (131)[1]	19	0.886	0.244	0.727
	overall	**120**	**(TG) 3–27 (34)[1]**	**68–117 (131)[1]**	**20**	**0.801**	**0.117**	**0.855**
CSg53	Pop1	36	(AC) 7–9	84–88	2	0.155	0.055	0.645
	Pop2	43	(AC) 7–15	84–100	3	0.402	0.256	0.367
	Pop3	41	(AC) 5–19	80–108	13	0.896	0.342	0.621
	overall	**120**	**(AC) 5–19**	**80–108**	**13**	**0.758**	**0.222**	**0.708**
CSg55	Pop1	36	(TG) 16–21	103–113	6	0.714	0.583	0.186
	Pop2	43	(TG) 8–13	87–97	3	0.046	0.046	−0.006
	Pop3	41	(TG) 8–20	87–111	11	0.666	0.175	0.740
	overall	**120**	**(TG) 8–21**	**87–113**	**13**	**0.685**	**0.252**	**0.633**
CSg59	Pop1	36	(TC) 6–8	94–98	3	0.394	0.417	−0.058
	Pop2	43	(TC) 7–8	96–98	2	0.492	0.643	−0.312
	Pop3	41	(TC) 3–10	88–102	7	0.780	0.341	0.565
	overall	**120**	**(TC) 3–10**	**88–102**	**7**	**0.646**	**0.471**	**0.273**
7G	Pop1	36	(AC) 5–9	88–96	3	0.133	0.139	−0.042
	Pop2	43	(AC) 5–17	88–100	4	0.389	0.349	0.104
	Pop3	41	(AC) 0–24	78–126	16	0.885	0.585	0.342
	overall	**120**	**(AC) 0–24**	**78–126**	**17**	**0.798**	**0.367**	**0.542**
11H	Pop1	36	(GT) 7–14	86–100	6	0.605	0.056	0.909
	Pop2	43	(GT) 7–11	86–94	3	0.298	0.116	0.613
	Pop3	41	(GT) 5–23	82–118	17	0.910	0.540	0.410
	overall	**120**	**(GT) 5–23**	**82–118**	**17**	**0.796**	**0.233**	**0.708**
11C	Pop1	36	(TG) 5–8	90–96	4	0.540	0.528	0.023
	Pop2	43	(TG) 3–10	86–100	4	0.525	0.395	0.249
	Pop3	41	(TG) 1–29 (42)[1]	82–138 (164)[1]	16	0.906	0.750	0.175
	overall	**120**	**(TG) 1–29 (42)[1]**	**82–138 (164)[1]**	**17**	**0.812**	**0.548**	**0.327**
6F	Pop1	36	(AC) 7–10	83–89	3	0.206	0.111	0.465
	Pop2	43	(AC) 7–12	83–93	3	0.353	0.349	0.012
	Pop3	41	(AC) 5–24	79–117	15	0.913	0.474	0.485
	overall	**120**	**(AC) 5–24**	**79–117**	**16**	**0.790**	**0.316**	**0.601**
10F	Pop1	36	(CA) 15–16	97–99	2	0.366	0.361	0.013
	Pop2	43	(CA) 13–21	93–109	5	0.364	0.419	−0.152
	Pop3	41	(CA) 12–21	91–109	9	0.839	0.400	0.527
	overall	**120**	**(CA) 12–21**	**91–109**	**9**	**0.733**	**0.395**	**0.462**
B6F	Pop1	36	(AC) 7–9	81–85	3	0.469	0.111	0.766
	Pop2	43	(AC) 6–20	79–107	11	0.690	0.738	−0.071
	Pop3	41	(AC) 2–21	71–109	14	0.855	0.300	0.652
	overall	**120**	**(AC) 2–21**	**71–109**	**16**	**0.855**	**0.398**	**0.535**
B3H	Pop1	36	(AC) 7–10	65–71	4	0.626	0.611	0.023

Table 2. Cont.

Marker	Population *K*3	n	Repeat array	Fragment sizearray [bp]	A	H_e	H_o	F_{IS}
	Pop2	43	(AC) 9–18	69–87	6	0.670	0.953	−0.431
	Pop3	41	(AC) 4–22	59–95	14	0.878	0.297	0.664
	overall	**120**	**(AC) 4–22**	**59–95**	**14**	**0.857**	**0.638**	**0.256**
AC01R	Pop1	36	(CA) 11–17	105–117	5	0.580	0.429	0.264
	Pop2	43	(CA) 4–8	91–99	2	0.023	0.023	0.000
	Pop3	41	(CA) 2–36 (46)[1]	87–155 (175)[1]	15	0.781	0.375	0.523
	overall	**120**	**(CA) 2–36 (46)[1]**	**87–155 (175)[1]**	**16**	**0.699**	**0.263**	**0.625**
AC16R	Pop1	36	(TG) 12–15	91–97	4	0.639	0.583	0.089
	Pop2	43	(TG) 9–14	85–95	4	0.497	0.651	−0.315
	Pop3	41	(TG) 9–26	85–119	14	0.867	0.475	0.456
	overall	**120**	**(TG) 9–26**	**85–119**	**14**	**0.756**	**0.571**	**0.245**

N, number of strains; A, number of alleles; H_o, observed heterozygosity; H_e, expected heterozygosity; F_{IS}, inbreeding coefficient;
[1]Alleles which occurred in exceptional single cases are given in brackets.
doi:10.1371/journal.pntd.0002490.t002

Population 1 did not show meaningful subdivision, but Population 2 was split in two sub-populations with the majority of *L. (V.) braziliensis* strains from Pernambuco and Bahia (northeast Brazil) found in 2A and strains from Minas Gerais, Espírito Santo, Rio de Janeiro and Paraná (southeast Brazil) in 2B (Table S1).

The distance analysis based on the MLMT profiles obtained for all individual strains and the inferred neighbor-joining tree (Figure 1, Figure S1) corroborated the results of FCA analysis. Populations 1 and 2 formed separate clusters corresponding to the populations observed by STRUCTURE. Strains of Population 3, however, did not appear as a monophyletic group in the tree. All strains of *L. (V.) lainsoni* were found in a small well-separated cluster. A distinct cluster was also obtained for the strains of *L. naiffi* plus the single strains of *L. (V.) utingensis* and *L. (V.) lindenbergi*, albeit with long branches. Interestingly, these strains were most closely related to five strains of *L. (V.) braziliensis*, three from Pernambuco and the two from Acre with mixed memberships for sub-populations 3A and 3D (Figure S2). The strains of *L. (V.) shawi* grouped together with the *L. (V.) braziliensis* strain L-0326 from Pará. The *L. (V.) guyanensis* strain L-2493 did not group with the other strains of *L. (V.) guyanensis*, but was closest to the strains of *L. (V.) shawi*. Another interesting cluster consisted of six relatively diverse *L. (V.) braziliensis* strains from Acre. The remaining strains of Population 3 were intermediate between the "Acre" cluster and Population 2. Most of these strains had mixed memberships in Populations 2 and 3 in the STRUCTURE analyses (Table S1) and might represent hybrids or mixed infections as cloned isolates were not used.

The phylogenetic NeighborNet network (Figure 4) largely confirmed the results described above. It clearly showed the tremendous diversity of the strains assigned to Population 3 and its four sub-populations. However, conflicting splits represented by boxes can be seen between and within the three main populations. The same analysis was carried out for each of the populations separately. The obtained phylogenetic networks confirm most of the sub-structures found previously by Bayesian analysis (Figures S3, S4, S5).

Population genetics characterization of identified populations

The difference in the degree of microsatellite polymorphism between the three main populations was also reflected by mean number of alleles (MNA) which fluctuated from 4.1 to 4.9 in Populations 1 and 2, respectively, up to 14.2 in population 3 (Table 4). In Populations 1 and 2 the observed heterozygosities (Table 4) were close to those expected under Hardy-Weinberg equilibrium (HWE), whilst global F_{IS} values were only moderately positive for Population 1 and Population 2 (Fig. S6). Significant ($P<0.05$) excess of heterozygosity indicated by negative F_{IS} values was detected for only two loci, locus CSg47 in population 1 (−0.551) and locus B3H in population 2. Another 8 loci presented significant deficit of heterozygosity in a least one of the two populations. Both global F_{IS} values were significant (0.184 and 0.088 respectively; $P=0.0017$ and $P=0.0167$) and therefore panmixia must be rejected. Testing HWE within population 1 and 2 supported this scenario since both populations significantly departed from the null-hypothesis ($P<0.002$ and $P<0.001$, respectively). The Maynard Smith index of association, I_A, [36,39], which assesses multilocus linkage disequilibrium, was calculated. Another estimator, r_d, was implemented because the I_A tends to increase with the number of loci, a trend corrected by this statistic. To summarise, both populations displayed positive I_A and r_d values (1.295 and 0.097 for population 1 and 1.586 and 0.123 for populations 2, respectively) departing significantly ($P<0.001$) from panmixia (I_A *and* $r_d=0$).

However, the F_{IS} values gathered from populations 1 and 2 are consistently lower than those observed in previous reports [40,41], indicating that gene conversion or recombination may play a substantial role in *Viannia* species present in the Amazon Basin and along the Atlantic coast and to a lesser extent westward from the Andes. In contrast, in Population 3 the observed heterozygosity was much less than the expected resulting in a high F_{IS} value, most probably due to population subdivision (Wahlund effect) although high rates of gene conversion or inbreeding cannot be excluded. This group was not tested for all population genetic parameters, since it represents a composite and artificial unit.

In order to test whether associations between the 15 microsatellite loci are in gametic equilibrium in populations 1 and 2, as expected for random-mating populations, we have applied both the composite disequilibrium test and the exact test to all 105 pairwise comparisons. For population 1, only one strain representing the microsatellite profile that was shared by two strains (Table S1) was included in the data matrix. Using Chi-square test

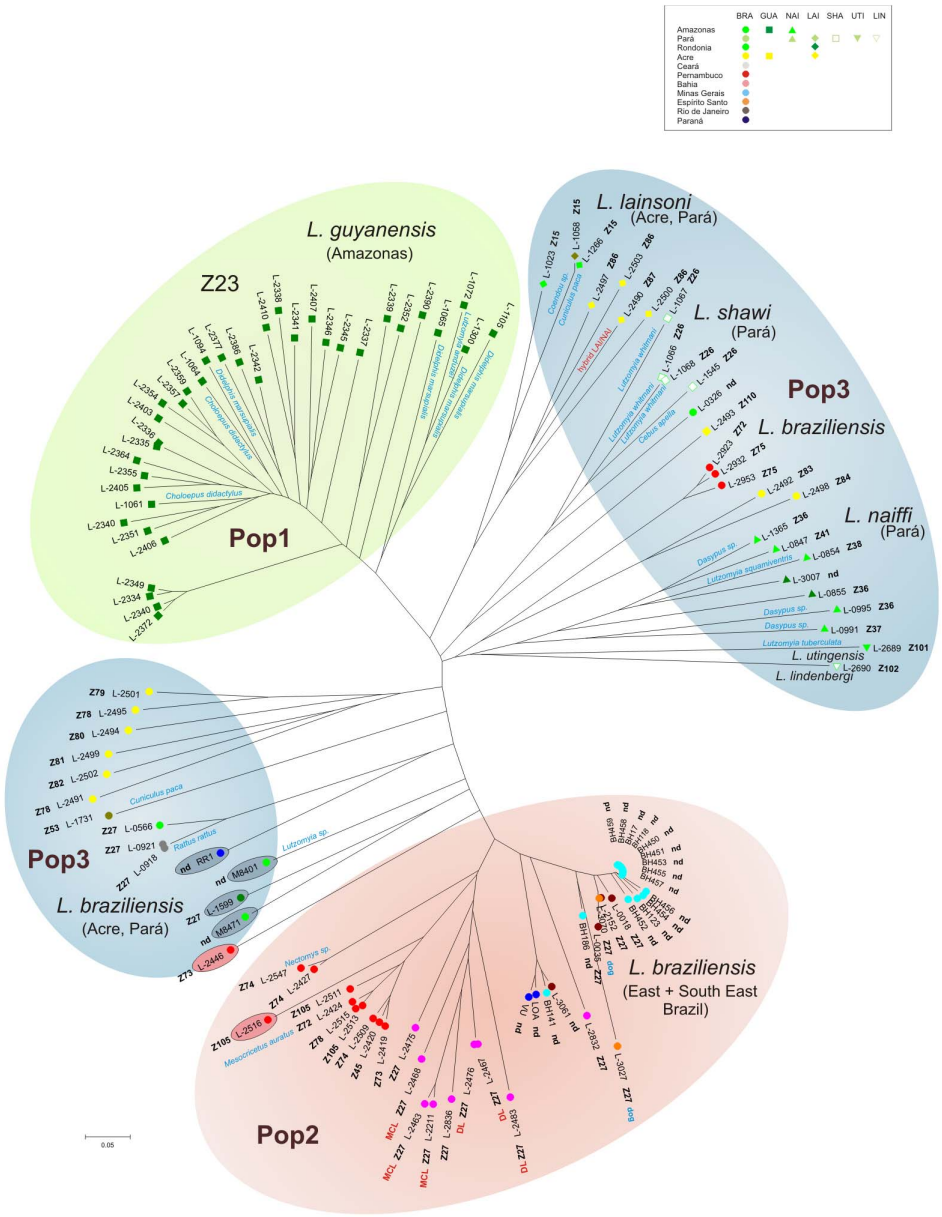

Figure 1. Populations and subpopulations of Brazilian strains of the subgenus of *Leishmania* (*Viannia*). A midpoint rooted Neighbour-joining (NJ) tree (radial version, rectangular version in Figure S1) was calculated for the MLMT profiles of 120 strains of different species of subgenus *L.* (*Viannia*), based on 15 microsatellite markers and using the Chord distance measure. The assignment of these strains to three main populations by the Bayesian model-based clustering approach implemented in STRUCTURE is indicated by coloured circles: population 1 (green), population 2 (red) and population 3 (blue). Strains belonging to these populations are listed in Table S1. Population 1 comprises all but one strain of *L. (V.) guyanensis* analysed in this study. Population 2 consists of 43 strains of *L. (V.) braziliensis* mainly from east Brazil. Population 3 is very diverse and includes all investigated strains of *L. (V.) lainsoni*, *L. (V.) naiffi*, *L. (V.) shawi*, *L. (V.) utingensis*, *L. (V.) lindenbergi*, 20 strains of *L. (V.) braziliensis* mainly from the north of Brazil as well as one strain of *L. (V.) guyanensis* from Acre. Putative hybrids are indicated by red or blue circles, according to their population assignment. Strain origins are indicated in the window alongside.
doi:10.1371/journal.pntd.0002490.g001

and exact test, 24 and 37 significant associations were respectively detected. However, after Bonferroni correction, those numbers dropped to 6 and 10. In population 2 nine strains of *L. braziliensis* from Minas Gerais shared an identical genotype, Lbra26, and further four strains had highly related genotypes, Lbra 28, 29, 30 and 31. These strains most probably represent an epidemic clone and only one Lbra 26 strain was therefore included in the data matrix for linkage disequilibrium calculations. The Chi-square test and the exact test revealed 30 and 49 significant associations (28 and 45 after Bonferroni correction), respectively for population 2. The three populations detected according to the Bayesian algorithm were also significantly separate from each other with highly significant F_{st} values ranging from 0.249 to 0.521 (Table 3).

Discussion

Because of its high resolution potential, its reproducibility and the possibility of data storage, multilocus microsatellite typing

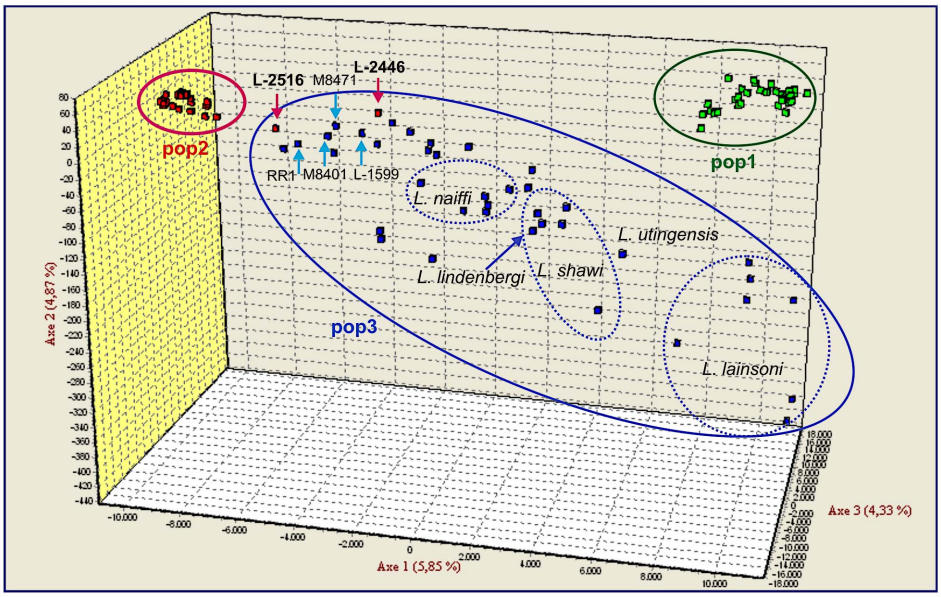

Figure 2. Factorial correspondence analysis (FCA) of 120 Brazilian strains of the subgenus *L. (Viannia)*. The strains labelled in green, red and blue correspond to those that were assigned, by STRUCTURE, to populations 1, 2 and 3, respectively. The two strains of Population 2, L2516 and L2446 that grouped within the blue cloud are indicated by a pink arrow. Four strains of mixed population membership, with predominating traits of population 3, are indicated by blue arrows.
doi:10.1371/journal.pntd.0002490.g002

(MLMT) is currently the most widely used approach for strain level differentiation in the genus *Leishmania*. Analysis of length polymorphisms of microsatellite-containing sequences has recently revealed geographical and hierarchical population structures in different *Leishmania* species such as *L. (L.) tropica*, *L. (L.) major* and the *L. (L.) donovani* complex (for review see [14]), the predominance of inbreeding in *L. (V.) braziliensis* and *L. (V.) guyanensis* [40,41], and confirmed that the agent of VL in the NW is *L. (L.) infantum*, which has been recently imported multiple times from southwest Europe to the New World [42]. In the present study, a MLMT approach employing 15 microsatellite markers previously shown to be highly discriminatory for strains of the subgenus *L. (Viannia)* [17], was used to explore the genetic diversity of 120 strains from Brazil in order to unravel discrete populations.

Population structure of Brazilian strains of the subgenus *L.* (*Viannia*)

Different types of population genetic analyses, including Bayesian inference (as implemented in STRUCTURE), distance-based (NJ and Neighbor Net in SplitsTree) and factorial correspondence analysis as well as *F* statistics revealed the existence of two well-defined populations in the sample set, namely Population 1 consisting of all but one strain of *L. (V.) guyanensis* from the Amazonas state and Population 2 comprising 43 strains of *L. (V.) braziliensis* from Eastern Brazil.

STRUCTURE identified a third population including 20 strains of *L. (V.) braziliensis*, one of *L. (V.) guyanensis* and the strains belonging to other species, most of which were isolated in the north of the country. When the strains of *L. (V.) lainsoni*, *L. (V.) naiffi*, *L. (V.) shawi*, *L. (V.) utingensis* and *L. (V.) lindenbergi* were excluded from the data set, the strains of *L. (V.) braziliensis* were assigned to the same Populations 2 and 3 (data not shown). The Population 3 identified in the complete data set was, however, not well supported by the distance and FCA analyses which showed that the strains of this population are highly diverse and only distantly related to each other as can be clearly seen in the NeighborNet network produced in SplitsTree (Figure 4). When re-analysed by STRUCTURE, Population 3 split into four sub-populations (Figure S2) and strains of *L. (V.) lainsoni* (3C) and *L. (V.) naiffi* (3D) were assigned to separate sub-populations. Two strains of *L. (V.) braziliensis* from Acre (L-2492 and L-2498) were found to be putative *L. (V.) braziliensis*/*L. (V.) naiffi* hybrids, as previously suggested [38], and were also assigned to the "*L. (V.) naiffi*" sub-population. However, strains of *L. (V.) shawi* grouped together with ten strains of *L. (V.) braziliensis*, mainly from Acre, strain L-2493 of *L. (V.) guyanensis* and the strains of *L. (V.) utingensis* and *L. (V.) lindenbergi* (3A), and the fourth sub-population (3B) consisted of *L. (V.) braziliensis* from Pernambuco, Ceará, Pará, Paraná and Amazonas. This last group seems, however to be rather artificial, since seven of the ten strains have hybrid genotypes sharing alleles that are specific for Population 2 and 3, respectively. Strains of *L. (V.) braziliensis* of population 3 are considerably different from each other and very distinct from those that were assigned to Population 2, except those having mixed memberships to Populations 2 and 3. Whether such strains represent outliers, as stated in a different study that used AFLP for typing strains of *L. (V.) braziliensis* and *L. (V.) peruviana* mainly from Peru and Bolivia [43], or different taxa requires further investigations including additional strains and using DNA sequence-based comparisons.

Table 3. F_{ST} values and corresponding *p*-values for the main three populations found by STRUCTURE.

F_{ST}-values	Pop1	Pop2	Pop3
Pop1 (36)	0	0.521	0.250
Pop2 (43)	0.0001	0	0.249
Pop3 (41)	0.0001	0.0001	0

F_{ST} values are in the upper triangle, *p*-values in the lower triangle. Number of strains belonging to each population is given in brackets.
doi:10.1371/journal.pntd.0002490.t003

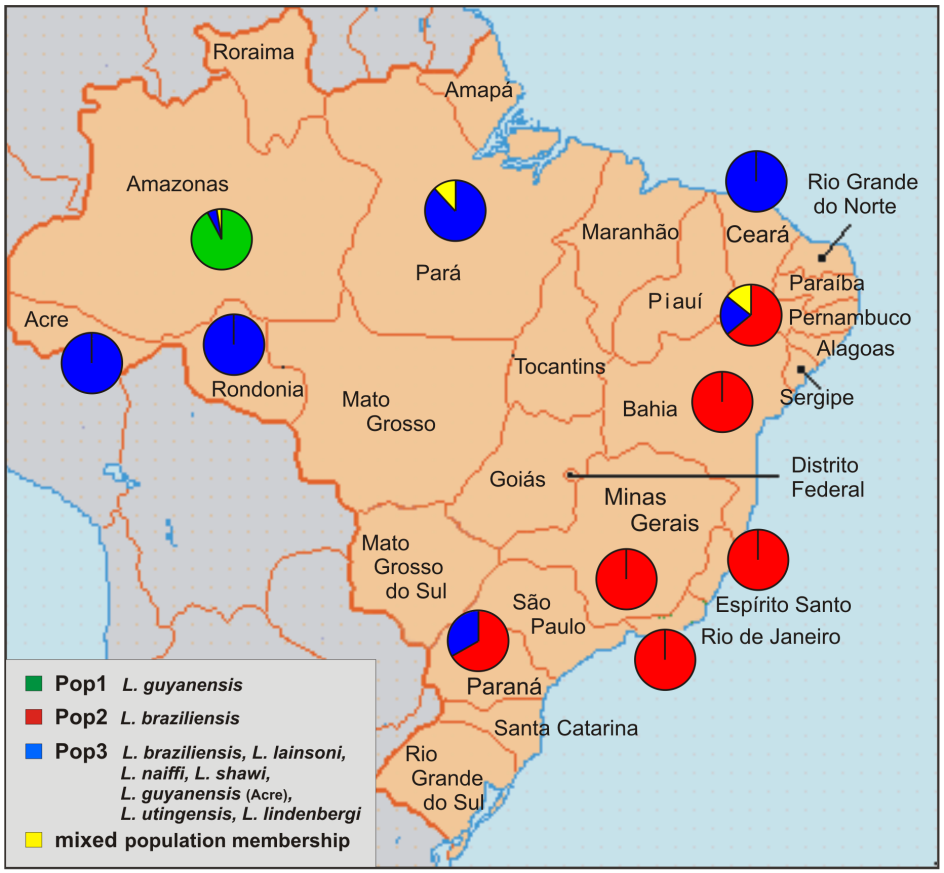

Figure 3. Geographical distribution of the three main populations inferred by STRUCTURE. According to the Bayesian clustering algorithm (STRUCTURE) the Brazilian strains of *L.* (*Viannia*) were assigned to three different populations, shown in green, red and blue. Pie-charts show the proportion of each population sampled in the respective geographical region. The distribution of the main populations correlates, at least partially, with the geographical origin of the strains.
doi:10.1371/journal.pntd.0002490.g003

NeighborNet analyses provide a snapshot overview of the general structure in the data and are useful as a guide for further analysis [44]. The phylogenetic network obtained here for the full sample set is distinctly non-treelike and demonstrates marked ambiguity in the signal. Only Populations 1 and 2 formed distinct clusters in the network. The reticulate patterns seen in the network between (Figure 4), and within the three main populations (Figures S3, S4, S5) could result from hybridization, recombination events or gene conversion.

For the strains assigned by STRUCTURE to Population 3 no clear structuring is seen in the NeighborNet network obtained for the whole data set (Figure 4), maybe with the exception of the strains of *L. (V.) lainsoni* which seem to represent a distinct lineage in the network. When the NeighborNet analysis was performed for Population 3 only, strains of *L. (V.) lainsoni*, *L. (V.) shawi* and *L. (V.) naiffi* were assigned to distinct clusters although conflicting signals were still detectable (Figure S5). The independent status of *L. shawi* and its apparent affiliation with *L. (V.) braziliensis* by this high resolution microsatellite analysis does not agree with results of MLST analysis, in which *L. (V.) shawi* and *L. (V.) guyanensis* are not resolved as separate entities [45]. MLEE, PCR-RFLP of the ribosomal ITS and sequencing of the *hsp*70 had already previously suggested that *L. (V.) guyanensis and L. (V.) shawi* were closely related [12,13,38], strains of *L. (V.) braziliensis* from northeastern Brazil belonging to zymodeme Z75 were however, found to be related to *L. (V.) shawi* [46] which is in agreement with the results of our MLMT study. Assessment of the taxonomic status of *L. shawi* thus warrants further investigation with more extensive DNA sequence comparisons. All strains of population 3 and only those are found on long branches in the overall network. Because NeighborNet is prone to long-branch attraction, rapidly evolving lineages can be inferred as being closely related regardless of their evolutionary relationships. Whether these strains have a high mutation rate leading to numerous homoplasies or to convergence, which could be misinterpreted as having evolved once in a common ancestor, remains to be established.

Given its high genetic diversity, Population 3 could represent the ancestral lineage and might have given rise to two new populations through bottleneck events (Populations 1 and 2). The Amazon forest seems to be the central distribution area with secondary spreads to the northeast, east and south. This would be consistent with the hypothesis of an Amazon origin of CL in Brazil, with later spread to other regions, most probably through human migrations [6]. However, we cannot exclude that sampling biases are responsible for the weak resolution of the strains in Population 3. More extensive sampling in the north of Brazil, where most of these strains were isolated, is needed to address these questions.

Genetic diversity of Brazilian strains of the subgenus *L.* (*Viannia*)

Almost all strains investigated in this study presented unique microsatellite profiles, except 13 strains from Minas Gerais that

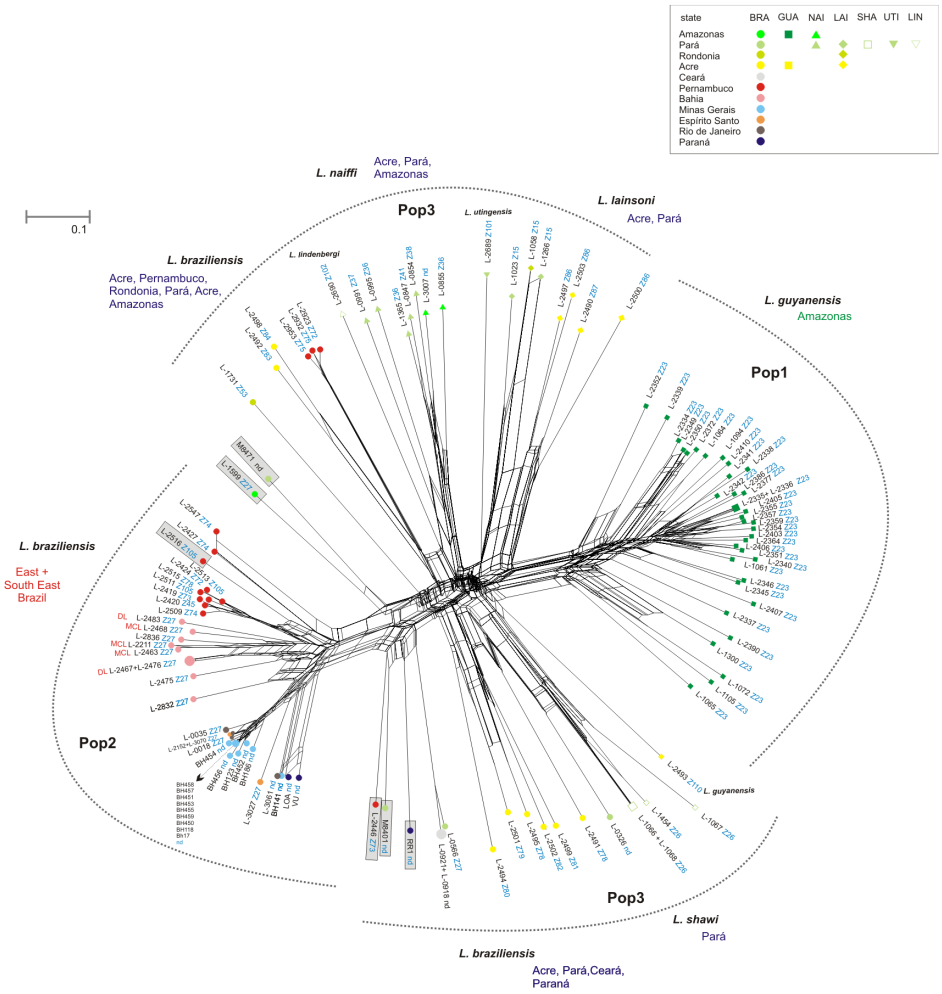

Figure 4. NeighborNet network based on the MLMT profiles of 120 Brazilian *L. (Viannia)* strains. The network was obtained using SplitsTree4 software and calculation of Chord distances for the 15 microsatellite markers used. The assignment of the strains to the sub-populations as inferred by STRUCTURE is indicated.
doi:10.1371/journal.pntd.0002490.g004

had identical or highly similar microsatellite profiles and might have been isolated during an outbreak of CL in this area.

Previous studies using isoenzyme typing [12] or PCR-RFLP of the ribosomal internal transcribed spacer [13] have already demonstrated that *L. (V.) braziliensis* is much more polymorphic than *L. (V.) guyanensis*. Furthermore, strains of *L. (V.) braziliensis* from the Amazon region and from Pernambuco were shown to present the highest level of genetic diversity whereas those from Rio de Janeiro were more homogeneous [11,46]. In our study the MNA (Table 4) was similar for strains of *L. (V.) guyanensis* (Population 1) and of *L. (V.) braziliensis* from east Brazil (Population 2) indicating a similar level of genetic diversity for these groups of strains. For *L. (V.) braziliensis* strains of Population 3, MNA and thus genetic heterogeneity was however, considerably higher. This confirms that the previously described greater diversity of *L. (V.) braziliensis* is related to the genetically heterogeneous strains from the north of Brazil.

The strains of *L. (V.) guyanensis* investigated herein were more diverse than expected considering that they all belonged to the same zymodeme Z23 which is in agreement the results of a recent MLST study [45]. Only two strains, MHOM/BR/1997/203P and MHOM/BR/1997/203G, isolated from the skin (P) and a lymph node (G) of the same patient, shared an identical MLMT profile (Figure S1), all other strains presented different patterns of microsatellite variation. Previous studies using monoclonal antibodies had already pointed to the existence of two distinct sub-populations in the Brazilian Amazon region [47]. In our study, the existence of two *L. (V.) guyanensis* sub-populations in the Amazonas state (data not shown) was not supported by the genetic distance analyses. Our MLMT approach confirmed however, the existence of a new *L. (V.) guyanensis* genotype in Acre, where this species is not commonly found. This strain was very closely related to *L. (V.) shawi* as previously suggested [38]. In French Guyana, two distinct populations of *L. (V.) guyanensis* had been studied using a PCR-RFLP approach targeting ribosomal DNA sequences and found to have originated from two ecologically different regions and to differ in clinical manifestations of CL [48]. In a previous study comparing microsatellite variation in a limited number of strains of *L. (V.) guyanensis* from Brazil, Peru, Suriname and French Guyana, the strains grouped according to their geographical origin [17]. Future investigations should include strains sampled in different locations, since it has been speculated recently that strains from the eastern and southern Brazilian Amazon region might represent different genetic groups of *L. (V.) guyanensis* [8].

Table 4. Population genetics indices of the three main populations detected by STRUCTURE.

Pop	Species	Region	N	P	MNA	H_e	H_o	F_{IS}[a]	F_{IS}[b]
Pop1	*L. guyanensis* (36/37)	Amazonas (36/36)	36	1	4.1	0.468	0.384	0.181	0.184
Pop2	*L. braziliensis* (43/63)	**Bahia (9/9),**	43	1	4.9	0.415	0.399	0.040	0.088
		Pernambuco (11/14),							
		Rio de Janeiro (4/4)							
		Espírito Santo (2/2)							
		Minas Gerais (15/15),							
		Paraná (2/3)							
Pop3	*L. guyanensis* (1/37),	**Acre (1/1)**	41	1	14.2	0.856	0.414	0.519	n.d.
	L. braziliensis (20/63),	**Rondonia (1/1),**							
		Amazonas (1/1),							
		Ceará (2/2),							
		Pernambuco (3/14),							
		Paraná (1/3)							
		Pará (4/4)							
		Acre (8/8)							
	L. shawi (4/4),	**Pará**							
	L. naiffi (7/7),	**Pará, Amazonas**							
	L. lainsoni (6/6),	**Acre, Pará, Rondonia**							
	*L.nai./L.lain.*hybr. (1/1)	**Acre**							
	L. utingensis (1/1),	**Pará**							
	L. lindenbergi (1/1)	**Pará**							
Overall			120	1	15.5	0.781	0.399	0.490	n.d.

N, number of strains; P, proportion of polymorphic loci; MNA, mean number of alleles; H_o, observed heterozygosity; H_e, expected heterozygosity; F_{is}, inbreeding coefficient, F_{is}[a] all strains of the data set were included in the calculations, F_{is}[b] only one strain per genotype was included in the calculations; n.d., not done. Predominating regions are marked by bold letters, normal letters are used for regions for which single strains are found in the respective population.
doi:10.1371/journal.pntd.0002490.t004

We found only weak correlations between the MLMT profiles and the results of previous isoenzyme typing (Table S1, Table S2). As already mentioned above, the strains of *L. (V.) guyanensis*, despite all being of zymodeme Z23, could be individualised by MLMT but were all grouped in Population 1, with the exception of the single strain from Acre. Strains of *L. (V.) braziliensis* with identical isoenzyme patterns were assigned to different populations or genetic groups by MLMT. The majority of the strains presenting the predominant zymodeme Z27 grouped in Population 2 but those with hybrid genotypes were found in Population 3B. This implies that zymodeme Z27 is paraphyletic and does not reflect the genetic diversity of the strains which was also shown by the MLST study [45].

We did not find any correlation between a particular MLMT profile and the clinical presentation of the disease. The two strains isolated from MCL patients in Bahia grouped together with strains from CL cases from the same area, although the long term outcome of those CL cases is not known. This is consistent with previous studies which suggested that the clinical outcome of the disease caused by *L. (Viannia)* parasites is also influenced by host genetic and/or immune factors [49–51] which could possibly be stimulated through pre-exposure to sand fly saliva [52]. In conclusion, the only correlation found for MLMT patterns of the *L. (Viannia)* strains studied herein was that to their geographical origin. Similar observations were published earlier for *L. (V.) braziliensis* [11] and *L. (L.) infantum* in Brazil [53] and might be associated with different transmission cycles with different sand fly vectors and/or animal reservoirs involved in those areas.

Reproductive strategies among Brazilian strains of the subgenus *L.* (*Viannia*)

Despite the fact that recombination in *Leishmania* has been proved to occur *in vitro* in the sand fly hosts [54,55] and the growing evidence of gene flow coming from different population genetic studies using MLST and MLMT approaches (reviewed in [14]), *Leishmania* species are still considered as predominantly clonal organisms [56]. Especially in the case of strains of the subgenus *L.* (*Viannia*) this hypothesis has been challenged by the frequent detection of hybrids involving different species of the *L. (Viannia)* subgenus indicating that recombination events are much more frequent in these parasites than previously thought [12,57]. *L. (V.) braziliensis/L. (V.) peruviana* hybrids have been identified by microsatellite typing and found to be quite common in Peruvian foci where both species can occur sympatrically [58] and *L. (V.) braziliensis/L. (V.) guyanensis* hybrids are not uncommon [59–61]. Hybrid MLEE profiles have been observed for *L. (V.) laisoni/L. (V.) naiffi*, *L. (V.) braziliensis/L. (V.) naiffi*, *L. (V.) braziliensis/L. (V.) guyanensis* in Brazil [62], and these hybrids had been mostly isolated from patients living in areas with sympatric circulation of both putative parental species. In our study, some strains had mixed membership of the different populations identified and were considered to be putative hybrids, although this will have to be confirmed by analysis of cloned parasites. A recently published MLSA study of subgenus *L.* (*Viannia*) strains has also provided evidence for recombination occurring in both *L. braziliensis* and *L. guyanensis* [45].

Strong linkage disequilibrium, or non-random association of genotypes at different loci, and a distinct phylogenetic signal are

the criteria for the identification of clonality [56]. To start with the latter, strong phylogenetic signals are clearly absent in both the NJ distance tree (Figure 1 and Figure S1) and the NeighborNet network (Figure 4). Calculation of linkage disequilibrium which does not depend on the ploidy status [56] revealed a higher number of significant associations between loci for both populations 1 and 2 than what would be expected due to chance for a random mating population. On the other hand, the majority of the pairwise comparisons were not significant and many more loci appear to be recombining than would be expected for a strictly clonal population. We can, however not exclude that population subdivision (Wahlund effect) accounts, at least partially, for the amount of disequilibrium found for populations 1 and 2. Taking in consideration the limited linkage disequilibrium, the absence of overrepresented genotypes and the weak phylogenetic signal observed in this study, we would conclude that recombination has an important impact in populations 1 and 2. The MLMT approach used here, however also identified an epidemic clone consisting of 13 strains of *L. braziliensis* isolated between 1986 and 1992 from human CL cases in Minas Gerais. Nine of these strains shared an identical microsatellite profile (Lgua26) and the profiles of the other four strains differed from profile Lgua26 in only one locus each.

Different levels of clonality versus recombination have been earlier suggested to occur within some bacterial and protozoan species due to variation in geographic sampling [36] and this might be also the case for species of the subgenus *L.* (*Viannia*). The substantial heterozygote deficiency and extreme inbreeding found by MLMT analysis of *L. (V.) braziliensis* from Bolivia and Peru was consistent with a predominantly endogamic mode of reproduction (mating with relatives) with occasional recombination events between individuals of different genotypes [40]. In contrast, significant homozygosity and only little linkage disequilibrium was observed for populations of *L. (V.) guyanensis* from French Guyana suggesting a high level of sexual recombination and substantial endogamy [41]. Ramirez et al. [63] stated that heterozygosity statistics at microsatellite loci has to be interpreted with caution in the context of parasite sexuality because strong linkage disequilibrium can be accompanied by negative and positive *Fis* values. In our study, mild linkage disequilibrium was observed together with relatively low *Fis* values, compared to those previously published [40,41]. Indeed the *Fis* values observed for the *L. guyanensis* strains from the Brazilian Amazon region (0.184) and for the *L. braziliensis* strains from Eastern Brazil (0.088) are below the values observed in former publications (0.278 and 0.307) and this is without partitioning steps to correct for a potential Wahlund effect, which might deflate this index. Consequently, the selfing rates for the considered populations are likely to be $\ll 0.50$. This implies that clonality, selfing and random union of gametes contribute to the shaping of *Viannia*'s natural populations. However, one of the future challenges is to understand why the contribution of sex is more significant in the Amazon Basin.

The analyses used for calculations of heterozygosity are based on the assumption of diploidy. Recently, whole genome sequencing and FISH analyses have, however confirmed significant chromosomal copy number variations for different species of *Leishmania* [64–68]. In the only *L. braziliensis* strain, MHOM/BR/75//M2904, that has been fully sequenced so far, 30 of the 35 chromosomes were clearly triploid, three were tetrasomic and one hexasomic [66]. Whether other strains of *L. braziliensis* show similar or different ploidy patterns, as shown for all other *Leishmania* species examined so far [64,66,67], remains to be established. More than two alleles have been observed for only 1.7% of the microsatellite loci analysed in this study. This could possibly be due to aneuploidy [64,66], although other reasons such as mixed strains, duplication or stutter bands cannot be excluded. Sterkers at al. ([64,65] were able to show that in *L. major* chromosomal content varies not only from strain to strain but also from cell to cell creating 'mosaic aneuploidy'. This leads to high karyotypic diversity and conserved intra-strain genetic heterogeneity combined with loss of heterozygosity per cell. The total number of alleles can, however be maintained in a strain. As a consequence, DNA-based typing methods, including the microsatellite typing approach used herein, cannot decide if a cell population (or strain) consists of heterozygous cells or of homozygous cells presenting different allelic and ploidy content [65].

In conclusion, this study showed the extensive microsatellite diversity present in the subgenus *L. (Viannia)* and indicated that *L. (V.) braziliensis* and, to a lesser extent, *L. (V.) guyanensis* exhibit features indicative of both clonality and recombination. Recombination could explain the tremendous genetic diversity and limited population structure. The genetic heterogeneity of Brazilian strains of different species of the subgenus *L. (Viannia)* was found to be higher than previously shown by techniques such as isoenzyme typing and PCR-RFLP approaches. The different clustering approaches used in this study identified two different genetic groups or populations, one consisting of *L. (V.) guyanensis* strains from the Amazon region and the other of *L. (V.) braziliensis* strains from the southeast of Brazil, clearly differentiated from the other investigated strains. *L. (V.) braziliensis* strains from the north of Brazil did not group with those from the Atlantic coast but were found to be very polymorphic. These strains seemed to be more closely related to the strains of *L. (V.) shawi*, *L. (V.) naiffi*, and *L. (V.) lainsoni* also isolated in northern Brazilian CL foci. All findings concerning the strains from northern Brazil may, however, be subject to bias due to an inadequate sampling strategy. More strains need to be sampled from this region in order to fine tune the population structure of these parasites and their mode of reproduction.

Supporting Information

Figure S1 Rectangular NJ tree showing the populations and subpopulations of 120 Brazilian *L.* (*Viannia*) strains. A midpoint rooted Neighbour-joining (NJ) tree (rectangular version) was calculated for the MLMT profiles of 120 strains of different species of the subgenus *L.* (*Viannia*), based on 15 microsatellite markers and using the Chord distance measure. The assignment of these strains to three main populations by the Bayesian model-based clustering approach implemented in STRUCTURE is indicated by colored branches: population 1 (green), population 2 (red) and population 3 (blue). Strains belonging to these populations are listed in Table S1. Population 1 comprises all but one strain of *L. (V.) guyanensis* analysed in this study. Population 2 consists of 43 strains of *L. (V.) braziliensis* mainly from east Brazil. Population 3 is very diverse and includes all investigated strains of *L. (V.) lainsoni*, *L. (V.) naiffi*, *L. (V.) shawi*, *L. (V.) utingensis*, *L. (V.) lindenbergi*, 20 strains of *L. (V.) braziliensis* mainly from the north of Brazil as well as one strain of *L. (V.) guyanensis* from Acre. Putative hybrids are indicated by red or blue circles, according to their population assignment. Strain origins are indicated in the window alongside.
(TIF)

Figure S2 Population structure of the 120 strains inferred by Bayesian analysis with STRUCTURE. The Bayesian algorithm assigned the 120 Brazilian strains of subgenus *L. (Viannia)* to three populations. Population 1 (green) comprises all but one strain of *L.*

(V.) guyanensis analysed in this study. Population 2 (red) consists of 43 strains of *L. (V.) braziliensis* mainly from eastern Brazil. Population 3 (blue) is very diverse and includes all investigated strains of *L. (V.) lainsoni*, *L. (V.) naiffi*, *L. (V.) shawi*, *L. (V.) utingensis*, *L. (V.) lindenbergi*, 20 strains of *L. (V.) braziliensis* mainly from the north of Brazil as well as one strain of *L. (V.) guyanensis* from Acre. Four sub-populations are distinguished in Population 3 when STRUCTURE was re-run separately for the strains of this population. Sub-population 3A comprises all strains of *L. (V.) shawi*, eight strains of *L. (V.) braziliensis* (6 from Acre, one from Pará and one from Rondonia), the single strain of *L. (V.) guyanensis* from Acre, *L. (V.) utingensis* and *L. (V.) lindenbergi*; 3B ten strains of *L. (V.) braziliensis* (3 from Pernambuco, 3 from Pará, 2 from Ceará and single strains from Paraná and Amazonas); 3C all strains of *L. (V.) lainsoni*; and 3D all strains of *L. (V.) naiffi* and two putative *L. (V.) braziliensis*/*L. (V.) naiffi* hybrids.
(TIF)

Figure S3 NeighborNet network of Population 1 as inferred by STRUCTURE. Six strains were isolated from animal hosts, two from *Choloepus didactylus* and four from *Didelphis marsupialis*, and one from a sand fly, *Lutzomyia anduzei*, all other strains were isolated from human CL cases.
(TIF)

Figure S4 NeighborNet network of Population 2 as inferred by STRUCTURE. Four strains were isolated from animal hosts, two from dogs and one each from *Nectomys sp.* and *Mesocricetus auratus*, three from human MCL and three from human DL cases, all other strains were isolated from human CL cases. The assignment of the strains to the two sub-populations of Population 2, A and B, is indicated. Strains presenting mixed membership coefficients in two sub-populations are highlighted in grey. The two sub-populations are largely confirmed by the phylogenetic network albeit some strains occur at intermediate positions.
(TIF)

Figure S5 NeighborNet network of Population 3 as inferred by STRUCTURE. Eight strains were isolated from animal hosts, three from *Dasypus sp.*, two from *Cuniculus paca* and one each from *Coendou sp.*, *Cebus apella* and *Rattus rattus*, five from sand flies, three from *Lutzomyia whitmani* and one each from *L. tuberculata, L. squamiventris* and *Lutzomyia sp.*, all other strains were isolated from human CL cases. The four sub-populations of Population 3, A, B, C and D, are indicated. The phylogenetic network confirms the assignment of strains of *L. (V.) lainsoni* to sub-population C and that of strains of *L. (V.) naiffi* to sub-population D. The sub-populations A and B are not well supported in this NeighborNet network. Strains of *L.* (V.) *shawi* were found on a separate branch together with one strain of *L. (V.) braziliensis* from Pará, L-0326, and the single strain of *L. (V.) guyanensis* from Acre, L-2493.
(TIF)

Figure S6 Calculation of inbreeding coefficients for Populations 1 and 2. F_{is} values were calculated for each of the 15 loci and over all loci. Strains with identical genotypes were excluded from the analyses to avoid medical driven sampling bias and clones over-representation. For each locus, 95% confidence intervals (CI) were obtained by bootstrapping over loci (GENETIX). *P* values are indicated by stars ($P<0.05$).
(PDF)

Table S1 Designation, characteristics and MLMT profiles of the Brazilian strains of subgenus *Leishmania (Viannia)* used in this study. [T] – reference strain of the species; [1] – zymodemes according to the CLIOC system – IOC/Z [11]; [2] – population assignment according to STRUCTURE analysis; [3] – normalization of microsatellite fragment sizes in described as Materials and Methods. VL – visceral leishmaniasis; CL – cutaneous leishmaniasis; MCL – mucocutaneous leishmaniasis; DL – disseminated cutaneous leishmaniasis; CLIOC - Coleção de Leishmania do Instituto Oswaldo Cruz (IOC-L-code); BH - Coleção de Universidade Federal de Minas Gerais, Belo Horizonte, Brazil.
(XLSX)

Table S2 Distribution of zymodemes and animal hosts for the populations and sub-populations found by MLMT.
(DOCX)

Acknowledgments

We want to thank our reviewers for their critical comments which helped us to improve the manuscript significantly.

Author Contributions

Conceived and designed the experiments: GS KK EC MM IM TW MNM. Performed the experiments: KK CS SOS MCB. Analyzed the data: KK CS TW SOS EC MCB GS. Contributed reagents/materials/analysis tools: EC MNM IM TW GS. Wrote the paper: GS KK MM EC TW.

References

1. Lainson R, Shaw, J.J. (2005) New world leishmaniases. In: FEGCDWSHGDD D, editor. Topley & Wilson's microbiology and microbial infections. 10th ed. London: Wiley & Blackwell. pp. 313–349.
2. Lessa MM, Lessa HA, Castro TW, Oliveira A, Scherifer A, et al. (2007) Mucosal leishmaniasis: epidemiological and clinical aspects. Braz J Otorhinolaryngol 73: 843–847.
3. Grimaldi G, Jr., Tesh RB (1993) Leishmaniases of the New World: current concepts and implications for future research. Clin Microbiol Rev 6: 230–250.
4. Schonian G, Mauricio I, Cupolillo E (2010) Is it time to revise the nomenclature of *Leishmania*? Trends Parasitol 26: 466–469.
5. Fraga J, Montalvo AM, De Doncker S, Dujardin JC, Van der Auwera G (2010) Phylogeny of *Leishmania* species based on the heat-shock protein 70 gene. Infect Genet Evol 10: 238–245.
6. Bedoya-Pacheco SJ, Araujo-Melo MH, Valete-Rosalino CM, Pimentel MI, Conceicao-Silva F, et al. (2011) Endemic tegumentary leishmaniasis in Brazil: correlation between level of endemicity and number of cases of mucosal disease. Am J Trop Med Hyg 84: 901–905.
7. Amato VS, Tuon FF, Siqueira AM, Nicodemo AC, Neto VA (2007) Treatment of mucosal leishmaniasis in Latin America: systematic review. Am J Trop Med Hyg 77: 266–274.
8. Guerra JA, Prestes SR, Silveira H, Coelho LI, Gama P, et al. (2011) Mucosal Leishmaniasis caused by *Leishmania* (*Viannia*) *braziliensis* and *Leishmania* (*Viannia*) *guyanensis* in the Brazilian Amazon. PLoS Negl Trop Dis 5: e980.
9. Camara Coelho LI, Paes M, Guerra JA, Barbosa MG, Coelho C, et al. (2011) Characterization of *Leishmania* spp. causing cutaneous leishmaniasis in Manaus, Amazonas, Brazil. Parasitol Res 108: 671–677.
10. Rangel EF, Lainson R (2009) Proven and putative vectors of American cutaneous leishmaniasis in Brazil: aspects of their biology and vectorial competence. Mem Inst Oswaldo Cruz 104: 937–954.
11. Cupolillo E, Brahim LR, Toaldo CB, de Oliveira-Neto MP, de Brito ME, et al. (2003) Genetic polymorphism and molecular epidemiology of *Leishmania* (*Viannia*) *braziliensis* from different hosts and geographic areas in Brazil. J Clin Microbiol 41: 3126–3132.
12. Cupolillo E, Grimaldi G, Jr., Momen H (1994) A general classification of New World *Leishmania* using numerical zymotaxonomy. Am J Trop Med Hyg 50: 296–311.
13. Cupolillo E, Grimaldi Junior G, Momen H, Beverley SM (1995) Intergenic region typing (IRT): a rapid molecular approach to the characterization and evolution of *Leishmania*. Mol Biochem Parasitol 73: 145–155.
14. Schonian G, Kuhls K, Mauricio IL (2011) Molecular approaches for a better understanding of the epidemiology and population genetics of *Leishmania*. Parasitology 138: 405–425.
15. Russell R, Iribar MP, Lambson B, Brewster S, Blackwell JM, et al. (1999) Intra and inter-specific microsatellite variation in the *Leishmania* subgenus *Viannia*. Mol Biochem Parasitol 103: 71–77.
16. Fakhar M, Motazedian MH, Daly D, Lowe CD, Kemp SJ, et al. (2008) An integrated pipeline for the development of novel panels of mapped microsatellite

markers for *Leishmania donovani* complex, *Leishmania braziliensis* and *Leishmania major*. Parasitology 135: 567–574.
17. Oddone R, Schweynoch C, Schonian G, de Sousa Cdos S, Cupolillo E, et al. (2009) Development of a multilocus microsatellite typing approach for discriminating strains of *Leishmania* (*Viannia*) species. J Clin Microbiol 47: 2818–2825.
18. Rougeron V, Waleckx E., Hide M., De Meeûs, T. Arevalo, J., Llanos-Cuentas, A., Banuls, A.L. (2008) A set of 12 microsatellite loci for genetic studies of *Leishmania braziliensis*. Molecular Ecology Resources 8: 351–353.
19. Rougeron V, De Meeus T, Hide M, Waleckx E, Dereure J, et al. (2010) A battery of 12 microsatellite markers for genetic analysis of the *Leishmania* (*Viannia*) *guyanensis* complex. Parasitology 137: 1879–1884.
20. Schonian G, Schweynoch C, Zlateva K, Oskam L, Kroon N, et al. (1996) Identification and determination of the relationships of species and strains within the genus *Leishmania* using single primers in the polymerase chain reaction. Mol Biochem Parasitol 77: 19–29.
21. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155: 945–959.
22. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14: 2611–2620.
23. Dawson KJ, Belkhir K (2001) A Bayesian approach to the identification of panmictic populations and the assignment of individuals. Genet Res 78: 59–77.
24. Cavalli-Sforza LL, Edwards AW (1967) Phylogenetic analysis. Models and estimation procedures. Am J Hum Genet 19: 233–257.
25. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24: 1596–1599.
26. Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 23: 254–267.
27. Goudet J, Perrin N, Waser P (2002) Tests for sex-biased dispersal using bi-parentally inherited genetic markers. Mol Ecol 11: 1103–1114.
28. Weir BS, Cockerman CC (1984) Estimating F statistics for the analysis of population structure. Evolution 38: 1358–1370.
29. Wright S (1978) Evolution and the Genetics of Populations. Variability within and among Natural Populations. Chicago: University of Chicago Press.
30. Cockerham CC, Weir BS (1984) Covariances of relatives stemming from a population undergoing mixed self and random mating. Biometrics 40: 157–164.
31. Dieringer D, Schlötterer(2003) Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite sets. Molec Ecol Notes 3: 167–169.
32. Excoffier L. LG, Schneider S. (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1: 47–50.
33. Rice TK, Schork NJ, Rao DC (2008) Methods for handling multiple testing. Adv Genet 60: 293–308.
34. Weir BS (1990) Genetic data analysis. Sunderland, MA: Sinnauer Associates.
35. Zaykin D, Zhivotovsky L, Weir BS (1995) Exact tests for association between alleles at arbitrary numbers of loci. Genetica 96: 169–178.
36. Smith JM, Smith NH, O'Rourke M, Spratt BG (1993) How clonal are bacteria? Proc Natl Acad Sci U S A 90: 4384–4388.
37. Agapow PM, Burt A (2001) Indices of multilocus linkage disequilibrium. Molecular Ecology Notes 1: 101–102.
38. da Silva LA, de Sousa Cdos S, da Graca GC, Porrozzi R, Cupolillo E (2010) Sequence analysis and PCR-RFLP profiling of the hsp70 gene as a valuable tool for identifying *Leishmania* species associated with human leishmaniasis in Brazil. Infect Genet Evol 10: 77–83.
39. Haubold B, Travisano M, Rainey PB, Hudson RR (1998) Detecting linkage disequilibrium in bacterial populations. Genetics 150: 1341–1348.
40. Rougeron V, De Meeus T, Hide M, Waleckx E, Bermudez H, et al. (2009) Extreme inbreeding in *Leishmania braziliensis*. Proc Natl Acad Sci U S A 106: 10224–10229.
41. Rougeron V, Banuls AL, Carme B, Simon S, Couppie P, et al. (2011) Reproductive strategies and population structure in Leishmania: substantial amount of sex in *Leishmania Viannia guyanensis*. Mol Ecol 20: 3116–3127.
42. Kuhls K, Alam MZ, Cupolillo E, Ferreira GE, Mauricio IL, et al. (2011) Comparative microsatellite typing of New World *Leishmania infantum* reveals low heterogeneity among populations and its recent old world origin. PLoS Negl Trop Dis 5: e1155.
43. Odiwuor S, Veland N, Maes I, Arevalo J, Dujardin JC, et al. (2012) Evolution of the *Leishmania braziliensis* species complex from amplified fragment length polymorphisms, and clinical implications. Infect Genet Evol ;12(8):1994–2002.
44. Bryant D, Moulton V (2004) Neighbor-net: an agglomerative method for the construction of phylogenetic networks. Mol Biol Evol 21: 255–265.
45. Boite MC, Mauricio IL, Miles MA, Cupolillo E (2012) New insights on taxonomy, phylogeny and population genetics of *Leishmania* (*Viannia*) parasites based on multilocus sequence analysis. PLoS Negl Trop Dis 6: e1888.
46. Brito ME, Andrade MS, Mendonca MG, Silva CJ, Almeida EL, et al. (2009) Species diversity of *Leishmania* (*Viannia*) parasites circulating in an endemic area for cutaneous leishmaniasis located in the Atlantic rainforest region of northeastern Brazil. Trop Med Int Health 14: 1278–1286.
47. Romero GA, Ishikawa E, Cupolillo E, Toaldo CB, Guerra MV, et al. (2002) Identification of antigenically distinct populations of *Leishmania* (*Viannia*) *guyanensis* from Manaus, Brazil, using monoclonal antibodies. Acta Trop 82: 25–29.
48. Rotureau B, Ravel C, Nacher M, Couppie P, Curtet I, et al. (2006) Molecular epidemiology of *Leishmania* (*Viannia*) *guyanensis* in French Guiana. J Clin Microbiol 44: 468–473.
49. Barral A, Teixeira M, Reis P, Vinhas V, Costa J, et al. (1995) Transforming growth factor-beta in human cutaneous leishmaniasis. Am J Pathol 147: 947–954.
50. Blackwell JM (1999) Tumour necrosis factor alpha and mucocutaneous leishmaniasis. Parasitol Today 15: 73–75.
51. Castellucci L, Menezes E, Oliveira J, Magalhaes A, Guimaraes LH, et al. (2006) IL6 -174 G/C promoter polymorphism influences susceptibility to mucosal but not localized cutaneous leishmaniasis in Brazil. J Infect Dis 194: 519–527.
52. de Moura TR, Oliveira F, Novais FO, Miranda JC, Clarencio J, et al. (2007) Enhanced *Leishmania braziliensis* infection following pre-exposure to sandfly saliva. PLoS Negl Trop Dis 1: e84.
53. Ferreira GE, dos Santos BN, Dorval ME, Ramos TP, Porrozzi R, et al. (2012) The genetic structure of *Leishmania infantum* populations in Brazil and its possible association with the transmission cycle of visceral leishmaniasis. PLoS One 7: e36242.
54. Akopyants NS, Kimblin N, Secundino N, Patrick R, Peters N, et al. (2009) Demonstration of genetic exchange during cyclical development of *Leishmania* in the sand fly vector. Science 324: 265–268.
55. Sadlova J, Yeo M, Seblova V, Lewis MD, Mauricio I, et al. (2011) Visualisation of *Leishmania donovani* fluorescent hybrids during early stage development in the sand fly vector. PLoS One 6: e19851.
56. Tibayrenc M, Ayala FJ (2012) Reproductive clonality of pathogens: a perspective on pathogenic viruses, bacteria, fungi, and parasitic protozoa. Proc Natl Acad Sci U S A 109: E3305–3313.
57. Banuls AL, Hide M, Prugnolle F (2007) *Leishmania* and the leishmaniases: a parasite genetic update and advances in taxonomy, epidemiology and pathogenicity in humans. Adv Parasitol 64: 1–109.
58. Nolder D, Roncal N, Davies CR, Llanos-Cuentas A, Miles MA (2007) Multiple hybrid genotypes of *Leishmania* (*Viannia*) in a focus of mucocutaneous leishmaniasis. Am J Trop Med Hyg 76: 573–578.
59. Bonfante-Garrido R, Melendez E, Barroeta S, de Alejos MA, Momen H, et al. (1992) Cutaneous leishmaniasis in western Venezuela caused by infection with *Leishmania venezuelensis* and *L. braziliensis* variants. Trans R Soc Trop Med Hyg 86: 141–148.
60. Delgado O, Cupolillo E, Bonfante-Garrido R, Silva S, Belfort E, et al. (1997) Cutaneous leishmaniasis in Venezuela caused by infection with a new hybrid between *Leishmania* (*Viannia*) *braziliensis* and *L.* (*V.*) *guyanensis*. Mem Inst Oswaldo Cruz 92: 581–582.
61. Banuls AL, Guerrini F, Le Pont F, Barrera C, Espinel I, et al. (1997) Evidence for hybridization by multilocus enzyme electrophoresis and random amplified polymorphic DNA between *Leishmania braziliensis* and *Leishmania panamensis/guyanensis* in Ecuador. J Eukaryot Microbiol 44: 408–411.
62. Tojal da Silva AC, Cupolillo E, Volpini AC, Almeida R, Romero GA (2006) Species diversity causing human cutaneous leishmaniasis in Rio Branco, state of Acre, Brazil. Trop Med Int Health 11: 1388–1398.
63. Ramirez JD, Guhl F, Messenger LA, Lewis MD, Montilla M, et al. (2012) Contemporary cryptic sexuality in *Trypanosoma cruzi*. Mol Ecol 21: 4216–4226.
64. Sterkers Y, Lachaud L, Crobu L, Bastien P, Pages M (2011) FISH analysis reveals aneuploidy and continual generation of chromosomal mosaicism in *Leishmania major*. Cell Microbiol 13: 274–283.
65. Sterkers Y, Lachaud L, Bourgeois N, Crobu L, Bastien P, et al. (2012) Novel insights into genome plasticity in Eukaryotes: mosaic aneuploidy in *Leishmania*. Mol Microbiol 86: 15–23.
66. Rogers MB, Hilley JD, Dickens NJ, Wilkes J, Bates PA, et al. (2011) Chromosome and gene copy number variation allow major structural change between species and strains of *Leishmania*. Genome Res 21: 2129–2142.
67. Downing T, Imamura H, Decuypere S, Clark TG, Coombs GH, et al. (2011) Whole genome sequencing of multiple *Leishmania donovani* clinical isolates provides insights into population structure and mechanisms of drug resistance. Genome Res 21: 2143–2156.
68. Mannaert A, Downing T, Imamura H, Dujardin JC (2012) Adaptive mechanisms in pathogens: universal aneuploidy in *Leishmania*. Trends Parasitol 28: 370–376.

4.4. Capítulo IV

Polymorphisms and ambiguous sites present in DNA sequences of *Leishmania* clones: looking closer.

Boité MC, Oliveira T, Ferreira GEM, Porrozzi R, Trannin MA, Santos BN, Cupolillo E
Infection, Genetics and Evolution
2014, doi: 10.1016/j.meegid.2014.04.011. Epub 2014 Apr
Artigo publicado

Neste artigo os sítios ambíguos encontrados nas sequencias de DNA obtidas durante a execução do primeiro trabalho (Capítulo I) foram avaliados à luz de experimentos adicionais. A metodologia incluiu a clonagem das cepas contendo os sítios ambíguos para obtenção de culturas monoclonais. A partir desse material, DNA foi isolado e novas sequencias obtidas diretamente dos produtos da PCR e também após a clonagem gênica destes produtos. Desta forma foi possível comparar sequencias de DNA obtidas de uma mesma cepa não clonada, mantida em cultura, com suas culturas monoclonais associadas ou não à clonagem genica. O resultado demonstrou que a diversidade genética deste parasita esta presente dentro de um mesmo isolado e, portanto não é completamente retratada a partir de análises convencionais baseadas em sequenciamento de DNA. Os achados podem refletir a plasticidade genética descrita para *Leishmania* e tais aspectos devem ser considerados durantes estudos baseados no sequenciamento de DNA.

Infection, Genetics and Evolution 25 (2014) 110–116

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Polymorphisms and ambiguous sites present in DNA sequences of *Leishmania* clones: Looking closer

Mariana Côrtes Boité *, Taíse Salgado de Oliveira, Gabriel Eduardo Melim Ferreira, Marcos Trannin, Barbara Neves dos Santos, Renato Porrozzi, Elisa Cupolillo

Laboratório de Pesquisa em Leishmaniose, Instituto Oswaldo Cruz (IOC), Fundação Oswaldo Cruz (FIOCRUZ), Rio de Janeiro, Brazil

ARTICLE INFO

Article history:
Received 26 August 2013
Received in revised form 9 April 2014
Accepted 13 April 2014
Available online 21 April 2014

Keywords:
Leishmania
Polyclonal
Aneuploidy
Heterozygosis

ABSTRACT

In genetic studies of *Leishmania* parasites, co-dominant markers are chosen for their ability to detect heterozygous polymorphisms, to infer the occurrence of inbreeding and to resolve genetic variability. The majority of DNA sequence based reports perform conventional dye terminator cycle sequencing where perfectly ambiguous sites or double peaks in the chromatogram are interpreted as heterozygous strains. However, molecular peculiarities of the parasite such as aneuploidy, mixed populations and homologous recombination advise that data from regular DNA sequence analysis should be carefully evaluated. We report here a closer look at ambiguous sites observed in *6pgd* DNA sequences obtained for a multilocus sequence analysis project on *Leishmania* (*Viannia*) strains. After comparing 286 DNA sequences from biological and molecular clones of six *L.* (*Viannia*) strains we could distinguish events that contribute to genetic variation in *Leishmania* (recombination, mutation, chromosomal mosaics). Also, the results suggest how diversity might not be completely revealed through regular DNA sequence analysis and demonstrate the importance for molecular epidemiology research to be aware of such possibilities while choosing samples for studies.

1. Introduction

The *Leishmania* parasite is the causative agent of the disease leishmaniasis, which affects millions of people annually and represents the third highest global burden of disease among the neglected tropical diseases. Despite observed morphological homogeneity across most specimens, more than 20 genetically distinct species have been described for the *Leishmania* genus. Phenotypic diversity can be observed between species, even among clones (Garin et al., 2001), which may be a causative agent in the presentation of different clinical forms of the disease and the variation in responses to conventional drug treatment. Therefore, ability to identify *Leishmania* species and to characterize genetic variants could increase physicians' ability to target treatment and therapy. With that in mind, many approaches have been developed, improved and applied aiming to contribute to population genetics, parasite identification and phylogeny. Among these approaches are multilocus enzyme electrophoresis (MLEE) (Cupolillo et al., 1994) and molecular markers based on DNA sequence analysis, such as multilocus sequence analysis (MLSA) (Zemanova et al., 2007; El Baidouri et al., 2013; Boité et al., 2012) and multi locus microsatellite typing (Alam et al., 2009; Oddone et al., 2009; Rougeron et al., 2010).

Traditionally, co-dominant markers used for these genetic based analyses are chosen by their ability to detect heterozygous polymorphisms, infer the presence of inbreeding, and determine genetic variability. The majority of DNA sequence based protocols utilize conventional dye terminator cycle sequencing, which assume standard procedures for interpreting inconsistencies in results, such as double vs single peaks. In MLSA sequences, double peaks in the chromatogram are normally considered to result from heterozygous strains. However, four major aspects of *Leishmania* biology raise concern over the interpretation of such DNA sequence based analysis. First, *Leishmania* presents mosaic aneuploidy in vitro (Sterkers et al., 2010; Rogers et al., 2011), which has also been detected in clinical isolates (Rogers et al., 2011). Mosaic aneuploidy creates cells with different karyotype contents and may result in overlapping the reads of different nucleotides, actually present in different cells, leading to misinterpretation or loss of information in the sequence analysis. Next, the occurrence of polyclonality and infrapopulation mixture in *Leishmania* isolates could cause the observation of more than one sequence in the

* Corresponding author. Address: Av Brasil, 4365, Pavilhão Leonidas Daene sala 509, Manguinhos, Rio de Janeiro, RJ, Brazil. Tel.: +55 21 38658226.
E-mail address: boitemc@ioc.fiocruz.br (M.C. Boité).

http://dx.doi.org/10.1016/j.meegid.2014.04.011
1567-1348/

chromatogram. This can occur simply by natural mixing or can be created by karyotype diversity (Dujardin et al., 2007) generated after asymmetric chromosomal allotment (ACA) (Sterkers et al., 2012). Third, clonal propagation is considered the predominant mode of reproduction of this parasite. This phenomenon may overlook cryptic sexuality and could explain the presence of the double peak rather than being representative of the clone. Finally, the clonal concept in *Leishmania* is more plausible from a biological point of view rather than a genetic one (Tibayrenc et al., 1990; Tibayrenc and Ayala, 2012).

Given that misinterpretation of sequencing results may have an important impact on all *Leishmania* typing and characterization studies, we investigated ambiguous sites represented by double peaks in *6pgd* DNA sequences obtained by MLSA for *Leishmania* (*Viannia*) strains (Boité et al., 2012) to observe how the above described parasite characteristics may interfere in DNA sequences generated for genetic variability analysis.

2. Material and methods

2.1. Strains, biological cloning and MLEE characterization

Six *L.* (*Viannia*) strains (Table 1) presenting double peaks in previously obtained *6pgd* DNA sequences (Boité et al., 2012) were retrieved from the Coleção de *Leishmania* do Instituto Oswaldo Cruz (CLIOC – WDCM 731). More details about the samples can be obtained by accessing the CLIOC catalogue (http://clioc.fiocruz.br). The parasites were cultivated in Schneider's medium and the culture volume divided for both procedures: biological cloning and direct DNA isolation. For biological cloning 10^4 parasites in 600 µl of Schneider's medium were added to 2.4 ml of Low Melting Point Agarose (Sigma) 1% warmed solution. The mixture was distributed over a Petri Plate containing NNN medium. After 7 days, the colonies were harvested and cultured separately in tubes containing biphasic medium (NNN + Schneider's) and processed for DNA isolation and MLEE assay. Enzymatic MLEE systems were assayed for the characterization of biological clones as previously described (Cupolillo et al., 1994).

2.2. Molecular procedures and analysis

DNA purification, PCRs for the *6pgd* locus and PCR product purification was carried out as described by Boité et al. (2012). Sequences were obtained from two parasite states: non cloned culture (NC) and culture from biological clone (BC) as pointed in Fig. 1; and through two distinct approaches: PCR products subjected to direct sequencing (DS) and to molecular cloning followed by sequencing (MC) (Fig. 1A). These conditions combined generated the following DNA sequences: non-cloned directly sequenced (NCDS), biological clones directly sequenced (BCDS) and biological clones subjected to molecular cloning (BC_MC) (Fig. 1B).

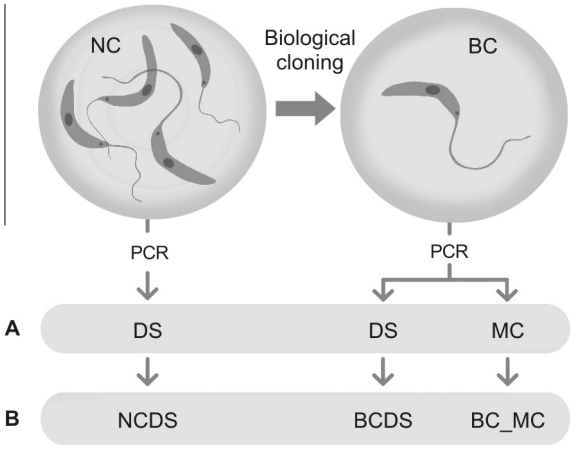

Fig. 1. Procedures to obtain the DNA sequences compared in the present study. Cultures used for DNA isolation: non-cloned (NC) and after biological cloning (BC). A = PCR products obtained from each culture either subjected to direct sequencing (DS) or molecular cloning (MC). B = Sequences obtained from NC and BC after MC or DS: the "non-cloned directly sequenced" (NCDS) represents a sequence obtained from the PCR product of DNA isolated from a non-cloned population of cells; the "biological clone directly sequenced" (BCDS) represents a sequence obtained from the PCR product of DNA isolated from cloned cultures; the "biological clone_molecular clone" (BC_MC) represents a sequence obtained after molecular cloning of the PCR product of DNA isolated from cloned cultures.

Purified PCR products from BC were molecularly cloned using pGEM®-T Easy Vectors System (Promega) and sequences obtained in an ABI 3730 DNA Sequencer by the PDTIS (FIOCRUZ, Rio de Janeiro, Brazil) sequencing service (Otto et al., 2008). Consensus sequences were generated and edited in Phred/Phrap/Consed Version: 0.020425.c from forward and reverse strands. Only sequence segments with Phred values above twenty were used for contig construction. The chromatogram was manually tracked for double peaks (equivalent presence of nucleotides, suggesting a 1:1 proportion of alleles) in at least one strand, and the respective sites coded using the standard IUPAC for combinations of two or more bases.

Contigs from all samples were manually assembled and aligned in MEGA5 software (Tamura et al., 2011). All new sequences obtained were compared with those from Boité et al. (2012) named here as "previous NCDS".

Recombination analysis was performed in RDP (Martin et al., 2010), using different recombination detection algorithms, with other *L.* (*Viannia*) sequences from Boité et al. (2012) and sequences available in GenBank®.

3. Results and discussion

Results demonstrated that *Leishmania* parasite biology may play a significant role in the outcome of sequence analysis. We found

Table 1
Leishmania (*Viannia*) strains analyzed for a fragment of *6pgd* DNA sequence of 716 bp, including species identification as determined by MLEE, number of biological clones and molecular clones, different alleles (sequence types) obtained and type of event observed.

Species by MLEE	IOC/L	Total of DNA sequences compared*	BCDS	MC	Sequence types**	Events observed
L. braziliensis	918	50	8	40	10	Hot spot sites
L. braziliensis	2538	52	9	41	9	Homologous recombination
L. braziliensis	2823	134	20	112	14	Infrapopulation of nonequivalent heterozygous cells
L guyanensis	2966	7	5	0	2	ACAs; polymorphisms
L. lainsoni/L. naiffi	2490	23	4	17	3	Monoclonal heterozygous
L. guyanensis	2957	20	4	15	2	ACAs

MLEE = multilocus enzyme electrophoresis; BCDS = biological clones directly sequenced; MC = molecular clones; ACAs = asymmetric chromosomal allotments.
* Including the previous and recent NCDS.
** Sequence types were determined considering a *6pgd* gene region as described previously (Boité et al., 2012).

that mosaic aneuploidy generated by ACA occurring during parasite culture played an important part in the interpretations of results. These results have a considerable impact on past and future molecular epidemiologic studies of leishmaniasis since we can no longer assume heterozygosis as the unique cause of ambiguous sites in sequencing based methods.

3.1. Loss of ambiguous site by asymmetric chromosomal allotments (ACA)

The double peak previously detected for two strains IOC/L 2957 and 2966 in site 602 was no longer observed, and NCDS and BCDS were identical, suggesting a monoclonal population (Table 2). The loss of heterozygous cells and the alternative allele once detected can be explained by ACAs during the axenic culture (Sterkers et al., 2012). The clonal propagation of the parasites presenting the homozygous chromosomal arrangement might have leaden to the detection of only DNA sequences without the double peaks in the chromatogram. It is not possible, however, to be certain about ploidy of the cell since ACAs explain the monomorphic DNA sequences among BCDS and MC, but ACAs also generates karyotype diversity.

In addition to the double peak site, polymorphisms were noticed between the previous sequence and the new sequences (sites 640, 672, 685, Table 2). These polymorphisms could be the result of point mutations during the culture procedure. Double peak and polymorphic sites, associated with preponderant clonal evolution (PCE) (Tibayrenc and Ayala, 2013) represents one more complicating factor in *Leishmania* population structure analysis; nevertheless some strictly sexual organisms might also challenge models dependent on whether ability to prevent whole chromosome elimination is present (Benatti et al., 2010).

3.2. Less frequent alleles can only be detected after molecular cloning in conventional dye terminator sequencing methods

The previously observed double peak for the strain IOC/L 2823 represented an equivalent proportion of nucleotides (1:1) but

Table 2
Sequence polymorphisms observed in a 6*pgd* gene region for non-cloned culture, and for biological and molecular clones of six *Leishmania* strains.

IOC	Sequence origin (origin)$_{number\ of\ sequences}$	Site position																												
		442	445	464	485	501	502	606	689	707	719	732	819	829	858	891	900	944	960	966	968	1038	1052	1054	1057	1602	1630	1640	1666	1672	1685
L 2966	Previous NCDS																									**R**		A		G	G
	(NCDS)$_1$;(BCDS)$_5$																									A		G		C	T
L 2957	Previous NCDS																									**R**					
	(NCDS)$_1$; (BCDS)$_4$; (MC)$_{17}$																									A					
L 2823	Previous NCDS		T		C	A		C	T	C	A	A	T	T	T		T		Y		A	A	A	A		A		A		
	(NCDS)$_1$; (BCDS)$_{17}$; (BC_MC)$_{90}$		.		.	.		.	.	.	.	.	.	.	.		.		T		.	.	.	.		.		.		
	(BCDS)$_3$; (BC_MC)$_1$		C		.	.		.	.	.	.	.	.	.	.		.		T		.	.	.	.		.		.		
	(BC_MC)$_2$		.		.	.		.	.	.	G	.	.	.	.		.		T		.	.	.	.		.		.		
	(BC_MC)$_2$		.		.	.		T	.	.	.	.	.	.	.		.		T		.	G	.	.		.		.		
	(BC_MC)$_2$		.		.	.		.	.	.	.	.	.	.	.		.		T		.	.	.	G		.		.		
	(BC_MC)$_3$		.		T	.		.	.	.	.	.	.	.	C		A		T		.	.	.	.		G		.		
	(BC_MC)$_3$		.		.	.		.	.	.	.	G	.	.	.		.		T		.	.	.	.		.		.		
	(BC_MC)$_3$		.		.	.		.	.	.	.	.	C	C	.		.		T		.	.	.	.		.		.		
	(BC_MC)$_1$		.		.	.		.	.	.	.	.	.	.	.		.		C		.	.	.	.		.		.		
	(BC_MC)$_2$		.		.	.		.	C	T	.	.	.	.	.		.		T		.	.	.	.		.		.		
	(BC_MC)$_2$		.		.	T		.	.	.	.	.	.	.	.		.		T		.	.	.	.		.		G		
	(BC_MC)$_1$		.		.	.		.	.	.	.	.	.	.	.		.		T		G	.	G	.		.		.		
	(BC_MC)$_1$		.		.	.		.	.	.	.	.	.	.	.		.		T		G	.	.	.		.		.		
L 2538	Previous NCDS; (BCDS)$_3$			R												S			Y	Y										
	(NCDS)$_1$; (BCDS)$_4$			A												.			.	.										
	(BCDS)$_2$			.												.			.	C										
	(BC_MC)$_{14}$			G												C			C	C										
	(BC_MC)$_5$			G												G			T	T										
	(BC_MC)$_{13}$			A												G			T	T										
	(BC_MC)$_7$			A												C			T	C										
	(BC_MC)$_2$			A												C			C	C										
	(BC_MC)$_1$			A												C			T	T										
L 2490	Previous NCDS; (NCDS)$_1$; (BCDS)$_4$						R											K												
	(BC_MC)$_9$						A											G												
	(BC_MC)$_1$						G											G												
	(BC_MC)$_7$						G											T												
L 918	Previous NCDS; (NCDS)$_1$	T		G												S			Y	Y										
	(BCDS)$_3$	.		A												.			.	.										
	(BCDS)$_1$	C		A												.			C	.										
	(BCDS)$_1$	C		.												.			C	.										
	(BCDS)$_1$	.		C												.			.	.										
	(BCDS)$_2$	.		.												.			C	.										
	(BC_MC)$_{17}$	.		A												G			T	T										
	(BC_MC)$_5$	.		A												C			C	C										
	(BC_MC)$_{17}$	.		.												C			C	C										
	(BC_MC)$_1$	.		.												G			T	T										

"Previous NCDS" = sequence from Boité et al. (2012); NCDS = non cloned directly sequenced; BCDS = biological clone directly sequenced; BC_MC = molecular clones of a biological clone; R = A (adenine) or G (guanine); Y = C (citosine) or T (timine); S = G or C; K = G/T.

was no longer identified among BCDS (Table 2, site 460); nonetheless one MC presented the alternative nucleotide "C". From that is possible to infer a change has occurred in allele's proportion, and only biological and molecular cloning allowed both nucleotide possibilities to be detected.

The heterozygous concept applied for *Leishmania* is the 1:1 nucleotides ratio in chromatogram. However, for polyploidy organisms, it is appropriated to determine a heterozygous individual by the identification of different copies for a given target gene or locus (van Belkum et al., 2001). The present result represents exactly such situation. The detection of less frequent alleles in non-cloned culture can be performed either through molecular cloning or, very cautiously, in chromatogram (Fig. 2). For the last option, however, the approach should be first validated since the smaller peak can be misinterpreted as background. Because of this, in the present study we only considered the sequences from molecular cloning to determine different copies of the targeted allele.

Aside the ambiguous sites targeted as the main objective of study, polymorphisms were also noticed along the alignment of all sequences obtained for this strain. In site 45, for instance, three BCDS and one BC_MC differ from their correspondent MCs and BC (Table 2). This suggests that biological clones represent a heterozygous infrapopulation, and the alternative allele (T) is among the respective BC_MC, but not in equal proportion to generate overlapped peaks in BCDS or NCDS.

The polymorphisms can be explained either by: (i) point mutations, which have generated infrapopulation during the culture or (ii) initial aneuploidy, with different heterozygote cells: the sequences with the polymorphism in position 45 would represent the least frequent alleles and therefore could be detected only after biological and molecular cloning. Although mutations might occur frequently, option (ii) is more plausible because it is corroborated by the detection of MC presenting alleles different from their corresponding BCDS. Homozygous cells might also be present (after MC many presented just one allele). Nonetheless these findings demonstrate how less frequent alleles might often be undetected or excluded from DNA sequence based analysis. It also points that the removal of samples presenting multiple peaks would not avoid this bias (Camara et al., 2010).

3.3. Detection of polyclonal heterozygous sample and homologous recombination

Previous NCDS of IOC/L 2538 presented four double peaks and identical sequence types were observed, after biological cloning, among three BCDS (Table 2). However, the pattern was not kept for all other sequences: the NCDS differed in site 64 from the previous obtained sequence; the BCDS were either (i) identical to the previous sequence; (ii) identical to the NCDS or (iii) a third variation, with no double peak in the site 466 (Table 2). These results suggest at least three possible heterozygous cell populations occur within the IOC/L 2538 strain. However, when each site is observed, only positions 391 and 460 appear as classical heterozygous, with the possible nucleotides detected among the MCs. Such pattern would be expected for all sites since they are located in the same locus.

For sequence types from case (ii), the corresponding MCs presented both nucleotide possibilities in the four sites, including position 64, which had no double peak. Therefore, it might represent a heterozygous infrapopulation in the sample, containing alleles with different nucleotides only in position 64. This confirms

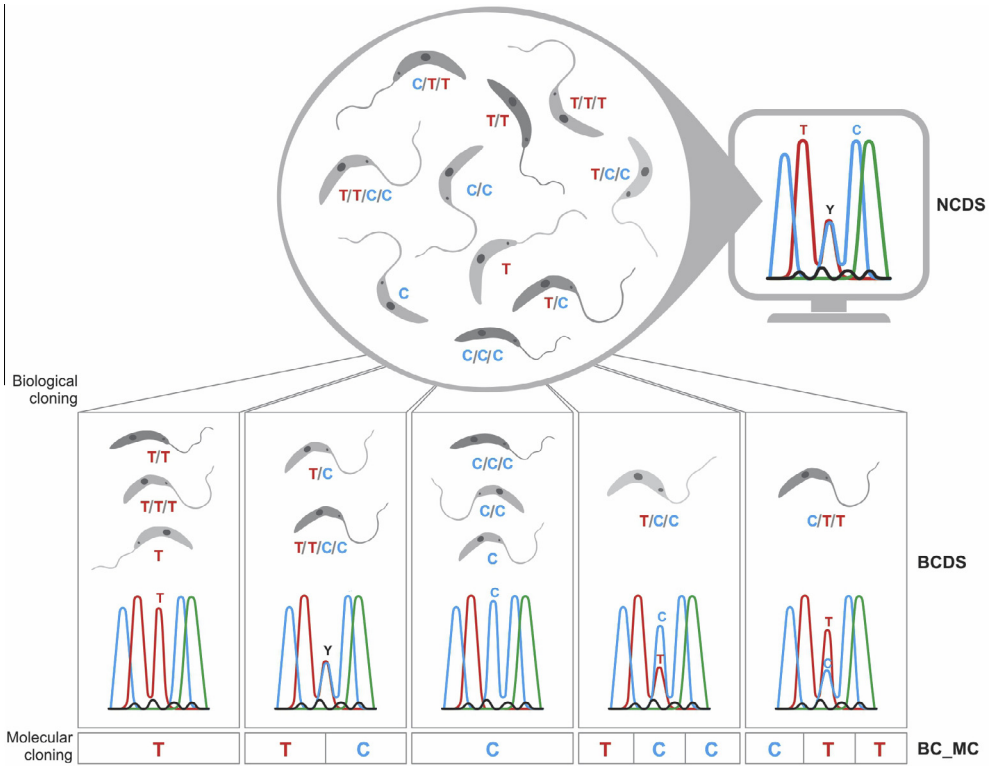

Fig. 2. Graphical representation of the possibilities of karyotype content, chromatogram profiles and polymorphisms that can be observed after biological and molecular procedures. Possibilities of karyotype content within a non-cloned *Leishmania* culture generating perfectly overlapped peaks in chromatogram profiles after direct sequencing. The overlapped peaks at the chromatogram presenting the ambiguous site (Y – IUPAC symbols: T/C) may result from different combination of heterozygous and homozygous cells. Biological cloning – karyotypes contents in cloned cultures and what can be observed at the resulted chromatogram of a biological clone directly sequenced – BCDS (from the left): homozygous cells presenting variable ploidy states; equivalent heterozygous cells; nonequivalent heterozygous cells. Molecular cloning – polymorphisms detected at the ambiguous sites after sequencing DNA isolated from biological clones and submitted to molecular cloning (BC_MC).

a third heterozygous allele within the culture observed only after BC and MC.

Upon molecular cloning in a diploid organism, the four ambiguous sites could generate at most 16 possible haplotypes. Here, a total of six different haplotypes were detected among 42 MCs, obtained from nine BCs from the same strain IOC/L 2538. This number of alleles demonstrates the intra-strain sequence type diversity. Although PCR recombination and point mutation cannot be excluded, the origin of the supernumerary heterozygous alleles could also be consequence of homologous recombination. In most organisms, the primary function of homologous recombination is to allow genome protection by the faithful repair of DNA double-strand breaks and search for sequence homology. However, *Trypanosoma brucei* and *Leishmania* use homologous recombination as a beneficial mechanism for antigenic variation (Boothroyd et al., 2009) or for drug resistance (Ubeda et al., 2008), respectively.

These results reflect the aneuploidy state and homologous recombination. As an aneuploidy organism, a gene previously described as one-copy can in fact present a variable number of copies. Moreover, even between BCs, ACAs do occur during parasite growth (mitosis). Thus, if a BC is a heterozygous cell initially, this state may not be sustained during the necessary period of culture to obtain DNA. Nevertheless, as pointed out by Sterkers et al. (2012), the alleles will still be present in the same initial proportion (the process is stable instead of fluctuating) although the number of heterozygous cells may decrease. This means if the heterozygote was not detected at first, it will be after MC (see point 3.2).

3.4. Monoclonal heterozygous strain profiling in DNA sequences cannot exclude karyotype diversity

The strain IOC/L 2490 preserved the two double peaks in NCDS and BCDS (Table 2). The NCDS and all four BCDS identical to the previous sequence suggest, therefore, a monoclonal heterozygous strain. However, the clonal concept in *Leishmania* is only suitable from a biological point of view, but not genetically (Dujardin et al., 2007) since karyotype diversity is generated from one parasite cell (Sterkers et al., 2012, 2011). Indeed, the clonal theory for *Leishmania* does not refer to any precise cytological mechanism, but instead to the genetic consequences of clonality. In the present study we cannot guarantee the cells present the same karyotype, but it is possible to guarantee the equivalent proportion of alleles through the perfect 1:1 overlapping peaks observed (Fig. 2). Considering the two polymorphic sites (Table 2) and recombination occurrence, four possible haplotypes at most could be detected for a diploid cell. Three alleles were detected within 17 MCs, suggesting more than two homologous chromosomes are present.

3.5. The double peaks and polymorphisms presented coincident sites between the strains analyzed

Strain IOC/L 918 NCDS and previously NCDS were identical, with the three double peaks still present (Table 2). For the BCDS, two sites (391, 466) kept the ambiguity. Two other sites (42, 64), not double peaks initially, were included in the analysis because, curiously, among the respective MCs both nucleotide alternatives were observed. Moreover, these sites were also double peaks for other strains (IOC/L 2494 and 2538).

Five haplotypes (Table 2) where detected among the eight BCDS. This may represents the occurrence of populations with different heterozygous karyotypes. Additionally, it suggests the polymorphic sites, which were double peaks for other strains but not for IOC/L 918, might be related to hot spots of mutation/recombination.

3.6. Recombination detection

After RDP analysis including 175 *L.* (*Viannia*) non-cloned sequences plus all 286 sequences obtained in the present study, one recombination event was detected for one *Leishmania braziliensis* strain (IOC/L 3089, from Bahia, Northeastern Brazil). The major and minor parental populations indicated were, respectively, IOC/L 3072 (*L. braziliensis*, from Rio de Janeiro), and the BCDS and NCDS (which were identical) of IOC/L 2490 *Leishmania naiffi*/*Leishmania lainsoni* – (hybrid profile in 6PGDH isoenzyme) by MaxChi and 3Seq algorithms ($P < 0.05$). The breakpoint beginning/end was wide enough to encompass almost the whole sequence, but the MaxChi graph representation indicates a window between positions 350 and 492, which includes the main polymorphic and double peaks sites evaluated in the present study.

Recombination is usually hard to detect within species, mainly because of low inter-strain diversity, or apparent low diversity due to inappropriate sampling (Prugnolle and De Meeus, 2010). However, some authors were able to present such data: population genetics studies have described inbreeding in *L. braziliensis* as well as in other *Leishmania* species (Rougeron et al., 2010; Ferreira et al., 2012). Recombination signals appear also in MLSA approaches (Boité et al., 2012) and in natural hybrids that were widely described (Akopyants et al., 2009; Belli et al., 1994; Brito et al., 2009) or even created in vitro (Coelho et al., 2012). RDP software recognizes IUPAC symbols, enhancing the ability of the algorithms to detect recombination. The possibility of PCR recombination cannot be excluded, but random PCR repetition of some sequences (data not shown) validated the polymorphisms detected. The present results indicate that recombination may occur between specific clones, and the molecular and biological cloning can contribute to the detection of a specific recombinant allele. Well-structured clonal complexes in *Leishmania* (Boité et al., 2012), could be the consequence of homologous recombination, as described in bacteria (Gonzalez-Gonzalez et al., 2013; Paul et al., 2013; Jinkerson et al., 2013) and other microorganisms (Vink et al., 2011).

4. Further comments

The main concern that has driven this study is the discussion regarding intra strain variation detected here in DNA sequence based analysis in *Leishmania* and elsewhere in *Trypanosoma cruzi* (Camara et al., 2010). The exact determination of allele frequency in individuals and, eventually, in populations is essential in association studies (Tibayrenc and Ayala, 2012, 2013); therefore, the quantitative accuracy of genotyping is critical. Multiple peaks in chromatograms, for instance, might point to: (i) different clones in the sample; (ii) heterozygotes or both. After the present results we observe that to exclude strains which present multiple peaks from the analysis is not the best way to deal with these facts since such diversity might be present even when the multiple peaks are not detected.

The occurrence of polyclonal samples is a constant if one considers the aneuploidy state described for *Leishmania*. The ACAs during mitosis seem to be quite frequent, generating daughter cells with different karyotype profiles, even after cellular cloning (infrapopulation) (Sterkers et al., 2012). Therefore, the (genetic) clonal concept is quite intricate to be applied in this parasite (Tibayrenc and Ayala, 2012). We demonstrate here that there is DNA sequence diversity within one strain kept in culture and/or cryopreserved; and changes can be detected along time, in different DNA batches. The allele frequency, however, usually does not change within the mosaic population. If multiple peaks are detected once and the two most prominent considered, that means the most frequent alleles (prevalent infrapopulation in the sample) are being distinguished.

Nevertheless, as presented in this project, some alleles will only be observed after BC and MC.

In terms of allele frequency, DNA sequence based approaches are able to reflect (partially) the biological reality, especially if there is a good sample representation in terms of diversity and geographic origin. A strain will still appear as a 'heterozygote' when global analysis methods are used because all the alleles present at first (in a real heterozygous cell) are still present in the population. This process seems to be stable, so, despite the mosaic of cells, a pattern can still be determined as representative for that population. From a population genetics point of view, if a heterozygous signal is detected, it does represent heterozygous strains at some point and reflects the two most frequent alleles in that sample. Nevertheless, the real diversity will always be underestimated (Prugnolle and De Meeus, 2010) because some alleles will remain undetected. The bias of underestimated diversity could be diminished through a proper representation of the circulating strains, such that the greatest number of different alleles would be detected. The differences found between NCDS and the previous sequences are the most disquieting findings in terms of DNA based genetic studies.

The ratio of overlapping peaks in electropherograms can vary in polyploidy genomes (Fig. 2) (Rickert et al., 2002). In tetraploid chromosomes, for instance, five allelic ratios are distinguishable: 4:0, 3:1, 1:1, 1:3, and 0:4. The present targeted *6pgd* sequences, after Genome Blast (http://www.ncbi.nlm.nih.gov/) and search in http://tritrypdb.org, were described as located in chromosome 35 (Rogers et al., 2011) and 34/20 (chromosome fusion), respectively. Either way, these chromosomes are considered non-supernumerary in *L. braziliensis*, which would mean a triploid state previously described (Rogers et al., 2011). Only perfectly double peaks were considered in the *L.* (*Viannia*) strains analyzed here suggesting a 1:1 proportion of alleles, which is an unexpected result for triploids (1:2; 3:0). If the triploid state is indeed real for *L.* (*Viannia*) strains, the results could reflect either the presence of other gene copies or the occurrence of ACAs, e.g. triploid heterozygous cells with an equal proportion of the alleles in the population (Fig. 2). Thus the strain would still appear as a "perfect monoclonal heterozygote" when sequencing methods are used.

Strains with a non-equivalent proportion for different alleles (such as 3:1 – multiple peaks) are usually analyzed as homozygous favoring the most prevalent allele. On the other hand, those strains considered heterozygous probably were either real heterozygotes at some point, or are composed by homozygous cells presenting different alleles in exactly equivalent proportions for the locus under study. If the last option occurs, it is unlikely to be due to occurrence of mixed populations; the most reasonable explanation is occurrence of cells under ACA events, descendent from a real heterozygote. Even for clones, ACAs do occur during parasite replication (mitosis). So, if a clone is a heterozygous cell initially, such a situation may not be sustained during the period of culture required to obtain DNA.

5. Conclusions

Results obtained after DNA sequence based analysis reflect events that contribute to genetic variation in *Leishmania* (recombination, mutation, chromosomal mosaics). The DNA sequences variety observed within the same strain demonstrates how diversity might not be completely represented through regular DNA sequence analysis and signal the importance for molecular epidemiology research to be aware of such possibilities while choosing the samples for studies. Moreover, to exclude samples with multiple peaks from analysis is not the best way to deal with the occurrence of infra-populations in *Leishmania* because such a phenomenon occurs frequently. Instead, tests based on diploidy, often used for *Leishmania* – an aneuploid organism, can be performed as segregation tests that explore allele distribution, thus representing a good tool for population genetics, even for loci with undetermined ploidy. The bias of underestimated diversity could be diminished through a proper representation of the circulating strains, such that the greatest number of different alleles would be detected. A deeper knowledge over the karyotypes profiles of the different strains is of major importance.

Acknowledgments

We thank the Plataforma de Sequenciamento do Instituto Oswaldo Cruz (PDTIS) for performing all the sequencing; Instituto Oswaldo Cruz – FIOCRUZ, CNPq/Fiocruz (PIBIC) and Faperj for financial support. We also would like to thanks Global Science Editing, UK, Mariel Marlow and Hugo Caire Castro de Faria Neto for English revision.

References

Akopyants, N.S., Kimblin, N., Secundino, N., Patrick, R., Peters, N., et al., 2009. Demonstration of genetic exchange during cyclical development of *Leishmania* in the sand fly vector. Science 324, 265–268.

Alam, M.Z., Kuhls, K., Schweynoch, C., Sundar, S., Rijal, S., et al., 2009. Multilocus microsatellite typing (MLMT) reveals genetic homogeneity of *Leishmania donovani* strains in the Indian subcontinent. InfectGenetEvol 9, 24–31.

Belli, A.A., Miles, M.A., Kelly, J.M., 1994. A putative *Leishmania panamensis*/*Leishmania braziliensis* hybrid is a causative agent of human cutaneous leishmaniasis in Nicaragua. Parasitology 109 (Pt. 4), 435–442.

Benatti, T.R., Valicente, F.H., Aggarwal, R., Zhao, C., Walling, J.G., et al., 2010. A neo-sex chromosome that drives postzygotic sex determination in the hessian fly (*Mayetiola destructor*). Genetics 184, 769–777.

Boité, M.C., Mauricio, I.L., Miles, M.A., Cupolillo, E., 2012. New insights on taxonomy, phylogeny and population genetics of *Leishmania* (*Viannia*) parasites based on multilocus sequence analysis. PLoS Negl. Trop. Dis. 6, e1888.

Boothroyd, C.E., Dreesen, O., Leonova, T., Ly, K.I., Figueiredo, L.M., et al., 2009. A yeast-endonuclease-generated DNA break induces antigenic switching in *Trypanosoma brucei*. Nature 459, 278–281.

Brito, M.E., Andrade, M.S., Mendonca, M.G., Silva, C.J., Almeida, E.L., et al., 2009. Species diversity of *Leishmania* (*Viannia*) parasites circulating in an endemic area for cutaneous leishmaniasis located in the Atlantic rainforest region of northeastern Brazil. Trop. Med. Int. Health 14, 1278–1286.

Camara, A.C., Varela-Freire, A.A., Valadares, H.M., Macedo, A.M., D'Avila, D.A., et al., 2010. Genetic analyses of *Trypanosoma cruzi* isolates from naturally infected triatomines and humans in northeastern Brazil. Acta Trop. 115, 205–211.

Coelho, A.C., Leprohon, P., Ouellette, M., 2012. Generation of *Leishmania* hybrids by whole genomic DNA transformation. PLoS Negl. Trop. Dis. 6, e1817.

Cupolillo, E., Grimaldi, G., Momen, H., 1994. A general classification of new-world *Leishmania* using numerical zymotaxonomy. Am. J. Trop. Med. Hyg. 50, 296–311.

Dujardin, J.C., De Doncker, S., Jacquet, D., Banuls, A.L., Balavoine, M., et al., 2007. Clonal propagation and the fast generation of karyotype diversity: an in vitro *Leishmania* model. Parasitology 134, 33–39.

El Baidouri, F., Diancourt, L., Berry, V., Chevenet, F., Pratlong, F., et al., 2013. Genetic structure and evolution of the *Leishmania* genus in Africa and Eurasia: what does MLSA tell us. PLoS Negl. Trop. Dis. 7, e2255.

Ferreira, G.E., dos Santos, B.N., Dorval, M.E., Ramos, T.P., Porrozzi, R., et al., 2012. The genetic structure of *Leishmania infantum* populations in Brazil and its possible association with the transmission cycle of visceral leishmaniasis. PLoS One 7, e36242.

Garin, Y.J., Sulahian, A., Pratlong, F., Meneceur, P., Gangneux, J.P., et al., 2001. Virulence of *Leishmania infantum* is expressed as a clonal and dominant phenotype in experimental infections. Infect. Immun. 69, 7365–7373.

Gonzalez-Gonzalez, A., Sanchez-Reyes, L.L., Delgado Sapien, G., Eguiarte, L.E., Souza, V., 2013. Hierarchical clustering of genetic diversity associated to different levels of mutation and recombination in *Escherichia coli*: a study based on Mexican isolates. Infect. Genet. Evol. 13, 187–197.

Jinkerson, R.E., Radakovits, R., Posewitz, M.C., 2013. Genomic insights from the oleaginous model alga *Nannochloropsis gaditana*. Bioengineered 4, 37–43.

Martin, D.P., Lemey, P., Lott, M., Moulton, V., Posada, D., et al., 2010. RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26, 2462–2463.

Oddone, R., Schweynoch, C., Schonian, G., de Sousa, C.S., Cupolillo, E., et al., 2009. Development of a multilocus microsatellite typing approach for discriminating strains of *Leishmania* (*Viannia*) species. JClinMicrobiol 47, 2818–2825.

Otto, T.D., Vasconcellos, E.A., Gomes, L.H., Moreira, A.S., Degrave, W.M., et al., 2008. ChromaPipe: a pipeline for analysis, quality control and management for a DNA sequencing facility. Genet. Mol. Res. 7, 861–871.

Paul, S., Linardopoulou, E.V., Billig, M., Tchesnokova, V., Price, L.B., et al., 2013. Role of homologous recombination in adaptive diversification of extraintestinal *Escherichia coli*. J. Bacteriol. 195, 231–242.

Prugnolle, F., De Meeus, T., 2010. Apparent high recombination rates in clonal parasitic organisms due to inappropriate sampling design. Heredity (Edinb.) 104, 135–140.

Rickert, A.M., Premstaller, A., Gebhardt, C., Oefner, P.J., 2002. Genotyping of Snps in a polyploid genome by pyrosequencing. Biotechniques 32 (592–593), 596–598, 600 passim.

Rogers, M.B., Hilley, J.D., Dickens, N.J., Wilkes, J., Bates, P.A., et al., 2011. Chromosome and gene copy number variation allow major structural change between species and strains of *Leishmania*. Genome Res. 21, 2129–2142.

Rougeron, V., De Meeus, T., Hide, M., Waleckx, E., Dereure, J., et al., 2010. A battery of 12 microsatellite markers for genetic analysis of the *Leishmania* (*Viannia*) *guyanensis* complex. Parasitology 137, 1879–1884.

Sterkers, Y., Lachaud, L., Crobu, L., Bastien, P., Pages, M., 2011. FISH analysis reveals aneuploidy and continual generation of chromosomal mosaicism in *Leishmania major*. Cell Microbiol. 13, 274–283.

Sterkers, Y., Lachaud, L., Bourgeois, N., Crobu, L., Bastien, P., et al., 2012. Novel insights into genome plasticity in Eukaryotes: mosaic aneuploidy in *Leishmania*. Mol. Microbiol. 86, 15–23.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., et al., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.

Tibayrenc, M., Ayala, F.J., 2012. Reproductive clonality of pathogens: a perspective on pathogenic viruses, bacteria, fungi, and parasitic protozoa. Proc. Natl. Acad. Sci. U.S.A. 109, E3305–E3313.

Tibayrenc, M., Ayala, F.J., 2013. How clonal are *Trypanosoma* and *Leishmania*? Trends Parasitol. 29, 264–269.

Tibayrenc, M., Kjellberg, F., Ayala, F.J., 1990. A clonal theory of parasitic protozoa: the population structures of *Entamoeba*, *Giardia*, *Leishmania*, *Naegleria*, *Plasmodium*, *Trichomonas*, and *Trypanosoma* and their medical and taxonomical consequences. ProcNatlAcadSciUSA 87, 2414–2418.

Ubeda, J.M., Legare, D., Raymond, F., Ouameur, A.A., Boisvert, S., et al., 2008. Modulation of gene expression in drug resistant *Leishmania* is associated with gene amplification, gene deletion and chromosome aneuploidy. Genome Biol. 9, R115.

van Belkum, A., Struelens, M., de Visser, A., Verbrugh, H., Tibayrenc, M., 2001. Role of genomic typing in taxonomy, evolutionary genetics, and microbial epidemiology. Clin. Microbiol. Rev. 14, 547–560.

Vink, C., Rudenko, G., Seifert, H.S., 2011. Microbial antigenic variation mediated by homologous DNA recombination. FEMS Microbiol. Rev..

Zemanova, E., Jirku, M., Mauricio, I.L., Horak, A., Miles, M.A., et al., 2007. The Leishmania donovani complex: genotypes of five metabolic enzymes (ICD, ME, MPI, G6PDH, and FH), new targets for multilocus sequence typing. Int. J. Parasitol. 37, 149–160.

5. Discussão

É provável que características intrínsecas aos parasitas do gênero *Leishmania* representem parte das razões que dificultam avanços nos estudos direcionados para estratégias de combate, tanto aqueles relacionados ao tratamento da doença como os voltados para medidas profiláticas. Uma dessas características pode ser a variabilidade genética elevada encontrada em *Leishmania*. Portanto, uma base taxonômica adequada, gerada a partir do conhecimento da estrutura de população e diversidade filogenética do parasita é essencial para entender alterações epidemiológicas da doença (Banuls et al., 1999). Como apontado por Wilson, 2005 (Wilson, 2005) "todo o tremendo desenvolvimento nas áreas de biologia celular e molecular podem ser prejudicados se não forem balanceados por um progresso similar na compreensão da diversidade biológica".

Leishmania é um gênero de protozoário parasita notável pelo grande número de espécies descritas, sendo que estudos empregando métodos moleculares para análises filogenéticas têm sugerido que realmente existe uma variabilidade elevada para este gênero e para algumas de suas espécies (Boite et al., 2012; Kuhls et al., 2013). Esta variabilidade tem impacto direto na saúde publica, principalmente em países como o Brasil, onde variadas espécies e cepas coexistem em certas regiões (Cupolillo et al., 2003) e levam a formas clínicas e respostas ao tratamento diferentes (Grimaldi & Tesh, 1993), o que invalida o diagnóstico apenas por orientação geográfica (Grimaldi & Tesh, 1993). Assim, para alguns casos de leishmaniose cutânea, a identificação do parasita deveria preceder a prescrição médica (Monge-Maillo & Lopez-Velez, 2013; Monge-Maillo et al., 2013; Rioux et al., 1990).

A identificação de patógenos e estudos de sua diversidade faz uso de ferramentas e metodologias que precisam ser bem estabelecidas com base na classificação taxonômica. Em *Leishmania* a validade taxonômica de várias espécies tem sido constantemente discutida, indicando a necessidade de uma revisão taxonômica para o gênero (Fraga et al., 2010; Schonian et al., 2010). Para tal um método único e padronizado de tipagem e identificação ainda é necessário, sendo que um marcador molecular é a melhor opção para atender tal demanda por razões já discutidas. Ao desenvolver um marcador molecular *multi locus* é preciso: i) considerar características moleculares do organismo estudado; ii) avaliar a aplicabilidade deste marcador através da comparação com outros; iii) validá-lo com um painel apropriado de amostras que represente a diversidade do gênero; iv) definir quais *loci* combinados formarão a unidade taxonômica.

É neste cenário que os objetivos do presente trabalho estão inseridos. Serão discutidos, em seguida, os resultados obtidos e a contribuição dos mesmos para o estudo das leishmanioses.

5.1. Marcadores *multi locus* – MLST e MLMT

MLST e MLMT representam dois marcadores com as vantagens da abordagem molecular e da codominância, o que permite a detecção de três possibilidades alélicas em um organismo diploide. Cada um possui graus diferentes de resolução e por isso podem revelar graus diferentes de variabilidade genética. Juntos contribuem para tipagem, filogenia, genética de populações e epidemiologia molecular das leishmanioses (Schonian et al., 2010; Schonian et al., 2011). Por essa razão foram utilizados para o estudo de cepas de *L.* (*Viannia*) neste trabalho.

Com o objetivo de desenvolver um painel MLST foram selecionadas para sequenciamento quatro regiões polimórficas, codificantes para enzimas metabólicas utilizadas nos ensaios de MLEE. Cepas de diferentes espécies, zimodemas e origem geográfica foram incluídas no estudo. Os fragmentos escolhidos, representando diferentes *loci*, apresentaram variados graus de diversidade entre os grupos de espécie, o que demonstra a aplicabilidade deste painel para inferências inter e intra-específicas. O MLST já foi aplicado previamente em cepas *Leishmania* (*Viannia*) (Tsukayama et al., 2009), mas o estudo não incluiu diversidade de cepas suficiente e a proposta foi diagnóstica, não incluindo análises filogenéticas robustas sobre os dados gerados. Por isso as análises do presente trabalho representam uma abordagem diferencial. São elas: i) o uso de programas como SplitsTree (Huson & Bryant, 2006), que acomoda sítios ambíguos e duplica as possibilidades de sequências quando estas são concatenadas para a construção da rede; ii) e do BURST, que gera complexos clonais (CC) entre cepas geneticamente relacionadas, mas não idênticas, o que permite uma melhor compreensão da relação entre os *diployd sequence types* (DST) gerados; iii) os CC tornam os dados mais apropriados para inferências epidemiológicas; iv) a diversidade de cepas utilizadas permitiu uma cobertura ampla da variabilidade genética. Apesar de não ser ainda a construção final de um sistema de tipagem *multi locus*, representa o ponto de início de um painel que pode ser ampliado em espécies, cepas e *loci*. Assim, o trabalho pôde contribuir para o estudo de marcadores moleculares e de sistemática de *Leishmania* (*Viannia*), e mais especificamente para as cepas que circulam no Brasil.

Dentro da discussão sobre MLST para tipagem de *Leishmania* surge como argumento contrário ao seu uso a crescente popularização do sequenciamento completo de genomas

(*Whole Genome Sequencing* – WGS). Este poderia se tornar a forma definitiva de tipagem de patógenos (*Whole Genome Sequencing Typing* – WGST) e, portanto fazer da elaboração de um MLST um esforço desnecessário (Pallen et al., 2010; Perez-Losada et al., 2013). Entretanto, este não é o cenário atual. Assim como o MLST, a análise por WGS teve inicio com bactérias, e pelo grande volume de dados gerados é imprescindível o uso de sistemas computacionais e critérios para sumarizar, identificar e quantificar a variabilidade WGS entre espécimes. Apesar dos inúmeros sistemas desenvolvidos pelos bacteriologistas, ainda não há consenso e muitos estudos continuam sendo realizados pelo tradicional MLST (Maiden et al., 2013). Uma das dificuldades no uso do WGST é lidar com o grande volume de dados gerados, além da ocorrência de erros durante a reunião e anotação dos genomas, principalmente em relação aos SNPs (Feng et al., 2014; Thomas & Robinson, 2014). Erros que podem ser confundidos com ocorrência de uma amostra variante (Perez-Losada et al., 2013). Por isso, foi sugerido por Maiden e colaboradores (Maiden et al., 2013) uma abordagem gene-por-gene, similar ao que já é feito em esquemas MLST, mas utilizando a informação do WGS. Perez-Lozada e colaboradores (Perez-Losada et al., 2013) apresentam proposta similar: utilizar as informações disponíveis nos painéis MLST durante a anotação dos genomas para minimizar erros. Esse sistema WG-MLST se beneficiaria da cobertura completa do genoma, associado às facilidades de aplicação e compartilhamento de informações do MLST. A proposta permite ainda a utilização das plataformas MLST e informações já disponíveis. Fica claro também na revisão de Maiden a importância da combinação e do número de genes incluídos, dependendo da pergunta em questão. Neste ponto, o sequenciamento completo de genomas pode contribuir ao ampliar o número dos genes em painéis MLST e permitir combinações de *loci* importantes para perguntas distintas, como por exemplo, para abordar aspectos filogenéticos e biológicos, como resistência a drogas. Outra contribuição marcante do WGS são as informações sobre número de cópias cromossômicas. Downing e colaboradores (Downing et al., 2011) encontraram estruturação entre 16 isolados clínicos de *L*. (*L*.) *donovani* de uma mesma região que apresentaram distinta susceptibilidade a droga *in vitro*. Foi observada uma variação de SNPs baixa entre os isolados, e a correlação obtida deveu-se mais às informações estruturais do genoma obtidas por WGS do que a diversidade das sequencias de DNA propriamente. Mas ainda assim, este exemplo demonstra que, pela sua elevada resolução, o WGS permite a identificação de *loci* para análises de variabilidade intraespecífica. Uma vez identificados, os *loci* podem ser incluídos em painéis MLST. Porém, para utilização mais ampla do MLST, nos vários níveis taxonômicos, é essencial cobrir a diversidade do patógeno, incluindo nos estudos cepas de

grupos distintos, mas próximos filogeneticamente, identificáveis por outras análises e de origem variada (Davis M. Hillis, 1996). O sequenciamento completo do genoma de poucos isolados, ou de um grupo não representativo da variabilidade pode incorrer em subestimação da diversidade e erro tipo II (Davis M. Hillis, 1996). Outro ponto importante que ainda dificulta o uso de WGS em larga escala para *Leishmania* é a maior complexidade da estrutura de seu genoma (Capítulo IV) (Dujardin et al., 2007; Peacock et al., 2007; Sterkers et al., 2011) quando comparado ao de bactérias. Sendo assim, peculiaridades como o mosaicismo cromossomal (Sterkers et al., 2011) precisam ser mais bem compreendidas antes que um sistema WGST de larga escala seja estabelecido. Diante dos relatos no campo da tipagem bacteriana, pode-se concluir que a construção de um sistema MLST para *Leishmania* ainda merece esforços, e que os resultados poderão trazer benefícios imediatos e serem associados futuramente aos dados gerados por WGS.

O MLMT, com seu alto poder discriminatório, foi utilizado neste trabalho para avaliação da estrutura da população de cepas de *L*. (*Viannia*) que circulam no Brasil. Os iniciadores utilizados já haviam sido previamente validados para cepas de *L*. (*Viannia*) (Oddone et al., 2009). Devido ao elevado polimorfismo dos microssatélites, esta abordagem detecta variação entre espécies próximas, o que representa uma característica importante para o estudo desses parasitas. A análise pelo programa STRUCTURE gerou três populações pelos cálculos de ΔK, entre as 120 cepas analisadas. As populações POP1 e POP2, formadas por cepas de *L*. (*V*.) *guyanensis* e *L*. (*V*.) *braziliensis* da região costeira, respectivamente, foram bem suportadas. A população POP3, entretanto, foi mais heterogênea, incluindo as outras espécies e cepas de *L*. (*V*.) *braziliensis* do norte do Brasil. Ficou demonstrado que a heterogeneidade genética de cepas de *L*. (*V*.) *braziliensis* isoladas no Brasil é maior do que havia sido descrito. Adicionalmente, o estudo permitiu a observação de populações entre as cepas estudadas; porém a falta de estruturação clara pode indicar panmixia, o que justifica a grande diversidade genética encontrada em isolados de *L*. (*Viannia*) no Brasil.

O presente resultado corroborou a informação que, apesar das suas contribuições, o MLMT não é indicado como técnica ouro para identificação das espécies de *Leishmania* (Schonian et al., 2011). Isso se deve a evolução rápida dos microssatélites que faz com que um conjunto de marcadores desenvolvidos para uma espécie funcione apenas nas espécies mais próximas (Schonian et al., 2011). Sendo assim, o mesmo painel de alvos não pode cobrir os dois subgêneros, e por isso os marcadores diferem dos estudos com subgênero *L*. (*Leishmania*) (Alam et al., 2014; Ferreira et al., 2012).

Foi possível com este trabalho demonstrar a elevada diversidade em microssatélites para *L*. (*Viannia*) e indicar que *L*. (*V*.) *braziliensis*, e *L*. (*V*.) *guyanensis* (em menor grau) apresentam uma estrutura clonal, mas com eventos de recombinação. Após a conclusão deste capítulo, foi possível realizar comparações entre os resultados com os dois marcadores, que serão discutidas em seguida.

A utilização e comparação de dados obtidos por diferentes marcadores já foi feita em diversos organismos, como em fungos (L'Ollivier et al., 2012), toxoplasma (Su et al., 2012) e bactérias. Naqueles estudos os resultados dos diferentes métodos apresentaram correlação elevada, e a combinação dos dados permitiu inferências mais completas sobre grupos de cepas / espécies controversas. Técnicas de genotipagem com diferentes níveis de resolução aplicados sobre as mesmas cepas tendem a revelar diferentes graus de variabilidade. Para compreender essa variação é importante compreender o processo evolutivo envolvido (mutação, recombinação e seleção) (van Belkum et al., 2001).

Em relação ao presente trabalho, a comparação dos resultados obtidos com MLSA e MLMT demonstra como a associação dessas duas abordagens é valiosa. Para ambos a diversidade encontrada foi elevada, corroborando a literatura. Quase todas as cepas apresentaram um perfil de microssatélites único, exceto pelas 13 cepas de Minas Gerais, que provavelmente foram isoladas durante um surto neste estado. Também foi elevado o número de DSTs únicos gerados. A divisão das cepas de *L*. (*V*.) *braziliensis* em duas populações por MLMT corrobora o achado de dois CC para esta espécie pelo MLSA. A formação de um grupo único, distinto com as cepas de *L*. (*V*.) *guyanensis* foi outro achado dos dois estudos (exceto quando cepas de *L*. (*V*). *panamensis* e *L*. (*V*.) *shawi* foram incluídas para construção de rede de haplótipos e em MSLA). Os resultados sobre *L*. (*V*.) *guyanensis* de forma geral corrobora a literatura, porém com a diversidade intraespecífica maior do que já foi relatado (Brito et al., 2009; Cupolillo et al., 1997; Cupolillo et al., 2003; de Brito et al., 2012; Oddone et al., 2009). Em relação às outras espécies incluídas (*L*. (*V*.) *lainsoni*, *L*. (*V*.) *naiffi*, *L*. (*V*.) *utingensis* e *L*. (*V*.) *lindenbergi*) os achados para os dois marcadores estão de acordo com a classificação por MLEE (Cupolillo et al., 1994).

Alguns pontos foram discordantes entre as abordagens, como a observação de cepas de *L*. (*V*.) *braziliensis* agrupadas no CC4 pelo MLSA, mas incluídas na POP2 ou na POP3 em MLMT. O CC4 aparece como grupo menos heterogêneo, representado a expansão de um clone adaptado. Esperava-se que as cepas incluídas neste grupo apresentassem o perfil da POP2, mais homogênea e bem estruturada. Apesar de só parte das cepas de *L*. (*V*.) *braziliensis* terem resultados para os dois métodos (N=23), é possível inferir que as diferenças

no perfil de microssatélites entre POP2 e POP3 não possam ser extrapolados para os *loci* conservados codificantes utilizados no MLSA. Por isso, esses resultados aparentemente conflitantes são na verdade reflexos das diferenças intrínsecas aos marcadores escolhidos, como o grau de polimorfismo, taxa de mutação, ocorrência de recombinação e pressão de seleção sobre os *loci* codificantes do MLSA. Assim é possível reforçar a importância de associar diferentes marcadores.

Uma informação controversa foi em relação às cepas de *L.* (*V.*) *shawi*. Por MLSA a validade da espécie pôde ser questionada, uma vez que as três cepas estudadas se apresentaram como um complexo clonal distinto, mas ainda assim dentro do grupo de *L.* (*V.*) *guyanensis* na NeighborNet. Por MLMT, no entanto, as cepas *L.* (*V.*) *shawi* foram agrupadas na heterogênea POP3, e ainda subestruturadas como POP3A com outras espécies. Tal achado demonstrou que as cepas de *L.* (*V.*) *shawi* parecem tratar-se de um grupo distinto de *L.* (*V.*) *guyanensis*, mas ainda assim não estruturado de forma individualizada como unidade de espécie. Tal incoerência pode ser devido ao alto poder discriminatório do MLMT, quando comparado ao MLSA. Este último, contudo, se apresenta como melhor fonte de informação filogenética (Schonian et al., 2011). Portanto, no que se refere à revisão taxonômica, o resultado com MLSA precisa ser considerado, principalmente por ser corroborado por outras análises (Schonian et al., 2011).

Ao total, 40 cepas foram analisadas, tanto para MLSA quanto para MLMT (Anexo 1). A comparação das informações gerou uma perspectiva para esta tese: completar a sobreposição de cepas para os dois métodos, de forma que uma análise comparativa de todo o painel possa ser feita. Isso permitirá, por exemplo, compreender se as *L.* (*V.*) *braziliensis* agrupadas na heterogênea POP3 representam *outliers* ou um grupo distinto dentro dessa espécie.

Os resultados obtidos por MLSA e MLMT com isolados de *Leishmania* (*Viannia*) indicam que ambos marcadores apresentam potencial para aplicação no estudo intraespecífico de cepas de *L.* (*V.*) *braziliensis*. Em cepas de *Leishmania* (*Viannia*), o MLSA mostrou alto poder resolutivo, capaz de diferenciar as espécies do subgênero de acordo com classificação taxonômica por MLEE (Cupolillo et al., 1994, 1995). Através de MLSA foi possível detectar complexos clonais distintos para cepas de *L.* (*V.*) *braziliensis*, o que sugere que esta abordagem pode ser útil para análises epidemiológicas e de estrutura de população através da determinação de haplótipos/diplótipos estáveis (Boite et al., 2012).

MLSA demonstrou ser uma ferramenta epidemiológica eficaz em estudo do surto de leishmaniose cutânea ocorrido no estado de Santa Catarina em 2006, sendo capaz de agrupar

casos autóctones e importados, além de permitir comparação com cepas circulantes em outras áreas do país (Capítulo III). Por outro lado, através do MLMT, a ocorrência de recombinação intraespecífica e a separação de cepas de *L.* (*V.*) *braziliensis* em duas populações foi detectada (Kuhls et al., 2013). Sendo assim, o MLSA e MLMT provaram sua contribuição no estudo de genética de populações em cepas de *Leishmania* (*Viannia*), mesmo com o advento atual do sequenciamento completo de genomas. Os resultados indicam que esses marcadores podem ser bastante informativos também para o estudo de variabilidade intraespecífica, especialmente quando há diversidade representada de forma suficiente nas amostras e com a associação de dados obtidos pelas duas abordagens.

As vantagens de associação de marcadores são claras, entretanto, ainda existe debate de qual seria o melhor método individualizado. Klaassen (Klaassen, 2009) apresentou uma revisão comparando MLST e MLMT para tipagem de *Aspergillus fumigatus*. Nela o autor conclui que MLMT é a melhor opção por considerá-lo mais rápido, de alto desempenho, mais barato e com maior poder discriminatório. Levando em consideração os resultados desta tese, para *Lesihmania* (*Viannia*) o MLMT apresenta sim maior poder discriminatório, mas não é mais rápido ou barato que o MLST, além da reprodutibilidade daquele ser questionável. MLMT é uma ferramenta valiosa para o estudo de genética de populações, mas para tipagem de *Leishmania* e avaliações taxonômicas o MLSA apresenta-se como marcador de escolha.

5.2. Aplicação epidemiológica

Com a emergência das leishmanioses em novas regiões cresce cada vez mais a necessidade de métodos moleculares para estudo e rastreamento de surtos. Para a vigilância epidemiológica desses eventos, metodologias com resolução elevada, que permitam a detecção de variações intraespecíficas, são de extrema importância. Em contrapartida, um método com menor resolução que permita a detecção do grupo / complexo clonal ao qual a cepa envolvida pertence também é relevante para compreensão da origem do surto. Uma resolução ainda menor pode ser suficiente para a determinação da espécie envolvida. Em situações epidêmicas, torna-se muito valiosa uma ferramenta que permita que todos esses níveis sejam alcançados em conjunto, utilizando a mesma abordagem para registros, compreensão da origem dos casos e para tomada de medidas de controle (Maiden et al., 2013).

Marcadores moleculares *multi locus* são frequentemente utilizados para estudo de surtos epidêmicos causados por diferentes agentes, como bactérias, plasmódio e vírus (Barco et al., 2014; Spanakos et al., 2013) e para vigilância epidemiológica (Brehony et al., 2007).

Também em tripanosomatídeos são usados para tentar compreender, por exemplo, a correlação entre genótipos, aspectos clínicos e ciclos de transmissão (Guhl & Ramirez, 2013; Ramirez et al., 2013).

Em *Leishmania*, a complexidade do genoma e ausência de sistemas moleculares padronizados / marcadores epidemiológicos dificultam a comparação de dados e estudos, em larga escala, de caracterização da emergência de casos em regiões livres da doença. Diante disso, um dos objetivos desta tese foi avaliar a aplicabilidade do uso dos marcadores moleculares descritos nos Capítulos I e III em situações epidemiológicas práticas.

Os resultados obtidos nos trabalhos I e III permitiram a detecção de variação intraespecífica, com informações relevantes, sobretudo para *L*. (*V*.) *braziliensis*, agente principal da LTA no Brasil. O MLSA foi a metodologia escolhida para aplicação prática, apesar da maior resolução do MLMT, por permitir concomitantemente a identificação da espécie e a realização de inferências filogenéticas em diferentes níveis taxonômicos. Ao ampliar o painel de *loci* sequenciados (três novos marcadores) e de cepas de *L*. (*V*.) *braziliensis*, a abordagem MLSA atingiu nível de resolução adequado para a proposta. Suas características de reprodutibilidade (maior quando comparada ao MLMT) e alto rendimento também contribuíram para sua escolha como abordagem a ser testada como ferramenta epidemiológica.

No início de 2006, um surto de leishmaniose cutânea ocorreu em Santa Catarina, onde a doença ainda não havia sido descrita (Marlow et al., 2013). A espécie apontada como responsável foi *L*. (*V*.) *braziliensis*, mas algumas questões sobre este surto persistiram, como por exemplo: a introdução de *L*. (*V*.) *braziliensis* no Estado foi um evento recente? Qual seria a relação das cepas envolvidas com as outras *L*. (*V*.) *braziliensis* que circulam no Brasil? O cenário representou a oportunidade para demonstrar a aplicabilidade do MLSA para responder tais questões, uma vez que as cepas foram eficientemente isoladas e as informações epidemiológicas eram completas.

A metodologia MLSA foi capaz de detectar variação intra-específica compatível com as características epidemiológicas das cepas isoladas do surto. Os dados de sequencia de DNA e geração de complexos clonais seguiu o racional descrito no Capítulo I, porém, o programa de agrupamento STRUCTURE, usualmente aplicado para análises microssatélites, foi utilizado para forçar as sequencias únicas obtidas em um grupo, e assim permitir análises estatísticas. Neste caso, a associação entre a forma clínica e grupos formados pelo STRUCTURE foi encontrada, demonstrando o potencial epidemiológico da metodologia.

Os alvos MLSA escolhidos para o presente trabalho são regiões codificantes de genes conservados (*housekeeping*) e representam bons marcadores taxonômicos. Para expandir e aprimorar o MLSA como ferramenta epidemiológica, alvos potencialmente envolvidos em aspectos como virulência podem ser incorporados ao painel MLSA, como feito em bactérias (Enright et al., 1999; Enright & Spratt, 1999).

5.3. Características moleculares das leishmânias e o desenvolvimento de marcadores

Aneuploidia e diversidade cariotípica vêm sido discutidas em *Leishmania* por muitos anos. No entanto, só recentemente o moisacismo aneuploide foi caracterizado experimentalmente (Bastien et al., 1990; Dujardin et al., 2007). Cepas de *Leishmaia* (*L.*) *major* e seus clones e subclones tiveram o conteúdo cromossômico avaliado individualmente, em nível celular (Rogers et al., 2011; Sterkers et al., 2011), o que permitiu aos autores concluir que o mosaicismo aneuploide é regra para a espécie avaliada, e representa uma característica constitutiva do parasito (Sterkers et al., 2011). O mecanismo que gera o mosaico ainda não é conhecido, mas as hipóteses são segregação assimétrica dos cromossomos ou duplicação aberrante dos mesmos durante a mitose (Sterkers et al., 2012). Tal padrão de plasticidade genômica gera diversidade fenotípica, que pode ser determinante para a bem sucedida adaptação do gênero *Leishmania* a diferentes meios (Dujardin et al., 2007), para virulência e ocorrência de falha no tratamento das leishmanioses (Mannaert et al., 2012).

Como apontado nos itens 1.3 e 5.1, marcadores moleculares codominantes foram escolhidos para este trabalho por, entre outros aspectos, permitirem a detecção de heterozigotos. Essa detecção é feita pela observação de sítios ambíguos (picos sobrepostos no eletroferograma) em MLSA e picos de mesma intensidade com diferença de mais de um nucleotídeo em MLMT. Entretanto, esta detecção pode ser afetada pelas características moleculares descritas. Isso se deve ao fato do sequenciamento de DNA a partir de uma cepa não clonada não conseguir determinar com precisão se uma população é composta: i) por células heterozigotas; ii) por células homozigotas com alelos distintos; iii) por um mosaico que gere os sítios ambíguos.

O conceito de heterozigoto normalmente aplicado em *Leishmania* é o apresentado por Odds e Jacobsen (Odds & Jacobsen, 2008), que considera picos perfeitamente sobrepostos de intensidade menor quando comparado aos picos adjacentes. Além disso, o mesmo padrão precisa estar presente nas duas fitas sequenciadas. Porém, para organismos poliploides, o conceito apresentado na revisão de Van Belkum e colaboradores (van Belkum et al., 2001), onde as cópias identificadas para um gene não são idênticas, é mais apropriado.

Diante do que foi acima exposto, era imprescindível para completar este trabalho, avaliar o efeito da diversidade cariotípica sobre tipos de sequencias de DNA obtidas a partir de clones e de cultura não clonada de mesma cepa. A comparação das sequencias demonstrou que, após clonagem das células e clonagem molecular do alvo 6PGDH, foi possível detectar novos alelos. A razão para estes não serem observados antes das etapas de clonagem podem ser: i) estavam presentes em cromossomos supranumerários; ii) pertencerem a uma infrapopulação celular.

O trabalho demonstrou que a diversidade genética em *Leishmania* é subestimada pelos estudos moleculares, uma vez que nem todos os alelos são detectados. Isso tem impacto em diferentes estudos sobre as leishmanioses, inclusive sobre os estudos de genética de populações. O mosaicismo pode levar a um falso déficit de heterozigose, uma vez que alelos menos frequentes em cromossomos supranumerários podem não ser detectados.

Tal impacto pode ser minimizado por uma apropriada representação das cepas circulantes incluídas nos trabalhos, uma vez que maior número de alelos distintos seria detectado. Outro ponto que minimiza esse impacto é a estabilidade do processo, descrita por Sterkers e colaboradores (Mannaert et al., 2012; Sterkers et al., 2012). Os autores mostraram que a maioria dos clones apresenta padrão de mosaico quase idêntico ao padrão da parental. Ainda assim, é imprescindível que futuros projetos considerem e busquem compreender o mosaico de cariótipos nas diferentes cepas e seus clones.

A plasticidade genômica, descrita também para outros tripanosomatídeos (Lima et al., 2013; Minning et al., 2011), associada a publicações que demonstram ocorrência de recombinação (Akopyants et al., 2009) desafiou, ou ao menos relativizou, o conceito de clonalidade apresentado para protozoários por Tibayrenc (Tibayrenc et al., 1990; Tibayrenc & Ayala, 1999). Realmente, no presente trabalho observa-se uma diversidade cariotípica intra-cepa, refletida nas sequencias de DNA. Esta diversidade pode ser gerada após mitose da célula parasitária pela duplicação aberrante e/ou segregação assimétrica de cromossomos (Sterkers et al., 2012). Mesmo que a cepa contenha apenas células homozigotas, as mesmas podem apresentar diversidade de cariótipos que contribuirá para a variabilidade de fenótipos (Mannaert et al., 2012).

Diante de tal discussão, Tibayrenc afirma que o conceito de clonalidade não leva em consideração mecanismos citológicos, mas sim as consequências genéticas da clonalidade. E que, portanto, é perfeitamente compatível com o moisacismo aneuploide. Comenta ainda que a teoria da clonalidade não significa ausência total de recombinação, mas que a ocorrência desta não é frequente o bastante para alterar o padrão da estrutura da população.

O resultado apresentado no Capítulo IV associado à literatura recente aponta para a necessidade de estudos de associação entre o padrão de mosaico e fenótipos como virulência e resistência a drogas. Como perspectiva dentro da linha de pesquisa onde a presente tese se insere é possível apontar também como estudo relevante a determinação da origem da aneuploidia descrita para *Tripanosma cruzi* e *Leishmania*: teria uma origem comum, e ocorreria também nos gêneros de tripanosomatídeos monoxênicos, como *Crithidia* e *Leptomonas*?

6. Conclusão

O painel MLSA construído e validado apesenta-se como boa alternativa molecular para tipagem, estudos taxonômicos e epidemiológicos em *L*. (*Viannia*). Como tal, pode representar o substituto molecular para o MLEE e o marcador padrão em uma revisão taxonômica do gênero. A contribuição desta abordagem ainda é relevante, mesmo com o advento do WGS. O MLSA apresenta potencial como marcador epidemiológico, apesar da inclusão de novos genes ser necessária.

O MLMT é a ferramenta de escolha para estudos de genética de populações em *L*. (*Viannia*), mas não para identificação e avaliação taxonômica do subgênero. Alguns achados do método devem ser confrontados com resultados obtidos por outros marcadores, como o MLSA, para validação da informação gerada.

A plasticidade genômica das leishmânias gera diversidade em tipos de sequencia de DNA e, portanto deve ser sempre considerada em estudos moleculares deste parasita.

7. Referências

1. Ait-Oudhia K, Gazanion E, Vergnes B, Oury B, Sereno D. Leishmania antimony resistance: what we know what we can learn from the field. Parasitology research. [Research Support, Non-U.S. Gov't Review]. 2011 Nov;109(5):1225-32.

2. 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):265-8.

3. Alam MZ, Nakao R, Sakurai T, Kato H, Qu JQ, Chai JJ, Chang KP, Schonian G, Katakura K. Genetic diversity of Leishmania donovani/infantum complex in China through microsatellite analysis. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases. 2014 Jan 27.

4. 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. [Research Support, Non-U.S. Gov't]. 2012;7(5):e35671.

5. Arevalo J, Ramirez L, Adaui V, Zimic M, Tulliano G, Miranda-Verastegui C, Lazo M, Loayza-Muro R, De Doncker S, Maurer A, Chappuis F, Dujardin JC, Llanos-Cuentas A. Influence of Leishmania (Viannia) species on the response to antimonial treatment in patients with American tegumentary leishmaniasis. J Infect Dis. 2007 Jun 15;195(12):1846-51.

6. Azeredo-Coutinho RB, Conceicao-Silva F, Schubach A, Cupolillo E, Quintella LP, Madeira MF, Pacheco RS, Valete-Rosalino CM, Mendonca SC. First report of diffuse cutaneous leishmaniasis and Leishmania amazonensis infection in Rio de Janeiro State, Brazil. Trans R Soc Trop Med Hyg. 2007 Jul;101(7):735-7.

7. Banuls AL, Hide M, Tibayrenc M. Molecular epidemiology and evolutionary genetics of Leischmania parasites. Int J Parasitol. 1999 Aug;29(8):1137-47.

8. Banuls AL, Dujardin JC, Guerrini F, De Doncker S, Jacquet D, Arevalo J, Noel S, Le Ray D, Tibayrenc M. Is Leishmania (Viannia) peruviana a distinct species? A MLEE/RAPD evolutionary genetics answer. J Eukaryot Microbiol. 2000 May-Jun;47(3):197-207.

9. Baptista C, Schubach AO, Madeira MF, Leal CA, Pires MQ, Oliveira FS, Conceicao-Silva F, Rosalino CM, Salgueiro MM, Pacheco RS. Leishmania (Viannia) braziliensis genotypes identified in lesions of patients with atypical or typical manifestations of tegumentary leishmaniasis: evaluation by two molecular markers. Experimental parasitology. [Research Support, Non-U.S. Gov't]. 2009 Apr;121(4):317-22.

10. Baptista C, Schubach Ade O, Madeira Mde F, de Freitas Campos Miranda L, Guimaraes de Souza Pinto A, Helena da Silva Barros J, Conceicao-Silva F, Fernandes Pimentel MI, da Silva Pacheco R. Evaluation of Genetic Polymorphism of Leishmania (V.) braziliensis Isolates Obtained from the Same Patient before and after Therapeutic Failure or Reactivation of Cutaneous Lesions. Journal of tropical medicine. 2012;2012:808132.

11. Barco L, Ramon E, Cortini E, Longo A, Dalla Pozza MC, Lettini AA, Dionisi AM, Olsen JE, Ricci A. Molecular Characterization of Salmonella enterica Serovar 4,[5],12:i:-DT193 ASSuT Strains from Two Outbreaks in Italy. Foodborne pathogens and disease. 2014 Feb;11(2):138-44.

12. Barnabe C, De Meeus T, Noireau F, Bosseno MF, Monje EM, Renaud F, Breniere SF. Trypanosoma cruzi discrete typing units (DTUs): microsatellite loci and population genetics of DTUs TcV and TcI in Bolivia and Peru. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases. [Research Support, Non-U.S. Gov't]. 2011 Oct;11(7):1752-60.

13. Bastien P, Blaineau C, Taminh M, Rioux JA, Roizes G, Pages M. Interclonal variations in molecular karyotype in Leishmania infantum imply a 'mosaic' strain structure. Mol Biochem Parasitol. 1990 Apr;40(1):53-61.

14. Bastien P, Blaineau C, Pages M. Molecular karyotype analysis in Leishmania. Sub-cellular biochemistry. [Research Support, Non-U.S. Gov't Review]. 1992;18:131-87.

15. Bastien P, Blaineau C, Pages M. Leishmania: sex, lies and karyotype. Parasitol Today. 1992 May;8(5):174-7.

16. Belazzoug S, Pratlong F, Rioux JA. [A new zymodeme of Leishmania tropica, agent of Aleppo boil (Syria)]. Archives de l'Institut Pasteur d'Algerie Institut Pasteur d'Algerie. [Comparative Study]. 1988;56:95-9.

17. Boite MC, Mauricio IL, Miles MA, Cupolillo E. New insights on taxonomy, phylogeny and population genetics of leishmania (viannia) parasites based on multilocus sequence analysis. PLoS Negl Trop Dis. 2012 Nov;6(11):e1888.

18. Brehony C, Jolley KA, Maiden MC. Multilocus sequence typing for global surveillance of meningococcal disease. FEMS Microbiol Rev. 2007 Jan;31(1):15-26.

19. Brito ME, Andrade MS, Mendonca MG, Silva CJ, Almeida EL, Lima BS, Felix SM, Abath FG, da Graca GC, Porrozzi R, Ishikawa EA, Shaw JJ, Cupolillo E, Brandao-Filho SP. Species diversity of Leishmania (Viannia) parasites circulating in an endemic area for cutaneous leishmaniasis located in the Atlantic rainforest region of northeastern Brazil. Tropical medicine & international health : TM & IH. [Research Support, Non-U.S. Gov't]. 2009 Oct;14(10):1278-86.

20. Bulle B, Millon L, Bart JM, Gallego M, Gambarelli F, Portus M, Schnur L, Jaffe CL, Fernandez-Barredo S, Alunda JM, Piarroux R. Practical approach for typing strains of *Leishmania infantum* by microsatellite analysis. Journal of clinical microbiology. 2002 Sep;40(9):3391-7.

21. Calvopina M, Armijos RX, Hashiguchi Y. Epidemiology of leishmaniasis in Ecuador: current status of knowledge -- a review. Memorias do Instituto Oswaldo Cruz. [Review]. 2004 Nov;99(7):663-72.

22. Chakravarty J, Sundar S. Drug resistance in leishmaniasis. Journal of global infectious diseases. 2010 May;2(2):167-76.

23. Cortes S, Esteves C, Mauricio I, Maia C, Cristovao JM, Miles M, Campino L. In vitro and in vivo behaviour of sympatric Leishmania (V.) braziliensis, L. (V.) peruviana and their hybrids. Parasitology. [Research Support, Non-U.S. Gov't]. 2012 Feb;139(2):191-9.

24. Coura JR. Dinâmica das Doenças Infecciosas e Parasitárias. Rio de Janeiro: Guanabara Koogan; 2005.

25. Cupolillo E, Grimaldi G, Jr., Momen H. A general classification of New World Leishmania using numerical zymotaxonomy. The American journal of tropical medicine and hygiene. [Research Support, Non-U.S. Gov't]. 1994 Mar;50(3):296-311.

26. Cupolillo E, Grimaldi G, Jr., Momen H. Discrimination of Leishmania isolates using a limited set of enzymatic loci. Annals of tropical medicine and parasitology. [Research Support, Non-U.S. Gov't]. 1995 Feb;89(1):17-23.

27. Cupolillo E, Grimaldi G, Jr., Momen H. Genetic diversity among Leishmania (Viannia) parasites. Annals of tropical medicine and parasitology. [Research Support, Non-U.S. Gov't]. 1997 Sep;91(6):617-26.

28. Cupolillo E, Momen H, Grimaldi G, Jr. Genetic diversity in natural populations of New World Leishmania. Memorias do Instituto Oswaldo Cruz. 1998 Sep-Oct;93(5):663-8.

29. Cupolillo E, Brahim LR, Toaldo CB, de Oliveira-Neto MP, de Brito ME, Falqueto A, de Farias Naiff M, Grimaldi G, Jr. Genetic polymorphism and molecular epidemiology of Leishmania (Viannia) braziliensis from different hosts and geographic areas in Brazil. Journal of clinical microbiology. 2003 Jul;41(7):3126-32.

30. da Silva LA, de Sousa Cdos S, da Graca GC, Porrozzi R, Cupolillo E. Sequence analysis and PCR-RFLP profiling of the hsp70 gene as a valuable tool for identifying Leishmania species associated with human leishmaniasis in Brazil. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases. 2010 Jan;10(1):77-83.

31. Dantas-Torres F, Brandao-Filho SP. Visceral leishmaniasis in Brazil: revisiting paradigms of epidemiology and control. Rev Inst Med Trop Sao Paulo. 2006 May-Jun;48(3):151-6.

32. Davis M. Hillis CM, Barbara K. mable. Molecular Systematics. second edition ed. Sunderland: Sinauer Associates; 1996.

33. de Brito ME, Andrade MS, Dantas-Torres F, Rodrigues EH, Cavalcanti Mde P, de Almeida AM, Brandao-Filho SP. Cutaneous leishmaniasis in northeastern Brazil: a critical appraisal of studies conducted in State of Pernambuco. Revista da Sociedade Brasileira de Medicina Tropical. [Review]. 2012 Jul-Aug;45(4):425-9.

34. Debourgogne A, Gueidan C, Hennequin C, Contet-Audonneau N, de Hoog S, Machouart M. Development of a new MLST scheme for differentiation of Fusarium solani Species Complex (FSSC) isolates. J Microbiol Methods. 2010 Sep;82(3):319-23.

35. Dobzhansky T. Drosophila Miranda, a New Species. Genetics. 1935 Jul;20(4):377-91.

36. Downing T, Imamura H, Decuypere S, Clark TG, Coombs GH, Cotton JA, Hilley JD, de Doncker S, Maes I, Mottram JC, Quail MA, Rijal S, Sanders M, Schonian G, Stark O, Sundar S, Vanaerschot M, Hertz-Fowler C, Dujardin JC, Berriman M. Whole genome sequencing of multiple Leishmania donovani clinical isolates provides insights into population structure and mechanisms of drug resistance. Genome Res. 2011 Dec;21(12):2143-56.

37. Dujardin JC, De Doncker S, Jacquet D, Banuls AL, Balavoine M, Van Bockstaele D, Tibayrenc M, Arevalo J, Le Ray D. Clonal propagation and the fast generation of karyotype diversity: An *in vitro Leishmania* model. Parasitology. 2007 Jan;134(Pt 1):33-9.

38. Eichner S, Thoma-Uszynski S, Herrgott I, Sebald H, Debus A, Tsianakas A, Ehrchen J, Harms G, Simon M, Sunderkotter C, Bogdan C. Clinical complexity of Leishmania (Viannia) braziliensis infections amongst travelers. European journal of dermatology : EJD. 2013 Apr 1;23(2):218-23.

39. 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 Jun;7(6):e2255.

40. Enright MC, Fenoll A, Griffiths D, Spratt BG. The three major Spanish clones of penicillin-resistant Streptococcus pneumoniae are the most common clones recovered in recent cases of meningitis in Spain. Journal of clinical microbiology. 1999 Oct;37(10):3210-6.

41. Enright MC, Spratt BG. Multilocus sequence typing. Trends in microbiology. [Research Support, Non-U.S. Gov't Review]. 1999 Dec;7(12):482-7.

42. Feng Y, Tiao N, Li N, Hlavsa M, Xiao L. Multilocus Sequence Typing of an Emerging Cryptosporidium hominis Subtype in the United States. Journal of clinical microbiology. 2014 Feb;52(2):524-30.

43. Ferreira GE, dos Santos BN, Dorval ME, Ramos TP, Porrozzi R, Peixoto AA, Cupolillo E. The genetic structure of Leishmania infantum populations in Brazil and its possible association with the transmission cycle of visceral leishmaniasis. PloS one. [Research Support, Non-U.S. Gov't]. 2012;7(5):e36242.

44. Fraga J, Montalvo AM, De Doncker S, Dujardin JC, Van der Auwera G. Phylogeny of Leishmania species based on the heat-shock protein 70 gene. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases. [Research Support, Non-U.S. Gov't]. 2010 Mar;10(2):238-45.

45. 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. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases. [Research Support, Non-U.S. Gov't]. 2013 Aug;18:229-37.

46. Goncalves Neto VS, Barros Filho AK, Santos AM, Prazeres MP, Bezerril AC, Fonseca AV, Rebelo JM. An analysis of the spatiotemporal distribution of American cutaneous leishmaniasis in counties located along road and railway corridors in the State of Maranhao, Brazil. Revista da Sociedade Brasileira de Medicina Tropical. 2013 May-Jun;46(3):322-8.

47. Graca GC, Volpini AC, Romero GA, Oliveira Neto MP, Hueb M, Porrozzi R, Boite MC, Cupolillo E. Development and validation of PCR-based assays for diagnosis of American cutaneous leishmaniasis and identification of the parasite species. Memorias do Instituto Oswaldo Cruz. 2012 Aug;107(5):664-74.

48. Grimaldi G, Jr., Tesh RB, McMahon-Pratt D. A review of the geographic distribution and epidemiology of leishmaniasis in the New World. The American journal of tropical medicine and hygiene. [Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, P.H.S. Review]. 1989 Dec;41(6):687-725.

49. Grimaldi G, Jr., Tesh RB. Leishmaniases of the New World: current concepts and implications for future research. Clinical microbiology reviews. [Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, P.H.S. Review]. 1993 Jul;6(3):230-50.

50. Guerbouj S, Guizani I, Speybroeck N, Le Ray D, Dujardin JC. Genomic polymorphism of Leishmania infantum: a relationship with clinical pleomorphism? Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases. [Research Support, Non-U.S. Gov't]. 2001 Jul;1(1):49-59.

51. Guhl F, Ramirez JD. Retrospective molecular integrated epidemiology of Chagas disease in Colombia. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases. [Research Support, Non-U.S. Gov't]. 2013 Dec;20:148-54.

52. Huson DH, Bryant D. Application of phylogenetic networks in evolutionary studies. Mol Biol Evol. 2006 Feb;23(2):254-67.

53. 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. Annals of tropical medicine and parasitology. 2002 Apr;96(3):265-70.

54. Jamjoom MB, Ashford RW, Bates PA, Chance ML, Kemp SJ, Watts PC, Noyes HA. Leishmania donovani is the only cause of visceral leishmaniasis in East Africa; previous descriptions of L. infantum and "L. archibaldi" from this region are a consequence of convergent evolution in the isoenzyme data. Parasitology. 2004 Oct;129(Pt 4):399-409.

55. Karagiannis-Voules DA, Scholte RG, Guimaraes LH, Utzinger J, Vounatsou P. Bayesian geostatistical modeling of leishmaniasis incidence in Brazil. PLoS Negl Trop Dis. [Research Support, Non-U.S. Gov't]. 2013;7(5):e2213.

56. Kazemi B. Genomic organization of leishmania species. Iran J Parasitol. 2011 Aug;6(3):1-18.

57. Killick-Kendrick WPaR. The Leishmaniasis in Biology and Medicine. London: Academic Press inc.; 1987.

58. Klaassen CH. MLST versus microsatellites for typing Aspergillus fumigatus isolates. Med Mycol. 2009;47 Suppl 1:S27-33.

59. Kuhls K, Alam MZ, Cupolillo E, Ferreira GE, Mauricio IL, Oddone R, Feliciangeli MD, Wirth T, Miles MA, Schonian G. Comparative microsatellite typing of new world leishmania infantum reveals low heterogeneity among populations and its recent old world origin. PLoS Negl Trop Dis. [Comparative Study Research Support, Non-U.S. Gov't]. 2011 Jun;5(6):e1155.

60. Kuhls K, Cupolillo E, Silva SO, Schweynoch C, Cortes Boite M, Mello MN, Mauricio I, Miles M, Wirth T, Schonian G. Population Structure and Evidence for Both Clonality and Recombination among Brazilian Strains of the Subgenus Leishmania (Viannia). PLoS Negl Trop Dis. 2013;7(10):e2490.

61. L'Ollivier C, Labruere C, Jebrane A, Bougnoux ME, d'Enfert C, Bonnin A, Dalle F. Using a Multi-Locus Microsatellite Typing method improved phylogenetic distribution of Candida albicans isolates but failed to demonstrate association of some genotype with the commensal or clinical origin of the isolates. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases. [Research Support, Non-U.S. Gov't]. 2012 Dec;12(8):1949-57.

62. Lai Y, Shinde D, Arnheim N, Sun F. The mutation process of microsatellites during the polymerase chain reaction. J Comput Biol. 2003;10(2):143-55.

63. Lainson R, Shaw JJ. Leishmaniasis of the New World: taxonomic problems. Br Med Bull. 1972 Jan;28(1):44-8.

64. Lainson R, Shaw JJ, Silveira FT, Braga RR. American visceral leishmaniasis: on the origin of Leishmania (Leishmania) chagasi. Trans R Soc Trop Med Hyg. 1987;81(3):517.

65. Laurent T, Rijal S, Yardley V, Croft S, De Doncker S, Decuypere S, Khanal B, Singh R, Schonian G, Kuhls K, Chappuis F, Dujardin JC. Epidemiological dynamics of antimonial resistance in Leishmania donovani: genotyping reveals a polyclonal population structure among naturally-resistant clinical isolates from Nepal. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases. 2007 Mar;7(2):206-12.

66. Lauthier JJ, Tomasini N, Barnabe C, Rumi MM, D'Amato AM, Ragone PG, Yeo M, Lewis MD, Llewellyn MS, Basombrio MA, Miles MA, Tibayrenc M, Diosque P. Candidate targets for Multilocus Sequence Typing of Trypanosoma cruzi: validation using parasite stocks from the Chaco Region and a set of reference strains. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases. 2012 Mar;12(2):350-8.

67. Leblois R, Kuhls K, Francois O, Schonian G, Wirth T. Guns, germs and dogs: On the origin of *Leishmania chagasi*. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases. 2011 Jul;11(5):1091-5.

68. Lima FM, Souza RT, Santori FR, Santos MF, Cortez DR, Barros RM, Cano MI, Valadares HM, Macedo AM, Mortara RA, da Silveira JF. Interclonal variations in the molecular karyotype of Trypanosoma cruzi: chromosome rearrangements in a single cell-derived clone of the G strain. PloS one. [Research Support, Non-U.S. Gov't]. 2013;8(5):e63738.

69. Lucas CM, Franke ED, Cachay MI, Tejada A, Cruz ME, Kreutzer RD, Barker DC, McCann SH, Watts DM. Geographic distribution and clinical description of leishmaniasis cases in Peru. The American journal of tropical medicine and hygiene. [Research Support, U.S. Gov't, Non-P.H.S.]. 1998 Aug;59(2):312-7.

70. Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou J, Zurth K, Caugant DA, Feavers IM, Achtman M, Spratt BG. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci U S A. [Comparative Study
Research Support, Non-U.S. Gov't]. 1998 Mar 17;95(6):3140-5.

71. Maiden MC, van Rensburg MJ, Bray JE, Earle SG, Ford SA, Jolley KA, McCarthy ND. MLST revisited: the gene-by-gene approach to bacterial genomics. Nature reviews Microbiology. [Research Support, Non-U.S. Gov't]. 2013 Oct;11(10):728-36.

72. Maltezou HC. Drug resistance in visceral leishmaniasis. Journal of biomedicine & biotechnology. [Review]. 2010;2010:617521.

73. Mannaert A, Downing T, Imamura H, Dujardin JC. Adaptive mechanisms in pathogens: universal aneuploidy in Leishmania. Trends Parasitol. 2012 Sep;28(9):370-6.

74. Marlow MA, da Silva Mattos M, Makowiecky ME, Eger I, Rossetto AL, Grisard EC, Steindel M. Divergent profile of emerging cutaneous leishmaniasis in subtropical Brazil: new endemic areas in the southern frontier. PloS one. [Research Support, Non-U.S. Gov't]. 2013;8(2):e56177.

75. Mauricio IL, Gaunt MW, Stothard JR, Miles MA. Genetic typing and phylogeny of the Leishmania donovani complex by restriction analysis of PCR amplified gp63 intergenic regions. Parasitology. 2001 Apr;122(Pt 4):393-403.

76. Mauricio IL, Yeo M, Baghaei M, Doto D, Pratlong F, Zemanova E, Dedet JP, Lukes J, Miles MA. Towards multilocus sequence typing of the *Leishmania donovani* complex: resolving genotypes and haplotypes for five polymorphic metabolic enzymes (ASAT, GPI, NH1, NH2, PGD). Int J Parasitol. 2006 Jun;36(7):757-69.

77. McGwire BS, Satoskar AR. Leishmaniasis: clinical syndromes and treatment. QJM : monthly journal of the Association of Physicians. 2014 Jan;107(1):7-14.

78. Minning TA, Weatherly DB, Flibotte S, Tarleton RL. Widespread, focal copy number variations (CNV) and whole chromosome aneuploidies in Trypanosoma cruzi strains revealed by array comparative genomic hybridization. BMC genomics. [Research Support, N.I.H., Extramural]. 2011;12:139.

79. Monge-Maillo B, Lopez-Velez R. Therapeutic options for old world cutaneous leishmaniasis and new world cutaneous and mucocutaneous leishmaniasis. Drugs. [Research Support, Non-U.S. Gov't]. 2013 Nov;73(17):1889-920.

80. Monge-Maillo B, Perez-Molina JA, Norman FF, Lopez-Velez R. Concerns about topical treatment for new world cutaneous leishmanisis. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. [Comment Letter Research Support, Non-U.S. Gov't]. 2013 Nov;57(10):1502-3.

81. Motoie G, Ferreira GE, Cupolillo E, Canavez F, Pereira-Chioccola VL. Spatial distribution and population genetics of Leishmania infantum genotypes in Sao Paulo State, Brazil, employing multilocus microsatellite typing directly in dog infected tissues. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases. 2013 May 9.

82. Noyes HA, Chance ML, Croan DG, Ellis JT. Leishmania (sauroleishmania): a comment on classification. Parasitol Today. 1998 Apr;14(4):167.

83. 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. Journal of clinical microbiology. [Research Support, Non-U.S. Gov't]. 2009 Sep;47(9):2818-25.

84. Odds FC, Jacobsen MD. Multilocus sequence typing of pathogenic Candida species. Eukaryot Cell. 2008 Jul;7(7):1075-84.

85. Odiwuor S, Vuylsteke M, De Doncker S, Maes I, Mbuchi M, Dujardin JC, Van der Auwera G. Leishmania AFLP: paving the way towards improved molecular assays and markers of diversity. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases. 2011 Jul;11(5):960-7.

86. Pacheco RS, Martinez JE, Valderrama L, Momen H, Saravia NG. Genotypic polymorphisms in experimental metastatic dermal leishmaniasis. Mol Biochem Parasitol. 1995 Feb;69(2):197-209.

87. Pallen MJ, Loman NJ, Penn CW. High-throughput sequencing and clinical microbiology: progress, opportunities and challenges. Current opinion in microbiology. [Research Support, Non-U.S. Gov't Review]. 2010 Oct;13(5):625-31.

88. Peacock CS, Seeger K, Harris D, Murphy L, Ruiz JC, Quail MA, Peters N, Adlem E, Tivey A, Aslett M, Kerhornou A, Ivens A, Fraser A, Rajandream MA, Carver T, Norbertczak H, Chillingworth T, Hance Z, Jagels K, Moule S, Ormond D, Rutter S, Squares R, Whitehead S, Rabbinowitsch E, Arrowsmith C, White B, Thurston S, Bringaud F, Baldauf SL, Faulconbridge A, Jeffares D, Depledge DP, Oyola SO, Hilley JD, Brito LO, Tosi LR, Barrell B, Cruz AK, Mottram JC, Smith DF, Berriman M. Comparative genomic analysis of three Leishmania species that cause diverse human disease. Nature genetics. [Comparative Study Research Support, Non-U.S. Gov't]. 2007 Jul;39(7):839-47.

89. Perez-Losada M, Cabezas P, Castro-Nallar E, Crandall KA. Pathogen typing in the genomics era: MLST and the future of molecular epidemiology. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases. [Research Support, Non-U.S. Gov't]. 2013 Jun;16:38-53.

90. Polakova S, Blume C, Zarate JA, Mentel M, Jorck-Ramberg D, Stenderup J, Piskur J. Formation of new chromosomes as a virulence mechanism in yeast Candida glabrata. Proc Natl Acad Sci U S A. [Research Support, Non-U.S. Gov't]. 2009 Feb 24;106(8):2688-93.

91. Ramirez JD, Tapia-Calle G, Guhl F. Genetic structure of Trypanosoma cruzi in Colombia revealed by a High-throughput Nuclear Multilocus Sequence Typing (nMLST) approach. BMC genetics. [Research Support, Non-U.S. Gov't]. 2013;14:96.

92. 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):111-25.

93. Rogers MB, Hilley JD, Dickens NJ, Wilkes J, Bates PA, Depledge DP, Harris D, Her Y, Herzyk P, Imamura H, Otto TD, Sanders M, Seeger K, Dujardin JC, Berriman M, Smith DF, Hertz-Fowler C, Mottram JC. Chromosome and gene copy number variation allow major structural change between species and strains of Leishmania. Genome Res. 2011 Dec;21(12):2129-42.

94. Rossi V, Wincker P, Ravel C, Blaineau C, Pages M, Bastien P. Structural organisation of microsatellite families in the Leishmania genome and polymorphisms at two (CA)n loci. Mol Biochem Parasitol. 1994 Jun;65(2):271-82.

95. Rougeron V, Waleckx E, Hide M, T DEM, Arevalo J, Llanos-Cuentas A, Banuls AL. PERMANENT GENETIC RESOURCES: A set of 12 microsatellite loci for genetic studies of *Leishmania braziliensis*. Mol Ecol Resour. 2008 Mar;8(2):351-3.

96. Russell R, Iribar MP, Lambson B, Brewster S, Blackwell JM, Dye C, Ajioka JW. Intra and inter-specific microsatellite variation in the Leishmania subgenus Viannia. Mol Biochem Parasitol. 1999 Sep 20;103(1):71-7.

97. Santos MF, Ribolla PE, Alonso DP, Andrade-Filho JD, Casaril AE, Ferreira AM, Fernandes CE, Brazil RP, Oliveira AG. Genetic structure of Lutzomyia longipalpis populations in Mato Grosso Do Sul, Brazil, based on microsatellite markers. PloS one. 2013;8(9):e74268.

98. Scholler JK, Reed SG, Stuart K. Molecular karyotype of species and subspecies of Leishmania. Mol Biochem Parasitol. [Comparative Study
Research Support, U.S. Gov't, P.H.S.]. 1986 Sep;20(3):279-93.

99. Schonian G, Mauricio I, Gramiccia M, Canavate C, Boelaert M, Dujardin JC. Leishmaniases in the Mediterranean in the era of molecular epidemiology. Trends Parasitol. 2008 Mar;24(3):135-42.

100. Schonian G, Mauricio I, Cupolillo E. Is it time to revise the nomenclature of Leishmania? Trends Parasitol. 2010 Oct;26(10):466-9.

101. Schonian G, Kuhls K, Mauricio IL. Molecular approaches for a better understanding of the epidemiology and population genetics of Leishmania. Parasitology. [Review]. 2011 Apr;138(4):405-25.

102. Schriefer A, Guimaraes LH, Machado PR, Lessa M, Lessa HA, Lago E, Ritt G, Goes-Neto A, Schriefer AL, Riley LW, Carvalho EM. Geographic clustering of leishmaniasis in northeastern Brazil. Emerging infectious diseases. [Research Support, N.I.H., Extramural]. 2009 Jun;15(6):871-6.

103. Spanakos G, Alifrangis M, Schousboe ML, Patsoula E, Tegos N, Hansson HH, Bygbjerg IC, Vakalis NC, Tseroni M, Kremastinou J, Hadjichristodoulou C. Genotyping Plasmodium vivax isolates from the 2011 outbreak in Greece. Malaria journal. [Research Support, Non-U.S. Gov't]. 2013;12:463.

104. Sterkers Y, Lachaud L, Crobu L, Bastien P, Pages M. FISH analysis reveals aneuploidy and continual generation of chromosomal mosaicism in Leishmania major. Cell Microbiol. 2011 Feb;13(2):274-83.

105. Sterkers Y, Lachaud L, Bourgeois N, Crobu L, Bastien P, Pages M. Novel insights into genome plasticity in Eukaryotes: mosaic aneuploidy in Leishmania. Molecular microbiology. [Review]. 2012 Oct;86(1):15-23.

106. Su C, Khan A, Zhou P, Majumdar D, Ajzenberg D, Darde ML, Zhu XQ, Ajioka JW, Rosenthal BM, Dubey JP, Sibley LD. Globally diverse Toxoplasma gondii isolates comprise six major clades originating from a small number of distinct ancestral lineages. Proc Natl Acad Sci U S A. [Research Support, N.I.H., Extramural]. 2012 Apr 10;109(15):5844-9.

107. Thomas JC, Robinson DA. Multilocus sequence typing of Staphylococcus epidermidis. Methods Mol Biol. 2014;1106:61-9.

108. Tibayrenc M, Kjellberg F, Ayala FJ. A clonal theory of parasitic protozoa: the population structures of *Entamoeba*, *Giardia*, *Leishmania*, *Naegleria*, *Plasmodium*, *Trichomonas*, and *Trypanosoma* and their medical and taxonomical consequences. Proc Natl Acad Sci U S A. 1990 Apr;87(7):2414-8.

109. Tibayrenc M, Ayala FJ. Evolutionary genetics of Trypanosoma and Leishmania. Microbes and infection / Institut Pasteur. [Review]. 1999 May;1(6):465-72.

110. Tibayrenc M, Ayala FJ. Reproductive clonality of pathogens: a perspective on pathogenic viruses, bacteria, fungi, and parasitic protozoa. Proc Natl Acad Sci U S A. 2012 Nov 27;109(48):E3305-13.

111. Tibayrenc M, Ayala FJ. How clonal are Trypanosoma and Leishmania? Trends Parasitol. 2013 Jun;29(6):264-9.

112. Toth G, Gaspari Z, Jurka J. Microsatellites in different eukaryotic genomes: survey and analysis. Genome Res. [Research Support, Non-U.S. Gov't]. 2000 Jul;10(7):967-81.

113. Tsukayama P, Lucas C, Bacon DJ. Typing of four genetic loci discriminates among closely related species of New World Leishmania. IntJ Parasitol. 2009;39(3):355-62.

114. Ubeda JM, Legare D, Raymond F, Ouameur AA, Boisvert S, Rigault P, Corbeil J, Tremblay MJ, Olivier M, Papadopoulou B, Ouellette M. Modulation of gene expression in drug resistant *Leishmania* is associated with gene amplification, gene deletion and chromosome aneuploidy. Genome Biol. 2008;9(7):R115.

115. van Belkum A, Struelens M, de Visser A, Verbrugh H, Tibayrenc M. Role of genomic typing in taxonomy, evolutionary genetics, and microbial epidemiology. Clinical microbiology reviews. 2001 Jul;14(3):547-60.

116. WHO. Control of the leishmaniases: report of a meeting of the WHO Expert Commitee on the Control of Leishmaniases. Geneva: World Health Organization2010 22-26 March 2010.

117. Williams BR, Prabhu VR, Hunter KE, Glazier CM, Whittaker CA, Housman DE, Amon A. Aneuploidy affects proliferation and spontaneous immortalization in mammalian cells. Science. 2008 Oct 31;322(5902):703-9.

118. Wilson EO. Systematics and the future of biology. Proc Natl Acad Sci U S A. 2005 May 3;102 Suppl 1:6520-1.

119. Yasinzai M, Khan M, Nadhman A, Shahnaz G. Drug resistance in leishmaniasis: current drug-delivery systems and future perspectives. Future medicinal chemistry. 2013 Oct;5(15):1877-88.

120. Zemanova E, Jirku M, Mauricio IL, Horak A, Miles MA, Lukes J. The Leishmania donovani complex: genotypes of five metabolic enzymes (ICD, ME, MPI, G6PDH, and FH), new targets for multilocus sequence typing. IntJ Parasitol. 2007;37(2):149-60.

121. Zhang CY, Lu XJ, Du XQ, Jian J, Shu L, Ma Y. Phylogenetic and evolutionary analysis of Chinese Leishmania isolates based on multilocus sequence typing. PloS one. 2013;8(4):e63124.

Anexo 1. Cepas utilizadas nos trabalhos apresentados nos capítulos I, II e IV com respectivos dados obtidos.

Espécie	IOCL	Código internacional	Região	Perfil MLMT	DST	Pop. K3[2]	Sub-pop.[2]	CC
L. braziliensis	18	MHOM/BR/1982/RPF	Rio de Janeiro	Lbra24	x	2	2B	x
L. braziliensis	35	MHOM/BR/1980/3309	Rio de Janeiro	Lbra23	x	2	2B	x
L. braziliensis	326	MHOM/BR/1983/HCM	Pará	Lbra34	x	3	3A	x
L. guyanensis	565	MHOM/BR/1975/M4147	Pará	x	37	x	**x**	CC6
L. braziliensis	566	MHOM/BR/1975/M2903^{T}	Pará	Lbra49	61	3	3B	CC4
L. naiffi	847	MHOM/BR/1987/AXEL	Pará	Lnai5	x	3	3D	x
L. braziliensis	849	MHOM/BR/1987/J.CARLOS	Rio de Janeiro	x	12	x	**x**	CC4
L. naiffi	854	ISQU/BR/1985/IM2264	Pará	Lnai3	5	3	3D	CC2
L. naiffi	855	MHOM/BR/1986/IM2773	Amazonas	Lnai2	3	3	3D	singl
L. braziliensis	918	MRAT/BR/1987/C18.454	Ceará	Lbra48	63	3	3B	htz
L. braziliensis	921	MHOM/BR/1987/H-210	Ceará	Lbra48	62	3	3B	CC4
L. naiffi	991	MDAS/BR/1987/IM3307	Pará	Lnai6	x	3	3D	x
L. naiffi	992	MDAS/BR/1987/IM3280	Pará	x	70	x	**x**	singl
L. naiffi	995	MDAS/BR/1987/IM3292	Pará	Lnai4	71	3	3D	htz
L. lainsoni	1023	MHOM/BR/81/M6426^{T}	Pará	Llai1	1	3	3C	CC1
L. lainsoni	1058	MCOE/BR/1983/IM1367	Rondonia	Llai5	x	3	3C	x
L. guyanensis	1061	MCHO/BR/1985/IM2183	Amazonas	Lgua22	x	1	x	x
L. guyanensis	1064	MCHO/BR/1985/IM2285	Amazonas	Lgua23	x	1	x	x
L. guyanensis	1065	MDID/BR/1985/IM2295	Amazonas	Lgua33	x	1	x	x
L. shawi	1066	IWHI/BR/1985/IM2322	Pará	Lsha3	x	3	3A	x
L. shawi	1067	IWHI/BR/1985/IM2324	Pará	Lsha2	x	3	3A	x
L. shawi	1068	IWHI/BR/1985/IM2326	Pará	Lsha3	x	3	3A	x
L. guyanensis	1072	IAND/BR/1985/IM2513	Amazonas	Lgua34	x	1	x	x
L. guyanensis	1094	MDID/BR/1986/IM2845	Amazonas	Lgua24	x	1	x	x
L. guyanensis	1105	MDID/BR/1987/IM3188	Amazonas	Lgua35	x	1	x	x
L. lainsoni	1266	MAGO/BR/1983/IM1721	Pará	Llai6	2	3	3C	CC1
L. guyanensis	1300	MDID/BR/1987/IM3189	Amazonas	Lgua36	x	1	x	x
L. naiffi	1365	MDAS/BR/79/M5533^{T}	Pará	Lnai1	4	3	3D	htz
L. shawi	1545	MCEB/BR/84/M8408^{T}	Pará	Lsha1	8	3	3A	CC3
L. braziliensis	1599	MHOM/BR/1990/IM3690	Amazonas	Lbra52	x	3 (2/3)	3B	x
L. braziliensis	1731	MAGO/BR/1992/IM154	Rondônia	Lbra43	11	3	3A	singl
L. braziliensis	1734	MHOM/BR/1991/IM3713	Amazonas	x	26	x	**x**	htz
L. naiffi	1939	ISQU/BR/1994/IM3936	Amazonas	x	6	x	**x**	CC2
L. naiffi	1953	MHOM/BR/1994/IM4000	Amazonas	x	7	x	**x**	CC2
L. braziliensis	2139	MHOM/BR/1996/GBS	Paraíba	x	12	x	**x**	CC4
L. braziliensis	2148	MHOM/BR/2001/VLNC	Pernambuco	x	13	x	**x**	CC4
L. braziliensis	2152	MHOM/BR/1996/AFN	Rio de Janeiro	Lbra22	36	2	2B	singl
L. braziliensis	2159	MHOM/BR/1996/JPS	Rio de Janeiro	x	12	x	**x**	CC4
L. braziliensis	2211	MHOM/BR/00/LTB300	Bahia	Lbra1	x	2	2A	x
L. braziliensis	2287	MHOM/BR/1998/AFS	Pernambuco	x	12	x	**x**	CC4
L. braziliensis	2288	MHOM/BR/1996/MAS	Pernambuco	x	12	x	**x**	CC4
L. braziliensis	2291	MHOM/BR/1997/ASB	Pernambuco	x	12	x	**x**	CC4
L. guyanensis	2334	MHOM/BR/1997/202P	Amazonas	Lgua12	42	1	x	CC6
L. guyanensis	2335	MHOM/BR/1997/203P	Amazonas	Lgua13	x	1	x	x
L. guyanensis	2336	MHOM/BR/1997/203G	Amazonas	Lgua13	x	1	x	x
L. guyanensis	2337	MHOM/BR/1997/210P	Amazonas	Lgua30	x	1	x	x
L. guyanensis	2338	MHOM/BR/1997/212P	Amazonas	Lgua14	x	1	x	x
L. guyanensis	2339	MHOM/BR/1997/213-P	Amazonas	Lgua28	x	1	x	x
L. guyanensis	2340	MHOM/BR/1997/215-P	Amazonas	Lgua10	x	1	x	x
L. guyanensis	2341	MHOM/BR/1997/223P	Amazonas	Lgua15	43	1	x	CC6
L. guyanensis	2342	MHOM/BR/1997/227P	Amazonas	Lgua16	x	1	x	x
L. braziliensis	2344	MHOM/BR/1997/NMT-MAO 229P	Amazonas	x	21	x	**x**	CC4
L. guyanensis	2345	MHOM/BR/1997/230-P	Amazonas	Lgua2	x	1	x	x
L. guyanensis	2346	MHOM/BR/1997/233P	Amazonas	Lgua31	x	1	x	x
L. guyanensis	2349	MHOM/BR/1997/236P	Amazonas	Lgua17	x	1	x	x
L. guyanensis	2350	MHOM/BR/1997/237-P	Amazonas	Lgua3	44	1	x	CC6
L. guyanensis	2351	MHOM/BR/1997/240-P	Amazonas	Lgua9	x	1	x	x

L. guyanensis	2352	MHOM/BR/1997/241P	Amazonas	Lgua32	x	1	x	x
L. guyanensis	2354	MHOM/BR/1997/243-P	Amazonas	Lgua11	x	1	x	x
L. guyanensis	2357	MHOM/BR/1997/247P	Amazonas	Lgua19	x	1	x	x
L. guyanensis	2359	MHOM/BR/1997/249P	Amazonas	Lgua20	x	1	x	x
L. guyanensis	2364	MHOM/BR/1997/257-P	Amazonas	Lgua21	x	1	x	x
L. guyanensis	2372	MHOM/BR/1997/292-G	Amazonas	Lgua1	x	1	x	x
L. guyanensis	2377	MHOM/BR/1997/298-P	Amazonas	Lgua5	x	1	x	x
L. guyanensis	2386	MHOM/BR/1997/304-P	Amazonas	Lgua7	x	1	x	x
L. guyanensis	2390	MHOM/BR/1997/308-P	Amazonas	Lgua29	x	1	x	x
L. guyanensis	2403	MHOM/BR/1997/324-P	Amazonas	Lgua8	x	1	x	x
L. guyanensis	2405	MHOM/BR/1997/325-P	Amazonas	Lgua6	x	1	x	x
L. guyanensis	2406	MHOM/BR/1997/327-P	Amazonas	Lgua4	x	1	x	x
L. guyanensis	2407	MHOM/BR/1997/328-P	Amazonas	Lgua25	x	1	x	x
L. guyanensis	2410	MHOM/BR/1997/331-P	Amazonas	Lgua27	x	1	x	x
L. braziliensis	2419	MHOM/BR/1999/JAS	Pernambuco	Lbra9	12	2	2A	CC4
L. braziliensis	2420	MHOM/BR/1999/PCS	Pernambuco	Lbra18	12	2	2A	CC4
L. braziliensis	2424	MMES/BR/1999/SENTINELAI	Pernambuco	Lbra14	x	2	2A	x
L. braziliensis	2427	MHOM/BR/1999/SJB	Pernambuco	Lbra15	12	2	2A	CC4
L. braziliensis	2446	MHOM/BR/2002/JCS	Pernambuco	Lbra17	x	2 (2/3)	2A	x
L. braziliensis	2463	MHOM/BR/2001/JOLIVAL	Bahia	Lbra7	74	2	2A	htz
L. braziliensis	2466	MHOM/BR/2001/LTCP13183	Bahia	x	35	x	**x**	CC4
L. braziliensis	2467	MHOM/BR/2001/LTCP14182	Bahia	Lbra6	12	2	2A	CC4
L. braziliensis	2468	MHOM/BR/2001/LTCP14183	Bahia	x	33	x	**x**	htz
L. braziliensis	2468	MHOM/BR/2001/LTCP14183	Bahia	Lbra3	x	2	2A	x
L. braziliensis	2472	MHOM/BR/2001/LTCP13455	Bahia	x	12	x	**x**	CC4
L. braziliensis	2475	MHOM/BR/2001/LTCP14214	Bahia	Lbra4	12	2	2A	CC4
L. braziliensis	2476	MHOM/BR/2001/LTCP14278	Bahia	Lbra6	35	2	2A	CC4
L. braziliensis	2480	MHOM/BR/2001/LTCP13980	Bahia	x	12	x	**x**	CC4
L. braziliensis	2481	MHOM/BR/2000/LTCP13490	Bahia	x	14	x	**x**	CC4
L. braziliensis	2483	MHOM/BR/2000/LTCP13396	Bahia	Lbra2	x	2	2A	x
L. naiffi/L. lainsoni	2490	MHOM/BR/2002/NMT-RBO004	Acre (Rio Branco)	Llai/Lnai1	54	3	3C	singl
L. braziliensis	2491	MHOM/BR/2002/NMT-RBO005	Acre	Lbra35	19	3	3A	singl
L. braziliensis	2492	MHOM/BR/2002/NMT-RBO011	Acre	Lbra36	x	3	3A/D	x
L. guyanensis	2493	MHOM/BR/2002/NMT-RBO013	Acre	Lgua26	55	3	3A	singl
L. braziliensis	2494	MHOM/BR/2002/NMT-RBO018	Acre	Lbra37	25	3	3A	htz
L. braziliensis	2495	MHOM/BR/2002/NMT-RBO025	Acre	Lbra38	23	3	3A	CC4
L. lainsoni	2497	MHOM/BR/2002/NMT-RBO 027P	Acre	Llai2	56	3	3C	htz
L. braziliensis	2498	MHOM/BR/2002/NMT-RBO029	Acre	Lbra39	57	3	3A/D	singl
L. braziliensis	2499	MHOM/BR/2002/NMT-RBO035	Acre	Lbra41	32	3	3A	htz
L. lainsoni	2500	MHOM/BR/2002/NMT-RBO036	Acre	Llai3	58	3	3C	singl
L. braziliensis	2501	MHOM/BR/2002/NMT-RBO037	Acre	Lbra42	59	3	3A	singl
L. braziliensis	2502	MHOM/BR/2002/NMT-RBO040	Acre	Lbra40	15	3	3A	CC4
L. lainsoni	2503	MHOM/BR/2002/NMT-RBO044	Acre	Llai4	60	3	3C	singl
L. braziliensis	2509	MHOM/BR/2001/CRFN	Pernambuco	Lbra19	27	2	2A	CC4
L. braziliensis	2510	MHOM/BR/2000/CEA	Pernambuco	x	12	x	**x**	CC4
L. braziliensis	2511	MHOM/BR/2001/JS	Pernambuco	Lbra13	31	2	2A	htz
L. braziliensis	2513	MHOM/BR/2000/LMG	Pernambuco	Lbra10	x	2	2A	x
L. braziliensis	2515	MHOM/BR/2001/TSS	Pernambuco	Lbra11	12	2	2A	CC4
L. braziliensis	2516	MHOM/BR/2001/WJS	Pernambuco	Lbra12	x	2 (2/3)	2A	x
L. braziliensis	2535	MHOM/BR/2001/HC-JS	Espírito Santo	x	20	x	**x**	htz
L. braziliensis	2538	MHOM/BR/2002/EMM	Rio de Janeiro	x	64	x	**x**	htz
L. braziliensis	2541	MHOM/BR/1999/MJAA-II	Pernambuco	x	65	x	**x**	htz
L. braziliensis	2547	MNEC/BR/2002/CPqAM-191	Pernambuco	Lbra16	x	2	2A	x
L. braziliensis	2571	MHOM/BR/2003/NJS	Mato Grosso do Sul	x	75	x	**x**	htz
L. braziliensis	2660	MHOM/BR/2004/ARARAQUARA-1	São Pauo	x	12	x	**x**	CC4
L. utingensis	2689	ITUB/BR/1977/4964^{T}	Pará	Luti1	66	3	3A	singl
L. lindenbergi	2690	MHOM/BR/1966/M15733^{T}	Pará	Llin1	67	3	3A	singl
L. braziliensis	2693	MNEC/BR/2003/NECTOMYS	Pernambuco	x	22	x	**x**	CC4
L. braziliensis	2832	MHOM/BR/1995/LTCP9783	Bahia	Lbra8	x	2	2A	x
L. braziliensis	2833	MHOM/BR/2001/LTCP14349	Bahia	x	29	x	**x**	htz
L. braziliensis	2836	MHOM/BR/21994/LTCP9845	Bahia	Lbra5	12	2	2A	CC4
L. braziliensis	2838	MHOM/BR/1997/LTCP11245	Bahia	x	12	x	**x**	CC4

L. braziliensis	2847	MHOM/BR/2002/LTCP15476	Bahia	x	35	x	**x**	CC4
L. braziliensis	2923	MHOM/BR/2006/AASR	Pernambuco	Lbra44	x	3	3B	x
L. braziliensis	2929	MHOM/BR/2005/NMT-LTCP16011-P	Bahia	x	27	x	**x**	CC4
L. braziliensis	2932	MHOM/BR/2006/HBO	Pernambuco	Lbra45	x	3	3B	x
L. guyanensis	2936	MHOM/BR/2007/069	Amazonas	x	46	x	**x**	CC6
L. guyanensis	2937	MHOM/BR/2007/065	Amazonas	x	39	x	**x**	CC6
L. guyanensis	2938	MHOM/BR/2007/063	Amazonas	x	47	x	**x**	htz
L. braziliensis	2950	MHOM/BR/2006/CEN	Pernambuco	x	24	x	**x**	CC5
L. braziliensis	2951	MHOM/BR/2006/CM	Pernambuco	x	30	x	**x**	CC5
L. braziliensis	2953	MHOM/BR/2006/MVRGR	Pernambuco	Lbra46	x	3	3B	x
L. guyanensis	2956	MHOM/BR/2007/011	Amazonas	x	48	x	**x**	CC6
L. guyanensis	2957	MHOM/BR/2007/014-JIS	Amazonas	x	49	x	**x**	htz
L. guyanensis	2960	MHOM/BR/2007/029-ZAV	Amazonas	x	68	x	**x**	singl
L. guyanensis	2961	MHOM/BR/2007/019-WDSN	Amazonas	x	50	x	**x**	singl
L. guyanensis	2962	MHOM/BR/2007/031-LOP	Amazonas	x	51	x	**x**	singl
L. guyanensis	2963	MHOM/BR/2007/033-MECM	Amazonas	x	38	x	**x**	CC6
L. guyanensis	2964	MHOM/BR/2007/034-MFPS	Amazonas	x	69	x	**x**	singl
L. guyanensis	2966	MHOM/BR/2007/039	Amazonas	x	52	x	**x**	htz
L. guyanensis	2969	MHOM/BR/2007/AC	Amazonas	x	41	x	**x**	CC6
L. guyanensis	2970	MHOM/BR/2007/021-HMB	Amazonas	x	45	x	**x**	CC6
L. guyanensis	2971	MHOM/BR/2007/021-R	Amazonas	x	40	x	**x**	CC6
L. naiffi	3007	MHOM/BR/2003/IRCF	Amazonas	Lnai7	x	3	3D	x
L. braziliensis	3027	MCAN/BR/2007/HG-04	Espírito Santo	Lbra20	x	2	2A/B	x
L. braziliensis	3061	MHOM/BR/2008/OPS	Rio de Janeiro	Lbra21	x	2	2B	x
L. braziliensis	3070	MCAN/BR/2008/CBG-15	Espírito Santo	Lbra22	x	2	2B	x
L. shawi	3199	MHOM/BR/1999/M17997	Pará	x	9	x	**x**	CC3
L. shawi	3200	MHOM/BR/1999/M17998	Pará	x	10	x	**x**	CC3
L. braziliensis	BH118	MHOM/BR/1991/BH118	Minas Gerais	Lbra26	x	2	2B	x
L. braziliensis	BH123	MHOM/BR/1981/BH123	Minas Gerais	Lbra28	x	2	2B	x
L. braziliensis	BH141	MHOM/BR/1986/BH141	Minas Gerais	Lbra25	x	2	2B	x
L. braziliensis	BH17	MHOM/BR/1988/BH17	Minas Gerais	Lbra26	x	2	2B	x
L. braziliensis	BH186	MHOM/BR/1988/BH186	Minas Gerais	Lbra27	x	2	2B	x
L. braziliensis	BH450	MHOM/BR/1991/BH450	Minas Gerais	Lbra26	x	2	2B	x
L. braziliensis	BH451	MHOM/BR/1991/BH451	Minas Gerais	Lbra26	x	2	2B	x
L. braziliensis	BH452	MHOM/BR/1991/BH452	Minas Gerais	Lbra29	x	2	2B	x
L. braziliensis	BH453	MHOM/BR/1991/BH453	Minas Gerais	Lbra26	x	2	2B	x
L. braziliensis	BH454	MHOM/BR/1991/BH454	Minas Gerais	Lbra30	x	2	2B	x
L. braziliensis	BH455	MHOM/BR/1991/BH455	Minas Gerais	Lbra26	x	2	2B	x
L. braziliensis	BH456	MHOM/BR/1992/BH456	Minas Gerais	Lbra31	x	2	2B	x
L. braziliensis	BH457	MHOM/BR/1992/BH457	Minas Gerais	Lbra26	x	2	2B	x
L. braziliensis	BH458	MHOM/BR/1992/BH458	Minas Gerais	Lbra26	x	2	2B	x
L. braziliensis	BH459	MHOM/BR/1992/BH459	Minas Gerais	Lbra26	x	2	2B	x
L. guyanensis	iz 18	MCHO/BR/80/M6202	*	x	34	x	**x**	CC6
L. lainsoni	iz 25	MHOM/BR/82/M6887	Pará	x	73	x	**x**	htz
L. braziliensis	iz 26	MORY/PE/84/AO23	-	x	16	x	**x**	htz
L. braziliensis	iz 27	MORY/PE/84/AO8	-	x	18	x	**x**	htz
L. braziliensis	iz 28	MRAT/PE/84/A1	-	x	16	x	**x**	htz
L. braziliensis	iz 31	MORY/PE/84/AC20	-	x	17	x	**x**	CC4
L. braziliensis	iz 33	MORY/PE/84/ABR23	-	x	28	x	**x**	htz
L. guyanensis	iz 34	MCHO/BR/80/M6200	*	x	53	x	**x**	htz
L. lainsoni	iz 35	IUBI/BR/00/M12025	Pará	x	72	x	**x**	singl
L. guyanensis	2355	MHOM/BR/1997/245P	Amazonas	Lgua18	x	1	x	x
L. braziliensis	LOA	MHOM/BR/1993/709-LOA	Paraná	Lbra32	x	2	2B	x
L. braziliensis	M8401	IPSY/BR/1981/M8401	Pará	Lbra50	x	3 (2/3)	3B	x
L. braziliensis	M8471	MHOM/BR/1984/M8471	Pará	Lbra51	x	3 (2/3)	3B	x
L. braziliensis	RR1	MHOM/BR/2000/RR1	Paraná	Lbra47	x	3 (2/3)	3B	x
L. braziliensis	VU	MHOM/BR/1993/769-VU	Paraná	Lbra33	x	2	2B	x

Htz=heterozigoto; CC=complex clonal; POP=população; DST=*diploid sequence type*; singl=*singleton*.