



Full length article

Standardization of DNA extraction from sand flies: Application to genotyping by next generation sequencing



Aline Etelvina Casaril ^{a,b}, Liliane Prado de Oliveira ^c, Diego Peres Alonso ^d,
 Everton Falcão de Oliveira ^e, Suellem Petilim Gomes Barrios ^a,
 Jucelei de Oliveira Moura Infran ^b, Wagner de Souza Fernandes ^a, Elisa Teruya Oshiro ^b,
 Alda Maria Teixeira Ferreira ^{a,f}, Paulo Eduardo Martins Ribolla ^d,
 Alessandra Gutierrez de Oliveira ^{a,b,*}

^a Programa de Pós-Graduação em Doenças Infecciosas e Parasitárias, Faculdade de Medicina, Universidade Federal de Mato Grosso do Sul, Av. Costa e Silva s/n°, Campus Universitário, Campo Grande, MS, Brazil

^b Laboratório de Parasitologia Humana, Universidade Federal de Mato Grosso do Sul, Campus Universitário s/n, 79070-900 Campo Grande, MS, Brazil

^c Curso de Ciências Biológicas, Universidade Federal de Mato Grosso do Sul, Campus Universitário s/n, 79070-900 Campo Grande, MS, Brazil

^d Laboratório de Pesquisas e Análises Genéticas, Universidade Estadual Paulista, Júlio de Mesquita Filho, Campus de Botucatu, Distrito de Rubião Junior s/n, 18618-970 Botucatu, SP, Brazil

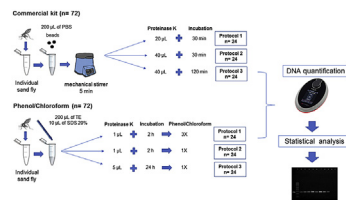
^e Departamento de Estatística, Universidade Federal do Rio de Janeiro, Campus Macaé, Macaé, RJ, Brazil

^f Laboratório de Imunologia Básica, Universidade Federal de Mato Grosso do Sul, Campus Universitário s/n, 79070-900 Campo Grande, MS, Brazil

HIGHLIGHTS

- Three protocols developed for the commercial kit were found to be more practical.
- Phenol/chloroform protocols resulted in significantly higher quantities of sand fly gDNA.
- Higher quantities of proteinase K resulted in higher concentrations of extracted sand fly DNA.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 16 November 2016

Received in revised form

30 March 2017

Accepted 25 April 2017

Available online 26 April 2017

Keywords:

Vector

Visceral leishmaniasis

Molecular biology

ABSTRACT

Standardization of the methods for extraction of DNA from sand flies is essential for obtaining high efficiency during subsequent molecular analyses, such as the new sequencing methods. Information obtained using these methods may contribute substantially to taxonomic, evolutionary, and eco-epidemiological studies. The aim of the present study was to standardize and compare two methods for the extraction of genomic DNA from sand flies for obtaining DNA in sufficient quantities for next-generation sequencing. Sand flies were collected from the municipalities of Campo Grande, Camapuã, Corumbá and Miranda, state of Mato Grosso do Sul, Brazil. Three protocols using a silica column-based commercial kit (*ReliaPrep™ Blood gDNA Miniprep System kit*, Promega®), and three protocols based on the classical phenol-chloroform extraction method (Uliana et al., 1991), were compared with respect to the yield and quality of the extracted DNA. DNA was quantified using a Qubit 2.0 fluorometer. The presence of sand fly DNA was confirmed by PCR amplification of the IVS6 region (constitutive gene), followed by electrophoresis on a 1.5% agarose gel. A total of 144 male specimens were analyzed, 72 per method. Significant differences were observed between the two methods tested. Protocols 2 and 3 of phenol-chloroform extraction presented significantly better performance than all commercial kit extraction protocols tested. For phenol-chloroform extraction, protocol 3 presented significantly better

* Corresponding author. Programa de Pós-Graduação em Doenças Infecciosas e Parasitárias, Faculdade de Medicina, Universidade Federal de Mato Grosso do Sul, Av. Costa e Silva s/n°, Campus Universitário, Campo Grande, MS, Brazil.

E-mail address: alessandra.oliveira@ufms.br (A.G. de Oliveira).

performance than protocols 1 and 2. The IVS6 region was detected in 70 of 72 (97.22%) samples extracted with phenol, including all samples for protocols 2 and 3. This is the first study on the standardization of methods for the extraction of DNA from sand flies for application to next-generation sequencing, which is a promising tool for entomological and molecular studies of sand flies.

© 2017 Elsevier Inc. All rights reserved.

1. Introduction

Sand flies (Diptera:Psychodidae) are vectors of several human and animal pathogens, such as *Bartonella bacilliformis*, viruses of the families Bunyaviridae and Rhabdoviridae, and protozoa of the genera *Endotrypanum* and *Leishmania* (Adler and Theodor, 1957; Lainson and Shaw, 2005; Lutz and Neiva, 1912; Young and Duncan, 1994).

Cryptic sand fly species belonging to the species complexes *longipalpis*, *cortezii*, and *intermedia* present few distinguishable taxonomic characters, making their identification difficult, especially when living in sympatry (Bauzer et al., 2007; Galati et al., 1989; Marcondes, 1996). This may affect the identification of the species and of their parasite-vector-host relationship for a given area. Molecular tools can be used to help taxonomic identification of sand flies.

The number of studies on the use of molecular markers, based on mitochondrial DNA (Folmer et al., 1994; Zhang and Hewitt, 1997), microsatellites (Queller et al., 1993; Santos et al., 2013) and, more recently, Single Nucleotide Polymorphism (SNPs) (Brumfield et al., 2003; Soria-Carrasco et al., 2014; Vignal et al., 2002), for the identification and evaluation of intra- and inter-population diversity is increasing. Genotyping by traditional Sanger sequencing has also been used for the analysis of target regions but this approach proves costly for population genetics studies. Novel genotyping techniques termed Next-Generation Sequencing (Illumina, Roche 454, and AB SOLiD platforms), which can generate and analyze thousands of DNA fragments in short intervals of time, can be used to sequence dipteran genomes in studies with limited financial resources (Davey and Blaxter, 2011; Etter et al., 2011; Rubin et al., 2012).

It is important to obtain DNA free from contaminants and in sufficient quantities to allow high efficiency of subsequent molecular analyses (Oliveira et al., 2007). The aims of the present study were to compare two methods for the extraction of genomic DNA from sand flies (Diptera, Psychodidae, Phlebotominae), and to standardize the protocol for obtaining DNA with better quality and quantity.

2. Material and methods

2.1. Study area and period

Sand flies were captured between August 2014 and July 2015 from the municipalities of Camapuã (19° 31' 51" S; 54° 02' 38" W), Campo Grande (20° 27' 5,5" S; 54° 37' 25,7" W), Corumbá (18° 59' 44" S; 57° 19' 36" W) and Miranda (20° 14' 27" S; 56° 22' 40"), state of Mato Grosso do Sul, Brazil (IBGE, 2016).

2.2. Identification

Male *Lutzomyia cruzi* and *Lutzomyia longipalpis* flies were selected, and identified in wet mount slides by observing the last abdominal tergite according to the classification proposed by Galati (2014). Head, thorax and the remaining part of the abdomen of the

sand flies were then stored separately at -20°C .

2.3. Extraction methods

Two extraction methods were tested: a commercial kit (*ReliaPrep™ Blood gDNA Miniprep System kit*, Promega®) and the classical phenol-chloroform extraction method as described by Uliana et al. (1991); two modifications were tested per method.

2.3.1. Genomic DNA extraction by commercial kit (Method 1)

Extraction of genomic DNA (gDNA) was performed using a commercial kit at the Genetic Research and Analysis Laboratory of the State University of São Paulo, Botucatu campus (PANGENE/UNESP). Twenty-four male *L. longipalpis* and *L. cruzi* sand flies were used for each protocol. The first protocol was based on the recommendations of the manufacturer of the commercial kit (*ReliaPrep™ Blood gDNA Miniprep System kit*, Promega®): Zirconium oxide microbeads and 200 μL of PBS were added to tubes containing one male sand fly, and homogenized in a mechanical stirrer (Bullet Blender® Homogenizer) for 5 min at ambient temperature. Twenty microliters of proteinase K and 200 μL of cellular lysis buffer were added; the mix was stirred for 10 s and incubated at 56°C for 30 min. Following incubation, 250 μL of binding buffer were added and stirred for 10 s. The contents were transferred to a kit column and centrifuged for 1 min at 13000 rpm. The kit column was transferred into a new collection tube, 500 μL of washing solution were added, and centrifuged for 3 min at 13000 rpm. This stage was repeated three times. Fifty microliters of nuclease-free water were added to the column, and centrifuged for 1 min at 13000 rpm. The column was discarded and the pellet was stored at -20°C .

The remaining protocols were developed by modifying the first protocol (i.e., the standard protocol). The modifications for each protocol have been presented in Table 1.

2.3.2. Genomic DNA extraction by Phenol-Chloroform (Method 2)

Extraction of gDNA by phenol-chloroform method was performed at the laboratory of Molecular Biology and Cell Culture of the Federal University of Mato Grosso do Sul. Twenty-four sand fly specimens were analyzed per protocol. The first protocol was based on Uliana et al. (1991). Ten microliters of 20% sodium dodecyl sulfate (SDS) and 200 μL of Tris-EDTA (TE) were added to tubes containing the specimens, and homogenized using plastic pestles. One microliter of proteinase K (20 mg/mL) was added and the contents homogenized. The tubes were placed in a dry-bath at 42°C for 2 h, and 200 μL of phenol:chloroform:isoamyl alcohol (25:24:1) were added. The samples were centrifuged for 5 min at

Table 1

Protocols of sand flies' genomic DNA extraction using a commercial kit according to the manufacturer and modified protocols (n = 72).

	Protocol 1 (n = 24)	Protocol 2 (n = 24)	Protocol 3 (n = 24)
V proteinase K (μL)	20	40	40
Incubation period (min)	30	30	120

12000 rpm, and the aqueous phase containing the DNA was collected. This was repeated three times. 200 μL of chloroform were added to each tube and centrifuged for 5 min at 12000 rpm to remove phenol residues. This was repeated three times. Twenty microliters of 3M sodium acetate and 400 μL of ice-cold 100% ethanol were added, and the tubes were incubated overnight at $-20\text{ }^{\circ}\text{C}$. The samples were then centrifuged for 10 min at 12000 rpm, at $4\text{ }^{\circ}\text{C}$. The supernatant was discarded, and the pellet was washed three times with 500 μL of 70% ethanol, followed by centrifugation for 10 min at 12000 rpm, at $4\text{ }^{\circ}\text{C}$. Following ethanol evaporation at ambient temperature, the samples were re-suspended in 50 μL of $1 \times \text{TE}$, and stored at $-20\text{ }^{\circ}\text{C}$.

The remaining protocols were developed by modifying the first protocol (i.e., the standard protocol). The modifications for each protocol are presented in Table 2.

2.4. Quantification

The extracted DNA was quantified using a second-generation Qubit 2.0 fluorometer, using a protocol based on the manufacturer's recommendations for the Qubit dsDNA HS kit (High Sensitivity, Invitrogen). The working solution was prepared by diluting (1:200) the fluorophore in the buffer solution included in the kit. A calibration curve was obtained using two standards, with concentrations 0 and 200 ng/ μL . Ten microliters of each standard were diluted in 190 μL of working solution. For quantification, 2 μL of the extracted DNA were added to 198 μL of work solution, and the fluorescence was recorded.

2.5. Statistical analysis

Averages and standard deviations of the DNA quantities obtained for the three protocols of the two extraction methods tested were calculated.

Since the required quantity of DNA could not be obtained using protocol 1 of Method 1, the unequal variances *t*-test (Welch Two Sample *t*-test) was used to compare protocols 2 and 3. Variances for Method 2 were found to be non-homogeneous (Bartlett test of homogeneity of variances), and the Kruskal-Wallis test was used to test for significant differences between protocols, with 3 groups and 24 observations per group.

All analyses were performed using the R 3.3.0 software and all inferences were drawn at $p \leq 0.05$ (R Core Team, 2016).

2.6. Evaluation of extracted DNA

Integrity and purity of the extracted DNA was evaluated by PCR, using region IVS6 (sand fly constitutive cacophony gene) as endogenous control. The following pair of primers specific for region IVS6 was used: 5Llac 5'-GTGGCCGAACATAATGTTAG-3' and 3Llac 5'-CCACGAACAAGTTCAACATC-3' (Lins et al., 2002).

PCR reactions were performed in a final volume of 25 μL ,

containing 5 μL of gDNA, 12.5 μL of GoTaq MasterMix (Promega®), 5.5 μL of milli-Q water, and 1 μL of each primer. Autoclaved milli-Q water and DNA obtained from an *in vitro* culture of *Leishmania (Leishmania) amazonensis* (IFLA/BR/1967/PH8), provided by the Laboratory of Parasitology of the René Rachou/Fiocruz Minas Gerais Research Center, were used as negative controls to confirm the absence of contaminants, and to validate the PCR, respectively.

The PCR program consisted of one initial denaturation cycle at $94\text{ }^{\circ}\text{C}$ for 12 min; followed by 35 cycles of $94\text{ }^{\circ}\text{C}$ for 30 s, $55\text{ }^{\circ}\text{C}$ for 30 s, and $72\text{ }^{\circ}\text{C}$ for 30 s; and final extension at $72\text{ }^{\circ}\text{C}$ for 10 min (Pita-Pereira et al., 2005).

The amplification products were resolved in 1.5% agarose in TBE buffer, and stained with GelRed (Biotium, USA) for visualization. A 100 bp ladder was also loaded in well for confirmation of the size of the amplification products.

3. Results

A total of 144 male sand flies were used, 72 for each extraction method, divided among the three protocols tested per method (Table 3). For the protocols 1, 2 and 3 using ReliaPrep Blood gDNA Miniprep System kit 6 samples from each locality were used, totaling 24 samples per protocol ($n = 72$). For protocols 1, 2 and 3 using phenol-chloroform were used *Lutzomyia longipalpis* from Campo Grande ($n = 72$).

Means and standard deviations of the DNA concentrations (ng/ μL) obtained by each extraction method and protocol are shown in Table 4. Significant differences were observed between the original and modified protocols (protocols 2 and 3) of each method; protocols 2 and 3 of method 2 (phenol-chloroform) performed significantly better than all other protocols.

The quantities of DNA obtained by the three protocols of each extraction method are shown for comparison in Figs. 1 and 2. Under method 1, protocol 3 presented higher yield than protocol 2 ($p < 0.001$) (Fig. 1).

Significant differences were observed between the yields of DNA obtained with the three protocols of method 2 ($p < 0.001$; Fig. 2). The distribution of the results indicates descriptively that protocol 3 presented higher yield.

PCR amplification of the *L. longipalpis* constitutive cacophony (*cac*) gene was observed for all protocols of the phenol-chloroform extraction method (Fig. 3). The IVS6 region was detected in 70 of the 72 (97.22%) samples extracted with phenol. Amplification products 220 bp in size were confirmed by agarose gel electrophoresis for 22 of the 24 (91.66%) samples under protocol 1, and for all samples under protocols 2 and 3.

4. Discussion

New sequencing methods, called Next-Generation Sequencing (NGS), have raised expectations for genomic studies and new applications previously not explored by traditional Sanger sequencing (Morozova and Marra, 2008). These new tools may help clarify biosystematics, taxonomy, population dynamics, and phylogenetic relationships, and may be applied to studies of sand flies. We

Table 2
Protocols of sand flies' genomic DNA extraction by phenol-chloroform (Uliana et al., 1991) and modified protocols ($n = 72$).

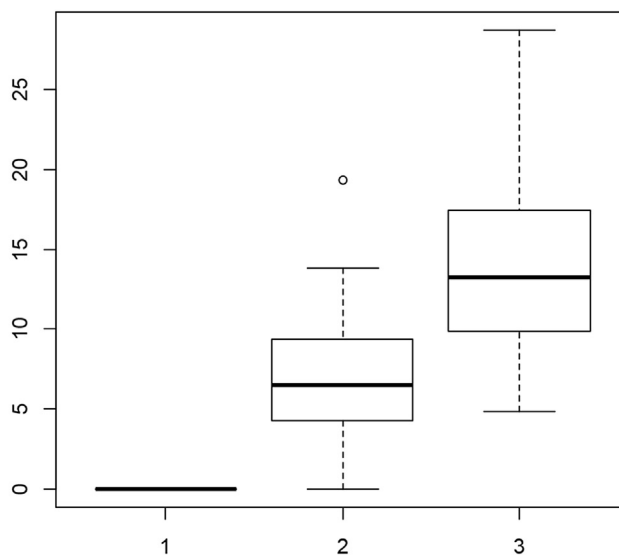
	Protocol 1 ($n = 24$)	Protocol 2 ($n = 24$)	Protocol 3 ($n = 24$)
200 μL of phenol: chloroform: isoamyl alcohol	3X	1X	1X
200 μL of chloroform	3X	–	–
500 μL of ethanol 70%	3X	3X	1X
V proteinase K	1 μL	1 μL	5 μL
Incubation period (hours)	2	2	24

Table 3
Sand fly specimens collected per species and municipality, Mato Grosso do Sul, Brazil, 2015.

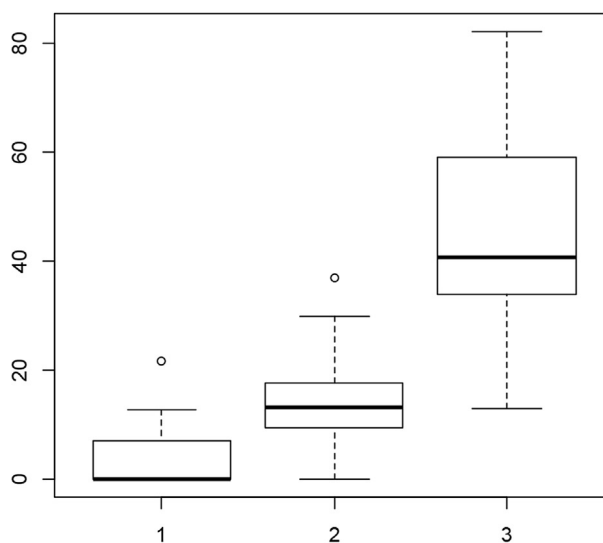
	<i>Lu. cruzi</i>	<i>Lu. longipalpis</i>
Camapuã	18	–
Campo Grande	–	90
Corumbá	18	–
Miranda	–	18

Table 4DNA concentration (ng/ μ L) average and standard deviation for the different protocols (method 1 = commercial kit, method 2 = phenol-chloroform).

Method	Protocol 1		Protocol 2		Protocol 3	
	Mean (SD)	t (p-value)	Mean (SD)	t (p-value)	Mean (SD)	t (p-value)
1	–	–	6.99 (4.59)	3.28 (0.002)	13.88 (5.86)	8.22 (<0.001)
2	3.44 (5.80)	–	13.83 (9.12)	–	45.03 (17.62)	–

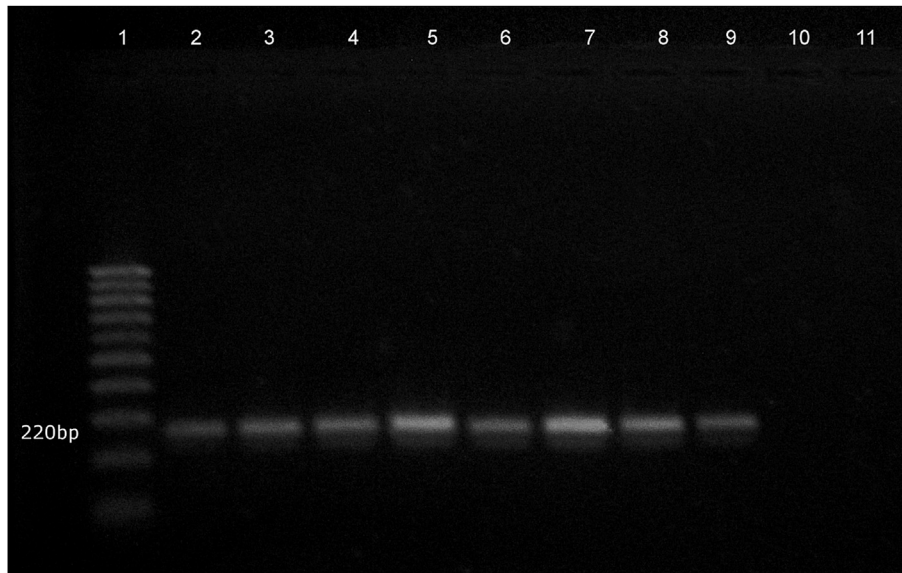
SD = standard deviation; t = *t*-test (Welch Two Sample *t*-test).

t = -4.54 (protocol 2 and 3); degrees of freedom = 43.496; $p < 0.001$.

Fig. 1. Box plot for DNA quantity (ng/ μ L) for the three protocols under method 1 (commercial kit).t = -4.54 (protocol 2 and 3); degrees of freedom = 43.496; $p < 0.001$.

Kruskal-Wallis chi-squared = 52.36; degrees of freedom = 2, $p < 0.001$

Fig. 2. Box plot for DNA quantity (ng/ μ L) for the three protocols under method 2 (phenol).Kruskal-Wallis chi-squared = 52.36; degrees of freedom = 2, $p < 0.001$.



Sample 1: 100 bp marker; Samples 2 to 9: amplification products of approximately 220 bp size; Sample 10: milli-Q water; sample 11: *Leishmania amazonensis* (IFLA/BR/1967/PH8).

Fig. 3. Amplification product for DNA encoding region IVS6 of the constitutive cacophony gene of *Lutzomyia longipalpis*. Sample 1: 100 bp marker; Samples 2 to 9: amplification products of approximately 220 bp size; Sample 10: milli-Q water; sample 11: *Leishmania amazonensis* (IFLA/BR/1967/PH8).

analyzed two methods with the aim of standardizing DNA extraction from sand flies for NGS and observed significant differences in the concentration of the extracted DNA.

An ideal DNA extraction protocol should meet some criteria such as concentrated yield, elimination of PCR inhibitors, implementable as laboratory technique, low cost, reproducibility, and safety (Rantakokko-Javala and Javala, 2002). Researchers have used different extraction methods to extract insect DNA, such as phenol-chloroform-isoamyl alcohol, detergents like Triton X-100 and Triton X-100/DDT, silica column solid phase extraction, and Chelex[®] Molecular Biology Grade Resin (Bio-Rad Laboratories, Hercules, USA) (Cova et al., 2015; Oliveira et al., 2015; Paiva et al., 2007; Pinto et al., 2015; Santos et al., 2013).

Most molecular studies involving sand flies focus on the detection of DNA of *Leishmania* spp., but may as well be used for studying the vector itself by using approaches such as microsatellite analysis (Santos et al., 2013) and cytochrome c oxidase analysis (Arrivillaga et al., 2002; Gajapathy et al., 2016). Irrespective of these analyses being used to study the parasite or its vector, it is essential to evaluate whether they can be applied to more complex genomic studies, such as the new sequencing methods demanding DNA of high purity and integrity.

NGS studies involving RAD sequencing require from 7.5 ng to 1 µg of DNA (Baird et al., 2008; Baxter et al., 2011; Emerson et al., 2015) whereas establishing a genomic library using the Nextera[®] DNA prep kit (Illumina; manufacturer's instructions) requires 20 µL of volume containing 50 ng of DNA (2.5 ng/µL). Extraction methods, therefore, need to be optimized so that the required DNA quantities can be obtained. In the present study, the three protocols developed for the commercial kit were found to be more practical. Although the protocols modified from Uliana et al. (1991) were more laborious, taking up to 48 h and requiring greater care due to the toxicity of the phenol/chloroform used for purification, they resulted in significantly higher quantities of gDNA.

In traditional extraction methods, the main stages are cell lysis (using a detergent such as sodium dodecyl sulfate (SDS), with or without proteinase K); purification (e.g., phenol/chloroform); and

precipitation (using sodium acetate and ice-cold 100% ethanol). Following the advent of commercial kits, these methods have been modified with the main differences being observed at the stages of purification and precipitation, using solid phases such as silica columns, magnetic beads, and filter paper (Clements et al., 2008).

Extraction methods using silica are based on the adsorption of DNA to silica particles in the presence of specific salts at specific pH. Cell remains and contaminants are removed during the washing stages, and DNA is re-suspended in an elution buffer (Yang et al., 1998). The advantage of this procedure is that silica can be included in spin columns, enabling the automation of the extraction and resulting in high quality purified DNA. Three protocols were tested for the commercial kit (*ReliaPrep[™] Blood gDNA Mini-prep System kit*, Promega[®]) using silica columns as a solid phase. This method may be recommended for obtaining DNA for high sensitivity molecular methods such as PCR because it is more practicable and less toxic, and has lower chances of contamination. However, for NGS, which requires higher volumes and concentration of DNA, our results with this method were not satisfactory.

An important factor that should be considered in DNA extraction is the chemical composition of sand flies. Their bodies are mostly constituted of exoskeletal chitin and parasitic microorganisms may be present (Romoser, 2004). These components may interfere with subsequent analyses, acting as PCR inhibitors. Therefore, we tested different times of digestion with proteinase K, both for the protocols of commercial kit extraction and of phenol-chloroform extraction.

For both extraction methods, higher quantities of proteinase K and longer incubation times resulted in higher concentrations of extracted DNA. The aim of the enzymatic digestion with proteinase K is digestion of the contaminating proteins, which tend to co-precipitate DNA with them. Isolda et al. (1994) and Mesquita et al. (2001) obtained better quantity and quality of DNA using low concentrations of proteinase K during long periods of digestion (3–5 days). The time during which the insects are in contact with proteinase K may, therefore, determine the final quantity of the extracted DNA.

Organic extraction with phenol-chloroform is the conventional method for extraction of DNA from different biological samples. Phenol-chloroform allows the separation of polynucleotides (DNA and RNA) from proteins and cell remains. The phenol-chloroform method has been widely used for DNA extraction due of its high purification capacity. The main limitation of this method is the possibility of residual phenol or chloroform in the extracted DNA. In addition, there is evidence that besides inhibiting Taq polymerase activity, phenol negatively affects PCR sensitivity (Aransay et al., 2000; Arez et al., 2000). For protocols 1 and 2 of phenol-chloroform extraction, the number of steps using this solution may have resulted in DNA loss, explaining the low concentrations of DNA obtained as well as the two negative PCR results received with protocol 1.

The best results were obtained under the standardized conditions of protocol 3, which resulted in the highest quantities of extracted DNA, with positive PCR results for all samples and absence of non-specific amplification. Although fluorometric quantification is an important indicator of the presence of DNA, the best way to ascertain the presence of nucleic acids is by PCR. The recommended target for sand flies is the cacophony (*cac*) gene, which codes for subunit alpha-1 of a voltage-dependent calcium channel and is related to the control of the sound emitted by the wings of males during courtship (Oliveira et al., 2001). Amplification of this target sequence indicates the presence of sand fly DNA, and has been used by several authors as an endogenous control in PCR (Cova et al., 2015; Lins et al., 2002; Neitzke-Abreu et al., 2014; Saraiva et al., 2010; Pita-Pereira et al., 2011).

The “gold standard” for species identification of sand flies is the use of a taxonomic key based on the works of Young and Duncan (1994) or Galati (2014), because of its high specificity. Of the more than 270 species reported from Brazil, the following species present high similarity: *Evandromyia cortelezzii*, *Evandromyia salesi*, and *Evandromyia corumbaensis*, which constitute the *cortelezzii* complex (Galati et al., 1989); *Nyssomyia neivai* and *Nyssomyia intermedia*, which constitute the *intermedia* complex (Lutz and Neiva, 1912); and *Lutzomyia cruzi*, *Lutzomyia gaminarai*, *Lutzomyia longipalpis*, and *Lutzomyia pseudolongipalpis*, included in the *longipalpis* complex (Arrivillaga and Feliciangeli, 2001; Bauzer et al., 2007). It should be highlighted that these cryptic species exhibit different vector capacity and competence, but only slight differences in taxonomic characters, making species identification and investigation of *Leishmania* spp. transmission cycle difficult (Ready, 2011). Molecular data may, therefore, help in taxonomic studies of sand flies, especially in areas where leishmaniasis is endemic with different species of sand flies existing in allopatry or sympatry.

Knowledge of the genetic structure of vector populations is essential for the eco-epidemiology of vector diseases and leishmaniasis. NGS studies have been performed previously on *Drosophila* spp. and *Anopheles* spp., suggesting their utility in studies of other insects, such as sand flies (which will be presented in future studies). Rubin et al. (2012) determined the phylogeny of *Drosophila* spp. using orthologous sequences, indicating that RAD-seq (a type of NGS) is a promising tool for the reconstruction of phylogenetic relationships among younger clades. Emerson et al. (2015) analyzed 12 populations of *Anopheles darlingi*, and found significant levels of genetic divergence in three clusters. In the study on *Drosophila* spp., gene sequences were obtained from secondary data, using gene banks, and in the *A. darlingi* study, DNA extraction was performed using the Qiagen DNEasy kit, which is also based on silica columns. Assuming that the quantity of DNA obtained is directly proportional to the size of the animal body, extractions from *Anopheles* spp. can be expected to yield larger quantities of DNA, since they can be up to 3 times bigger than a sand fly.

Standardization studies are necessary to optimize DNA extraction from different biological samples, especially considering that Next-Generation Sequencing is a promising tool for entomological and molecular studies of sand flies. The present study is a pioneer study on the extraction of DNA from sand flies, aimed at its application to NGS. Our results indicate that protocol 3 of the phenol-chloroform extraction method yields higher quantities of sand fly DNA, and is therefore best suited for studies on sand flies using NGS.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

We are grateful to the local residents, who assisted us in a patient and courteous way; to the staff of the Laboratory of Parasitology-UFMS for their collaboration and to UFMS for logistical support.

References

- Adler, S., Theodor, O., 1957. Transmission of disease agents by phlebotomine sand flies. *Annu. Rev. Entomol.* 2, 203–226.
- Aransay, A.M., Scoulica, E., Tselentis, Y., 2000. Detection and identification of *Leishmania* DNA within naturally infected sand flies by seminested PCR on minicircle kinetoplastic DNA. *Appl. Environ. Microbiol.* 66, 1933–1938.
- Arez, A.P., Lopes, D., Pinto, J., Franco, A.S., Snounou, G., Rosário, V.E., 2000. *Plasmodium* sp.: optimal protocols for PCR detection of low parasite numbers from mosquito (*Anopheles* sp.) samples. *Exp. Parasitol.* 94, 269–272.
- Arrivillaga, J.C., Feliciangeli, M.D., 2001. *Lutzomyia pseudolongipalpis*: the first new species within the *longipalpis* (Diptera: Psychodidae: Phlebotominae) complex from La Rinconada, Curarigua, Lara state, Venezuela. *J. Med. Entomol.* 38, 783–790.
- Arrivillaga, J.C., Norris, D.R., Feliciangeli, M.D., Lanzaro, G.C., 2002. Phylogeography of the neotropical sand fly *Lutzomyia longipalpis* inferred from mitochondrial DNA sequences. *Infect. Genet. Evol.* 2, 83–95.
- Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A., Johnson, E.A., 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* 3, 1–7.
- Bauzer, L.G., Souza, N.A., Maingon, R.D., Peixoto, A.A., 2007. *Lutzomyia longipalpis* in Brazil: a complex or a single species? A mini-review. *Mem. Inst. Oswaldo Cruz* 102, 1–12.
- Baxter, S.W., Davey, J.W., Johnston, J.S., Shelton, A.M., Heckel, D.G., Jiggins, C.D., Blaxter, M.L., 2011. Linkage mapping and comparative genomics using next-generation RAD sequencing of a non-model organism. *PLoS One* 6, 1–11.
- Brumfield, R.T., Beerli, P., Nickerson, D.A., Edwards, S.V., 2003. The utility of single nucleotide polymorphism in inferences of population history. *Trends Ecol. Evol.* 18, 249–256.
- Clements, D.N., Wood, S., Carter, S.D., Ollier, W.E.R., 2008. Assessment of the quality and quantity of genomic DNA recovered from canine blood samples by three different extraction methods. *Res. Vet. Sci.* 85, 74–79.
- Cova, B.O., Ferreira, F.S., Sousa, O.M.F., Santos, C.G.S., Fonseca, E.O.L., Ribas, L.L.R., 2015. Padronização das condições para a extração automatizada do DNA de Flebotomíneos no Laboratório Central de Saúde Pública da Bahia. *Rev. Baiana Saúde Pública* 39, 125–1352.
- Davey, J.W., Blaxter, M., 2011. RADSeq: next-generation population genetics. *Brief. Funct. Genomics* 9, 416–423.
- Emerson, K.J., Conn, J.E., Bergo, E.S., Randel, M.A., Sallum, M.A.M., 2015. Brazilian *Anopheles darlingi* root (Diptera: Culicidae) clusters by major biogeographical region. *PLoS One* 10, 1–15.
- Etter, P.D., Bassham, S., Hohenlohe, P., Johnson, E.A., Cresko, W., 2011. SNP discovery and genotyping for evolutionary genetics using RAD sequencing. *Methods Mol. Biol.* 157–178.
- Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.* 3, 294–299.
- Gajapathy, K., Tharmasegaram, T., Eswaramohan, T., Peries, L.B.S.L., Jayanetti, R., Surendran, S.N., 2016. DNA barcoding of Sri Lankan phlebotomine sand flies using cytochrome c oxidase subunit I reveals the presence of cryptic species. *Acta Trop.* 161, 1–7.
- Galati, E.A.B., Nunes, V.L.B., Oshiro, E.T., Rego Jr., F.A., 1989. Nova espécie de Phlebotominae, *Lutzomyia corumbaensis*, sp. n. (Diptera, Psychodidae) do complexo *Lutzomyia cortelezzii*. *Rev. Bras. Entomol.* 33, 765–775.
- Galati, E.A.B., 2014. Phlebotominae (Diptera, Psychodidae): classificação, morfologia, terminologia e identificação de adultos. http://www.fsp.usp.br/egalati/ApostilaPhlebotominae_vol_I.pdf.

- IBGE, 2016. Instituto Brasileiro De Geografia E Estatística. IBGE@Cidades.
- Isolda, J., De Vries, S., Chu, L., Ghazvini, S., Waldman, F., 1994. Analysis of changes in DNA sequence copy number by comparative genomic hybridization in archival paraffin-embedded tumor samples. *Am. J. Pathol.* 145, 1301–1308.
- Lainson, R., Shaw, J.J., 2005. New world leishmaniasis. In: Cox, F.E.G., Kreier, J.P., Wakelin, D. (Eds.), *Microbiology and Microbial Infections. Parasitology*, pp. 313–349.
- Lins, R.M.M.A., Oliveira, S.G., Souza, N.A., Queiroz, R.G., Justiniano, S.C.B., Ward, D., Kyriacou, C.P., Peixoto, A.A., 2002. Molecular evolution of the cacophony IVS6 region in sandflies. *Insect Mol. Biol.* 11, 117–122.
- Lutz, A., Neiva, A., 1912. Contribuição para o conhecimento das espécies do gênero *Phlebotomus* existente no Brasil. *Mem. Inst. Oswaldo Cruz* 4, 84–95.
- Marcondes, C.B., 1996. A redescription of *Lutzomyia* (*Nyssomyia*) *intermedia* (Lutz & Neiva, 1912), and resurrection of *L. neivai* (Pinto, 1926) (Diptera, Phlebotominae). *Mem. Inst. Oswaldo Cruz* 91, 457–462.
- Mesquita, R.A., Anzai, E.K., Oliveira, R.N., Nunes, F.D., 2001. Avaliação de três métodos de extração de DNA de material parafinado para amplificação de DNA genômico pela técnica da PCR. *Pesqui. Odontol. Bras.* 15, 314–319.
- Morozova, O., Marra, M.A., 2008. Applications of next-generation sequencing Technologies in functional genomics. *Genomics* 92, 255–264.
- Neitzke-Abreu, H.C., Reinhold-Castro, K.R., Venazzi, M.S., Scodro, R.B.L., Dias, A.C., Silveira, T.G.V., Teodoro, U., Leonardi, M.V.C., 2014. Detection of *Leishmania* (*Viannia*) in *Nyssomyia neivai* and *Nyssomyia whitmani* by multiplex polymerase chain reaction, in southern Brazil. *Rev. Inst. Med. Trop. São Paulo* 56, 391–395.
- Oliveira, S.G., Bottecchia, M., Bauzer, L.G.S.R., Souza, N.A., Ward, R.D., Kyriacou, C.P., Peixoto, A.A., 2001. Courtship song genes and speciation in sand flies. *Mem. Inst. Oswaldo Cruz* 96, 403–405.
- Oliveira, M.C.S., Regitano, L.C.A., Roese, A.D., Anthonisen, D.G., Patrocínio, E., Parma, M.M., Scagliusi, S.M.M., Timóteo, W.H.B., Jardim, S.N., 2007. Fundamentos teórico-práticos e protocolos de extração e de amplificação de DNA por meio da técnica de reação em cadeia de polimerase. *Embrapa Pecuária Sudeste*.
- Oliveira, E.F., Casaril, A.E., Mateus, N.L.F., Murat, P.G., Fernandes, W.S., Oshiro, E.T., Oliveira, A.G., Galati, E.A.B., 2015. *Leishmania amazonensis* DNA in wild females of *Lutzomyia cruzi* (Diptera: Psychodidae) in the state of Mato Grosso do Sul, Brazil. *Mem. Inst. Oswaldo Cruz* 1–7.
- Paiva, B.R., Secundino, N.F.C., Pimenta, P.F.P., Galati, E.A.B., Junior, H.F.A., Malafrente, R.S., 2007. Padronização de condições para detecção de DNA de *Leishmania* spp. em flebotomíneos (Diptera, Psychodidae) pela reação em cadeia da polimerase. *Cad. Saúde Pública* 23, 87–94.
- Pinto, I.S., Chagas, B.D., Rodrigues, A.A.F., Ferreira, A.L., Rezende, H.R., Bruno, R.V., Falqueto, A., Andrade-Filho, J.D., Galati, E.A.B., Shimabukuro, P.H.F., Brazil, R.P., Peixoto, A.A., 2015. DNA barcoding of neotropical sand flies (Diptera, Psychodidae, Phlebotominae): species identification and discovery within Brazil. *PLoS One* 10, 1–18.
- Pita-Pereira, D., Alves, C.R., Souza, M.B., Brazil, R.P., Bertho, A.L., Barbosa, A.F., Britto, C.C., 2005. Identification of naturally infected *Lutzomyia intermedia* and *Lutzomyia migonei* with *Leishmania* (*Viannia*) *braziliensis* in Rio de Janeiro (Brazil) revealed by a PCR multiplex non-isotopic hybridization assay. *Trans. R. Soc. Trop. Med. H.* 99, 905–913.
- Pita-Pereira, D., Souza, G.D., Pereira, T.A., Zwetsch, A., Britto, C., Rangel, E.F., 2011. *Lutzomyia* (*Pintomyia*) *fischeri* (Diptera: Psychodidae: Phlebotominae), a probable vector of american cutaneous leishmaniasis: detection of natural infection by *Leishmania* (*Viannia*) DNA in specimens from the municipality of porto alegre (RS), Brazil, using multiplex PCR assay. *Acta Trop.* 120, 273–275.
- Queller, D.C., Strassmann, J.E., Hughes, C.R., 1993. Microsatellites and kinship. *TREE* 8, 285–288.
- Rantakokko-Javala, K., Javala, J., 2002. Optimal DNA isolation method for detection of bacteria in clinical specimens by broad-range PCR. *J. Clin. Microbiol.* 40, 4211–4217.
- R Core Team, 2016. R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Disponível em: <https://www.R-project.org/>.
- Ready, P.D., 2011. Should sand fly taxonomy predict vectorial and ecological traits. *J. Vector Ecol.* 36, 17–22.
- Romoser, W.S., 2004. Introductions to arthropodes: structure, function and development. Eldridge, B. F.; Edman, J. *Med. Entomol.* Cap 2, 15.
- Rubin, B.E.R., Ree, R.H., Moreau, C.S., 2012. Inferring phylogenies from RAD sequence data. *PLoS One* 7, 1–12.
- Santos, M.F.C., Ribolla, P.E.M., Alonso, D.P., Andrade-Filho, J.D., Casaril, A.E., Ferreira, A.M.T., Fernandes, C.E.S., Brazil, R.P., Oliveira, A.G., 2013. Genetic structure of *Lutzomyia longipalpis* populations in Mato Grosso do Sul, Brazil, based on microsatellite markers. *PLoS One* 8, 1–7.
- Saraiva, L., Andrade Filho, J.D., Silva, S.O., Andrade, A.S.R., Melo, M.N., 2010. The molecular detection of different *Leishmania* species within sand flies from a cutaneous and visceral leishmaniasis sympatric area in Southeastern Brazil. *Mem. Inst. Oswaldo Cruz* 105, 1033–1039.
- Soria-Carrasco, V., Gompert, Z., Comeault, A.A., Farkas, T.E., Parchman, T.L., Johnston, J.S., Buerkle, C.A., Feder, J.L., Bast, J., Schwander, T., Egan, S.P., Crespi, B.J., Nosil, P., 2014. Stick insect genomes reveal natural Selection's role in parallel speciation. *Science* 344, 738–742.
- Uliana, S.R.B., Affonso, M.H.T., Camargo, E.P., Floeter-Winter, L.M., 1991. *Leishmania*: genus identification based on a specific sequence of the 18S ribosomal RNA sequence. *Exp. Parasitol.* 72, 157–163.
- Vignal, A., Milan, D., Sancristobal, M., Eggen, A., 2002. A review on SNP and other types of molecular markers and their use in animal genetics. *Genet. Sel. Evol.* 34, 275–305.
- Yang, D.Y., Eng, B., Waye, J.S., Dudar, J.C., Saunders, S.R., 1998. Technical note: Improved DNA extraction from ancient bones using silica-based spin columns. *Am. J. Phys. Anthropol.* 105, 539–543.
- Young, D.G., Duncan, M.A., 1994. Guide to the Identification and Geographical Distribution of *Lutzomyia* Sand Flies in Mexico, the West Indies, Central and South America (Diptera: Psychodidae). American Entomological Institute, Gainesville.
- Zhang, D., Hewitt, G.M., 1997. Insect mitochondrial control region: a review of its Structure, Evolution and Usefulness in evolutionary studies. *Biochem. Syst. Ecol.* 25, 99–120.