# Gp63 gene polymorphism and population structure of *Leishmania donovani* complex: influence of the host selection pressure?

S. GUERBOUJ[1,2], K. VICTOIR[2], I. GUIZANI[1], N. SERIDI[3], N. NUWAYRI-SALTI[4], M. BELKAID[3], R. BEN ISMAIL[1], D. LE RAY[2] *and* J. C. DUJARDIN[2]*

[1] *Institut Pasteur de Tunis, Laboratoire d'Epidémiologie et Ecologie Parasitaire, 13 Place Pasteur BP74, 1002 Tunis Belvédère, Tunisia*
[2] *Prins Leopold Instituut voor Tropische Geneeskunde, Protozoology Unit, 155 Nationalestraat, B-2000 Antwerp, Belgium*
[3] *Institut Pasteur d'Algérie, rue du Docteur Laveran, 16000 El Hamma, Algiers, Algeria*
[4] *American University of Beirut, Bliss street, P.O. Box: 11-0236, Beirut, Lebanon*

(*Received 2 March 2000; revised 28 June 2000; accepted 6 July 2000*)

SUMMARY

The gp63 encoding genes were characterized by PCR–RFLP in 35 isolates representative of the *Leishmania donovani* complex (*L. infantum*, *L. donovani*, *L. archibaldi* and *L. chagasi*), with special attention to Mediterranean *L. infantum* from different geographical origins, and in separate groups from Old World *Leishmania* (*L. major*, *L. tropica* and *L. aethiopica*). The aim was to evaluate how the possible selective pressure by the host on these important surface proteins would influence structuring of our sample. Comparison was carried out with the structure obtained (i) from reported isoenzyme data, characters supposed to vary neutrally, and (ii) from PCR–RFLP analysis of gp63 inter-genic regions, containing non-translated spacers and regulatory genes. Polymorphism within the gp63-encoding region, was much higher than in gp63 inter-genic regions. In the gp63 intra-genic dendrogram, the 4 species of *L. donovani* complex were discriminated and quite distinct from outgroups. Within *L. infantum*, geographical structuring was observed and did not overlap with the structure built-up from isoenzymes and inter-genic data. These results support the idea of a strong host-selection on gp63, at vector level but most of all at vertebrate (human or dog) immunological level. Furthermore, they illustrate how the nature of genetic characters may influence the perception of population structuring.

Key words: immunogen, polymorphism, selection, diagnosis.

## INTRODUCTION

Population structure of parasites is influenced by many factors, either selective (immune pressure, vector specificity …) or random (drift, migration …). Genetic characterization of *Leishmania* and other Trypanosomatids is mostly based on multilocus enzyme analysis (MLEE, Rioux *et al*. 1990). However, it is classically considered that variations due to isoenzymes reflect neutral mutations (Nei, 1987). Another evaluation of population structure might result from analysis of genes subject to a different selective pressure.

In this respect, gp63 genes are a potentially informative target. Gp63, the major surface glycoprotein of the parasite, is a strong protective antigen (further called immunogen, Rivier *et al*. 1999), and plays an important role in the adhesion to the macrophage and the intra-phagolysosomal survival (Medina-Acosta, Beverley & Russel, 1993). As such gp63 is likely to be exposed to strong selective pressure. Gp63 encoding genes are tandemly repeated and sequence heterogeneity has been described within a stock and also at inter- and intra-specific levels (Button *et al*. 1989; Webb, Button & McMaster, 1991; Roberts *et al*. 1993; Medina-Acosta *et al*. 1993; Victoir *et al*. 1995; Gonzàlez-Aseguinolaza *et al*. 1997). A PCR–RFLP assay was developed on the basis of these genomic properties, and this permitted the discrimination of the major species of subgenus *Viannia* (Victoir *et al*. 1998).

The aim of present work was to explore gp63 gene polymorphism in the *L. donovani* complex. This group of species is distributed over a large territory, and is responsible for a variable clinical pattern (Moreno *et al*. 1986). In addition, this group has been particularly well studied at isoenzyme level (Moreno *et al*. 1986, Rioux *et al*. 1990; Thomaz-Soccol *et al*. 1993) and thus constitutes an excellent model to test the influence of the selective pressure on the population structure. Thus, 35 stocks, from different geographical origins, representing the *L. donovani* complex (*L. infantum*, *L. donovani*, *L. archibaldi* and *L. chagasi*) were selected. They had previously been characterized by MLEE analysis

* Corresponding author: Prins Leopold Instituut voor Tropische Geneeskunde, Protozoology Unit, 155 Nationalestraat, B-2000 Antwerp, Belgium. Tel: +32 3 2476355. Fax: +32 3 2476362.
E-mail: jcdujard@proto.itg.be

 
https://doi.org/10.1017/S0031182000007125 Published online by Cambridge University Press

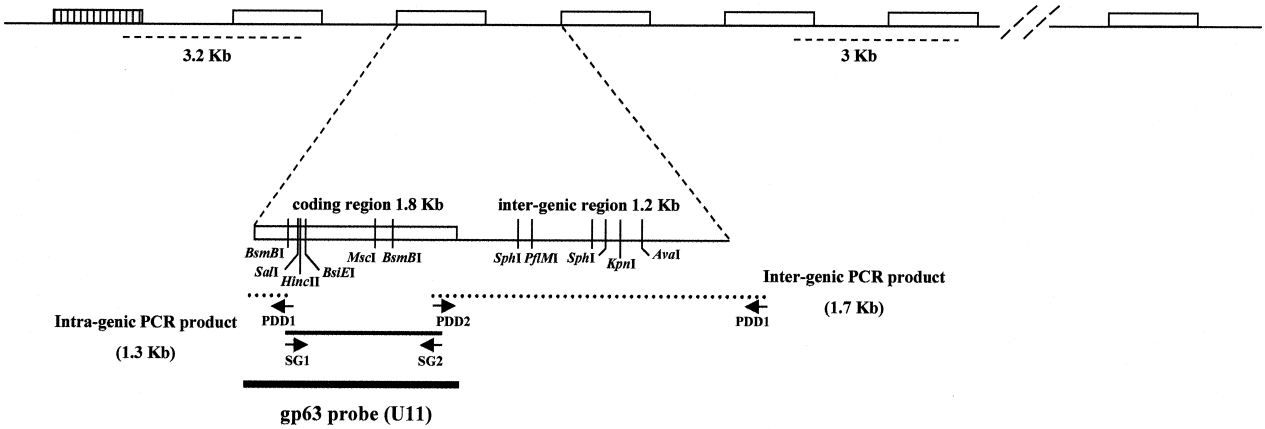


Fig. 1. Gp63 gene PCR amplification scheme. Arrows indicate the target of oligonucleotide primers used for intra-genic (SG1 and SG2) and inter-genic (PDD1 and PDD2) amplification. Localization of the target for the gp63 probe (U11) is shown. Slash marks indicate a gap within the gene cluster in which the number of additional genes is not known (from data according to Gonzàlez-Aseguinolaza *et al*. 1997). Positions of all the restriction enzymes used are indicated.

and were here analysed by gp63 PCR–RFLP, both at intra-genic (glycoprotein coding region) and inter-genic (spacers and regulatory sequences) levels. Species representing the other Old World complexes (*L. major*, *L. tropica* and *L. aethiopica*) were included as separate groups. Our results gave evidence for genetic polymorphism within gp63 genes and provided a powerful tool to differentiate species within the *L. donovani* complex. Geographical structuring was observed within *L. infantum*. The role of host (vertebrate and/or invertebrate) as a source of selective pressure on polymorphic antigens and structuring of the parasite populations is discussed.

## MATERIALS AND METHODS

### *Parasites*

Promastigotes were cultivated at 26 °C either in RPMI 1640 supplemented with 10% inactivated foetal calf serum or on blood agar (Tobie, Von Brand & Mehlman, 1950). *Leishmania* isolates used in this study (Table 1) were 26 *L. infantum* isolates originating from Tunisia (8), Algeria (9), France (4), Spain (4), and Lebanon (1) as well as 7 *L. donovani* from India (2), Ethiopia (2), Kenya (2) and Saudi Arabia (1). *L. chagasi*, *L. major*, *L. tropica*, *L. aethiopica* and *L. archibaldi* reference strains were also included. All but 3 of these isolates have been previously characterized by isoenzymatic starch gel electrophoresis with 15 enzyme systems according to Rioux *et al*. (1990).

### *DNA extraction, labelling and hybridization*

High molecular weight genomic DNA was isolated as described elsewhere (Van Eys *et al*. 1991). DNA from amplification products was extracted with an equal volume of phenol/chloroform and ethanol-precipitated. After depurination and denaturation with alkali, DNA was transferred to a Nylon Hybond N+ Amersham membrane according to the 'pocket blotting' method (Cuny, Veas & Roizès, 1991). A 2 kb *Eco*R I/*Xba* I fragment representing the total coding region of the gp63 sequence of *L. chagasi* (Ramamoorthy *et al*. 1992; Roberts *et al*. 1993) was used as a probe (U11 probe, gift of M. E. Wilson, University of Iowa, USA). The insert was gel-excised and purified using a QIAEX II purification kit (Qiagen). The probe was labelled with [$^{32}$P]dCTP by random primer labelling (Amersham) and used on all blots. Hybridization and high stringency washings (0·1 × SSC) were performed at 65 °C according to the manufacturer's instructions (Amersham).

### *Polymerase chain reaction (PCR)*

Two PCR assays were developed aiming at the amplification of the coding and the inter-genic region (untranslated region) of the gp63 genes. Published sequences corresponding to the gp63 genes from *L. infantum*, *L. donovani*, *L. chagasi* and *L. major* were aligned using algorithms and databases, part of the Wisconsin Package Version 9·1 (Genetics Computer Group (GCG), Madison, WI). Regions found to be more conserved among the different sequences were used to design primers for amplification of the coding region of the gp63 genes (intra-genic gp63 PCR): 5′ GTCTCCACCGAGGACCTCACCGA 3′ (SG1) and 5′ TGATGTAGCCGCCCTCCTC-GAAG 3′ (SG2). These primers correspond to positions 823–846 and 2133–2156 respectively in the gp63 sequence of *L. infantum* according to Gonzàlez-Aseguinolaza *et al*. (1997) (Fig. 1). Reactions (50 or 100 μl) contained 50 ng template DNA, 10 mM

https://doi.org/10.1017/S0031182000007125 Published online by Cambridge University Press

Table 1. Origin and code of the stocks

| International code | Geographical origin | Zymodeme | Pathology* |
|---|---|---|---|
| *L. infantum* | | | |
| MHOM/TN/80/IPT1† | Tunisia | MON 1 | VL |
| MHOM/TN/87/CN65 | Tunisia | MON 1 | VL |
| MHOM/TN/87/ CN64 | Tunisia | MON 1 | VL |
| MCAN/TN/89/ALM220 | Tunisia | — | CVL |
| MHOM/TN/88/Nabil | Tunisia | MON 1 | VL |
| MHOM/TN/87/KA412 | Tunisia | MON 1 | VL |
| MHOM/TN/87/KA413 | Tunisia | MON 1 | VL |
| MHOM/TN/89/AFEF | Tunisia | MON 1 | VL |
| MCAN/LB/96/Trip6 | Lebanon | — | CVL |
| MHOM/FR/90/LEM2154 | France | MON 1 | VL |
| MHOM/FR/87/LEM1121 | France | MON 1 | VL |
| MHOM/FR/91/LEM2234 | France | MON 1 | VL |
| MCAN/FR/82/LEM382 | France | MON 1 | CVL |
| MHOM/ES/97/LLM637 | Spain | MON 34 | VL |
| MHOM/ES/97/LLM662 | Spain | MON 1 | VL |
| MHOM/ES/98/LLM778 | Spain | MON 24 | VL |
| MCAN/ES/98/LLM805 | Spain | MON 1 | CVL |
| MHOM/DZ/95/LIPA439 | Algeria | MON 24 | CL |
| MHOM/DZ/95/LIPA441 | Algeria | MON 24 | CL |
| MHOM/DZ/95/LIPA447 | Algeria | MON 24 | CL |
| MHOM/DZ/95/LIPA451 | Algeria | MON 24 | VL |
| MHOM/DZ/96/LIPA469 | Algeria | MON 24 | VL |
| MHOM/DZ/96/LIPA474 | Algeria | MON 24 | VL |
| MHOM/DZ/96/LIPA487 | Algeria | — | VL |
| MHOM/DZ/96/LIPA489 | Algeria | MON 1 | VL |
| MHOM/DZ/96/LIPA516 | Algeria | MON 1 | VL |
| *L. donovani* | | | |
| MHOM/IN/00/DEVI | India | MON 2 | VL |
| MHOM/IN/80/DD8† | India | MON 2 | VL |
| MHOM/ET/67/HU3 | Ethiopia | MON 18 | VL |
| MHOM/ET/84/ADDIS164 | Ethiopia | MON 83 | VL |
| IMRAT/KE/62/LRC-L57 | Kenya | MON 37 | NA |
| MHOM/KE/75/H9 | Kenya | MON 32 | VL |
| MHOM/SA/81/JEDDAH-KA | Saudi Arabia | MON 31 | VL |
| *L. chagasi* | | | |
| MHOM/BR/74/pp75† | Brazil | MON 1 | VL |
| *L. archibaldi* | | | |
| MHOM/ET/72/GEBRE1† | Ethiopia | MON 82 | VL |
| *L. major* | | | |
| MHOM/SU/73/5-ASKH† | USSR | MON 4 | CL |
| *L. tropica* | | | |
| MHOM/SU/74/SAF-K27† | USSR | MON 60 | CL |
| *L. aethiopica* | | | |
| MHOM/ET/72/L100† | Ethiopia | MON 14 | CL |

* VL, Visceral leishmaniasis; CL, Cutaneous leishmaniasis; CVL, Canine visceral leishmaniasis.
† WHO reference stocks, – not provided.

Tris–HCl, pH 8, 1 mM $MgCl_2$, 200 $\mu$M each of dATP, dCTP, dGTP and dTTP (Boehringer), 1·5 U *Taq* DNA polymerase (Eurogentec) and 400 nM of each primer. Thermal cycling parameters were: initial denaturation of 4 min at 94 °C, denaturation at 94 °C for 1 min, annealing at 69 °C for 1 min 30 sec and extension at 72 °C for 2 min for 30 cycles, followed by a final extension at 72 °C for 8 min. Oligonucleotides corresponding to the reverse sequences of SG1 and SG2 primers were used to amplify the inter-genic region of gp63 genes (inter-genic gp63 PCR). The sequences of these oligonucleotides (PDD1 and PDD2) are the following: 5′ TCGGTGAGGTCCTCGGTGGAGAC 3′ (PDD1) and 5′ CTTCGAGGAGGGCGGCTAC-ATCA 3′ (PDD2). Because these primers are located within the coding region at 327 and 137 bp from the start and stop codons respectively, a part of the coding region was also amplified with the inter-genic region (Fig. 1). Amplification reactions of the inter-genic genes were performed in the same conditions as for the intra-genic gp63 PCR, except for $MgCl_2$ (1·5 mM final concentration was necessary) and dimethyl sulfoxide (DMSO) (10% added to all reactions, in order to optimize amplifications). Thermal cycling parameters were: initial

https://doi.org/10.1017/S0031182000007125 Published online by Cambridge University Press

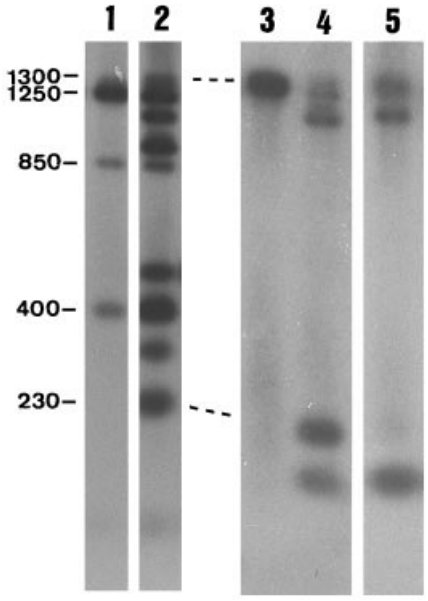


Fig. 2. Intra-genic gp63 PCR/*Hinc*II patterns of Old World complex species. Lane 1, *Leishmania infantum* (IPT1); Lane 2, *L. donovani* (DD8); Lane 3, *L. major* (5-ASKH); Lane 4, *L. tropica* (SAF-K27); Lane 5, *L. aethiopica* (L100). (All marked sizes are in bp.)

denaturation of 4 min at 94 °C, denaturation at 94 °C for 1 min, annealing at 65 °C for 2 min and extension at 72 °C for 2 min 30 sec for 35 cycles, followed by a final extension at 72 °C for 8 min. In order to evaluate sensitivity of the intra-genic gp63 PCR, serial dilutions of genomic DNA (500, 250, 100, 10, 1, 0·1, 0·01 and 0·001 pg) were subjected to PCR. Increase of the number of cycles (40) as well as addition of 5% of DMSO were necessary. In all PCR assays, contamination was monitored by a negative control at every PCR run, the product size was checked by electrophoresis in 1·5% agarose gel (Eurogentec) and reaction tubes were stored at 4 °C prior to analysis. Amplification products were digested overnight with an excess of enzymes with the following restriction enzymes, according to the manufacturer's instructions (Eurogentec): *BsiE*I, *BsmB*I, *Msc*I, *Hinc*II and *Sal*I for intra-genic amplicons and *Ava*I, *Kpn*I, *PflM*I and *Sph*I for inter-genic amplicons. Electrophoresis was performed overnight in 3% small fragment agarose gels (Eurogentec) for PCR–RFLP analyses. DNA fragments were sized using 1 kb, 100 bp and 200 bp DNA ladders (MBI and Promega). In order to verify the reproducibility of the patterns obtained, at least 2 digestions of each amplification product were performed.

### *Cluster analysis*

Jaccard distance (Jaccard, 1908) was calculated from RFLP data according to the following formula: $D_{ij} = 1-[a/(a+b+c)]$ where $a$ = number of bands common to $i$ and $j$ stocks, $b$ = number of bands present in $i$ stock and absent in $j$ stock, $c$ = number of bands present in $j$ stock and absent in $i$ stock. Jaccard distances and file transformation for agglomeration were computed with the program package TAXO (E. Serres, Laboratoire d'Ecologie Médicale et de Pathologie Parasitaire, Faculté de Medecine, Montpellier, France). Agglomeration of the distances and drawing of the dendrograms were performed with the Fitch-Margoliash (Felsenstein, 1984) and Drawtree programs, respectively, part of the PHYLIP program package (version 3.573c, Felsenstein, University of Washington, 1984). Statistical significance of each cluster was tested by comparing the means of genetic distance (i) between all pairs of stocks of that cluster, and (ii) between all these stocks and those of other clusters. Therefore, the Kruskal–Wallis method (EPI INFO program) was used.

## RESULTS

### *Intra-genic gp63 PCR–RFLP*

PCR amplification of the gp63 coding region (intra-genic gp63 PCR) was applied to the stocks belonging to the *L. donovani* complex (26 *L. infantum*, 7 *L. donovani*, 1 *L. chagasi* and 1 *L. archibaldi*), as well as to 3 reference strains of *L. major*, *L. tropica* and *L. aethiopica* (see Table 1). A single 1·3 kb band was observed in all stocks (not shown) that corresponded to the length expected between the annealing sites of the 2 primers on the known gp63 gene sequence of *L. infantum* (Gonzàlez-Aseguinolaza *et al.* 1997; Fig. 1). Sensitivity of the detection was found to reach 0·01 pg (not shown). Specificity of the target was confirmed by the strong hybridization under high stringency conditions with the *L. chagasi* gp63 probe.

In order to explore polymorphism within the 1·3 kb amplification fragment, the latter was cleaved (PCR–RFLP) with 5 restriction enzymes (*Sal*I, *Msc*I, *Hinc*II, *BsmB*I and *BsiE*I) known from data base sequences to cut either once or twice within the gp63 gene coding region of *L. infantum* (Gonzàlez-Aseguinolaza *et al.* 1997; Fig. 1). In most cases, an uncut 1·3 kb band was found together with cleaved fragments following restriction. In order to exclude incomplete digestion, the uncut 1·3 kb band was excised, purified, re-amplified and re-incubated with the same restriction enzyme. Complex PCR–RFLP patterns were obtained and showed to be discriminative at 3 taxonomic levels.

First, the 35 stocks of the *L. donovani* complex were quite distinct from the *L. major*, *L. tropica* and *L. aethiopica* reference strains. With *Hinc*II for instance, all *L. donovani* s. l. stocks shared an 850 bp fragment absent in the separate groups (Fig. 2). The amplification product from *L. major* stock remained uncut (track 3 in Fig. 2). Secondly, within the

https://doi.org/10.1017/S0031182000007125 Published online by Cambridge University Press

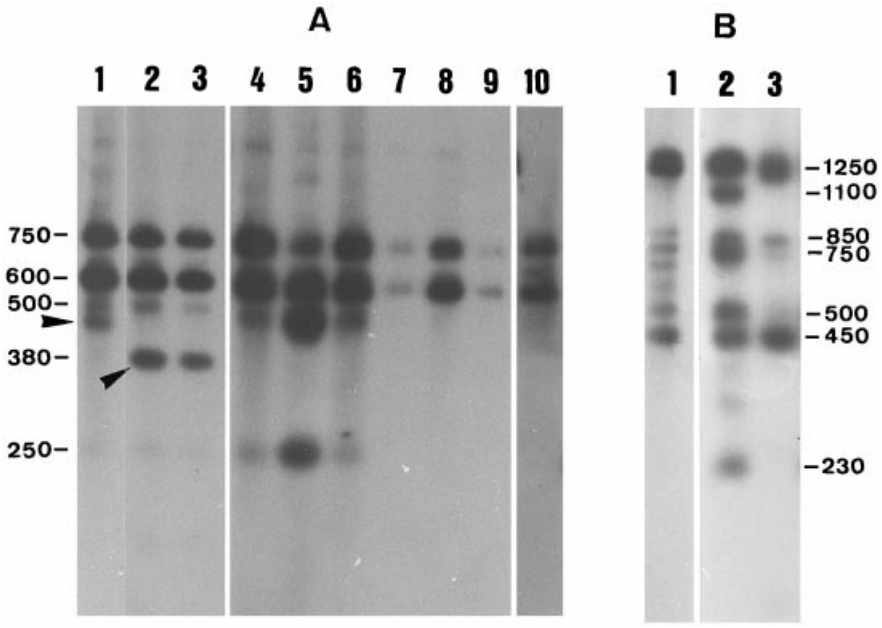


Fig. 3. Intra-genic gp63 PCR–RFLP in *Leishmania donovani* complex. (A) *Msc*I patterns. Lane 1, *L. chagasi* (pp75); Lanes 2–3, *L. infantum* IPT1 and KA412; Lanes 4–10, *L. donovani* DEVI, HU3, DD8, JEDDAH-KA, H9, ADDIS164 and LRC-L57. Arrow heads indicate *L. chagasi*-specific band (480 bp) and *L. infantum*-specific band (380 bp) respectively. (B) *Sal*I patterns. Lane 1, *L. infantum* (IPT1); Lane 2, *L. archibaldi* (GEBRE1); Lane 3, *L. donovani* (LRC-L57). (All marked sizes are in bp.)

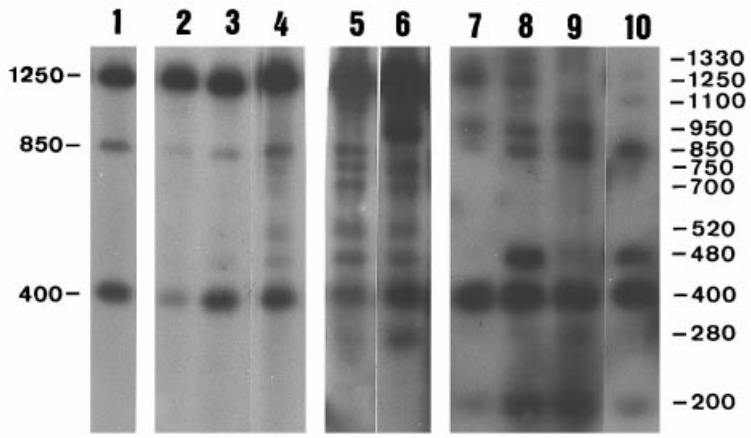


Fig. 4. Intra-specific polymorphism within *Leishmania infantum* stocks as shown by intra-genic gp63 PCR/*Hinc*II patterns. Lane 1, *L. infantum* from Tunisia (IPT1); Lanes 2–4, *L. infantum* from France (LEM2234, LEM2154 and LEM1121); Lanes 5–6, *L. infantum* from Spain (LLM662 and LLM637); Lanes 7–10 *L. infantum* from Algeria (Lipa474, Lipa487, Lipa 441 and Lipa 439). The 80 bp-sized band is not shown. (All marked sizes are in bp.)

*L. donovani* complex, specific differences were displayed by the 4 species examined. Using *Msc*I for instance, *L. infantum* and *L. donovani* could be distinguished by a 380 bp band specific to *L. infantum* (Fig. 3A). This fragment was absent in the *L. chagasi* reference strain, which showed an additional specific fragment at 480 bp (Fig. 3A). *L. archibaldi* was distinguished from *L. infantum* and *L. donovani* by 2 specific *Sal*I fragments (1100 and 230 bp, Fig. 3B).

Thirdly, among both the *L. infantum* and *L. donovani* s. s. species, an important intra-specific polymorphism was observed. Among *L. donovani* s. s. 2 groups were distinguished: *Msc*I patterns, for example, showed 500 and 250 bp-sized fragments which were only present in some stocks and absent in the others (see Lanes 4–10 in Fig. 3A). The same clustering was observed with the other restriction enzymes used in this analysis, except for *BsmB*I enzyme (data not shown). Among *L. infantum* stocks intra-specific polymorphism was best illustrated by the *Hinc*II patterns: further to the 80 (not shown), 400, 850 and 1250 bp bands encountered in most stocks, several additional fragments (750, 700, 520 and 480 bp) were present in some stocks from France and Spain (Fig. 4). In the case of 1 stock from

https://doi.org/10.1017/S0031182000007125 Published online by Cambridge University Press

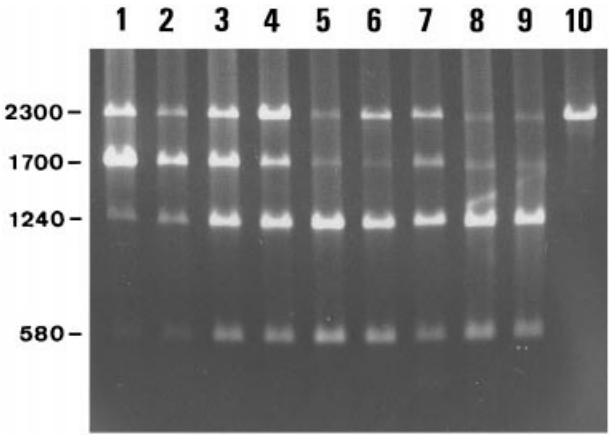


Fig. 5. Intraspecific polymorphism within *Leishmania infantum* as shown by inter-genic gp63 PCR/*Pfl*MI patterns. Lanes 1–2, Tunisian stocks KA413 and CN64; Lanes 3–4, French stocks LEM 2154 and 1121; Lanes 5–6, Spanish stocks LLM 637 and 805; Lanes 7–9, Algerian stocks LIPA 469, 487 and 451; Lane 10, Lebanese stock Trip6. (All marked sizes are in bp.)

Spain, the 850 bp band was absent while 950 and 280 bp-sized fragments were additionally present. Among the Algerian stocks, bands of 1100, 950, 480 and 200 bp were either present or absent (Fig. 4).

### *Inter-genic gp63 PCR–RFLP*

The regions between gp63 tandem repeats contain untranslated sequences as well as regulatory genes (McCoy *et al*. 1998), and might be under a different selective pressure than the gp63 encoding regions. Therefore, inter-genic PCR–RFLP was performed in all species included in this study.

Within all stocks 2 fragments were amplified: a band at 1·7 kb corresponding to the expected size between the annealing sites of the 2 primers on the sequence of *L. infantum* (according to Gonzàlez-Aseguinolaza *et al*. 1997; Fig. 1) and a second band of 2·3 kb, suggesting the presence of larger spacers separating some gene copies. The *L. infantum* stock from Lebanon displayed a unique pattern with 1 single amplification product band of 2·3 kb (not shown). Patterns were reproducible and identity to gp63 of both amplification products was demonstrated by hybridization with the *L. chagasi* gp63 probe.

The amplified fragments were then cleaved with *Ava*I, *Kpn*I, *Pfl*MI and *Sph*I enzymes, known to cleave either once (*Ava*I, *Kpn*I and *Pfl*MI) or 3 times (*Sph*I) within the inter-genic, but not the intra-genic regions, according to *L. infantum* sequence (Gonzàlez-Aseguinolaza *et al*. 1997; Fig. 1). In contrast to the extensive polymorphism revealed by intra-genic gp63 PCR–RFLP, less variable patterns have been found with inter-genic gp63 PCR–RFLPs. The same *Pfl*MI patterns, for instance, were found in all *L. infantum* stocks analysed, but the Lebanese one (Fig. 5). With *Ava*I, *Kpn*I and *Sph*I, polymorphic profiles were only encountered within the Algerian stocks (not shown).

### *Cluster analysis of PCR–RFLP data*

In order to integrate all previous data and to analyse the genetic relationships between the parasite stocks considered in this study, cluster analysis of intra-genic PCR–RFLP patterns was performed with the Fitch–Margoliash method. On the resulting intra-genic dendrogram (Fig. 6), the *L. donovani* complex clustered separately from the 3 separate groups: *L. major*, *L. tropica* and *L. aethiopica* whose reference stocks branched together. Within the *L. donovani* complex, a *L. donovani* s. s. cluster was individualized ($P < 0{\cdot}05$), with further subdivision into the 2 groups mentioned above. *L. archibaldi* species was found to be branched between these two groups in this *L. donovani* s. s. cluster. Within *L. infantum*, 2 major clusters appeared. One of these clusters was constituted by all but 1 Algerian stock and branched between *L. donovani* s. s. cluster and other *L. infantum* stocks (similar distance *versus* each of these 2 groups, $P = 0{\cdot}39$). The second *L. infantum* cluster was extensively spread out into groups that corresponded to the geographical origin of the parasites (Tunisia, France and Spain) with 2 exceptions (1 French and 1 Algerian stock). Interestingly, the representatives of each of the 2 major *L. infantum* zymodemes here considered (MON1 and MON24) did not cluster together, but were dispersed among the different *L. infantum* clusters. *L. chagasi* was positioned on a separate branch well individualized from the other *L. infantum* clusters.

When considering *L. infantum* stocks only, the resulting intra-genic dendrogram (Fig. 7A) showed again well-clustered groups corresponding to the geographical origin of the stocks ($P < 0{\cdot}05$). In contrast, such a distribution was not found within the inter-genic dendrogram (Fig. 7B). Indeed, only 2 clusters were individualized: on the one hand, Tunisian, French and Spanish stocks grouped together and on the other hand, Algerian stocks positioned on a separate branch (with the exception of 2 stocks; see Fig 7B). In addition, the Lebanese stock, which clustered within the same group as the Tunisian ones in the intra-genic dendrogram, was well individualized in a separate branch within the inter-genic dendrogram (see Fig. 7B). In order to compare the level of polymorphism revealed by both PCR methods, the number of different *L. infantum* genotypes was counted. This value was higher for the intra- (13 for 5 restriction enzymes, average of 2·6 per restriction enzyme) than the inter- (6 for 4 restriction enzymes, average of 1·5 per restriction enzyme) genic gp63 PCR–RFLP.

https://doi.org/10.1017/S0031182000007125 Published online by Cambridge University Press

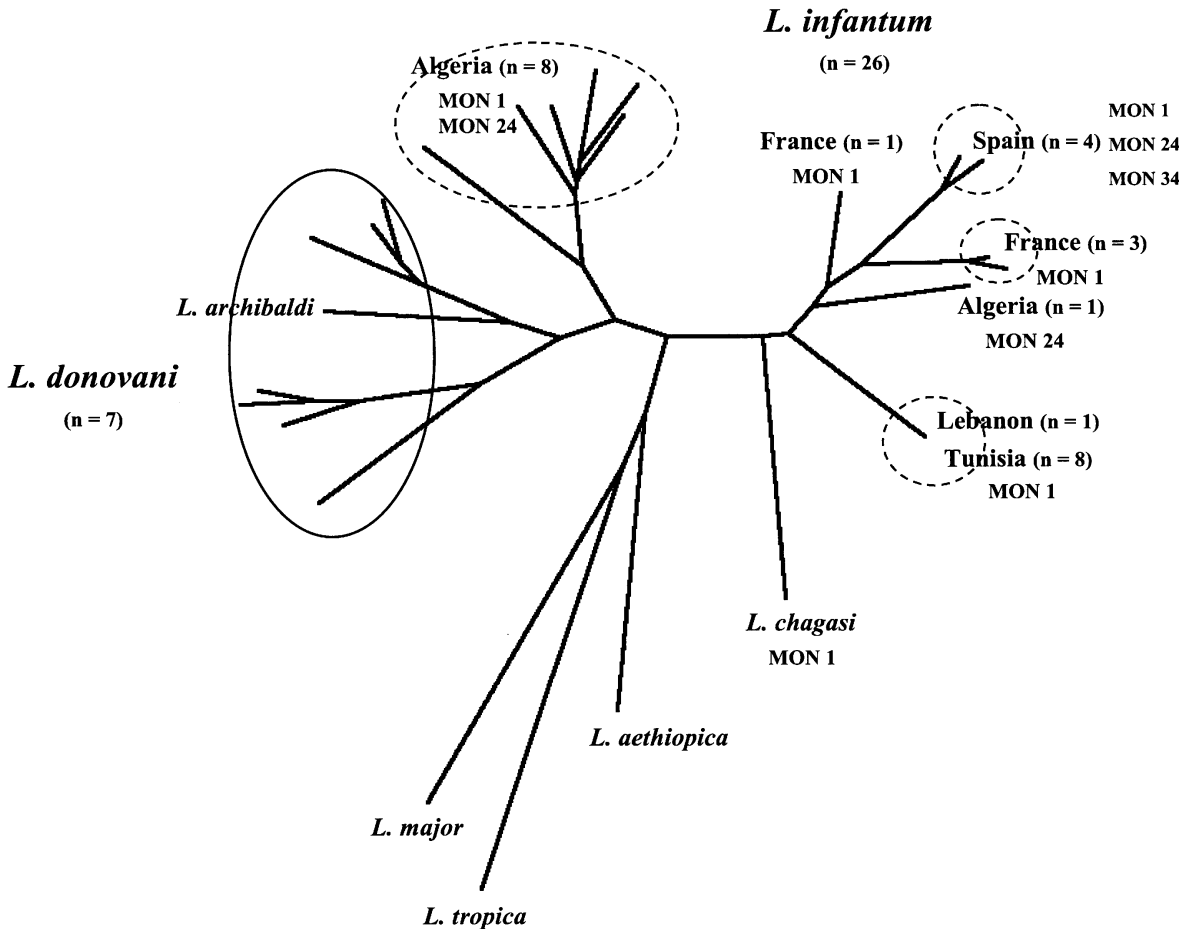


Fig. 6. Fitch-Margoliash dendrogram constructed from matrix of Jaccard distances calculated from data of intra-genic gp63 PCR–RFLP with *BsmB*I, *BsiE*I, *Hinc*II, *Msc*I and *Sal*I. *n*, Number of isolates; MON, zymodeme according to the Montpellier classification, (-----) *Leishmania infantum*; (——) *L. donovani*.

## DISCUSSION

The aim of the present study was to analyse a panel of isolates of the *L. donovani* complex, using genetic characters supposed to be under strong host-selective pressure, and to evaluate how the ensuing sample structuring would correlate with that obtained elsewhere from isoenzymes, characters which are considered to vary neutrally. Therefore, the genes coding for the major surface glycoprotein of *Leishmania* (gp63) were amplified by PCR and cleaved with restriction enzymes (intra-genic gp63 PCR–RFLP). An additional reference point was introduced by PCR–RFLP analysis of gp63 inter-genic regions, which are expected to be under a different selective pressure than the gp63 encoding regions (non-translated spacers and regulatory genes, McCoy *et al*. 1998).

A higher polymorphism was revealed within the gp63 encoding region than in the inter-genic regions. Three results deserved particular attention on dendrograms built-up from intra-genic data. (1) A taxonomic structure was observed in our sample. On the one hand, the whole *L. donovani* complex was distinct from the outgroups (*L. major*, *L. tropica* and *L. aethiopica*). On the other hand, the 4 species constituting the *L. donovani* s. l. complex were discriminated, with *L. infantum* clustering together with *L. chagasi*, and *L. archibaldi* with *L. donovani* s. s., as in MLEE studies (Rioux *et al*. 1990; Thomaz-Soccol *et al*. 1993). The discrimination of *L. infantum* and *L. chagasi*, considered as synonymous (Beverley, Ismach & McMahon Pratt 1987; Rioux *et al*. 1990; Van Eys *et al*. 1991; Fernandez *et al*. 1994; Mauricio *et al*. 1999), should be further evaluated by testing additional *L. chagasi* stocks. (2) Within *L. infantum* stocks a geographical structuring appeared on the dendrogram; interestingly, Algerian stocks clustered closer to *L. donovani* s. s. than to other *L. infantum* clusters. (3) Representatives of each of the main *L. infantum* zymodemes (MON1 and MON24) did not cluster together, but were scattered across the different branches of the dendrogram. Our intra-genic data should not be used for phylogenetic interpretation, as (i) they correspond to single-locus analyses, and (ii) some of the genetic variation might be adaptive, as suggested by the geographical structuring of gp63-encoding genes and the low overlapping with the structure of zymodemes and gp63 inter-genic regions. Considering the nature of gp63, two selective factors might theoretically play an important role.

https://doi.org/10.1017/S0031182000007125 Published online by Cambridge University Press

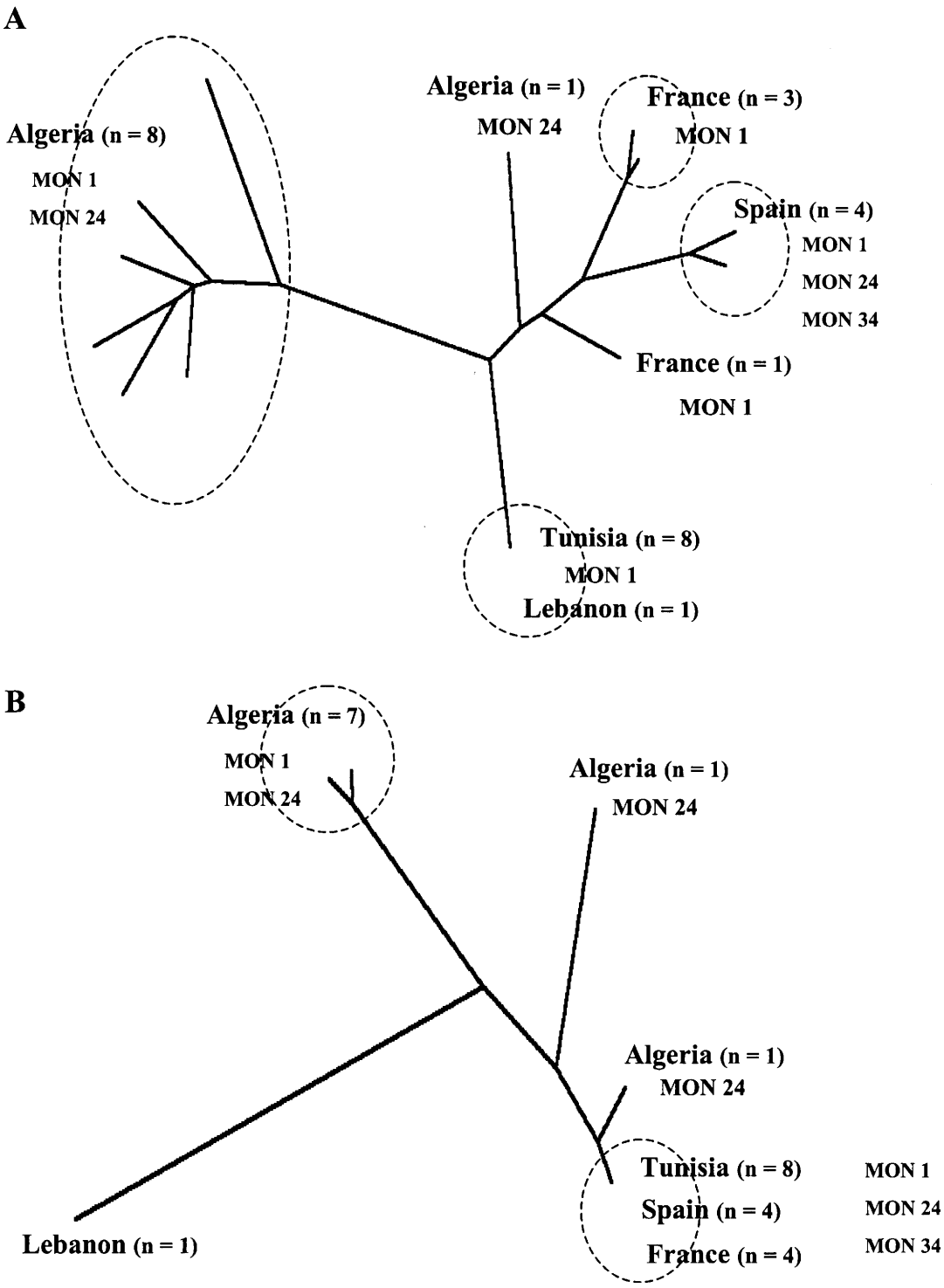


Fig. 7. Fitch-Margoliash dendrogram of *Leishmania infantum* stocks constructed from matrix of Jaccard distances calculated from data of (A) intra-genic gp63 PCR–RFLP with *BsmB*I, *BsiE*I, *Hinc*II, *Msc*I and *Sal*I and (B) inter-genic gp63 PCR–RFLP with *Ava*I, *Kpn*I, *PflM*I and *Sph*I. *n*, Number of isolates; MON, zymodeme according to the Montpellier classification.

First, as proteinases similar to gp63 have been encountered in monoxenic insect parasites like *Crithidia* and *Herpetomonas* (Medina-Acosta *et al.* 1993), a specific interaction between gp63 of promastigotes and vector components might be considered. *Phlebotomus perniciosus* is an important vector species of *L. infantum* in Algeria, Tunisia, France and Spain, but other vectors have been reported: *P. ariasi* (France (Killick-Kendrick, 1987) and Spain (Arnedo *et al.* 1994)) and *P. perfiliewi* (Algeria; Izri & Belazzoug, 1993). The potential selection by the vector has already been proposed for the New World species *L.* (*V.*) *peruviana*, in which eco-geographical populations were demonstrated by

https://doi.org/10.1017/S0031182000007125 Published online by Cambridge University Press

gp63 PCR–RFLP, and correlated with the presence of different sandfly vectors (Victoir *et al*. 1998). In addition, for other surface molecules like lipophosphoglycan (LPG), it was shown that the vectorial competence of phlebotomine sandflies for different species of *Leishmania* was controlled by structural polymorphism of these molecules (Pimenta *et al*. 1994). Further studies should be done on parasites obtained from different sandfly samples in order to document such an association in the case of gp63 proteins.

Secondly, because of the strong immunogenic potential of the gp63 antigen (Rivier *et al*. 1999), immune selection might play an important role. Gp63 has shown inter- and intra-specific polymorphism at amino acid and structural levels, mainly in the surface residues (Schlagenhauf, Etges & Metcalf, 1998), and some of our polymorphic restriction sites were found to correspond to these regions highly accessible to immune attack (not shown). Furthermore, according to the mathematical models of Gupta & Anderson (1999), a difference in the population structure generated by the analysis of (polymorphic) antigen and housekeeping genes would be indicative of the polymorphic nature of gp63 immunogens. This is supported by previous studies of canine visceral leishmaniasis, which showed recognition of gp63 by canine sera to be heterogeneous, reactivity being preferentially observed against most divergent regions of the glycoprotein (Morales *et al*. 1997). DNA sequencing of gp63 genes and immunogenetic analyses among human and canine populations should further confirm our results. Definitive answers to these questions are most important for the design of gp63 vaccines (Rivier *et al*. 1999) and the assessment of vaccine trials.

In the context of PCR-characterization of *Leishmania*, our results also highlight the innovating nature of gp63 PCR. The most commonly used PCR assays have been targeting either kDNA minicircles or specific nuclear genomic sequences. Major applications were (i) detection through the high sensitivity (1 fg) of kDNA minicircle PCR (De Bruijn & Barker, 1992), (ii) discrimination of species complexes (rDNA intergenic regions (Cupolillo *et al*. 1995) or mini-exon spacers (Fernandez *et al*. 1994)) or New world species (Cupolillo *et al*. 1995), or (iii) parasite fingerprinting (Noyes *et al*. 1998). GP63 intra-genic PCR combines several advantages of these methods: sensitivity was estimated to 0·01 pg (the theoretical equivalent of 1/10 parasite (Leon, Fontes & Manning, 1978)) and, in our sample, all species could be discriminated. However, the additional advantage is the capacity of probing – directly in the biopsies – genetic variability of molecules involved in immunopathology, thus to have a new insight on clinical polymorphism. This was already documented in the New World where significant gp63 gene differences were found between the metastasizing *L*. (*V*.) *braziliensis* and the strictly cutaneous *L*. (*V*.) *peruviana* (Victoir *et al*. 1995) parasites.

Finally, our results may also shed light in the genomic organization of gp63 genes, especially in *L*. *infantum*. Indeed, in that species, 6 copies of the same gene have been reported (Gonzàlez-Aseguinolaza *et al*. 1997). In many of the stocks here analysed, when summing the size of intra-genic PCR–RFLP fragments for a given restriction enzyme (with all the necessary controls for completion of digestion), multiples of the size of the amplicon were obtained, suggesting a higher gene heterogeneity than expected for that species, but equivalent to the variability described in *L*. *chagasi* (Roberts *et al*. 1993). Further work is required in order to determine whether the variants identified by our method actually correspond to different gp63 gene families expressed at different moments of the parasite's life-cycle, as reported in *L*. *chagasi* (Ramamoorthy *et al*. 1995).

We are grateful to Professor M. E. Wilson, University of Iowa, USA, for kindly providing the gp63 probe of *L*. *chagasi*; Professor J. P. Dedet and Dr C. Blaineau, Laboratoire de Parasitologie, Faculté de Médecine, Montpellier, France and Dr J. Alvar, Instituto de Salud Carlos III, Spain, who kindly provided *L*. *donovani* and French and Spanish *L*. *infantum* parasites, respectively. This work received financial support from the INCO-DC DGXII-program of the European Community (Contract No. IC18-CT98-0256), from the SERST-PNM96 program (Tunisia) contract No. P96BSP46 and from the Agence Universitaire de le Francophonie (AUPELF-UREF, programme PFR-PDR).

## REFERENCES

ARNEDO, P. A., BELLIDO BLASCO, J. B., GONZALEZ MORAN, F., ARIAS SANCHEZ, A., CALVO MAS, C., SAFONT ADSUARA, L., FABRA PEIRAT, E., CRIADO JUAREZ, J. & PONS ROIG, P. (1994). Leishmaniasis in Castellon: an epidemiological study of human cases, the vector and the canine reservoir. *Revista de Sanidad e Higiene Publica, Madrid* **68**, 481–491.

BEVERLEY, S. M., ISMACH, R. B. & MCMAHON PRATT, D. (1987). Evolution of the genus *Leishmania* as revealed by comparison of nuclear DNA restriction fragment patterns. *Proceedings of the National Academy of Sciences, USA* **84**, 484–488.

BUTTON, L. L., RUSSELL, D. G., KLEIN, H. L., MEDINA-ACOSTA, E., KARESS, R. E. & MCMASTER, W. R. (1989). Genes encoding the major surface glycoprotein in *Leishmania* are tandemly linked at a single chromosomal locus and are constitutively transcribed. *Molecular and Biochemical Parasitology* **32**, 271–284.

CUNY, G., VEAS, F. & ROIZÈS, G. (1991). 'Pocket blotting': a method for transferring nucleic acids onto nylon membranes. *Annals of Biochemistry* **193**, 45–48.

CUPOLILLO, E., GRIMALDI, G., MOMEN, H. & BEVERLEY, S. M. (1995). Intergenic region typing (IRT): A rapid

molecular approach to the characterisation and evolution of *Leishmania*. *Molecular and Biochemical Parasitology* **73**, 145–155.

DE BRUIJN, M. H. & BARKER, D. C. (1992). Diagnosis of New World leishmaniasis: specific detection of species of the *Leishmania braziliensis* complex by amplification of kinetoplast DNA. *Acta Tropica* **52**, 45–58.

FELSENSTEIN, J. (1984). Distance methods for inferring phylogenies: a justification. *Evolution* **38**, 16–24.

FERNANDEZ, O., MURTHY, V., KURATH, U., DEGRAVE, W. & CAMPBELL, D. (1994). Mini-exon gene variation in human pathogenic *Leishmania* species. *Molecular and Biochemical Parasitology* **66**, 261–271.

GONZÀLEZ-ASEGUINOLAZA, G., ALMAZAN, F., RODRIGUEZ, J. F., MARQUET, A. & LARRAGA, V. (1997). Cloning of the gp63 surface protease of *Leishmania infantum* differential post-translational modifications correlated with different infective forms. *Biochemica et Biophysica Acta* **1361**, 92–102.

GUPTA, S. & ANDERSON, R. (1999). Population structure of pathogens: the role of immune selection. *Parasitology Today* **15**, 497–501.

IZRI, M. A. & BELAZZOUG, S. (1993). *Phlebotomus* (*Laroussius*) *perfiliewi* naturally infected with dermotropic *Leishmania infantum* at Tenes, Algeria. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **87**, 399.

JACCARD, P. (1908). Nouvelles recherches sur la distribution florale. *Bulletin de la Societé Vaudoise de Sciences Naturelles* **44**, 223–270.

KILLICK-KENDRICK, R. (1987). Breeding places of *Phlebotomus ariasi* in the Cévennes focus of leishmaniasis in the south of France. *Parasitologia* **29**, 181–191.

LEON, W., FONTES, D. L. & MANNING, J. (1978). Sequence arrangement of the 16S and 26S rRNA genes in the pathogenic haemoflagellate *Leishmania donovani*. *Nucleic Acids Research* **5**, 491–503.

MAURICIO, I. L., HOWARD, M. K., STOTHARD, J. R. & MILES, M. A. (1999). Genomic diversity in the *Leishmania donovani* complex. *Parasitology* **119**, 237–246.

MCCOY, J. J., BEETHAM, J. K., OCHS, D. E., DONELSON, J. E. & WILSON, M. E. (1998). Regulatory sequences and a novel gene in the msp (GP63) gene cluster of *Leishmania chagasi*. *Molecular and Biochemical Parasitology* **95**, 251–265.

MEDINA-ACOSTA, E., BEVERLEY, S. M. & RUSSELL, D. G. (1993). Evolution and expression of the *Leishmania* surface proteinase (gp63) gene locus. *Infectious Agents and Disease* **2**, 25–34.

MORALES, G., CARRILLO, G., REQUENA, J. M., GUZMAN, F., GOMEZ, L. C., PATARROYO, M. E. & ALONSO, C. (1997). Mapping of the antigenic determinants of the *Leishmania infantum* gp63 protein recognized by antibodies elicited during canine visceral leishmaniasis. *Parasitology* **114**, 507–516.

MORENO, G., RIOUX, J. A., LANOTTE, G., PRATLONG, F. & SERRES, E. (1986). Le complexe *Leishmania donovani* s. l. Analyse enzymatique et traitement numérique. Individualisation du complexe *Leishmania infantum*. Corollaires biogéographique et phylétique. A propos de 146 souches originaires de l'Ancien et du Nouveau Monde. In *Leishmania. Taxonomie et phylogenèse* (ed. Rioux, J. A.), pp. 105–117. IMEEE, Montpellier.

NEI, M. (1987). *Molecular Evolutionary Genetics*. Colombia University Press, New York.

NOYES, H. A., REYBURN, H., BAILEY, J. W. & SMITH, D. (1998). A nested-PCR-based schizodeme method for identifying *Leishmania* kinetoplast minicircles classes directly from clinical samples and its application to the study of the epidemiology of *Leishmania tropica* in Pakistan. *Jouranal of Clinical Microbiology* **36**, 2877–2881.

PIMENTA, P. F., SARAIVA, E. M., ROWTON, E., MODI, G. B., GARRAWAY, L. A., BEVERLEY, S. M., TURCO, S. J. & SACKS, D. L. (1994). Evidence that the vectorial competence of phlebotomine sandflies for different species of *Leishmania* is controlled by structural polymorphisms in the surface lipophosphoglycan. *Proceedings of the National Academy of Sciences, USA* **91**, 9155–9156.

RAMAMOORTHY, R., DONELSON, J. E., PAETZ, K. E., MAYBODI, M., ROBERTS, S. C. & WILSON, M. E. (1992). Three distinct RNAs for the surface protease gp63 are differentially expressed during development of *Leishmania donovani chagasi* promastigotes to an infectious form. *Journal of Biological Chemistry* **267**, 1888–1895.

RAMAMOORTHY, R., SWIHART, K. G., MCCOY, J. J., WILSON, M. E. & DONELSON, J. E. (1995). Intergenic regions between tandem gp63 genes influence the differential expression of gp63 RNAs in *Leishmania chagasi* promastigotes. *Journal of Biological Chemistry* **270**, 12133–12139.

RIOUX, J. A., LANOTTE, G., SERRES, E., PRATLONG, F., BASTIEN, P. & PERIERES, J. (1990). Taxonomy of *Leishmania*. Use of isoenzymes. Suggestions for a new classification. *Annales de Parasitologie Humaine et Comparée* **65**, 111–125.

RIVIER, D., BOVOY, P., SHAH, R., DIDISHEIM, S. & MAUEL, J. (1999). Vaccination against *Leishmania major* in a CBA mouse model of infection: role of adjuvant and mechanism of protection. *Parasite Immunology* **21**, 461–473.

ROBERTS, S. C., SWIHART, K. G., AGEY, M. W., RAMAMOORTHY, R., WILSON, M. E. & DONELSON, J. E. (1993). Sequence diversity and organization of the msp gene family encoding gp63 of *Leishmania chagasi*. *Molecular and Biochemical Parasitology* **62**, 157–172.

SCHLAGENHAUF, E., ETGES, R. & METCALF, P. (1998). The crystal structure of the *Leishmania major* surface proteinase leishmanolysin (gp63). *Structure* **6**, 1035–1046.

THOMAZ-SOCCOL, V., LANOTTE, G., RIOUX, J. A., PRATLONG, F., MARTINI-DUMAS, A. & SERRES, E. (1993). Monophyletic origin of the genus *Leishmania* Ross, 1903. *Annales de Parasitologie Humaine et Comparée* **68**, 104–106.

TOBIE, E. J., VON BRAND, T. & MEHLMAN, B. (1950). Cultural and physiological observations on *Trypanosoma rhodesiense* and *Trypanosoma gambiense*. *Journal of Parasitology* **36**, 48–54.

VAN EYS, G. J. J. M., GUIZANI, I., LIGTHART, G. S. & DELLAGI, K. (1991). A nuclear DNA probe for identification of strains within the *Leishmania donovani* complex. *Experimental Parasitology* **72**, 459–463.

VICTOIR, K., DUJARDIN, J. C., DE DONCKER, S., BARKER,

D. C., AREVALO, J., HAMERS, R. & LE RAY, D. (1995). Plasticity of gp63 gene organization in *Leishmania* (*Viannia*) *braziliensis* and *Leishmania* (*Viannia*) *peruviana*. *Parasitology* **111**, 265–273.

VICTOIR, K., BANULS, A. N., AREVALO, J., LLANOS-CUENTAS, A., HAMERS, R., NOËL, S., DE DONCKER, S., LE RAY, D., TIBAYRENC, M. & DUJARDIN, J. C. (1998). The gp63 gene locus, a target for genetic characterization of *Leishmania* belonging to subgenus *Viannia*. *Parasitology* **117**, 1–13.

WEBB, J. R., BUTTON, L. L. & MCMASTER, W. R. (1991). Heterogeneity of the genes encoding the major surface glycoprotein of *Leishmania donovani*. *Molecular and Biochemical Parasitology* **48**, 173–184.