J Vector Borne Dis 48, June 2011, pp. 67–71

Research Articles

Preliminary study on the galectin molecular diversity in Moroccoan *Phlebotomus papatasi* sandfly populations

S. Guernaoui[1], D. Garcia[1], A. Boumezzough[2], D. Fontenille[1] & D. Sereno[1]

[1]*MIVEGEC, Maladies Infectieuses: Vecteur, Evolution, Génétique, Ecologie Contrôle, Montpellier Cedex 5, France;* [2]*Laboratoire Ecologie-Environnement, Faculté des Sciences Semlalia, Université Cadi Ayyad, Département de Biologie, Marrakech, Morocco.*

ABSTRACT

Context: Galactose binding protein (PpGalec) plays an important role in the specificity of *Phlebotomus papatasi* sandfly for *Leishmania major*. The molecular diversity of this ligand is currently unknown but might have some influence on the ability of PpGalec to efficiently recognize *L. major* in natural sandfly populations.

Objective: To explore the molecular diversity of the *P. papatasi* Galectin gene (PpGalec) in natural sandfly population of Morocco.

Results & Conclusions: Sequence variations of PpGalec was analyzed in 31 *P. papatasi* specimens collected from endemic and non-endemic zoonotic cutaneous leishmaniasis foci of Morocco. Among the 211 amino acid positions analyzed, 11 are subjected to mutation. Interestingly, we observe that one mutation directly affect an amino acid known to be involved in the substrate recognition by galectin. The repercussion of this polymorphism on the capacity of the galectin to efficiently bind the *L. major* Lipophosphoglycane (LPG) awaits further investigations.

Key words Galectin binding protein; *Leishmania major;* molecular diversity; *Phlebotomus paptasi*

INTRODUCTION

Intraspecific genetic variations of sandfly populations linked to vector capacity and competence for *Leishmania* remained currently poorly studied. Approximately, 700 species of sandflies were described to date; of these 10% have been incriminated as *Leishmania* vectors. In many, if not all cases, sandfly distribution does not correlate well with leishmaniasis distribution. Indeed, *Phlebotomus sergenti* (Diptera: Psychodidae) has an extensive geographical distribution, wider than that of the anthroponotic cutaneous leishmaniasis (ACL). The presence of *P. sergenti* in *L. tropica* free areas might be related to the existence of some cryptic vector species differing in their vectorial capacity[1]. The distribution of *P. papatasi* extends from southern Europe and eastern regions to the Indian subcontinent. It thus, largely exceeds those of *L. major* and its reservoir hosts. Even locally, in Egypt for example, zoonotic cutaneous leishmaniasis (ZCL) is primarily present in northern Sinai while the vector distribution is more extensive. In Morocco, despite the wide distribution of both *P. papatasi* and *Meriones shawi*, ZCL caused by *L. major* are restricted to the pre-Saharan areas. The underlying factors that trigger such discrepancies might include ones linked to the capacity of *Leishmania* to accomplish its intra-vectorial developmental life cycle.

Leishmania resides continuously in the sandfly gut that represents an important interface for vector-parasite interaction. The developmental life cycle of *Leishmania* in their insect host includes a first initial step; the attachment of *Leishmania* to gut epithelial cells and then several rounds of proliferation and differentiation that ends with the appearance of a transmissible population. Among the sandflies transmitting *Leishmania*, *P. papatasi* is considered as highly genetically homogeneous, but the debate remains open[2-5]. *Phlebotomus papatasi* is a "selective" vector, which transmits only *L. major*. In this vector, parasite specificity is controlled by a stage-specific modification in the LPG (Lipophosphoglycane) adhesin. Modified PG repeats of *L. major* bearing β1,3 galactosyl residues side chain, are specifically recognized by a midgut LPG-receptor, PpGalec[6]. Conversely, variation in PG motif in natural *L. major* populations has consequences on the ability of the parasite to survive within *P. papatasi*[7]. Since specific attachment and development of *L. major* within its specific vector involved a highly specific ligand-receptor interaction, it can be hypothesized that in natural *P.*

papatasi population, molecular variation in the ligand (PpGalec) might affect its ability to efficiently recognize specific pattern of the *Leishmania* LPG. Currently, the molecular diversity of this important LPG receptor has not yet been investigated. We, therefore, explored the diversity of the galectin gene in natural populations of *P. papatasi* originating from the endemic, and non-endemic ZCL foci of Morocco.

MATERIAL & METHODS

Sandfly collection and dissection

Phlebotomus papatasi were collected using CDC trap in localities differing in their incidence for ZCL. Geographic references of the stations are given in Table 1. For species identification of *P. papatasi*, genitalia of males and females were mounted in Canada balsam.

DNA extraction and galectin gene amplification

DNA was extracted from 31 specimen, under sterile conditions, using Qiamp kit (Qiagen), according to the manufacturer's protocol. Purity and quantity of DNA were ascertained with a NanoDrop instrument (Thermo Scientific) and 40 ng of DNA was used as template for amplification with primers specific for the galectin gene, PpGalecFW-GTCAATTGGCATGTCCTGTG and PpGalecRev-CACATGATCAACCTCCGTGA. Amplifications were carried out using 50 µL reaction volume and the PCR mix containing 40 ng genomic DNA, 1X PCR buffer, 0.5 µM of each primer, 0.2 mM of each dNTP, 3% DMSO, and 0.02 U Phusion (Fininzyme). A negative PCR control was set up, by omitting the DNA. Samples were amplified using the following program: 98°C for 45 sec, and 35 cycles of 98°C for 10 sec, 55°C for 30 sec, 72°C for 30 sec and 72°C for 7 min last extension. Aliquots of 5 µl were analyzed by electrophoresis on 1% agarose gel containing ethidium bromide, in order to check the sizes of the amplicon after which samples were sent for sequencing.

Table 1: Distribution of sandfly sampling stations in Morocco

Site	Nb	Geo Ref	Alt (m)
Ouarzazate	12	30°55′24" N-6°53'12" O	1140
Oukaemiden	1	31°12′36" N-7°52′8" O	2500
N'Fifa	6	31°09′36" N-8°54′36" O	ND
Sti-Fadma (Setti-Fatma)	1	31°17′45" N-7°42′36" O	1500
Taferiat	1	31°31′58" N-7°36′29" O	760
Chouiter	2	31°33′35" N-7°49′0" O	543
Lijoukak	1	30°55′00"N-8°15′00" O	1200
Amizmiz	4	31°13′00" N-8°15′00" O	1004
Labrouj (El Borouj)	3	32°30′17" N-7°11′33" O	388

ND- Not determined.

Sequence analysis

For each nucleic acid sequence, the conceptual translation into protein sequence was performed with the application "translate to protein" of the CLC work package (http://www.clcbio.com/). Protein sequences were then truncated to a minimal of 211 amino acids common to all samples. They were aligned with other identified galectin proteins using the program pipe-align (*http://bips.u-strasbg.fr/PipeAlign/*). To visualize the diversity of the protein sequences, a neighbor-joining tree was built using the online program (*http://www.phylogeny.fr/*) with 500 bootstraps and the substitution matrix model of Dayhoff PAM matrix.

RESULTS & DISCUSSION

The galectin gene was amplified in a set of 31 *P. papatasi,* 13 from the southern and 18 from the northern part of the atlas. The amplification was also performed in another proven vector for *L. major*, *P. duboscqi*. Sequence analysis revealed the presence of an intron in the genomic copy of the gene, it was therefore removed before conceptual translation of the sequences. Among the 211 amino acids examined, 11 were subjected to mutations demonstrating that various isoforms of the galectin gene might be present in natural *P. papatasi* population (Fig. 1).

The *P. papatasi* galectin protein contains two Carbohydrate Recognition Domains (CRD) (conserved domain, galectin/galactose-binding lectin cd00070). In these domains, 8 amino acids are known to have a crucial role in the galectin function (dark grey bars in Fig. 1). Among the 11 substitutions observed, 2 of them that substitute a tyrosine (Y) into a phenylalanine (F) at position 26 and a serine (S) into a leucine (L) at position 37 (Fig. 1) impact directly an amino acid involved in the binding function of the galectin. These substitutions correspond to the amino acid positions 68 and 74 in the reference protein sequence AY538600. Interestingly, one of these two point mutations impact directly an amino acid involved in the discrimination of the carbohydrate nature of the galactin ligand Y68–>F68[8]. To what extent such substitution could modify the recognition or the affinity of the galectin for the *L. major* LPG is currently unknown and have to be further investigated.

To visualize more easily the diversity of the galectin

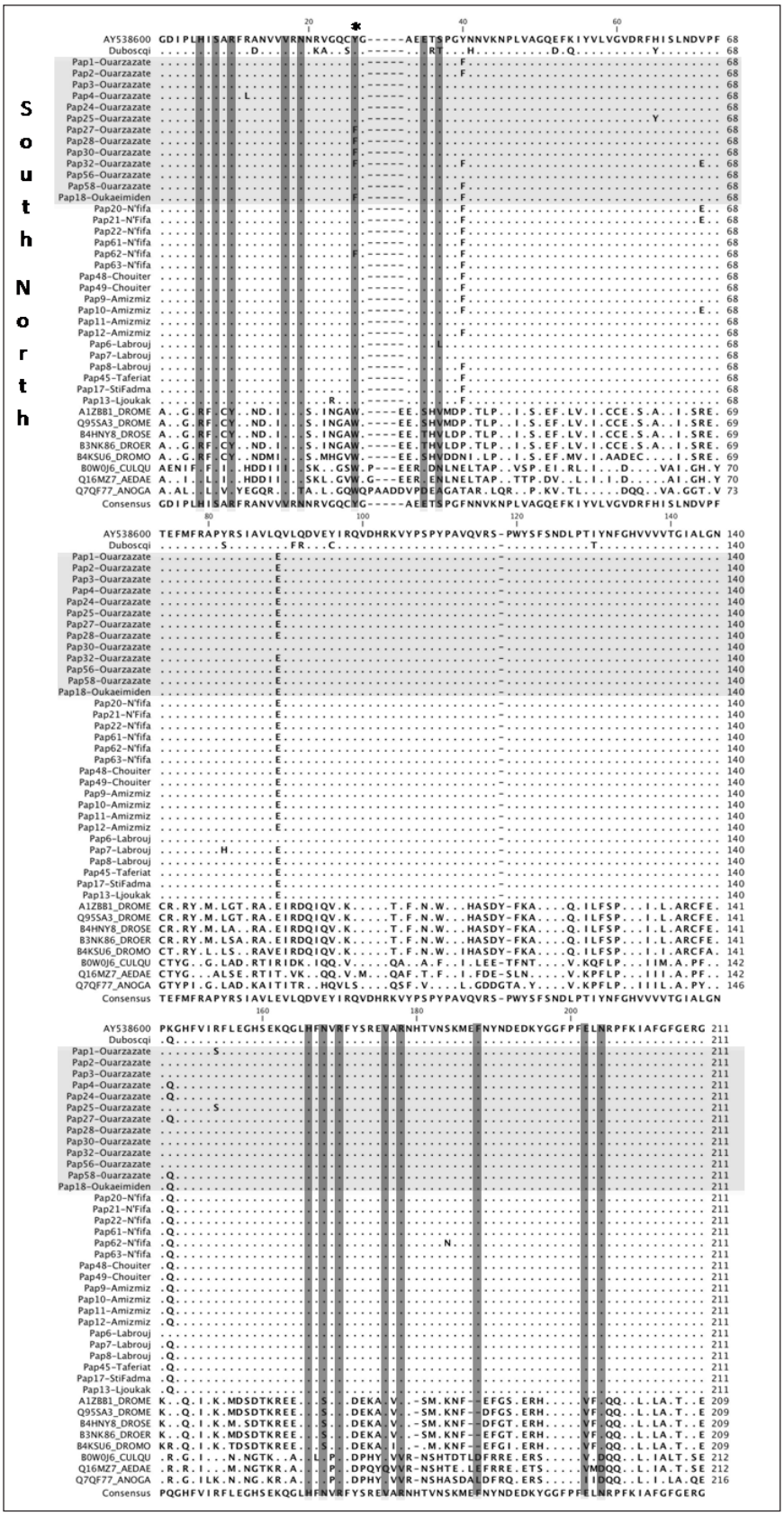

Fig. 1: Multiple alignments of the galectin sequences from different *P. papatasi* individuals and another species in comparison to the reference sequence AY538600. Only mutations are displayed. Residues identical to their corresponding ones in the reference sequence are replaced by dots. Numbers above the alignments indicate the residue positions. Dark grey bars correspond to the putative residues playing role in the galactose recognition.

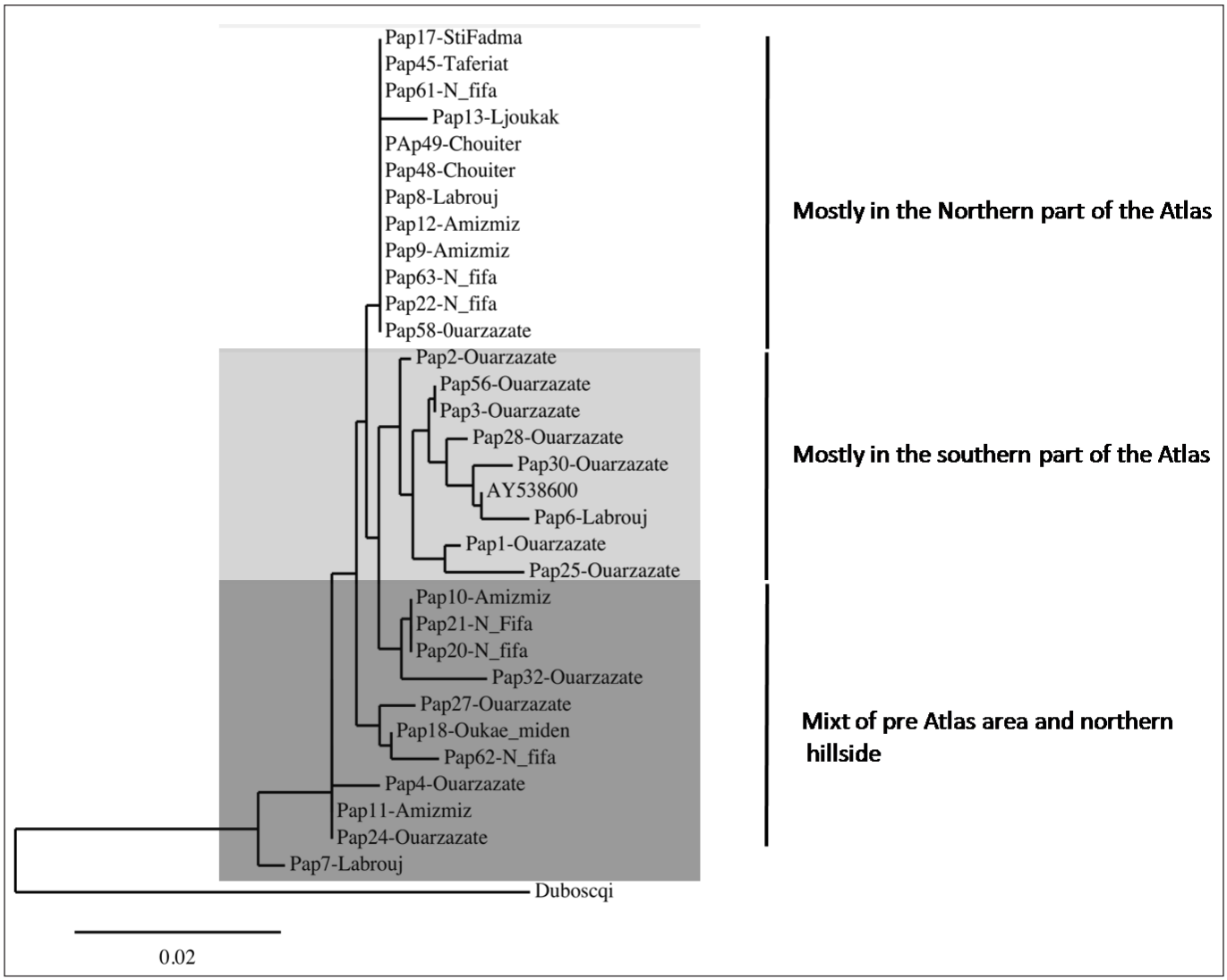

Fig. 2: Neighbor-joining tree derived from the distance matrix and based on the galectin protein sequence. *P. duboscqi* galectin sequence was used as an out-group.

proteins, a neighbor-joining tree was built. As shown in Fig. 2, roughly 3 groups of sequences can be individualized according to the sampling area. The molecular diversity of the galectin protein appears to be more important in sandfly population originating from the pre-atlas region (endemic zoonotic cutaneous leishmaniasis foci) than those collected in the northern part. Additionally, a third group, mostly composed by sandflies collected on the northern hillside of the Atlas is evidenced (i.e. Oukaemiden, Amizmiz, N'fifa). Further, in our sample, the substitutions of the tyrosine into phenylalanine at postion 68 appear to be more frequent in sandflies that originates from the southern part of Morocco (i.e. historical Moroccoan foci of ZCL). The origin of such distribution is unknown but might be the result of the spreading of *P. papatasi* populations across the atlas mountain.

Intraspecific differences in the vectorial capacity of natural sandfly populations are poorly studied. Our data have evidenced that various isoforms of galectin are present in sandfly populations. Knowing that this gene is directly involved in the vectorial competence of *P. papatasi*, it is therefore tempting to speculate that the mutations evidenced might affect the vectorial capacity of natural *P. papatasi* populations. However, this assertion remains purely speculative, since no information is available on the binding affinity of the various galectin isoforms for *L. major* LPG or on the vectorial capacity of sandflies carrying these isoforms.

Altogether, our study has highlightened the molecular diversity of an essential gene involved in the recognition of *L. major* by its specific vector *P. papatasi*. Futher investigations are now required to further address the func-

tionality of these mutations: (i) on the affinity and specificity of the galectin for *L. major* LPG; and (ii) on the vectorial capacity of *P. papatasi* populations carrying these specific isoforms.

REFERENCES

1. Depaquit J, Ferte H, Leger N, Lefrac F, Alves-Pires C, Hanalfi H, Maroli M, Morias-Marquez F, Rioux JA, Svobodova M, Volf P. ITS2 sequences heterogeneity in *Phlebotomus sergenti* and *Phlebotomus similis* (Diptera: Psychodidae): possible consequences in their ability to transmit *Leishmania tropica*. *Int J Parasitol* 2002; *32:* 1123–31.
2. Depaquit J, Lienard E, Verzeaux-Griffon A, Ferté H, Bounamous A, Gantier JC, Hanafi AH, Jacobson RL, Maroli M, Moin-Vaziri V, Müller F, Özbel Y, Svobodova M, Volf P, Leger N. Molecular homogeneity in diverse geographical populations of *Phlebotomus papatasi* (Diptera: Psychodidae) inferred from ND4 and mtDNA and ITS2 rDNA: Epidemiological consequences. *Infect Genet Evol* 2008; *8:* 159–70.
3. Hamarsheh O, Presber W, Al-Jawabreh A, Abdeen Z, Amro A, Schönian G. Molecular markers for *Phlebotomus papatasi* (Diptera: Psychodidae) and their usefulness for population genetic analysis. *Trans R Soc Trop Med Hyg* 2009; *11:* 1085–6.
4. Hamarsheh O, Presber W, Yaghoobi-Ershadi MR, Amro A, Al-Jawabreh A, Sawalha S, Al-Lahem A, Das ML, Guernaoui S, Seridi N, Dhiman RC, Hashiguchi Y, Ghrab J, Hassan M, Schönian G. Population structure and geographical subdivision of the *Leishmania major* vector *Phlebotomus papatasi* as revealed by microsatellite variation. *Med Vet Entomol* 2009; *1:* 69–77.
5. Dantas-Torres F, Latrofa MS, Otranto D. Occurrence and genetic variability of *Phlebotomus papatasi* in an urban area of southern Italy. *Parasite Vectors* 2010; *3:* 77.
6. Kamhawi S, Ramalho-Ortigao M, Pham VM, Kumar S, Lawyer PG, Turco SJ, Barillas-Mury C, Sacks DL, Valenzuella JG. A role for insect galectin in parasite survival. *Cell* 2004; *119:* 329–41.
7. Dobson DE, Kamhawi S, Lawer P, Turco SJ, Beverley SM, Sacks DL. *Leishmania major* survival in selective *Phlebotomus papatasi* sandfly vector requires a specific SCG-encoded lipophosphoglycan galactosylation pattern. *PLoS Pathog* 2010; *6:* e1001185.
8. Meynier C, Guerlesquin F, Roche P. Computational studies of human galectin-1: role of conserved tryptophan residue in stacking interaction with carbohydrate ligands. *J Biomol Struct Dyn* 2009; *27:* 1–9.

Correspondence to: Denis Sereno, Département Santé, Représentation IRD au Maroc, 15 rue Abou Derr, 10 000 Rabat.
E-mail: denis.sereno@ird.fr

Received: 7 March 2011 *Accepted in revised form:* 21 May 2011