Univerzita Karlova

Přírodovědecká fakulta

Doktorský studijní program: Parazitologie

Ph.D. study program: Parasitology

Mgr. Michal Šíma

Identifikace, exprese a charakterizace antigenů

slinných žláz *Phlebotomus orientalis*

Identification, expression, and characterization

of *Phlebotomus orientalis* salivary antigens

Ph.D. Thesis / Dizertační práce

Supervisor / Školitel: RNDr. Iva Kolářová, Ph.D.

Co-supervisor / Konzultant: doc. RNDr. Ivan Hrdý, Ph.D.

Prague / Praha 2016

Declaration of the author / Prohlášení autora

I declare that this Ph.D. thesis was written by myself and that the literary sources were cited properly. Neither this work nor the obtained data have been used to reach any other academic degree.

Prohlašuji, že jsem předkládanou dizertační práci vypracoval samostatně a že použité literární zdroje jsou pečlivě uvedeny. Dále prohlašuji, že jak samotná práce, tak výsledky experimentů v ní zahrnutých nebyly použity k získání jiného nebo stejného akademického titulu.

Prague, November 1st, 2016
Praha 1. 11. 2016

_____

Mgr. Michal Šíma

Declaration of the supervisor / Prohlášení školitele

I declare that the data presented in this Ph.D. thesis are results of Michal Šíma work. Michal participated on the projects and experiments ongoing in our laboratory and substantially contributed to the writting of the manuscripts.

Prohlašuji, že data prezentovaná v předkládané dizertační práci jsou výsledky práce Michala Šímy. Michal se významně podílel na projektech probíhajících v naší laboratoři i na sepsání uvedených publikací.

Prague, November 1st, 2016
Praha 1. 11. 2016

_____

RNDr. Iva Kolářová, Ph.D.

Acknowledgments / Poděkování

Table of contents

LIST OF ABBREVIATIONS

ABP	Amine-binding protein
ATP/ADP/AMP	Adenosine tri/di/mono phosphate
cDNA	Complementary deoxyribonucleic acid
DTH	Delayed-type hypersensitivity
ELISA	Enzyme-linked immunosorbent assay
HBP	Histamine-binding protein
Ig	Immunoglobulin
L.	*Lutzomyia*
Le.	*Leishmania*
OBP	Odorant-binding proteins
P.	*Phlebotomus*
SGH	Salivary gland homogenate
VL	Visceral leishmaniasis
YRP	Yellow-related protein

ABSTRACT

Sand flies (Diptera: Psychodidae) are vectors of *Leishmania* spp. (Kinetoplastida: Trypanosomatidae), medically and veterinary important parasites causing leishmaniasis. These protozoans are inoculated into the host during blood-feeding together with the vector saliva. Salivary components are crucial for sand fly female blood-sucking but also play a major role in the initial stage of parasite development. It was demonstrated that this rich mixture of proteins performs different functions after its injection into the host like hydrolyzing ATP and ADP, binding biogenic amines, and others which inhibits the haemostatic reaction and modulate the immune system thus enables the successful feeding.

So far, cDNA salivary transcriptomes have been constructed for twelve sand fly species and we added another one – *Phlebotomus orientalis*, the main vector of *Leishmania donovani*, a causative agent of human visceral leishmaniasis in East African countries. We described members of thirteen protein families which corresponded with previously published salivary cDNA libraries of other sand fly species and showed the main salivary antigens using sera of experimentally bitten mice and naturally exposed dogs.

Specific IgG antibody response is induced against sand fly salivary proteins in repeatedly bitten hosts. Preceding studies demonstrated that antigens from saliva (or their recombinant counterparts) can be utilized for detection of this antibody reaction, thus might indicate the feeding preferences, estimate the risk of *Leishmania* transmission, or evaluate the effectiveness of vector control campaigns. We identified five antigenic proteins from *P. orientalis* saliva based on the reaction with antibodies from naturally exposed dogs from Ethiopia and prepared them in a recombinant form. Out of the five recombinant proteins we expressed, the recombinant yellow-related protein showed the most promising results, indicating that this protein could be used as a reliable epidemiological tool for detection of anti-*P. orientalis* antibodies across different host species.

Recently, the function of yellow-related proteins was described in *Lutzomyia longipalpis*. They serve as high affinity binders of prohaemostatic and proinflammatory biogenic ammines. Based on the known structure of one of these proteins, we were able to construct 3D models of all up to date identified proteins from this family in sand flies. Our study suggests that binding ability of yellow-related proteins can differ within as well as among sand fly species.

ABSTRAKT

Flebotomové (Diptera: Psychodidae) jsou přenašeči leishmanií (Kinetoplastida: Trypanosomatidae), medicínsky a veterinárně významných parazitů způsobujících leishmaniózu. Tito prvoci jsou inokulováni společně s obsahem slinných žláz do hostitele během sání krve. Slinné proteiny hrají důležitou roli v počáteční fázi vývoje parazita v hostiteli, ale jsou zásadní i pro samice flebotomů a jejich sání krve. Bylo prokázáno, že proteiny ze slin flebotomů jsou mimo jiné schopny hydrolýzy ATP a ADP, vazby biogenních aminů a dalších funkcí, díky nimž dochází k inhibici hemostatických reakcí a modulaci reakcí imunitního systému, což ve svém důsledku umožňuje úspěšné nasátí krve.

V minulosti bylo složení slin popsáno u dvanácti druhů flebotomů a my přidali další – *Phlebotomus orientalis*, který je hlavním přenašečem *Leishmania donovani*, původce lidské viscerální leishmaniózy ve východoafrických zemích. Z jeho slinných žláz jsme popsali třináct proteinových rodin, což odpovídá dříve publikovaným studiím z ostatních druhů flebotomů, a také jsme identifikovali hlavní antigeny pro dva druhy hostitelů - myši a psy.

U opakovaně poštípaných hostitelů dochází k tvorbě specifických IgG protilátek proti slinných proteinům flebotomů. Antigeny ze slinných žláz (nebo jejich varianty připravené v rekombinantní formě) lze využít k detekci těchto protilátek a tím je možné odhalit spektrum obratlovčích hostitelů, odhadnout riziko přenosu leishmaniózy nebo vyhodnotit účinnost preventivních opatření. Celkem jsme u *P. orientalis* identifikovali pět antigenních proteinů, které jsme posléze exprimovali v bakteriích *E. coli*. Nejslibnějších výsledků bylo dosaženo s yellow-related proteinem, který by se tak v budoucnu dal využít pro detekci protilátek proti *P. orientalis* v epidemiologických studiích.

Yellow-related proteiny ze slin flebotomů se podílí se na vazbě prohemostatických a prozánětlivých biogenních aminů. Na základě známé struktury jednoho z nich jsme byli schopni vytvořit 3D modely dalších proteinů z této skupiny i u ostatních flebotomů. Naše výsledky naznačují, že schopnost vázat aminy se mezi těmito proteiny může lišit vnitrodruhově i mezidruhově.

INTRODUCTION

Subfamily Phlebotominae (sand flies) belongs to the Psychodidae family and the order Diptera. The two genera - *Phlebotomus* and *Lutzomyia* – are vectors of the protozoan parasites of the genus *Leishmania* (Trypanosomatidae, Kinetoplastida), the causative agents of neglected tropical/subtropical disease called leishmaniasis. Infected sand fly females inject *Leishmania* parasites into the host skin during taking a bloodmeal. Infection of the mammalian host usually begins as a lesion in the skin (cutaneous leishmaniasis) but it can spread to secondary foci in the mucosa (muco-cutaneous leishmaniasis), to the bone marrow, spleen, or liver (visceral leishmaniasis), or be asymptomatic (Ready, 2010). The exact outcome of this disease is individual; it is influenced e.g. by the virulence of the parasite strain, host's immune status, and the genetic background of both the parasite and the host (reviewed in Antinori et al., 2012; Savoia, 2015). The estimated incidence is 1.3 million new cases of human leishmaniasis with mortality about 20 000-30 000 deaths annually (WHO, 2016).

During the sand fly blood-sucking, the content of salivary glands is inoculated into the host. For more than three decades, researchers have investigated the composition and biological functions of saliva, as well as its possible use in anti-*Leishmania* vaccine development (reviewed in Abdeladhim et al., 2014). Several salivary proteins showed highly antigenic properties, which resulted in eliciting strong antibody mediated response in repeatedly bitten hosts. Those specific IgG antibodies can be used as markers of exposure measured by the whole salivary gland homogenate (Rohousova et al., 2005; de Moura et al., 2007; Vlkova et al., 2011; Gidwani et al., 2011; Marzouki et al., 2011) or individual recombinant proteins (Teixeira et al., 2010; Souza et al., 2010; Marzouki et al., 2012; Vlkova et al., 2012; Drahota et al., 2014; Martin-Martin et al., 2014; Kostalova et al., 2015; Marzouki et al., 2015; Mondragon-Shem et al., 2015).

For better understanding of the saliva function and for expression of recombinant proteins, it is necessary to know nucleotide and amino acid sequences of salivary proteins. To date, cDNA libraries from salivary glands of twelve sand fly species have been constructed. Yellow-related proteins, odorant-binding proteins, antigen 5-related proteins, apyrases, hyaluronidases, and other enzymes and proteins were identified among all tested sand fly species (Charlab et al., 1999; Valenzuela et al., 2001; Valenzuela et al., 2004; Anderson et al., 2006; Oliveira et al., 2006; Kato et al., 2006; Hostomska et al., 2009;

Abdeladhim et al., 2012; Rohousova et al., 2012a; Kato et al., 2013; de Moura et al., 2013; Abdeladhim et al., 2016).

This thesis is composed of detailed characteristics of selected sand fly salivary proteins, description of their properties and possible use in anti-*Leishmania* vaccination or detection of specific IgG antibodies as exposure markers. In the next chapters, I will focus on salivary cDNA libraries of sand flies, recombinant proteins expressed based on the main antigens, and on one of the most important salivary protein functions – binding of biogenic amines and the associated inhibition of haemostatic reactions during blood-feeding.

1. *Phlebotomus orientalis*

Phlebotomus orientalis belongs to the subgenus *Larroussius* which consists of the main vectors of *Leishmania donovani* and *L. infantum* (Seccombe et al., 1993). It was found in several East African countries, such as Kenya, Sudan, or Ethiopia (Fig 1) (reviewed in Elnaiem et al., 2011).

This highly opportunistic sand fly species feeds on various mammals depending on host availability (Gebre-Michael et al., 2010; Gebresilassie et al., 2015a; Gebressilassie et al., 2015b) and plays a major role as a part of human visceral leishmaniasis cycle in these countries (reviewed in Elnaiem et al., 2011). In Sudan, it is confirmed as *L. donovani* vector in several regions (Schorscher and Goris, 1992; Elnaiem and Osman, 1998; Hassan et al., 2004; Hassan et al., 2008). In Ethiopia, a participation of *P. orientalis* in *Leishmania* transmission has never been confirmed, but because of its abundance in VL foci (Gebre-Michael et al., 2007; Gebre-Michael et al., 2010; Gebresilassie et al., 2015c) or missing confirmed vector (*P. martini*, reviewed in Elnaiem et al., 2011) in these areas (Ngumbi et al., 2010), *P. orientalis* is highly suspected as Ethiopian *Le. donovani* vector, as well.
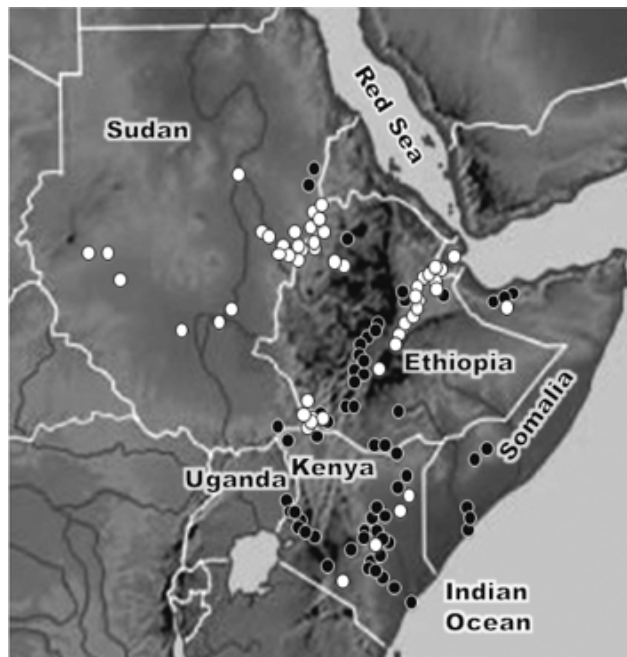
Fig 1. A distribution of *P. orientalis* in East African countries.
This map indicates the distribution of *P. orientalis* (white circles) and *P. martini* (black circles) in Sudan, Ethiopia, Somalia, Kenya and Uganda (adapted from Elnaiem et al., 2011).

The opportunistic blood-feeding behavior of *P. orientalis* females was confirmed by blood identification from blood-fed females (Gebresilassie et al., 2015a) and also

indirectly by ELISA based on the presence of specific IgG antibodies in naturally bitten hosts from Ethiopia against sand fly salivary proteins. A positive reaction was observed in domestic animals – dogs, donkeys, sheep, goats, and cows – using salivary gland homogenate as an antigen (Rohousova et al., 2015).

2. Sand fly saliva

2.1. Properties of sand fly saliva

After the host skin is lacerated by sand fly proboscis during blood-feeding, the host body is forced to minimize this harm by starting the haemostatic processes. The vasoconstriction, blood coagulation, and platelet aggregation are the necessary steps, which help the host to prevent loss of blood. In this step, proteins contained in the sand fly saliva start to perform their function by preventing the defense reactions of the host (reviewed in Ribeiro and Francischetti, 2003). This is possible due to two main functions of salivary components: 1) to stop the host haemostatic reactions and 2) to modulate host immunity.

At first, haemostatic reactions can be inhibited by platelet deactivation, inhibition of blood coagulation, and vasodilatation (reviewed in Ribeiro and Francischetti, 2003). In sand fly saliva many proteins with various functions are present, e.g. inhibitors of platelet aggregation, inhibitors of blood coagulation, and vasodilatators (reviewed in Abdeladhim et al., 2014).

Besides antihaemostatic properties of salivary components, it was also discovered, that sand fly saliva is able to modulate the host immune reactions by for example inhibiting of classical pathway of the complement (reviewed in Abdeladhim et al., 2014). Sand fly saliva is highly chemotactic for various immune cells which in case of saliva from *P. duboscqi, P. papatasi*, or *L. longipalpis* led to macrophages chemotaxis followed by accelerated entrance of *Leishmania* parasites into these cells (Anjili et al., 1995; Zer et al., 2001). In *P. papatasi* and *L. longipalpis*, it was described that saliva of these species inhibits nitric oxid production in macrophages (Hall and Titus, 1995). Inoculation of *P. papatasi* saliva resulted in decreasing of proinflammatory cytokines secretion and increasing production of anti-inflammatory cytokines (Mbow et al., 1998; Rogers and Titus, 2003). Saliva of *L. longipalpis* also enhanced the apoptosis of neutrophils, but in cells preincubated with anti-saliva antibodies, this effect was abrogated (Prates et al., 2011).

In repeatedly bitten hosts, sand fly salivary proteins induce strong antibody reaction, which can be utilized in epidemiological studies. Previously it was shown, that level of anti-saliva antibodies persists in mice, rabbits, and dogs for several weeks or months (Hostomska et al., 2008; Vlkova et al., 2012, Martin-Martin et al., 2015) but it decreased in humans after the last exposure to sand flies (Vinhas et al., 2007; Clements et al., 2010). Salivary antigens can be used for measuring the markers of exposure – specific IgG antibodies that could be detected with the whole salivary gland homogenate (Rohousova et al., 2005; de Moura et al., 2007; Vlkova et al., 2011; Gidwani et al., 2011; Marzouki et al., 2011) or with individual recombinant proteins. Up to date, the later strategy was used for three sand fly species – *L. longipalpis, P. papatasi,* and *P. perniciosus* (Teixeira et al., 2010; Souza et al., 2010; Marzouki et al., 2012; Vlkova et al., 2012; Drahota et al., 2014; Martin-Martin et al., 2014; Kostalova et al., 2015; Marzouki et al., 2015; Mondragon-Shem et al., 2015). It was indicated, that recognition of individual antigens from sand fly saliva could be host-species specific (Teixeira et al., 2010; Martin-Martin et al., 2014). This detection of specific anti-sand fly saliva antibodies is also possible because of none (or very low) cross-reactivity between antigens derived from a sand fly species and hosts bitten by another species (Souza et al., 2010; Teixeira et al., 2010; Marzouki et al., 2015).

Previous works also showed that salivary components play a key role in *Leishmania* infection development. The effect, so called "enhancing effect", was described. It was proved, that *Leishmania* infection was exacerbated in naive hosts, where parasites were inoculated into the mouse ear together with the saliva. Larger lesions were described in mice challenged with *Le. major* together with *P. papatasi* or *L. longipalpis* salivary gland homogenate in comparison to mice inoculated only by parasites (Belkaid et al., 1998; Theodos et al., 1991; Titus and Ribeiro, 1988; Samuelson et al., 1991). Partial protection against *Leishmania* parasites was observed in hosts, which were immunized with saliva or by sand fly bites (Belkaid et al. 1998; Kamhawi et al., 2000; Teixeira et al., 2014), as well as with inoculation of plasmids coding salivary proteins or recombinant proteins themselves (Collin et al., 2009; Xu et al., 2011; Gomes et al., 2012; Fiuza et al., 2016). All these experiments suggested a possible use of sand fly salivary proteins in anti-*Leishmania* vaccine development (reviewed in Abdeladhim et al., 2014).

2.2. Composition of sand fly saliva

Sand fly saliva and its components have been studied for several decades for many reasons mentioned in previous chapter. For better identification and characterization of salivary proteins, it is advantageous to know the nucleotide and amino acid sequences of these proteins. Their phylogenetic analysis showed close relationships among sand flies from the same subgenera (*Phlebotomus, Paraphlebotomus,* and *Larrousius*) (Rohousova et al., 2012a) which corresponds with previous analysis based on small subunit nuclear ribosomal DNA (Aransay et al., 2000). In both works, *Phlebotomus* and *Paraphlebotomus* subgenera clustered together and left other subgenera (e.g. *Euphlebotomus, Larroussius, and Adlerius*) in one other branch.

The initial characterization of sand fly salivary proteins started in 1999, when Charlab et al. identified several proteins in *L. longipalpis* saliva by cloning combined with biochemical approaches. The complete cDNA library of salivary glands of this *Leishmania infantum chagasi* vector was obtained five years later, when *L. longipalpis* salivary proteins were identified by cDNA sequencing, proteomics and customized computational biology approaches (Valenzuela et al., 2004). Meanwhile, transcriptomic analysis of salivary proteins of *P. papatasi*, the *Le. major* vector, was published (Valenzuela et al., 2001) and updated later (Abdeladhim et al., 2012).

Up to date, approximately 800 sand fly species is known worldwide and 166 are suspected or proven *Leishmania* vectors. However, salivary cDNA library from only twelve sand fly species have been constructed; for eight species from ***Phlebotomus*** genus and for four species from the genus ***Lutzomyia*** (Table 1).

More than 20 various proteins have been identified in each cDNA library and several protein families are shared among all of them usually with more than one homolog. Protein families, which were detected in *Phlebotomus* as well as in *Lutzomyia* species are: antigen 5-related proteins, apyrases, odorant-binding proteins (with D7-related proteins and PpSP15-like proteins), yellow-related proteins, silk-related proteins, and lufaxin-like proteins (Abdeladhim et al., 2016), but of course, there are proteins, which were detected only in few sand fly species. The complete list of sand fly protein families and their members is reviewed in Abdeladhim et al. (2014). For purpose of this thesis, only some protein families will be further discussed in detail concerning their biological functions and antigenic properties.

Table 1. Overview of sand fly species with published salivary glands cDNA libraries.
Genus, subgenus, species, and references are indicated for each published cDNA library of sand fly salivary glands.

Genus	Subgenus	Species	Reference
Phlebotomus	*Phlebotomus*	*P. papatasi*	Valenzuela et al., 2001; Abdeladhim et al., 2012
		P. duboscqi	Kato et al., 2006
	Paraphlebotomus	*P. sergenti*	Rohousova et al., 2012a
	Euphlebotomus	*P. argentipes*	Anderson et al., 2006
	Adlerius	*P. arabicus*	Hostomska et al., 2009
	Larroussius	*P. perniciosus*	Anderson et al., 2006
		P. ariasi	Oliveira et al., 2006
		P. tobbi	Rohousova et al., 2012a
Lutzomyia	*Lutzomyia*	*L. longipalpis*	Charlab et al., 1999; Valenzuela et al., 2004
	Helcocyrtomyia	*L. ayacuchensis*	Kato et al., 2013
	Nyssomyia	*L. intermedia*	de Moura et al., 2013
		L. olmeca	Abdeladhim et al., 2016

Antigen 5-related proteins are present in sand fly saliva but also in saliva of most other blood-sucking insects (Assumpcao et al., 2013), in venom of Hymenoptera (reviewed in King and Spangfort, 2000) and among secreted proteins of hookworms (Ma et al., 2015). These proteins belong to CAP family composed of Cysteine-rich secretory proteins, Antigen 5, and Pathogenesis-related 1 proteins. Proteins from CAP family are common also in prokaryotes in their cell wall, periplasm or as secreted proteins (Yeats et al., 2003) and in non-vertebrate eukaryotes like in venom of cone snail *Conus textile* (Milne et al., 2003). Homologous proteins identified in kissing bugs showed copper-dependent antioxidant properties inhibiting platelet aggregation and oxidative burst of neutrophils (Assumpcao et al., 2013) but these functions have not been confirmed yet in their sand fly counterparts. Bands that corresponded to antigen 5-related family in sand flies were recognized by specific antibodies from sera of repeatedly bitten hosts. Reaction was detected in mice bitten by *P. papatasi* (Vlkova et al., 2012) or *P. arabicus* (Hostomska et al., 2009), hamsters exposed to *P. argentipes* (Martin-Martin et al., 2013), rabbits to *P. tobbi* (Rohousova et al., 2012a), and dogs bitten by *P. perniciosus* (Vlkova et al., 2011).

Apyrases are the best described enzymes in sand fly saliva. They are the main enzymes which inhibit platelet aggregation by hydrolyzing ATP and ADP to AMP and orthophosphate. It causes the blocking of physiological signal of damaged cells and tissues. Blood-sucking insect apyrases can be divided into three families: GTPase/CD-39, 59-nucleotidase, and *Cimex* type (reviewed in Ribeiro et al., 2010). Apyrases of sand flies

belong to the last mentioned group (Valenzuela et al., 1998) and are strictly calcium dependent (Hamasaki et al., 2009). Besides of all blood-sucking insects, they were also detected in *Drosophila melanogaster* (Fenckova et al., 2011) or moth *Helicoverpa zea* (Wu et al., 2012). Sand fly apyrases also showed their antigenic properties by being recognized with IgG antibodies from dogs bitten by *P. perniciosus* (Vlkova et al., 2011; Martin-Martin et al., 2012) or hamsters exposed to *P. argentipes* (Martin-Martin et al., 2013).

Odorant-binding protein family belongs to the bigger group of proteins containing pheromone-binding proteins (PBP) and general odorant-binding proteins (GOBP). In sand fly saliva, it is represented with two groups of proteins – PpSP15-like proteins and D7-related proteins. PpSP15-like proteins have approximately 15 kDa and are sand fly specific. It was shown, that *P. papatasi* PpSP15-like protein elicited humoral and cellular immune response. It resulted in protection of immunized mice against *Le. major* infection (Valenzuela et al., 2001; Oliveira et al., 2008). A PpSP15-like protein from *P. duboscqi* was shown to inhibit activation of the coagulation cascade and to bind heparin (Alvarenga et al., 2013). These proteins are very abundant and highly variable in their amino acid sequences (Anderson et al., 2006; Hostomska et al., 2009 Rohousova et al., 2012a; Kato et al., 2013), which could result in different functions of this protein family in other sand flies. It was also suggested, that they likely occur in multiple gene copies (Elnaiem et al., 2005).

The other group of OBPs, D7-related proteins, has on the other hand homologs in various blood-sucking insects. Their functions in sand fly saliva remains unclear but similar proteins in mosquitoes are proven binders of biogenic amines or eicosanoids (Mans et al., 2007; Calvo et al., 2009) and play a role as anticoagulants (Isawa et al., 2002; Alvarenga et al., 2010). D7-related proteins of sand flies have approximately 27 kDa and are well known antigens recognized by various hosts: mice and humans bitten by *P. papatasi* (Rohousova et al., 2005; Marzouki et al., 2011), and dogs bitten by *L. longipalpis* or *P. perniciosus* (Bahia et al., 2007; Vlkova et al., 2011).

Yellow-related proteins are insect specific and are characterized by the presence of the MRJP domain (major royal jelly protein domain), which was identified also in tissues of non-haematophagus insects as in *Drosophila melanogaster* (Nash et al., 1983) or in honeybees (Hanes et al., 1992). It has been shown that *L. longipalpis* yellow-related proteins act as high affinity binders of prohaemostatic and proinflammatory biogenic amines such as serotonin, histamine, and catecholamines (Xu et al., 2011). High conservancy in the ligand-binding pocket suggests a similar function of these proteins in

other sand flies (Xu et al., 2011). Yellow-related proteins are also main antigens in sand fly saliva and possible candidates for anti-*Leishmania* vaccine (Collin et al., 2009; Xu et al., 2011; Gomes et al., 2012). Strong antibody responses to sand fly salivary yellow-related proteins was previously demonstrated in many studies, mainly focused on canines and humans exposed to various sand fly species, e.g. *L. longipalpis, P. papatasi*, or *P. perniciosus* (Gomes et al., 2002; Vinhas et al., 2007; Gomes et al, 2007; Bahia et al., 2007; Hostomska et al., 2008; Teixeira et al., 2010; Gidwani et al., 2011; Marzouki et al., 2011; Vlkova et al., 2011; Vlkova et al., 2012).

In contrast to previously mentioned protein families, ParSP25-like proteins seem to be Old World sand fly species specific. They have not been detected yet in other blood-sucking Diptera or in New World sand fly species (Charlab et al., 1999; Valenzuela et al., 2004; Abdeladhim et al., 2012; de Moura et al., 2013; Kato et al., 2013; Abdeladhim et al., 2016) but they have been found in cDNA libraries of several *Phlebotomus* species, e.g. *P. ariasi, P. perniciosus, P. tobbi, P. arabicus,* and *P. papatasi* (Anderson et al., 2006; Oliveira et al., 2006; Hostomska et al., 2009; Rohousova et al., 2012a; Abdeladhim et al., 2012). Exact functions of these proteins still remains unclear, but it was demonstrated, that antibodies from mice, dogs, and hamsters repeatedly bitten by *P. perniciosus* reacted with proteins identified as members of ParSP25-like family (Vlkova et al., 2011; Martin-Martin et al., 2012).

In summary, in sand fly salivary glands there is a rich mixture of proteins with various functions. In some cases, this function still remains undiscovered. It is possible, that construction of cDNA libraries from more sand fly species or preparing recombinant proteins could give us more answers and could reveal, if there are chances of using these proteins for detecting antibody response in repeatedly bitten hosts or incorporating them in anti-*Leishmania* vaccine.

3. Detection of anti-saliva antibodies

As mentioned above, sand fly saliva in repeatedly bitten hosts induce cellular and antibody mediated immune responses. It was discovered, that the increased level of specific IgG antibodies persist in bitten hosts (mice, rabbits, and dogs) for a relatively long time – weeks or months (Hostomska et al., 2008; Vlkova et al., 2012; Martin-Martin et al., 2015),

but in humans it decreased rapidly after the last exposure to sand flies (Vinhas et al., 2007; Clements et al., 2010).

Screening for anti-sand fly saliva antibodies has its limitation due to high amount of work, which is necessary for obtaining the sufficient amount of salivary gland homogenate (SGH), traditionally used in screening tests as antigen. At first, the colony of sand flies has to be maintained under special conditions (Volf and Volfova, 2011) and the dissection of salivary glands has to be managed precisely by well-trained person. Finally, there is a risk of potential cross-reactivity with non-vector species if whole salivary gland homogenate is used as an antigen (Clements et al., 2010). Most of these disadvantages of SGH can be circumvented by using recombinant proteins.

It has been shown previously, that reaction of antibodies with sand fly salivary recombinant proteins is highly host- and vector-species specific. The cross-reactivity between salivary antigens from different sand fly species is none or very low (Souza et al., 2010; Teixeira et al., 2010; Marzouki et al., 2015). Although the antibody response to sand fly saliva is host-specific, it was also demonstrated, that some recombinant proteins can be used to detect anti-saliva antibodies across various host species (Teixeira et al., 2010; Martin-Martin et al., 2014). Antibody response can be utilized for indication the feeding preferences of sand flies (Gomes et al., 2007; Rohousova et al., 2015), for evaluation of the effectiveness of vector control campaigns (Gidwani et al., 2011), or for estimation the risk of *Leishmania* transmission (Rohousova et al., 2005; de Moura et al., 2007; Vlkova et al., 2011; Marzouki et al., 2011). Altogether, it suggests that detection of specific antibodies against sand fly saliva by using recombinant proteins as antigens can be used as a reliable epidemiological tool.

3.1. Recombinant proteins

While salivary proteins and antigens has been characterized from twelve sand fly species (Table 1), recombinant proteins from only three species – *L. longipalpis, P. perniciosus,* and *P. papatasi* – have been tested as antigens for detecting antibodies in epidemiological studies (Teixeira et al., 2010; Souza et al., 2010; Vlkova et al., 2012; Marzouki et al., 2012; Drahota et al., 2014; Martin-Martin et al., 2014; Kostalova et al., 2015; Marzouki et al, 2015; Mondragon-Shem et al., 2015). Description of them is listed bellow and the best performing recombinant proteins are summarized in Table 2.

Table 2. Candidate recombinant proteins as antigens for detecting anti-sand fly saliva antibodies.

Recombinant proteins used as antigens for detection of specific anti-sand fly saliva IgG antibodies are listed together with the following details: sand fly species, host species with a positive antibody reaction, and a reference.

Recombinant protein identifier	Sand fly species	Host species	Reference
LJM17	*L. longipalpis*	dog, fox, human	Teixeira et al., 2010; Souza et al., 2010
LJM11	*L. longipalpis*	human, dog	Teixeira et al., 2010; Souza et al., 2010
LJM111	*L. longipalpis*	human, dog	Teixeira et al., 2010
LJL23	*L. longipalpis*	dog	Teixeira et al., 2010
LJL13	*L. longipalpis*	dog	Teixeira et al., 2010
LJM04	*L. longipalpis*	dog	Teixeira et al., 2010
rSP03B	*P. perniciosus*	mouse, dog, hare, rabbit	Drahota et al., 2014; Martin-Martin et al., 2014; Kostalova et al., 2014
rSP01	*P. perniciosus*	mouse, dog	Drahota et al., 2014
rSP01B	*P. perniciosus*	mouse, dog, hare, rabbit	Drahota et al., 2014; Martin-Martin et al., 2014
rPpSP28	*P. papatasi*	mouse	Vlkova et al., 2012
rPpSP30	*P. papatasi*	mouse	Vlkova et al., 2012
rPpSP42	*P. papatasi*	mouse	Vlkova et al., 2012
rPpSP32	*P. papatasi*	human	Marzouki et al., 2012; Marzouki et al., 2015; Mondragon-Shem et al., 2015

First studies were published for *L. longipalpis*, the vector of *Le. infantum chagasi* – a causative agent of visceral leishmaniasis in Latin America. Teixeira et al. (2010) expressed nine different recombinant proteins in HEK-293F mammalian cells. They tested three members of yellow-related proteins (LJM17, LJM11, and LJM111), an apyrase (LJL23), a D7-related protein (LJL13), a SL1 protein (LJM04), an endonuclease (LJL138), a 5'nucleotidase (LJL11), and a lufaxin (LJL143) with dog, fox, and human sera from endemic and non-endemic areas in Brazil. The most universal recombinant protein from this study was the yellow-related protein LJM17, which was recognized by sera from all three different hosts, while the other two yellow-related proteins – LJM11 and LJM111 - were recognized only by two (humans and dogs) and one (humans) host species, respectively. Reaction only with dog sera was observed for recombinant proteins LJL23, LJL13, LJM04, and LJL143. LJL138 (endonuclease) and LJL11 (5'nucleotidase) were not recognized by any of the host sera. Recombinant yellow-related proteins LJM17 and LJM11 were validated as an antigen with larger set of human samples from Brazil, thus it

seems that these two proteins may substitute whole salivary gland homogenate for testing human samples from this area (Souza et al., 2010).

Later, there were three studies focused on recombinant proteins based on antigens from *P. perniciosus* saliva (Drahota et al., 2014; Martin-Martin et al., 2014; Kostalova et al., 2015). This sand fly species transmits *Le. infantum*, which causes human and canine visceral leishmaniasis in Mediterranean region. Drahota et al. (2014) used six different bacterially expressed recombinant proteins - yellow-related protein (rSP03B), two apyrases (rSP01B and rSP01), antigen 5-related (rSP07), ParSP25-like protein (rSP08), and D7-related protein (rSP04) for detecting IgG antibodies in experimentally bitten mice and dogs. Only yellow-related protein and both apyrases showed high correlation with the whole *P. perniciosus* SGH for both host species. Those three proteins (rSP03B, rSP01, and rSP01B) were further tested with sera from naturally bitten hosts - hares, rabbits, and dogs from Spain (Martin-Martin et al., 2014). Positive correlation between all three recombinant proteins and *P. perniciosus* SGH was detected for dogs, while in hare and rabbit sera only yellow-related protein rSP03B and apyrase rSP01B showed correlation higher than 0.65 (Martin-Martin et al., 2014). Large set of sera from naturally bitten dogs from Italy, the area endemic for *P. perniciosus*, was tested by Kostalova et al. (2014). The experiment with the two most promising recombinant proteins from previous studies (rSP03B and rSP01) showed, that only yellow-related protein rSP03B could be used in these experimental settings to replace whole *P. perniciosus* SGH.

Last sand fly species, from which salivary recombinant proteins were expressed, is *P. papatasi*, the vector of *Leishmania major* (etiological agent of cutaneous leishmaniasis). In the first study, Vlkova et al. (2012) expressed two yellow-related proteins (rPpSP42 and rPpSP44) and two D7-related proteins (rPpSP28 and rPpSP30) in *E. coli* and tested them with sera of experimentally bitten mice. Murine sera reacted with three recombinant proteins (rPpSP42, rPpSP30, and rPpSP28), but authors declared that these reactions could be mouse strain-specific (Vlkova et al., 2012). Anti-*P. papatasi* IgG antibodies in humans were detected with two recombinant proteins – rPpSP30 (D7-related protein) and rPpSP32 (silk-related protein) expressed in mammalian cells HEK-293F. Marzouki et al. (2012) tested these recombinants with human sera from Tunisia and only rPpSP32 was recognized with these samples. Later, the usage of this recombinant protein was confirmed and validated for antibody detection using larger set of human samples from Tunisia (Marzouki et al., 2015) and Saudi Arabia (Mondragon-Shem et al., 2015).

In conclusion, several recombinant proteins were confirmed, that they could replace a whole salivary gland homogenate as antigens in epidemiological studies. Mainly yellow-related proteins proved to be a very effective tool for detecting anti-sand fly saliva IgG antibodies in various hosts. Although all these authors published interesting results for different recombinant proteins with various host species, only two of them (Teixeira et al., 2010; Marzouki et al., 2015) tested a possible cross-reactivity between recombinant proteins derived from a sand fly species and hosts bitten by another species. They showed, that there was no cross-reactivity between recombinants from *L. longipalpis* and hosts bitten by *L. intermedia* (Teixeira et al., 2010) or a recombinant protein derived from *P. papatasi* and hosts bitten by *P. perniciosus* (Marzouki et al., 2015). Both studies suggest the suitability of these proteins for detecting specific IgG antibodies as markers of exposure to sand flies. Further analyses with other recombinant proteins derived from more sand fly species are necessary to validate, that recombinant proteins might be more specific than the whole salivary gland homogenate.

3.2. Vaccine candidates

Previously, it was showed, that immunization of hosts by sand fly saliva results in protection against *Leishmania* parasites. First experiments were performed with *P. papatasi* SGH inoculated together with *Le. major* promastigotes into ear dermis of mice pre-immunized with *P. papatasi* saliva. The dermal lesions appeared later, were less destructive, and contained lower number of parasites (Belkaid et al., 1998). Similar protective effect was observed when mice were pre-exposed to uninfected sand flies prior to *Le. major* infection (Kamhawi et al., 2000). This protection can be due to inhibition of exacerbative effect of sand fly saliva (Belkaid et al., 1998) or due to induction of delayed-type hypersensitivity response (DTH) in form of Th1 cellular response (Kamhawi et al., 2000). Most of the follow-up studies support this protective feature of sand fly saliva or individual salivary proteins (Valenzuela et al., 2001; Morris et al., 2001; Oliveira et al., 2008; Gomes et al., 2008; Collin et al., 2009; Xu et al., 2011; Tavares et al., 2011; Gomes et al., 2012; Fiuza et al., 2016), but in the New World species, *L. intermedia*, pre-exposition to sand fly saliva and subsequent challenge with *L. braziliensis* led to exacerbation of the infection (de Moura et al., 2010; Weinkopff et al., 2014).

In 2001, it was suggested, that a single protein - PpSP15 from *P. papatasi* saliva - can be a potential candidate for anti-*Leishmania* vaccine, since immunization of mice with

its plasmid induced DTH which led to protection against *Le. major* (Valenzuela et al., 2001). This protection was cell-mediated and antibody-independent. Later, PpSP15 was validated as a candidate protein for anti-*Le. major* protection and it was discovered that, besides induction of DTH, immunized mice expressed IFN-γ and IL-12 (proinflammatory Th1 response) after exposition to sand flies which further boosted an anti-*Leishmania* immunity (Oliveira et al., 2008).

Immunization of mice with *L. longipalpis* yellow-related protein LJM11 also led to induction of Th1 response which resulted in long-lasting immunity and ulcer-free protection of animals challenged by sand fly inoculation of *Le. major* (Gomes et al., 2012). Similar effect was observed by immunization of mice with plasmid for the same protein (Xu et al., 2011) as well as in dogs immunized with plasmids of another yellow-related protein, LJM17, from *L. longipalpis* and challenged by *Le. infantum chagasi* (Collin et al., 2009).

Several other recombinant proteins or their plasmids from *L. longipalpis* were successfully tested to protect the host against *Leishmania* parasites. Maxadilan induced Th1 response and protected mice against *Le. major* (Morris et al., 2001). Similarly, LJM19 protected hamsters against *Le. infantum chagasi* and *Le. braziliensis* (Gomes et al., 2008; Tavares et al., 2011) and lufaxin (LJL143) protected dogs against *Le. infantum chagasi* (Collin et al. 2009).

Overall, these experiments suggest that protection against *Leishmania* parasites is hidden behind delayed type hypersensitivity reaction and Th1 immune response with production of IFN-γ. Unfortunately, vaccine against any *Leishmania* parasite is still not available, neither with sand fly salivary protein(s) nor without.

4. Salivary amine-binding proteins

4.1. Amine-binding proteins of blood-feeding arthropods

Many proteins with various functions were detected in blood-feeding arthropods saliva. One of the key mechanisms for successful blood-feeding is removing biogenic amines from feeding site e.g. by binding them into the proteins, commonly named as kratagonists. Binding of these small molecules (such as serotonin, histamine, and others) leads to prevention of inflammation and haemostasis thus allowing the blood-feeding process (reviewed in Ribeiro and Arca, 2009; Ribeiro et al., 2010). Histamine is derived from

amino acid histidine and is present in granules of mast cells and basophils, from where it can be released. It increases the permeability of blood capillaries and contractility of smooth muscles in small blood vessels and bronchioles. Serotonin is formed from tryptophan and can be detected in human platelets, digestive tract, or the central nervous system. Its role is to induce smooth muscle contractions, enhance the permeability of small vessels, or induce vasoconstriction in large vessels (reviewed in Klein and Horejsi, 1997). Blocking those amines result in vasodilatation, platelet deactivation and decreased vascular permeability (Oliveira et al., 2007; Xanthos et al., 2008).

Salivary proteins with proven amine-binding function were described e.g. in ticks (Paesen et al., 1999; Mans et al., 2008), Hemiptera (Andersen et al., 2003, Xu et al., 2013), mosquitoes (Mans et al., 2007; Calvo et al., 2006; Calvo et al., 2009), and sand flies (Xu et al., 2011). They can be divided in three groups – 1) lipocalins, 2) D7 proteins, and 3) yellow-related proteins. The common feature of all three groups is their shape – hollow barrel with two possible entrances and the ligand-binding pocket inside of this structure (Paesen et al., 1999; Mans et al., 2007; Mans et al., 2008; Calvo et al., 2009; Xu et al., 2011; Xu et al., 2013) (Fig 2).

Lipocalins were described e.g. in saliva of soft and hard ticks and Hemiptera (Triatominae). They are usually extracellular proteins that transport small hydrophobic ligands with various specificities. Their structure consists of eight-stranded antiparallel β-barrel (reviewed in Flower, 1996). Three different lipocalins were obtained from the hard tick *Rhipicephalus appendiculatus* and all shared the same shape and bound two molecules of histamine, each on opposite ends of the barrel. One of the binding sites has higher affinity than the other. Two of these lipocalins are 20 kDa proteins (Ra-HBP1 and Ra-HBP2), which are female specific, monomeric, and non-glycosylated. The third lipocalin, dimeric 45 kDa protein (Ra-HBP3), was characterized as male-specific in feeding period and occurs also in feeding larvae and nymphs. The lowest histamine-binding affinity was detected in Ra-HBP1. The variations among HBPs suggested different feeding behavior between males and females (Paesen et al., 1999).

Two lipocalins, monomine and monotonin were discovered in the soft tick *Argas monolakensis*. They have the same fold as lipocalins from the hard tick and they are stabilized with three disulfide bonds. In comparison to Ra-HBPs, monomine and monotonin bind only single molecule of histamine and 5-hydroxytryptamine, respectivelly. The binding site is comparable to the binding site of Ra-HBP2 with lower affinity to bioamine, the other one is missing in these two lipocalins (Mans et al., 2008).
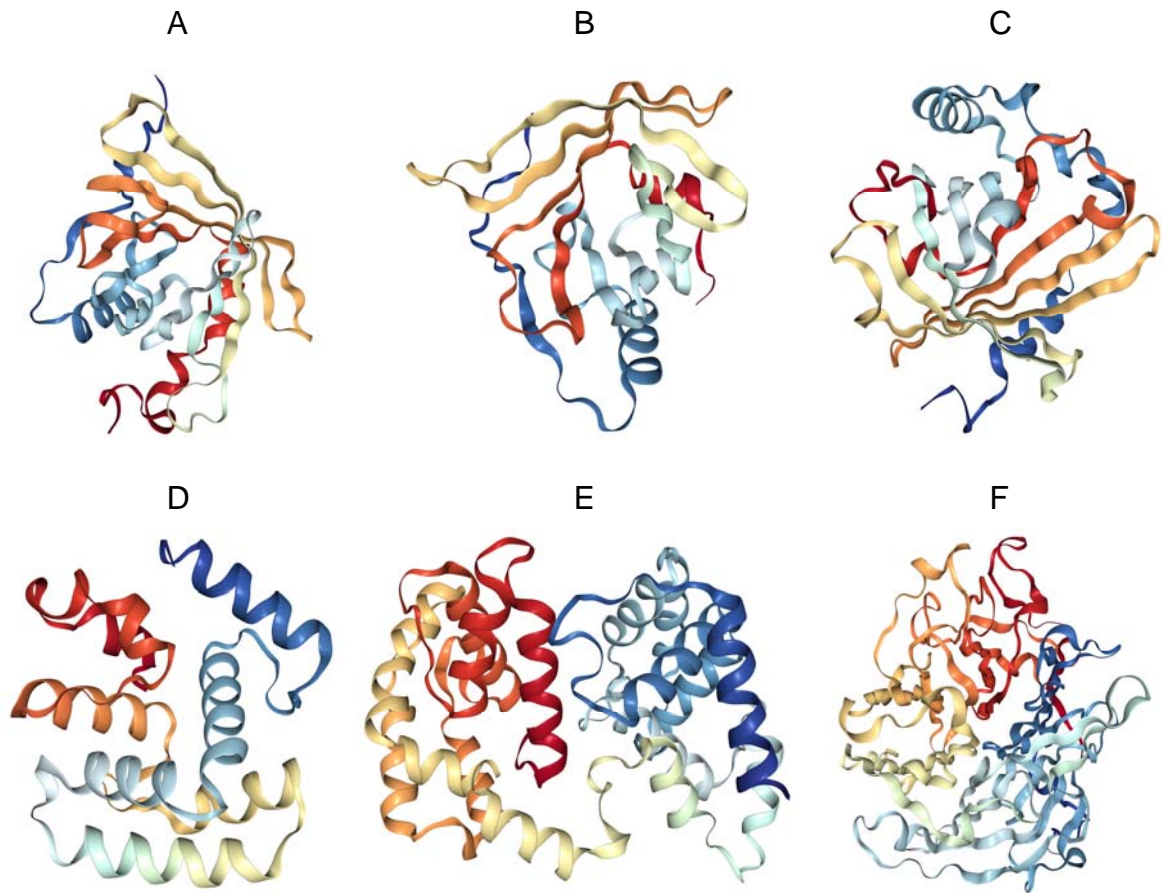
Fig 2. Structures of amine-binding proteins in blood-sucking arthropods. Published structures of amine-binding proteins were downloaded from Protein Data Bank (Berman et al., 2000) using the NGL viewer (Rose and Hildebrand, 2015) and composed into this figure. Each structure represents a binding protein from different blood-sucking species: A – a lipocalin Ra-HBP2 from *Rhipicephalus appendiculatus* (PDB ID: 1QFT), B – a lipocalin monomine from *Argas monolakensis* (PDB ID: 3BU1), C – a lipocalin amine-binding protein from *Rhodnius prolixus* (PDB ID: 4GE1), D – the D7 protein D7r4 from *Anopheles gambiae* (PDB ID: 2QEB), E – the D7 protein AeD7 from *Aedes aegypti* (PDB ID: 3DYE), and F – a yellow-related protein LJM11 from *Lutzomyia longipalpis* (PDB ID: 3Q6K). The color saturation is decreasing from N- (red) to C-terminus (blue).

The other blood-feeding arthropod with salivary lipocalin is *Rhodnius prolixus*, a representative of Hemiptera (Andersen et al., 2003). Its amine-binding protein (ABP) is able to bind serotonin, epinephrine, and norepinephrine. The binding site is probably shared for all possible ligands. Based on the dissociation constants for each of the ligands, ABP showed the highest affinity to norepinephrine, thus inhibiting vasoconstriction in the feeding site. (Andersen et al., 2003). *R. prolixus* lipocalin is closely related to nitrophorin group of lipocalins found in the same insect species (Weichsel et al., 1998). Nitrophorins are heme-binding nitric oxide transporters and after releasing into the blood and skin, they cause vasodilatation and inhibition of platelet aggregation by production of nitric oxide

(Ribeiro et al., 1993; Andersen and Montfort, 2000). In 2013, the 3D structure of this ABP was determined as eight-stranded antiparallel β-barrel with a ligand-binding pocket (Xu et al., 2013), which corresponded with the nitrophorin 2 having a similar fold and containing a ferric heme (Andersen and Montfort, 2000).

D7 and D7-related proteins have been described in saliva of various insect species, for example black flies (e.g. Andersen et al., 2009), mosquitoes (e.g. Jariiyapan, et al., 2012), biting midges (e.g. Campbell et al., 2005), or all tested sand flies (Table 1). In blood-sucking Diptera, amine-binding properties of D7 proteins were confirmed only for two mosquito species – *Anopheles gambiae* and *Aedes aegypti* (Calvo et al., 2006). Salivary D7 proteins are related to odorant-binding proteins (OBPs) (Arca et al., 2002). In mosquitoes, two forms of D7 proteins (short, around 15 kDa; and long, around 25 kDa) are present (Arca et al., 2005) and both of them can bind biogenic amines (Calvo et al., 2006). Calvo et al. (2006) expressed five short D7 proteins of *A. gambiae* and one long D7 protein of *A. aegypti*. All of them (except one short-form protein) showed high affinity to serotonin (with K_d around 1 nM) and most of them bound other biogenic amines (histamine, epinephrine, and norepinephrine) as well but with lower affinity (Calvo et al., 2006). Calvo et al. (2006) suggested that differences in amine-binding ability among D7 proteins have implications in evolution of this protein family. Later, the crystal structures of one short-form D7 protein of *A. gambiae* (D7r4) and one long-form D7 protein of *A. aegypti* (AeD7) were determined and their binding abilities were described in more details (Mans et al., 2007; Calvo et al., 2009).

For D7r4, a fold arranged of eight α-helices is stabilized with three disulfide bonds. The binding pocket is created by hydrophobic side chains with polar and charged groups which can be connected with aliphatic amino groups of the ligands (Mans et al., 2007). High affinities for serotonin and tryptamine to this protein were confirmed and weaker binding of norepinephrine was explained by different structure of this ligand allowing to create another type of binding (Mans et al., 2007). None of the three main acidic residues involved in ligand-binding (Asp 111, Glu 114, and Asp 139) were identified in other insects having short-forms of D7 proteins (culicine mosquitoes, sand flies, and *Culicoudes* sp.) which suggests that binding function of these proteins occurs only in Anophelinae (Mans et al., 2007). D7r4 is structurally similar to other OBPs, however it contains two other C-terminal helices and that can play a major role in binding ability of these proteins (Mans et al., 2007).

AeD7 from *A. aegypti* was described as two domains protein, where both domains are made up of seven α-helices. The C-terminal domain, which binds biogenic amines, is stabilized by three hydrogen bonds while the N-terminal domain, binding cysteinyl leukotrienes, by only two hydrogen bonds (Calvo et al., 2009). The amine-binding domain is sequentially similar to the domain of short-form D7 proteins from *A. gambiae* and only one binding site was detected as well. Amine-binding does not affect the binding of cysteinyl leukotrienes to the other protein domain. A comparison to *Anopheles* two domain D7 proteins showed sequence conservation in the N-terminal domain (suggesting possible binding of cysteinyl leukotrienes) but it revealed differences in essential binding residues in C-terminal binding domain which can lead to lose the binding ability in other mosquito species (Calvo et al., 2009).

Yellow-related proteins (or their homologs) were identified in non-haematophagus insects like *Drosophila melanogaster* (Nash et al., 1983) or honeybees (Hanes et al., 1992) as well as in other blood-feeders – for example tse-tse flies (Alves-Silva et al., 2010) and mosquitoes (Johnson et el., 2001). They were also found in all published salivary cDNA libraries of sand flies (Table 1). These proteins are named after a gene mutation in *Drosophila*, where mutation in the gene for yellow protein caused color change in the fruit fly phenotype (Albert et al., 1999). Yellow-related proteins are characterized with the entire major royal jelly protein (MRJP) domain which was identified in royal jelly proteins of honeybees (Hanes et al., 1992). Properties and amine-binding function of proteins from this family in sand fly will be described in the following chapter.

4.5. Amine-binding proteins in sand flies

In sand flies, two protein families can serve as putative kratagonists – D7-related proteins and yellow-related proteins (YRPs). So far, lipocalins have not been described in any of the 12 salivary transcriptomes of sand fly species. On the other hand, D7-related proteins and yellow-related proteins are present in all sand fly species tested so far (Table 1).

Despite detection of OBPs in sand flies, so far, amine-binding ability was described only for YRPs (Xu et al., 2011). They are highly conserved and have similar molecular mass; between 40-42 kDa. There is a high variability in number of YRPs among different sand fly species, for example only one member of YRPs was found in *P. arabicus* (Hostomska et al., 2009), *P. argentipes* (Anderson et al., 2006), and *L. intermedia* (de Moura et al., 2013) but on contrary five YRPs were detected in *P. sergenti* (Rohousova et

al., 2012a). This variability might by attributed to the sensitivity of sequencing method, however, it is not likely for those species/cDNA libraries that were constructed in the same laboratory under the same conditions (Anderson et al., 2006; Rohousova et al., 2012a). Thus, the occurrence of various numbers of YRPs in sand fly could be caused by differential gene expression, because genes with lower expression might be recorded less frequently as it was proven in other sand fly protein families (Rohousova et al., 2012b). YRPs can occur in N-glycosylated, O-glycosylated, C-glycosylated, or non-glycosylated forms (Rohousova et al., 2012a). Their lectin-like activity and swallowing of YRPs into the midgut together with saliva was described for *P. duboscqi* (Volf et al., 2000).

In 2011, Xu et al. expressed three recombinant YRPs of *L. longipalpis* in *Escherichia coli* system. They determined their binding abilities to six different biogenic amines and characterized a crystal structure of one of these YRPs (LJM11) (Xu et al., 2011). They proved that there is one ligand-binding site in the protein structure and that all three YRPs bind five different biogenic amines (norepinephrine, epinephrine, serotonin, dopamine, octopamine, or histamine) with various affinities. The highest affinity was observed for serotonin with all three proteins and no detectable binding was discovered for histamine with LJM11 and LJM111 and for epinephrine with LJM17 (Xu et al., 2011). Crystallographic analysis showed that LJM11 has a six-bladed β-propeller fold with a central, solvent-accessible tunnel with two entrances on opposite sides of the molecule. This structure is stabilized by two disulfide bonds (Xu et al., 2011). Homologous modeling of LJM111, the other *L. longipalpis* YRP, suggested different electrostatic surface charges (and isoelectric points − 9.3 for LJM11 and 4.7 for LJM11), which can cause changes of immunological properties between these proteins (Xu et al., 2011). Sequence analysis of ligand-binding pocket revealed highly conserved amino acid motif among other sand flies. Five out of eleven binding amino acids in this pocket are identical for all sand fly YRPs. The most variable position is annotated for LJM11 as Phe 344 and among all YRPs is occupied by six different amino acids (His, Gln, Tyr, Phe, Met, or Lys) (Xu et al., 2011; de Moura et al., 2013; Abdeladhim et al., 2016). Three of them (in LJM11 Thr 327, Asn 342, and Phe 344) are crucial for ligand-binding because of possible hydrogen bonds between the protein and ligands. The other eight interact with biogenic amines by Van der Waals or hydrophobic bonds (Xu et al., 2011).

As it was shown, several amine-binding proteins were found across various groups of blood-feeding arthropods and the affinity to small ligands was measured in these species.

Three different protein families showed high affinity to biogenic amines – lipocalins, D7 proteins, and yellow-related proteins. For sand flies, only yellow-related proteins are confirmed amine-binders and those experiments were performed with only one sand fly species – *L. longipalpis*. Based on high sequence conservancy, it was suggested, that yellow-related proteins of other sand fly species share this binding ability but it has not been experimentally confirmed, yet.

OBJECTIVES

cDNA libraries of sand fly salivary glands have been published for main *Leishmania* vectors across the world – four New World *Lutzomyia* species and eight Old World *Phlebotomus* species. But still, there are a lot of medically and veterinary important sand flies, whose salivary proteins have not been characterized yet and it is also a case of the main *Le. donovani* vector in East African countries – *P. orientalis*.

Some sand fly salivary proteins have been tested in anti-*Leishmania* vaccination and the most promising results were obtained with candidates from a yellow-related protein family. Proteins from this group are binders of biogenic amines, thus inhibiting host haemostasis during sand fly blood-feeding. However, detailed characterization has been performed only with YRPs from *Lutzomyia longipalpis*.

Yellow-related proteins are, as well, highly valuable in recombinant form for detecting anti-sand fly saliva antibodies in repeatedly bitten hosts. Their use in larger epidemiological studies could help to evaluate the effectiveness of vector control campaigns, to indicate sand fly feeding preferences or to estimate the risk of *Leishmania* transmission. But this approach was so far applied only for three sand fly species and mainly with human and dog hosts. However, other domestic animals could also play a significant role in *Leishmania* transmission cycle, either as a reservoir host or as a source of blood.

Altogether, better characterization of this sand fly salivary protein family would be of great importance.

The main objectives of this thesis are:

1) to analyze salivary gland transcriptome and to describe the main salivary proteins and antigens of *P. orientalis*,
2) to express and evaluate *P. orientalis* salivary antigens for detection of specific IgG antibodies as markers of exposure,
3) to compare sand fly salivary yellow-related proteins and determine their phylogenetic relationships, putative glycosylation pattern, and theoretical biogenic amines binding ability.

PUBLICATIONS

1) Vlkova M, Sima M, Rohousova I, Kostalova T, Sumova P, Volfova V, Jaske EL, Barbian KD, Gebre-Michael T, Hailu A, Warburg A, Ribeiro JMC, Valenzuela JG, Jochim RC, Volf P. Comparative analysis of salivary gland transcriptomes of *Phlebotomus orientalis* sand flies from endemic and non-endemic foci of visceral leishmaniasis. PLoS Negl Trop Dis. 2014; 8: e2709.

2) Sima M, Ferencova B, Warburg A, Rohousova I, Volf P. Recombinant salivary proteins of *Phlebotomus orientalis* are suitable antigens to measure exposure of domestic animals to sand fly bites. PLoS Negl Trop Dis. 2016a; 10: e0004553.

3) Sima M, Novotny M, Pravda L, Sumova P, Rohousova I, Volf P. The diversity of yellow-related proteins in sand flies (Diptera: Psychodidae). PloS One. 2016b; Accepted.

Comparative analysis of salivary gland transcriptomes of *Phlebotomus orientalis* sand flies from endemic and non-endemic foci of visceral leishmaniasis

Vlkova M, **Sima M**, Rohousova I, Kostalova T, Sumova P, Volfova V, Jaske EL, Barbian KD, Gebre-Michael T, Hailu A, Warburg A, Ribeiro JMC, Valenzuela JG, Jochim RC, Volf P.

(2014)

Comparative Analysis of Salivary Gland Transcriptomes of *Phlebotomus orientalis* Sand Flies from Endemic and Non-endemic Foci of Visceral Leishmaniasis

Michaela Vlkova[1], Michal Sima[1], Iva Rohousova[1], Tatiana Kostalova[1], Petra Sumova[1], Vera Volfova[1], Erin L. Jaske[2], Kent D. Barbian[2], Teshome Gebre-Michael[3], Asrat Hailu[4], Alon Warburg[5], Jose M. C. Ribeiro[6], Jesus G. Valenzuela[7]*, Ryan C. Jochim[7¤]*, Petr Volf[1]*

1 Department of Parasitology, Faculty of Science, Charles University, Prague, Czech Republic, 2 Genomics Unit, Research Technologies Section, Rocky Mountain Laboratories, Hamilton, Montana, United States of America, 3 Aklilu Lemma Institute of Pathobiology, Addis Ababa University, Addis Ababa, Ethiopia, 4 Department of Microbiology, Immunology & Parasitology, Faculty of Medicine, Addis Ababa University, Addis Ababa, Ethiopia, 5 Department of Parasitology, The Kuvin Centre for the Study of Infectious and Tropical Diseases, Hadassah Medical School, The Hebrew University of Jerusalem, Jerusalem, Israel, 6 Vector Biology Section, Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, Maryland, United States of America, 7 Vector Molecular Biology Section, Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, Maryland, United States of America

Abstract

Background: In East Africa, *Phlebotomus orientalis* serves as the main vector of *Leishmania donovani*, the causative agent of visceral leishmaniasis (VL). *Phlebotomus orientalis* is present at two distant localities in Ethiopia; Addis Zemen where VL is endemic and Melka Werer where transmission of VL does not occur. To find out whether the difference in epidemiology of VL is due to distant compositions of *P. orientalis* saliva we established colonies from Addis Zemen and Melka Werer, analyzed and compared the transcriptomes, proteomes and enzymatic activity of the salivary glands.

Methodology/Principal Findings: Two cDNA libraries were constructed from the female salivary glands of *P. orientalis* from Addis Zemen and Melka Werer. Clones of each *P. orientalis* library were randomly selected, sequenced and analyzed. In *P. orientalis* transcriptomes, we identified members of 13 main protein families. Phylogenetic analysis and multiple sequence alignments were performed to evaluate differences between the *P. orientalis* colonies and to show the relationship with other sand fly species from the subgenus *Larroussius*. To further compare both colonies, we investigated the humoral antigenicity and cross-reactivity of the salivary proteins and the activity of salivary apyrase and hyaluronidase.

Conclusions: This is the first report of the salivary components of *P. orientalis*, an important vector sand fly. Our study expanded the knowledge of salivary gland compounds of sand fly species in the subgenus *Larroussius*. Based on the phylogenetic analysis, we showed that *P. orientalis* is closely related to *Phlebotomus tobbi* and *Phlebotomus perniciosus*, whereas *Phlebotomus ariasi* is evolutionarily more distinct species. We also demonstrated that there is no significant difference between the transcriptomes, proteomes or enzymatic properties of the salivary components of Addis Zemen (endemic area) and Melka Werer (non-endemic area) *P. orientalis* colonies. Thus, the different epidemiology of VL in these Ethiopian foci cannot be attributed to the salivary gland composition.

Introduction

Protozoan parasites belonging to the genus *Leishmania* are the pathogenic agents causing a broad range of diseases commonly known as leishmaniasis. Sand fly vectors (Diptera: Phlebotominae) spread leishmaniasis among the vertebrate hosts during the bloodfeeding when infected sand fly females eject parasites into the wound along with their saliva. Salivary compounds possess powerful anti-hemostatic and immunomodulatory properties (reviewed in [1]); nonetheless, the salivary proteins are highly antigenic. As the repeated exposure to sand fly bites was shown to be protective against leishmaniasis (e.g. [2]), the immune profiles elicited by single salivary proteins are of major scientific interest.

Author Summary

Phlebotomus orientalis is the vector of visceral leishmaniasis (VL) caused by *Leishmania donovani* in Northeast Africa. Immunization with sand fly saliva or with individual salivary proteins has been shown to protect against leishmaniasis in different hosts, warranting the intensive study of salivary proteins of sand fly vectors. In our study, we characterize the salivary compounds of *P. orientalis*, thereby broadening the repertoire of salivary proteins of sand fly species belonging to the subgenus *Larroussius*. In order to find out whether there is any connection between the composition of *P. orientalis* saliva and the epidemiology of VL in two distinct Ethiopian foci, Addis Zemen and Melka Werer, we studied the transcriptomes, proteomes, enzymatic activities, and the main salivary antigens in two *P. orientalis* colonies originating from these areas. We did not detect any significant difference between the saliva of female sand flies originating in Addis Zemen (endemic area) and Melka Werer (non-endemic area). Therefore, the different epidemiology of VL in these Ethiopian foci cannot be related to the distant salivary gland protein composition. Identifying the sand fly salivary gland compounds will be useful for future research focused on characterizing suitable salivary proteins as potential anti-*Leishmania* vaccine candidates.

To date, the intensive investigation of salivary proteins in certain sand fly species has allowed the generation of individual recombinant salivary proteins that have been employed as reliable markers of exposure to sand fly bites [3–5] or as the protective agent against cutaneous and visceral leishmaniases (CL and VL, respectively) under laboratory conditions [6–13]. However, most of the experiments were performed using New World VL vector *Lutzomyia longipalpis*. As the composition of salivary glands and the protective effect conferred by sand fly saliva is species-specific [14–19], it is vital to continue with detailed characterization of the salivary proteins with a special focus on sand fly species causing lethal VL.

Phlebotomus orientalis is a member of the subgenus *Larroussius* and represents the main sand fly species transmitting *Leishmania donovani* within the countries of East Africa (reviewed in [20]) as well as in Saudi Arabia [21] and Yemen [22]. At two distinct localities in Ethiopia, Addis Zemen and Melka Werer, we observed different epidemiology of VL, although *P. orientalis* was present in both places. While in Addis Zemen, human VL caused by *Le. donovani* with high mortality rate was reported [23], Melka Werer is considered to be a non-endemic area with no human cases. A recently published study compared various molecular aspects of colonies from both foci and showed that the susceptibility of Addis Zemen and Melka Werer colonies to *Le. donovani* infection was identical [24]. As Warburg et al. described the possible connection of the salivary gland composition with varying pathologies of CL [25] and sand fly saliva is known to play a crucial role in transmission of *Leishmania* spp. (e.g. [2]), we hypothesized that the composition of salivary glands may explain the different epidemiology in these Ethiopian foci. Therefore, we studied the transcriptomes, proteomes and the enzymatic activities (apyrase and hyaluronidase) in the saliva of female sand flies from Addis Zemen (VL endemic) and from Melka Werer (non-endemic). Furthermore, we characterized the main salivary antigens in both colonies and determined the level of glycosylation of *P. orientalis* salivary proteins. Importantly, we compared our data with other sand fly species from the subgenus *Larroussius*, whose cDNA libraries have already been constructed [26–28], and used

sequences of the New World sand fly species *L. longipalpis* as an outgroup.

Methods

Ethics statement

BALB/c mice were maintained and handled in the animal facility of Charles University in Prague in accordance with institutional guidelines and Czech legislation (Act No. 246/1992 coll. on Protection of Animals against Cruelty in present statutes at large). The experiments were approved by the Committee on the Ethics of Animal Experiments of the Charles University in Prague (Permit Number: 24773/2008-10001) and were performed under the Certificate of Competency (Registration Numbers: CZU 934/ 05, CZU 307/09) in accordance with the Examination Order approved by Central Commission for Animal Welfare of the Czech Republic.

Sand flies and salivary gland dissections

Two colonies of *P. orientalis* were established; one from a non-endemic lowland area in central Ethiopia, Melka Werer (MW) (altitude of 800 m), the later one from an endemic focus of VL in the highlands of Northwest Ethiopia, Addis Zemen (AZ) (altitude of 1800–2000 m), and then transferred to Czech Republic. Both sand fly colonies were kept in the insectary of Charles University in Prague and were reared under standard conditions as described in [29]. For the experiments, the sand flies from F5–F6 generation were used. Salivary glands of 1-day old adult females were dissected; mRNA was extracted and stored in RNA later (Ambion). For proteome analysis, western blot, affinity blot, and hyaluronidase assay, salivary glands from 5- to 8-day old *P. orientalis* adult females were dissected and stored in Tris buffer (20 mM Tris, 150 mM NaCl, pH 7.7). For the apyrase assay, 8-day old adult female salivary glands were dissected into Tris buffer containing 0.005% Triton X-100 and stored at −80°C.

Construction of salivary gland cDNA libraries

Salivary gland mRNA was isolated separately from 45 pairs each of MW and AZ glands using Micro-FastTrack mRNA isolation kit (Invitrogen). Both cDNA libraries were constructed following the manufacturer's instructions for SMART cDNA Library Construction Kit (BD Clontech) with some modifications as described in [30]. Each library was fractionated into large, medium, and small cDNA fragments. Gigapack III Gold Packaging Extract (Stratagene) was used for packaging the phage. Both libraries were then plated by infecting log-phase XL-1 blue *Escherichia coli* (Clontech). Transfected plaques were randomly selected and a PCR reaction with vector primers flanking the inserted cDNA was made. The presence of recombinants was checked by visualization the PCR products on 1.1% agarose gel with SYBR Safe (Invitrogen). Inserts were sequenced as previously described [31] using a ABI 3730XL DNA Sequencer (Applied Biosystems).

Bioinformatics

Detailed description of the bioinformatics analysis can be found elsewhere [28]. Briefly, expression sequence tags (ESTs) were analyzed using a customized program based on the Phred algorithm [32,33]. Sequences with Phred quality scores lower than 25 were removed, as well as vector sequences and primers. Resulting sequences were grouped based on nucleotide homology of 90% identity over 100 residues and aligned into consensus transcript sequences (contigs) using the CAP3 sequence assembly program. BLAST programs were used to compare contigs and

singletons (contigs with a single sequence) to the non-redundant protein database of the NCBI, the Gene Ontology database (GO) [34], to COG conserved domains database [35], Protein Family database (Pfam) [36], SimpleModular Architecture Tool database (SMART) [37], and to rRNA Nucleotide Sequences, and Mitochondrial and Plastid Sequence (MITPLA) databases available from NCBI. The three frame translations of each dataset were submitted to the SignalP server [38] to find signal sequences. The grouped and assembled sequences, BLAST results, and SignalP results, combined by dCAS software [39] in an Excel spreadsheet, were manually verified and annotated. N- and O-Glycosylation sites on the proteins were predicted using NetNGlyc 1.0 and NetOGlyc 3.1 software (www.cbs.dtu.dk/services/NetNGlyc, www.cbs.dtu.dk/services/NetOGlyc) [40].

Phylogenetic analysis

Protein sequences were aligned using ClustalX (version 2.0) [41] and manually refined in BioEdit 7.1.3.0 editing software. For each alignment, best substitution matrix was determined by ProtTest software 2.0 [42]. This matrix was subsequently used by TREE-PUZZLE 5.2 [43] to reconstruct maximum likelihood phylogenetic trees from the protein alignments using quartet puzzling with 1000 puzzling steps in each phylogenetic analysis. Resulting trees were visualized in MEGA 4 [44].

Proteome analysis

For mass spectrometry analysis, salivary glands of both AZ and MW *P. orientalis* colonies were dissolved in non-reducing sample buffer and electrophoretically separated in 12.5% SDS gel. Proteins within the gel were visualized by staining with Coomassie Brilliant Blue R-250 (Serva). The individual bands were cut and incubated with 10 mM dithiothreitol (Sigma) and then treated with 55 mM iodoacetamide (Sigma). Washed and dried bands were digested with trypsin (Promega). The tryptic peptides were separated by liquid chromatography using an Ultimate 3000 HPLC system (Dionex). The peptide samples diluted in 0.3% trichloroacetic acid (TCA) with 10% acetonitrile (ACN) were loaded onto a PepMap 100 C18 RP column (Dionex) at a flow rate of 300 nl per minute. The peptides were eluted by a 45-min linear gradient of 5–80% (v/v) ACN in 0.1% (v/v) TCA over a period of 20 min. The eluent was mixed 1:3 with matrix solution (20 mg/ml a-cyano-4-hydroxycinnamic acid in 80% ACN) and subsequently spotted onto MALDI target plates using a Probot microfraction collector (Dionex). Spectra were acquired on 4800 Plus MALDI TOF/TOF analyzer (Applied Biosystems/MDS Sciex) equipped with a Nd: YAG laser (355 nm, firing rate 200 Hz) as described in detail in [28].

Hyaluronidase activity analysis

Hyaluronidase activity in salivary glands of both *P. orientalis* colonies was quantified using a sensitive assay in microtitration plates coupled with biotinylated hyaluronic acid (bHA). Salivary glands were homogenized by three freeze-thaw cycles and salivary gland extract (SGE) was obtained by centrifugation at 17000 g (5 min, 2°C). Biotinylated HA, prepared as described in [45], was immobilized onto Covalink NH microtiter plates (NUNC) using the method by Frost and Stern [46] modified by [28] at a final concentration of 1 µg/well bHA. The plates were incubated overnight at 4°C and washed three times in PBS, pH 7.2 containing 2 M NaCl and 50 mM MgSO$_4$. The plates with immobilized bHA were coated for 45 min with 1% BSA in PBS, then washed and equilibrated with assay buffer (0.1 M acetate buffer, pH 5.0, 0.1 M NaCl, 0.1% Triton X-100) to adjust the pH for optimum sand fly salivary hyaluronidase activity. Four SGE

samples for each colony were pipetted into the plates in triplicate at a final concentration of 0.5 salivary gland per well and incubated for 45 min at 37°C. To obtain a standard curve ranging from 0.5 to 7.8×10^3 rTRU, hyaluronidase from bovine testes (Sigma), at a concentration of 0.01 TRU/µl, was diluted by two-fold serial dilution in 0.1 M acetate buffer, pH 4.5, 0.1 M NaCl, 0.1% Triton X-100. Wells without bHA or enzyme were used as controls. The reaction was terminated by the addition of 200 µl/ well of 6 M guanidine. After washing, avidin-peroxidase (Sigma, 2 µg/ml) was added at a final concentration of 0.2 µg/well and incubated for 30 min at room temperature. Color reaction was developed with o-phenylenediamine substrate in 0.1 M citrate-phosphate buffer, pH 5.5. Absorbance was measured at 492 nm using Infinite M 200 fluorometer (Schoeller Instruments). Raw data were evaluated by Measurement Parameters Editor Magellan 6 (Tecan) and the standard curve created using a 4-parameter logistic fit.

Apyrase activity analysis

Apyrase activity was determined using the Fiske and Subbarow method for measuring inorganic phosphate (Pi) released from ADP or ATP [47], with some modifications. Salivary glands were homogenized by one freeze-thaw cycle combined with a mechanical homogenization. Two µl of salivary gland homogenate (SGH) diluted 1:25 in assay buffer (50 mM TRIS 150 mM NaCl, pH 8.5 with 5 mM CaCl$_2$ or 5 mM MgCl$_2$) were mixed in wells with 78 µl of assay buffer and 20 µl of substrate to obtain a final concentration of 2 mM ATP or ADP and 1/25 of gland pair per well. SGH samples were pipetted into the microtiter plate in series of six. Wells containing only assay buffer were used as negative controls. Plates were incubated for 15 min at 37°C. Then the enzymatic reaction was stopped by addition of 25 µl of 1.25% ammonium molybdate in 1.25 M sulfuric acid and 5 µl of Fiske-Subbarow reducer (25 mg/ml, F5428 Sigma) per well. The colorimetric reaction was read after 15 min by Tecan Infinite M 200 fluorometer (Schoeller Instruments) at 665 nm. The amount of Pi released from substrate was determined using potassium dihydrogen phosphate as a standard. The study of pH optimum was carried out within a range of pH 6.0–9.5. Salivary glands of *P. papatasi*, the species with previously described apyrase activity [48], were used as a positive control. Amount of proteins within SGHs was determined using Bio-Rad *DC* Protein Assay with BSA as a standard according to the manufacturer's instructions.

Western blotting

Salivary glands of both *P. orientalis* colonies were separated by SDS-PAGE on 10% gel under non-reducing conditions using Mini-Protean III apparatus (Biorad). Salivary proteins were transferred from gel to nitrocellulose membrane (NC) by Semi-Phor equipment (Hoefer Scientific Instruments) and cut into strips. The strips were then blocked with 5% low fat dry milk in Tris-buffered saline with 0.05% Tween 20 (TBS-Tw) and subsequently incubated with BALB/c mice sera (AZ – mice bitten 18 times in a week interval; MW – mice bitten 17 times in a week interval), diluted 1:100 in TBS-Tw, for 1 hour. After the washing with TBS-Tw, the strips were incubated for 1 hour with peroxidase-conjugated goat anti-mouse IgG (Serotec) diluted 1:1000 in TBS-Tw. The chromogenic reaction was developed using a substrate solution containing diaminobenzidine and H$_2$O$_2$.

Affinity blotting

Affinity blotting was performed using salivary glands from MW *P. orientalis* colony separated by SDS-PAGE as described above. After transfer, free binding sites on NC membrane were blocked

with 5% bovine serum albumin in 20 mM TBS-Tw overnight at 4°C. The strips were then incubated for 1.5 hour on the shaker at room temperature with biotinylated lectins from *Dolichos biflorus* (DBA, Vector), *Glycine max* (SBA, Vector), *Ulex europaeus* (UEA-I, Vector), *Tetragonolobus purpureus* (LTA, Sigma), *Canavalia ensiformis* (ConA, Sigma), and *Pisum sativum* (PSA, Vector). Based on the preliminary experiments with different lectin concentrations, the lectins were diluted: 5 µg/ml, 10 µg/ml, 10 µg/ml, 0.2 µg/ml, 0.1 µg/ml and 10 µg/ml in TBS-Tw, respectively. To control the reaction specificity the aforementioned lectins were pre-incubated for 30 min with the appropriate saccharide inhibitors (Sigma) as follows: 0.25 M N-acetyl-D-galactosamine for DBA and SBA, 0.5 M L-fucose for UEA-I and LTA, 0.5 M methyl-α-D-mannopyranoside for ConA and PSA, and subsequently applied on the strips. After the washing with TBS-Tw, streptavidin-peroxidase (Sigma) was added to strips at a final concentration of 1 µg/ml and incubated for 1 h on the shaker at room temperature. The chromogenic reaction was developed as mentioned above.

Results and Discussion

Sequencing of *P. orientalis* salivary gland cDNA libraries

Two cDNA libraries were constructed from salivary glands of *P. orientalis* colonies originating in Addis Zemen and Melka Werer, Ethiopia. For each cDNA library, 940 clones were randomly selected and sequenced, which resulted in 835 and 749 high quality sequences from AZ and MW, respectively. Based on nucleotide homology, sequences were clustered into contigs, analyzed using the dCAS cDNA annotation software [39] and subsequently verified by manual annotation. From the AZ cDNA library, sequences were assembled into 263 contigs, where 185 of them were singletons (one sequence per contig). From the MW cDNA library, we obtained 242 contigs, including 171 singletons. In accordance with previously published cDNA libraries from sand fly salivary glands, the most abundant transcripts were those coding for putative salivary proteins (607 out of 835 in AZ; 567 out of 749 in MW). Of the nucleotide sequences encoding putative salivary proteins, 574 (AZ) and 506 (MW) salivary transcripts encoded a predicted signal peptide sequence. Those that did not possess sequences encoding a signal peptide were truncated at the 5′ end. Most of the contigs coding for putative salivary proteins were comprised of more than one sequence (averaging 7.14 sequences per contig in AZ and 6.23 in MW), whereas housekeeping proteins or proteins with unknown function were mostly represented by singletons. All obtained ESTs were deposited in the NCBI dbEST database under accession numbers JZ479238–JZ480094 for AZ colony and JZ480095–JZ480885 for MW colony.

Members of 13 main protein families were found among the putative salivary proteins of the two *P. orientalis* colonies: apyrase, yellow-related protein, antigen 5-related protein, odorant-binding proteins (D7-related and PpSP15-like proteins), hyaluronidase, endonuclease, phospholipase, pyrophosphatase, amylase, PpSP32-like protein, ParSP25-like protein, SP16-like protein, and Lufaxin (SP34-like protein). Detailed descriptions of each protein family are listed in the following paragraphs. Interestingly, we did not detect any sequences coding for adenosin deaminase in either *P. orientalis* cDNA library. Thus, we expect that *P. orientalis* saliva contains adenosin and ADP/AMP; leaving only *P. duboscqi*, *L. longipalpis*, and *L. intermedia* [49–52] as the sand fly species identified to produce adenosine deaminase, to date.

BLAST comparison of translated nucleotide sequences with the non-redundant (NR) protein database showed high similarity with salivary proteins of *P. perniciosus* and *P. tobbi* (both subgenus *Larroussius*). Sporadically, the best match was found with salivary proteins of *P. arabicus* (subgenus *Adlerius*) or *P. argentipes* (subgenus *Euphlebotomus*). Representative sequences of putative salivary proteins from both *P. orientalis* colonies that were deposited into NCBI GenBank database are listed in Table 1 and Table 2. Both tables include GenBank accession numbers, the predicted molecular weight, isoelectric point, best match to the NR database, the sand fly species with the highest homology, and presence in the proteome.

Proteome analysis

Salivary proteins presented in the proteome were identified by mass spectrometry and are shown in Figure 1. In both cDNA libraries, 12 salivary proteins were determined to be present in proteome. In Addis Zemen colony, the identified proteins were two yellow-related proteins (PorASP2/KC170933; PorASP4/KC 170934), three apyrases (PorASP11/KC170935; PorASP14/KC170936; PorASP15/KC170937), two D7-related proteins (Por ASP48/KC170943; PorASP122/KC170954), two antigen 5-related proteins (PorASP74/KC170947; PorASP76/KC170948), and three PpSP15-like proteins (PorASP28/KC170938; PorASP37/KC170940; PorASP61/KC170944) (Figure 1). In Melka Werer colony, the identified proteins were two yellow-related proteins (PorMSP23/KC170966; PorMSP24/KC170967), two apyrases (PorMSP3/KC170960; PorMSP4/KC170961), three D7-related proteins (PorMSP28/KC170969; PorMSP38/KC170970; Por MSP67/KC170973), two antigen 5-related proteins (PorMSP6/KC170962; PorMSP8/KC170963), and three PpSP15-like proteins (PorMSP12/KC170964; PorMSP74/KC170974; PorMSP96/KC170978) (Figure 1). Except for apyrase, none of the salivary enzymes identified in *P. orientalis* transcriptomes were detected in proteome analysis, even though all of the nucleotide sequences coding for these salivary proteins possessed signal peptides. It might be explained by the fact that extremely active enzymes do not need a huge amount of protein to be effective.

Yellow-related proteins

Yellow-related proteins are abundantly expressed in the sand fly salivary glands and have been detected in the saliva of all sand fly species tested, to date [7,26–28,31,50,52–56]. Two yellow-related proteins were found in the cDNA library of the AZ (PorASP2/KC170933; PorASP4/KC170934) as well as the MW (PorMSP23/KC170966; PorMSP24/KC170967) *P. orientalis* colony). All four *P. orientalis* yellow-related proteins had similar predicted molecular mass (41.5–42.3 kDa) and wide range of pI (6.1–8.1) (Table 1, Table 2). All obtained sequences contained the entire major royal jelly protein (MRJP) domain, which is characteristic for the yellow-related proteins. Some advances have been also made in describing the function of sand fly yellow-related proteins. It was shown that recombinant yellow-related proteins from *L. longipalpis* saliva (AAD32198, AAS05318) act as high affinity binders of prohemostatic and proinflammatory biogenic amines such as serotonin, catecholamines and histamine [12]. Similarly, the amino acid motif present in the ligand binding pocket of *L. longipalpis* (T-x(52,63)-Y-Q-x(85,90)-[FY]-x(44,46)-F-x(54)-[IVL]-x(45,46)-[FY]-x-[TS]-D-x(13)-[NT]-x-[QHFL]) was discovered in the yellow-related proteins of *L. ayacuchensis* (BAM69111, BAM69185, BAM69109, BAM69110) [56] and *L. intermedia* (AFP99235) [52], but also in *P. orientalis* and other sand fly species from the subgenus *Larroussius* tested (Figure 2). These findings suggest similar anti-inflammatory function of these salivary proteins in other *Lutzomyia* and *Phlebotomus* sand fly species [12] and could potentially explain the lectin-like properties of 42 kDa yellow-related protein from *P. duboscqi* saliva

Table 1. Salivary gland transcripts of *Phlebotomus orientalis* – Addis Zemen colony.

Cluster	Sequence name	Accession number	Proteome	MW	pI	Best match to NR protein database		
						Accession number	Species	E-value
PorASP2	42 kDa yellow-related salivary protein	KC170933	Y	41.54	6.09	ABA43049	*Phlebotomus perniciosus*	0.0
PorASP4	42.6 kDa yellow-related salivary protein	KC170934	Y	42.31	8.07	ADJ54080	*Phlebotomus tobbi*	0.0
PorASP11	35.5 kDa salivary apyrase	KC170935	Y	35.53	9.95	ABB00906	*Phlebotomus perniciosus*	0.0
PorASP14	35.2 kDa salivary apyrase	KC170936	Y	35.08	8.99	ADJ54077	*Phlebotomus tobbi*	0.0
PorASP15	35.2 kDa salivary apyrase	KC170937	Y	35.33	9.16	ADJ54077	*Phlebotomus tobbi*	0.0
PorASP28	14.6 kDa PpSP15-like salivary protein	KC170938	Y	14.53	8.88	ADJ54089	*Phlebotomus tobbi*	2e-75
PorASP31	14.4 kDa PpSP15-like salivary protein	KC170939		14.32	8.73	ADJ54088	*Phlebotomus tobbi*	6e-77
PorASP37	14.9 kDa PpSP15-like salivary protein	KC170940	Y	14.91	8.77	ADJ54084	*Phlebotomus tobbi*	3e-73
PorASP40	3.7 kDa-like salivary protein	KC170941		3.93	9.16	ADJ54106	*Phlebotomus tobbi*	2e-07
PorASP46	27 kDa D7-related salivary protein	KC170942		26.68	6.36	ABA43052	*Phlebotomus perniciosus*	4e-151
PorASP48	27.1 kDa D7-related salivary protein	KC170943	Y	26.93	8.26	ADJ54095	*Phlebotomus tobbi*	9e-162
PorASP61	13.8 kDa PpSP15-like salivary protein	KC170944	Y	13.88	9.07	ADJ54086	*Phlebotomus tobbi*	1e-68
PorASP64	14.7 kDa PpSP15-like salivary protein	KC170945		14.70	7.99	ADJ54085	*Phlebotomus tobbi*	8e-62
PorASP68	5.0 kDa-like salivary protein	KC170946		4.89	9.84	ADJ54105	*Phlebotomus tobbi*	5e-15
PorASP74	28.8 kDa antigen 5-related salivary protein	KC170947	Y	28.78	8.94	ADJ54083	*Phlebotomus tobbi*	3e-151
PorASP76	30 kDa antigen 5-related salivary protein	KC170948	Y	28.78	8.94	ABA43055	*Phlebotomus perniciosus*	1e-179
PorASP80	30 kDa salivary phospholipase A2	KC170949		29.66	8.44	ABA43062	*Phlebotomus perniciosus*	0.0
PorASP86	24.53 kDa PpSP32-like salivary protein	KC170950		24.97	10.14	ADJ54102	*Phlebotomus tobbi*	2e-125
PorASP98	4.5 kDa-like salivary protein	KC170952		5.63	10.51	ADJ54097	*Phlebotomus tobbi*	3e-18
PorASP106	38.8 kDa ParSP25-like salivary protein	KC170953		27.61	4.72	ADJ54098	*Phlebotomus tobbi*	1e-140
PorASP112	salivary hyaluronidase	KC170958		37.22	6.50	ACS93505	*Phlebotomus arabicus*	1e-178
PorASP122	27 kDa D7-related salivary protein SP10	KC170954	Y	26.76	9.20	ABA43058	*Phlebotomus perniciosus*	6e-155
PorASP139	41 kDa salivary endonuclease	KC170955		41.66	9.27	ABA43064	*Phlebotomus perniciosus*	0.0
PorASP150	16 kDa salivary protein A	KC170956		16.04	5.04	ACS93506	*Phlebotomus arabicus*	1e-42
PorASP262	47 kDa pyrophosphatase-like salivary protein SP132	KC170959		32.88	7.18	ABA12155	*Phlebotomus argentipes*	8e-163

Putatively secreted salivary proteins from AZ *Phlebotomus orientalis* colony with the number of cluster, GenBank accession number, presence in proteome, putative mature protein features (MW- molecular weight, pI- isoelectric point), and best match to NR protein database.
doi:10.1371/journal.pntd.0002709.t001

[57]. Sand fly yellow-related proteins share homology with the yellow protein of *Drosophila melanogaster* and to the MRJPs of honeybees. Similarly, sequences with homology to *D. melanogaster* yellow protein were also found in other bloodsucking insects; for example, the mosquito *Aedes aegypti* [58] and the tsetse fly *Glossina morsitans morsitans* [59].

Phylogenetic analysis shows that yellow-related proteins from *P. orientalis* saliva are divided into two clades (Figure 3). Both clades are represented by two yellow-related salivary proteins, one from each *P. orientalis* cDNA library (clade I - PorASP2, PorMSP23; clade II - PorASP4, PorMSP24). *Phlebotomus orientalis* sequences within the same clade revealed high degree of identity (99 and 100%, respectively), while comparison between clades showed 77% identity (Figure 2). Yellow-related proteins of other sand fly species from subgenus *Larroussius* were also split into two clades and these sequences are closely related to *P. orientalis* proteins (83–91% identity) (Figures 2 and 3).

Yellow-related proteins were shown to be highly immunogenic. These proteins were recognized by sera of repeatedly bitten hosts such as mice [5,15,60], hamsters [60], dogs [4,61–63], foxes [4,64], and humans [4,15,65–67]. Furthermore, recombinant yellow-related salivary proteins (AAD32198, AAS05318) were

successfully employed as the markers of sand fly exposure for individuals in endemic areas [3,4]. Importantly, *L. longipalpis* salivary yellow-related proteins seem to be promising candidates for anti-*Leishmania* vaccine. Inoculation of plasmids coding for *L. longipalpis* yellow-related salivary proteins (AAD32198, AAS05318) into the skin elicited a strong delayed type hypersensitivity (DTH) reaction in various hosts [8,12,68], which resulted in efficient killing of *Le. infantum chagasi* parasites *in vitro* [68] and protection against *Le. major* infection *in vivo* [12,13].

According to the glycosylation prediction servers (NetNGlyc and NetOGlyc) we found out that PorASP4 and PorMSP24 are likely N-glycosylated and have three threonine sites for potential O-glycosylation. PorASP2 and PorMSP23 have four threonines for potential O-glycosylation and no N-glycosylation was predicted.

Apyrase

Sequences coding for apyrase were detected in the cDNA libraries of both the AZ (PorASP11/KC170935; PorASP14/KC170936; PorASP15/KC170937) and the MW (PorMSP3/KC170960; PorMSP4/KC170961) *P. orientalis* colonies. All sequences had similar predicted molecular mass (33.2–35.6 kDa) and

Table 2. Salivary gland transcripts of *Phlebotomus orientalis* – Melka Werer colony.

Cluster	Sequence name	Accession number	Proteome	MW	pI	Best match to NR protein database		
						Accession number	Species	E-value
PorMSP3	35.5 kDa salivary apyrase	KC170960	Y	35.63	8.83	ABB00906	*Phlebotomus perniciosus*	0.0
PorMSP4	35.2 kDa salivary apyrase	KC170961	Y	33.22	8.89	ADJ54077	*Phlebotomus tobbi*	0.0
PorMSP6	30 kDa antigen 5-related salivary protein	KC170962	Y	28.78	8.94	ABA43055	*Phlebotomus perniciosus*	1e-179
PorMSP8	28.8 kDa antigen 5-related salivary protein	KC170963	Y	28.78	8.94	ADJ54083	*Phlebotomus tobbi*	3e-151
PorMSP12	14.9 kDa PpSP15-like salivary protein	KC170964	Y	14.9	8.77	ADJ54084	*Phlebotomus tobbi*	3e-73
PorMSP15	24.53 kDa PpSP32-like salivary protein	KC170965		25.02	10.24	ADJ54102	*Phlebotomus tobbi*	1e-127
PorMSP23	42 kDa yellow-related salivary protein	KC170966	Y	41.59	6.09	ABA43049	*Phlebotomus perniciosus*	0.0
PorMSP24	42.6 kDa yellow-related salivary protein	KC170967	Y	42.31	8.07	ADJ54080	*Phlebotomus tobbi*	0.0
PorMSP27	putative alpha-amylase	KC170968		33.4	5.75	ACS93490	*Phlebotomus arabicus*	4e-178
PorMSP28	27.0 kDa D7-related salivary protein	KC170969	Y	27.27	7.53	ADJ54096	*Phlebotomus tobbi*	1e-156
PorMSP38	27.1 kDa D7-related salivary protein	KC170970	Y	26.94	8.26	ADJ54095	*Phlebotomus tobbi*	3e-162
PorMSP43	27 kDa D7-related salivary protein SP04B	KC170971		26.7	6.71	ABA43052	*Phlebotomus perniciosus*	1e-151
PorMSP65	38.8 kDa ParSP25-like salivary protein	KC170972		27.56	4.78	ADJ54098	*Phlebotomus tobbi*	6e-140
PorMSP67	27 kDa D7-related salivary protein	KC170973	Y	26.76	9.2	ABA43058	*Phlebotomus perniciosus*	6e-155
PorMSP74	13.8 kDa PpSP15-like salivary protein	KC170974	Y	13.92	9.18	ADJ54086	*Phlebotomus tobbi*	4e-70
PorMSP75	14.7 kDa PpSP15-like salivary protein	KC170975		14.7	7.99	ADJ54085	*Phlebotomus tobbi*	8e-62
PorMSP78	33 kDa salivary lufaxin	KC170976		18.78	8.4	ABA43054	*Phlebotomus perniciosus*	4e-99
PorMSP90	14.4 kDa PpSP15-like salivary protein	KC170977		14.32	8.73	ADJ54088	*Phlebotomus tobbi*	6e-77
PorMSP96	14.6 kDa PpSP15-like salivary protein	KC170978	Y	14.5	8.88	ADJ54089	*Phlebotomus tobbi*	1e-74
PorMSP101	41 kDa salivary endonuclease	KC170979		41.7	9.44	ABA43064	*Phlebotomus perniciosus*	0.0
PorMSP104	4.5 kDa-like salivary protein	KC170980		5.63	10.97	ADJ54097	*Phlebotomus tobbi*	3e-18
PorMSP108	salivary hyaluronidase	KC170981		35.6	7.98	ACS93505	*Phlebotomus arabicus*	2e-163
PorMSP129	30 kDa salivary phospholipase A2	KC170982		29.72	8.31	ABA43062	*Phlebotomus perniciosus*	0.0
PorMSP162	16 kDa salivary protein A	KC170983		15.97	5.04	ACS93506	*Phlebotomus arabicus*	1e-41
PorMSP169	3.7 kDa-like salivary protein	KC170984		3.93	9.16	ADJ54106	*Phlebotomus tobbi*	2e-7
PorMSP196	5.0 kDa-like salivary protein	KC170985		4.97	10.18	ADJ54105	*Phlebotomus tobbi*	2e-14

Putatively secreted salivary proteins from MW *Phlebotomus orientalis* colony with the number of cluster, GenBank accession number, presence in proteome, putative mature protein features (MW- molecular weight, pI- isoelectric point), and best match to NR protein database.
doi:10.1371/journal.pntd.0002709.t002

pI ranged from 8.8 to 10.0. Apyrase is the principal anti-platelet aggregation enzyme which hydrolyses ATP and ADP to AMP and orthophosphate, thereby blocks the physiological signal of damaged cells and tissues. This enzyme has been found in the saliva of all sand fly species tested, to date [7,26–28,31,50,52–56], but also in the saliva of other medically important hematophagous insect such as triatomine bugs (e.g. [69]), mosquitoes (e.g. [70]), fleas (e.g. [71]), tsetse flies (e.g. [72]), biting midges (e.g. [73]), and horseflies (e.g. [74]). Interestingly, apyrase has also been described in non-bloodsucking insects; for example, *Helicoverpa zea* [75] and *D. melanogaster* [76], indicating that apyrase may have a broader functional potential than only the facilitation of blood acquisition.

Apyrases of the bloodfeeding insect are divided into three families: GTPase/CD-39, 5′- nucleotidase, and *Cimex* type (reviewed in [77]). Apyrases from *P. orientalis* colonies, as well as from other sand fly species, are homologous to the bed bug apyrase which defined the *Cimex* type family [78]. Phylogenic analysis showed that salivary apyrases from *P. orientalis* colonies are separated into two clades (Figure 4). Clade I includes two apyrases from the AZ colony and one from the MW colony (PorASP14, PorASP15, and PorMSP4), as the analogue of the second apyrase

in MW colony was excluded from the phylogenetic analysis due to absence of signal peptide and the low quality of sequence. Clade II contains two apyrases, one from each colony (PorASP11 and PorMSP3). Sequences of *P. orientalis* apyrase within the same clade revealed high degree of identity (95–99%), whereas the comparison between the clades showed an identity of 66% (Figure S1). Comparison with other sand fly species from the subgenus *Larroussius* showed that apyrases from *P. tobbi* (ADJ54077, ADJ54078) and *P. perniciosus* (ABB00906, ABB00907) saliva are closely related to *P. orientalis*, while apyrase from *P. ariasi* (AAX56357) saliva is more distinct (Figures 4 and S1).

Apyrase activity has been demonstrated in the saliva of *L. longipalpis* [53], *P. papatasi* [48,79], *P. duboscqi* [80], *P. perniciosus*, *P. argentipes*, and *P. colabaensis* [48]. In our experiments, apyrase activity was measured in the saliva of both AZ and MW *P. orientalis* colonies (Table 3). ATPase as well as ADPase activity, determined per the pair of salivary glands, was slightly higher in AZ colony, but recalculation of enzymatic activity per milligram of total proteins showed that apyrase activity in both colonies is comparable. Also, the ATPase/ADPase ratio was equal in both colonies (Table 3). Comparison of *P. orientalis* apyrase activity with
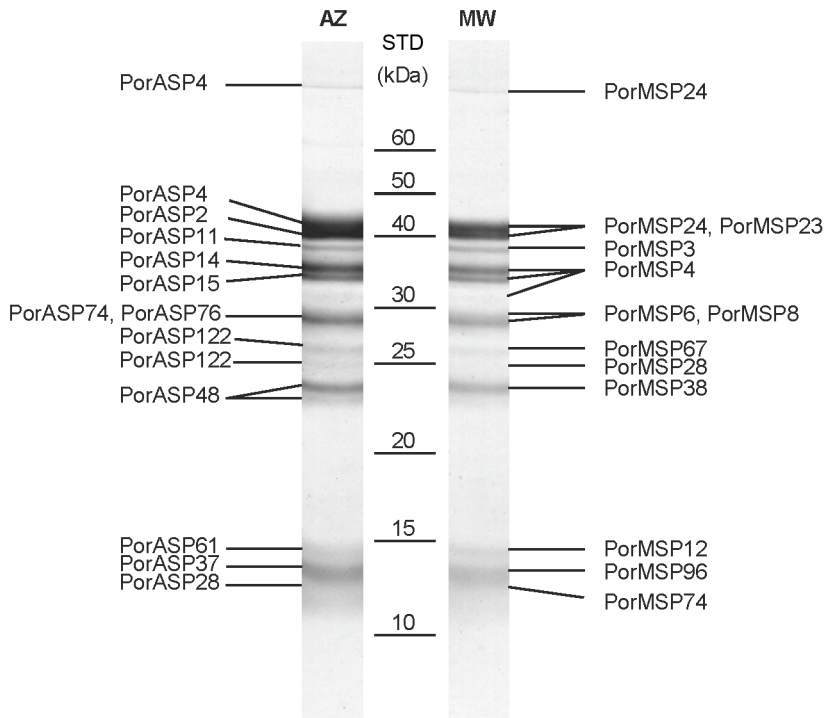
Figure 1. Proteomic analysis of salivary gland homogenates from *Phlebotomus orientalis.* *Phlebotomus orientalis* salivary proteins from Addis Zemen (AZ) and Melka Werer (MW) colonies (Ethiopia) were identified using Mass Spectrometry. The name of sequences contained in each protein band and molecular weight in kDa (STD/kDa) are indicated.
doi:10.1371/journal.pntd.0002709.g001

P. perniciosus [48] revealed that ATPase and ADPase activities determined per pair of glands are comparable. Additionally, in accordance with previous data [48,53,79–81], we showed that *P. orientalis* apyrase activity is dependent on presence of Ca^{2+} but not on Mg^{2+} ions.

Besides the anti-hemostatic effect of this enzyme, apyrase is also known as a powerful antigen. Specific antibodies from dogs bitten by *P. perniciosus* in the field, as well as under laboratory conditions, reacted strongly with two salivary apyrases [63]. Apyrases from *P. perniciosus*, *P. papatasi*, and *P. argentipes* saliva were also recognized by sera of mice and hamsters immunized by homologous antigen [5,60,82]. Furthermore, bacterially expressed recombinant *P. duboscqi* apyrase (ABI20147) was also recognized by specific antibodies from mice immunized with *P. duboscqi* saliva [80], suggesting that antibody recognition is not solely targeted to the glycosylated parts of the antigen. On the other hand, inoculation of bacterially expressed recombinant *L. longipalpis* apyrase (AAD33513) into C57BL/6 mice did not elicit either antibody response or DTH reaction [12]. These data indicates that the immunogenicity of the protein or saccharidic part of antigen may vary in different sand fly species. According to the glycosylation prediction servers (NetNGlyc and NetOGlyc), *P. orientalis* apyrases PorASP14, PorASP15, and PorMSP4 are N-glycosylated, while no O-glycosylation sites were predicted.

Hyaluronidase

Hyaluronidase is an enzyme that degrades hyaluronic acid and other glycosaminoglycan constituents abundantly present in the vertebrate extracellular matrix. It is a well-known allergen occurring in the venom of bees, hornets, wasps, spiders, and snakes (reviewed in [83,84]), but hyaluronidase activity was also observed in the saliva of various bloodsucking Diptera [28,31,45,85,86]. Previously

published data showed that hyaluronidase is able to promote the spreading of other components of bloodfeeding insect saliva within the skin, as well as to enhance the success of potential parasite transmission [86]. Although positive enzymatic activity was detected in all sand fly species tested to date [28,31,45,53,85,86], transcripts coding for putative hyaluronidase were ascertained only in four of them, namely *P. arabicus* (ACS93505), *P. tobbi* (AEK98519), *L. longipalpis* (AAD32195), *L. intermedia* (AFP99265) [28,31,52–54], and in both *P. orientalis* colonies (PorASP112/KC170958; PorMSP108/ KC170981) (Table 1, Table 2). The predicted molecular mass of AZ and MW hyaluronidase was 37.2 and 35.6 kDa, respectively, and the pI was 6.5 and 8.0, respectively.

Hyaluronidase activity measured in the *P. orientalis* saliva was found to be lower than the activity of other *Larroussius* species tested by [28]. While hyaluronidase activity expressed in the relative Turbidity Reducing Units (rRTU) reached approximately 0.62 rTRU/gland in *P. tobbi* and 0.48 rTRU/gland in *P. perniciosus* [28], enzymatic activity in *P. orientalis* saliva was 0.22 rTRU per gland (0.22±0.036 rTRU in AZ and 0.215±0.045 rTRU in MW).

Phlebotomus orientalis salivary hyaluronidase of AZ and MW colonies revealed identity reaching 94% (Figure 5). High degree of identity was achieved with *P. tobbi* sequences (AEK98519) (89–93%), followed by *P. arabicus* (ACS93505) (80–83%), *L. longipalpis* (AAD32195) (56–58%), and *L. intermedia* (AFP99265) (47–48%).

Moreover, glycosylation prediction servers (NetNGlyc and NetO-Glyc) showed that salivary hyaluronidase is the most glycosylated protein in both colonies, with seventeen predicted N-glycosylation sites in the AZ and sixteen in the MW colony (Figure 5).

Other enzymes

Another enzyme that was identified among the transcripts from both *P. orientalis* cDNA libraries is a putative endonuclease

```
              ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
PorASP4       FHVEREYAWKNITYEGIDPASYNIENSIPTAFAHDAASKKIFITVPRRN-QIPFTLTELDTTKHP----ERSPPLSKFPG  75
PorMSP24      FHVEREYAWKNITYEGIDPASYNIENSIPTAFAHDAASKKIFITVPRRN-QIPFTLTELDTTKHP----ERSPPLSKFPG  75
PtobSP38      FHVEREYAWKNISYEGIDPASFNIENIIPTGFAHDAAIKKIFFAVPRRFPQIPFTLTELDTAKHP----ERSPPLSKFPG  76
PperSP03B     FHVEREYAWKNISYEGVDPALFNIDNIIPTGFVHDAINKKIFIAVPRRSPQIPFTLTELDTTKHP----ERSPPLSKFPG  76
PariSP04B     YHVEREYAWKNITFEGIDQASYNIENSIPTAFVHDALSKKIILAIPRLYPQVPITLTQLDTTKHP----ERSPPLEKFPG  76
PorASP2       YHIEREYAWKNIIYEGINPASYNIENSIPTAFAHDATSKKIFITVPRIN-PVPITLTEFDTTKHP----EGSPSLSKFPG  75
PorMSP23      YHIEREYAWKNIIYEGINPASYNIENSIPTAFAHDATSKKIFITVPRIN-PVPITLTEFDTTKHP----EGSPSLSKFPG  75
PtobSP37      YHIEREYAWKNIIYEGIDPASYNVKNGIATGFAHDAASKKIFITVPRLN-PVPFTLTELDISKHP----GGSPSLSKFPG  75
PperSP03      YDIEREYAWKNISFEGIDPASYSVKNSIVTGFAHDADSKKIFITIPRLN-PVPITLTELDTTKHP----EGSPSLSKFPG  75
PariSP04      SQIEREYAWKNIIYEGIDQGSYNIENSIPTAFAHDAASKKIFITIPRIN-QVPITLTEFDSIKYP----GGSPPLSKFPG  75
LuloYELLOW    -ADTQGYKWKQLLYNNVTPGSYNPDNMISTAFAYDAEGEKLFLAVPRKLPRVPYTLAEVDTKNSLGVKGKHSPLLNKFSG  79
                                                                      ~

              ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
PorASP4       SDK---LISVYQPVIDECRRLWIADVGQVEYKG-DEQKYPKQKAAIIAYDLTKDNYPEIDRYEIPNNVAG-NSLGFGGFA  150
PorMSP24      SDK---LISVYQPVIDECRRLWIADVGQVEYKG-DEQKYPKQKAAIIAYDLTKDNYPEIDRYEIPNNVAG-NSLGFGGFA  150
PtobSP38      SDK---LINVYQPVIDECRRLWIADVGRINYKG-DEQKYPTQKAVIIAYDLTKDNYPEIDRYEIPSKIAGPNPIGFGGFA  152
PperSP03B     SDK---LINVYQPVIDECRRLWIADVGRVDYKG-DEQKYPNQNAVLIAYDLTKENYPEIHRYEIPSKIAGSNTIPFGGFA  152
PariSP04B     SDK---LTSVYQPMLDECRRLWIVDVGQVEYKG-DEQKYPKKNPAIIAYDLTKDNYPEIDRYEIPINIAG-NQIGFGGFT  151
PorASP2       SDK---IISVYQPVIDECRRLWIADVGQVEYKE-DEKKFPKQNAAIIAYDLTKDNYPEIDRYEIPSIVAG-NPLGFGGFA  150
PorMSP23      SDK---IISVYQPVIDECRRLWIADVGQVEYKE-DEKKFPKQNAAIIAYDLTKDNYPEIDRYEIPSIVAG-NPLGFGGFA  150
PtobSP37      GDK---LISVYQPVIDECRRLWIVDVGQVEYKD-DEQKFPKQNAAIIAYDLTKANYPEIDRYEIPSNVAC-DSLGFGGFT  150
PperSP03      SDK---LISVYQPVIDECRRLWIVDAGQVEYKG-DEQKIPKKNAIIAYDLTKDNYPEIDRYEIPNNVAG-NPLGFGGFA  150
PariSP04      SDN---IISVYQPVIDECRRLWIVDAGQVEYKG-DEQKYPKKNPAIIAYDLTKDNYPEIDRYEIPINIAG-NPLGFGGFT  150
LuloYELLOW    HKTGKELTSIYQPVIDDCRRLWVVDIGSVEYRSRGAKDYPSHRPAIVAYDLKQPNYPEVVRYYFPTRLVE-KPTYFGGFA  158
                        ~~

              ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
PorASP4       VDVTNPKEGCCNTFVYITNFEDNTLIVYDQEKKDSWKISHGSFKPEHDSTLSHNGEQYKYRVGIFGIALGDRDPEGNRPA  230
PorMSP24      VDVTNPKEGCCNTFVYITNFEDNTLIVYDQEKKDSWKISHGSFKPEHDSTLSHNGEQYKYRVGIFGIALGDRDPEGNRPA  230
PtobSP38      VDVTNPKEGCCKTFIYITNFDDNTLIVYDQEKKDSWKISHGSFKPEHDSTLSHNGEQYKYRVGIFGITLGDRDPEGNRPA  232
PperSP03B     VDVTNPKEGCCKTFVYITNFEDNTLIVYDQEKKDSWKISHGSFKPEHDSTLSHDGKQYKYRVGLFGITLGDRDPEGNRPA  232
PariSP04B     VDVTNPKEGCCKTFIYITNFEDNTLIVYDQEKKDSWKISHGSFKPEHESNFSHNGAQYKYAGIFGITLGDRDPEGNRPA  231
PorASP2       VDVTNPKGGCGKTFIYIANFNDNTLIVYDQEKKDSWKISHGTFKPEHESILTHNGAQHILKLGIFGITLGDLDAEGNRPA  230
PorMSP23      VDVTNPKGGCGKTFIYIANFNDNTLIVYDQEKKDSWKISHGTFKPEHESILTHNGAQHILKLGIFGITLGDLDAEGNRPA  230
PtobSP37      VDVTNPKEGCAKTFVYIANFVDNALIVYDQEKKDSWKISHGSFKPEHDSILIHDGAQHILKLGIFGITLGDLDAEGNRPA  230
PperSP03      VDVTNPKEGCGKTFVYITNFEDNTLIVYDQEKKDSWKISHDSFKPEHESILTHNGAQHILKLGIFGITLGDLDEEGNRQA  230
PariSP04      VDVTNPKEGCGKTFIYITNFEDNTLIVYDQEKKDSWKISHGSFKPEHESILIHNGVDHILKLGIFGITLGDRDSEGNRPA  230
LuloYELLOW    VDVANPKGDCSETFVYITNFLRGALFIYDHKKQDSWNVTHPTFKAERPTKFDYGGKEYEFKAGIFGITLGDRDSEGNRPA  238
                ~                                               ~

              ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
PorASP4       YYIAGSSTKLFEISTKILKQKGAKFDPVNLGNRGPHSEAIALAYDPKTKVIFFAEADSRQISCWNIQKP-LNHKNTDVIY  309
PorMSP24      YYIAGSSTKLFEISTKILKQKGAKFDPVNLGNRGPHSEAIALAYDPKTKVIFFAEADSRQISCWNIQKP-LNHKNTDVIY  309
PtobSP38      YYIAGSSTKLFEISTKILKQKGAKFDPVNLGNRGPHSEAVALAYDPKTKVIFFAESDSRQISCWNIQKP-LNHKNIDVIY  311
PperSP03B     YYIAGSSTKLFEISTKILKEKGAKFDPVNLGNRGPHTEAVALVYDPKTKVIFFAESDSRQVSCWNTQKP-LNHKNTDVIF  311
PariSP04B     YYLGGSSTKLFEVSTEALKKKGAKFDPVRLGDRGRHTEAIALVYDPKTKVIFFAESDSRQISCWNTQKP-LNHKNTDVIY  310
PorASP2       YYLGGSSTKLFRVNTKDLKKKDGQIEPTPLGDRGSHSEALALAYDPKTKVIFFTEYNSKRISCWNTQKS-LNHDSIDAIY  309
PorMSP23      YYLGGSSTKLFRVNTKDLKKKDGQIEPTPLGDRGSHSEALALAYDPKTKVIFFTEYNSKRISCWNTQKS-LNHDSIDVIY  309
PtobSP37      YFLGGSSTKLFKVNTKDLKKKNGQIEPIPLGDRGSNSEALALAYDPKTKVIFFIEYNSKRISCWNTQNS-LNHENIDVIY  309
PperSP03      YYLGGSSTKLFRVNTKDLKKKAGQIEFTPLGDRGSHSEALALAYDPKTKVIFFIEYNSKRISCWNTQKS-LNPDNIDVIY  309
PariSP04      YYLGGSSTKLFEVNTKALKKKEGEIEPITLGDRGPHSEAIALAYDPKTKVIFFTEYNSKKISCWNIKKP-LIHDNMDKIY  309
LuloYELLOW    YYLAGSAIKVYSVNTKELKQKGGKLNPELLGNRGKYNDAIALAYDPKTKVIFFAEANTKQVSCWNTQKMPLRMKNTDVVY  318
                                                                      ~

              ....|....|....|....|....|....|....|....|....|....|....|....|....|....|...
PorASP4       ASSKFIFGTDISVDSESQLWFLSNGQPPIDN-LKLTFDKPHIRLMR-VDTAKAIRRTKCEVKRVKP-  373
PorMSP24      ASSKFIFGTDISVDSESQLWFLSNGQPPIDN-LKLTFDKPHIRLMR-VDTAKAIRRTKCEVKRVKP-  373
PtobSP38      ASSKFIYGSDISVDSESQLWFLSNGQPPIDN-LKLTFDKPHIRLMR-VDTAKAIRRTKCEVKPPKP-  375
PperSP03B     ASAKFIYGSDISVDSESQLWFLSTGHPPIPN-LKLTFDKPHIRLMR-VDTAKAIRRTRCEVKPRKP-  375
PariSP04B     ASSKFIFGTDIQIDSDSQLWFLSNGQPPIDN-LKLTFDKPHIRLMR-VDTKNSIRRTRCEVKPIKKP  375
PorASP2       DGPDFIFGTDISVDSDSKLWFFSNGHPPIEN-LQLTYDKPHIRLIS-MDTEKSIHGTKCEVK-----  369
PorMSP23      DSPDFIFGTDISVDSDSKLWFFSNGHPPIEN-LQLTYDKPHIRLIS-MDTEKSIHGTKCEVK-----  369
PtobSP37      DNPDFIFGTDISVDSESKLWFFSNGHPPIEN-VQLTHDKPHFRLIS-MDTEKSIHGTKCEVKS----  370
PperSP03      HSPDFIFGTDISMDSESKLWFFSNGHPPIEN-VQLTFDKPHFRLIS-MDTKKSIHGTKCEVKP----  370
PariSP04      ASPEFIFGTDISVDSESKLWFFSNGHPPIEN-LQLSSDKPHIHLIS-VDTEKAIRGTKCEVKA----  370
LuloYELLOW    TSSRFVFGTDISVDSKGGLWFMSNGFPPIRKSEKFKYDFPRYRLMRIMDTQEAIAGTACDMNA----  381
                ~    ~~           ~ ~
```

Figure 2. Multiple sequence alignment of the sand fly yellow-related protein family. Multiple sequence alignment of yellow-related salivary proteins from *Phlebotomus ariasi* (Pari), *Phlebotomus perniciosus* (Pper), *Phlebotomus orientalis* Addis Zemen colony (PorA), *P. orientalis* Melka Werer colony (PorM), *Phlebotomus tobbi* (Ptob), and *Lutzomyia longipalpis* (Lulo). Sequence names and the number of amino acids per line are indicated. Identical amino acid residues are highlighted black and similar residues grey. Specific symbols indicate: ~ amino acid motif binding prohemostatic and proinflammatory biogenic amines. The symbols refer to the lines above.
doi:10.1371/journal.pntd.0002709.g002

(PorASP139/KC170955; PorMSP101/KC170979) (Table 1, Table 2). Addis Zemen, as well as Melka Werer, sequences contained the NUC Smart motif, which is typical for DNA/RNA non-specific endonucleases and phosphodiesterases. Predicted molecular mass of both AZ and MW endonucleases was 41.7 kDa and predicted pI was 9.3 and 9.4, respectively. Endonuclease function in sand fly saliva is still unclear; however, properties that facilitate blood acquisition are assumed. Endonucleases were detected in salivary gland cDNA libraries of some sand flies species tested [26,27,31,52,54–56], but also in another bloodsucking Diptera [87,88]. Sequences of AZ and MW *P. orientalis* colony coding for endonuclease revealed 97% identity and furthermore, there was no difference in the numbers and positions of the active sites, Mg^{2+} binding sites, and substrate binding sites (Figure 6). *Phlebotomus perniciosus* salivary endonuclease (ABA43064) was found to be the most relative sequence (92% identity), while homology of *P. orientalis* enzymes with other sand fly endonucleases ranged between 44–80%. Endonuclease was also shown to have antigenic properties; sera of dogs from an endemic area of VL in Italy, as well as dogs experimentally bitten by *P. perniciosus*, reacted with a 41 kDa salivary protein identified as the endonuclease (ABA43064) [63].

Transcripts coding for a putative phospholipase A2 (PLA2) were detected in both *P. orientalis* cDNA libraries (PorASP80/KC170949; PorMSP129/KC170982) (Table 1, Table 2). In AZ and MW sequences, the whole PLA2 domain was present. The predicted molecular mass of PLA2 was 29.7 kDa and pI was 8.4 and 8.3 for AZ and MW, respectively. PLA2 was described as the main allergen in the hymenopteran venom (reviewed in [84]), however, its allergenic effect in sand flies remains to be elucidated. Sequences coding for PLA2 revealed a high degree of conservancy between AZ and MW colonies as well as among various sand fly species. The AZ and MW *P. orientalis* PLA2 were almost identical (99%) and the metal binding sites were present on the same positions (Figure 7). Similarly, the catalytic sites were detected on the same positions and amino acids in both colonies, with the exception of the catalytic site on 215th amino acid, where glycine present in AZ colony was in MW replaced by aspartic acid (Figure 7). Within the *Larroussius* subgenus, the homology of *P. orientalis* PLA2 reached 99% with *P. perniciosus* (ABA43062) and 94% with *P. ariasi* (AAX54852). Moreover, comparing the PLA2 enzymes of *P. orientalis* and *P. arabicus* (ACS93491), subgenus *Adlerius*, showed 88% identity.

A single 3′truncated transcript coding for a putative α-amylase was detected in the salivary gland cDNA library of MW *P. orientalis* colony (PorMSP27/KC170968) (Table 2), but no homologous sequences were found in the AZ colony. Amylase is an enzyme which is likely not involved in the bloodfeeding process, but participates in dietary sugar digestion (reviewed in [89]). Transcripts coding for α-amylase were detected in the salivary gland cDNA libraries of *L. longipalpis* (AAD32192) [53], *P. arabicus* (ACS93490) [31], and in *P. papatasi* (AAD32192) [55]. Alpha-amylase activity was detected in the sand fly salivary glands [53,90,91]. The predicted molecular weight of the *P. orientalis* MW amylase was 33.4 kDa and the predicted pI was, 5.8 (Table 2). Amino acid sequence alignment of *P. orientalis* α-amylase shows an 88% identity with the *P. arabicus* α-amylase (ACS93490) and 82% identity with *L. longipalpis* α-amylase (AAD32192).

A single sequence containing signal peptide, truncated in the 3′region (missing stop codon), coding for a putative salivary pyrophosphatase (PorASP262/KC170959) was ascertained in the AZ cDNA library (Table 1). The predicted molecular mass was 32.9 kDa and the predicted pI was 7.2. Pyrophosphatase was also detected in the MW colony, but these sequences did not contain signal peptides. Nonetheless, the identity of AZ and MW pyrophosphatases reached 99%. Salivary pyrophosphatase was found also in saliva of other sand fly species from the genus *Phlebotomus* such as *P. duboscqi* (ABI20154) [50], *P. argentipes*
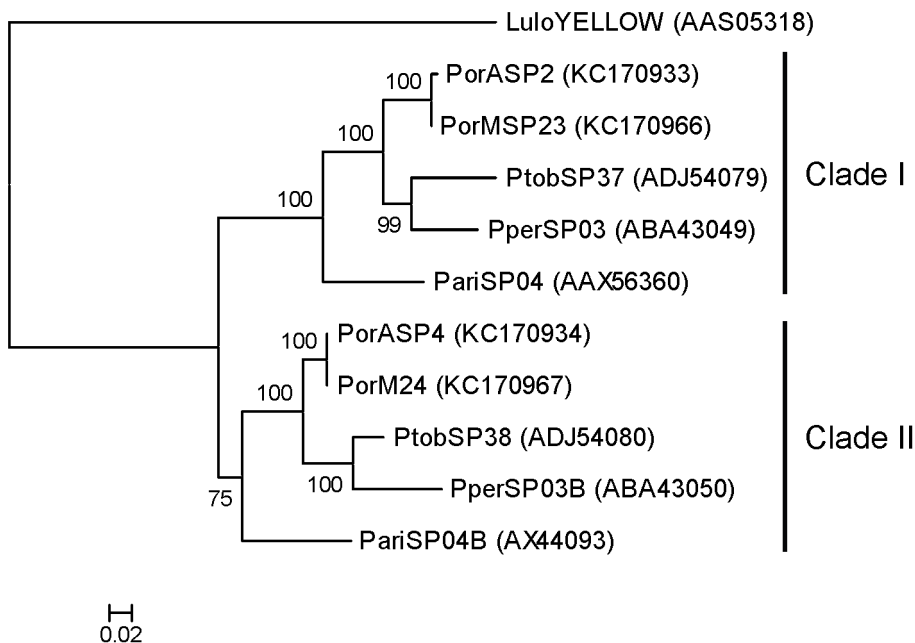
Figure 3. Phylogenetic analysis of the yellow-related family of sand fly salivary proteins. Phylogenetic analysis of yellow-related salivary proteins from *Phlebotomus ariasi* (Pari), *Phlebotomus perniciosus* (Pper), *Phlebotomus orientalis* Addis Zemen colony (PorA), *P. orientalis* Melka Werer colony (PorM), *Phlebotomus tobbi* (Ptob), and *Lutzomyia longipalpis* (Lulo). The JTT model was used for this phylogenic analysis. Sequence names, GenBank accession numbers and branch values are indicated. Yellow-related salivary proteins from *Larroussius* sand fly species are divided into two distinct clades (Clade I, II).
doi:10.1371/journal.pntd.0002709.g003
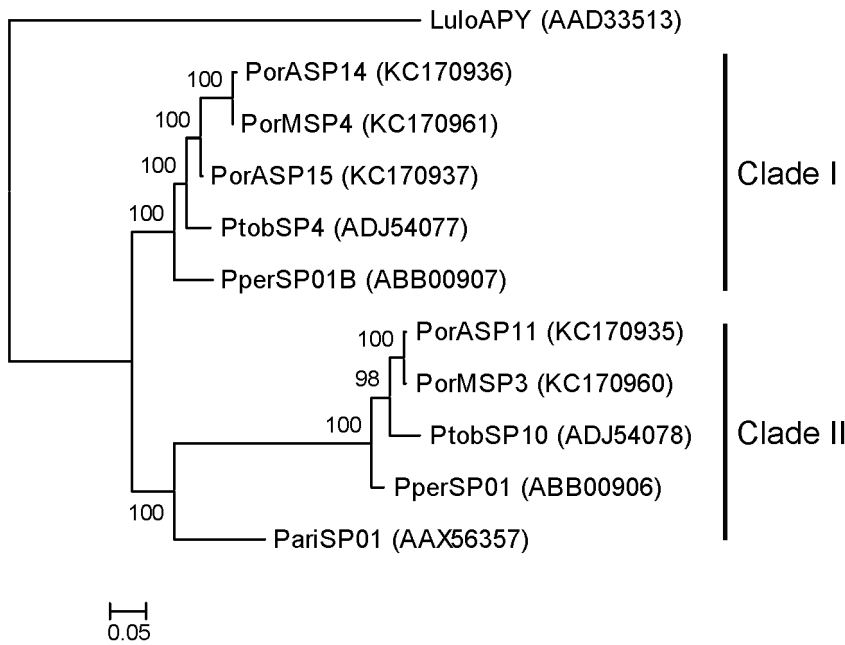
Figure 4. Phylogenetic analysis of the apyrase family of sand fly salivary proteins. Phylogenetic analysis of salivary apyrases from *Phlebotomus ariasi* (Pari), *Phlebotomus perniciosus* (Pper), *Phlebotomus orientalis* Addis Zemen colony (PorA), *P. orientalis* Melka Werer colony (PorM), *Phlebotomus tobbi* (Ptob), and *Lutzomyia longipalpis* (Lulo). The WAG model was used for this phylogenic analysis. Sequence names, GenBank accession numbers and branch values are indicated. Apyrases from *Larroussius* sand fly species are divided into two distinct clades (Clade I, II).
doi:10.1371/journal.pntd.0002709.g004

(ABA12155) [26], and *P. arabicus* (ACS93498) [31]. Transcripts coding for pyrophosphatase did not reveal a high degree of conservancy, as homology of AZ or MW *P. orientalis* enzymes with the aforementioned sequences ranged between 39–74%.

D7-related proteins

D7-related proteins belong to the odorant-binding protein superfamily, which is composed of pheromone-binding proteins (PBP) and general-odorant-binding proteins (GOBP). D7 proteins are commonly present in the salivary glands of various blood-feeding insect such as mosquitoes (e.g. [70]), black flies (e.g. [92]), biting midges (e.g. [93]) and sand flies [7,26–28,31,50,52,54–56,94]. Moreover, proteins belonging to the insect odorant

binding protein family were recently detected in *L. longipalpis* pheromone glands [95].

In the *P. orientalis* cDNA libraries we found three different D7-related proteins in the AZ colony (PorASP46/KC170942, PorASP48/KC170943, PorASP122/KC170954) and four in the MW colony (PorMSP28/KC170969, PorMSP38/KC170970, PorMSP43/KC170971, PorMSP67/KC170973) (Table 1, Table 2). They all had a similar predicted molecular mass (26.7–27.3 kDa) and wide range of pI (6.4–9.2).

The function of sand fly salivary D7-related proteins remains unknown, although it might be similar to mosquito D7 proteins; either as a binder of biogenic amines and/or eicosanoids [96,97] or as an anticoagulant [98,99].

Table 3. Salivary apyrase in two *P. orientalis* colonies originated from Melka Werer (MW) and Addis Zemen (AZ), Ethiopia.

		P. orientalis **AZ**	*P. orientalis* **MW**	*P. papatasi*
Total protein in µg/gland pair		0.61±0.05	0.52±0.06	0.68±0.07
Mean specific apyrase activity* at 37°C, pH 8.5:				
mUnits/pair of glands **	ATPase	87.9±2.9	74.3±3.9	77.9±5.3
	ADPase	99.3±6.7	84.2±6.7	89.80±6.9
Units/mg of total protein	ATPase	144	143	115
	ADPase	163	162	132
ATPase/ADPase ratio		0.88	0.88	0,87
pH optimum		8.5	8.5	nd
Activator cation		Ca^{2+}	Ca^{2+}	Ca^{2+}

Phlebotomus papatasi was used as the control.
*One unit of enzyme activity is defined as the amount of enzyme that releases one micromole of orthophosphate per minute from the nucleotide substrate at specified assay conditions.
**Individual specific activity was calculated per gland pair as *P. papatasi* is characterized by dissimilar size of salivary glands [45].
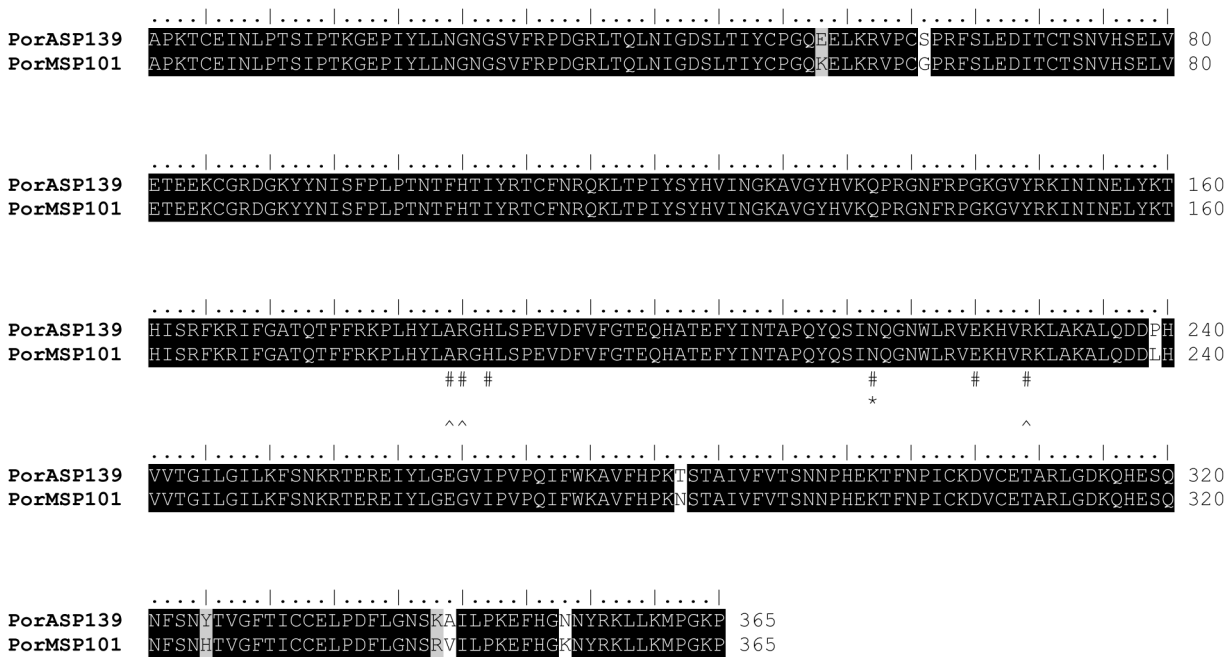doi:10.1371/journal.pntd.0002709.t003

```
              _              _   _            _          _                           _     _
       ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
PoriASP112  ENPEKNFTIYWNVPTDQCNRHNYTNETKPNFPELLTNLSIVHNVNGSFRGEEFRILYSPGLWPSMEHNKTENGTHGGMPH  80
PoriMSP108  ENSEKNFTIYWNVPTDQCNRHNYTNVTKPNFPELLTNLSIVHNVNGSFRGEGFRILYSPGLWPSMEHNKTENGTHEGMPQ  80
                _              _               _          _                           _   _

                  _              _             _          _                           _     _
       ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
PoriASP112  HGNLTEHLEQLEKDINNCSHINYIPNNFTGMAVIDMESWRPVFRQNTGWMEIYRKLVFREIDSNATLRNLVYNDSSGITC  160
PoriMSP108  HGNLTKHLEQLEKDINNCSHINYIPENFTGMAVIDMESWRPVFRQNTGWMKIYRKLVFIEIDKNETLRKLVYNDSSGITC  160
              _              _               _          _

                                                          _               _
       ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
PoriASP112  EKNRTNLCFKEAAKIFEPMAIDYMNKSIAKLRELRPEAHWGYYGFPYCFNIRKDNRSESCAELVQKENDNTTWLFQSYNS  240
PoriMSP108  EKNRTNDCFKVAAKIFEPMAIDYMNKSIAKLRELRSEAHWGYYGFPYCFNIRKNRNESCAELVQKENDNTTWLFQSYSS  240
                                                                                    _

                         _                              _       _
       ....|....|....|....|....|....|....|....|....|....|....|....|..
PoriASP112  WYPSVYISHDNFTEEDRLKLVRGRVQEYNRLRNLTNQNATIYPYIWLLYNLD  292
PoriMSP108  WYPSVYISHDNFTEEDRLKLVRGRVQEYNRLRNLTNKNATIYPYIWLLYNLD  292
                         _                              _
```

Figure 5. Sequence alignment of the *P. orientalis* hyaluronidase. Sequence alignment of salivary hyaluronidase from *Phlebotomus orientalis* Addis Zemen (PorA) and *Phlebotomus orientalis* Melka Werer (PorM) colonies represented by hyaluronidase protein domains. Sequence names and the number of amino acids per line are indicated. Identical amino acid residues are highlighted black and similar residues grey. Specific symbols indicate: - N-glycosylation prediction sites. The symbols above and under the lines refer to the Addis Zemen and Melka Werer sequences, respectively. doi:10.1371/journal.pntd.0002709.g005

Phylogenetic analysis showed that *P. orientalis* D7-related salivary proteins are divided into three clades (Figure 8). Clade I contains two 100% identical *P. orientalis* D7-related proteins, one from each cDNA library (PorASP122, PorMSP67). Clade II contains only one *P. orientalis* protein from the MW colony (PorMSP28), as the analogue from AZ colony did not contain the signal peptide sequence and therefore was not included into the phylogenetic

analysis. Clade III includes two proteins from each library, PorASP46 and PorASP48 from the AZ colony and PorMSP38 and PorMSP43 from the MW colony. *Phlebotomus orientalis* D7-related proteins within clade III form two distinct subclades (PorASP48, PorMSP38 and PorASP46, PorMSP43), where the identity reached 99%. Overall, the identity in the clade III was 92%. Sequences coding for salivary D7-related proteins in *P.*

```
       ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
PorASP139  APKTCEINLPTSIPTKGEPIYLLNGNGSVFRPDGRLTQLNIGDSLTIYCPGQEELKRVPCSPRFSLEDITCTSNVHSELV  80
PorMSP101  APKTCEINLPTSIPTKGEPIYLLNGNGSVFRPDGRLTQLNIGDSLTIYCPGQKELKRVPCGPRFSLEDITCTSNVHSELV  80

       ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
PorASP139  ETEEKCGRDGKYYNISFPLPTNTFHTIYRTCFNRQKLTPIYSYHVINGKAVGYHVKQPRGNFRPGKGVYRKININELYKT  160
PorMSP101  ETEEKCGRDGKYYNISFPLPTNTFHTIYRTCFNRQKLTPIYSYHVINGKAVGYHVKQPRGNFRPGKGVYRKININELYKT  160

       ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
PorASP139  HISRFKRIFGATQTFFRKPLHYLARGHLSPEVDFVFGTEQHATEFYINTAPQYQSINQGNWLRVEKHVRKLAKALQDDPH  240
PorMSP101  HISRFKRIFGATQTFFRKPLHYLARGHLSPEVDFVFGTEQHATEFYINTAPQYQSINQGNWLRVEKHVRKLAKALQDDLH  240
                        ## #                                        #         #    #
                                                                     *
                         ^^                                                   ^
       ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
PorASP139  VVTGILGILKFSNKRTEREIYLGEGVIPVPQIFWKAVFHPKTSTAIVFVTSNNPHEKTFNPICKDVCETARLGDKQHESQ  320
PorMSP101  VVTGILGILKFSNKRTEREIYLGEGVIPVPQIFWKAVFHPKNSTAIVFVTSNNPHEKTFNPICKDVCETARLGDKQHESQ  320

       ....|....|....|....|....|....|....|....|....|....|....|....|
PorASP139  NFSNYTVGFTICCELPDFLGNSKAILPKEFHGNNYRKLLKMPGKP  365
PorMSP101  NFSNHTVGFTICCELPDFLGNSRVILPKEFHGKNYRKLLKMPGKP  365
```

Figure 6. Sequence alignment of the *P. orientalis* endonuclease. Sequence alignment of endonucleases from *Phlebotomus orientalis* Addis Zemen (PorA) and *Phlebotomus orientalis* Melka Werer (PorM) colonies. Sequence names and the number of amino acids per line are indicated. Identical amino acid residues are highlighted black and similar residues grey. Specific symbols indicate: # enzyme active sites, * Mg^{2+} binding site, and ^ substrate binding sites. The specific symbol refers to the sequence line above. doi:10.1371/journal.pntd.0002709.g006

orientalis species did not reveal high degree of conservancy as the alignment of all D7-related proteins from both colonies reached only 31% identity (Figure S2). Comparison with other sand fly species from subgenus *Larroussius* showed that *P. orientalis* D7-related proteins are more related to *P. tobbi* and *P. perniciosus* than to *P. ariasi* (Figures 8 and S2).

Based on the glycosylation prediction servers (NetNGlyc and NetOGlyc), we found out that D7-related proteins have very limited glycosylation sites. Only those D7-related proteins occurring within clade I (PorASP122 and PorMSP67) were predicted to have N-glycosylation, while the others are likely not glycosylated at all. Similarly, only those D7-related proteins of *P. perniciosus*, *P. tobbi*, and *P. ariasi* included in the clade I are likely N-glycosylated, moreover, the glycosylation sites are predicted in all *Larroussius* sand fly species on the same positions. These data strengthens the idea that the proteins from different clades might have different molecular functions even though they are all D7-related proteins. Furthermore, mixtures of glycosylated and non-glycosylated D7-related proteins were previously detected in other sand fly species such as *P. arabicus* or *P. tobbi* [28,31].

D7-related proteins are highly antigenic and were recognized by specific antibodies from the sera of repeatedly bitten hosts, regardless of natural [15,61,63,67] or experimental exposure [5,60,62,63,82]. Recombinant *P. ariasi* D7-related protein (AAX55749) elicited the production of specific humoral response in immunized mice [27]. Anti-*P. papatasi* saliva antibodies reacted with the 30 kDa recombinant *P. papatasi* D7-related protein (AAL11049) [5], but the same protein was not recognized by the human sera from an endemic area of CL in Tunisia [100]. Moreover, recombinant 28 kDa D7-related protein from *P. papatasi* saliva (AAL11048) was not targeted by the specific antibodies of immunized mice [5]. Thus, a broad use of D7-related salivary proteins as the reliable marker of sand fly exposure is not likely. Importantly, no significant cellular immunity was observed in various hosts after the inoculation of DNA plasmids coding for D7-related sand fly salivary proteins [12,27,68].

PpSP15-like proteins

Transcripts coding for PpSP15-like proteins represented the most abundant family in *P. orientalis* cDNA libraries. PpSP15-like proteins were detected in both the AZ (PorASP28/KC170938;

PorASP31/KC170939; PorASP37/KC170940; PorASP61/KC170944; PorASP64/KC170945) and the MW colony (PorMSP12/KC170964; PorMSP74/KC170974; PorMSP75/KC170975; PorMSP90/KC170977; PorMSP96/KC170978) (Table 1, Table 2). The predicted molecular mass ranged from 13.9 to 14.9 and the isoelectric point was slightly basic (8.0–9.2).

Phylogenetic analysis showed that *P. orientalis* PpSP15-like proteins are divided into three clades. Clade I contains two *P. orientalis* PpSP15-like proteins, one from each library (PorASP37, PorMSP12), which have an identity of 100%. Clades II and III each contain four *P. orientalis* proteins, two from each library (clade II: PorASP61, PorASP64, PorMSP74, PorMSP75; clade III: PorASP28, PorASP31, PorMSP90, PorMSP96) (Figure 9). Alignment of known *Larroussius* and *P. orientalis* PpSP15-like proteins revealed high degree of divergence (overall identity 24%) (Figure S3). Comparison of *P. orientalis* and other *Larroussius* species PpSP15-like proteins within each clade showed identity ranging from 61 to 96%. Our results comply well with previous reports [26,28,31,56], where PpSP15-like proteins of various sand fly species were described as extremely variable proteins, likely occurring in multiple gene copies [101].

PpSP15-like proteins belong to the odorant-binding protein family but, so far, the exact function of these proteins in sand flies remains unknown. However, SP15 protein from *P. papatasi* saliva (AAL11047) was shown to elicit specific humoral and cellular immunity, which resulted in the protection of immunized mice against *Leishmania major* infection [7,9]. Similarly, a DTH reaction was also observed in mice immunized by the inoculation of a *P. ariasi* DNA plasmid coding for SP15-like salivary protein (AAX56359) [27]. On the other hand, DNA plasmids coding for *L. longipalpis* SL1 protein (AAD32197) failed to promote the cellular immunity in experimental mice [12], hamsters [8], and dogs [68]. Glycosylation prediction servers (NetNGlyc and NetOGlyc) revealed that *P. orientalis* PpSP15-like proteins are likely not glycosylated.

Antigen 5-related proteins

Antigen 5-related proteins (Ag5r) belong to the CAP family of proteins which is composed of **C**ysteine-rich secretory proteins, **A**ntigen 5, and **P**athogenesis-related 1 proteins. Proteins with the CAP domain are commonly present in various organisms that
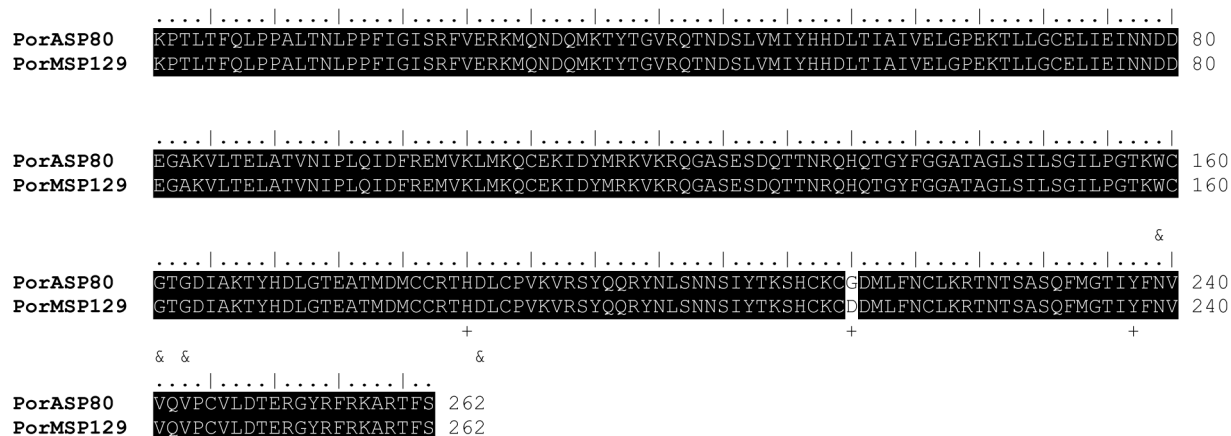
Figure 7. Sequence alignment of the *P. orientalis* phospholipase. Sequence alignment of phospholipases from *Phlebotomus orientalis* Addis Zemen (PorA) and *Phlebotomus orientalis* Melka Werer (PorM) colonies. Sequence names and the number of amino acids per line are indicated. Identical amino acid residues are highlighted black. Specific symbols indicate: + catalytic sites, & metal binding sites. The specific symbol refers to the sequence line above.
doi:10.1371/journal.pntd.0002709.g007
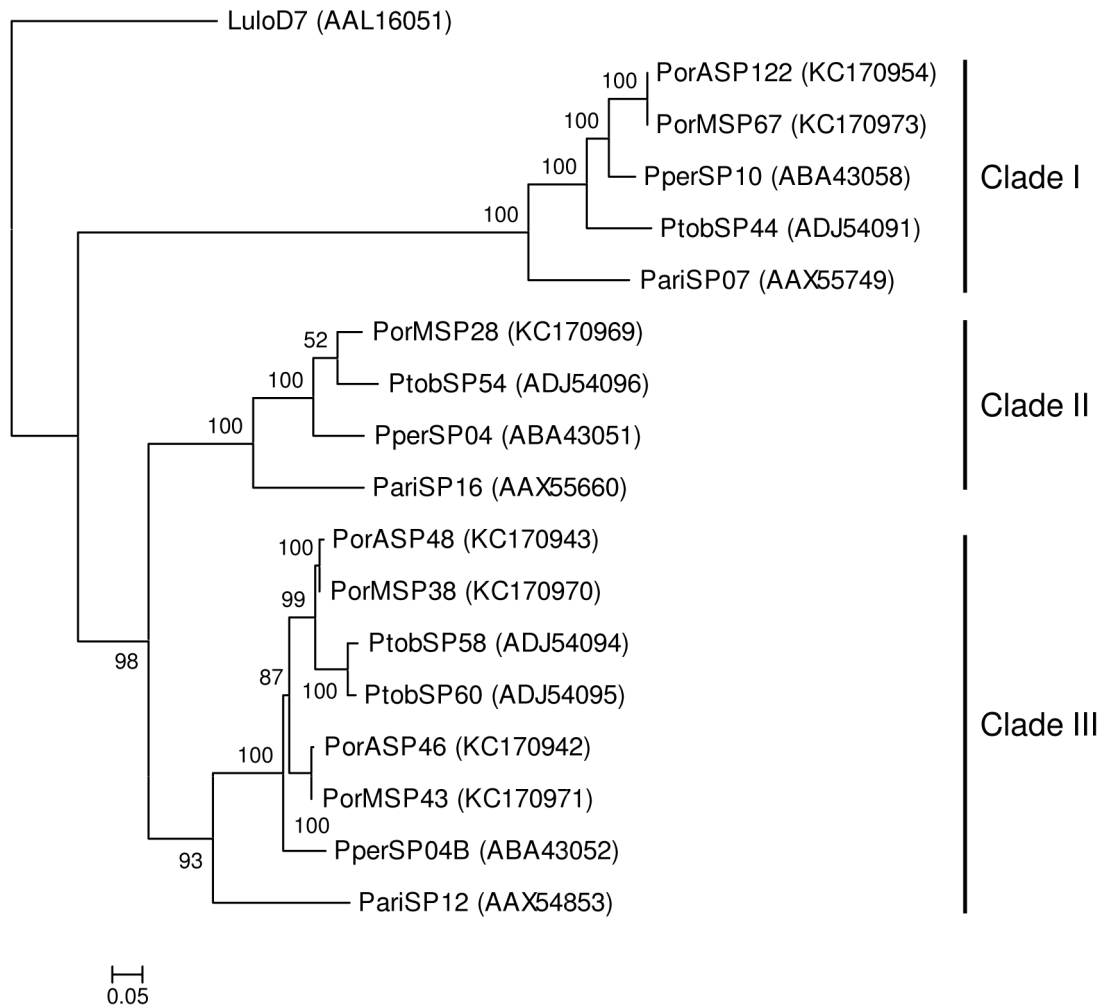
Figure 8. Phylogenetic analysis of the D7-related family of sand fly salivary proteins. Phylogenetic analysis of D7-related salivary proteins from *Phlebotomus ariasi* (Pari), *Phlebotomus perniciosus* (Pper), *Phlebotomus orientalis* Addis Zemen colony (PorA), *P. orientalis* Melka Werer colony (PorM), *Phlebotomus tobbi* (Ptob), and *Lutzomyia longipalpis* (Lulo). The WAG model was used for this phylogenic analysis. Sequence names, GenBank accession numbers and branch values are indicated. D7-related proteins from *Larroussius* sand fly species are divided into three distinct clades (Clade I–III).

doi:10.1371/journal.pntd.0002709.g008

include prokaryotes and non-vertebrate eukaryotes [102,103]. Ag5r proteins were described from the venom of ants, wasps and other Hymenoptera [104–106], but were also found in salivary glands of various bloodsucking insects, including sand flies [7,26–28,31,50,52–56]. The exact function of Ag5r in sand flies is still unknown although biological properties of other proteins from the same family may give us some clue. The X-ray structure of NA-ASP-2 protein (pathogenesis –related 1 protein) from the human hookworm, *Necator americanus*, reveals structural and charge similarities to chemokines, suggesting that these proteins could potentially modulate the host immune response [107]; more recently, a triatomine salivary member of the family was shown to have superoxide dismutase activity and to exert anti-neutrophil activity [108].

Sequences coding for salivary Ag5r proteins were found in cDNA library from the AZ (PorASP74/KC170947; PorASP76/KC170948) and the MW (PorMSP6/KC170962; PorMSP8/KC170963) *P. orientalis* colonies (Table 1, Table 2). The predicted molecular weight was 28.8 kDa and pI was slightly basic (8.9). Phylogenetic analysis showed that Ag5r proteins from the saliva of

sand fly species from the subgenus *Larroussius* are separated into two clades (Figure 10). The first clade contains only Ag5r protein from *P. ariasi* (AAX44092), whereas the second clade includes proteins of *P. tobbi* (ADJ54082, ADJ54083), *P. perniciosus* (ABA43055), and *P. orientalis* (Figure 10). *Phlebotomus orientalis* Ag5r proteins are represented by four salivary transcripts; two from each colony (PorASP74, PorASP76, and PorMSP6, PorMSP8). Phylogenetic analysis assembled *P. orientalis* Ag5r proteins into two subclades, with an identity of 100%, the first one represented by PorASP74 and PorMSP8, the later one by PorASP76 and PorMSP6. The homology among the Ag5r proteins from different subclades reached 99% as these sequences differed in those amino acids on position 47 and 50 (Figure S4). Sequences from *P. perniciosus* (ABA43055) and *P. tobbi* (ADJ54082, ADJ54083) coding for Ag5r proteins were ascertained to be the closest relatives with the identity of 92% and 88 to 93%, respectively, while identity with *P. ariasi* protein (AAX44092) reached only 77% (Figure S4).

Antigenic properties of Ag5r proteins were demonstrated in various sand fly - host combinations. Salivary Ag5r proteins were
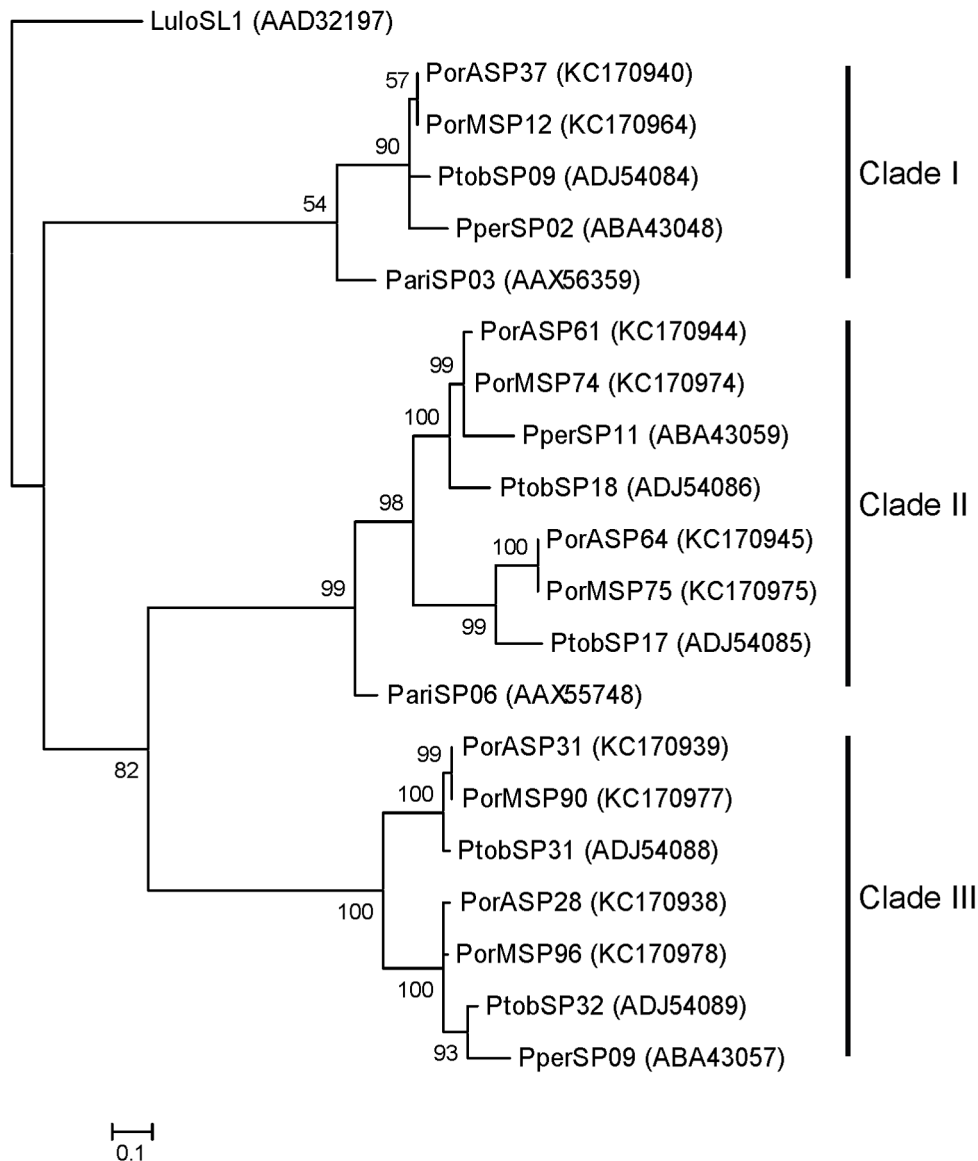
Figure 9. Phylogenetic analysis of the PpSP15-like family of sand fly salivary proteins. Phylogenetic analysis of PpSP15-like salivary proteins from *Phlebotomus ariasi* (Pari), *Phlebotomus perniciosus* (Pper), *Phlebotomus orientalis* Addis Zemen colony (PorA), *P. orientalis* Melka Werer colony (PorM), *Phlebotomus tobbi* (Ptob), and *Lutzomyia longipalpis* (Lulo). The JTT model was used for this phylogenic analysis. Sequence names, GenBank accession numbers and branch values are indicated. PpSP15-like proteins from *Larroussius* sand fly species are divided into three distinct clades (Clade I–III).
doi:10.1371/journal.pntd.0002709.g009

recognized by sera of mice repeatedly bitten by *P. papatasi* [5] or *P. arabicus* [31], by sera of dogs bitten by *P. perniciosus* [63], as well as by sera of hamsters exposed to *P. tobbi* [28] or *P. argentipes* [82]. On the other hand, inoculation of DNA plasmids coding for Ag5r protein from saliva of *P. ariasi* (AAX44092) or *L. longipalpis* (AAD32191) did not elicit a specific humoral response but did induce a cell-mediated immune response [12,27]. Glycosylation prediction servers (NetNGlyc and NetOGlyc) showed that all *P. orientalis* Ag5r proteins are N- and O-glycosylated.

PpSP32-like proteins

The PpSP32-like protein family was described for the first time in the saliva of *P. papatasi* [7]. These proteins occur solely in sand fly saliva and their exact function is unknown. PpSP32-like proteins were found in the transcriptomes of various sand flies

[7,26–28,31,50,54–56] and sequences coding for these proteins were also found in both *P. orientalis* cDNA libraries (PorASP86/KC170950; PorMSP15/KC170964) (Table 1, Table 2). The predicted molecular mass was 25 kDa and the pI was very basic (10.1–10.2). PpSP32-like proteins of AZ and MW colony revealed high degree of identity (98%); high identity was also obtained by comparing *P. orientalis* with other *Larroussius* sand fly species; 85–87% with *P. perniciosus* (ABA43053) and 81–83% with *P. tobbi* (ADJ54102). Glycosylation prediction servers (NetNGlyc and NetOGlyc) showed a high degree of glycosylation of *P.orientalis* PpSP32-like proteins, which could be potentially responsible for their immunogenicity. Sera of mice experimentally bitten by *P. papatasi* recognized *P. papatasi* SP32 protein [5] and human sera from endemic area of CL in Tunisia reacted preferentially with recombinant PpSP32 (AAL11050) prepared in mammalian
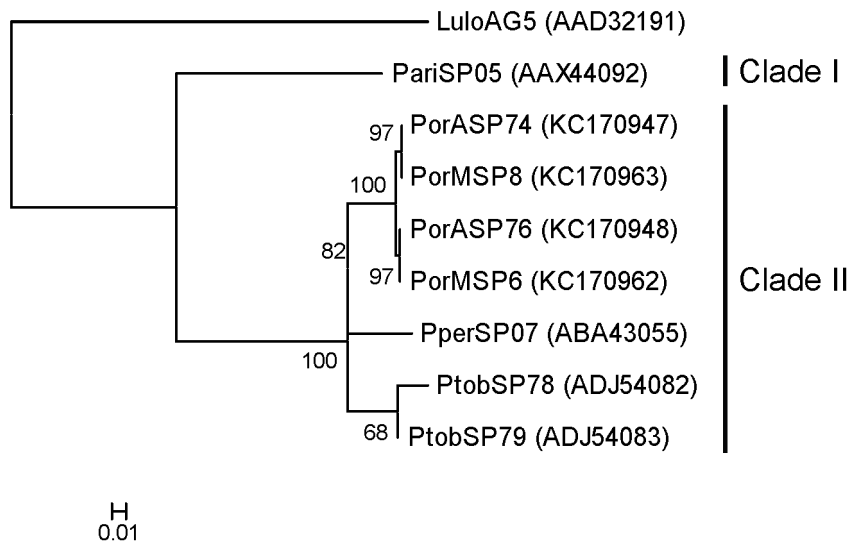
Figure 10. Phylogenetic analysis of the antigen 5-related family of sand fly salivary proteins. Phylogenetic analysis of antigen 5-related salivary proteins from *Phlebotomus ariasi* (Pari), *Phlebotomus perniciosus* (Pper), *Phlebotomus orientalis* Addis Zemen colony (PorA), *P. orientalis* Melka Werer colony (PorM), *Phlebotomus tobbi* (Ptob), and *Lutzomyia longipalpis* (Lulo). The Dayhoff model was used for this phylogenic analysis. Sequence names, GenBank accession numbers and branch values are indicated. Antigen 5-related proteins from *Larroussius* sand fly species are divided into two distinct clades (Clade I, II).
doi:10.1371/journal.pntd.0002709.g010

expressing system [100]. On the other hand, bacterially-expressed recombinant PpSP32-like protein from *L. longipalpis* (AAS16906) did not elicit either specific humoral or cellular response [12].

ParSP25-like proteins

Transcripts coding for ParSP25-like proteins were identified in the cDNA library from the AZ (PorASP106/KC170953) and the MW (PorMSP65/KC170972) *P. orientalis* colony (Table 1, Table 2). The predicted molecular mass was 27.6 kDa and, due to the high proportion of acidic residues present in the amino acid sequences, the pI was very acidic (4.7–4.8). ParSP25-like proteins were detected in the saliva of sand flies from the subgenus *Larroussius* (*P. ariasi*, *P. perniciosus*, *P. tobbi*), *Adlerius* (*P. arabicus*), and *Phlebotomus* (*P. papatasi*) [26–28,31,55]. ParSP25-like proteins have not yet been found in New World sand fly species [52–56]. The ParSP25-like proteins of AZ and MW colonies are almost identical (98%). Homology of *P. orientalis* proteins with other *Larroussius* species reached 85–86% for *P. tobbi* (ADJ54100), followed by 73–74% for *P. perniciosus* (ABA43056) and 64% for *P. ariasi* (AAX55664). Although the exact function of these proteins remains unknown, some ParSP25-like proteins were demonstrated to be immunogenic. Sera from dogs, hamsters and mice bitten by *P. perniciosus* reacted with salivary protein identified as the member of ParSP25-like family [60,63]. Similarly to other sand fly species [28], ParSP15-like proteins of *P. orientalis* are not predicted to be glycosylated.

Lufaxin-like proteins

A 32.4 kDa protein from *L. longipalpis* saliva belongs to a novel family of slow-tight factor Xa inhibitors, displays anti-thrombotic and anti-inflammatory activities, and is named Lufaxin (<u>Lu</u>tzomyia longipalpis <u>F</u>actor <u>Xa</u> inhibitor) [109]. Members of the Lufaxin family were detected in saliva of various sand flies [26–28,31,50,52,54–56], but not in other bloodsucking insects. Sequences coding for a Lufaxin-like protein, were detected in the cDNA library of MW *P. orientalis* colony (PorMSP78/KC170976) (Table 2). Transcripts similar to Lufaxin were also found in AZ colony, but these sequences had low quality scores. The predicted molecular mass of MW Lufaxin-like protein was 18.8 kDa, the pI was 8.4. *Phlebotomus*

orientalis Lufaxin-like protein was found to be highly homologous with *P. perniciosus* (ABA43054) (88% identity) and *P. tobbi* (ADJ54104) (87% identity) Lufaxin-like proteins. According to the glycosylation prediction servers (NetNGlyc and NetOGlyc) *P. orientalis* Lufaxin-like protein is N-glycosylated.

Lufaxin was previously shown to have antigenic properties. Sera of repeatedly bitten dogs recognized Lufaxin and the Lufaxin homologue from *P. perniciosus* [4,63]. Similarly, sera of hamsters experimentally bitten by *P. argentipes* reacted with Lufaxin-like salivary protein [82]. Recombinant Lufaxin (AAS05319) was also demonstrated to promote strong cellular immunity [12,68] and therefore was suggested as the promising candidate for vaccine against canine leishmaniasis [68].

Other putative salivary proteins

Several other putative salivary proteins were found in both cDNA libraries from *P. orientalis* saliva. Transcripts encoding a 16 kDa salivary protein, with a pI of 5.0 and unknown function, were found in the AZ (PorASP150/KC170956) and MW (PorMSP162/KC170983) colonies (Table 1, Table 2). PorASP150 and PorMSP162 are closely related to 16 kDa salivary protein A (ACS93506) and protein B (ACS93507) from *P. arabicus* saliva. A high degree of homology was also found with salivary proteins from *P. argentipes* (ABA12153) and *P. sergenti* (ADJ54127). A related protein was recently identified in saliva of *P. papatasi* (ADJ54127). *Phlebotomus orientalis* 16 kDa proteins are likely not glycosylated.

Three clusters, encoding small salivary proteins with unknown function, were identified in each *P. orientalis* cDNA library: 3.9 kDa protein (PorASP40/KC170941; PorMSP169/KC170984), 4.9 kDa protein (PorASP68/KC170946; PorMSP196/KC170985), and 5.6 kDa protein (PorASP98/KC170952; PorMSP104/KC170980) (Table 1, Table 2). The proteins had small predicted molecular mass (3.9–5.6 kDa) and basic pI (9.2–11.0). *Phlebotomus orientalis* 3.9 kDa protein and 4.9 kDa protein were found to be closely related to the 3.7 kDa (ADJ54106) and 5 kDa (ADJ54105) *P. tobbi* proteins, respectively. Transcripts coding for a 5.6 kDa *P. orientalis* proteins share predicted sequence homology with the 4.5 kDa protein of *P. tobbi* (ADJ54097), 7 kDa protein of *P. perniciosus* (ABA43060), and the

5 kDa protein of *P. ariasi* (AAX55658). Based on the glycosylation prediction servers (NetNGlyc and NetOGlyc) we found that all *P. orientalis* small salivary proteins are likely O-glycosylated.

Antigens and glycoproteins

To identify the salivary antigens in both *P. orientalis* colonies and the degree of cross-reactivity between them, electrophoretically separated salivary proteins of each colony were incubated with sera from mice experimentally bitten by either the AZ or MW colony. By comparing the western blot analysis with the *P. orientalis* proteomes (Figure 1), we predict that the most intensive reactions detected the yellow-related proteins (AZ: PorASP2, PorASP4; MW: PorMSP23, PorMSP24), apyrases (AZ: PorASP11, PorASP14, PorASP15; MW: PorMSP3, PorMSP4), and antigen 5-related proteins (AZ: PorASP74, PorASP76; MW: PorMSP6, PorMSP8). All these proteins were recognized by all AZ and MW mice sera tested, while D7-related proteins (AZ: PorASP48, PorASP122; MW: PorMSP28, PorMSP38, PorMSP67) and PpSP15-like proteins (AZ: PorASP28, PorASP37, PorASP61; MW: PorMSP12, PorMSP74, PorMSP96) were recognized only by some sera (Figure 11). Strong cross-reactivity was detected between AZ and MW *P. orientalis* colonies. The small differences in the intensity of reaction or the number of recognized protein bands were probably caused by the individual variability between mice. These data suggest that the salivary proteins in both colonies share similar antibody epitopes.

Due to the near predicted amino acid sequences, we chose only MW colony to study the level of glycosylation of *P. orientalis* saliva. Separated MW salivary proteins were incubated with biotinylated lectins (DBA, SBA, UEA-I, LTA, ConA, PSA) to detect mainly the N- and O-glycosylation sites. To control the specificity of the reactions each lectin was preincubated with the appropriate saccharide inhibitor. The specific reaction was observed only with ConA, the other lectins did not bind specifically or they possessed higher affinity for the glycoprotein, than for the saccharide inhibitor. We detected the specific binding of ConA to the protein bands corresponding to the 42 kDa yellow-related protein (PorMSP24), 36 kDa hyaluronidase (PorMSP108), 33 kDa salivary apyrase (PorMSP4), 29 kDa antigen 5-related salivary proteins (PorMSP6, PorMSP8), and 27 kDa D7-related salivary protein (PorMSP67), suggesting that these proteins are N-glycosylated (Figure 12). In accordance with the NetNGlyc glycosylation prediction server, the strongest reaction was detected with salivary hyaluronidase and yellow-related protein indicating that these proteins are the most glycosylated.

Conclusions

The parasites from the *Le. donovani* complex can cause lethal VL with approximately 60 000 new cases per year [110]. Therefore, it is crucial to continue the search for the salivary proteins in relevant vector species in order to find suitable candidates of anti-*Leishmania* vaccines or markers of host exposure to sand flies. Our study provides the first detailed description of the salivary proteins of *P. orientalis*, the most important vector of VL in Northeast Africa. We made a broader comparison of the salivary gland transcriptomes, proteomes, and enzymatic activities of salivary hyaluronidase and apyrase of two laboratory reared *P. orientalis* colonies originating from an endemic focus of VL, Addis Zemen, and from a non-endemic area, Melka Werer, Ethiopia.

We revealed a high degree of homology between the AZ and MW transcripts with the overall identity of the appropriate sequences ranging from 94 to 100%. As the mitochondrial genes Cyt b and CO-I, commonly used for the molecular identification of species, reached 100% identity in these *P. orientalis* colonies [24],

we assume that the slight differences in both cDNA libraries are due to the faster evolution in the genes coding for the salivary proteins. Thus, we do not consider the differences in AZ and MW colony as significant ones. Moreover, the absence of some transcripts in any of the cDNA libraries could be likely caused either by the low quality of some sequences or by the low occurrence of the transcripts in the number of randomly sequenced phages. Importantly, the equivalence of compounds and properties in AZ and MW *P. orientalis* salivary glands was also supported by the equal proteomes and enzymatic activities as well as by the powerful antigenic cross-reactivity.

Our data suggests that the composition of the salivary glands is likely not responsible for the different epidemiology of leishmaniases caused by *Le. donovani* observed in Addis Zemen and Melka Werer, although we are aware that we did not quantitatively compared the expression of various salivary proteins. Furthermore, recent study showed that also the susceptibility of both colonies to *Le. donovani* infection is identical [24]. Therefore, we can assume that there are likely other factors affecting the circulation of *Leishmania* parasites causing VL in these foci. In East Africa, the transmission and the life cycle of *Le. donovani* is not fully understood and several wild animals are suspected of being zoonotic reservoir hosts [20]. Thus, we can not exclude the possibility that the presence of putative reservoir hosts in Addis Zemen and their absence from Melka Werer may explain the different epidemiology.

Our study expanded the knowledge of the salivary proteins of sand fly species from the subgenus *Larroussius* and confirmed that *P. orientalis* is closely related to *P. tobbi* and *P. perniciosus*, two vectors causing *Le. infantum* derived CL and VL, respectively [111]. On the other hand, phylogenetic analysis determined *P. ariasi*, an important vector of visceral *Le. infantum* infection, as the
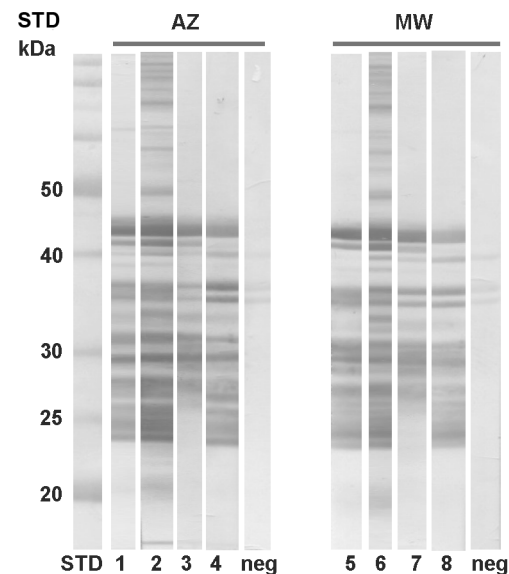
Figure 11. Humoral response to salivary gland antigens of Addis Zemen and Melka Werer *Phlebotomus orientalis* **colony.** Salivary proteins of Addis Zemen (AZ) and Melka Werer (MW) *P. orientalis* colony were separated under non-reducing conditions by SDS-PAGE electrophoresis. Western blot analysis was performed by two different sera of BALB/c mice experimentally bitten by AZ (the same sera used in the lanes 1, 5, and 2, 6) and two sera of mice bitten by MW (the same sera used in the lanes 3, 7, and 4, 8) colony. Serum from a naive mouse was used as the negative control (Neg). Molecular weight standard (STD), stained by amido black, labeled with the corresponding molecular weights (kDa).
doi:10.1371/journal.pntd.0002709.g011
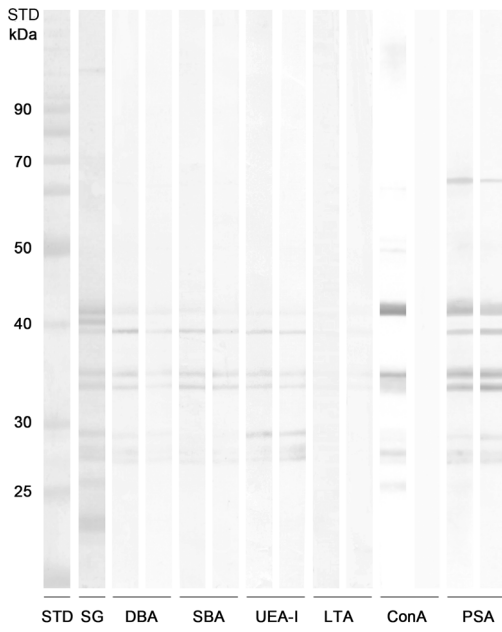
Figure 12. *Phlebotomus orientalis* salivary gland glycoproteins. Salivary proteins of Melka Werer *P. orientalis* colony (SG) were separated under non-reducing conditions by SDS-PAGE electrophoresis and incubated with biotinylated lectin from *Dolichos biflorus* (DBA), *Glycine max* (SBA), *Ulex europaeus* (UEA-I), *Tetragonolobus purpureas* (LTA), *Canavalia ensiformis* (ConA), and *Pisum sativum* (PSA). Doublets were used to test reactivity of each lectin; the first line represents the reaction of lectin with SG, in the later one the lectins were pre-incubated with the appropriate saccharide inhibitors to prove the specificity of reaction. Molecular weight standard (STD), stained by amido black, labeled with the corresponding molecular weights (kDa).
doi:10.1371/journal.pntd.0002709.g012

evolutionarily more distinct species. Importantly, a similar relationship of *Larroussius* sand fly species was also achieved in previously published studies showing that *P. ariasi* is a more phylogenetically distinct member of the subgenus using the nuclear and mitochondrial genes ITS2, EF-α, or Cyt b [112,113].

Overall, *P. orientalis* salivary proteins identified by transcriptome and proteome analysis can be further tested in order to explore their biological and pharmacological properties and to find out whether these salivary proteins could, in the recombinant form, be the suitable vaccine candidates. The identification of the antigenic properties of salivary proteins in several sand fly species would also indicate the feasibility of cross-protection between closely related and more distant sand fly species as promisingly demonstrated by [11,12]. Furthermore, the humoral immune response elicited by the powerful salivary antigens would allow us to predict the intensity of exposure to sand fly bites [5,15,62,63,114] and, consequently, to estimate the risk of *Leishmania* transmission in hosts bitten by sand flies in endemic areas [15,63,65,67,115–117].

Supporting Information

Figure S1 Multiple sequence alignment of the sand fly apyrase protein family. Multiple sequence alignment of salivary apyrases from *Phlebotomus ariasi* (Pari), *Phlebotomus perniciosus* (Pper), *Phlebotomus orientalis* Addis Zemen colony (PorA), *P. orientalis* Melka Werer colony (PorM), *Phlebotomus tobbi* (Ptob), and *Lutzomyia longipalpis* (Lulo) was performed using Clustal X 2.0. Sequence names and the number of amino acids per line are indicated. Identical amino acid residues are highlighted black and similar residues grey.
(TIF)

Figure S2 Multiple sequence alignment of the sand fly D7-related protein family. Multiple sequence alignment of D7-related salivary proteins from *Phlebotomus ariasi* (Pari), *Phlebotomus perniciosus* (Pper), *Phlebotomus orientalis* Addis Zemen colony (PorA), *P. orientalis* Melka Werer colony (PorM), *Phlebotomus tobbi* (Ptob), and *Lutzomyia longipalpis* (Lulo) was performed using Clustal X 2.0. Sequence names and the number of amino acids per line are indicated. Identical amino acid residues are highlighted black and similar residues grey.
(TIF)

Figure S3 Multiple sequence alignment of the sand fly PpSP15-like protein family. Multiple sequence alignment of PpSP15-like salivary proteins from *Phlebotomus ariasi* (Pari), *Phlebotomus perniciosus* (Pper), *Phlebotomus orientalis* Addis Zemen colony (PorA), *P. orientalis* Melka Werer colony (PorM), *Phlebotomus tobbi* (Ptob), and *Lutzomyia longipalpis* (Lulo) was performed using Clustal X 2.0. Sequence names and the number of amino acids per line are indicated. Identical amino acid residues are highlighted black and similar residues grey.
(TIF)

Figure S4 Multiple sequence alignment of the sand fly antigen 5-related protein family. Multiple sequence alignment of antigen 5-related salivary proteins from *Phlebotomus ariasi* (Pari), *Phlebotomus perniciosus* (Pper), *Phlebotomus orientalis* Addis Zemen colony (PorA), *P. orientalis* Melka Werer colony (PorM), *Phlebotomus tobbi* (Ptob), and *Lutzomyia longipalpis* (Lulo) was performed using Clustal X 2.0. Sequence names and the number of amino acids per line are indicated. Identical amino acid residues are highlighted black and similar residues grey.
(TIF)

Author Contributions

Conceived and designed the experiments: JGV RCJ PV. Performed the experiments: MV MS TK PS VV ELJ KDB RCJ. Analyzed the data: MV MS TK VV JMCR. Contributed reagents/materials/analysis tools: MS IR TGM AH AW JGV RCJ PV. Wrote the paper: MV MS.

References

1. Fontaine A, Diouf I, Bakkali N, Misse D, Pages F, et al. (2011). Implication of haematophagous arthropod salivary proteins in host-vector interactions. Parasit Vectors 4: 187.

2. Kamhawi S, Belkaid Y, Modi G, Roeton E, Sacks D (2000). Protection against cutaneous leishmaniasis resulting from bites of uninfected sand flies. Science 290: 1351–1354.

3. Souza AP, Andrade BB, Aquino D, Entringer P, Miranda JC, et al. (2010). Using recombinant proteins from *Lutzomyia longipalpis* saliva to estimate human vector exposure in visceral leishmaniasis endemic areas. PLoS Negl Trop Dis 4: e649.

4. Teixeira C, Gomes R, Collin N, Reynoso D, Jochim R, et al. (2010). Discovery of markers of exposure specific to bites of *Lutzomyia longipalpis*, the vector of *Leishmania infantum chagasi* in Latin America. PLoS Negl Trop Dis 4: e638.

5. Vlkova M, Rohousova I, Hostomska J, Pohankova L, Zidkova L, et al. (2012). Kinetics of antibody response in BALB/c and C57BL/6 mice bitten by *Phlebotomus papatasi*. PLoS Negl Trop Dis 6: e1719.

6. Morris RV, Shoemaker CB, David JR, Lanzaro GC, Titus RG (2001). Sand fly maxadilan exacerbates infection with *Leishmania major* and vaccinating against it protects against *L. major* infection. J Immunol 167: 5226–5230.

7. Valenzuela JG, Belkaid Y, Garfield MK, Mendez S, Kamhawi S, et al. (2001). Toward a defined anti-*Leishmania* vaccine targeting vector antigens: characterization of a protective salivary protein. J Exp Med 194: 331–342.

8. Gomes R, Teixeira C, Teixeira MJ, Oliveira F, Menezes MJ, et al. (2008). Immunity to a salivary protein of a sand fly vector protects against the fatal outcome of visceral leishmaniasis in a hamster model. Proc Natl Acad Sci U S A 105: 7845–7850.

9. Oliveira F, Lawyer PG, Kamhawi S, Valenzuela JG (2008). Immunity to distinct sand fly salivary proteins primes the anti-*Leishmania* immune response towards protection or exacerbation of disease. PLOS Negl Trop Dis 2: e226.

10. da Silva RA, Tavares NM, Costa D, Pitombo M, Barbosa L, et al. (2011). DNA vaccination with KMP11 and *Lutzomyia longipalpis* salivary protein protects hamsters against visceral leishmaniasis. Acta Trop 120: 185–190.

11. Tavares NM, Silva RA, Costa DJ, Pitombo MA, Fukutani KF, et al. (2011). *Lutzomyia longipalpis* saliva or salivary protein LJM19 protects against *Leishmania braziliensis* and the saliva of its vector, *Lutzomyia intermedia*. PLoS Negl Trop Dis 5: e1169.

12. Xu X, Oliveira F, Chang BW, Collin N, Gomes R, et al. (2011). Structure and function of a "yellow" protein from saliva of the sand fly *Lutzomyia longipalpis* that confers protective immunity against *Leishmania major* infection. J Biol Chem 286: 32383–32393.

13. Gomes R, Oliveira F, Teixeira C, Meneses C, Gilmore DC, et al. (2012). Immunity to sand fly salivary protein LJM11 modulates host response to vector-transmitted *Leishmania* conferring ulcer-free protection. J Invest Dermatol 132: 2735–2743.

14. Volf P, Rohousova I (2001). Species-specific antigens in salivary glands of phlebotomine sandflies. Parasitology 122: 37–41.

15. Rohousova I, Ozensoy S, Ozbel Y, Volf P (2005). Detection of species-specific antibody response of humans and mice bitten by sand flies. Parasitology 130: 493–499.

16. Thiakaki M, Rohousova I, Volfova V, Volf P, Chang KP, et al. (2005). Sand fly specificity of saliva-mediated protective immunity in *Leishmania amazonensis*-BALB/c mouse model. Microbes Infect 7: 760–766.

17. Wahba M, Riera C (2006). Salivary gland composition of some Old World vector sand fly. J Egypt Soc Parasitol 36: 289–296.

18. Drahota J, Lipoldova M, Volf P, Rohousova I (2009). Specificity of anti-saliva immune response in mice repeatedly bitten by *Phlebotomus sergenti*. Parasite Immunol 31: 766–770.

19. Rohousova I, Volfova V, Nova S, Volf P (2012a). Individual variability of salivary gland proteins in three *Phlebotomus* species. Acta Trop 122: 80–86.

20. Elnaiem DE (2011). Ecology and control of the sand fly vectors of *Leishmania donovani* in East Africa, with special emphasis on *Phlebotomus orientalis*. J Vector Ecol 36: S23–31.

21. Doha SA, Samy AM (2010). Bionomics of phlebotomine sand flies (Diptera: Psychodidae) in the province of Al-Baha, Saudi Arabia. Mem Inst Oswaldo Cruz 105: 850–856.

22. Daoud W, Rioux JA, Delabre-Belmonte A, Dereure J, Rageh HA (1989). Eco-epidemiology of visceral and cutaneous leishmaniasis in the Arab Republic of Yemen. III. Inventory and dynamics of *Phlebotomus*. Bull Soc Pathol Exot Filiales 82: 669–677.

23. Herrero M, Orfanos G, Argaw D, Mulugeta A, Aparicio P, et al. (2009). Natural history of a visceral leishmaniasis outbreak in highland Ethiopia. Am J Trop Med Hyg 81: 373–377.

24. Seblova V, Volfova V, Dvorak V, Pruzinova K, Votypka J, et al. (2013). *Phlebotomus orientalis* sand flies from two geographically distant Ethiopian localities: biology, genetic analyses and susceptibility to *Leishmania donovani*. PLoS Negl Trop Dis 7: e2187.

25. Warburg A, Saraiva E, Lanzaro GC, Titus RG, Neva F (1994). Saliva of *Lutzomyia longipalpis* sibling species differs in its composition and capacity to enhance leishmaniasis. Philos Trans R Soc Lond B Biol Sci 345: 223–230.

26. Anderson JM, Oliveira F, Kamhawi S, Mans BJ, Reynoso D, et al. (2006). Comparative salivary gland transcriptomics of sandfly vectors of visceral leishmaniasis. BMC Genomics 7: a.n. 52.

27. Oliveira F, Kamhawi S, Seitz AE, Pham VM, Guigal PM, et al. (2006). From transcriptome to immunome: Identification of DTH inducing proteins from a *Phlebotomus ariasi* salivary gland cDNA library. Vaccine 24: 374–390.

28. Rohousova I, Subrahmanyam S, Volfova V, Mu J, Volf P, et al. (2012b). Salivary gland transcriptomes and proteomes of *Phlebotomus tobbi* and *Phlebotomus sergenti*, vectors of leishmaniasis. PLoS Negl Trop Dis 6: e1660.

29. Volf P, Volfova V (2011). Establishment and maintenance of sand fly colonies. J Vector Ecol 36: S1–S9.

30. Chmelar J, Anderson JM, Mu J, Jochim RC, Valenzuela JG, et al. (2008). Insight into the sialome of the castor bean tick, *Ixodes ricinus*. BMC Genomics 9: 233.

31. Hostomska J, Volfova V, Mu JB, Garfield M, Rohousova I, et al. (2009). Analysis of salivary transcripts and antigens of the sand fly *Phlebotomus arabicus*. BMC Genomics 10: 282.

32. Ewing B, Hillier L, Wendl MC, Green P (1998). Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res 8: 175–185.

33. Ewing B, Green P (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res 8: 186–194.

34. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. (2000). Gene Ontology: tool for the unification of biology. Nat Genet 25: 25–29.

35. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, et al. (2003). The COG database: an updated version includes eukaryotes. BMC Bioinformatics 4: 4.

36. Bateman A, Birney E, Durbin R, Eddy SR, Howe KL, et al. (2000). The Pfam protein families database. Nucleic Acids Res 28: 263–266.

37. Schultz J, Copley RR, Doerks T, Ponting CP, Bork P (2000). SMART: a web-based tool for the study of genetically mobile domains. Nucleic Acids Res 28: 231–234.

38. Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004). Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 340: 783–795.

39. Guo YJ, Ribeiro JMC, Anderson JM, Bour S (2009). dCAS: a desktop application for cDNA sequence annotation. Bioinformatics 25: 1195–1196.

40. Julenius K, Mølgaard A, Gupta R, Brunak S (2005). Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites. Glycobiology 15: 153–164.

41. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, et al. (2007). Clustal W and clustal X version 2.0. Bioinformatics 23: 2947–2948.

42. Abascal F, Zardoya R, Posada D (2005). ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21: 2104–2105.

43. Schmidt HA, Strimmer K, Vingron M, von Haeseler A (2002). TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics 18: 502–504.

44. Tamura K, Dudley J, Nei M, Kumar S (2007). MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24: 1596–1599.

45. Cerna P, Mikes L, Volf P (2002). Salivary gland hyaluronidase in various species of phlebotomine sand flies (Diptera : Psychodidae). Insect Biochem Mol Biol 32: 1691–1697.

46. Frost GI, Stern R (1997). A microtiter-based assay for hyaluronidase activity not requiring specialized reagents. Anal Biochem 251: 263–269.

47. Marinotti O, deBrito M, Moreira CK (1996). Apyrase and alpha-glucosidase in the salivary glands of *Aedes albopictus*. Comp Biochem Physiol 113B: 675–679.

48. Ribeiro JMC, Modi GB, Tesh RB (1989). Salivary apyrase activity of some Old World phlebotomine sand flies. Insect Biochemistry 19: 409–412.

49. Charlab R, Rowton ED, Ribeiro JM (2000). The salivary adenosine deaminase from the sand fly *Lutzomyia longipalpis*. Exp Parasitol 95: 45–53.

50. Kato H, Anderson JM, Kamhawi S, Oliveira F, Lawyer PG, et al. (2006). High degree of conservancy among secreted salivary gland proteins from two geographically distant *Phlebotomus duboscqi* sandflies populations (Mali and Kenya). BMC Genomics 7: 226.

51. Kato H, Jochim RC, Lawyer PG, Valenzuela JG (2007). Identification and characterization of a salivary adenosine deaminase from the sand fly *Phlebotomus duboscqi*, the vector of *Leishmania major* in sub-Saharan Africa. J Exp Biol 210: 733–740.

52. de Moura TR, Oliveira F, Carneiro MW, Miranda JC, Clarêncio J, et al. (2013). Functional transcriptomics of wild-caught *Lutzomyia intermedia* salivary glands: identification of a protective salivary protein against *Leishmania braziliensis* infection. PLoS Negl Trop Dis 7: e2242.

53. Charlab R, Valenzuela JG, Rowton ED, Ribeiro JMC (1999). Toward an understanding of the biochemical and pharmacological complexity of the saliva of a hematophagous sand fly *Lutzomyia longipalpis*. Proc Natl Acad Sci U S A 96: 15155–15160.

54. Valenzuela JG, Garfield M, Rowton ED, Pham VM (2004). Identification of the most abundant secreted proteins from the salivary glands of the sand fly *Lutzomyia longipalpis*, vector of *Leishmania chagasi*. J Exp Biol 207: 3717–3729.

55. Abdeladhim M, Jochim RC, Ben Ahmed M, Zhioua E, Chelbi I, et al. (2012). Updating the salivary gland transcriptome of *Phlebotomus papatasi* (Tunisian Strain): The search for sand fly-secreted immunogenic proteins for humans. PLoS One 7: e47347.

56. Kato H, Jochim RC, Gomez EA, Uezato H, Mimori T, et al. (2012). Analysis of salivary gland transcripts of the sand fly *Lutzomyia ayacuchensis*, a vector of Andean-type cutaneous leishmaniasis. Infect Genet Evol 13C: 56–66.

57. Volf P, Skarupova S, Man P (2002). Characterization of the lectin from females of *Phlebotomus duboscqi* sand flies. Eur J Biochem 269: 6294–6301.

58. Johnson JK, Li J, Christensen BM (2001). Cloning and characterization of a dopachrome conversion enzyme from the yellow fever mosquito, *Aedes aegypti*. Insect Biochem Mol Biol 31: 1125–1135.

59. Alves-Silva J, Ribeiro JMC, Van Den Abbeele J, Attardo G, Hao ZR, et al. (2010). An insight into the sialome of *Glossina morsitans morsitans*. BMC Genomics 11: 213.

60. Martin-Martin I, Molina R, Jimenez M (2012). An insight into the *Phlebotomus perniciosus* saliva by a proteomic approach. Acta Trop123: 22–30.

61. Bahia D, Gontijo NF, Leon IR, Perales J, Pereira MH, et al. (2007). Antibodies from dogs with canine visceral leishmaniasis recognise two proteins from the saliva of *Lutzomyia longipalpis*. Parasitol Res 100: 449–454.

62. Hostomska J, Rohousova I, Volfova V, Stanneck D, Mencke N, et al. (2008). Kinetics of canine antibody response to saliva of the sand fly *Lutzomyia longipalpis*. Vector Borne Zoonotic Dis 8: 443–450.

63. Vlkova M, Rohousova I, Drahota J, Stanneck D, Kruedewagen EM, et al. (2011). Canine antibody response to *Phlebotomus perniciosus* bites negatively correlates with the risk of *Leishmania infantum* transmission. PLoS Negl Trop Dis 5: e1344.

64. Gomes RB, Mendonça IL, Silva VC, Ruas J, Silva MB, et al. (2007). Antibodies against *Lutzomyia longipalpis* saliva in the fox *Cerdocyon thous* and the sylvatic cycle of *Leishmania chagasi*. Trans R Soc Trop Med Hyg 101: 127–133.

65. Gomes RB, Brodskyn U, de Oliveira CI, Costa J, Miranda JC, et al. (2002). Seroconversion against *Lutzomyia longipalpis* saliva concurrent with the development of anti-*Leishmania chagasi* delayed-type hypersensitivity. J Infect Dis 186: 1530–1534.

66. Vinhas V, Andrade BB, Paes F, Bomura A, Clarencio J, et al. (2007). Human anti-saliva immune response following experimental exposure to the visceral leishmaniasis vector, *Lutzomyia longipalpis*. Eur J Immunol 37: 3111–3121.

67. Marzouki S, Ahmed MB, Boussoffara T, Abdeladhim M, Aleya-Bouafif NB, et al. (2011). Characterization of the antibody response to the saliva of *Phlebotomus papatasi* in people living in endemic areas of cutaneous leishmaniasis. Am J Trop Med Hyg 84: 653–661.

68. Collin N, Gomes R, Teixeira C, Cheng L, Laughinghouse A, et al. (2009). Sand fly salivary proteins induce strong cellular immunity in a natural reservoir of visceral leishmaniasis with adverse consequences for *Leishmania*. PLoS Pathog 5: e1000441.

69. Santos A, Ribeiro JM, Lehane MJ, Gontijo NF, Veloso AB, et al. (2007). The sialotranscriptome of the blood-sucking bug *Triatoma brasiliensis* (Hemiptera: Triatominae). Insect Biochem Mol Biol 37: 702–712.

70. Jariyapan N, Roytrakul S, Paemanee A, Junkum A, Saeung A, et al. (2012). Proteomic analysis of salivary glands of female *Anopheles barbirostris* species A2 (Diptera: Culicidae) by two-dimensional gel electrophoresis and mass spectrometry. Parasitol Res 111: 1239–1249.

71. Andersen JF, Hinnebusch BJ, Lucas DA, Conrads TP, Veenstra TD, et al. (2007). An insight into the sialome of the oriental rat flea, *Xenopsylla cheopis* (Rots). BMC Genomics 8: 102.

72. Caljon G, De Ridder K, De Baetselier P, Coosemans M, Van Den Abbeele J (2010). Identification of a tsetse fly salivary protein with dual inhibitory action on human platelet aggregation. PLoS One 5: e9671.

73. Perez de Leon AA, Tabachnick WJ (1996). Apyrase activity and adenosine diphosphate induced platelet aggregation inhibition by the salivary gland proteins of *Culicoides variipennis*, the North American vector of bluetongue viruses. Vet Parasitol 61: 327–338.

74. Ma D, Wang Y, Yang H, Wu J, An S, et al. (2009). Anti-thrombosis repertoire of blood-feeding horsefly salivary glands. Mol Cell Proteomics 8: 2071–2079.

75. Wu S, Peiffer M, Luthe DS, Felton GW (2012). ATP hydrolyzing salivary enzymes of caterpillars suppress plant defenses. PLoS One 7: e41947.

76. Fenckova M, Hobizalova R, Fric ZF, Dolezal T (2011). Functional characterization of ecto-5′-nucleotidases and apyrases in *Drosophila melanogaster*. Insect Biochem Mol Biol 41: 956–967.

77. Ribeiro JM, Mans BJ, Arca B (2010). An insight into the sialome of blood-feeding Nematocera. Insect Biochem Mol Biol 40: 767–784.

78. Valenzuela JG, Charlab R, Galperin MY, Ribeiro JM (1998). Purification, cloning and expression of an apyrase from the bed bug *Cimex lectularius*. A new type of nucleotide-binding enzyme. J Biol Chem 273: 30583–30590.

79. Valenzuela JG, Belkaid Y, Rowton E, Ribeiro JM (2001b). The salivary apyrase of the blood-sucking sand fly *Phlebotomus papatasi* belongs to the novel *Cimex* family of apyrases. J Exp Biol 204: 229–237.

80. Hamasaki R, Kato H, Terayama Y, Iwata H, Valenzuela JG (2009). Functional characterization of a salivary apyrase from the sand fly, *Phlebotomus duboscqi*, a vector of *Leishmania major*. J Insect Physiol 55: 1044–1049.

81. Ribeiro JMC, Rossignol PA, Spielman A (1986). Blood-finding strategy of a capillary-feeding sandfly, *Lutzomyia longipalpis*. Comp Biochem Physiol 83A: 683–686.

82. Martin-Martin I, Molina R, Jimenez M (2013). Identifying salivary antigens of *Phlebotomus argentipes* by a 2DE approach. Acta Trop 126: 229–239.

83. Stern R, Jedrzejas MJ (2006). Hyaluronidases: their genomics, structures, and mechanisms of action. Chem Rev 106: 818–839.

84. Muller UR (2011). Hymenoptera venom proteins and peptides for diagnosis and treatment of venom allergic patients. Inflamm Allergy Drug Targets 10: 420–428.

85. Ribeiro JM, Charlab R, Rowton ED, Cupp EW (2000a). *Simulium vittatum* (Diptera: Simuliidae) and *Lutzomyia longipalpis* (Diptera: Psychodidae) salivary gland hyaluronidase activity. J Med Entomol 37: 743–747.

86. Volfova V, Hostomska J, Cerny M, Votypka J, Volf P (2008). Hyaluronidase of bloodsucking insects and its enhancing effect on *Leishmania* infection in mice. PLoS Negl Trop Dis 2: e294.

87. Ribeiro JM, Charlab R, Pham VM, Garfield M, Valenzuela JG (2004). An insight into the salivary transcriptome and proteome of the adult female mosquito *Culex pipiens quinquefasciatus*. Insect Biochem Mol Biol 34: 543–563.

88. Calvo E, Ribeiro JM (2006). A novel secreted endonuclease from *Culex quinquefasciatus* salivary glands. J Exp Biol 209: 2651–2659.

89. Jacobson RL, Schlein Y, Eisenberger CL (2001). The biological function of sand fly and *Leishmania* glycosidases. Med Microbiol Immunol 190: 51–55.

90. Ribeiro JM, Rowton ED, Charlab R (2000b). Salivary amylase activity of the phlebotomine sand fly, *Lutzomyia longipalpis*. Insect Biochem Mol Biol 30: 271–277.

91. Jacobson RL, Schlein Y (2001). *Phlebotomus papatasi* and *Leishmania major* parasites express alpha-amylase and alpha-glucosidase. Acta Trop 78: 41–49.

92. Andersen JF, Pham VM, Meng Z, Champagne DE, Ribeiro JM (2009). Insight into the sialome of the black fly, *Simulium vittatum*. J Proteome Res 8: 1474–1488.

93. Campbell CL, Vandyke KA, Letchworth GJ, Drolet BS, Hanekamp T, et al. (2005). Midgut and salivary gland transcriptomes of the arbovirus vector *Culicoides sonorensis* (Diptera:Ceratopogonidae). Insect Mol Biol 14: 121–136.

94. Valenzuela JG, Charlab R, Gonzalez EC, de Miranda-Santos IK, Marinotti O, et al. (2002). The D7 family of salivary proteins in blood sucking diptera. Insect Mol Biol 11: 149–55.

95. Gonzalez-Caballero N, Valenzuela JG, Ribeiro JM, Cuervo P, Brazil RP (2013). Transcriptome exploration of the sex pheromone gland of *Lutzomyia longipalpis* (Diptera: Psychodidae: Phlebotominae). Parasit Vectors 6: 56.

96. Calvo E, Mans BJ, Andersen JF, Ribeiro JMC (2006). Function and evolution of a mosquito salivary protein family. J Biol Chem 281: 1935–1942.

97. Mans BJ, Calvo E, Ribeiro JM, Andersen JF (2007). The crystal structure of D7r4, a salivary biogenic amine-binding protein from the malaria mosquito *Anopheles gambiae*. J Biol Chem 282: 36626–36633.

98. Isawa H, Yuda M, Orito Y, Chinzei Y (2002). A mosquito salivary protein inhibits activation of the plasma contact system by binding to factor XII and high molecular weight kininogen. J Biol Chem 277: 27651–27658.

99. Alvarenga PH, Francischetti IMB, Calvo E, Sa-Nunes A, Ribeiro JMC, et al. (2010). The function and three-dimensional structure of a tromboxane A(2)/cysteinyl leukotriene-binding protein from the saliva of a mosquito vector of the malaria parasite. Plos Biol 8: e1000547.

100. Marzouki S, Abdeladhim M, Abdessalem CB, Oliveira F, Ferjani B, et al. (2012). Salivary antigen SP32 is the immunodominant target of the antibody response to *Phlebotomus papatasi* bites in humans. PLoS Negl Trop Dis 6: e1911.

101. Elnaiem DE, Meneses C, Slotman M, Lanzaro GC (2005). Genetic variation in the sand fly salivary protein, SP-15, a potential vaccine candidate against *Leishmania major*. Insect Mol Biol 14: 145–150.

102. Milne TJ, Abbenante G, Tyndall JD, Halliday J, Lewis RJ (2003). Isolation and characterization of a cone snail protease with homology to CRISP proteins of the Pathogenesis-related Protein Superfamily. J Biol Chem 278: 31105–31110.

103. Yeats C, Bentley S, Bateman A (2003). New knowledge from old: In silico discovery of novel protein domains in *Streptomyces coelicolor*. BMC Microbiol 3: 3.

104. Hoffman DR (1993). Allergens in Hymenoptera venom. XXV: The amino acid sequences of antigen 5 molecules and the structural basis of antigenic cross-reactivity. J Allergy Clin Immunol 92: 707–716.

105. Lu G, Villalba M, Coscia MR, Hoffman DR, King TP (1993). Sequence analysis and antigenic cross-reactivity of a venom allergen, antigen 5, from hornets, wasps, and yellow jackets. J Immunol 150: 2823–2830.

106. King TP, Spangfort MD (2000). Structure and biology of stinging insect venom allergens. Int Arch Allergy Immunol 123: 99–106.

107. Asojo OA, Goud G, Dhar K, Loukas A, Zhan B, et al. (2005). X-ray structure of Na-ASP-2, a pathogenesis-related-1 protein from the nematode parasite, *Necator americanus*, and a vaccine antigen for human hookworm infection. J Mol Biol 346: 801–814.

108. Assumpção TC, Ma D, Schwarz A, Reiter K, Santana JM, et al. (2013). Salivary antigen-5/CAP family members are Cu2+-dependent antioxidant enzymes that scavenge O2- and inhibit collagen-induced platelet aggregation and neutrophil oxidative burst. J Biol Chem 288: 14341–14361.

109. Collin N, Assumpção TC, Mizurini DM, Gilmore DC, Dutra-Oliveira A, et al. (2012). Lufaxin, a novel factor Xa inhibitor from the salivary gland of the sand fly *Lutzomyia longipalpis* blocks protease-activated receptor 2 activation and inhibits inflammation and thrombosis in vivo. Arterioscler Thromb Vasc Biol 32: 2185–2198.

110. Alvar J, Velez ID, Bern C, Herrero M, Desjeux P, et al. (2012). Leishmaniasis worldwide and global estimates of its incidence. PLoS One 7: e35671.

111. Svobodova M, Alten B, Zidkova L, Dvorak V, Hlavackova J, et al. (2009). Cutaneous leishmaniasis caused by *Leishmania infantum* transmitted by *Phlebotomus tobbi*. Int J Parasitol 39: 251–256.

112. Di Muccio T, Marinucci M, Frusteri L, Maroli M, Pesson B, et al. (2000). Phylogenetic analysis of *Phlebotomus* species belonging to the subgenus *Larroussius* (Diptera, psychodidae) by ITS2 rDNA sequences. Insect Biochem Mol Biol 30: 387–393.

113. Esseghir S, Ready P, Ben-Ismail R (2000). Speciation of *Phlebotomus* sandflies of the subgenus *Larroussius* coincided with the late Miocene-Pliocene aridification of the Mediterranean subregion. Biol J Linn Soc Lond 70: 189–219.

114. Clements MF, Gidwani K, Kumar R, Hostomska J, Dinesh DS, et al. (2010). Measurement of recent exposure to *Phlebotomus argentipes*, the vector of indian visceral leishmaniasis, by using human antibody responses to sand fly saliva. Am J Trop Med Hyg 82: 801–807.

115. Barral A, Honda E, Caldas A, Costa J, Vinhas V, et al. (2000). Human immune response to sand fly salivary gland antigens: a useful epidemiological marker? Am J Trop Med Hyg 62: 740–745.

116. de Moura TR, Oliveira F, Novais FO, Miranda JC, Clarencio J, et al. (2007). Enhanced *Leishmania braziliensis* infection following pre-exposure to sandfly saliva. PLoS Negl Trop Dis 1: e84.

117. Aquino DM, Caldas AJ, Miranda JC, Silva AA, Barral-Netto M, et al. (2010). Epidemiological study of the association between anti-*Lutzomyia longipalpis* saliva antibodies and development of delayed-type hypersensitivity to *Leishmania* antigen. Am J Trop Med Hyg 83: 825–827.

Recombinant salivary proteins of *Phlebotomus orientalis* are suitable antigens to measure exposure of domestic animals to sand fly bites

Sima M, Ferencova B, Warburg A, Rohousova I, Volf P.

(2016a)

RESEARCH ARTICLE

Recombinant Salivary Proteins of *Phlebotomus orientalis* are Suitable Antigens to Measure Exposure of Domestic Animals to Sand Fly Bites

Michal Sima[1]*, **Blanka Ferencova**[1], **Alon Warburg**[2], **Iva Rohousova**[1], **Petr Volf**[1]

1 Department of Parasitology, Faculty of Science, Charles University in Prague, Prague, Czech Republic,
2 Department of Microbiology and Molecular Genetics, The Kuvin Centre for the Study of Infectious and Tropical Diseases, The Hebrew University - Hadassah Medical School, The Hebrew University of Jerusalem, Jerusalem, Israel

* michal-sima@seznam.cz

Abstract

Background

Certain salivary proteins of phlebotomine sand flies injected into the host skin during blood-feeding are highly antigenic and elicit strong antibody-mediated immune responses in repeatedly-exposed hosts. These antibodies can be measured by enzyme-linked immuno sorbent assays (ELISAs) using salivary gland homogenates (SGHs) as the source of antigens and serve as a markers for exposure to biting sand flies. Large-scale screening for anti-sand fly saliva antibodies requires replacement of SGH with recombinant salivary proteins. In East Africa, *Phlebotomus orientalis* is the main vector of *Leishmania donovani*, a trypanosomatid parasite causing visceral leishmaniasis. We tested recombinant salivary proteins derived from *Ph*. *orientalis* saliva to study exposure of domestic animals to this sand fly species.

Methodology/Principal Findings

Antigenic salivary proteins from *Ph*. *orientalis* were identified by immunoblot and mass spectrometry. Recombinant apyrase rPorSP15, yellow-related protein rPorSP24, ParSP25-like protein rPorSP65, D7-related protein rPorSP67, and antigen 5-related protein rPorSP76 were tested using ELISA with sera of domestic animals from *L. donovani* foci in Ethiopia where *Ph*. *orientalis* is present. Our results highlighted recombinant yellow-related protein rPorSP24 as the most promising antigen, displaying a high positive correlation coefficient as well as good sensitivity and specificity when compared to SGH. This recombinant protein was the most suitable one for testing sera of dogs, sheep, and goats. In addition, a different antigen, rPorSP65 was found efficacious for testing canine sera.

Conclusions/Significance

Recombinant salivary proteins of *Ph*. *orientalis*, specifically rPorSP24, were shown to successfully substitute SGH in serological experiments to measure exposure of domestic animals to *Ph*. *orientalis*, the vector of *L*. *donovani*. The results suggest that rPorSP24 might be a suitable antigen for detecting anti-*Ph*. *orientalis* antibody-mediated reactions also in other host species.

Author Summary

The sand fly *Phlebotomus orientalis* is the main vector of *Leishmania donovani*, the causative agent of visceral leishmaniasis in East Africa. During bloodfeeding, sand flies inject saliva into the host skin and repeated bites result in a specific antibody response in the bitten hosts. Antibody responses are directed against sand fly salivary proteins and the levels of these antibodies reflect the intensity of exposure to biting sand flies. The antibody reactions can be measured using salivary gland homogenates (SGHs), but for large-scale testing its use is impractical because of the amount of work required to obtain sufficient quantities of SGH. Recombinant proteins prepared based on the antigens in the sand fly saliva can substitute whole SGH in large-scale studies. We tested five recombinant proteins from *Ph*. *orientalis* saliva expressed in *Escherichia coli* and demonstrated that the yellow-related protein rPorSP24 can replace the SGH in estimating exposure to sand flies of dogs, goats, and sheep in Ethiopia. Immune reactions to vector saliva in endemic areas, provides useful information on levels of exposure and, thereby, on the effectiveness of vector control programs.

Introduction

Phlebotomine sand flies are the vectors of *Leishmania* parasites causing leishmaniasis, the disease responsible for an estimated 1.3 million new human cases and 20 000 to 30 000 deaths annually [1]. During blood-feeding, sand fly females inoculate saliva into the host skin. Over the last three decades, various research groups have investigated the composition and biological activities of saliva, as well the potential use of salivary antigens in an anti-*Leishmania* vaccine (reviewed in [2]).

Sand fly salivary molecules are also highly antigenic and elicit strong antibody-mediated response in repeatedly exposed hosts. This response can be utilized as a marker for exposure to biting sand flies. In animals experimentally-exposed to sand fly bites the production of specific anti-saliva IgG antibodies is positively correlated with the number of blood-fed sand flies [3,4]. The elevated antibody levels persisted in bitten hosts for weeks or even months [3–6] but decreased after the last exposure to sand flies, suggesting that screening for anti-saliva antibodies can be used also for estimating the timing of exposure [7,8]. As a reliable epidemiological tool, anti-sand fly saliva antibodies have already been successfully employed to evaluate the effectiveness of vector control interventions [4,9], to estimate the risk of *Leishmania* transmission [4, 10–12], and to indicate the feeding preferences of sand flies [13–15].

Screening for anti-sand fly saliva antibodies in large populations is impractical due to the amount of work required to obtain sufficient quantities of salivary gland homogenate (SGH). However, the use of recombinant salivary proteins enables to circumvent the necessity for sand

fly colony maintenance, laborious dissections of salivary glands and potential cross-reactivity with non-vector species [8]. The main salivary antigens in several sand fly species have already been characterized [4,12,16–18], however, recombinant salivary proteins from only three species—*Lutzomyia longipalpis*, *Ph. perniciosus*, and *Ph. papatasi*—have been tested so far in seroepidemiological studies [13,19–25].

Here, we focus on *Ph. orientalis*, the most important vector of human visceral leishmaniasis (VL) in East Africa [reviewed in [26]]. In Ethiopia, the main endemic areas of VL are located in the lowlands of southwestern Ethiopia and in the Metema-Humera plains in the northwest [27], where *Ph. orientalis* was found to be an abundant sand fly species [28]. This opportunistic sand fly feeds on different mammals, depending on the host availability [29–31]. Indeed, anti-*Ph. orientalis* antibodies have recently been detected in several domestic animal species in Ethiopia—dogs, donkeys, sheep, goats, and cows using SGH as antigen [15]. In the present study, five proteins from saliva of *Ph. orientalis* were expressed in *Esherichia coli* and evaluated as markers for exposure using sera of domestic animals, namely dogs, sheep, and goats, from *L. donovani* endemic foci in northern Ethiopia.

Methods

Ethical statement

BALB/c mice were maintained and handled in the animal facility of Charles University in Prague in accordance with institutional guidelines and the Czech legislation (Act No. 246/1992 coll. on Protection of Animals against Cruelty in present statutes at large), which complies with all relevant European Union and international guidelines for experimental animals. The experiments were approved by the Committee on the Ethics of Animal Experiments of the Charles University in Prague (Permit Number: 24773/2008-10001) and were performed under the Certificate of Competency (Registration Number: CZ 02439, CZ 02457) in accordance with the Examination Order approved by Central Commission for Animal Welfare of the Czech Republic. Sera of domestic animals were collected within the study by [15]. Their collection was approved by the Ethiopian National Research Ethics Review Committee (NRERC) under approval no. 3.10/3398/04. For more details see [15].

Host sera

Murine sera were obtained from animals exposed at least ten-times to about 150 insectary-bred sand fly females of a single species in two-week interval. Ten mice were exposed *to Ph. orientalis*, four to *Ph. papatasi*, and four to *Sergentomyia schwetzi*. Four mice served as non-exposed controls. Serum samples of Ethiopian domestic animals, obtained during the previous study by [15] included 179 sheep, 36 dog, and 233 goat sera. Sera from 30 sheep, 14 dogs, and 15 goats non-exposed to sand flies were used as negative controls. More details about samples from domestic animals (both of Ethiopian origin and controls) are provided in [15].

Sand flies and salivary gland dissection

The colony of *Ph. orientalis* (originating from Ethiopia, Melka Werer) was established in 2008 [32] and reared under standard conditions as described in [33]. Salivary glands were dissected from 4–6 day old female sand flies in 20 mM Tris buffer with 150 mM NaCl and stored at -20°C. Before use, salivary glands were disrupted by freeze-thawing three times in liquid nitrogen [34].

Immunoblot

Antigenic proteins were selected based on the reactivity of two pools of canine sera (five sera each) from endemic area in Ethiopia with SGH using one pool (five sera) of non-exposed dogs as control. Salivary proteins (equivalent to 20 glands per well) were separated by SDS-PAGE on 12% polyacrylamide gel under non-reducing conditions. Proteins were transferred from the gel to nitrocellulose membranes using an iBLOT dry system (Invitrogen). Membranes were cut into strips and blocked overnight with 5% nonfat dry milk in Tris buffer with 0.05% Tween (Tris-Tw) and incubated for 1 hour with dog sera diluted 1:50 in Tris-Tw. Then, the strips were incubated with peroxidase conjugated anti-dog IgG (Bethyl Laboratories) diluted 1:3000 in Tris-Tw. Antigenic protein bands were visualized using the substrate solution with diamino-benzidine. The protein profile was compared with *Ph. orientalis* SGH studied by [35] and the identity of antigenic bands was confirmed by proteome analysis and mass spectrometry; according to the protocol described in [35].

Recombinant proteins preparation

Five proteins from *Ph. orientalis* salivary glands were chosen for expression in *E. coli*: PorSP15, PorSP24, PorSP65, PorSP67, and PorSP76 (in [35] marked as PorASP15, PorMSP24, PorMSP65, and PorMSP67, and PorASP76, respectively) (Table 1). The PCR products from a previously constructed cDNA library [35] were used as the starting material. Products were amplified by PCR under the following conditions: initial incubation (3 minutes in 94°C), then 30 cycles of denaturation (30 seconds in 94°C), annealing (1 minute in 57°C), and elongation (1 minute in 72°C). The whole reaction was terminated by heating to 72°C for 10 minutes. Specific primers were synthesized according to the sequences of the mature protein (without the signal peptide) (Table 1). Thereafter, we followed the procedure described in [5]. Briefly, PCR products were ligated into *E. coli* pGEM-T Easy Vector (Promega) using TA cloning and the ligation products were transfected into *E. coli* competent cells TOP10 (Invitrogen). Vectors were replicated in bacteria and after that, the gene for yellow-related protein was restricted using Spe I and Xho I and the genes for the remaining four proteins were restricted using Nde I and Xho I enzymes. Restricted *E. coli* pET-42 Expression Vectors (Novagen) were ligated and ligation products were transformed into *E. coli* competent cells TOP10 (Invitrogen) again. Plasmids were isolated from the bacteria, and transfected into *E. coli* BL21 (DE3) gold (Agilent) for expression. *E. coli* lysates were prepared under denaturing conditions and His-tagged proteins (with six histidins) were purified under denaturing conditions with 8M urea in Ni-NTA column (731–1550: Bio-Rad, USA). Purity of the recombinant proteins was verified on immunoblot using the monoclonal anti-polyHistidine-peroxidase (A7058-1VL: Sigma Aldrich, UK) and protein concentration was measured by the Lowry method (Bio-Rad) following the manufacturer's protocol.

ELISA experiments

The ELISA protocol described in [15] was used with the following modification: The ELISA plates Immulon 4HBX (96w flat bottomed plate, 735–0465: VWR, USA) were coated in concentrations of 5 µg/ml (0.5 µg/well) for recombinant proteins and 28 ng/well for SGH (corresponding to 0.2 of salivary gland/well). In all ELISA tests, serum samples were tested individually.

In the first series of experiments (evaluation step), sera from ten experimentally-bitten mice and sera of 10 dogs, 10 goats, and 35 sheep with the highest anti-*Ph. orientalis* SGH titer values found by [15] were used to evaluate the antigenicity of the five recombinant proteins with anti-

Table 1. *Phlebotomus orientalis* salivary proteins expressed in *Escherichia coli*.

Sequence name	Protein family	GenBank ACCN	Forward primer	Reverse primer
PorSP15	Apyrase	AGT96431	CATATGGCTCCTAGAGCAACAAAAT	CTCGAGCTTAATGCCTTTGGGAT
PorSP24	Yellow-related protein	AGT96428	ACTAGTTTTCACGTTGAAAGAGAAT	CTCGAGCTTTGTCTTGGGATCATA
PorSP65	ParSP25-like protein	AGT96466	CATATGGATCGGGGAGTGGATGG	CTCGAGGTGCAATCGGTTGTTTATG
PorSP67	D7-related protein	AGT96467	CATATGCTGCAATTCCCTCGTGAT	CTCGAGTTTTGCCGATATCTCATCC
PorSP76	Antigen 5-related protein	AGT96441	CATATGGCTAACGACTATTGCCAGC	CTCGAGTGTCCTGGGCTTCTTGAG

Sequence name, the protein family, GenBank accession number, and forward and reverse primers for PCR amplification are provided for each protein.

doi:10.1371/journal.pntd.0004553.t001

Ph. orientalis saliva IgG. Sera from non-exposed animals (3 mouse, 3 dogs, 3 goats, and 8 sheep) were used as negative controls.

In the second series of experiments, 16 murine sera (4 exposed to *Ph. orientalis*, 4 to *Ph. papatasi*, 4 to *Se. schwetzi*, and 4 non-exposed controls) were used to verify specificity of selected recombinant proteins. Based on the results of evaluation experiments with murine sera, three recombinant proteins with significant correlation with SGH (rPorSP24, rPorSP67, and rPorSP76) were selected.

In the third series of experiments (validation step), selected recombinant proteins with correlation coefficient higher than 0.7 from evaluation experiments (rPorSP15, rPorSP24, rPorSP65, and rPorSP67) were tested with the whole set of serum samples from Ethiopia (179 sheep, 36 dogs, and 233 goats) and an appropriate number of non-exposed controls (30 sheep, 14 dogs, and 15 goats).

Statistical analysis

The non-parametric Spearman test was used to assess correlations between total anti-SGH and anti-recombinant protein IgG levels using GraphPad Prism version 6 (GraphPad Software, Inc., San Diego, CA). For evaluating the possible cross-reactivity with other sand fly species non-parametric Wilcoxon Rank-Sum test in GraphPad Prism version 6 (GraphPad Software, Inc., San Diego, CA) was used. Statistical significance was considered when the p-value was<0.05. Cut-off values were calculated from the mean optical density of control sera plus 3 standard deviations. The optical density values of anti-SGH antibodies were used as the gold standard to validate recombinant proteins in ELISA tests using positive and negative predictive values, sensitivity, and specificity.

Accesion numbers

Accession numbers of proteins used in this study: AGT96431, AGT96428, AGT96466, AGT96467, AGT96441. Accession numbers of proteins discussed in this study: AAL16051, AHA49643, AAL11049, AAL11048, AFY13224, ABI20147, AHF48995, AHF48996, AAD32198, AAS05318, AHF49000

Results

Identification of *Ph*. *orientalis* salivary antigens by immunoblot

To identify antigenic proteins in *Ph. orientalis* salivary glands, two pools of canine sera (five sera each) from naturally-exposed dogs were tested with SGH of *Ph. orientalis*. Individual bands were identified based on the proteomic analysis, immunoblot, and mass spectrometry of *Ph. orientalis* SGH. Canine sera reacted with at least 10 protein bands (Fig 1); five of them were
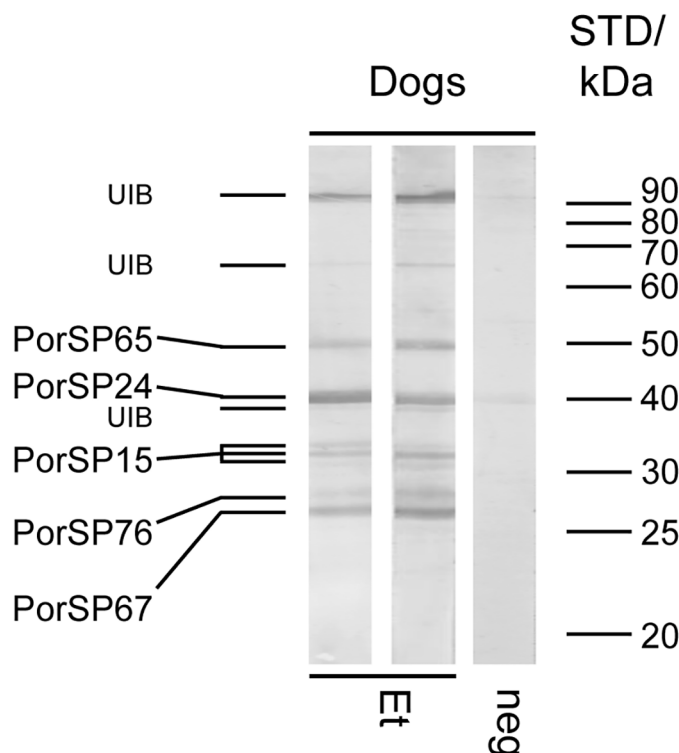
Fig 1. Identification of *Phlebotomus orientalis* salivary antigens in dogs. *Ph. orientalis* salivary proteins were separated under non-reducing conditions by SDS-PAGE on a 12% gel and incubated with two different pools of sera from naturally-exposed Ethiopian dogs (Et), and one pooled sera from non-exposed control dogs (neg). Each pool was a mixture of 5 individual samples. Five antigenic proteins (PorSP65, PorSP24, PorSP15, PorSP76, and PorSP67) were identified by successive proteome analysis and mass spectrometry. Molecular weights of standard (STD) are indicated. UIB means unidentified bands.

doi:10.1371/journal.pntd.0004553.g001

identified as ParSP25-like protein PorSP65, yellow-related protein PorSP24, apyrase PorSP15, antigen 5-related protein PorSP76, and D7-related protein PorSP67 (for GenBank ACCN refer to Table 1). These five antigenic proteins were chosen for expression in *E. coli*. Very weak reaction was observed between SGH and negative control sera around 30, 40, and 90 kDa (Fig 1).

Evaluation of recombinant proteins with anti-*Ph. orientalis* IgG

To evaluate the reactivity of anti-*Ph. orientalis* saliva IgG with five recombinant proteins, we screened them first with selected sera of naturally-exposed domestic animals from Ethiopia (dogs, sheep, and goats), mice experimentally-bitten by *Ph. orientalis*, and non-exposed controls. The antigenicity of recombinant proteins was evaluated based on the correlation of antibody reactions with SGH for each of the twenty combinations between recombinant proteins and animal species (Table 2).

In canine sera, a significant correlation was achieved for all tested proteins; the highest correlation coefficient was found for rPorSP24 ($\rho = 0.868$), followed by rPorSP65, and rPorSP15. Similarly, in sheep and goat sera the best correlation (above 0.8) was found for rPorSP24, other recombinant proteins with correlation coefficient above 0.75 were rPorSP67 and rPorSP15 for goats and rPorSP65 for sheep (Table 2). Sera from experimentally-bitten mice showed significant correlation with three out of five proteins tested; the highest correlation coefficient was achieved for rPorSP24 ($\rho = 0.857$).

**Table 2. Evaluation of recombinant proteins by correlation analysis.**

HOST		SGH	rPorSP15	rPorSP24	rPorSP65	rPorSP67	rPorSP76
Dog	controls (n = 3)	0.396 ± 0.049	**0.329 ± 0.037**	**0.289 ± 0.008**	**0.278 ± 0.038**	0.341± 0.050	0.332 ± 0.053
	exposed (n = 10)	1.129 ± 0.336	**0.709 ± 0.181**	**0.734 ± 0.232**	**1.008 ± 0.304**	0.742 ± 0.227	0.755 ± 0.230
	ρ	N.A.	**0.830 *****	**0.868 *****	**0.852 *****	0.687 **	0.599 *
Goat	Controls (n = 3)	0.143 ± 0.002	0.215 ± 0.021	**0.150 ± 0.023**	0.161 ± 0.030	**0.213 ± 0.031**	0.354 ± 0.058
	exposed (n = 10)	0.265 ± 0.224	0.559 ± 0.244	**0.329 ± 0.234**	0.411 ± 0.234	**0.422 ± 0.202**	0.737 ± 0.230
	ρ	N.A.	0.797 ***	**0.835 *****	0.412	**0.802 *****	0.643*
Sheep	controls (n = 8)	0.133 ± 0.014	0.294 ± 0.041	**0.213 ± 0.048**	0.166 ± 0.027	0.288 ± 0.047	0.520 ± 0.113
	exposed (n = 35)	0.218 ± 0.141	0.486 ± 0.138	**0.238 ± 0.165**	0.313 ± 0.180	0.353 ± 0.084	0.573 ± 0.091
	ρ	N.A.	0.629 ***	**0.806 *****	0.777 ***	0.373 *	0.349 *
Mouse	controls (n = 3)	0.086 ± 0.006	0.158 ± 0.022	**0.142 ± 0.037**	0.154 ± 0.034	0.136 ± 0.029	**0.214 ± 0.059**
	exposed (n = 10)	1.395 ± 0.562	0.245 ± 0.186	**0.394 ± 0.329**	0.753 ± 0.761	0.334 ± 0.314	**0.391 ± 0.136**
	ρ	N.A.	0.308	**0.857 *****	0.456	0.560 *	**0.802 *****

Spearman-Rank Correlation Matrix test for optical densities between sera tested against *Ph. orientalis* SGH and against each salivary recombinant protein. Mean values of exposed animals and non-exposed controls ± standard deviation, correlation coefficient (ρ) are indicated. N.A. = not applicable, asterisk (*) indicates significant correlations: *p<0.05, **p<0.01, ***p<0.001. Combinations of significant correlation and ρ>0.8 are in bold.

Specificity of recombinant proteins

Specificity of recombinant proteins was tested only with murine sera due to the absence of positive control samples from other host species. Three recombinant proteins with significant correlation to SGH in evaluation experiments with murine sera (Table 2) were selected to verify their specific reaction with anti-*Ph. orientalis* IgG. Fig 2 shows the strong reactions of sera from mice exposed to *Ph. orientalis* bites with SGH and with all tested recombinant proteins (rPorSP67, rPorSP76, and rPorSP24), while the recognition of these antigens by sera from mice exposed to solely to *Ph. papatasi* or *Se. schwetzi* were similar to the negative controls (sera of unexposed mice).
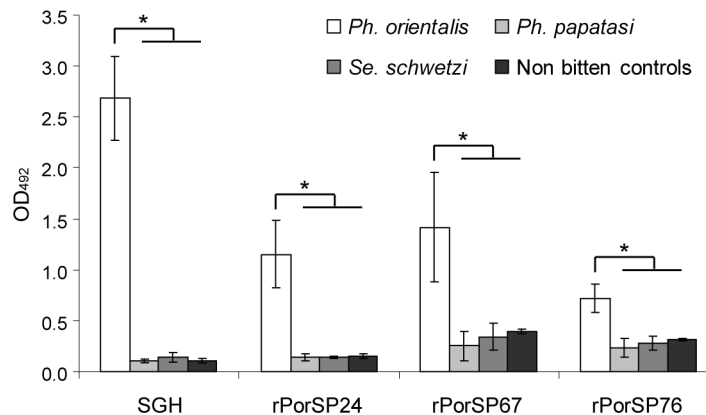
Fig 2. Specificity of recombinant proteins. Results from ELISA are presented as mean optical densities ± standard deviation of IgG antibody reaction with *P. orientalis* salivary gland homogenate (SGH) and three recombinant proteins (rPorSP24, rPorSP67, and rPorSP76) in mice experimentally bitten by *Ph. orientalis*, *Ph. papatasi*, or *Se. schwetzi*, and non-exposed control mice. Four mice per sand fly colony and four non-exposed controls were used. Asterisks (*) indicate significant differences (p<0.05, calculated with non-parametric Wilcoxon Rank-Sum Test) of IgG levels between mice bitten by *Ph. orientalis* and mice bitten by other sand fly species or non-bitten controls.

Validation of selected recombinant proteins

Four recombinant proteins with the highest correlation from evaluation experiments (rPorSP15, rPorSP24 rPorSP65, and rPorSP67) were chosen for further validation using the whole set of Ethiopian serum samples (179 sheep, 36 dogs, and 233 goats) and non-exposed controls.

For canine sera, the highest correlation coefficient was achieved with rPorSP65 (ρ = 0.906) followed by rPorSP24 and PorSP15. For sheep as well as for goats, the highest correlation coefficient was detected with rPorSP24 (ρ = 0.818 and ρ = 0.522, respectively) followed with rPorSP65 for sheep and with rPorSP15 and rPorSP67 for goats (Fig 3). All results from correlation analyses between SGH and four recombinant proteins were highly significant and cut-off values for individual recombinant proteins were the lowest for the proteins with the highest correlation coefficient (Fig 3). Additionally, rPorSP24 reached the highest values of positive and negative predictive values (PPV and NPV) in all host species as well as the sensitivity in

HOST		SGH	rPorSP15	rPorSP24	rPorSP65	rPorSP67
Dog	Cut-off	0.353	0.382	0.310	0.279	Correlation coefficient in evaluation experiments was lower than 0.7
	Exposed	0.504 ± 0.339	0.434 ± 0.140	0.423 ± 0.176	0.573 ± 0.282	
	ρ	N.A.	0.726 ***	0.790 ***	0.906 ***	
	PPV		0.650	0.655	0.545	
	NPV		0.625	1.000	0.667	
	Sensitivity		0.684	1.000	0.947	
	Specificity		0.588	0.412	0.118	
Goat	Cut-off	0.113	0.434	0.259	Correlation coefficient in evaluation experiments was lower than 0.7	0.411
	Exposed	0.127 ± 0.082	0.401 ± 0.103	0.231 ± 0.088		0.434 ± 0.113
	ρ	N.A.	0.513 ***	0.522 ***		0.283 ***
	PPV		0.701	0.727		0.452
	NPV		0.665	0.667		0.604
	Sensitivity		0.475	0.404		0.576
	Specificity		0.837	0.887		0.481
Sheep	Cut-off	0.156	Correlation coefficient in evaluation experiments was lower than 0.7	0.265	0.273	Correlation coefficient in evaluation experiments was lower than 0.7
	Exposed	0.190 ± 0.082		0.280 ± 0.104	0.326 ± 0.135	
	ρ	N.A.		0.818 ***	0.678 ***	
	PPV			0.932	0.822	
	NPV			0.582	0.556	
	Sensitivity			0.683	0.733	
	Specificity			0.898	0.678	

Fig 3. Validation of recombinant proteins. Selected recombinant proteins were validated in ELISA tests with sera of indicated domestic animals naturally-exposed to *Phlebotomus orientalis*. The analysis was based on comparison with anti-*Ph. orientalis* SGH IgG as a standard. The table provides cut-off values, mean values of optical density ± standard deviation of IgG levels in animals exposed to *Ph. orientalis*, correlation coefficients between IgG levels against SGH and a recombinant protein (ρ), positive predictive values (PPV), negative predictive values (NPV), sensitivity, and specificity. Asterisks (*) indicate significant correlations: *p<0.05, **p<0.01, ***p<0.001. Combinations with the correlation coefficient lower than 0.7 in evaluation experiments (Table 2) were excluded from the validation experiments. For each combination, the lowest cut-off value, the highest correlation coefficient, and the highest PPV, NPV, sensitivity, and specificity values are shaded grey. N.A. means not applicable.

doi:10.1371/journal.pntd.0004553.g003
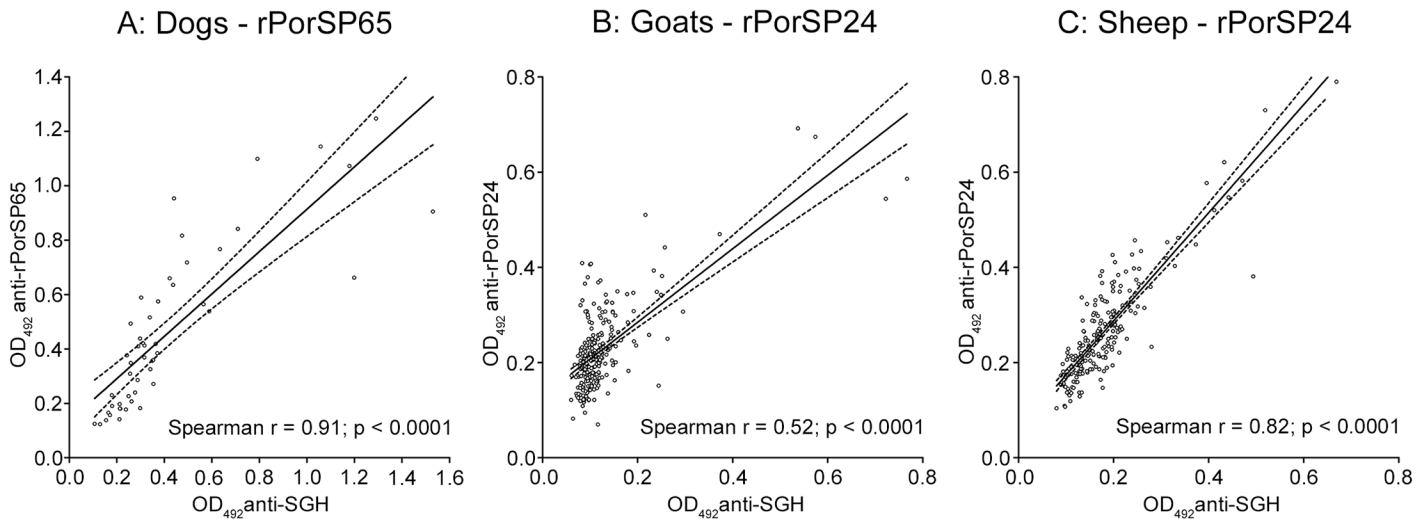
Fig 4. Correlation analyzes between IgG antibodies against SGH and recombinants rPorSP65 and rPorSP24 in ELISA. For each animal species from validation experiments (Fig 3), the protein with the highest positive correlation was displayed: A: rPorSP65 (ParSP25-like protein) tested with canine sera (n = 50), B: rPorSP24 (yellow-related protein) tested with goat sera (n = 248), C: rPorSP24 (yellow-related protein) tested with sheep sera (n = 209). Sera from naturally-exposed Ethiopian animals together with sera from non-exposed controls were included in this analysis. Correlation coefficients and p-values from Spearman-Rank analysis are indicated.

doi:10.1371/journal.pntd.0004553.g004

dogs and the specificity in goats and sheep. The specificity in dogs was the best with rPorSP15 and the sensitivity in goats with rPorSP67. The best combinations (the lowest cut-off value, the highest correlation coefficient, PPV, NPV, specificity, and sensitivity values) between SGH and recombinant protein for each animal species tested are shown in Fig 4.

Discussion

We have studied antigenic salivary proteins of *Ph. orientalis*, the most important vector of VL in Ethiopia and Sudan, using sera of naturally-exposed hosts. Antigenic proteins were identified on immunoblot based on their recognition by canine sera. These were: D7-related protein (PorSP67), antigen 5-related protein (PorSP76), apyrase (PorSP15), yellow-related protein (PorSP24), and ParSP25-like protein (PorSP65). The antigenicity of their recombinant counterparts expressed in *E. coli* was validated in large-scale tests using sera from naturally-exposed dogs, sheep, and goats. The utilization of recombinant proteins as markers for exposure may help to highlight the specificity of the reaction and to evade nonspecific binding as observed when SGH was recognized by negative control. sera.

Sand fly salivary proteins from the D7-related family are well known antigens; they were recognized by sera of mice bitten by *Ph. papatasi* [10], dogs bitten by *Lu. longipalpis* and *Ph. perniciosus* [4,36], and humans bitten by *Ph. papatasi* [12]. As far as we are aware, five recombinant D7-related (rD7) proteins from sand fly saliva have already been tested as exposure markers; one from *Lu. longipalpis* (AAL16051, also known as LJL13) [13], another from *Ph. perniciosus* (AHA49643) [21], and three from *Ph. papatasi* (AAL11049, AAL11048, and AFY13224) [5,20]. However, only some of them bound anti-saliva IgG and their antigenicity was host-specific. The rD7 protein (AAL11049) was specifically recognized by sera from mice bitten by *Ph. papatasi* [5], but the same protein was not recognized by human sera from Tunisia [20]. Another rD7 protein (AAL11048) from *Ph. papatasi* did not react with sera from mice immunized by sand flies [5]. The rD7 protein LJL13 was recognized by dogs naturally-exposed to *Lu. longipalpis* but not by sera from foxes and humans from endemic focus of *L. infantum*

[13]. In the present study, rPorSP67 showed promising results with limited number of goat sera during the evaluation test but was not validated in a broader test, when medium or low correlation coefficient, NPV, PPV, sensitivity, and specificity with values ranging between 0.28 and 0.6 were observed. This suggests that this recombinant protein would not be useful as an exposure marker to *Ph. orientalis* bites.

Proteins of the antigen 5-related family from various sand fly species were repeatedly shown to be potent salivary antigens, being recognized by sera of mice bitten by *Ph. papatasi* and *Ph. arabicus* [5,16], dogs bitten by *Ph. perniciosus* [4], rabbits exposed to *Ph. tobbi* [18], and hamsters bitten by *Ph. argentipes* [37]. In our study, rPorSP76 showed high correlation with SGH only with sera from experimentally-bitten mice (ρ = 0.8), thus its use in field studies with domestic animals is not justified.

Apyrases are well-known salivary enzymes with anti-haemostatic properties [38]. Antigenic properties of apyrases were described in SGHs of various vector-host models such as *Ph. perniciosus* and dogs [4, 17] or *Ph. argentipes* and hamsters [37]. Recombinant apyrase (ABI20147) was recognised by sera of mice immunized with *Ph. duboscqi* saliva [39]. Two apyrases in a recombinant form (AHF48995, AHF48996) were used as exposure markers for dogs, hares, and rabbits bitten by *Ph. perniciosus* [21,22], however, the most recent work by Kostalova et al. [23] revealed that neither of these recombinant proteins gave optimal ELISA results in large-scale tests with naturally-exposed dogs. Our results showed a significant correlation between antibody response against SGH of *Ph. orientalis* and recombinant apyrase rPorSP15 in sera of all domestic animals tested but not in mice sera. The best correlation (ρ = 0.7) was observed with canine sera with medium values of NPV, PPV, sensitivity, and specificity (ranging between 0.59 and 0.68) suggesting necessity of further validation before its utilization as antigen for detecting dog exposure to *Ph. orientalis*.

The most promising and universal candidate for an exposure marker to sand fly bites belongs to the family of yellow-related proteins. Strong antibody responses to these proteins were previously demonstrated for various sand fly and host species, including dogs and humans [3–7,10,12–14,17,36,40]. *Ph. perniciosus* recombinant yellow-related protein AHF49000 was successfully used as an antigen both in ELISA and immunoblot reacting well with sera of mice and dogs experimentally-bitten by *Ph. perniciosus* [21]. Yellow-related recombinant proteins were also validated as exposure markers to sand fly bites in endemic areas; *Lu. longipalpis* AAD32198 and AAS05318 for humans [19,13] and *Ph. perniciosus* AHF49000 for dogs, rabbits, and hares [22,23]. In the present study, rPorSP24 from *Ph. orientalis* confirmed the high reactivity of yellow-related protein family with antibodies from sera of bitten hosts and this advocates for its use as an exposure marker in large-scale field studies. It reached high correlation with SGH in mice (ρ = 0.9), sheep (ρ = 0.8), and dogs (ρ = 0.8) and the best in goats (ρ = 0.5). Similarly, high correlation between SGH and recombinant yellow-related protein was previously attained with dogs (ρ > 0.7 [21–23]), hares (ρ = 0.9 [22]), and rabbits (ρ = 0.7 [22]). Moreover, rPorSP24 achieved the highest values of specificity, sensitivity, PPV and NPV with majority of the host species tested. In dogs, this recombinant protein showed 100% NPV and sensitivity but lower values of PPV and specificity (0.66 and 0.41, respectively), indicating higher probability of false positivity among non-exposed dogs. On the other hand, in sheep, very high values of PPV and specificity were achieved (both 0.9) but medium values of NPV and sensitivity (0.58 and 0.68, respectively) could indicate possible false negative results. However, this statistical analysis is based on data from naturally-exposed hosts and negative controls; further validation is needed using sera of experimentally-bitten animals as a positive control.

The fifth recombinant protein tested belongs to the ParSP25-like family. Antigenicity of this protein family was previously demonstrated in *Ph. perniciosus* [4,17]. So far as we are aware,

no recombinant protein from this group was prepared and used for measuring the antibody reaction with sera of bitten hosts. Our results suggest significant correlation between rPorSP65 and antibody response against SGH of *Ph. orientalis*. The highest correlation coefficient was observed with canine sera ($\rho = 0.9$), accompanied by high degree of sensitivity (0.95). Nevertheless, the specificity of the test with rPorSP65 was very low (0.1) suggesting high probability of false positivity among non-exposed dogs.

Antigenic specificity of recombinant proteins was confirmed by using murine sera experimentally-exposed to *Sergentomyia schwetzi* or *Phlebotomus papatasi*. These two sand fly species are present in Ethiopia, in some places sympatrically with *Ph. orientalis* [27]. Antibodies from sera of mice bitten by *Ph. papatasi* or *Se. schwetzi* did not react with the recombinant proteins with significant correlation from evaluation experiments (rPorSP24, rPorSP67, and rPorSP76) and confirmed that they are species-specific.

The reactivity of recombinant proteins might be affected by the expression system or conditions of protein purification. Antibodies from bitten hosts can be targeted also to the glycosylated parts of the antigen that are lacking in proteins from the *E. coli* expression system, therefore some authors prefer to express recombinant proteins in mammalian cells [20,24]. Nevertheless, some of the recombinant proteins without posttranslational modifications proved to be as efficient markers of exposure as native antigens [5,21–23]. Similarly, several proteins prepared in this study were found as suitable antigens, despite their being expressed in *E. coli*.

In conclusion, our study suggests rPorSP24, the recombinant protein from the yellow-related family, as the most reliable and universally efficacious antigen for measuring exposure of dogs, sheep and goats to *Ph. orientalis* bites. In addition, the recombinant protein rPorSP65 from ParSP25-like group was found as a good antigen to screen for canine exposure but its low specificity suggests possible false positivity in some specimens. Serological tests using these proteins could be a highly practical and economical tool for screening of domestic animals for exposure to the main vector of *L. donovani* in East Africa. Moreover, the promising characteristics of rPorSP24 suggest a potential use of this antigen for screening sera of other hosts, including humans. The availability of recombinant salivary proteins should enable to measure anti-*Ph. orientalis* antibodies in large-scale experiments to evaluate vector control programs in areas affected by VL in East Africa. However, further studies are needed to validate such recombinant protein-based test for routine use.

Acknowledgments

Author Contributions

Conceived and designed the experiments: PV IR. Performed the experiments: MS BF. Analyzed the data: MS. Contributed reagents/materials/analysis tools: PV IR AW. Wrote the paper: MS IR PV AW.

References

1. WHO 2015. http://www.who.int/mediacentre/factsheets/fs375/en/. Retrieved on November 4th 2015.

2. Abdeladhim M, Kamhawi S, Valenzuela JG. What's behind a sand fly bite? The profound effect of sand fly saliva on host hemostasis, inflammation and immunity. Infect Genet Evol. 2014; 28: 691–703. doi: 10.1016/j.meegid.2014.07.028 PMID: 25117872

3. Hostomska J, Rohousova I, Volfova V, Stanneck D, Mencke N, Volf P. Kinetics of canine antibody response to saliva of the sand fly *Lutzomyia longipalpis*. Vector Borne Zoonotic Dis. 2008; 8: 443–450. doi: 10.1089/vbz.2007.0214 PMID: 18260789

4. Vlkova M, Rohousova I, Drahota J, Stanneck D, Kruedewagen EM, Mencke N, et al. Canine antibody response to *Phlebotomus perniciosus* bites negatively correlates with the risk of *Leishmania infantum* transmission. PLoS Negl Trop Dis. 2011; 5: e1344. doi: 10.1371/journal.pntd.0001344 PMID: 22022626

5. Vlkova M, Rohousova I, Hostomska J, Pohankova L, Zidkova L, Drahota J, et al. Kinetics of antibody response in BALB/c and C57BL/6 mice bitten by *Phlebotomus papatasi*. PLoS Negl Trop Dis. 2012; 6: e1719. doi: 10.1371/journal.pntd.0001719 PMID: 22802977

6. Martin-Martin I, Molina R, Jimenez M. Kinetics of Anti-*Phlebotomus perniciosus* Saliva Antibodies in Experimentally Bitten Mice and Rabbits. PLoS ONE. 2015; 10: e0140722. doi: 10.1371/journal.pone.0140722 PMID: 26569103

7. Vinhas V, Andrade BB, Paes F, Bomura A, Clarencio J, Miranda JC, et al. Human anti-saliva immune response following experimental exposure to the visceral leishmaniasis vector, *Lutzomyia longipalpis*. Eur J Immunol. 2007; 37: 3111–3121. PMID: 17935072

8. Clements MF, Gidwani K, Kumar R, Hostomska J, Dinesh DS, Kumar V, et al. Measurement of recent exposure to *Phlebotomus argentipes*, the vector of Indian visceral leishmaniasis, by using human antibody responses to sand fly saliva. Am J Trop Med Hyg. 2010; 82: 801–807. doi: 10.4269/ajtmh.2010.09-0336 PMID: 20439958

9. Gidwani K, Picado A, Rijal S, Singh SP, Roy L, Volf P, et al. Antibody response to sand fly saliva to evaluate human visceral leishmaniasis vector exposure in India and Nepal: effect of long-lasting insecticidal nets. Trop Med Int Health. 2011; 16: 66–66

10. Rohousova I, Ozensoy S, Ozbel Y, Volf P. Detection of species-specific antibody response of humans and mice bitten by sand flies. Parasitology. 2005; 130: 493–499. PMID: 15991492

11. de Moura TR, Oliveira F, Novais FO, Miranda JC, Clarêncio J, Follador I, et al. Enhanced *Leishmania braziliensis* infection following pre-exposure to sandfly saliva. PLoS Negl Trop Dis. 2007; 1: e84. PMID: 18060088

12. Marzouki S, Ben Ahmed M, Boussoffara T, Abdeladhim M, Ben Aleya-Bouafif N, Namane A, et al. Characterization of the antibody response to the saliva of *Phlebotomus papatasi* in people living in endemic areas of cutaneous leishmaniasis. Am J Trop Med Hyg. 2011; 84: 653–661. doi: 10.4269/ajtmh.2011.10-0598 PMID: 21540371

13. Teixeira C, Gomes R, Collin N, Reynoso D, Jochim R, Oliveira F, et al. Discovery of markers of exposure specific to bites of *Lutzomyia longipalpis*, the vector of *Leishmania infantum* chagasi in Latin America. PLoS Negl Trop Dis. 2010; 4: e638. doi: 10.1371/journal.pntd.0000638 PMID: 20351786

14. Gomes RB, Mendonça IL, Silva VC, Ruas J, Silva MB, Cruz MS, et al. Antibodies against *Lutzomyia longipalpis* saliva in the fox *Cerdocyon thous* and the sylvatic cycle of *Leishmania chagasi*. Trans R Soc Trop Med Hyg. 2007; 101: 127–133. PMID: 16887159

15. Rohousova I, Talmi-Frank D, Kostalova T, Polanska N, Lestinova T, Kassahun A, et al. Exposure to *Leishmania* spp. and sand flies in domestic animals in northwestern Ethiopia. Parasit Vectors. 2015; 8: 360. doi: 10.1186/s13071-015-0976-1 PMID: 26152578

16. Hostomska J, Volfova V, Mu JB, Garfield M, Rohousova I, Volf P, et al. Analysis of salivary transcripts and antigens of the sand fly *Phlebotomus arabicus*. BMC Genomics. 2009; 10: 282. doi: 10.1186/1471-2164-10-282 PMID: 19555500

17. Martin-Martin I, Molina R, Jimenez M. An insight into the *Phlebotomus perniciosus* saliva by a proteomic approach. Acta Trop. 2012; 123: 22–30. doi: 10.1016/j.actatropica.2012.03.003 PMID: 22445778

18. Rohousova I, Volfova V, Nova S, Volf P. Individual variability of salivary gland proteins in three *Phlebotomus* species. Acta Trop. 2012; 122: 80–86. doi: 10.1016/j.actatropica.2011.12.004 PMID: 22197318

19. Souza AP, Andrade BB, Aquino D, Entringer P, Miranda JC, Alcántara R, et al. Using recombinant proteins from *Lutzomyia longipalpis* saliva to estimate human vector exposure in visceral leishmaniasis endemic areas. PLoS Negl Trop Dis. 2010; 4: e649. doi: 10.1371/journal.pntd.0000649 PMID: 20351785

20. Marzouki S, Abdeladhim M, Abdessalem CB, Oliveira F, Ferjani B, Gilmore D, et al. Salivary antigen SP32 is the immunodominant target of the antibody response to *Phlebotomus papatasi* bites in humans. PLoS Negl Trop Dis. 2012; 6: e1911. doi: 10.1371/journal.pntd.0001911 PMID: 23209854

21. Drahota J, Martin-Martin I, Sumova P, Rohousova I, Jimenez M, Molina R, et al. Recombinant antigens from *Phlebotomus perniciosus* saliva as markers of canine exposure to visceral leishmaniases vector. PLoS Negl Trop Dis. 2014; 8: e2597. doi: 10.1371/journal.pntd.0002597 PMID: 24392167

22. Martin-Martin I, Molina R, Rohousova I, Drahota J, Volf P, Jimenez M. High levels of anti-*Phlebotomus perniciosus* saliva antibodies in different vertebrate hosts from the re-emerging leishmaniosis focus in Madrid, Spain. Vet Parasitol. 2014; 202: 207–216. doi: 10.1016/j.vetpar.2014.02.045 PMID: 24629428

23. Kostalova T, Lestinova T, Sumova P, Vlkova M, Rohousova I, Berriatua E, et al. Canine antibodies against salivary recombinant proteins of *Phlebotomus perniciosus*: A longitudinal study in an endemic focus of canine leishmaniasis. PLoS Negl Trop Dis. 2015; 9: e0003855. doi: 10.1371/journal.pntd.0003855 PMID: 26111018

24. Marzouki S, Kammoun-Rebai W, Bettaieb J, Abdeladhim M, Hadj Kacem S, Abdelkader R, et al. Validation of recombinant salivary protein PpSP32 as a suitable marker of human exposure to *Phlebotomus papatasi*, the vector of *Leishmania major* in Tunisia. PLoS Negl Trop Dis. 2015; 9: e0003991. doi: 10.1371/journal.pntd.0003991 PMID: 26368935

25. Mondragon-Shem K, Al-Salem WS, Kelly-Hope L, Abdeladhim M, Al-Zahrani M, Valenzuela JG, et al. Severity of Old World Cutaneous Leishmaniasis Is Influenced by Previous Exposure to Sandfly Bites in Saudi Arabia. PLoS Negl Trop Dis. 2015; 9: e3449.

26. Elnaiem DE. Ecology and control of the sand fly vectors of *Leishmania donovani* in East Africa, with special emphasis on *Phlebotomus orientalis*. J Vector Ecol. 2011; 36: S23–S31. doi: 10.1111/j.1948-7134.2011.00109.x PMID: 21366778

27. Elnaiem DA, Ward R, Hassan KH, Miles MA, Frame IA. Infection rates of *Leishmania donovani* in *Phlebotomus orientalis* from a focus of visceral leishmaniasis in eastern Sudan. Ann Trop Med Parasitol. 1998; 92: 229–232. PMID: 9625920

28. Gebresilassie A, Kirstein OD, Yared S, Aklilu E, Moncaz A, Tekie H, et al. Species composition of phlebotomine sand flies and bionomics of *Phlebotomus orientalis* (Diptera: Psychodidae) in an endemic focus of visceral leishmaniasis in Tahtay Adiyabo district, Northern Ethiopia. Parasit Vectors. 2015; 8: 248. doi: 10.1186/s13071-015-0849-7 PMID: 25907252

29. Gebre-Michael T, Balkew M, Berhe N, Hailu A, Mekonnen Y. Further studies on the phlebotomine sandflies of the kala-azar endemic lowlands of Humera-Metema (north-west Ethiopia) with observations on their natural blood meal sources. Parasit Vectors. 2010; 3: 6. doi: 10.1186/1756-3305-3-6 PMID: 20181077

30. Gebresilassie A, Yared S, Aklilu E, Kirstein OD, Moncaz A, Tekie H, et al. Host choice of *Phlebotomus orientalis* (Diptera: Psychodidae) in animal baited experiments: a field study in Tahtay Adiyabo district, northern Ethiopia. Parasit Vectors. 2015; 8: 190. doi: 10.1186/s13071-015-0807-4 PMID: 25885333

31. Gebresilassie A, Abbasi I, Aklilu E, Yared S, Kirstein OD, Moncaz A, et al. Host-feeding preference of *Phlebotomus orientalis* (Diptera: Psychodidae) in an endemic focus of visceral leishmaniasis in northern Ethiopia. Parasit Vectors. 2015; 8: 270. doi: 10.1186/s13071-015-0883-5 PMID: 25963759

32. Seblova V, Volfova V, Dvorak V, Pruzinova K, Votypka J, Kassahun A, et al. *Phlebotomus orientalis* Sand Flies from Two Geographically Distant Ethiopian Localities: Biology, Genetic Analyses and Susceptibility to *Leishmania donovani*. PLoS Negl Trop Dis. 2013; 7: e2187. doi: 10.1371/journal.pntd.0002187 PMID: 23638207

33. Volf P, Volfova V. Establishment and maintenance of sand fly colonies. J Vector Ecol. 2011; 36: S1–S9. doi: 10.1111/j.1948-7134.2011.00106.x PMID: 21366760

34. Volf P, Rohousova I. Species-specific antigens in salivary glands of phlebotomine sandflies. Parasitology. 2001; 122: 37–41 PMID: 11197762

35. Vlkova M, Sima M, Rohousova I, Kostalova T, Sumova P, Volfova, et al. Comparative analysis of salivary gland transcriptomes of *Phlebotomus orientalis* sand flies from endemic and non-endemic foci of visceral leishmaniasis. PLoS Negl Trop Dis. 2014; 8: e2709. doi: 10.1371/journal.pntd.0002709 PMID: 24587463

36. Bahia D, Gontijo NF, Leon IR, Perales J, Pereira MH, Oliveira G, et al. Antibodies from dogs with canine visceral leishmaniasis recognise two proteins from the saliva of *Lutzomyia longipalpis*. Parasitol Res. 2007; 100: 449–454. PMID: 17058112

37. Martin-Martin I, Molina R, Jimenez M. Identifying salivary antigens of *Phlebotomus argentipes* by a 2DE approach. Acta Trop. 2013; 126: 229–239. doi: 10.1016/j.actatropica.2013.02.008 PMID: 23422341

38. Ribeiro JMC, Arca B. From sialomes to the sialoverse: An insight into the salivary potion of blood feeding insects. Adv In Insect Phys. 2009; 37: 59–118.

39. Hamasaki R, Kato H, Terayama Y, Iwata H, Valenzuela JG. Functional characterization of a salivary apyrase from the sand fly, *Phlebotomus duboscqi*, a vector of *Leishmania major*. J Insect Physiol. 2009. 55: 1044–1049. doi: 10.1016/j.jinsphys.2009.07.010 PMID: 19651132

40. Gomes RB, Brodskyn C, de Oliveira CI, Costa J, Miranda JC, Caldas A, et al. Seroconversion against *Lutzomyia longipalpis* saliva concurrent with the development of anti-*Leishmania chagasi* delayed-type hypersensitivity. J Infect Dis. 2002; 186: 1530–1534. PMID: 12404176

The diversity of yellow-related proteins in sand flies (Diptera: Psychodidae)

Sima M, Novotny M, Pravda L, Sumova P, Rohousova I, Volf P.

(2016b)

1 **Title:**

2 The diversity of yellow-related proteins in sand flies (Diptera: Psychodidae)

3

4 **Author names and affiliations:**

5 Michal Sima[1*], Marian Novotny[2], Lukas Pravda[3,4], Petra Sumova[1], Iva Rohousova[1], and Petr

6 Volf[1]

7

8

9 [1] Department of Parasitology, Faculty of Science, Charles University, Prague, Czech Republic

10 [2] Department of Cell Biology, Faculty of Science, Charles University, Prague, Czech

11 Republic

12 [3] CEITEC - Central European Institute of Technology, Masaryk University, Brno, Czech

13 Republic

14 [4] National Centre for Biomolecular Research, Faculty of Science, Masaryk University, Brno,

15 Czech Republic

16

17

18 ***Corresponding author**

19 E-mail: michal-sima@seznam.cz

Abstract

Yellow-related proteins (YRPs) present in sand fly saliva act as affinity binders of bioamines, and help the fly to complete a bloodmeal by scavenging the physiological signals of damaged cells. They are also the main antigens in sand fly saliva and their recombinant form is used as a marker of host exposure to sand flies. Moreover, several salivary proteins and plasmids coding these proteins induce strong immune response in hosts bitten by sand flies and are being used to design protecting vaccines against *Leishmania* parasites. In this study, thirty two 3D models of different yellow-related proteins from thirteen sand fly species of two genera were constructed based on the known protein structure from *Lutzomyia longipalpis*. We also studied evolutionary relationships among species based on protein sequences as well as sequence and structural variability of their ligand-binding site. All of these 33 sand fly YRPs shared a similar structure, including a unique tunnel that connects the ligand-binding site with the solvent by two independent paths. However, intraspecific modifications found among these proteins affects the charges of the entrances to the tunnel, the length of the tunnel and its hydrophobicity. We suggest that these structural and sequential differences influence the ligand-binding abilities of these proteins and provide sand flies with a greater number of YRP paralogs with more nuanced answers to bioamines. All these characteristics allow us to better evaluate these proteins with respect to their potential use as part of anti-*Leishmania* vaccines or as an antigen to measure host exposure to sand flies.

Introduction

During intake of a bloodmeal, sand flies (Diptera: Phlebotominae) and other bloodsucking insects inject saliva into the host skin. This saliva contains a mixture of various proteins, which play a major role in preventing host haemostatic and inflammatory responses of different pathways, e.g. platelet activation, coagulation, inflammation, mast cell function, and vasoconstriction (reviewed in [1]). In sand flies, the vectors of *Leishmania* protozoan parasites, these salivary proteins have been studied for decades due to their biological activities and possible use in anti-*Leishmania* vaccines (reviewed in [2]). In repeatedly bitten hosts, several salivary proteins elicit a strong antibody response, which can be utilized for the detection of exposure to sand flies in epidemiological studies [3–7].

Transcripts of yellow-related proteins (YRPs) have been present in the salivary cDNA libraries of all sand fly species tested to date [8–21]. Usually, they are found in more than one homolog, which may occur in N-glycosylated, O-glycosylated, C-glycosylated, or non-glycosylated forms [16, 20].

YRPs are known as (1) kratagonists that remove small molecule mediators of haemostasis by high affinity ligand-binding proteins, and as (2) antigens that elicit a host immune response, both antibody and cell-mediated. All sequences of sand fly YRPs contain the entire insect-specific MRJP (major royal jelly protein) domain, which defines this protein family across several insect orders/families including *Drosophila melanogaster* [22], honeybees [23], mosquitoes [24], and tse-tse flies [25]. In *Lutzomyia longipalpis*, Xu et al. [26] demonstrated that YRPs are high affinity binders of pro-haemostatic and pro-inflammatory biogenic amines, such as serotonin, histamine and catecholamines. Blocking of these small molecules by YRPs results in vasodilatation, platelet deactivation, and a decrease in vascular permeability [27, 28]. The YRP of *Phlebotomus duboscqi* was shown to have lectin-like activity and is swallowed into the midgut, together with saliva [29].

Sand fly salivary recombinant YRPs are the most promising antigens for measuring exposure in naturally bitten hosts, and have been the subject of large epidemiological studies [30–34]. Besides eliciting an antibody response, YRPs as well as plasmids coding these proteins of *Lutzomyia longipalpis* induce a strong delayed type hypersensitivity (DTH) reaction, which leads to protection against *Leishmania major* in vivo [26, 35] and against *L. infantum* in vitro [36]. This suggests a possible use of these proteins in an anti-*Leishmania* vaccine.

72   In 2011, the crystal structure of *L. longipalpis* YRP LJM11 (GenBank ACCN:
73   AAS05318) was published as 3Q6K (Protein Data Bank ID) with a description of the ligand-
74   binding pocket [26]. Based on this structure and available amino acid sequences obtained
75   from GenBank, we constructed 3D models of all YRPs identified so far in sand fly sialomes.
76   We predicted their phylogenetic relationships, glycosylation sites, surface electrostatic
77   potentials, compared their sequences and characterized the ligand-binding tunnel. Our results
78   show differences among individual proteins within one species as well as differences among
79   various species. Our results may lead to a better understanding of the biological function of
80   YRPs.
81
82   **Methods**
83
84   **Phylogenetic analysis**
85
86   Amino acid sequences of YRPs were identified in public databases at NCBI using BLAST
87   [37] based on similarity with *Lutzomyia longipalpis* LJM11 (Protein Data Bank: 3Q6K, for
88   the purpose of this study called 3Q6K_Llon1), the best explored protein from this group, [26].
89   All these analyses were performed for the sequences without a signal peptide, which was
90   identified using SignalP 4.0 [38]. Sequences were consequently aligned using ClustalX
91   (version 2.0) [39]. The best substitution matrix for creating a phylogenetic tree of sand fly
92   salivary YRPs was determined in ProtTest software 2.0 [40]. TREEPUZZLE 5.2 [41] was
93   used to create a maximum likelihood phylogenetic tree from the protein alignment using the
94   WAG model [42] and quartet puzzling with 10000 puzzling steps. The resulting tree for all 31
95   proteins from 11 sand fly species in two genera was visualized in MEGA 4 [43] and rooted by
96   the related protein from *Drosophila melanogaster* (ACCN: NP650247). Clustal Omega [44]
97   was used with default settings to calculate a Percent Identity Matrix among all sand fly
98   salivary YRPs.
99
100   **Prediction of glycosylation**
101
102   Putative N-, O-, and C- glycosylation sites for all 31 protein sequences were determined using
103   NetNGlyc 1.0, NetOGlyc 4.0 [45], and NetCGlyc 1.0 [46] servers with default settings.
104

3D models construction

All proteins were modeled using 3Q6K, the only available structure of sand fly salivary YRP in Protein Data Bank (PDB, [47]), from where its PDB file and fasta sequence were downloaded. The alignment of template and target sequence was done in Clustal Omega [44] for all proteins. Scripts for Python version 2.7 (Python Software Foundation) were prepared in a txt file and ran in MODELLER [48]. Five models were calculated for each protein, with the best one chosen based on the lowest energy levels of molpdf and the DOPE score. All models were displayed and analyzed in PyMOL (The PyMOL Molecular Graphics System, Version 1.5 Schrödinger, LLC.). Electrostatic surface potentials were calculated using the APBS Tools2 plugin [49] in PyMOL.

Tunnel analysis

Tunnels in the protein structures were detected and characterized using the channel analysis tool MOLE 2.0 [50]. All proteins were superimposed using PyMOL in order to use the same settings, and therefore obtain comparable tunnels. In brief, MOLE 2.0 calculates a Delaunay triangulation/Voronoi diagram of the atomic centers. Next, tunnels are identified between every tuple of user defined end points within a protein structure using Dijkstra's algorithm. The resulting tunnels are defined by their centerline and are uniformly divided into layers. Each layer is defined by the residues lining it. A new layer starts whenever there is a change in residues lining it along its length. Additionally, MOLE 2.0 inferrs basic physicochemical properties for each tunnel as well as layers. These were calculated from a unique set of lining residues averaging tabulated values for hydrophobicity [51]. WebLogo 2.8.2. [52] and Clustal Omega [44] were used to visualize differences among protein tunnels.

Results

Identification of sand fly salivary YRPs

Thirty two sand fly salivary YRPs were identified based on their similarity with PDB ID: 3Q6K from *L. longipalpis* [26], the only known 3D structure of YRPs of phlebotominae sand flies. These YRPs were detected in all 13 sand fly species with a published salivary gland

transciptome: 9 *Phlebotomus* species from 5 subgenera (*Phlebotomus*, *Paraphlebotomus*, *Larroussius*, *Adlerius*, and *Euphlebotomus*), and 4 *Lutzomyia* species from 3 subgenera (*Lutzomyia, Nyssomyia,* and *Helcocyrtomyia*). In summary, there are currently 33 different YRPs described in the Phlebotominae subfamily (Table 1). There are large differences in the number of these proteins found in various species, ranging from 1 in *P. arabicus* and *P. argentipes* to 5 in *P. sergenti*.

Table 1. Identified sand fly salivary YRPs

ACCN or name	Sand fly species	Subgenus	Published by	Identifier
BAM69109	*L. ayacuchensis*	*Helcocyrtomyia*	[19]	Laya1
BAM69110	*L. ayacuchensis*	*Helcocyrtomyia*	[19]	Laya2
BAM69111	*L. ayacuchensis*	*Helcocyrtomyia*	[19]	Laya3
BAM69185	*L. ayacuchensis*	*Helcocyrtomyia*	[19]	Laya4
AFP99235	*L. intermedia*	*Nyssomyia*	[17]	Lint
AAS05318, LJM11, 3Q6K	*L. longipalpis*	*Lutzomyia*	[10, 25]	3Q6K_Llon1
AAD32198, LJM17	*L. longipalpis*	*Lutzomyia*	[8]	Llon2
ABB00904, LJM111	*L. longipalpis*	*Lutzomyia*	[10]	Llon3
ANW11467	*L. olmeca*	*Nyssomyia*	[21]	Lolm1
ANW11468	*L. olmeca*	*Nyssomyia*	[21]	Lolm2
ANW11469	*L. olmeca*	*Nyssomyia*	[21]	Lolm3
ACS93501	*P. arabicus*	*Adlerius*	[14]	Para
ABA12136	*P. argentipes*	*Euphlebotomus*	[12]	Parg
AAX44093	*P. ariasi*	*Larroussius*	[11]	Pari1
AAX56360	*P. ariasi*	*Larroussius*	[11]	Pari2
ABI15938	*P. duboscqi*	*Phlebotomus*	[13]	Pdub1
ABI15941	*P. duboscqi*	*Phlebotomus*	[13]	Pdub2
ABI20172	*P. duboscqi*	*Phlebotomus*	[13]	Pdub3
AGT96460	*P. orientalis*	*Larroussius*	[20]	Pori1
AGT96461	*P. orientalis*	*Larroussius*	[20]	Pori2
AAL11051	*P. papatasi*	*Phlebotomus*	[9]	Ppap1
AAL11052	*P. papatasi*	*Phlebotomus*	[9]	Ppap2
AGE83094	*P. papatasi*	*Phlebotomus*	[15]	Ppap3
AGE83095	*P. papatasi*	*Phlebotomus*	[15]	Ppap4
ABA43049	*P. perniciosus*	*Larroussius*	[12]	Pper1
ABA43050	*P. perniciosus*	*Larroussius*	[12]	Pper2
ADJ54114	*P. sergenti*	*Paraphlebotomus*	[16]	Pser1
ADJ54115	*P. sergenti*	*Paraphlebotomus*	[16]	Pser2
ADJ54116	*P. sergenti*	*Paraphlebotomus*	[16]	Pser3

ADJ54122	*P. sergenti*	*Paraphlebotomus*	[16]	Pser4
ADJ54123	*P. sergenti*	*Paraphlebotomus*	[16]	Pser5
ADJ54079	*P. tobbi*	*Larroussius*	[16]	Ptob1
ADJ54080	*P. tobbi*	*Larroussius*	[16]	Ptob2

147 Sequence GenBank accession numbers (or names), sand fly species (ordered alphabetically),

148 references, and identifier names used in this study are provided for each protein.

149

150

151

152 **Phylogenetic analysis**

153

154 The phylogenetic analysis was performed using the maximum likelihood phylogenetic tree of

155 all 33 YRPs from 13 sand fly species. A higher variability in *Lutzomyia* YRPs compared to

156 *Phlebotomus* proteins was found: two clusters from seven *Lutzomyia* YRPs were detected but

157 four other *Lutzomyia* proteins created own branches. Twenty two *Phlebotomus* proteins

158 clustered together in one branch with high bootstrap support. A closer phylogenetic

159 relationship was discovered between the subgenera *Phlebotomus* (*P. papatasi* and *P.*

160 *duboscqi*) and *Paraphlebotomus* (*P. sergenti*), as well as between *Larrousius* (*P. tobbi, P.*

161 *orientalis, P. perniciosus,* and *P. ariasi*) and *Adlerius* (*P. arabicus*) species. *Euphlebotomus*

162 (*P. argentipes*) protein created a single branch, which is closer to species from subgenera

163 *Larroussius* and *Adlerius* (Fig 1).

164

165

166

167 **Fig 1. Phylogenetic tree of 33 sand fly salivary YRPs with putative glycosylation sites**

168 **marked.** The maximum likelihood phylogenetic tree was created in TREEPUZZLE with

169 WAG model using quartet puzzling with 10000 puzzling steps. For rooting, the related

170 protein (ACCN: NP650247) from *Drosophila melanogaster* (Dmel) was used. Bootstraps

171 with support for branching are shown. The letters N, O, and C indicate putative N-, O-, and C-

172 glycosylation sites, respectively. Protein codes refer to Table 1.

173

174

175

Based on the identity matrix from Clustal Omega, the similarity among all 33 YRPs

177   (528 combinations) varied from 44 to 99% (S1 Table).

178

179   **Glycosylation**

180

181   For all 33 proteins, potential N-, O- and C-glycosylation sites were predicted. In 11 proteins,

182   only N-glycosylation sites were predicted. In four YRPs of *P. sergenti* (Pser1-Pser4), only O-

183   glycosylation sites were identified, and C-glycosylation sites were predicted for only two

184   proteins from *P. duboscqi* (Table 2). In eleven proteins, more than one type of putative

185   glycosylation was found.

186   The predicted glycosylation sites for the majority of YRPs (77% and 53% of putative

187   N-glycans and O-glycans, respectively) share the same positions (Table 2). Importantly,

188   proteins clustering together in the phylogenetic tree (Fig 1) had similar predicted

189   glycosylation patterns.

190

191

192   **Table 2. Putative glycosylation sites**

Identifier	N-glycosylation	O-glycosylation	C-glycosylation
Laya1	Asn 164, Asn 196	Ser 208	-
Laya2	Asn 164, Asn 196	Ser 113, Ser 208	-
Laya3	Asn 164, Asn 196	Ser 208	-
Laya4	Asn 11, Asn 264	Ser 262	-
Lint	Asn 11	Thr 262, Thr 263	-
3Q6K_Llon1	Asn 195	-	-
Llon2	Asn 11	-	-
Llon3	Asn 123	Ser 300	-
Lolm1	Asn 122	Thr 262	-
Lolm2	Asn 194	Thr 261	-
Lolm3	Asn 11	Thr 197	-
Para	Asn 11	-	-
Parg	Asn 11, Asn 18, Asn 307	-	-
Pari1	Asn 11, Asn 201	-	-
Pari2	-	-	-
Pdub1	Asn 11, Asn 65	Thr 367	-
Pdub2	-	Ser 208	Trp 338
Pdub3	-	Ser 208	Trp 338
Pori1	-	-	-

Pori2	Asn 11	-	-
Ppap1	-	-	-
Ppap2	Asn 11, Asn 65	-	-
Ppap3	-	-	-
Ppap4	Asn 11, Asn 65	Ser 264	-
Pper1	Asn 11	-	-
Pper2	Asn 11	-	-
Pser1	-	Ser 81, Ser 208, Thr 263	-
Pser2	-	Ser 208, Thr 263	-
Pser3	-	Ser 208, Thr 263	-
Pser4	-	Ser 208	-
Pser5	Asn 11, Asn 65, Asn 250	-	-
Ptob1	-	-	-
Ptob2	Asn 11	-	-

193 N-, O- and C-glycosylation sites were predicted for each protein using glycosylation servers.
194 Numbers and three-letter abbreviations indicate the positions of predicted glycosylation and
195 the amino acid where the glycosylation occurs (Asn − asparagine, Thr − threonine, Ser −
196 serine, and Trp − tryptophan). **-** shows cases where no glycosylation sites were identified.
197 Protein codes refer to Table 1.

198

199

200

201 **3D model analysis**

202

203 The amino acid sequence alignment of 33 sand fly salivary YRPs showed high conservancy in
204 several regions, with 57 invariant sites identified among all sequences, representing
205 approximately 15% of the total sequence. Five out of eleven amino acids creating the ligand-
206 binding pocket known from Xu et al. [26] were identical across all proteins. The most
207 variable position in the ligand-binding site was the one annotated as Phe 344 in the crystal
208 structure of YRP 3Q6K - this position is occupied by six different amino acids (His, Gln, Tyr,
209 Phe, Met, or Lys) in our studied set of proteins. Two most common amino acids on this
210 position are histidine and glutamine (Fig 2).

211

212

213

214   **Fig 2. Amino acid alignment within the YRP ligand-binding sites.** The alignment was

215   created in ClustalX, and shows high conservancy in binding sites (letters in brackets below

216   the sequences, slashes indicate amino acid substitutions in appropriate positions). Ligand

217   binding sites are highlighted in black, three bold stars below the sequences represent the main

218   binding sites where hydrogen bonds between amino acids and the ligand are predicted.

219   Protein codes refer to Table 1.

220

221

222

223   Models for 32 different proteins were created based on the template sequence 3Q6K

224   from *L. longipalpis* (S1 Fig). All of them folded as a six-bladed beta-propeller, similarly as in

225   3Q6K [26]. In 3Q6K, 11 ligand-binding amino acids were identified [26]. Three of them (in

226   the case of 3Q6K – Thr 327 with 2 bonds, Asn 342 with 2 bonds, and Phe 344 with 1 bond)

227   play a major role in binding abilities, because of the hydrogen bonds between the protein

228   amino acids and the ligand. The other 8 amino acids bind serotonin by Van der Waals or

229   hydrophobic interactions. These three amino acids were visualized in the 3D alignment of all

230   33 proteins. Nine different amino acids were found at three positions that create hydrogen

231   bonds with serotonin from 3Q6K (Fig 3).

232

233

234

235   **Fig 3. Hydrogen-binding between serotonin and ligand-binding amino acids in sand fly**

236   **salivary YRPs.** The figure shows the variability in hydrogen bonds between serotonin and

237   3Q6K (provided by Thr 327, Asn 342, and Phe 344) described in Xu et al. [26] and other

238   YRP models created during this study and visualized in PyMOL. The red bold molecule

239   symbolizes the ligand serotonin and the black dashed line show the hydrogen bonds between

240   amino acids and serotonin in 3Q6K. Other colors indicate one representative member of one

241   protein group. A group was defined as having at least one unique mutation from the template

242   structure in one of the three described amino acids binding the ligand. From each group, the

243   protein with the most abundant amino acid rotamer was chosen as being representative for

244   this visualization. Protein codes refer to Table 1.

245

246

247

248   As well as in 3Q6K_Llon1, in Laya1-3 and Lolm1, the amino acids creating hydrogen

249 bonds with serotonin are Thr 327, Asn 342, and Phe 344 (using the numbering for 3Q6K).

250 The largest group includes 11 proteins (Llon2, Para, Parg, Pari, Pdub1, Pori2, Ppap2, Ppap4,

251 Pper1, Pser5, and Ptob1). In comparison with the group containing 3Q6K_Llon1, there is a

252 substitution of histidine instead of phenylalanine in the last binding position. A substitution in

253 the same position (glutamine instead of phenylalanine) also occurs in the second biggest

254 group, which includes 10 proteins (Pari1, Pdub2, Pdub3, Pori2, Ppap1, Ppap3, Pser1, Pser2,

255 Pser3, and Pser4). The last group with more than one protein (Laya4, Lint, and Lolm3) has

256 substitution in the last binding position, as well. In this case, there is tyrosine instead of

257 phenylalanine. Pper2 has substitutions in all three positions (serine, threonine, and histidine).

258 Because of these substitutions (mainly due to the presence of Thr instead of Asn) these

259 proteins are likely to have different binding ability to biamines from 3Q6K_Llon1. In the last

260 binding positions of Llon3 and Lolm1 there are leucine and methionine, respectively, instead

261 of phenylalanine. Ptob2 contains a substitution in the first binding position, with threonine

262 replaced by serine. The distribution of YRPs in these groups partially corresponds with

263 branches in the phylogenetic tree (S2 Fig).

264

265 Tunnel in YRPs

266 In Xu et al. [26], YRP 3Q6K from *L. longipalpis* was described as a hollow barrel with two

267 possible entrances. Here, based on the models created using other sand fly YRPs and MOLE

268 analysis, we found a very similar hollow barrel in all proteins we studied (Table 3). It is

269 therefore likely that all studied proteins retain the ability to bind a ligand. The differences in a

270 length and minimal radius of the identified tunnels have limited relevance due to the fact that

271 they are calculated on theoretical models. On contrary, differences in hydrophobicity should

272 be less influenced by homology modeling process and could therefore indicate (where

273 present) changes in binding affinities among paralogs within species.

274

275

276 **Table 3. A comparison of tunnels in sand fly salivary YRPs.**

Identificator	Length	Min. Radius	Hydrophobicity	Ligand-side	Opposite-side
Laya1	38.320	1.769	-0.13	Positive	Negative
Laya2	38.707	1.813	-0.22	Positive	Negative
Laya3	41.045	0.962	-0.17	Positive	Negative
Laya4	41.295	1.785	-0.25	Neutral (negative)	Neutral (negative)

	Length	Radius	Hydrophobicity	Ligand-side	Opposite-side
Lint	39.194	1.539	-0.32	Negative	Negative
3Q6K_Llon1	41.245	1.892	-0.23	Positive	Neutral (negative)
Llon2	39.722	1.361	-0.07	Negative	Negative
Llon3	43.634	1.831	-0.29	Negative	Negative
Lolm1	38.595	2.055	-0.21	Negative (neutral)	Neutral
Lolm2	41.696	1.997	-0.45	Positive (neutral)	Neutral (negative)
Lolm3	43.098	1.416	-0.86	Neutral (negative)	Negative (neutral)
Para	45.294	1.229	-0.53	Negative	Negative
Parg	40.506	1.262	-0.20	Negative	Negative
Pari1	45.068	1.673	-0.02	Negative	Negative
Pari2	41.717	1.880	-0.29	Negative	Negative
Pdub1	37.292	1.744	-0.14	Negative	Neutral (negative)
Pdub2	39.912	1.433	0.03	Neutral (positive)	Neutral (negative)
Pdub3	35.967	2.021	-0.34	Negative	Neutral (positive, negative)
Pori1	42.203	1.677	-0.22	Negative	Negative
Pori2	32.368	1.824	-0.18	Negative	Neutral (negative)
Ppap1	38.163	1.834	-0.23	Negative	Neutral (positive)
Ppap2	40.508	1.780	-0.05	Neutral (negative)	Negative (neutral)
Ppap3	40.849	2.053	-0.08	Negative	Neutral (positive)
Ppap4	46.137	1.518	-0.27	Neutral	Neutral (negative, positive)
Pper1	41.652	1.419	-0.18	Negative	Negative
Pper2	41.944	1.948	-0.25	Neutral (negative)	Neutral (negative)
Pser1	50.023	2.047	-0.27	Negative	Neutral
Pser2	37.843	1.755	-0.05	Negative	Neutral
Pser3	38.013	1.757	0.02	Neutral	Negative
Pser4	31.949	1.812	0.01	Negative	Negative
Pser5	40.761	1.853	-0.01	Negative	Neutral
Ptob1	33.603	1.740	-0.31	Negative	Negative
Ptob2	34.118	1.822	-0.22	Neutral (positive)	Neutral (negative)

277   The length and minimum radius (in Ångströms) and hydrophobicity are shown for the tunnels

278   of each protein. All values were calculated using MOLE. Surface electrostatic potentials

279   around protein entry sides were calculated in PyMOL using the APBS tool. "Ligand-side"

280   refers to the side with the shorter distance from the beginning of the tunnel to the ligand-

281   binding site, while "Opposite-side" is the other side of the tunnel. The most common

282   electrostatic surface potentials of both entrances to the tunnel are shown. Charges in brackets

283   indicate that the charge around the entrance is not uniform, and small parts of the entrance are

284   charged differently from the majority of the entrance. Protein codes refer to Table 1.

285

286

287

288   The ligand-binding site for serotonin in 3Q6K is located approximately one third of

289   the way along this tunnel (Fig 4). The comparison based on the APBS tools in PyMol of all

290   models and the 3Q6K structure revealed that the surface electrostatic potential of the entrance

291   closer to the binding site varies from negative to positive (S2 Fig), and the size of this

292   entrance is smaller than the opposite entrance. In contrast, the entrance more distant from the

293   binding site is larger and neutrally or negatively charged in all these proteins (Table 3). The

294   distribution of the surface potentials in most of these proteins corresponds with the

295   phylogenetic relationships (Fig 1) and predicted glycosylation sites (Table 2).

296

297

298

299   **Fig 4. Visualization of the tunnel in 3Q6K_Llon1.** The structure was visualized in PyMOL

300   using the MOLE script for calculating the tunnel. The 3Q6K_Llon1 protein is drawn in green,

301   the tunnel through this protein structure is in black mesh and the red-stick molecule represents

302   serotonin.

303

304

305

306   The sequence conservation of the amino acids lining the tunnels was explored using

307   WebLogo server (Fig 5). The WebLogo was based on the multiple sequence alignment (S3

308   Fig) of all amino acids lining the accessible tunnels towards the ligand-binding site, as

309   described by MOLE. The analysis showed that there are not any absolutely conserved amino

310   acids, but both entrances to the tunnel are dominated by negatively-charged residues;

311   however, there are proteins (e.g. Pdub1, Pari2, Ppar1, Ppap4, and Pser5) that are positively

312   charged at least at one of the entrances. The middle section of the tunnel lining is occupied

313   mostly by hydrophobic and small polar amino acids.

314

315

316

317   **Fig 5. Sequence conservation of the amino acids lining the ligand-binding tunnel.** The

318   WebLogo server was used to visualize the sequence conservation of the amino acids lining

319   the ligand-binding tunnel in 31 YRPs. The amino acids are colored according to their

320   physical-chemical properties. LS and OS below the x-axis indicate the side closer to the

321   ligand-binding site and the opposite side of the tunnel, respectively.

322

323

324

325   **Discussion**

326

327   Here we focused on the bioinformatic characterization of sand fly salivary YRPs.

328   These proteins have high affinity binding properties for pro-haemostatic and pro-

329   inflammatory biogenic amines [26], but they are also highly antigenic for hosts bitten by sand

330   flies [30–34] and can be used for anti-*Leishmania* vaccine development [53]. We searched

331   published transcriptomes and recombinant proteins and found 33 different proteins from this

332   family among 13 species from the two genera – Old World *Phlebotomus* and New World

333   *Lutzomyia*. We analyzed their phylogenetic relationships, predicted glycosylation sites,

334   constructed 3D models, identified tunnels passing through these structures, described the

335   ligand-binding site, and compared their surface electrostatic potentials. The number of

336   detected YRPs in these 13 sand fly species varied from one in *P. arabicus* [14], *P. argentipes*

337   [12], and *L. intermedia* [17] to five in *P. sergenti* [16]. Such interspecies variability might be

338   partially attributed to the sensitivity of sequencing methods; however, this is not likely the

339   case of YRPs that were analyzed in the same laboratory using the same method. For example,

340   Anderson et al. [12] reported a single YRP in *P. argentipes* but two in *P. perniciosus*, while

341   Rohousova et al. [16] reported two in *P. tobbi* but five in *P. sergenti*. It is also possible that

342   occurence of various numbers of YRPs is caused by differential gene expression. Genes with

343   lower expression might be recorded less frequently as it was shown for other sand fly protein

344   families [54].

345   Our phylogenetic analysis suggested a higher variability in *Lutzomyia* than in

346   *Phlebotomus* YRPs. Sequence identity among YRPs ranged between 44 and 99 percent, with

347   a majority of proteins having sequence identity around 50 percent. Further analysis divided

348   these proteins into several branches, with closer relationships within sand flies from the same

349   subgenera (*Phlebotomus, Paraphlebotomus,* and *Larrousius*). This confirmed previous

350   analyses based on amino acid sequences [16, 20] and also corresponds with an analysis based

351   on small subunit nuclear ribosomal DNA [55] showing two main clades – *Phlebotomus* with

352   *Paraphlebotomus* subgenera clustering together and leaving other subgenera (e.g.
353   *Euphlebotomus, Larroussius, and Adlerius*) in the second clade.

354   The biogenic function of proteins may be affected by the presence or absence of
355   posttranslational modifications [56, 57]. For example, the antibody response against some
356   glycoproteins can be directed strictly to the glycan part [58]. In accordance with our previous
357   studies [16, 20, 54, 59], we showed both inter- as well as intra-species variability in the
358   glycosylation of sand fly YRPs, which corresponded well with the phylogenetic analysis.
359   Moreover, the same positions of putative glycosylation sites among YRPs mirrored the
360   presumed effects of glycosylation on protein folding and stability, suggesting a highly
361   conserved tertiary structure. It would be interesting to evaluate similarities in the antigenic
362   potential of the proteins belonging to the same clusters.

363   The only protein with an experimentally solved 3D structure, LJM11 (3Q6K) from
364   *Lutzomyia longipalpis*, has a unique six-bladed beta propeller fold that has only been
365   identified in this protein. It is characterized by a long ligand-binding tunnel with two
366   entrances – one on each side of the protein. Our 3D models of other YRPs showed similar
367   structures as 3Q6K, and the tunnel is present in all studied proteins. Its length or minimum
368   width vary between and within species but it remains uninterrupted in all modeled proteins,
369   suggesting it is very important feature of these proteins. LJM11 is known to bind positively
370   charged biogenic amines [26], and our structural analysis of YRPs from other sand fly species
371   showed that in all tested sequences the ligand-binding tunnels were negatively charged and
372   could therefore bind positively charged ligands of bioamine size. Nevertheless, we observed
373   that the proteins and more specifically the paralogs within one species differ in the charge of
374   tunnel entrances. Our data indicate that paralogs within one species will bind similar ligands
375   with different affinities, and the ligands may travel different paths to the binding site
376   according to the charge of the entrances. YRPs also differed within and among species in the
377   length of their tunnels, the minimum radius of the tunnel and its hydrophobicity. We expect
378   that all these parameters affect the nature and the affinity of the preferred ligand, which
379   results in a greater binding versatility of the host bioamines.

380   In the 3Q6K YRP structure, the serotonin ligand is stabilized by hydrogen bonds from
381   three amino acids in the main chain and side chain and by a number of additional hydrophobic
382   interactions [26]. Both our sequence and structure analyses showed the diversity of this
383   crucial site among the 33 proteins studied here. The largest diversity was observed on the
384   position denoted in 3Q6K as Phe 344 (see Fig 3). The amino acid at this position participates
385   in ligand-binding through a hydrogen bond from the main chain. Substitutions at this position

386 should not play a major role in ligand-binding, yet it is the only amino acid of the ligand-
387 binding-site that is different among three *L. longipalis* yellow-related proteins that have been
388 shown to have different affinities for serotonin and other tested bioamines [26]. Our models
389 of the other YRPs suggest that at least in the case of the Phe344Gln mutation, the glutamine
390 side chain could provide a further hydrogen bond to the ligand (serotonin). Interestingly, our
391 studied YRPs can be divided into several groups according to their ligand-binding
392 composition, and some groups correlate with the branches identified in phylogenetic
393 distribution as well as with the predicted glycosylation patterns. For example, Ppap2, Ppap4,
394 Pdub1, and Pser5 (proteins from the largest group with binding amino acids Thr-Asn-His) or
395 Pser1-4 (proteins from the second largest group with binding amino acids Thr-Asn-Glu) are
396 located in two branches in the phylogenetic tree and share the same N-glycosylation and O-
397 glycosylation pattern, respectively (see S2 Fig).

398   The lining of the tunnel is not conserved, but the entrances to the tunnel are composed
399 mostly of negatively charged residues and the inner part of the tunnel is lined by hydrophobic
400 and small polar amino acids (threonines). There are however several proteins (Pdub1, Ppap4,
401 Pser5 etc.) that have positively charged residues forming one of the entrances (typically the
402 one closer to ligand-binding site within the tunnel) – this might allow interaction with
403 different ligands.

404   In summary, we found interesting diversity among these important proteins. All 33
405 sand fly salivary YRPs studied here share similar folding and contain a ligand-binding tunnel.
406 Nevertheless, modifications found among these proteins influence the charge of entrances to
407 the tunnel, the length of the tunnel and/or its hydrophobicity. We suggest that these
408 modifications provide sand fly species with multiple paralogs of YRPs with more nuanced
409 answers to host bioamines. Further experiments are needed to validate these models, but we
410 assume, that our results should allow a better understanding of the biological role of these
411 important proteins with respect to their potential use in anti-*Leishmania* vaccines or as host
412 exposure markers to sand flies.

413

414 **Acknowledgements**

415

418

References

1. Ribeiro JMC, Mans BJ, Arca B. An insight into the sialome of bloodfeeding Nematocera. Insect Biochem Mol Biol. 2010;40: 767–784.

2. Abdeladhim M, Kamhawi S, Valenzuela JG. What's behind a sand fly bite? The profound effect of sand fly saliva on host hemostasis, inflammation and immunity. Infect Genet Evol. 2014;28: 691–703.

3. Rohousova I, Ozensoy S, Ozbel Y, Volf P. Detection of species-specific antibody response of humans and mice bitten by sand flies. Parasitology. 2005;130: 493–499.

4. de Moura TR, Oliveira F, Novais FO, Miranda JC, Clarencio J, Follador I, et al. Enhanced *Leishmania braziliensis* infection following pre-exposure to sandfly saliva. PLoS Negl Trop Dis. 2007;1: e84.

5. Vlkova M, Rohousova I, Drahota J, Stanneck D, Kruedewagen EM, Mencke N, et a. Canine antibody response to *Phlebotomus perniciosus* bites negatively correlates with the risk of *Leishmania infantum* transmission. PLoS Negl Trop Dis. 2011;5: e1344.

6. Gidwani K, Picado A, Rijal S, Singh SP, Roy L, Volfova V, et al. Serological markers of sand fly exposure to evaluate insecticidal nets against visceral leishmaniasis in India and Nepal: A cluster-randomized trial. PLoS Negl Trop Dis. 2011;5: e1296.

7. Marzouki S, Ben Ahmed M, Boussoffara T, Abdeladhim M, Ben Aleya-Bouafif N, Namane A, et al. Characterization of the antibody response to the saliva of *Phlebotomus papatasi* in people living in endemic areas of cutaneous leishmaniasis. Am J Trop Med Hyg. 2011;84: 653–661.

8. Charlab R, Valenzuela JG, Rowton ED, Ribeiro JMC. Toward an understanding of the biochemical and pharmacological complexity of the saliva of a hematophagous sand fly *Lutzomyia longipalpis*. Proc Natl Acad Sci USA. 1999;96: 15155–15160.

9. Valenzuela JG, Belkaid Y, Garfield MK, Mendez S, Kamhawi S, Rowton ED. Toward a defined anti-*Leishmania* vaccine targeting vector antigens: characterization of a protective salivary protein. J Exp Med. 2011;194: 331–342.

10. Valenzuela JG, Garfield M, Rowton ED, Pham, VM. Identification of the most abundant secreted proteins from the salivary glands of the sand fly *Lutzomyia longipalpis*, vector of *Leishmania chagasi*. J Exp Biol. 2004;207: 3717–3729.

11. Oliveira F, Kamhawi S, Seitz AE, Pham VM, Guigal PM, Fisher L, et al. From transcriptome to immunome: Identification of DTH inducing proteins from a *Phlebotomus ariasi* salivary gland cDNA library. Vaccine. 2006;24: 374–390.

452   12. Anderson JM, Oliveira F, Kamhawi S, Mans BJ, Reynoso D, Seitz AE, et al. Comparative
453   salivary gland transcriptomics of sandfly vectors of visceral leishmaniasis. BMC
454   Genomics. 2006;7: 52.

455   13. Kato H, Anderson JM, Kamhawi S, Oliveira F, Lawyer PG, Pham VM, et al. High degree
456   of conservancy among secreted salivary gland proteins from two geographically distant
457   *Phlebotomus duboscqi* sandflies populations (Mali and Kenya). BMC Genomics. 2006;7:
458   226.

459   14. Hostomska J, Volfova V, Mu J, Garfield M, Rohousova I, Volf P, et al. Analysis of
460   salivary transcripts and antigens of the sand fly *Phlebotomus arabicus*. BMC Genomics.
461   2009;10: 282.

462   15. Abdeladhim M, Jochim RC, Ben Ahmed M, Zhioua E, Chelbi I, Cherni S, et al. Updating
463   the salivary gland transcriptome of *Phlebotomus papatasi* (Tunisian Strain): The search
464   for sand fly-secreted immunogenic proteins for humans. PLoS One. 2012;7: e47347.

465   16. Rohousova I, Subrahmanyam S, Volfova V, Mu J, Volf P, Valenzuela JG, et al. Salivary
466   gland transcriptomes and proteomes of *Phlebotomus tobbi* and *Phlebotomus sergenti*,
467   vectors of leishmaniasis. PLoS Negl Trop Dis. 2012;6: e1660.

468   17. de Moura TR, Oliveira F, Carneiro MW, Miranda JC, Clarencio J, Barral-Netto M, et al.
469   Functional transcriptomics of wild-caught *Lutzomyia intermedia* salivary glands:
470   identification of a protective salivary protein against *Leishmania braziliensis* infection.
471   PLoS Negl Trop Dis. 2013;7: e2242.

472   18. Abrudan J, Ramalho-Ortigao M, O'Neil S, Stayback G, Wadsworth M, Bernard M, et al.
473   The characterization of the *Phlebotomus papatasi* transcriptome. Insect Mol Biol.
474   2013;22: 211–232.

475   19. Kato H, Jochim RC, Gomez EA, Uezato H, Mimori T, Korenaga M, et al. Analysis of
476   salivary gland transcripts of the sand fly *Lutzomyia ayacuchensis*, a vector of Andean-type
477   cutaneous leishmaniasis. Infect Genet Evol. 2013;13: 56–66.

478   20. Vlkova M, Sima M, Rohousova I, Kostalova T, Sumova P, Volfova V, et al. Comparative
479   analysis of salivary gland transcriptomes of *Phlebotomus orientalis* sand flies from
480   endemic and non-endemic foci of visceral leishmaniasis. PLoS Negl Trop Dis. 2014;8:
481   e2709.

482   21. Adeladhim M, Coutinho-Abreau IV, Townsend S, Pasos-Pinto S, Sanchez L, Rasouli M,
483   et al. Molecular diversity between salivary proteins from New World and Old World sand
484   flies with emphasis on *Bichromomyia olmeca*, the sand fly vector of *Leishmania mexicana*
485   in Mesoamerica. PLoS Negl Trop Dis. 2016;10: e0004771.

486 22. Nash WG, Kamerow HN, Merril CR. Identification of a yellow gene-specific protein in
487   *Drosophila melanogaster* by two-dimensional gel-electrophoresis. Biochem Genet.
488   1983;21: 1135–1142.

489 23. Hanes J, Simuth J. Identification and partial characterization of the major royal jelly
490   protein of the honey-bee (*Apis mellifera L*). J Apic Res. 1992;31: 22–26.

491 24. Johnson JK, Li J, Christensen BM. Cloning and characterization of a dopachrome
492   conversion enzyme from the yellow fever mosquito, *Aedes aegypti*. Insect Biochem Mol
493   Biol. 2001;31: 1125–1135.

494 25. Alves-Silva J, Ribeiro JMC, Van Den Abbeele J, Attardo G, Hao R, Haines LR, et al. An
495   insight into the sialome of *Glossina morsitans morsitans*. BMC Genomics. 2010;11: e213.

496 26. Xu X, Oliveira F, Chang BW, Collin N, Gomes R, Teixeira C, et al. Structure and
497   function of a ''yellow'' protein from saliva of the sand fly *Lutzomyia longipalpis* that
498   confers protective immunity against *Leishmania major* infection. J Biol Chem. 2011;286:
499   32383–32393.

500 27. Oliveira MCG, Pelegrini-Da-Silva A, Parada CA, Tambeli CH. 5-HT acts on nociceptive
501   primary afferents through an indirect mechanism to induce hyperalgesia in the
502   subcutaneous tissue. Neuroscience. 2007;145: 708–714.

503 28. Xanthos DN, Bennett GJ, Coderre TJ. Norepinephrine-induced nociception and
504   vasoconstrictor hypersensitivity in rats with chronic post-ischemia pain. Pain. 2008;137:
505   640–651.

506 29. Volf P, Skarupova S, Man P. Characterization of the lectin from females of *Phlebotomus*
507   *duboscqi* sand flies. Eur J Biochem. 2002;269: 6294–6301.

508 30. Teixeira C, Gomes R, Collin N, Reynoso D, Jochim RC, Oliveira F, et al. Discovery of
509   markers of exposure specific to bites of *Lutzomyia longipalpis*, the vector of *Leishmania*
510   *infantum* chagasi in Latin America. PLoS Negl Trop Dis. 2010;4: e638.

511 31. Souza AP, Andrade BB, Aquino D, Entringer P, Miranda JC, Alcantara R, et al. Using
512   recombinant proteins from *Lutzomyia longipalpis* saliva to estimate human vector
513   exposure in visceral leishmaniasis endemic areas. PLoS Negl Trop Dis. 2010;4: e649.

514 32. Martin-Martin I, Molina R, Rohousova I, Drahota J, Volf P, Jimenez M. High levels of
515   anti-*Phlebotomus perniciosus* saliva antibodies in different vertebrate hosts from the re-
516   emerging leishmaniosis focus in Madrid, Spain. Vet Parasitol. 2014;202: 207–216.

517 33. Kostalova T, Lestinova T, Sumova P, Vlkova M, Rohousova I, Berriatua E, et al. Canine
518   antibodies against salivary recombinant proteins of *Phlebotomus perniciosus*: A

519   longitudinal study in an endemic focus of canine leishmaniasis. PLoS Negl Trop Dis.
520   2015;9: e0003855.

521   34. Sima M, Ferencova B, Warburg A, Rohousova I, Volf P. Recombinant salivary proteins
522   of *Phlebotomus orientalis* are suitable antigens to measure exposure of domestic animals
523   to sand fly bites. PLoS Negl Trop Dis. 2016;10: e0004553.

524   35. Gomes R, Oliveira F, Teixeira C, Meneses C, Gilmore DC, Elnaiem DE, et al. Immunity
525   to sand fly salivary protein LJM11 modulates host response to vector-transmitted
526   *Leishmania* conferring ulcer-free protection. J Invest Dermatol. 2012;132: 2735–2743.

527   36. Collin N, Gomes R, Teixeira C, Cheng L, Laughinghouse A, Ward JM, et al. Sand fly
528   salivary proteins induce strong cellular immunity in a natural reservoir of visceral
529   leishmaniasis with adverse consequences for *Leishmania*. PLoS Pathog. 2009;5:
530   e1000441.

531   37. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search
532   tool. J Mol Biol. 1990;215: 403–410.

533   38. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal
534   peptides from transmembrane regions. Nat Methods. 2001;8: 785–786.

535   39. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al.
536   Clustal W and clustal X version 2.0. Bioinformatics. 2007;23: 2947–2948.

537   40. Abascal F, Zardoya R, Posada D. ProtTest: selection of best-fit models of protein
538   evolution. Bioinformatics. 2005;21: 2104–2105.

539   41. Schmidt HA, Strimmer K, Vingron M, von Haeseler A. TREE-PUZZLE: maximum
540   likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics.
541   2002;18: 502–504.

542   42. Goldman N, Whelan S. Statistical tests of gamma-distributed rate heterogeneity in models
543   of sequence evolution in phylogenetics. Mol Biol Evol. 2000;17: 975–978.

544   43. Tamura K, Dudley J, Ne, M, Kumar S. MEGA4: Molecular evolutionary genetics analysis
545   (MEGA) software version 4.0. Mol Biol Evol. 2007;24: 1596–1599.

546   44. Sievers F, Wilm A, Dineen DG, Gibson TJ, Karplus K, Li W, et al. Fast, scalable
547   generation of high-quality protein multiple sequence alignments using Clustal Omega.
548   Mol Syst Biol. 2011;7: 539.

549   45. Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT,
550   et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell
551   technology. EMBO J. 2013;32: 1478–1488.

552   46. Julenius K. NetCGlyc 1.0: Prediction of mammalian C-mannosylation sites.
553   Glycobiology. 2007;17: 868–876.

554   47. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein
555   Data Bank. Nucleic Acids Res. 2000;28: 235–242.

556   48. Sali A, Overington JP. Derivation of rules for comparative protein modeling from a
557   database of protein-structure alignments. Protein Sci. 1994;3: 1582–1596.

558   49. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems:
559   application to microtubules and the ribosome. Proc Natl Acad Sci USA. 2001;98: 10037–
560   10041.

561   50. Sehnal D, Svobodova Varekova R, Berka K, Pravda L, Navratilova V, et al. MOLE 2.0:
562   advanced approach for analysis of biomacromolecular channels. J Cheminform. 2013;5:
563   39.

564   51. Cid H, Bunster M, Canales M, Gazitua F. Hydrophobicity and structural classes in
565   proteins. Protein Eng Des Sel. 1992;5: 373–375.

566   52. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: A sequence logo generator.
567   Genome Res. 2004;14: 1188–1190.

568   53. Fiuza JA, Dey R, Davenport D, Abdeladhim M, Meneses C, Oliveira F, et al. Intradermal
569   immunization of *Leishmania donovani* centrin knock-out parasites in combination with
570   salivary protein LJM19 from sand fly vector induces a durable protective immune
571   response in hamsters. PLoS Negl Trop Dis. 2016;10: e0004322.

572   54. Rohousova I, Volfova V, Nova S, Volf P. Individual variability of salivary gland proteins
573   in three *Phlebotomus* species. Acta Trop. 2012; 122, 80-86.

574   55. Aransay AM, Scoulica E, Tselentis Y, Ready PD. Phylogenetic relationships of
575   Phlebotomine sandflies inferred from small subunit nuclear ribosomal DNA. Insect Mol
576   Biol. 2000;9: 157–168.

577   56. Hofinger ESA, Spickenreither M, Oschmann J, Bernhardt G, Rudolph R, Buschauer A.
578   Recombinant human hyaluronidase Hyal-1: insect cells versus *Escherichia coli* as
579   expression system and identification of low molecular weight inhibitors. Glycobiology.
580   2007; 17: 444–453.

581   57. Reddy VB, Kounga K, Mariano F, Lerner EA. Chrysoptin is a potent glycoprotein IIb/IIIa
582   fibrinogen receptor antagonist present in salivary gland extracts of the deerfly. J Biol
583   Chem. 2000;275: 15861–15867.

584   58. Atta A, Colossi R, Sousa-Atta MLB, Jeronimo SMB, Nascimento MDSB, Bezerra GF, et
585   al. Antileishmanial IgG and IgE antibodies recognize predominantly carbohydrate

586   epitopes of glycosylated antigens in visceral leishmaniasis. Mem Inst Oswaldo Cruz.

587   2004;99: 525–530.

588 59. Volf P, Tesarova P, Nohynkova E. Salivary proteins and glycoproteins in phlebotomine

589   sandflies of various species, sex and age. Med Vet Entomol. 2000;14: 251–256.

590

Supporting Information

592

S1 Table. Percent Identity Matrix of sand fly yellow-related proteins.

Clustal Omega was used to calculate this sequence identity matrix among all identified yellow-related proteins. The similarity between the proteins is shown in percents. Different colors represent 10 percent similarity intervals. Protein codes refer to Table 1.

597

S1 Fig. Modeled 3D structures of sand fly yellow-related proteins. Zipped PDB files of 32 sand fly yellow-related proteins calculated in MODELLER based on 3Q6K as the template structure. Protein codes refers to Table 1.

601

S2 Fig. Visual summary of the diversity of yellow-related proteins in sand flies. The maximum likelihood phylogenetic tree was created in TREEPUZZLE with WAG model using quartet puzzling with 10000 puzzling steps. For rooting, the related protein (ACCN: NP650247) from *Drosophila melanogaster* (Dmel) was used. Bootstraps with support for branching are shown. Protein codes refer to Table 1. Symbols preceding the protein codes indicate the most common surface electrostatic potential of the entrance to the protein tunnel closer to the ligand-binding site (LS) as shown in Table 3 (◊ positive, ○ neutral, and □ negative). Colors in symbols indicate the affiliation to protein groups with the same ligand-binding amino acids as shown in Figure 3. The letters N, O, and C indicate putative N-, O-, and C-glycosylation sites, respectively, based on Table 2.

612

S3 Fig. Alignment of amino acids creating tunnels in yellow-related proteins. For each protein, amino acids creating tunnel were determined using MOLE. Clustal Omega was used to visualize differences of these protein tunnels. In sequences, small letters indicate main chain interaction and capital letters represent side chain interactions. Above the sequences, LS and OS describe the side closer to the ligand binding site and the opposite side of the tunnel, respectively. Protein codes refer to Table 1.
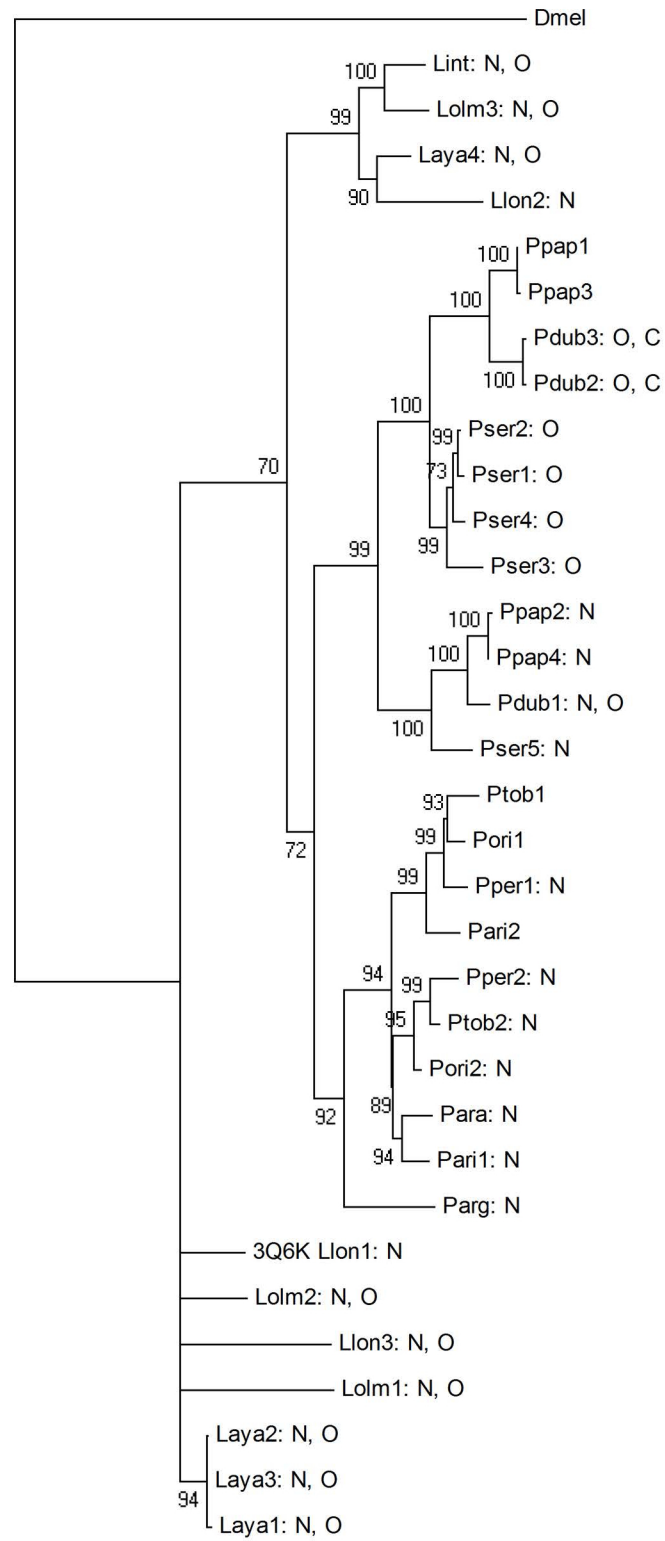
Figure 1

Figure 2

```
                  27    33     91    98    183   189   229   235   284   290   331      340   348     356
                  ...|...     ....|...    ..|...     .|....|    .|....|    ....|....|    ..|....|.
Lint         NVPTAFA     VSAYQPVI    ITNEGEN    VGLFGIT    TDAIALA    RFIFGTDISI    FLSNGYPPI
Lolm1        ILPTAFA     TSIYQPVI    ITNFDTN    FGLFAIT    SDAIALV    DFIFMTDISV    FMANGMPPI
Lolm2        NIPTAFA     TSVYQPVI    VTNFLSN    AGIFGIT    SDAIALV    KFIFGTDISV    FMSNGFPPV
Lolm3        NVPTAFA     VPVYQPVI    ITNFEEN    VGIFGIT    TDAIALA    RFIFGTDISI    FLSNGYPPI
Para         TIPTAFA     INVYQPVI    ITNFEEN    VGIFGIT    TEAIALA    KFIYGTDISI    FLSNGHPPV
Parg         SIPTGFA     ISIYQPVI    ITNYEEN    VGIFGIT    TEAITLV    RFLFGTDISV    VMANGHPPV
Pari1        SIPTAFV     TSVYQPML    ITNFEDN    AGIFGIT    TEAIALV    KFIFGTDIQI    FLSNGQPPI
Pari2        SIPTAFA     ISVYQPVI    ITNFEDN    LGIFGIT    SEAIALA    EFIFGTDISV    FFSNGHPPI
Pdub3        IIPTGVA     TSVYQPVI    IANFVEN    TGIFGIA    TEAIALA    KLNFATDMMV    FMSNGQPPF
Pdub1        VIPTGVT     TSIYQPVI    ITNFIDN    VGIFGIT    TEAIALA    YFVFGTDIMV    FMSNAHPPT
Pdub2        IIPTGVA     TSVYQPVI    IANFVEN    TGIFGIA    TEAIALA    KLNFATDMMV    FMSNGQPPF
Pori1        SIPTAFA     ISVYQPVI    IANFNDN    LGIFGIT    SEAIALA    DFIFGTDISV    FFSNGHPPI
Pori2        SIPTAFA     ISVYQPVI    ITNFEDN    VGIFGIA    SEAIALA    KFIFGTDISV    FLSNGQPPI
Ppap2        VIPTGVT     TSIYQPVI    ITNFIDN    AGIFGIT    TEAIALA    YFVFGTDIMV    FMSNAHPPT
Ppap4        VIPTGVT     TSIYQPVI    ITNFIDN    AGIFGIT    TEAIALA    YFVFGTDIMV    FMSNAHPPT
Ppap1        IVPTGVA     TSVYQPVI    IANFDEN    AGIFGIA    TEAIALA    NFNFGTDIMV    FMSNGQPPI
Ppap3        IVPTGVA     TSVYQPVI    IANFDEN    AGIFGIA    TEAIALA    NFNFGTDIMV    FLSNGQPPI
Pper1        SIVTGFA     ISVYQPVI    ITNFEDN    LGIFGIT    SEAIALA    DFIFGTDISM    FFSNGHPPI
Pper2        IIPTGFV     INVYQPVI    ITNFEDN    VGLFGIT    TEAVALV    KFIYGSDISV    FLSTGHPPI
Pser5        IIPTGVT     TSIYQPVI    ITNFIDN    TGIFGIA    TEAIALA    YFVFGTDIMV    FMANGHPPI
Pser4        VIPTGVA     TSVYQPVI    IANFVEN    AGIFGIA    TEAIALA    RFNFGTDIMV    FMANGQPPI
Pser3        VIPTGVV     TSVYQPVI    IANFVEN    AGIFGIT    TEAIALA    DFNFGTDLMV    FMANGQPPI
Pser2        VIPTGVA     TSVYQPVI    IANFVEN    AGIFGIA    TEAIALA    RFNFGTDIMV    FMANGQPPI
Pser1        VIPTGVA     TSVYQPVI    IANFVEN    AGIFGIA    TEAIALA    RFNFGTDIMV    FMANGQPPI
Ptob2        IIPTGFA     INVYQPVI    ITNFDDN    VGIFGIT    SEAVALA    KFIYGSDISV    FLSNGQPPI
Ptob1        GIATGFA     ISVYQPVI    IANFVDN    LGIFGIT    SEAIALA    DFIFGTDISV    FFSNGHPPI
Laya4        NIPTAFA     VSVYQPVI    ITNFEEN    VGLFGIT    TDAVALA    RFIFGTDISV    FMSNGYPPI
Laya2        NIPTAFA     TSVYQPVI    ITNFLSN    AGLFGIT    NDAIALA    RFVFGTDISV    FMSNGFPPV
Laya1        NIPTAFA     TSVYQPVI    ITNFLSN    AGLFGIT    NDAIALA    RFVFGTDISV    FMSNGFPPV
Laya3        NIPTAFA     TSVYQPVI    ITNFLSN    AGLFGIT    NDAIALA    RFVFGTDISV    FMSNGFPPV
Llon3        NILTAFA     TSVYQPVI    ITNFLTN    AGIFGIT    TDAIGLV    DFSFGTDISI    FLANGLPPL
Llon2        NIPTGLA     VNVYQPVI    LTNFKDN    VGLFGIA    TDAIALA    RFTFGTDILV    IMANGHPPV
3Q6K_Llon1   MISTAFA     TSIYQPVI    ITNFLRG    AGIFGIT    NDAIALA    RFVFGTDISV    FMSNGFPPI
                                                                         *              *  *
                  (T)          (YQ)       (F/Y)      (F)       (I/L/V)    (Y/F) (TD/SD)  (N/T) (H/Q/Y/
                                                                                              F/L/M)
```

Figure 3

threonine, 327

asparagine, 342

phenylalanine, 344

● 3Q6K_Llon1
● Pdub1
● Pser2
● Pper2
● Llon3
● Laya4
● Ptob2
● Lolm1

Figure 4

Figure 5

Table S1

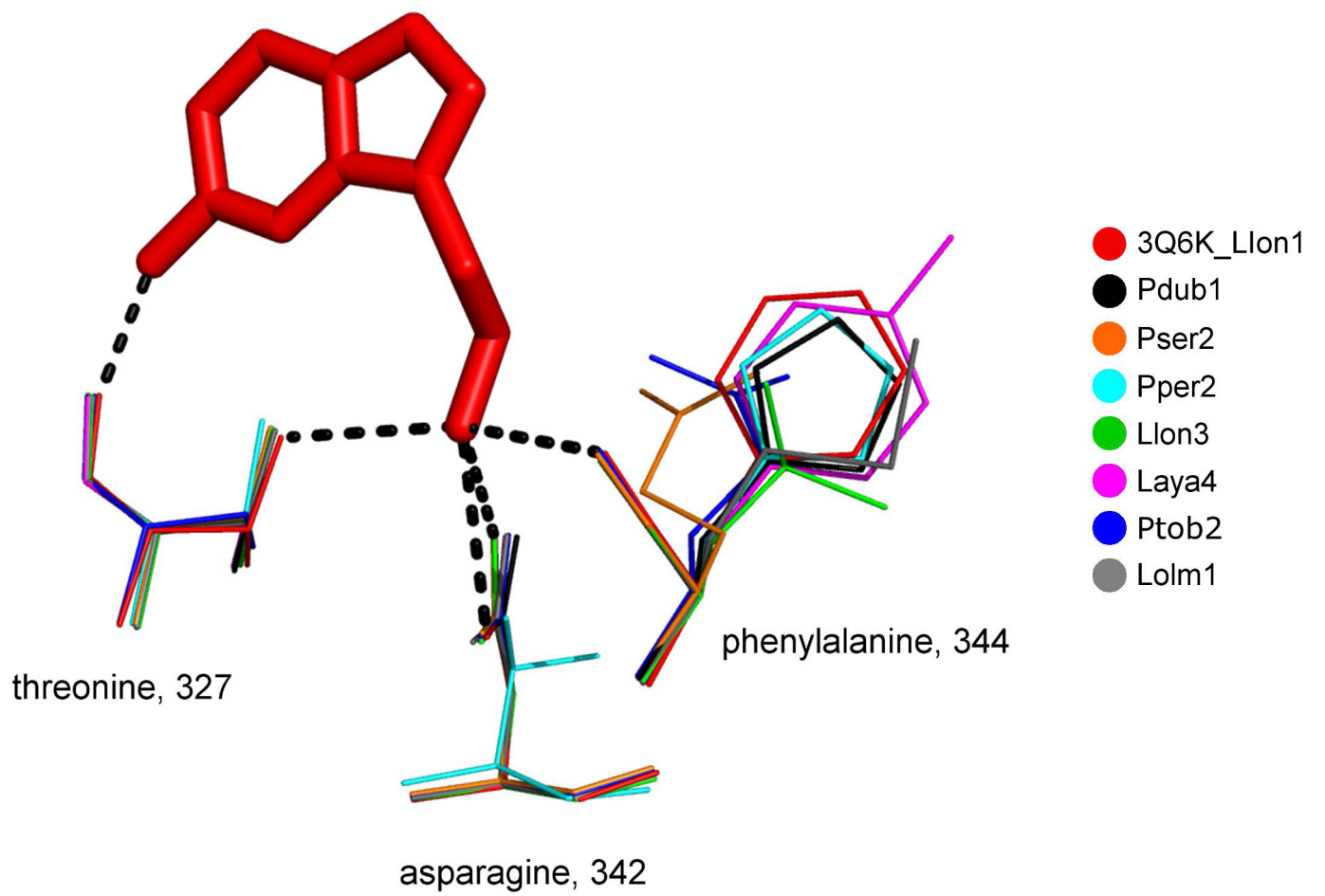|	Parg	Pari2	Ptob1	Pori1	Pper1	Pori2	Pper2	Ptob2	Para	Pari1	Pser5	Pdub1	Ppap2	Ppap4	Pser3	Pser4	Pser2	Pser1	Ppap1	Ppap3	Pdub3	Pdub2	Llon2	Laya4	Lint	Lolm3	Llon3	Lolm1	3Q6K_Llon1	Lolm2	Laya1	Laya2	Laya3
Parg	100.00	57.84	57.84	58.81	58.11	63.27	62.57	60.96	64.97	63.64	54.93	55.47	56.38	56.91	54.45	53.64	53.62	53.62	48.66	53.85	54.11	55.97	48.28	52.16	51.89	51.35	54.72	55.53	53.64	54.45	54.13	54.13	54.13
Pari2	57.84	100.00	80.54	84.01	82.43	78.38	74.59	70.54	78.11	77.84	53.26	51.89	52.16	52.70	56.68	56.95	53.66	53.12	49.59	49.86	53.39	53.55	45.80	52.46	52.46	53.55	56.40	56.40	56.95	56.68	52.03	52.03	52.03
Ptob1	57.84	80.54	100.00	87.26	87.30	75.68	72.16	72.16	72.70	72.16	51.90	50.27	50.27	50.55	55.04	54.50	54.22	51.76	51.23	49.05	50.95	50.55	44.17	50.55	50.55	50.95	54.22	54.50	54.50	55.04	51.49	51.49	51.49
Pori1	58.81	84.01	87.26	100.00	88.35	77.78	74.53	74.53	75.61	76.76	53.13	50.95	50.95	50.86	55.86	56.13	55.86	53.26	52.86	49.73	52.45	51.64	45.38	51.91	51.64	52.19	55.59	56.13	55.86	55.86	52.45	52.45	52.45
Pper1	58.11	82.43	87.30	88.35	100.00	77.03	75.14	75.14	76.76	76.76	54.35	51.62	52.16	52.16	56.40	56.13	56.13	55.01	50.00	49.32	52.85	53.01	46.34	53.01	53.01	52.46	56.13	56.40	56.13	56.16	51.49	51.49	51.49
Pori2	63.27	78.38	75.68	77.78	77.03	100.00	86.86	83.65	84.45	81.55	57.77	54.72	56.18	57.22	56.68	55.74	56.68	57.77	50.68	52.73	58.60	59.14	49.46	55.19	55.46	55.19	54.50	54.60	56.68	55.46	54.32	54.32	54.32
Pper2	62.57	74.59	72.16	74.53	75.14	86.86	100.00	90.13	83.65	80.21	53.26	54.35	56.13	56.18	55.43	54.35	54.35	57.07	50.95	51.76	59.88	56.57	49.33	55.74	56.72	51.23	54.35	54.35	53.26	54.35	53.10	53.10	53.10
Ptob2	60.96	70.54	72.16	74.53	75.14	83.65	90.13	100.00	81.82	81.55	54.35	55.65	55.50	55.50	54.89	54.35	54.35	55.98	51.23	52.30	56.84	56.84	47.45	53.13	53.23	53.95	55.43	55.43	54.89	54.35	52.56	52.56	52.56
Para	64.97	78.11	72.70	75.61	76.76	84.45	83.65	81.82	100.00	81.82	57.07	53.49	53.35	53.35	55.43	54.50	54.50	57.07	50.00	52.85	58.29	56.68	48.93	54.22	54.77	54.50	57.07	57.07	55.98	55.43	53.64	53.64	53.64
Pari1	63.64	77.84	72.16	76.76	76.76	81.55	80.21	81.55	81.82	100.00	56.30	53.49	54.30	54.30	54.89	54.35	54.77	57.07	50.27	52.59	55.61	55.50	48.40	53.95	54.22	53.95	55.98	55.98	54.89	54.35	52.56	52.56	52.56
Pser5	54.93	53.26	51.90	53.13	54.35	57.77	53.26	54.35	57.07	56.30	100.00	76.78	77.11	77.63	66.67	66.40	75.53	75.00	75.53	75.27	75.27	74.73	49.47	48.81	51.99	51.34	48.25	49.60	49.87	48.38	53.42	54.47	53.60
Pdub1	55.47	51.89	50.27	50.95	51.62	54.72	54.35	55.65	53.49	53.49	76.78	100.00	87.66	88.45	64.40	64.53	62.60	62.86	74.47	74.47	75.00	75.00	48.92	48.94	51.74	50.94	48.39	50.67	49.74	48.11	49.46	50.27	55.17
Ppap2	56.38	52.16	50.27	50.95	52.16	56.18	56.13	55.50	53.35	54.30	77.11	87.66	100.00	98.95	59.95	60.21	62.07	62.33	57.18	57.18	56.12	57.18	48.94	48.94	51.88	50.54	44.05	48.25	50.67	48.38	50.55	50.27	54.38
Ppap4	56.91	52.70	50.55	50.86	52.16	57.22	56.18	55.50	53.35	54.30	77.63	88.45	98.95	100.00	60.48	60.74	62.07	62.33	57.45	57.18	56.12	57.98	48.94	48.94	52.12	51.32	44.05	47.17	51.21	48.25	50.81	50.27	54.38
Pser3	54.45	56.68	55.04	55.86	56.40	56.68	55.43	54.89	55.43	54.89	66.67	64.40	59.95	60.48	100.00	89.66	87.27	87.00	69.41	72.61	70.48	72.61	50.21	51.61	54.16	51.74	47.98	50.94	50.67	49.87	54.96	55.23	55.50
Pser4	53.64	56.95	54.50	56.13	56.13	55.74	54.35	54.35	54.50	54.35	66.40	64.53	60.21	60.74	89.66	100.00	94.96	94.16	72.87	72.61	72.87	72.61	50.94	54.42	54.16	53.17	47.75	50.67	52.80	49.06	55.23	55.23	55.23
Pser2	53.62	53.66	54.22	55.86	56.13	56.68	54.35	54.35	54.50	54.77	75.53	62.60	62.07	62.07	87.27	94.96	100.00	97.61	75.00	75.27	75.00	75.00	51.74	55.50	53.62	53.08	46.90	50.67	52.30	50.00	55.53	56.03	56.03
Pser1	53.62	53.12	51.76	53.26	55.01	57.77	57.07	55.98	57.07	57.07	75.00	62.86	62.33	62.33	87.00	94.16	97.61	100.00	74.47	75.00	75.00	74.73	51.74	54.42	54.18	52.55	47.17	49.60	52.55	49.60	54.72	55.76	55.76
Ppap1	48.66	49.59	49.05	49.73	50.00	50.68	50.95	51.23	50.00	50.27	75.53	74.47	57.18	57.45	69.41	72.87	75.00	74.47	100.00	99.20	83.78	84.04	57.71	57.18	60.16	54.50	53.95	51.23	55.19	52.46	50.55	53.55	50.27
Ppap3	53.85	49.86	49.05	49.73	49.32	52.73	51.76	52.30	52.85	52.59	75.27	74.47	57.18	57.18	72.61	72.61	75.27	75.00	99.20	100.00	83.51	83.24	57.18	57.18	59.36	54.77	54.22	51.50	55.46	52.46	50.55	53.55	50.00
Pdub3	54.11	53.39	50.95	52.45	52.85	58.60	59.88	56.84	58.29	55.61	75.27	75.00	56.12	56.12	70.48	72.87	75.00	75.00	83.78	83.51	100.00	98.41	51.34	51.61	59.63	54.22	53.13	52.04	55.74	52.46	50.55	51.89	51.61
Pdub2	55.97	53.55	50.55	51.64	53.01	59.14	56.57	56.84	56.68	55.50	74.73	75.00	57.18	57.98	72.61	72.61	75.00	74.73	84.04	83.24	98.41	100.00	51.61	51.61	59.89	54.22	52.86	51.77	55.19	52.16	50.55	51.61	51.61
Llon2	48.28	45.80	44.17	45.38	46.34	49.46	49.33	47.45	48.93	48.40	49.47	48.92	48.94	48.94	50.21	50.94	51.74	51.74	57.71	57.18	51.34	51.61	100.00	68.22	63.14	49.07	49.19	48.52	49.46	49.74	48.28	48.28	48.28
Laya4	52.16	52.46	50.55	51.91	53.01	55.19	55.74	53.13	54.22	53.95	48.81	48.94	48.94	48.94	51.61	54.42	55.50	54.42	57.18	57.18	51.61	51.61	68.22	100.00	68.22	51.61	51.34	51.08	52.12	52.25	54.74	54.47	55.79
Lint	51.89	52.46	50.55	51.64	53.01	55.46	56.72	53.23	54.77	54.22	51.99	51.74	51.88	52.12	54.16	54.16	53.62	54.18	60.16	59.36	59.63	59.89	63.14	68.22	100.00	74.68	63.14	51.88	51.59	54.18	53.12	54.47	55.79
Lolm3	51.35	53.55	50.95	52.19	52.46	55.19	51.23	53.95	54.50	53.95	51.34	50.94	50.54	51.32	51.74	53.17	53.08	52.55	54.50	54.77	54.22	54.22	49.07	51.61	74.68	100.00	72.61	61.76	50.81	54.18	53.39	53.85	55.53
Llon3	54.72	56.40	54.22	55.59	56.13	54.50	54.35	55.43	57.07	55.98	48.25	48.39	44.05	44.05	47.98	47.75	46.90	47.17	53.95	54.22	53.13	52.86	49.19	51.34	63.14	72.61	100.00	49.60	47.75	44.30	49.59	48.86	63.06
Lolm1	55.53	56.40	54.50	56.13	56.40	54.60	54.35	55.43	57.07	55.98	49.60	50.67	48.25	47.17	50.94	50.67	50.67	49.60	51.23	51.50	52.04	51.77	48.52	51.08	51.88	61.76	49.60	100.00	54.40	49.87	50.82	50.95	62.86
3Q6K_Llon1	53.64	56.95	54.50	55.86	56.13	56.68	53.26	54.89	55.98	54.89	49.87	49.74	50.67	51.21	50.67	52.30	52.30	52.55	55.19	55.46	55.74	55.19	49.46	52.12	51.59	50.81	47.75	54.40	100.00	57.98	60.05	60.05	79.74
Lolm2	54.45	56.68	55.04	55.86	56.16	55.46	54.35	54.35	55.43	54.35	48.38	48.11	48.38	48.25	49.87	49.06	50.00	49.60	52.46	52.46	52.46	52.16	49.74	52.25	54.18	54.18	44.30	49.87	57.98	100.00	50.27	50.27	79.68
Laya1	54.13	52.03	51.49	52.45	51.49	54.32	53.10	52.56	53.64	52.56	53.42	49.46	50.55	50.81	54.96	55.23	55.53	54.72	50.55	50.55	50.55	50.55	48.28	54.74	53.12	53.39	49.59	50.82	60.05	50.27	100.00	98.43	78.42
Laya2	54.13	52.03	51.49	52.45	51.49	54.32	53.10	52.56	53.64	52.56	54.47	50.27	50.27	50.27	55.23	55.23	56.03	55.76	53.55	53.55	51.89	51.61	48.28	54.47	54.47	53.85	48.86	50.95	60.05	50.27	98.43	100.00	78.89
Laya3	54.13	52.03	51.49	52.45	51.49	54.32	53.10	52.56	53.64	52.56	53.60	55.17	54.38	54.38	55.50	55.23	56.03	55.76	50.27	50.00	51.61	51.61	48.28	55.79	55.79	55.53	63.06	62.86	79.74	79.68	78.42	78.89	100.00

Figure S2

Figure S3

```
                        OS                                              LS
Lolm3           -----EAidyalRiTVSfii-lQgADfigFT-fYFPYRTEKE-------
Pper1           ------dEmiPRylTVSfii-lQgADflFgy-YTFPHLiEeE-------
Pser3           -----DPvdVyiMlVAtifTQllgAgfitIgFyYTFPKLipK-------
Llon2           -----DPVdyRLilvVAfaiDQggyYFfF--------vRVEEveEK---
Parg            ----iERPevylTAfVVaSCiQDTtfgcFtT-yYfYhHRpP-vDLd---
Llon3           ---EPdseiiRylTVVSfiliQDgfgitFTt-TlKYP-LFFElpEY---
Pdub2           -----DPVdyil--MVAfaigQDAftIFgty-T--YTPFKf--------
Lolm1           -----ESVdiylVTvTVfiSilDgAQftgFt-TTIYmKPFFEipd----
Laya2           -----DPVdiyT--VSvfiilQDgAftFgIT-tTfPKFYvpk-------
Lolm2           -----DPVdiyT--VSvV-fiQDgAftFgtT-YfFFKPvYpdD------
Pper2           -----EPvdyli--TVVSfi-QDgAsfgFSt-TYfFHPREPiIN-----
Laya3           -----DPVdiylTVvSAifi-lDQggFfYFP-FKvYk------------
Laya1           -----DPvdyil--VSTifi-lDgAQftIgF-fYFPFKYvpK------
Para            -----EPdsiyilSTTVfii-lQgADfitFt-TITYhHRpPpvE-----
Pori1           -----EPsdvVi-yVTSfii-lDgAQflLFt-tTTFhPYHFELpieE--
3Q6K_Llon1      -----DPvdyli--TVSfii-lDgAQfitFt-TITfFYFPKYp-irR--
Ppap3           -----DPVdiylVMvAAfii-lDgAQtfIgT-FyFPqYIpEKiDLD---
Laya4           -----DPVKiyl-TVSvfii-lDgAQfvg-V-tFtTTYfPYFRpiEK--
Ptob1           -----EPvdyil-TVSvfii-lDgAQtfg-F-tyTYfFl----------
Ppap2           -----DPVd-iy-TVMvfii-lDgAQfFg-T-fYFlRDLPteIE-----
Ptob2           -----EPVd-iy-TVSvfii-lDgAQsfg-F-TfYFiRI----------
Lint            ---El-RySIiV--iStTiyilDgAQftIFg-T--fYFYPRp--iEnN-
Pori2           --------ilyV--SVAifaliDgAQfFgYf-F--RQPLLipdD-----
Pser5           -----DRRPlLv--MAaifaliDgAQftgFT-f--YFPHLiDLIdD---
Pari1           EPKRldelyiMV--tTiTfQligDAQftFgT-f--YFiIPLED-LKen-
Ppap4           ---DR-RPlMyT--LA-itfligDAQftIgF-t--ThPHRpptPleEdL
Pari2           ---ER-RPlyiV--TVTSifliDgAQftFtT-T--hHPYFLLEipeEE-
Pdub1           ---DR-RPlLvM--TAitfaligDQAftFtT-T--hYPHFL--------
Pdub3           ----idRPvVyl--MAA-limfDgAQtfFtT-TqPQYKfDp--------
Ppap1           -----DRRPlyM--LAA-liifDgAQtfFTI-FqPY-QIppEi--d---
Pser2           --------DlRy--MVA-ilfiDgAtQf-gF-T-yYFPQiLVd--D---
Pser1           DKEsPdvViyVM--AlA-ifliDgAQtfFgy-Y--yFPQKiDp------
Pser4           --------DvLM--Ala-ifliDgAQtfIgF-TtTYyPFQIi-------
```
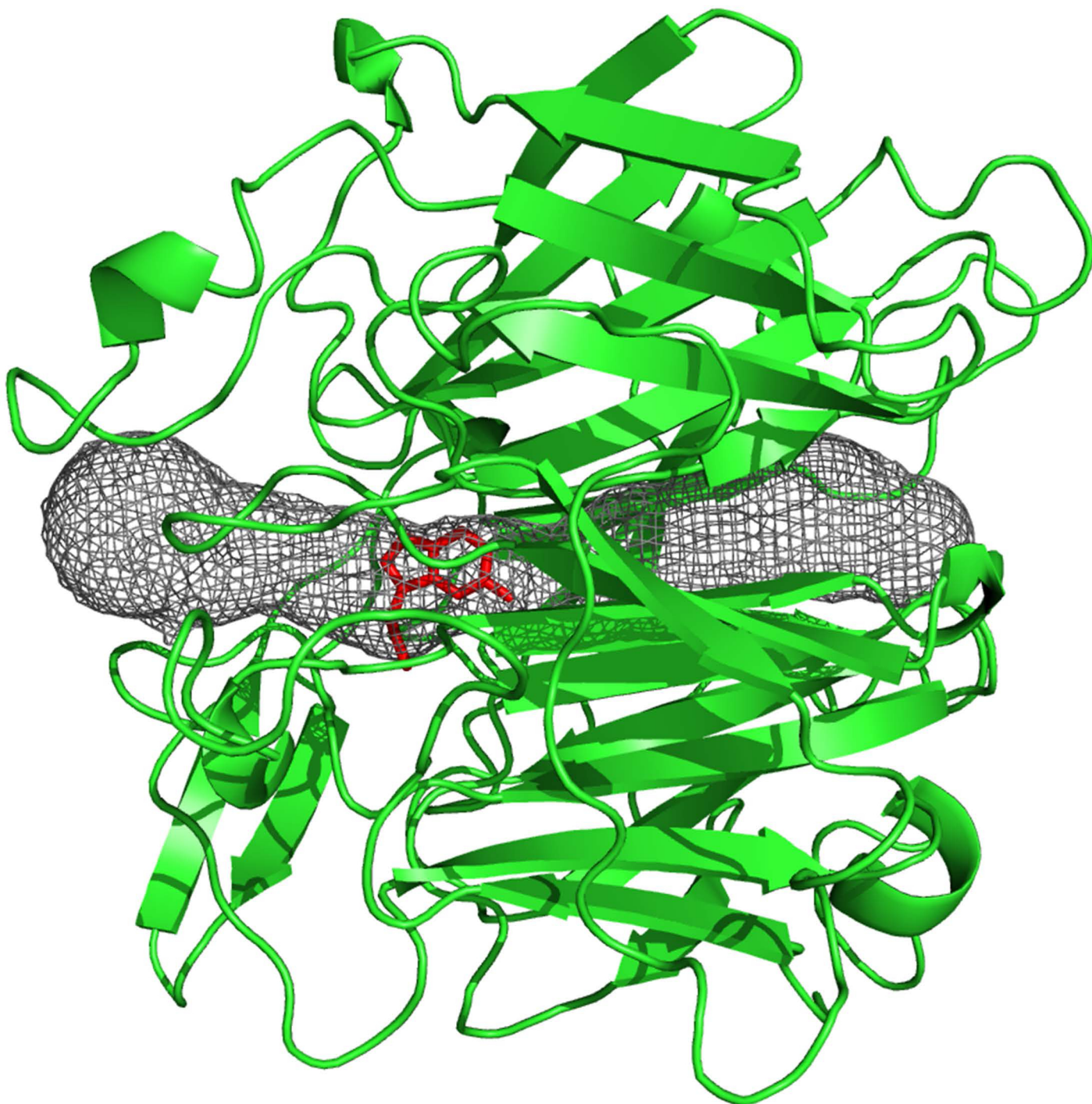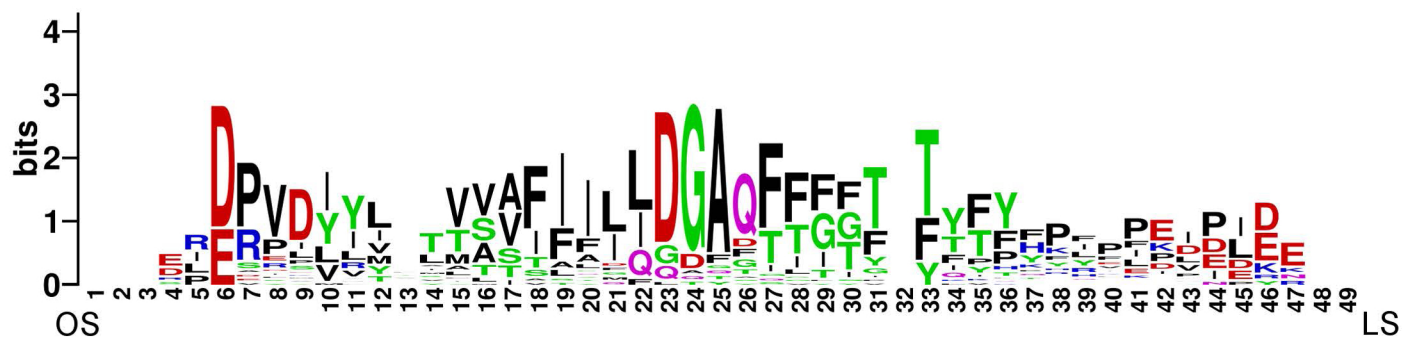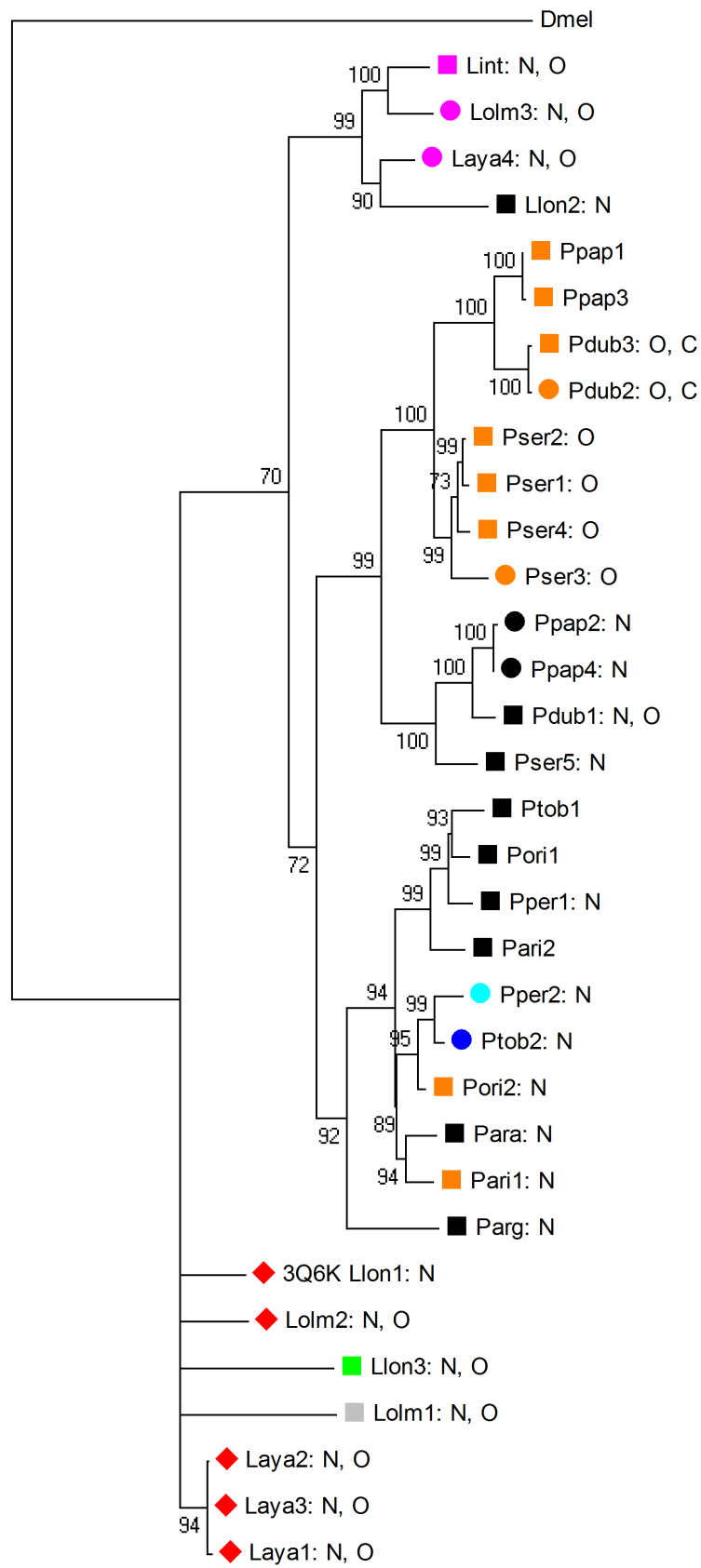
DISCUSSION

Three original research articles are included in this thesis. First two presented studies are strictly focused on *Phlebotomus orientalis*, the important vector of *Leishmania donovani* in East African countries (reviewed in Elnaiem et al., 2011), and its main salivary proteins and antigens. The third study deals with salivary yellow-related proteins from different sand fly species, including those from *P. orientalis*. It has been previously shown, that these proteins are highly antigenic (e.g. Marzouki et al., 2011), can be used in recombinant form for detecting anti-sand fly saliva antibodies (e.g. Kostalova et al., 2015), are the most promising candidates for anti-*Leishmania* vaccine based on salivary proteins (e.g. Collin et al., 2009), and act as high affinity binders of prohaemostatic and proinflammatory biogenic amines (Xu et al., 2011). Results of those three papers will be discussed below in the following order: 1) cDNA library of *P. orientalis* salivary glands, 2) expression and using recombinant salivary proteins derived from main antigens of *P. orientalis* for detecting specific IgG antibodies in exposed animals, and 3) the detailed bioinformatic analyses of sand fly salivary yellow-related protein family.

We had a unique opportunity to study and compare two different colonies of *P. orientalis* from Ethiopia. One originates from Addis Zemen area with several human visceral leishmaniasis outbreaks in the past (Herrero et al., 2009) and the other from Melka Werer area which is non-endemic for VL. Addis Zemen is a mountain region (altitude 1800-2000 m) in northwestern Ethiopia and Melka Werer is, on the contrary, located in Rift Valley in southeast of the country at an altitude 800 m. Both of these colonies were established in our laboratory in 2008 and reared under standard conditions as described in Volf and Volfova (2011). Our colleagues, Seblova et al. (2013), performed experimental infection of both colonies with *Le. donovani* promastigotes and tried to find out, if there are some differences in susceptibility to *Leishmania* parasites, which could explain the various epidemiology of VL in these two aforementioned endemic areas. Their results suggest that both colonies are comparable in terms of susceptibility to *Le. donovani*.

Both colonies of *P. orientalis* were used to construct salivary gland cDNA libraries and to analyse obtained sequences. Amino acid sequences of all salivary proteins were subjected to phylogenetic analysis and putative glycosylation pattern was determined. Mass spectrometry was used for proteome analysis of salivary glands from both colonies. Hyaluronidase and apyrase activities were measured as well as the protein antigenicity and

cross-reactivity. In the colonies from Addis Zemen and Melka Werer, we obtained 607 and 567, high quality nucleotide sequences, respectively. We did not detect any significant differences in protein composition or their sequences between these two colonies. High degree of cross-reactivity was observed between antigens from both colonies using sera of experimentally bitten mice. Thus, our results suggest that both colonies are comparable also in terms of salivary protein composition and antigenicity.

P. orientalis salivary proteins could be divided into 13 different protein families. Members of five main protein families (apyrases, D7-related proteins as members of OBP protein superfamily, antigen 5-related proteins, ParSP25-like proteins, and yellow-related proteins) will be discussed in detail in following paragraphs since proteins from these families were further explored in the other two publications included in this thesis. Besides, another eight protein families were detected in both colonies of *P. orientalis* (PpSP15-like proteins – members of OBP protein superfamily, hyaluronidases, endonucleases, phospholipases, pyrophosphatases, amylases, PpSP32-like proteins, SP16-like proteins, and a lufaxin) (Vlkova et al., 2014). Interestingly, we did not found any sequences of adenosine deaminase, so far detected only in one *Phlebotomus* species – *P. duboscqi* (Kato et al., 2006) and in three out of four tested New World sand fly species (Valenzuela et al., 2004; de Moura et al., 2013; Abdeladhim et al., 2016).

Apyrases are the main anti-platelet aggregation enzymes and were identified in all other sand fly salivary cDNA libraries studied (Table 1) as well as in other blood-feeding insects including mosquitoes, fleas, tsetse flies, triatomine bugs, horseflies, biting midges (e.g. Perez de Leon and Tabachnick, 1996; Andersen et al., 2007; Ma et al., 2009; Caljon et al., 2010; Jariyapan et al, 2012), and several non-haematophagus species (e.g. Fenckova et al., 2011; Wu et al., 2012). In *P. orientalis*, two apyrases were detected in colony originating from Melka Werer and three from Addis Zemen. All had similar molecular mass (around 34 kDa) and pI varying between 8.8-10.0. They all clustered into two clades with 66% similarity between clades and more than 95% within clades (Vlkova et al., 2014). The most closely related apyrases from *Larroussius* subgenus were detected from *P. perniciosus* and *P. tobbi*. ATPase as well as ADPase activities were comparable for both colonies (Vlkova et al., 2014). All five apyrases were detected in proteome analysis and all reacted with specific antibodies (Vlkova et al., 2014), thus confirming antigenic properties of this protein family (Vlkova et al., 2011; Martin-Martin et al., 2012; Martin-Martin et al.,

2013). Glycosylation of apyrases was predicted and in one case confirmed experimentally by ConA binding (Vlkova et al., 2014).

D7-related proteins belong to odorant-binding protein superfamily. Three members were detected in *P. orientalis* colony from Addis Zemen and four in Melka Werer colony. Their pI varied between 6.4-9.2 and their molecular mass was approximately 27 kDa (Vlkova et al., 2014). The function of these proteins still remains unclear in sand flies, but they are suspected to be binders of biogenic amines like homologous D7 proteins in mosquitoes (Calvo et al., 2006; Mans et al., 2007; Calvo et al., 2009). In the phylogenetic tree they were divided into three clades with low overall similarity among them (31%) which showed high diversity in this protein family. D7-related proteins of *P. orientalis* were recognized only by some murine sera (Vlkova et al., 2014) but in previous studies, recognition of these proteins by various host sera was observed (Rohousova et al., 2005; Bahia et al., 2007; Marzouki et al., 2011; Vlkova et al., 2011). Although the limited number of glycosylation sites was predicted for these proteins, they reacted with ConA (Vlkova et al., 2014). The putative glycosylation pattern corresponded with previous studies (Hostomska et al., 2009; Rohousova et al., 2012a) and showed, that all glycosylated D7-related proteins clustered together in the phylogenetic tree (Vlkova et al., 2014), which could suggest, that proteins from different clades might have different molecular functions even though they are all from one protein family.

Two homologs of antigen 5-related proteins were detected in each cDNA library and proteome of *P. orientalis*. All these proteins had predicted molecular mass 28.8 kDa and were slightly basic (8.9). Two clades with similarity more than 99% were determined in the phylogenetic tree (Vlkova et al., 2014). Proteins belonging to the same family were previously described in most blood-feeding insects (Asumpcao et al, 2013) as well as in other organisms including Hymenoptera, hookworms, snails, or bacteria (reviewed in King and Spangfort, 2000; Milne et al., 2003 Ma et al., 2015). Their function in sand flies is still unclear, but pathogenesis-related 1 proteins (from the same superfamily) from the hookworm *Necator americanus* showed sequential similarities to chemokines, which could suggest a host immune system modulatory function (Asojo et al., 2013). In triatominae bugs, superoxide dismutase activity of these proteins was described (Assumpcao et al., 2013). In *P. orientalis*, antigenicity of antigen 5-related proteins to mice was confirmed (Vlkova et al., 2014) which corresponds with generally high antigenicity of these proteins across all sand fly species tested so far (Hostomska et al., 2009; Vlkova et al., 2011; Vlkova et al., 2012; Rohousova et al., 2012a; Martin-Martin et al., 2013). Putative

glycosylation of these proteins was predicted and confirmed experimentally by binding of ConA to proteins with bands with appropriate molecular mass (Vlkova et al, 2014).

ParSP25-like proteins seem to be specific for some Old World sand fly species. Previously, they have been detected in only five out of the eight species: *P. ariasi, P. perniciosus, P. tobbi, P. arabicus,* and *P. papatasi* (Anderson et al., 2006; Oliveira et al., 2006; Hostomska et al., 2009; Rohousova et al., 2012a; Abdeladhim et al., 2012). We were able to obtain sequences of only one ParSP25-like protein for each colony of *P. orientalis*. They shared molecular mass 27.6 kDa, acidic pI around 4.7, and their sequential similarity was 98% (Vlkova et al., 2014). Exact function of this protein family is unknown, in *P. perniciosus* ParSP25-like proteins showed antigenic properties to various hosts bitten by *P. perniciosus* (Vlkova et al., 2011; Martin-Martin et al., 2012), however, we did not see any positive reaction with sera of mice bitten by *P. orientalis*. No type of glycosylation was predicted for these two proteins (Vlkova et al., 2014).

Yellow-related proteins are highly abundant and were discovered in all up to date published cDNA libraries of sand fly salivary glands (Table 1) as well as in other blood-feeding and non-haematophagus insects (for example Nash et al., 1983; Hanes et al., 1992; Johnson et el., 2001; Alves-Silva et al., 2010). In each colony, two YRPs were identified and detected in proteomes, all of them with similar molecular mass (around 42 kDa) but with wide range of pI (6.1-8.1). Phylogenetic analysis divided these proteins into two clades with sequential similarity more than 99% in each clade and 77% between clades. High degree of similarity among all sand fly YRPs indicate the same anti-inflammatory function, so far experimentally proven only for *L. longipalpis* YRPs (Xu et al., 2011). Their reaction with antibodies from sera of experimentally bitten mice confirmed very high antigenicity of these proteins described in previous studies (Gomes et al., 2002; Vinhas et al., 2007; Gomes et al., 2007; Bahia et al., 2007; Hostomska et al., 2008; Teixeira et al., 2010; Gidwani et al., 2011; Marzouki et al., 2011; Vlkova et al., 2011; Vlkova et al., 2012). Rohousova et al. (2012a) showed that glycosylation sites occur in YRPs in *P. tobbi* and *P. sergenti*. Putative glycosylation sites were predicted also for *P. orientalis* YRPs and experimentally confirmed by specific binding of ConA to proteins with molecular mass of approximately 42 kDa (Vlkova et al., 2014).

For purpose of the second study, we expressed major salivary antigens in the recombinant form and used them for detecting antibodies against *P. orientalis* in sera of domestic animals collected in Ethiopia, in the area endemic for human visceral

leishmaniasis. Sera from various host species were previously tested by Rohousova et al. (2015) by ELISA, using the whole salivary gland homogenate as an antigen. Here we focused on antigens recognized by canine sera, expressed them in *E. coli* system, and tested them first with a relatively small number of serum samples of four host species. Then, we chose the best performing proteins to validate them in a large-scale study. We also tested the specificity of the reaction by using sera of mice exposed to other sand fly species.

Proteomic analysis and mass spectrometry were used to identify main salivary antigens for dogs. Five antigens were recognized by specific IgG antibodies: a ParSP25-like protein, a yellow-related protein, an antigen 5-related protein, an apyrase, and a D7-related protein (Sima et al., 2016a). Members of these protein families have been previously shown to react with sera from dogs bitten by *L. longipalpis* or *P. perniciosus* (Bahia et al., 2007; Hostomska et al., 2008; Vlkova et al., 2011). After this initial selection, all five proteins were expressed in bacterial expression system and purified in Ni-NTA column.

So far, recombinant proteins have been expressed only from 3 sand fly species (*P. perniciosus, P. papatasi*, and *L. longipalpis*) and tested for detecting anti-sand fly saliva antibodies in experimentally as well as naturally bitten hosts with a partial success (Table 2). It has been shown, that several individual recombinant proteins are host species-specific. For example, recombinant apyrase, D7-related protein, and SL1 protein from *L. longipalpis* were recognized only by canine sera (Teixeira et al., 2010). On the other hand, recombinant yellow-related proteins (LJM17 from *L. longipalpis* and rSP03B from *P. perniciosus*) showed strong reaction with sera of all host species tested (Teixeira et al., 2010; Drahota et al., 2014; Martin-Martin et al., 2014; Kostalova et al., 2015). These studies suggested that yellow-related proteins might be the best candidate for an antigen in epidemiological studies involving different host species.

In the preliminary experiment, five recombinant proteins from *P. orientalis* were tested with small number of sera from experimentally bitten mice and naturally exposed dogs, sheep, and goats from Ethiopia. Results with antigen 5-related recombinant protein were not satisfactory. Although we detected strong reaction with a band corresponding to antigen 5-related protein in the immunoblot with the whole salivary gland homogenate, the correlation coefficient between SGH and the recombinant protein was very low for Ethiopian animals so we decided not to use it in large-scale experiments (Sima et al.,

2016a). Recombinant antigen 5-related protein from *P. perniciosus* reacted neither with mice nor with canine sera (Drahota et al., 2014).

The other four recombinant proteins showed promising results at least for one hosts species. Therefore, we decided to test them by ELISA using all available serum samples, i.e. 36 canine sera, 179 sheep sera, 233 goat sera, and appropriate number of sera from non-exposed controls. These ELISA experiments revealed the recombinant yellow-related protein as the most universal candidate for further prospective testing; compared to SGH, it showed the best correlation coefficient in sheep and goats and the second best in dogs. The highest correlation coefficient for dogs was observed with recombinant ParSP25-like protein, but the specificity of the test was very low suggesting high probability of false positivity among non-exposed dogs (Sima et al., 2016a). These results confirmed the possibility of using recombinant yellow-related proteins for detecting specific anti-sand fly saliva antibodies in various host species, including naturally exposed dogs, foxes, hares, rabbits, sheep, goats, or humans (Teixeira et al., 2010; Souza et al., 2010; Drahota et al., 2014; Martin-Martin et al., 2014; Kostalova et al., 2015, Sima et al., 2016a).

The reactivity of individual recombinants could be influenced by the system used for expression or protein purification and therefore some authors prefer using mammalian cells, which enable posttranslational modifications (e.g. glycosylation − antibodies can be also targeted to the glycosylated part of the antigen) (Marzouki et al., 2012; Marzouki et al., 2015). On contrary, recombinants expressed in *E. coli* proved to be efficient enough to detect anti-sand fly saliva IgG (Vlkova et al., 2012; Drahota et al., 2014; Martin-Martin et al., 2014; Kostalova et al., 2015). Moreover, the specificity of *P. orientalis* recombinant antigens was proved by using murine sera experimentally exposed to sand fly species *(P. papatasi, Sergentomyia schwetzi*), which are sympatric with *P. orientalis* in Ethiopia. Antibodies from sera of mice bitten by *P. papatasi* and *S. schwetzi* did not react with recombinant proteins from *P. orientalis* (Sima et al., 2016a). So far as I am aware, only two other studies tested and confirmed the specificity of recombinant proteins. Recombinants from *L. longipalpis* were not recognized by antibodies from sera of humans exposed to *L. intermedia* (Teixeira et al., 2010) and rPpSP32 from *P. papatasi* was not recognized by dogs exposed to *P. perniciosus* (Marzouki et al., 2015).

Because of the promising antigenic properties of recombinant yellow-related proteins mentioned above, their protective effect against *Leishmania* parasites (Collin et al., 2009; Xu et al., 2011; Gomes et al., 2012), and their known structure and function in *L.*

longipalpis (Xu et al., 2011), we decided to focus on this important protein family in the third publication included in this thesis. We identified proteins from this family among all published cDNA libraries of sand fly salivary glands, determined their phylogenetic relationship, predicted glycosylation sites, constructed models of their 3D structures, showed differences in surface electrostatic potentials, and described the binding site for biogenic amines.

Thirty three different yellow-related proteins were found out across thirteen sand fly species with published cDNA libraries of salivary glands (Sima et al., 2016b). The number of YRPs in individual sand fly species varied from one to five (e.g. Anderson et al., 2006; Rohousova et al., 2012a), probably because of different sensitivity of sequencing methods or differential gene expression, when genes with lower expression could be recorded less frequently (Rohousova et al., 2012b). High sequential similarity (44-99%) was detected among all identified YRPs. Our phylogenetic analysis divided these proteins into several branches. Closer relationship was discovered among proteins from *Phlebotomus* species and mainly proteins from the same subgenera (*Phlebotomus, Paraphlebotomus,* and *Larroussius*) than from *Lutzomyia* YRPs (Sima et al., 2016b). It confirmed the previously published sand fly phylogenetic relationships based on YRP amino acid sequences (Rohousova et al., 2012a; Vlkova et al., 2014) as well as on small subunit nuclear ribosomal DNA (Aransay et al., 2000).

Xu et al. (2011) solved the 3D structure of *L. longipalpis* YRP and confirmed its ability to bind biogenic amines such as serotonin, histamine, and catecholamines with different affinities. Blocking of these small molecules leads to vasodilatation, platelet deactivation, and a decrease in vascular permeability (Oliveira et al., 2007; Xanthos et al., 2008) which help the sand fly female to complete its blood meal. Similar function of various salivary proteins was described in other blood-sucking arthropods as well (Paesen et al., 1999; Andersen et al., 2003; Calvo et al., 2006; Mans et al., 2007; Mans et al., 2008; Calvo et al., 2009), which indicates the importance of amine-binders for blood-feeders.

In sand flies, only LJM11, the YRP from *L. longipalpis*, has known 3D structure. It is folded as a six-bladed β-propeller, which is a unique fold identified only in this protein. It is characterized by a ligand-binding tunnel with two entrances (Xu et al., 2011). Based on the homologous modeling of other sand fly YRPs, we showed that these proteins share similar fold among all sand fly species. The length or minimum width of the tunnels vary among YRPs but it remains uninterrupted in all modeled proteins (Sima et al., 2016b). Therefore, we suggested the possibility that also other members of this sand fly protein

family bind biogenic amines as has been shown for LJM11 (Xu et al., 2011). Our data also indicate probable differences in the charge of tunnel entrances, thus paralogs within one species could bind same ligands with different affinities (Sima et al., 2016b).

In the LJM11 structure, the biogenic amine serotonin is stabilized by hydrogen bonds from three amino acids in the main chain and the side chain and by a number of additional hydrophobic interactions (Xu et al., 2011). We presented that there is a variability among YRPs in these binding amino acids, mainly on the positions denoted in LJM11 as Phe 344, which binds ligands through a hydrogen bond. Substitution on this ligand-binding position was the only one described in three YRPs from *L. longipalpis* which indicate different affinities for serotonin and other tested bioamines (Xu et al., 2011). Our studied proteins can be divided into eight groups based on their ligand-binding amino acid composition. Distribution of YRPs in several of these groups corresponds with the branches identified in phylogenetic tree as well as with the predicted glycosylation patterns (Sima et al., 2016b).

CONCLUSIONS

Three publications dealing with a similar topic – salivary proteins of *P. orientalis* and other sand flies - were included in this thesis.

In the first study, we constructed cDNA library of salivary glands of two colonies of *P. orientalis*, the main vector of *Le. donovani* in East African countries. We detected a rich mixture of proteins in each library, which corresponded to all other published transcriptomes. Particularly, we detected and described members of 13 different protein families and discussed their biological properties. We also confirmed that ParSP25-like protein family is probably specific only for some sand fly species. We did not detect any significant differences in protein composition, sequences, putative glycosylation, or antigenicity between colonies originating from areas endemic (Addis Zemen) and non-endemic (Melka Werer) for *Le. donovani* (Vlkova et al., 2014).

Up to date, thirteen cDNA libraries of sand fly salivary glands have been published. They were all focused on important *Leishmania* vectors and neglected the other sand flies. Although some sand fly species do not transmit *Leishmania* parasites infectious for humans or domestic animals, it would be interesting to know, which proteins occur in their saliva and if their antigenic and enzymatic activities reflect the adaptation to different host species.

In the second study, we expressed five recombinant antigens from *P. orientalis* salivary glands and selected four of them for large-scale experiments, detecting specific anti-saliva antibodies in repeatedly bitten hosts. The recombinant yellow-related protein was suggested as the most promising antigen for large-scale studies based on the high correlation with the antibody response against the whole salivary gland homogenate and the highest values of specificity, sensitivity, positive and negative predictive values with majority of the host species tested. We also demonstrated that there was no cross-reactivity with antibodies from mice bitten by other sympatric sand fly species (Sima et al., 2016a).

Using recombinant proteins in such studies could replace salivary gland homogenate and circumvent the disadvantages of this whole protein mixture, such as high amount of work required to obtain sufficient quantities of the proteins, the necessity for sand fly colony maintenance, laborious dissections of salivary glands and potential cross-reactivity with non-vector species. Despite positive results from previous studies and as well from our experiments, further investigations with another proteins and another sand fly or host species are needed to validate such a recombinant protein-based test for routine

So far, none of the recombinant proteins have been tested with sera from hosts bitten by other blood-sucking insects/arthropods to show the sand fly species-specificity of the reaction

Salivary recombinant proteins have been validated in various vector-host models as a reliable epidemiological tool. However, antibodies from sera of exposed host recognize only some amino acid parts, so called B-cell epitopes, on the surface of the antigens. For tse-tse flies, it has been shown that these epitopes, when prepared synthetically, are suitable for antibodies detection as well (Dama et al., 2013). We suggest that epitopes from sand fly salivary antigens could be utilized in similar way. There is a possibility, that using these epitopes in areas with sympatric occurrence of related species might increase the specificity of the reaction as it was shown for different strains of *Trypanosoma cruzi* in Latin America (Bhattacharyya et al., 2014). Sand fly salivary yellow-related proteins should be the first choice for designing B-cell epitopes because of their suitability as antigens in various hosts. Utilizing B-cell epitopes instead of recombinant proteins could also substantially reduce the cost of the test.

In the third study, we identified 33 different yellow-related proteins among thirteen sand fly species with published cDNA libraries of salivary glands. All identified 33 YRPs share similar folding and contain a ligand-binding tunnel. Based on the constructed 3D models, we suggested similar function for all of these proteins – binding of biogenic amines. Nevertheless, diversity in amino acid composition and modifications in the charge of entrances to the tunnel may affect ligand-binding properties of a particular YRP. However, experimental data which would confirm this hypothesis are necessary for better understanding of the biological role of these important proteins. We found out that distribution of amine-binding amino acids corresponds with the position of the protein in the phylogenetic tree as well as with the predicted glycosylation patterns (Sima et al., 2016b).

Yellow-related proteins have many interesting characteristics that deserve further research. By binding biogenic amines, these proteins help the sand fly female to overcome the haemostatic reaction of the host during blood-feeding, they also have potential to be used in anti-*Leishmania* vaccines and, last but not least, are most universal antigens for detecting anti-sand fly salivary antibodies in repeatedly bitten hosts in large epidemiological studies.

REFERENCES

1. Abdeladhim M, Jochim RC, Ben Ahmed M, Zhioua E, Chelbi I, Cherni S, et al. Updating the salivary gland transcriptome of *Phlebotomus papatasi* (Tunisian Strain): The search for sand fly-secreted immunogenic proteins for humans. PLoS One. 2012; 7: e47347.

2. Abdeladhim M, Kamhawi S, Valenzuela JG. What's behind a sand fly bite? The profound effect of sand fly saliva on host hemostasis, inflammation and immunity. Infect Genet Evol. 2014; 28: 691–703.

3. Adeladhim M, Coutinho-Abreu IV, Townsend S, Pasos-Pinto S, Sanchez L, Rasouli M, et al. Molecular diversity between salivary proteins from New World and Old World sand flies with emphasis on *Bichromomyia olmeca*, the sand fly vector of *Leishmania mexicana* in Mesoamerica. PLoS Negl Trop Dis. 2016; 10: e0004771.

4. Albert S, Bhattacharya D, Klaudiny J, Schmitzova J, Simuth J. The family of major royal jelly proteins and its evolution. J Mol Evol. 1999; 49: 290–7.

5. Alvarenga PH, Francischetti IMB, Calvo E, Sa-Nunes A, Ribeiro JMC, Andersen JF. The function and three-dimensional structure of a tromboxane A(2)/cysteinyl leukotriene-binding protein from the saliva of a mosquito vector of the malaria parasite. PLoS Biol. 2010; 8: e1000547.

6. Alvarenga PH, Xu X, Oliveira F, Chagas AC, Nascimento CR, Francischetti IMB, et al. Novel family of insect salivary inhibitors blocks contact pathway activation by binding to polyphosphate, heparin, and dextran sulfate. Arterioscler Thromb Vasc Biol. 2013; 33: 2759–2770.

7. Alves-Silva J, Ribeiro JMC, Van Den Abbeele J, Attardo G, Hao ZR, Haines LR, et al. An insight into the sialome of *Glossina morsitans morsitans*. BMC Genomics. 2010; 11: 213.

8. Andersen JF, Francischetti IMB, Valenzuela JG, Schuck P, Ribeiro JMC. Inhibition of hemostasis by a high affinity biogenic amine-binding protein from the saliva of a blood-feeding insect. J Biol Chem. 2003; 278: 4611–4617.

9. Andersen JF, Hinnebusch BJ, Lucas DA, Conrads TP, Veenstra TD, Pham CM, et al. An insight into the sialome of the oriental rat flea, *Xenopsylla cheopis* (Rots). BMC Genomics. 2007; 8: 102.

10. Andersen JF, Montfort WR. The crystal structure of nitrophorin 2: A trifunctional antihemostatic protein form the saliva of *Rhodnius prolixus*. J Biol Chem. 2000; 275: 30496–30503.

11. Andersen JF, Pham VM, Meng Z, Champagne DE, Ribeiro JM. Insight into the sialome of the black fly, *Simulium vittatum*. J Proteome Res. 2009; 8: 1474–1488.

12. Anderson JM, Oliveira F, Kamhawi S, Mans BJ, Reynoso D, Seitz AE, et al. Comparative salivary gland transcriptomics of sandfly vectors of visceral leishmaniasis. BMC Genomics. 2006; 7: 52.

13. Anjili CO, Mbati PA, Mwangi RW, Githure JI, Olobo JO, Robert LL, et al. The chemotactic effect of *Phlebotomus duboscqi* (Diptera, Psychodidae) salivary gland lysates to murine monocytes. Acta Trop. 1995; 60: 97–100.

14. Antinori S, Schifanella L, Corbellino M. Leishmaniasis: new insights from an old and neglected disease. Eur J Clin Microbiol Infect Dis. 2012; 31: 109–118.

15. Aransay AM, Scoulica E, Tselentis Y, Ready PD. Phylogenetic relationships of Phlebotomine sandflies inferred from small subunit nuclear ribosomal DNA. Insect Mol Biol. 2000; 9: 157–168.

16. Arca B, Lombardo F, Lanfrancotti A, Spanos L, Veneri M, Louis C; et al. A cluster of four D7-related genes is expressed in the salivary glands of the African malaria vector *Anopheles gambiae*. Insect Mol Biol. 2002; 11: 47–55.

17. Arca B, Lombardo F, Valenzuela JG, Francischetti IM, Marinotti O, Coluzzi M, et al. An updated catalogue of salivary gland transcripts in the adult female mosquito, *Anopheles gambiae*. J Exp Biol. 2005; 208: 3971–86.

18. Asojo OA, Goud G, Dhar K, Loukas A, Zhan B, Deumic V, et al. X-ray structure of Na-ASP-2, a pathogenesis-related-1 protein from the nematode parasite, *Necator americanus*, and a vaccine antigen for human hookworm infection. J Mol Biol. 2005; 346: 801–814.

19. Assumpcao TC, Ma D, Schwarz A, Reiter K, Santana JM, Andersen JF, et al. Salivary antigen-5/CAP family members are Cu2+-dependent antioxidant enzymes that scavenge O2- and inhibit collagen-induced platelet aggregation and neutrophil oxidative burst. J Biol Chem. 2013; 288: 14341–14361.

20. Bahia D, Gontijo NF, Leon IR, Perales J, Pereira MH, Oliveira G, et al. Antibodies from dogs with canine visceral leishmaniasis recognise two proteins from the saliva of *Lutzomyia longipalpis*. Parasitol Res. 2007; 100: 449–454.

21. Belkaid Y, Kamhawi S, Modi G, Valenzuela J, Noben-Trauth N, Rowton E, et al. Development of a natural model of cutaneous leishmaniasis: powerful effects of vector saliva and saliva preexposure on the long-term outcome of *Leishmania major* infection in the mouse ear dermis. J Exp Med. 1998; 188: 1941–1953.

22. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. Nucleic Acids Res. 2000; 28: 235–242.

23. Bhattacharyya T, Falconar AK, Luquetti AO, Costales JA, Grijalva MJ, Lewis MD, et al. Development of peptide-based lineage-specific serology for chronic Chagas disease: geographical and clinical distribution of epitope recognition. PLoS Negl Trop Dis. 2014; 8: e2892.

24. Caljon G, De Ridder K, De Baetselier P, Coosemans M, Van Den Abbeele J. Identification of a tsetse fly salivary protein with dual inhibitory action on human platelet aggregation. PLoS One. 2010; 5: e9671.

25. Calvo E, Mans BJ, Andersen JF, Ribeiro JM. Function and evolution of a mosquito salivary protein family. J Boil Chem. 2006; 28: 1935–42.

26. Calvo E, Pham VM, Marinotti O, Andersen JF, Ribeiro JM. The salivary gland transcriptome of the neotropical malaria vector *Anopheles darlingi* reveals accelerated evolution of genes relevant to hematophagy. BMC Genomics. 2009; 10: 57.

27. Campbell CL, Vandyke KA, Letchworth GJ, Drolet BS, Hanekamp T, Wilson WC. Midgut and salivary gland transcriptomes of the arbovirus vector *Culicoides sonorensis* (Diptera: Ceratopogonidae). Insect Mol Biol. 2005; 14: 121–136.

28. Charlab R, Valenzuela JG, Rowton ED, Ribeiro JMC. Toward an understanding of the biochemical and pharmacological complexity of the saliva of a hematophagous sand fly *Lutzomyia longipalpis*. Proc Natl Acad Sci USA. 1999; 96: 15155–15160.

29. Clements MF, Gidwani K, Kumar R, Hostomska J, Dinesh DS, Kumar V, et al. Measurement of recent exposure to *Phlebotomus argentipes*, the vector of Indian visceral leishmaniasis, by using human antibody responses to sand fly saliva. Am J Trop Med Hyg. 2010; 82: 801–807.

30. Collin N, Gomes R, Teixeira C, Cheng L, Laughinghouse A, Ward JM, et al. Sand fly salivary proteins induce strong cellular immunity in a natural reservoir of visceral leishmaniasis with adverse consequences for *Leishmania*. PLoS Pathog. 2009; 5: e1000441.

31. Dama E, Cornelie S, Camara M, Somda MB, Poinsignon A, Ilboudo H, et al. *In Silico* identification of a candidate synthetic peptide (Tsgf1 $_{18-43}$) to monitor human exposure to tsetse flies in west Africa. PLoS Negl Trop Dis. 2013; 7: e2455.

32. de Moura TR, Oliveira F, Carneiro MW, Miranda JC, Clarencio J, Barral-Netto M, et al. Functional transcriptomics of wild-caught *Lutzomyia intermedia* salivary glands: identification of a protective salivary protein against *Leishmania braziliensis* infection. PLoS Negl Trop Dis. 2013; 7: e2242.

33. de Moura TR, Oliveira F, Novais FO, Miranda JC, Clarencio J, Follador I, et al. Enhanced *Leishmania braziliensis* infection following pre-exposure to sandfly saliva. PLoS Negl Trop Dis. 2007; 1: e84.

34. de Moura TR, Oliveira F, Rodrigues GC, Carneiro MW, Fukutani KF, Novais FO, et al. Immunity to *Lutzomyia intermedia* saliva modulates the inflammatory environment induced by *Leishmania braziliensis*. PLoS Negl Trop Dis. 2010; 4: e712.

35. Drahota J, Martin-Martin I, Sumova P, Rohousova I, Jimenez M, Molina R, et al. Recombinant antigens from *Phlebotomus perniciosus* saliva as markers of canine exposure to visceral leishmaniases vector. PLoS Negl Trop Dis. 2014; 8: e2597.

36. Elnaiem DA, Osman OF. Evidence for active transmission of visceral leishmaniasis within a village in eastern Sudan. Acta Trop. 1998; 71: 305–309.

37. Elnaiem DA. Ecology and control of the sand fly vectors of *Leishmania donovani* in East Africa, with special emphasis on *Phlebotomus orientalis*. J Vector Ecol. 2011; 36: S23–S31.

38. Elnaiem DE, Meneses C, Slotman M. Lanzaro GC. Genetic variation in the sand fly salivary protein, SP-15, a potential vaccine candidate against *Leishmania major*. Insect Mol Biol. 2005; 14: 145–150.

39. Fenckova M, Hobizalova R, Fric ZF, Dolezal T. Functional characterization of ecto-5'-nucleotidases and apyrases in *Drosophila melanogaster*. Insect Biochem Mol Biol. 2011; 41: 956–967.

40. Fiuza JA, Dey R, Davenport D, Abdeladhim M, Meneses C, Oliveira F, et al. Intradermal immunization of *Leishmania donovani* centrin knock-out parasites in combination with salivary protein LJM19 from sand fly vector induces a durable protective immune response in hamsters. PLoS Negl Trop Dis. 2016; 10: e0004322.

41. Flower DR. The lipocalin protein family: structure and function. Biochemical J. 1996; 318: 1–14.

42. Gebre-Michael T, Balkew M, Alamirew T, Gudeta N, Reta M. Preliminary entomological observations in a highland area of Amhara region, northern Ethiopia, with epidemic visceral leishmaniasis. Ann Trop Med Parasitol. 2007; 101: 367–70.

43. Gebre-Michael T, Balkew M, Berhe N, Hailu A, Mekonnen Y. Further studies on the phlebotomine sandflies of the kala-azar endemic lowlands of Humera-Metema (north-west Ethiopia) with observations on their natural blood meal sources. Parasit Vectors. 2010; 3: 6.

44. Gebresilassie A, Abbasi I, Aklilu E, Yared S, Kirstein OD, Moncaz A, et al. Host-feeding preference of *Phlebotomus orientalis* (Diptera: Psychodidae) in an endemic focus of visceral leishmaniasis in northern Ethiopia. Parasit Vectors. 2015a; 8: 270.

45. Gebresilassie A, Yared S, Aklilu E, Kirstein OD, Moncaz A, Tekie H, et al. Host choice of *Phlebotomus orientalis* (Diptera: Psychodidae) in animal baited experiments: a field study in Tahtay Adiyabo district, northern Ethiopia. Parasit Vectors. 2015b; 8: 190.

46. Gebresilassie A, Kirstein OD, Yared S, Aklilu E, Moncaz A, Tekie H, et al. Species composition of phlebotomine sand flies and bionomics of *Phlebotomus orientalis* (Diptera: Psychodidae) in an endemic focus of visceral leishmaniasis in Tahtay Adiyabo district, Northern Ethiopia. Parasit Vectors. 2015c; 8: 248.

47. Gidwani K, Picado A, Rijal S, Singh SP, Roy L, Volf P, et al. Antibody response to sand fly saliva to evaluate human visceral leishmaniasis vector exposure in India and Nepal: effect of long-lasting insecticidal nets. Trop Med Int Health. 2011; 16: 66–66.

48. Gomes R, Oliveira F, Teixeira C, Meneses C, Gilmore DC, Elnaiem DE, et al. Immunity to sand fly salivary protein LJM11 modulates host response to vector-transmitted *Leishmania* conferring ulcer-free protection. J Invest Dermatol. 2012; 132: 2735–2743.

49. Gomes R, Teixeira, C, Teixeira MJ, Oliveira F, Menezes MJ, Silva C, et al. Immunity to a salivary protein of a sand fly vector protects against the fatal outcome of visceral leishmaniasis in a hamster model. Proc Natl Acad Sci USA. 2008; 105: 7845–7850.

50. Gomes RB, Brodskyn C, de Oliveira CI, Costa J, Miranda JC, Caldas A, et al. Seroconversion against *Lutzomyia longipalpis* saliva concurrent with the development of anti-*Leishmania chagasi* delayed-type hypersensitivity. J Infect Dis. 2002; 186: 1530–1534.

51. Gomes RB, Mendonca IL, Silva VC, Ruas J, Silva MB, Cruz MS, et al. Antibodies against *Lutzomyia longipalpis* saliva in the fox *Cerdocyon thous* and the sylvatic cycle of *Leishmania chagasi*. Trans R Soc Trop Med Hyg. 2007; 101: 127–133.

52. Hall LR, Titus RG. Sand fly vector saliva selectively modulates macrophage functions that inhibit killing of *Leishmania major* and nitric oxide production. J Immunol. 1995; 155: 3501–3506.

53. Hamasaki R, Kato H, Terayama Y, Iwata H, Valenzuela JG. Functional characterization of a salivary apyrase from the sand fly, *Phlebotomus duboscqi*, a vector of *Leishmania major*. J Insect Physiol. 2009; 55: 1044–1049.

54. Hanes J, Simuth J. Identification and partial characterization of the major royal jelly protein of the honey-bee (*Apis mellifera L*). J Apic Res. 1992; 31: 22–26.

55. Hassan MM, Elamin EM, Mukhtar MM. Isolation and identification of *Leishmania donovani* from *Phlebotomus orientalis*, in an area of eastern Sudan with endemic visceral leishmaniasis. Ann Trop Med Parasitol. 2008; 102: 553–555.

56. Hassan MM, Elraba'a FM, Ward RD, Maingon RD, Elnaiem DA. Detection of high rates of in-village transmission of *Leishmania donovani* in eastern Sudan. Acta Trop. 2004; 92: 77–82.

57. Herrero M, Orfanos G, Argaw D, Mulugeta A, Aparicio P, Parreno F, et al. Natural history of a visceral leishmaniasis outbreak in highland Ethiopia. Am J Trop Med Hyg. 2009; 81: 373–377.

58. Hostomska J, Rohousova I, Volfova V, Stanneck D, Mencke N, Volf P. Kinetics of canine antibody response to saliva of the sand fly *Lutzomyia longipalpis*. Vector Borne Zoonotic Dis. 2008; 8: 443–450.

59. Hostomska J, Volfova V, Mu J, Garfield M, Rohousova I, Volf P, et al. Analysis of salivary transcripts and antigens of the sand fly *Phlebotomus arabicus*. BMC Genomics. 2009; 10: 282.

60. Isawa H, Yuda M, Orito Y, Chinzei Y. A mosquito salivary protein inhibits activation of the plasma contact system by binding to factor XII and high molecular weight kininogen. J Biol Chem. 2002; 277: 27651–27658.

61. Jariyapan N, Roytrakul S, Paemanee A, Junkum A, Saeung A, Thongsahuan S, et al. Proteomic analysis of salivary glands of female *Anopheles barbirostris* species A2 (Diptera: Culicidae) by two-dimensional gel electrophoresis and mass spectrometry. Parasitol Res. 2012; 111: 1239–1249.

62. Johnson JK, Li J, Christensen BM. Cloning and characterization of a dopachrome conversion enzyme from the yellow fever mosquito, *Aedes aegypti*. Insect Biochem Mol Biol. 2001; 31: 1125–1135.

63. Kamhawi S, Belkaid Y, Modi G, Rowton E, Sacks D. Protection against cutaneous leishmaniasis resulting from bites of uninfected sand flies. Science. 2000; 290: 1351–1354.

64. Kato H, Anderson JM, Kamhawi S, Oliveira F, Lawyer PG, Pham VM, et al. High degree of conservancy among secreted salivary gland proteins from two geographically distant *Phlebotomus duboscqi* sandflies populations (Mali and Kenya). BMC Genomics. 2006; 7: 226.

65. Kato H, Jochim RC, Gomez EA, Uezato H, Mimori T, Korenaga M, et al. Analysis of salivary gland transcripts of the sand fly *Lutzomyia ayacuchensis*, a vector of Andean-type cutaneous leishmaniasis. Infect Genet Evol. 2013; 13: 56–66.

66. King TP, Spangfort MD. Structure and biology of stinging insect venom allergens. Int Arch Allergy Immunol. 2000; 123: 99–106.

67. Klein J, Horejsi V. Immunology. 2nd ed. Oxford: Blackwell Science, 1997; ISBN 0-632-04228-1.

68. Kostalova T, Lestinova T, Sumova P, Vlkova M, Rohousova I, Berriatua E, et al. Canine antibodies against salivary recombinant proteins of *Phlebotomus perniciosus*: A longitudinal study in an endemic focus of canine leishmaniasis. PLoS Negl Trop Dis. 2015; 9: e0003855.

69. Ma D, Francischetti IM, Ribeiro JM, Andersen JF. The structure of hookworm platelet inhibitor (HPI), a CAP superfamily member from *Ancylostoma caninum*. Acta Crystallogr F Struct Biol Commun. 2015; 71: 643–649.

70. Ma D, Wang Y, Yang H, Wu J, An S, Gao L, et al. Anti-thrombosis repertoire of blood-feeding horsefly salivary glands. Mol Cell Proteomic. 2009; 8: 2071–2079.

71. Mans BJ, Calvo E, Ribeiro JM, Andersen JF. The crystal structure of D7r4, a salivary biogenic amine-binding protein from the malaria mosquito *Anopheles gambiae*. J Biol Chem. 2007; 282: 36626–36633.

72. Mans BJ, Ribeiro JM, Andersen JF. Structure, function, and evolution of biogenic amine-binding proteins in soft ticks. J Biol Chem. 2008; 283: 18721–18733.

73. Martin-Martin I, Molina R, Jimenez M. An insight into the *Phlebotomus perniciosus* saliva by a proteomic approach. Acta Trop. 2012; 123: 22–30.

74. Martin-Martin I, Molina R, Jimenez M. Identifying salivary antigens of *Phlebotomus argentipes* by a 2DE approach. Acta Trop. 2013; 126: 229–239.

75. Martin-Martin I, Molina R, Jimenez M. Kinetics of anti-*Phlebotomus perniciosus* saliva antibodies in experimentally bitten mice and rabbits. PLoS One. 2015; 10: e0140722.

76. Martin-Martin I, Molina R, Rohousova I, Drahota J, Volf P, Jimenez M. High levels of anti-*Phlebotomus perniciosus* saliva antibodies in different vertebrate hosts from the re-emerging leishmaniosis focus in Madrid, Spain. Vet Parasitol. 2014; 202: 207-216.

77. Marzouki S, Abdeladhim M, Abdessalem CB, Oliveira F, Ferjani B, Gilmore D, et al. Salivary antigen SP32 is the immunodominant target of the antibody response to *Phlebotomus papatasi* bites in humans. PLoS Negl Trop Dis. 2012; 6: e1911.

78. Marzouki S, Ben Ahmed M, Boussoffara T, Abdeladhim M, Ben Aleya-Bouafif N, Namane A, et al. Characterization of the antibody response to the saliva of *Phlebotomus papatasi* in people living in endemic areas of cutaneous leishmaniasis. Am J Trop Med Hyg. 2011; 84: 653–661.

79. Marzouki S, Kammoun-Rebai W, Bettaieb J, Abdeladhim M, Hadj Kacem S, Abdelkader R, et al. Validation of recombinant salivary protein PpSP32 as a suitable marker of human exposure to *Phlebotomus papatasi*, the vector of *Leishmania major* in Tunisia. PLoS Negl Trop Dis. 2015; 9: e0003991.

80. Mbow ML, Bleyenberg JA, Hall LR, Titus RG. *Phlebotomus papatasi* sand fly salivary gland lysate down-regulates a Th1, but up-regulates a Th2, response in mice infected with *Leishmania major*. J Immunol. 1998; 161: 5571–5577.

81. Milne TJ, Abbenante G, Tyndall JD, Halliday J, Lewis RJ. Isolation and characterization of a cone snail protease with homology to CRISP proteins of the Pathogenesis-related Protein Superfamily. J Biol Chem. 2003; 278: 31105–31110.

82. Mondragon-Shem K, Al-Salem WS, Kelly-Hope L, Abdeladhim M, Al-Zahrani M, Valenzuela JG, et al. Severity of Old World cutaneous leishmaniasis is influenced by previous exposure to sandfly bites in Saudi Arabia. PLoS Negl Trop Dis. 2015; 9: e3449.

83. Morris RV, Shoemaker CB, David JR, Lanzaro GC, Titus RG. Sandfly maxadilan exacerbates infection with *Leishmania major* and vaccinating against it protects against *L. major* infection. J Immunol. 2001; 167: 5226–5230.

84. Nash WG, Kamerow HN, Merril CR. Identification of a yellow gene-specific protein in *Drosophila melanogaster* by two-dimensional gel-electrophoresis. Biochem Genet. 1983; 21: 1135–1142.

85. Ngumbi PM, Kaburi JC, Anjili CO, Haas F. *Phlebotomus* (*Larroussius*) *orientalis* (Diptera: Psychodidae) as a probable secondary vector of visceral leishmaniasis in Kenya. J Vector Borne Dis. 2010; 47: 58–60.

86. Oliveira F, Kamhawi S, Seitz AE, Pham VM, Guigal PM, Fisher L, et al. From transcriptome to immunome: Identification of DTH inducing proteins from a *Phlebotomus ariasi* salivary gland cDNA library. Vaccine. 2006; 24: 374–390.

87. Oliveira F, Lawyer PG, Kamhawi S, Valenzuela JG. Immunity to distinct sand fly salivary proteins primes the anti-*Leishmania* immune response towards protection or exacerbation of disease. PLOS Negl Trop Dis. 2008; 2: e226.

88. Oliveira MCG, Pelegrini-Da-Silva A, Parada CA, Tambeli CH. 5-HT acts on nociceptive primary afferents through an indirect mechanism to induce hyperalgesia in the subcutaneous tissue. Neuroscience. 2007; 145: 708–714.

89. Paesen G, Adams P, Harlos K, Nuttall P, Stuart D. Tick histamine-binding proteins: isolation, cloning, and three-dimensional structure. Mol Cell. 1999; 3: 661–71.

90. Perez de Leon AA, Tabachnick WJ. Apyrase activity and adenosine diphosphate induced platelet aggregation inhibition by the salivary gland proteins of *Culicoides variipennis*, the North American vector of bluetongue viruses. Vet Parasitol. 1996; 61: 327–338.

91. Prates DB, Araujo-Santos T, Luz NF, Andrade BB, Franca-Costa J, Afonso L, et al. *Lutzomyia longipalpis* saliva drives apoptosis and enhances parasite burden in neutrophils. J Leukoc Biol. 2011; 90: 575–582.

92. Ready PD. Leishmaniasis emergence in Europe. Euro Surveill. 2010; 15: 29–39.

93. Ribeiro JM, Arca B. From sialomes to the sialoverse: An insight into salivary potion of blood-feeding insects. Adv in Insect Phys. 2009; 37: 59–118.

94. Ribeiro JM, Francischetti IM. Role of arthropod saliva in blood feeding: sialome and post-sialome perspectives. Annu Rev Entomol. 2003; 48: 73–88.

95. Ribeiro JM, Mans BJ, Arca B. An insight into the sialome of bloodfeeding Nematocera. Insect Biochem Mol Biol. 2010; 40: 767–784.

96. Ribeiro JMC, Hazzard JMH, Nussenzveig RH, Champagne DE, Walker FA. Reversible binding of nitric oxide by a salivary heme protein by a bloodsucking insect. Science. 1993; 260: 539−541.

97. Rogers KA, Titus RG. Immunomodulatory effects of Maxadilan and *Phlebotomus papatasi* sand fly salivary gland lysates on human primary in vitro immune responses. Parasite Immunol. 2003; 25: 127–134.

98. Rohousova I, Ozensoy S, Ozbel Y, Volf P. Detection of species-specific antibody response of humans and mice bitten by sand flies. Parasitology. 2005; 130: 493–499.

99. Rohousova I, Subrahmanyam S, Volfova V, Mu J, Volf P, Valenzuela JG, et al. Salivary gland transcriptomes and proteomes of *Phlebotomus tobbi* and *Phlebotomus sergenti*, vectors of leishmaniasis. PLoS Negl Trop Dis. 2012a; 6: e1660.

100. Rohousova I, Talmi-Frank D, Kostalova T, Polanska N, Lestinova T, Kassahun A, et al. Exposure to *Leishmania* spp. and sand flies in domestic animals in northwestern Ethiopia. Parasit Vectors. 2015; 8: 360.

101. Rohousova I, Volfova V, Nova S, Volf P. Individual variability of salivary gland proteins in three *Phlebotomus* species. Acta Trop. 2012b; 122: 80–86.

102. Rose AS, Hildebrand PW. NGL Viewer: a web application for molecular visualization. Nucleic Acids Res. 2015; 43: W576–W579.

103. Samuelson J, Lerner E, Tesh R, Titus R. A mouse model of *Leishmania braziliensis braziliensis* infection produced by coinjection with sand fly saliva. J Exp Med. 1991; 173: 49–54.

104. Savoia D. Recent updates and perspectives on leishmaniasis. J Infect Dev Ctries. 2015; 9: 588–596.

105. Schorscher JA, Goris M. Incrimination of *Phlebotomus* (*Larroussius*) *orientalis* as a vector of visceral leishmaniasis in western Upper Nile Province, southern Sudan. Trans R Soc Trop Med Hyg. 1992; 86: 622–623.

106. Seblova V, Volfova V, Dvorak V, Pruzinova K, Votypka J, Kassahun A, et al. *Phlebotomus orientalis* sand flies from two geographically distant Ethiopian localities: biology, genetic analyses and susceptibility to *Leishmania donovani*. PLoS Negl Trop Dis. 2013; 7: e2187.

107. Seccombe AK, Ready PD, Huddleston LM. A Catalogue of Old World Phlebotomine sandflies (Diptera: Psychodidae, Phlebotominae). Occas Pap Syst Entomol. 1993; 8: 1–57.

108. Sima M, Ferencova B, Warburg A, Rohousova I, Volf P. Recombinant salivary proteins of *Phlebotomus orientalis* are suitable antigens to measure exposure of domestic animals to sand fly bites. PLoS Negl Trop Dis. 2016a; 10: e0004553.

109. Sima M, Novotny M, Pravda L, Sumova P, Rohousova I, Volf P. The diversity of yellow-related proteins in sand flies (Diptera: Psychodidae). PloS One. 2016b; Accepted.

110. Souza AP, Andrade BB, Aquino D, Entringer P, Miranda JC, Alcantara R, et al. Using recombinant proteins from *Lutzomyia longipalpis* saliva to estimate human vector exposure in visceral leishmaniasis endemic areas. PLoS Negl Trop Dis. 2010; 4: e649.

111. Tavares NM, Silva RA, Costa DJ, Pitombo MA, Fukutani KF, Miranda JC, et al. *Lutzomyia longipalpis* saliva or salivary protein LJM19 protects against *Leishmania braziliensis* and the saliva of its vector, *Lutzomyia intermedia*. PLoS Negl Trop Dis. 2011; 5: e1169.

112. Teixeira C, Gomes R, Collin N, Reynoso D, Jochim R, Oliveira F, et al. Discovery of markers of exposure specific to bites of *Lutzomyia longipalpis*, the vector of *Leishmania infantum chagasi* in Latin America. PLoS Negl Trop Dis. 2010; 4: e638.

113. Teixeira C, Gomes R, Oliveira F, Meneses C, Gilmore DC, Elnaiem DE, et al. Characterization of the early inflammatory infiltrate at the feeding site of infected sand flies in mice protected from vector-transmitted *Leishmania major* by exposure to uninfected bites. PLoS Negl Trop Dis. 2014; 8: e2781.

114. Theodos CM, Ribeiro JM, Titus RG. Analysis of enhancing effect of sand fly saliva on *Leishmania* infection in mice. Infect Immun. 1991; 59: 1592–1598.

115. Titus RG, Ribeiro JM. Salivary gland lysates from the sand fly *Lutzomyia longipalpis* enhance *Leishmania* infectivity. Science. 1988; 239: 1306–1308.

116. Valenzuela JG, Belkaid Y, Garfield MK, Mendez S, Kamhawi S, Rowton ED. Toward a defined anti-*Leishmania* vaccine targeting vector antigens: characterization of a protective salivary protein. J Exp Med. 2001; 194: 331–342.

117. Valenzuela JG, Charlab R, Galperin MY, Ribeiro JM. Purification, cloning and expression of an apyrase from the bed bug *Cimex lectularius*. A new type of nucleotide-binding enzyme. J Biol Chem. 1998; 273: 30583–30590.

118. Valenzuela JG, Garfield M, Rowton ED, Pham, VM. Identification of the most abundant secreted proteins from the salivary glands of the sand fly *Lutzomyia longipalpis*, vector of *Leishmania chagasi*. J Exp Biol. 2004; 207: 3717–3729.

119. Vinhas V, Andrade BB, Paes F, Bomura A, Clarencio J, Miranda JC, et al. Human anti-saliva immune response following experimental exposure to the visceral leishmaniasis vector, *Lutzomyia longipalpis*. Eur J Immunol. 2007; 37: 3111–3121.

120. Vlkova M, Rohousova I, Drahota J, Stanneck D, Kruedewagen EM, Mencke N, et al. Canine antibody response to *Phlebotomus perniciosus* bites negatively correlates with the risk of *Leishmania infantum* transmission. PLoS Negl Trop Dis. 2011; 5: e1344.

121. Vlkova M, Rohousova I, Hostomska J, Pohankova L, Zidkova L, Drahota J, et al. Kinetics of antibody response in BALB/c and C57BL/6 mice bitten by *Phlebotomus papatasi*. PLoS Negl Trop Dis. 2012; 6: e1719.

122. Vlkova M, Sima M, Rohousova I, Kostalova T, Sumova P, Volfova V, et al. Comparative analysis of salivary gland transcriptomes of *Phlebotomus orientalis* sand flies from endemic and non-endemic foci of visceral leishmaniasis. PLoS Negl Trop Dis. 2014; 8: e2709.

123. Volf P, Tesarova P, Nohynkova E. Salivary proteins and glycoproteins in phlebotomine sandflies of various species, sex and age. Med Vet Entomol. 2000; 14: 251–256.

124. Volf P, Volfova V. Establishment and maintenance of sand fly colonies. J Vector Ecol. 2011; 36: S1–S9.

125. Weichsel A, Andersen JF, Champagne DE, Walker FA, Montfort WR. Crystal structures of a nitric oxide transport protein from a blood-sucking insect. Nat Struct Biol. 1998; 5: 304–309.

126. Weinkopff T, de Oliveira CI, de Carvalho AM, Hauyon-La Torre Y, Muniz AC, Miranda JC, et al. Repeated exposure to *Lutzomyia intermedia* sand fly saliva induces local expression of interferon-inducible genes both at the site of injection in mice and in human blood. PLOS Negl Trop Dis. 2014; 8: e2627.

127. WHO 2016. http://www.who.int/mediacentre/factsheets/fs375/en/. Retrieved on October 19th 2016.

128. Wu S, Peiffer M, Luthe DS, Felton GW. ATP hydrolyzing salivary enzymes of caterpillars suppress plant defenses. PLoS One. 2012; 7: e41947.

129. Xanthos DN, Bennett GJ, Coderre TJ. Norepinephrine-induced nociception and vasoconstrictor hypersensitivity in rats with chronic post-ischemia pain. Pain. 2008; 137: 640–651.

130. Xu X, Chang BW, Mans BJ, Ribeiro JM, Andersen JF. Structure and ligand-binding properties of the biogenic amine-binding protein from the saliva of a blood-feeding insect vector of *Trypanosoma cruzi*. Acta Crystallogr. 2013; 69: 105–113.

131. Xu X, Oliveira F, Chang BW, Collin N, Gomes R, Teixeira C, et al. Structure and function of a ''yellow'' protein from saliva of the sand fly *Lutzomyia longipalpis* that confers protective immunity against *Leishmania major* infection. J Biol Chem. 2011; 286: 32383–32393.

132. Yeats C, Bentley S, Bateman A. New knowledge from old: In silico discovery of novel protein domains in *Streptomyces coelicolor*. BMC Microbiol. 2003; 3: 3.

133. Zer R, Yaroslavski I, Rosen L, Warburg A. Effect of sand fly saliva on *Leishmania* uptake by murine macrophages. J Int Parasitol. 2001; 31: 810–814.