

New Molecular Markers for Phlebotomine Sand Flies

A.A. Peixoto^{1*}, C. A. Gomes¹, P. R. de Amoretty¹, R.M.M.A. Lins¹, A.C.A. Meireles-Filho¹,
N. A. de Souza², C. P. Kyriacou³

¹ Department of Biochemistry and Molecular Biology, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

² Department of Entomology, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

³ Department of Genetics, University of Leicester, Leicester, U.K.

*Contact details corresponding author

Departamento de Bioquímica e Biologia Molecular

Fundação Oswaldo Cruz

Av. Brasil 4365 - Manguinhos

CEP 21045-900

Rio de Janeiro

Brazil

tel: (55)(21) 290-7549

fax: (55)(21) 590-3495

e-mail: apeixoto@gene.dbbm.fiocruz.br

Abstract

Using degenerate-primers PCR we isolated and sequenced fragments from the sand fly *Lutzomyia longipalpis* homologous to two behavioural genes in *Drosophila*, *cacophony* and *period*. In addition we identified a number of other gene fragments that show homology to genes previously cloned in *Drosophila*. A codon usage table for *L. longipalpis* based on these and other genes was calculated. These new molecular markers will be useful in population genetics and evolutionary studies in phlebotomine sand flies and in establishing a preliminary genetic map in these important leishmaniasis vectors.

Keywords: *Lutzomyia*, *cacophony*, *period*, molecular markers

Sand flies of the genus *Lutzomyia* (Diptera:Psychodidae:Phlebotominae) are vectors of leishmaniasis in the Americas (Young and Duncan, 1994; Tesh and Guzman, 1996). Despite their medical importance, there is a paucity of molecular markers in the species of this genus and related genera. So far, only a handful of genes such as those expressed in salivary glands have been characterised (Charlab et al., 1999). Behaviour is one aspect of sand fly biology where no molecular data is yet available. One area where behavioural genes are particularly important is in speciation. Studies of some of the most important vectors species, such as *Lutzomyia longipalpis*, suggest that they might represent complexes of cryptic species (Ward et al., 1988; Lanzaro et al., 1993; Lampo et al., 1999; Uribe, 1999). The occurrence of sibling species in sand flies has important epidemiological consequences (Lanzaro and Warburg, 1995) as they seem to differ in vector competence, habitat (peri-domestic x silvatic) and host preference (anthropophily x zoophily).

One of the most important forms of reproductive isolation between closely related species involves differences in courtship behaviour. In a great number of *Drosophila* species, males vibrate their wings during courtship producing a lovesong (Hall, 1994) that seems to increase the receptivity of females and has been implicated as one of the signals they used to recognise conspecific males (Kyriacou and Hall, 1982; 1986). This lovesong is therefore potentially important in the reproductive isolation between closely related species and, as a consequence, in the process of speciation. A few *Drosophila* genes controlling the courtship song have been identified and cloned (Hall, 1994).

cacophony is a voltage-gated calcium channel gene, also known as *Dmca1A* (Smith et al., 1996; 1998; Peixoto et al., 1997), that shows homology to mammalian calcium channel class A (Stea et al., 1995) and related channels B and E. *cacophony* is particularly interesting from an evolutionary point of view because one of its mutant alleles produces a song that resembles those of other *Drosophila* species (Peixoto and Hall, 1998). Using degenerate primer PCR we isolated a sand fly fragment homologous to the *Drosophila cacophony* gene. Fig. 1 shows the amino acid sequence encoded by a PCR fragment amplified from *L. longipalpis*, aligned to the corresponding region of the *Drosophila melanogaster cacophony* and other calcium channels. This fragment is ~270 bp and encodes the IVS6 domain of the channel. It also includes an intron (~100 bp). Fig. 2 shows a phylogenetic tree where the amino acid sequence of the *L. longipalpis* fragment is compared to other mammalian and *Drosophila* calcium channels, and clusters with the sand fly homologue of *Drosophila's cac*.

Another locus that also has an effect in courtship is *period*, a gene encoding a clock protein that has been shown to control species specific courtship song rhythm differences between *D. melanogaster* and its sibling species *Drosophila simulans* (Kyriacou and Hall, 1986, Wheeler et al., 1991). *period* also controls circadian rhythms of activity in *Drosophila* (Konopka and Benzer,

1971). The behaviour of insect vector species plays an important role in the dynamics of the diseases they transmit (Klowden, 1996). The rhythmic patterns of activity and feeding presented by haematophagous insects, which are so important to disease transmission, are certainly controlled by the internal biological clock (Saunders, 1982), which is in turn under genetic control (Dunlap, 1999). Using the same strategy mentioned above we isolated a sand fly fragment homologous to the *Drosophila period* (*per*) gene. Fig. 3 shows an alignment of the putative amino acid sequences encoded by PCR fragments obtained from *Lutzomyia longipalpis* and the same region of the *per* gene of *Drosophila virilis*. The sand fly fragment is nearly 1 Kb long, includes an intron and was obtained using two pairs of degenerate primers that amplified two overlapping fragments. Its predicted amino acid sequence extends from the end of the PAS protein dimerisation domain to the Thr-Gly repeat region. These two domains are particularly important to *period* function and evolution (Costa and Kyriacou, 1998; Dunlap, 1999).

In addition to these two behavioural genes, a number of other fragments were isolated which show homology to genes previously cloned in *Drosophila*. Table 1 gives a list of nine sand fly molecular markers (including *period* and *cacophony*) for which highly significant similarity to *Drosophila* genes were observed in BLASTX (Altschul et al., 1997) searches using the NCBI (National Center for Biotechnology Information) web page. Other fragments were also obtained, but not included as no clear putative *Drosophila* homologues were found. The table also gives the size of fragments, level of amino-acid identity and similarity of predicted encoded proteins and their putative functions in sand flies based on their role in *Drosophila*.

One molecular marker particularly interesting is fragment 6, which shows homology to *purity of essence*, a *Drosophila* gene coding for a calmodulin binding protein. Mutations in this locus affect behaviour and synaptic transmission at the neuromuscular junction. Mutant flies are sluggish and uncoordinated (Richards et al, 1996). In addition, mutant alleles also cause male sterility (Fabrizio et al., 1998).

Of the nine markers listed, three were obtained from *Lutzomyia intermedia*, an important vector of cutaneous leishmaniasis in Brazil (Young and Duncan, 1994). One of these shares homology to *Dmca1D* (Zheng et al, 1995), a *Drosophila* voltage-gated calcium channel different from *cacophony* and more similar to mammalian channel class D (see figs. 1 and 2). Like *L. longipalpis cacophony* homologue, this fragment encodes the IVS6 domain of the channel and it also includes an intron. *Dmca1D* is highly expressed in the embryonic nervous system (Zheng et al, 1995). Some more severe mutations at this channel are lethal recessives and mutants die at the late embryonic stage, while hypomorphic alleles cause flies to have difficulty in eclosing from pupae (Eberl et al, 1998). It is not yet known whether this gene might have also an effect in behaviour.

Using the available *L. longipalpis* sequences from our work and the ones from Charlab et al.

(1999), we calculated the codon usage for this sand fly species. Table 2 shows the frequencies of the different codons used by *L. longipalpis* compared with the values observed in *Drosophila melanogaster*. Note that the majority of the amino acids, 15 out of 18, show differences in codon preference between the two species. This is consistent with the variation observed among different *Drosophila* species (Powell, 1997). Once more sequences are available, it will be possible to look at variation among genes as observed in other organisms.

We are currently attempting to isolate other sand fly molecular markers controlling different aspects of behaviour. These new molecular markers will be useful in population genetics and evolutionary studies. For example, based on the sequence of the *L. longipalpis* homologues of *period* and *cacophony*, we designed more specific primers to this species and we are now using them to study the differentiation among putative cryptic species of the *L. longipalpis* complex in Brazil (Oliveira et al, in press). We are also considering using the molecular markers available to initiate the construction of a preliminary genetic map in these important leishmaniasis vectors.

Acknowledgments

This work was supported by an International Research Development Award from the Wellcome Trust, with additional support from Faperj, Fiocruz, CNPq, and UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR). We would like to thank Dr. Yara Traub-Cseko for support in the initial stages of this work and Robson Costa da Silva for his technical assistance.

References

- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Ac. Res.* 25, 3389-3402.
- Charlab, R., Valenzuela, J.G., Rowton, E.D., Ribeiro, J.M., 1999. Toward an understanding of the biochemical and pharmacological complexity of the saliva of a hematophagous sand fly *Lutzomyia longipalpis*. *Proc. Natl. Acad. Sci. USA* 96, 15155-15160.
- Costa, R., Kyriacou, C.P., 1998. Functional and evolutionary implications of natural variation in clock genes. *Curr. Opin. Neurobiol.* 8, 659-664.
- Dunlap, J.C., 1999. Molecular Bases for Circadian Clocks. *Cell* 96, 271-290.
- Eberl, D.F., Ren, D., Feng, G., Lorenz, L.J., Vactor D.V., Hall, L.M., 1998. Genetic and developmental characterization of Dmca1D, a calcium channel alpha1 subunit gene in *Drosophila melanogaster*. *Genetics* 148, 1159-1169.

- Fabrizio, J.J., Hime, G., Lemmon, S.K., Bazinet, C., 1998. Genetic dissection of sperm individualization in *Drosophila melanogaster*. *Development* 125, 1833-1843
- Hall, J.C., 1994. The mating of a fly. *Science* 264, 1702-1714.
- Jowett, T., 1998. Preparation of nucleic acids. In: Roberts, D.B., (Ed.). *Drosophila: A practical approach*. IRL press, Oxford, pp. 347-371.
- Klowden, M., 1996. Vector Behavior. In: Beaty, B.J., Marquardt, W.C., (Ed.). *The biology of disease vectors*. University Press of Colorado, Colorado, pp. 34-50.
- Konopka, R.J., Benzer, S., 1971. Clock mutants of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* 68, 2112-2116.
- Kyriacou, C.P., Hall, J.C., 1982. The function of courtship song rhythms in *Drosophila*. *Anim. Behav.* 30, 784-801.
- Kyriacou, C.P., Hall, J.C., 1986. Interspecific genetic control of courtship song production and reception in *Drosophila*. *Science* 232, 494-497.
- Lampo, M., Torgerson, D., Marquez, L.M., Rinaldi, M., Garcia, C.Z., Arab, A., 1999. Occurrence of sibling species of *Lutzomyia longipalpis* (Diptera: Psychodidae) in Venezuela: first evidence from reproductively isolated sympatric populations. *Am. J. Trop. Med. Hyg.* 61, 1004-1009.
- Lanzaro, G.C., Ostrovska, K., Herrero, M.V., Lawyer, P.G., Warburg, A. *Lutzomyia longipalpis* is a species complex: genetic divergence and interspecific hybrid sterility among three populations. *Am. J. Trop. Med. Hyg.* 1993; 48, 839-847.
- Lanzaro, G.C., Warburg, A., 1995. Genetic variability in phlebotomine sandflies: Possible implications for leishmaniasis epidemiology. *Parasitol. Today* 11, 151-154.
- Oliveira, S.G., Bottecchia, M., Bauzer, L.G.S.R., de Souza, N.A., Ward, R.D., Kyriacou, C.P., Peixoto, A.A., in press. Courtship Song Genes and Speciation in Sand Flies. *Mem. Inst. Oswaldo Cruz*.
- Peixoto, A.A., Hall, J.C., 1998. Analysis of temperature-sensitive mutants reveals new genes involved in the courtship song of *Drosophila*. *Genetics* 148, 827-838.
- Peixoto, A.A., Smith, L.A., Hall, J.C., 1997. Genomic organization and evolution of alternative exons in a *Drosophila* calcium channel gene. *Genetics* 145, 1003-1013.
- Powell, J.R. 1997. *Progress and Prospects in Evolutionary Biology: The Drosophila Model*. Oxford University Press, Oxford.
- Richards, S., Hillman, T., Stern, M., 1996. Mutations in the *Drosophila* pushover gene confer increased neuronal excitability and spontaneous synaptic vesicle fusion. *Genetics* 142, 1215-1223
- Saunders, D.S., 1982. *Insect Clocks*. 2nd ed. Pergamon press, Oxford.
- Smith, L.A., Peixoto, A.A., Kramer, E.M., Villella, A., Hall, J.C., 1998. Courtship and visual defects of cacophony mutants reveal functional complexity of a calcium- channel α -1 subunit in *Drosophila*. *Genetics* 149, 1407-1426.

- Smith, L.A., Wang, X.J., Peixoto, A.A., Neumann, E.K., Hall, L.M., Hall, J.C., 1996. A *Drosophila* calcium channel α -1 subunit gene maps to a genetic locus associated with behavioral and visual defects. *J. Neuroscience* 16, 7868-7879.
- Stea, A., Soong, T.W., Snutch, T.P., 1995. Voltage-Gated Calcium Channels. In: North, R.A., (Ed.) *Ligand- and Voltage-gated Ion Channels*. CRC Press, Boca Raton, pp. 114-151.
- Tesh, R.B., Guzman H., 1995. Sand Flies and the agents they transmit. In: Beaty, B.J., Marquardt, W.C., (Ed). *The biology of disease vectors*. University Press of Colorado, Colorado, pp. 117-127.
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Ac. Res.* 24, 4876-4882.
- Uribe, S., 1999. The status of the *Lutzomyia longipalpis* species complex and possible implications for Leishmania transmission. *Mem. Inst. Oswaldo Cruz* 94, 729-734
- Ward, R.D., Phillips, A., Burnet, B., Marcondes, C.B., 1988. The *Lutzomyia longipalpis* complex: reproduction and distribution. In: M.W. Service, editor. *Biosystematics of Haematophagous Insects*. Oxford University Press, Oxford, pp. 258-269.
- Wheeler, D.A., Kyriacou, C.P., Greenacre, M.L., Yu, Q., Rutila, J.E., Rosbash, M., Hall, J.C., 1991. Molecular transfer of a species-specific behavior from *Drosophila simulans* to *Drosophila melanogaster*. *Science* 251, 1082-1085.
- Young, D.G., Duncan, M.A., 1994. Guide to the identification and geographic distribution of *Lutzomyia* sand flies in Mexico, the West Indies, Central and South America (Diptera:Psychodidae). *Mem. Amer. Ent. Inst.* 54, 1-881.
- Zheng, W., Feng, G., Ren, D., Eberl, D.F., Hannan, F., Dubald, M., Hall, L.M., 1995. Cloning and characterization of a calcium channel alpha 1 subunit from *Drosophila melanogaster* with similarity to the rat brain type D isoform. *J. Neuroscience* 15, 1132-1143.

Table 1.
New sand fly molecular markers.

Fragment [#]	Species	Size (bp)	Putative <i>Drosophila</i> Homologue	Protein sequence similarity and identity	Putative function
1	<i>L. longipalpis</i>	921*	<i>period</i>	76% / 63%	transcriptional regulator (clock gene)
2	<i>L. longipalpis</i>	225*	<i>cacophony (Dmca1A)</i>	87% / 82%	voltage-gated calcium channel α -1 subunit
3	<i>L. longipalpis</i>	470*	CG10046	100% / 100%	RNA binding protein
4	<i>L. longipalpis</i>	432	CG15828	51% / 34%	ligand binding or carrier
5	<i>L. longipalpis</i>	363	CG9649	51% / 36%	endopeptidase
6	<i>L. longipalpis</i>	273	<i>purity of essence</i>	94% / 84%	calmodulin binding protein
7	<i>L. intermedia</i>	291*	CG3003	60% / 43%	unknown
8	<i>L. intermedia</i>	249	CG6726	80% / 67%	peptidase
9	<i>L. intermedia</i>	226*	Dmca1D	81% / 71%	voltage-gated calcium channel α -1 subunit

[#] Genomic DNA of *Lutzomyia longipalpis* and *L. intermedia* was prepared according to Jowett (1998) or using the GenomicPrep from Amersham Pharmacia Biotech. From 50 ng to 100 ng of genomic DNA were used in the PCR. The reactions were carried out in a Perkin-Elmer GeneAmp PCR system 9600 for 35 cycles (95°C for 0.5 min; 37°C to 55°C for 1.0 min; 72°C to 1.0 min). In some cases reamplifications were also carried out. Different combinations of degenerate primers were used. PCR products were electrophoresed in 2% TAE agarose gels, bands were purified using the Sephaglas BandPrep kit Band (Amersham Pharmacia Biotech) and cloned into pMOSBlue vector. Sequencing was done automatically at Fundação Oswaldo Cruz (Rio de Janeiro, Brazil) and at University of Leicester (UK). The GCG software (Version 7, 1991) was used for DNA sequence editing and analysis. * These fragments include putative intron sequences. Protein sequence similarity and identity values were obtained from BLASTX (Altschul et al., 1997) searches and they are based on comparisons to *Drosophila melanogaster* in all cases, except for *period* where the sequence of *D. virilis* was used.

Table 2
Codon usage in *L. longipalpis*

Amino Acid	Codon	Frequency*	Amino Acid	Codon	Frequency
Gly	GGG	0.21 (0.06)	Ser	AGU	0.18 (0.12)
Gly	GGA	<i>0.43 (0.28)</i>	Ser	AGC	0.18 (<i>0.24</i>)
Gly	GGU	0.26 (0.23)	Ser	UCG	0.08 (0.22)
Gly	GGC	0.10 (<i>0.43</i>)	Ser	UCA	<i>0.22 (0.09)</i>
Glu	GAG	0.35 (<i>0.71</i>)	Ser	UCU	0.14 (0.08)
Glu	GAA	<i>0.65 (0.29)</i>	Ser	UCC	0.19 (0.25)
Asp	GAU	<i>0.66 (0.52)</i>	Arg	AGG	0.17 (0.10)
Asp	GAC	0.34 (0.48)	Arg	AGA	<i>0.25 (0.08)</i>
Val	GUG	0.22 (<i>0.47</i>)	Arg	CGG	0.04 (0.14)
Val	GUA	0.16 (0.10)	Arg	CGA	0.15 (0.14)
Val	GUU	<i>0.48 (0.18)</i>	Arg	CGU	0.22 (0.18)
Val	GUC	0.14 (0.25)	Arg	CGC	0.16 (<i>0.35</i>)
Ala	GCG	0.09 (0.18)	Gln	CAG	0.47 (<i>0.73</i>)
Ala	GCA	<i>0.36 (0.16)</i>	Gln	CAA	<i>0.53 (0.27)</i>
Ala	GCU	0.34 (0.19)	His	CAU	<i>0.71 (0.34)</i>
Ala	GCC	0.21 (<i>0.47</i>)	His	CAC	0.29 (<i>0.66</i>)
Lys	AAG	0.47 (<i>0.75</i>)	Leu	UUG	<i>0.21 (0.18)</i>
Lys	AAA	<i>0.53 (0.25)</i>	Leu	UUA	0.11 (0.04)
Asn	AAU	<i>0.74 (0.42)</i>	Leu	CUG	0.15 (<i>0.45</i>)
Asn	AAC	0.26 (<i>0.58</i>)	Leu	CUA	0.13 (0.08)
Ile	AUA	0.17 (0.16)	Leu	CUU	0.20 (0.09)
Ile	AUU	<i>0.54 (0.32)</i>	Leu	CUC	0.20 (0.16)
Ile	AUC	0.30 (<i>0.52</i>)	Pro	CCG	0.15 (0.30)
Thr	ACG	0.25 (0.24)	Pro	CCA	<i>0.44 (0.24)</i>
Thr	ACA	<i>0.35 (0.19)</i>	Pro	CCU	0.24 (0.11)
Thr	ACU	0.23 (0.15)	Pro	CCC	0.17 (<i>0.35</i>)
Thr	ACC	0.17 (<i>0.42</i>)	Tyr	UAU	0.45 (0.34)
Cys	UGU	0.49 (0.27)	Tyr	UAC	<i>0.55 (0.66)</i>
Cys	UGC	<i>0.51 (0.73)</i>	Phe	UUU	0.52 (0.33)
Met	AUG	1.00 (1.00)	Phe	UUC	0.48 (<i>0.67</i>)
Trp	UGG	1.00 (1.00)	End	UAG	0.10 (0.28)
			End	UAA	<i>0.60 (0.53)</i>
			End	UGA	0.30 (0.18)

* The frequencies are calculated from a total of 3683 codons. Values in brackets are from *D. melanogaster* (Powell, 1997) and they are shown here for comparison. The most frequent used codons in both species are in italics.

Legends to Figures

Fig. 1. Alignment of the amino acid sequence encoded by a fragment obtained from *L. longipalpis* (called here Llca1A) compared to the IVS6 domain of mammalian (rat alpha-1A, human alpha-1D and rat alpha-1G) and *Drosophila* (Dmca1A and Dmca1D) voltage-gated calcium channels. The amino acid sequence encoded by a fragment obtained from *L. intermedia* (called here Lica1D) is also included (see also Table 1). The software ClustalX (Thompson et al., 1997) was used for the alignment. Accession numbers of protein sequences used: P54282 (rat alpha-1A), Q01668 (human alpha-1D), O54898 (rat alpha-1G), P91645 (Dmca1A or *cacophony*), Q24270 (Dmca1D).

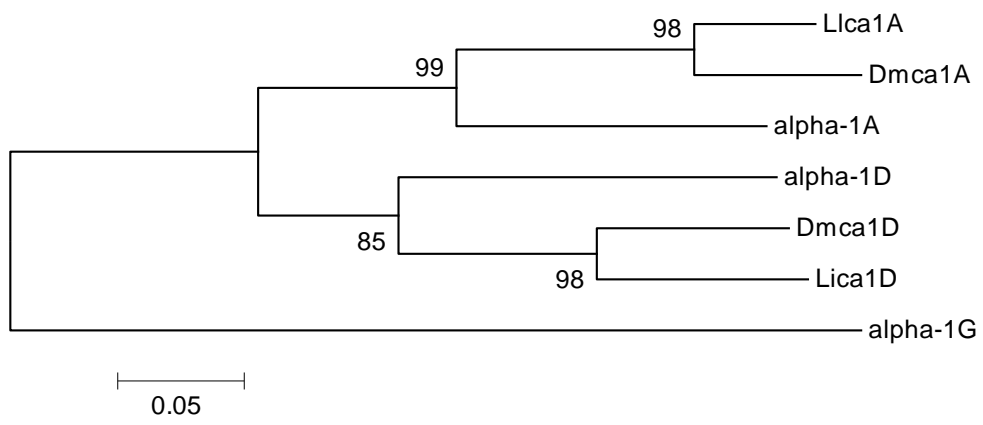
Fig. 2. Phylogenetic tree using sequences of Fig. 1. The “neighbour-joining” method available at ClustalX software (Thompson et al., 1997) was used to construct the tree. The numbers on nodes represent the “bootstrap” percentage values based on a 1000 replicates.

Fig. 3. Alignment of the amino acid sequence encoded by a fragment obtained from *L. longipalpis* compared to the protein sequence encoded by the *period* gene of *Drosophila virilis* (accession number S02035).

```

Llca1A      PNIMLACLKGR---PCDERAGKEPHE--TC-GSSLAYAYFVSFIFFCSFLMLNLFVAV
Dmca1A      PNIMLACLK GK---ACDDDAEKAPGE--YC-GSTLAYAYFVSFIFFCSFLMLNLFVAV
alpha-1A    HNIMLSCLSGK---PCDKNSGIQKPE---C-GNEFAYFYFVSFIFLCSFLMLNLFVAV
alpha-1D    QEIMLACLPGK---LCDPESDYNPGEETHC-GSNFAIVYFISFYMLCAFLIINLFVAV
Dmca1D      QEIMMSCSAQP-DVKCDMNSDTPG---EPC-GSSIAYPYFISFYVLC SFLIINLFVAV
Lica1D      QDIMMDCSSRPGEVNCDDRSDDRGSK-DGC-GSSIAPYFISFYVLC SFLIINLFVAV
alpha-1G    NGIMKDTLR-----DCDQEST-----CYNTVISPIYFVSFVLTAQFVLVNVVIAV
          **           **      :           * .. ::  **:** . . *::*:**

```



D. virilis LIGRSILDFYHHEDLSDIKDIYEKVVKKQTVGATFCSKPFRLFQNGCYILLETEWTSF
L. longipalpis MIGRSIMDFYHPEDFSYLREVIYETVMRVGKTAGASFCSKPYRFLAHNGFYITLETEWTSF
:*****:**** *:* ::::*. *.: *:*.*:*****:*** :** ** *****

L. longipalpis VNPWSRQLEFVIGHHRVLRGSPNPQVFAST-LVNQQFSEDVLNDAKINQEKILCLLTFV
D. virilis VNPWSRKLEFVVGHRVFGPKQCDVFEMSPNVTPIPEDEQNRNACIKEDILKMMTETV
*****:****:*****:*. : ** : * . :. ** * :*. ** :*. *

D. virilis TRPSDTPVKQEVSRRCQALASFMETLMDEVARGDLKLDLPHETELTVSERDSVMLGEISPH
L. longipalpis SKDIDTVKQQVSKRCLALASFMETLMDEVTRPDLKLDLPQETELTISERDSVMLGEISPH
: : *****:*. * *****: * *****:*****:*****

D. virilis HDYYDSKSSSTETPPSYNQLNYNENLLRFFNSKPVTAPVDTDPPKMDSSYVSSAR-EDALS
L. longipalpis HDYYDSKSSSETPPSYNQLNYNENLQRFFESKPITIGPDEAMKVEHTEPESTGDPQNSLS
*****:*****:***** * * * * . : . * : . : : **

D. virilis PVHGFEGSGGSGSSGNLTTASNVRMSSVTNTSNTGTGTSGGENSAGSSNPLPVNMTLTE
L. longipalpis PVQCF-GSG-SGSAGNLSSGSNIQMSMT--SNTGTGTS-----SGSYQP-P---ALTE
* : * ** * * : * * : . * * : * . *

D. virilis ILLN
L. longipalpis SLLS
* * .