HAL
open science

Development of Exon-Primed Intron-Crossing (EPIC) PCR primers for the malaria vector Anopheles pseudopunctipennis (Diptera : Culicidae)

Frédéric Lardeux, Claudia Aliaga, Rosenka Tejerina, Raul Ursic-Bedoya

To cite this version:

Frédéric Lardeux, Claudia Aliaga, Rosenka Tejerina, Raul Ursic-Bedoya. Development of Exon-Primed Intron-Crossing (EPIC) PCR primers for the malaria vector Anopheles pseudopunctipennis (Diptera : Culicidae). Comptes Rendus Biologies, 2012, 335 (6), pp.398-405. 10.1016/j.crvi.2012.05.002 . hal-01254144

HAL Id: hal-01254144

https://hal.science/hal-01254144v1

Submitted on 13 Jan 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Post-print version of te article published in COMPTES RENDUS BIOLOGIES - 2012- vol 335: 398-405

Development of Exon-Primed Intron-Crossing (EPIC) PCR Primers for the Malaria Vector *Anopheles pseudopunctipennis* (Diptera: Culicidae)

Développement d'Amorces PCR Exon-Primed Intron-Crossing (EPIC) pour le vecteur de paludisme *Anopheles pseudopunctipennis* (Diptera: Culicidae)

Frédéric Lardeux,[1,2,*] Claudia Aliaga,[1,2] Rosenka Tejerina,[1,2], Raùl Ursic-Bedoya[3]

[1] Institut de Recherche pour le Développement (IRD), C.P. 9214, La Paz, Bolivia

[2] Instituto Nacional de Laboratorios de Salud (INLASA), Laboratorio de Entomologia Medica, Rafael Zubieta n° 1889, Casilla M-10019, Miraflores, La Paz, Bolivia

[3] Simon Fraser University, Department of Biological Sciences, 8888 University Dr., Burnaby, BC V5A1S6, Canada.

* Corresponding Author: IRD, C.P. 9214, La Paz, Bolivia.
e.mail: frederic.lardeux@ird.fr
Tel : +591-2-222 52 80
Fax : +591-2-278 29 44

RESUME Utilisant le génome *d'Anopheles gambiae* Giles comme modèle, nous avons identifié 14 nouvelles Exon-Primed Intron-Crossing (EPIC)-paires d'amorces PCR pour *Anopheles pseudopunctipennis* Theobald 1901, un vecteur majeur de *Plasmodium sp*. humains en Amérique du Sud. Ces amorces ont été conçues afin de cibler les régions conservées flanquant les exons consécutifs de différents gènes et ont permis l'amplification de 17 loci dont neuf étaient polymorphes. Le polymorphisme de ces loci varie de deux à quatre allèles. L'analyse du polymorphisme de longueur d'intron est un outil utile qui permettra l'étude de la structure de la population de cette espèce de moustique qui demeure mal comprise.

MOTS-CLES EPIC-PCR, *Anopheles pseudopunctipennis,* paludisme, amorce, intron

ABSTRACT Using the *Anopheles gambiae* Giles genome as a template, we designed, screened and identified 14 novel Exon-Primed Intron-Crossing (EPIC)- PCR primer pairs for *Anopheles pseudopunctipennis* Theobald 1901, a major vector of human *Plasmodium sp*. in South America. These primers were designed to target the conserved regions flanking consecutive exons of different genes and enabled the amplification of 17 loci of which nine were polymorphic. Polymorphisms at these loci ranged from two to four alleles. Intron length polymorphism analysis is a useful tool which will allow the study of the population structure of this mosquito species which remains poorly understood.

KEY WORDS EPIC-PCR, *Anopheles pseudopunctipennis,* malaria, primer, intron

1. Introduction

The mosquito *Anopheles pseudopunctipennis* Theobald 1901 (Diptera: Culicidae) is a major vector of human *Plasmodium sp.* in South America. It is a difficult species to characterize as demonstrated by its variable behavior, habits and ecological needs [1], its inconsistencies as a malaria vector in its wide distribution range [2] and the several morphological sub-species described [3]. Crossmating experiments and cytogenetic pointed out evidence that these mosquitoes are comprised of a species complex [4, 5]. Although the genetic variability and population structure of *An. pseudopunctipennis* has been studied biochemically, many aspects remain poorly known [6]. This species lacks specific molecular tools, and because of its high level of variability more data are urgently needed to better understand how the population genetic structure of this mosquito is related to malaria transmission. Such data will help to better target vector control strategies by Public Health authorities.

Variability of non coding sequences, particularly intron sequences, is a valuable marker of population variation and subdivision, and can be assayed by PCR amplification using conserved exon primers. Intron-targeted PCR was pioneered by [7]. This approach, called Exon-Primed Intron-Crossing (EPIC)-PCR [8, 9], has been shown identify substantial variability, mainly from intron length polymorphism. Intron constitute suitable markers for analyzing population structure within a species [10, 11, 12, 13, 14, 15] as well as for reconstructing relationships among closely related species [16, 17, 18]. EPICs have been also used in gene mapping [19, 20, 21, 22, 23], and phylogeography [24, 25, 26, 27, 28] where they are the most widely used nuclear markers for such studies [29]. EPIC-PCR has several advantages in populations genetic studies: (i) by using primers from heterologous genes, cloning and sequencing steps can be avoided [30, 31]; (ii) cross-species amplification is easier than when primers are designed from highly conserved exon sequences, (iii) for the same

reason, within species, PCR artefacts such as null alleles are expected to be less frequent. Furthers advantages are that intron systems do not require previous knowledge of the genome to be analyzed; they are generally polymorphic and sometimes hypervariable; are expected to be codominant and selectively neutral; are easily amplifiable by PCR; can be revealed on simple agarose or acrylamide gels; and can be obtained at low cost [32]. Moreover, having both the exon and intron fragments, EPICs can be useful for examining genetic variation at the intraspecific and interspecific level simultaneously, a feature that is particularly useful when studying species complexes [27]. EPIC markers are becoming more popular for use in population genetic studies in insects [13, 14, 15, 33, 34] and do not require assumptions about a particular model of evolution that is often required for microsatellites [14]. In the present paper, primers for EPIC amplification of intron sequences for *An. pseudopunctipennis* are designed, of which several pairs amplify (length)-polymorphic loci that can be used in population genetic studies of this malaria vector.

2. Materials and methods

2.1. Mosquitoes

In October 2006, females *An. pseudopunctipennis* were captured by one of us (FL) from a natural population in Mataral (S 18.6024, W 65.1444, altitude 1500 m), a small village situated in the dry inter-Andean valleys in the centre of Bolivia, using human bait collection technique outside houses. In the field, collected insects were chloroform killed and stored over desiccant (silica gel) in small vials. In the laboratory, mosquitoes were identified using [35] and kept at -20°C in their individual vials with silica gel until DNA extraction.

For the various EPIC-PCRs (see paragraph below), positive controls consisted in females *An. gambiae* from a laboratory strain and were provided by our main laboratory at IRD-

Montpellier (France). These mosquitoes were stored using the same conditions as for *An. pseudopunctipennis* until processing.

2.2. Selection of introns and design of primers

Primers for *An. pseudopunctipennis* were designed from the conserved regions of consecutive exons of different genes from the closely related species *Anopheles gambiae* Giles. Exon sequences were downloaded in 2005 at http://www.anobase.com in Excel format from the *An. gambiae* genome database. Gene candidates that were dispersed amongst the *An. gambiae* genome were selected, and close genes on the same chromosome were avoided. Gene candidates with a higher percentage of similarity with genes from Diptera *Apis mellifera* and *Drosophila melanogaster*, were first selected to enhance the chance of similarity to genes from *An. pseudopucntipennis*.

Intron length's ranged from 100 to 500 bp. Primers pairs were designed to target the flanking exon sequences taking into account their stability (in terms of CG content and ending with CG or GC), their size (18 – 20 pb), their close annealing temperatures, and the low probability of primer-dimer formation during the PCR. Possible primers adjacent to the intron sequences were discarded. Fifty four primer pairs were initially designed and screened by PCR using *An. pseudopunctipennis* DNA as a template (Table 1).

2.3. DNA extraction and amplification

An. pseudopunctipennis and *An. gambiae* DNA extractions were carried out on mosquito legs using a slightly modified CTAB (cetyltrimethylammonium bromide) based protocol from [36]. The protocol was as followed: mosquito legs were homogenized in 200µl lysis CTAB solution (100 mmol/l Tris HCl pH 8.0; 10 mmol/l EDTA pH 8.0; 1.4 mol/l NaCl and CTAB 2%) in 1.5 ml Eppendorf microcentrifuge tubes. Incubation was carried out at 65°C for 15

min; the resulting extract was washed with 200 µl chloroform and centrifugated for 5 min at 12 000 rpm. The supernatant was precipitated in 200 µl isopropanol and centrifugated again at 12 000 rpm for 15 min. The pellet was washed with 200 µl 70% ethanol, centrifugated at 12 000 rpm for 5 min, dried at 37°C for one hour and suspended in 100 µl nuclease-free H_20. DNA amplifications were carried out immediately after extraction in volumes of 25 µl (1x*Taq* buffer, 2.5 mM $MgCl_2$, 0.4 mM dNTPs (Eurogentec, Angers, France), 0.5 UI *Taq* polymerase (Quiagen, Courtaboeuf, France), 20-25 ng of DNA template, and depending of the locus 0.04 µM or 0.4 µM of each primer (Eurogentec, Angers, France) (Table 2). The optimum annealing temperatures for each primer pair are listed in Table 2. PCRs were performed on a Perkin Elmer DNA Thermal Cycler 480 (US Instrument Division, Norwalk, CT, USA) and conditions were: 1 min at 94°C, followed by 36 cycles of 30s at 94°C, 30s at annealing temperature, 30 s at 72°C, and a final extension step of 5 min at 72°C. The amplified products were first visualized on 1.5% agarose gels con ethidium bromide (Fig. 1.). Then, for allele size analysis, they were separated by electrophoresis on 8% polyacrylamide gels and visualized by silver-staining. In all PCR, negative control (H_2O) and positive controls (*An. gambiae*) were used.

2.4. Data analysis

Allele sizes were scored using the numerical procedure implemented in LabImage version 3.0 [37] using a size-standard 100 bp gene ruler. Observed (H_O) and expected heterozygosities (H_E), the inbreeding coefficients (*Fis* statistics) and the R^2 coefficient to estimate linkage disequilibrium for each locus pairs were computed using procedures implemented in the GENETIX package [38].

3. Results and Discussion

Of the 54 selected pairs of primers, 14 pairs successfully amplified 17 loci (Table 2). All the 54 pairs of primers successfully amplified loci in *An. gambiae.* Forty pairs of primer did not amplify any loci in *An. pseudopunctipennis* while they did in *An. gambiae.* As an example, Fig. 1 presents results of agarose gel electrophoresis for four primers pairs. Fig 1a. shows positive amplifications in both *An. pseudopunctipennis* and *An. gambiae* while Fig. 1 b. gives examples of negative amplifications in *An. pseudopunctipennis* but positive ones in *An. gambiae* (positive controls). Because EPICs are universal primers they may amplify introns in other species. Table 3 is based on nucleotide BLAST (NCBI web site) of the 54 studied primers pairs and exhibits those that might amplify in a variety of other species including other mosquitoes (*Aedes aegypti* and *Ae. albopictus,* other 5 species of *Anopheles, Culex quinquefasciatus*), other Diptera (12 species of *Drosophila*), other insects (the cricket species *Gryllus bimaculatus*, the moths of the Noctuidae family *Helicoverpa armigera* and *Heliotis viriplaca*), the deer tick *Ixodes scapularis*, vertebrates (Human, dog, mouse), a choanoflagelate *Monosiga brevicolis,* a cyanobacteria *Thermosynechococcus elongates,* a Dermatophytes zoophilic fungus *Trichophyton verrucosum* and even a plant, the common plantain *Plantago major*. Some primer pairs are more universal than others in the sense that, for example, primer pair n°4 (intron G3LEX28), primer pair n° 11 (intron G3LEX36) or primer pair n° 21 might amplify in 16, 11 and 7 species respectively. Of all the primer pairs studied, 11 might amplify in *Aedes aegypti*, the main vector of dengue viruses and 15 in *Culex quinquefasciatus*, an annoying species also vector of human filariasis and West Nile virus (Table 3).

All the 17 loci amplified in *An. pseudopunctipennis* were tested for their polymorphism level using 20 *An. pseudopunctipennis.* Nine of the 17 amplified loci were polymorphic, ranging from two to four alleles (Table 2).

Previous studies relying on EPIC-PCR reported that various loci might be scored for a given pair of primers [11, 32, 39, 40, 41]. This phenomenon may reflect former polyploidizations, tandem duplications, and other phenomena occurring during lineage evolution and producing pseudogenes. Since *An. pseudopunctipennis* is a diploid species, gene duplication processes are more likely involved. The simultaneous amplification of two or three loci (two for G2LEX1236 and three for GUKEX1859) can then reveal the presence of another gene or a pseudogene. These simultaneous amplifications did not seem to disturb the quality of PCRs and the reading of the genotypes because (1) the supplementary profiles did not present parasitic bands and (2) the variation of size between band systems was larger than the size polymorphism of each locus. Indeed, when intra-specific allelic variants occur, a locus can be evidenced by the presence of all the allelic combinations following a Mendelian inheritance (homozygotes and heterozygotes); if not, in order to circumvent erroneous allelic assignment, only loci where the allelic size-variation is lower between orthologous than paralogous loci, should be retained [11].

With EPIC-PCR, artifacts such as null alleles are expected to be less frequent than for example, with microsatellites [42]. However, Table 2 exhibits some differences in the number of mosquitoes that correctly responded to amplification, in particular for primer pairs n° 6 and 7 (Intron G3REX1037 and EX113 respectively, with 12 and 14 mosquitoes respectively instead of the 20 expected). A careful examination of results seemed to indicate the presence of null alleles for these primer pairs and moreover, no heterozygotes were encountered (*Fis*=1). Therefore, these two primer pairs may not be suitable for population genetic studies. The lack of amplification in two mosquitoes for primers pairs n°1, 2, 4 and 5 are more likely the fact of PCR manipulation problems (low DNA concentration ?) because the failure of amplification with the four above primer pairs occurred with the same two mosquitoes. Null alleles are therefore not suspected and these primer pairs may be proposed for genetic studies.

For the small population under study, the R^2 coefficients to estimate linkage disequilibrium for each locus pair are presented in Table 4. Strong linkage disequilibrium appeared between G2REX47 and most of the other loci.

Compared with other DNA-based techniques such as microsatellites, EPICPCR is the only technique based on universal primers that allows a fast screening even for cross-species studies. Because *An. pseudopunctipennis* is likely a species complex [4, 5], with at least five described subspecies [43] and two recognized morphological variants in Bolivian larvae [44], the above EPIC primers may therefore be used to better understand the species status.

Another interesting characteristic of EPIC –PCR is that, likely because gene duplication, one single pair of primer can produce more than one locus and can provide a number of polymorphisms, as it was the case with primer pair n°3 (intron GUKEX1859) in *An. pseudopunctipennis*.

Primer pairs n° 2 and 4 indicate (at least in the small sample of 20 mosquitoes from Mataral village) that the genotype frequencies conform to Hardy-Weinberg equilibrium. Therefore, polymorphism in these introns may be selectively neutral, as predicted for variation in most non-coding DNA sequences. If so, such markers are particularly powerful in population genetics studies.

Population genetics of *An. pseudopunctipennis* will benefit of the above EPIC markers, and using recently available methods [27, 42], other EPIC markers could be isolated more easily.

Disclosure of interest

The authors declare that they have no conflicts of interest concerning this article.

Acknowledgments

This project was partly supported by a PAL+ Grant from the Ministère de la Recherche (France) and a Contrat de Développement Grant from Ministère des Affaires Etrangères (France). The participation of R. Ursic-Bedoya in this study was possible thanks to the Association of Universities and Colleges of Canada (AUCC) and its Canada-Latin America and the Caribbean Research Exchange Grants. The authors would like to thank P. Kengne who encouraged this study and an anonymous referee whose comments helped us to greatly improve the manuscript.

References

[1] L.J. Bruce-Schwatt, Essential Malariology, 2nd Ed, John Wiley & Sons, New York, 1985.

[2] T.G.H. Aitken, Studies on the anopheline complex of Western America. Univ.Calif. Pub. Entomol. 7 (1945) 273-364.

[3] K.L. Knight, A. Stone, A Catalog of the Mosquitoes of the World (Diptera: Culicidae). 2nd Ed. Vol. 6. The Thomas Say Foundation, Lanham, MD, 1977.

[4] J.G. Estrada-Franco, M.C. Ma, R.W. Gwadz, R. Sakai, G.C. Lanzaro, A. Laughinghouse, C. Galvan-Sanchez, J.L. Cespedes, R. Vargas-Sagarnaga, Evidence through crossmating experiments of a species complex in *Anopheles pseudopunctipennis* sensu lato: a primary malaria vector of the American continent. Am. J. Trop. Med. Hyg. 49 (1993) 746-55.

[5] M. Coetzee, J.G. Estrada-Franco, C.A. Wunderlich, R.H. Hunt, Cytogenetic evidence for a species complex within *Anopheles pseudopunctipennis* Theobald (Diptera: Culicidae). Am. J. Trop. Med. Hyg. 60 (1999) 649-653.

[6] S. Manguin, D.R. Roberts, E.L. Peyton, I. Fernandez-Salas, M. Barreto, R. Fernandez-Loayza, R. Elqueta-Spinola, R. Martinez-Granaou, M.H. Rodriguez, Biochemical systematics and population genetic structure of *Anopheles pseudopunctipennis*, vector of malaria in Central and South America. Am. J. Trop. Med. Hyg. 53 (1995) 362-377.

[7] E.P. Lessa, Rapid survey of DNA sequence variation in natural populations. Mol. Biol. Evol. 9 (1992) 323-330.

[8] S.R. Palumbi, C.S. Baker, Contrasting population structure from nuclear intron sequences and mtDNA of humpback whales. Mol. Biol. Evol. 11 (1994) 426-435.

[9] R.W. Slade, C. Moritz, A. Heideman, P.T. Hale, Rapid assessment of single-copy nuclear DNA variation in diverse species. Mol. Ecol. 2 (1993) 359-373.

[10] P. Berrebi, X. Retif, F. Fang, C.G. Zhang, Population structure and systematics of *Opsariichthys bidens* (Osteichthyes: Cyprinidae) in south-east China using a new nuclear marker: the introns (EPIC-PCR). Biol. J. Linn. Soc. 87 (2006) 155-166.

[11] N. Hubert, F. Duponchelle, H. Nunez, R. Rivera, J.-F. Renno, Evidence of reproductive isolation among closely related sympatric species of *Serrasalmus* (Ostariophysii, Characidae) from the Upper Madeira River, Amazon, Bolivia. J. Fish Biol. 68 (2006) 1–21.

[12] N. Backström, S. Fagerberg, H. Ellegren, Genomics of natural bird populations: a gene-based set of reference markers evenly spread across the avian genome. Mol. Ecol. 17 (2008) 964-980.

[13] W.T. Tay, G.T. Behere, D.G. Heckel, S.F. Lee, P. Batterham, Exon-primed intron crossing (EPIC) PCR markers of *Helicoverpa armigera* (Lepidoptera: Noctuidae). Bull. Ent. Res. 98 (2008) 509-518.

[14] N.M. Endersby, A.A. Hoffmann, V.L. White, S. Lowenstein, S. Ritchie, P.H. Johnson, L.P. Rapley, P.A. Ryan, V.S. Nam, N.T. Yen, P. Kittiyapong, A.R. Weeks, Genetic structure of *Aedes aegypti* in Australia and Vietnam revealed by microsatellite and Exon Primed Intron Crossing markers suggests feasibility of local control options. J. Med. Entomol. 46 (2009) 1074-1083.

[15] J. Dixit, H. Srivastava, O.P. Singh, D.N. Saksena, A. Das, Multilocus nuclear DNA markers and genetic parameters in an Indian *Anopheles minimus* population. Infec. Genet. Evol. 11 (2011) 572-579.

[16] M. He, D.S. Haymer, Polymorphic intron sequences detected within and between populations of the Oriental Fruit Fly (Diptera: Tephritidae). Ann. Ent. Soc. Am. 90 (1997) 825-831.

[17] S. Creer, A. Malhotra, R.S. Thorpe, C.E. Pook, Targeting optimal introns for phylogenetic analyses in non-model taxa : experimental results in Asian pitvipers. Cladistics 21 (2005) 390-395.

[18] S. Creer, Choosing and using introns in molecular phylogenetics. Evol. Bioinfo. 3(2007) 99-108.

[19] K.S. Wydner, J.L. Sechler, C.D. Boyd, H.C. Passmore, Use of an intron polymorphism to localize the tropoelastin gene to mouse chromosome 5 in a region of linkage conservation with human chromosome 7. Genomics 23 (1994) 125-131.

[20] Y. Yasukochi, L.A. Ashakumary, K. Baba, A. Yoshido, K. Sahara, A second generation integrated map of the silkworm reveals synteny and conserved gene order between lepidopteran insects. Genetics 173 (2006) 1319–1328.

[21] E.G. Pringle, S.W. Baxter, C.L. Webster, A. Papanicolaou, S.F. Lee, C.D. Jiggins, Synteny and chromosome evolution in the Lepidoptera: evidence from mapping in *Heliconius melpomene*. Genetics 177 (2007) 417–426.

[22] N. Backström, M. Brandström, L. Gustafsson, A. Qvarnström, H. Cheng, H. Ellegren. A gene-based genetic linkage map of the collared flycatcher (*Ficedula albicollis*) reveals extensive synteny and gene-order conservation during 100 million years of Avian evolution. Genetics 179 (2008) 1479-1495.

[23] J. Slate, J. Gratten, D. Beraldi, J. Stapley, M. Hale, J.M. Pemberton, Gene mapping in the wild with SNPs: guidelines and future directions. Genetica 136 (2009) 97-107.

[24] J.A. Wilder, H. Hollocher, Recent radiation of endemic Caribbean *Drosophila* of the *dumni* subgroup inferred from multilocus DNA sequence variation. Evolution 57 (2003) 2566-2579.

[25] A. Das, S. Mohanty, W. Stephan, Inferring the population structure and demography of *Drosophila ananasae* from multilocus data. Genetics 168 (2004) 1975-1985.

[26] P. Berrebi, E. Boissin, F. Fang, G. Berrebi-Cattaneo, Intron polymorphism (EPIC-PCR) reveals phylogeographic structure of *Zacco platypus* in China: a possible target for aquaculture development. Heredity 94 (2005) 589-598.

[27] C. Li, J.J. Riethoven, L. Ma, Exon-primed intron crossing (EPIC) markers for non-model teleost fishes. BMC Evol. Biol. 10 (2010) 90. doi: 10.1186/1471-2148-10-90.

[28] K. Lohse, B. Sharanowski, M. Blaxter, J.A. Nicholls, G. N. Stone, Developping EPIC markers for chalcidoid Hymenoptera from EST and genomic data. Mol. Ecol. Res. 11 (2011) 521-529.

[29] R.C. Thomson, I.J. Wang, J.R. Johnson, Genome-enabled development of DNA markers for ecology, evolution and conservation. Mol. Ecol. 19 (2010) 2184-2195.

[30] N. Bierne, S.A. Lehnert, E. Bédier, F. Bonhomme, S.S. Moore, Screening for intron-length polymorphism in paenid shrimps using exon-primed intron-crossing (EPIC)-PCR. Mol. Ecol. 9 (2000) 233–235.

[31] H.B.S.M. Côrte-Real, D.R. Dixon, P.W.H. Holland, Intron-targeted PCR: a new approach to survey neutral DNA polymorphism in bivalve populations. Mar. Biol. 120 (1994) 407–413.

[32] T. Atarhouch, M. Rami, G. Cattaneo-Berrebi, C. Ibanez, S. Augros, E. Boissin, A.L. Dakkak, P. Berrebi, Primers for EPIC amplification of intron sequences for fish and other vertebrate population genetic studies. BioTech. 3 (2003) 676-682.

[33] L.M. Gomulski, K. Bourtzis, S. Brogna, P.A. Morandi, C. Bonvicini, F. Sebastiani, C. Torti, C.R. Guglielmino, C. Savakis, G. Gasperi, A.R. Malacrida, Intron size polymorphism of the Adh1 gene parallels the worldwide colonization history of the Mediterranean fruit fly, *Ceratitis capitata*. Mol. Ecol., 7 (1998) 1729–1741.

[34] A.J. Bohonak, N. Davies, F.X. Villablanca, G.K. Roderick, Invasions genetics of New World medflies: testing alternative colonization scenarios. Biol. Invasions 3 (2001) 103-111.

[35] J.R. Gorham, C.J. Stojanovich, H.G. Scott, Clave illustrada para los mosquitos Anofelinos de Sudamerica Occidental (Illustrated key to the Anopheline mosquitoes of Western South America). Mosq. Syst. 5 (1973) 97-156.

[36] J.K. Edwards, Miniprep procedures for the isolation of plant DNA, in: A. Karp, P. G. Issac, Ingram (eds), Molecular tools for screening biodiversity, Chapman & Hall, London, 1998, pp.22-24.

[37] LabImage, Gel Analysis Software. Halle, Kapelan GmbH, 2005.

[38] K. Belkhir, P. Borsa, J. Goudet, L. Chikhi, F. Bonhomme, GENETIX, Logiciel sous WindowsTM pour la Génétique des Populations, Laboratoire Génome et Populations, CNRS UPR 9060. Univ.Montpellier II, Montpellier, France, 1998.

[39] M. Hassan, M. Harmelin-Vivien, F. Bonhomme, Lessepsian invasion without bottleneck: example of two rabbitfish species (*Siganus rivulatus* and *Siganus luridus*). J. Exp. Mar. Biol. Ecol. 291 (2003) 219-232.

[40] S.C. France, N. Tachino, T.F. Duda, R.A. Schleser, S.R. Palumbi, Intraspecific genetic diversity in the marine shrimp *Penaeus vannamei*: multiple polymorphic elongation factor-1a loci revealed by intron sequencing. Mar. Biotech. 1 (1999) 261–268.

[41] P. Borsa, A. Collet, J.-D. Durand, Nuclear-DNA markers confirm the presence of two anchovy species in the Mediterranean. C. R. Biologies 327 (2004) 1113–1123.

[42] A. Chenuil, T.B. Hoareau, E. Egea, G. Penant, C. Rocher, D. Aurelle, K. Mokhtar-Jamai, J.D.D. Bishop, E. Boissin, A. Diaz, M. Krakau, P.C. Luttikhuizen, F.P. Patti, N. Blavet, S. Mousset, An efficient method to find potentially universal genetic markers,

applied to metazoans. BMC Evol. Biol. 10 (2010) 276. doi: 10.1186/1471-2148-10-276.

[43] Walter Reed Biosystematics Unit website: http://www.mosquitocatalog.org/taxon_descr.aspx?ID=18300

[44] F. Lardeux, T. Chavez, R. Rodriguez, L. Torrez, Anopheles of Bolivia: new records with an updated and annotated checklist. C.R. Biologies 332 (2009) 489-499.

Table legends

Table 1. The fifty four selected genes and their accession number.

Primers pairs n° 1 to 14 successfully amplified intron sequences in both *An. pseudopunctipennis* (see also table 2) and *An. gambiae* (see also Fig. 1 a). Primers designed in the others genes (n° 15 to 54) did not amplify intron sequences in *An. pseudopunctipennis* also they did in *An. gambiae* (see also Fig. 1b).

Table 2. Sequences of the 14 pairs of primers which successfully amplified 17 loci from *Anopheles pseudopunctipennis* and their characterization.

Tm, annealing temperature in °C, H_E, expected heterozygosity; H_O, observed heterosigosity, *Fis*, inbreeding coefficient The amplification conditions MM1 and MM2 correspond to 0.04 µM and 0.4µM of primers in the master mix respectively. The overall computations for H_E, H_O and *Fis* were carried out taking into account the polymorphic alleles only

Table 3. Primer pairs and species in which they may potentially amplify, according to nucleotide sequence alignments using BLAST (NCBI)

Table 4. R^2 coefficients for linkage disequilibrium computed for each pair of polymorphic loci.

Figure captions

Figure 1. Examples of agarose gel electrophoresis results for four primer pairs

a. Positive amplifications with *An. pseudopunctipennis*. The size marker (lane 8) is a 1000 bp ladder with first band of 1000 bp and last band of 100 bp. Lanes 7 and 14 are negative controls (H_20).

   Lanes at the left side of the ladder correspond to results for primer pair n° 3. Lanes 3, 4, 5 are *An. pseudopunctipennis*, while lanes 1, 2, 6 are *An. gambiae* (positive controls). Lanes at the right side of the ladder correspond to results for primer 9. Lanes 11, 12, 13 are *An. pseudopunctipennis* while lanes 9, 10 are A*n. gambiae* (positive controls).

b. Negative amplifications with *An. pseudopunctipennis*. The size marker (lane 7) is a 1000 bp ladder first band of1000 bp and last band of 100 bp. Lane 6 is a negative control (H_20).

   Lanes at the left side of the ladder correspond to results for primer pair n°26. Lanes 1, 2, 3, 4 are *An. pseudopunctipennis*, while lane 5 is *An. gambiae* (positive control). Lanes at the right side of the ladder correspond to results for primer pair n°27. Lanes 8, 9, 10, 11, 12, 14 are *An. pseudopunctipennis* while lane 13 is *An. gambiae* (positive control).

Table 1

Primer pair n°	Gene	Accession number	Primer pair n°	Gene	Accession number
1	AgaP_AGAP005839	XM_315863.4	28	AgaP_AGAP011438	XM_554781
2	AgaP_AGAP003128	XM_001237495.2	29	AgaP_AGAP011717	XM_320795.4
3	AgaP AGAP012571	XM_307301.3	30	AgaP_AGAP010343	XM_311599.4
4	AgaP_AGAP011936	XM_320597.4	31	AgaP_AGAP010725	XM_559235.3
5	AgaP_AGAP001573	XM_551238.3	32	AgaP_AGAP001813	XM_321242.4
6	AgaP_AGAP008026	XM_555438.3	33	AgaP_AGAP003857	XM_310416.6
7	AgaP_AGAP004780	XM_318036.4	34	AgaP_AGAP004298	XM_313573.4
8	AgaP_AGAP005961	XM_316001.4	35	AgaP_AGAP002956	XM_311943.3
9	AgaP AGAP012014	XM_320516.2	36	AgaP_AGAP002301	XM_312670.1
10	AgaP AGAP011363	XM_001238009.2	37	AgaP_AGAP003360	XM_314262.4
11	AgaP_AGAP011166	XM_309483.4	38	AgaP_AGAP001407	XM_321726.4
12	AgaP_AGAP001874	XM_550942.3	39	AgaP_AGAP003437	XM_311723.4
13	AgaP_AGAP007887	XM_317605.4	40	AgaP AGAP007738	XM_001689008.1
14	AgaP_AGAP009824	XM_318932.3	41	AgaP AGAP008288	XM_001688954.1
15	AgaP_AGAP004745	XM_318073	42	AgaP AGAP009200	XM_319976.3
16	AgaP_AGAP004934	XM_315024.3	43	AgaP_AGAP008938	XM_319692.3
17	AgaP_AGAP005622	XM_315632.3	44	AgaP_AGAP009835	XM_553715.3
18	AgaP_AGAP005693	XM_315704.4	45	AgaP_AGAP009785	XM_318880.4
19	AgaP_AGAP005806	XM_315822.4	46	AgaP_AGAP009856	XM_318967.4
20	AgaP AGAP006809	XM_308938.4	47	AgaP_AGAP004698	XM_318145.4
21	AgaP_AGAP006825	XM_308919.3	48	AgaP_AGAP004775	XM_318039.4
22	AgaP_AGAP007640	XM_308229.4	49	AgaP_AGAP004841	XM_314353.3
23	AgaP_AGAP007720	XM_574504.3	50	AgaP_AGAP004717	XM_318113.4
24	AgaP_AGAP008526	XM_316916.3	51	AgaP_AGAP004862	XM_314327.3
25	AgaP_AGAP008527	XM_316915.4	52	AgaP_AGAP004915	XM_315006.4
26	AgaP_AGAP012345	XM_320207.4	53	AgaP_AGAP004692	XM_001231109.2
27	AgaP_AGAP011730	XM_320779.3	54	AgaP_AGAP005948	XM_315983.4

Table 2

Primer pair n°	Intron Name	Chromosome	Accession number	Primer sequences (5' – 3')	*Tm* (°C)	No. of alleles	Sizes of alleles (bp)	Amplification conditions	No. mosquitoes	H_E	H_O	F_{IS}
1	G2LEX1236-1	2L	XM_315863.4	F:TGGCTGGCTTCACGTCCG R:CGAGTGCAGGAACGGTGA	55	4	114, 151, 165, 180	MM1	18	0.4	0.3	0.25
1	G2LEX1236-2	2L	XM_315863.4	F:TGGCTGGCTTCACGTCCG R:CGAGTGCAGGAACGGTGA	55	1	297	MM1	18	-	-	-
2	GUKEX1858	2R	XM_001237495.2	F: GCCTGTGATCGTGCGTTTCG R:GGCATACCAGCAGCGTGACG	55	4	794, 857, 870, 876	MM1	18	0.3	0.3	0.09
3	GUKEX1859-1	UNKN	XM_307301.3	F:CGAGGAGGGTGTACAAACGC R:GGTGTCGCCTAGCTCGCCCG	55	2	715, 766	MM1	20	0.5	0.7	-0.3
3	GUKEX1859-2	UNKN	XM_307301.3	F:CGAGGAGGGTGTACAAACGC R:GGTGTCGCCTAGCTCGCCCG	55	2	306, 320	MM1	20	0.1	0.1	0.65
3	GUKEX1859-3	UNKN	XM_307301.3	F:CGAGGAGGGTGTACAAACGC R:GGTGTCGCCTAGCTCGCCCG	55	3	183, 193, 197	MM1	20	0.2	0.1	0.79
4	G3LEX28	3L	XM_320597.4	F:CCAACTACTCGGCCGTGC R:GCCGGCCATCTCCTTCGC	60	3	246, 253, 261	MM2	18	0.4	0.4	0.02
5	G2REX47	2R	XM_551238.3	F:GGCACGGTGGGGAAGACG R:CCGTCCACCACCATCGGG	60	4	220, 228, 257, 273	MM2	18	0.3	0.4	-0.1
6	G3REX1037	3R	XM_555438.3	F:GCAAACGCGAAAGAACCG R:GCCTGGTAGCGCTTCTCG	60	4	281, 315, 345, 405	MM2	12	0.5	0	1
7	EX113	2L	XM_318036.4	F:CATCTATCTGCTGAACTCGC R:CGTCGGTCACATTCCACATC	60	3	541, 557, 574	MM2	14	0.4	0	1
8	EX1358	2L	XM_31600104	F:CATGCCTCCAATGGTGCC R:CCGTACGTTCCTTCGCCA	60	1	249	MM2	20	-	-	-
9	G3LEX3	3L	XM_320516.2	F:CCGAAGATGAGCTCAGAGATGC R:CCTAGCTTGTCGGTGATTTCTG	55	1	185	MM1	20	-	-	-
10	G3LEX9	3L	XM_001238009.2	F:CGCCCTGCCTGGCATGGATTCG R:GCAGGCACAGCCACCTTCCGGG	55	1	656	MM1	20	-	-	-
11	G3LEX36	3L	XM_309483.4	F:CGCGGCAATCATGAGTGCGCC R:CCACCGGCAGACAGTTGAAGC	55	1	191	MM1	20	-	-	-
12	G2REX46	2R	XM_550942.3	F:CCGACGATAGAGGACAGC R:GTTGAAGGTCGACTGTGC	60	1	497	MM2	20	-	-	-
13	G3REX491	3R	XM_317605.4	F:CGTTGGAGCAGCAACAACAGC R:GGTAATGATTCCTGATATTGC	55	1	136	MM2	20	-	-	-
14	G3REX1062	3R	XM_318932.3	F:CGATCTGCTGGCCGACTTCC R:CCATCGCCCTTGCGCTCACC	55	1	177	MM1	20	-	-	-
									Overall*	0.35	0.24	0.35

Table 3

Especie	Primer pair n°																									number or primer pairs that might amplify in the species
	1	2	4	8	9	10	11	12	13	14	16	18	19	20	21	22	29	31	34	35	36	37	38	43	50	56	
Aedes aegypti	x	x	x		x	x				x			x		x		x	x					x				11
Ae. albopictus														x													1
Anopheles arabiensis											x	x															2
An. bwambae												x															1
An. christyi												x															1
An. merus												x															1
An. quadrianulatus												x															1
An. stephensi																										x	1
Canis familiaris							x																				1
Culex quinquefasciatus		x	x	x		x	x			x	x		x	x	x	x		x	x			x				x	15
Drosophila melanogaster			x				x								x			x	x		x	x		x			8
D. ananassae			x				x								x	x											4
D. erecta			x												x						x	x					4
D. grimshawi			x		x		x								x												4
D. mojavensis			x				x								x												3
D. persimilis			x				x										x										3
D. pseudoobscura			x				x										x										3
D. sechellia			x				x														x	x					4
D. simulans			x				x															x					3
D. virilis									x																		1
D. willistoni									x																		1
D. yakuba			x						x										x		x	x		x			6
Gryllus bimaculatus			x																						x		2
Helicoverpa armigera			x																								1
Heliotis viriplaca			x																								1
Homo sapiens						x											x										2
Ixodes scapularis			x			x	x											x									4
Monosiga brevicolis								x	x																		2
Mus musculus																			x								1
Plantago major																				x							1
Thermosynechococcus elongatus																									x		1
Trichophyton verrucosum						x																					1
Number of species that might be recognized	1	2	16	1	2	5	11	1	4	2	2	5	2	2	7	2	4	4	4	1	4	6	1	2	2	2	

Table 4.

Intron names	G2LEX1858	GUKEX1859-1	GUKEX1859-2	GUKEX1859-3	G3LEX28	G2REX47	G3REX1037	EX113
G2LEX1236-1	0.28	0.42	0.65	0.01	0.10	0.91	0.78	0.04
G2LEX1235-1		0.79	0.92	0.74	0.62	0.21	0.59	0.89
GUKEX1859-1			0.30	0.30	0.29	0.85	0.71	0.20
GUKEX1859-2				0.96	0.67	0.89	0.83	-
GUKEX1859-3					0.10	0.99	0.06	0.10
G3LEX28						0.90	0.21	0.01
G2REX47							0.02	0.52
G3REX1037								-
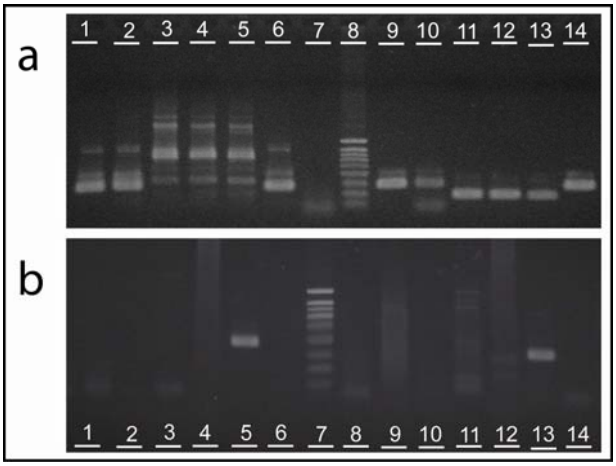
Figure 1.