When population and evolutionary genetics met behaviour

Rodolfo Costa[1], Ralf Stanewsky[2]/+

[1]Department of Biology, University of Padova, Padova, Italy
[2]Department of Cell and Developmental Biology, Rockefeller Building, University College London, London, UK

In this review, we analyse the impact of a population and evolutionary genetics approach on the study of insect behaviour. Our attention is focused on the model organism Drosophila melanogaster *and several other insect species. In particular, we explore the relationship between rhythmic behaviours and the molecular evolution of clock and ion channel genes.*

Key words: population genetics - evolutionary genetics - behaviour -
Lutzomyia longipalpis - *period* gene - *cacophony* - *Drosophila melanogaster*

The first clock gene in Drosophila - In the 1980s, we were at the dawn of the molecular dissection of the circadian machinery; at the time, only one clock gene, *period* (*per*), had been identified and cloned (Konopka & Benzer 1971, Bargiello & Young 1984, Reddy et al. 1984). With this limited molecular knowledge, the organisation and functioning of the circadian machinery was a complete mystery and it was six more years before the first experimental circadian molecular model, the negative feedback loop, was generated for the fly (Hardin et al. 1990).

The population and evolutionary genetics approach - At the beginning of 1991, in the Department of Genetics at the University of Leicester (UK), Alexandre Peixoto, Bambos Kyriacou and myself (RC) began to speculate on the possibility of investigating the *Drosophila per* gene from a population and evolutionary perspective. We first explored the naturally occurring genetic variability at the *per locus* in *Drosophila melanogaster*. We examined dozens of population samples obtained from different European localities (including flies collected during our summer holidays on fabulous Greek islands) and we focussed our attention on a naturally occurring length polymorphism in the so-called "threonine-glycine encoding repeat region" (the Thr-Gly region) in the fifth exon of the *per* gene. At this *locus*, a variable number of cassettes made by the "*ACN GGN*" codon motif generated a stretch of Thr-Gly pairs at the protein level while retaining important molecular signatures at the third silent nucleotide position. It turned out that the flies from natural populations encoded several length variants, which were most commonly stretches of 14, 17, 20 or 23 Thr-Gly pairs (Costa et al. 1991). This repetitive coding region had features similar to minisatellite sequences and we decided to undertake a systematic geographical analysis of this polymorphism within the European and North African natural populations. We observed a large-scale, robust latitudinal cline for the main allelic length variants (Costa et al. 1992), which suggested that the length polymorphism could be maintained by climate-related selection. We then extended our studies on the geographical distribution of the Thr-Gly length polymorphism in *D. melanogaster*. In fact, we analysed several population samples from the Southern Hemisphere. In these samples, we observed an inverse latitudinal cline for the $(Thr\text{-}Gly)_{20}$ allele compared with the European samples and a higher level of the Thr-Gly allelic variants that are quite rare in European populations, which were present particularly at tropical latitudes in the southern hemisphere (Sawyer et al. 2006). Overall, these results suggested that the clinal distribution of the major Thr-Gly length alleles in Australia is less robust than the distribution described in Europe and North Africa, most likely due to the much more recent colonisation of Australia by *D. melanogaster* (David & Capy 1988). Meanwhile, our hypothesis that temperature-related selection is the main force shaping the Thr-Gly length polymorphism in *D. melanogaster* in nature received significant support from the results of our study on the effects of temperature on the free-running period associated with the locomotor activity behaviour of flies with variable Thr-Gly length alleles (Sawyer et al. 1997). We examined flies homozygous for various common or rare natural Thr-Gly alleles and per^0 null mutant flies that were transformed with constructs of artificially manipulated Thr-Gly encoding regions [deletion of the complete region, Δ(Thr-Gly), 1 cassette encoding for one Thr-Gly pair, $(Thr\text{-}Gly)_1$ or the naturally occurring variants, $(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_{20}$]. We found that the transgenic Δ(Thr-Gly) and $(Thr\text{-}Gly)_1$ homozygous flies lost temperature compensation, while the transgenic or natural flies homozygous for the major European alleles $(Thr\text{-}Gly)_{14}$, $(Thr\text{-}Gly)_{17}$, $(Thr\text{-}Gly)_{20}$ and $(Thr\text{-}Gly)_{23}$ showed an approximately linear relationship between the Thr-Gly repeat length and temperature compensation. Where the (Thr-Gly)20 homozygous flies exhibited the best temperature-compensated circadian rhythmicity, which, consequently, is potentially favoured in latitudes with more variable temperatures. As temperature varies systematically with latitude, we hypothesised that the different Thr-Gly length variants are involved in the mechanisms conferring behavioural temperature

doi: 10.1590/0074-0276130498
+ Corresponding author: r.stanewsky@ucl.ac.uk
Received 13 October 2013
Accepted 26 November 2013

compensation, a cardinal property of the circadian clock (Sawyer et al. 1997). At the time, it was well established that circadian clocks maintain a stable ~24-h periodicity over a large range of temperatures (Pittendrigh 1954). Therefore, the latitudinal clinal structure for the Thr-Gly length polymorphism might be adaptive. Indeed, using statistical analysis of the DNA sequences around the Thr-Gly region, we were able to conclude that the nucleotide variation observed was consistent with a balancing selection scenario (Rosato et al. 1996, 1997). This finding was very important, as it verified the signature of natural selection at the DNA level, providing an independent assessment that this region was of adaptive significance. We also extended this type of statistical analysis to the Thr-Gly region in *Drosophila simulans,* the cosmopolitan sibling species that also revealed a signature of balancing selection (Rosato et al. 1994); we followed this analysis up with functional studies of temperature compensation in the Thr-Gly *D. simulans* variants (Rogers et al. 2004a, b).

During these investigations, we became particularly interested in the comparison of the Thr-Gly allelic variants of identical lengths (isolength alleles), which apparently were carrying high numbers of synonymous substitutions. For instance, in the case of two different alleles encoding 23 Thr-Gly pairs, putative silent substitutions were detected at 21 of the 46 silent sites. At first glance, this finding indicated an ancient time of separation of approximately 25 Myr since a common ancestor, a time interval that was completely inconsistent with the estimation that *D. melanogaster* originated approximately 2 Myr ago in tropical Africa (Lachaise et al. 1988). However, a detailed examination of the sequences of several natural-length variants indicated that this divergence could be understood in terms of four deletion/insertion events: from an ancestral $(Thr\text{-}Gly)_{23}$, through a deletion of five pairs to a $(Thr\text{-}Gly)_{18}$, another deletion to a $(Thr\text{-}Gly)_{17}$, then successive duplications of the same (Thr-Gly)3 cassette to generate a $(Thr\text{-}Gly)_{20}$ and then the derived $(Thr\text{-}Gly)_{23}$, a much more parsimonious evolutionary scenario (Costa et al. 1991, Rosato et al. 1996)! Later, with E Rosato, we were able to measure the approximate mutation rate of the Thr-Gly repeat (~10^4 times faster than the point mutation rate) (Rosato et al. 1997), which, when applied to the sequences we had studied, suggested that their evolution could have taken at most approximately 1,000-5,000 years. The Thr-Gly region thus appeared to be a rapidly evolving coding sequence resembling the non-coding minisatellites that are characterised by high rates of mutational events, which were being studied at about the same time a few meters from our bench in Leicester by Sir A Jeffreys et al. (1990).

We then extended the evolutionary analysis to the DNA sequence surrounding and including the Thr-Gly encoding region of the *per* gene from the eight species that form the *D. melanogaster* subgroup. This region can be divided into the Thr-Gly repeat and the less-repetitive flanking sequences. The number of amino acids encoded in the variable region ranged from 40 in *Drosophila teissieri* to 69 in *Drosophila mauritiana*. Even in this case, it was possible to derive a specific Thr-Gly allele of one species from that of another by invoking only one or a few deletion/insertion events or a few hypothetical Thr-Gly intermediates (Peixoto et al. 1992). Indeed, the Thr-Gly region could be used as a proxy for the phylogeny of the species and mirrored more classical phylogenies based on other *loci* (Coyne & Kreitman 1986, Lemeunier et al. 1986).

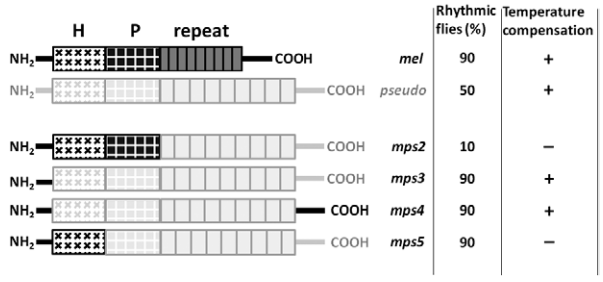
Later, we further extended our analyses to compare the Thr-Gly region from other *Drosophila* species within the *Drosophila* and *Sophophora* subgenera and found that this repetitive region exhibits enormous variability in both DNA sequence and length in these species (Peixoto et al. 1993). For instance, *Drosophila pseudoobscura* has approximately 35 copies of a 5-amino-acid degenerate repeat (rich in serine, glycine and asparagine or threonine), which appears to be derived through replicative slippage from a Thr-Gly repeat sequence. This corresponding repeat region in *D. pseudoobscura* is greater than 200 residues in length, which is approximately four times longer than the repeat region of *D. melanogaster* (Colot et al. 1988). Other species have very short repeats with little more than 10 residues (Peixoto et al. 1993). Moreover, in spite of the differences between the amino acid sequences of the repeats, their predicted secondary structures (a stretch of flexible turns that separates 2 globular domains) were conserved, suggesting that evolutionary and mechanistic constraints shaped the PERIOD protein of these two species (Costa et al. 1991). More importantly, a comparative analysis of the DNA sequences of the species under investigation indicated that the changes in length of the so-called "Thr-Gly variable region" were associated with amino acid replacements in the more conserved flanking sequences (Peixoto et al. 1993, Nielsen et al. 1994). A Peixoto immediately realised that this finding implied that the repeat length was co-evolving with the flanking region to maintain the protein conformation of that region. To challenge this coevolution hypothesis, we designed a functional test in which several chimeric *per* transgenes from *D. melanogaster* and *D. pseudoobscura* were generated; we used various chimeric junctions to maintain or disrupt the species-specific contiguity of the "Thr-Gly repetitive region" with either all or part of the co-evolved 5' flanking region that encodes the 55 amino acids immediately upstream of the Thr-Gly regions of the two species (Peixoto et al. 1998). The analysis of the locomotor behaviours of the different transgenic flies provided dramatic experimental support for the coevolutionary interpretation: disrupting the coevolution of the repeat with its flanking region led to an almost arrhythmic clock, whereas maintaining the species-specific contiguity of the two regions resulted in wild-type phenotypes. A partial disruption of the species-specific contiguity generated a highly temperature-sensitive period, revealing a disturbance of the temperature compensation of the clock (Peixoto et al. 1998) (Figure).

At present, the molecular mechanism of temperature compensation and the role of the Thr-Gly region remain largely elusive. We suspect that PERIOD temperature-sensitive degradation may be part of the puzzle. Nevertheless, the molecular evolutionary approach initiated in the 1990s and consolidated in the first decade of the new

century (Kyriacou et al. 2008) added an important and novel dimension to the field of modern molecular chronobiology and has opened the way for similar studies in other organisms (Bauzer et al. 2002a, b, Mazzoni et al. 2002, Meireles-Filho et al. 2006a, b, Gentile et al. 2009, Chahad-Ehlers et al. 2013).

Background on the role of ion channels in the speciation of Lutzomyia longipalpis - In the laboratory of J Hall, the *D. melanogaster cacophony* (cac) gene was identified in a mutant hunt for flies exhibiting an aberrant male courtship song (Schilcher 1977, Kulkarni & Hall 1987, Smith et al. 1998). *D. melanogaster* males extend their wings during courtship and, through high frequency vibration, produce a species-specific "song", which consists of sound "pulses" (2-3 cycles each) that are repeated between two and 30 times (train of pulses) roughly every 35 msec and are interspersed by "inter pulse intervals" (IPIs) (Hall 1994). The *cac* mutant earned its name because the song produced by *cac* males has (i) an increased number of cycles per pulse, (ii) pulses with larger amplitudes and (iii) longer IPIs compared to wild-type flies (Schilcher 1977, Kulkarni & Hall 1987, Smith et al. 1998). Interestingly, the song produced by *cac* males is more similar to the songs of other *Drosophila* species (Peixoto & Hall 1998), suggesting that the song characteristics contribute to the reproductive isolation between closely related species (an idea later adopted in sandfly studies, see below).

The cloning of *cac* revealed that the gene encodes the ion-conducting α1-subunit of a voltage-gated N-type Ca^{2+}-channel and that the original *cac* mutant carries a single-nucleotide substitution resulting in a missense mutation in one of its transmembrane domains (a Phe to Ile change, most likely affecting the pore of the channel) (Smith et al. 1996, 1998, Peixoto et al. 1997). In addition to being among the first Ca^{2+}-channel α1-subunits cloned from invertebrates, *cac* was the first example of a behaviourally relevant gene that is modified by RNA editing: an adenosine nucleotide in the DNA is changed to a guanosine in the mRNA by a double-stranded RNA deamination reaction (Smith et al. 1996, Peixoto et al. 1997). In the case of *cac*, RNA editing and alternative splicing resulted in the generation of multiple versions of the α1-subunit, potentially differing in voltage sensitivity and β-subunit selection (Smith et al. 1996).



period (*per*) transgenes of *Drosophila melanogaster* (*mel*) (black) and *Drosophila pseudoobscura* (*pseudo*) (light gray). H [25 amino acid (aa)] and P (30 aa) domains represent the protein region immediately N-terminal to the repetitive region in both species. *mps2*, *mps3*, *mps4* and *mps5* are chimeric *melanogaster-pseudoobscura* transgenes in which different portions of the two *mel* and *pseudo* genes are combined, so that to maintain or disrupt the conspecific coevolving aa blocks. The repeat region corresponds to the threonine-glycine encoding repeat region (Thr-Gly) dipeptide run in *D. melanogaster* and to the pentapeptide run in *D. pseudoobscura*. Levels of rescue of rhythmicity (% of rhythmic flies) for the different transgenes in a per^{01} genetic background and presence (+) or absence (-) of temperature compensation are indicated.

Again, the diversity of the *cac* isoforms and their role in courtship song quality implicate this gene in speciation, suggesting that other ion channel genes could have similar roles. Indeed, the voltage-gated Na^{+}-channel *paralytic* (*para*) and the Ca^{2+}-gated K^{+}-channel *slowpoke* (*slo*) also affect the courtship song. Similar to the mutations in *para* and *slo, cac* mutant flies also exhibit locomotor defects (seizures) at high temperatures, which is not surprising given that these genes affect neuronal membrane excitability (Peixoto & Hall 1998) and also demonstrates that ion channel mutants have pleiotropic effects (but not all ion channel mutants exhibit song phenotypes) (Peixoto & Hall 1998). In summary, the results obtained from studying ion channel mutations in *D. melanogaster* suggest that the subtle and cumulative changes in individual or multiple ion channel genes could contribute to the generation of species-specific love songs and could therefore also drive speciation.

Cacophony, paralytic and love songs in sandflies - Phlebotomine sandflies also produce love songs, although in contrast to *D. melanogaster,* the males sing to the females during copulation rather than before. Do ion channels and love songs also influence speciation in these and other insects, some of which are vectors of epidemic diseases? To address this question, sequences from the *cac* gene in *Lu. longipalpis,* the main vector of American visceral leishmaniasis in the New World, were cloned (Peixoto et al. 2001). *cac* sequences from other *Lutzomyia* species were also isolated and it was determined that a particular intron sequence (not present in *D. melanogaster*) undergoes rapid alteration between different sandfly species while the coding region for a transmembrane domain of the channel is highly conserved (Lins et al. 2002). Overall, the sequences confirmed the classification of the various Brazilian *Lutzomyia* species into the *Lutzomyia*, *Nyssomia* and *Migonei* subgroups, demonstrating that these *cac* sequences are useful molecular phylogenetic markers.

Using the same approach, *cac* sequences (and sequences of other genes with potential roles in love song generation: the *L. longipalpis* homologues of the *D. melanogaster per* and *para* genes) were compared between various populations of *L. longipalpis* in Brazil. These studies were partially driven by the increasing doubts in the field that the various *L. longipalpis* populations belong to the same species, as had been concluded mainly from isoenzyme comparisons [reviewed in Bauzer et al. (2007)]. Over a time span of approximately 10 years, a careful analysis of approximately 14 different Brazilian sandfly populations was conducted in several laborato-

ries (Table and references therein). In addition to a comparison of the *cac, per* and *para* sequences isolated from individuals of these populations, love songs from a majority of the populations (those that were possible to keep in the lab) were also recorded. The recording technique developed in the Hall lab for *Drosophila* was adapted to record the love songs of sandfly males (Peixoto & Hall 1998, Souza et al. 2004); a virgin male and female (from the F1 of a female collected in the wild) were placed in the recording chamber of an INSECTAVOX (Gorczyca & Hall 1987) to record the sounds produced by the male during copulation. In parallel, the pair was filmed with a video camera, which enabled the correlation of the flies' behaviours with the songs they produced. The results obtained were both interesting and surprising and are broadly summarised in the Table. Males from the populations native to the coastal regions of the North and Northeast down to the Southeast of Brazil produce a burst-type song, which is characterised by trains of extremely polycyclic pulses ("bursts") with irregular frequencies and amplitudes. In addition, all burst-type populations produce the same sex pheromone (cembrene-type 1), strongly suggesting that they all belong to the same species. In contrast, populations from more central parts of Brazil produce "pulse" songs that are more similar to the songs described above for *D. melanogaster*. Interestingly, the pulse songs between populations differ reproducibly in several parameters (leading to their classification into P1-P5 pulse song types) (Table), suggesting that they contribute to the reproductive isolation between different populations. In addition, most of the pulse singers produce different sex pheromones (Table), strongly supporting the idea that they indeed represent different sibling species rather than a single species.

The existence of an *L. longipalpis* species complex is also strongly supported by the comparison of the *per*, *cac* and *para* genes between the different populations (Table). Phylogenetic trees based on the DNA sequences of all three genes for the populations producing different pulse or burst songs indicate that they represent divergent species, which, in the case of the two Sobral populations, separated approximately 1 Myr ago (Bauzer et al. 2002a) (Table). Moreover, this DNA analysis revealed some fixed differences between the sibling species that were analysed for *para* polymorphisms, which can be used as reliable species indicators for sandflies collected in the field. Perhaps the most interesting examples with respect to reproductive isolation are the three sympatric populations that were analysed (Sobral, Estrala and Jaiba). Morphologically, these populations can be distinguished by the number of spots on their abdominal tergites; 1S populations carry a single pair of spots on the fourth tergite, while 2S populations carry an additional pair of spots on the third tergite. Considering all *L. longipalpis* populations, these spots are not a reliable marker for indicating different species, but in the sympatric species they are a reliable species indicator. The 2S populations from the three locations all belong

TABLE

Summary of Brazilian *Lutzomyia longipalpis* population characteristics

Species	Population	Song type	Pheromone	DNA phylogeny
1	Marajó	burst[e,f]	cemb-1[f,i]	*per*[i], ***para***[k]
	Sobral 2S	burst[e,f]	cemb-1[f,g,i]	*per*[b,i], *cac*[d4], ***para***[h,k]
	Natal	burst[e,f]	cemb-1[f,i]	*per*[a,c,i], *cac*[a,d], ***para***[k]
	Estrela 2S	burst[i]	cemb-1[i]	*per*[i], ***para***[k]
	Jaiba 2S	burst[i]	cemb-1[i]	*per*[i], ***para***[k]
	Pancas	burst[i]	cemb-1[i]	*per*[i], ***para***[k]
2	Sobral 1S	P3[e-g]	germ[f,g,i]	*per*[b,i], *cac*[d], ***para***[h,k]
	Teresina	P3[i]	germ[i]	*per*[i], ***para***[k]
3	Estrela 1S	P5[i]	cemb-1[i]	*per*[i], ***para***[k]
4	Jacobina	P1[e-g]	hima[f,g,i]	*per*[a,c,i], *cac*[a,d], ***para***[k]
5	Jaiba 1S	P4[i]	cemb-2[i]	***per***[i], ***para***[k]
6	Lapinha	P2[e-g]	germ[f,g,i]	*per*[a,c,i], *cac*[a,d], ***para***[k]
7	Mesquita	mix[i]	?	***per***[i]
8	Barra de Guaratiba	?	germ and cemb[i]	*per*[i]
9	*Lutzomyia cruzi*	burst[j]	germ[j]	*per*[j], ***para***[k]

summary of evidence for a *Lu. longipalpis* species complex and the grouping of the populations that have been analysed mainly by the Peixoto's lab. For a detailed description of song types "burst", "pulse P1-P5" and mix (mix of burst and pulse), sex pheromones and phylogenetic DNA analysis (DNA phylogeny) (*cac*: *cacophony*; *para*: *paralytic*; *per*: *period*) see references indicated in main text. Entries in bold in the DNA phylogeny column indicate fixed differences between populations. Note that most of the sex pheromone determinations were performed in various labs, as cited in the references [*a*: Oliveira et al. (2001); *b*: Bauzer et al. (2002a); *c*: Bauzer et al. (2002b); *d*: Bottecchia et al. (2004); *e*: Souza et al. (2004); *f*: Bauzer et al. (2007); *g*: Maingon et al. (2007); *h*: Lins et al. (2008); *i*: Araki et al. (2009); *j*: Vigoder et al. (2010); *k*: Lins et al. (2012)].

to the same species according to their song type (burst), sex pheromone (cembrane-type 1) and DNA phylogeny (Table). In contrast, the 1S populations each belong to a different sibling species based on the same criteria (including a fixed polymorphism in the *per* gene between Jaiba 1S and 2S) (Table).

The ultimate test to determine if the sibling species are indeed reproductively isolated is the outcome of crosses between them. This test has partly been accomplished; the copulation rates in heterospecific crosses between Jacobina, Natal and Laphina populations were reduced compared with the rates in homospecific crosses and no offspring were obtained in the heterospecific crosses (Souza et al. 2008). Interestingly, when crosses between the Sobral 1S and 2S populations were attempted, not even copulation took place, indicating that pre-mating behaviour may contribute to reproductive isolation.

The work summarised here beautifully demonstrates how an integrated approach led to the firm establishment of the *L. longipalpis* species complex. Importantly, this conclusion would not have been possible by following only a single trait (e.g., relying only on pheromone differences would not have distinguished between Estrela 1S and 2S populations, as both produce the cembrane-type 1 pheromone) (Table). Hopefully these studies will be continued and extended to investigate whether the different sibling species also differ in their vectorial capacity. It would also be interesting to try functionally linking *L. longipalpis* homologues of the *D. melanogaster* "song genes" *cac* and *para* to love songs in sandflies by RNAi methodology. If these genes also influence song in the sandflies (which is likely, based on the results of the phylogenetic DNA analysis), this finding would ultimately link back to the idea that subtle changes in these conserved ion channels result in behavioural changes that are a driving force for speciation.

REFERENCES

Araki AS, Vigoder FM, Bauzer LGSR, Ferreira GEM, Souza NA, Araujo IB, Hamilton JGC, Brazil RP, Peixoto AA 2009. Molecular and behavioral differentiation among Brazilian populations of *Lutzomyia longipalpis* (Diptera: Psychodidae: Phlebotominae). *PLoS Negl Trop Dis 3*: e365.

Bargiello TA, Young MW 1984. Molecular genetics of a biological clock in *Drosophila*. *Proc Natl Acad Sci USA 81*: 2142-2146.

Bauzer LGSR, Gesto JSM, Souza NA, Ward RD, Hamilton JGC, Kyriacou CP, Peixoto AA 2002a. Molecular divergence in the *period* gene between two putative sympatric species of the *Lutzomyia longipalpis* complex. *Mol Biol Evol 19*: 1624-1627.

Bauzer LGSR, Souza NA, Maingon RDC, Peixoto AA 2007. *Lutzomyia longipalpis* in Brazil: a complex or a single species? A mini-review. *Mem Inst Oswaldo Cruz 102*: 1-12.

Bauzer LGSR, Souza NA, Ward RD, Kyriacou CP, Peixoto AA 2002b. The *period* gene and genetic differentiation between three Brazilian populations of *Lutzomyia longipalpis*. *Insect Mol Biol 11*: 315-323.

Bottecchia M, Oliveira SG, Bauzer LGSR, Souza NA, Ward RD, Garner KJ, Kyriacou CP, Peixoto AA 2004. Genetic divergence in the *cacophony* IVS6 intron among five Brazilian populations of *Lutzomyia longipalpis*. *J Mol Evol 58*: 754-761.

Chahad-Ehlers S, Gentile C, Lima JB, Peixoto AA, Bruno RV 2013. Analysis of *cycle* gene expression in *Aedes aegypti* brains by in situ hybridization. *PLoS ONE 8*: e52559.

Colot HV, Hall JC, Rosbash M 1988. Interspecific comparison of the *period* gene of *Drosophila* reveals large blocks of non-conserved coding DNA. *EMBO J 7*: 3929-3937.

Costa R, Peixoto AA, Barbujani G, Kyriacou CP 1992. A latitudinal cline in a *Drosophila* clock gene. *Proc Biol Sci 250*: 43-49.

Costa R, Peixoto AA, Thackeray JR, Dalgleish R, Kyriacou CP 1991. Length polymorphism in the threonine-glycine-encoding repeat region of the *period* gene in *Drosophila*. *J Mol Evol 32*: 238-246.

Coyne JA, Kreitman M 1986. Evolutionary genetics of two sibling species, *Drosophila simulans* and *D. sechellia*. *Evolution 40*: 673-691.

David JR, Capy P 1988. Genetic variation of *Drosophila melanogaster* natural populations. *Trends Genet 4*: 106-111.

Gentile C, Rivas GB, Meireles-Filho AC, Lima JB, Peixoto AA 2009. Circadian expression of clock genes in two mosquito disease vectors: *cry2* is different. *J Biol Rhythms 24*: 444-451.

Gorczyca M, Hall JC 1987. The INSECTAVOX, an integrated device for recording and amplifying courtship songs. *Drosoph Inf Serv 66*: 157-160.

Hall JC 1994. The mating of a fly. *Science 264*: 1702-1714.

Hardin PE, Hall JC, Rosbash M 1990. Feedback of the *Drosophila period* gene product on circadian cycling of its messenger RNA levels. *Nature 343*: 536-540.

Jeffreys AJ, Neumann R, Wilson V 1990. Repeat unit sequence variation in minisatellites: a novel source of DNA polymorphism for studying variation and mutation by single molecule analysis. *Cell 60*: 473-485.

Konopka RJ, Benzer S 1971. Clock mutants of *Drosophila melanogaster*. *Proc Natl Acad Sci USA 68*: 2112-2116.

Kulkarni SJ, Hall JC 1987. Behavioral and cytogenetic analysis of the *cacophony* courtship song mutant and interacting genetic variants in *Drosophila melanogaster*. *Genetics 115*: 461-475.

Kyriacou CP, Peixoto AA, Sandrelli F, Costa R, Tauber E 2008. Clines in clock genes: fine-tuning circadian rhythms to the environment. *Trends Genet 24*: 124-132.

Lachaise D, Cariou ML, Lemeunier F, Tsakas L, Ashburner M 1988. Historical biogeography of the *Drosophila melanogaster* species subgroup. *Evol Biol 22*: 159-225.

Lemeunier F, David JR, Tsakas L, Ashburner M 1986. The *melanogaster* species group. In M Ashburner, HL Carson, JN Thompson (eds.), *The genetics and biology of Drosophila*, Vol. 3, Academic Press, New York, p. 147-256.

Lins RMMA, Oliveira SA, Souza NA, de Queiroz RG, Justiniano SCB, Ward RD, Kyriacou CP, Peixoto AA 2002. Molecular evolution of the *cacophony* IVS6 region in sandflies. *Insect Mol Biol 11*: 117-122.

Lins RMMA, Souza NA, Brazil RP, Maingon RDC, Peixoto AA 2012. Fixed differences in the *paralytic* gene define two lineages within the *Lutzomyia longipalpis* complex producing different types of courtship songs. *PLoS ONE 9*: e44323.

Lins RMMA, Souza NA, Peixoto AA 2008. Genetic divergence between two sympatric species of the *Lutzomyia longipalpis* complex in the *paralytic* gene, a *locus* associated with insecticide resistance and lovesong production. *Mem Inst Oswaldo Cruz 103*: 736-740.

Maingon RDC, Ward RD, Hamilton JGC, Bauzer LGSR, Peixoto AA 2007. The *Lutzomyia longipalpis* complex: does population substructure matter to *Leishmania* transmission? *Trends Parasitol 24*: 12-17.

Mazzoni CJ, Gomes CA, Souza NA, de Queiroz RG, Justiniano SC, Ward RD, Kyriacou CP, Peixoto AA 2002. Molecular evolution of the *period* gene in sandflies. *J Mol Evol 55*: 553-562.

Meireles-Filho AC, Amoretty PR, Souza NA, Kyriacou CP, Peixoto AA 2006a. Rhythmic expression of the *cycle* gene in a hematophagous insect vector. *BMC Mol Biol 7*: 38.

Meireles-Filho AC, Rivas GBS, Gesto JS, Machado RC, Britto C, de Souza NA, Peixoto AA 2006b. The biological clock of an hematophagous insect: locomotor activity rhythms, circadian expression and downregulation after a blood meal. *FEBS Lett 580*: 2-8.

Nielsen J, Peixoto AA, Piccin A, Costa R, Kyriacou CP, Chalmers D 1994. Big flies, small repeats: the "Thr-Gly" region of the *period* gene in Diptera. *Mol Biol Evol 11*: 839-853.

Oliveira SG, Bottecchia M, Bauzer LGSR, Souza NA, Ward RD, Kyriacou CP, Peixoto AA 2001. Courtship song genes and speciation in sand flies. *Mem Inst Oswaldo Cruz 96*: 403-405.

Peixoto AA, Campesan S, Costa R, Kyriacou CP 1993. Molecular evolution of a repetitive region within the *per* gene of *Drosophila*. *Mol Biol Evol 10*: 127-139.

Peixoto AA, Costa R, Wheeler DA, Hall JC, Kyriacou CP 1992. Evolution of the threonine-glycine repeat region of the *period* gene in the *melanogaster* species subgroup of *Drosophila*. *J Mol Evol 35*: 411-419. Erratum in *J Mol Evol 36*: 199.

Peixoto AA, Gomes CA, Amoretty PR, Lins RMMA, Meireles-Filho ACA, Souza NA, Kyriacou CP 2001. New molecular markers for phlebotomine sand flies. *Int J Parasitol 31*: 635-639.

Peixoto AA, Hall JC 1998. Analysis of temperature-sensitive mutants reveals new genes involved in the courtship song of *Drosophila*. *Genetics 148*: 827-838.

Peixoto AA, Hennessy JM, Townson I, Hasan G, Rosbash M, Costa R, Kyriacou CP 1998. Molecular coevolution within a *Drosophila clock* gene. *Proc Natl Acad Sci USA 95*: 4475-4480.

Peixoto AA, Smith LA, Hall JC 1997. Genomic organization and evolution of alternative exons in a *Drosophila* calcium channel gene. *Genetics 145*: 1003-1013.

Pittendrigh CS 1954. On temperature independence in the clock system controlling emergence time in *Drosophila*. *Proc Natl Acad Sci USA 40*: 1018-1029.

Reddy P, Zehring WA, Wheeler DA, Pirrotta V, Hadfield C, Hall JC, Rosbash M 1984. Molecular analysis of the *period* locus in *Drosophila melanogaster* and identification of a transcript involved in biological rhythms. *Cell 38*: 701-710.

Rogers AS, Escher SA, Pasetto C, Rosato E, Costa R, Kyriacou CP 2004a. A mutation in *Drosophila simulans* that lengthens the circadian period of locomotor activity. *Genetica 120*: 223-232.

Rogers AS, Rosato E, Costa R, Kyriacou CP 2004b. Molecular analysis of circadian clocks in *Drosophila simulans*. *Genetica 120*: 213-222.

Rosato E, Peixoto AA, Barbujani G, Costa R, Kyriacou CP 1994. Molecular polymorphism in the *period* gene of *Drosophila simulans*. *Genetics 138*: 693-707.

Rosato E, Peixoto AA, Costa R, Kyriacou CP 1997. Linkage disequilibrium, mutational analysis and natural selection in the repetitive region of the *clock* gene, *period*, in *Drosophila melanogaster*. *Genet Res 69*: 89-99.

Rosato E, Peixoto AA, Gallippi A, Kyriacou CP, Costa R 1996. Mutational mechanisms, phylogeny and evolution of a repetitive region within a *clock* gene of *Drosophila melanogaster*. *J Mol Evol 42*: 392-408.

Sawyer LA, Hennessy JM, Peixoto AA, Rosato E, Parkinson H, Costa R, Kyriacou CP 1997. Natural variation in a *Drosophila clock* gene and temperature compensation. *Science 278*: 2117-2120.

Sawyer LA, Sandrelli F, Pasetto C, Peixoto AA, Rosato E, Costa R, Kyriacou CP 2006. The *period* gene Thr-Gly polymorphism in Australian and African *Drosophila melanogaster* populations: implications for selection. *Genetics 174*: 465-480.

Schilcher FV 1977. A mutant which changes courtship song in *Drosophila melanogaster*. *Behav Genet 7*: 251-259.

Smith LA, Peixoto AA, Kramer EM, Villella A, Hall JC 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 LA, Wang XJ, Peixoto AA, Neumann EK, Hall LM, Hall JC 1996. A *Drosophila* Calcium channel α1 subunit gene maps to a genetic locus associated with behavioral and visual defects. *J Neurosci 16*: 7686-7879.

Souza NA, Andrade-Coelho CA, Vigoder FM, Ward RD, Peixoto AA 2008. Reproductive isolation between sympatric and allopatric Brazilian populations of *Lutzomyia longipalpis s.l.* (Diptera: Psychodidae). *Mem Inst Oswaldo Cruz 103*: 216-219.

Souza NA, Vigoder FM, Araki AS, Ward RD, Kyriacou CP, Peixoto AA 2004. Analysis of the copulatory courtship songs of *Lutzomyia longipalpis* in six populations from Brazil. *J Med Entomol 41*: 906-913.

Vigoder FM, Araki AS, Bauzer LGSR, Souza NA, Brazil RP, Peixoto AA 2010. Lovesongs and *period* gene polymorphisms indicate *Lutzomyia cruzi* (Mangabeira, 1938) as a sibling species of the *Lutzomyia longipalpis* (Lutz and Neiva, 1912) complex. *Infect Genet Evol 10*: 734-737.