

Molecular evidence to suggest the origin of a colonization: *Drosophila subobscura* in America

Pedro A. Araúz · Francesc Peris-Bondia ·
Amparo Latorre · Luís Serra · Francesc Mestres

Received: 2 November 2011 / Accepted: 26 March 2012 / Published online: 6 April 2012
© Springer Science+Business Media B.V. 2012

Abstract The recent colonization of America by *Drosophila subobscura* represents a great opportunity for evolutionary biology studies. Knowledge of the populations from which the colonization started would provide an understanding of how genetic composition changed during adaptation to the new environment. Thus, a 793 nucleotide fragment of the *Odh* (Octanol dehydrogenase) gene was sequenced in 66 chromosomal lines from Barcelona (western Mediterranean) and in 66 from Mt. Parnes (Greece, eastern Mediterranean). No sequence of *Odh* fragment in Barcelona or Mt. Parnes was identical to any of those previously detected in America. However, an *Odh* sequence from Barcelona differed in only one nucleotide from another found in American populations. In both cases, the chromosomal lines presented the same inversion: O₇, and the *Odh* gene was located within this inversion. This evidence suggests a possible western Mediterranean origin for the colonization. Finally, the molecular and inversion

data indicate that the colonization was not characterized by multiple reintroductions.

Keywords *D. subobscura* · *Odh* gene · Chromosomal inversion · Origin of colonization · Gene flow · Reintroduction

Introduction

Colonizations and invasions are often an undesirable element in ecosystem integrity and biodiversity, and result in substantial economic costs due to their impact on agriculture, marine aquaculture and human health (Davies et al. 1999; Lee 2002; Hess et al. 2009). However, they offer an excellent opportunity for evolutionary biology, as they enable researchers to study, for instance, the speed and predictability of evolution in nature (Lee 2002; Huey et al. 2005). A key issue in any colonization or invasion process is to determine their origin, since this makes it possible to prevent re-introductions. It also helps us understand the genetic variability of the source population when predicting the evolutionary potential of new established populations (Davies et al. 1999; Dlugosh and Parker 2008). In our opinion, colonizations or invasions can be classified into three groups: those which occurred a long time ago (millions or thousands of years), such as postglacial or volcanic island occupations (Taberlet et al. 1998; Hewitt 1999, 2000; Capi and Gibert 2004; Liggins et al. 2008), those which took place several centuries ago, probably due to journeys or other activities carried out by man (Fontdevila 1989; Gouin et al. 2003; Capi and Gibert 2004; Keller 2007; Tollenaere et al. 2010) and recent examples (less than 60–70 years ago), most of which are due to human global activity (Davies et al. 1999; Reiland et al. 2002;

Pedro A. Araúz and Francesc Peris-Bondia contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s10709-012-9647-8) contains supplementary material, which is available to authorized users.

P. A. Araúz · L. Serra · F. Mestres (✉)
Departament de Genètica, Facultat de Biologia, Universitat de
Barcelona, Av. Diagonal, 643, 08071 Barcelona, Spain
e-mail: fmestres@ub.edu

F. Peris-Bondia · A. Latorre
Centro Superior de Investigación en Salud Pública,
CSISP Generalitat Valenciana, Valencia, Spain

A. Latorre
Institut Cavanilles de Biodiversitat i Biologia Evolutiva,
Universitat de València, Valencia, Spain

Capy and Gibert 2004; Nardon et al. 2005; Rius et al. 2008; Hess et al. 2009). In the case of the latter group, it is more likely that the origin of the process can be determined with greater precision. In this article, we focus on the origin of the American colonization by *Drosophila subobscura*, which is a recent event and thus belongs to the third group described.

The colonization of the Americas by *D. subobscura* represented an excellent opportunity to analyse how different evolutionary mechanisms act in nature. This invasion took place on the west coasts of both North and South America, and was probably analysed -in both cases- from its earliest stages (Brncic et al. 1981; Beckenbach and Prevosti 1986). For this reason, this double colonization was considered a grand natural experiment with two replicates and a unique research opportunity (Ayala et al. 1989). Not only were the process and mechanisms of the colonization success studied in depth, but this was also a magnificent opportunity to analyse the ecology (ecological preferences and competitive ability with other American species of the *Drosophila* genus) and evolution of natural *D. subobscura* American populations. Fundamental information on these colonization events was obtained by classical genetic markers, such as chromosomal inversions, lethal genes and allozyme loci (for a summary see Ayala et al. 1989; Prevosti et al. 1989; Mestres et al. 2005). The magnitude of the bottlenecks produced was measured using different genetic markers (chromosomal polymorphism, allozymes, lethal genes and restriction-size variation of the *rp49* region and microsatellite loci) and it was concluded that the initial number of colonizers was between 8 and 15 (Brncic et al. 1981; Prevosti et al. 1989; Mestres et al. 1990; Rozas and Aguadé 1991; Pascual et al. 2001). The same composition in chromosomal arrangements and allozymes was detected in both colonized areas (Prevosti et al. 1988, 1989; Balanyà et al. 1994) and the same associations between lethal genes and chromosomal inversions were observed in North and South America, leading to the conclusion that both colonizations were not independent events (Mestres et al. 1992). Molecular markers such as mtDNA (Latorre et al. 1986; Rozas et al. 1990) and restriction-size variation of the *rp49* region (Rozas and Aguadé 1991) were in agreement with this finding. However, the population from which the colonization originated remained obscure: the chromosomal polymorphism of American samples resembled those generally obtained in the western Mediterranean region, with the dramatic exception of the O_5 inversion (Ayala et al. 1989; Prevosti et al. 1989). This inversion is relatively abundant (around 10 %) in Scandinavian *D. subobscura* populations (Sperlich 1964; Pinsker and Sperlich 1981; Mestres et al. 1994), but it has not been observed in the western Mediterranean (Prevosti et al. 1984; Solé et al.

2002; Mestres et al. 2005) and it has been reported only in old samples from the eastern Mediterranean (Krimbas 1964). However, the chromosomal composition of Scandinavian populations was not compatible with those detected in American populations of *D. subobscura*. The early use of molecular markers (such as mtDNA and restriction-map analysis of the *rp49* region) failed to provide any new insight into the origin of the colonization (Latorre et al. 1986; Rozas et al. 1990; Rozas and Aguadé 1991). Some indirect evidence suggested a Mediterranean origin: the analyses of evolution rates for quantitative traits in the American population are consistent with the pattern of enhanced evolution observed in northern latitudes (Gilchrist et al. 2001). Microsatellite analyses also support this hypothesis (Pascual et al. 2001), as well as predicting the direction of the double colonization: from the Palearctic region to South America and finally to North America (Pascual et al. 2007).

However, the puzzle with the O_5 inversion persisted. Although it is only relatively abundant in Scandinavia, its distribution in the remaining Palearctic region is erratic and presents negligible frequencies (Zivanovic and Mestres 2000). It has not been reported from the western Mediterranean area (which is a probable source of the colonization), and it is seldom found in low frequencies in the eastern Mediterranean region (for a revision see Krimbas 1993; Araúz et al. 2009a). It was even found sporadically in Israel (Goldschmidt 1956; Malogolowkin-Cohen and Sperlich 1981). In contrast, this inversion is relatively abundant in the American populations, presenting a significant latitudinal cline in both American hemispheres (Prevosti et al. 1988; Balanyà et al. 2003). Another peculiarity of the American O_5 inversions is their complete association with a lethal gene (Mestres et al. 1990, 1992, 1995, 2005, 2009), though it has proved to be heterotic in these populations (Mestres et al. 2001). In an early phase of research we characterized this inversion by sequencing the *Odh* (Octanol dehydrogenase) gene, which is located inside the O_5 inversion close to its proximal break point (Mestres et al. 2004). This gene is also located within the chromosomal inversions O_7 , O_1 and O_{22} , and outside O_2 (but close to its proximal break point). The American colonizing *Odh* sequences obtained (34 and 51 from North and South America, respectively) confirmed the small number of colonizers, the resemblance between both colonized areas, and the fact that only one O_5 inversion reached the American continent (Mestres et al. 2004; Gómez-Baldó et al. 2008). Many strong associations between the *Odh* haplotypes and chromosomal inversions were observed, but different recombinants were also detected, indicating that the historical (due to the founder event) associations were breaking. Only the adaptive associations remained through the generations (Gómez-Baldó et al. 2008).

To date, only non-Palearctic populations have been analysed at this genetic level to try to ascertain the origin of the colonization. The main aim of this study is to obtain *Odh* haplotypes from two Mediterranean populations supposed to be the most probable area from which the colonization started according to previous data, and to compare them with those previously obtained in American populations. One of these two populations is Barcelona (Spain) located in the western Mediterranean region, and the other is Mt. Parnes (Greece) in the eastern part of the Mediterranean. These two populations were chosen because they are well studied populations of *D. subobscura* (for a revision see Krimbas 1992, 1993 and Mestres et al. 2005) and characteristic of the Mediterranean areas tested in our hypotheses on the origin of the colonization. We examined whether any American haplotype for this gene is present in any of these Palearctic populations. We also studied the presence of associations of *Odh* haplotypes with chromosomal inversions, to deduce their evolutionary consequences. The study of the chromosomal inversion associations with the *Odh* gene sequences is very informative, but it is a laborious task. For this reason we focused on the Barcelona and Mt. Parnes populations. Finally, we examined whether repeated invasions from the Palearctic region might have occurred in this colonization.

Materials and methods

Populations and chromosomal lines

The Barcelona population was collected in the foothills of the Tibidabo mountain (located at the edge of Barcelona at approximately 400 m above sea level) in October 2004, whereas the Mt. Parnes population (about 25 km from Athens at 1,100 m above sea level) was sampled in May 2006 (Araúz et al. 2009a).

The homokaryotypic lines and lethal chromosomal lines were obtained by appropriate crosses using the *chcu* (*cherry curled*), homokaryotypic strain and *Va/Ba* (*Vari-cose/Bare*) balanced-lethal strain, as described in Mestres et al. (1990) and Araúz et al. (2009a).

Finally, the *Odh* sequences of the chromosomal lines from Barcelona and Mt. Parnes were compared with those available from America: 34 from USA, samples obtained in Gilroy (California, 37°33'N 121°31'W), Bellingham (Washington state, 48°45'N 122°29'W) and Centralia (Washington state, 46°43'N 122°58'W) populations and 51 from Chile, samples collected in Santiago de Chile (33°30'S 70°40'W) and Puerto Montt (41°28'S 73°00'W) populations (Mestres et al. 2004; Gómez-Baldó et al. 2008). The experimental procedure for obtaining these American chromosomal lines and their *Odh* sequences was

the same as that described below in the present manuscript (for more details see Mestres et al. 2004; Gómez-Baldó et al. 2008).

DNA extraction, PCR amplification and sequencing

Total DNA was isolated from a single fly using the protocol of Pascual et al. (1997). To amplify the *Odh* gene, the primers ODH-F and CD4 were used (described in Mestres et al. 2004). PCR conditions were: 94 °C for 5 min; 35 cycles of 94 °C for 1 min, 55 °C for 1 min, 72 °C for 1 min; with a final extension of 4 min at 72 °C. The QIAquick PCR Purification Kit (QIAGEN) was used to purify this PCR product, while direct sequencing was carried out using the following primers: ODH-F, ODHseq-R, C2 and CD6 (Mestres et al. 2004). Cycling conditions were: 96 °C for 1 min; 25 cycles of 96 °C for 10 s, 55 °C (45 °C for ODH-F primer) for 5 s, 60 °C for 4 min; and a final extension of 1 min at 4 °C. *Odh* genes were sequenced using an ABI PRISM™ 3700 DNA Analyser in the “Unitat de Genòmica, Serveis Científicotècnics” of the Universitat de Barcelona.

Sequences alignment and analysis

Sequence alignments were carried out with SeqMan™ II v. 4.03 (DNA Star Inc. 1999) and BioEdit v. 4.8.6 (Hall 1999). DnaSP v. 4 was used to analyse DNA polymorphism (Rozas et al. 2003). With this software, *h* (haplotype diversity), π (nucleotide diversity), θ (expected average number of nucleotide differences) and *k* (average number of nucleotide differences) were estimated. Finally, gene trees were reconstructed using the maximum likelihood composite method of the MEGA 5.02 Software (Tamura et al. 2011), applying the Tamura-Nei model, with gamma parameter and 500 bootstrap replicates.

Results

Nucleotide variation

A total of 132 sequences from a 793-nucleotide fragment of the *Odh* gene (containing intron 2, exon 3, intron 3 and part of exon 2 and exon 4) were obtained; 66 of them were from Barcelona (54 from homokaryotypic lines and 12 from lethal chromosomal lines), and 66 from Mt. Parnes (41 from homokaryotypic lines and 25 from lethal chromosomal lines). The descriptions of all chromosomal lines sequenced, including their GenBank/EMBL accession numbers, are shown in the Supplementary Table 1. Despite the existence of two introns, no indels (insertions or deletions) were found in any sequence. The estimates of

parameters that describe the nucleotide polymorphism of the *Odh* gene in Barcelona and Mt. Parnes populations are summarized in Table 1. In this table, the same parameters are also presented for the North and South American populations. In total, we have observed 48 nucleotide polymorphic sites in Barcelona and 45 in Mt. Parnes, but 23 are different in the two populations (most of them being singletons). As expected, the number of polymorphic sites in both populations is significantly higher in introns than in exons ($\chi^2 = 15.64$, d.f. = 1, $P = 0.0001$, with Yates correction and $\chi^2 = 7.20$, d.f. = 1, $P = 0.0073$, with Yates correction, for Barcelona and Mt. Parnes, respectively). For the coding region according to a test based on the binomial distribution (Mestres et al. 2001; Gómez-Baldó et al. 2008), the number of nucleotide changes in third codon positions was significantly higher than in the other positions (for both populations $k \geq 16$, $P = 0.000$). It is interesting to compare the polymorphic sites observed with those from American samples (Mestres et al. 2004; Gómez-Baldó et al. 2008): several sites were found in the American populations but not in the Mediterranean populations (Table 2). It could probably mean that there is a lot of variability at the level of these sites. However, in the studied populations several sites have been detected only once.

Nucleotide sequences and chromosomal arrangements

Considering all sequences together most haplotypes appear only once (116), with some sequences being shared between different chromosomal lines from the same population or even between Barcelona and Mt. Parnes (Table 3). There is one haplotype associated with the O_{3+4+1} arrangement, both in Barcelona and Mt. Parnes. However, this association is incomplete as other chromosomal lines O_{3+4+1} presented different haplotypes. In

Barcelona and Mt. Parnes, no other associations between *Odh* haplotypes and chromosomal arrangements were detected. Most importantly, no haplotypes found in American colonizing populations were observed in both Mediterranean populations analysed. However, the haplotype of the chromosomal line BC43 (from Barcelona) is almost identical to that found in American chromosomal lines S49, PM110 and PM57 (Mestres et al. 2004; Gómez-Baldó et al. 2008). All these American chromosomal lines come from Chile (S49 from Santiago de Chile; PM110 and PM57 from Puerto Montt). There is only one change in nucleotide 254, a third position in exon 3. Chromosomal line BC43 has a C in this position, while S49, PM110 and PM57 have a T. In both cases, the amino acid coded is the same: glycine. It is worth pointing out that BC43, S49 and PM110 lines present the same chromosomal arrangement (O_{3+4+7}), while PM57 has a derivate of it, O_7 , the product of an infrequent recombination event between O_{3+4+7} and O_{ST} chromosomes (Gómez-Baldó et al. 2008; Mestres et al. 2009). Another attempt to analyse the similarity in sequences and associations between haplotypes and

Table 2 Nucleotide polymorphic sites characteristic of American populations of *D. subobscura* and other populations where they were detected

American nucleotide site	Population in which nucleotide site was found
20	Mt. Parnes
87	Mt. Parnes
119	Barcelona
213	Only in North America
311	Only in North and South America
329	Only in North and South America
616	Barcelona

Table 1 Nucleotide polymorphism estimates for the *Odh* gene from Barcelona and Mt. Parnes populations

	Barcelona	Mt. Parnes	North America	South America
n	66	66	44	46
Nhap	59	62	11	11
h \pm SD	0.996 \pm 0.004	0.998 \pm 0.003	0.802 \pm 0.046	0.867 \pm 0.023
S	48	45	31	29
$\pi \pm$ SD	0.01263 \pm 0.00040	0.01303 \pm 0.00042	0.01030 \pm 0.00094	0.01177 \pm 0.00061
$\theta \pm$ SD (per site)	0.01273 \pm 0.00375	0.01194 \pm 0.00354	0.00891 \pm 0.00160	0.00833 \pm 0.00275
k \pm SD	10.007 \pm 1.854	10.317 \pm 1.882	8.237 \pm 1.695	9.372 \pm 1.830
$\theta \pm$ SD (per sequence)	10.086 \pm 1.456	9.455 \pm 1.450	7.126 \pm 1.280	6.598 \pm 1.225

For comparison, the values obtained in North and South American *D. subobscura* populations are also presented (Mestres et al. 2004; Gómez-Baldó et al. 2008)

n = number of sequences, Nhap = number of haplotypes, h = haplotype diversity, S = number of polymorphic sites, π = nucleotide diversity; θ = expected average number of nucleotide differences; k = average number of nucleotide differences

Table 3 Chromosomal lines whose haplotypes are shared inside one population or between populations or both

Barcelona	Mt. Parnes	Barcelona–Mt. Parnes
BC8 (O_{3+4+8}) and BC31 (O_{3+4})	MP3 (O_{3+4}) and MP4 (O_{3+4})	BC21 (O_{3+4+7}) and MP25 (O_{3+4+22})
BC9 (O_{3+4+1}) and BC20 (O_{3+4+1})	MP10 (O_{3+4}) and MP38 (O_{3+4})	BC55 (O_{ST}) and MP27 (O_{3+4+8})
FBC7 (O_{3+4}) and FBC58 (O_{ST})	MP24 (O_{3+4}) and MP91 (O_{3+4+22})	BC14 (O_{3+4+8}), MP24 (O_{3+4}) and MP91 (O_{3+4+22})
FBC22 (O_{3+4+8}) and FBC46 (O_{3+4+7})	MP36 (O_{3+4+1}) and MP78 (O_{3+4+1})	FBC7 (O_{3+4}), FBC58 (O_{ST}) and MP47 (O_{3+4+22})
FBC38 (O_{3+4+8}) and FBC62 (O_{3+4+8})		BC9 (O_{3+4+1}), BC20 (O_{3+4+1}), MP36 (O_{3+4+1}) and MP78 (O_{3+4+1})
FBC23 (O_{3+4+8}), FBC25 (O_{3+4+8}) and FBC51 (O_{3+4+8})		

The arrangement for each chromosomal line is shown in brackets. BC and FBC stand for chromosomal lines obtained from a wild male and from an individual son of a wild female offspring from Barcelona, respectively. MP stands for chromosomal lines obtained from a wild male from Mt. Parnes

chromosomal arrangements was carried out by creating gene trees. These trees for Barcelona and Mt. Parnes are presented in Fig. 1 and Fig. 2 of Supplementary material, respectively. In both cases, no clusters are detected and all sequences are mixed. Furthermore, main nodes are poorly supported, because very low bootstrap values are obtained. Additional gene trees, for instance using the sequences of both Palearctic populations together or those from Mt. Parnes and America, do not provide any valuable information (data not shown). The only exception is a joint analysis of the Barcelona and American sequences (Fig. 1), where the cluster of BC43, S49, PM110 and PM57 sequences can be observed, which is supported by a valid bootstrap value (85 %). They cluster because BC43 differs in only one nucleotide with regard to S49, PM110 and PM57 sequences.

Also interesting is the number of different *Odh* haplotypes observed in Barcelona and Mt. Parnes: 59 out of 66 sequenced chromosomal lines and 62 out of 66, respectively. It seems that both populations could present a large effective population size (N_e). As shown by Wright et al. (1942), the allelism of lethal genes is high when N_e is small, so most lethal genes in the population are identical by descendant. On the contrary, the allelism of lethal genes is low when N_e is large. In this population, few lethal genes are identical by descendant, and thus the allelic cases are scarce. A similar concept can be applied to the nucleotide sequences: the “allelism of sequences”, that is, determining how many are identical in all possible comparisons between two nucleotide sequences from the same populations. We would expect low values in populations with high N_e and vice versa. For the *Odh* gene, the estimated values of “allelism of sequences” for Barcelona and Mt. Parnes were 0.0037 ± 0.0013 and 0.0019 ± 0.0009 , respectively. However, these values are lower than those

computed using the data of Gómez-Baldó et al. (2008) from Chilean populations: 0.1176 ± 0.0260 and 0.1905 ± 0.0202 for Santiago de Chile and Puerto Montt, respectively. The low values from Chilean populations of *D. suboscuro* are most probably due to the founder effect (Ayala et al. 1989; Prevosti et al. 1989; Mestres et al. 2005).

Amino acid sequences

Non-coding regions were studied, but without providing any valuable information. However, the amino acid sequences derived from the nucleotide sequences gave us new insights into the connections between Barcelona, Mt. Parnes and American populations. The positions where amino acid changes were observed are presented in Table 4. Changes were detected in four, seven and five positions in Barcelona, Mt. Parnes and American populations, respectively. It is worth noting that amino acid positions 50 and 88 are exclusive to America, and that position number 7 is shared between America and Mt. Parnes. When analyzing Barcelona, Mt. Parnes and American populations, these amino acid changes define 16 different haplotypes (Supplementary Table 2). Their relative abundance (in percentage terms) is summarized in Supplementary Table 3. The distribution of haplotypes is similar in both European populations, with haplotype 6 being very abundant (61.29 and 57.14 % in Barcelona and Mt. Parnes, respectively), haplotype 1 presenting relatively high frequency (32.26 and 33.33 % in Barcelona and Mt. Parnes, respectively) and different haplotypes with very low frequencies. In American populations, haplotype 1 is the most abundant (62.07 %), haplotypes 2 (10.34 %) and 3 (20.69 %) are relatively frequent and haplotypes 4 and 5 are relatively infrequent.

Fig. 1 Gene tree obtained using the *Odh* nucleotide sequences from Barcelona and America. Only bootstrap values over 70 are presented. The cluster of BC43, S49, PM110 and PM57 sequences is marked with a square. Abbreviations for Barcelona population are as follows: BC and FBC stand for chromosomal lines obtained from a wild male and from an individual son of a wild female offspring from Barcelona, respectively. Abbreviations from American populations are as follows: BF and BM (Bellingham, Washington state USA); C (Centralia, Washington state, USA); G, GM, FGF (Gilroy, California, USA); S and SC (Santiago de Chile, Chile); PM (Puerto Montt, Chile). The arrangement of each chromosomal line sequenced is also presented after its abbreviation

Discussion

The origin of the colonization

Knowledge of the origin of a colonization (or invasion) would help evolutionary biologists to understand the amount of genetic variability reduction in the newly-established populations in comparison to the population from which the colonization started. This information could grant us a general overview of the quantity and kind of genetic variability needed to succeed in the colonization of a particular environment (Lee 2002; Dlugosh and Parker 2008). Unfortunately, it is difficult to track with any precision the origin of a colonization that took place in the distant past. However, recent colonizing events provide more accurate information (genetic markers have still accumulated few changes), which would tend to help us to ascertain the population or specific geographical region of origin. The *Drosophila subobscura* colonization of the American continent is generally considered to have taken place in the late 1970s (for a review see Ayala et al. 1989; Prevosti et al. 1989; Huey et al. 2005; Mestres et al. 2005). Ascertaining the origin of New World populations was a priority from the earliest studies (Brncic et al. 1981). Different genetic markers provided different levels of information: chromosomal inversion polymorphism indicated a western Mediterranean origin, but the presence of the O_5 inversion in American populations was difficult to explain (Brncic et al. 1981; Ayala et al. 1989; Prevosti et al. 1989). The O_5 inversion is found with a frequency of around a 10 % in Scandinavia (for a revision see Krimbas 1993), but the origin of the colonization cannot be from Northern Europe due to the chromosomal composition of American samples. Several chromosomal arrangements found in the American continent have never been detected in this European region (Prevosti et al. 1989; Mestres et al. 1990; Zivanovic and Mestres 2000). The *D. subobscura* mtDNA haplotypes and microsatellites were not conclusive, but they were compatible with a western Mediterranean origin (Latorre et al. 1986; Pascual et al. 2001). Joint analysis of *Odh* sequences and chromosomal inversions gives new insight into the origin of this colonization. While

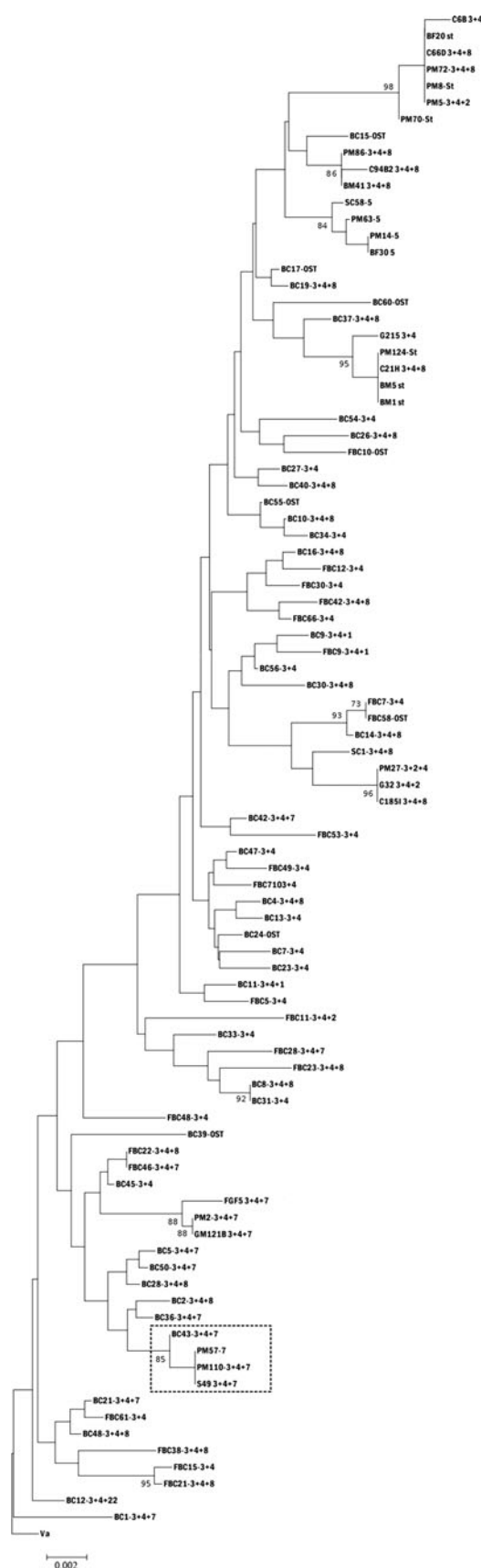


Table 4 Positions of amino acid change, type of amino acid change and number of times that the same change was detected in the Barcelona, Mt. Parnes and American populations for the ODH peptide

	Changes in amino acid positions											
	1	7	7	34	50	88	120	148	149	216	218	218
<i>Va</i> (reference)	E	V	V	D	V	F	G	E	T	E	A	A
Change	G	M	E	E	I	L	S	D	S	K	D	V
America			7		2	7		3	29			
Barcelona	1							1	21		2	
Mt. Parnes		1	1	1			1	1	23	1		1
Total	1	1	8	1	2	7	1	5	73	1	2	1

The ODH peptide of the *Va* chromosome (balancer chromosome used for obtaining the chromosomal lines) is used as reference for the amino acid changes. Positions 7 and 218 appear twice because two different amino acid changes had taken place in them (V to M and V to E for position 7; A to D and A to V for position 218)

E = Glutamic acid; V = Valine; D = Aspartic acid; F = Phenylalanine; G = Glycine; T = Threonine; A = Alanine; M = Methionine; I = Isoleucine; L = Leucine; S = Serine; K = Lysine

none of the Palearctic *Odh* sequences from Barcelona or Mt. Parnes has been found in America, one sequence from Barcelona (BC43) was almost identical to one found in the New World, differing only in one nucleotide (American chromosomal lines S49, PM110 and PM57). And most importantly, both sequences were located inside the same inversion, the O_7 . This evidence strongly supports a possible origin from the western Mediterranean region. However, this hypothesis has a drawback: the O_5 inversion has not been reported in the Iberian Peninsula, despite the fact that chromosomal inversion polymorphism has been studied extensively (for a review see Krimbas 1993; Mestres et al. 2005). In all probability, the O_5 inversion is adaptive to cold conditions, as in the Palearctic region it is found mainly in Scandinavia (Krimbas 1993; Ruiz-Martín 2006), and in the New World it presents significant latitudinal clines in North and South America (Prevosti et al. 1985, 1988, 1990) which persisted over time (Balanyà et al. 2003), despite being completely associated with a particular lethal gene (Mestres et al. 2001). As molecular data show, this inversion probably appeared recently (Araúz et al. 2009b). Although it is adapted to cold conditions, it may well have spread from Scandinavia to other warmer Palearctic regions by gene flow, and would then be eliminated by selection or genetic drift if this inversion reaches these warmer areas. Thus, with the exception of Scandinavia, its distribution is erratic in the Old World and presents negligible frequency (for a revision see Zivanovic and Mestres, 2000). Hence, one possible explanation is that at a certain moment, it could have reached the western Mediterranean population from which the colonization originated and be included in the sample of colonizers. It is true that this inversion was found in Greece a long time ago (Krimbas 1964, 1967; Krimbas and Alevizos, 1973), but since 1975 (just before the probable beginning of the American colonization) it has not been reported again

(Loukas et al. 1979, 1980, 1981; Araúz et al. 2009a). Furthermore, the inversion composition of the Greek populations, though possible, makes it unlikely that any of them originated the colonizing process. For instance, Mt. Parnes population (and other Greek populations) presented a non-negligible frequency of O_{3+4+1} (between 14.66 and 28.00 %) and O_{3+4+22} (between 2.38 and 11.47 %) chromosomal arrangements (Krimbas 1967; Araúz et al. 2009a), but neither has been found in American *D. subobscura* populations (Prevosti et al. 1985, 1988, 1989, 1990; Balanyà et al. 2003). Furthermore, different O_{3+4+7} arrangements reached the American continent (Mestres et al. 1990, 1995, 2004; Gómez-Baldó et al. 2008), whereas its frequency is very low in Mt. Parnes and other Greek populations, with values ranging from 0 to 6.65 % (Krimbas 1967; Araúz et al. 2009a). For these reasons an origin in an eastern Mediterranean population seems unlikely. Thus, the most probable explanation is that the colonization originated from a western Mediterranean population which fortuitously presented the sporadic O_5 inversion (due to gene flow) and, by chance, it was included in the sample of colonizers. However, more analyses of chromosomal polymorphism and *Odh* sequences will be needed to pinpoint more accurately the particular area from which the colonization started.

Repeated invasions that superimpose onto one another have been described in some species (Davies et al. 1999; Ellstrand and Schierenbek 2000; Bossdorf et al. 2005; Frankham 2005; Krafur et al. 2005; Novak and Mack 2005; Dlugosh and Parker 2008). In the case of the American colonization by *D. subobscura*, chromosomal inversion data accumulated from the beginning of the colonization indicate that the colonization was a unique event, since new genetic variability at this level has not been found since the earlier studies (Prevosti et al. 1988, 1989; Ayala et al. 1989; Balanyà et al. 2003). With regard

to the *Odh* nucleotide and amino acid sequence, their variability in Barcelona and Mt. Parnes compared with that of American populations indicates that the founder event was considerable (Mestres et al. 2004; Gómez-Baldó et al. 2008). Repeated invasions would produce an increment of new *Odh* sequences in the American continent. However, no new *Odh* sequences have been reported in the American continent, with the exception of those that appeared there through recombination (Gómez-Baldó et al. 2008). In the future, more studies will be needed to ascertain if new genetic variability is introduced in America from Palearctic populations.

Nucleotide polymorphism in the Mediterranean populations

The nucleotide polymorphism for the *Odh* gene presents very similar values in Barcelona and Mt. Parnes populations and slightly lower values of this polymorphism were observed for the *Odh* gene in American populations of *D. subobscura*, a consequence of the founder event (Mestres et al. 2004; Gómez-Baldó et al. 2008). In general, the values π (nucleotide diversity) and θ (expected average number of nucleotide differences) are equivalent to those found in the Palearctic region when other nuclear genes of *D. subobscura* were studied (Rozas and Aguadé 1990; Rozas et al. 1995, 1999; Cirera and Aguadé 1998; Navarro-Sabaté et al. 1999; Llopart and Aguadé, 2000; Munté et al. 2005). In these studies, the π estimates range between 0.004 and 0.012, and those for θ between 0.010 and 0.021.

Barcelona and Mt. Parnes populations appear to have a large effective population size (N_e) due to the high number of different *Odh* haplotypes observed and the low values obtained in the “allelism of sequences”. N_e was estimated using the lethal genes allelism, and the values ranged between 6,964–13,004 for Barcelona and 11,874–26,828 for Mt. Parnes (Araúz et al. 2009a), confirming the large effective population size of both populations. However, although Barcelona and Mt. Parnes are large populations with a high number of different *Odh* haplotypes, a high gene flow was observed at this molecular level. Indeed, different haplotypes are shared by both populations and in one case they even presented the same chromosomal arrangement, O_{3+4+1} (Table 3). The gene flow between distant Palearctic populations has been observed using different genetic markers, such as mtDNA (Latorre et al. 1992), microsatellites (Pascual et al. 2001) and lethal genes (Zivanovic et al. 2007). Finally, most haplotypes were not associated with inversions and only one seems to be partially associated with the O_{3+4+1} arrangement. This observation is in clear contrast with the associations between *Odh* haplotypes and chromosomal arrangements

in North and South America, a product of the colonization (Mestres et al. 2004; Gómez-Baldó et al. 2008).

In summary, the American colonization by *D. subobscura* was a unique event and the colonizers most probably came from the western Mediterranean region. Our results are in agreement with some previous indirect evidence (Ayala et al. 1989; Prevosti et al. 1989). Secondary colonizations explain the expansion of the species from the west coast of South America to the west coast of North America and the eastern coast of Argentina (Pascual et al. 2007; Fernández-Iriarte et al. 2009). Although the initial sample of colonizers was small, it contained the genetic variability needed to achieve the colonization success. As proposed in Lee (2002), the initial sample of colonizers was influenced by the genetic architecture of the original Palearctic population, and natural selection operated rapidly on this genetic basis to allow adaptation to the new environments. Finally, it is worth pointing out that molecular markers in association with chromosomal inversions are useful tools to locate the origin of recent colonizations and invasions. Thus, it may be possible to study the evolutionary potential of the new established populations and the possibility of reintroductions.

Acknowledgments We thank Mr. R. Rycroft and Mr. M. Bunn (S.A.L. Universitat de Barcelona) for corrections to the English text. This study was supported by the following grants: CGL2006-13423-C02-02 from the Ministerio de Ciencia y Tecnología (MCYT, Spain), BFU2009-07564 and BFU2009-12895-CO2-01 from the Ministerio de Ciencia e Innovación (Spain) and 2009 SGR 636 from the Generalitat de Catalunya (Spain). FM is member of IRBio (Institut de Recerca de la Biodiversitat).

References

- Araúz PA, Mestres F, Pegueroles C, Arenas C, Tzannidakis G, Krimbas CB, Serra L (2009a) Tracking the origin of the American colonization by *D. subobscura*: Genetic comparison between Eastern and Western Mediterranean populations. *J Zool Syst Evol Res* 47:25–34
- Araúz PA, Pegueroles C, Zivanovic G, Fernández-Iriarte PJ, Ruiz-Martín H, Balanyà J, Serra L, Mestres F (2009b) Recent origin of a chromosomal inversion revealed by its association with nucleotide haplotypes. *Hereditas* 146:257–259
- Ayala FJ, Serra L, Prevosti A (1989) A grand experiment in evolution: the *D. subobscura* colonization of the Americas. *Genome* 31:246–255
- Balanyà J, Segarra C, Prevosti A, Serra L (1994) Colonization of America by *D. subobscura*: the founder event and a rapid expansion. *J Hered* 85:427–432
- Balanyà J, Serra L, Gilchrist GW, Huey RB, Pascual M, Mestres F, Solé E (2003) Evolutionary pace of chromosomal polymorphism in colonizing populations of *D. subobscura*: an evolutionary time series. *Evolution* 57:1837–1845
- Beckenbach AT, Prevosti A (1986) Colonization of North America by the European species *D. subobscura* and *D. ambigua*. *The Am Midl Nat* 115:10–18

- Bossdorf O, Auge H, Lafuma L, Rogers WE, Siemann E, Prati D (2005) Phenotypic and genetic differentiation between native and introduced plant populations. *Oecologia* 144:1–11
- Brncic D, Prevosti A, Budnik M, Monclús M, Ocaña J (1981) Colonization of *D. subobscura* in Chile. I. First population and cytogenetic studies. *Genetica* 56:3–9
- Capy P, Gibert P (2004) *Drosophila melanogaster*, *Drosophila simulans*: so similar yet so different. *Genetica* 120:5–16
- Cirera S, Aguadé M (1998) Molecular evolution of duplication: the sex-peptide (*Acp70A*) gene region of *D. subobscura* and *D. madeirensis*. *Mol Biol Evol* 15:988–996
- Davies N, Villablanca FX, Roderick GK (1999) Bioinvasions of the Medfly *Ceratitis capitata*: source estimation using DNA sequences at multiple intron loci. *Genetics* 153:351–360
- Dlugosh KM, Parker IM (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Mol Ecol* 17:431–449
- Ellstrand NC, Schierenbek KA (2000) Hybridization as a stimulus for the evolution of invasiveness in plants? *Proc Natl Acad Sci USA* 97:7043–7050
- Fernández-Iriarte PJ, Balanyà J, Pascual M, Mestres F, Hasson E, Fontdevila A, Serra L (2009) Tracking the origin of an invasive species: *D. subobscura* in Argentina. *J Evol Biol* 22:650–658
- Fontdevila A (1989) Founder effects in colonizing populations: the case of *Drosophila buzzatii*. In: Fontdevila A (ed) *Evolutionary biology of transient unstable populations*. Springer, Berlin, pp 74–95
- Frankham R (2005) Resolving the genetic paradox in invasive species. *Heredity* 94:385
- Gilchrist GW, Huey RB, Serra L (2001) Rapid evolution of wing size clines in *D. subobscura*. *Genetica* 112–113:273–286
- Goldschmidt E (1956) Chromosomal polymorphism in a population of *D. subobscura* from Israel. *J Genet* 54:474–496
- Gómez-Baldó L, Latorre A, Serra L, Mestres F (2008) Molecular variation in the *Odh* gene in Chilean natural populations of *D. subobscura*. *Hereditas* 145:154–162
- Gouin N, Granjean F, Pain S, Souty-Grosset C, Reynolds J (2003) Origin and colonization history of the white-clawed crayfish, *Austropotamobius pallipes*, in Ireland. *Heredity* 91:70–77
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acid Ser* 41:95–98
- Hess JE, Swalla BJ, Moran P (2009) New molecular markers to genetically differentiate populations of *Didemnum vesillum* (Kott, 2002): an invasive ascidian species. *Aquat Invasions* 4:299–310
- Hewitt GM (1999) Post-glacial re-colonization of European biota. *Biol J Linn Soc* 68:87–112
- Hewitt GM (2000) The genetic legacy of the quaternary ice ages. *Nature* 405:907–913
- Huey RB, Gilchrist GW, Hendry AP (2005) Using invasive species to study evolution. Case studies with *Drosophila* and Salmon. In: Sax DF, Gaines SD, Stachowicz JJ (eds) *Species invasions: insights to ecology evolution and biogeography*. Sinauer Associates, Sunderland, MA, pp 139–164
- Keller A (2007) *Drosophila melanogaster*'s history as a human commensal. *Current Biol* 17:R77–R81
- Krafsur ES, Cummings MA, Endsley MA, Marquez JG, Nason JD (2005) Geographic differentiation in the house fly estimated by microsatellite and mitochondrial variation. *J Hered* 96:502–512
- Krimbas CB (1964) The genetics of *D. subobscura* populations. I. Inversion polymorphism in natural populations of Southern Greece. *Evolution* 18:541–552
- Krimbas CB (1967) The genetics of *D. subobscura* populations. III. Inversion polymorphism and climatic factors. *Mol Gen Genet* 99:133–150
- Krimbas CB (1992) The inversion polymorphism in *D. subobscura*. In Krimbas CB, Powell JR (eds) *Drosophila inversion polymorphism*. CRC Press Inc., Boca Raton, pp 127–220
- Krimbas CB (1993) *D. subobscura*: Biology, Genetics and Inversion polymorphism. Verlag Dr. Kovac, Hamburg
- Krimbas CB, Alevizos V (1973) The genetics of *D. subobscura* populations. IV. Further data on inversion polymorphism in Greece: evidence of microdifferentiation. *Egypt J Genet Cytol* 2:121–132
- Latorre A, Moya A, Ayala FJ (1986) Evolution of mitochondrial DNA in *D. subobscura*. *Proc Natl Acad Sci USA* 83:8649–8653
- Latorre A, Hernandez C, Martinez D, Castro JA, Ramon M, Moya A (1992) Population structure and mitochondrial DNA gene flow in Old World populations of *Drosophila subobscura*. *Heredity* 68:15–24
- Lee CE (2002) Evolutionary genetics of invasive species. *Trends Ecol Biol* 17:386–391
- Liggins L, Chapple DG, Daugherty CH, Ritchie PA (2008) Origin and post-colonization evolution of the Chatham Islands skink (*Oligosoma nigroplantare nigroplantare*). *Mol Ecol* 17:3290–3305
- Llopart A, Aguadé M (2000) Nucleotide polymorphism at the *RpII215* gene in *D. subobscura*: Weak selection on synonymous mutations. *Genetics* 155:1245–1252
- Loukas M, Krimbas CB, Vergini Y (1979) The genetics of *D. subobscura* populations. IX. Studies on linkage disequilibrium in four natural populations. *Genetics* 93:497–523
- Loukas M, Krimbas CB, Morgan K (1980) The genetics of *D. subobscura* populations. XIV. Further data on linkage disequilibria. *Genetics* 95:757–768
- Loukas M, Vergini Y, Krimbas CB (1981) The genetics of *D. subobscura* populations. XVII. Further genetic heterogeneity within electromorphs by urea denaturation and the effect of the increased genic variability on linkage disequilibrium studies. *Genetics* 97:429–441
- Malogolowkin-Cohen C, Sperlich D (1981) The effect of isolation and marginality on the inversion polymorphism of *D. subobscura* in Israel. *Rev Bras Genet* 2:213–230
- Mestres F, Pegueroles G, Prevosti A, Serra L (1990) Colonization of America by *D. subobscura*: lethal genes and the problem of the O₅ inversion. *Evolution* 44:1823–1836
- Mestres F, Balanyà J, Segarra C, Prevosti A, Serra L (1992) Colonization of America by *D. subobscura*: analysis of the O₅ inversions from Europe and America and their implications for the colonizing process. *Evolution* 46:1564–1568
- Mestres F, Balanyà J, Segarra C, Prevosti A, Serra L (1994) O chromosome inversion polymorphism in Northern and Atlantic Europe and its implications in the American colonization of *D. subobscura*. *Z Zool Syst Evol Forsch* 32:108–116
- Mestres F, Serra L, Ayala FJ (1995) Colonization of the Americas by *D. subobscura*: lethal-gene allelism and association with chromosomal arrangements. *Genetics* 140:1297–1305
- Mestres F, Balanyà J, Arenas C, Solé E, Serra L (2001) Colonization of America by *D. subobscura*: heterotic effect of chromosomal arrangements revealed by the persistence of lethal genes. *Proc Natl Acad Sci USA* 98:9167–9170
- Mestres F, Abad L, Sabater-Muñoz B, Latorre A, Serra L (2004) Colonization of America by *D. subobscura*: Association between *Odh* gene haplotypes, lethal genes and chromosomal arrangements. *Genes Genet Syst* 79:233–244
- Mestres F, Balanyà J, Pascual M, Arenas C, Solé E, Serra L (2005) Lethal genes and the colonization of America by *D. subobscura*. *Curr Top Genet* 1:31–55
- Mestres F, Balanyà J, Pascual M, Arenas C, Gilchrist GW, Huey RB, Serra L (2009) Evolution of Chilean colonizing populations of *D. subobscura*: lethal genes and chromosomal arrangements. *Genetica* 136:37–48

- Munté A, Rozas J, Aguadé M, Segarra C (2005) Chromosomal inversion polymorphism leads to extensive genetic structure: a multilocus survey in *D. subobscura*. *Genetics* 169:1573–1581
- Nardon C, Deceliere G, Loevenbruck C, Weiss M, Vieira C, Biémont C (2005) Is genome size influenced by colonization of new environments in dipteran species? *Mol Ecol* 14:869–878
- Navarro-Sabaté A, Aguadé M, Segarra C (1999) The relationship between allozyme and chromosomal polymorphism inferred from nucleotide variation at the *Acph-1* gene region of *D. subobscura*. *Genetics* 153:871–889
- Novak SJ, Mack RN (2005) Genetic bottlenecks in alien plant species: influence of mating systems and introduction dynamics. In: Sax DF, Gaines SD, Stachowicz JJ (eds) *Exotic species—bane to conservation and boon to understanding: ecology, evolution and biogeography*. Sinauer Associates, Sunderland, MA, pp 95–122
- Pascual M, Balanyà J, Latorre A, Serra L (1997) Analysis of the variability of *D. azteca* and *D. athabasca* populations revealed by randomly amplified polymorphic DNA. *J Zool Syst Evol Res* 35:159–164
- Pascual M, Aquadro CF, Soto V, Serra L (2001) Microsatellite variation in colonizing and Palearctic populations of *D. subobscura*. *Mol Biol Evol* 18:731–740
- Pascual M, Chapuis MP, Mestres F, Balanyà J, Huey RB, Gilchrist GW, Serra L, Estoup A (2007) Introduction history of *D. subobscura* in the New World: a microsatellite-based survey using ABC methods. *Mol Ecol* 16:3069–3083
- Pinsker W, Sperlich D (1981) Geographic pattern of allozyme and inversion polymorphism in chromosome O of *D. subobscura*. *Genetica* 57:51–64
- Prevosti A, de Frutos R, Alonso G, Latorre A, Monclús M, Martínez MJ (1984) Genetic differentiation between natural populations of *D. subobscura* in the Western Mediterranean area with respect to chromosomal variation. *Génét Séle Evol* 16:143–156
- Prevosti A, Serra L, Ribó G, Aguadé M, Segarra E, Monclús M, García MP (1985) The colonization of *D. subobscura* in Chile. II. Clines in the chromosomal arrangements. *Evolution* 39:838–844
- Prevosti A, Ribó G, Serra L, Aguadé M, Balanyà J, Monclús M, Mestres F (1988) Colonization of America by *D. subobscura*: experiment in natural populations that supports the adaptive role of chromosomal-inversion polymorphism. *Proc Natl Acad Sci USA* 85:5597–5600
- Prevosti A, Serra L, Aguadé M, Ribó G, Mestres F, Balañá J, Monclús M (1989) Colonization and establishment of the palearctic species *D. subobscura* in North and South America. In: Fontdevila A (ed) *Evolutionary biology of transient unstable populations*. Springer, Berlin, pp 114–129
- Prevosti A, Serra L, Segarra C, Aguadé M, Ribó G, Monclús M (1990) Clines of chromosomal arrangements of *D. subobscura* in South America evolve closer to Old World patterns. *Evolution* 44:218–221
- Reiland J, Hodge S, Noor MAF (2002) Strong founder effect in *Drosophila pseudoobscura* colonizing New Zealand from North America. *J Hered* 93:415–420
- Rius M, Pascual P, Turon X (2008) Phylogeography of the widespread marine invader *Microcosmus squamiger* (Ascidacea) reveals high genetic diversity of introduced populations and non-independent colonizations. *Divers Distrib* 14:818–828
- Rozas J, Aguadé M (1990) Evidence of extensive genetic exchange in the *rp49* region among polymorphic chromosome inversions in *D. subobscura*. *Genetics* 126:417–426
- Rozas J, Aguadé M (1991) Using restriction-map analysis to characterize the colonization process of *D. subobscura* on the American continent. I. *rp49* region. *Mol Biol Evol* 8:447–457
- Rozas J, Hernández M, Cabrera CV, Prevosti A (1990) Colonization of America by *D. subobscura*: Effect of the founder event on the mitochondrial DNA polymorphism. *Mol Biol Evol* 7:103–109
- Rozas J, Segarra C, Zapata C, Alvarez G, Aguadé M (1995) Nucleotide polymorphism at the *rp49* region of *D. subobscura*: lack of geographic subdivision within chromosomal arrangements in Europe. *J Evol Biol* 8:355–367
- Rozas J, Segarra C, Ribó G, Aguadé M (1999) Molecular population genetics of the *rp49* gene region in different chromosomal inversions of *D. subobscura*. *Genetics* 151:189–202
- Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19:2496–2497
- Ruiz-Martín H (2006) Estudi de les associacions entre al·lels de loci microsatel·lits i inversions cromosòmiques en poblacions europees i colonitzadores de *D. subobscura*. Diploma d'Estudis Avançats. Dept. Genètica, Universitat de Barcelona, Spain
- Solé E, Balanyà J, Sperlich D, Serra L (2002) Long-term changes in the chromosomal inversion polymorphism of *D. subobscura*. I. Mediterranean populations from Southwestern Europe. *Evolution* 56:830–835
- Sperlich D (1964) Chromosomale Strukturanalyse und Fertilitätsprüfung an einer Marginalpopulation von *D. subobscura*. *Z Vererbungslehre* 95:73–81
- Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Mol Ecol* 7:453–464
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28:2731–2739
- Tollenaere C, Brouart C, Duplantier J-M, Rahalison L, Rahelinirina S, Pascal M, Moné H, Mouahid G, Leirs H, Cosson J-F (2010) Phylogeography of the introduced species *Rattus rattus* in the western Indian Ocean, with special emphasis on the colonization history of Madagascar. *J Biogeogr* 37:398–410
- Wright S, Dobzhansky Th, Hovanitz W (1942) Genetics of natural populations. VII. The allelism of lethals in the third chromosome of *D. pseudoobscura*. *Genetics* 27:363–394
- Zivanovic G, Mestres F (2000) Lethal genes in O₅ chromosomes of *D. subobscura* from Europe and America. *J Zool Syst Evol Res* 38:123–126
- Zivanovic G, Arenas C, Mestres F (2007) The genetic structure of Balkan populations of *D. subobscura*. *Hereditas* 144:120–128