

Research article

Open Access

## Distribution of the transposable elements *bilbo* and *gypsy* in original and colonizing populations of *Drosophila subobscura*

María Pilar García Guerreiro\*<sup>1</sup>, Blanca E Chávez-Sandoval<sup>1</sup>, Joan Balanyà<sup>2</sup>, Lluís Serra<sup>2</sup> and Antonio Fontdevila<sup>1</sup>

Address: <sup>1</sup>Grup de Biologia Evolutiva. Departament de Genètica i Microbiologia, Facultat de Biociències, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain. and <sup>2</sup>Grup de Biologia Evolutiva. Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, 08071 Barcelona, Spain.

Email: María Pilar García Guerreiro\* - mariapilar.garcia.guerreiro@uab.es; Blanca E Chávez-Sandoval - blancaestela.chavez@uab.es; Joan Balanyà - jbalanya@ub.edu; Lluís Serra - lserra@ub.edu; Antonio Fontdevila - antonio.fontdevila@uab.es

\* Corresponding author

Published: 14 August 2008

Received: 15 February 2008

BMC Evolutionary Biology 2008, 8:234 doi:10.1186/1471-2148-8-234

Accepted: 14 August 2008

This article is available from: <http://www.biomedcentral.com/1471-2148/8/234>

© 2008 García Guerreiro et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Background:** Transposable elements (TEs) constitute a substantial amount of all eukaryotic genomes. They induce an important proportion of deleterious mutations by insertion into genes or gene regulatory regions. However, their mutational capabilities are not always adverse but can contribute to the genetic diversity and evolution of organisms. Knowledge of their distribution and activity in the genomes of populations under different environmental and demographic regimes, is important to understand their role in species evolution. In this work we study the chromosomal distribution of two TEs, *gypsy* and *bilbo*, in original and colonizing populations of *Drosophila subobscura* to reveal the putative effect of colonization on their insertion profile.

**Results:** Chromosomal frequency distribution of two TEs in one original and three colonizing populations of *D. subobscura*, is different. Whereas the original population shows a low insertion frequency in most TE sites, colonizing populations have a mixture of high (frequency  $\geq 10\%$ ) and low insertion sites for both TEs. Most highly occupied sites are coincident among colonizing populations and some of them are correlated to chromosomal arrangements. Comparisons of TE copy number between the X chromosome and autosomes show that *gypsy* occupancy seems to be controlled by negative selection, but *bilbo* one does not.

**Conclusion:** These results are in accordance that TEs in *Drosophila subobscura* colonizing populations are submitted to a founder effect followed by genetic drift as a consequence of colonization. This would explain the high insertion frequencies of *bilbo* and *gypsy* in coincident sites of colonizing populations. High occupancy sites would represent insertion events prior to colonization. Sites of low frequency would be insertions that occurred after colonization and/or copies from the original population whose frequency is decreasing in colonizing populations. This work is a pioneer attempt to explain the chromosomal distribution of TEs in a colonizing species with high inversion polymorphism to reveal the putative effect of arrangements in TE insertion profiles. In general no associations between arrangements and TE have been found, except in a few cases where the association is very strong. Alternatively, founder drift effects, seem to play a leading role in TE genome distribution in colonizing populations.

## Background

TEs are widely distributed in eukaryotes, representing 50% of the human genome [1], 15% of the *Drosophila* genome, and up to 70% in *Zea mays* [2]. Because of their capacity of transposition they are able to invade the genome and promote insertional mutations and chromosomal rearrangements. Recurrent mobility allows them to persist in spite of their harmful effects in the host [3]. Most of the proposed models in population dynamic studies [4-8] suggest that TEs are able to invade the genome if their transposition rate is enough to balance out opposing forces as excision and selection against deleterious insertions and chromosomal arrangements. Yet, these models, often too general, do not consider that each element behaves depending on both its own characteristics and the history of the population to which it belongs. This challenge to standard reasoning is most relevant in colonizing populations [9]. Several authors have suggested that bursts of transposition could be induced in colonization by the foreign, often stressful, environment faced by the founders of colonizing populations [10,11]. Moreover, colonizing populations are subjected to well documented founder, drift effects [12]. Both processes generate population instabilities that may incorporate new variables to the interpretation of TE occupancy profiles in colonizing populations. These considerations qualify the study of TEs in colonization as of prime interest to understanding their invasive dynamics and putative evolutionary role in populations.

Colonization effects on TEs were studied in *Drosophila* species [9,11,13] showing that this process plays an important role in the TE chromosomal distribution. In particular studies in colonizing populations of *D. buzzatii* showed a TE bimodal distribution with sites either highly occupied, in a few cases, or showing low insertion occupancy, in most cases. Molecular studies of TE copies from high and low occupied sites [14] strongly indicated that the most reliable explanation of the observed bimodal distribution is that a founder effect followed by genetic drift occurred during the colonization process. These results notwithstanding, valid for *D. buzzatii*, cannot be generalized to other colonizing *Drosophila* species, with different genomic characteristics, and subjected to different environmental pressures.

*D. subobscura*, a Palearctic species belonging to the obscura group [15] and characterized by a rich inversion polymorphism [16], has colonized North and South America almost 30 years ago [17,18]. It was found for the first time in Puerto Montt (Chile) in 1978 [19] and later near Port Townsend in Washington (USA) in 1982 [20]. Thereafter this species showed a rapid spread and adaptation to the new colonized environment in form of latitudinal clines

for chromosomal polymorphism and body size that paralleled the Palearctic clines [17,18,21]. Main after-colonization population effects were the presence of allelic lethal genes in different populations [22], the low genetic variability of mtDNA [23,24] and the reduction of microsatellite allele numbers [25] compared to original founder populations. These are expected outcomes of the founder drift effect of colonization. However nothing is known of the impact of colonization on the TE chromosomal distribution in this species.

Here we present the study of the distribution of two TEs, *gypsy* and *bilbo*, in original and colonizing populations of *D. subobscura*. Results show that TE frequency distribution differ between original and colonizing populations in a way that colonization, chromosomal inversion polymorphism and particular characteristics inherent to each element can provide a sufficient likely explanation. In this paper we particularly emphasize the importance of population structure and history to explain TE distribution in natural populations.

## Results

### Chromosomal distribution of *bilbo* and *gypsy*

We analyzed the distribution of *bilbo* and *gypsy* in polytene chromosomes of *D. subobscura*. Fig. 1 shows two examples of chromosomal distribution: *bilbo* in chromosome O and *gypsy* in chromosome U. A different distribution pattern is observed, in general, when we compare colonizing and original populations. Colonizing populations present insertion frequencies of *bilbo* and *gypsy* higher than those of the original population. In general the same distribution pattern is observed for the rest of chromosomes. Ten sites (7A, 16A, 20A, 45C, 58D, 74D 82A, 83C, 85A, 89C) show a *bilbo* insertion polymorphism greater than 32% in at least one colonizing population. *Gypsy* insertion frequencies are lower than those of *bilbo* with an occupancy of more than 10% in eight chromosomal sites (39D, 41C, 43D, 49D, 52D, 63C, 71B, 74D). Differences in occupancy profiles between original and colonizing populations are represented in Table 1 that shows the distribution of the number of times that each site is occupied in the studied sample. Thus, in *bilbo* the occupancy frequency ranges from 1 to 51 times in colonizing populations and only from 1 to 19 in the original population. Although *gypsy* shows a low occupancy profile compared to *bilbo* (colonizing populations range: 1–15; original population range: 1–5), the occupancy rate of both TEs in colonizing populations is greater than in the original population. The highest *bilbo* insertion frequencies are observed, in decreasing order, in Bellingham, Maipú and Davis. In the original population of Bordils, the highest insertion frequency corresponds to one site observed 16 times.

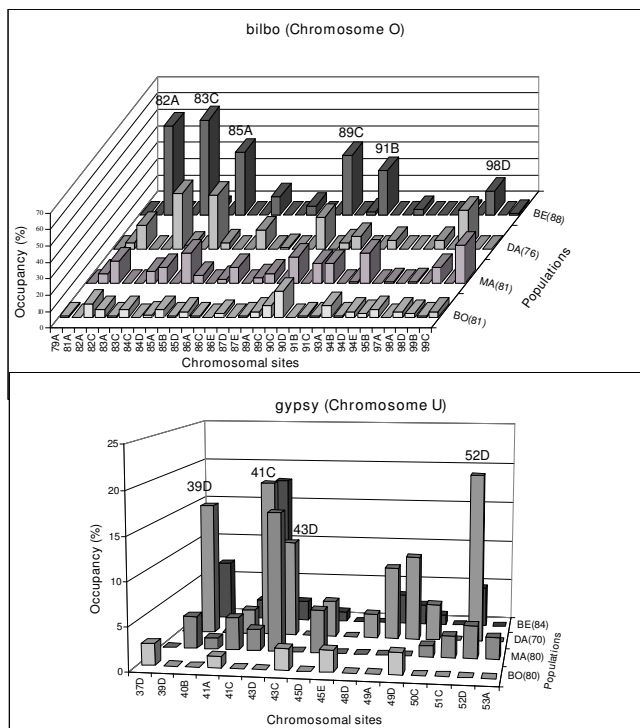
**Table 1: Occupancy profiles of euchromatic sites in original and colonizing populations**

TE	Populations	Occupancy profiles																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17-51	
<b>bilbo</b>	<b>Colonizing</b>	DA	7	6	9	4	1	7	0	2	2	1	5	1	1	0	1	0	7
		BE	3	5	3	2	4	2	1	2	1	2	0	0	4	1	2	2	10
		MA	8	8	4	7	4	2	2	5	2	3	2	3	2	1	2	1	12
	<b>Original</b>	BO	32	19	21	11	6	2	5	1	2	0	1	0	1	0	3	1	1
<b>gypsy</b>	<b>Colonizing</b>	DA	8	4	4	0	1	2	2	2	0	0	1	0	1	1	0		
		BE	11	5	1	2	2	2	1	0	0	0	0	0	0	0	1		
		MA	9	10	6	3	2	1	1	0	0	0	0	0	2	0	0		
	<b>Original</b>	BO	10	12	5	2	1	0	0	0	0	0	0	0	0	0	0		

Population origin: Davis (DA), Bellingham (BE), Maipú (MA), Bordils (BO).  
Occupancy profile: number of times that each site is occupied in populations.

Table 2 lists the means and variances of copy number for *bilbo* and *gypsy* per chromosome and haploid genome. The mean copy number of both TEs for the whole genome (HG) is always higher in colonizer populations than in the original one. The *bilbo* mean copy number differs greatly among chromosomes ranging from 2.58 copies in

chromosome O from Bellingham to 0.55 in chromosome J from Davis. In fact chromosome J hosts the lowest number of *bilbo* in all populations. A different scenario is found for *gypsy* in which the A (X) chromosome contains the lowest number of insertions, in colonizer and original populations alike. However, among the autosomes J is the least occupied in all populations. Deviation from Poisson distribution was tested by chi-square goodness of fit tests (for details, see additional files 1 and 2) pooling adjacent classes with low expected numbers. In colonizing populations *bilbo* distribution in each chromosome fits a Poisson distribution. *Gypsy* deviates from a Poisson distribution in E chromosome from Davis and Bellingham and in U chromosome from Maipú. When the whole genome is considered both TEs follow a Poisson distribution in the original population and deviate in all colonizing ones, except for *bilbo* in Bellingham and *gypsy* in Maipú. For this element the general trend in colonizing populations is a lower than expected number of genomes with a single copy and an excess of genomes with three or more copies (see Table 1). An alternative test was performed using dispersion coefficients (DC), which measure the ratio between the variance ( $V_n$ ) and the mean ( $m$ ) ( $DC = V_n/m$ , see table 2). DC of 1 indicates that TE distribution is Poisson, and  $DC > 1$  or  $DC < 1$  indicates contagious or repulsive distributions, respectively. When the haploid genome is considered, there is a general tendency towards DCs  $> 1$  for both elements in all populations except for *gypsy* in Maipú (these results are due to the greater effect of some chromosomes in the final result of the test).



**Figure 1**  
**Distribution of *bilbo* and *gypsy* in chromosomes O and U, respectively, from colonizing (DA: Davis, BE: Bellingham, MA: Maipú) and original populations (BO: Bordils) of *D. subobscura*. Number of haploid genomes analyzed are given in parenthesis.**

Because in some cases TE sites seem to be distributed in a contagious way ( $DC > 1$ ), linkage disequilibrium was computed for each pair of sites by way of  $2 \times 2$  contingency tables [26]. Linkage disequilibrium between TE sites could be responsible of the non-random distribution detected in some cases. The observed distribution of correlation coefficients between all paired sites was com-

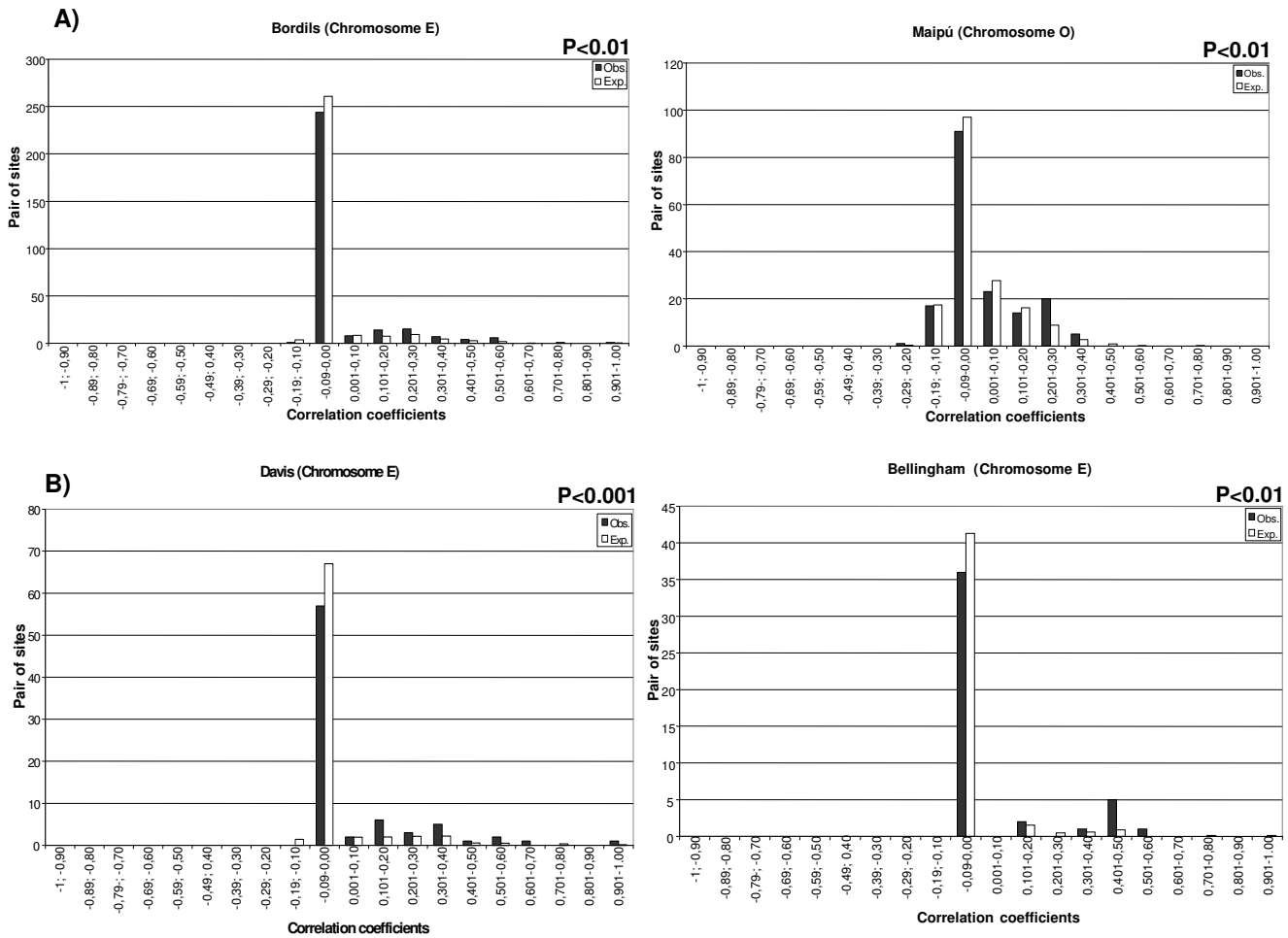
**Table 2: Tests of the Poisson distribution of bilbo and gypsy per chromosome and haploid genome.**

		Populations									
		DA (76)					BE (88)				
TE	Ch.	m	V <sub>n</sub>	DC	χ <sup>2</sup>	df	m	V <sub>n</sub>	DC	χ <sup>2</sup>	df
bilbo	A	0.84	0.72	0.86	0.61	2	1.07	0.77	0.73	2.68	2
	J	0.55	0.49	0.89	0.11	1	0.76	0.67	0.88	1.61	2
	U	1.01	0.87	0.86	5.74	3	1.11	1.00	0.90	1.46	3
	E	1.38	1.73	1.25	4.70	4	0.82	0.84	1.03	0.69	3
	O	1.71	1.67	0.98	3.47	5	2.58	1.80	0.70	7.06	4
	HG	5.50	9.40	1.71**	32.56**	11	6.34	7.79	1.23	18.61	11
		MA (81)					BO (81)				
		m	V <sub>n</sub>	DC	χ <sup>2</sup>	df	m	V <sub>n</sub>	DC	χ <sup>2</sup>	df
bilbo	A	1.75	1.69	0.96	1.54	5	1.06	0.71	0.67	8.11	2
	J	0.91	0.93	1.02	0.25	3	0.75	0.81	1.08	13.65*	4
	U	1.54	1.28	0.83	4.15	4	0.84	0.69	0.82	1.14	2
	E	1.80	1.66	0.92	1.70	4	1.09	1.70	1.57**	29.79**	5
	O	1.79	1.99	1.11	6.53	5	0.96	1.28	1.34	9.58	3
	HG	7.80	19.51	2.50**	254.34**	18	4.70	5.76	1.22	11.88	10
		DA (70)					BE (84)				
		m	V <sub>n</sub>	DC	χ <sup>2</sup>	df	m	V <sub>n</sub>	DC	χ <sup>2</sup>	df
gypsy	A	0	0	-	-	-	0.02	0.02	1.00	-	-
	J	0	0	-	-	-	0.02	0.02	1.00	-	-
	U	0.98	1.26	1.28	5.53	3	0.43	0.49	1.14	1.76	3
	E	0.58	0.97	1.66**	44.10**	4	0.33	0.56	1.69**	38.10	3
	O	0.04	0.04	0.97	-	-	0.09	0.11	1.17	1.38	1
	HG	1.61	2.44	1.51*	14.89	5	0.90	1.60	1.77**	33.18**	4
		MA (80)					BO (80)				
		m	V <sub>n</sub>	DC	χ <sup>2</sup>	df	m	V <sub>n</sub>	DC	χ <sup>2</sup>	df
gypsy	A	0.12	0.16	1.29	4.46	1	0.02	0.02	1.00	-	-
	J	0.14	0.12	0.87	-	-	0.10	0.12	1.16	1.26	1
	U	0.42	0.27	0.64	8.49*	1	0.11	0.10	0.90	-	-
	E	0.39	0.32	0.82	1.43	1	0.30	0.34	1.13	2.27	1
	O	0.27	0.28	1.01	0.12	1	0.24	0.23	0.99	0.00	1
	HG	1.35	1.27	0.94	11.48	4	0.77	0.88	1.14	3.44	2

TE: Transposable elements; Ch: chromosome; HG: haploid genome; m: mean copy number; V<sub>n</sub>: variance of copy number; Numbers of haploid genomes are in parenthesis; DC: dispersion coefficient (V<sub>n</sub>/m); \*P < 0.05; \*\*P < 0.01. Bonferroni's correction was applied. See Table 1 for population origin

pared to the expected distribution in absence of linkage disequilibrium using Fisher's hypergeometric formula [27]. Figure 2 depicts, as an example, the correlation coefficient distributions (pooled in intervals of 0.1) of bilbo in chromosome E from Bordils and in chromosome O from Maipú; and of gypsy in chromosome E from Davis and Bellingham. Tests were significant in most cases where a deviation from Poisson distribution was observed. More-

over we also found significant results in some cases where departures from Poisson distribution were not detected (e.g bilbo on chromosome O of Maipú). The general trend is a defect of class (-0.09-0.00) and an excess of some positive correlation classes. This indicates that some sites tend to stay together, as indicated by a DC > 1. This tendency was observed in all cases where deviations from Poisson distribution were observed, except for gypsy in



**Figure 2**  
**Observed and expected frequency distributions of correlation coefficients between all pairs of sites in natural populations: A) *bilbo* in chromosome E from Bordils and O from Maipú. B) *gypsy* in chromosome E from Davis and Bellingham.**

chromosome U from Maipú where there is an overabundance in class (-0.09–0.00) and the DC is lower than 1.

**Copy number comparisons among chromosomes**

Montgomery et al [28] proposed that selection against TE insertions would lead to a lower number of TE copies in chromosome X than in autosomes due to the stronger deleterious effect of recessive insertional mutations in the X chromosome of hemizygous males. In order to test this hypothesis we compared the copy number in the A (X) chromosome with that in autosomes. To estimate the expected number of insertions we multiply the relative proportion of chromatin of each chromosome by the number of total insertions in the population. The relative proportion of chromatin is that reported by Stumm-Zollinger and Goldschmidt [29] corrected by eliminating

the dot chromosome, not included in our analyses. If TEs are randomly distributed, we expect a TE copy number per chromosome proportional to the amount of chromatin.

Observed and expected proportions were compared by a G test [30] among all chromosomes ( $G_a$ ), between the A (X) chromosome and autosomes ( $G_b$ ), and among autosomes ( $G_c$ ), as indicated in Table 3.  $G_b$  values were significant for *gypsy* in all populations, and for *bilbo* only in Maipú and Bordils. Because some differences may be due to high insertion sites, additional analyses were done after eliminating these sites. After elimination the significance was maintained for *gypsy* in all populations except in Maipú, and removed for *bilbo*. In general *gypsy* shows a low copy number in the A (X) chromosome compared to autosomes. However, this is not the rule for *bilbo* where

**Table 3: Comparison of the proportion of *gypsy* and *bilbo* sites among chromosomes, autosomes and between chromosome A and autosomes**

TE		<i>gypsy</i>				<i>bilbo</i>			
Ch.	P. chromat	DA	BE	MA	BO	DA	BE	MA	BO
<b>A(X)</b>	0.16	0.00	0.03	0.09	0.03	0.15	0.17	0.22	0.23
<b>J</b>	0.20	0.01	0.03	0.10	0.13	0.10	0.12	0.12	0.16
<b>U</b>	0.19	0.61	0.48	0.31	0.14	0.18	0.17	0.20	0.18
<b>E</b>	0.20	0.36	0.37	0.29	0.39	0.25	0.13	0.23	0.23
<b>O</b>	0.25	0.03	0.10	0.20	0.31	0.31	0.40	0.23	0.20
<b>Df</b>									
<b>Ga</b>	4	197.23** (115.82**)	71.18** (61.31**)	22.26** (5.08)	21.40** (21.40**)	36.98** (0.08)	87.16** (16.98**)	41.82** (4.80)	15.67 (8.66)
<b>Gb</b>	1	40.48** (26.15**)	15.24** (17.20**)	4.63* (1.14)	11.06**	0.36 (4.8)	0.08 (3.47)	15.56** (2.9)	9.69** (0.52)
<b>Gc</b>	3	156.75** (89.66**)	55.93** (44.11**)	17.63** 3.94	10.34*	36.61** (3.47)	87.08** (13.50*)	26.26** (1.87)	5.97 (8.10*)
<b>Ga</b>									
Total	12	290.67** (182.21**)				165.96** (30.06**)			
Pooled	4	230.15* (110.30**)				102.01** (1.8)			
H	8	60.52** (71.90**)				63.95** (28.23**)			
<b>Gb</b>									
Total	3	60.36** (44.50**)				16.00** (11.20**)			
Pooled	1	44.97** (23.78**)				5.77* (1.09)			
H	2	15.38** (20.71**)				10.24* (10.10*)			
<b>Gc</b>									
Total	9	230.32** (137.71**)				149.95** (18.85)			
Pooled	3	185.18** (86.52**)				96.25** (0.72)			
H	6	45.13** (51.19**)				53.71** (18.12*)			

P. chromat: Proportion of chromatin. Ga: Comparison of the proportion of TEs among chromosomes. Gb: Comparison of the proportion of TEs between chromosome A (X) and autosomes. Gc: Comparison of the proportion of TEs among autosomes; H: Heterogeneity test between colonizing populations; Df: Degrees of freedom; Pooled: Only colonizing populations; \*P < 0.05; \*\*P < 0.001. Bonferroni's correction was applied. Test values excluding high insertion frequency sites are in parenthesis.

Maipú and Bordils show a high copy number in A (X). Interestingly, those populations that display *gypsy* copy number differences between A (X) and autosomes, show also significant differences among autosomes (G<sub>c</sub>), specially in colonizer populations where chromosomes E and O show a higher copy number than expected.

In general, copy number tend to be higher for *bilbo* in chromosome O and for *gypsy* in chromosome U in all colonizing populations, (except *bilbo* in Maipú), whereas in the original population the E chromosome hosts the highest proportion of *gypsy* and *bilbo*. In order to determine if chromosomal differences have the same tendency in colonizing populations, heterogeneity (H) tests were performed for comparisons among chromosomes, between A (X) and autosomes and among autosomes. Table 3 shows that all cases were heterogeneous for both TEs. However when Maipú is excluded from the analyses and high insertion frequency sites are eliminated, Bellingham and Davis become homogeneous for *bilbo* (data not shown).

**Correlation studies between high frequency sites and chromosomal arrangements**

All five pairs of acrocentric chromosomes of *D. subobscura* are polymorphic for inversions. Frequencies of chromosomal arrangements show clinal variation correlated with latitude in Palearctic populations [31,32] and clines that follow the same latitudinal gradient evolved in recent colonizing populations in both hemispheres of the Americas [17,18]. These parallel observations across continents provided a natural experiment that supports the adaptive role of the chromosomal inversion polymorphism.

Frequencies of chromosomal arrangements in the analyzed populations of this work are summarized in Table 4. Each arrangement is conventionally designed by the letter of the chromosome in which it occurs, followed by a combination of digits that identify the set of inversions included in it [16]. Arrangement frequencies are of the same order of magnitude as those previously reported, including the North-South latitudinal variation of most arrangements [21,31,32]. However it is interesting to note

**Table 4: Frequencies of chromosomal arrangements in natural populations**

Arrangements	Populations			
	DA	BE	MA	BO
<b>Ast</b>	0.57	0.67	0.55	0.53
<b>A1</b>	-	-	-	0.17
<b>A2</b>	0.43	0.33	0.45	0.30
<b>Jst</b>	0.41	0.45	0.26	0.40
<b>J1</b>	0.59	0.55	0.74	0.60
<b>Ust</b>	0.32	0.42	0.50	0.17
<b>U1</b>	-	-	-	0.03
<b>U1+2</b>	0.39	0.40	0.20	0.72
<b>U1+2+8</b>	0.29	0.18	0.30	0.07
<b>U1+2+3</b>	-	-	-	0.01
<b>Est</b>	0.54	0.74	0.59	0.53
<b>E8</b>	-	-	-	0.01
<b>E1+2</b>	0.07	0.05	-	0.30
<b>E1+2+9</b>	0.12	0.01	0.08	0.03
<b>E1+2+9+12</b>	0.05	0.15	0.23	0.12
<b>E1+2+9+3</b>	0.22	0.05	0.10	0.01
<b>Ost</b>	0.08	0.23	0.22	0.22
<b>O2</b>	-	-	-	0.01
<b>O5</b>	0.01	0.14	-	-
<b>O7</b>	-	-	0.01	0.04
<b>O3+4</b>	0.20	0.11	0.30	0.32
<b>O3+4+7</b>	0.17	0.01	0.21	0.04
<b>O3+4+2</b>	0.34	0.19	0.22	0.06
<b>O3+4+8</b>	0.20	0.32	0.04	0.30
<b>O3+4+23+2</b>	-	-	-	0.01

-: Arrangement absent

that Maipú presents a higher  $O_{St}$  frequency than expected according to its latitude.

Some authors consider recombination as the main factor determining the chromosomal distribution of TEs [33,34], but see [34]. The model of ectopic exchange, predicts a negative correlation between recombination rate and TE copy number if ectopic exchange is reduced in parallel with regular meiotic recombination rate [35,36]. Under this model, TEs are expected to be more abundant in regions of low recombination as inversions or inversion break-points. In these regions the probability of induction of deleterious rearrangements produced by unequal recombination between TEs, is low because most of the time inversions will be found in heterozygous state (recombination is suppressed inside). Experimental evidences [37-39] suggest that TEs are responsible of chromosomal inversions in natural populations of Diptera and are particularly abundant inside and near inversion break-points [6,7,40].

In order to know whether an association between high insertion sites and arrangements exist, we computed the product-moment correlation coefficient ( $r$ ) for high-frequency sites (Table 5). We observed two *bilbo* sites of particular interest (67A and 89C) that show the highest correlation coefficients. The 67A site is located inside the breakpoint of arrangement  $E_{12}$  and is significantly associated with  $E_{1+2+9+12}$  in Davis ( $r = 0.64$ ) and Maipú ( $r = 0.85$ ). The 89C site is located near the break-point of  $O_8$  arrangement and is significantly correlated with arrangement  $O_{3+4+8}$  in Davis ( $r = 0.58$ ) and Bellingham ( $r = 0.34$ ), and only marginally ( $r = 0.26$ ) in Maipú. Other instances of significant associations are not so easily explained because sites are external to inversion breakpoints. Thus, highly occupied 74D *bilbo* site is located outside of chromosomal inversions, yet, it is also significantly associated with  $E_{1+2+9+12}$  in Davis ( $r = 0.33$ ) and Maipú ( $r = 0.24$ ). This site is also highly occupied by *gypsy* but in this case associations are not significant. In other cases we observe associations of sites inside highly frequent inversions where the crossing-over is not reduced. This is the case of 11B *bilbo* site, for example, negatively associated to  $A_2$  arrangement in all populations except Bellingham but located inside it.

## Discussion

### **Bilbo and gypsy distributions are different in original and colonizer populations**

Results show a clear differential TE distribution in original and colonizing populations. While in the original population most sites have low insertion frequencies, colonizing populations present some highly occupied sites, with frequencies higher than 50% for *bilbo* and close to 20% for *gypsy*. Interestingly, most of them are common to all populations. Mean copy number of both elements is higher in colonizing populations than in the original one due to the presence of these highly occupied sites.

Low occupied sites would represent insertions occurred after colonization and/or copies from the original population whose frequency is decreasing in colonizing populations. An argument in favour of the former hypothesis is the existence of unique sites that would correspond to new transpositions (i.e. site 48D of *gypsy*), while the latter hypothesis explains the existence of low-occupancy original sites common to different populations (i.e. site 41A of *gypsy* or 85B of *bilbo*).

High insertion frequency sites are most likely due to a founder event during the colonization process (the founder hypothesis), as previously reported in other *Drosophila* species [9,11,13]. In *D. buzzatii* this hypothesis was also verified by molecular studies showing identical *Oswaldo* retrotransposon structures and flanking genomic

**Table 5: Correlation coefficients between chromosomal arrangements and high insertion frequency (HF) sites**

HF sites of bilbo	Arrang.	Populations							
		DA		BE		MA		BO	
		r	q-value	r	q-value	r	q-value	r	q-value
11B	A2	-0.13	(0.42)	0.35**	(10 <sup>-7</sup> )	0.28*	0.02	-0.07	(0.99)
20A	JI	0.24	(0.11)	-	-	-0.06	(0.39)	-0.04	(0.99)
43B	Ust	-0.08	(0.46)	0.04	0.44	0.23*	(0.04)	0.24	(0.99)
45C	Ust	0.11	(0.56)	0.20	(0.08)	0.40**	(3.10 <sup>-7</sup> )	0.05	(0.99)
45D	U1+2	-0.13	(0.59)	0.22*	(0.04)	-0.19	0.08	0.02	(0.99)
	Ust	0.15	(0.42)	0.49**	(3.10 <sup>-7</sup> )	--	--	--	--
53A	U1+2	-0.02	(0.86)	-0.34**	(2.10 <sup>-3</sup> )	--	--	--	--
	Ust	-0.20	(0.15)	-0.35**	(3.10 <sup>-7</sup> )	-0.22*	(0.04)	-0.05	(0.99)
57D	U1+2	0.26	(0.11)	0.05	0.44	0.18	(0.08)	0.07	(0.99)
	El+2+9	-	-	0.25	(0.12)	-0.10	(0.44)	0.57	(0.41)
59C	El+2+9+3	-0.06	(0.72)	0.17	(0.15)	0.28*	(0.04)	-0.02	(0.99)
	Est	0.36**	(5.10 <sup>-3</sup> )	-	-	-	-	-0.02	(0.80)
67A	Est	-0.15	(0.42)	-0.25*	(0.04)	-0.69**	(4.10 <sup>-8</sup> )	-0.03	(0.99)
	El+2+9+12	0.64**	(1.10 <sup>-4</sup> )	-	-	0.85**	(4.10 <sup>-8</sup> )	0.22	(0.51)
74D	El+2+9+12	0.33*	(0.04)	-	-	0.24*	(0.04)	0.13	(0.99)
82A	Ost	0.12	(0.63)	-0.10	(0.28)	0.29*	(0.02)	-0.16	(0.99)
83C	O3+4+7	0.00	(0.72)	0.09	(0.08)	-0.14	(0.17)	0.56	(0.11)
85A	O3+4	-0.14	(0.46)	0.27*	(0.02)	0.10	(0.17)	0.03	(0.99)
89C	O3+4+7	-0.17	(0.41)	0.08	0.53	-0.24*	(0.04)	-0.04	(0.99)
	O3+4	-0.16	0.42	-0.25*	(0.03)	-0.21	(0.06)	-0.09	(0.99)
	O3+4+2	-0.36**	(5.10 <sup>-3</sup> )	-0.31**	(7.10 <sup>-3</sup> )	0.09	(0.22)	0.07	(0.99)
91B	O3+4+8	0.58**	(1.10 <sup>-4</sup> )	0.34**	(4.10 <sup>-3</sup> )	0.26	(0.05)	0.02	(0.99)
	Ost	0.28	(0.16)	0.02	(0.12)	0.32*	(0.03)	-0.06	(0.99)
98D	O5	-0.03	(0.86)	0.28*	(0.02)	--	--	--	--
	O3+4+2	0.05	(0.72)	0.12	(0.18)	0.32*	(0.03)	0.29	(0.66)
<b>HF sites of gypsy</b>									
41C	U1+2	0.35*		0.52**		0.25*		--	
52D	U1+2	0.32*		0.28*		-0.12		--	
63C	Est	--		--		-0.46**		--	
	El+2+9+12	--		--		0.45**		--	
74D	El+2+9+3	0.31		-0.09		0.34		-0.04	

Only high insertion sites showing correlation coefficient values either significant or higher than 0.20 at least in one population, are considered. Arrang: arrangement; r: Correlation coefficient; -: indicates cases where correlations cannot be computed because of low ETs copy number; --: indicates the lack of a site or an inversion in the population). See Table 1 for population origin. \*P < 0.05; \*\*P < 0.01. Q-value and Bonferroni corrections were applied to *bilbo* and *gypsy* respectively

sequences in high insertion frequency sites from different colonizing populations [14].

In this study two lines of evidence support the founder hypothesis. First, the two studied TEs belong to different subclasses, yet they show a similar population behaviour. Second, most highly occupied sites are located in colonizing population chromosomes, although some exceptions occur for *bilbo* whose insertion frequency exceeds 10% in 9 sites in the original population of Bordils. All these sites correspond also to high insertion sites in colonizing populations, except 90C site and 21A, which are, respectively, free of insertions or occupied at low frequency in America.

The presence of high frequency sites in the original population could be a consequence of the transposition mechanism of *bilbo*, a LINE element. It has been shown that LINE elements (L1) make 5' truncated copies during their transposition mechanism indicating that 5' sequences are not absolutely necessary to insertion [41-43]. In fact, the majority of the L1 copies present in mammalian genomes are 5' truncated with a length of not more than 1 kb [1,44]. We can think that selection against truncated, "dead-on-arrival" (DOA) copies should be weak because they are not transcribed, potentially immobile and shorter than full copies. Thus, deleted copies could persist in some genomic regions without being completely elimi-



nated by natural selection. In fact, some *Drosophila* TE families (most of them LINE like elements) seem to be only marginally affected by purifying selection, reaching high insertion frequencies in euchromatin [45].

On the other hand, some of *bilbo* high frequency sites from Bordils could be explained by the dragging effect from the rich inversion polymorphism of *D. subobscura*. For example the 67A site located in the break-point of  $E_{1,2}$  arrangement presents highly significant correlations with this arrangement in 2 out of 3 colonizing populations. In Bordils, this correlation is not significant because of the lower frequency of this arrangement in this population. Arrangements of chromosome E cover approximately 75% of its length and it is not rare to find this kind of associations. In this chromosome another high insertion frequency site (74D) shows association with the same  $E_{1+2+9+12}$  chromosomal arrangement. This site corresponds to a heterochromatic telomeric site where it is not rare to find an accumulation of TE insertions. In fact *gypsy* is inserted also in this chromosomal site at occupation rates that range from 1.3 to 11.4%. Accumulation of TEs in heterochromatin is well documented in *D. melanogaster* where a significant excess of insertions were reported in heterochromatin, dot and Y chromosomes alike [46-49].

Seasonal fluctuations in population frequency of chromosomal rearrangements can modify recombination rates and associations between arrangements and genes. In *D. subobscura* no seasonal fluctuations were reported in some works [50,51], but fluctuations and seasonal changes of associations between chromosomal inversions and allozymes were reported in others, specially in the O chromosome from original populations [52,53]. In the present case, we observe no associations between insertions and specific chromosomal arrangements in the original population, but we do detect this kind of associations in colonizing populations (where fluctuations were not studied). However, changes in associations between chromosomal arrangements and chromosomal sites do not follow a definite trend. As an example, the  $U_{ST}$  arrangement, whose frequency has increased in all colonizing populations, shows a positive association to 43B and 45C sites but a negative one to 53A site in Maipu. This is a rather odd outcome since increase of rearrangement frequency is always expected to break down associations due to an increase of recombination rate. So, the likely explanation would be that fluctuations do not affect associations or at least not in the same way for every studied rearrangement polymorphism.

On the other hand, we favor the general idea that the positive correlation between arrangements and TE copies is not due to an inversion effect but, most probably, to the

founder event [19,25,54,55]. This could explain why arrangement  $E_{1+2+9+12}$  and the 74D site, which is located outside of the inversion, show a positive association and also why an excess of classes including positive correlation coefficients between chromosomal sites was observed in some chromosomes like E. Genetic estimates suggest that the number of founders ranged from 10 to 150 [25,56]. If some founders carried together this site and this arrangement, both will appear together in all populations because they are identical by descent. The founder hypothesis is favored by the fact that all correlations between sites and arrangements are significant only in colonizing populations. In the original population in spite of having correlation coefficients of 0.57 (in 57D) and 0.56 (in 83C) with  $E_{1+2+9}$  and  $O_{3+4+7}$  respectively, these are not significant. In fact, these two arrangements are currently decreasing in frequency in the Mediterranean populations and perhaps these combinations descend also from a few individuals. All these considerations suggest that most of the associations detected are due to a founder effect.

The general rule, as reported in *D. melanogaster* [45,57], is that TEs are spread and have low insertion frequencies in euchromatin. In some cases, however, accumulations of TEs in some chromosomal sites have been reported, as in the 42B [58], 87C [59] and 38 [60] regions, of *D. melanogaster*, and the 85D region of *D. subobscura* [61,62], and even fixation has occurred, as in the 42C site in natural populations of *D. simulans* [63]. Preferential insertion sites (hotspots) have been suggested for some *Drosophila* elements [64-66] and we cannot completely discard the possibility of an activation of transposition to specific hotspots during the colonization process. This hypothesis could be verified if a process affecting equally the two TEs studied occurred, as shown in *D. melanogaster*. In this species some proteins are involved in RNA-silencing mechanisms for retrotransposable elements repression [67-69]. We cannot discard the existence of a similar mechanism in *D. subobscura* that was de-repressed as a consequence of the colonization process contributing simultaneously to an increase of transposition of different transposable elements.

#### **Factors affecting TEs distribution in *D. subobscura***

In *Drosophila*, TEs seem to be maintained in populations as the result of a balance between transposition and opposing forces that reduce their copy number. In this way selection can act either directly against deleterious insertions or indirectly against deleterious chromosomal rearrangements produced by ectopic recombination between TEs [4,5,36,70]. In this work a test of selection against deleterious insertions was done by comparing copy numbers between X and autosomes, selection being more effective in the former than in the latter.

For *gypsy* we observe a clear tendency to follow a selection model, except in Maipú. This result is in concordance with that observed in a natural population of *D. melanogaster* with this element [71]. For *bilbo* the data do not fit a selection model against deleterious insertions; even in those cases where the test is significant, a higher copy number on A (X), compared to autosomes, is observed. A possible explanation of this result is that *bilbo* could have a differential transposition rate between X and autosomes. Some examples of transposition restricted to female or male *D. melanogaster* germ line have been reported [72,73] and they should be taken into account when X and autosomes are compared. On the other hand, the discrepancies observed between the two elements may be accounted for by the different factors that control copy numbers in each of these elements. In *D. melanogaster gypsy* is a retrovirus [74] submitted probably to a strong selection effect, its transposition depending on the presence of permissive alleles most likely segregating in natural populations. In *D. subobscura* this retrotransposon seems to be non infectious because current available copies have an apparently inactive *env* region [75], but this does not discard the putative presence of alleles that control its transposition. On the other hand *bilbo* is a LINE element and could be submitted to a soft selection pressure due to its DOA transposition mechanism. Most of the copies are probably deleted and its deleterious capability by transposition is diminished. The model of selection against deleterious insertions has been questioned by some authors [28,48] because neither all ETs nor all populations had a lower insertion frequency on X chromosomes compared to autosomes. However in a later work [76], where the authors reanalyze the data including more results from other species, selection against insertions is considered as the major mechanism of TE copy number control. On the other hand, values of selection coefficients against deleterious mutations could not be comparable to mutations associated to TE insertions. Moreover, deleterious effects of TEs can be species specific and populations may also sometimes suffer TE mobilizations that mask selection effects on TE distribution.

In this work each element presents a different behavior probably due to their distinct transposition mechanisms. Moreover we should not forget that elements which are stable in some genome conditions could be unstable in others. Recently mobilized TEs and/or colonization events, in populations, could lead to a differential copy distribution between chromosomes, rendering the selection undetected. This could be the case of Maipú, a new colonizing Argentinian population, which shows a distribution pattern for *gypsy* and *bilbo* quite different from the other colonizing populations. In particular, some high insertion frequency sites are more represented, or even exclusive, in this population. It is possible that Maipú was

established through a bottleneck of founder flies from Chile as a consequence of a secondary colonization. In this case, we cannot discard the existence of new transposition events in founders induced by the new environmental conditions encountered as previously proposed by other authors [10,11]. If this colonization occurred recently, as indicated by collecting records, selection has not had enough time to act, explaining the discrepancies in this population when comparing A (X) and autosome copy numbers in Table 3 or when this population is included in heterogeneity tests. In addition if TEs are not at equilibrium, departures from random distribution across chromosomes could reflect the insertion pattern rather than the effect of natural selection.

Another model proposed to explain the TE dynamics is the selection against deleterious arrangements produced by ectopic recombination between TEs. In *D. subobscura* accurate measures of recombination rate are not available and it is not possible to calculate a correlation between TE copies and recombination rates. This species has a rich inversion polymorphism in all chromosomes and recombination is reduced in heterokaryotypes. Under this model we expect accumulation of TEs in inverted segments, and in inversion break points or near them. In some cases arrangements include overlapping inverted fragments, often reaching frequencies higher than the standard arrangements, but in other cases, of low frequency arrangements, TE copy number is too low to allow statistical tests. Also recombination between non-overlapping inversions or inversion complexes may also be prevented [77].

We looked for accumulations of *bilbo* and *gypsy* in break-points of inversions but only one high insertion frequency *bilbo* site, 67A, coincides with an inversion breakpoint ( $E_{1+2+9+12}$ ). In another case the 89C high frequency site of *bilbo* is located near the inversion  $O_8$  and shows a significant correlation with  $O_{3+4+8}$  arrangement. This is in concordance with several unsuccessful attempts to localize *in situ* hybridization middle repeated sequences in *D. subobscura* inversions breakpoints [61,78]. These data notwithstanding, we cannot discard that other elements may be responsible of chromosomal inversion induction as reported in other *Drosophila* species [37,38].

## Conclusion

We conclude that the differential distribution of *bilbo* and *gypsy* between original and colonizing *D. subobscura* populations, is mainly due to a founder effect occurred during the colonization process of this species. We have shown that both founder effect and inversion polymorphism contribute notably to an excess of positive correlations between site pairs. Moreover the two transposable elements show a different pattern of distribution in popula-

tions that might be due to their differences in transposition and copy number regulatory mechanisms. This paper is also an attempt to emphasize the importance of population structure and history to explain the TE chromosomal distribution. We highlight the fact that comparisons in TE copy number between X and autosomes have to be interpreted cautiously. Sometimes TE mobilizations can mask the effect of selection on TE distribution.

## Methods

### *Drosophila* strains

The control strain *chcu* carries the recessive markers cherry eyes and curled wings and is homokaryotypic for chromosomal arrangements  $A_{st}$ ,  $J_{st}$ ,  $U_{st}$ ,  $E_{st}$  and  $O_{3+4}$ . It is kept by mass-culturing to maintain its viability. *In situ* hybridization for insertions of *bilbo* and *gypsy* displayed high stability over generations in 19C, 46A, 46C, 73A, 81D, 84A, 96A for *bilbo* and in 7C and 52A for *gypsy*.

The original population was sampled in Spring 2005 in Bordils (42.30°N, Girona, Spain). The colonizing populations were sampled in Spring 2004 in Davis (38.33°N, California, USA) and Bellingham (48.45°N, Washington, USA), and in Spring 2005 in Maipú (36.52°S, Argentina).

### Mating system (prior to "in situ" hybridization)

Individual males of natural populations were crossed with virgin females of the control line *chcu*. Insertion profiles were analyzed in F<sub>1</sub> female larval progeny to include the X chromosome. The TE insertion profile of each male was deduced by subtracting the TE insertion profile of the control line from that of the F<sub>1</sub> larva.

### In situ hybridization and DNA probes

Polytene chromosome [16] squashes from salivary glands of third-instar larvae, prepared as described in [79], were hybridized with digoxigenin labelled probes of *bilbo* and *gypsy*. The probes consisted of PCR fragments (2.6 and 2.8 kb long) which included the reverse transcriptase region. Prehybridization solutions and posthybridization washes were done following a protocol by Roche [80]. PCR reactions were carried out in a final volume of 25 µl, including 1× activity buffer (Ecogen), 1.6 mM MgCl<sub>2</sub>, 0.2 mM of each dNTP (Roche), 0.4 µM primer (Roche), 10–20 ng of genomic template DNA, and 0.04 units per µl of Taq polymerase (Ecotaq from Ecogen). Amplifications were run in a MJ Research Inc. thermocycler programmed as follows: 5 min preliminary denaturation at 94°, 30 cycles of 45 s at 94° (denaturation), 45 s at specific PCR annealing temperatures, 1.5 min at 72° (extension) and a final extension for 10 min at 72°. PCR products were gel purified with a GeneClean kit (BIO 101) and labelled using the random primer method. After hybridization signal devel-

opment was done using an anti-digoxigenin antibody conjugated with alkaline phosphatase (Roche).

*In situ* hybridization is the more suitable method used in localization of TEs on chromosomal arms. However, the power of resolution of this technique allow us neither discriminate between closely neighbouring sites, nor between elements that diverge below 10%.

### Statistical analyses

Statistical analyses were performed excluding centromeric and pericentromeric TE insertions. The statistical software SPSS version 14.0 was used for most of the statistical data analyses.

In cases of multiple testing, corrections were achieved measuring the significance of False Discovery Rates [81] through q values. To get the q-value we used the software QVALUE [82] on the p values obtained from the multiple test. When this test could not be applied, Bonferroni's correction was performed [83].

### Authors' contributions

MPGG participated in the design, the chromosomal slides, some statistical analyses and the writing of the manuscript. BECS collected the Bordils population, performed most of the technical work, the reading of slides and some statistical analyses. JB collected Davis and Bellingham populations, supervised all arrangement readings and performed the data set analyses. LS collected the Bordils population, contributed to the design and thoroughly revised the manuscript. AF participated in the design, directed the project, coordinated the data analyses, contributed to the writing of the manuscript and collected the Maipú and Bordils populations. All authors read and approved the final manuscript

### Additional material

#### Additional file 1

*Poisson distribution: raw data of bilbo copy number per chromosome and population, P values and chi tests. A table of detailed tests of Poisson distribution of bilbo per chromosome and haploid genome.*

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2148-8-234-S1.xls>]

#### Additional file 2

*Poisson distribution: raw data of gypsy copy number per chromosome and population, P values and chi tests. A table of detailed tests of Poisson distribution of gypsy per chromosome and haploid genome.*

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2148-8-234-S2.xls>]

## Acknowledgements

This work was funded by projects BOS2003-05904-C02-01 and CGL2006-13423-C02-01/02 from the Ministerio de Educación y Ciencia, Spain, and project 2005SGR 00995 from Generalitat de Catalunya, Spain to the Grup de Biologia Evolutiva (Principal Investigator: AF); and a doctoral fellowship from UAB to BECS. We acknowledge technical help by Montse Peiró and field assistance by Dr. E. Hasson, Dr. P. Fernández-Iriarte, Dr. F. García-Franco and D. Vela.

## References

- IHGSC: **Initial sequencing and analysis of the human genome.** *Nature* 2001, **409**:860-921.
- Deininger PL, Roy-Engel AM: **Mobile elements in animals and plant genomes.** In *Mobile DNA II* Edited by: Craig NL, Craigie R, Gellert M, Lambowitz A. Washington D.C. , American Society for Microbiology; 2002:1074-1092.
- Le Rouzic A, Capy P: **Population genetics models of competition between transposable elements subfamilies.** *Genetics* 2006, **174**:785-793.
- Charlesworth B, Charlesworth D: **The population dynamics of transposable elements.** *Genet Res* 1983, **42**:1-27.
- Charlesworth B, Langley CH: **The population genetics of Drosophila transposable elements.** *Annu Rev Genet* 1989, **23**:251-287.
- Charlesworth B, Lapid A, Canada D: **The distribution of transposable elements within and between chromosomes in a population of Drosophila melanogaster. I. Element frequencies and distribution.** *Genet Res* 1992, **60**:103-114.
- Charlesworth B, Lapid A, Canada D: **The distribution of transposable elements within and between chromosomes in a population of Drosophila melanogaster. II. Inferences on the nature of selection against elements.** *Genet Res* 1992, **60**:115-130.
- Decelière G, Charles S, Biémont C: **The dynamics of transposable elements in structured populations.** *Genetics* 2005, **169**:467-474.
- García Guerreiro MP, Fontdevila A: **Chromosomal distribution of the transposable elements Osvaldo and blanco in original and colonizer populations of Drosophila buzzatii.** *Genet Res* 2001, **77**:227-238.
- Vieira C, Nardon C, Arpin C, Lepetit D, Biémont C: **Evolution of the genome size in Drosophila. Is the invader's genome being invaded by transposable elements?** *Mol Biol Evol* 2002, **19**:1154-1161.
- Wisotzkey RG, Felger I, Hunt JA: **Biogeographic analysis of the Uhu and LOA elements in the Hawaiian Drosophila.** *Chromosoma* 1997, **106**:465-477.
- Fontdevila A: **Colonizing Species of Drosophila.** In *Molecular Techniques in Taxonomy Volume 57.* Issue H Edited by: Hewitt GM, Johnston AWB, Young JPW. Heidelberg , Springer-Verlag ; 1991:249-269.
- Labrador M, Seleme MDC, Fontdevila A: **The evolutionary history of Drosophila buzzatii. XXXIV. The distribution of the retrotransposon Osvaldo in original and colonizing populations.** *Mol Biol Evol* 1998, **15**:1532-1547.
- García Guerreiro MP, Fontdevila A: **The evolutionary history of Drosophila buzzatii. XXXVI. Molecular structural analysis of Osvaldo retrotransposon insertions in colonizing populations unveils drift effects in founder events.** *Genetics* 2007, **175**:301-310.
- Sturtevant AH: **The classification of the genus Drosophila with the description of nine new species.** *Univ Texas Publ* 1942, **42**:13:5.
- Krimbas CB: **Drosophila subobscura.** Hamburg , Verlag Dr. Kovac; 1993.
- Ayala FJ, Serra L, Prevosti A: **A grand experiment in evolution: the Drosophila subobscura colonization of the Americas.** *Genome* 1989, **31**:246-255.
- Prevosti A, Ribó G, Serra L, Aguadé M, Balañá J, Monclús M, Mestres F: **Colonization of America by Drosophila subobscura: experiment in natural populations that supports the adaptive role of chromosomal inversion polymorphism.** *Proc Natl Acad Sci USA* 1988, **85**:5597-5600.
- Brcnic D, Prevosti A, Budnik M, Monclús M, Ocaña J: **Colonization of Drosophila subobscura in Chile I. First population and cytogenetic studies.** *Genetica* 1981, **56**:3-9.
- Beckenbach AT, Prevosti A, Serra L: **Colonization of North America by the European species Drosophila subobscura and D. ambigua.** *Am Midl Nat* 1986, **115**:10-18.
- Menozzi P, Krimbas CB: **The inversion polymorphism of D. subobscura revisited: Synthetic maps of gene arrangement frequencies and their interpretation.** *J Evol Biol* 1992, **5**:625-641.
- Mestres F, Balañá J, Segarra C, Prevosti A, Serra L: **Colonization of America by Drosophila subobscura: Analysis of the O5 inversions from Europe and America and their implications for the colonizing process.** *Evolution* 1992, **46**:1564-1568.
- Latorre A, Moya A, Ayala FJ: **Evolution of mitochondrial DNA in Drosophila subobscura.** *Proc Natl Acad Sci U S A* 1986, **83**:8649-8653.
- Rozas J, Hernández M, Cabrera VM, Prevosti A: **Colonization of America by Drosophila subobscura: Effect of the founder event on the mitochondrial DNA polymorphism.** *Mol Biol Evol* 1990, **7**:103-109.
- Pascual M, Aquadro CF, Soto V, Serra L: **Microsatellite variation in colonizing and Palearctic populations of Drosophila subobscura.** *Mol Biol Evol* 2001, **18**:731-740.
- Charlesworth B, Lapid A: **A study of ten families of transposable elements on X chromosome from a population of Drosophila melanogaster.** *Genet Res* 1989, **54**:113-125.
- Fisher RA: **Statistical methods for research workers.** 13rd edition. New York , Hafner; 1958.
- Montgomery EA, Charlesworth B, Langley CH: **A test for the role of a natural selection in the stabilization of transposable element copy number in a population of Drosophila melanogaster.** *Genet Res* 1987, **49**:31-41.
- Stumm-Zollinger E, Goldschmidt E: **Geographical differentiation of inversion systems in Drosophila subobscura.** *Evolution* 1959, **13**:89-98.
- Sokal RR, Rohlf FJ: **Biometry.** 3rd edition. New York , Freeman; 1995.
- Balanyà J, Serra L, Gilchrist GW, Huey RB, Pascual M, Mestres F, Solé E: **Evolutionary pace of chromosomal polymorphism in colonizing populations of Drosophila subobscura: An evolutionary time series.** *Evolution* 2003, **57**:1837-1845.
- Balanyà J, Solé E, Oller JM, Sperlich D, Serra L: **Long-term changes in the chromosomal inversion polymorphism of Drosophila subobscura. II. European populations.** *J Zool Syst Evol Res* 2004, **42**:191-201.
- Bartolomé C, Maside X, Charlesworth B: **On the abundance and distribution of transposable elements in the genome of Drosophila melanogaster.** *Mol Biol Evol* 2002, **19**:926-937.
- Rizzon C, Marais G, Gouy M, Biémont C: **Recombination rate and the distribution of transposable elements in the Drosophila melanogaster genome.** *Genome Res* 2002, **12**:400-407.
- Goldman ASH, Lichten M: **The efficiency of meiotic recombination between dispersed sequences in Saccharomyces cerevisiae depends upon their chromosomal location.** *Genetics* 1996, **144**:43-55.
- Langley CH, Montgomery EA, Hudson R, Kaplan N, Charlesworth B: **On the role of unequal exchange in the containment of transposable element copy number.** *Genet Res* 1988, **52**:223-235.
- Cáceres M, Ranz JM, Barbadilla A, Long M, Ruiz A: **Generation of a Widespread Drosophila inversion by a Transposable Element.** *Science* 1999, **285**:415-418.
- Lyttle TW, Haymer DS: **The role of the transposable element hobo in the origin of the endemic inversions in wild populations of Drosophila melanogaster.** *Genetical Research* 1992, **86**:113-126.
- Mathiopoulos KD, Della Torre A, Predazzi V, Petrarca V, Coluzzi M: **Cloning of inversion breakpoints in the Anopheles gambiae complex traces a transposable element at the inversion junction.** *Proc Natl Acad Sci USA* 1998, **95**:12444-12449.
- Sniegowsky PD, Charlesworth B: **Transposable element numbers in cosmopolitan inversions from a natural population of Drosophila melanogaster.** *Genetics* 1994, **137**:815-827.
- Belancio VP, Whelton M, Deininger P: **Requirements for polyadenylation at the 3' end of LINE-1 elements.** *Gene* 2007, **390**:98-107.

42. Busseau I, Pélisson A, Bucheton A: **Characterization of 5' truncated transposed copies of the I factor in *Drosophila melanogaster*.** *Nucleic Acids Research* 1989, **17**:6939-6945.
43. Fanning TG, Singer MF: **LINE I: A mammalian transposable element.** *Biochem Biophys Acta* 1987, **910**:203-212.
44. Medstrand P, van de Lagemaat L N, Mager DL: **Retroelement distributions in the human genome: Variations associated with age and proximity to genes.** *Genome Res* 2002, **12**:1483-1495.
45. Petrov D, Aminetzach YT, Davis JC, Bensasson D, Hirsh AE: **Size matters: Non-LTR retrotransposable elements and ectopic recombination in *Drosophila*.** *Mol Biol Evol* 2003, **20**:880-892.
46. Bartolomé C, Maside X: **The lack of recombination drives the fixation of transposable elements on the fourth chromosome of *Drosophila melanogaster*.** *Genetical Research* 2004, **83**:91-100.
47. Carmena M, González C: **Transposable elements map in a conserved pattern of distribution extending from beta-heterochromatin to centromeres in *Drosophila melanogaster*.** *Chromosoma* 1995, **103**:676-684.
48. Charlesworth B, Sniegowsky PD, Stephan W: **The evolutionary dynamics of repetitive DNA in eukaryotes.** *Nature* 1994, **371**:215-220.
49. Dimitri P, Junakovic N, Arca B: **Colonization of heterochromatic genes by transposable elements in *Drosophila*.** *Mol Biol Evol* 2003, **20**:503-512.
50. Krimbas CB: **The genetics of *Drosophila subobscura* populations. III. Inversion polymorphism and climatic factors.** *Mol Gen Genet* 1967, **99**:133-150.
51. Prevosti A: **Chromosomal polymorphism in *Drosophila subobscura* populations from Barcelona (Spain).** *Genet Res* 1964, **5**:27-38.
52. Fontdevila A, Zapata C, Alvarez C, Sánchez L, Méndez J, Enriquez I: **Genetic coadaptation in the chromosomal polymorphism of *Drosophila subobscura*. I. Seasonal changes of gametic disequilibrium in a natural population.** *Genetics* 1983, **105**:935-955.
53. Rodríguez-Trelles F: **Seasonal cycles of allozyme-by-chromosomal-inversion genetic disequilibrium in *Drosophila subobscura*.** *Evolution* 2003, **57**(4):839-848.
54. Balanyà J, Segarra C, Prevosti A, Serra L: **Colonization of America by *Drosophila subobscura*: The founder event and a rapid expansion.** *J Hered* 1994, **85**:427-432.
55. Prevosti A, Serra L, Ribó G, Aguadé M, Segarra E, Balanyà J, Monclús M, García MP: **The colonization of *Drosophila subobscura* in Chile. II. Clines in the chromosomal arrangements.** *Evolution* 1985, **39**:838-844.
56. Mestres F, Pegueroles G, Prevosti A, Serra L: **Colonization of America by *Drosophila subobscura*: lethal genes and the problem of the O5 inversion.** *Evolution* 1990, **44**:1823-1836.
57. Biémont C: **Population genetics of transposable DNA elements. A *Drosophila* point of view.** *Genetica* 1992, **86**:67-84.
58. Biémont C, Gautier C: **Interactions between transposable elements for insertion in the *Drosophila melanogaster* genome.** *Heredity* 1989, **63**:125-133.
59. Maside X, Bartolomé C, Charlesworth B: **S-element insertions are associated with the evolution of the Hsp70 genes in *Drosophila melanogaster*.** *Current Biol* 2002, **12**:1686-1691.
60. Bergman CM, Quesneville H, Anxolabéhère D, Ashburner M: **Recurrent insertion and duplication generate networks of transposable element sequences in the *Drosophila melanogaster* genome.** *Genome Biol* 2006, **7**(11):.
61. Felger I, Sperlich D: **Cytological localization and organization of dispersed middle repetitive DNA sequences of *Drosophila subobscura*.** *Chromosoma* 1989, **98**:342-350.
62. Vivas MV, García-Planells J, Ruiz C, Marfany G, Paricio N, González-Duarte R, de Frutos R: **GEM, a cluster of repetitive sequences in the *Drosophila subobscura* genome.** *Genet Res* 1999, **229**:47-57.
63. Vieira C, Biémont C: **Geographical variation in insertion site number of retrotransposon 412 in *Drosophila simulans*.** *J Mol Evol* 1996, **42**:443-451.
64. Belyaeva ES, Ananiev EV, Gvozdev VA: **Distribution of mobile dispersed genes (mdg-1 and mdg-3) in the chromosomes of *Drosophila melanogaster*.** *Chromosoma* 1984, **90**:16-19.
65. Gvozdev VA, Belyaeva ES, Ilyin IV, Amosova IS, Kaidanov LZ: **Selection and transposition of mobile dispersed genes in *Drosophila melanogaster*.** *Cold Spring Harbor Symp Quant Biol* 1980, **45**:673-685.
66. Pasyukova E, Belyaeva ES, Kogan GL, Kaidanov LZ, Gvozdev VA: **Concerted transpositions of mobile genetic elements coupled with fitness changes in *Drosophila melanogaster*.** *Mol Biol Evol* 1986, **3**:299-312.
67. Mével-Ninio M, Pélisson , Kinder J, Campos AR, Bucheton A: **The flamenco locus controls the gypsy and ZAM retroviruses and is required for *Drosophila* oogenesis.** *Genetics* 2007, **175**:1615-1624.
68. Pélisson A, Sarot E, Payen-Groschêne G, Bucheton A: **A novel repeat-associated small interfering RNA-mediated silencing pathway downregulates complementary sense gypsy transcripts in somatic cells of the *Drosophila* ovary.** *Journal of Virology* 2007, **81**(4):1951-1960.
69. Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD: **A distinct small RNA pathway silences selfish genetic elements in the germline.** *Science* 2006, **313**:320-324.
70. Kaplan NL, Brookfield JFY: **Transposable elements in mendelian populations. III. Statistical results.** *Genetics* 1983, **104**:485-95.
71. Biémont C, Lemeunier F, García Guerreiro MP, Brookfield JF, Gautier C, Aulard S, Pasyukova EG: **Population dynamics of the copia, mdg1, mdg3, gypsy and P transposable elements in a natural population of *Drosophila melanogaster*.** *Genet Res* 1994, **63**:197-212.
72. Bucheton A, Busseau I, Teninges A: **I elements in *Drosophila melanogaster*.** In *Mobile DNA II* Edited by: Craig NL, Craigie R, Gellert M, Lambowitz AM. Washington DC , ASM press; 2002:796-813.
73. Pasyukova E, Nuzhdin WL, Flavell AJ: **Germ line transposition of the copia retrotransposon in *Drosophila melanogaster* is restricted to males by tissue-specific control of copia RNA levels.** *Mol Gen Genet* 1997, **255**:115-124.
74. Kim A, Terzian C, Santamaría P, Pélisson A, Prud'homme N, Bucheton A: **Retroviruses in invertebrates: The gypsy retrotransposon is apparently an infectious retrovirus of *Drosophila melanogaster*.** *Proc Nat Acad Sci USA* 1994, **91**:1285-1289.
75. Alberola TM, de Frutos R: **Molecular structure of a gypsy element of *Drosophila subobscura* (gypsy Ds) constituting a degenerate form of insect retroviruses.** *Nucleic Acids Res* 1996, **24**:914-923.
76. Biémont C, Vieira C, Hoogland C, Cizeron G, Loevenbruck C, Arnault C, Carante JP: **Maintenance of transposable element copy number in natural populations of *Drosophila melanogaster* and *D. simulans*.** *Genetica* 1997, **100**:161-166.
77. Mestres F, Sanz J, Serra L: **Chromosomal structure and recombination between inversions in *Drosophila subobscura*.** *Hereditas* 1998, **128**:105-113.
78. Cirera S, Martín-Campos JM, Segarra C, Aguadé M: **Molecular characterization of the breakpoints of an inversion fixed between *Drosophila melanogaster* and *D. subobscura*.** *Genetics* 1995, **139**:321-326.
79. Labrador M, Naveira H, Fontdevila A: **Genetic mapping of the Adh locus in the repleta group of *Drosophila* by in situ hybridization.** *J Hered* 1990, **81**:83-86.
80. Smidt ER: **A simplified and efficient protocol for non-radioactive in situ hybridization to polytene chromosomes with a DIG-labeled DNA probe.** In *Non-radioactive In situ hybridization Application Manual* Edited by: Grünwald-Janho S, Keeseey J, Leous M, van Miltenburg R, Schroeder C. Mannheim , Roche; 1996:97-107.
81. Benjamini Y, Hochberg Y: **Controlling the false discovery rate. A practical and powerful approach to multiple testing.** *J R Stat Soc Ser B Stat Meth* 1995, **57**:289-300.
82. Storey JD: **A direct approach to false discovery rates.** *J R Stat Soc Ser B Stat Meth* 2002, **64**:479-498.
83. Holm S: **A simple sequentially rejective multiple test procedure.** *Scand J Stat* 1979, **6**:65-70.