

ORIGINAL ARTICLE

The *mariner* transposable element in natural populations of *Drosophila simulans*

S Picot^{1,2}, GL Wallau³, ELS Loreto³, FO Heredia¹, A Hua-Van^{1,2} and P Capy^{1,2}

¹Laboratoire Evolution Génomes et Spéciation, UPR9034, CNRS, Gif-sur-Yvette Cedex, France; ²Université Paris-Sud, Orsay Cedex, France and ³Departamento de Biologia—CCNE, Universidade Federal de Santa Maria, Santa Maria, RS, Brazil

In cosmopolitan species, geographical variations in copy number and/or level of transposition activity have been observed for several transposable elements (TEs). Environment, history and population structure can contribute to such variation in ways that are difficult to tease apart. For the *mariner* element, previous studies of the geographic variation of its somatic activity in natural populations of *Drosophila simulans* have shown contradictory results (latitudinal clines of divergent orientations or no apparent structure). To try and resolve these inconsistencies, we gathered all available data on the *mariner* somatic activity of worldwide natural populations. This includes previously published results by different groups and also new data. The correlations between the level of activity and several

geoclimatic factors were tested. Although no general effect of temperature was found, a relationship with the invasion history was detected. It was also shown that recent invasive populations have a higher level of activity than the putative ancestral ones. Our results strongly suggest that variability of the *mariner* somatic activity among natural populations of *D. simulans* is mainly due to populational and historical factors probably related to the recent world colonization of this species. Indeed, this activity is correlated to the main route out of Africa (the Nile route) and the recent colonization of continents such as Australia and South America.

Heredity (2008) **101**, 53–59; doi:10.1038/hdy.2008.27; published online 7 May 2008

Keywords: transposable element; *mariner*; somatic activity; *Drosophila simulans*; invasive populations

Introduction

Relationships between transposable elements (TEs) and genomes can be affected by different factors, such as transposition rates, number of autonomous and non-autonomous copies, environmental influences, history and population structure. Most of these factors have been discussed from a theoretical point of view and very few data are available from natural populations. The only exceptions come from the *412* retroelement and *P. hobo* and *mariner* elements (Anxolabéhère *et al.*, 1985; Vieira and Biémont, 1996; Giraud and Capy, 1996; Russell and Woodruff, 1999; Bonnivard *et al.*, 2002). In several cases, geographic variations (for instance, copy number, transposition rate, and so on.) have been observed. Some of them are clearly related to the invasion history, that is to the diffusion of the element after a horizontal transfer or a reinvasion of the species (*P. hobo* and *I* elements), others, such as the *412* retroelement, show a pattern that could be linked to the colonization process, whereas for the *mariner* element contradictory results have been published.

This study will focus on the activity of the *mariner* TE in *Drosophila simulans*. The ancestral populations of this species are thought to originate in central and eastern

Africa and on the nearby islands in the Indian Ocean (Madagascar, Seychelles and Mascarene Islands; Lachaise *et al.*, 1988). The worldwide expansion of the species is relatively recent compared to that of its sibling species *Drosophila melanogaster* (Singh *et al.*, 1987). For instance, it invaded Japanese islands as recently as the last century (Watanabe and Kawanishi, 1976), and also Australia (Bock and Parsons, 1981) and North and South America, accompanying human colonization.

Mariner is a widespread class II element from the *Tc1/mariner* superfamily. Copies belonging to the *mauritiana* subfamily have been found in all natural populations of *D. simulans* tested (Giraud and Capy, 1996; Russell and Woodruff, 1999). Previous studies on a few natural populations of *D. simulans* revealed that *mariner* activity increases with its breeding temperature in the laboratory (Chakrani *et al.*, 1993). This suggests that temperature may affect *mariner* activity. Two analyses of its somatic activity carried out on a large sample of natural populations showed conflicting results. On one hand, Giraud and Capy (1996) found a positive correlation ($R^2 = 0.526$, P -value < 0.01) between the level of activity and latitude along a transect between Africa and Europe. As temperature generally decreases with latitude, this suggests that populations from warmer areas have a lesser activity than those from colder areas when tested in standard laboratory conditions. On the other hand, Russell and Woodruff (1999) showed a slight negative correlation ($R^2 = 0.089$, P -value $= 0.0527$) between latitude and level of activity, and detected a significant negative correlation for the Australian strains ($R^2 = 0.445$,

Correspondence: Professor P Capy, Laboratoire Evolution Génomes et Spéciation, 1 avenue de la terrasse, BP1, UPR9034, CNRS, Gif-sur-Yvette Cedex 91198, France.

E-mail: pierre.capy@legs.cnrs-gif.fr

Received 31 August 2007; revised 1 February 2008; accepted 17 March 2008; published online 7 May 2008

P -value = 0.0003). Both analyses suggest the existence of limited or local correlations between latitude and somatic activity of the element, but no global trend can be detected.

The main objective of this study is to pool all the available data to get a global view of the dynamic of the *mariner* element in *D simulans*. To this end, several new data from the South American and European continents have been added, along with geographic and climatic data. The main question addressed here is: does the level of *mariner* activity depend on temperature or on other geographical, populational and/or historical parameters?

Materials and methods

Data: PMM and geoclimatic variables

Our data combine all those from the previous analyses (Giraud and Capy, 1996; Russell and Woodruff, 1999), with the addition of 61 new natural populations. A total of 395 lines, including mass and isofemale strains, were therefore available to us. From the South American continent, 21 new data from Loreto and Wallau were added. All the populations from Loreto and Wallau were tested shortly after their introduction into the laboratory, that is less than 10 generations. Most other populations were tested less than 3 years after establishing laboratory cultures. However, it was shown (partly published in Giraud and Capy, 1996, see also Table 1 in Supplementary Data) that the percentage of mosaic males (PMM) of populations kept in laboratory conditions at a temperature between 20 and 25 °C is relatively stable. In the present set of populations or lines, this concerns only 49 out of 395 samples.

When available, the latitude and longitude were determined using Google Earth. The 24-h average temperature ('mean temperature'), the 24-h average temperature for the coldest ('minimum temperature') and the warmest ('maximum temperature') month were collected from the nearest meteorological station (World Meteorological Organization, WMO Clino or

www.worldweather.org or from the www.worldclimate.com site after checking the supplied data for errors). If a location was unknown, the central point or capital of the country was chosen arbitrarily. Latitude and longitude are given according to the decimal system. Latitude values for the South hemisphere and the West longitude are negative. The absolute latitude (that is the absolute value of latitude) was used for clinal analyses. In both works previously published by Giraud and Capy (1996) and by Russell and Woodruff (1999), the same protocol defined by Bryan *et al.* (1990) and Garza *et al.* (1991) was used to estimate the somatic activity of the *mariner* element from a strain or a population. This activity was estimated by the PMM obtained from a cross involving 5–10 females from a reference strain (w^{pch}) with 5–10 males from the tested population. The w^{pch} strain is stable and has a single inactive *mariner* copy called peach, inserted in the white gene on the X chromosome. This is responsible for w^{pch} mutation, leading to peach-coloured eyes. Although non-autonomous, the peach copy can be *trans*-mobilized from the transposase produced by paternally inherited copies. The peach excisions from the white gene generally restore the wild-type phenotype. If the excision occurs in somatic cells during development, it leads to mosaic flies (w^{pch} eyes with red spots). In all cases, no distinction between different categories of mosaic flies was made, that is from flies with a single small spot to flies with several spots of different sizes. Table 1 summarizes the basic statistics on the level of activity for the main population groups.

Statistical analyses

For the data coming from populations kept as isofemale lines, the mean value of the PMM calculated on all the lines was used. Therefore each population is characterized by a single value for its somatic activity. All the analyses were done with *R* (Ihaka and Gentleman, 1996). Non-parametric tests, such as Kendall's test for correlation and Wilcoxon's test (Mann–Whitney's test), were preferentially used for mean comparison. We also employed principal components analysis to infer the

Table 1 Somatic activity of the *mariner* TE (PMM)

Geographic clusters	PMM				
	N	Minimum	Maximum	Mean	Standard error
All data	172	0.0	100	37.1	1.9
Africa	45	0.0	80.0	23.3	3.5
Asia	15	0.0	100	33.9	8.2
Europe	31	0.1	85.0	41.6	3.9
North America	20	0.0	66.3	34.4	4.0
Oceania	32	16.8	90.5	49.8	3.8
South and Central America	33	10.7	83.3	41.5	4.0
Australia	30	16.8	90.5	51.5	3.9
Sub-Saharan Africa	38	0.0	66.1	18.4	3.8
Northeast Africa with Near East	7	0.1	29.0	15.4	3.6
Northwest Africa	6	36.7	80.0	55.8	8.1
Western Europe	19	74.8	85.0	46.2	4.5
South America	27	10.7	83.3	38.9	4.1
Central America	6	17.0	83.1	53.3	11.5
South America (Loreto, Wallau)	21	10.7	83.3	41.9	4.6

Abbreviations: N, number of population samples; PMM, percentage of mosaic males.

The geographic clusters are those used in the present analyses. This table shows the PMM values for each cluster (minimum, maximum, mean and standard error).

relationship between the PMM and the geoclimatic variables (princomp function, on the correlation matrix). The mantel test was performed with Euclidian distances with the mantel.randtest function from the ade4 package (Chessel *et al.*, 2004), with 999 permutations.

Results

Global analysis

The significant correlations (with Kendall's τ) between the level of activity and absolute latitude, longitude, temperature (mean temperature, minimum and maximum temperature) are given in Table 2 (see also Figure 1 in Supplementary Data for the scatterplots). A significant positive correlation coefficient is detected for absolute latitude ($\tau = 0.2$, P -value = $4.5e-5$). Next, an analysis by continent (Africa, Europe, South-, North- and Central-America, Australia, Asia) was carried out. For African populations, positive correlations were found between the PMM and the maximum temperature ($\tau = 0.37$, P -value = $6.0e-5$) or with the absolute latitude ($\tau = 0.43$, P -value = $3.9e-5$). The last correlation is in agreement with that found by Giraud and Capy (1996). For Europe, there are negative correlations between the level of activity and minimum temperature ($\tau = -0.42$, P -value = $3.5e-2$) or longitude ($\tau = -0.29$, P -value = $2.6e-2$), and positive correlations between the level of activity and absolute latitude ($\tau = 0.32$, P -value = $1.3e-2$). Australian populations showed a negative correlation between the level of activity and absolute latitude ($\tau = -0.49$, P -value = $2.2e-4$), and a positive correlation between the level of activity and longitude ($\tau = 0.40$, P -value = $2.5e-3$). The

latitudinal cline observed here conforms with that reported by Russell and Woodruff (1999).

The samples recently collected by Loreto *et al.* (21 samples from South America) show a positive correlation between the PMM and the mean temperature ($\tau = 0.36$, P -value = $2.4e-2$) and latitude ($\tau = 0.38$, P -value = $1.6e-2$), and negative correlations between the level of activity and absolute longitude ($\tau = -0.42$, P -value = $4.1e-3$).

To sum up, the PMM seems to be frequently correlated with the absolute latitude, longitude and latitude. But the correlation coefficients can be positive or negative for absolute latitude and longitude. If this is not surprising for longitude, it is more puzzling for absolute latitude. Similarly, the PMM can be positively or negatively correlated to temperature. Thus, in conclusion, no general tendency seems to emerge.

As the geoclimatic variables are not independent (Rouault *et al.*, 2004), a principal component analysis was performed in order to get a global view of all the variables. The results are summarized in Figure 1. Again, the projection of the variable on the first two axes (66.3 % of the total variance, Figure 1a) do not show any clear relationship between the level of activity (PMM) and climatic variables (see Figure 1b).

The Nile route

From our data, we tested the Nile Route hypothesis (Figure 2). This hypothesis states that sub-Saharan populations invaded North Africa by ascending the Nile (Lachaise *et al.*, 1988) and from there invaded Europe. Different groups were established: sub-Saharan Africa

Table 2 Correlations between the level of somatic activity (PMM) and geoclimatic variables for the different geographic clusters

Geographic clusters		Mean temperature	Minimum temperature	Maximum temperature	Absolute latitude	Latitude	Longitude
All data	τ	NS	NS	NS	0.21	NS	NS
	P -value	NS	NS	NS	$4.7E-5$	NS	NS
Africa	τ	NS	NS	0.37	0.43	NS	NS
	P -value	NS	NS	$7.9E-3$	$3.9E-5$	NS	NS
Asia	τ	NS	NS	NS	NS	NS	NS
	P -value	NS	NS	NS	NS	NS	NS
Europe	τ	NS	-0.42	NS	0.32	Ibid	-0.29
	P -value	NS	$3.5E-2$	NS	$1.3E-2$	Ibid	$2.6E-2$
North America	τ	NS	NS	NS	NS	NS	NS
	P -value	NS	NS	NS	NS	NS	NS
South and Central America	τ	NS	NS	NS	NS	NS	NS
	P -value	NS	NS	NS	NS	NS	NS
Australia	τ	NS	NS	NS	-0.49	0.49	0.41
	P -value	NS	NS	NS	$2.2E-4$	$2.2E-4$	$2.5E-3$
Sub-Saharan Africa	τ	0.35	NS	0.43	0.33	-0.32	NS
	P -value	$2.3E-2$	NS	$5.8E-3$	$3.5E-3$	$4.3E-3$	NS
South America (Loreto, Wallau)	τ	0.36	NS	NS	-0.38	0.38	0.47
	P -value	$2.4E-2$	NS	NS	$1.6E-2$	$1.6E-2$	$4.1E-3$
Europe and northwest Africa	τ	-0.30	NS	-0.41	NS	NS	-0.29
	P -value	$6.7E-2$	NS	$1.1E-2$	NS	NS	$4.9E-2$

Only significant correlations are presented. For the Europe cluster, ibid refers to the values found for absolute latitude, because for this data set, absolute latitude = latitude. The statistical test is a non-parametric one, Kendall's τ .

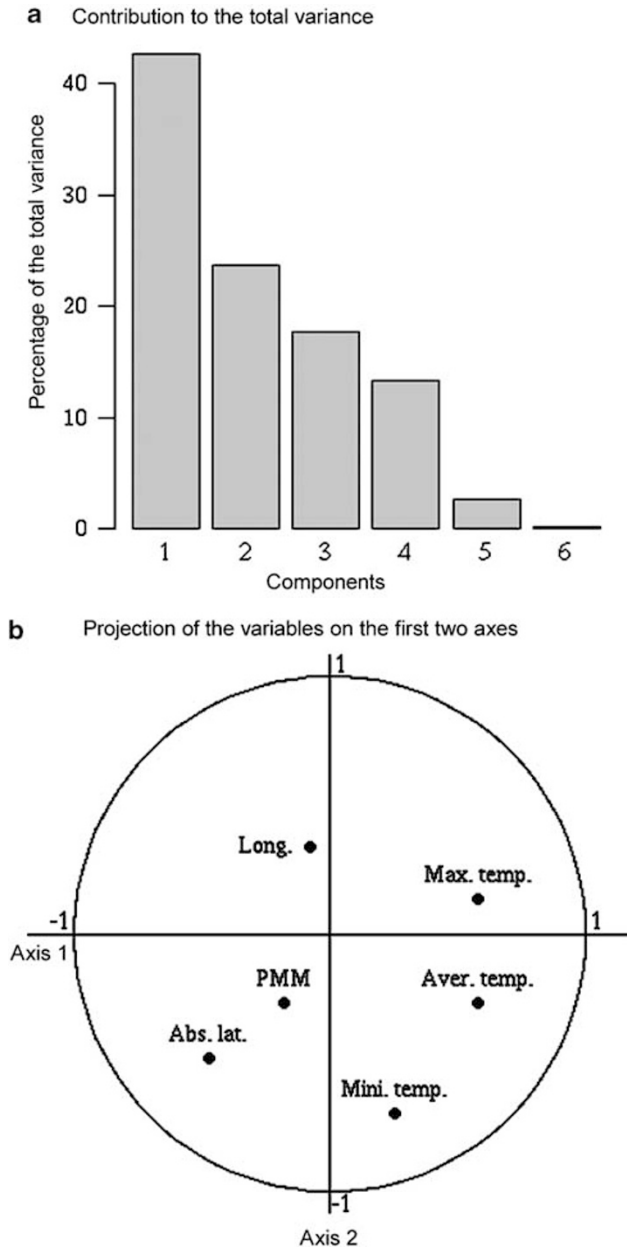


Figure 1 Principal component analysis on percentage of mosaic males (PMM), minimum, maximum and average temperature, absolute longitude and latitude. (a) Histogram representing the contribution of the component to the total variance. (b) Projection of the variables on the first two axes.

(African populations considered as ancestral, latitude < 10, 38 locations), western Europe (Longitude > 20, 19 locations), northwest Africa (African population, latitude > 25, longitude < 10, 6 locations), northeast Africa and the Near East (Latitude from 35 to 24, Longitude > 10, 7 locations). We compared the mean PMM between these groups with the Wilcoxon's test (Table 3a). Two groups can be discriminated: northwest Africa and western Europe are not significantly different (P -value = 2.7×10^{-1}); nor are northeast Africa, the Near East and sub-Saharan Africa (P -value = 8.5×10^{-1}). Both components of the first group show a significantly higher PMM than either of the members of the second group (northwest Africa vs northeast Africa, P -value = 3.4×10^{-3} ,

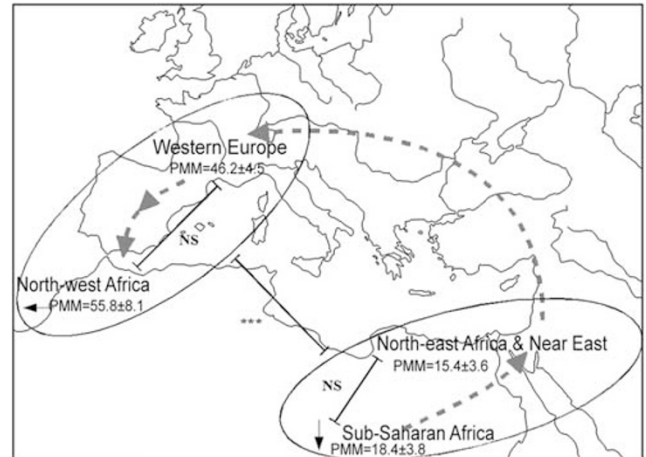


Figure 2 The Nile route hypothesis. The black ellipses correspond to the two groups defined by the statistical analyses: on top, West Europe and northwestern Africa (Morocco), with a higher level of activity and at the bottom, northeastern Africa + the Near East and sub-Saharan Africa, with a lower level of activity. The percentage of mosaic males (PMM) is given with standard error. In dashed grey, the hypothetical invasion route. Not represented here, shipping trade across the Mediterranean Sea could also be important in exchanges in populations between the Near East, northern Africa and Europe.

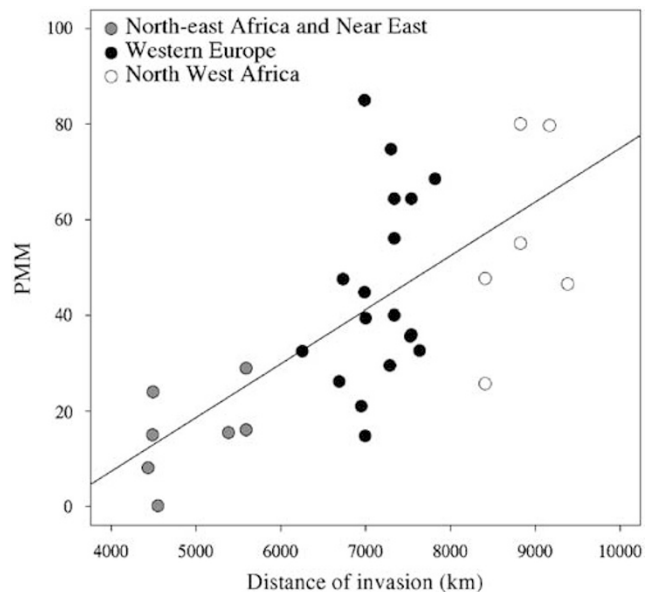


Figure 3 Variation of percentage of mosaic males (PMM) according to the distance of the invasion. The estimation of distances is based on the route drawn in Figure 2. The regression line is drawn in grey.

$d = 38.0$; northwest Africa vs sub-Saharan Africa, P -value = 1.7×10^{-3} , $d = 35.7$; western Europe vs northeast Africa, P -value = 2.2×10^{-5} , $d = 29.0$; western Europe vs sub-Saharan Africa, P -value = 6.6×10^{-5} , $d = 28.3$). The relationship between the PMM and the distance of invasion from sub-Saharan Africa is presented in Figure 3. Distances were calculated according to the Nile Route hypothesis drawn in Figure 2. This is an Euclidian distance. A significant correlation was detected between the level of activity and the distance of invasion

Table 3 Differences in the level of activity between geographic clusters

(a) Geographic clusters		Northwest Africa	Northeast Africa and Near East	Western Europe
Northeast Africa and Near East	P-value	3.4E-03	—	—
	Difference	-38.0	—	—
Western Europe	P-value	2.7E-01	2.2E-05	—
	Difference	-10.5	29.0	—
Sub-Saharan Africa	P-value	1.7E-03	8.5E-01	6.5E-05
	Difference	-35.7	-2.4	-28.3
(b) Geographic clusters		Sub-Saharan Africa	Australia	
Australia	P-value	2.7E-07	—	—
	Difference	33.3	—	—
South America	P-value	1.4E-04	3.6E-02	—
	Difference	18.8	-12.3	—

A non-parametric test, Wilcoxon's test, was used. (a) Tests on the four geographic clusters involved in the Nile Route hypothesis. The results show two clusters, western Europe+northwestern Africa and sub-Saharan Africa+northeastern Africa and the Near East. (b) Tests on the three sub-equatorial geographical clusters: Australia and South America (invasive populations) and sub-Saharan Africa (ancestral populations). The PMM in Australia and South America are significantly higher than in sub-Saharan Africa (respectively, P -value = 2.7E-07 and P -value = 1.4E-4). The PMM in Australia is significantly higher than in South America (P -value = 3.6E-2).

(P -value = 1.1e-4 and Kendall's $\tau = 0.485$; Figure 4). The mantel test on the level of activity and the distance of invasion shows P -value = 0.001 with $r = 0.352$ (Figure 2 in Supplementary Data).

Ancestral vs invasive

To look for effects (historical, populational, and so on) independent of climatic or geographical parameters, we have also compared different natural populations from the same latitude but on different continents: sub-Saharan, Australian and South American populations (Table 3b). These populations are located near the Equator, mainly in the southern hemisphere. South America and Australia have been colonized by *D. simulans* relatively recently, whereas the sub-Saharan populations are thought to represent the ancestral type and to have a large population size (Lachaise et al., 1988; Hamblin and Veuille, 1999; Veuille et al., 2004; Schöfl and Schlötterer, 2006). The comparisons show that, on the one hand, the mean levels of activity in Australia and South America are significantly higher than those of sub-Saharan populations (Australia vs sub-Saharan Africa, $d = 33.3$, P -value = 2.7e-7; South America vs sub-Saharan Africa, $d = 18.8$, P -value = 1.4e-4). On the other hand, South America's PMM is slightly, but significantly lower than Australia's ($d = -12.3$, P -value = 3.6e-2).

Discussion

Copies belonging to the *mauritiana* subfamily have been found in all natural populations of *D. simulans* tested (Giraud and Capy, 1996; Russell and Woodruff, 1999) and also in the members of the *simulans* complex (*D. simulans*, *D. sechellia* and *D. mauritiana*). Moreover, the phylogeny of this element is consistent with that of the species (Brunet et al., 1999). This strongly suggests that the element was present in the common ancestor of this complex. Therefore, it is likely that the *mariner* genome invasion was more ancient compared to the species'

world expansion and thus does not interfere with the distribution of the activity level observed today.

From our data compilation, no general effect on the PMM can be assigned to temperature or latitude. A slight positive correlation was observed between absolute latitude and the level of activity for the whole data, but opposite values at the continental level were found, that is positive for Europe and Africa and negative for Australia and South America. Moreover, a positive correlation was seen between maximum temperature and PMM in Africa and between mean temperature and PMM in South America, whereas a negative value was found between minimum temperature and PMM in Europe. These observations suggest that the hypothetical effect of temperature on somatic activity suggested by Giraud and Capy (1996) cannot be generalized.

Giraud and Capy (1996) suggested a hypothesis to explain the correlation they found between absolute latitude and *mariner* somatic activity. Previous experiments have shown that for several *D. simulans* strains, *mariner* somatic activity increases with developmental temperature in laboratory conditions (Chakrani et al., 1993). Moreover, transposition activity is generally viewed as deleterious. Thus, selection against transposition (regulation or selection for less copies or less active copies) is expected to be stronger in warmer areas. With the phenotypic test being carried out at 25 °C, the African populations should have a lower level of activity than those from Europe.

The PMM may also be affected by several parameters, including the effective population sizes (N_e ; Lynch and Conery, 2003), and for instance *via* the active copy number, the genetic status and/or the stressful conditions of the invasive populations facing new environments (Biémont et al., 1997; Vieira et al., 1999). The effect of these parameters remains difficult to assess because generally we do not know (1) whether a population is a recent invasive or a long-established one; (2) the adaptation latency for a natural population; (3) the

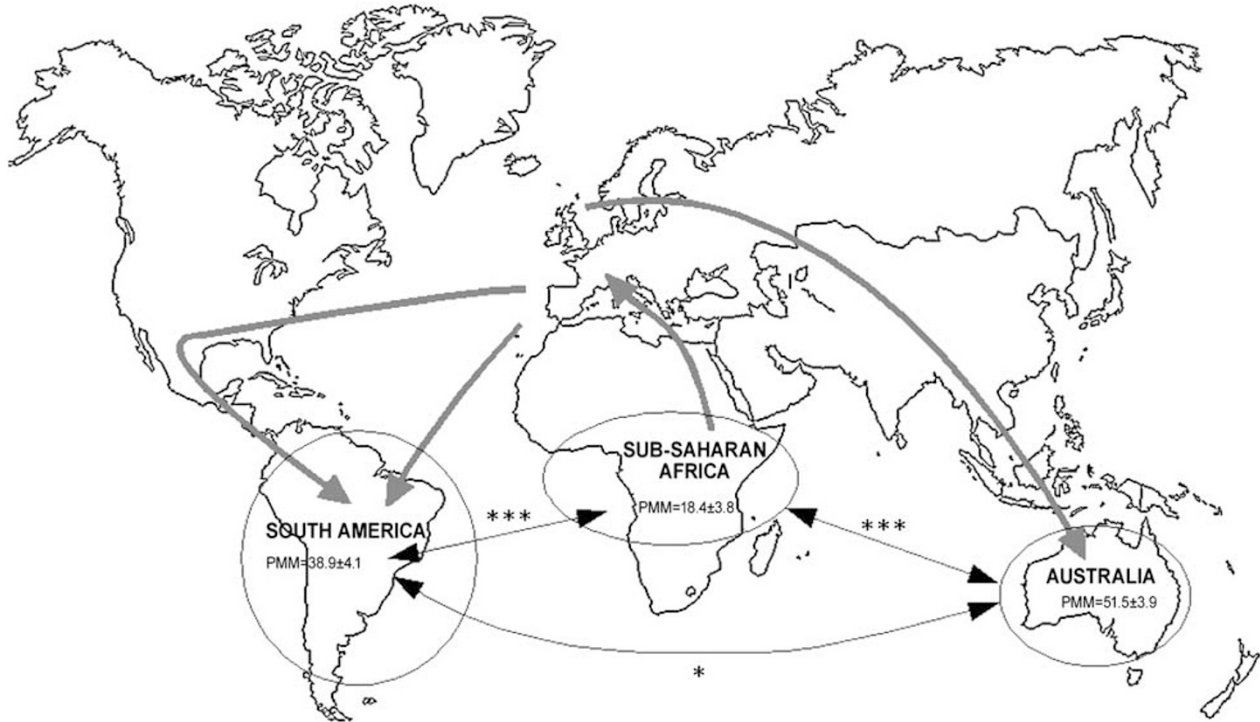


Figure 4 Comparison between two set of invasive populations (in Australia and South America) and the ancestral populations (sub-Saharan Africa). The percentage of mosaic males (PMM) is given with the standard error. The results of the tests show that the level of activity is significantly higher in Australian and South American populations than in sub-Saharan ones ($d = 33.3$, $P\text{-value} = 2.7e-7$ and $d = 18.8$, $P\text{-value} = 1.4e-4$). The PMM of the Australian population is higher than the South American population ($d = 12.3$, $P\text{-value} = 3.6e-2$). The arrows in grey show the suggested population movements.

effective population size and (4) to what extent selection affects and decreases the activity level of a population.

From a geographical point of view, European and northwest African populations exhibit a higher level of activity than sub-Saharan or northeast African ones. According to the Nile Route hypothesis, it is quite possible that the former correspond to invasives originating from the latter. The association between Europe and northwest Africa is not surprising given the geographical proximity of Spain and Morocco and the extensive trade that is likely to have promoted a relatively high migration rate. Schöfl and Schlötterer (2006) have suggested a scenario of substantial migration across the Mediterranean Sea to explain the little genetic differentiation found between European and north-African populations from Tunisia. South America and Australia also show a higher level of activity when compared to the sub-Saharan ancestral populations. Our analysis thus clearly suggests that all recent invasive populations show a higher level of *mariner* activity. The term 'recent' is somewhat imprecise, but it is currently impossible to determine the exact date of invasion of several parts of the world by *D. simulans*. We know that this species was described for the first time in the Japanese islands in 1976 (Watanabe and Kawanishi, 1976), but for the other continents, we are restricted to the hypothesis that massive population displacements and trade are probably at the origin of the migration of many insects species over the last few centuries.

Previous studies have emphasized that sub-Saharan populations can be viewed as the ancestors of *D. simulans*. The speciation events that led to the evolution of *D. simulans* are assumed to have taken place on the

east coast of central Africa and on the Indian Ocean islands (for example Madagascar). The populations would then have dispersed to the sub-Saharan area, but were stopped by the Cameroon cordillera. Later, *D. simulans* would have become a cosmopolitan invasive, using the Nile route for its way out of Africa (Lachaise et al., 1988; Lachaise and Silvain, 2004). Regarding *mariner*, it has been shown that ancestral African populations exhibit less *mariner* somatic activity and a reduced copy number compared to non-African populations (Giraud and Capy, 1996; Russell and Woodruff, 1999). In this study, we also show that the low level of activity can be extended to the northeast African and Near-Eastern populations which probably correspond to the oldest invasive populations (about 6500–5000 years old according to Lachaise et al., 1988).

These results suggest that the level of *mariner* somatic activity is probably linked to the process of invasion by *D. simulans*. However, prudence is necessary as such a conclusion remains difficult to test. The PMM is only slightly and probably not significantly related to climatic parameters, but populational parameters could be more important. What might these parameters be? If we assume that invasion is the most relevant hypothesis to explain the *mariner* somatic activity, we must try to understand why.

A first possible explanation could be that during the invasion process a population undergoes repetitive and heavy bottlenecks, that is drastic decreases in population size. In population genetics, it is known that the fluctuation of the N_e modifies the equilibrium between genetic drift and natural selection. In other words, when N_e increases, natural selection supplants genetic drift,

whereas a weak N_e reinforces the genetic drift effect. In this context, Lynch and Conery (2003) propose that an accumulation of TEs (and a possible increase in their transposition activity) could be the result of low N_e .

Environmental and genomic stresses increase the level of transcription and transposition activity of TEs. So a second explanation could be that migrants that face new and stressful environments show a deregulation of the activity of the *mariner* elements and other TEs, leading to genome instability (Capy et al., 2000). This instability may produce genetic variation and novelties from which new genetic combinations could be selected and adaptive (Biémont et al., 1997; Vieira et al., 1999). Such a deregulation should be temporary to avoid a deleterious effect on selected genetic combinations. In this respect, epigenetic phenomena could be involved in the activity variation of TEs. Today, several works suggest that 'epigenetic mechanisms have evolved in eukaryotic cells to silence the expression and mobility of transposable elements' (see Slotkin and Martienssen, 2007 for a review). Therefore, invasions of new and stressful environments may lead to modifications of the epigenome and then to the activity of TEs. Many other phenomena could be also involved in the deregulation of TEs' activity as discussed by Capy et al. (2000). For instance, there are transcription activator binding sites in TE's regulatory region also found in the regulatory regions of host defence genes (Grandbastien et al., 1997).

A third hypothesis is that the higher activity of *mariner* could be partly due to functional hitchhiking. In this case, we have to assume that a *mariner* copy able to produce a functional transposase is genetically linked to a selected allele or region.

These different hypotheses, adaptive, stochastic and functional hitchhiking can lead to an increase of the *mariner* activity, and to a relationship between the level of TE's activity and the invasive history of the species. They are not mutually exclusive and they are difficult to tease apart.

Acknowledgements

This work was supported by grants from GDR 2157 from the CNRS and from the COFECUB project 492/05. We thank Malcolm Eden and Shane McEvey for the English review of the text.

References

Anxolabéhère D, Nouaud D, Périquet G, Tchen P (1985). *P-element* distribution in Eurasian populations of *Drosophila melanogaster*: a genetic and molecular analysis. *Proc Natl Acad Sci USA* **82**: 5418–5422.

Biémont C, Vieira C, Hoogland C, Cizeron G, Loevenbruck C, Arnault C et al. (1997). Maintenance of transposable element copy number in natural populations of *Drosophila melanogaster* and *D. simulans*. *Genetica* **100**: 161–166.

Bock IR, Parsons PA (1981). Species of Australia and New Zealand. In: Ashburner M, Carson HL, Thompson JN Jr (eds). *The Genetics and Biology of Drosophila*, vol. 3a, Academic Press: London. pp 291–308.

Bonnivard E, Bazin C, Higué D (2002). High polymorphism of TPE repeats within natural populations of *Drosophila melanogaster*: a gradient of the 5TPE hobo element in Western Europe. *Mol Biol Evol* **19**: 2277–2284.

Brunet F, Godin F, Bazin C, Capy P (1999). Phylogenetic analysis of *Mos1*-like transposable elements in the *Drosophilidae*. *J Mol Evol* **49**: 760–768.

Bryan G, Garza D, Hartl D (1990). Insertion and excision of the transposable element *mariner* in *Drosophila*. *Genetics* **125**: 103–114.

Capy P, Gasperi G, Biémont C, Bazin C (2000). Stress and transposable elements: co-evolution or useful parasites? *Heredity* **85** (Part 2): 101–106.

Chakrani F, Capy P, David JR (1993). Developmental temperature and somatic excision rate of *mariner* transposable element in three natural populations of *Drosophila simulans*. *Genet Sel Evol* **25**: 121–132.

Chessel D, Dufour A, Thioulouse J (2004). The ade4 package-I: one-table methods. *R News* **4**: 5–10.

Garza D, Medhora M, Koga A, Hartl D (1991). Introduction of the transposable element *mariner* into the germline of *Drosophila melanogaster*. *Genetics* **128**: 303–310.

Giraud T, Capy P (1996). Somatic activity of the *mariner* transposable element in natural populations of *Drosophila simulans*. *Proc R Soc Lond B Biol Sci* **263**: 1481–1486.

Grandbastien MA, Lucas H, Morel JB, Mhiri C, Vernhettes S, Casacuberta JM (1997). The expression of the tobacco *Tnt1* retrotransposon is linked to plant defense responses. *Genetica* **100**: 241–252.

Hamblin MT, Veuille M (1999). Population structure among African and derived populations of *Drosophila simulans*: evidence for ancient subdivision and recent admixture. *Genetics* **153**: 305–317.

Ihaka R, Gentleman R (1996). R: a language for data analysis and graphics. *J Comput Graph Stat* **5**: 299–314.

Lachaise D, Cariou M, David JR, Lemeunier F, Tsacas L, Ashburner M (1988). Historical biogeography of the *Drosophila melanogaster* species subgroup. In: Hecht KM, Wallace B, Ghilleant TP (eds). *Evolutionary Biology*, vol. 22, Plenum Press: New York. pp 159–225.

Lachaise D, Silvain J (2004). How two Afrotropical endemics made two cosmopolitan human commensals: the *Drosophila melanogaster*-*D. simulans* paleogeographic riddle. *Genetica* **120**: 17–39.

Lynch M, Conery JS (2003). The origins of genome complexity. *Science* **302**: 1401–1404.

Rouault J, Marican C, Wicker-Thomas C, Jallon J (2004). Relations between cuticular hydrocarbon (HC) polymorphism, resistance against desiccation and breeding temperature; a model for HC evolution in *D. melanogaster* and *D. simulans*. *Genetica* **120**: 195–212.

Russell A, Woodruff R (1999). The genetics and evolution of the *mariner* transposable element in *Drosophila simulans*: worldwide distribution and experimental population dynamics. *Genetica* **105**: 149–164.

Schöfl G, Schlötterer C (2006). Microsatellite variation and differentiation in African and non-African populations of *Drosophila simulans*. *Mol Ecol* **15**: 3895–3905.

Singh RS, Choudhary M, David JR (1987). Contrasting patterns of geographic variation in the cosmopolitan sibling species *Drosophila melanogaster* and *Drosophila simulans*. *Biochem Genet* **25**: 27–40.

Slotkin RK, Martienssen R (2007). Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet* **8**: 272–285.

Veuille M, Baudry E, Cobb M, Derome N, Gravot E (2004). Historicity and the population genetics of *Drosophila melanogaster* and *Drosophila simulans*. *Genetica* **120**: 61–70.

Vieira C, Biémont C (1996). Geographical variation in insertion site number of retrotransposon 412 in *Drosophila simulans*. *J Mol Evol* **42**: 443–451.

Vieira C, Lepetit D, Dumont S, Biémont C (1999). Wake up of transposable elements following *Drosophila simulans* worldwide colonization. *Mol Biol Evol* **16**: 1251–1255.

Watanabe TK, Kawanishi M (1976). Colonization of *Drosophila simulans* in Japan. *Proc Jpn Acad* **52**: 191–194.

Supplementary Information accompanies the paper on Heredity website (<http://www.nature.com/hdy>)