

RESEARCH ARTICLE

Open Access



Patterns of geographic variation of thermal adapted candidate genes in *Drosophila subobscura* sex chromosome arrangements

Pedro Simões^{1,2*}  and Marta Pascual¹

Abstract

Background: The role of chromosomal arrangements in adaptation is supported by the repeatable clinal variation in inversion frequencies across continents in colonizing species such as *Drosophila subobscura*. However, there is a lack of knowledge on the genetic variation in genes within inversions, possibly targets of climatic selection, across a geographic latitudinal gradient. In the present study we analysed four candidate loci for thermal adaptation, located close to the breakpoints, in two chromosomal arrangements of the sex (A) chromosome of *Drosophila subobscura* with different thermal preferences. Individual chromosomes with A₂ (the inverted arrangement considered warm adapted) or A_{ST} (the standard ancestral arrangement considered cold adapted) were sequenced across four European localities at varying latitudes, up to ~ 2500 Kms apart.

Results: Importantly, we found very low differentiation for each specific arrangement across populations as well as no clinal patterns of genomic variation. This suggests wide gene exchange along the cline. Differentiation between the sex chromosome arrangements was significant in the two more proximal regions relative to the A_{ST} orientation but not in the distal ones, independently of their location inside or outside the inversion. This can be possibly due to variation in the levels of gene flux and/or selection acting in these regions.

Conclusions: Gene flow appears to have homogenized the genetic content within-arrangement at a wide geographical scale, despite the expected diverse selective pressures in the specific natural environments of the different populations sampled. It is thus likely that the inversion frequency clines in this species are being maintained by local adaptation in face of gene flow. The differences between arrangements at non-coding regions might be associated with the previously observed differential gene expression in different thermal regimes. Higher resolution genomic scans for individual chromosomal arrangements performed over a large environmental gradient are needed to find the targets of selection and further elucidate the adaptive mechanisms maintaining chromosomal inversion polymorphisms.

Keywords: Chromosomal arrangements, *Drosophila subobscura*, UTR variation, Genetic differentiation, Gene flow, Geographic variation, Clinal variation, Climatic selection, Thermal adaptation

Background

There is increasing evidence of the role of chromosomal inversion polymorphisms in adaptation and speciation [1, 2]. Several lines of evidence attest this namely: (1) the repeated occurrence of latitudinal clinal variation across

multiple invaded continents, particularly in *Drosophila* species ([3–6]; (2) consistent associations between inversions and several relevant phenotypic traits [7–13]; and (3) inversion frequency changes beyond genetic drift expectations across replicated populations adapting to imposed laboratory conditions [14–16].

Several hypotheses have been advanced to explain the adaptive character of inversions (see [1]), with the most widely accepted hypotheses relying on recombination reduction in inversion heterokaryotypes and consequent protection of genes with positive effects on fitness [17, 18].

* Correspondence: pmsimoes@fc.ul.pt

¹Departament de Genètica, Microbiologia i Estadística and IRBio, Facultat de Biologia, Universitat de Barcelona, 08028 Barcelona, Spain

²Present Address: cE3c - Centre for Ecology, Evolution and Environmental Changes, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal



The coadaptation hypothesis [17] postulates that specific favourable combinations of alleles interacting epistatically within a particular inversion arise as an adaptive response to a given environment. The local adaptation hypothesis [18] states that the spread of a new inversion results from advantageous locally adapted alleles captured within the inverted segment, without epistasis being necessarily involved. Inversion frequencies changes will result from a balance between selection and migration. A general expectation in both these scenarios is that there is genetic differentiation between inversion types, as a result of local adaptation shaping the genetic content of different inversions. Coalescent models suggest that this differentiation is likely higher near breakpoints – where recombination is greatly reduced – and near locally adapted alleles within the inversion [19]. In addition, under the coadaptation hypothesis there is a clear expectation of genetic differentiation within inversions sampled from different geographical populations, as a result of population-specific epistatic interactions evolving in response to different environmental conditions [20]. However, this pattern might not be exclusive of the co-adaptation model as some within-inversion differentiation might occur depending on the balance between migration and selection on locally adapted alleles (with or without epistasis involved) as postulated by the local adaptation.

Empirical studies on the molecular variation associated with inversions have frequently detected high levels of genetic differentiation between chromosomal arrangements ([21–27] but see [28]). This pattern is consistent with the aforementioned hypotheses and is expected as a result of reduced recombination in inversion heterokaryotypes leading to distinct genetic pools. Studies in the O chromosome of *D. subobscura* found an extensive genetic differentiation between arrangements occurring regardless of the distance to the nearest breakpoint, with selection likely implicated [29, 30]. On the other hand, several studies have reported a general lack of measurable genetic differentiation within arrangements sampled from different populations ([20, 31–33], contrary to the expectations of the coadaptation hypothesis. An important exception is the genetic clinal variation within the inversion *In(3R)Payne* of *Drosophila melanogaster*, both in North American and Australian localities, implicating a role of this inversion in the response to climatic selection [6, 31, 34]. In addition, there is some evidence for epistasis – as expected under the coadaptation model – based on patterns of linkage disequilibrium (LD) with regions within arrangements showing high levels of LD and differentiation interspersed with other regions of low LD [20, 31] as well as LD between arrangements and genes located outside them [30].

One of the most emblematic examples of the adaptive role of chromosomal polymorphisms comes from *Drosophila subobscura*, a species presenting repeatable clinal patterns for inversion frequencies across Europe, the ancestral area, and North and South America [3]. Thermal adaptation or associated factors might be implicated in such patterns. Evidence for this hypothesis comes from the association between temporal frequency changes of *D. subobscura* chromosomal arrangements and increased temperature due to climate warming at a worldwide scale [5]. Also, seasonal inversion frequency changes in *D. subobscura* have been reported, even shortly responding to heat waves [35] as well as associations between inversions and several phenotypic traits, including thermal preference and tolerance [36–38] and body size [10, 39]. Importantly, differences in gene expression patterns were also observed in *D. subobscura* from a colonizing South American population evolving in the laboratory at different thermal regimes (13, 18 and 22 °C); with an enrichment for candidate genes located inside arrangements [40]. This suggests that different inversions may contribute to the maintenance of genes with differential expression and potential impact on fitness. Indeed, this was the case in *D. pseudoobscura* with multiple genes presenting expression differences among arrangements [41]. Consequently, it is relevant to analyse overall patterns of molecular variation in these candidate genes in the vicinity/within the *D. subobscura* inversions as well as address their variation across the species geographic gradient to search for targets of climatic selection.

Pegueroles et al. [33] analysed 6 of these genes previously associated with thermal adaptation located in the O chromosome of *D. subobscura* and found no differentiation for the same arrangement between Barcelona and Greece indicating high levels of gene flow and the absence of geographic varying selection at least in those two southern regions. In a study in *D. subobscura* involving several localities across western Europe, Simões et al. [32] using microsatellite loci also detected low within-arrangement differentiation in several inversions of two autosomes (J and U) and the sex chromosome (A) confirming an extensive gene flow in this species (see also [42, 43]). Nevertheless, significant clinal patterns at few specific microsatellite loci within arrangements emerged (one in the distal region of A₂), suggesting adaptation at neighbouring loci.

The aim of the present study is to examine the genetic variation associated with two arrangements of the sexual chromosome of *Drosophila subobscura* across a geographic gradient. Specifically, we intend to (1) characterize genetic variation and differentiation between arrangements, addressing the impact of inverted regions at the genetic level; (2) analyse patterns of within-arrangement differentiation, searching for evidence of different epistatic

combinations in different populations; (3) identify latitudinal patterns of genetic variation, which might indicate targets of climatic selection at a geographical scale. In this study we targeted A_{ST} and A_2 , the two most frequent inversions of the A chromosome in our sampled localities. These two arrangements present clinal variation of frequencies in Europe, as well as in the two recently colonized American hemispheres, with A_2 increasing frequency towards warmer locations and A_{ST} showing the corresponding opposite pattern [32, 44]. We analysed DNA sequence variation in four gene regions (*PhKgamma*, *Ubc-E2H*, *Hfw* and *Nipsnap*). These genes are good candidates to search for evidences of climatic latitudinal selection associated with the *Drosophila subobscura* inversion clines as they (1) have been previously implicated in thermal adaptation ([40], see also above) and (2) are located relatively close to inversion breakpoints, with reduced recombination in these regions possibly protecting genetic variants and showing the footprint of selection.

Methods

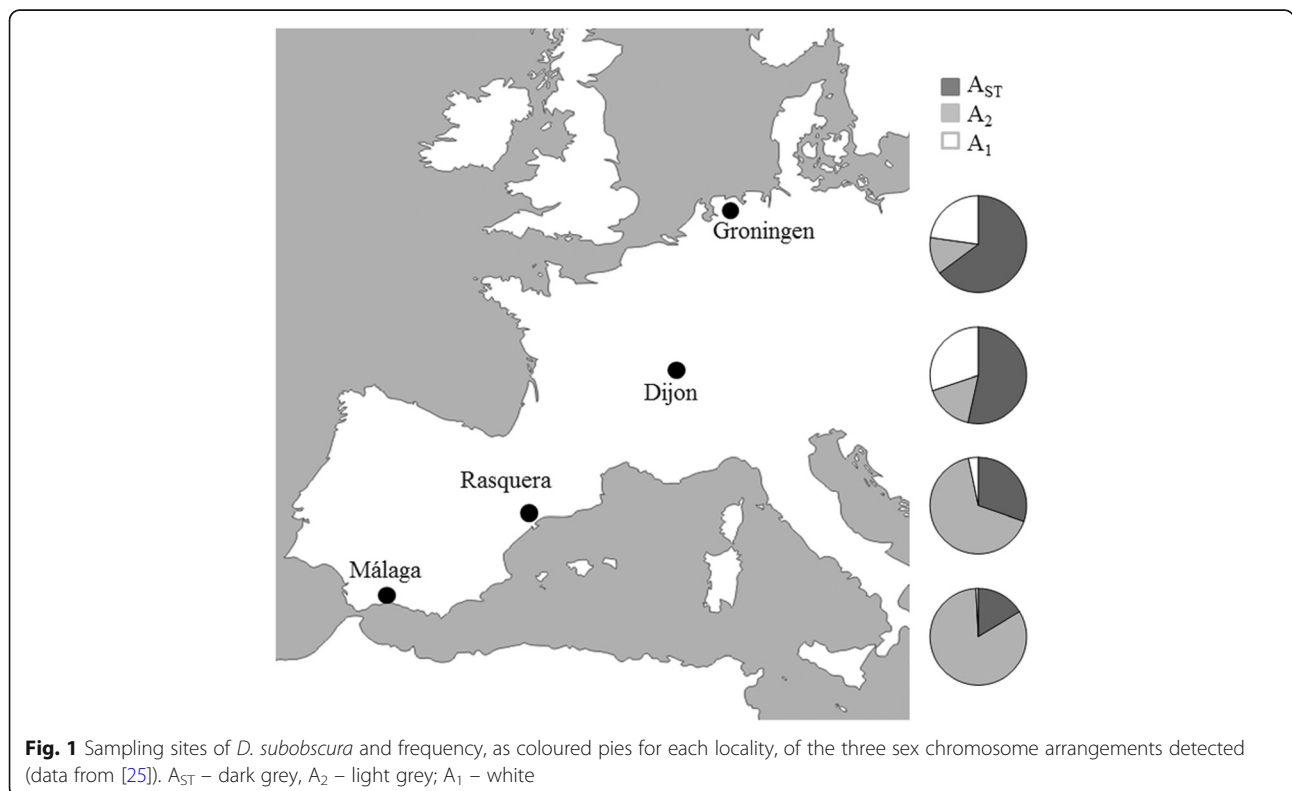
Fly samples and chromosomal inversion scoring

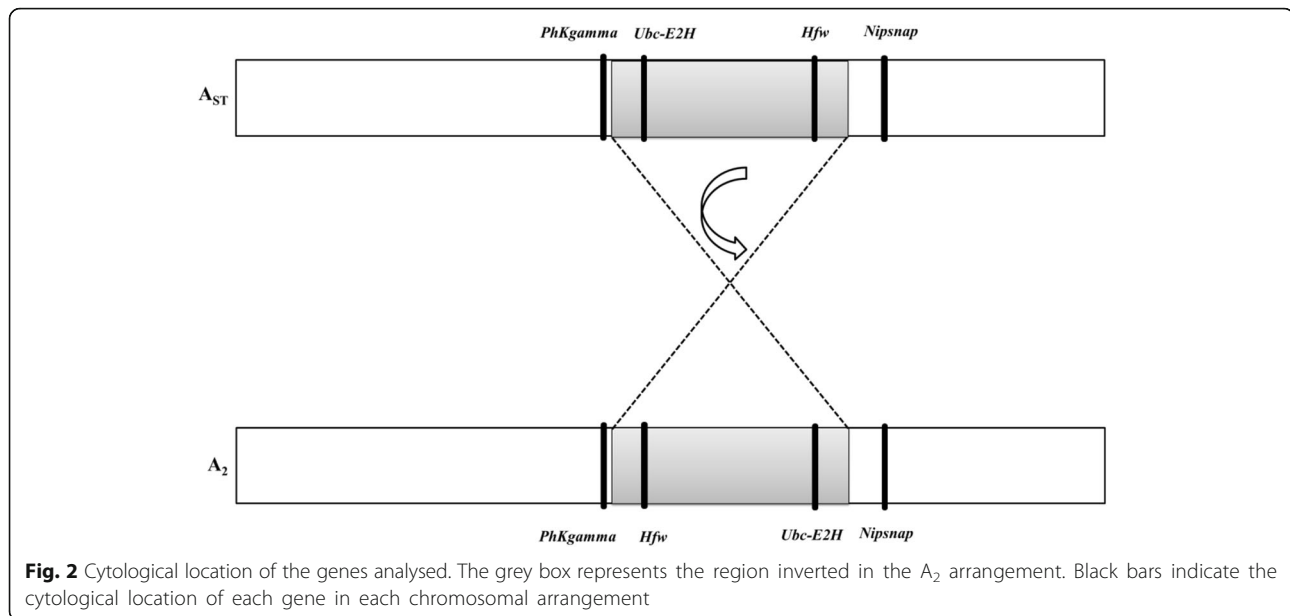
Wild *Drosophila subobscura* samples were collected from four European locations (Fig. 1). Samples from Málaga (Mal, 36° 43'N, Spain) and Rasquera (Ras, 40° 57'N, Spain) were obtained in October 2008 with a total of 169 and 152 individuals, respectively. Samples from Dijon (Dij, 47° 18'N, France) and Groningen (Gro, 57° 13'N,

Netherlands) were obtained in August/September 2009 with a total of 344 and 326 individuals, respectively [45]. These locations vary greatly in environmental conditions, for instance the two most extreme sampled locations have presented clearly contrasting range of yearly average temperatures (7–31 °C in Málaga vs. 0–23 °C in Groningen) and also average precipitation values (44 mm in Málaga vs. 75 mm in Groningen) between 2000 and 2012. We performed the collections in the late summer/early fall to capture the second peak of abundance for the species in the different locations allowing a more reliable comparisons between warmer and colder populations [46].

Our study complies with National and International guidelines, either involving field collections or experimentation in animals. We use *Drosophila subobscura*, which is not a threatened species nor requires legal ethical consent for experimentation.

This study focused on the two most frequent arrangements present on the acrocentric A (sexual) chromosome (A_{ST}/A_2), and the only two arrangements present in all analysed locations (Fig. 1). A_{ST} is considered an ancestral chromosome arrangement and A_2 a derived medium-sized inversion (~7.1 Mb) with an estimated age of 160.000 years [23]. The breakpoints of A_2 (Fig. 2) are 8C/D-12C/D [47]. In order to score the chromosomal arrangements of the A chromosome, wild-caught males or F_1 male descendants from wild isofemale lines were individually crossed with virgin females of the *chcu*





strain, homokaryotypic for the A_{ST} arrangement. Polytene chromosomes from salivary glands of one female larva descendant per cross were stained with 2% orcein in 60% acetic/lactic acid (50:50) (see details in [32]). This protocol allowed identifying the chromosomal arrangement of each parental male from the chromosomal configuration of their female larva progeny.

Genomic regions, DNA extraction and sequencing

DNA was individually extracted from wild males (or male descendants of wild females) whose chromosomal arrangements (A_2 or A_{ST}) had been previously scored as described above. By amplifying and directly sequencing males we avoided the problems associated with sequencing of diploid data. The karyotype of these males was inferred from their descendants (see above) making it possible to obtain both genetic and karyotypic data for the same individual. Genomic DNA extraction was performed following the protocol described in [48]. The four genomic regions analysed were localized again by in situ hybridization with digoxigenin following standard protocols [49], as some genes in Laayouni et al. [40] have proved to be mislocated [33]. All probes hybridized in the same band previously reported [40] with the exception of *Nipsnap* that was localized nearby but outside the inversion (see Fig. 2 and Additional file 1). Thus, *PhKgamma* (cytological location 8C; CG1830) and *Nipsnap* (13B; CG9212) are located outside the A_2 inversion, and *Ubc-E2H* (9A; CG2257) and *Hfw* (12A; CG3095) within that inversion (Fig. 2). Following the A_{ST} arrangement, being the ancestral one, we will refer to *PhKgamma* and *Ubc-E2H* as the proximal loci, the region being closer to the centromere, and *Hfw* and

Nipsnap as the distal loci. *Ubc-E2H* (Ubiquitin conjugating enzyme E2H) is involved in lateral inhibition and protein polyubiquitination, *PhKgamma* (Phosphorylase kinase enzyme) is involved in embryonic morphogenesis and glycogen metabolism and *Hfw* (Halfway) is involved in intracellular signal transduction and pupariation and no involvement in biological processes is yet ascribed to the *Nipsnap* gene although a putative role in vesicular transport has been proposed (the function of each gene is according to Flybase). Primers used for amplification and sequencing reactions are listed in Table 1. The fragments amplified correspond to partial sequences, not comprising the totality of each gene (see Table 1 for details). Distance to the nearest breakpoint (Mb) was estimated by counting the number of cytological bands in the map of Kunze-Mühl and Müller [50] assuming that all bands contain the same DNA content (82 Kb) as in [51]. Distances were around 0.082 Mb for *PhKgamma* and around 0.574 Mb for the other 3 genes (Fig. 2).

DNA amplification and sequencing was carried out as in [45] with an annealing temperature of 5 s at 54 °C during the amplification reaction. For each individual two sequences were obtained, from forward and reverse primers (see primer sequences in Table 1). Both sequences were assembled using SeqMan II from the Lasergene program (DNASTAR) to get one final sequence per individual. All sequences were then aligned with Clustal W, in BioEdit v.7.2 [52]. The sequence length of the total alignment for each of the four analysed genomic regions is given in Table 1 with the 5'UTR, exonic and intronic regions assigned by comparison with the orthologous sequences of *Drosophila pseudoobscura* retrieved from Flybase and visual inspection

Table 1 Chromosomal location, length and primers used for each of the 4 genes studied

Gene ^a	Alignment ^b	Primers (Forward; Reverse)	5'UTR	Exons	Introns
<i>PhKgamma</i> (CG1830, GA14880)	831 bp (895 bp)	5' - CATGACCTGGCGCAGTATTG - 3' 5' - TAACAGCGGAGCGAGCAGTC - 3'	1–240	241–332	333–895
<i>Ubc-E2H</i> (CG2257, GA15327)	1085 bp (1100 bp)	5' - TCACTGTAGTCGGACATGCT - 3' 5' - AGGAGAGCAACGTCACAGAT - 3'	1–662	663–1100	–
<i>Hfw</i> (CG3095, GA15922)	1073 bp (1079 bp)	5' - GCATTCAAGCGGTCCGTAA - 3' 5' - TATGTTGAGGCACTTGAGCG - 3'		349–1079	1–348
<i>Nipsnap</i> (CG9212, GA21617)	1095 bp (1168 bp)	5' - CATGCGACTGTGAGCCTCTT - 3' 5' - CGCAGCAATACAACAAGTGG - 3'	1–411	412–470; 1149–1168	471–1148

^aGene symbol of the homologous gene in *D. melanogaster* and *D. pseudoobscura* respectively in parentheses

^balignment length with *D.pseudoobscura* in parentheses. 5' UTR, introns and exon positions are referred relative to the *D. pseudoobscura* alignment

of the resulting expected proteins. The number of males analysed per locality and inversion are presented for each gene fragment in Table 2. Our sampling strategy targeted a comparable number of males to be analysed for each population and arrangement, and as such the inversion frequencies in our samples do not match natural inversion frequencies as seen in Fig. 1.

Data analysis and statistical methods

Several standard parameters for each locality and arrangement were estimated using DnaSP v5 [53]: the number of haplotypes (*h*), number of polymorphic sites (*S*), the nucleotide diversity (π), nucleotide diversity in synonymous sites and non-coding positions (π_{sil}), heterozygosity in silent sites (θ_{sil}) and divergence per silent site (K_{sil}) using *D. pseudoobscura* as outgroup [54], with

the Jukes-Cantor correction. Comparisons of the diversity values across localities and arrangements were performed with the Wilcoxon matched pairs test using Statistica v10. Genetic differentiation between or within arrangements was assessed using F_{ST} [55] and D_{xy} [56] as well as significant levels for Snn [57] were obtained with 10,000 replicates. False Discovery rate - FDR - correction [58] was applied to account for multiple comparisons in the differentiation analyses. Analyses were performed excluding gaps. For a visual representation of the genetic differences between arrangements and populations in the *PhKgamma* and *Ubc-E2H* gene regions, pairwise F_{ST} values were used to compute a Principal Coordinate Analysis (pcoa in R package ape version 3.5 (<http://ape-package.ird.fr/>)). A locus-by-locus analysis of Molecular variance (AMOVA) was performed on Arlequin v3.5.2 [59] to test for genetic

Table 2 Nucleotide variation for each gene and arrangement across all localities

	<i>PhKgamma</i>				<i>Ubc-E2H</i>				<i>Hfw</i>		<i>Nipsnap</i>	
	Mal	Ras	Dij	Gro	Mal	Ras	Dij	Gro	Mal	Gro	Mal	Gro
<i>A_{ST}</i>												
<i>n</i>	9	9	10	10	9	10	10	8	7	9	9	10
<i>h</i>	6	9	10	8	9	10	10	8	7	9	9	10
<i>S</i>	9	13	14	9	12	15	16	11	34	24	27	35
π	0.0031	0.0057	0.0055	0.0031	0.0038	0.0040	0.0047	0.0031	0.0105	0.0070	0.0084	0.0097
π_{sil}	0.0034	0.0052	0.0042	0.0020	0.0038	0.0058	0.0055	0.0034	0.0142	0.0103	0.0090	0.0102
θ_{sil}	0.0044	0.0055	0.0048	0.0027	0.0043	0.0071	0.0073	0.0044	0.0175	0.0120	0.0099	0.0124
K_{sil}	0.1912	0.1926	0.1910	0.1887	0.3865	0.3780	0.3914	0.3799	0.2548	0.2474	0.4104	0.4138
<i>A₂</i>												
<i>n</i>	10	8	9	10	10	9	10	10	10	10	10	10
<i>h</i>	5	3	3	5	8	9	8	6	10	10	9	10
<i>S</i>	5	8	2	4	12	11	8	6	19	18	24	22
π	0.0014	0.0025	0.0008	0.0012	0.0028	0.0030	0.0020	0.0015	0.0053	0.0052	0.0068	0.0056
π_{sil}	0.0012	0.0026	0.0009	0.0008	0.0033	0.0028	0.0011	0.0014	0.0085	0.0095	0.0070	0.0057
θ_{sil}	0.0021	0.0040	0.0011	0.0011	0.0050	0.0042	0.0020	0.0020	0.0102	0.0108	0.0084	0.0076
K_{sil}	0.1988	0.1981	0.1999	0.1988	0.3728	0.3728	0.3714	0.3716	0.2469	0.2483	0.4144	0.4177

Note: *n* - sample size; *h* - number of haplotypes; *S* - number of polymorphic sites; π - nucleotide diversity

π_{sil} - nucleotide diversity in silent sites (synonymous and non-coding positions); θ_{sil} - heterozygosity in silent sites

K_{sil} - divergence per silent site using *D. pseudoobscura* as outgroup

differentiation (F_{CT}) between arrangements at each specific polymorphic site. A hierarchical structure was considered with two groups corresponding to the two arrangements (A_2 vs A_{ST}), with the different localities and individuals sampled within these groups.

Gene conversion tracts (GCTs) were detected using the algorithm defined by [60]. Linkage disequilibrium (LD) between pairs of informative sites was assessed and the statistical significance analysed with a chi-square test after Bonferroni correction. These analyses were performed for each arrangement and gene region, as well as for the concatenated dataset. The overall linkage disequilibrium statistic ZnS was estimated for informative sites. LD was also estimated between polymorphic informative sites through r^2 estimates. Deviations from neutrality were tested using Tajima's D [61] for each chromosomal arrangement and locality. Population size changes were also determined using R^2 [62]. Statistical significance was obtained by coalescent simulations (5000 replicates) considering free recombination, as it seems most suitable when analysing *D. subobscura* genes whether located inside or outside chromosomal arrangements (see [33]). All neutrality tests were performed after removing recombinant individuals (with indication of GCTs). The DnaSP v5 program [53] was used to perform the GCTs and linkage disequilibrium analyses as well as neutrality tests.

Results

Nucleotide variability and genetic differentiation

For the two arrangements (A_{ST} and A_2), we analysed the four genomic regions in the two most geographically distant localities (Málaga and Groningen), and two gene regions for the two central localities (Rasquera and Dijon) (Table 2). The two gene fragments being analysed in all localities were *PhKgamma* and *Ubc-E2H* as they showed genetic differentiation between arrangements (see Figs. 3 and 4 and Additional file 2). The genetic variation in those regions is likely to be affected by inversions to a greater extent and thus could possibly present clinal variation.

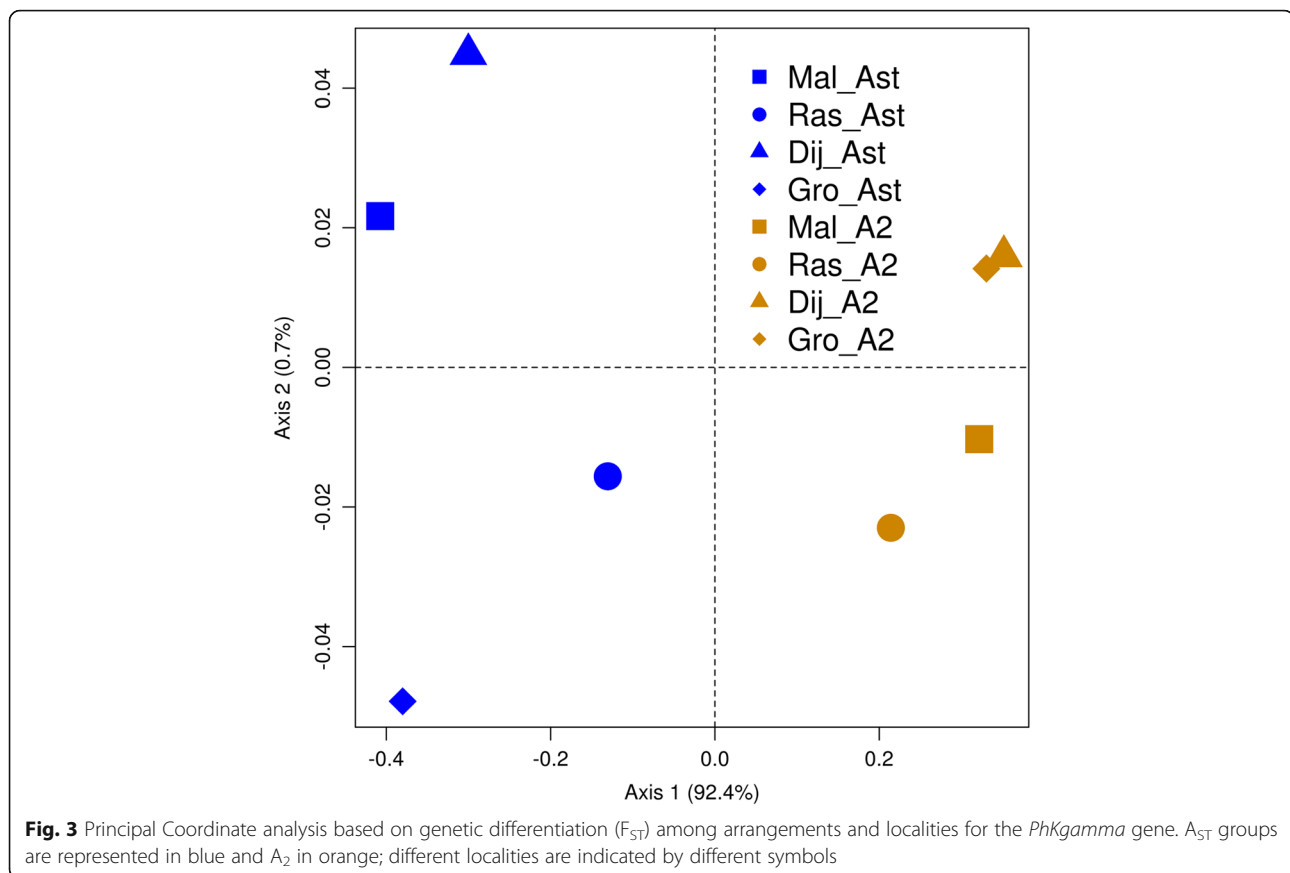
Diversity values were in general quite similar between localities for the same gene and chromosomal arrangement (see Additional file 3). Higher nucleotide diversity was found in A_{ST} than A_2 arrangements at all gene regions (Table 2 and Additional file 3), with significant differences when including information from all regions and localities for all diversity measures (Wilcoxon matched pairs test; $Z = 3.06$, $P = 0.002$ for π and π_{sil} ; $Z = 2.93$, $P = 0.003$ for S ; $Z = 2.98$, $P = 0.003$ for Θ_{sil}). No differences in the divergence per silent site values (K_{sil}) relative to *D. pseudoobscura* were found between the two arrangements (Wilcoxon matched pairs test; $Z = 0.24$, $P = 0.814$). Higher nucleotide diversity was found for

Hfw and *Nipsnap* (Table 2 and Additional file 3), which might possibly indicate a higher substitution rate in these two more distally located regions.

High significant genetic differentiation between the A_{ST} and A_2 arrangements was detected within each locality and also between localities for *Ubc-E2H* and *PhKgamma* (see Figs. 3 and 4 and Additional file 2) and also the concatenated dataset of all genes (see Additional file 4). On the other hand, *Hfw* and *Nipsnap* did not show significant differentiation between arrangements whether from the same or different localities (see Additional file 4). In general, no significant differentiation was observed when analysing groups of individuals carrying the same arrangement across localities with the exception of the comparison of A_{ST} individuals from Málaga and Rasquera in *PhKgamma* (see Additional file 2 and Additional file 4).

As low levels of within-population variation – as is the case in *PhKgamma* and *Ubc-E2H* – can inflate F_{ST} estimates [63, 64] we have used D_{xy} , which is not affected by such low variation. The results for each gene region were qualitatively similar to those reported above for F_{ST} , with consistently higher between-arrangement than within-arrangement differentiation for *PhKgamma* and *Ubc-E2H*, although the magnitude of such differences was lower than that obtained with F_{ST} estimates (Additional file 2). However, both statistics are highly correlated as shown by a Mantel test for *PhKgamma* ($r^2 = 0.754$, $P = 0.003$) and *Ubc-E2H* ($r = 0.741$, $P = 0.020$).

Analysis of genetic differentiation, done separately in coding (exon) and non-coding (non-exon) regions for the four genes concatenated data set for Málaga and Groningen (the two geographically distant localities), showed that the significant differentiation between groups of individuals carrying different arrangements was due solely to variation in non-coding regions (see Additional file 5). Interestingly, these analyses indicated low but significant differentiation for A_{ST} individuals from different locations in non-coding regions, but not for A_2 individuals. Considering each gene region, we found high differentiation between arrangements in upstream regions for *Ubc-E2H* and both upstream and intronic regions for *PhKgamma* while no relevant differentiation was observed for the other two genes in any region analysed (Fig. 5, see also Additional file 5). Curiously, in the intronic region of *PhKgamma* significant genetic differentiation between A_{ST} individuals from different localities was observed (see Additional file 5). A locus-by-locus analysis of Molecular variance (AMOVA) was performed to test the genetic differentiation (F_{CT}) between arrangements at each specific polymorphic site using a hierarchical structure (see Methods). Again, differentiation between arrangements was observed exclusively in non-coding regions (see Additional file 6).



On average F_{ST} values between arrangements found in the same locality or in different localities showed similar values in both *PhKgamma* and *Ubc-E2H* and the concatenated dataset (see Additional file 2 and Additional file 4; see also Figs. 3 and 4). In these regions, differentiation between arrangements was similar across localities with the exception of Rasquera for *PhKgamma* (see Additional file 7). However, this exception was due to the recombinant individuals from Rasquera (four individuals presenting gene conversion tracks, see below).

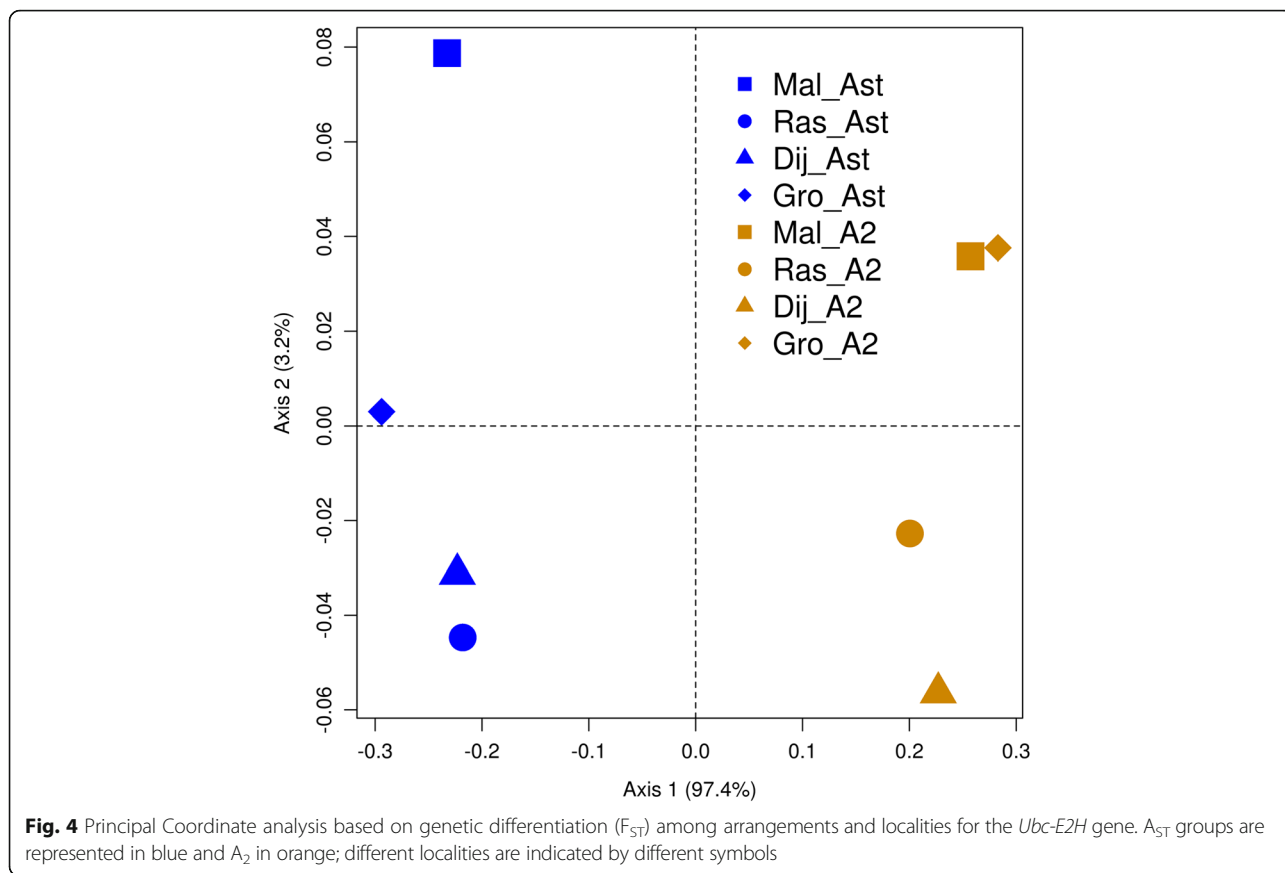
Pairwise genetic distances, based on the concatenated dataset of *PhKgamma* and *Ubc-E2H*, were correlated with their geographic distance with Mantel tests. No significant association between genetic and geographic distances was found for individuals carrying A_2 ($r^2 = 0.004$, $P = 0.462$), or A_{ST} ($r^2 = 0.017$, $P = 0.284$). Similar results are also obtained for the two genes separately (see Figs. 3 and 4). In agreement with this, no particular haplotype in any arrangement showed consistent indications of directional changes in frequency with latitude.

Gene conversion tracts (GCT) were detected only in *PhKgamma* and in four individuals, all of them from Rasquera. When removing individuals presenting GCTs in the *PhKgamma* gene, four fixed differences were detected between arrangements (position 205 in the

5'UTR, and positions 358–360 in the intron). For *Ubc-E2H*, one fixed difference was found between the two arrangements (position 60 in the 5'UTR). No fixed differences were found in the other two genes.

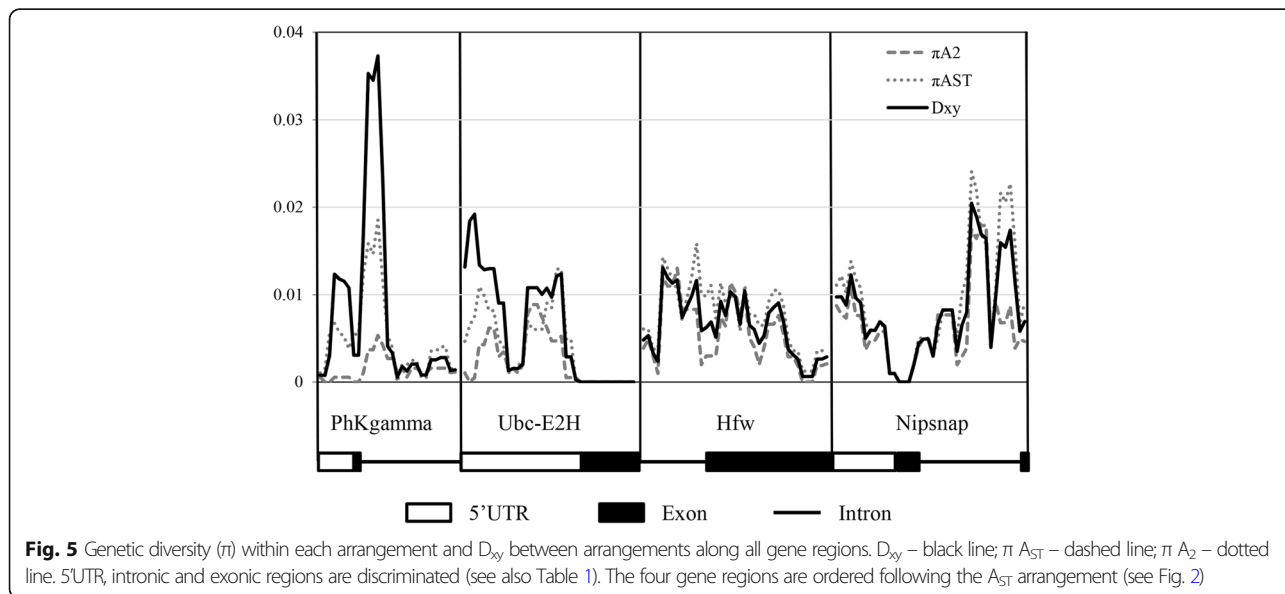
Linkage disequilibrium

In general, levels of linkage disequilibrium (LD) were low in all gene regions and arrangements analysed (see Additional file 8). Using the concatenated dataset of all four genes and considering both arrangements, LD between pairs of polymorphic sites (R^2) showed only 35 significant comparisons out of a total of 2415 ($P < 0.001$, chi-square with Bonferroni correction). Importantly, the majority of the significant LD comparisons were observed between sites from the same gene region, indicating short-range LD (see also Additional file 9). In fact, only 12 of the 35 significant comparisons corresponded to LD between distinct gene regions - 6 of them between *PhKgamma* and *Ubc-E2H* (namely between positions 358–360 bp from *PhKgamma* and position 60 bp from *Ubc-E2H*, see below). In the analysis including only the individuals carrying the A_2 inversion just 3 out of a total of 496 comparisons indicated significant LD after Bonferroni correction ($P < 0.001$, chi-square with Bonferroni correction). However, in these three cases, the significant LD



comparisons were always between sites within the same gene region (see Additional file 9). No significant comparisons were found between polymorphic sites for A_{ST} individuals. Importantly, these results show that within-arrangement LD was not observed across populations.

However, fixed associations (complete LD) were found between arrangements and sites in *PhKgamma* (positions 205, 358–360 of the total alignment with *D. pseudoobscura*; $P < 0.001$, chi-square with Bonferroni correction) and also in *Ubc-E2H* (position 60; $P < 0.001$, chi-square with



Bonferroni correction). Interestingly, all positions in *PhKgamma* are also in complete LD among themselves ($P < 0.001$, chi-square with Bonferroni correction). All these sites are located in non-coding/intronic regions. The fixed position in the *Ubc-E2H* gene is located 547 bp upstream from the first exon. The fixed positions at 358–360 bp in the *PhKgamma* gene are located in an intronic region just 26 bp downstream of the first exon. Comparison against *D. pseudoobscura* orthologous gene indicates more similarity to A_{ST} sequences. The fixed position between arrangements at 205 bp is within the 5'UTR region 36 bp upstream of the start coding site of *PhKgamma*, and possibly within the core promoter region, despite the absence of characteristic sequence motifs such as TATA box and the TFIIB recognition element [65].

Neutrality tests

All Tajima's D values were negative for each gene, arrangement and locality, and also for all locations combined (see Additional file 10). Tajima's D values were generally more negative in the A_2 arrangement. R_2 values also indicated stronger deviations from neutrality in the A_2 arrangement (see Additional file 10). The general scenario points to an excess of low frequency variants, consistent with a population expansion scenario, particularly in the A_2 arrangement.

Discussion

A_{ST} and A_2 are respectively a cold and a warm-adapted arrangement: while A_{ST} shows increased frequencies towards higher, colder latitudes, the opposite pattern is found for A_2 [32, 44], with both gene flow and local adaptation likely implicated [7]. In this study we analysed patterns of genetic variation associated with these arrangements and searched for patterns of geographical selection at the genetic level in candidate genes for thermal adaptation [40], located within or near these inverted regions.

Our study shows that these arrangements shape the patterns of genetic variation likely by reducing recombination between certain regions within chromosomes. In fact, the two proximally located regions in A_{ST} (*PhKgamma* and *Ubc-E2H*) presented significant differentiation between arrangements particularly in non-coding regions, regardless of the arrangement's geographical origin. Several molecular studies in *Drosophila* have documented significant genetic differentiation between chromosomes carrying different arrangements, particularly within inverted regions [22, 23, 25, 29, 33, 66] and also specifically near inversion breakpoints (e.g. [14, 21, 23]). The differentiation between arrangements that we observe could result from locally adapted alleles protected within the inversions which would be in agreement with the expectations of both

the coadaptation and local adaptation hypotheses [17, 18]. Furthermore, the fact that these differences are observed in 5'UTR and intronic regions might indicate that regulatory regions could be mostly the targets of selection (see [67]). In addition, the low differentiation between arrangements at coding regions - and the low variability observed - might be an effect of purifying selection acting in these regions. However, interpreting these patterns as being due to local adaptation is not straightforward, as neutral processes might also be involved. Non-adaptive associations between the inversion and neutral genetic variants generated during the origin and/or initial spread of the inversion could lead to longstanding differentiation and linkage disequilibrium [68], particularly near breakpoints where recombination is expectedly very low [19, 69]. This signal is more likely to be maintained when population size is high and the inversion is young - see [19]. In our case, despite the high effective size of the species [70] the age of A_2 - around 160,000 years [23], makes this expectation less likely. Moreover, the fact that microsatellite loci located in the same (proximal) regions as *PhKgamma* and *Ubc-E2H* did not show similar patterns [32] - namely the observed depletion in genetic variation and increase in differentiation between arrangements - further suggests that selection might be involved in the regions analysed here.

In the specific context of the A_{ST}/A_2 complex, Nóbrega et al. [23] also found that the extent of genetic exchange between arrangements was limited, and not sufficient to homogenize nucleotide variation in the centre of A_2 . In our study, we show that the reduction of recombination is not homogeneous in regions around the vicinity of the breakpoints of the A_{ST}/A_2 complex, with significant genetic differentiation and linkage disequilibrium between inversions occurring only in the proximal region of A_{ST} . Interestingly, linkage disequilibrium occurred in one region (*PhKgamma*) located outside the inversion, although near the breakpoint. Other studies have suggested that suppressed recombination can extend to 2.5 Mb or more outside inversion breakpoints [25, 71]. Functional analyses might elucidate if the non-coding variation regions found in *PhKgamma* and *Ubc-E2H* is in fact adaptive.

We did not find evidence for geographical (i.e. climatic) selection within arrangements, as very low within-arrangement genetic differentiation was detected across localities. It is possible that the differences in gene expression described previously for these genomic regions between thermal regimes [40] were controlled by other genes/sequences besides the upstream and coding regions analysed here. Importantly, the differential expression [40] was observed at the population level - between populations evolving under cold (13 °C) vs warm (22 °C) regimes - and not addressed at the karyotype

level. It is possible that the gene expression in a particular karyotype does not present clinal variation, with differential gene expression occurring between distinct karyotypes. Most interestingly, the expression of *PhKgamma* and *Nipsnap* have shown to be positively associated to latitude in semi natural populations of *D. subobscura* from a similar geographic gradient, but not when the same populations were maintained in a common garden [72]. Thus, although evidence is not conclusive, these gene regions could be involved in local adaptation to temperature.

Other studies in *D. subobscura* found very limited evidence for geographical molecular variation within inversions. Pegueroles et al. [33] also analysed candidate genes for thermal adaptation located in the O chromosome of *D. subobscura* and found no significant differentiation within each arrangement for individuals sampled in Barcelona (Spain) and Mount Parnes (Greece). Evidence for similar patterns of generally low within-arrangement differentiation was also described in other *D. subobscura* studies [73, 74] as well as studies in *D. melanogaster* ([31], but see [75]) and *D. pseudoobscura* [20]. However, recent evidence in *D. melanogaster* indicates several instances of genetic clinal variation in regions linked with inversions particularly associated with the *In(3R)Payne* [6, 34], perhaps due to the higher resolution power of these later studies in terms of the genomic scan performed and number of sampling sites studied.

Within-arrangement differentiation across localities would be expected if different epistatic interactions evolved (and eventually got fixed) in different localities due to adaptation to the specific local conditions as proposed by the Dobzhansky's coadaptation model [17]. The lack of within-arrangement differentiation of the selected genes across populations suggests that population-specific epistatic interactions are not occurring in the candidate genes for thermal adaptation that we sampled. It is of course possible that other loci within the inversion might present such varying epistatic interactions in the different localities. Some studies have reported evidence of epistasis based on patterns of low LD between neighbouring sites and high LD between distant sites within the inversion [20, 31]. Recently, indications of epistatic selection have also been obtained from observations of regions with significant genetic differentiation outside arrangements presenting strong linkage disequilibrium with the arrangements. In *Atpa* gene, another candidate gene for thermal adaptation, Pegueroles et al. [30] found strong genetic differentiation associated to the analysed arrangements of the O chromosome despite this gene being located outside them. Higher resolution genomic scans of adaptive variation as well as functional analyses are needed to conclusively assess the presence of relevant epistatic interactions and if these differ across populations.

Conclusions

Our genetic analysis of candidate loci for thermal adaptation [40] across geographically distinct *D. subobscura* localities of the European cline found very low within-arrangement differentiation across localities as well as no clinal patterns of genetic variation. This indicates that gene flow across the European continent has homogenized the genetic content of these potentially selected regions at a wide geographical scale. Given the strong effect of gene flow in this species, it is possible that considerable within-arrangement differentiation at a geographical scale is only maintained at relatively few loci subjected to stronger selective pressures in the different environments. In this scenario, one can argue that restricted gene flow cannot explain the observed inversion frequency clines and that local adaptation is maintaining the clinal variation in this species, counteracting the homogenizing effects of the high gene flow. We also found an important effect of the A_{ST}/A_2 complex in shaping genetic variation and preventing relevant gene exchange in different genomic regions, although its impact was not uniform across the different regions analysed. More generally, our study adds to a body of evidence showing that inversions act as barriers to recombination even outside inversions and can potentially have a role in the spread of favourable genic combinations and consequent adaptation to local habitats.

Additional files

Additional file 1: New hybridization location for *Nipsnap* gene. (PDF 108 kb)

Additional file 2: Measures of genetic differentiation within and between arrangements for *PhKgamma* and *Ubc-E2H*. A) Genetic differentiation (F_{ST}) between all localities and arrangements in *PhKgamma* (below diagonal) and *Ubc-E2H* (above diagonal). B) Genetic differentiation (D_{xy}) between all localities and arrangements in *PhKgamma* (below diagonal) and *Ubc-E2H* (above diagonal). (XLSX 12 kb)

Additional file 3: Nucleotide diversity (π) for all genes, arrangements and localities. (PDF 178 kb)

Additional file 4: Genetic differentiation (F_{ST}) between Malaga and Groningen in *Hfw*, *Nipsnap* and the concatenated dataset (four loci). (XLSX 10 kb)

Additional file 5: Genetic differentiation (F_{ST}) for coding and non-coding regions in the concatenated dataset for all 4 genes (A) and for coding, intronic and 5' UTR regions for *PhKgamma* (B). (XLSX 13 kb)

Additional file 6: Genetic differentiation between arrangements (F_{CT}) for all genes. 5'UTR, intronic and exonic regions are discriminated (see also Table 1). The four gene regions depicted are not contiguous. (PDF 238 kb)

Additional file 7: Genetic differentiation between arrangements for each locality in *PhKgamma* (A) and *Ubc-E2H* (B). (PDF 344 kb)

Additional file 8: Linkage Disequilibrium parameters for all regions and the concatenated dataset. (XLSX 10 kb)

Additional file 9: Linkage disequilibrium (LD) between pairs of polymorphic sites (R^2) against the nucleotide distance between compared sites. Significant comparisons after Bonferroni correction are presented in red (see Methods for additional details). A) LD between arrangements; B) LD within A_{ST} ; C) LD within A_2 . (PDF 248 kb)

Additional file 10: Neutrality tests and test of population expansion for each arrangement and gene. (XLSX 36 kb)

Abbreviations

AMOVA: Analysis of Molecular Variance; FDR: False discovery rate; GCT: Gene conversion track; LD: Linkage disequilibrium; PCR: Polymerase chain reaction; UTR: Untranslated region

Acknowledgments

The authors thank Margarida Matos, Inês Fragata and Josiane Santos for helpful discussions and suggestions. We also thank Gemma Calabria, João Picão-Osório and Olga Dolgova for help in inversion polymorphism scoring, Neus Perez for help in the genetic analyses and Montse Peiró for help in the in situ hybridizations.

Funding

This research was funded by projects CTM2013–48163 and CTM2017–88080 (AEI/FEDER, UE) of the Spanish Government. PS had a FCT ('Fundação para a Ciência e Tecnologia') post-doctoral grant (SFRH/BPD/36829/2007) and has currently another post-doctoral Grant from FCT (SFRH/BPD/86186/2012). MP is part of the research group 2017SGR1120 of the Generalitat de Catalunya. The funding body had no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Availability of data and materials

Data supporting the conclusions of this article are included within the article and in its additional files. All gene sequences are deposited in GenBank with accession numbers: MG877856 - MG877894 (*Nipsnap*), MG877895 - MG877930 (*Hfw*), MG877931 - MG878006 (*Phkgamma*) and MG878007 - MG878083 (*Ubc-E2H*).

Authors' contributions

PS participated in the experiment design, sample and data collection, data analysis and writing of the manuscript. MP participated in the experiment design, sample collection and revised the manuscript. Both authors have read and approved the final version of the manuscript.

Ethics approval and consent to participate

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 26 July 2017 Accepted: 11 April 2018

Published online: 24 April 2018

References

- Hoffmann AA, Rieseberg LH. Revisiting the impact of inversions in evolution: from population genetic markers to drivers of adaptive shifts and speciation? *Annu Rev Ecol Syst.* 2008;39:21–42.
- Kirkpatrick M. How and why chromosome inversions evolve. *PLoS Biol.* 2010;8(9):e1000501.
- Prevosti A, Ribo G, Serra L, Aguade M, Balanyà J, Monclus M, et al. Colonization of America by *Drosophila subobscura*: experiment in natural populations that supports the adaptive role of chromosomal-inversion polymorphism. *Proc Natl Acad Sci U S A.* 1988;85:5597–600.
- Umina PA, Weeks AR, Kearney MR, McKechnie SW, Hoffmann AA. A rapid shift in a classic clinal pattern in *Drosophila* reflecting climate change. *Science.* 2005;308:691–3.
- Balanyà J, Oller JM, Huey RB, Gilchrist GW, Serra L. Global genetic change tracks global climate warming in *Drosophila subobscura*. *Science.* 2006;313:1773–5.
- Kapun M, Fabian DK, Goudet J, Flatt T, Flatt T. Genomic evidence for adaptive inversion clines in *Drosophila melanogaster*. *Mol Biol Evol.* 2016;33:1317–36.
- Rezende E, Balanyà J, Rodríguez-Trelles F, Rego C, Fragata I, Matos M, et al. Climate change and chromosomal inversions in *Drosophila subobscura*. *Clim Res.* 2010;43:103–14.
- Ayala D, Ullastres A, Gonzalez J. Adaptation through chromosomal inversions in *Anopheles*. *Front Genet.* 2014;5:129.
- Lowry DB, Willis JH. A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biol.* 2010;8(9):e1000500.
- Simões P, Fragata I, Lopes-Cunha M, Lima M, Kellen B, Bárbaro M, et al. Wing trait-inversion associations in *Drosophila subobscura* can be generalized within continents, but may change through time. *J Evol Biol.* 2015;28:2163–74.
- Hoffmann AA, Weeks AR. Climatic selection on genes and traits after a 100 year-old invasion: a critical look at the temperate-tropical clines in *Drosophila melanogaster* from eastern Australia. *Genetica.* 2007;129:133–47.
- Kennington WJ, Hoffmann AA, Partridge L. Mapping regions within cosmopolitan inversion *In(3R)Payne* associated with natural variation in body size in *Drosophila melanogaster*. *Genetics.* 2007;177:549–56.
- Rako L, Anderson AR, Sgrò CM, Stocker AJ, Hoffmann AA. The association between inversion *In(3R)Payne* and clinally varying traits in *Drosophila melanogaster*. *Genetica.* 2006;128:373–84.
- Kapun M, Van Schalkwyk H, McAllister B, Flatt T, Schlötterer C. Inference of chromosomal inversion dynamics from pool-Seq data in natural and laboratory populations of *Drosophila melanogaster*. *Mol Ecol.* 2014;23:1813–27.
- Fragata I, Lopes-Cunha M, Bárbaro M, Kellen B, Lima M, Santos MA, et al. How much can history constrain adaptive evolution? A real-time evolutionary approach of inversion polymorphisms in *Drosophila subobscura*. *J Evol Biol.* 2014;27:2727–38.
- Santos J, Pascual M, Fragata I, Simões P, Santos MA, Lima M, et al. Tracking changes in chromosomal arrangements and their genetic content during adaptation. *J Evol Biol.* 2016;29:1151–67.
- Dobzhansky T. Genetics of natural populations. XIX. Origin of heterosis through natural selection in populations of *Drosophila pseudoobscura*. *Genetics.* 1950;35:288–302.
- Kirkpatrick M, Barton N. Chromosome inversions, local adaptation and speciation. *Genetics.* 2006;173:419–34.
- Guerrero RF, Rousset F, Kirkpatrick M. Coalescent patterns for chromosomal inversions in divergent populations. *Phil Trans R Soc B.* 2012;367:430–8.
- Schaeffer SW, Goetting-Minesky MP, Kovacevic M, Peoples JR, Graybill JL, Miller JM, et al. Evolutionary genomics of inversions in *Drosophila pseudoobscura*: evidence for epistasis. *Proc Natl Acad Sci U S A.* 2003;100:8319–24.
- Laayouni H, Hasson E, Santos M, Fontdevila A. The evolutionary history of *Drosophila buzzatii*. XXXV. Inversion polymorphism and nucleotide variability in different regions of the second chromosome. *Mol Biol Evol.* 2003;20:931–44.
- Schaeffer SW, Anderson WW. Mechanisms of genetic exchange within the chromosomal inversions of *Drosophila pseudoobscura*. *Genetics.* 2005;171:1729–39.
- Nóbrega C, Khadem M, Aguadé M, Segarra C. Genetic exchange versus genetic differentiation in a medium-sized inversion of *Drosophila*: the A_{27}/A_{57} arrangements of *Drosophila subobscura*. *Mol Biol Evol.* 2008;25:1534–43.
- Corbett-Detig RB, Hartl DL. Population genomics of inversion polymorphisms in *Drosophila melanogaster*. *PLoS Genet.* 2012;8(12):e1003056.
- Mcgaugh SE, Noor MAF. Genomic impacts of chromosomal inversions in parapatric *Drosophila* species. *Phil Trans R Soc B.* 2012;367:422–9.
- Cheng C, White BJ, Kamdem C, Mockaitis K, Costantini C, Hahn MW, et al. Ecological genomics of *Anopheles gambiae* along a latitudinal cline: a population-resequencing approach. *Genetics.* 2012;190:1417–32.
- Stump AD, Pombi M, Goeddel L, Ribeiro JMC, Wilder JA, Torre AD, et al. Genetic exchange in *2La* inversion heterokaryotypes of *Anopheles gambiae*. *Insect Mol Biol.* 2007;16:703–9.
- White BJ, Cheng C, Sangaré D, Lobo NF, Collins FH, Besansky NJ. The population genomics of trans-specific inversion polymorphisms in *Anopheles gambiae*. *Genetics.* 2009;183:275–88.
- Munté A, Rozas J, Aguadé M, Segarra C. Chromosomal inversion polymorphism leads to extensive genetic structure: a multilocus survey in *Drosophila subobscura*. *Genetics.* 2005;169:1573–81.
- Pegueroles C, Ferrés-Coy A, Martí-Solano M, Aquadro CF, Pascual M, Mestres F. Inversions and adaptation to the plant toxin ouabain shape DNA sequence variation within and between chromosomal inversions of *Drosophila subobscura*. *Sci Rep.* 2016;6:23754.
- Kennington WJ, Partridge L, Hoffmann AA. Patterns of diversity and linkage disequilibrium within the cosmopolitan inversion *In(3R)Payne* in *Drosophila melanogaster* are indicative of coadaptation. *Genetics.* 2006;172:1655–63.
- Simões P, Calabria G, Picão-Osório J, Balanyà J, Pascual M. The genetic content of chromosomal inversions across a wide latitudinal gradient. *PLoS One.* 2012;7:e51625.
- Pegueroles C, Aquadro CF, Mestres F, Pascual M. Gene flow and gene flux shape evolutionary patterns of variation in *Drosophila subobscura*. *Heredity.* 2013;110:520–9.
- Rane RV, Rako L, Kapun M, Lee SF, Hoffmann AA. Genomic evidence for role of inversion *3RP* of *Drosophila melanogaster* in facilitating climate change adaptation. *Mol Ecol.* 2015;24:2423–32.

35. Rodríguez-Trelles F, Tarrío R, Santos M. Genome-wide evolutionary response to a heat wave in *Drosophila*. *Biol Lett*. 2013;9:20130228.
36. Rego C, Balanya J, Fragata I, Matos M, Rezende E, Santos M. Clinal patterns of chromosomal inversion polymorphisms in *Drosophila subobscura* are partly associated with thermal preferences and heat stress resistance. *Evolution*. 2010;64:385–97.
37. Dolgova O, Rego C, Calabria G, Balanya J, Pascual M, Rezende E, et al. Genetic constraints for thermal coadaptation in *Drosophila subobscura*. *BMC Evol Biol*. 2010;10:363.
38. Calabria G, Dolgova O, Rego C, Castañeda LE, Rezende EL, Balanya J, et al. Hsp70 protein levels and thermotolerance in *Drosophila subobscura*: a reassessment of the thermal co-adaptation hypothesis. *J Evol Biol*. 2012;25:691–700.
39. Fragata I, Balanya J, Rego C, Matos M, Rezende EL, Santos M. Contrasting patterns of phenotypic variation linked to chromosomal inversions in native and colonizing populations of *Drosophila subobscura*. *J Evol Biol*. 2010;23:112–23.
40. Laayouni H, García-Franco F, Chávez-Sandoval BE, Trotta V, Beltran S, Corominas M, et al. Thermal evolution of gene expression profiles in *Drosophila subobscura*. *BMC Evol Biol*. 2007;7:42.
41. Fuller ZL, Haynes GD, Richards S, Schaeffer SW. Genomics of natural populations: how differentially expressed genes shape the evolution of chromosomal inversions in *Drosophila pseudoobscura*. *Genetics*. 2016;204:287–301.
42. Pascual M, Aquadro CF, Soto V, Serra L. Microsatellite variation in colonizing and paleartic populations of *Drosophila subobscura*. *Mol Biol Evol*. 2001;18:731–40.
43. Pascual M, Chapuis MP, Mestres F, Balanya J, Huey RB, Gilchrist GW, et al. Introduction history of *Drosophila subobscura* in the new world: a microsatellite-based survey using ABC methods. *Mol Ecol*. 2007;16:3069–83.
44. Balanya J, Serra L, Gilchrist GW, Huey RB, Pascual M, Mestres F, et al. Evolutionary pace of chromosomal polymorphism in colonizing populations of *Drosophila subobscura*: an evolutionary time series. *Evolution*. 2003;57:1837–45.
45. Calabria G, Máca J, Bachli G, Serra L, Pascual M. First records of the potential pest species *Drosophila suzukii* (Diptera: Drosophilidae) in Europe. *J Appl Entomol*. 2012;136:139–47.
46. Huey RB, Pascual M. Partial thermoregulatory compensation by a rapidly evolving invasive species along a latitudinal cline. *Ecology*. 2009;90:1715–20.
47. Krimbas C. The inversion polymorphism of *Drosophila subobscura*. In: Krimbas C, Powell J, editors. *Drosophila inversion polymorphism*. Boca Raton (FL): CRC press; 1992. p. 127–220.
48. Pascual M, Balanya J, Latorre A, Serra L. Analysis of the variability of *Drosophila azteca* and *D. athabasca* populations revealed by randomly amplified polymorphic DNA. *J Zool Syst Evol Res*. 1997;35:159–64.
49. Laayouni H, Santos M, Fontdevila A. Toward a physical map of *Drosophila buzzatii*: use of randomly amplified polymorphic DNA polymorphisms and sequence tagged site landmarks. *Genetics*. 2000;156:1797–816.
50. Kunze-Mühl E, Müller E. Weitere Untersuchungen über die chromosomale Struktur und die natürlichen Strukturtypen von *Drosophila subobscura* coll. *Chromosoma*. 1957;9:559–70.
51. Pegueroles C, Araúz PA, Pascual M, Mestres F. A recombination survey using microsatellites: the O chromosome of *Drosophila subobscura*. *Genetica*. 2010;138:795–804.
52. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symp*. 1999;41:95–8.
53. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*. 2009;25:1451–2.
54. Nei M, Gojobori T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol*. 1986;3:418–26.
55. Hudson R, Slatkin M, Maddison W. Estimation of levels of gene flow from DNA sequence data. *Genetics*. 1992;132:583–9.
56. Nei M. *Molecular evolutionary genetics*. New York: Columbia University Press; 1987.
57. Hudson R. A new statistic for detecting genetic differentiation. *Genetics*. 2000;155:2011–4.
58. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *Ann Stat*. 2001;29:1165–88.
59. Excoffier L, Lischer H. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and windows. *Mol Ecol Resour*. 2010;10:564–7.
60. Betrán E, Rozas J, Navarro A, Barbadilla A. The estimation of the number and the length distribution of gene conversion tracts from population DNA sequence data. *Genetics*. 1997;146:89–99.
61. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*. 1989;123:585–95.
62. Ramos-Onsins SE, Rozas J. Statistical properties of new neutrality tests against population growth. *Mol Biol Evol*. 2002;19:2092–100.
63. Charlesworth B. Measures of divergence between populations and the effect of forces that reduce variability. *Mol Biol Evol*. 1998;15:538–43.
64. Cruickshank TE, Hahn MW. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol Ecol*. 2014;23:3133–57.
65. Butler JEF, Kadonaga JT. The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev*. 2002;16:2583–92.
66. Corbett-Detig RB, Cardeno C, Langley CH. Sequence-based detection and breakpoint assembly of polymorphic inversions. *Genetics*. 2012;192:131–7.
67. Palsson A, Wesolowska N, Reynisdóttir S, Ludwig MZ, Kreitman M. Naturally occurring deletions of hunchback binding sites in the even-skipped stripe 3+7 enhancer. *PLoS One*. 2014;9:e91924.
68. Ishii K, Charlesworth B. Associations between allozyme loci and gene arrangements due to hitch-hiking effects of new inversions. *Genet Res*. 1977;30:93–106.
69. Navarro A, Barbadilla A, Ruiz A. Effect of inversion polymorphism on the neutral nucleotide variability of linked chromosomal regions in *Drosophila*. *Genetics*. 2000;155:685–98.
70. Pascual M, Schug MD, Aquadro CF. High density of long dinucleotide microsatellites in *Drosophila subobscura*. *Mol Biol Evol*. 2000;17:1259–67. <https://doi.org/10.1093/oxfordjournals.molbev.a026409>.
71. Pegueroles C, Ordóñez V, Mestres F, Pascual M. Recombination and selection in the maintenance of the adaptive value of inversions. *J Evol Biol*. 2010;23:2709–17.
72. Porcelli D, Westram AM, Pascual M, Gaston KJ, Butlin RK, Snook RR. Gene expression clines reveal local adaptation and associated trade-offs at a continental scale. *Sci Rep*. 2016;6:32975.
73. Rozas J, Segarra C, Zapata C, Alvarez G, Aguadé M. Nucleotide polymorphism at the rp49 region of *Drosophila subobscura*: lack of geographic subdivision within chromosomal arrangements in Europe. *J Evol Biol*. 1995;367:355–67.
74. Sánchez-Gracia A, Rozas J. Molecular population genetics of the OBP83 genomic region in *Drosophila subobscura* and *D. guanche*: contrasting the effects of natural selection and gene arrangement expansion in the patterns of nucleotide variation. *Heredity*. 2011;106:191–201.
75. Kennington WJ, Hoffmann AA. Patterns of genetic variation across inversions: geographic variation in the In(2L)t inversion in populations of *Drosophila melanogaster* from eastern Australia. *BMC Evol Biol*. 2013;13:100.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

