RESEARCH ARTICLE   Open Access

Basal *hsp70* expression levels do not explain adaptive variation of the warm- and cold-climate O_{3+4+7} and O_{ST} gene arrangements of *Drosophila subobscura*

Marta Puig Giribets, Mauro Santos and María Pilar García Guerreiro[*] ![ORCID]

Abstract

Background: *Drosophila subobscura* exhibits a rich inversion polymorphism, with some adaptive inversions showing repeatable spatiotemporal patterns in frequencies related to temperature. Previous studies reported increased basal HSP70 protein levels in homokaryotypic strains for a warm-climate arrangement compared to a cold-climate one. These findings do not match the similar *hsp70* genomic organization between arrangements, where gene expression levels are expected to be similar. In order to test this hypothesis and understand the molecular basis for *hsp70* expression, we compared basal *hsp70* mRNA levels in males and females, and analysed the 5′ and 3′ regulatory regions of *hsp70* genes in warm- and cold-climate isochromosomal O_{3+4+7} and O_{ST} lines of *D. subobscura*.

Results: We observed comparable mRNA levels between the two arrangements and a sex-biased *hsp70* gene expression. The number of heat-shock elements (HSEs) and GAGA sites on the promoters were identical amongst the O_{ST} and O_{3+4+7} lines analysed. This is also true for 3′ AU-rich elements where most A and B copies of *hsp70* have, respectively, two and one element in both arrangements. Beyond the regulatory elements, the only notable difference between both arrangements is the presence in 3′ UTR of a 14 bp additional fragment after the stop codon in the *hsp70*A copy in five O_{3+4+7} lines, which was not found in any of the six O_{ST} lines.

Conclusions: The equivalent *hsp70* mRNA amounts in O_{ST} and O_{3+4+7} arrangements provide the first evidence of a parallelism between gene expression and genetic organization in *D. subobscura* lines having these arrangements. This is reinforced by the lack of important differential features in the number and structure of regulatory elements between both arrangements, despite the genetic differentiation observed when the complete 5′ and 3′ regulatory regions were considered. Therefore, the basal levels of *hsp70* mRNA cannot account, in principle, for the adaptive variation of the two arrangements studied. Consequently, further studies are necessary to understand the intricate molecular mechanisms of *hsp70* gene regulation in *D. subobscura*.

Keywords: *Drosophila subobscura*, *hsp70* expression, Chromosomal arrangements, Thermal adaptation

* Correspondence: mariapilar.garcia.guerreiro@uab.es
Grup de Genòmica, Bioinformàtica i Biologia Evolutiva (GGBE), Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain

Background

Inversions can foster local adaptation as they protect chromosomal inverted sequences from recombination in heterokaryotypes. *Drosophila subobscura* exhibits a rich inversion polymorphism [1], with the O chromosome (Muller element E that also corresponds to chromosome arm 3R in *Drosophila melanogaster* ([2], p., 307) having the highest number of inversions that form complex gene arrangements of overlapping and non-overlapping inversions [3]. The overlapping inversions cluster in the distal segment of the chromosome and form the O_{3+4} phylad, whose different gene arrangements have higher frequencies at lower latitudes [1, 4]. Thus, the gene arrangements on chromosome O can be divided into two groups: group 1 (cold-climate) includes O_{ST} that shows a negative correlation coefficient with maximum temperatures along the latitudinal clines, or a positive correlation coefficient with latitude; and group 2 (warm-climate) that includes gene arrangements showing the reverse pattern (e.g., O_{3+4}, $O_{3+4} + 7$ and O_{3+4+8}). In agreement with the clinal patterns, warm-climate inversions increase in frequency during summer and decrease during winter [5, 6]. Furthermore, they show long-term trends of increasing frequency related to global warming [7, 8]. Overall, these spatiotemporal patterns strongly suggest that temperature is the main selective factor driving inversion frequencies, but the underlying genetic causes are still poorly understood.

The heat-shock HSP70 protein is involved in protein folding, assembly, and can be induced by several stressors such as a sudden increase in temperature [9, 10]. It appears to be the primary protein involved in thermotolerance in *D. melanogaster* [11, 12], and allele frequencies show latitudinal clines and change in response to thermal evolution in the laboratory [13]. In a previous work with *D. subobscura*, expression levels of HSP70 protein in flies carrying cold- and warm-climate gene arrangements were analysed [14]. It was found that homokaryotypic O_{3+4} flies had increased HSP70 basal protein levels that did not boost after a heat-shock, in contrast to what happened to their O_{ST} homokaryotypic counterparts. This result was somehow surprising because HSP70 protein expression entails fitness costs [15–17]. The interpretation was that low-latitude populations with relatively high O_{3+4} frequencies are more likely to be exposed to occasionally higher temperatures than populations located at higher latitudes, and the benefits of mounting the stress response might outweigh the costs. Actually, an extreme heat wave in Europe in spring 2011 caused frequencies of warm-climate $O_{3+4} + 7$ to shift

$+5.4\sigma$ from the expected average frequency (in close agreement with the temperature anomaly), and high selection coefficients are needed to explain this shift [18].

This interpretation is not without caveats. A trend to a decrease in basal HSP70 protein expression was observed in strains of *D. melanogaster* [19, 20] and *Drosophila buzzatii* [21] continually exposed to high temperatures, probably to avoid the deleterious effects of increased HSP70 levels. On the other hand, increased basal HSP70 protein and mRNA levels have been observed in certain strains of *Drosophila virilis* with increased thermotolerance [22]. Therefore, basal levels of HSP70 protein expression might be contingent on the thermal history and the species under study. A positive link between *hsp70* copy number, expression levels, and thermal resistance in artificial [23–25] and natural [26] *Drosophila* strains has been reported, but this does not apply to *D. subobscura*. In this species, we have recently shown that the genomic sequence organization of the *hsp70* gene family (located in the 94A subsection of the O chromosome, which is inside the region covered by O_3 and O_4 inversions [27];), is essentially similar in cold- (O_{ST}) and warm-climate (O_{3+4}, O_{3+4+8} and O_{3+4+16}) gene arrangements, and consists of only two genes, each with four heat-shock elements and three GAGA sites on its promoter [27]. To date, no clear evidence has been found to link the genomic sequence organization of the *hsp70* gene family with the different HSP70 protein expression patters previously detected [14].

Knowing that proteins are produced from mRNA which, in turn, depends on gene regulatory regions, we extended the former approach to gauge if *hsp70* mRNA levels are similar in cold- and warm-climate arrangements as expected by their similar genomic organization. We quantified basal *hsp70* mRNA expression levels in males and females, and included a higher number of isogenic O_{ST} and $O_{3+4} + 7$ strains than in our previous study [27]. We expanded the study to the 5′ and 3′ *hsp70* regulatory regions, which are known to be involved in transcriptome variation in *D. melanogaster* [28]. A clarification is in order. The arrangement $O_{3+4} + 7$ has the overlapping inversions O_{3+4} in the distal segment and the single O_7 inversion in the proximal segment of chromosome O. Because our lines derived from a natural population where both gene arrangements (O_{ST} and $O_{3+4} + 7$) have a relatively high frequency

(Berbikiz, north of Spain [18]) the heterokaryotype O_{3+4} /$O_{3+4} + 7$ freely recombines in the distal segment and, therefore, the gene content in the O_{3+4} region is expected to be the same in both arrangements. The reason for using $O_{3+4} + 7$ was because the integrity of the wild chromosome is better preserved during the isogenization process with the available balancer stock in *D. subobscura* (see below).

We found equivalent mRNA amounts in O_{ST} and $O_{3+4} + 7$ arrangements, a sex-biased *hsp70* mRNA gene expression compatible with the high degree of conservation of the cis-regulatory elements observed, and an identical genomic organization of the gene family previously reported in our study.
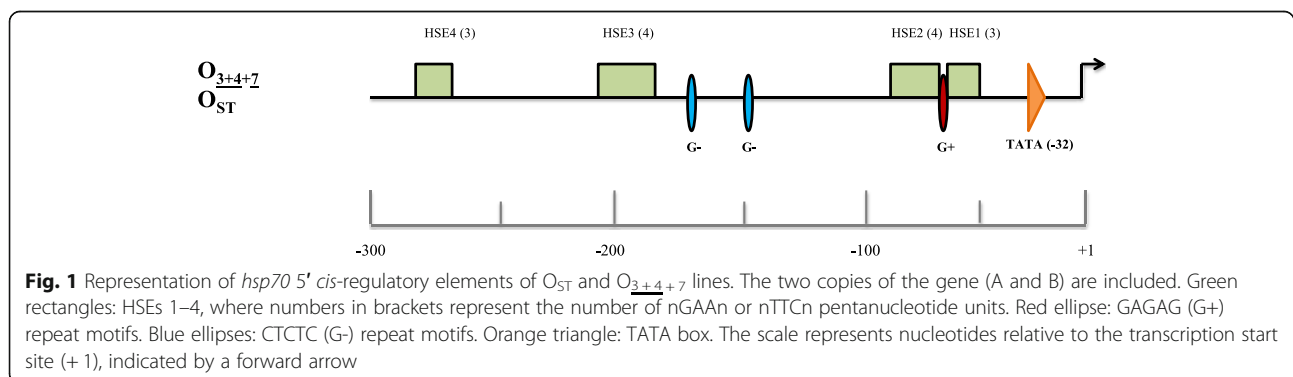
Results

High degree of conservation in the *hsp70* proximal and distal CRE elements between O_{ST} and $O_{3+4} + 7$ arrangements

We have sequenced the proximal and distal regulatory regions of the *hsp70* locus in 12 isogenic lines (6 lines for O_{ST} and 6 for $O_{3+4} + 7$ gene arrangements). The proximal 5′ region consists of a ~ 550 bp, of which ~ 290 bp correspond to the region upstream of the transcription start site (TSS) in the two *hsp70* gene copies (*hsp70A* and *hsp70B*). Multiple sequence alignments indicated high CRE (Cis-Regulatory Element) conservation between O_{ST} and $O_{3+4} + 7$. As previously found [27], four highly conserved heat shock elements (HSEs) were identified (Fig. 1), and no fixed nucleotide differences between arrangements were observed (Additional files 1 and 2). The genomic organization of HSEs resembles that of other *Drosophila* species [29] and consists of at least three continuous alternate nGAAn/nTTCn pentanucleotide subunits representing binding sites for the heat shock factor (HSF) monomer. Although the number of repetitions is identical in the 12 isogenic lines, it varies between HSE sequences: three in the most distal

and proximal (HSE1, HSE4) and four in HSE2 and HSE3 (Fig. 1). Within each HSE, polymorphisms appear to be overrepresented in the first and fifth position of nGAAn/nTTCn pentanucleotides, which are the least constrained [29]. Some insertion/deletion (INDEL) polymorphisms were also detected in the presumably unconstrained regions separating the more distal HSEs. These polymorphisms tend to be shared between the two *hsp70* paralogs of the same line as a result of ectopic gene conversion [27]. We did not observe differences between isogenic lines either in the two core proximal promoter elements (TATA box and TSS), or in the total number of GAGA sites in the proximal promoters. The latter is known to be recognized by the transcription factor GAF and to participate, like other promoters, in the regulation of global gene expression [30].

The distal regulatory region corresponds to the 3′ untranslated mRNA region (3'UTR), which includes a region spanning 198 ± 4 bp in all *hsp70A* and *hsp70B* copies, except in one *hsp70B* copy from O_{ST} (4*th* line) that is 381 bp long. The in silico screening let us identify conserved regulators consisting in AU-rich elements (AREs), known to be involved in the post-transcriptional regulation of gene expression in *Drosophila* [31] by recruiting ARE-binding proteins that trigger mRNA deadenylation and turnover. In most lines we identified two and one ARE motifs, consisting in a single AUUUA pentamer in *hsp70A* and *hsp70B* copies, respectively. There are a few exceptions; namely, the 1*st* and 6*th* O_{ST} lines that have an unique motif in *hsp70A* (Fig. 2 and Additional file 3) and the 4*th* O_{ST} line, where the 3'UTR region of the *hsp70B* copy is longer and has a second ARE motif (Fig. 2 and Additional file 4). We found that ARE motifs are, in general, highly conserved in both copies of O_{ST} and $O_{3+4} + 7$, but the *hsp70A* copy in five $O_{3+4} + 7$ lines has a 14 bp additional fragment after the stop codon, which was not found in O_{ST} lines (Additional file 3).

Table 1 summarizes the level of genetic differentiation between chromosomal rearrangements for the complete

Fig. 1 Representation of *hsp70* 5′ cis-regulatory elements of O_{ST} and O_{3+4+7} lines. The two copies of the gene (A and B) are included. Green rectangles: HSEs 1–4, where numbers in brackets represent the number of nGAAn or nTTCn pentanucleotide units. Red ellipse: GAGAG (G+) repeat motifs. Blue ellipses: CTCTC (G-) repeat motifs. Orange triangle: TATA box. The scale represents nucleotides relative to the transcription start site (+ 1), indicated by a forward arrow

Fig. 2 Schematic representation *hsp70* 3′ UTR region of O_{ST} and O_{3+4+7} lines. The two copies of the gene (A and B) are included. The region represented spans from the stop codon to the polyadenylation signal. Green rectangles: ARE motifs. In red: polyadenylation signal. Square brackets: number of nucleotides from the stop codon and between ARE motifs

nucleotide sequence of 5′ and 3′ regulatory regions of A and B copies. Regulatory regions in both copies, except 3′B, exhibited significant genetic differentiation between the two arrangements studied. However, whereas differentiation levels were similar in 5′ regulatory regions for A and B copies, they were significantly higher in 3′ regulatory regions from A copy, where the average number of differences per site (D_{xy}) was much higher than in the B copy.

Basal *hsp70* mRNA expression levels

hsp70 mRNA levels have been quantified in 67 (out of 72) *D. subobscura* samples consisting of 3 biological replicates per isogenic line (6 O_{ST} and 6 $O_{3+4} + 7$) and sex, except five lines where only two replicates for females and/or males could be analysed. The ANOVA is shown

in Table 2, and results are plotted in Fig. 3. Global *hsp70* mRNA levels were about the same between arrangements ($P = 0.473$) but were higher in females than in males ($P = 0.005$), particularly for the O_{3+4+7} gene arrangement; but the arrangement × sex interaction was not statistically significant ($P = 0.296$).

Discussion

Here we have sequenced the proximal and distal regulatory regions of the *hsp70* loci and have analysed mRNA expression levels in males and females from 6 isochromosomal lines of warm- ($O_{3+4} + 7$) and 6 isochromosomal lines of cold-climate (O_{ST}) gene arrangements in *D. subobscura*. Our aim was to shed some light on the basal amounts of *hsp70* mRNA in these arrangements that show an almost identical genetic organization in the

Table 1 Genetic differentiation between O_{3+4+7} and O_{ST} arrangements for 5′ and 3′ regulatory regions

Region	Comparisons	Gen copy	$SP_{ST}SF_{347}$	$SP_{347}SF_{ST}$	Shared	D_{xy}	F_{st}	P
5′	O_{ST} vs O_{3+4+7}	Conc	17	35	2	0.031	0.344	0.000***
		A (6)	13	33	2	0.034	0.306	0.008**
		B (6)	10	28	2	0.030	0.330	0.045*
3′	O_{ST} vs O_{3+4+7}	Conc	17	38	7	0.127	0.045	0.000***
		A (6)	16	5	0	0.075	0.713	0.007**
		B (6)	1	9	0	0.014	0.081	0.068

Sample sizes are in parenthesis. F_{st}: Proportion of nucleotide diversity due to differences between populations; D_{xy}: Average number of nucleotide substitutions per site between populations; $SP_{ST}SF_{347}$: Sites polymorphic in O_{ST} and fixed in O_{3+4+7}; $SP_{347}SF_{ST}$: Sites polymorphic in O_{3+4+7} and fixed in O_{ST}; Conc: concatened data set; Statistical significance was assessed using the *P*-value of Snn statistic; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Table 2 ANOVA for the effects of chromosome gene arrangements ($O_{\underline{3+4}+7}$ and O_{ST}), isochromosomal line, and sex on mRNA relative expression ($2^{-\Delta C_T}$). Data plotted in Fig. 3
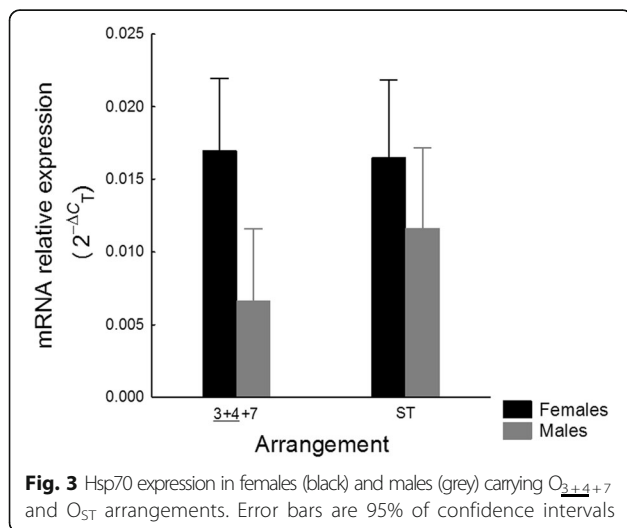
Source	d.f.	MS	F	P (parametric)	P (permutation)
Arrangement	1	0.000084	0.56	0.473	0.471
Line (Arrangement)	10	0.000152	1.36	0.223	0.222
Sex	1	0.000955	8.55	0.005**	0.004**
Arrangement × Sex	1	0.000125	1.11	0.296	0.294
Error	53	0.000112			

Arrangement was nested within lines. MS: Mean squares, d.f.: degrees of freedom; ** $P < 0.01$.

hsp70 locus [27]. We have now performed for the first time a thorough analysis of 5′ and 3′ *hsp70* regulatory regions together with the quantification of mRNA levels in O_{ST} and O_{3+4+7} arrangements. We found similar *hsp70* mRNA levels in both arrangements, in accordance with the identical gene configuration found in our previous work [27], but also found differences due to sex (Table 2). Since transcript levels are under the influence of 5′ and 3′ regulatory elements, we analysed these regions and found that the number of 5′ HSE elements in both $O_{3+4}+7$ and O_{ST} arrangements was also identical, as previously observed in other chromosomal arrangements as O_{3+4+8} and O_{3+4+16}, where these elements evolve in concert through gene conversion [27]. Analysis of 3′ UTR regions also showed no noticeable differences between chromosomal arrangements in the number of ARE elements, considered to be critical regulators of mRNA turnover [31]. However, differences were observed between paralogous genes: in most lines, two ARE motifs were found in *hsp70A* but only one in *hsp70B*, which could point to post-transcriptional regulation events responsible for RNA shortening and degradation. Interestingly, the 3′UTRs were already considered as more divergent between paralogs than the coding and proximal promoter regions [27].

Our main aim here was to correlate the genome organization of *hsp70* loci between cold- (O_{ST}) and a warm-climate ($O_{3+4}+7$) gene arrangements with mRNA basal expression patterns. The use of isochromosomal lines precludes any extrapolation to wild, outbred flies. mRNA and protein levels are not always correlated [32], and we did not subject our flies to a thermal stress because inbreeding can interfere with the heat-shock response [33]. Besides, inbred individuals might express conditional lethality under heat stress [34]. Therefore, the present results cannot be directly comparable to those obtained by Calabria et al. [14] because these authors measured both basal and heat-induced HSP70 protein levels in outbred strains (obtained from cyclically permuted reciprocal crosses between isochromosomal lines). The differences observed here in the number of ARE motifs among the two *hsp70* copies could contribute to ensure basal amounts of HSP70 protein, basically from one of the gene copies. Moreover, the existence in *D. melanogaster* of 3 and five ARE motifs, respectively, in the two *hsp70* gene clusters compared to *D. subobscura* could contribute to explain why in the previous work [14] certain basal protein amounts were detected. In this regard, a recent study suggested that differences in acclimation capacity between *D. melanogaster* and *D. subobscura* correlate with structural differences in the *hsp70* regulatory regions [35]. However, in view of the highly conserved regulatory elements between chromosomal arrangements found here, we cannot easily explain the differences observed in protein amounts between arrangements [14]. These regulatory elements seem to be particularly highly constrained even though significant genetic differentiation between arrangements was detected when regulatory 5′ and 3′ were considered as a whole (Table 1). This is in concordance with a previous study where differences between arrangements at non-coding regions were also observed [36].

Previous findings pointed out that changes in the expression of particular genes adjacent to inversion breakpoints could contribute to an inversion evolutionary success [37, 38]. However, studies using genome wide transcriptome analyses in *D. melanogaster* [39] and *D. pseudoobscura* [40] concluded that, whereas inversions affect the expression of numerous genes, little evidence of direct position effects was globally detected at the chromosomal level. In *D. subobscura* the *hsp70* loci are

Fig. 3 Hsp70 expression in females (black) and males (grey) carrying O_{3+4+7} and O_{ST} arrangements. Error bars are 95% of confidence intervals

located in subsection 94A of the hypothetical O_3 ancestral arrangement. This subsection is inside the derived O_4 and O_{ST} inversions that gave rise to present day O_{3+4} and O_{ST} gene arrangements, respectively [27]. Using the recently assembled genome of *D. subobscura* [41] we have estimated that the *hsp70* loci are around 0.8 Mb from the closest O_4 distal breakpoint and likely too far away as to be directly affected by this breakpoint, an inference that is in agreement with the similar RNA amounts found in the two chromosomal arrangements analysed here. Four genes (*Pxd*, *Set 8*, *CG5225* and *Acf*) are located near O_4 inversion breakpoints [42], but at present we do not know whether potential position effects affecting these genes play any role in the in the adaptive value of this inversion. In any case, fitness interactions among loci (epistasis) seem to be important because a recombination load (i.e., the loss of fitness caused by recombination between beneficial combinations of interacting alleles) was detected in $O_{ST}/O_{3+4} + 7$ heterokaryotypes [43].

A potential limitation of our study is that we cannot extrapolate the results to mRNA expression levels during the larval stage. Insects with complex life cycles as *D. subobscura* might evolve different thermal sensitivities across life stages, particularly between mobile adults (with a higher capacity for thermoregulatory behaviour) and more sessile (eggs, larvae, pupae) life stages. A previous study assessing the thermal evolution of total mRNA expression levels in *D. subobscura* larvae did not detect differential gene expression for *hsp70* among thermal environments, though the highest temperature used in the experiments was likely not high enough as to induce a heat-shock response [44]. We cannot unequivocally generalize our present results to other life stages, but we think that there are not compelling reasons to believe that the basal mRNA expression of *hsp70* would have been different between O_{3+4} and O_{ST} gene arrangements had we used larvae instead of adults.

An important finding is the trend towards higher mRNA amounts in females compared to males. Differences between sexes are, in principle, not surprising because in *Drosophila* a substantial fraction of the transcriptome shows sex-dependent regulation [45]. In our case, the differences could be somehow understandable if the HSP70 protein played sex-specific functions. In *D. melanogaster*, this protein is expressed during spermatogenesis under heat-shock conditions in adults [46], and at normal rearing temperatures in the larval stage and onwards [16, 47]. In the case of females, a novel role for the molecular chaperone HSP70 in the regulation of border cell migration during *Drosophila* oogenesis has also been reported [48].

Conclusions

The mRNA levels from isogenic lines for O_{ST} and $O_{\underline{3+4}+7}$ gene arrangements in *D. subobscura* are quantitatively

equivalent. This provides some evidence of a parallelism between *hsp70* mRNA expression and its genetic organization. Therefore, the basal levels of *hsp70* mRNA cannot, in principle, explain the spatiotemporal patterns observed for these two gene arrangements. A potential drawback of our study is that we cannot extrapolate the results to wild, outbred flies. Further experiments are necessary to understand the actual role (if any) played by *hsp70* genes in the adaptive value of the rich chromosomal polymorphism on the distal segment of chromosome O, where the *hsp70* genes are located.

Methods

Base stocks and fly handling

Wild *D. subobscura* males were collected by net-sweeping over banana baits fermented with baker's yeast in Berbikiz (north of Spain; 43°11′20.31″ N, 3°5′23.74″ W) in 2012. They were individually crossed to three or four virgin females from the *ch-cu* marker strain, which carries the recessive markers on the O chromosome cherry eyes (*ch*) and curled wings (*cu*) and has a highly homogeneous genetic background, to identify the gene arrangements of one set of the five major wild-type chromosomes (the other set of homologous chromosomes coming from the *ch-cu* strain). Thereafter, offspring males from each line were crossed to females from the *Va/Ba* (*Varicose/Bare*) balancer stock [49], which has the same genetic background than the *ch-cu* strain, to obtain isochromosomal lines for the O chromosome in an otherwise highly homogeneous genetic background.

A total of 6 O_{ST} and 6 $O_{3+4} + 7$ isogenic lines were used in the experiments. All 12 lines were reared at 18 °C (the optimum temperature [50]) with a 12:12 light/dark cycle on David's killed-yeast *Drosophila* medium [51]. To obtain the experimental flies, eggs from the different lines were placed in 2×8 cm vials (45 eggs/vial) with 6 mL of food, which guaranteed low or negligible larval crowding that could have triggered the *hsp70* stress response [20]. For each isogenic line tested, two independent sets of an average of three biological replicates, each containing three 7-day old adult males and females, were used for mRNA analyses.

Characterization and sequence analysis of the *hsp70* 5′ and 3′ regulatory regions

For each isogenic line, two genomic regions expanding ~ 300 bp upstream the transcription starting site and ~ 205 bp downstream the stop codon were sequenced for the two *hsp70* gene copies (*hsp70B* and *hsp70A*; see [27]). The primers used for amplification of 3′ *hsp70B* and *hsp70A* regulatory regions were, respectively: P154-proximal (5′- ACC CGC AAA ATT GAA CCC AA-3′),

CD_1_F (5′-ACC ATA CAG AAC GAC AAG GGT- 3′) and DMT_proximal (5′-AGT CGG AAT TGT GAA GCC TT-3′). Due to the SNPs (Single nucleotide polymorphisms) found in 5′regulatory regions between lines, different primer sets were used: a) *hsp70B*: ARN_1R (5′-TGT GAT GCT TTG GCC CAG AT-3′) and D5_NP (5′-GAT CCT CAA GGA ATG CGA AA-3′)/D7_NP(5′-CTG TAG ATC TTC AGG CAT AAG CTG-3′)/NP-1-FW (5′ GTA TTG TGG CTT CTT AAC GAG GTTC C-3′); b) *hsp70A*: ARN_1R and and D1_NP (5′-CTT CGG TGG GTT GGA TTG TAG ATT C-3′)/ D3_NP(5′-CTA CTT TCT CTC CTG TGT ATG TCT GG-3′)/D4_NP(5′-CTC TCC TGT GTA CAT ATG TCT TGC G-3′). To identify several CREs (Cis-Regulatory Elements) that are typically found in the proximal 5′ flanking region, the sequences of the two *hsp70* genes were first separately aligned using the MAFFT E-INS-i algorithm [52] available at: https://mafft.cbrc.jp/alignment/server/. Next, we used the information available on the *hsp70* proximal promoters of *D. subobscura* [27] and other *Drosophila* species [29] to identify the conserved heat shock elements in the sequences of the 12 isogenic lines (HSEs 1–4), and the Eukaryotic Neural Network Promoter Prediction software (http://www.fruitfly.org/seq_tools/promoter.html) to find other CREs such as GAGA binding sites and some core promoter elements, including the transcription start site (TSS) and the TATA box. 3′ UTR sequences were aligned using the MAFFT E-INS-i algorithm [52] and manually refined using the Genedoc alignment editor (version 2.7) [53]. The web-based ARE score application (http://arescore.dkfz.de/arescore.pl) was used to identify post-transcriptional regulators of mRNA levels such as AU-rich elements [54]. The DNASp version v6.12 program [55] was used for genetic differentiation analyses between arrangements, assessed using D_{xy} [56] and F_{st} [57], and the significance for Snn [58] was obtained with 1000 replicates. Analyses were performed excluding gaps.

Quantification of *hsp70* transcripts by qRTPCR

Total RNA extraction was performed separately in adult males and females using the Trizol Lysis reagent (Thermo Fisher Scientific Inc.) following manufacturer's instructions, and then treated with 1 µl (2 U/µl) of Ambion DNase I (Thermo Fisher Scientific Inc.) 4 h at 37 °C. Concentration of DNA-free samples was checked in a Nanodrop-2000 spectrophotometer (Thermo Scientific Nanodrop) and adjusted to ~ 1 µg/µL and then cDNA synthesis was carried out with anchored-Oligo (dT)18 primers using the Transcriptor First Strand cDNA Synthesis kit (Roche Diagnostics SL, SPAIN). Gene expression was quantified by measuring fluorescence intensity using iQ SYBR Green Supermix (BioRad, Hercules CA, USA) on a CFX96 BioRad Real-Time

lightcycler at 59 °C using *hsp70* specific primers. All assays were performed in three technical replicates, and the *rp49* housekeeping gene was used as an endogenous control because this gene showed to be expressed equally in all samples analysed ($P = 0.772$, see Additional file 5). Relative expression of *hsp70* was then calculated with the comparative C_T method [59] and ΔC_T values for the set of samples are provided in Additional file 6.

For each isogenic line, we amplified a conserved *hsp70* coding fragment of 226 bp, common to both *hsp70* copies described in *D. subobscura* [27]. The primers used for amplification were: HSP70_1L (5′-CAC AGT CTT TGA CGC CAA GC-3′) and HSP70_1R (5′-TGT GAT GCT TTG GCC CAG AT-3′). For the *rp49* housekeeping gene, the primers RP49F (5′-ACA TCG GTT ATG GCT CCA C-3′) and RP49R (5′-GAT TTC CTT GCG CTT CTT TG-3′) were designed from a *D. subobscura* GenBank sequence (Accession: AJ228921), amplifying a 164 bp segment from the gene's second exon. Primer efficiencies of *hsp70* and *rp49*, calculated on a standard curve slope using serially diluted templates, were 98 and 99.6%, respectively.

Statistical methods

Hsp70 mRNA expression levels were analysed using the linear model.

$$y_{ijkl} = \mu + A_i + L_{j(i)} + S_k + AS_{ik} + \varepsilon_{ijkl}, (1)$$

where μ is the overall grand mean, A_i is the fixed effect of gene arrangement (O_{ST}, $O_{\underline{3+4}} + 7$), $L_{j(i)}$ is the random effect of the *j*th isochromosomal line ($j = 1, 2, \cdots, 6$) within arrangement *i*, S_k is the fixed effect of sex, and ε_{ijkl} is the residual error associated with the mRNA expression levels of the *l*th replicate from the *k*th sex with the *i*th gene arrangement from the *j*th isochromosomal line. Following Schmittgen and Livak [59], statistical tests for mRNA relative expression were performed after the $2^{-\Delta C_T}$ transformation.

To avoid potential problems with the standard assumption that errors are independent, and identically distributed as normal random variables with common variance and zero expectation, we also performed permutation tests [60, 61] to check the validity of parametric *P*-values. In model (1) the permutation test is a four-stage process. First, random permutation among lines, within arrangements, for the among-line $F_{\text{line(arrangement)}}$; second, random permutation among lines and arrangements for the between-arrangements $F_{\text{arrangement}}$; third, random permutation between sexes for the between-sex F_{sex}; and fourth, random permutation between arrangements and sexes for the arrangement × sex interaction $F_{\text{arrangement} \times \text{sex}}$. In each case, 9999 random allocations plus the observed *F*-values were obtained, which gave us

Puig Giribets *et al. BMC Evolutionary Biology*   (2020) 20:17

Page 8 of 9

10,000 values of each *F*-statistics for each dependent variable. We then calculate the percentage of replications under the null hypothesis that provide as large a value of each *F* as the obtained *F*.

The equal expression of *rp49* gene was tested on Ct values using one-way ANOVA with four sample groups: females O_{ST}, males O_{ST}, females $O_{3+4} + 7$ and males $O_{3+4} + 7$. The Levene's test was used to test the homogeneity of variances between groups.

Analyses were performed using the statistical software STATISTICA 8.0 [62]. Permutation tests were implemented in the algebra environment MATLAB using tools supplied by the Statistics Toolbox [63].

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12862-020-1584-z.

Additional file 1. Multiple sequence alignment of 5'proximal promoters of *hsp70*A.

Additional file 2. Multiple sequence alignment of 5'proximal promoters of *hsp70*B.

Additional file 3. Multiple sequence alignment of *hsp70*A 3'UTR.

Additional file 4. Multiple sequence alignment of *hsp70*B 3'UTR.

Additional file 5. Statistical results for the homogeneous expression of *rp49* housekeeping gene.

Additional file 6. Raw gene expression data.

Abbreviations
ARE: AU-rich element; CREs: Cis-regulatory elements; HSEs: Heat shock elements; HSF: Heat shock factor; SNPs: Single nucleotide polymorphisms; UTR: Untranslated region

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

References
1. Menozzi P, Krimbas CB. The inversion polymorphism of *D. subobscura* revisited: Synthetic maps of gene arrangement frequencies and their interpretation. J Evol Biol. 1992;5:625–41.
2. . Powell JR: Progress and Prospects in Evolutionary Biology. The Drosophila Model, New York: Oxford Univ Press; 1997.
3. Krimbas CB. Drosophila subobscura, Biology, Genetics and Inversion Polymorphism. Hamburg: Verlag Dr. Kovač; 1993.
4. Balanyà J, Serra L, Gilchrist GW, Huey RB, Pascual M, Mestres F, Solé E. evolutionary pace of chromosomal polymorphism in colonizing populations of *Drosophila subobscura*: an evolutionary time series. Evol. 2003;57:1837–45.
5. Fontdevila A, Zapata C, Alvarez C, Sánchez L, Méndez J, Enriquez I. Genetic coadaptation in the chromosomal polymorphism of *Drosophila subobscura*. I. Seasonal changes of gametic disequilibrium in a natural population. Genetics. 1983;105:935–55.
6. Rodríguez-Trelles F, Alvarez G, Zapata C. time-series analysis of seasonal changes of the O inversion polymorphism of *Drosophila subobscura*. Genetics. 1996;142:179–87.
7. Rodríguez-Trelles F, Rodríguez MA. rapid micro-evolution and loss of chromosomal diversity in *Drosophila* in response to climate warming. Evol Ecol. 1998;12:829–38.
8. Balanyá J, Oller JM, Huey RB, Gilchrist GW, Serra L. global genetic change tracks global climate warming in *Drosophila subobscura*. Sci. 2006;313:1773–5.
9. Lindquist S, Craig EA. the heat-shock proteins. Ann Rev Genet. 1988;22:631–77.
10. Lindquist S. the heat-shock response. Annu Rev Biochem. 1986;55:1151–91.
11. Parsell DA, Taulien J, Lindquist S. the role of heat shock proteins in thermotolerance. Phil Trans Royal Soc London Series B. 1993;339:279–86.
12. Feder ME, Blair N, Figueras H. natural thermal stress and heat-shock protein expression in *Drosophila* larvae and pupae. Funct Ecol. 1997;11:90–100.
13. Bettencourt BR, Kim I, Hoffmann AA, Feder ME. response to natural and laboratory selection at the *Drosophila hsp70* genes. Evol. 2002;56:1796–801.
14. Calabria G, Dolgova O, Rego C, Castañeda LE, Rezende EL, Balanyà J, Pascual M, Sørensen JG, Loeschcke V, Santos M. Hsp70 protein levels and thermotolerance in *Drosophila subobscura*: a reassessment of the thermal co-adaptation hypothesis. J Evol Biol. 2012;25:691–700.
15. Krebs RA, Loeschcke V. acclimation and selection for increased resistance to thermal stress in *Drosophila buzzatii*. Genetics. 1996;142:471–9.
16. Krebs RA, Feder ME. tissue-specific variation in Hsp70 expression and thermal damage in *Drosophila melanogaster* larvae. J Exp Biol. 1997;200:2007–15.
17. Feder JH, Rossi JM, Solomon J, Solomon N, Lindquist S. the consequences of expressing hsp70 in *Drosophila* cells at normal temperatures. Genes Dev. 1992;6:1402–13.
18. Rodriguez-Trelles F, Tarrio R, Santos M. Genome-wide evolutionary response to a heat wave in *Drosophila*. Biol Lett. 2013;9:20130228.
19. Bettencourt BR, Feder ME, Cavicchi S. experimental evolution of Hsp70 expression and thermotolerance in *Drosophila melanogaster*. Evol. 1999;53:484.
20. Sørensen JG, Loeschcke V. larval crowding in *Drosophila melanogaster* induces Hsp70 expression, and leads to increased adult longevity and adult thermal stress resistance. J Insect Physiol. 2001;47:1301–7.
21. Sørensen JG, Michalak P, Justesen J, Loeschcke V. expression of the heat-shock protein HSP70 in *Drosophila buzzatii* lines selected for thermal resistance. Hereditas. 1999;131:155–64.
22. Garbuz D, Evgenev MB, Feder ME, Zatsepina OG. Evolution of thermotolerance and the heat-shock response: evidence from inter/intraspecific comparison and interspecific hybridization in the virilis species group of *Drosophila*. I. Thermal phenotype. J Exp Biol. 2003;206:2399–408.
23. Gong WJ, Golic KG. loss of Hsp70 in *Drosophila* is pleiotropic, with effects on thermotolerance, recovery from heat shock and neurodegeneration. Genetics. 2006;172:275–86.
24. Feder ME, Cartaño NV, Milos L, Krebs RA, Lindquist SL. effect of engineering Hsp70 copy number on Hsp70 expression and tolerance of ecologically

relevant heat shock in larvae and pupae of *Drosophila melanogaster*. J Exp Biol. 1996;199:1837–44.

25. Bettencourt BR, Hogan CC, Nimali M, Drohan BW. Inducible and constitutive heat shock gene expression responds to modification of *Hsp70* copy number in *Drosophila melanogaster* but does not compensate for loss of thermotolerance in *Hsp70* null flies. BMC Biol. 2007;6:5.

26. Evgen'ev MB, Zatsepina OG, Garbuz D, Lerman DN, Velikodvorskaya V, Zelentsova E, Feder ME. evolution and arrangement of the *hsp70* gene cluster in two closely related species of the virilis group of *Drosophila*. Chromosoma. 2004;113:223–32.

27. Puig Giribets M, García Guerreiro MP, Santos M, Ayala FJ, Tarrío R, Rodríguez-Trelles F. chromosomal inversions promote genomic islands of concerted evolution of Hsp70 genes in the *Drosophila subobscura* species subgroup. Mol Ecol. 2019;28:1316–32.

28. Genissel A, LM MI, Wayne ML, Nuzhdin SV. Cis and trans regulatory effects contribute to natural variation in transcriptome of *Drosophila melanogaster*. Mol Biol Evol. 2008;25:101–10.

29. Tian S, Haney RA, Feder ME. Phylogeny disambiguates the evolution of heat-shock cis-regulatory elements in *Drosophila*. PLoS One. 2010;5:e10669.

30. Omichinski JG, Pedone PV, Felsenfeld G, Gronenborn AM, Clore GM. the solution structure of a specific GAGA factor-DNA complex reveals a modular binding mode. Nat Struct Biol. 1997;4:122–32.

31. Cairrao F, Halees AS, KSA K, Morello D, Vanzo N. AU-rich elements regulate *Drosophila* gene expression. Mol Cell Biol. 2009;29:2636–43.

32. Becker K, Bluhm A, Casas-Vila N, Dinges N, Dejung M, Sayols S, Kreutz C, Roignant JY, Butter F, Legewie S. Quantifying post-transcriptional regulation in the development of *Drosophila melanogaster*. Nat Commun. 2018;9:4970.

33. Franke K, Fischer K. inbreeding interferes with the heat-shock response. Heredity. 2015;114:80–4.

34. Vermeulen CJ, Bijlsma R. characterization of conditionally expressed mutants affecting age specific survival in inbred lines of *Drosophila melanogaster* lethal conditions and temperature-sensitive periods. Genetics. 2004;167:1241–8.

35. Sørensen JG, Giribets MP, Tarrío R, Rodríguez-Trelles F, Schou MF, Loeschcke V. expression of thermal tolerance genes in two *Drosophila* species with different acclimation capacities. J Therm Biol. 2019;84:200–7.

36. Simões P, Pascual M. Patterns of geographic variation of thermal adapted candidate genes in *Drosophila subobscura* sex chromosome arrangements. BMC Evol Biol. 2018;18:60.

37. Puig M, Cáceres M, Ruiz A. silencing of a gene adjacent to the breakpoint of a widespread *Drosophila* inversion by a transposon-induced antisense RNA. Proc Natl Acad Sci. 2004;101:9013–8.

38. Huang YC, Dang VD, ChangNC WJ. Multiple large inversions and breakpoint rewiring of gene expression in the evolution of the fire ant social supergene. Proc R Soc B. 2018;285:20180221.

39. Lavington E, Kern AD. The effect of common inversion polymorphisms *In (2L) t* and *In (3R) Mo* on patterns of transcriptional variation in *Drosophila melanogaster*. G3 (Bethesda). 2017;7:3659–68.

40. 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.

41. Karageorgiou C, Gámez-Visairas V, Tarrío R, Rodríguez-Trelles F. Long-read based assembly and synteny analysis of a reference *Drosophila subobscura* genome reveals signatures of structural evolution driven by inversions recombination-suppression effects. BMC Genomics. 2019;20:223.

42. Puerma E, Orengo DJ, Aguadé M. Multiple and diverse structural changes affect the breakpoint regions of polymorphic inversions across the *Drosophila* genus. Sci Rep. 2016;6:36248.

43. Santos M. recombination load in a chromosomal inversion polymorphism of *Drosophila subobscura*. Genetics. 2009;181:803–9.

44. Laayouni H, García-Franco F, Chávez-Sandoval BE, Trotta V, Beltran S, Corominas M, Santos M. Thermal evolution of gene expression profiles in *Drosophila subobscura*. BMC Evol Biol. 2007;10:363.

45. Ranz JM, Castillo-Davis CI, Meiklejohn CD, Hartl DL. sex-dependent gene expression and evolution of the *Drosophila* transcriptome. Sci. 2003;300:1742–5.

46. Michaud S, Marin R, Westwood JT, Tanguay RM. cell-specific expression and heat-shock induction of Hsps during spermatogenesis in *Drosophila melanogaster*. J Cell Sci. 1997;110:1989–97.

47. Lakhotia SC, Prasanth KV. tissue- and development-specific induction and turnover of *hsp70* transcripts from loci 87A and 87C after heat shock and during recovery in *Drosophila melanogaster*. J Exp Biol. 2002;205:345–58.

48. Cobreros L, Fernández-Miñán A, Luque CM, González-Reyes A, Martín-Bermudo MD. a role for the chaperone Hsp70 in the regulation of border cell migration in the *Drosophila* ovary. Mech Dev. 2008;125:1048–58.

49. Sperlich D, Feuerbach-Mravlag H, Lange P. genetic load and viability distribution in central and marginal populations of *Drosophila subobscura*. Genetics. 1977;86:835–48.

50. Rego C, Balanyà J, Fragata I, Matos M, Rezende EL, Santos M. Clinal patterns of chromosomal inversion polymorphisms in *Drosophila subobscura* are partly associated with thermal preferences and heat stress resistance. Evol. 2010;64:385–97.

51. David J. A new medium for rearing *Drosophila* in axenic conditions. Dros Inf Serv. 1962;36:128.

52. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013; 30:772–80.

53. Nicholas KB, Nicholas HBJ, DWI D. GeneDoc: Analysis and visualization of genetic variation. EMBNEW NEWS. 1997;4:14.

54. Spasic M, Friedel CC, Schott J, Kreth J, Leppek K, Hofmann S, Ozgur S, Stoecklin G. Genome-wide assessment of AU-rich elements by the AREScore algorithm. PLoS Genet. 2012;8:e1002433.

55. Rozas J, Ferrer-Mata A, JC S-DB, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol Biol Evol. 2017;34:3299–302.

56. Nei M. Molecular Evolutionary Genetics. New York: Univ Press; 1987.

57. Hudson RR, Slatkin M, Maddison WP. estimation of levels of gene flow from DNA sequence data. Genetics. 1992;132:583–9.

58. Hudson RR. a new statistic for detecting genetic differentiation. Genetics. 2000;155:2011–4.

59. Schmittgen TD, Livak KJ. analyzing real-time PCR data by the comparative CT method. Nat Protoc. 2008;3:1101–8.

60. Edgington ES. Randomization Tests. 3rd ed. New York: Marcel Dekker; 1995.

61. BFJ M. Randomization, Bootstrap and Monte Carlo Methods in Biology. 3rd ed. New York: Chapman & Hall/CRC; 2007.

62. StatSoft, Inc. 2007. STATISTICA (data analysis software system), version 8.0. www.statsoft.com.

63. MATLAB and Statistics Toolbox Release. 2016. The MathWorks, Inc., Natick.

Publisher's Note