Full genome gene expression analysis of the heat stress response in *Drosophila melanogaster*

Jesper G. Sørensen,¹ Morten M. Nielsen,^{1,2} Mogens Kruhøffer,³ Just Justesen,^{1,2} and Volker Loeschcke¹

¹Aarhus Centre for Environmental Stress Research (ACES), Department of Ecology and Genetics, University of Aarhus, Building 540, Ny Munkegade, DK-8000 Aarhus C, Denmark

²Department of Molecular Biology (MBI), University of Aarhus, C. F. Møllers Allé 130, DK-8000 Aarhus C, Denmark

³Molecular Diagnostic Laboratory, Aarhus University Hospital, Brendstrupgårdsvej 100, DK-8200 Aarhus N, Denmark

Abstract The availability of full genome sequences has allowed the construction of microarrays, with which screening of the full genome for changes in gene expression is possible. This method can provide a wealth of information about biology at the level of gene expression and is a powerful method to identify genes and pathways involved in various processes. In this study, we report a detailed analysis of the full heat stress response in *Drosophila melanogaster* females, using whole genome gene expression arrays (Affymetrix Inc, Santa Clara, CA, USA). The study focuses on up- as well as downregulation of genes from just before and at 8 time points after an application of short heat hardening (36°C for 1 hour). The expression changes were followed up to 64 hours after the heat stress, using 4 biological replicates. This study describes in detail the dramatic change in gene expression over time induced by a short-term heat treatment. We found both known stress responding genes and new candidate genes, and processes to be involved in the stress response. We identified 3 main groups of stress responsive genes that were early-upregulated, early-downregulated, and late-upregulated, respectively, among 1222 differentially expressed genes in the data set. Comparisons with stress sensitive genes identified by studies of responses to other types of stress allow the discussion of heat-specific and general stress responses in *Drosophila*. Several unexpected features were revealed by this analysis, which suggests that novel pathways and mechanisms are involved in the responses to heat stress and to stress in general. The majority of stress responsive genes identified in this and other studies were downregulated, and the degree of overlap among downregulated genes was relatively high, whereas genes responding by upregulation to heat and other stress factors were more specific to the stress applied or to the conditions of the particular study. As an expected exception, heat shock genes were generally found to be upregulated by stress in general.

INTRODUCTION

Cellular responses to stress are evolutionarily ancient, ubiquitous, and essential mechanisms for the continued survival and reproduction of organisms (Lindquist 1986; Feder and Hofmann 1999). These mechanisms protect the cells from cellular damage to proteins and membranes by environmental challenges. When exposed to such damaging stress conditions (eg, a nonlethal heat shock), major changes in gene expression and cell physiology take place

(Lindquist 1986; Feder and Hofmann 1999). These changes are known as the cellular stress response. The changes are believed to include inhibition of deoxyribonucleic acid synthesis and transcription, inhibition of ribonucleic acid (RNA) processing and translation, and enhanced degradation of proteins through proteosomal and lysosomal pathways, and at the same time, increased expression and activity of molecular chaperones and associated genes (Lindquist 1986; Feder and Hofmann 1999).

An important part of the cellular response to heat stress is constituted by a group of genes coding for heat shock proteins (Hsps) or stress proteins because their expression can be induced by high temperatures and a whole range of other stress factors (Sørensen et al 2003).

Correspondence to: Jesper G. Sørensen, E-mail: biojgs@biology.au.dk.

Note: Data set material is available at the "Aarhus stress group" homepage <http://mit.biology.au.dk/aces>

Received 8 April 2005; Revised 26 July 2005; Accepted 26 July 2005.

Hsps constitute an inducible part of molecular chaperones that play important roles in transport, folding, unfolding, assembly and disassembly of multistructured protein complexes, signal pathways, degradation of misfolded or aggregated proteins, and the activation of enzymes and receptors (Parsell and Lindquist 1993). The heat shock genes are upregulated after exposure to stressful, potentially damaging conditions and provide the organism with a temporary enhanced tolerance to stress (for review, see Lindquist 1986; Feder and Hofmann 1999; Sørensen et al 2003). Hsps, other molecular chaperones, and the stress response in general have been widely studied in many fields of biology, and a large number of publications on their molecular and physiological functions are available (see Parsell and Lindquist 1993; Feder and Hofmann 1999; Morimoto et al 1999; Pockley 2003; Sørensen et al 2003; Prohászka and Füst 2004). Although much has been learned about the heat shock response since the discovery by Ritossa (1962, 1964), only a limited number of genes affected are known, and only little is known about the majority of genes regulated in association with this response. For example, little information exists on genes affected by stress exposure at longer time scales. Such genes might play a vital role in induced resistance because the time course of improved resistance matches poorly with the expression profiles of known molecular chaperones, eg, Hsp70 (Dahlgaard et al 1998; Sørensen et al 2003). This suggests that other genes so far unidentified must be involved. Identification of new genes of importance for the responses to environmental stress has relevance for many fields of biology. Stress responsive genes are of interest for studying and understanding not only the environmental stress resistance and the stress response in general but also protein folding-mediated diseases in humans, immunological responses, animal breeding, genetic stress, protein quality control, developmental biology, and gene regulation. Thus, the heat shock response is of continued interest and relevance (see Pauli et al 1992). The questions asked are related to which genes are up- or downregulated in the heat stress response and in which sequence, how these genes are regulated, the time course of induced changes in gene expression, the function of genes involved, the induction of heat tolerance, and other changes to the phenotype.

In this study, we use the microarray technique to investigate the stress response at the level of gene expression in *D. melanogaster*. This technique provides the possibility to reinvestigate the heat stress response in the full *Drosophila* genome on a detailed time scale and in greater detail than has been possible before. We have investigated the genomic expression profile replicated in different genetic backgrounds of *D. melanogaster* using microarrays in unstressed flies and in a time series from immediately

after to 64 hours after application of a short nonlethal heat stress.

The aim of the study was 2-fold. First, to describe the heat shock response in temporal details with attention to genes, which respond after the known *hsp* genes expression levels are back to unstressed levels, and to do this in a complex whole organism to include potentially important genes overlooked in cell culture investigations. Moreover, by comparing with the results of other expression studies of responses to a range of stressful conditions, more knowledge of specific and general stress responses could be gained. Second, to identify new heat shock genes or other new groups of genes regulated after heat stress induction, as potential candidates for the response and tolerance to heat stress and stress in general. The analyzed data provide an overview of the genes and the processes involved in the response to heat stress over time and has a number of unexpected features, which suggest novel pathways and mechanisms.

MATERIALS AND METHODS

Origin of experimental flies

The selection and control lines used in this experiment originated from a mass population. Flies from different geographical regions and background were mixed to increase the genetic diversity in the starting material. The mass population was created in September 2002 by mixing 600–700 flies from each of several preexisting stocks from the laboratory. The origin of flies and the setup of the mass population are described in more detail in Bublly and Loeschcke (2005).

Maintenance regimen

The mass population was maintained as 1 interbreeding population for 4 generations before a number of selection regimens were established. In this study, flies from 2 maintenance regimens were used (control and selection for induced heat survival). Independent replicate lines were established for each selection regimen. Each replicate line was maintained in 5 culture bottles with a minimum population size of 60 pairs in each (in total, a population size of 300 pairs). The 5 bottles, within a replicate line, were mixed each generation. The lines were kept at this relatively high population size to reduce the effect of drift. Replicate lines of the control regimen were kept at standard laboratory conditions at 25°C and a 12/12 hours light/dark cycle on standard agar-sugar-yeast-oatmeal medium, unless otherwise stated. Flies from the induced heat survival regimen were heat shocked every second generation when 5 days old. Flies were hardened for 30 minutes at 36°C in a preheated waterbath. Subsequently,

Table 1 Design of experiment and number of microarrays analyzed. The heat treatment was 36°C for 1 h in the interval between –1 and 0 h. Before and after heat treatment, flies were kept at 25°C. Samples from Co1 (control line 1) and Hs1 (heat selected line 1) were analyzed in the first array run and Co2 and Hs2 in the second array run

Sampling times (h)	–1	0.25	1	2	4	8	16	32	64	=
Co1	1	1	1	1	1	1	1	1	1	9
Co2	1	1	1	1	1	1	1	1	1	9
Hs1	1	1	1	1	1	1	1	1	1	9
Hs2	1	1	1	1	1	1	1	1	1	9
Total per time point	4	4	4	4	4	4	4	4	4	36

1 hour later, flies were heat shocked for 1 hour at 38°C. The heat shock temperature was gradually increased (to 38.5°C for 65 minutes) between selection events to maintain the mortality constant because the lines improved the resistance due to the selection. For the array experiment, the lines had passed 20 generations, ie, 10 selected generations. The flies used for experimentation were offspring from an unselected generation to avoid any cross-generational effects of the selection procedure. For a more detailed description of origin and selection regimens of flies, see Bubli and Loeschke (2005).

Induction of the stress response

Flies were set up in 200-mL bottles under uncrowded conditions (25 pairs were allowed egg laying for 24 hours); on the start of emergence, the bottles were emptied, and the next morning, flies were collected under CO₂ anesthesia. We collected 10 females per vial and 60 vials from each replicate line. After recovery from CO₂ anesthesia and handling, flies that were dead or injured were replaced with fresh ones. The flies were aged to 3–4 days before the onset of the experiment. Immediately before the hardening treatment, all flies were transferred to 5-mL plastic vials that allowed for continuous ventilation when partly open. During hardening and heat shock, the vials were closed but otherwise partly open at all times, allowing ventilation. The vials had a layer of agar mixed with sugar at the bottom to provide moisture and food during the experiment. Before the experiment, it was determined that adult flies could be maintained in these vials in excess of 4 days so long as ventilation, water, and basic food was provided under the experimental conditions. Heat hardening took place at 36°C for 1 hour in preheated waterbaths (from time –1 to 0 hour), and subsequently, the flies were returned to 25°C. One set of flies was not hardened and served as untreated controls. These were sampled immediately before heat hardening commenced at time –1 hour. The remaining vials with hardened flies were collected and frozen in liquid nitrogen and quickly stored at –80°C at different time points after hardening. Samples were taken at 0.25, 1, 2, 4, 8, 16, 32, 64 hours after hardening in addition to the untreated control (–1 hour).

For the experiment described here, the time series was replicated 4 times (4 × 9 chips). Two time series on flies were from the control regimen, and 2 on flies were from the induced heat survival selected regimen. Each time series (of 9 chips) was done on flies from an independent replicate line. Extreme care was taken to treat all samples equally and to be very precise with the timing of all steps throughout the process from collection and handling of flies to freezing of the treated samples and running the microarrays. Use of 10 female flies for each RNA sample hybridized, further reduced variation arising from individual transcription differences. The microarray analysis (see below) was for practical reasons done in 2 steps of 18 chips each. In each of the batches, all time points (9; from –1 to 64 hours) from both selection regimens (control and heat selected) were present. The design is graphically represented in Table 1.

RNA extraction and complementary DNA creation

The Affymetrix array contains 13 966 probe sets representing approximately 13 000 unique genes. Flies from each vial (10 females) were homogenized with a FP-120 Fast Prep bead beater, according to manufacturer's protocols (Bio-101, Carlsbad, CA, USA), in 1.5 mL Trizol Reagent (Invitrogen, Carlsbad, CA, USA) and 150 µL chloroform. Labeling, hybridization, and staining were performed essentially as described by Dyrskjot et al (2003). In brief, double-stranded complementary DNA was prepared from 5 µg of total RNA using the SuperScript Choice System (Invitrogen), according to the manufacturer's instructions, except using an oligo-dT primer containing a T7 RNA polymerase promoter site. Biotin-labeled complementary RNA (cRNA) was prepared using the BioArray High Yield RNA Transcript Labelling Kit (Enzo, Farmingdale, NY, USA). After the *in vitro* transcription (IVT) reaction, the unincorporated nucleotides were removed using RNeasy columns (Qiagen, Hilden, Germany).

Array hybridization and scanning

Fifteen micrograms of cRNA was fragmented at 94°C for 35 minutes in a fragmentation buffer containing 40 mM

Tris-acetate, pH 8.1, 100 mM potassium acetate (KOAc) KOAc, 30 mM magnesium acetate (MgOAc) MgOAc. Before hybridization, the fragmented cRNA in a 6× SSPE + 0.01% Tween 20 (SSPE-T) hybridization buffer (1 M NaCl, 10 mM Tris, pH 7.6, 0.005% Triton X-100) was heated to 95°C for 5 minutes and subsequently to 45°C for 5 minutes before loading onto the Affymetrix probe array cartridge (Drosophila Genome Array, 13 966 probe set). The probe array was then incubated for 16 hours at 45°C at constant rotation (60 rpm). The washing and staining procedure was performed in the Affymetrix Fluidics Station. The probe array was exposed to 10 washes in 6× SSPE-T at 25°C, followed by 4 washes in 0.5× SSPE-T at 50°C. The biotinylated cRNA was stained with a streptavidin-phycoerythrin conjugate, final concentration 2 µg/L (Molecular Probes, Eugene, OR, USA), in 6× SSPE-T for 30 minutes at 25°C, followed by 10 washes in 6× SSPE-T at 25°C. An antibody-amplification step followed using normal goat immunoglobulin G as a blocking reagent, final concentration 0.1 mg/mL (Sigma), and biotinylated antistreptavidin antibody (goat), final concentration 3 µg/mL (Invitrogen). This was followed by a staining step with a streptavidin-phycoerythrin conjugate, final concentration 2 µg/µL (Molecular Probes), in 6× SSPE-T for 30 minutes at 25°C and 10 washes in 6× SSPE-T at 25°C. The probe arrays were scanned at 560 nm using a confocal laser-scanning microscope (Hewlett Packard GeneArray Scanner G2500A).

Statistical analysis

The raw data was GC-RMA normalized with the BIOCONDUCTOR application for R (Wu and Irizarry 2004). We used Multi Class SAM to select probe sets differentially expressed over time for further analyses (Tusher et al 2001). SAM was run on the full data set; thus, no initial filtering was performed. This procedure allows the identification of genes that are differentially expressed consistently across replicates while controlling for multiple testing by estimating the false discovery rate (FDR) (Benjamini and Hochberg 1995). For the SAM analysis, we used 100 permutations, which gave a stable number of significant genes and estimates of FDR among runs. We selected 1222 significant genes, which corresponded to a median FDR of <1% (FDR 90 percentile around 6%).

Significant genes were further analyzed by applying clustering to identify coexpressed genes (Eisen et al 1998; Spellman et al 1998; Gasch et al 2000). To reduce the data set further and to focus on general and consistent patterns of gene expression, we calculated the mean expression of the 4 replicates of each time point and used these 9 means for each gene in the cluster analysis. Two clustering methods were used, K-mean and hierarchical clustering, using average linkage clustering for both cases on

the means of the significantly differentially expressed genes (1222). For the K-mean clustering, the genes were mean-centered, so that each gene had mean 0 and variance 1 over the 9 time points, to be able to compare expression profiles among genes. This clustering method has been successfully used to identify coregulated genes (Tavazoie et al 1999). Repeated K-mean clustering gave consistently very similar clusters with approximately equal number of genes in each, also when initial filtering criteria and number of clusters chosen was varied. We chose to produce 20 clusters because this produced a high degree of separation and a low number of clusters with no functional meaning among repeated runs. Analysis of variance, K-mean clustering, and hierarchical trees were done using the computer package TIGR MeV v. 3.0.1 (Saeed et al 2003). Subsequently, subsets of clusters were selected for more detailed investigation. Because the K-mean method does not provide information about expression levels or fold changes, we plotted the genes from some clusters using the log2 mean expression (all times) vs the fold change from time -1 (control) to each time point. The genes from each cluster subset were analyzed further to establish functional groups of genes that increased or decreased over time and to establish the functional link to heat exposure. Information from the Gene Ontology (GO) (Gene-Ontology-Consortium 2001) database was combined with expression data, using the EASE application on the DAVID homepage (<http://david.niaid.nih.gov/david/ease.htm>) (Hosack et al 2003). In brief, each probe set was assigned, when 1 was available, to its GO annotation. The number of probe sets for the different GO terms were computed for each K-mean cluster and for groups of similar clusters. A probability for recovering this number, given the number of genes in the data set, was assigned to each represented GO term (Hosack et al 2003).

RESULTS AND DISCUSSION

The aim of this experiment was to investigate the heat stress response by investigating gene expression before and after a heat stress in temporal detail and to identify new candidate genes for heat stress resistance in this response using microarrays. In analyzing microarray data, it is crucial to identify and separate relevant or "real" differentially expressed genes from randomly varying genes. This has been traditionally done using various criteria such as fold change and thresholds of minimum expression levels; however, there are usually no statistical properties tied to them. In this study, we used multiclass SAM to select probe sets for further analyses. This procedure allows the identification of genes differentially expressed consistently across replicates while controlling for multiple testing by calculating the FDR (Benjamini

Fig 1. Average linkage hierarchical cluster of 9 time points (each the average of 4 microarrays) with 1222 significant differentially expressed genes each. The prestress control (–1 hour) clusters with 4 hours indicate that the main response to heat stress has ended between 4 and 8 hours after heat exposure.

and Hochberg 1995). We used a cutoff value of 0.011, yielding 1222 significant genes, which is equal to a median FDR of <1%. This seems to be a quite conservative threshold, and genes with smaller effects are overlooked and dismissed. However, >1200 genes are in the analysis at this threshold, and selecting more genes did not contribute significant new information. The design also addresses any heat-induced expression differences between controls and flies selected for induced heat survival. Because the responses to heat in these 2 regimens are nearly identical, such an investigation requires a different analytical approach, which will be presented in another article, along with additional data to increase power of detection (M.M. Nielsen, J.G. Sørensen, J. Justesen, M. Kruhøffer, and V. Loeschcke, unpublished observations).

Hierarchical clustering on all genes of all 36 chips (data not shown) showed little pattern as to which time points or replicates were clustered together (the chips from the first and the second hybridization runs showed a tendency to cluster). When hierarchical clustering was done on the 1222 significantly expressed genes, a pattern with much more dependency on time was found, ie, there was a clear structure introduced by time after hardening (Fig 1, Supplemental Fig I). The late time points from 8 to 64 hours made up 1 major cluster, whereas the early time points (–1 to 4 hours) made up another. The 4 chips from each of the time points –1, 4, and 8 hours, respectively, all clustered together. All 4 chips from time 0.25 hour clustered together with 1 chip from 1 hour, and the remaining chips from 1 and 2 hours clustered with each

other. Time points –1 and 4 hours clustered as neighbors, suggesting that much of the hardening effect on RNA expression has passed at time 4 hours. Chips from 16 and up to 64 hours were mixed and showed no clear time trend. That these did not cluster close to –1 hour might be surprising; however, the differentially expressed genes were expected to consist, in addition to the direct responses to the heat treatment, of primarily 2 categories of genes that could affect the cluster in this manner. One was circadian rhythm genes because the 4 late time points were spaced with 24- and 48-hour intervals (8 to 32 hours and 16 to 64 hours, respectively). It is well known that a large number of genes are regulated as a consequence of circadian rhythm (McDonald and Rosbash 2001). The other category was genes involved in reproduction. These genes were expected because female flies aged and matured during the time course of the experiment. To identify heat stress-responsive genes, it is important that these 2 categories were consistently clustered and identified because they should not be considered stress responsive. Examination of the K-mean clusters indeed showed that this was the case. When examining means of the replicates over time points (Fig 1) the time-dependent pattern became even stronger. The late time points from 8 to 64 hours again formed 1 cluster, with 16 and 64 hours on one side and 8 and 32 hours on the other forming separate branches. In the branch including the early time points, –1 and 4 hours on the one hand and 0.25, 1, and 2 hours on the other formed groups together. Selecting even fewer genes, by applying a higher significance threshold in the SAM analysis, did not improve structure attributable to time in the hierarchical trees. The pattern provided by the genes picked up by the SAM analyses suggests that we to a large extent identified the truly differentially expressed genes with a biological basis and not random genes.

K-mean clusters

The expression profile over the 9 time points for the 20 K-mean clusters is shown in Supplemental Figure II. As stated earlier, this procedure produced consistent results among runs. The complete gene list for each cluster is presented in the supplementary material (Supplemental Table I). The EASE scores for each cluster are given in supplemental Table II. Three cluster subsets were selected for further analyses. These represented 3 main stress responsive patterns found among the 20 K-mean clusters (Supplemental Fig II). These were early-upregulated genes (clusters 1, 2, 11, 12), early-downregulated genes (clusters 3, 4, 7, 13, 18), and late-upregulated genes (clusters 10, 19, 20). The expression profiles of these 3 main clusters are presented in Figure 2, and the EASE scores

Fig 2. Normalized expression profile over time in 3 main groups of heat stress-responsive genes identified by K-mean clustering. (a) Early-upregulated genes (265). (b) Early-downregulated genes (508). (c) Late-upregulated genes (226). These 3 main clusters constituted total 999 of 1222 significant differentially expressed genes (82%).

most highly significant for these 3 main clusters are presented in Table 2.

The K-mean procedure clustered differentially expressed genes not expected to respond to heat stress, particularly clusters 8, 16, and 20 (Supplemental Fig II). However, because this study only covered 2 time points (8 hours apart) repeatedly, we did not expect to detect all circadian rhythm genes. Genes regulated at other time points would not be detected. We have compared the genes differentially regulated over time in this study with known circadian rhythm genes. A study using similar microarrays found 392 genes to be rhythmically expressed in *D. melanogaster* males (McDonald and Rosbash 2001). Of these, 120 genes are found among our 1222 significantly differentially expressed genes. The clusters with highest proportion of known circadian rhythm genes were cluster 8 (48%) and cluster 17 (24%). That

cluster 17 had a high fraction of circadian genes was a surprise because this cluster does not show the predicted zigzag pattern among the 4 last time points. Thus, this information would have been missed if visual inspection was the only criterion for selection and analysis of clusters. Equally surprising, cluster 16 only had 2 known rhythm genes (3%), although this cluster showed a circadianlike pattern among the 4 last time points. The high proportion of rhythm genes in clusters 17 and especially 8 suggests that yet unknown rhythmically expressed genes may be found among unannotated genes in these clusters. The remaining circadian genes identified by McDonald and Rosbash (2001) and this study were distributed about equally in the rest of the clusters with occurrences of circadian genes between 0% and 17%, with most clusters below 10%. The occurrence in the early-up, early-down, and late-up main clusters were between 5%

Table 2 Ease scores for genes in each main cluster. (A) Early-up; (B) Early-down; and (C) Late-up. The genes belonging to each of the 3 main clusters were up loaded to and evaluated by the EASE application at the National Institutes of Health homepage. The genes were grouped by "System" (molecular function, biological process, or cellular component) and "Category" within each of these systems. "LH" represents number of genes in gene list assigned to category; "LT" represents number of genes in gene list assigned to system; "PH" represents number of all known genes assigned to category; and "PT" represents number of all known genes in that system (at present the number of genes assigned to each system is molecular function, 6508; biological process, 4221; cellular component, 3883). The representation of each category of genes (≥ 2) was evaluated by 2 criteria (Ease score and Fisher's exact test). Each test calculates the probability of detecting the actual detected number of genes in a category, by evaluating the proportion of genes in each gene list (here cluster) belonging to a category vs the proportion of genes belonging to this category out of all known genes. Thus, the ratio LH/LT is compared with the ratio PH/PT

System	Category	LH	LT	PH	PT	EASE score	Fisher's exact
A							
Molecular.function	chaperone activity	15	110	87	6508	1.15E-10	9.53E-12
Molecular.function	heat shock protein activity	10	110	37	6508	6.25E-09	2.99E-10
Biological.process	response to temperature	9	79	34	4221	1.16E-07	6.37E-09
Biological.process	response to stress	17	79	211	4221	9.4E-07	2.03E-07
Biological.process	response to heat	8	79	31	4221	1.03E-06	5.86E-08
Molecular.function	glutathione transferase activity	8	110	35	6508	1.39E-06	8.31E-08
Molecular.function	transferase activity, transferring alkyl or aryl (other than methyl) groups	9	110	56	6508	3.49E-06	3.16E-07
Biological.process	gluconeogenesis	3	79	3	4221	0.000999	6.31E-06
Biological.process	protein kinase cascade	6	79	47	4221	0.00158	0.000207
Biological.process	response to abiotic stimulus	10	79	159	4221	0.00239	0.000662
B							
Molecular.function	hydrolase activity	130	270	1430	6508	3.96E-22	1.71E-22
Molecular.function	catalytic activity	183	270	2998	6508	2.54E-13	1.69E-13
Molecular.function	peptidase activity	56	270	540	6508	1.37E-10	5.08E-11
Molecular.function	serine-type peptidase activity	36	270	261	6508	3.78E-10	1.02E-10
Cellular.component	membrane fraction	17	107	90	3883	1.23E-09	1.50E-10
Cellular.component	cell fraction	17	107	91	3883	1.46E-09	1.81E-10
Biological.process	catabolism	30	122	294	4221	1.57E-09	4.10E-10
Cellular.component	extracellular	25	107	228	3883	4.57E-09	1.05E-09
Molecular.function	serine-type endopeptidase activity	31	270	229	6508	1.32E-08	3.64E-09
Cellular.component	microsome	15	107	82	3883	2.53E-08	3.22E-09
Cellular.component	vesicular fraction	15	107	82	3883	2.53E-08	3.22E-09
Biological.process	proteolysis and peptidolysis	23	122	223	4221	2.1E-07	5.39E-08
Molecular.function	monosaccharide transporter activity	9	270	20	6508	6.28E-07	3.59E-08
Molecular.function	trypsin activity	14	270	97	6508	0.000159	4.04E-05
Biological.process	amino acid and derivative metabolism	12	122	107	4221	0.000205	4.78E-05
Biological.process	immune response	11	122	90	4221	0.000213	0.000045
Cellular.component	lysosome	5	107	17	3883	0.000949	6.88E-05
Molecular.function	glucose transporter activity	5	270	15	6508	0.00271	0.000252
Biological.process	eye pigmentation	5	122	28	4221	0.00772	0.00107
C							
Biological.process	catabolism	17	66	294	4221	5.81E-06	1.49E-06
Molecular.function	carrier activity	23	136	426	6508	5.54E-05	2.03E-05
Biological.process	macromolecule catabolism	14	66	243	4221	6.17E-05	1.58E-05
Biological.process	proteolysis and peptidolysis	13	66	223	4221	0.000117	2.96E-05
Biological.process	protein catabolism	13	66	229	4221	0.000151	3.92E-05
Molecular.function	peptidase activity	25	136	540	6508	0.000259	0.000111
Molecular.function	serine-type peptidase activity	16	136	261	6508	0.000304	9.75E-05
Molecular.function	endopeptidase activity	19	136	361	6508	0.000433	0.000163
Molecular.function	electrochemical potential-driven transporter activity	12	136	161	6508	0.000479	0.000125
Molecular.function	porter activity	12	136	161	6508	0.000479	0.000125
Molecular.function	serine-type endopeptidase activity	14	136	229	6508	0.000875	0.00028
Molecular.function	transporter activity	30	136	771	6508	0.000925	0.000476
Biological.process	response to light	5	66	43	4221	0.00403	0.000488

and 10% and thus equal to or below the overall mean of gene overlap (120 of 1222 genes). No significant functional groups were identified among these genes in any of the 3 main clusters. It is not possible from these data to determine whether these genes were truly regulated by

both heat stress and circadian rhythm or if circadian regulation by chance matched heat-induced profiles for a few genes. Another possibility is that the genes were identified as false positive in any of these 2 studies; however, the FDR in this study was estimated to be very low

(<1%). Finally, because the sexes differed between these 2 studies, it is possible that circadian expression to a certain extent is gender specific, further complicating the comparison.

Genes involved in reproduction were primarily restricted to cluster 16. In this study, genes with known functions almost entirely are connected to reproduction (8 of 13 annotated probe sets), showing a huge overexpression of this GO category. Thus, based on coregulation, reproductive genes are expected to be identified from the so far unknown genes in cluster 16.

An additional cluster showing an interesting profile should be mentioned here. Cluster 17 showed a large decrease at the first time point after stress (0.25 hour), whereafter expression returned to prestress levels. This small cluster had 3 genes (of the 13 annotated) connected to immune responses. The connection between the immune response and heat stress is not clear. In the full data set of significant probe sets, 29 genes belong to the GO category "defense response." These probe sets show both increased and decreased expression at various time points, with no consistent pattern among them. Genes involved in immune response, such as antibacterial peptides, are often found to be differentially expressed and seem to respond fast and many fold to changes in experimental conditions (eg, this study, Kristensen et al 2005) and many are even under circadian regulation (McDonald and Rosbash 2001).

The rhythmically expressed genes and genes involved in reproduction do, in addition to the heat stress-expressed genes, explain the main patterns consistently identified by K-mean clustering and the hierarchical trees. Most K-mean clusters include genes that are likely to be attributable to either heat exposure, circadian rhythm, or reproduction. Hierarchical trees show that the primary gene expression response to stress has returned to normal less than 8 hours after the heat treatment. The primary up- and downregulation closely mirror each other (Fig 2). The vast majority of up- or downregulated genes return to control levels within 4 hours after stress has ended (Fig 3). Another interesting feature seemingly general for the response to stresses is the proportion of genes up- and downregulated. As has been observed in other gene expression studies of the response to various other stressors, eg, old age and oxidative stress (Landis et al 2004), the number of genes downregulated after stress exceeded the number upregulated. However, this may be dependent on the organism studied, sex, and the stress type used (Jin et al 2001).

To test if the gene expression was affected by senescence or if a link between aging and stress resistance could be detected on gene expression level, we compared the 1222 significant genes from this study with the genes found to be affected during senescence by Jin et al (2001).

Many studies have addressed the connection between life span and the ability to tolerate stress (Tatar et al 1997; Norry and Loeschcke 2002; Hercus et al 2003), and it could be speculated that similar genes might be involved in responses induced by heat stress and age. Only 13 genes overlapped between the pure aging genes of Jin et al (2001) and the 1222 significant genes of this study. The 13 genes had a seemingly random distribution because maximum 2 genes per original K-mean cluster was found. The 3 main clusters shared 1 (early-up), 6 (early-down), and 5 (late-up) genes, respectively, and comprised less than 2% of the genes in all 3 clusters. No functional groups or genes with clear connection to stress responses were detected among these genes. This indicates that ageing had little effect on gene expression changes during the course of the experiment and provides no information about the connection between longevity and stress resistance.

The K-mean clusters have no information on actual expression values or fold changes. Therefore, we plotted the change in gene expression compared with time point -1 hour (before the heat treatment) against mean expression for the 3 major stress responsive clusters of genes (early-up, late-up, and early-down) (Fig 3). From these plots, it is evident that the genes that are upregulated generally are more changed than those downregulated. Figure 3 also shows that a few upregulated genes are not returning to near normal levels within 4 hours. After 8 hours post-heat stress, 8 probe sets from the early-up cluster remain more than 2-fold upregulated compared with before heat stress (CG6785, Hsp22, Hsp67Bc, Hsp68, Hsp70Bc, CG32130, EG:39E1.3, Hsp23). Not surprisingly, several well-known heat shock genes are found among these genes. Not much is known about the remaining 3 unannotated genes; however, 1 (CG32130) contains a BAG3 motif known to inhibit Hsp70 function, suggesting a function related to the stress response or chaperone activity. After 32 hours, only 2 probe sets remain more than 2 times upregulated (Hsp22, Hsp23). Thus, the expression of these 2 genes remain increased after other heat stress-upregulated genes (including other known *hsps*) have returned to prestress levels. For Hsp70, the gene expression profile after hardening match the protein level profile found by Dahlgaard et al (1998). If the gene expression levels reflect protein concentration for other Hsps as well, it is possible that Hsp22 and Hsp23 contribute to the long-lasting heat-induced heat resistance that lacks correlation with Hsp70 expression (Dahlgaard et al 1998; Sørensen et al 2003).

To evaluate the possible connections to heat stress exposure of the differentially expressed genes, we applied EASE to the 3 main clusters identified (Table 2). Furthermore, we compared our 3 main clusters with the genes found to be stress responsive in 3 other studies using the

Fig 4. Venn diagrams showing the distribution of genes identified as stress responsive in this study and the overlap with genes identified as stress responsive in 3 other array studies (Girardot et al 2004; Landis et al 2004; Kristensen et al 2005). Gray circle represents 1 of the 3 main clusters identified in this study (a. Early-up. b. Early-down. c. Late-up). The number of genes found in this and the compared study are given in parentheses. In the center of the gray circle, the number of genes found to be stress responsive in all 4 studies. A more detailed representation of the stress responsive genes shared among several stress studies can be found in the supplementary material (Fig III) and in Table 3.

Affymetrix platform for gene expression analyses. These studies are on gene expression induced by inbreeding (Kristensen et al 2005), oxidative stress, and aging (Girardot et al 2004; Landis et al 2004). The number of genes overlapping between our study and these 3 investigations are shown in Figure 4. A more detailed outline of the stress responsive genes identified by this and the 3 other studies can be found in the supplementary material (Supplemental Fig III). In the following analysis, we focused on genes found to be differentially expressed in this study (heat-specific response) and genes identified in this and in at least 2 of the 3 above-mentioned studies (general stress response) (Fig 4; Table 3). In total, 34 of 265 early-

upregulated, 123 of 508 early-downregulated, and 34 of 226 late-upregulated genes fulfilled this criterion for a general stress response (Table 3).

Early-up

The EASE scores of significantly heat early-upregulated genes are given in Table 2. Highest scoring GO included (number of genes in parentheses): chaperone activity (15), response to stress (17), glutathione transferase activity (8), protein kinase cascade (6), and gluconeogenesis (3). The genes from each of these categories are given in Table 4. The genes found to be early-upregulated by stress were

Fig 3. Time resolution of the 3 main groups of heat stress-responsive genes identified by K-mean clustering. Each cluster is plotted as mean expression (x-axis) vs fold change compared with time -1 hour (both axes are on a log2 scale). Red, green, and blue dots represent early-up, early-down, and late-up clusters, respectively. Each plot represents the fold change observed in 1 poststress time point compared with the -1 hour (control) situation.

Table 3 The stress responsive genes identified by this study and at least 2 of the 3 studies used for comparisons are shown (Girardot et al 2004; Landis et al 2004; Kristensen et al 2005). Table shows Affymetrix ID and corresponding gene symbol, maximum expression change after stress (increase for early- and late-up clusters and decrease for early-down cluster) in expression fold change and chromosomal position. Furthermore, the genes belonging to significant GO terms identified by EASE have been marked with appropriate GO term. Finally, known or suspected connection to (heat) stress has been indicated in italics

AFFYID	Symbol	FC	Position	Significant stress GO
A. Early-up				
143198.at	Hsp83	1.53	3L63B11	heat shock protein activity
142836.at	Hsp22	32.26	3L67B1	<i>Hsp</i>
153307.at	Hsp27	3.55	3L67B3	<i>Hsp</i>
153583.at	Hsp23	91.81	3L67B2	heat shock protein activity
153731.at	Hsp26	3.01	3L67B1-67B2	heat shock protein activity
149222.at	CG7130	3.96	3L79B2	heat shock protein activity
143197.at	Hsp68	53.93	3R95D11	heat shock protein activity
151036.f.at	Hsp70Bbb	4.51	3R87B14	<i>Hsp</i>
149782.at	Hsp70Bc	13.26	3R87B14-87B15	heat shock protein activity
153941.at	Fst	1.14	3R85E2	<i>response to cold</i>
152598.at	GstE1	3.47	2R55C6	glutathione transferase activity; <i>response to oxidative stress</i>
148411.at	CG6781	1.73	3L66D5	glutathione transferase activity
149756.at	GstD2	4.14	3R87B8	glutathione transferase activity
149759.at	GstD5	1.60	3R87B8	glutathione transferase activity
142479.at	Gst3-2	1.76	2R55D4	
151834.at	CG32130	20.76	3L70B2	<i>molecular chaperone regulator-3 (Mus musculus)</i>
143041.l.at	CG11700	1.77	X5E4-5E5	<i>stress response (Saccharomyces cerevisiae)</i>
143299.at	Pepck	1.86	2R55D3	<i>gluconeogenesis</i>
143062.at	ade3	1.23	2L27D4	
152056.at	ref(2)P	2.29	2L37F1	
143690.at	Fdxh	1.86	3L67B1	
143802.at	CG14906	1.22	3R89D6	
154053.at	Trxr-1	1.25	X7D22-7E1	<i>antioxidant activity</i>
154184.at	aay	1.87	3L67B4-67B5	
141233.at	CG5966	4.22	X5D1	
154938.at	CG1583	1.66	X7D16-7D17	
145054.at	CG32597	1.53	X12E4-12E5	
152901.at	CG10383	1.60	2L37A1	
146757.at	CG2065	4.40	2R43E9	
147118.at	CG12374	1.10	2R49D3	
147525.at	CG16898	3.02	2R56F4	
152228.at	Xbp1	1.45	2R57C3-57C4	<i>transcription factor activity</i>
153963.at	CG16971	1.18	3L61B2	
150837.at	Obp99b	1.83	3R99B8	
B. Early-down				
151533.at	CG31148	2.42	3R95A8	catalytic activity
151539.at	CG30359	1.10	2R44D1	
143147.at	Eip71CD	1.06	3L71D3-71D4	catalytic activity, carboxylic acid metabolism
143151.at	Est-6	2.07	3L69A1	catalytic activity
141472.at	Glt	2.00	2L29E3	
143242.at	LvpH	1.63	2R44D1	catalytic activity
152200.at	ry	1.44	3R87D9	catalytic activity; <i>eye pigmentation</i>
143401.at	Uro	1.55	2L28C3	catalytic activity
152559.at	v	1.79	X9F11	catalytic activity
151767.at	Nmdmc	1.26	3R85C3	catalytic activity, carboxylic acid metabolism
151960.at	Myo61F	1.23	3L61F6	
153423.at	Pbprp2	1.31	X19D1	
143775.at	iotaTry	1.73	2R47F5	catalytic activity, peptidase activity
143872.at	Aph-4	1.48	3R100B1	catalytic activity, alkaline phosphatase activity
153432.at	Thor	1.58	2L23F6	<i>immune response</i>
153260.at	Mipp1	1.34	3L73A7-73A9	catalytic activity
142950.at	CG7953	1.76	2L34D7	
142951.at	CG7916	2.02	2L34D7	
144310.at	CG8997	2.01	2L34D7	
152148.at	CG16887	1.64	2L34D6-34D7	
152743.at	CG14629	1.29	X1E1	
144369.at	CG11378	1.74	X1E3	
144614.at	CG14439	1.56	X6C11-6C12	
144675.at	CG2254	1.30	X7D5	catalytic activity
144806.at	CG32687	1.47	X9D2	
144853.at	CG15201	1.90	X10A8	
145153.at	CG33173	1.40	X13E15-13E17	catalytic activity

Table 3 Continued

AFFYID	Symbol	FC	Position	Significant stress GO
145202_at	CG9673	2.02	X_15A2-15A3	catalytic activity, peptidase activity
145235_at	CG5162	1.39	X_15F4	
145412_at	CG12656	1.76	X_19B1	
145449_at	NPC1b	1.73	X_19E5	
152999_at	Iris	1.66	2L_21F1	
145684_at	CG17224	1.38	2L_23C5	catalytic activity
142390_at	CG3246	1.39	2L_23F3	
145779_at	CG3244	1.62	2L_25A6	
152567_at	CG18585	1.50	2L_28C1	
145969_at	CG7025	1.06	2L_28C1	catalytic activity, peptidase activity
145970_at	TepII	1.39	2L_28C1	antibacterial peptide activity
152369_at	CG13095	1.64	2L_29D1-29D3	catalytic activity, peptidase activity
142147_at	CG9466	1.76	2L_29F1	catalytic activity; lysosome
146050_at	CG9468	3.11	2L_29F1	catalytic activity; lysosome
146103_at	CG17633	1.55	2L_30C8	catalytic activity, peptidase activity
146138_at	CG5390	1.31	2L_31D1	catalytic activity, peptidase activity
146184_at	CG17108	1.42	2L_32A4	
146214_at	CG16743	1.44	2L_32C1	
146233_at	hgo	1.24	2L_32D5	catalytic activity, carboxylic acid metabolism
146257_at	CG14935	1.21	2L_33A4	catalytic activity
146311_at	CG9928	1.37	2L_34A6	
146323_at	CG5945	1.96	2L_34A11	
146387_at	CG6012	1.29	2L_36B2	catalytic activity
146449_at	Irk3	1.30	2L_37A1	
146486_at	CG13086	1.83	2L_37D2	
141465_at	CG10026	1.43	2L_37E1	
152470_at	CG10189	1.18	2L_37E4	
146526_at	CG10680	1.49	2L_38A8	
146569_at	CG9259	1.99	2L_39A1	
146660_at	CG7882	1.68	2R_42A1	glucose transporter activity
151882_at	CG3270	1.41	2R_42C5	
146742_at	CG30502	1.66	2R_43C3	
146808_at	CG8693	1.67	2R_44D1	catalytic activity
154318_at	Cyp6a13	1.61	2R_44D3	cytochrome P450 activity
151956_at	CG1809	1.34	2R_45F3	catalytic activity, alkaline phosphatase act
147014_at	lambdaTry	1.92	2R_47F3	catalytic activity, peptidase activity
147028_Lat	CG30035	1.28	2R_48B6-48B7	glucose transporter activity
154772_at	sug	1.39	2R_49E1	
147235_at	CG8093	2.42	2R_51E7	catalytic activity; triacylglycerol lipase activity
147265_at	CG8249	1.40	2R_52C8	glucose transporter activity
147319_at	CG8317	1.59	2R_53C9	
147336_at	CG6426	1.17	2R_53D14	
151800_at	CG4847	1.38	2R_54C1	catalytic activity, peptidase activity
142386_at	CG6484	1.50	2R_54C10	glucose transporter activity
147434_at	CG17522	1.34	2R_55C6	catalytic activity; glutathione transferase activity
154747_at	Jheh3	1.21	2R_55F8	catalytic activity; microsome
147514_at	Obp56a	1.45	2R_56E2	
152785_at	CG11200	1.18	2R_56F16	catalytic activity
147544_at	Obp57c	1.39	2R_57A4	
147564_at	CG16799	1.33	2R_57A9	catalytic activity; immunity protein activity
152027_at	Ance-5	2.34	2R_60E5	catalytic activity, peptidase activity
152648_at	CG2736	1.60	2R_60E10-60E11	defense response; scavenger receptor activity
152830_at	CG3344	1.79	3L_61C8	catalytic activity, peptidase activity
148048_at	CG16762	1.83	3L_62E3	
148054_at	CG16985	1.45	3L_62E7	
148055_at	CG16986	1.22	3L_62E7	
148228_at	CG10592	1.77	3L_64D5	catalytic activity, alkaline phosphatase activity
151980_at	CG5150	1.87	3L_64D5	catalytic activity, alkaline phosphatase activity
148251_at	CG10477	1.62	3L_65A3	catalytic activity, peptidase activity
152006_at	CG10472	1.35	3L_65A5	catalytic activity, peptidase activity
148422_at	CG5804	2.00	3L_66E1	carboxylic acid metabolism
148428_at	CG13309	2.51	3L_66E1	
148678_at	CG14120	1.82	3L_69E8	
148963_at	CG6298	1.69	3L_74D3-74D4	catalytic activity, peptidase activity
148976_at	CG5582	1.99	3L_75A4	
151832_at	Oat	1.69	3L_76B11	catalytic activity, carboxylic acid metabolism
141437_at	CG5618	1.32	3L_77B5	catalytic activity

Table 3 Continued

AFFYID	Symbol	FC	Position	Significant stress GO
151894_at	CG18249	1.73	3R_84F4	
149491_at	CG7443	1.81	3R_84F12	
149610_at	CG12813	2.42	3R_85F8	
141327_at	CG8773	1.76	3R_87E4	catalytic activity, peptidase activity
152779_at	CG8774	1.80	3R_87E4-87E5	catalytic activity, peptidase activity
149938_at	CG3984	1.50	3R_88E2	
149964_at	CG14872	1.58	3R_88F7	
154786_at	CG18522	1.45	3R_88F7	
152227_at	CG6126	1.26	3R_89B7	
141483_at	CG17836	1.30	3R_91D4-91D5	
150191_at	CG3734	2.16	3R_92A3	catalytic activity, peptidase activity
152841_at	CG5023	1.35	3R_92D2-92D3	
151819_at	CG3301	1.36	3R_93D2-93D3	catalytic activity
150403_at	CG6660	1.18	3R_94D3-94D4	
150421_at	CG6726	1.42	3R_94D13	catalytic activity
150424_at	CG6733	1.69	3R_94D13	catalytic activity
150482_at	CG13607	1.57	3R_95E1	
152513_at	CG10513	2.07	3R_96C8-96C9	
150576_at	CG10514	2.08	3R_96C9	
150577_at	CG31104	1.07	3R_96C9	
151781_at	CG5107	1.43	3R_96E2	
150699_at	CG6295	1.49	3R_97D13	catalytic activity; triacylglycerol lipase activity
150732_at	Cyp6a18	1.86	3R_98A2	cytochrome P450 activity
142206_s.at	CG14527	1.77	3R_98F1	catalytic activity, peptidase activity
142413_at	CG11314	1.46	3R_100A3	
142414_at	CG11315	1.27	3R_100A3	
151069_at	CG5791	1.62	3R_94A2	
151094_at	CG3348	2.16	3R_97F1	
151209_at	CG15065	1.72	2R_55C6	
C. Late-up				
143201_at	ImpL3	1.90	3L_65A11	glycolysis
143243_at	LvpL	2.89	2R_44D1	alpha-glucosidase activity; glucose metabolism
143276_at	MtnA	7.13	3R_85E9	
151805_at	Reg-3	1.93	X_8D8-8D9	
144701_at	CG12116	2.70	X_7F1	
152289_at	CG3106	5.84	X_8F6	
142953_at	CG9675	1.86	X_15A1-15A2	peptidase activity
145816_at	Cyp4ac2	1.58	2L_25D2	cytochrome P450 activity; microsome
152813_at	CG9497	1.94	2L_26C3	
152958_at	Cyp4d21	3.41	2L_28B4	cytochrome P450 activity; microsome
153012_at	CG7300	2.66	2L_32A4	
146190_at	CG17124	2.32	2L_32A5	protein phosphatase inhibitor activity
146763_at	CG1946	5.34	2R_43E12	
146991_at	CG9080	4.23	2R_47E1	
147093_at	CG8834	1.62	2R_49A4	long-chain fatty acid transporter activity
147188_at	CG18327	2.47	2R_50E1	
147636_at	CG13492	1.31	2R_58A2	
155144_at	Cct1	2.08	3L_62A9	
148259_at	CG6602	3.76	3L_65A5	
148332_at	CG8562	2.12	3L_65F11	peptidase activity
152794_at	CG8560	1.95	3L_65F11	peptidase activity
142402_at	CG16749	40.84	3R_85D11	peptidase activity
154821_at	CG8147	1.50	3R_85D18	
149615_at	CG3940	2.13	3R_85F12	
149618_at	1(3)IX-14	2.42	3R_85F14-85F15	peptidase activity
152573_at	CG32919	3.78	3R_87B9	
149899_at	Cyp313a1	2.94	3R_88C4	cytochrome P450 activity; microsome
150192_at	CG18493	16.39	3R_92A3	
150193_at	CG3739	2.15	3R_92A3	peptidase activity
150412_at	CG13833	2.89	3R_94D10	
141409_at	CG6164	2.32	3R_95E5	
150791_at	CG14528	1.28	3R_98F1	peptidase activity
141310_at	CG9682	2.31	3R_99E3	
150901_at	CG9702		3R_99F9	

Go, Gene Ontology.

Table 4 List of most significant functional groups and the genes in those groups responding to heat stress in *Drosophila*. Functional ontology and the genes in each group are given for each of 3 main patterns of response to heat (early-up, early-down, and late-up). The functional ontology was identified by EASE, and the significant groups represent functional groups where genes are overrepresented compared with what could be expected in a random gene list. These 3 main responses of expression change are discussed in the text

Functional ontology (Early-up)	
Chaperone activity	CG7945, Heat shock gene 67Ba, Heat shock gene 67Bc, Heat shock protein 68, Heat shock protein 83, Cdc37, DnaJ-like-1, Hsp70/Hsp90 organizing protein homolog, CCS, CG7130, Heat shock protein 70Bc, Heat shock protein 23, Heat shock protein 26, Hsc70Cb, CG11035
Response to stress	Heat shock factor, Heat shock gene 67Ba, Heat shock gene 67Bc, Heat shock protein 68, Heat shock protein 83, Rad51D, CG4049, PEK, Pten, Heat shock protein 70Bc, Glutathione S transferase E1, immune deficiency, Rev1, Heat shock protein 23, Heat shock protein 26, Frost, CG11403
Glutathione transferase activity	CG17524, CG17531, CG6776, CG6781, Glutathione S transferase D10, Glutathione S transferase D2, Glutathione S transferase D5, Glutathione S transferase E1
Protein kinase cascade	anterior open, Cdc37, raw, Signal transducing adaptor molecule, Jun-related antigen, Src oncogene at 42A
Gluconeogenesis	Phosphoenolpyruvate carboxykinase, CG10924, Phosphoglucose isomerase
Functional ontology (Early-down)	
Monosaccharide transporter activity	CG15406, CG15408, CG7882, CG30035, CG30035, CG8249, CG6484, CG1213, CG14606, CG31100
Lysosome	CG9466, CG9468, G5322, CG6206, CG17119
Microsome (Cytochrome P450 activity)	Cytochrome P450-4d1, Cytochrome P50-4e3, Cyp9h1, Cyp4d14, Cyp6d2, Cyp4d20, Cyp6a18, Resistance to DDT, 152937_at, Cytochrome P450-4ae1, 153031_at, 154318_at, Juvenile hormone epoxide hydrolase 3
Response to biotic stimulus	broad, Gram-negative bacteria binding protein 3, Lysozyme P, Immune-induced protein 2, Peptidoglycan recognition protein SA, Persephone, Thiolester containing protein II, necrotic, Gram-negative bacteria binding protein 1, Immune-induced protein 4, Gram-negative bacteria binding protein 2, CG12789, CG16799, CG2736, CG1102
Eye pigmentation	karmoisin, rosy, vermilion, scarlet, Punch
Functional ontology (Late-up)	
Electrochemical potential-driven transporter activity	CG10806, Picot, CG7255, CG7571, CG7442, CG11739, CG31090, CG9702, CG2196, 1(2)08717, Equilibrative nucleoside transporter 2, Na ⁺ -driven anion exchanger 1
Response to light	timeless, Rhodopsin 6, Guanyl cyclase alpha-subunit at 99B, neither in-activation nor afterpotential A, Rhodopsin 5
Transporter activity	Ent2, Ndae1, Syx1A, Vha26, CG1628, CG1718, CG4822, CG10806, CG6870, CG11163, CG8834, CG18327, CG18324, Ih, Picot, CG4797, CG7255, CG7571, CG7442, CG11739, CG31116, CG8790, CG4562, CG31090, CG5646, CG9702, CG2196, CG2165, Vha68-1, 1(2)08717
Catabolism	CG6048, CG32834, CG9676, CG1304, Cyp4ac2, CG5045, CG8329, CG16749, CG5896, CG7829, Ecdysone-inducible gene L3, zeta-Trypsin, Serine protease 12, Serine protease 6, Serine protease like 89E, Chitinase-like, Glutamate dehydrogenase

to a larger degree than those expected, ie, genes known to be heat responsive and to belong to the group of genes having chaperone activity. However, several chaperone genes not annotated to be heat or stress responsive were also identified. The GO term response to heat only includes 34 genes, whereas this study suggests that more than 200 are immediately upregulated as a response to heat exposure. The GOs identified as significant among genes early-upregulated are also among those found in other studies of stress-induced gene expression. Genes found to be differentially expressed in this and at least 2 of the 3 other comparable studies included the ontologies "stress/chaperone" genes, and "glutathione transferase" genes (Fig 4; Table 3). However, no genes from the "pro-

tein kinase cascade" and only 1 of the 3 "gluconeogenesis" genes were found in the comparison with other stress studies, although these categories were much more frequently represented among the original heat early-up cluster (Table 2). This suggests these 2 processes may be more specific for heat stress conditions, although the former processes could be considered to be more general stress responses. The significance of the gluconeogenesis genes is underlined by the fact that only 3 genes are found in this category, all of which were found to be upregulated by heat stress in this study. The relation between this process and stress remains unclear. Although metabolism in general is downregulated after stress, the process of gluconeogenesis, which produces glucose from

pyruvate, seems upregulated. Glucose is needed as an energy source for nerve tissues, and possibly gluconeogenesis is upregulated by stress while the remaining metabolism is downregulated to secure an energy source for important and sensitive nerve tissue.

Early-down

The genes found to be early-downregulated by heat were primarily related to metabolism, with many genes in the GO catalytic activity (188), hydrolase activity (129), and peptidase activity (57). The EASE scores of significantly early-downregulated genes are given in Table 2. Highest scoring GOs include monosaccharide transporter activity (9), lysosome (5), microsome (cytochrome P450 activity) (13), response to biotic stimulus (15), eye pigmentation (5). The genes from each of these categories are given in Table 4.

There is a large degree of overlap between genes found to be downregulated after stress in this study and the genes found to respond to stress in general. The number of genes commonly differentially expressed in comparisons between this and the other studies is much larger with down- than with upregulated genes (Fig 3). This indicates that the effects of stress in general are more uniform among stress types in terms of what is closed down (decreased expression), whereas the response in terms of increased expression includes more specific groups of genes. Interestingly, a high number of "eye pigmentation" genes are found in this study. This corresponds with data suggesting that the same category of genes respond to selection for increased heat resistance (M.M. Nielsen, J.G. Sørensen, J. Justesen, M. Kruhøffer, and V. Loeschcke, unpublished observations).

Late-up

The genes found to be late-upregulated were also related to metabolism to a large degree. The EASE scores of significantly late-upregulated genes are given in Table 2. Highest scoring GOs include: electrochemical potential-driven transporter activity (12), response to light (5), transporter activity (30), catabolism (17). The genes from each of these categories are given in Table 4.

Late-upregulated genes seem to share characteristics with early-downregulated genes, with many genes involved in metabolism. In Figure 3, one can see that the genes defined as early-down and late-up to some degree resemble each other. However, also specific groups of genes are found among the late-up genes that are not found to the same degree in the early-down. The similarity between late-up and early-downregulated genes is especially evident in the genes shared with the other stress studies. Here, 2 GOs were significant: "peptidase

activity" and "cytochrome P450 activity/microsome." These 2 are also significant in the early-down cluster of genes. Still, when all late-up genes are considered, some additional categories of genes show up. Here, interestingly 5 additional genes were connected to the detection and response to light showed up.

From the comparisons with other stress studies, some general patterns were observed. In all 3 comparisons (Girardot et al 2004; Landis et al 2004; Kristensen et al 2005), only relatively few genes were found to be shared among the early-upregulated genes, whereas relatively more genes are found to be shared among the late-up and early-downregulated genes (Fig 4). A few well-known Hsps are often, if not always represented, as upregulated in the stress response, but in addition to these, the genes expressed seem not always to be the same. The lack of a large common response seems to be a general feature of the stress response of more complex organisms. For example, Murray et al (2004) found an absence of a common stress response to multiple types of stress in human cells in contrast to a large common response in yeast (Gasch et al 2000). Gasch et al (2000) investigated the gene expression responses of a large number of genes in yeast and found ~900 to respond to a range of environmental perturbations. Although the response to environmental stress in yeast and heat stress in *Drosophila* seems to differ in terms of the specificity and number of genes expressed, some similarities are striking. In both cases, the immediate response consists of surprisingly parallel up- and downregulated genes, with around twice the number of genes repressed as induced (Gasch et al 2000 and Fig 2 a,b). Also, the physiological functions of the genes are similar (eg, carbohydrate metabolism, cellular defense, protein folding, energy production etc.). The seemingly more specialized response in *Drosophila* probably arose because complex organisms (like whole organism *Drosophila*) have specialized tissue and organs and thus also cell-to-cell communication, which is expected to play an important role in maintaining homeostasis. Thus, using whole organism *Drosophila*, the results are averages across many tissues, and the specific reactions of small tissues can be easily overlooked. However, we regard this as an advantage, because the aim here was to identify general responses, ie, responses taking place in several tissues, thus not being diluted below detection limits. Also, using whole organisms (compared with cell culture or unicellular organism studies, where specific/single tissues can be investigated) can account for more complex features, eg, signaling, and traits linked to reproduction and behavior. These features are likely to be overlooked in assays on in vitro cell cultures compared with whole organism studies (Murray et al 2004).

Apart from changes in gene expression level, stress induces other changes as well. Several metabolites and hor-

mones are involved in stress responses and confer tolerance to stress exposures. Trehalose has been shown to protect from damage of a range of environmental stresses in yeast, including heat, and also accumulates after exposure to some types of stress exposures (Alexandre et al 2001). In *Drosophila*, trehalose protects against anoxia stress (Chen et al 2002), and it could be hypothesized that it could also protect against heat stress. An amino acid that plays an important role in the *Drosophila* stress response is tyrosine, which acts as a precursor of several stress hormones in insects including dopamine, octopamine, and tyramine (Sukhanova et al 1997; Hirashima et al 2000). Changes in tyrosine levels have been shown to correlate with the changes in the levels of stress hormones during heat stress (Hirashima et al 2000). In a study of multiple metabolites by nuclear magnetic resonance, the changes after mild stress, severe stress, and mild stress (hardening) followed by severe stress were investigated (A. Malmendal, J. Overgaard, J.G. Bundy, J.G. Sørensen, N.C. Nielsen, V. Loeschcke, and M. Holmstrup, unpublished observations). Trehalose showed no clear tendency for regulation after the heat stress. However, tyrosine was markedly elevated in the severely stressed flies although the levels were intermediate after mild stress. We identified 15 genes found to be involved in trehalose (Treh, CG5171, CG5177, Tps1, CG6262, CG16965) and tyrosine (Aph-4, Tbh, ple, Bc, amd, Ddc, Hn, CG11796, CG9674) synthesis and removal. Four of these genes were found to be stress responsive (amd: dopa decarboxylase, ple: Tyrosine hydroxylase, Aph-4: Alkaline phosphatase 4 and CG16965: trehalase activity). One (Aph-4) was found in the early-down cluster, whereas none were identified in the early-up, and late-up clusters. However, all 4 genes showed an immediate downregulation after stress, followed by less clear expression patterns at the later time points. Thus, although this study did not detect a clear response to stress in the processes connected to trehalose and tyrosine synthesis and metabolism, the observed changes might indicate responses to stress in these particular genes.

This study describes the pattern of up- and downregulation after heat stress in *Drosophila*. Numerous genes are up- or downregulated, at various points in time after stress, but the vast majority of the genes responding to stress returned to prestress levels within 8 hours after stress. The genes that were found to respond to heat stress were both genes known to respond to heat stress and new candidate genes. The gene expression analysis, by microarrays, used in this study successfully identified genes responding to heat stress with high certainty. This technique has proven to be a powerful technique to identify genes and processes involved in physiological processes and to provide candidates for further studies,

which is expected to yield important biological information in the future.

ACKNOWLEDGMENTS

We are grateful to Doth Andersen, Mia Skov Jensen, and Bente Devantié for excellent technical assistance, to Leif Schausser and Jens Ledet Jensen for helpful discussion of statistics, and to Torsten N. Kristensen and 2 anonymous reviewers for helpful suggestions to the manuscript. We thank Hinnerk Boriss for many discussions during the planning phase of the experiment and to and many other postdocs and students in the "Aarhus stress group" who gave a hand during peak times of the experimental part of this work. The work was supported by the Danish Natural Sciences Research Council by a Centre grant.

REFERENCES

- Alexandre H, Ansany-Galeote V, Dequin S, Blondin B. 2001. Global gene expression during short term ethanol stress in *Saccharomyces cerevisiae*. *Yeast* 18: S177.
- Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate—a practical and powerful approach to multiple testing. *J R Stat Soc B* 57: 289–300.
- Bublioy OA, Loeschcke V. 2005. Correlated responses to selection for stress resistance and longevity in a laboratory population of *Drosophila melanogaster*. *J Evol Biol* 18: 789–803.
- Chen QF, Ma E, Behar KL, Xu T, Haddad GG. 2002. Role of trehalose phosphate synthase in anoxia tolerance and development in *Drosophila melanogaster*. *J Biol Chem* 277: 3274–3279.
- Dahlgaard J, Loeschcke V, Michalak P, Justesen J. 1998. Induced thermotolerance and associated expression of the heat-shock protein Hsp70 in adult *Drosophila melanogaster*. *Funct Ecol* 12: 786–793.
- Dyrskjot L, Thykjaer T, Kruhoffer M, et al. 2003. Identifying distinct classes of bladder carcinoma using microarrays. *Nat Genet* 33: 90–96.
- Eisen MB, Spellman PT, Brown PO, Botstein D. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 95: 14863–14868.
- Feder ME, Hofmann GE. 1999. Heat-shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. *Annu Rev Physiol* 61: 243–282.
- Gasch AP, Spellman PT, Kao CM, et al. 2000. Genomic expression programs in the response of yeast cells to environmental changes. *Mol Biol Cell* 11: 4241–4257.
- Gene-Ontology-Consortium. 2001. Creating the gene ontology resource: design and implementation. *Genome Res* 11: 1425–1433.
- Girardot F, Monnier V, Tricoire H. 2004. Genome wide analysis of common and specific stress responses in adult *Drosophila melanogaster*. *BMC Genomics* 5: 74.
- Hercus MJ, Loeschcke V, Rattan SIS. 2003. Lifespan extension of *Drosophila melanogaster* through hormesis by repeated mild heat stress. *Biogerontology* 4: 149–156.
- Hirashima A, Sukhanova MJ, Rauschenbach IY. 2000. Biogenic amines in *Drosophila virilis* under stress conditions. *Biosci Biotechnol Biochem* 64: 2625–2630.
- Hosack DA, Dennis G, Sherman BT, Lane HC, Lempicki RA. 2003. Identifying biological themes within lists of genes with EASE. *Genome Biol* 4: P4.

- Jin W, Riley RM, Wolfinger RD, White KP, Passador-Gurgel G, Gibson G. 2001. The contributions of sex, genotype and age to transcriptional variance in *Drosophila melanogaster*. *Nat Genet* 29: 389–395.
- Kristensen TN, Sørensen P, Kruhøffer M, Pedersen KS, Loeschcke V. Genome wide analysis on inbreeding and effects on gene expression in *Drosophila melanogaster*. *Genetics* 171: 157–167.
- Landis GN, Abdueva D, Skvortsov D, et al. 2004. Similar gene expression patterns characterize aging and oxidative stress in *Drosophila melanogaster*. *Proc Natl Acad Sci U S A* 101: 7663–7668.
- Lindquist SL. 1986. The heat-shock response. *Annu Rev Biochem* 55: 1151–1191.
- McDonald MJ, Rosbash M. 2001. Microarray analysis and organization of circadian gene expression *Drosophila*. *Cell* 107: 567–578.
- Morimoto RI, Jolly C, Satyal S, Mathew A, Shi Y, Kitagawa K. 1999. Molecular chaperones and the heat shock response. *Br J Cancer* 80: S18.
- Murray JI, Whitfield ML, Trinklein ND, Myers RM, Brown PO, Botstein D. 2004. Diverse and specific gene expression responses to stresses in cultured human cells. *Mol Biol Cell* 15: 2361–2374.
- Norry FM, Loeschcke V. 2002. Longevity and resistance to cold stress in cold-stress selected lines and their controls in *Drosophila melanogaster*. *J Evol Biol* 15: 775–783.
- Parsell DA, Lindquist S. 1993. The function of heat-shock proteins in stress tolerance: degradation and reactivation of damaged proteins. *Annu Rev Genet* 27: 437–496.
- Pauli D, Arrigo AP, Tissieres A. 1992. Heat shock response in *Drosophila*. *Experientia* 48: 623–629.
- Pockley AG. 2003. Heat shock proteins as regulators of the immune response. *Lancet* 362: 469–476.
- Prohászka Z, Füst G. 2004. Immunological aspects of heat-shock proteins—the optimum stress of life. *Mol Immunol* 41: 29–44.
- Ritossa F. 1962. A new puffing pattern induced by temperature shock and DNP in *Drosophila*. *Experientia* 18: 571–573.
- Ritossa FM. 1964. Experimental activation of specific loci in polytene chromosomes of *Drosophila*. *Exp Cell Res* 35: 601–607.
- Saeed AI, Sharov V, White J, et al. 2003. TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* 34: 374–378.
- Sørensen JG, Kristensen TN, Loeschcke V. 2003. The evolutionary and ecological role of heat shock proteins. *Ecol Lett* 6: 1025–1037.
- Spellman PT, Sherlock G, Zhang MQ, et al. 1998. Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. *Mol Biol Cell* 9: 3273–3297.
- Sukhanova MJ, Shumnaya LV, Grenback LG, Gruntenko NE, Khlebodarova TM, Rauschenbach IY. 1997. Tyrosine decarboxylase and dopa decarboxylase in *Drosophila virilis* under normal conditions and heat stress: genetic and physiological aspects. *Biochem Genet* 35: 91–103.
- Tatar M, Khazaeli AA, Curtsinger JW. 1997. Chaperoning extended life. *Nature* 390: 30.
- Tavazoie S, Hughes JD, Campbell MJ, Cho RJ, Church GM. 1999. Systematic determination of genetic network architecture. *Nat Genet* 22: 281–285.
- Tusher VG, Tibshirani R, Chu G. 2001. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 98: 5116–5121.
- Wu ZJ, Irizarry RA. 2004. Preprocessing of oligonucleotide array data. *Nat Biotechnol* 22: 656–658.