EVOLUTIONARY BIOLOGY

Increased proteomic complexity in *Drosophila* hybrids during development

Casimir Bamberger,[1] Salvador Martínez-Bartolomé,[1] Miranda Montgomery,[2] Mathieu Lavallée-Adam,[3] John R. Yates III[1]*

Cellular proteomes are thought to be optimized for function, leaving no room for proteome plasticity and, thus, evolution. However, hybrid animals that result from a viable cross of two different species harbor hybrid proteomes of unknown complexity. We charted the hybrid proteome of a viable cross between *Drosophila melanogaster* females and *Drosophila simulans* males with bottom-up proteomics. Developing hybrids harbored 20% novel proteins in addition to proteins that were also present in either parental species. In contrast, adult hybrids and developmentally failing embryos of the reciprocal cross showed less additional proteins (5 and 6%, respectively). High levels of heat shock proteins, proteasome-associated proteins, and proteasomal subunits indicated that proteostasis sustains the expanded complexity of the proteome in developing hybrids. We conclude that increased proteostasis gives way to proteomic plasticity and thus opens up additional space for rapid phenotypic variation during embryonic development.

INTRODUCTION

The human genome encodes a surprisingly high number of proteins (*1*) and proteoforms (*2*). However, the number of proteoforms simultaneously observed in cells and tissues (*3*, *4*) is lower than the complexity encoded in the genome, a result believed to stem from a constant optimization of the proteome to enable the highest efficiency and functionality. This ceaseless optimization leaves limited space for the proteome to evolve. One possible way to maintain a minimal plasticity for the proteome may be through chaperone proteins that fold and stabilize other proteins (*5*). Together with protein trafficking and degradation in the proteasome, protein folding monitors the proteome for any presence of misfolded proteins. This specialized network of proteins secures overall proteostasis of the proteome (*6*, *7*). For example, inbreeding of *Drosophila* is accompanied by an up-regulation of heat shock protein Hsp70, which is most likely needed to stabilize proteins with recessive mutations that are otherwise deleterious (*8*, *9*). The proteostasis network and the heat shock protein Hsp90 specifically can support the continued presence of increasingly mutated proteins in tumorigenic cells, thus providing additional, unwanted plasticity (*10*).

It is difficult to test the upper limits of proteome plasticity in whole organisms because it is experimentally challenging to preserve viability while increasing the diversity of proteins and proteoforms. To improve the chances for viability, two closely related species can be crossed to combine two different but related genomes in a single zygote so that one species-specific haploid genome of each parent contributes to the diploid genome in the hybrid organism. Depending on the species, the resultant hybrids may be viable and survive into adulthood with a proteome that is of principally higher yet unknown complexity.

Not all interspecies crosses are successful, and many hybrid organisms fail to develop. For example, recent research has shown that an uncontrolled expression of transposons in the germ line of adult hybrid fruit flies establishes a species barrier (*11*, *12*). Whereas incomplete transposon suppression is an effective mechanism to perpetuate a species barrier, extensive genetic screens have identified individual genes and encoded proteins that confer lack of viability to hybrid sons, but not hybrid daughters, of a cross between two closely related fruit flies, *Drosophila melanogaster* (*D. mel.*) and *Drosophila simulans* (*D. sim.*) (*13*). Male offspring produced by crossing *D. mel.* (females) ♀ × *D. sim.* (males) ♂ die early in larval instar stage 3, but death can be delayed by mutations in one of two genes, *Lhr* (lethal hybrid rescue) or *Hmr* (hybrid male rescue), which form a pair of Dobzhansky-Muller genes. However, even after rescue at developmental stage 3 in Hmr^{-} or Lhr^{-} hybrids, termination of development still occurs at later developmental stages. These genetic screens highlight singular causes for failed development due to an incompatibility in the two species-specific proteomes but do not reveal the level of increased proteome complexity that is tolerated in surviving female hybrids. The simultaneous expression of species-specific proteins with high-sequence similarity in hybrids may represent an additional source of proteome incompatibility during development and in adulthood.

The *Drosophila* proteome undergoes marked changes during early development. In oogenesis, a maternally derived transcriptome and proteome are expressed in nurse cells and stored in oocytes to guide and control the initial development of the *Drosophila* embryo. The expression of genes from the paternal and maternal genome in the zygote is initiated during early embryogenesis ~2 hours after fertilization (developmental stage 4), with the earliest transcription events occurring at the 8th to 10th nuclear division in *D. melanogaster* (*14*). After this stage, the compatibility of the RNAs and proteins expressed from the two genomes determines the stability of the hybrid organism during further development. While the ratio of species-specific alleles (orthologs) in hybrids is 1:1, excluding gene duplications or deletions, RNA expression from both genomes depends on cis and trans transcriptional regulation and results in substantial variations of RNA transcript levels between parents and hybrids (*15*–*17*). In addition, protein translation, folding, and degradation may also influence the complexity of the proteome in hybrids.

We used bottom-up proteomics to determine species-specific protein levels in *Drosophila* hybrids during development and adulthood. Previous proteomic studies of a limited number of species-specific proteins in *Xenopus* hybrids focused on RNA transcription and translation initiation during embryonic development (*18*–*20*). Here, we measured and compared the proteome of *Drosophila* hybrids to find out whether plasticity of the proteome is different between tissue-determining early

[1]Department of Molecular Medicine, Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA. [2]Altman Clinical and Translation Research Institute, 9452 Medical Center Drive, University of California San Diego, La Jolla, CA 92037, USA. [3]Department of Biochemistry, Microbiology and Immunology and Ottawa Institute of Systems Biology, Faculty of Medicine, University of Ottawa, 451 Smyth Road, Ottawa, Ontario K1H 8M5, Canada.
*Corresponding author. Email: jyates@scripps.edu

embryogenesis and phenotype-maintaining adult tissues. We show that developing hybrids harbor an increased proportion of novel proteins that are not found in either parent at the corresponding developmental stage. This proportion of novel proteins is subsequently reduced in adult hybrids.

RESULTS

Drosophila hybrids were obtained by crossing isogenic inbred *D. melanogaster* and *D. simulans* fly strains whose genomes have been previously sequenced (*21*). Female hybrid embryos of the *D. mel.* ♀ × *D. sim.* ♂ cross-develop into sterile daughter flies, whereas male hybrids die at larval stage 3 (fig. S1A). Neither male nor female embryos of the reciprocal cross *D. sim.* ♀ × *D. mel.* ♂ developed beyond a very early developmental stage (*13*).

Hybrid proteomes

Hybrid proteomes were analyzed for any change in complexity using quantitative bottom-up proteomics (Materials and Methods). To compare proteomes, we used an internal standard that combines parental *D. melanogaster* and *D. simulans* embryos in a 1:1 ratio. The internal standard included the same number of embryos as in the sample and reflects a hypothetical hybrid in which all orthologous proteins are expressed to the exact same level as observed in each parental species at the same developmental stage (fig. S1B). Our analysis revealed that, on average, 65.8% of proteins overlapped between actual developing hybrid embryos and the hypothetical hybrid because 65.8% of identified and quantified peptides were present in both developing hybrid embryos of the *D. mel.* ♀ × *D. sim.* ♂ cross and in the internal standard (Fig. 1A and network S1_*Rc*).

In addition, Fig. 1A shows that 13.3% of the peptides were found solely in hybrid embryos and 7.7% were present solely in the internal standard, corresponding to 1.7-fold more peptides unique to hybrids than to the internal standard (13.2% of peptides not quantified; networks S1a and S1b). On average, we saw a 20% increase in complexity of the hybrid proteome (65.8% + 13.3%) over the subset of the hybrid proteome that is identical to the internal standard (65.8%). When the hybrid proteome (65.8% + 13.3%) and internal standard (65.8% + 7.7%) were directly compared, the complexity of the hybrid proteome was 10% increased. Although errors of measurement were high, the hybrid proteome of developing embryos was more complex than a hypothetical proteome in which both species-specific proteomes were simply combined.

Mapping peptides back to their species of origin revealed that *D. simulans*–specific peptides (4.0%) were 1.9-fold more abundant than *D. melanogaster*–specific peptides (2.1%) within the 13.3% fraction of peptides that were solely found in developing hybrids (fig. S2A). Peptides (7.2%) remained elusive with regard to species specificity because their amino acid sequences are identical in *D. melanogaster* and *D. simulans*. Overall, *D. simulans*–specific peptides were 10% less abundant than *D. melanogaster*–specific peptides within all proteins in the hybrid proteome [(9.1% + 4.0%):(12.3% + 2.1%) or 1:1.1 for *D. simulans*/*D. melanogaster*; fig. S2A). The theoretical ratio of *D. simulans*– to *D. melanogaster*–specific peptides obtained with an in silico digest of both databases with endoproteinase LysC is 1:1.1, ruling out a strong bias toward one of the two species due to an overt difference in the complexity of the two databases (fig. S3). Thus, mainly an aberrant expression of *D. simulans*–specific proteins derived from the paternal genome contributed to the increase in proteome complexity.

The in silico digest of both protein databases with endoprotease LysC also showed that approximately two-thirds of peptide sequences of both species are unambiguously assigned to either *D. melanogaster* or *D. simulans* (fig. S3). This unambiguous assignment of peptides enabled us to identify and quantify all pairs of proteins that are derived from the same gene in two different species. Figure 1A shows which pairs of species-specific orthologs were present in hybrids and parental controls. As displayed in Fig. 1A, 71% of ortholog protein pairs were found in the hybrid proteome and in the internal standard. Within the nonoverlapping proteomes, twofold more ortholog protein pairs were present in hybrids (19%) only than in the internal standard only (10%). The result indicated that the fraction of novel proteins in the developing hybrids includes more pairs of orthologs than that present in the hypothetical hybrid and thus synonymous expression of orthologs.

To determine whether the composition of the hybrid proteome in developing embryos persists into adulthood, we measured the species-specific contributions to the proteome in heads of adult *D. mel.* ♀ × *D. sim.* ♂ hybrid female flies (Fig. 1B). Peptide sequences (78.1%) were present in both the hybrid sample and the corresponding internal standard (network S2_*Rc*). The 4% peptides unique to hybrids was only a 5% contribution overall [(78.1% + 4%)/78.1%] and fourfold less than peptides that were unique to the internal standard (15.9%; networks S2a and S2b). The complexity of the proteome in hybrid heads (78.1% +

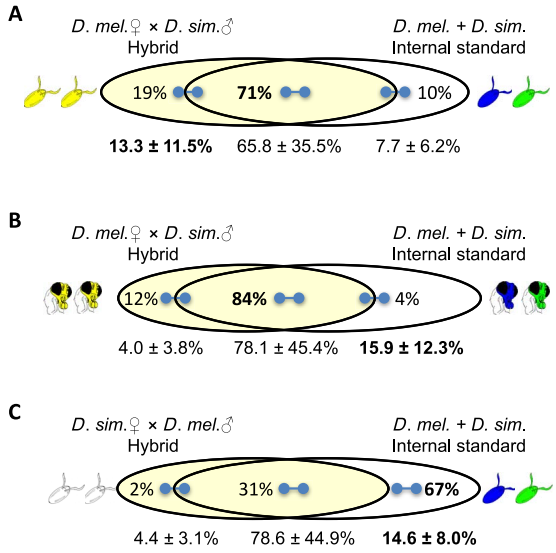

Fig. 1. The proteome in developing hybrid embryos is more complex than the proteome of both species combined. Peptides were grouped according to their presence in either the hybrid or internal standard sample. (**A**) The proteome of developing hybrid embryos (yellow) from a cross of *D. mel.* ♀ with *D. sim.* ♂ was analyzed with bottom-up proteomics and compared to an internal standard that is a 1:1 mixture of the respective *D. melanogaster* (blue) and *D. simulans* (green) embryos. The relative percentage and standard deviation (SD) of peptides (experimental triplicates) are indicated below the nonproportional Venn diagrams, and the larger, sample-specific fraction of the proteome in either sample is highlighted (bold). Handles within each Venn diagram indicate pairs of two species-specific orthologs that were detected in both or hybrids or internal standard. The relative contribution (percentage) to all pairs of orthologs identified in a proteome that comprises all biological replicates is depicted next to the handle. The largest contribution is highlighted in bold. (**B**) In addition, the head-specific proteome of adult hybrid female flies of the same cross (*D. mel.* ♀ × *D. sim.* ♂) was compared to the corresponding internal standard. (**C**) Hybrid embryos of the reverse cross (*D. sim.* ♀ × *D. mel.* ♂) are not viable (white). The proteome of these nondeveloping hybrid embryos was compared to the same internal standard as in (A).

4.0%) was overall 10% lower than the complexity of the internal standard (78.1% + 15.9%), which is in contrast to the 10% increase in complexity of the hybrid proteome in developing embryos of the same cross. Specifically, there was a 3.3-fold reduction in novel peptides present only in adult hybrids (4.0%) over developing embryos (13.3%).

Developing hybrid embryos harbored twofold more *D. simulans*– than *D. melanogaster*–specific peptides in the proteome unique to hybrids only, as described above. In contrast, 1.4-fold more *D. melanogaster*– (1.1%) than *D. simulans*–specific peptides (0.8%) were found exclusively in the proteome of adult hybrids (fig. S2B). This trend was even more pronounced when looking at the complete proteome, where *D. melanogaster*–specific peptides outnumbered *D. simulans*–specific peptides by an additional 100% [(9.5% + 0.8%):(19.6% + 1.1%) or 1:2 for *D. simulans*/*D. melanogaster*]. The predominance of *D. melanogaster*–specific peptides in adult hybrids is opposite to the complete proteome of developing hybrid embryos, which harbored overall 10% less *D. simulans*–specific peptides. It showed that *D. melanogaster*–specific proteins provided a greater relative contribution to a stable hybrid proteome than *D. simulans*–specific proteins in adult flies than in developing hybrids. *D. simulans*–specific proteins were more deregulated in developing hybrid embryos than in adult hybrids.

The proportion of orthologous protein pairs was threefold higher in adult hybrid flies (12%) than in the internal standard (4%) despite a reduced complexity of the hybrid proteome. The remaining 84% of all orthologs were present in both hybrids and internal standard. As in developing hybrid embryos, pairs of orthologous proteins were more abundant in hybrids (12% + 84%) than in the internal standard (4% + 84%).

As an additional reference, we analyzed embryos of the reciprocal cross *D. sim.* ♀ × *D. mel.* ♂ that failed to develop past very early larval stages (Fig. 1C and network S3_*Rc*). Like in adult hybrids, these nondeveloping hybrid embryos harbored 3.3-fold less peptides that were unique to the hybrid (4.4%; network S3a) than to the internal standard (14.6%; network S3b). The complexity in the hybrid proteome (78.6% + 4.4%) increased by only 6% over the proteome that was present in both hybrids and internal standard (78.6%). Overall, the complete proteome of nondeveloping hybrids (78.6% + 4.4%) was 10% less complex than the proteome of the internal standard (78.6% + 15.9%).

The proteome unique to nondeveloping hybrids was almost completely depleted of *D. melanogaster*–specific peptides (0.2%). *D. simulans*–specific peptides were 14-fold more abundant (2.8%; fig. S2C). The complete proteome of nondeveloping hybrid embryos harbored 40% more *D. simulans*– than *D. melanogaster*–specific peptides [(12.6% + 2.8%):(10.9% + 0.2%) or 1.4:1 for *D. simulans*/*D. melanogaster*). When directly compared to developing hybrid embryos, this relative predominance of proteins derived from the *D. simulans* genome in nondeveloping hybrid embryos is rather due to an absence of *D. melanogaster*– (2.1% versus 0.2% in developing and nondeveloping hybrids, respectively; 10.5-fold reduction) than of *D. simulans*–specific peptides (4.0% versus 2.8% in developing and nondeveloping hybrids, respectively; 1.43-fold reduction; fig. S2, A and C). It indicated that hybrids of the reciprocal cross failed to effectively transcribe and translate the paternally derived *D. melanogaster* genome in a zygote that initially consists of a *D. simulans*–specific proteome.

Notably, the majority of orthologous protein pairs (67%) were exclusively located in the internal standard and not in the hybrids and internal standard (Fig. 1C). The proteome shared between the hybrids and internal standard included 31% of orthologous protein pairs, whereas the proteome solely found in the nondeveloping hybrids included only 2% of orthologous protein pairs. The 3.5-fold increase in orthologous protein pairs in the internal standard (67% + 31%) over nondeveloping hybrids (31% + 2%) indicated that the proteome of nondeveloping hybrids was not only less complex but also harbored reduced translation or expression of orthologous genes that were present otherwise in developing hybrids.

In summary, developing *D. mel.* ♀ × *D. sim.* ♂ hybrid embryos showed increased proteome complexity during early development and an overproportional contribution of proteins from the paternal *D. simulans* species among newly present proteins. Proteome complexity was reduced in adult hybrid flies, and maternally derived *D. melanogaster*–specific peptides became overrepresented. A large number of orthologous protein pairs supported the postdevelopmental life of adult hybrids. In contrast, nondeveloping hybrid embryos of the reverse cross *D. sim.* ♀ × *D. mel.* ♂ showed an overall reduced proteome complexity and a pronounced absence of *D. melanogaster*–specific proteins.

Changes in relative protein abundance in hybrids

Our initial data analysis determined the presence or absence of peptides in the hybrid proteome based on replicate experiments (fig. S4). To increase robustness of data analysis, we decided to quantify the relative change in abundance of peptides that were detected in both actual and hypothetical hybrids, that is, sample and internal standard. Having subsumed peptides into peptide nodes in cases where proteins shared the identical peptide sequences, we further included only peptide nodes with at least two isobaric isotopolog ratio measurements [isobaric ratio (*Ri*)] across all three biological replicates (Materials and Methods).

The peptide nodes were differentiated according to species specificity as indistinguishable (DmDs), *D. melanogaster*–specific (Dm), and *D. simulans*–specific (Ds), and the relative frequency of binned ratio values is plotted. Because the internal standard was added 1:1 to the sample, the average ratio of all quantified peptide nodes was expected to be 1:1 (or 0 when $\log_2$-transformed), assuming that the size of *D. melanogaster* and *D. simulans* zygotes is similar and that the hybrid proteome is similar to the proteome of both parental species combined (internal standard). Values of >0 or <0 of the $\log_2$-transformed ratios indicated enrichment or depletion in hybrids, respectively (Fig. 2, left panels). Hybrid proteomes were similar but not identical to the internal standard with respect to the relative abundance of individual peptide nodes. DmDs peptide nodes were enriched by 25% in developing (mean, 0.32) and by 6% in adult hybrids (mean, 0.09), respectively, and depleted by 12% in nondeveloping hybrid embryos (mean, –0.19).

When considering peptide nodes that were specific to a single species, the subset of *D. simulans*–specific peptide nodes increased overproportionally in developing hybrid embryos (Fig. 2A and network S1_*Ri*). The mean of *D. simulans*–specific peptide node ratios (0.90) was at least 1.4-fold higher than the mean of the *D. melanogaster*–specific (0.43) or DmDs (0.32) peptide node ratios. In addition, ratio values of *D. simulans*–specific peptide nodes showed a wider distribution (SD = ±1.05) than *D. melanogaster*–specific (SD = ±0.85) and DmDs peptide nodes (SD = ±0.79), hinting at a loosened control or overall dysregulation of proteins that are encoded in the paternally derived *D. simulans* genome.

The differences between the species-specific proteomes disappeared in adult hybrid tissue (Fig. 2B). The mean of the ratio values of either species-specific proteome were closer to 1:1 (or 0 when $\log_2$-transformed) in adults than in developing hybrid embryos (0.09 for DmDs, 0.19 for Dm, and 0.23 for Ds; network S3_*Ri*). Thus, a selective enrichment of *D. simulans* genome-encoded proteins, as observed in

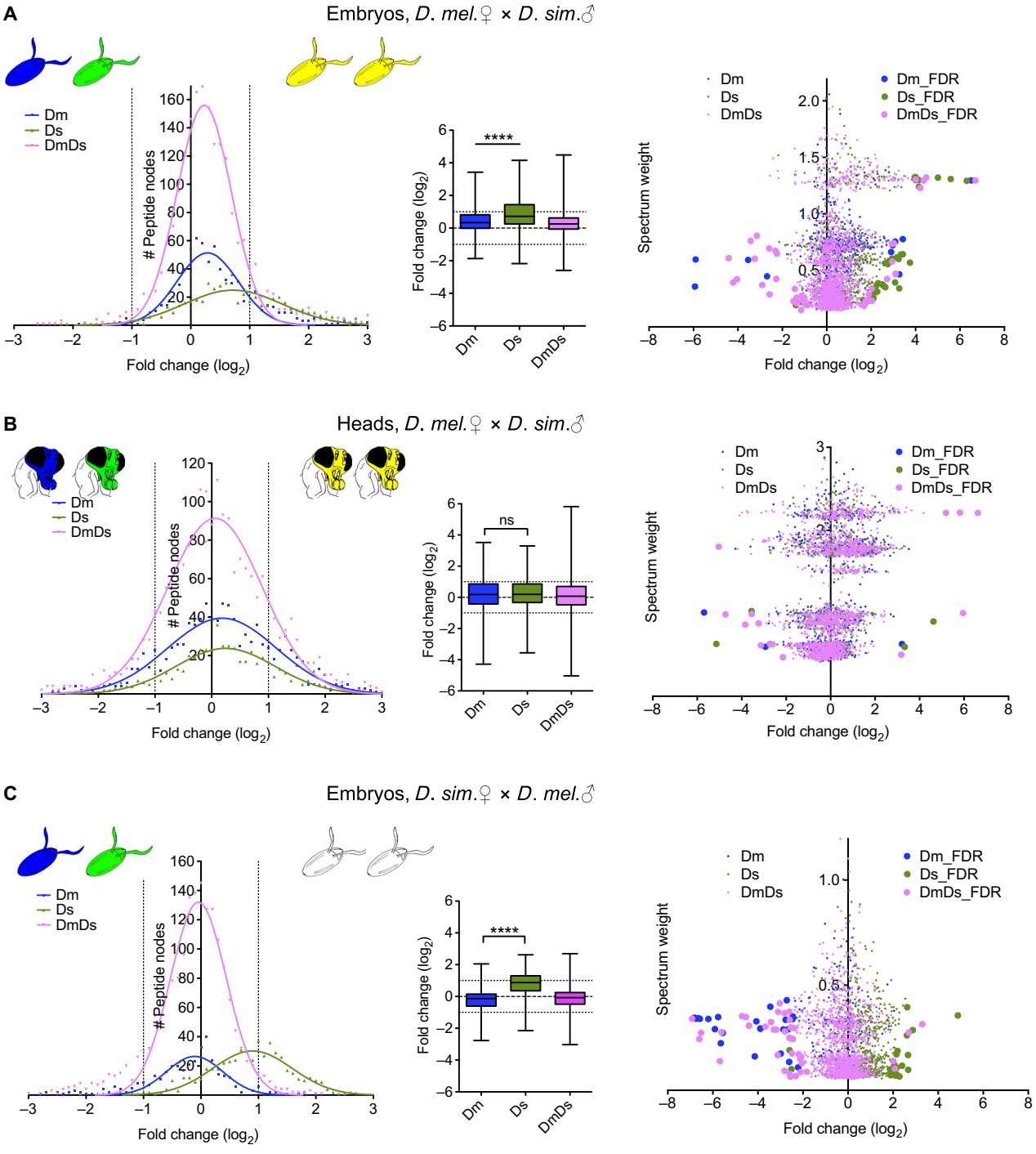

Fig. 2. Protein levels in hybrids differ between the proteomes that are specific to the genome of the father and mother. (**A**) Peptide nodes were classified according to species specificity in *D. melanogaster* (Dm) or *D. simulans* (Ds) or indistinguishable (DmDs) and the distribution of peptide node ratios [isobaric isotopolog ratio (*Ri*)] of developing embryos (*D. mel.* ♀ × *D. sim.* ♂ cross) versus the internal standard plotted. Individual points in the graph (left) indicate the relative frequency of peptide nodes per ratio bin, which is the fold change in peptide abundance in hybrids in comparison to the internal standard. The best fit of a Gaussian (line) for each subset of peptide nodes is included. The distribution of all measurements is shown in a box-and-whisker graph (middle). The statistical significance between *D. melanogaster*– and *D. simulans*–specific peptide nodes was determined with unpaired *t* test with Welch's correction for unequal variances, assuming that groups are independent (****$P < 0.0001$). A scatterplot (right) of ratio versus spectrum weight highlights peptide nodes (large dots) that were significantly regulated according to an FDR ≤ 0.05. The spectrum weight (*y* axis) is inversely proportional to the variance of measurement of each peptide node. All three plots are repeated for the proteome in heads of adult hybrid flies of the *D. mel.* ♀ × *D. sim.* ♂ cross (**B**) and in non-developing embryos of the *D. sim.* ♀ × *D. mel.* ♂ cross (**C**).

developing hybrids, was lost in adulthood. *D. melanogaster*– and *D. simulans*–specific peptide node ratios spread out as wide as species-indistinguishable peptide nodes (SD = ±1.03, ±1.05, and ±1.06, respectively). Although the overall proteome in adult somatic tissue of hybrids was consolidated with regard to its species-specific origin, alternations in relative protein levels from the internal standard were more pronounced and did not show any species specificity. This dysregulation might reflect a potentially more general characteristic of hybrid proteomes.

Nondeveloping hybrid embryos displayed similar discrepancies in protein levels as did developing hybrid embryos (Fig. 2C and network S2_*Ri*). Indistinguishable (mean, −0.19) and *D. melanogaster*–specific peptide nodes (mean, −0.47) were depleted, whereas *D. simulans*–specific peptide nodes (mean, 0.82) were 2.4-fold enriched in nondeveloping hybrids. The relative enrichment of *D. simulans*–specific proteins is most likely a consequence of a profound overall depletion of the proteome encoded by the paternal *D. melanogaster* genome. Nonetheless, peptide nodes of the incoming, paternal genome–derived *D. melanogaster*–specific proteome preserved a wide distribution of ratio values (SD = ±1.24), whereas peptide nodes of the maternal genome–derived proteome and DmDs peptide nodes displayed tight ratio value distributions (SD = ±0.80 for both). A wide spread of ratio values (SD > 1.0) was a characteristic of the incoming paternal genome–encoded proteome in developing hybrid embryos regardless of the direction of the cross because it was already observed in the developing hybrids.

Thus, the proteome that is newly synthesized from the parental genome had wide ratio distributions in hybrids in early developmental stages, whereas the ratio distributions of the initially maternally deposited proteome remained tight. Proteins encoded in the paternally inherited genome were overall either enriched in developing hybrids or depleted in nondeveloping hybrids. Whereas the former reflected successful but aberrant translation of the proteome that is encoded in the paternally inherited genome, the latter pointed to a failure to fully establish this parent-specific proteome. Therefore, nondeveloping hybrids displayed a significant relative enrichment of the proteome that is encoded in the maternally inherited genome. Finally, the proteome of adult hybrid daughters revealed overall higher variability in protein expression levels but without a bias for either species.

Quantitative protein set enrichment analysis for enrichment of *D. melanogaster* Gene Ontology terms in hybrids

Next, we asked whether the increased complexity in the proteome of hybrids is broadly distributed over all biological processes or preferentially alters specific ones. Using protein set enrichment analysis (PSEA)–Quant (*22*, *23*), we found that 10 Gene Ontology (GO) (*24*) terms with the key phrases "proteasome regulatory particle," "translation initiation," and "mitotic spindle elongation" were selectively attributed to proteins enriched in developing hybrid embryos (table S1). This indicated for the first time that proteins involved in protein turnover were selectively increased in developing hybrid embryos.

Proteins with GO terms "regulation of organ growth" and "drug binding" were more enriched in adult hybrid flies than internal standard, whereas "actin-myosin structure organization" and "chitin-based cuticle development" were enriched for proteins more abundant in the internal standard than hybrid flies. Moreover, the GO term "non-recombinational repair" (Ku70/Ku80 protein complex) was enriched in proteins that were more abundant in nondeveloping hybrid embryos than in the internal standard, which suggested that a potentially high load of DNA damage in embryos of the reverse cross might be part of the reason for developmental failure. Whereas GO terms of surviving hybrid embryos revealed altered protein turnover and those associated with nondeveloping hybrids suggested difficulties in maintaining the fidelity of the genome, GO terms broadly indicative of an altered cell and tissue homeostasis were associated with the proteome in adult hybrids.

Dysregulated proteins in hybrids

In an attempt to understand the differences in proteome complexity, we took a closer look at individual proteins that were significantly dysregulated in hybrids. To this end, peptide nodes with a false discovery rate (FDR) of ≤5% were selected, and the corresponding proteins were further analyzed (Fig. 2, right panel, and table S2).

Twenty-nine peptide nodes were significantly depleted (network S5), and 57 peptide nodes were significantly enriched (network S4) in developing hybrid embryos. Enriched peptide nodes mapped more often to proteins of *D. simulans* (31 proteins) than to proteins of *D. melanogaster* (6 proteins). The *D. simulans*–specific proteins covered a variety of different cellular processes, whereas *D. melanogaster*–specific proteins were involved in protein turnover. Eukaryotic translation initiation factor 2 subunit 2 (eIF2β, P41375) and Hsp27 as well as proteasome subunits Prosα1, Prosα6, Prosβ2, and Prosα2R1 play important roles in protein translation, folding, and degradation and were found in higher abundance in the *D. melanogaster*–specific proteome.

Hybrid embryos of the nondeveloping, reciprocal cross displayed more depleted (61; network S7) than enriched (26; network S6) peptide nodes, and depleted peptide nodes were almost exclusively *D. melanogaster*–specific, except cystatin-like protein (B4QZT8) and ribosomal L1 domain–containing–like protein (B4Q6V5), which were *D. simulans*–specific. Conversely, 22 of 26 significantly enriched peptide nodes were *D. simulans*–specific, and none were *D. melanogaster*–specific. Up-regulated proteins included Sod1 (superoxid dismutase 1), which counteracts high levels of reactive oxygen species that are a known source of oxidative DNA damage in *D. simulans* infected with the bacterial endosymbiont *Wolbachia pipientis* (*25*).

Only a few significantly regulated peptide nodes were observed in hybrid heads, a result that is consistent with previous observations in hybrid mammalian somatic cells (*26*). Eight peptide nodes were enriched (network S8) in hybrids including *D. simulans* yolk protein 3 (Yp3) as well as CG31087, which is involved in neurogenesis, and Pebp1, which is part of the antimicrobial humoral response. Only one peptide node was species-specific and mapped to the *D. melanogaster* protein tobi, a glycoside hydrolase involved in carbohydrate metabolic processes. Fourteen additional peptide nodes were significantly depleted in adult hybrid heads (network S9), two of which were *D. melanogaster*–specific. One of the three depleted and *D. simulans*–specific proteins was TurandotA1, which is involved in stress tolerance. None of the significantly regulated proteins in nondeveloping hybrids and adult hybrids mapped to the proteostasis network.

Expression of heterochromatin protein 1

We determined an increased complexity of the proteome in developing hybrids and found that proteins that were enriched in developing hybrids participate in protein translation, folding, and degradation. Previous studies identified two proteins in large-scale genetic screens based on mutations that rescue male hybrid lethality: hybrid male rescue (Hmr) (*27*, *28*) and lethal hybrid rescue (Lhr) (*29*). Hmr and Lhr associate with the centromere and ensure proper mitotic division but are not essential for viability. Previous experiments also revealed that Hmr

levels are higher in *D. melanogaster*, whereas Lhr is more abundant in *D. simulans* (*30*). In our data, Hmr protein was detected in only one of the replicate experiments of nondeveloping hybrid embryos, whereas Lhr was found in only one replicate experiment of hybrid heads of the *D. mel.* ♀ × *D. sim.* ♂ cross, limiting a more detailed analysis of their relative protein levels in hybrids.

Recent work showed that Hmr and Lhr are part of a stable protein complex that consists of 25 proteins including heterochromatin protein 1 (HP1a) that binds to centromeric and pericentromeric regions at chromosomes. The protein complex was identified with tandem affinity purification, which strongly suggested that HP1a is a protein-protein interactor of Lhr and Hmr (*30*). HP1a is essential for viability, and the gene locus *HP1a* is also named *su(var)2-5* because previous genetic screens identified it as a suppressor of variegation mutations (*31*). HP1a orthologs in hybrid embryos originated from the maternal genome (Fig. 3): *D. melanogaster* HP1a (P05205) was present in developing hybrid embryos, whereas *D. simulans* HP1a proteoforms were detected in nondeveloping hybrid embryos. Furthermore, the *D. melanogaster* HP1a ortholog was depleted in nondeveloping hybrid embryos, whereas *D. simulans* HP1a proteoforms were enriched in developing hybrid embryos and then depleted in adult hybrids, which, in turn, maintained increased levels of *D. melanogaster* HP1a protein. Thus, surviving hybrid embryos harbored high levels of HP1a protein potentially to engage Hmr and Lhr in a protein complex.

Expression of transposon-associated proteins

Uncontrolled transposon activity has been shown to be critical for hybrid dysgenesis and creation of a species barrier in flies (*11*). To control transposon activity in flies, piRNAs (Piwi-interacting RNAs) associate with Piwi proteins to guide its endonuclease activity to complementary RNA sequences of transposons. Recently, the Hsp80 cochaperone Hsp83 was identified as essential for loading piRNAs to Piwi, which is an important mechanism of transposable element suppression (*32*). Absence of the Hsp83 gene *shutdown* increases mobility of transposable elements, and therefore, its presence contributes to developmental robustness by stabilizing the genome. In developing hybrid embryos, we found that the *D. melanogaster* Hsp83 was significantly depleted, but its *D. simulans* ortholog remained as high as in the internal standard. The inverse was observed in developmentally failing hybrid embryos of the reverse cross: The paternal genome–encoded *D. melanogaster* Hsp83 protein was absent and thus was not available to assist in Piwi-mediated suppression of newly arising transposon activity. Finally, both Hsp83 orthologs were enriched in the heads of adult hybrid flies. Thus, survival of embryos correlated with induction and maintenance of paternally derived Hsp83 protein. Paternal Hsp83 protein might be needed to suppress transposon activity newly arising from the paternal genome in an initially maternally conditioned embryonic environment.

Because we observed differential expression of Hsp83 in developing and nondeveloping hybrid embryos and because suppression of either *Lhr* or *Hmr* results in elevated levels of RNA transcripts that are derived from transposable elements (*30*, *32*), we searched the data for the presence of any transposon proteins. On the basis of spectrum counts (SpC), we found low levels of *Blood* (1 SpC), *Opus* (1 SpC), and *Roo* (2 SpC) retrotransposons in the internal standard (parents) only; *Het-A* (1 SpC) in adult hybrids only; and the viral capsid protein gag (1 SpC) in hybrids and the internal standard. Thus, we did not obtain direct evidence that transposable proteins were expressed in developmentally failing hybrid embryos.

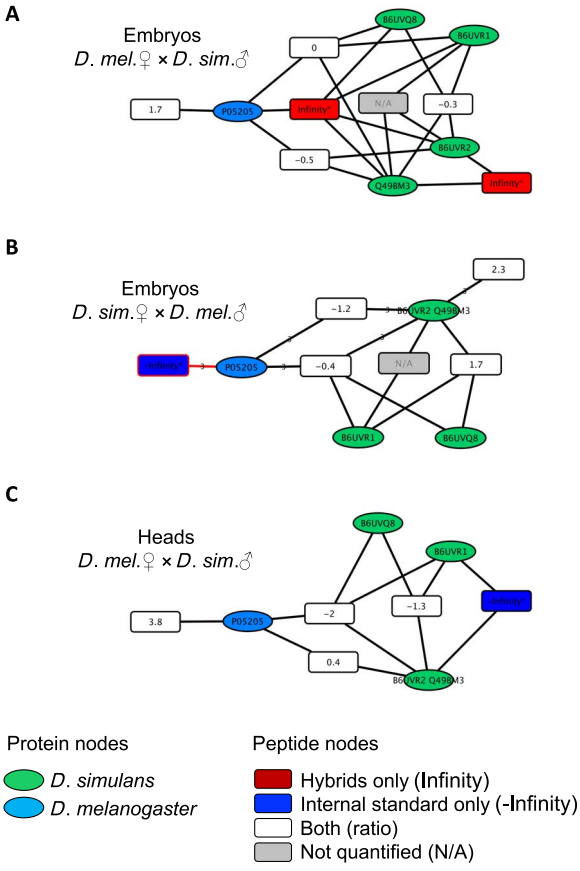

Fig. 3. HP1a in hybrid embryos and adult flies shows distinct, species-specific protein levels. All peptide nodes (rectangles) that are part of specific HP1a proteoforms and orthologous proteins (ovals) are connected by edges that show the relationship between peptide nodes and protein nodes in a bipartite protein-peptide network. Different HP1a proteoforms and orthologs are indicated by their respective UniProt identifiers in developing embryos of the *D. mel.* ♀ × *D. sim.* ♂ cross (**A**), nondeveloping embryos of the *D. sim.* ♀ × *D. mel.* ♂ cross (**B**), and in adult fly heads of the *D. mel.* ♀ × *D. sim.* ♂ cross (**C**). Colors indicate either species specificity (light blue and green) or relative abundance wherein a peptide node is present in both samples (hybrids and internal standard) and measured with a relative ratio (white) or present in either hybrids or internal standard (red and dark blue, respectively) or was identified but not quantified (gray). The $\log_2$-transformed value of the fold change (*Rc* value) per peptide node or its presence in hybrids only (Infinity; red rectangles) or the internal standard (–Infinity; dark blue rectangles) is indicated. Peptides subsumed in peptide nodes with the value "N/A" were identified, but the relative abundance of the peptide node was not quantified because the number of ratio measurements was too low (<3 independent measurements). The asterisk indicates a ±Infinity value (*Rc*-based data analysis). An outline of a peptide node in red as well as an edge in red indicates that the peptide node (>4-fold or highest difference value) deviates from the two additional peptide nodes in a pair of two proteins with unique peptide nodes. The pair of peptide nodes is indicated by a classification number attributed to the edge, and the classification number "3" indicates that unique peptide nodes are different (>4-fold or highest difference value) from the remaining peptide nodes in the protein pair.

Hybrid proteome complexity in view of the proteostasis network

Hybrid daughters are sensitive to elevated temperatures (*34*), and this phenotype can be rescued by depleting Hmr (*35*). Lethality at elevated temperatures argues for dissociation of the previously identified heterospecies Lhr-Hmr-HP1a protein complex (*31*, *33*). Protein-protein interactions in the Lhr-Lmr-HP1a protein complexes of different species

vary widely (*33*), potentially causing a nonassociated Hmr protein to perform a different, neomorphic, and most likely lethal function in hybrid males (*35*, *36*). Because it is conceivable that an increase in complexity of the proteome accelerates the propensity of proteins to engage in new or unstable protein complexes, and because GO terms proteasome regulatory particle and translation initiation were enriched in developing hybrid embryos, we decided to analyze proteins that support protein turnover in more detail.

Translational initiation

Developing hybrid embryos showed a considerable increase in select *D. simulans* and *D. melanogaster* orthologs involved in translation initiation, protein folding, and protein degradation including ubiquitination and proteasomal degradation (Fig. 4). The *D. melanogaster* ortholog eIF2β was overexpressed in developing hybrid embryos. EIF2β was previously identified as a candidate gene implicated in the endoplasmic reticulum (ER) stress response (*37*) during which its dimerization partner, eIF2α, is phosphorylated in response to ER stress by protein kinase–like ER kinase (PERK). This phosphorylation limits global protein translation and facilitates noncanonical translational initiation, which is mediated by eIF2A. In developing embryos, the alternative translation initiation factor eIF2A (Q9VNX8) was enriched, which suggests that alternative translation initiation might cause preferential mRNA translation during early embryogenesis of *Drosophila* hybrids and thus an enrichment of specific proteins (such as chaperones) that are involved in proteostasis.

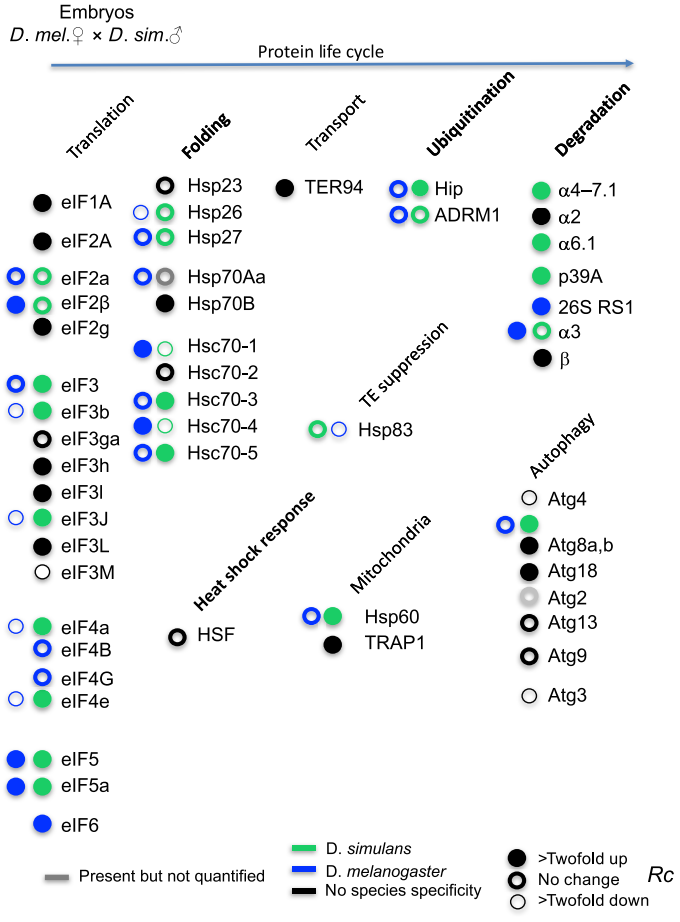

Fig. 4. *D. simulans*–specific orthologs of the proteostasis network are increased in developing hybrid embryos. The graph displays presence and relative expression levels of orthologs and groups these according to their contributions to a protein's life cycle in specific cell biological pathways. Pathways that are part of the proteostasis network are printed in bold. Species specificity is indicated by color (blue, *D. melanogaster*; green, *D. simulans*; black, no species specificity), and relative abundance is indicated by the color fill of the circle (*Rc*-values, light circle, >2-fold down; bold circle, 2-fold down ≤ *R* ≤ 2-fold up; filled circle, >2-fold up).

Expression of proteostasis-associated proteins

Co- and posttranslational protein folding play a critical role in maintaining protein stability and, therefore, a functional proteome. Misfolded proteins trigger an up-regulation of the unfolded protein response (UPR) during which heat shock factor (HSF) increases transcription and translation of heat shock proteins. HSF was more than 1.5-fold enriched in developing hybrid embryos, suggesting a possible activation of the UPR that is part of the proteostasis network. We found that individual members of the UPR were up-regulated in developing hybrids. Hsp70Aa was detected but not quantified, and Hsp70B was increased in developing hybrid embryos. Higher levels of *D. simulans* Hsp26 and Hsp27 indicated a select up-regulation of de novo transcription or translation of stress-associated proteins. The *D. melanogaster* Hsp26 ortholog was down-regulated, although its mRNAs are deposited in oocytes during oogenesis (*38*). Not all heat shock proteins changed in abundance: The cytosolic small heat shock protein Hsp23 remained unaltered.

The heat shock cognate proteins of the Hsc70 family that associate with Hsp70 to assist protein folding (*39*) were specifically enriched in developing hybrid embryos. Hsc70-1 and Hsc70-5 were doubled in relative abundance, Hsc70-3 and Hsc70-4 were enriched, and Hsc70-2 remained unaltered. None of the Hsc70 proteoforms were up-regulated in the developmentally failing reciprocal cross or in adult hybrid flies.

Hsc70 proteins interact with Hsc70-interacting protein 1 (HIP, C4NYP8), a cochaperone that removes proteins that fail to fold correctly. HIP transfers unfolded proteins from the folding machinery into protein degradation and thereby facilitates ubiquitination and clearance of misfolded proteins (*40*). In developing hybrid embryos, *D. melanogaster* HIP was reduced, whereas its *D. simulans* ortholog was up-regulated, and it remained at high levels in adult hybrid heads.

Misfolded membrane proteins at the ER also undergo assisted protein folding or are eventually degraded. The transitional ER adenosine triphosphatase Ter94 [valosin-containing protein (VCP)] plays a critical role in targeting misfolded proteins to the ER-associated degradation pathway ERAD (*41*). VCP was significantly up-regulated in developing hybrid embryos and adult hybrids with a higher relative contribution of *D. simulans*– than *D. melanogaster*–specific orthologs. VCP remained unaltered in nondeveloping hybrid embryos of the reciprocal cross. Thus, independent of developmental stage, surviving embryo and adult hybrids maintained high levels of HIP and VCP, which channel misfolded proteins for proteasomal degradation.

Autophagy removes protein aggregates and damaged organelles in cells (*42*). A survey of autophagy-associated proteins did not reveal alterations in protein levels for the autophagy-associated proteins Atg1, Atg3, Atg9, Atg13, and Atg18a. The cysteine protease Atg4 was depleted in developing hybrid embryos, whereas its substrate Atg8 and *D. simulans* E1-like ubiquitin-conjugating enzyme Atg7 were detected only in developing hybrid embryos.

Mitochondrial proteins are under the control of a separate proteostasis network that is located within the mitochondria. The mitochondrial Hsp90-related protein TRAP1 (A1Z6L9) and the *D. simulans* ortholog of mitochondrial Hsp60 (O02649) were enriched in developing

hybrid embryos. Although TRAP1, the mitochondrial analog of Hsp75, remained unaltered in developing hybrids, it was more than fourfold enriched in adult hybrid flies, suggesting that proteostasis in mitochondria was selectively altered during the development and adulthood of hybrids.

Expression of proteasome and associated proteins

Misfolded proteins are ubiquitinated and subsequently targeted for proteasomal degradation. Protein levels of the *D. simulans* proteasomal ubiquitin receptor ADRM1 (adhesion regulating molecule 1) homolog Rpn13 were almost twofold elevated in hybrid embryos, whereas its *D. melanogaster* ortholog remained unaltered. Again, the *D. simulans*–specific ADRM1 ortholog might be required for a species-selective transfer of proteins to the proteasome.

In addition, several species-specific proteasomal subunits and associated regulatory factors were detected exclusively in developing hybrid embryos (networks S10 to S12). Specifically, *D. melanogaster* 26*S* subunit α3, *D. melanogaster* 26*S* regulatory subunit 1, *D. simulans* proteasomal subunit α7 type 1, *D. simulans* 26*S* regulatory subunit p39A, and proteasomal subunits α2 and β of both species were identified in developing hybrid embryos only and not in parental controls.

Proteasomal subunit α2 protein (Pros25) was the most abundant protein within the subset of proteins that were exclusively present in developing hybrid embryos and not the internal standard, suggesting that high protein levels of proteasomal subunit α2 are important during development of *D. mel.* ♀ × *D. sim.* ♂ hybrid embryos. Proteasomal subunit α2 lacks a proteasome-specific regulatory protein domain and a TATA box in the promoter of the gene and thus represents a member of the housekeeping gene family. In contrast, proteasomal subunits α6.1 and α7.1 are not housekeeping genes (*43*).

Both proteasomal subunits α6.1 and α7.1 had a species-specific protein expression pattern. The *D. simulans* proteasomal subunits α6.1 (B4Q8Q8) and α4-7.1 (Q6QH37) were enriched in developing hybrid embryos, making it likely that they are disproportionally needed to stabilize a hybrid proteome, whereas the *D. melanogaster* proteasomal subunits α6.1 (Pros35, P12881) and α4-7.1 (Pros28.1, P22769) were depleted or unaltered, respectively.

The high levels of proteasomal subunit α2 in embryos were not retained into adulthood, whereas *D. simulans* proteasomal subunit α6.1 remained overexpressed. The differential regulation of species-specific proteasomal subunits and proteasome-associated proteoforms might fine-tune protein degradation in response to an increase in the complexity of the proteome in developing hybrid embryos. It indicated that turnover of proteins is explicitly regulated during development of hybrids and, to a lesser extent, in adult hybrids because only a specific subset of proteasomal subunits remained enriched in adult hybrids.

In conclusion, differences between the proteome of hybrids and the pooled parental proteome (internal standard or hypothetical hybrid) were most pronounced during embryonic development. Developing hybrid embryos tolerated an increase in complexity of the proteome based on proteins exclusively found in hybrids, most of which originated from the paternally inherited genome. Proteins derived from the paternally inherited genome showed overall higher fluctuations in protein abundance. Although the complexity of the proteome was ultimately reduced in adult hybrids and differences in species-specific protein abundances diminished, an overall dysregulation of relative protein abundances remained (Fig. 5).

In an attempt to understand the molecular underpinnings that keep the hybrid proteome functional, we found a selective increase in proteostasis-associated proteins, which predominantly originated from the paternal genome. In developing hybrid embryos, as well as adult hybrid flies, proteostasis-associated *D. simulans* orthologs were preferentially overexpressed. It included heat shock proteins and proteins that are indirectly or directly involved in proteasomal degradation. The results suggest that a successful turnover or clearance of potentially nonfunctional proteins or protein complexes might be required to stabilize a hybrid proteome.

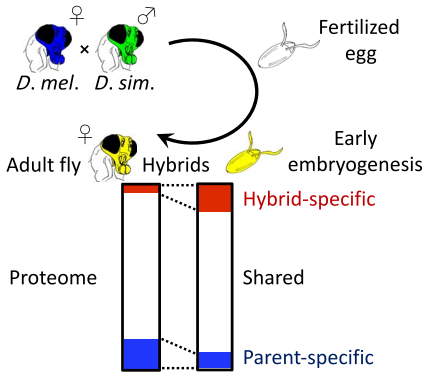

Fig. 5. The hybrid proteome is increased in complexity during early embryonic development and decreased in adult hybrid flies. Hybrid embryos undergo early development with an increased proteome complexity following a cross of *D. melanogaster* with *D. simulans*. Proteome complexity returns to a minimal proteome in adult hybrid flies that is closer to the proteomes of each parental species.

DISCUSSION

Our results show that the maternally deposited *D. melanogaster*–specific proteome correctly guides its protein levels in *D. mel.* ♀ × *D. sim.* ♂ hybrid embryos but initially falls short of accurately controlling *D. simulans*–specific protein levels: Proteins derived from the *D. simulans* genome contributed more novel proteins and showed higher variation in relative abundance. Conversely, *D. sim.* ♀ × *D. mel.* ♂ hybrid embryos showed a marked absence of *D. melanogaster*–specific proteins and failed to develop, maybe because of increased noncombinatorial DNA repair. Here, we found that *D. melanogaster* Hsp83 was absent in nondeveloping hybrid embryos, which might preclude an efficient suppression of *D. melanogaster*–derived transposable elements. In contrast, *D. simulans* Hsp83 was present in developing hybrid embryos of the *D. mel.* ♀ × *D. sim.* ♂ cross.

Quantitative GO term enrichment analysis indicated increased cellular stress in developing hybrid embryos. Several abundant proteins were enriched in developing hybrids that might boost folding and degradation as a molecular mechanism to accommodate the expansion in complexity of the hybrid proteome. High levels of proteasomal proteins, even in adult hybrids, point to a persistent induction of protein degradation. A differential expression of proteasomal subunits has been previously identified in Hmr^- male *D. mel.* ♀ × *D. sim.* ♂ hybrids, which develop past the third larval instar stage (*44*). Nine different proteasomal subunits were up-regulated in the Hmr^+ males of the *D. mel.* ♀ × *D. sim.* ♂ cross, but increased expression was lost in surviving Hmr^- males. Further genetic studies showed that heterozygotes for any one of 3 of 14 different proteasomal subunits rescued male lethality (Pros25, Prosβ2R2, and Prosβ7). Thus, an increased abundance of proteasomal subunit α2 (Pros25) might lead to an increased degradation or turnover of male-specific proteins such as Hmr, which are otherwise critical for transposon suppression in hybrids.

Classifying the hybrid proteome by species specificity revealed that the incoming paternally encoded proteome is more dysregulated than the maternally encoded proteome, which is at least in part already deposited during oogenesis. Surprisingly, an increase in overall proteome complexity by up to 10% is not deleterious as long as both genomes remain intact, arguing for an unexpected high plasticity of the proteome during early development. Morphogenetic features of the developing embryo are set during early development (*45*), and any variation of these requires additional plasticity of the proteome. Given the plasticity of the hybrid proteome observed here, our results indicate that there is an unexpected space for additional proteome complexity during early embryogenesis. It may allow the proteome to expand during early embryogenesis and thus may provide a hitherto unknown space for molecular explorations that might enable rapid phenotypic variation.

MATERIALS AND METHODS

Fly husbandry and collecting embryos

Flies were raised and maintained according to standard husbandry (*46*). The following isogenic fly strains were obtained from the University of California San Diego (UCSD) *Drosophila* Stock Center and were used throughout all experiments: *D. melanogaster* (y[1]; Gr22b[1] Gr22d[1] cn[1] CG33964[R4.2] bw[1] sp[1]; LysC[1] MstProx[1] GstD5[1] Rh6[1]) and *D. simulans* (w[501]). To collect embryos from crosses, we placed parental flies on apple juice or grapefruit juice agar plates. Embryos were collected individually from the agar plates with a thin fused silica column and transferred into an Eppendorf tube. Alternatively, agar plates were rinsed with water to dislodge and collect many embryos into a 50-ml Falcon tube. We collected embryos between 0 and 6 hours after egg laying to capture the hybrid proteome present during early embryonic development. The embryos were assigned the approximate developmental stage by microscopic inspection and comparison to a developmental staging table (*47*).

Proteomics overview

Proteomes were enzymatically digested to peptides with endoproteinase LysC, and peptides in the hybrid and internal standard were dimethylated with light or heavy variants of isobaric isotopologs. Equal aliquots were combined and subsequently analyzed with multidimensional protein identification technology (MudPIT) (*48*, *49*) on an Orbitrap Velos mass spectrometer (Thermo Fisher Scientific; fig. S5A) in biological triplicate. Relative ion abundance was extracted with Census (*50*) using tandem mass spectra of peptides identified with ProLuCID (*51*) in a search against a combined database of *D. melanogaster* and *D. simulans* proteomes (fig. S5B). All peptide spectrum matches and associated quantitative values as well as identified proteins were stored in Proteome INTegrator (PINT). ProteinClusterQuant (PCQ) (https://github.com/proteomicsyates/ProteinClusterQuant) assembled and analyzed the resultant proteomes wherein small, species-specific sequence differences in peptides determined species-specific contributions to the hybrid proteome. For accurate quantification, PCQ condensed subset peptides and subset proteins in single nodes, respectively. The final bipartite protein-peptide network can be visualized in Cytoscape (*52*).

Sample preparation

Drosophila embryos were collected into Eppendorf tubes, washed with water, and then dechorionated for 20 s in 50% household bleach (Vons). The embryos were then washed three more times with water, lysed in 2× TNI buffer [250 mM NaCl, 50 mM tris (pH 7.5), 0.5% Igepal CA-630, 1 mM EDTA, and 1× complete ULTRA EDTA-free protease inhibitor cocktail (Roche)], and crushed with a pestle in an Eppendorf tube to release proteins. The proteins were then precipitated using methanol/chloroform precipitation (sample/methanol/chloroform, 1:4:1, v/v/v), and the protein pellet was resuspended in digestion buffer [0.1 M Hepes buffer (pH 8.0)] supplemented with 1% RapiGest (Waters). Disulfide bonds were reduced with Bond-Breaker TCEP solution (2.5 mM at 37°C for 20 min; Thermo Fisher Scientific) and free sulfhydryl groups acetylated with iodoacetamide (500 mM at 37°C for 30 min; Thermo Fisher Scientific). The proteins were digested into peptides with endoproteinase LysC (1.5 μg of LysC per embryo; Promega).

Peptides were either light (hybrid) or heavy (synthetic mix), dimethylated at primary amines with isotope-defined formaldehyde, as described by Bamberger *et al.* (*53*). Briefly, the samples were labeled as either heavy with deuterium or light with ^{13}C. Primary amines were methylated with formaldehyde (40 mM), and the resulting Schiff's base was reduced with sodium cyanoborohydride (25 mM). CH_2O and $NaBD_3CN$ were used for the synthetic mixture, and $^{13}CH_2O$ and $NaBH_3CN$ were used for hybrids. After labeling the samples (for 1 hour at room temperature), the dimethyl labeling reaction was inactivated by adding an excess of free amines (1% ammonium bicarbonate for 10 min at room temperature). Twenty *D. melanogaster* embryos were combined with 20 *D. simulans* embryos to create an internal standard reflecting a hypothetical hybrid that was labeled heavy, as described. Hybrids were labeled light, and the synthetic mixture was labeled heavy to differentiate protein expression between the F_1 generation of the cross and the inbred strains. For comparison of embryos, an identical aliquot of the heavy-labeled synthetic mixture was added to each light-labeled hybrid sample for direct comparison. The labeled synthetic mixture served as an external standard.

Mass spectrometry

Heavy- and light-labeled samples were mixed in a 1:1 ratio, loaded onto a MudPIT column (*49*), and analyzed on an Orbitrap Velos (Thermo Fisher Scientific). For MudPIT, the following preparative and analytical columns were prepared: 2 cm of strong cation exchange resin (SCX; Partisphere, Waters) followed by 2 cm of reversed-phase resin (Aqua 5, Phenomenex) were packed into a fused silica column [250 μm intradermally (id)], which was attached to an analytical column (100 μm id) with a pulled microtip (~1.0-μm opening) filled with 15 cm of Aqua 3 (Phenomenex). Both columns were equilibrated in buffer A (5% acetonitrile and 0.1% formic acid).

A 10-step MudPIT experiment was performed on an Orbitrap Velos mass spectrometer (*48*). Briefly, peptides were eluted from the column with a linear gradient of 0 to 55% buffer B (80% acetonitrile and 0.1% formic acid in water, v/v/v) over a 98-min time window and injected directly into the mass spectrometer by electrospray ionization at 2.5 kV. Full MS1 scans were collected at a resolution of *R* 60,000 within 400 to 2000 mass/charge ratio (*m*/*z*) and an automatic gain control set to 1×10^6 counts for survey mass spectra (MS1) and 5×10^4 counts for tandem mass spectrometry (MS/MS) fragment ion mass spectra. Each MS1 was followed by 10 MS/MS experiments for the top 10 most abundant precursor ions with a charge state of $z \geq +2$ and an inclusion window of $-0.51 \leq m/z \leq +1.5$. Selected precursor ion masses were dynamically excluded with an exclusion window of $-0.51 \leq m/z \leq +1.5$, a repeat count of 1, and a repeat duration of 30 s with an exclusion list size of 500 and a maximum duration of 120 s. Each precursor ion was fragmented twice, once with standard collisional dissociation in the

LTQ (linear trap quadrupole) and once with higher-energy collisional dissociation in the Orbitrap (*R* 30,000).

Proteome analysis

Mass spectra were searched with ProLuCID against a combined *D. melanogaster* and *D. simulans* UniProtKB protein database (04_2014), which included the reverse sequences of both species-specific databases to allow determination of an FDR at the peptide level with DTASelect2 (*54*). Search results were uploaded to IP2 (Integrated Proteomics Solutions), and high-resolution MS/MS mass spectra were used to quantify relative protein expression based on isobaric isotopolog fragment ions present in high-resolution MS/MS scans. Final results were uploaded in PINT and subsequently queried for peptides identified and quantified to determine the relative number of peptide spectrum matches for each experimental condition.

PSEA-Quant analysis and PCQ

Using PSEA-Quant (*23*), the most abundant peptide nodes (>7 isobaric peak counts) were extracted and further analyzed for the enrichment of GO terms (*24*) to identify terms that are enriched for proteins that are overexpressed in either the hybrids or the internal standard. The number of random samplings to estimate the statistical significance of the enrichment of each GO term among the overexpressed proteins (FDR) was set to 10^6, and the annotation type was set to GO. An interdependency of protein abundance was assumed during sampling, the coefficient of variation tolerance factor was set to 0.5, and the bias in protein annotation was set to "true." GO terms for abundant proteins with an FDR (*q* value) less than 5% were further considered during analysis (table S2). GO terms that associated with fewer than 50 proteins were further considered to increase specificity of GO terms reported.

PCQ was used to assemble complete proteomes and quantify relative differences between hybrids and synthetic mixture. Additional statistical testing was performed in Prism 5 (GraphPad Software Inc.).

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/4/2/eaao3424/DC1

table S1. Significantly regulated peptide nodes (FDR < 0.05) and associated proteins in *D. mel.* ♀ × *D. sim.* ♂ embryos, *D. sim.* ♀ × *D. mel.* ♂ embryos, and *D. mel.* ♀ × *D. sim.* ♂ heads.

table S2. PSEA-Quant results obtained for *D. mel.* ♀ × *D. sim.* ♂ embryos, *D. sim.* ♀ × *D. mel.* ♂ embryos, and heads of the *D. mel.* ♀ × *D. sim.* ♂ cross are presented in individual tabs as indicated in the Excel file.

network S1. *D. mel.* ♀ × *D. sim.* ♂ embryos; >2 isobaric peptide counts per peptide node.

network S2. *D. sim.* ♀ × *D. mel.* ♂ embryos; >2 isobaric peptide counts per peptide node.

network S3. *D. mel.* ♀ × *D. sim.* ♂ heads; >2 isobaric peptide counts per peptide node.

network S4. *D. mel.* ♀ × *D. sim.* ♂ embryos; FDR < 0.01 and *Ri* > 0.0.

network S5. *D. mel.* ♀ × *D. sim.* ♂ embryos; FDR < 0.01 and *Ri* < 0.0.

network S6. *D. sim.* ♀ × *D. mel.* ♂ embryos; FDR < 0.01 and *Ri* > 0.0.

network S7. *D. sim.* ♀ × *D. mel.* ♂ embryos; FDR < 0.01 and *Ri* < 0.0.

network S8. *D. mel.* ♀ × *D. sim.* ♂ heads; FDR < 0.01 and *Ri* > 0.0.

network S9. *D. mel.* ♀ × *D. sim.* ♂ heads; FDR < 0.01 and *Ri* < 0.0.

network S10. *D. mel.* ♀ × *D. sim.* ♂ embryos; proteasomal proteins.

network S11. *D. sim.* ♀ × *D. mel.* ♂ embryos; proteasomal proteins.

network S12. *D. mel.* ♀ × *D. sim.* ♂ hybrid heads; proteasomal proteins.

Data repositories

fig. S1. Only hybrids of *D.* melanogaster females crossed with *D. simulans* males are viable.

fig. S2. Relative abundance of peptide nodes according to species specificity and sample origin.

fig. S3. The Venn diagram displays the relative overlap of peptides retrieved for the *D. melanogaster* and *D. simulans* proteome databases when digested in silico with the endoprotease LysC.

fig. S4. Venn diagram of proteomes identified in each of the replicate experiments.

fig. S5. Sample analysis workflow and detection of species-specific peptides.

REFERENCES AND NOTES

1. E. A. Ponomarenko, E. V. Poverennaya, E. V. Ilgisonis, M. A. Pyatnitskiy, A. T. Kopylov, V. G. Zgoda, A. V. Lisitsa, A. I. Archakov, The size of the human proteome: The width and depth. *Int. J. Anal. Chem.* **2016**, 7436849 (2016).
2. L. M. Smith, N. L. Kelleher; Consortium for Top Down Proteomics, Proteoform: A single term describing protein complexity. *Nat. Methods* **10**, 186–187 (2013).
3. X. Yang, J. Coulombe-Huntington, S. Kang, G. M. Sheynkman, T. Hao, A. Richardson, S. Sun, F. Yang, Y. A. Shen, R. R. Murray, K. Spirohn, B. E. Begg, M. Duran-Frigola, A. MacWilliams, S. J. Pevzner, Q. Zhong, S. A. Trigg, S. Tam, L. Ghamsari, N. Sahni, S. Yi, M. D. Rodriguez, D. Balcha, G. Tan, M. Costanzo, B. Andrews, C. Boone, X. J. Zhou, K. Salehi-Ashtiani, B. Charloteaux, A. A. Chen, M. A. Calderwood, P. Aloy, F. P. Roth, D. E. Hill, L. M. Iakoucheva, Y. Xia, M. Vidal, Widespread expansion of protein interaction capabilities by alternative splicing. *Cell* **164**, 805–817 (2016).
4. S. S. Thakur, T. Geiger, B. Chatterjee, P. Bandilla, F. Fröhlich, J. Cox, M. Mann, Deep and highly sensitive proteome coverage by LC-MS/MS without prefractionation. *Mol. Cell. Proteomics* **10**, M110.003699 (2011).
5. C. Queitsch, T. A. Sangster, S. Lindquist, Hsp90 as a capacitor of phenotypic variation. *Nature* **417**, 618–624 (2002).
6. E. T. Powers, W. E. Balch, Diversity in the origins of proteostasis networks—A driver for protein function in evolution. *Nat. Rev. Mol. Cell Biol.* **14**, 237–248 (2013).
7. S. Pechmann, J. Frydman, Interplay between chaperones and protein disorder promotes the evolution of protein networks. *PLOS Comput. Biol.* **10**, e1003674 (2014).
8. K. S. Pedersen, T. N. Kristensen, V. Loeschcke, Effects of inbreeding and rate of inbreeding in *Drosophila melanogaster*– Hsp70 expression and fitness. *J. Evol. Biol.* **18**, 756–762 (2005).
9. J. Dahlgaard, R. A. Krebs, V. Loeschcke, Heat-shock tolerance and inbreeding in *Drosophila buzzatii*. *Heredity (Edinbg)* **74**, 157–163 (1995).
10. L. Whitesell, S. Santagata, M. L. Mendillo, N. U. Lin, D. A. Proia, S. Lindquist, HSP90 empowers evolution of resistance to hormonal therapy in human breast cancer models. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 18297–18302 (2014).
11. J. Brennecke, C. D. Malone, A. A. Aravin, R. Sachidanandam, A. Stark, G. J. Hannon, An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science* **322**, 1387–1392 (2008).
12. J. S. Khurana, J. Wang, J. Xu, B. S. Koppetsch, T. C. Thomson, A. Nowosielska, C. Li, P. D. Zamore, Z. Weng, W. E. Theurkauf, Adaptation to *P* element transposon invasion in *Drosophila melanogaster*. *Cell* **147**, 1551–1563 (2011).
13. A. H. Sturtevant, Genetic studies on *Drosophila simulans*. I. Introduction. Hybrids with *Drosophila melanogaster*. *Genetics* **5**, 488–500 (1920).
14. D. K. Pritchard, G. Schubiger, Activation of transcription in *Drosophila* embryos is a gradual process mediated by the nucleocytoplasmic ratio. *Genes Dev.* **10**, 1131–1142 (1996).
15. R. M. Graze, L. M. McIntyre, B. J. Main, M. L. Wayne, S. V. Nuzhdin, Regulatory divergence in *Drosophila melanogaster* and *D. simulans*, a genomewide analysis of allele-specific expression. *Genetics* **183**, 547–561 (2009).
16. C. J. McManus, J. D. Coolon, M. O. Duff, J. Eipper-Mains, B. R. Graveley, P. J. Wittkopp, Regulatory divergence in *Drosophila* revealed by mRNA-seq. *Genome Res.* **20**, 816–825 (2010).
17. P. J. Wittkopp, B. K. Haerum, A. G. Clark, Regulatory changes underlying expression differences within and between *Drosophila* species. *Nat. Genet.* **40**, 346–350 (2008).
18. E. M. de Robertis, P. Black, Hybrids of *Xenopus laevis* and *Xenopus borealis* express proteins from both parents. *Dev. Biol.* **68**, 334–339 (1979).
19. H. R. Woodland, J. E. M. Ballantine, Paternal gene expression in developing hybrid embryos of *Xenopus laevis* and *Xenopus borealis*. *J. Embryol. Exp. Morphol.* **60**, 359–372 (1980).
20. H. R. Woodland, J. M. Flynn, A. J. Wyllie, Utilization of stored mRNA in Xenopus embryos and its replacement by newly synthesized transcripts: Histone H1 synthesis using interspecies hybrids. *Cell* **18**, 165–171 (1979).
21. A. Stark, M. F. Lin, P. Kheradpour, J. S. Pedersen, L. Parts, J. W. Carlson, M. A. Crosby, M. D. Rasmussen, S. Roy, A. N. Deoras, J. G. Ruby, J. Brennecke; Harvard FlyBase curators; Berkeley Drosophila Genome Project, E. Hodges, A. S. Hinrichs, A. Caspi, B. Paten, S.-W. Park, M. V. Han, M. L. Maeder, B. J. Polansky, B. E. Robson, S. Aerts, J. van Helden, B. Hassan, D. G. Gilbert, D. A. Eastman, M. Rice, M. Weir, M. W. Hahn, Y. Park, C. N. Dewey, L. Pachter, W. J. Kent, D. Haussler, E. C. Lai, D. P. Bartel, G. J. Hannon, T. C. Kaufman, M. B. Eisen, A. G. Clark, D. Smith, S. E. Celniker, W. M. Gelbart, M. Kellis, Discovery of functional elements in 12 *Drosophila* genomes using evolutionary signatures. *Nature* **450**, 219–232 (2007).
22. M. Lavallée-Adam, J. R. Yates III, Using PSEA-quant for protein set enrichment analysis of quantitative mass spectrometry-based proteomics. *Curr. Protoc. Bioinformatics* **53**, 13.28.1–13.28.16 (2016).
23. M. Lavallée-Adam, N. Rauniyar, D. B. McClatchy, J. R. Yates III, PSEA-Quant: A protein set enrichment analysis on label-free and label-based protein quantification data. *J. Proteome Res.* **13**, 5496–5509 (2014).
24. The Gene Ontology Consortium, M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill,

L. Issel-Tarver, A. Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin, G. Sherlock, Gene ontology: Tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).
25. L. J. Brennan, J. A. Haukedal, J. C. Earle, B. Keddie, H. L. Harris, Disruption of redox homeostasis leads to oxidative DNA damage in spermatocytes of *Wolbachia*-infected *Drosophila simulans*. *Insect Mol. Biol.* **21**, 510–520 (2012).
26. J. Hou, X. Wang, E. McShane, H. Zauber, W. Sun, M. Selbach, W. Chen, Extensive allele-specific translational regulation in hybrid mice. *Mol. Syst. Biol.* **11**, 825 (2015).
27. A. W. Davis, J. Roote, T. Morley, K. Sawamura, S. Herrmann, M. Ashburner, Rescue of hybrid sterility in crosses between *D. melanogaster* and *D. simulans*. *Nature* **380**, 157–159 (1996).
28. P. Hutter, M. Ashburner, Genetic rescue of inviable hybrids between *Drosophila melanogaster* and its sibling species. *Nature* **327**, 331–333 (1987).
29. N. J. Brideau, H. A. Flores, J. Wang, S. Maheshwari, X. Wang, D. A. Barbash, Two Dobzhansky-Muller genes interact to cause hybrid lethality in *Drosophila*. *Science* **314**, 1292–1295 (2006).
30. A. W. Thomae, G. O. M. Schade, J. Padeken, M. Borath, I. Vetter, E. Kremmer, P. Heun, A. Imhof, A pair of centromeric proteins mediates reproductive isolation in *Drosophila* species. *Dev. Cell* **27**, 412–424 (2013).
31. J. C. Eissenberg, G. D. Morris, G. Reuter, T. Hartnett, The heterochromatin-associated protein HP-1 is an essential protein in Drosophila with dosage-dependent effects on position-effect variegation. *Genetics* **131**, 345–352 (1992).
32. D. Olivieri, K.-A. Senti, S. Subramanian, R. Sachidanandam, J. Brennecke, The cochaperone shutdown defines a group of biogenesis factors essential for all piRNA populations in *Drosophila*. *Mol. Cell* **47**, 954–969 (2012).
33. P. R. V. Satyaki, T. N. Cuykendall, K. H.-C. Wei, N. J. Brideau, H. Kwak, S. Aruna, P. M. Ferree, S. Ji, D. A. Barbash, The *Hmr* and *Lhr* hybrid incompatibility genes suppress a broad range of heterochromatic repeats. *PLOS Genet.* **10**, e1004240 (2014).
34. C. J. J. Miller, D. R. Matute, The effect of temperature on *Drosophila* hybrid fitness. *G3 (Bethesda)* **7**, 377–385 (2017).
35. D. A. Barbash, J. Roote, M. Ashburner, The *Drosophila melanogaster Hybrid male rescue* gene causes inviability in male and female species hybrids. *Genetics* **154**, 1747–1771 (2000).
36. H. A. Orr, S. Irving, Genetic analysis of the *Hybrid male rescue* locus of Drosophila. *Genetics* **155**, 225–231 (2000).
37. C. Y. Chow, M. F. Wolfner, A. G. Clark, Using natural variation in *Drosophila* to discover previously unknown endoplasmic reticulum stress genes. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 9013–9018 (2013).
38. J. L. Zimmerman, W. Petri, M. Meselson, Accumulation of a specific subset of D. melanogaster heat shock mRNAs in normal development without heat shock. *Cell* **32**, 1161–1170 (1983).
39. K. B. Palter, M. Watanabe, L. Stinson, A. P. Mahowald, E. A. Craig, Expression and localization of Drosophila melanogaster hsp70 cognate proteins. *Mol. Cell. Biol.* **6**, 1187–1203 (1986).
40. A. M. Wang, Y. Miyata, S. Klinedinst, H.-M. Peng, J. P. Chua, T. Komiyama, X. Li, Y. Morishima, D. E. Merry, W. B. Pratt, Y. Osawa, C. A. Collins, J. E. Gestwicki, A. P. Lieberman, Activation of Hsp70 reduces neurotoxicity by promoting polyglutamine protein degradation. *Nat. Chem. Biol.* **9**, 112–118 (2013).
41. A. E. Johnson, H. Shu, A. G. Hauswirth, A. Tong, G. W. Davis, VCP-dependent muscle degeneration is linked to defects in a dynamic tubular lysosomal network in vivo. *eLife* **4**, e07366 (2015).
42. C. K. McPhee, E. H. Baehrecke, Autophagy in *Drosophila melanogaster*. *Biochim. Biophys. Acta* **1793**, 1452–1460 (2009).
43. A. Seelig, M. Troxell, P.-M. Kloetzel, Sequence and genomic organization of the *Drosophila* proteasome PROS-Dm25 gene. *Biochim. Biophys. Acta* **1174**, 215–217 (1993).
44. D. A. Barbash, J. G. Lorigan, Lethality in *Drosophila melanogaster*/*Drosophila simulans* species hybrids is not associated with substantial transcriptional misregulation. *J. Exp. Zool. B Mol. Dev. Evol.* **308**, 74–84 (2007).
45. W. Driever, C. Nüsslein-Volhard, A gradient of *bicoid* protein in Drosophila embryos. *Cell* **54**, 83–93 (1988).
46. W. Sullivan, M. Ashburner, R. S. Hawley, *Drosophila Protocols* (Cold Spring Harbor Laboratory Press, 2000), p. 697.
47. M. Ashburner, K. G. Golic, R. S. Hawley, *Drosophila: A Laboratory Handbook* (Cold Spring Harbor Laboratory Press, ed. 2, 2005).
48. D. A. Wolters, M. P. Washburn, J. R. Yates III, An automated multidimensional protein identification technology for shotgun proteomics. *Anal. Chem.* **73**, 5683–5690 (2001).
49. M. P. Washburn, D. Wolters, J. R. Yates III, Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **19**, 242–247 (2001).
50. S. K. Park, J. D. Venable, T. Xu, J. R. Yates III, A quantitative analysis software tool for mass spectrometry–based proteomics. *Nat. Methods* **5**, 319–322 (2008).
51. T. Xu, S. K. Park, J. D. Venable, J. A. Wohlschlegel, J. K. Diedrich, D. Cociorva, B. Lu, L. Liao, J. Hewel, X. Han, C. C. L. Wong, B. Fonslow, C. Delahunty, Y. Gao, H. Shah, J. R. Yates III, ProLuCID: An improved SEQUEST-like algorithm with enhanced sensitivity and specificity. *J. Proteomics* **129**, 16–24 (2015).
52. C. T. Lopes, M. Franz, F. Kazi, S. L. Donaldson, Q. Morris, G. D. Bader, Cytoscape Web: An interactive web-based network browser. *Bioinformatics* **26**, 2347–2348 (2010).
53. C. Bamberger, S. Pankow, S. K. R. Park, J. R. Yates III, Interference-free proteome quantification with MS/MS-based isobaric isotopologue detection. *J. Proteome Res.* **13**, 1494–1501 (2014).
54. D. L. Tabb, W. H. McDonald, J. R. Yates III, DTASelect and Contrast: Tools for assembling and comparing protein identifications from shotgun proteomics. *J. Proteome Res.* **1**, 21–26 (2002).

Acknowledgments: We would like to thank S. Pankow and C. Delahunty for reading and commenting on the manuscript, and we are grateful to the UC San Diego *Drosophila* Stock Center for providing the *Drosophila* species used in the study. **Funding:** This work was supported by grants 5R01HL079442-08, P01AG031097, P41GM103533, and HHSN268201000035C to J.R.Y. and a Natural Sciences and Engineering Research Council of Canada Discovery Grant to M.L.-A. M.L.-A. also held a postdoctoral fellowship from Fonds de Recherche du Québec-Nature et Technologies. **Author contributions:** C.B. conceived the experiment. C.B. and M.M. executed the experiments. C.B., M.M., S.M.-B., and M.L.-A. analyzed the data. C.B. prepared the figures and wrote the manuscript. C.B., M.M., S.M.-B., M.L.-A., and J.R.Y. read and revised the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 10 July 2017
Accepted 11 January 2018
Published 7 February 2018
10.1126/sciadv.aao3424

Citation: C. Bamberger, S. Martínez-Bartolomé, M. Montgomery, M. Lavallée-Adam, J. R. Yates, Increased proteomic complexity in *Drosophila* hybrids during development. *Sci. Adv.* **4**, eaao3424 (2018).