Supporting Information

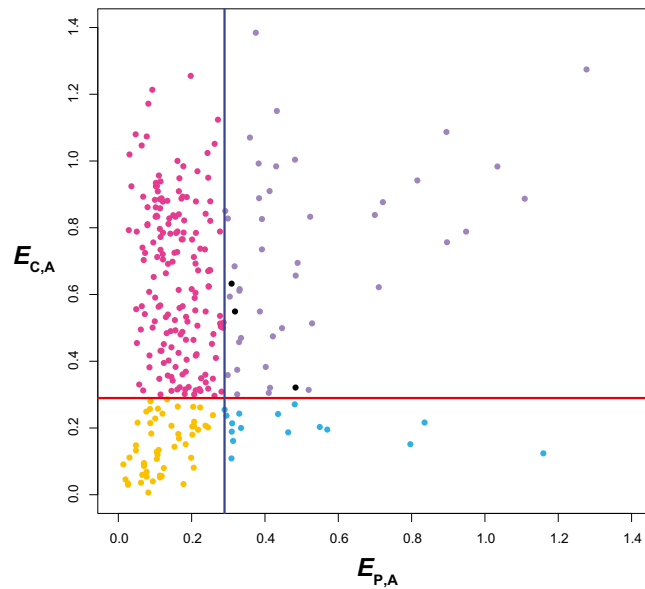## Assis and Bachtrog 10.1073/pnas.1313759110

Fig. S1. Classification of evolutionary processes maintaining individual pairs of duplicate genes. Euclidian distances between expression profiles of parent and ancestral copies ($E_{P,A}$) is plotted against Euclidian distances between expression profiles of child and ancestral copies ($E_{C,A}$) for all duplicates, which are colored by class (yellow, conservation; blue, neofunctionalization of parent copy; pink, neofunctionalization of child copy; black, subfunctionalization; purple, specialization). The cutoff used to define gene expression divergence is represented by a vertical blue line for $E_{P,A}$ and by a horizontal red line for $E_{C,A}$.
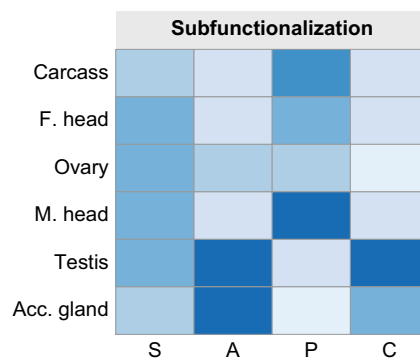

Fig. S2. Relative expression patterns of the three duplicate gene pairs classified as subfunctionalized. Mean relative expression levels in six tissues of single-copy (S), ancestral (A), parent (P), and child (C) genes. Relative expression ranges from 0% to 46%, with darker colors indicating higher values.

Table S1. Classifications resulting from application of various cutoffs to Euclidian distances

Cutoff	Value	Conservation	Neofunctionalization of parent copy	Neofunctionalization of child copy	Subfunctionalization	Specialization
Mean	0.262	49	15	160	7	50
Mean + standard deviation	0.482	119	7	141	0	14
Median	0.194	30	9	127	12	103
Median + median absolute deviation, $c = 0.303$	0.219	40	13	143	10	75
Median + median absolute deviation, $c = 1$	0.279	51	16	165	5	44
Median + median absolute deviation, $c = 1.4826$	0.320	72	13	161	2	33
Median + semi-interquartile range	0.290	53	16	167	4	41
55% quantile	0.211	36	12	138	13	82
60% quantile	0.230	41	13	147	10	70
70% quantile	0.279	51	16	165	5	44
75% quantile	0.313	68	12	163	2	36

Table S2. Classification of evolutionary processes maintaining *Drosophila melanogaster* duplicate genes that arose after different evolutionary divergence times

Divergence time	Conservation	Neofunctionalization of parent copy	Neofunctionalization of child copy	Subfunctionalization	Specialization	P
Drosophila simulans/ Drosophila sechellia	2 (3)	2 (1.33)	12 (10.17)	0 (0.17)	2 (3.33)	0.7905
Drosophila yakuba/ Drosophila erecta	0 (1.67)	0 (0.74)	9 (5.65)	0 (0.09)	1 (1.85)	0.3002
Drosophila ananassae	9 (8.50)	5 (3.78)	26 (28.81)	0 (0.47)	11 (9.44)	0.8398
Drosophila pseudoobscura	7 (4.83)	1 (2.15)	14 (16.38)	1 (0.27)	6 (5.37)	0.4082

Divergence times are listed from most to least recent. Expectations are shown in parentheses.

Table S3. Classification of evolutionary processes maintaining *D. pseudoobscura* duplicate genes that arose after different evolutionary divergence times

Divergence time	Conservation	Neofunctionalization of parent copy	Neofunctionalization of child copy	Subfunctionalization	Specialization	P
Drosophila persimilis	16 (12.34)	3 (2.82)	34 (37.38)	2 (0.71)	6 (7.76)	0.3858
D. ananassae	11 (16.39)	2 (3.75)	55 (49.63)	0 (0.94)	13 (10.30)	0.3071
D. yakuba/D. erecta	1 (1.42)	2 (0.32)	4 (4.29)	0 (0.08)	0 (0.89)	NC
D. melanogaster	7 (4.86)	1 (1.11)	13 (14.71)	0 (0.89)	3 (3.05)	0.8402

Divergence times are listed from most to least recent. Expectations are shown in parentheses. NC, not calculated because expectations were all less than five.

Table S4. Rules for classifying evolutionary processes by divergence in gene expression localization patterns

Classification	$D_{P,A}$	$D_{C,A}$	$D_{P+C,A}$
Conservation	0	0	—
Neofunctionalization of parent copy	1	0	—
Neofunctionalization of child copy	0	1	—
Subfunctionalization	1	1	0
Specialization	1	1	1

$D_{P,A}$, divergence between expression localization patterns of parent and ancestral copies; $D_{C,A}$, divergence between expression localization patterns of child and ancestral copies; $D_{P+C,A}$, divergence between the combined parent–child expression localization patterns and those of the ancestral copy.