Supporting Information

Kapelnikov *et al.* 10.1073/pnas.0710997105

SI Methods

Microarray Analysis. To identify genes with significant changes in expression between groups, we first used MAS 5.0 software (Affymetrix). We selected genes that met the following criteria: i) were significantly present on at least two arrays of either unmated or mated groups and ii) had at least two expression value >50. This generated a dataset of 5,615 genes (see Fig. S1*A*). We then used Robust Multichip Average analysis (1) to simultaneously normalize data from all groups and used *t* tests to identify differentially expressed genes (ScoreGenes software tools, http://compbio.cs.huji.ac.il/scoregenes/) (2). Genes with a *p* value <0.05 were considered differentially expressed (432 genes, Table S2).

Results from real-time or semiquantitative PCR validated our reported microarray results for 16 individual genes identified in the oviduct as: (i) present/no change (*CG7777*, *CG12251*, *CG4019*, *period*, *cycle*, *vrille*, *clock*; *pyruvate dehydrogenase phosphatase*, *cryptochrome*, *upheld*, *armadillo*, and *dystropin*), (ii) absent (*limpet*), or (iii) increased after mating (*juvenile hormone acid methyltransferase*, *cecropin*, *drip*, *timeless*).

Functional analysis was performed by using DAVID (http://david.abcc.ncifcrf.gov/gene2gene.jsp), FlyMine (http://www.flymine.org/release10.0/aspect.do?name=Gene%20Ontology), and FuncAssociate (3) (http://llama.med.harvard.edu/cgi/func/ funcassociate), which calculate the statistical probability for the representation of genes within a given Gene Ontology (GO) category relative to the total number of genes annotated with the same function. GO terms were considered as significantly enriched after correction for multiple hypothesis testing (Table S1).

Hierarchical Clustering. To examine whether the experimental groups (unmated and mated arrays) classify to different groups on the basis of their expression profiles, we used DoublePCluster (part of the ScoreGenes package). DoublePCluster is an agglomerative hierarchical model-based unsupervised biclustering approach in which the genes and the samples are both clustered according to their expression profiles (http://leibniz.cs.huji.ac.il/tr/acc/2003/HUJI-CSE-LTR-2003-80_pcluster.pdf). Briefly, this procedure uses a generative probabilistic model, which creates a hierarchy where genes (resp. samples) are combined together if they show similar expression levels in each cluster of samples (resp. genes). This analysis demonstrates that unmated and mated arrays cluster separately ($P < 4.62 \times 10^{-17}$; see Fig. S1*B*).

We next used PCluster (also in the ScoreGenes package) to cluster genes based on differences in expression patterns between mated and unmated samples (for example, in refs. 4 and 5). PCluster, also a model-based agglomerative clustering approach, assigns genes into the same cluster if they have similar distributions of expression values across preclassified groups of samples (here unmated and mated classes; see Fig. S1*C*). We highlight in the manuscript two clusters that exhibited the greatest differences in expression between unmated and mated females (fold change ≥1.5). The output files were visualized with Java Treeview (http://taxonomy.zoology.gla.ac.uk/rod/treeview.html).

Comparison to Other Datasets. To determine which mating components regulate gene expression changes in the oviduct, we compared our oviduct dataset (432 significantly up- or down-regulated genes with a whole-body dataset of 1780 genes) (Table S2) (6). To test whether the percentage of genes shared between the two datasets (4%, 66/1,780) is significant, we compared the whole-body dataset and the oviduct dataset with the following datasets: 1) nuclear morphogenesis (575 genes) (7); 2) transcripts enriched in the mushroom body (113 genes) (8) that are not related to our datasets; and 3) transcripts responsive to ecdysone at the onset of morphogenesis (743 genes) (9), which may be of relevance to the mating response. We found that the whole-body dataset and the oviduct dataset shared a small percentage of genes with dataset 2 (2.6%, 3/113; 2.6%, 3/113 genes, respectively) and with dataset 3 (17.6%, 131/743 and 5.5%, 41/743 genes, respectively). Moreover, only 10 of the 66 genes that are shared between the oviduct and the whole-body datasets are also shared with dataset 3 but not with datasets 1 and 2.

To evaluate the significance representation of the genes differentially expressed after mating in the female oviduct, we compared our dataset of differentially expressed transcripts and proteins (449; 432 genes and 17 proteins) with the following datasets: 1) testis soma, ovary soma, testis, and ovary (856, 893, 1692, and 453 genes, respectively) (10); 2) transcripts enriched in the lower reproductive tract at 3 h after mating (1603 genes) (11); 3) Malpighian tubule (1,456 genes) (12); and 4) mouse oviduct (43 genes) (13). We also mined ESTs available in UniGene of different relevant body parts [head (8,139 genes), fat body (3,033 genes), hemocytes (2,365 genes) and salivary gland (1,370 genes)] and immune-responsive genes (400 genes) (14). We used PCluster in ScoreGenes (2) (http://compbio.cs.huji.ac.il/scoregenes/) to compare annotation patterns as described above.

Construction and Visualization of Interactome Networks. A composite *Drosophila* protein–protein interaction map, with 7,590 links between 5,394 proteins, was created from the union of three published interaction maps (15–17). These studies each reported protein–protein interactions from high-throughput yeast two-hybrid and coimmunoprecipitation assays that passed a series of quality control measures as described. We excluded from our composite network low quality interactions from these studies [specifically, we used only the subset with confidence score ≥0.5 from (15) and classes A–D from (17)]. The contributions of individual datasets to the full network are listed in Table S4*A*. Table S4*B* provides a complete list of interactions and sources for each data point contained in the composite network with quality scores from the original publications where available.

We also constructed an augmented network including interologs from published protein–protein interaction datasets for human (18, 19), worm (20) (excluding the "Noncore " dataset), and budding yeast (21) ("FYI" dataset), for a total of 10,084 links between 6,019 proteins (totals for each are listed in Table S4*A*). Interologs are inferred protein–protein interactions based on experimentally determined interactions between homologous protein pairs in other species. Homologous groups were identified by using InParanoid (22) (http://inparanoid.cgb.ki.se/; April 15, 2005 release) and were not further filtered for joint percent identity or percent length along the proteins. A complete list of the interologs and their data sources is included in Table S4*B*.

Integrated networks of protein–protein and interolog interactions, supplemented with genetic interactions between pairs of *Drosophila* genes (extracted from FlyBase http://flybase.bio.indiana.edu/), were visualized (Fig. 3 and Fig. S2 *C* and *D*) by using Cytoscape (23) version 2.5.2 (http://www.cytoscape.org/). *SI File 1* provides a Cytoscape session file containing the network data shown in Fig. S2 *C* and *D*. Table S4*B* includes a complete list of all genetic interactions used. The full network can be

explored interactively online by using the web-based network browser N-Browse (24) (http://www.gnetbrowse.org), which also provides additional network data such as predicted miRNA-target relationships from PicTar (25).

Analysis of Interactome Networks. We used protein interaction data from *Drosophila* and inferred interactions from other species (interologs) to characterize the functional relatedness of subnetworks containing the different sets of oviduct proteins and transcripts that we identified by proteomic and microarray analyses (Table S5). For each set, we calculated the average shortest distance, characteristic path length (CPL), and mutual clustering coefficient (C_{vw}) between all possible pairs of proteins in the subnetwork. Mutual clustering is a measure of neighborhood cohesiveness, and reports the degree of mutual association between neighbors of any pair of proteins (26). Proximity in the protein interaction map (low CPL) and neighborhood cohesiveness (high C_{vw}) between pairs of proteins are indicative of functional relatedness. CPL was calculated by using the Floyd-Warshall algorithm (27); C_{vw} was calculated as described (26). To determine the significance of topological parameters and to control for any potential false positive interactions or systematic bias within the composite network, we performed statistical analysis of observed parameters relative to cohorts of randomly sampled gene sets from the full network (described below).

For each set of proteins (present, up-, or down-regulated) or transcripts (differentially expressed, up-, or down-regulated) we analyzed two subnetworks: one comprising all direct interactions within the set and the nearest neighbor network (NNN) for each set (comprising all set members, their first-degree neighbors, and all interactions between any of these). We repeated these analyses for four variants of the full interactome network: the *Drosophila* compendium interactome network, the same network supplemented with interologs, and each of these with ribosomal and vacuolar ATPase genes removed.

The statistical significance of topological parameters for each oviduct subnetwork was evaluated against comparable gene sets selected at random from the full interactome map for each network variant. A *z*-score was calculated to determine significance of the observed parameter value vs. the distribution in 1,000 random sets. For NNNs, random sets were generated by selecting from the interactome the same number of genes as the starting test set and drawing in all interaction partners as well as interactions between these neighbors. To control for potential bias in the yeast two-hybrid (Y2H) data, we repeated the analysis by using random sets containing the same ratio of Y2H "baits" and "preys" as present in the test sets for the *Drosophila* interactome (we could not perform this control on the interactome including interologs because we did not have directional information for all edges from the expanded network). Table S5 provides complete summary statistics and Fig. S2 *A* and *B* shows a graphical summary of these data. For reference, CPL and C_{vw} across the full networks were in the range of 5.9–6.4 and $<10^{-7}$, respectively.

Both CPL and C_{vw} revealed significant connectivity among subnetworks for the two sets of proteins identified by multidimensional protein identification technology (MudPIT) (Table S5). For the first set (the set of all 89 expressed proteins), 56 were present in the interactome map. These showed a CPL of 5.16 vs. 6.41 for comparable random sets (*z*-score = −3.55), and C_{vw} was significantly higher compared with random sets of 56 (0.16 vs. 0.03; *z*-score = 8.25). For the second set (the set of 16 up-regulated proteins), for the 12 present in the interactome vs. random sets of 12, CPL was 3.91 vs. 6.36 (*z*-score = −3.11), and C_{vw} was 0.69 vs. 0.03 (*z*-score = 12.23). Subnetworks expanded to include interologs gave similar results (Table S5). NNNs, with or without interologs, showed similar CPL but less difference in C_{vw} relative to random sets, indicating less functional cohesiveness than the core sets (Table S5). Excluding the 11 ribosomal and vacuolar ATPase subunits gave essentially the same results (Table S5).

For the sets analyzed from the 432 differentially expressed transcripts identified by microarray analysis, no significant trend was observed compared with random sets in any of the network variants (Table S5).

Proteomic Assays. Proteins were isolated from the phenol-ethanol supernatant obtained after precipitation of DNA from TRIzol according to the manufacturer's protocol (three independent biological samples, each consisting of pooled tissues from 400–500 females, were created for unmated and mated replicates). Total protein extract (100 μg in 8 M urea and 100 mM ammonium bicarbonate) was reduced (10 mM DTT, incubated at 60°C for 30 min) and alkylated (10 mM iodoacetamide, at room temperature for 30 min). The denatured and alkylated protein mixture was diluted 3-fold with water, followed by digestion overnight at 37°C by using modified trypsin (Promega) at a 1:100 enzyme-to-substrate ratio. Trypsinized total protein extracts were desalted with a C_{18} tip (Harvard), eluted with 90% acrylonitrile (ACN), dried, and dissolved in 0.1% formic acid. Aliquots of the peptides were analyzed by MudPIT (28) with off-line SCX column (LC Packings) by using nine salt steps of 10, 20, 40, 60, 80, 100, 150, 250, and 500 mM ammonium acetate in 2% ACN with 0.1% acetic acid. The eluted fractions were dried again by vacuum centrifugation and dissolved in 0.1% acetic acid. These fraction were further resolved by reverse-phase chromatography on 0.1 × 300-mm fused silica capillaries (J&W, 100 micrometer ID) packed with POROS R2–10 reversed phase material (Applied Biosystems). The peptides were eluted with linear 80 min gradients from 5 to 95% acetonitrile containing 0.1% acetic acid at flow rates of ≈0.4 μl/min. Mass spectrometry was performed with ion-trap mass spectrometers (LCQ-DecaXP, Thermo) operated in the positive mode by using repetition of full MS scan followed by fragmentation of the three most dominant ions selected from the first full MS scan.

The mass spectrometry data were clustered and analyzed and identified by Pep-Miner (29) and Sequest 3.1 software (Thermo) against the Insects section of the NCBI NR database. Pep-Miner organized the data, reduced its volume, improved the spectra quality, and increased the confidence in identification results. MS/MS data from the different chromatographs were treated in a unified manner. Moreover, by clustering at the raw data level, Pep-Miner enabled us to perform a precise comparison between unmated and mated oviducts, both at the raw data level and at the identified peptide level.

A peptide was considered as high quality if its Pep-Miner identification score was >80, the Sequest Xcore of 1.5 for singly charged peptides, 2.5 for doubly charged peptides, and 3 for triply charged peptides, and a normalized correlation score (ΔCn) ≥ 0.1). In addition, the correctness of identification of individual peptides was assessed visually by a trained operator. Using these tools, we detected a total of 178 proteins in unmated and mated oviducts.

Proteomic Analysis. Some parameters, such as the peptide hit number, spectral sampling, and the number of peptides per protein (30–33), can be considered as indicators for protein abundance in the analyzed sample. To determine the direction (increase or decrease) of changes in protein abundance we used the number of peptides per identified protein as a semiquantitative measure of protein abundance. Our approach was to perform a pair-wise comparison, that is, to compare the relative abundance of a protein in mated vs. unmated oviducts.

We first filtered the detected proteins for reproducibility and high quality. We selected proteins that met the following criteria: (1) had at least two different peptides with Pep-Miner identification score >80, the Sequest Xcore of 1.5 for singly charged peptides, 2.5

for doubly charged peptides, and 3 for triply charged peptides, and ΔCn ≥0.1 per identified protein in at least two replicates; and (2) showed the same direction of regulation in at least two replicates from either unmated or mated groups. This generated a dataset of 89 proteins of 178 that were identified in the oviducts of unmated and mated females (Table S3*B*).

We further filtered the dataset to select proteins that at least 2/3 replicates of unmated and at least 2/3 replicates of mated showed the same direction of change (i.e., increase vs. decrease). We compared the relative abundance of each protein in our dataset (89 proteins) in mated vs. unmated oviducts. We then average the relative abundance for each replicate that showed the same direction of regulation as follows:

$$\frac{1}{3}\sum_{i=1}^{3}\frac{mated_i}{unmated_i}, \quad \text{where i = replicate number 1,2,3.} \quad \textbf{[1]}$$

Proteins that showed (1) the same direction of regulation in at least two replicates of unmated and mated groups (>4/6) and (2) an increase or decrease of at least 2-fold were considered differentially expressed proteins. This generated a dataset of 17 differentially expressed proteins of 89 that were identified in the oviducts of unmated and mated females and met the above criteria (Table S3*A*).

SDS/PAGE and Western Blot Analysis. To validate that the proteins detected by MudPIT are differentially expressed, extracts of 60 oviducts from unmated and mated females were loaded and separated by SDS/PAGE. Proteins were blotted by using primary antibodies against: Coracle (Developmental Studies Hybridoma Bank), Adducin (DSHB), Neuroglian (M. Hortsch, University of Michigan, Ann Arbor, MI), Mlp84B (M. Beckerle, University of Utah, Salt Lake City), and Na pump α subunit (DSHB). Proteins were incubated with secondary antibodies and visualized by enhanced chemiluminescence. The membrane was then scanned, and the signal intensity of each band was determined by using Image J (1.37d, National Institutes of Health) software. Four independent biological preparations of unmated and mated oviducts were performed for each of the selected proteins. Relative protein levels in mated oviducts vs. unmated oviducts were then calculated (see Table S3C).

Comparison of Proteomic and Genomic Datasets. To match genes represented on microarrays with corresponding gene products identified in proteomic experiments, we compared the proteomic (89 proteins) and genomic (5,615 present transcripts) datasets. We then determined: 1) whether transcripts were observed for proteins detected in the oviduct; and 2) whether differentially expressed proteins also had differentially expressed transcripts as described (11). Briefly, we categorized the identified proteins as follows: proteins with transcript present in unmated and mated female were scored as positive, proteins that did not have a corresponding transcript in unmated or mated female were scored as negative, and proteins with transcript in unmated or mated female were scored as intermediate.

***In Situ* Hybridization.** Digoxigenin (DIG)-labeled sense and antisense riboprobes were prepared from linearized plasmids by using a T3 RNA transcription kit (Roche) according to the manufacturer's instructions. The *cecA1* riboprobe was derived from Bluescript plasmid k-7 bearing the *cecA1* cDNA insert (34) provided by D. Hultmark (Umeå University, Umeå, Sweden). Unmated and mated females at 3 h after mating were fixed in 4% paraformaldhyde in PBS containing 1% Triton X-100 for 1 h, rinsed 3 times (10 min each time) in PBS containing 0.1% Triton X-100 (PBST), and incubated overnight in 25% sucrose in PBS at 4°C. Fixed flies were rinsed in PBST and thereafter had their wings and legs removed before embedding into optimal cutting temperature freezing medium (TissueTek) and frozen at −20°C. Cryosections of 10 μm thick were thaw-mounted on SuperFrost plus slides (Menzel-Glaser), air dried for 20 min, and stored in −80°C until used. Hybridizations were carried out as described (35). Briefly, sections were covered with hybridization solution containing DIG-labeled riboprobe and incubated overnight. Hybridized sections were rinsed, blocked, and then incubated with anti-DIG alkaline phosphatase-conjugated antibody (Roche). Hybridized DIG-labeled probes were visualized with the HNPP Fast Red fluorescence detection set (Roche). Note that the *cecA1* RNA probe prepared can cross react with both *cecA2* and *cecB*.

Analysis of *Pep$_x$-GFP* Localization. Three-day-old unmated *cec-GFP* or other *Pep$_x$-GFP* females were rapidly screened under UV light and GFP filter to identify individuals with the most closely matched fluorescence intensity of the spermathecae (which fluoresce strongly and are easily seen through the abdomen cuticle), to normalize fluorescence intensity, and to ensure that flies did not encounter immune challenge. Selected females were used to examine the spatial distribution of antimicrobial peptides in the reproductive tract, or used for mating experiments as described in Material and Methods, sample preparation section. Reproductive tracts were dissected on ice in PBS, washed in a drop of PBS, and mounted in antifade media (36) on a multiwell glass slide (Hendley). Each slide held reproductive tracts from unmated and mated females.

Microscopy. Hybridized sections and reproductive tracts were viewed with a Zeiss 510 laser scanning confocal microscope by using an ×20 objective with additional zooming. Fluorescence was detected by using an argon excitation laser 488 nm and the appropriate filter sets: 650 nm for HNPP Fast Red fluorescence and 505–530 nm for GFP. Optical sections from different focal plans of each reproductive tract region (lateral oviducts, common oviduct, seminal receptacle and spermathecae) were collected and projected as a reconstructed three-dimensional image by using LSM image browser (version 3,5,0,376) software. Image collection and parameter settings were identical for each of the different reproductive tract regions analyzed.

Evaluation of Fluorescence Intensity Level. Fluorescence of *cecA1*-hybridized sections and *Pep$_x$-GFP* female reproductive tracts was quantified by using Image J by local mean fluorescence intensity level (LMFIL) approach as described (37). The LMFIL value quantifies the mean intensity at five different locations within the selected reproductive tract region. One-way ANOVA (SPSS 15.0) was used to measure the difference in *cec-GFP* intensity level between unmated and mated females as in (37). Same protocol was used to measure the difference in *cecA1* intensity level in hybridized sections of unmated and mated females.

SI File 1. This Cytoscape 2.5.2 session file, available at http://xerces.bio.nyu.edu/~kris/manuscripts;Kapelnikov_2008/, contains a description of the two NNN subnetworks shown in Fig. 3 and Fig. S2 *C* and *D* that can be viewed using Cytoscape network visualization software. See the Cytoscape website for downloads and other information, including Java version requirements (http://www.cytoscape.org).

1. Bolstad BM, Irizarry RA, Astrand M, Speed TP (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 19:185–193.
2. Kaminski N, Friedman N (2002) Practical approaches to analyzing results of microarray experiments. *Am J Respir cell Mol Biol* 27:125–132.
3. Berriz GF, King OD, Bryant B, Sander C, Roth FP (2003) Characterizing gene sets with FuncAssociate. *Bioinformatics* 19:2502–2504.
4. Dekel B, *et al.* (2003) Human and porcine early kidney precursors as a new source for transplantation. *Nat Med* 9:53–60.
5. Segman RH, *et al.* (2005) Peripheral blood mononuclear cell gene expression profiles identify emergent post-traumatic stress disorder among trauma survivors. *Mol Psychiatry* 10:421–422.
6. McGraw LA, Gibson G, Clark AG, Wolfner MF (2004) Genes regulated by mating, sperm, or seminal proteins in mated female *Drosophila melanogaster*. *Curr Biol* 14:1509–1514.
7. Pilot F, Philippe JM, Lemmers C, Chauvin JP, Lecuit T (2006) Developmental control of nuclear morphogenesis and anchoring by charleston, identified in a functional genomic screen of *Drosophila* cellularisation. *Development* 133:711–723.
8. Kobayashi M, *et al.* (2006) Differential microarray analysis of *Drosophila* mushroom body transcripts using chemical ablation. *Proc Natl Acad Sci USA* 103:14417–14422.
9. Beckstead RB, Lam G, Thummel CS (2005) The genomic response to 20-hydroxyecdysone at the onset of *Drosophila* metamorphosis. *Genome Biol* 6:R99–R99.13.
10. Parisi M, *et al.* (2004) A survey of ovary-, testis-, and soma-biased gene expression in *Drosophila melanogaster* adults. *Genome Biol* 5:R40–R40.18.
11. Mack PD, Kapelnikov A, Heifetz Y, Bender M (2006) Mating-responsive genes in reproductive tissues of female *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 103:10358–10363.
12. Wang J, *et al.* (2004) Function-informed transcriptome analysis of *Drosophila* renal tubule. *Genome Biol* 5:R69–69.21.
13. Fazeli A, Affara NA, Hubank M, Holt WV (2004) Sperm-induced modification of the oviductal gene expression profile after natural insemination in mice. *Biol Reprod* 71:60–65.
14. De Gregorio E, Spellman PT, Tzou P, Rubin GM, Lemaitre B (2002) The Toll and Imd pathways are the major regulators of the immune response in *Drosophila*. *EMBO J* 21:2568–2579.
15. Giot L, *et al.* (2003) A protein interaction map of *Drosophila melanogaster*. *Science* 302:1727–1736.
16. Stanyon CA, *et al.* (2004) A *Drosophila* protein-interaction map centered on cell-cycle regulators. *Genome Biol* 5:R96–R96.14.
17. Formstecher E, *et al.* (2005) Protein interaction mapping: A *Drosophila* case study. *Genome Res* 15:376–384.
18. Rual JF, *et al.* (2005) Towards a proteome-scale map of the human protein–protein interaction network. *Nature* 437:1173–1178.
19. Stelzl U, *et al.* (2005) A human protein–protein interaction network: A resource for annotating the proteome. *Cell* 122:957–968.
20. Li S, *et al.* (2004) A map of the interactome network of the metazoan *C. elegans*. *Science* 303:540–543.
21. Han JD, *et al.* (2004) Evidence for dynamically organized modularity in the yeast protein–protein interaction network. *Nature* 430:88–93.
22. O'Brien KP, Remm M, Sonnhammer EL (2005) InParanoid: A comprehensive database of eukaryotic orthologs. *Nucleic Acids Res* 33:D476–D480.
23. Shannon P, *et al.* (2003) Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504.
24. Lall S, *et al.* (2006) A genome-wide map of conserved microRNA targets in *C. elegans*. *Curr Biol* 16:460–471.
25. Grun D, Wang YL, Langenberger D, Gunsalus KC, Rajewsky N (2005) microRNA target predictions across seven *Drosophila* species and comparison to mammalian targets. *PLoS computational biology* 1:e13–e13.16.
26. Goldberg DS, Roth FP (2003) Assessing experimentally derived interactions in a small world. *Proc Natl Acad Sci USA* 100:4372–4376.
27. Cormen TH, Leiserson CE, Rivest RL, Stein C (2001) *Introduction to Algorithms*, (MIT Press & McGraw–Hill, Cambridge), 2nd Ed, p 1180.
28. Link AJ, *et al.* (1999) Direct analysis of protein complexes using mass spectrometry. *Nat Biotechnol* 17:676–682.
29. Beer I, Barnea E, Ziv T, Admon A (2004) Improving large-scale proteomics by clustering of mass spectrometry data. *Proteomics* 4:950–960.
30. Corbin RW, *et al.* (2003) Toward a protein profile of Escherichia coli: Comparison to its transcription profile. *Proc Natl Acad Sci USA* 100:9232–9237.
31. Gao J, Opiteck GJ, Friedrichs MS, Dongre AR, Hefta SA (2003) Changes in the protein expression of yeast as a function of carbon source. *J Proteome Res* 2:643–649.
32. Liu H, Sadygov RG, Yates JR, III (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem* 76:4193–4201.
33. Zhang B, *et al.* (2006) Detecting differential and correlated protein expression in label-free shotgun proteomics. *J Proteome Res* 5:2909–2918.
34. Kylsten P, Samakovlis C, Hultmark D (1990) The cecropin locus in *Drosophila*: A compact gene cluster involved in the response to infection. *EMBO J* 9:217–224.
35. Krieger J, *et al.* (2002) A divergent gene family encoding candidate olfactory receptors of the moth *Heliothis virescens*. *Eur J Neurosci* 16:619–628.
36. Giloh H, Sedat JW (1982) Fluorescence microscopy: Reduced photobleaching of rhodamine and fluorescein protein conjugates by *n-propyl gallate*. *Science* 217:1252–1255.
37. Heifetz Y, Wolfner MF (2004) Mating, seminal fluid components, and sperm cause changes in vesicle release in the *Drosophila* female reproductive tract. *Proc Natl Acad Sci USA* 101:6261–6266.

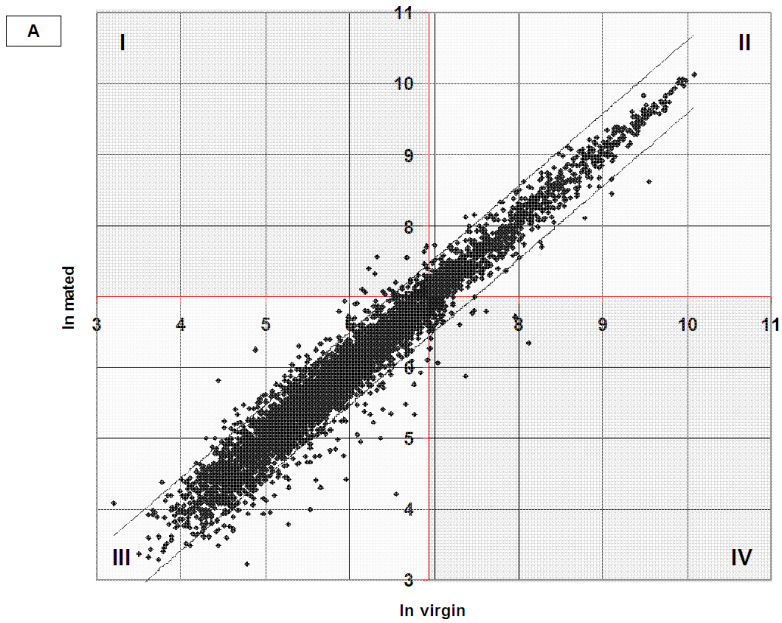


Fig. S1. Oviduct genes and gene expression pattern in the 432 oviduct mating-responsive genes. Scatter plot shows oviduct ''present'' calls of unmated and mated female. Data are pair-wise comparisons of natural log (Ln) of signals averaged from 3 biological replicates (*A*). Unsupervised hierarchical clustering dendogram of the experimental groups (unmated and mated arrays) on the basis of the similarity of their expression profiles, demonstrates that unmated and mated arrays cluster separately ($P < 4.62 \times 10^{-17}$) and that the expression profiles of the unmated arrays or the mated arrays are not significantly different from each other (*B*). Gene expression patterns in the oviducts of unmated and mated females (*C*). Shown is hierarchical clustering dendogram and expression profile of groups of genes with ''similar'' expression pattern. Highlighted are clusters of the most: (1) up-regulated genes and (2) down regulated genes in mated female ($P < 0.05$). Expression ratios range from -1.8 (purple) to $+$ 1.8 (yellow). See also Fig. 1. (*D*) Genes differentially expressed in oviduct and whole body assay. Mating-responsive genes showed more extreme changes in the oviduct than in whole body (highlighted cluster; see also Fig. 1). Genes from cluster shown in Fig. 1*C* are highlighted in red (immune-related genes).

B

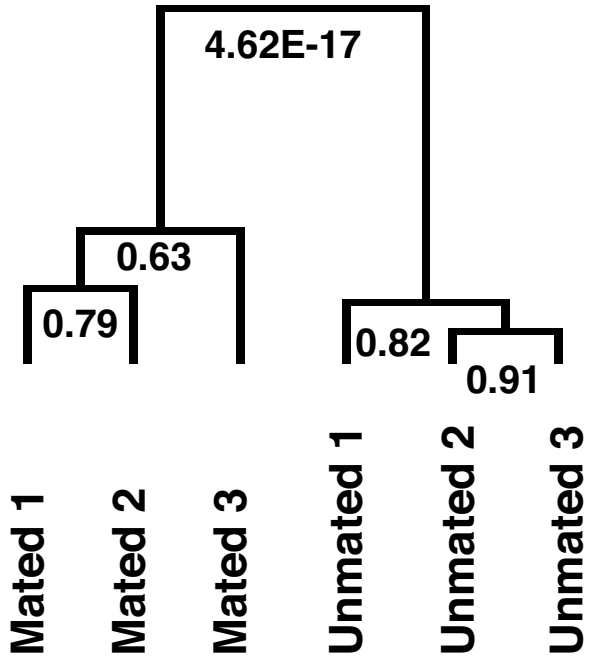

4.62E-17
0.63
0.79
0.82
0.91
Mated 1
Mated 2
Mated 3
Unmated 1
Unmated 2
Unmated 3

Fig. S1 cont.

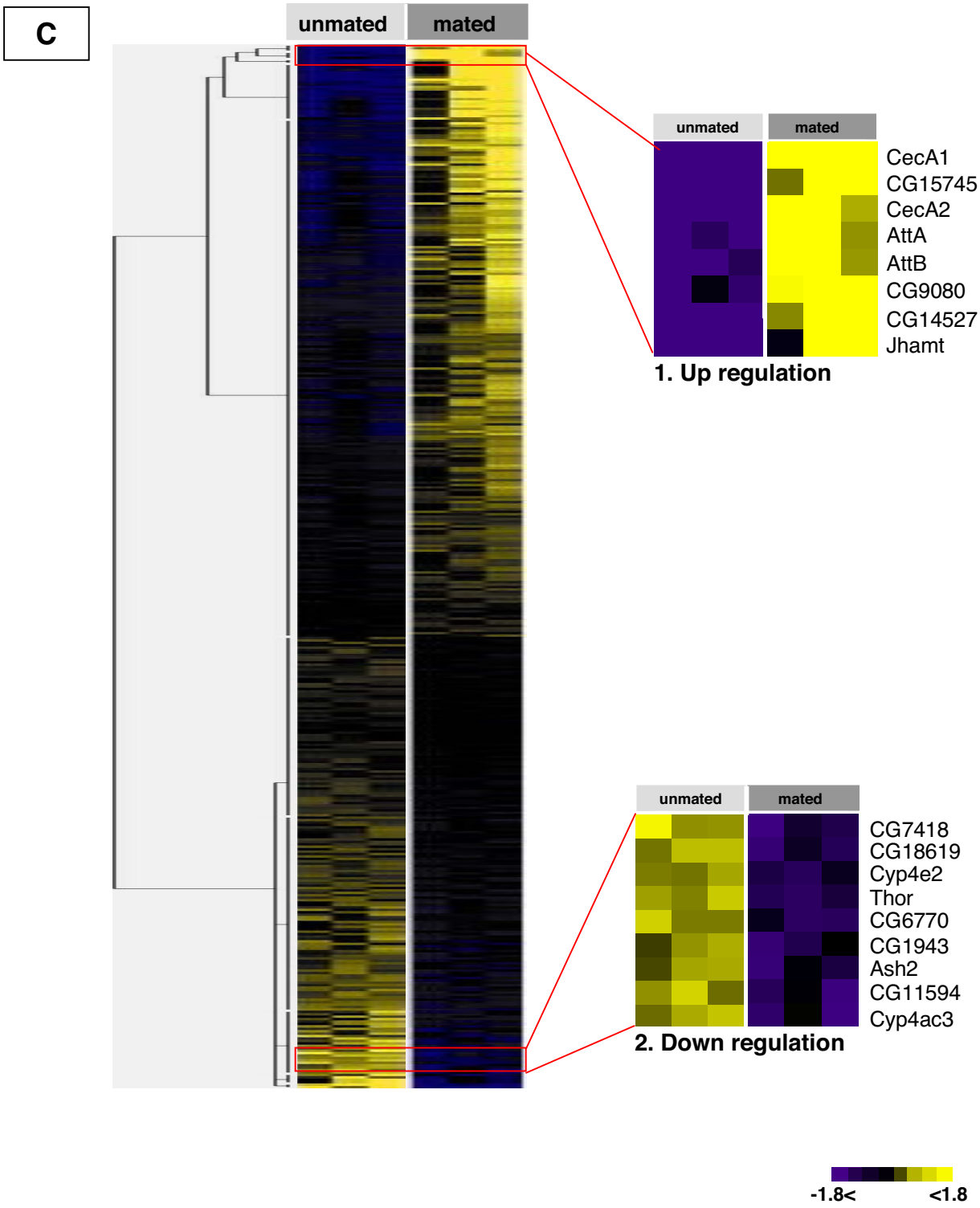
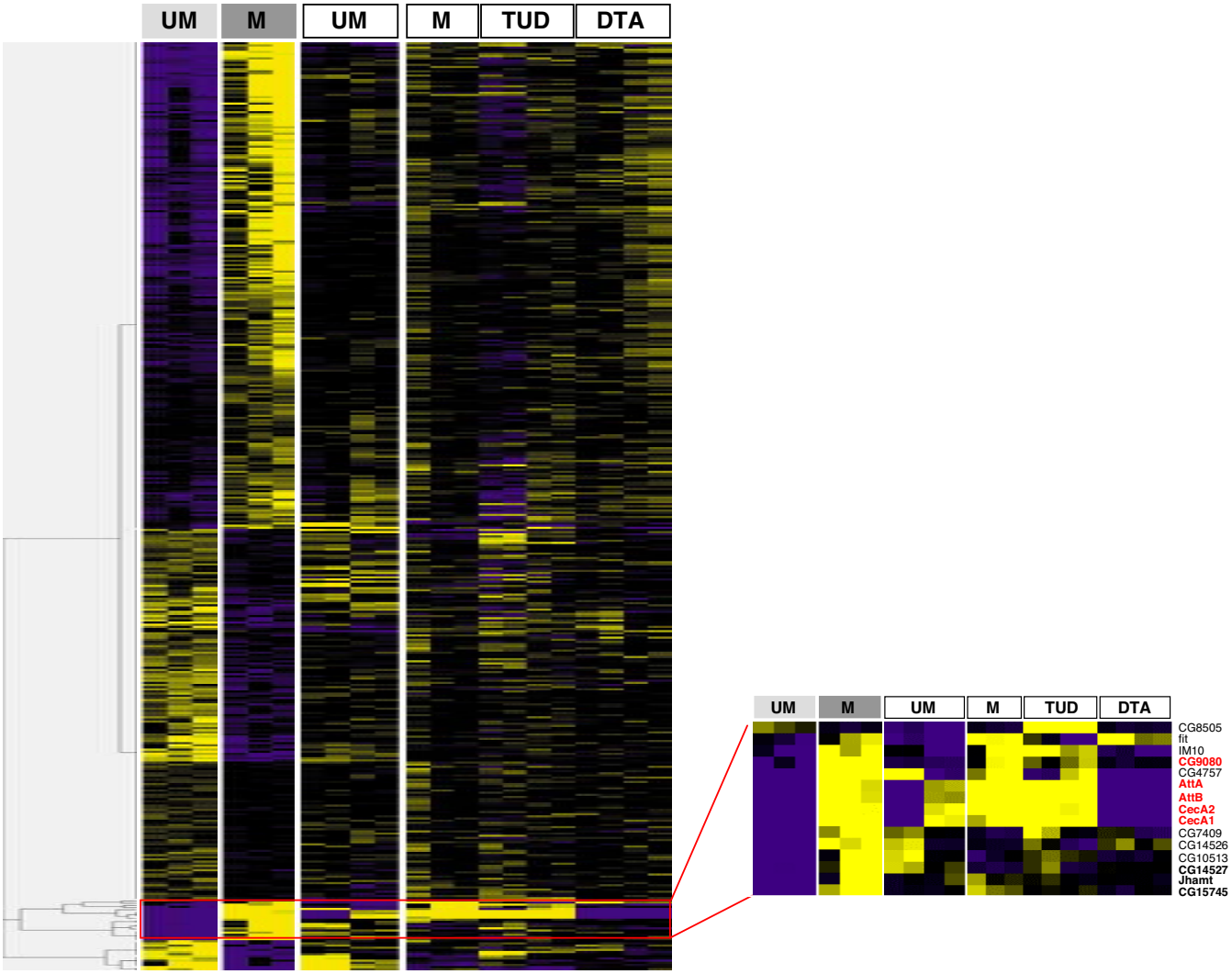

C
unmated
mated
unmated
mated
CecA1
CG15745
CecA2
AttA
AttB
CG9080
CG14527
Jhamt
1. Up regulation
unmated
mated
CG7418
CG18619
Cyp4e2
Thor
CG6770
CG1943
Ash2
CG11594
Cyp4ac3
2. Down regulation
-1.8<
<1.8

Fig. S1 cont.

D



UM	Oviducts of unmated females
M	Oviducts of females mated to wild type males
UM	Whole body of unmated female
M	Whole body of females mated to wild type males
TUD	Whole body of females mated to spermless males
DTA	Whole body of females mated to males lacking sperm and Acps

Fig. S1 cont.

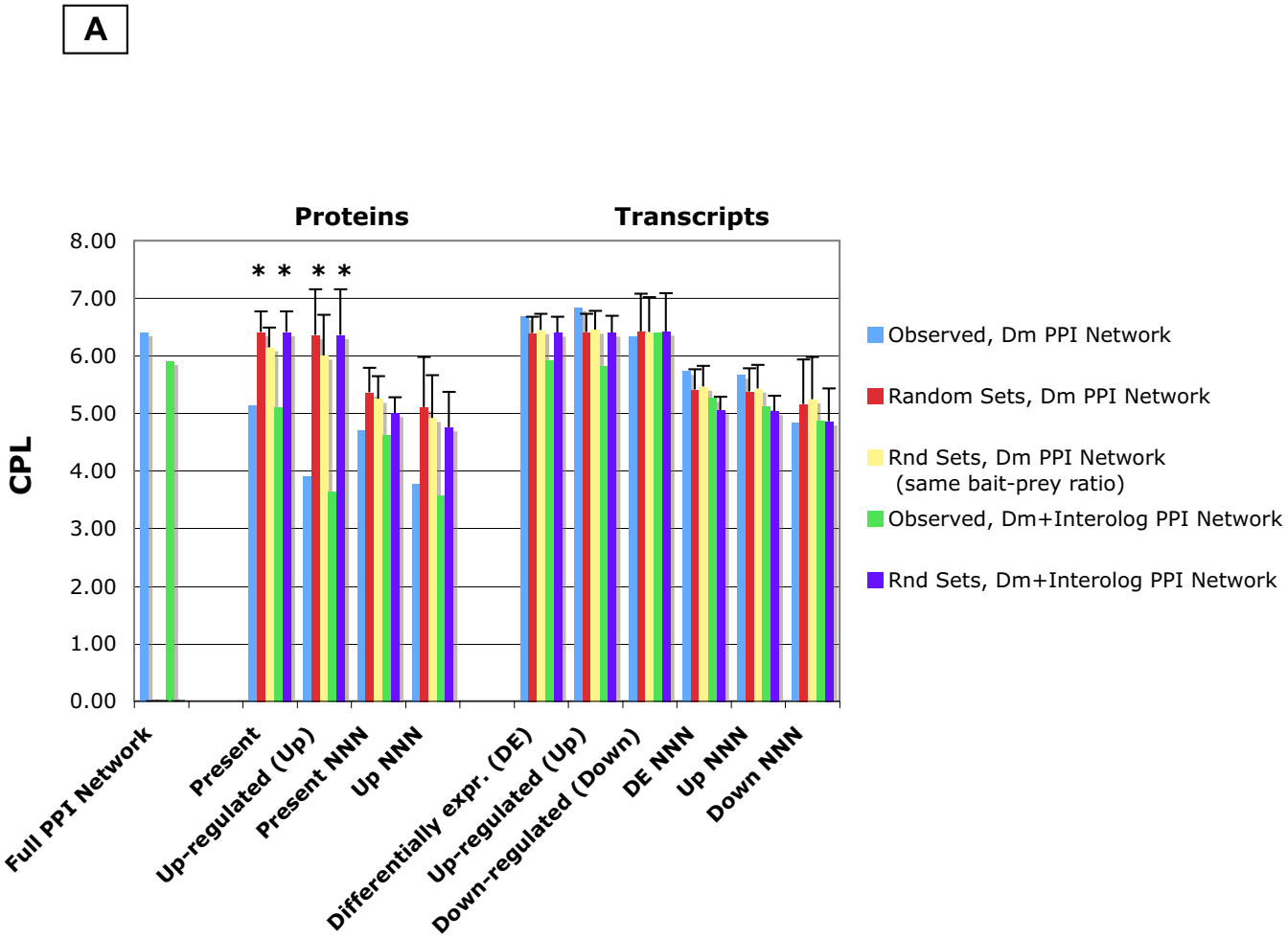


Fig. S2. Analysis of molecular networks for mating-responsive gene products in the oviduct based on a composite molecular interaction map for *Drosophila*. (*A* and *B*) Bar charts of (*A*) characteristic path length (CPL) and (*B*) mutual clustering coefficient (C_{vw}) in protein–protein interaction subnetworks for oviduct-expressed protein and transcript sets vs. random sets (see Table S5 for raw data). Table S5 provides raw data and statistics for these networks and variants that exclude ribosomal and vacuolar H^+ ATPase subunits (not shown here). Subnetworks were drawn from the *Drosophila* protein–protein interaction (''*Dm* PPI'') network or the *Dm* PPI network supplemented with interologs from human, worm, and yeast (''*Dm*+Interolog PPI''). See Table S4 for details on networks. Left-hand set of bars, statistics across each full PPI network. Middle set of bars, subnetworks of proteins detected by MudPIT. Right-hand set of bars, subnetworks of transcripts detected by microarray. All gene products (''Present''), just those whose expression changed significantly after mating [''Differentially Expr. (DE)'', ''Up-regulated (Up)'', ''Down-regulated (Down)''], and nearest neighbor networks (NNNs) for each of these were compared with 1000 subnetworks generated by using similarly sized sets of proteins selected at random from the corresponding interactome map. Asterisks indicate statistically significant deviation of observed vs. random sets (*z*-score >3, corresponding to a *P* value <0.001). (*C* and *D*) Nearest neighbor networks (NNNs) for mating-responsive gene products in the oviduct: (*C*) NNN for up-regulated proteins detected by MudPIT, (*D*) NNN for differentially expressed transcripts detected by microarray. Edges, Protein–protein interactions from *Drosophila* (blue), interologs (violet), or both (black) or genetic interactions (green). Nodes, proteins or genes. Node color, expression of mRNAs in the oviduct measured by microarray [reproducibly detected in either unmated or mated samples (yellow), up-regulated on mating (red), down-regulated on mating (green), not detected (gray)]. Node borders, proteins detected by MudPIT [not detected (circles), present (squares), up-regulated (diamonds)]. *SI File 1* (http://xerces.bio.nyu.edu/~kris/manuscripts/Kapelnikov_2008/) contains a Cytoscape-compatible version of these subnetworks. The full networks can be browsed interactively online by using N-Browse (http://www.gnetbrowse.org).

B

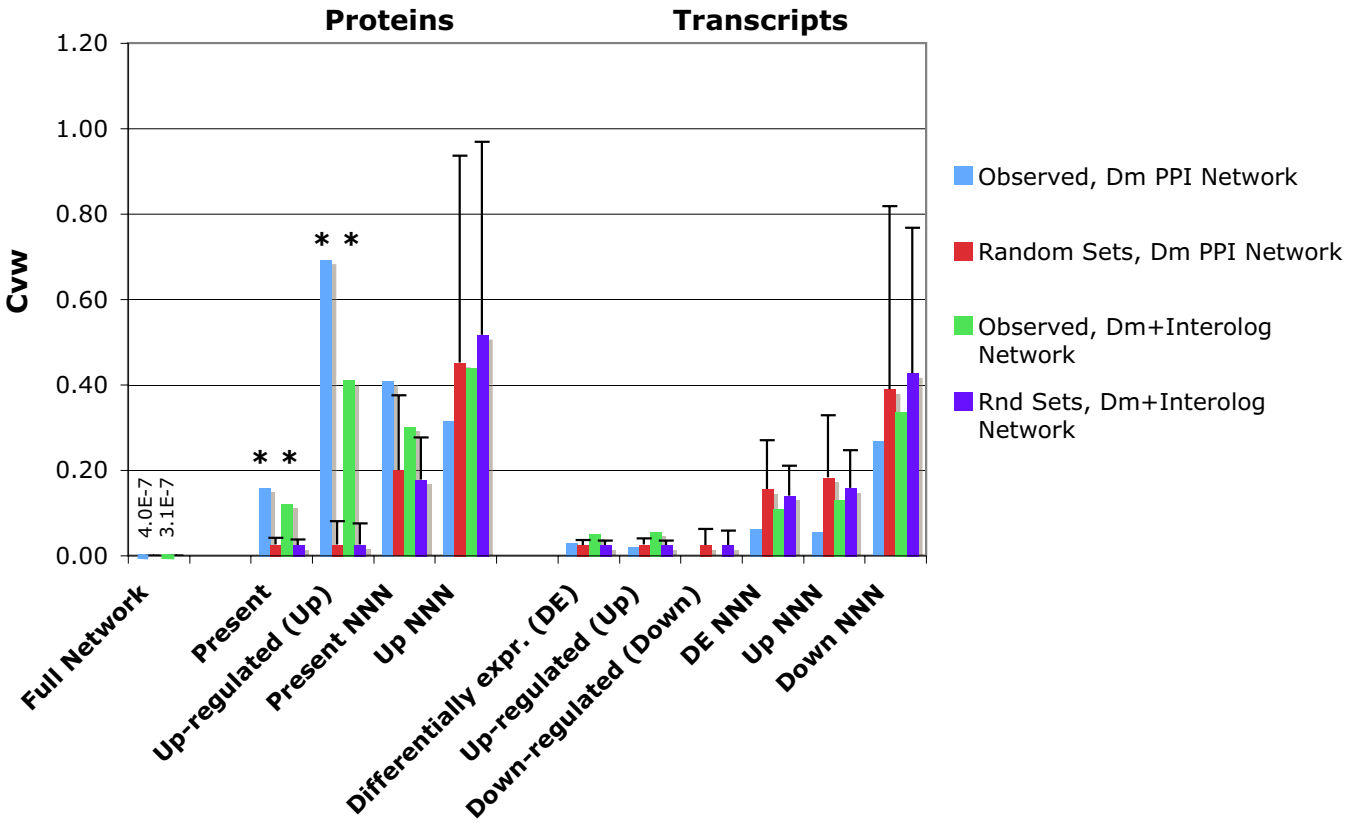

Proteins
Transcripts
1.20
1.00
0.80
0.60
0.40
0.20
0.00
Cvw
4.0E-7
3.1E-7
* *
* *
Full Network
Present
Up-regulated (Up)
Present NNN
Up NNN
Differentially expr. (DE)
Up-regulated (Up)
Down-regulated (Down)
DE NNN
Up NNN
Down NNN
Observed, Dm PPI Network
Random Sets, Dm PPI Network
Observed, Dm+Interolog Network
Rnd Sets, Dm+Interolog Network

Fig. S2 cont.

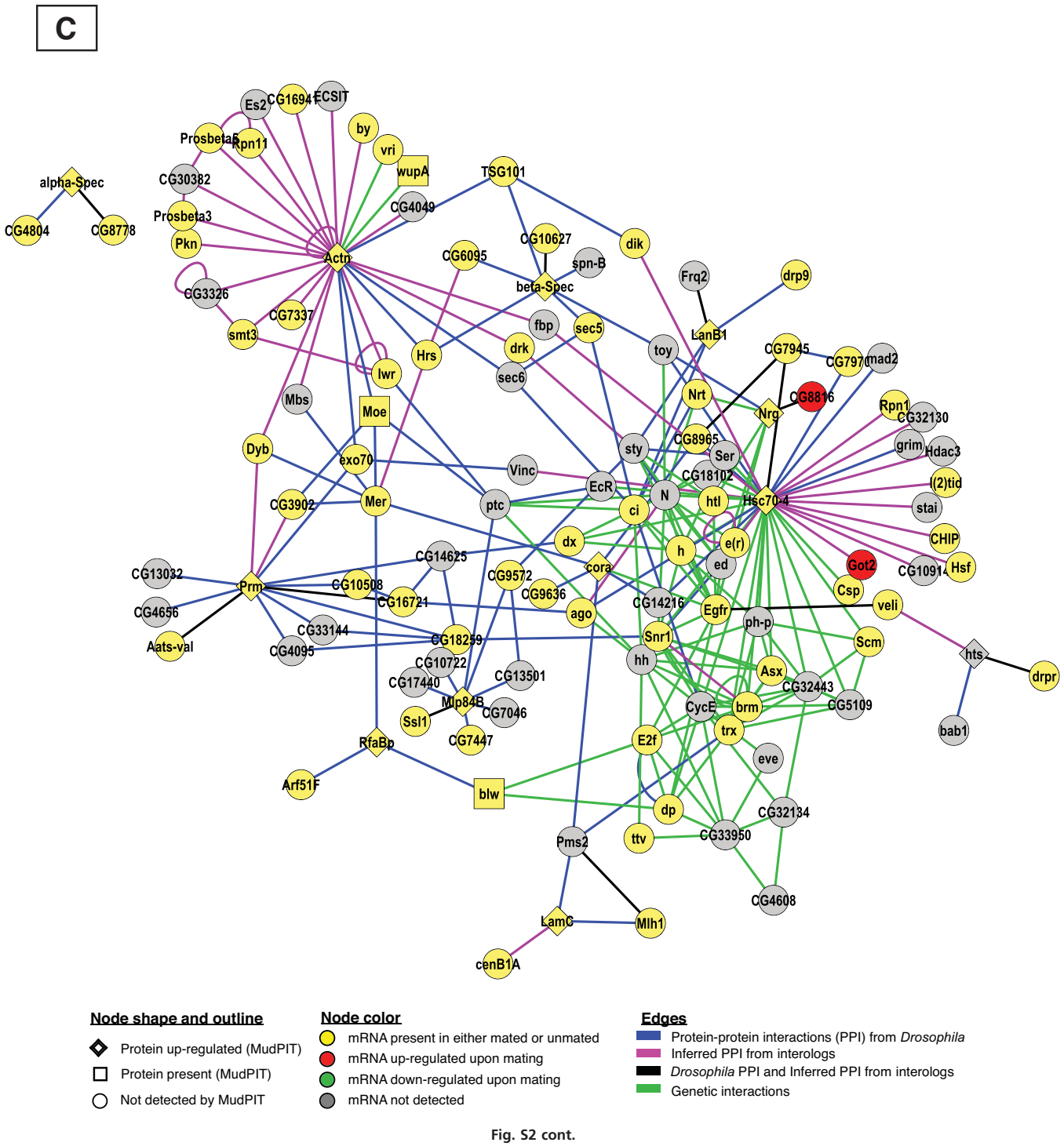

C
Es2
CG1694
ECSIT
by
vri
wupA
Prosbeta5
Rpn11
alpha-Spec
CG30382
TSG101
CG4049
Prosbeta3
CG4804
CG8778
Pkn
Actn
CG10627
dik
CG6095
spn-B
beta-Spec
Frq2
drp9
CG3326
CG7337
smt3
fbp
sec5
LanB1
Hrs
drk
toy
CG7945
CG797
mad2
lwr
sec6
Nrt
CG8816
Nrg
Rpn1
CG32130
Mbs
Moe
CG8965
grim
Hdac3
Dyb
exo70
sty
Ser
Vinc
CG18102
EcR
N
htl
Hsc70-4
I(2)tid
CG3902
Mer
ptc
ci
stai
CHIP
dx
h
e(r)
CG14625
ed
Got2
CG10911
Hsf
CG13032
cora
CG9572
Csp
Prm
CG10508
CG9636
CG16721
CG14216
Egfr
veli
CG4656
ago
ph-p
CG33144
Snr1
Scm
Aats-val
CG18259
CG4095
hh
hts
drpr
CG10722
Asx
CG17440
CG13501
CG32443
CycE
brm
CG5109
Mlp84B
CG7046
Ssl1
trx
bab1
RfaBp
CG7447
E2f
eve
Arf51F
blw
dp
CG32134
ttv
CG33950
Pms2
CG4608
LamC
Mlh1
cenB1A
Node shape and outline
Protein up-regulated (MudPIT)
Protein present (MudPIT)
Not detected by MudPIT
Node color
mRNA present in either mated or unmated
mRNA up-regulated upon mating
mRNA down-regulated upon mating
mRNA not detected
Edges
Protein-protein interactions (PPI) from Drosophila
Inferred PPI from interologs
Drosophila PPI and Inferred PPI from interologs
Genetic interactions

Fig. S2 cont.

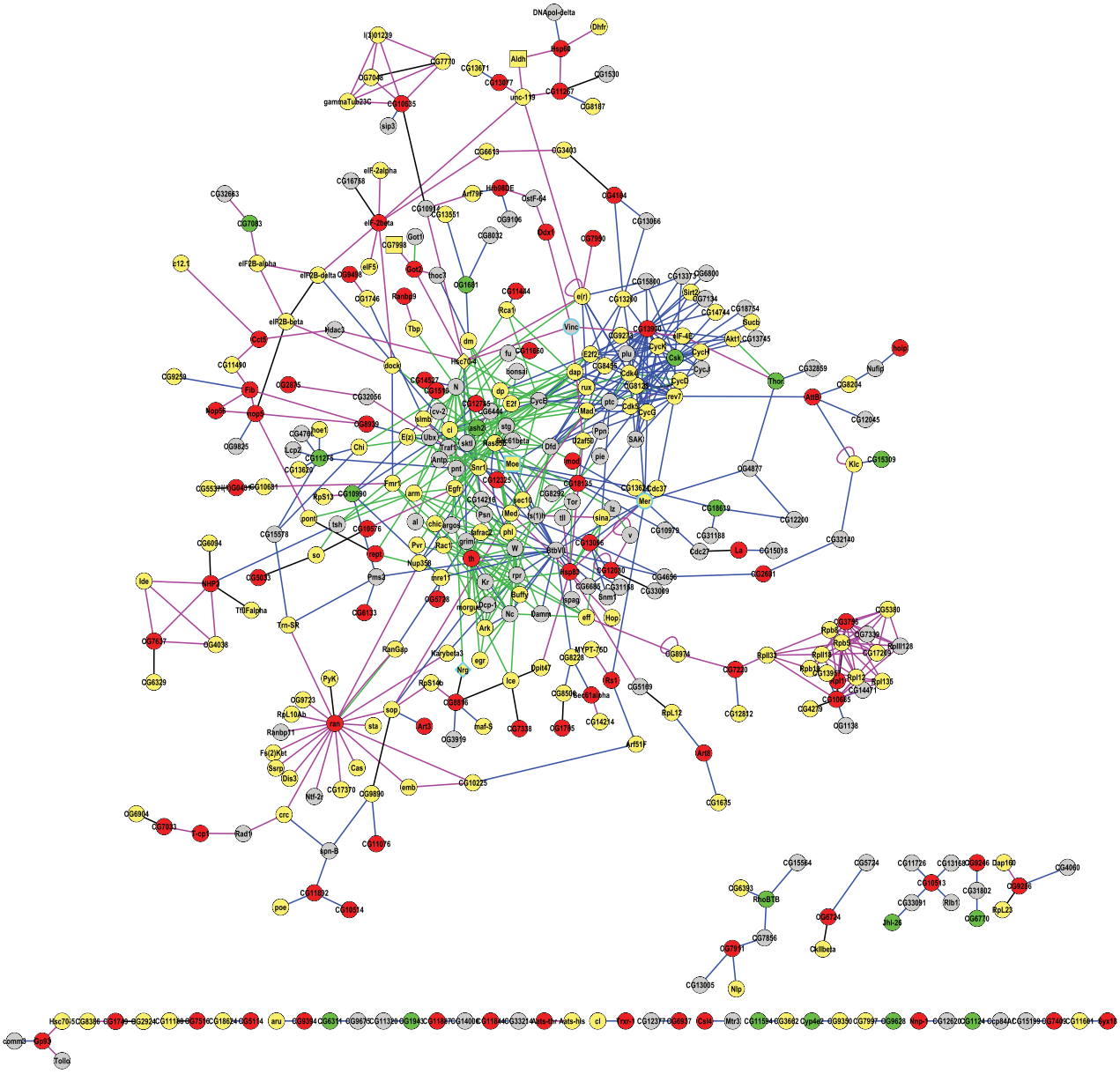

D
Node shape and outline
Protein present (MudPIT)
Not detected by MudPIT
Node color
mRNA present in either mated or unmated
mRNA up-regulated upon mating
mRNA down-regulated upon mating
mRNA not detected
Edges
Protein-protein interactions (PPI) from Drosophila
Inferred PPI from interologs
PPI from Drosophila and Inferred PPI from interologs
Genetic interactions

Fig. S2 cont.

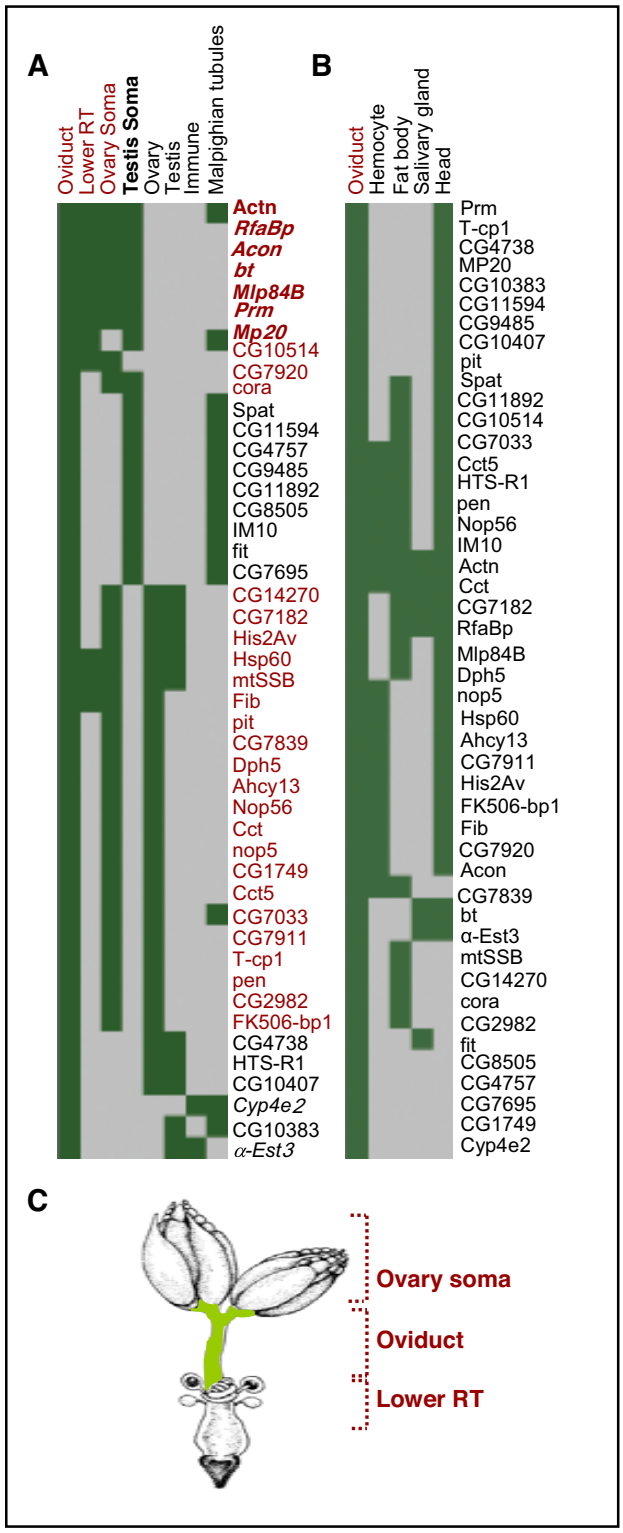


Fig. S3. Oviduct after mating signature is enriched in other reproductive tissues examined. (*A*) Clustering of oviduct after mating signature with genes enriched in reproductive or nonreproductive tissues, based on simple presence or absence in these tissues. Tissue samples (columns): Oviduct (this study); Lower RT (includes sperm storage organs and sperm stored therein); Ovary and Testis (gametes and somatic tissues that support and maintain the gametes); Malpighian tubules (nonreproductive soma); and Immune (genes induced by an immune challenge). (*B*) Clustering of oviduct after mating signature associated with other tissues from (*A*) with ESTs from hemocyte, fat body, salivary gland, and head. Highlighted in bold and italicized are genes enriched in the oviduct, ovary soma, lower RT, and testis soma; highlighted in red are genes that the oviduct shares with ovary soma and/or lower RT; genes shared between the oviduct soma and Malpighian tubules are italicized only. (*C*) Schematic representation of female reproductive tract showing the different somatic regions (ovary, oviduct, lower RT).

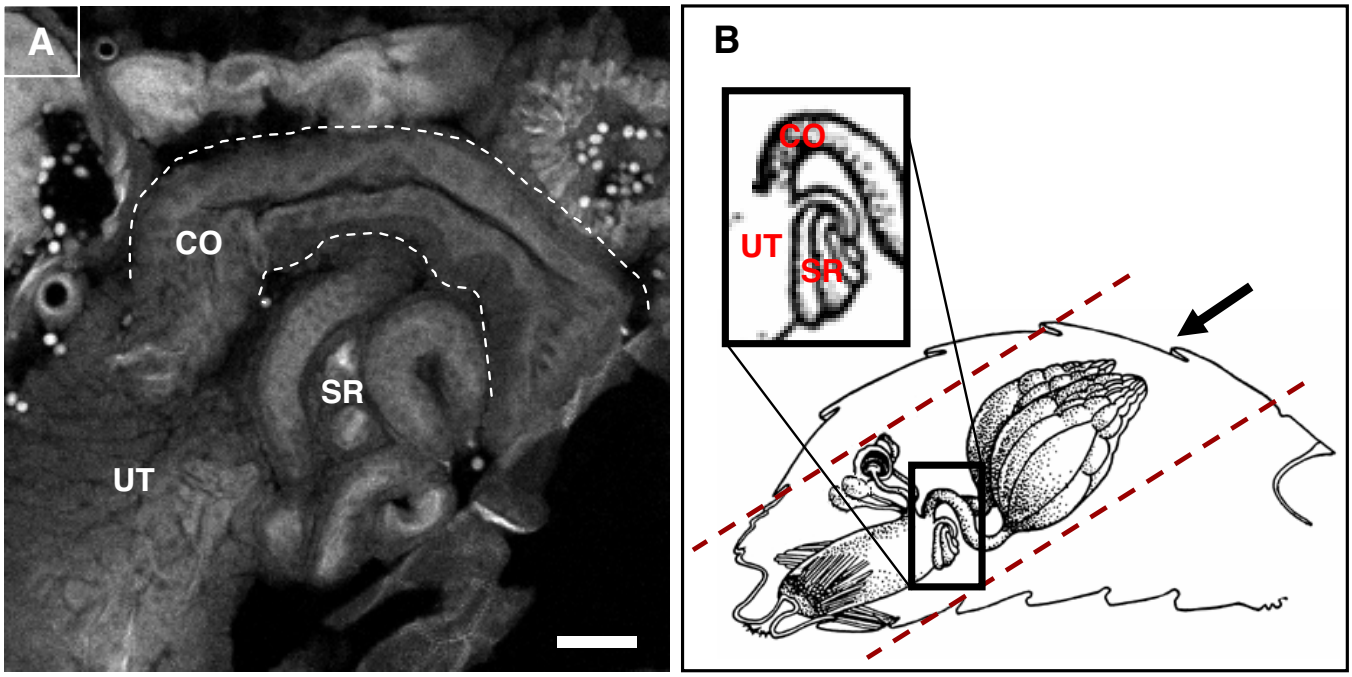


Fig. S4. *Cec* hybridization in the female reproductive tract. Control DIG-labeled sense *cecA1* probe hybridized to cryosection of unmated female (*A*); (Scale bar, 20 μm); $n = 10$ females; (*B*) Schematic of female reproductive system (adapted from FlyBase image FBim0000077) representing the direction of sectioning. UT, uterus; SR, seminal receptacle; CO, common oviduct.

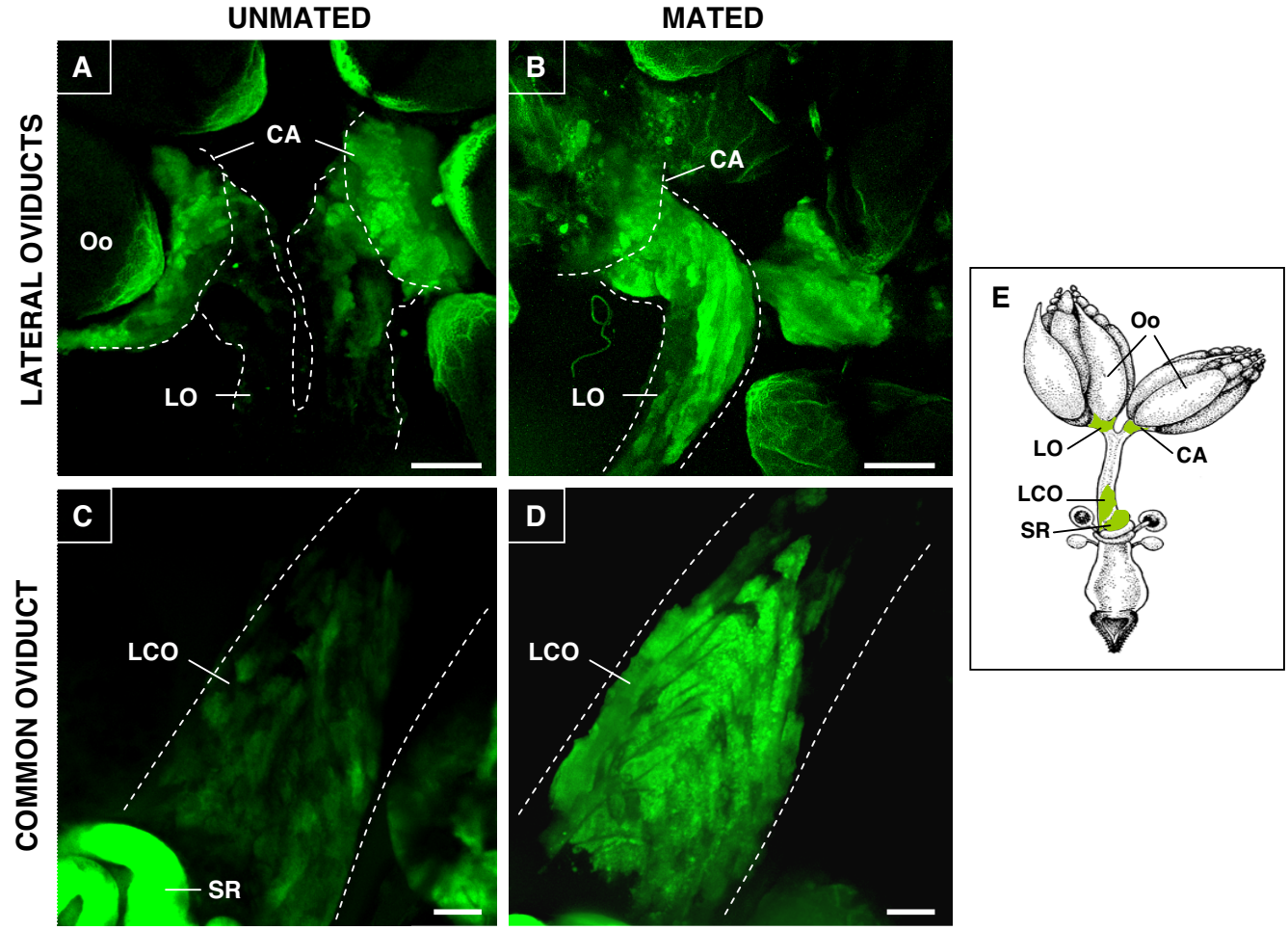


Fig. S5. The spatial localization of a reporter for Cec protein differs from that of the *cec* transcript, but shows the same after-mating induction pattern. Distribution of Cec-GFP in the lateral oviducts of (*A*) unmated and (*B*) mated female at 3 h after mating; and in the lower part (dorsal side) of the common oviduct of (*C*) unmated and (*D*) mated female. Cec-GFP is excluded from the middle part of the common oviduct (*C–E*), See also Fig. S6*A* (g) (zones I and II). Note the high intensity of Cec-GFP fluorescence in the lower common oviduct of mated vs. unmated female (unmated = 29.2 ± 2.8; mated = 44.9 ± 3.13; $P < 0.0001$); To rule out possible male contribution to the after mating fluorescence signal, we analyzed *cec-GFP* transgenic females mated to WT males, and WT females mated to *cec-GFP* transgenic males. The latter had no effect in female tissues (not shown), suggesting a solely female contribution to the after-mating oviduct immune response. (*E*) Schematic of female reproductive tract (adapted from FlyBase image FBim6932540). Oo, oocyte; CA, calyx; LO, lateral oviducts; LCO, lower common oviduct; SR, seminal receptacle. (Scale bar, 50 μm in *A–D*); $n > 36$ females for each treatment.

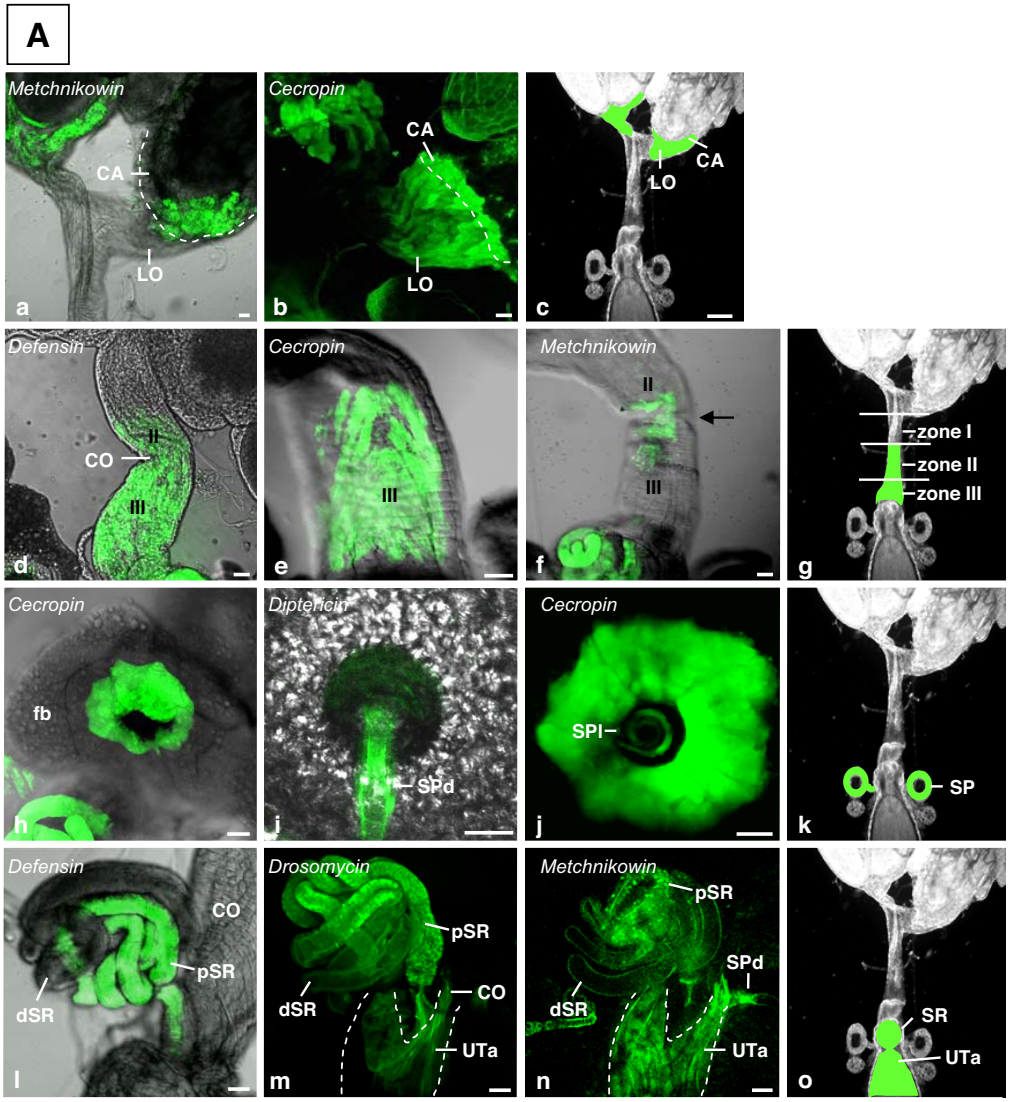


Fig. S6. Antimicrobial peptides are expressed throughout the female reproductive tract but each antimicrobial peptide has a unique expression pattern. (*A*) *Pep$_x$-GFP* flies express GFP in different regions of female reproductive tract. Selected confocal images shown represent the localization of antimicrobial peptides within: (a-c) the calyx (CA) and lateral oviducts (LO); (d-g) common oviduct (CO) and the three zones observed within the CO (I, II and III): (h-k) spermathecha (SP); and (l-o) seminal receptacle (SR) and uterus (UT). Also shown are models of reproductive tract that summarize the topographical distribution of antimicrobial peptides in the: CA and LO (c), different zones of the CO (g), SP (k) and SR and UT (o). Details on specific GFP patterns: (a) Metchnikowin-GFP fluorescence in the CA but not in the LO; (b) Cecropin-GFP fluorescence in the CA and the LO; (d) Defensin-GFP fluorescence in Zones II and III of the CO; (e) Cecropin-GFP fluorescence in a subpopulation of epithelial cells in Zone III of the CO; (f) Metchnikowin-GFP fluorescence in a subpopulation of epithelial cells in the CO between Zones II and III; (h) Cecropin-GFP fluorescence in the secretory epithelium covering the SP but not in the fat body (fb); (i) Diptericin-GFP fluorescence in the SP duct (SPd); (j) Cecropin-GFP fluorescence in the lumen of the SP (SPl); (l) Defensin-GFP fluorescence in the proximal part of the SR (pSR), note the lack of fluorescence in the distal part of the SR (dSR); and (m) Drosomycin-GFP fluorescence in the proximal part (pSR) and in the distal part (dSR) of the SR. Note the fluorescence in the anterior domain of the UT (UTa) (marked by doted line) where the pSR and the CO are connected. (n) Metchnikowin-GFP fluorescence in the pSR and dSR as shown for drosomycin in (m). Note that Metchnikowin fluorescence in the anterior part of the UT is excluded from the connection to the pSR (marked by dotted line). (*B*) Spatial localization of the different antimicrobial peptides within the female reproductive tract. A pair of fluorescent and transmitted confocal images is presented for each *Pep$_x$-GFP* fly line. (a,b) Cecropin-GFP fluorescence is localized to the CO, SR, and SP; (c,d) Drosocin-GFP fluorescence is localized to the LO, CO, SR, and SP; (e,f) Defensin-GFP fluorescence is localized to the CO and SR; (g,h) Attacin-GFP fluorescence is localized to the LO, CO, SR, and UT; note the egg within the uterus; (i,j) Metchnikowin-GFP fluorescence is localized to the CO, spermathecal duct (SPd), SR and UT; (k,l) Diptericin-GFP fluorescence is localized to the LO, CO, SR, SPd, and the UT; (m,n) Drosomycin-GFP fluorescence is localized to the CO, SR, SP, and mainly to the anterior part of the UT. For each *Pep$_x$-GFP* and mating status (unmated or mated), $n \geq 20$ females. (Scale bar for GFP images, 20 μm and for the summary models, 100 μm.) (*C*) Topographic map of different antimicrobial peptides in the female reproductive tract. Spatial modeling of antimicrobial peptides: i) The Color Scale Index represents the combination of antimicrobial peptides that can be found in each domain. The same color in different domains indicates that the same combination of antimicrobial peptides is present in these domains. ii) The Gray Scale Color Intensity Level Index represents the number of antimicrobial peptides localized to a specific domain within the reproductive tract. The highest intensity level indicates all antimicrobial peptides examined, with the lighter shades of gray (leading to white) indicating six, five, four, and one antimicrobial peptide examined. Note that it is only at the calyx and the distal part of the seminal receptacle marked as dark brown that we can find all of the antimicrobial peptides examined. Schematic adapted from FlyBase image FBim0000248.

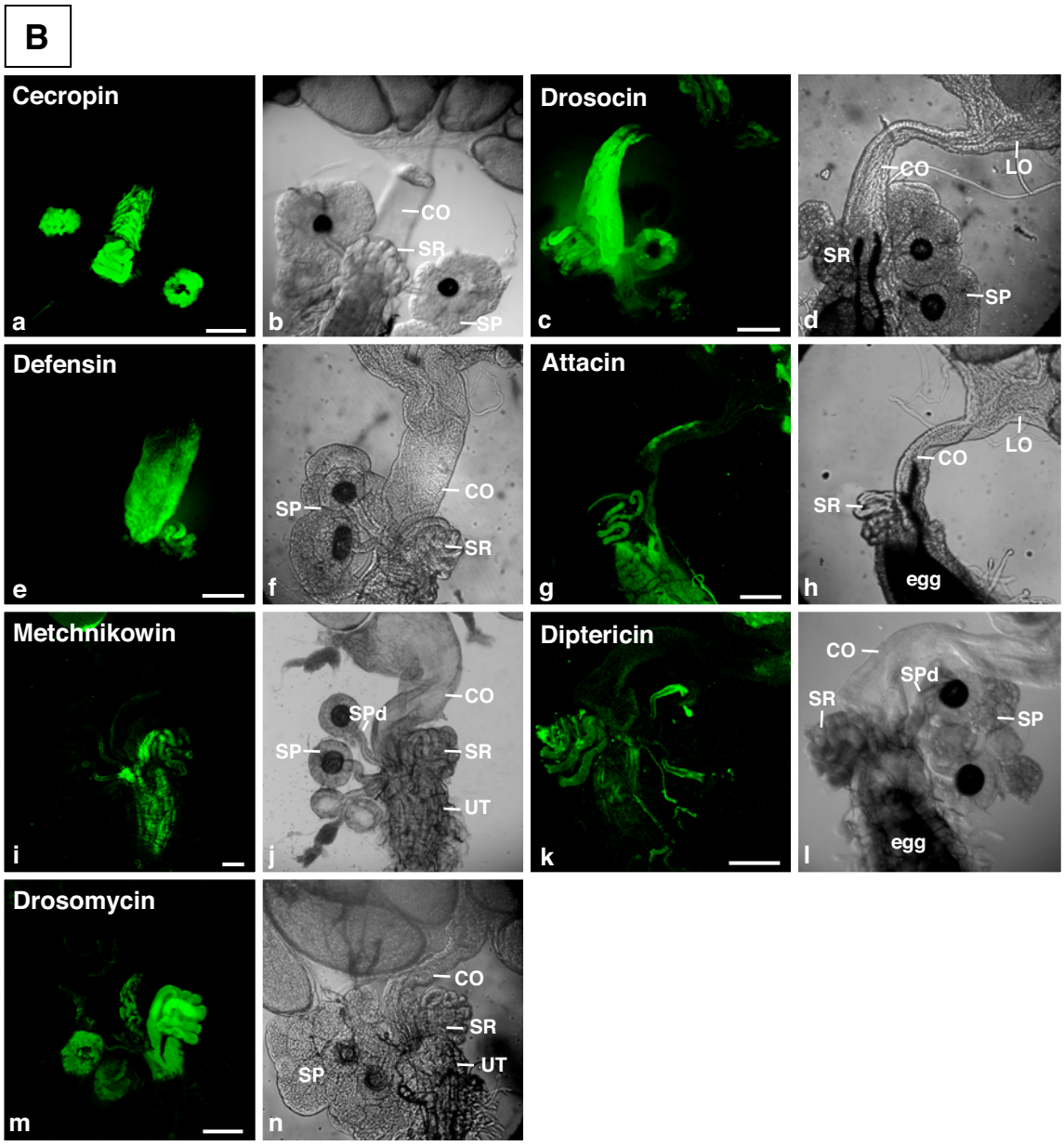

B
Cecropin
a
b
CO
SR
SP
Drosocin
c
CO
LO
SR
SP
d
Defensin
e
f
SP
CO
SR
Attacin
g
CO
LO
SR
egg
h
Metchnikowin
i
j
SPd
CO
SP
SR
UT
Diptericin
k
CO
SPd
SR
SP
egg
l
Drosomycin
m
n
CO
SR
SP
UT

Fig. S6 cont.

C

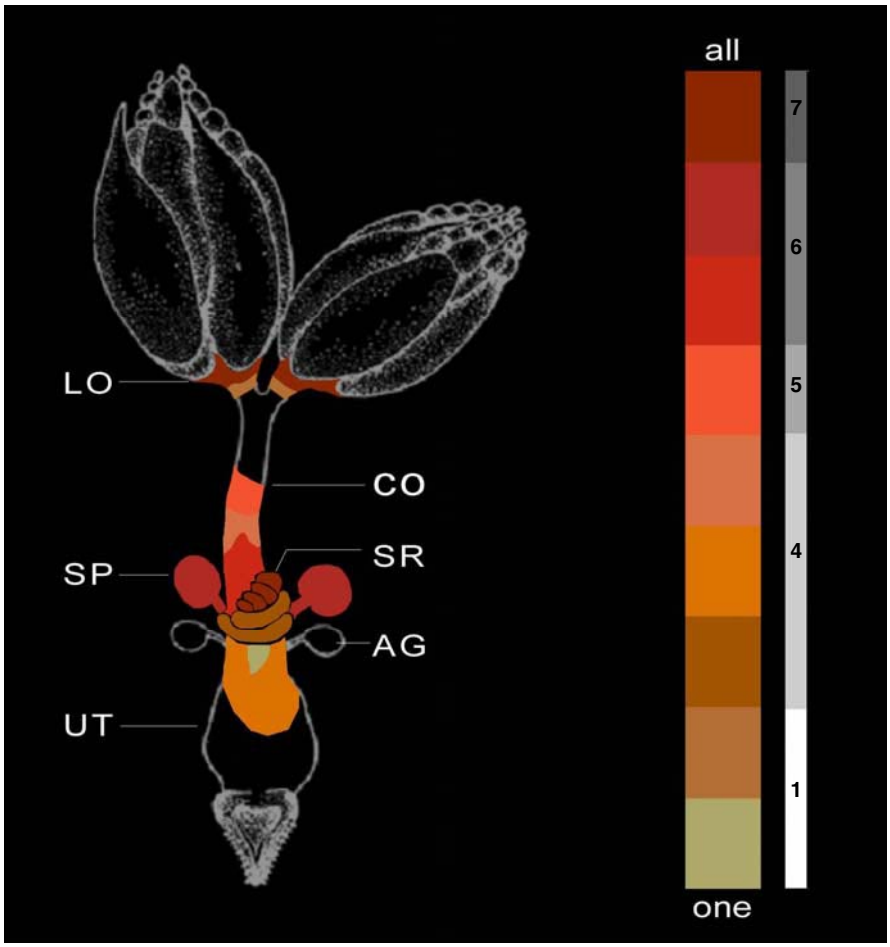


Fig. S6 cont.

Table S1. GO analysis of oviduct gene products identified as present or mating-responsive

Dataset	Molecular function	David *P*-value	FlyMine *P*-value	FuncAssociate *P-adj*
Genes present in unmated (5011)	structural constituents of the ribosome	1.70E-16	4.20E-63	<0.001
	RNA binding	5.50E-10	1.18E-36	<0.001
	protein binding	5.50E-12	3.18E-11	0.001
	Transferase activity	3.20E-08		
	actin binding	1.50E-04	2.74E-18	0.001
	Translation regulator activity	5.20E-05		0.001
	ATP binding	6.40E-07		
	oxidoreductase activity, acting on NADH or NADPH	1.30E-04		
	electron carrier activity	1.40E-04		
	transporter activity		2.65E-67	<0.001
	transcription activator activity		1.35E-14	
	regulation of cell shape			0.027
Genes that were detected only in unmated (53)	inositol or phosphatidylinositol phosphatase activity		0.007	
	serine-type endopeptidase activity*		0.08	
	sugar transmembrane transporter activity*		0.09	
Genes detected only in mated (198)	helicase activity	1.30E-03	5.90E-06	
	RNA binding	5.00E-04	0.0013	
	ATP-dependent helicase activity	7.20E-03	0.011	
Cluster of up-regulated genes (8)	antibacterial humoral response (BP)	1.80E-04	1.46E-09	<0.001
Proteins significantly expressed in oviduct (89)	structural constituent of cytoskeleton	1.70E-05		
	actin cytoskeleton organization and biogenesis	2.10E-04	3.44E-05	<0.001
	developmental process		1.41E-13	
	muscle contraction			<0.001
Up-regulated proteins (16)	structural constituent of cytoskeleton	1.50E-03		
	cytoskeletal protein binding		2.20E-06	<0.001
	actin binding	3.60E-02		<0.001
	ion binding		0.0023	
Genes expressed only in oviduct and not in other datasets (80)	negative regulation of neurogenesis	3.00E-02		
	protein tyrosine/serine/threonine phosphatase activity*	0.07	0.074	

GO terms were considered as significantly enriched after correction for multiple hypothesis testing analyzed by David, FlyMine, and FuncAssociate; David, FlyMine - multiple corrections by Bonferoni; *P-adj* is the adjusted *p*-value, corrected for multiple hypothesis testing (fraction of 1,000 null-hypothesis simulations having attributes with this single-hypothesis *P* value or smaller).
*Term was highly represented but not significant ($P > 0.05$).

Table S3A. Oviduct mating-responsive proteins

FBgn	CG#	Protein name	Molecular function	Relative abundance (M/UM)	Concordance Score
FBgn0005666	CG1479	Bent (Bt)	ATP binding	6.25	P
FBgn0003470	CG1977	α-Spectrin	Actin binding	6.00	I
FBgn0010434	CG11949	Coracle	Actin binding	5.83	P
FBgn0004873	CG9325	Adducin-like protein R1	Actin binding	5.50	N
FBgn0010100	CG9244	Mitochondrial aconitase	Aconitate hydratase activity	5.25	N
FBgn0000667	CG4376	α-Actinin	Actin binding	4.00	P
FBgn0003471	CG5870	β-Spectrin	Actin binding	4.00	P
FBgn0002968	CG1634	Neuroglian	Calcium ion binding	3.50	P
FBgn0010397	CG10119	Lamin C	Structural constituent of cytoskeleton	3.00	P
FBgn0016724	CG11064	Retinoid- and fatty-acid binding protein	Fatty acid binding	2.88	P
FBgn0014863	CG1019	Muscle LIM protein at 84B	Protein binding	2.75	P
FBgn0001402	CG33950	Terribly reduced optic lobes	DNA binding	2.63	N
FBgn0001219	CG4264	Heat shock protein cognate 4	ATP binding	2.56	P
FBgn0002527	CG7123	Laminin B1 chain	Structural molecule activity	2.50	P
FBgn0039737	CG7920	Ribosomal protein S8	4-hydroxbutyrate CoA-transferase activity	2.50	P
FBgn0003149	CG5939	Standard Paramyosin	Cytoskeletal protein binding	2.06	P
FBgn0002789	CG4696	Muscle protein 20	Actin binding	0.50	P

M, Mated female; UM, Unmated female; U, Up-regulated; D, Down-regulated; NC, No change; P, Positive (transcript was detected in unmated and mated female arrays); N, Negative (transcript was not detected in either unmated or mated female arrays); I, Intermediate (transcript was detected in mated or unmated female arrays).

Note that of the 16 up-regulated proteins, 12 had corresponding transcripts in both unmated and mated females arrays (P); for 3 proteins no transcript was observed in either unmated or mated females arrays (N); and for one protein the transcript was found in mated females arrays only (I); see also part C.

Table S3B. Proteins identified in the oviduct

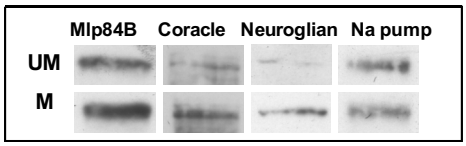
FBgn	GG#	Protein name	Number of peptides						Relative abundance	Regulation	mRNA Presence		Concordance score
			M1	M2	M3	UM1	UM2	UM3			M	UM	
FBgn0000667	CG4376	α-Actinin	6	10	8	0	9	4	4.00	U	1	1	P
FBgn0001219	CG4264	Heat shock protein cognate 4	7	5	13	2	5	8	2.56	U	1	1	P
FBgn0002527	CG7123	Laminin B1 chain	2	1	3	0	1	0	2.50	U	1	1	P
FBgn0002968	CG1634	Neuroglian	3	4	4	1	7	1	3.50	U	1	1	P
FBgn0003149	CG5939	Standard paramyosin	28	30	28	10	31	21	2.06	U	1	1	P
FBgn0003470	CG1977	α-Spectrin	7	8	5	0	8	0	6.00	U	0	1	I
FBgn0003471	CG5870	β-Spectrin	6	6	4	1	5	2	4.00	U	1	1	P
FBgn0004873	CG9325	Adducin-like protein R1	5	2	6	0	5	1	5.50	U	0	0	N
FBgn0005666	CG1479	Bent	6	6	13	0	7	2	6.25	U	1	1	P
FBgn0010100	CG9244	Mitochondrial aconitase	6	4	9	1	5	2	5.25	U	0	0	N
FBgn0010397	CG10119	Lamin C	4	3	4	1	4	2	3.00	U	1	1	P
FBgn0010434	CG11949	Coracle	7	4	14	0	10	3	5.83	U	1	1	P
FBgn0014863	CG1019	Muscle LIM protein	3	2	5	1	2	2	2.75	U	1	1	P
FBgn0016724	CG11064	Retinoid- and fatty-acid binding protein	4	0	7	0	5	4	2.88	U	1	1	P
FBgn0039737	CG7920	Ribosomal protein S8	3	4	6	1	6	3	2.50	U	1	1	P
FBgn0001402	CG33950	Terribly reduced optic lobes	4	2	10	0	6	8	2.63	U	0	0	N
FBgn0000046	CG18290	Actin 87E	14	14	15	14	12	11	1.27	NC	0	0	N
FBgn0000055	CG32954	ADH	0	0	4	0	0	2	1.00	NC	1	1	P
FBgn0000064	CG6058	Aldolase gamma	0	2	2	0	1	2	1.30	NC	1	1	P
FBgn0000261	CG6871	Catalase	0	1	6	0	2	1	0.75	NC	1	1	P
FBgn0000556	CG8280	Elongation factor 1α48D	5	6	8	2	6	6	1.91	NC	1	1	P
FBgn0000559	CG2238	Elongation factor 2b	0	0	2	0	2	1	1.00	NC	1	1	P
FBgn0000579	CG17654	Enolase	0	2	3	0	0	2	1.25	NC	1	1	P
FBgn0001092	CG8893	Glyceraldehyde 3 phosphate dehydrogenase 2	5	3	8	5	4	6	0.87	NC	1	1	P
FBgn0001098	CG5320	Glutamate dehydrogenase	0	0	2	0	0	3	0.89	NC	1	1	P
FBgn0002526	CG10236	Laminin A	1	2	7	0	2	1	1.00	NC	1	1	P
FBgn0002528	CG3322	Laminin B2	6	5	3	0	6	4	0.79	NC	1	1	P
FBgn0002607	CG2746	RpL19	1	2	1	0	2	2	0.84	NC	1	1	P
FBgn0002741	CG17927	Myosin heavy chain	92	84	104	56	91	70	1.56	NC	1	1	P
FBgn0002772	CG5596	Myosin alkali light chain 1	4	4	3	3	4	4	1.03	NC	1	1	P
FBgn0002773	CG2184	Myosin light chain 2	8	10	6	9	11	2	0.90	NC	1	1	P
FBgn0002921	CG5670	Na pump α subunit	13	17	19	6	15	12	1.62	NC	1	1	P
FBgn0003274	CG4918	R Ribosomal protein	0	0	4	0	0	2	1.00	NC	1	1	P
FBgn0003279	CG5502	Ribosomal protein L4	0	2	2	0	1	2	1.00	NC	1	1	P
FBgn0003360	CG16944	ADP/ATP translocase	0	2	4	2	1	6	0.58	NC	1	1	P
FBgn0003462	CG11793	Slow superoxide dismutase	0	0	2	0	0	3	0.89	NC	1	1	P
FBgn0003721	CG4898	Tropomyosin 1	15	14	20	28	17	18	0.67	NC	0	0	N
FBgn0003884	CG1913	α-Tubulin84B	7	7	5	2	9	8	0.70	NC	1	1	P
FBgn0003887	CG9277	Tubulin at 56D	4	6	13	3	7	8	1.09	NC	1	1	P
FBgn0003887	CG9277	Succinyl coenzyme A synthetase flavoprotein subunit	1	2	2	0	3	2	0.89	NC	1	1	P
FBgn0003942	CG5271	Ribosomal protein S27A	0	0	3	0	0	2	1.17	NC	1	1	P
FBgn0004028	CG7178	Troponin-I wings-up A	7	7	8	4	7	7	1.07	NC	1	1	P
FBgn0004047	CG11129	Yolk protein 3	1	1	2	0	2	2	0.84	NC	1	1	P
FBgn0004117	CG4843	Tropomyosin 2	15	14	20	20	19	18	0.75	NC	1	1	P
FBgn0004169	CG7107	Troponin T-1	13	11	6	15	16	5	0.68	NC	1	1	P
FBgn0004363	CG6647	Porin	4	6	1	4	5	4	1.10	NC	1	1	P
FBgn0004432	CG9916	Cyclophilin 1	0	1	3	2	1	2	1.25	NC	1	1	P
FBgn0004551	CG3725	Calcium ATPase at 60A	2	3	2	0	3	1	1.50	NC	1	1	P
FBgn0004922	CG10944	Ribosomal protein S6	0	1	3	0	2	2	1.00	NC	1	1	P
FBgn0005391	CG2979	Yolk protein 2	2	1	12	2	1	9	1.00	NC	1	1	P
FBgn0005671	CG17369	Vacuolar H + ATP synthase subunit B	0	2	3	0	2	3	1.00	NC	1	1	P
FBgn0010217	CG11154	ATP synthase β	7	13	3	9	7	7	0.61	NC	1	1	P
FBgn0010410	CG15442	Ribosomal protein L27A	1	2	2	1	1	2	1.30	NC	1	1	P
FBgn0010424	CG7930	Troponin C 73F	3	4	3	4	3	3	1.03	NC	1	1	P
FBgn0010620	CG10939	SRY interacting protein 1	3	4	3	2	3	0	1.42	NC	0	0	N
FBgn0011211	CG3612	Mitochondrial ATP synthase α subunit precursor	6	9	8	6	4	8	1.00	NC	1	1	P
FBgn0011272	CG4651	Ribosomal protein L13	1	0	3	0	1	2	1.16	NC	1	1	P

FBgn	GG#	Protein name	Number of peptides						Relative abundance	Regulation	mRNA Presence		Concordance score
			M1	M2	M3	UM1	UM2	UM3			M	UM	
FBgn0011643	CG3220	Muscle LIM protein 1	2	1	1	1	1	2	1.16	NC	1	0	I
FBgn0011661	CG10701	Moesin	4	3	2	2	4	2	0.88	NC	1	1	P
FBgn0012036	CG3752	Aldehyde dehydrogenase	0	0	2	0	0	2	1.00	NC	0	0	N
FBgn0013770	CG6692	Cysteine proteinase-1	0	0	3	1	0	3	1.00	NC	1	1	P
FBgn0014002	CG6988	Protein disulfide isomerase	1	1	5	1	1	4	1.08	NC	1	1	P
FBgn0017545	CG2168	Ribosomal protein S3A	1	0	2	0	1	2	1.00	NC	1	1	P
FBgn0017579	CG6253	Ribosomal protein L14	1	0	6	0	0	3	1.00	NC	1	1	P
FBgn0020367	CG3762	Vha68–2	2	6	4	1	5	7	1.60	NC	0	0	N
FBgn0020618	CG7111	Receptor of activated protein kinase C 1	1	0	4	0	1	2	1.00	NC	1	1	P
FBgn0020910	CG4863	Ribosomal protein L3	1	0	4	0	0	3	1.11	NC	1	1	P
FBgn0026403	CG12908	Nidogen/entactin	0	1	3	0	2	2	1.00	NC	0	0	N
FBgn0027527	CG1151	Osiris 6	2	2	1	1	2	1	1.40	NC	0	0	N
FBgn0028479	CG4389	CG4389	0	0	2	0	1	2	1.00	NC	1	1	P
FBgn0029176	CG11901	Elongation factor 1, gamma chain	0	0	4	0	1	2	1.00	NC	0	0	N
FBgn0032422	CG6579	CG6579	0	0	2	0	0	2	1.00	NC	1	1	P
FBgn0032518	CG9282	Ribosomal protein L24	1	0	2	0	1	2	1.00	NC	1	1	P
FBgn0032897	CG9336	CG9336	0	2	4	0	1	2	1.50	NC	1	1	P
FBgn0032899	CG9338	CG9338	1	0	4	4	1	3	1.17	NC	1	1	P
FBgn0033919	CG8547	CG8547	2	2	0	0	1	0	1.50	NC	1	1	P
FBgn0034603	CG9480	Glycogenin	3	5	2	2	6	2	1.11	NC	0	0	N
FBgn0035753	CG8615	Ribosomal protein L18	0	0	2	0	0	2	1.00	NC	1	1	P
FBgn0035917	CG6416	CG6416	2	2	3	4	1	2	1.75	NC	1	1	P
FBgn0036279	CG43567	Sodium chloride cotransporter 69	6	3	2	1	3	1	1.50	NC	1	1	P
FBgn0036682	CG11661	RE42354p	2	0	2	0	1	3	0.83	NC	1	1	P
FBgn0037245	CG14648	CG14648	1	2	1	0	2	0	1.00	NC	1	1	P
FBgn0038587	CG7998	CG7998	0	0	5	0	0	3	1.00	NC	1	1	P
FBgn0039713	CG7808	Ribosomal protein S8	1	2	5	0	3	3	0.83	NC	1	1	P
FBgn0040813	CG11051	Neuropeptide-like precursor 2	0	0	2	0	0	2	1.00	NC	1	1	P
FBgn0053859	CG33859	Histone 2A	0	1	3	0	0	2	1.17	NC	0	0	N
FBgn0053861	CG33861	Histone 1	1	1	2	0	2	0	0.75	NC	0	0	N
FBgn0053868	CG33868	Histone H2B	1	1	3	0	0	3	1.00	NC	0	0	N
FBgn0002789	CG4696	Muscle protein 20	0	0	2	2	2	1	0.50	D	1	1	P

M, Mated female; UM, Unmated female; U, Up-regulated; D, Down-regulated; NC, No change; P, Positive (transcript was detected in unmated and mated female arrays); N, Negative (transcript was not detected in either unmated or mated female arrays); I, Intermediate (transcript was detected in mated or unmated female arrays).

Table S3C. Validation

Representative Western blot images of oviduct mating-responsive proteins



Relative abundance of oviduct mating-responsive proteins in mated vs. unmated oviducts

Protein name	Relative abundance (MudPIT)	Relative abundance (Western blot)	Regulation	No. of identified peptides	Coverage, %
Coracle	5.8	1.9	U	25	22
Adducin	5.5	2	U	13	28
Neuroglian	3.5	3.2	U	11	14
Muscle LIM protein at 84B	2.75	2.79	U	6	24
Na pump α subunit	1.62	1	NC	32	38

M, Mated female; UM, Unmated female; U, Up-regulated; D, Down-regulated; NC, No change; P, Positive (transcript was detected in unmated and mated female arrays); N, Negative (transcript was not detected in either unmated or mated female arrays); I, Intermediate (transcript was detected in mated or unmated female arrays).

Note that of the 16 up-regulated proteins, 12 had corresponding transcripts in both unmated and mated females arrays (P); for 3 proteins no transcript was observed in either unmated or mated females arrays (N); and for one protein the transcript was found in mated females arrays only (I); see also part C.

Table S5. Characteristic path length (CPL) of protein–protein interaction (PPI) subnetworks for oviduct-expressed transcripts or proteins compared with random sets

Network and subnetwork	Oviduct gene set	Size of set	Observed CPL	Mean CPL of random sets (1,000)	SD of random sets (1,000)	*z* score	Mean CPL of random sets (1,000; same bait/prey ratio)	SD of random sets (1,000; same bait/prey ratio)	*z* score
Dm PPI map									
Full network		5394	6.41	NA	NA	NA	NA	NA	NA
Present	Proteins	56	5.16	6.41	0.35	**−3.55**	6.15	0.34	**−2.93**
Up-regulated (Up)	(detected by MudPIT)	12	3.91	6.36	0.79	**−3.11**	6.01	0.71	**−2.97**
Present NNN		223	4.71	5.36	0.42	−1.53	5.25	0.40	−1.35
Up NNN		61	3.78	5.11	0.87	−1.54	4.92	0.74	−1.54
Differentially expr. (DE)	Transcripts	84	6.70	6.39	0.29	1.06	6.45	0.28	0.89
Up	(detected by microarray)	68	6.84	6.41	0.32	1.36	6.46	0.32	1.20
Down-regulated (Down)		16	6.34	6.43	0.65	−0.14	6.41	0.61	−0.13
DE NNN		234	5.74	5.41	0.35	0.94	5.47	0.36	0.77
Up NNN		190	5.67	5.38	0.40	0.72	5.43	0.41	0.58
Down NNN		64	4.83	5.16	0.77	−0.42	5.24	0.73	−0.56
Dm PPI map plus interologs (inParanoid orthologs)									
Full Network		6019	5.91	NA	NA	NA	NA	NA	NA
Present	Proteins	65	5.12	6.41	0.35	**−3.65**	5.65	0.24	**−2.22**
Up	(detected by MudPIT)	12	3.65	6.36	0.79	**−3.43**	5.63	0.54	**−3.64**
Present NNN		333	4.62	5.01	0.28	−1.40	4.90	0.25	−1.08
Up NNN		94	3.57	4.76	0.61	−1.93	4.61	0.56	−1.84
DE	Transcripts	97	5.92	6.41	0.27	−1.76	5.90	0.20	0.14
Up	(detected by microarray)	79	5.83	6.41	0.29	−2.00	5.85	0.22	−0.10
Down		18	6.41	6.42	0.67	−0.02	6.11	0.46	0.65
DE NNN		324	5.27	5.06	0.23	0.91	5.04	0.22	1.04
Up NNN		273	5.14	5.04	0.26	0.35	4.97	0.24	0.69
Down NNN		71	4.88	4.86	0.58	0.04	4.91	0.52	−0.06
Dm PPI map without ribosome and vacuolar H + ATPase									
Full network		5320	6.25	NA	NA	NA	NA	NA	NA
Present	Proteins	56	5.17	6.40	0.34	**−3.65**	6.12	0.33	**−2.92**
Up	(detected by MudPIT)	12	3.91	6.38	0.77	**−3.22**	5.95	0.67	**−3.05**
Present NNN		218	4.75	5.36	0.44	−1.37	5.22	0.40	−1.19
Up NNN		59	3.79	5.04	0.84	−1.49	4.90	0.80	−1.38
DE	Transcripts	82	6.77	6.38	0.29	1.30	6.42	0.29	1.21
Up	(detected by microarray)	66	6.93	6.41	0.34	1.57	6.41	0.32	1.65
Down		16	6.34	6.41	0.69	−0.10	6.38	0.57	−0.07
DE NNN		226	5.78	5.40	0.38	0.98	5.43	0.35	1.01
Up NNN		182	5.71	5.37	0.40	0.83	5.39	0.41	0.77
Down NNN		64	4.83	5.13	0.75	−0.39	5.17	0.69	−0.49
Dm PPI interolog map without ribosome and vacuolar H + ATPase									
Full network		5929	5.93	NA	NA	NA	NA	NA	NA
Present	Proteins	64	5.09	6.41	0.34	**−3.83**	5.66	0.24	**−2.39**
Up	(detected by MudPIT)	12	3.65	6.36	0.79	**−3.43**	5.63	0.55	**−3.58**
Present NNN		326	4.63	5.03	0.29	−1.39	4.89	0.25	−1.07
Up NNN		92	3.57	4.77	0.64	−1.89	4.60	0.54	−1.91
DE	Transcripts	94	5.90	6.42	0.27	−1.93	5.89	0.20	0.07
Up	(detected by microarray)	76	5.80	6.41	0.31	−1.96	5.84	0.23	−0.18
Down		18	6.41	6.42	0.67	−0.02	6.12	0.45	0.63
DE NNN		314	5.27	5.08	0.24	0.79	5.04	0.21	1.06
Up NNN		263	5.13	5.05	0.27	0.28	4.97	0.24	0.65
Down NNN		71	4.88	4.85	0.53	0.07	4.91	0.54	−0.05

Table S5B. Average mutual clustering coefficient (Cvw) of PPI subnetworks for oviduct-expressed transcripts or proteins compared with random sets

Network and subnetwork	Oviduct gene set	Size of observed set	Mean random set size	SD of size of random sets (1,000)	Avg. Cvw	Mean Cvw of random sets (1,000)	SD Cvw of random sets (1000)	*z* score
Dm PPI map								
Full network		5394	NA	NA	0.00	NA	NA	NA
Present	Proteins	56	NA	NA	0.16	0.03	0.02	**8.25**
Up	(detected by MudPIT)	12	NA	NA	0.69	0.03	0.05	**12.23**
Present NNN		223	203.01	33.51	0.41	0.20	0.17	1.20
Up NNN		61	45.71	18.63	0.32	0.45	0.48	−0.28
DE	Transcripts	84	NA	NA	0.03	0.02	0.01	0.51
Up	(detected by microarray)	68	NA	NA	0.02	0.03	0.01	−0.43
Down		16	NA	NA	0.00	0.03	0.04	−0.69
DE NNN		234	299.83	39.60	0.06	0.16	0.11	−0.82
Up NNN		190	244.62	37.08	0.06	0.18	0.15	−0.88
Down NNN		64	60.34	19.43	0.27	0.39	0.43	−0.28
Dm PPI map plus interologs (inParanoid orthologs)								
Full network		6019	NA	NA	0.00	NA	NA	NA
Present	Proteins	65	NA	NA	0.12	0.02	0.01	**7.20**
Up	(detected by MudPIT)	12	NA	NA	0.41	0.03	0.05	**7.82**
Present NNN		333	266.97	36.79	0.30	0.18	0.10	1.26
Up NNN		94	50.83	17.45	0.44	0.52	0.45	−0.17
DE	Transcripts	97	NA	NA	0.05	0.03	0.01	2.54
Up	(detected by microarray)	79	NA	NA	0.06	0.02	0.01	2.83
Down		18	NA	NA	0.00	0.02	0.03	−0.73
DE NNN		324	388.04	43.43	0.11	0.14	0.07	−0.41
Up NNN		273	319.53	41.53	0.13	0.16	0.09	−0.32
Down NNN		71	77.51	23.62	0.34	0.43	0.34	−0.26
Dm PPI map without ribosome and vacuolar H + ATPase								
Full network		5320	NA	NA	0.00	NA	NA	NA
Present	Proteins	56	NA	NA	0.16	0.03	0.02	**8.98**
Up	(detected by MudPIT)	12	NA	NA	0.70	0.03	0.05	**12.96**
Present NNN		218	201.47	33.72	0.42	0.19	0.17	1.36
Up NNN		59	46.01	18.27	0.31	0.44	0.49	−0.27
DE	Transcripts	82	NA	NA	0.03	0.03	0.01	0.55
Up	(detected by microarray)	66	NA	NA	0.02	0.03	0.01	−0.21
Down		16	NA	NA	0.00	0.02	0.04	−0.65
DE NNN		226	290.35	40.86	0.07	0.16	0.12	−0.77
Up NNN		182	236.54	36.04	0.06	0.17	0.14	−0.83
Down NNN		64	60.46	19.05	0.27	0.39	0.43	−0.29
Dm PPI-interolog map without ribosome and vacuolar H + ATPase								
Full network		5929	NA	NA	0.00	NA	NA	NA
Present	Proteins	64	NA	NA	0.13	0.02	0.01	**8.18**
Up	(detected by MudPIT)	12	NA	NA	0.42	0.03	0.06	**7.04**
Present NNN		326	260.17	37.33	0.31	0.19	0.12	1.10
Up NNN		92	50.18	17.45	0.45	0.51	0.44	−0.13
DE	Transcripts	94	NA	NA	0.05	0.02	0.01	2.38
Up	(detected by microarray)	76	NA	NA	0.05	0.02	0.01	2.45
Down		18	NA	NA	0.00	0.03	0.04	−0.70
DE NNN		314	373.89	44.78	0.11	0.14	0.07	−0.43
Up NNN		263	306.03	40.81	0.13	0.16	0.09	−0.33
Down NNN		71	75.66	20.51	0.34	0.42	0.36	−0.25

Shown are the topological statistics for PPI subnetworks containing gene products expressed in the oviduct. Significant *z* scores are in bold type.

Other Supporting Information Files

Table S2 (XLS)
Table S4 (XLS)