Research Article

# Temporal and sequential order of nonoverlapping gene networks unraveled in mated female Drosophila

Claude Pasquier[1] , Alain Robichon[2]

In this study, we reanalyzed available datasets of gene expression changes in female Drosophila head induced by mating. Mated females present metabolic phenotypic changes and display behavioral characteristics that are not observed in virgin females, such as repulsion to male sexual aggressiveness, fidelity to food spots selected for oviposition, and restriction to the colonization of new niches. We characterize gene networks that play a role in female brain plasticity after mating using AMINE, a novel algorithm to find dysregulated modules of interacting genes. The uncovered networks of altered genes revealed a strong specificity for each successive period of life span after mating in the female head, with little conservation between them. This finding highlights a temporal order of recruitment of waves of interconnected genes which are apparently transiently modified: the first wave disappears before the emergence of the second wave in a reversible manner and ends with few consolidated gene expression changes at day 20. This analysis might document an extended field of a programmatic control of female phenotypic traits by male seminal fluid.

## Introduction

Phenotypic plasticity is associated with changes in gene expression and can be adaptive to fluctuating environmental conditions or nonadaptive independent of the context of natural selection pressure. Many acquired phenotypic traits appear irreversible in individuals of many insect species (DeWitt et al, 1998; Debat & David, 2001; Braendle & Felix, 2008, 2009; Dingemanse et al, 2010; Beldade et al, 2011). However, in insects like *Drosophila*, phenotypic plasticity operating in female after mating can produce the apparition of behavioral and physiological traits that can be reversible in relation to tissue specific temporal changes of transcriptome (Fedorka et al, 2007; Dalton et al, 2010; Ameku & Niwa, 2016; Delbare et al, 2017). Briefly, the female post-mating response includes increased egg laying and feeding (Carvalho et al, 2006), a preference for amino acids and for salt (Walker et al, 2015), decreased receptivity to

mating (Chapman et al, 2003), decreased daytime sleep (Isaac et al, 2010; Garbe et al, 2016; Dove et al, 2017), and a diminished immune response (Short & Lazzaro, 2010, 2013). Moreover, in *Drosophila* species, mated females show an increase in activity during the light phase compared with virgin females and the same low level activity in the dark (Isaac et al, 2010). An increased consumption of yeast in the post-mating period was observed that corresponds to intense oogenesis (Drummond-Barbosa & Spradling, 2001). Another report describes that some genes of the immune system was highly expressed after mating, in contrast with a severely diminished efficiency of the fly to fight infectious pathogens (Fedorka et al, 2007). These conclusions have been completed by another observation that presents a panel of modified genes in mated female, most of them involved in immune system (Lawniczak & Begun, 2004). All these elements argue in favor of complex gene networks under the control of seminal fluid that markedly change the female destiny.

At physiological level, mating was reported to modify the gustatory signaling leading to an increase in appetite and uptake of salt and yeast by activation of Sex Peptide Receptor located in neuronal terminals in female reproductive tract (Walker et al, 2015). In this regard, few male seminal peptides that regulate general female metabolism, behavior, and life span have been identified to act in post mating period (Peng et al, 2005; Isaac et al, 2010). The strength of female gametogenesis is regulated by environmental factors, temperature, food abundance, and also mating. Stem cell division in ovary was found partly under control of male-derived Sex Peptide (SP) acting through a neuroendocrine pathway, neuromediators, and signaling molecules (Ameku & Niwa, 2016).

Authors have equally reported time-dependent changes at transcriptional level in mated female (Mack et al, 2006; Dalton et al, 2010). In a related study, ion channel transcripts have been found significantly down-regulated, whereas most of up-regulated genes reside in the head fat body (Dalton et al, 2010). Furthermore, mating that changes the physiology and behavior of *Drosophila* female was shown to affect the secretory granule release at some pre synaptic nerve terminal spaces in reproductive female tract (Heifetz & Wolfner, 2004). On the other side, an article reports that not only some mRNAs but microRNAs and histone chemical modifications are markers associated with mating leading to the idea that an epigenetic component orchestrate the genes networks recruitment

---

[1]Université Côte d'Azur, CNRS, I3S, Nice, France    [2]Université Côte d'Azur, INRAE, CNRS, ISA, Nice, France

Correspondence: claude.pasquier@univ-cotedazur.fr

(Zhou et al, 2014). Moreover, strong differences in female post-mating transcriptome has been reported depending on female/male genetic interactions, which suggests a solid contribution of genetic background to modulate the post mating female physiology (Delbare et al, 2017). The gene expression changes in female head and central nervous system tissues that follow mating and seminal acquisition has been documented in some extent. Behavior changes induced as post mating response was shown to be dependent on neuronal circuitry requiring expressing of doublesex (dsx) gene (Rezával et al, 2012). Authors have shown that the changes in gene expression induced by mating are influenced by cross tissue interactions and the global effects impact sleep, food preference, and re-mating (Newell et al, 2020). Robust changes in mRNA patterns in tissues between virgin and mated female have indicated integrative and coordinated functional gene networks acting in a concerted manner upon mating (McDonough-Goldstein et al, 2021). However, our knowledge on the integrative and coordinated process involving numerous transcriptional changes triggered by mating in female *Drosophila* have not been fully completed. At present, we know little about the molecular mechanisms that affect female brain gene expression after mating.

In this study, we use a new approach to analyze the transcriptomic changes occurring in the female head after mating. This approach relies on (1) a novel method, AMINE (Pasquier et al, 2021 *Preprint*), for the identification of dysregulated gene modules, and (2) accurate RNA-Seq data from the head of both virgin and mated *Drosophila* females at three time periods after mating, made available by the modENCODE project. AMINE applies a greedy algorithm on a highly informative and compact representation combining general knowledge about gene interactions and measured gene variation. This strategy allows us to assess both the extent of gene variation and the topology of gene interactions, overcoming limitations of classical enrichment analysis of the most differentially expressed genes and earlier network-based approaches.

Our current analysis, along with many other reports, argues in favor of a new paradigm that underlies a coordinated process of female transcriptome modification after mating, which parallels the behavioral and metabolic changes to produce offspring. The broad transcriptional changes initiated by copulation can be followed by the differential landscape of gene modules that vary between the three post-mating periods (1, 4, and 20 d after mating). Assuming that the long distance between the head and spermathecae excludes sperm RNA contamination during the experimental process of dissection, the analysis uncovered waves of modified gene networks that show marked differences between the time periods of 1, 4, and 20 d. This hints that mating triggers temporal and ordered activation of gene networks, that is, the first wave inducing the interdependent second wave and so on up to the end of the process.

# Results

## Differential expression analysis

We considered only the RNA-Seq lists of head tissues of virgin versus mated females aged 1, 4, and 20 d, which eliminates any possible fertilized egg interference/contamination that occurs with the full carcass or other tissues, such as the gut. Under these conditions, the number of differentially expressed genes, which was at a maximal value 1 d post-mating (334 down-regulated and 291 up-regulated genes–Fig 1A and Table S1), decreased after 4 d (32 down-regulated and 50 up-regulated genes–Fig 1B and Table S2) until a very low number was observed after 20 d (one down-regulated and five up-regulated genes–Fig 1C and Table S3). The lists obtained for the three different time points mainly do not overlap. A single gene (FBgn0283437, prophenoloxidase 1 involved in the melanization reaction, notably in response to wounding and produced by crystal cells, a type of hemocyte in *Drosophila*) that is overexpressed at day 1 is also overexpressed at day 20 (Fig 2C). This gene is also known to enhance the mating process and reproductive success (Parkash et al, 2011). Only 10 genes were differentially expressed on both days 1 and 4 in the same orientation (Fig 2A–D). Of those, three genes were overexpressed in mated females (FBgn0030608/Lsd-2, FBgn0259937/Nop60B, and FBgn0032538/Vajk2), six were underexpressed (FBgn0260653/serp, FBgn0031089/aspr, FBgn0016120/ATPsynD, FBgn0033485/RpLP0-like, FBgn0033661/CG13185, and FBgn0011824/CG4038) in mated females, and one gene was overexpressed at day 1 but underexpressed at day 4 (FBgn0039670/CG7567). Little is known about the functions of this later gene, but the other nine genes are mainly implicated in rRNA processing and ribosome biogenesis. Fig 2 shows the evolution over the three time points of the $\log_2$-fold change ($\log_2$FC) values of the genes that were identified as differentially expressed at least at one time point.

## Enrichment analysis of differentially expressed genes

Enrichment analysis is a classical process to obtain a synthetic view of a gene list. It consists in associating to this list the most typical annotations originating from various databases. We therefore performed enrichment analyses for the differentially expressed genes identified at each time point. We did this for all differentially expressed genes, for underexpressed genes, and then for overexpressed genes (results are listed in Table S4).

Annotations characterizing the whole list of differentially expressed genes at day 1 are associated with relatively modest false discovery rates (FDRs) on the order of $10 \times 10^{-4}$. Among these annotations, we find several terms related to development (in particular cuticle development with a FDR of $1 \times 10^{-4}$ and chitin-based cuticle development with a FDR of $1.5 \times 10^{-4}$) and the chemical reactions involving carbohydrates (carbohydrate metabolic process with a FDR of $8.7 \times 10^{-4}$). By considering only overexpressed genes, we find annotations associated with much more significant FDR. Among them, many annotations are related to metabolic regulation and cell cycle. Among the annotations associated with underexpressed genes, we find again terms related to cuticle development and carbohydrate metabolism. Regarding day 4, we find new annotations that did not appear significantly in day 1 as rRNA processing (FDR = $2.3 \times 10^{-7}$) and ribosome biogenesis (FDR = $3.19 \times 10^{-7}$). We find these same annotations, with even more significant FDR, for the underexpressed genes, whrereas no annotation appears significantly for the overexpressed genes. Concerning the six genes that are differentially expressed at day 20, no term from Gene Ontology stands out significantly. We only obtain an overrepresentation of the term Oxidoreductase from UniProt Keywords.
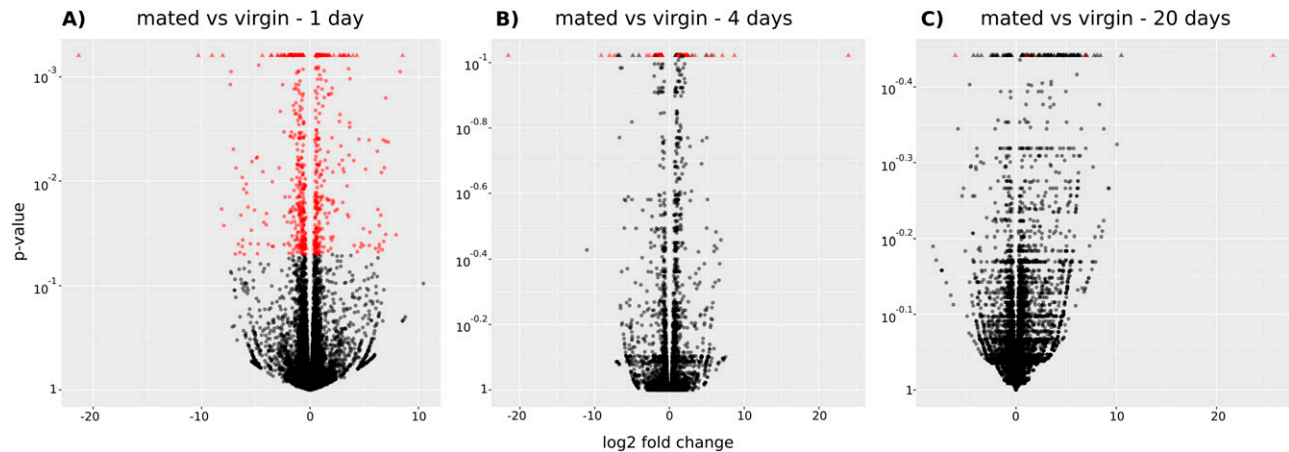
**Figure 1. Volcano plots of differentially expressed genes.**
Volcano plots showing the significance of the differential expression of genes between mated and virgin flies at day 1, 4, and 20. $Log_2$ FC is plotted on the x-axis, and adjusted $P$-values are plotted on the y-axis. Genes with a variation of at least a factor of 1.25 (i.e., a $log_2$-fold change greater than 0.32 or less than −0.32) and an adjusted $P$-value of 0.05 or less are considered differentially expressed and are represented in red. Triangles correspond to genes with a too-high adjusted $P$-value to be displayed on the plot or a $P$-value equal to NA. Genes appearing in the three plots are listed in Tables S1–S3. **(A)** Mated versus virgin females at day 1. The analysis highlighted 625 differentially expressed genes identified with high confidence (most of these genes are associated with an adjusted $P$-value well below 0.05). **(B)** Mated versus virgin females at day 4. Eighty-two genes are considered differentially expressed. Important variations were found between replicates because most of the adjusted $P$-values were higher than 0.1, as seen in the panel. **(C)** Mated versus virgin females at day 20. At this time period, the large variations between the quantifications performed with the different replicates only allowed the identification of six differentially expressed genes.

This transcriptome analysis highlights the observation that mating modifies the female biology and that these modifications affect temporal and sequential gene networks such that these show variations among days 1, 4, and 20. The number of genes with changes in the transcription level decreased markedly throughout the life span of female *Drosophila* after mating. Moreover, the analysis of the transcriptome at the three time periods revealed the specificity of the modified genes for each period, as revealed by the little overlap or conservation of these changes throughout the female fly life span after mating. This finding argues in favor of singular waves of gene networks acting at each time period after mating with little, if any, interference and acting as a cascade in which the first wave orchestrates the outcome of the second and so on along the fly life span.

## Limitations of conventional RNA-Seq data analysis pipeline

Differential gene expression analysis allows identification of the genes that vary the most. In clinical and pharmaceutical research, these genes can be valuable in identifying candidate biomarkers, therapeutic targets, and genetic signatures for diagnosis. However, the identification of specific changes in the expression of a few genes is usually not sufficient to reveal ongoing biological activities. The classical approach consisting in performing an enrichment analysis of the most differentially expressed genes does indeed hardly provide an accurate picture of the processes at stake. This is mainly due to the fact that in this approach, genes are considered as isolated entities. However, interactions between genes, is fundamentally important to understand the genetic pathways involved and the dynamics of complex genetic systems. As pinpointed by Rapaport et al (2007), "a small but coherent difference in the expression of all the genes in a pathway should be more significant

than a larger difference occurring in unrelated genes." Arising from this observation, many methods have been proposed for the analysis of gene activity based on our knowledge about their molecular interactions (Nguyen et al, 2019). The idea underlying active module identification is to identify pertinent modules of genes by simultaneously considering two criteria: one is based on a measurement of gene activity (differential expression) and the other reflects their molecular interactions. In their 2019 survey, Nguyen et al (2019) identified 22 computational tools for active module identification that they divided into six categories: greedy algorithms (six methods), evolutionary algorithms (five methods), diffusion-flow emulation models (five methods), random walk algorithms (two methods), maximum clique identification (two methods), and clustering-based methods (two methods). To this list, we add WMAXC (Amgalan & Lee, 2014), a method belonging to the category of maximum clique identification, GiGA (Breitling et al, 2004), which is based on a greedy algorithm and MRF (Robinson et al, 2017), which uses a diffusion-flow emulation model. This makes 25 methods, whose basic principles and results differ, but which also have differences from the point of view of usability; some are easy to use by a non-expert, others are more demanding, involving command line execution, complex setup, the need to properly format input data in a particular layout or requiring many steps to obtain the final result. Usability is a very important aspect for programs intended to be used by people who are not necessarily experts in computer science or data analysis. Difficulties may arise during installation, data preparation, or program execution.

### *Installation*
A local installation requires downloading a package and installing it; it may require installing other required libraries or entering some command lines. Applications in the form of a web service allow
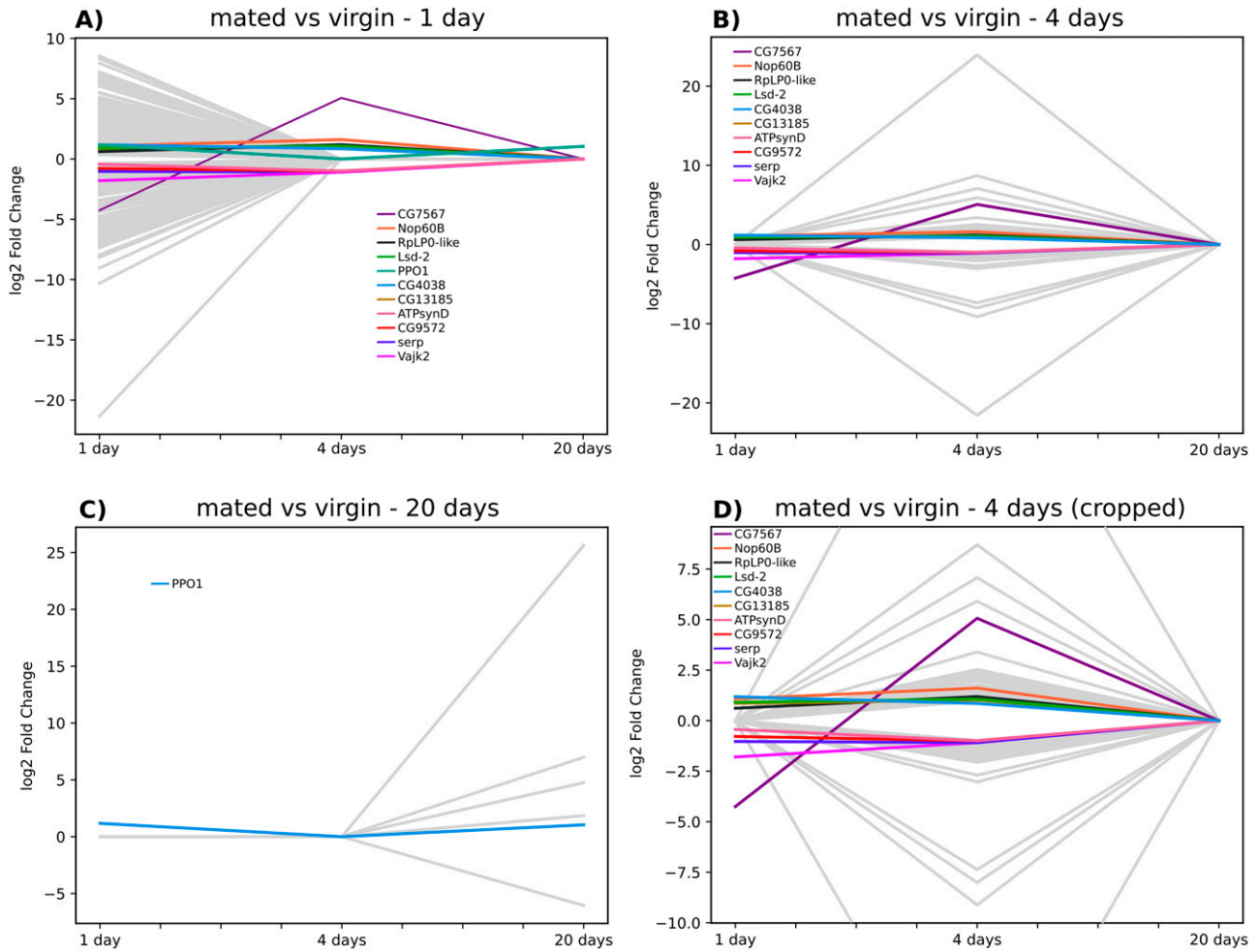
**Figure 2.  Line graph showing the evolution of log₂ FC values over the three time points (days 1, 4, and 20) for genes that were identified as differentially expressed.**
**(A)** Evolution of log₂ FC values for all genes identified as differentially expressed at day 1. The 11 colored lines correspond to the genes that were also differentially expressed at day 4 or 20. **(B)** Evolution of log₂ FC values for genes identified as differentially expressed at day 4. The colored lines represent the genes that were also differentially expressed at any of the other time points. **(C)** Evolution of log₂ FC values for genes identified as differentially expressed at day 20. The blue line illustrates the log₂ FC value of PPO1, which was the only gene that was also overexpressed at day 1. **(C, D)** Cropped version of the panel shown in (C), which allows better differentiation among the different colored lines.

access to a program simply by accessing a web address on a browser. A special case is represented by applications that come in the form of a plugin for Cytoscape (Shannon et al, 2003), a popular platform dedicated to network visualization and analysis. Cytoscape plugins require the use of Cytoscape but have the advantage of being easily installed through Cytoscape app store.

### Data preparation

A module identification program requires two essential data: an interaction network and the quantification of the differential expression of genes which can be specified either by the level of expression in each condition, by a fold change (or a log₂ fold change) or by a *P*-value. The interaction network is provided either by the application or by the user. If it is included in the application, it must be possible to link the identifiers of the genes or gene products whose expression has been measured to the identifiers of the nodes of the network. In the case where the interaction network

is to be provided by the user, it usually takes the form of a list of edges, that is, a list in which each line contains a connection between the identifier of a source node and the identifier of a target node, possibly associated with a numerical value that may represent the strength of the interaction or its plausibility. All other requested parameters are detrimental to usability, whether they are internal program parameters or information about the desired results (e.g., number of modules and minimum and maximum module size).

### Program execution

The most pleasant way to execute a program is to be in front of a graphical interface that allows us to enter the various input data. However, an execution in the form of commands to be entered on a command line interface, although less user-friendly, seems quite accessible. We consider programs that require numerous preprocessing steps, for example, for data conversion or parameter

evaluation, to be useable only by computer-literate people and therefore not suitable for an end user.

Tables 1 and S5 show a summary of the methods emphasizing their availability and ease of use. Our study reveals that, out of the 25 methods, three are no longer accessible (broken link), one has problems in the installation procedure (installation failure), one causes a runtime error, two seem to run but do not return any results, two of the methods identified by Nguyen et al (2019) are not suitable for processing transcriptomic data and one does not allow the analysis of Drosophila data. There is also the case of three methods, which are available for free but must run in the MATLAB environment which is a commercial product. This makes a total of 13 non-useable methods which are identified by rows with white background in Tables 1 and S5. In these tables, eight methods that are potentially useable but not in an easy way for a non-specialist are displayed with an orange background. The difficulties come, for half of the cases, from the running of the applications in a software environment that has become obsolete: two methods work with java five which was released in 2004 (Java is currently in version 11) with a graphical interface based on Swing whose development was abandoned in 2014; two methods are plugins for Cytoscape two whose latest version dates from 2012. It might be possible to run these applications by reinstalling old versions of Java or Cytoscape, but one can expect to face difficulties in running these programs on recent software environments. In any case, these applications cannot be considered easy to use, so they have not been tested in our review. The four other programs considered as being difficult to use require the execution of several commands or transformations on the data which require a fair amount of work. At the end of this filtering, there are only four methods that an end user can use: JActiveModules (Ideker et al, 2002), the oldest method that comes as a plugin for Cytoscape 3, GiGA (Breitling et al, 2004) an application written in Perl, ClustEx2 (Gu et al, 2010), an application written in C++ and DIAMOnD (Ghiassian et al, 2015), an application written in Python 2. These last three applications must be run from the command line, but their invocation is straightforward. It should be noted that DIAMOnD identifies only one set of genes involved, unlike other methods that identify several modules. JActiveModules requires specification of an additional parameter which is the number of modules to identify. The three other methods require indicating the maximum size of the modules to be identified.

The commonality between all the existing methods is that their efficiency is highly dependent on the network topology. For these methods, two genes or gene products are as likely to be part of the same functional module as they are close on the interaction network. This proximity is described by the existence of an inter-action or a chain of interactions between the two molecules. Un-fortunately, molecular interaction networks stored in public databases are known to be noisy and incomplete (Wang et al, 2009). It has also been shown that PPI networks are "small-world" networks, meaning that the neighbors of a given node are likely to be neighbors of each other, and most nodes can be reached from every other node by a small number of hops.

Over the past few years, network embedding has emerged as a powerful network analysis approach by generating a highly informative and compact vector representation for each vertex in the network (Cui et al, 2018). This approach maps nodes into a vector space in which the distances between nodes accurately reflect their proximity in the original network. We hypothesize that network embedding methods can provide the basis for a more reliable method by estimating distances between nodes that take into account the entire topology of the graph and that in addition are little affected by the proportion of missing edges. Based on this assumption, we have developed AMINE, a new efficient method for detecting active modules in the vector space generated by an embedding of the interaction network (Pasquier et al, 2021 Preprint).

## Identification of active modules

Using the web server at http://amine.i3s.unice.fr to execute the method AMINE, we proceeded to identify active modules in an independent manner for each of the three time points. Our method identified seven active modules of genes at day 1, 38 active modules at day 4 and four active modules at day 20 (Figs 3–5 and Tables S6–S8).

Drawing the modules identified at the different time points on the same picture showed few overlapping genes (only five genes, FBgn0038252/BigH1, FBgn0260653/serp, FBgn0061492/loj, FBgn0039805/Cpr100A, and FBgn0022770/PeriA, which belong to modules identified at days 1 and 4) but highlights the fact that many modules interact between the different time points. Fig 6 depicts a network that expands from seven to 38 modules from day 1–4 and then shrinks to the four modules that are still active at 20 d.

Similar to the findings from the differential analysis, the genes highlighted at each of the three time points were mostly different. This observation reinforces the hypothesis of successive waves of gene activation. To highlight interactions between modules at different time points, we used the same data for an exhaustive search: for each module we determined whether the interactors of the genes they contain are included in other modules (1-hop neighbors) or whether they interact with genes in other modules (2-hop neighbors). Fig 7 shows a visualization of the intermodule interactions in the form of a network in which each node is a module and each link represents an interaction between one or more genes of the corresponding modules. One can thus visualize a cascade of interactions in which most of the modules are involved.

## Enrichment analysis of active modules

The identification of active modules allows us to gain a clearer view of the studied process. On day 1, we retrieved a group of under-expressed genes enriched in functions related to cuticle development and chitin metabolic processes. The functions were already identified through a traditional enrichment analysis of differentially expressed genes, but in this study, we focused on a group of 22 genes in comparison to the 335 genes detected as underexpressed by DESeq2 (Love et al, 2014). We noted that despite its limited size, this module contains genes that cannot be detected via a standard differential analysis. FBgn0040582/BomBc3, FBgn0031940/CG7214, FBgn0032507/CG9377, and FBgn0033869/Cpr50Cb are indeed associated with FDR ranging from 0.10 to 0.90, which are well above the commonly used minimum threshold of 0.05.
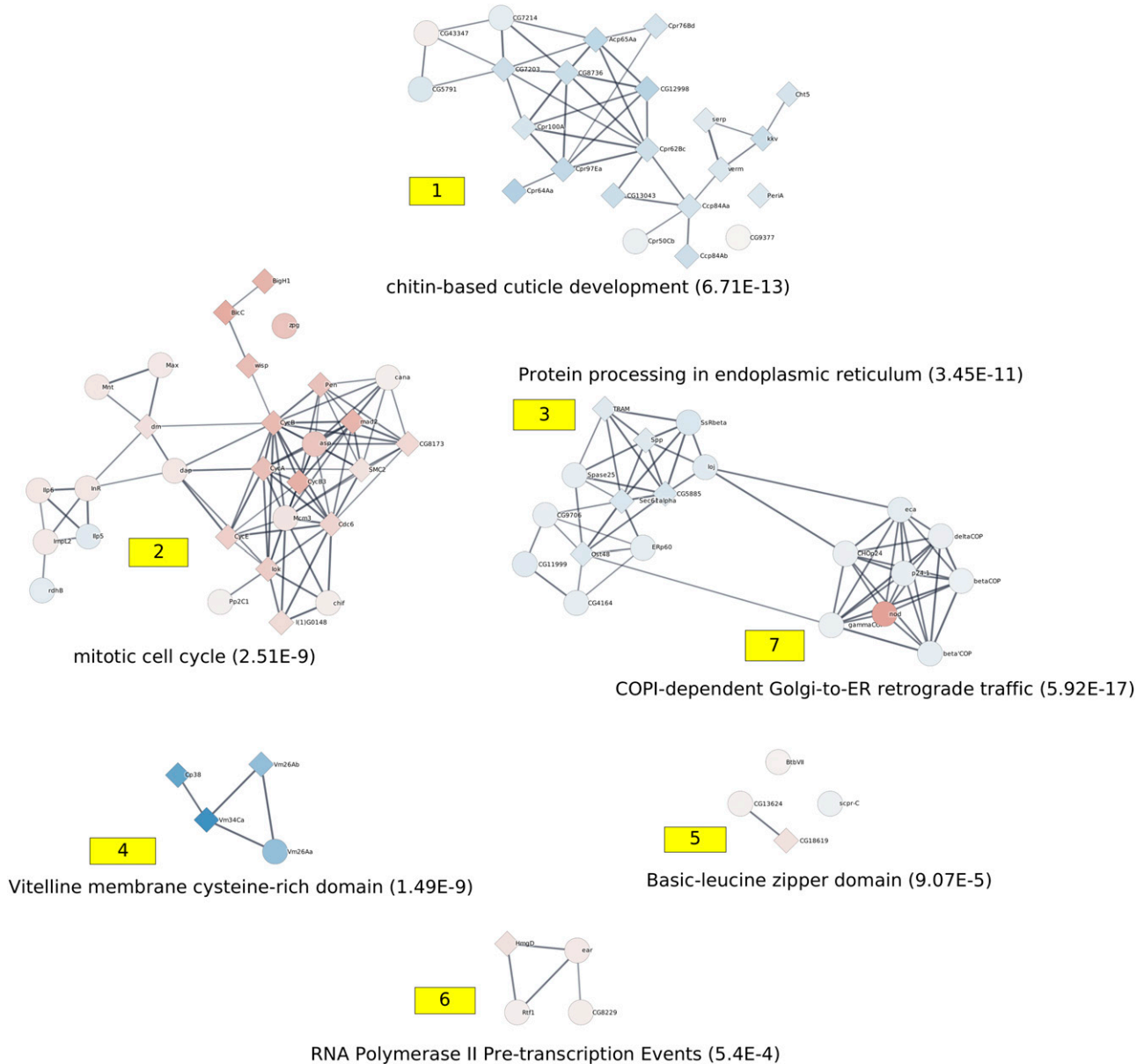
**Figure 3.  Representation of active modules identified at day 1.**

On the networks, the nodes correspond to genes, and the edges correspond to interactions reported in the String database with an evidence score ≥0.7. The number associated with each module corresponds to the module ID specified in Table S6. Each module is annotated with a representative enrichment. The complete lists of the enrichments of all the modules are shown in Table S9. The node colors represent the $\log_2$ FC values of the corresponding gene on a scale varying from blue (for the most underexpressed genes) to red (for the most overexpressed genes). The diamond-shaped nodes represent genes that are considered differentially expressed based on the DESeq2 method.

Based on data collected by FlyAtlas, FBgn0040582/BomBc3 is very highly expressed in the adult head, heart, fat body, and carcass. This gene is also very highly expressed in female spermathecae of both virgin and mated females (Chintapalli et al, 2007). BomBc3 is known to control the Toll pathway, which plays a key role in the innate immune system. FBgn0031940/CG7214 has been implicated in the immune response (Pal et al, 2008). FBgn0032507/CG9377 is up-regulated in the brains of *Drosophila melanogaster* courting males compared with noncourting males (Ellis & Carney, 2011). All of these genes are highly expressed in the head fat body. Changes in

expression in the fat body have been highlighted by a previous study in which the authors demonstrated that mating modifies fatty acid metabolism in the male brain (Ellis & Carney, 2010).

Among the other detected active modules, we found one module that contains overexpressed genes involved in the cell cycle and developmental processes, which are functions that have already been highlighted by the standard enrichment analysis procedure. This enrichment is also mostly related to the mitotic cell cycle and germ cell development and has been implied in metabolic processes. The third module was related to protein processing in the
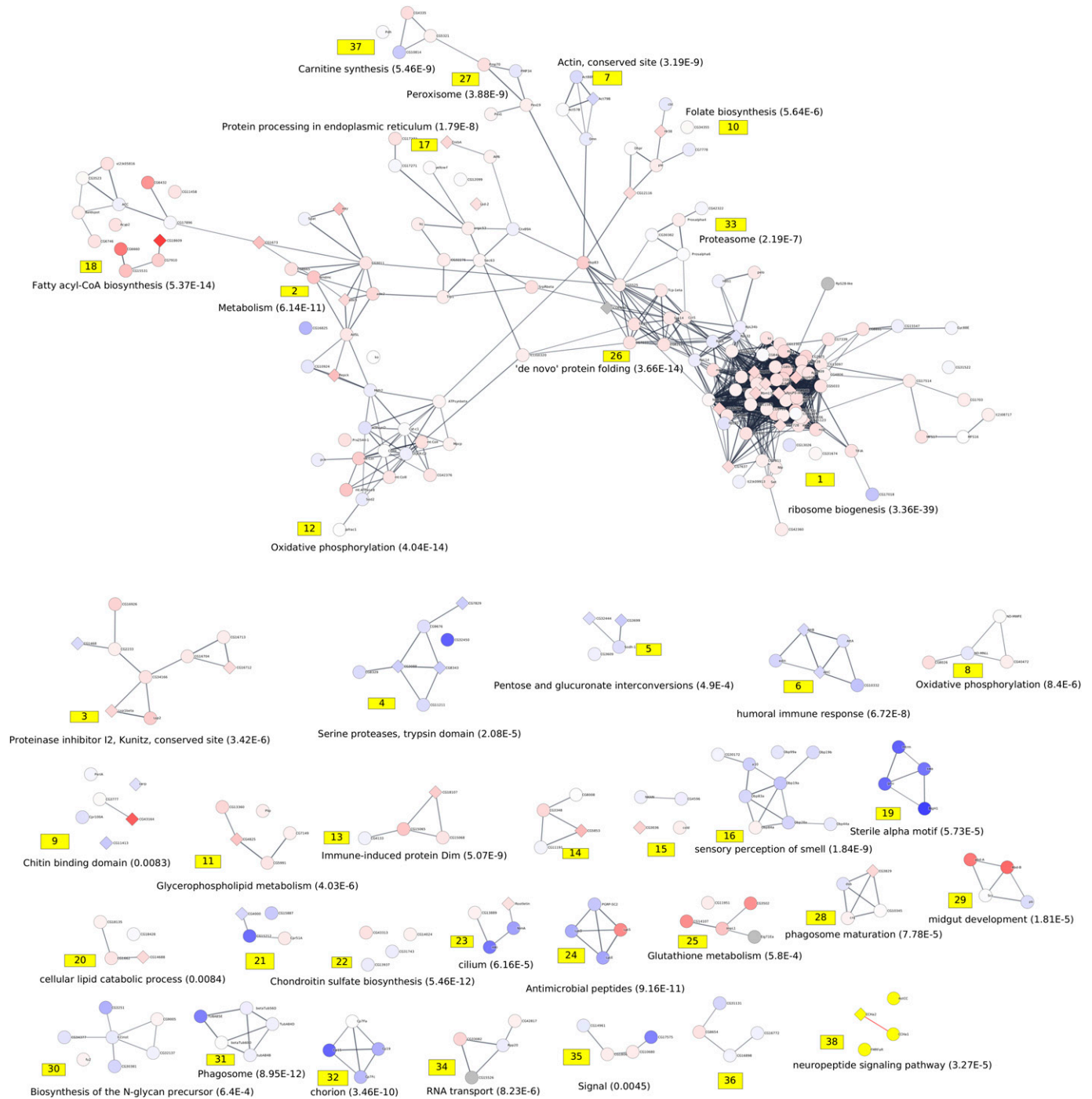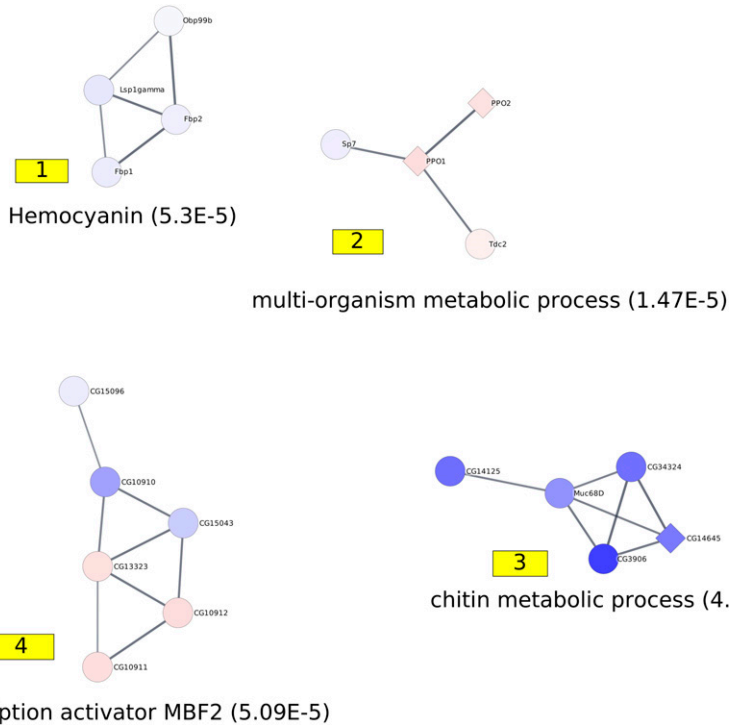
**Figure 4. Representation of active modules identified at day 4.**
On the networks, the nodes correspond to genes, and the edges correspond to interactions reported in the String database with an evidence score ≥0.7. The number associated with each module corresponds to the module ID specified in Table S7. Each module is annotated with a representative enrichment (when available). The complete lists of the enrichments of all the modules are shown in Table S10. The node colors represent the $\log_2$ FC values of the corresponding gene on a scale varying from blue (for the most underexpressed genes) to red (for the most overexpressed genes). The diamond-shaped nodes represent genes that are considered differentially expressed based on the DESeq2 method.

endoplasmic reticulum; the fourth module was involved in vitelline membrane formation; the fifth module was involved in the receptor signaling process; the sixth module was related to RNA polymerase II pretranscription events; and the seventh module was related to

COPI-dependent Golgi-to-ER retrograde trafficking (Fig 3 and Table S9).

At day 4, we retrieved a large module (module 1 of Fig 4) containing overexpressed genes linked to rRNA processing and

**Figure 5. Representation of active modules identified at day 20.**
On the networks, the nodes correspond to genes, and the edges correspond to interactions reported in the String database with an evidence score ≥0.7. The number associated with each module corresponds to the module ID specified in Table S8. Each module is annotated with a representative enrichment. The complete lists of the enrichments of all the modules are shown in Table S11. The node colors represent the $\log_2$ FC values of the corresponding gene on a scale varying from blue (for the most underexpressed genes) to red (for the most overexpressed genes). The diamond-shaped nodes represent genes that are considered differentially expressed based on the DESeq2 method.

ribosome biogenesis. We found 37 other active modules enriched in genes with implied functions investigated in other studies, such as changes in the immune response (Kapelnikov et al, 2008a, 2008b – modules 6 and 13), lowering the expression of genes implied in the perception of smell (Lebreton et al, 2014 – module 14), midgut development (White et al, 2021 – module 29), and other processes that have been uncovered (Fig 4 and Table S10). At 20 d, the transcriptomes of virgin and mated females were quite similar because only four active modules were highlighted. Among these modules, we found genes involved in chitin metabolism that were also identified as being differentially expressed at day 1 (Fig 5 and Table S11). Surprisingly, the bioinformatically extracted modules were very specific for each considered period. Rare overlaps were observed between periods. However, when all the genes within the different modules at the three periods were assembled, significant interconnected modules were found. This finding suggests that a chronological and temporal order of activation/inhibition occurs after mating and continues during the female life span. At day 1, numerous early genes were affected, and the vast majority of these early genes returned to normal at day 4. These temporal and sequential gene networks that are manipulated by mating hint at internetwork regulation: the first wave orchestrates the second wave and so on down the cascade. Because the analysis was restricted to head tissue, the predicted genes involved in egg development, processing food uptake, and general metabolism were not considered "a priori." However, the changes in the brain, which are likely related to changes in behavior, can be seen as considerable at days 1 and 4, although these changes appear to evaporate at the end of the life span.

## Comparison between AMINE, JActiveModules, ClustEx, DIAMOnD, and GiGA

The different programs were used to analyze the same data set, namely, the differential expression of genes at day 1. The execution of AMINE was performed as presented above. JActiveModules was executed using default parameters. ClustEx2 asks to specify the size of the largest module. There is a procedure to follow, described in the documentation, to determine the most appropriate size. Following the instructions, we have specified this size as 100. DIAMOnD generates a unique and ordered list of the most involved genes. The program asks for the size of this list which we set to 100. GiGA also uses the maximum module size as a parameter. The default value is 20; we have left it unchanged. On this dataset, AMINE finds seven modules ranging from 4 to 29 genes (Table S6), ClustEx2 finds 86 modules, of which some are very small and not really informative (Table S12). By keeping only the modules containing more than three genes, we obtain a list of 30 modules of size 4–100 (the maximum specified size). JActiveModules finds five modules (the number of modules specified in parameter) with size ranging from 23 to 228 (Table S13). DIAMOnD outputs an ordered list of 100 genes (Table S14). GiGA finds 21 modules of which two contain only two genes (Table S15). By keeping only the modules containing more than three genes, we obtain a list of 19 modules of size 4–19 (just below the specified maximum size).

The first thing we notice about jActiveModules is that although the method is able to identify several modules, they are very overlapping. Modules 1–3 contain a very similar group of genes. The content of modules 4 and 5 is more different, but mainly because these modules are larger. An Euler diagram displaying the five
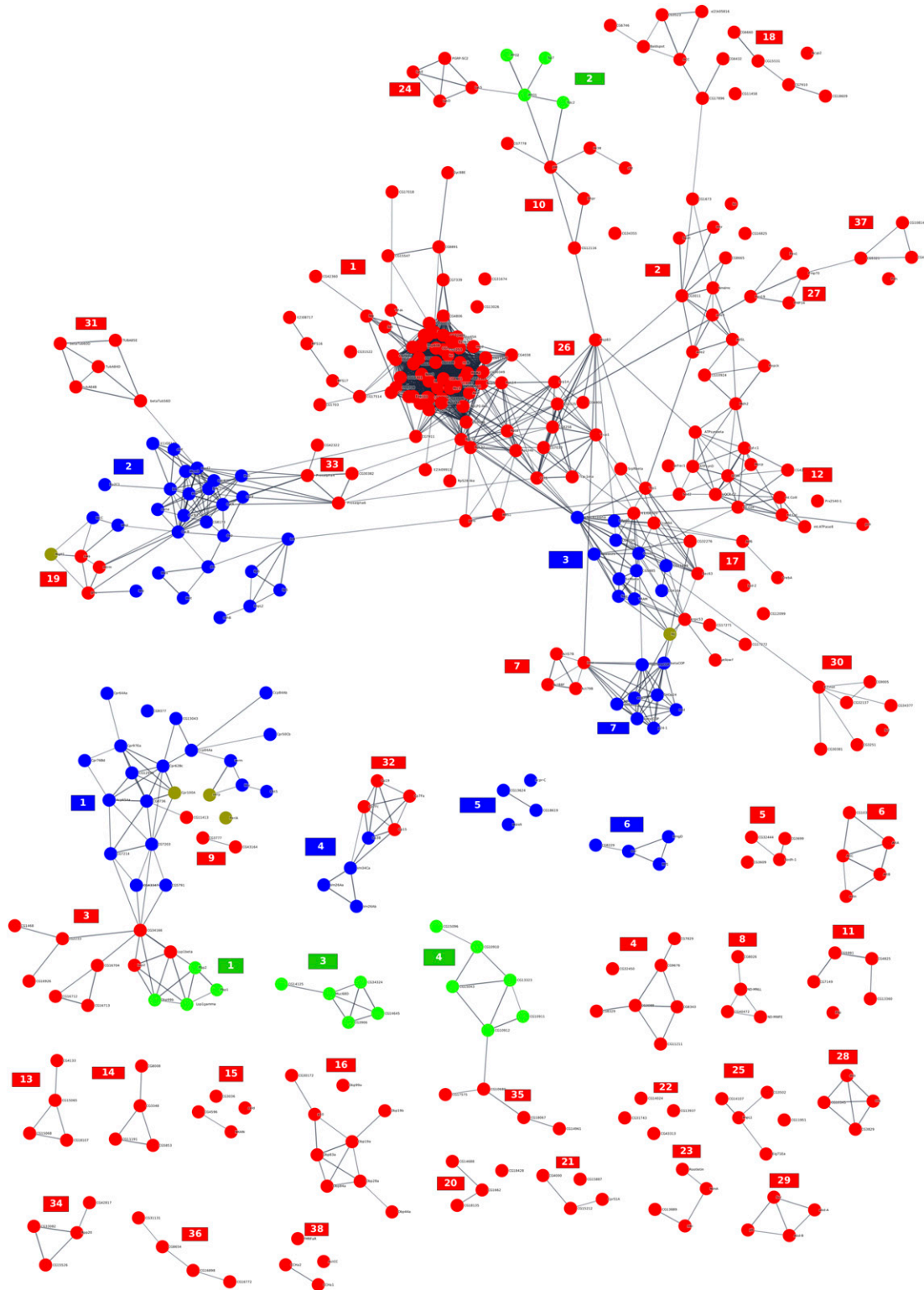
**Figure 6. Union of the active modules identified at each time point.**
The network combines the active modules identified at days 1 (in blue), 4 (in red), and 20 (in green). The module numbers correspond to those specified in Figs 3–5.

Active modules at 1 day          Active modules at 4 days          Active modules at 20 days
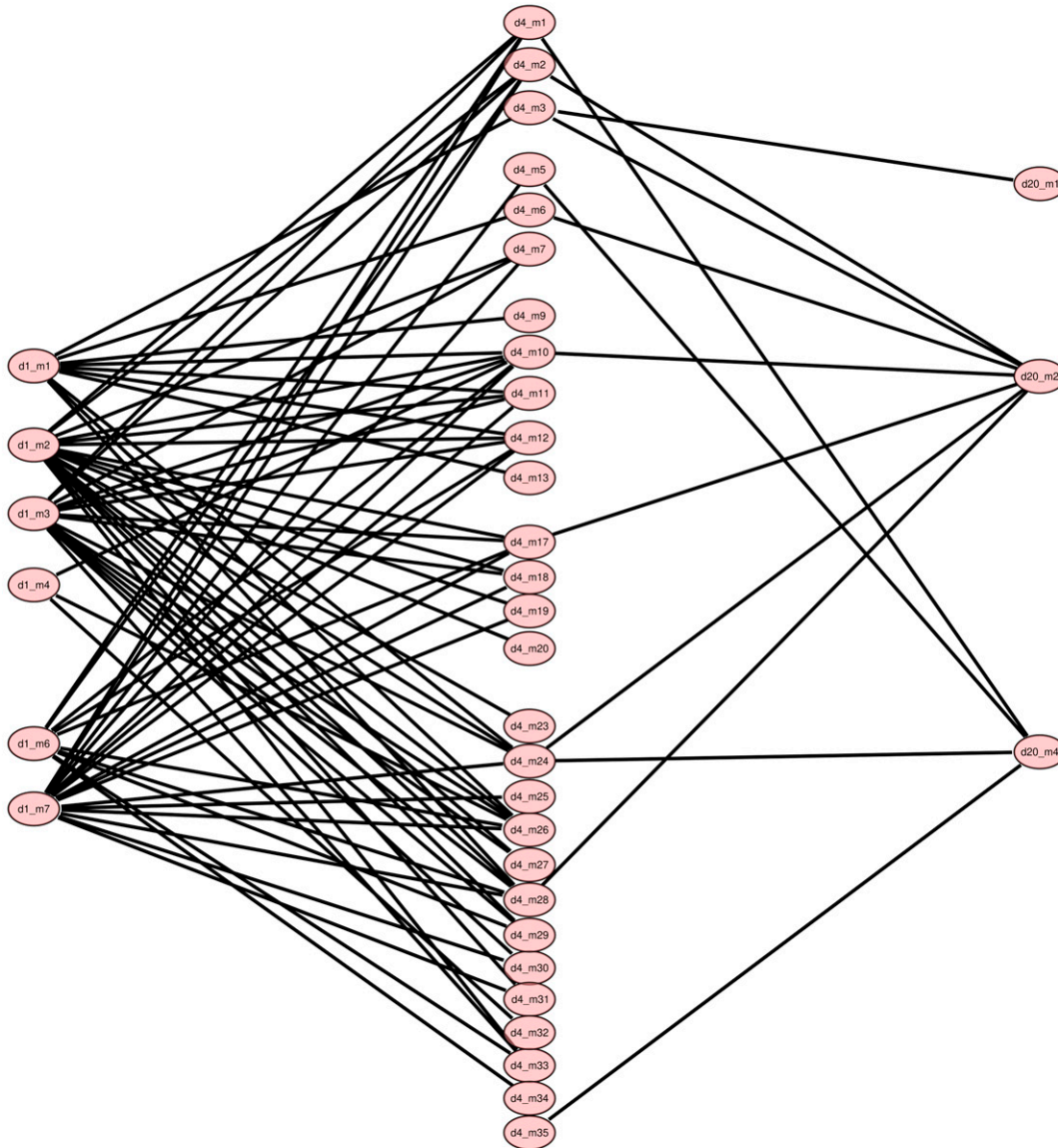


**Figure 7. Interactions between modules at different time points.**
Visualization of the intermodule interactions in the form of a network in which each node is a module and each link represents an interaction between one or more genes of the corresponding modules. The modules are labeled with the time point followed by the number of modules. Thus, the module name "d4_m1" refers to module number 1 identified at day 4.

modules found by JActiveModules is shown in Fig S1. The GO enrichment with biological process terms corresponding to the genes within the identified modules are presented in Table S16. Not surprisingly, the first three modules identified by JActiveModules correspond to the same function which is "chitin-based cuticle development." The two other modules correspond to "carbohydrate metabolic process." The 100 genes highlighted by DIAMOnD correspond to the very generic function "cell cycle." By selecting only the first genes appearing in the list, we obtain more specific functions: "cell cycle process" for the top 50 genes and "mitotic cell cycle process" for the top 25 genes. ClustEx2 and GiGA, like AMINE, generate non-overlapping gene modules. AMINE, ClustEx2, and GiGA

identify a different number of modules but it is difficult to decide which is better for the user. We can gain an idea of the relevance of the modules by looking at which ones are significantly enriched in terms from GO biological process. Among the seven modules identified by AMINE, six are enriched with GO Biological Process with a FDR of less than 0.05 (which accounts for 85% of the total). It is also the case for 15 of the 30 modules identified by ClustEx2 (50% of the total) and 9 of the 19 modules identified by GiGA (47%). These ratios are not proof in themselves, but the fact that the identified modules correspond to known processes is rather a sign that coherent and functional modules have been identified. If we adopt a more global view by looking at the high-level GO terms, we obtain,
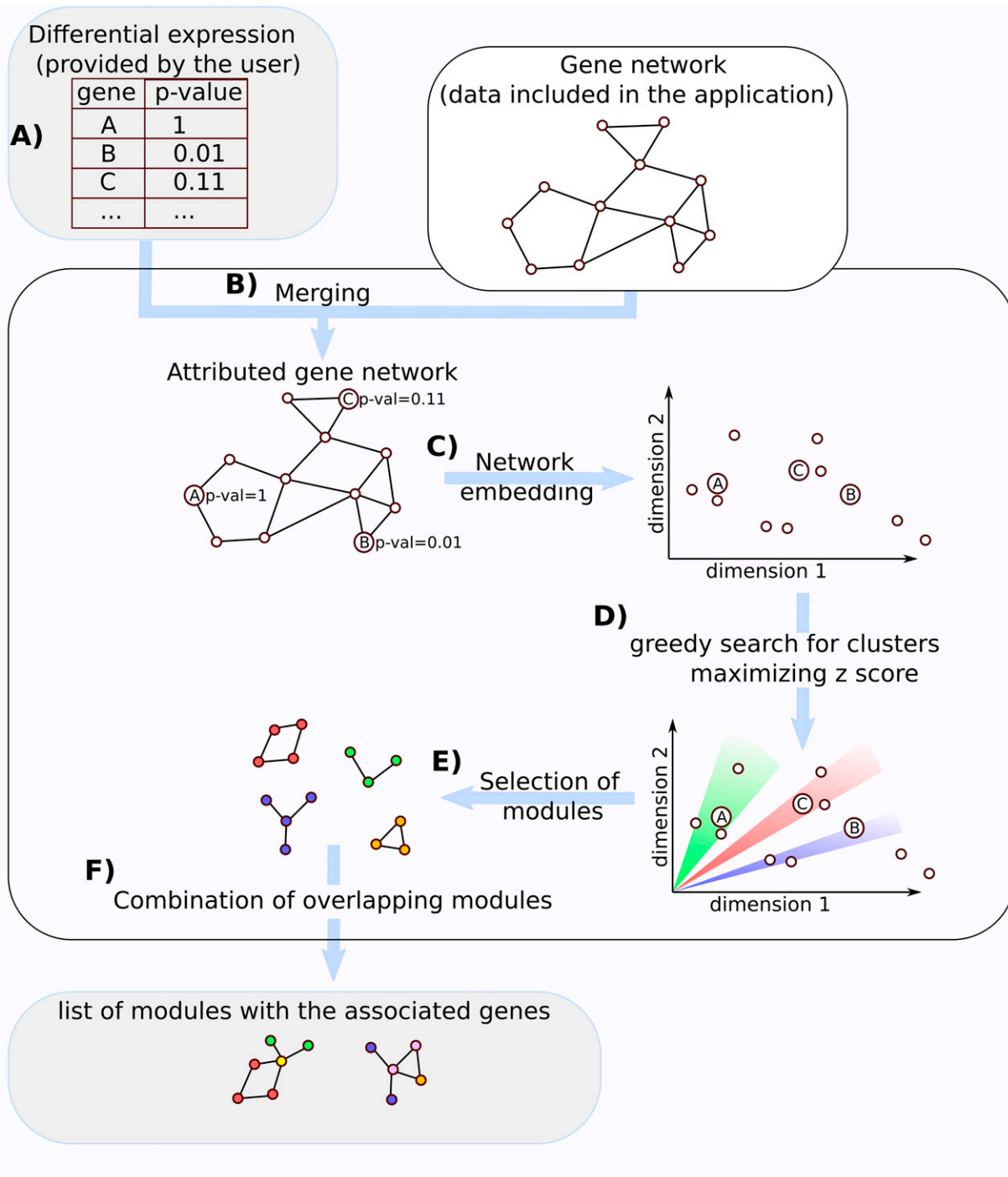
**Figure 8. Workflow of the AMINE method.**
**(A)** Input data are composed of a table storing the significance of the expression variation of genes between two conditions and a network representing known gene interactions. Gene variations must be provided by the user, whereas data concerning gene interactions are included in the application. **(B)** Data about gene interactions and gene variations are merged to generate an attributed gene network. **(C)** Nodes belonging to the attributed gene network are mapped to a low-dimensional space through the use of a biased Node2Vec method. **(D)** Sets of genes that are both cohesive and differentially expressed are identified in the embedded space by maximizing both the scores of the nodes and the cosine distance between the vectors representing the nodes. **(E)** Each gene from the gene network is associated with the cluster that maximizes the z-score. **(F)** Redundancy in the content of modules is ruled out by combining sets of nodes obtained in the previous step while ensuring that the result remains spatially cohesive.

**Table 1. Computational tools for active module identification.**

| name | Year | Reference | Type | Note |
|---|---|---|---|---|
| JActiveModules | 2002 | Ideker et al (2002) | Cytoscape plugin | OK |
| GiGA | 2004 | Breitling et al (2004) | cmd | OK |
| SAMBA | 2004 | Tanay et al (2004) | gui/cmd | Execution error |
| GXNA | 2007 | Nacu et al (2007) | cmd | Broken link |
| MATISSE | 2007 | Ulitsky and Shamir (2007) | gui | Difficult to use[a] |
| PinnacleZ | 2007 | Chuang et al (2007) | Cytoscape plugin | Difficult to use[b] |
| CEZANNE | 2009 | Ulitsky and Shamir (2009) | gui | Difficult to use[a] |
| BioNet | 2010 | Beisser et al (2010) | cmd | Difficult to use[c] |
| RegMOD | 2010 | Qiu et al (2010) | cmd | Not free to use[d] |
| ClustEx2 | 2010 | Gu et al (2010) | cmd | OK |
| RME | 2011 | Miller et al (2011) | cmd | Not for transcriptomics data[e] |
| COSINE | 2011 | Ma et al (2011) | cmd | Difficult to use[f,g] |
| EnrichNet | 2012 | Glaab et al (2012) | cmd | Unable to run on drosophila data |
| MEMo | 2012 | Ciriello et al (2012) | cmd | Difficult to use[h] |
| BMRF-net | 2013 | Chen et al (2013) | Cytoscape plugin | Difficult to use[b] |
| Walktrap-gm | 2013 | Petrochilos et al (2013) | cmd | Installation failure |
| TimeXNet | 2014 | Patil and Nakai (2014) | web/gui/cmd | No results[i] |
| WMAXC | 2014 | Amgalan and Lee (2014) | cmd | Broken link |
| DIAMOnD | 2015 | Ghiassian et al (2015) | cmd | OK[g] |
| Hotnet2 | 2015 | Leiserson et al (2015) | cmd | Difficult to use[j] |
| GLADIATOR | 2017 | Silberberg et al (2017) | cmd | Not for transcriptomics data[k] |
| MOEA | 2017 | Chen et al (2017) | cmd | Not free to use[d] |
| MRF | 2017 | Robinson et al (2017) | cmd | Not free to use[d] |
| ModuleDiscoverer | 2018 | Vlaic et al (2018) | cmd | Broken link |
| ResponseNet | 2019 | Basha et al (2019) | web | No results[i] |
| AMINE | 2021 | Pasquier et al (2021) *Preprint* | web | OK |

[a]Program written in Java 5 with a graphical interface based on Swing whose development was abandoned when Java 8 appeared in 2014.
[b]Not available for Cytoscape 3.
[c]Require to execute a dozen of lines to process the example data included with the program.
[d]Need MATLAB which is not free.
[e]The method identifies patterns of recurrent genomic aberration in tumor samples.
[f]Requires to execute multiple commands in R language, in particular it is necessary to compute a matrix containing the differential correlation between each pair of connected nodes on the network.
[g]Identifies only one module.
[h]The documentation contains the following warning: Generating the files needed by the program may require a fair amount of work, as several of them are the results of other algorithms.
[i]The program ran for several hours without providing any results.
[j]Requires to execute multiples commands ; the script used to analyze data of the associated paper is 68 lines long.
[k]The aim of the program is to identify relationship between diseases modules ; user need to suply known genes associated with diseases.

for AMINE, modules corresponding to "cuticle development," "cell cycle," "transport," "reproduction," and "signaling." For clustex2, 8 of the 15 modules found are associated with these same annotations except for "cell cycle," for JActiveModules, three modules are associated with "cuticle development." The set of genes highlighted by DIAMOnD are associated with "cell cycle" and GiGA reports modules associated with three annotations found by AMINE, 1 by ClustEx2, and 1 by DIAMOnD. In addition to the six main biological processes highlighted by AMINE, ClustEx2 identifies modules corresponding to "cell organisation," "cellular metabolic process," and "cell differentiation." The two large modules identified by JActiveModules correspond to "carbohydrate metabolic process" and GiGA finds modules associated with: "cellular metabolic process," "small molecule metabolic process," "response to radiation," and "system development." Overall, all processes revealed by amine are shared with at least one other method. Two other annotations were identified only by ClustEx2, one only by JActive-Modules and three only by GiGA (Fig S2). If we look at the set of

genes considered, by each of the methods, as being important in the process studied, we notice that there is no gene in common between the five methods and that a significant proportion of genes is only found by only one method (Fig S3A). However, this proportion of genes only identified by one method differs; it is 32% for amine, 36% for GiGA, 62% for JActiveModules, 69% for ClustEx2, and 86% for DIAMOnD. Overall, we note that the AMINE method identifies the fewest genes but has the most genes found in common with each of the other methods. If we compare with the genes identified by differential expression analysis using a *P*-value less than 0.05 and a $\log_2$FC greater than 0.32 or less than −0.32, we find that most of the genes with significant variation are not highlighted by either of the module detection methods. Conversely, 443 genes are only identified by module detection methods (Fig S3B). This means that all the reviewed module detection methods have the ability to pinpoint genes that may have an influence on the studied process but whose expression variations are very small (a $\log_2$FC of 0.32 corresponds to a variation of ±25%).

The Biological Process "chitin-based cuticle development" is an important process occurring at day 1. This annotation has been highlighted both by differential gene analysis and by four out of five module detection methods. AMINE found a cluster of 22 genes implied in this process, ClustEx2 highlighted a cluster of 100 genes, JActiveModules detected three overlapping modules representing a total of 37 genes, and GiGA found a cluster of 19 genes. Fig S4A shows the overlap between the modules associated with "chitin-based cuticle development" for the four different methods. Differential expression analysis reveals 625 genes with a $\log_2$FC greater than 0.32 or less than −0.32. Of these 625 genes, 63 (representing 10%) were also identified by one or more of the module detection methods. However, even when using a $\log_2$FC threshold as low as 0.32, 45 genes were only revealed using module detection methods. This again emphasizes that searching for groups of genes working together can pick out important genes that move very little during the experiment. The overlap between differentially expressed genes and modules associated with "chitin-based cuticle development" for AMINE, ClustEx2, JActiveModules, and GiGA is shown in Fig S4B. This illustrates that only using methods that combine both the differential expression of genes and their interactions can relate the process under study to particular genes that vary only slightly.

This comparative analysis on a real-world dataset highlights the advantages of AMINE over the few other easy-to-use methods. First of all, it does not require installation because it can be run via a web interface. It does not require any special formatting of the input data because it is sufficient to provide tabulated or Excel files generated by differential analysis pipelines. The use of AMINE is particularly simple because no parameters are required, unlike the other four tested methods which requires either the number of modules to search or the maximum size of the modules. In terms of results, they are rather more concise than the other methods because the modules are less numerous or contain fewer genes. This conciseness allows biologists to focus on a smaller number of genes that appear to be highly relevant to the process being studied because a large proportion of them are also identified by other methods.

# Discussion

Changes in gene expression in the female head appear unique at different time scales after mating that fills spermathecae with sperm and other molecules in the seminal fluid. These changes appear to be a continued developmental program occurring in adulthood. The *Drosophila* female phenotypic plasticity after mating is remarkable, as stimuli released and/or contained in sperm might orchestrate a programmatic outcome that underlies the transcriptomic temporal changes in mated female flies. Changes in the behavior of mated versus virgin female flies could be seen as an extension of phenotypic plasticity, and from this point of view, caste distribution in eusocial insects could be considered as the most spectacular and achieved modality. Sperm acquisition, as any environmental change that faces the *Drosophila* female, can alter some physiological functions and behavior over short and long time scales. Our analysis unraveled the transcriptomic modifications in the *Drosophila* model through which the genome shapes diverse physiological functions in relation to metabolism needed for fertilized egg laying. Unrelated to the genetic coding of spermatozoids, the effect of male seminal fluid proteins on the post-mating physiology and behavior of females has been documented in the past. The implicated molecules likely include peptides, proteins, small molecules, and RNAs that act to promote profound behavioral changes regarding oviposition choices, changes in feeding behavior and mating refractoriness in females (McGraw et al, 2004, 2008).

Mating-induced transcriptome changes in females have been also reported in diverse tissues or the entire body of many insects, including *D. melanogaster* (Lawniczak & Begun, 2004; McGraw et al, 2004, 2008; Innocenti & Morrow, 2009; Short & Lazzaro, 2013), *Apis mellifera* (Kocher et al, 2008), *Ceratitis capitate* (Gomulski et al, 2012), and *Anopheles gambiae* (Rogers et al, 2008). To this regard, the transcriptomic changes in the oviduct of *D. melanogaster* and *Aedes aegypti* have been extensively studied (Kapelnikov et al, 2008a, 2008b; Alfonso-Parra et al, 2016). The kinetics of the published RNA-Seq lists show that in these later species, few genes are affected after a few hours of copulation, whereas maximal gene expression changes are observed at 6 h (Heifetz et al, 2001, 2014; Heifetz & Wolfner, 2004; Carmel et al, 2016). In honeybee, the mating process induce marked changes in the expression of genes related to vision, chemoreception, metabolism, and immunity (Fahrbach et al, 1995; Richard et al, 2007; Kocher et al, 2009). In contrast, reverting the transcriptome from mated females back to virgin females has not been observed in honeybee species in which distinct phenotypes emerge from the same genotype, giving caste attributes to individuals (Page & Peng, 2001). The gene expression levels in the full-scale genome of eusocial insects like honeybees and aunts has provided the evidence for the irreversible programmatic changes associated with mating in mated versus the virgin queen. (Whitfield et al, 2003; Grozinger et al, 2007; Simola et al, 2016). The transcriptomic changes map known behavioral and physiological traits that characterize the transition from a virgin queen to a newly mated queen. If most queens start laying eggs after one mating flight, others will engage in multiple mating flights with a new set of multiple males at each time, which enrich the

seminal male compounds in relation to their genetic background (Woyke, 1964; Tarpy & Page, 2000). More interestingly, virgin queens are photophilic (Kocher et al, 2010) and aggressive towards other virgin queens (Manfredini et al, 2015). In contrast, mated queens are photophobic, lack aggressiveness, and move protected by clustered workers (Kocher et al, 2010; Manfredini et al, 2015). In social Hymenoptera (ants, some bees, and wasps), mated queens that present extended lifespans compared with nonreproductive workers or virgin queen show the same physiological and behavioral changes, which triggers a nest-bound egg-laying status (Page & Peng, 2001; Heinze & Schrempf, 2008; Castella et al, 2009; Simola et al, 2016). However, for most of insect species for which this phenomenon has been reported, the molecular mechanisms remain unelucidated because of the lack of powerful genetic tools for silencing genes.

In contrast, significant advances have been obtained with the *Drosophila* model for which the seminal fluid protein sex peptide (SP) was found to activate an identified receptor in female fly, which constitutes one rare known pathway in insect world (Yapici et al, 2008). Defects in this seminal fluid signaling process are hypothesized to provoke reproductive disorders in females, which eventually results in reduced offspring output or unhealthy offspring or, in an ultimate scenario, a reproductive barrier that might guide or lead to speciation (Delbare et al, 2017). Among the other known molecules in *Drosophila*, the seminal fluid proteins (Acps) produced by male accessory glands were found partly responsible for the post-mating transcriptomic changes (Kubli, 2003; Swanson, 2003; McGraw et al, 2004; Peng et al, 2005; Isaac et al, 2010). More than 100 potential Acps have been identified in male *Drosophila sperm* as potential candidates to induce changes in tissue specific transcriptome and a lack of one of them, SP (Acp70A), in seminal fluid is correlated with a reduced number of laid eggs (Kubli, 2003; Swanson, 2003), which triggers female remating (Chapman et al, 1996, 2003; Liu & Kubli, 2003) and lowers female food uptake (Carvalho et al, 2006). The same way, Acp29AB and Acp62F were found responsible for a large panel of modified genes in mated female (McGraw et al, 2004).

Although modified gene expression is expected to refer to egg formation and is observed in ovaries and reproductive tracts, researchers have noted that marked transcriptional changes occur in the brain, fat body, or other tissues (McGraw et al, 2008; Mack, et al, 2006; Kapelnikov et al, 2008a, 2008b; Prokupek et al, 2009). Data obtained using microarray techniques have revealed a transcriptional response to mating with a peak of intensity after 3 h and a sharp decline at later time points (McGraw et al, 2008). Other researchers using the same technology have confirmed a peak at 6 h with changes that involve more than 500 genes (Mack, et al, 2006; Kapelnikov et al, 2008a, 2008b; Prokupek et al, 2009). Genetic manipulation that causes defects in ejaculation have allowed to identify molecules involved in post-mating modifications and the role of individual seminal factors for each specific female physiological trait (Chapman et al, 2003, 1996; Liu & Kubli, 2003; Wigby & Chapman, 2005; Wigby et al, 2011).

A comprehensive review outlines the seminal proteins that have been identified in many insect species for which the physiological functions in relation with female post-mating responses remain elusive (Avila et al, 2011). However, to disentangle the roles of

spermatozoid component, the seminal Acps and other active molecules in seminal fluid will be a very challenging task for most of insect species.

In *Drosophila*, integrative signaling molecules present in seminal fluid coordinate global homeostasis and physiology involving the female nervous system, fat body, endocrine cells, gut/microbiome, and reproductive tissues. The role of seminal fluid as a contributor to the fine regulation in a continuum with the action of the couple juvenile hormone/ecdysone in young adult females has been investigated in the past on the basis of genetic screen of mutants (Toivonen & Partridge, 2009; Rajan & Perrimon, 2011; Droujinine & Perrimon, 2013, 2016).

In this report, we have restrained our analysis to female brain transcriptional modification induced by mating. We found that only two post-mating time periods presented considerable scale changes at 1 d (625 genes), which declined at 4 d (82 genes) to end at 20 d, when only a few genes showed lasting expression changes. Amazingly, most of the identified genes showing modified expression at each stage were unique with little, if any, overlap with those identified at other stages. This finding argues in favor of time-specific activation/inhibition of gene expression organized by the recruitment of successive waves of interconnected genes that occurs after mating. Obviously, the first cascade of changes induces specific changes in the second cascade, and so on up to the end of the process. The most striking fact is that each wave presents a unique set of implicated genes. Unfortunately, it is difficult to characterize these interaction cascades more precisely given the data used in the study. Indeed, we only have data obtained at 1, 4, and 20 d, but many changes, including the activation of other gene modules that cannot be detected, can occur between these time points.

It is clear that signaling molecules in seminal liquid and/or within spermatozoid cells act at distance through unknown processes, such as the hemolymph transport of soluble compounds or microvesicles traveling with RNA and lipids, to regulate egg production, homeostasis, and behavior. Many reports in the literature have described the molecular transport between tissues as intertissue dialogue, which attributes credibility to the present scenarios (Rajan & Perrimon, 2011; Droujinine & Perrimon, 2013, 2016). In this regard, this field of investigation has achieved solid advances in mammals where extracellular vesicles containing RNA molecules have been shown to be transported through the blood–brain barrier, which results in the delivery of functional RNAs to brain neurons, the origin of these RNAs being far remote, mainly located in peripheral tissues such as the genital tractus (Rassoulzadegan et al, 2006; Gapp et al, 2014; Skinner et al, 2015; Chen et al, 2016a, 2016b).

Recent findings have proven that RNAs are capable of conferring information regarding interphase environment/germline interactions to guide, to some extent, the destiny of offspring through modification of developmental processes (Rassoulzadegan et al, 2006; Gapp et al, 2014; Skinner et al, 2015; Chen et al, 2016a, 2016b). Recently, robust experimental data have documented a direct causal role of sperm RNAs in transferring deleterious phenotypic traits across generations in mammals, which adds a new paradigm to an already well documented heredity driven by chromosomal DNA mutation insertions/deletions that generate allele catalogs (Rassoulzadegan et al, 2006; Gapp et al, 2014; Skinner et al, 2015; Chen et al, 2016a, 2016b). These studies in mammals might suggest

that similar scenario might occur in insects like *Drosophila* where seminal compounds including RNAs could manipulate the phenotypes of offspring in parallel with their action on mated female brain. If "memorized" sperm RNA or other molecules turn to be epigenetic markers, we might hypothesize that the female *Drosophila* responses to sperm signals could be co-substantial to phenotypic changes in some offspring traits in a non-Mendelian manner as a response to life history and environmental cues experienced by parents.

Understanding how mating during *Drosophila* reproduction triggers a developmental program extension in the female life span constitutes conceptual and experimental challenges to understand the insect reproduction biology.

# Materials and Methods

### Mating protocol and fly maintenance

Flies were raised at 25° under 16:8 h light–dark cycle and kept in bottles with standard cornmeal food media. For the mating protocol, the emerged flies were collected on short period of time (1 h) immediately after eclosion. Males and females at 1:1 ratio were left for 24 h together in fresh bottles of food to achieve 100% mating. Females were then removed and placed in fresh bottle for 1, 4, and 20 d. Fly heads were isolated and anesthetized by C02 and flash frozen in liquid nitrogen for 1 min. Male and female were separated. The frozen heads were cut off with a dissecting forceps. A group of 100 flies were processed for each determination. The biological material were stored at –80°C until the RNA was purified. See Brown et al (2014) and Graveley et al (2011) for more details.

### Differential expression analysis

Data of gene expression in the head of both virgin and mated *Drosophila* females made available by the modENCODE project were downloaded as FASTQ files from the Sequence Read Archive with the accession ID: PRJNA75285. For the RNA isolation, Illumina RNA-Seq library construction and sequencing, Read mapping and filtering, Differential Gene Expression Analysis, and Statistical methods are developed in supplemental information of the two related canonical articles (Graveley et al, 2011; Brown et al, 2014). The data contain the results from RNA-Seq experiments performed at three different time points: 1, 4, and 20 d. Details of these datasets are presented in Table S17.

The transcript abundance was quantified with Salmon (Patro et al, 2017), and the differential gene expression at each time point was calculated using DESeq2. Genes were considered differentially expressed if they showed a variation of at least a factor of 1.25 (i.e., a log$_2$-fold change greater than 0.32 or less than –0.32) associated with an adjusted *P*-value of at most 0.05.

### Enrichment analysis of differentially expressed genes

Functional enrichments for Gene Ontology terms, KEGG and Reactome Pathways, protein domains and UniProt Keywords were retrieved using Cytoscape (Shannon et al, 2003) and the StringApp plugin (Doncheva et al, 2018). The annotations identified with an FDR of less than 0.05 are listed in Table S4.

### Identification of active modules

The search for active modules has been performed with AMINE (Pasquier et al, 2021 *Preprint*). The AMINE method was designed to identify the modules of genes that are triggered in a biological experiment. It uses as input the background knowledge on the interactions between genes and measurements representing, in the specific context of a given experiment, indicators of the involvement of genes in the studied process. Gene interactions data are retrieved from the String database (Szklarczyk et al, 2017) with a combined evidence score ≥0.7 and included in the application. Measurements on the activity of genes must be supplied by the user. They are represented by the *P*-value, adjusted *P*-values, or FDR associated with the fold changes of the genes. At a minimum, the data are composed of a tabular file consisting of two columns: one containing the name of the gene and the other, the associated *P*-value. A more convenient way to specify the data to be used is to simply input the result file produced by differential expression analysis methods such as DeSeq2 or EdgeR (Robinson et al, 2010). The input file can be in csv (comma-separated values) or xlsx format (Fig 8A).

The interaction data are merged with gene associated *P*-values to generate an attributed gene network in which vertices represent genes, edges represent interactions between genes, and each vertex is annotated with a numeric attribute reflecting its associated *P*-value (Fig 8B).

The method relies on Node2vec (Grover & Leskovec, 2016) to generate a highly informative and compact vector representation for each vertex in the network. This transformation is called network embedding (Fig 8C). In the resulting vector space, the cosine distance between the vectors representing the nodes accurately reflect their proximity in the original network. AMINE, then, uses a greedy approach to build, from each vertex, groups of vertices of increasing size based on the aggregation of the closest vertices; the closeness being calculated with the cosine distance of the vertices' encoding vectors. Each group of vertices is then evaluated using Stouffer's Z method (Ideker et al, 2002). To do this, we first transform the *P*-value associated with each vertex into a z-score. The z-score $z(v_i)$ corresponding to the vertex $v_i$ associated with the *P*-value $p_i$, is calculated with $z(v_i) = \Phi^{-1}(1 - p_i)$, where $\Phi$ is the standard normal cumulative distribution function. Let $lov(v_i)$ be the list of vertices in the network sorted by their cosine distance from $v_i$ (closest vertices fist, which implies that $v_0 = v_i$), the aggregate z-score $z(v_i, k)$ for a set of nodes composed of the $k$ closest vertices of $v_i$ in the vector space is computed with:

$$z(v_i, k) = \frac{1}{\sqrt{k+1}} \sum_{j=0}^{j<k+1} z(lov(v_i)[j])$$

We get the best cluster associated to each vertex by selecting $k$ such that it allows us to obtain the best z-score (Fig 8D and E).

The last phase of the method consists in combining the different clusters while ensuring that new merged clusters remain spatially cohesive (Fig 8F). We say that two clusters $M_i$ and $M_j$ are spatially

cohesive when $M_i \cap M_j \neq \varnothing$. Starting from the module with the higher z-score, the process consists in evaluating all possible clusters formed by $M_i \cup \wp(M_{j/j\neq i})$ using the z-score, with $\wp(M_i)$ denoting the powerset of $M_i$ and keeping the modules with the highest z-score. The workflow of the AMINE method is presented in Fig 8.

The details of the method and the demonstration of its effectiveness are described in an article by Pasquier et al (2021) *Preprint*. The great advantage of AMINE is that it does not require any settings; it is not even necessary to indicate the number of modules to be identified or the size of the modules.

## Data access

The transcriptomic datasets were retrieved from the Sequence Read Archive: https://www.ncbi.nlm.nih.gov/sra with accession numbers SRR070436, SRR070437, SRR100281, SRR070430, SRR100278, SRR100282, SRR070388, SRR070419, SRR100275, SRR070434, SRR070435, SRR100279, SRR070414, SRR070415, SRR111882, SRR070420, SRR116383, and SRR100274.

# Supplementary Information

Supplementary Information is available at https://doi.org/10.26508/lsa.202101119.

# Acknowledgements

## Author Contributions

C Pasquier: conceptualization, software, validation, methodology, and writing—original draft, review, and editin.
A Robichon: conceptualization, validation, methodology, and writing—original draft, review, and editing.

## Conflict of Interest Statement

The authors declare that they have no conflict of interest.

# References

Alfonso-Parra C, Ahmed-Braimah YH, Degner EC, Avila FW, Villarreal SM, Pleiss JA, Wolfner MF, Harrington LC (2016) Mating-induced transcriptome changes in the reproductive tract of female Aedes aegypti. *PLoS Negl Trop Dis* 10: e0004451. doi:10.1371/journal.pntd.0004451

Ameku T, Niwa R (2016) Mating-induced increase in germline stem cells via the neuroendocrine system in female Drosophila. *PLoS Genet* 12: e1006123. doi:10.1371/journal.pgen.1006123

Amgalan B, Lee H (2014) WMAXC: A weighted maximum clique method for identifying condition-specific sub-network. *PLoS One* 9: e104993. doi:10.1371/journal.pone.0104993

Avila FW, Sirot LK, Laflamme BA, Rubinstein CD, Wolfner MF (2011) Insect seminal fluid proteins: Identification and function. *Annu Rev Entomol* 56: 21–40. doi:10.1146/annurev-ento-120709-144823

Basha O, Mauer O, Simonovsky E, Shpringer R, Yeger-Lotem E (2019) ResponseNet v.3: Revealing signaling and regulatory pathways connecting your proteins and genes across human tissues. *Nucleic Acids Res* 47: W242–W247. doi:10.1093/nar/gkz421

Beisser D, Klau GW, Dandekar T, Müller T, Dittrich MT (2010) BioNet: An R-package for the functional analysis of biological networks. *Bioinformatics* 26: 1129–1130. doi:10.1093/bioinformatics/btq089

Beldade P, Mateus AR, Keller RA (2011) Evolution and molecular mechanisms of adaptive developmental plasticity. *Mol Ecol* 20: 1347–1363. doi:10.1111/j.1365-294X.2011.05016.x

Braendle C, Félix MA (2008) Plasticity and errors of a robust developmental system in different environments. *Dev Cell* 15: 714–724. doi:10.1016/j.devcel.2008.09.011

Braendle C, Felix MA (2009) The other side of phenotypic plasticity: A developmental system that generates an invariant phenotype despite environmental variation. *J Biosci* 34: 543–551. doi:10.1007/s12038-009-0073-8

Breitling R, Amtmann A, Herzyk P (2004) Graph-based iterative group analysis enhances microarray interpretation. *BMC bioinformatics* 5: 100–110. doi:10.1186/1471-2105-5-100

Brown JB, Boley N, Eisman R, May GE, Stoiber MH, Duff MO, Booth BW, Wen J, Park S, Suzuki AM, et al (2014) Diversity and dynamics of the Drosophila transcriptome. *Nature* 512: 393–399. doi:10.1038/nature12962

Carmel I, Tram U, Heifetz Y (2016) Mating induces developmental changes in the insect female reproductive tract. *Curr Opin Insect Sci* 13: 106–113. doi:10.1016/j.cois.2016.03.002

Carvalho GB, Kapahi P, Anderson DJ, Benzer S (2006) Allocrine modulation of feeding behavior by the sex peptide of Drosophila. *Curr Biol* 16: 692–696. doi:10.1016/j.cub.2006.02.064

Castella G, Christe P, Chapuisat M (2009) Mating triggers dynamic immune regulations in wood ant queens. *J Evol Biol* 22: 564–570. doi:10.1111/j.1420-9101.2008.01664.x

Chapman T, Bangham J, Vinti G, Seifried B, Lung O, Wolfner MF, Smith HK, Partridge L (2003) The sex peptide of Drosophila melanogaster: Female post-mating responses analyzed by using RNA interference. *Proc Natl Acad Sci U S A* 100: 9923–9928. doi:10.1073/pnas.1631635100

Chapman T, Choffat Y, Lucas WE, Kubli E, Partridge L (1996) Lack of response to sex-peptide results in increased cost of mating in dunce Drosophila melanogaster females. *J Insect Physiol* 42: 1007–1015. doi:10.1016/s0022-1910(96)00070-4

Chen L, Xuan J, Riggins RB, Wang Y, Clarke R (2013) Identifying protein interaction subnetworks by a bagging Markov random field-based method. *Nucleic Acids Res* 41: e42. doi:10.1093/nar/gks951

Chen Q, Yan M, Cao Z, Li X, Zhang Y, Shi J, Feng GH, Peng H, Zhang X, Zhang Y, et al (2016a) Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder. *Science* 351: 397–400. doi:10.1126/science.aad7977

Chen Q, Yan W, Duan E (2016b) Epigenetic inheritance of acquired traits through sperm RNAs and sperm RNA modifications. *Nat Rev Genet* 17: 733–743. doi:10.1038/nrg.2016.106

Chen W, Liu J, He S (2017) Prior knowledge guided active modules identification: An integrated multi-objective approach. *BMC Syst Biol* 11: 8. doi:10.1186/s12918-017-0388-2

Chintapalli VR, Wang J, Dow JA (2007) Using FlyAtlas to identify better Drosophila melanogaster models of human disease. *Nat Genet* 39: 715–720. doi:10.1038/ng2049

Chuang HY, Lee E, Liu YT, Lee D, Ideker T (2007) Network-based classification of breast cancer metastasis. *Mol Syst Biol* 3: 140. doi:10.1038/msb4100180

Ciriello G, Cerami E, Sander C, Schultz N (2012) Mutual exclusivity analysis identifies oncogenic network modules. *Genome Res* 22: 398–406. doi:10.1101/gr.125567.111

Cui P, Wang X, Pei J, Zhu W (2018) A survey on network embedding. *IEEE Trans Knowl Data Eng* 31: 833–852. doi:10.1109/TKDE.2018.2849727

Dalton JE, Kacheria TS, Knott SR, Lebo MS, Nishitani A, Sanders LE, Stirling EJ, Winbush A, Arbeitman MN (2010) Dynamic, mating-induced gene expression changes in female head and brain tissues of Drosophila melanogaster. *BMC Genomics* 11: 541. doi:10.1186/1471-2164-11-541

Debat V, David P (2001) Mapping phenotypes: Canalization, plasticity and developmental stability. *Trends Ecol Evol* 16: 555–561. doi:10.1016/s0169-5347(01)02266-2

Delbare SYN, Chow CY, Wolfner MF, Clark AG (2017) Roles of female and male genotype in post-mating responses in Drosophila melanogaster. *J Hered* 108: 740–753. doi:10.1093/jhered/esx081

DeWitt TJ, Sih A, Wilson DS (1998) Costs and limits of phenotypic plasticity. *Trends Ecol Evol* 13: 77–81. doi:10.1016/s0169-5347(97)01274-3

Dingemanse NJ, Kazem AJ, Réale D, Wright J (2010) Behavioural reaction norms: Animal personality meets individual plasticity. *Trends Ecol Evol* 25: 81–89. doi:10.1016/j.tree.2009.07.013

Doncheva NT, Morris JH, Gorodkin J, Jensen LJ (2018) Cytoscape StringApp: Network analysis and visualization of proteomics data. *J Proteome Res* 18: 623–632. doi:10.1021/acs.jproteome.8b00702

Dove AE, Cook BL, Irgebay Z, Vecsey CG (2017) Mechanisms of sleep plasticity due to sexual experience in Drosophila melanogaster. *Physiol Behav* 180: 146–158. doi:10.1016/j.physbeh.2017.08.020

Droujinine IA, Perrimon N (2013) Defining the interorgan communication network: Systemic coordination of organismal cellular processes under homeostasis and localized stress. *Front Cell Infect Microbiol* 3: 82. doi:10.3389/fcimb.2013.00082

Droujinine IA, Perrimon N (2016) Interorgan communication pathways in physiology: Focus on Drosophila. *Annu Rev Genet* 50: 539–570. doi:10.1146/annurev-genet-121415-122024

Drummond-Barbosa D, Spradling AC (2001) Stem cells and their progeny respond to nutritional changes during Drosophila oogenesis. *Dev Biol* 231: 265–278. doi:10.1006/dbio.2000.0135

Ellis LL, Carney GE (2011) Socially-responsive gene expression in male Drosophila melanogaster is influenced by the sex of the interacting partner. *Genetics* 187: 157–169. doi:10.1534/genetics.110.122754

Ellis LL, Carney GE (2010) Mating alters gene expression patterns in *Drosophila melanogaster* male heads. *BMC Genomics* 11: 558. doi:10.1186/1471-2164-11-558

Fahrbach SE, Giray T, Robinson GE (1995) Volume changes in the mushroom bodies of adult honey bee queens. *Neurobiol Learn Mem* 63: 181–191. doi:10.1006/nlme.1995.1019

Fedorka KM, Linder JE, Winterhalter W, Promislow D (2007) Post-mating disparity between potential and realized immune response in *Drosophila melanogaster*. *Proc Biol Sci* 274: 1211–1217. doi:10.1098/rspb.2006.0394

Gapp K, Jawaid A, Sarkies P, Bohacek J, Pelczar P, Prados J, Farinelli L, Miska E, Mansuy IM (2014) Implication of sperm RNAs in transgenerational inheritance of the effects of early trauma in mice. *Nat Neurosci* 17: 667–669. doi:10.1038/nn.3695

Garbe DS, Vigderman AS, Moscato E, Dove AE, Vecsey CG, Kayser MS, Sehgal A (2016) Changes in female Drosophila sleep following mating are mediated by SPSN-SAG neurons. *J Biol Rhythms* 31: 551–567. doi:10.1177/0748730416668048

Ghiassian SD, Menche J, Barabási AL (2015) A DIseAse MOdule detection (DIAMOnD) algorithm derived from a systematic analysis of connectivity patterns of disease proteins in the human interactome. *PLoS Comput Biol* 11: e1004120. doi:10.1371/journal.pcbi.1004120

Glaab E, Baudot A, Krasnogor N, Schneider R, Valencia A (2012) EnrichNet: Network-based gene set enrichment analysis. *Bioinformatics* 28: i451–i457. doi:10.1093/bioinformatics/bts389

Gomulski LM, Dimopoulos G, Xi Z, Scolari F, Gabrieli P, Siciliano P, Clarke AR, Malacrida AR, Gasperi G (2012) Transcriptome profiling of sexual maturation and mating in the mediterranean fruit fly, Ceratitis capitata. *PLoS One* 7: e30857. doi:10.1371/journal.pone.0030857

Graveley BR, Brooks AN, Carlson JW, Duff MO, Landolin JM, Yang L, Artieri CG, van Baren MJ, Boley N, Booth BW, et al (2011) The developmental transcriptome of Drosophila melanogaster. *Nature* 471: 473–479. doi:10.1038/nature09715

Grover A, Leskovec J (2016) node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*. San Francisco, CA, USA: Association for Computing Machinery (ACM): 855–864.

Grozinger CM, Fan Y, Hoover SE, Winston ML (2007) Genome-wide analysis reveals differences in brain gene expression patterns associated with caste and reproductive status in honey bees (Apis mellifera). *Mol Ecol* 16: 4837–4848. doi:10.1111/j.1365-294X.2007.03545.x

Gu J, Chen Y, Li S, Li Y (2010) Identification of responsive gene modules by network-based gene clustering and extending: Application to inflammation and angiogenesis. *BMC Syst Biol* 4: 47. doi:10.1186/1752-0509-4-47

Heifetz Y, Lindner M, Garini Y, Wolfner MF (2014) Mating regulates neuromodulator ensembles at nerve termini innervating the Drosophila reproductive tract. *Curr Biol* 24: 731–737. doi:10.1016/j.cub.2014.02.042

Heifetz Y, Tram U, Wolfner MF (2001) Male contributions to egg production: The role of accessory gland products and sperm in Drosophila melanogaster. *Proc Biol Sci* 268: 175–180. doi:10.1098/rspb.2000.1347

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 U S A* 101: 6261–6266. doi:10.1073/pnas.0401337101

Heinze J, Schrempf A (2008) Aging and reproduction in social insects–a mini-review. *Gerontology* 54: 160–167. doi:10.1159/000122472

Ideker T, Ozier O, Schwikowski B, Siegel AF (2002) Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics* 18: S233–S240. doi:10.1093/bioinformatics/18.suppl_1.s233

Innocenti P, Morrow EH (2009) Immunogenic males: A genome-wide analysis of reproduction and the cost of mating in Drosophila melanogaster females. *J Evol Biol* 22: 964–973. doi:10.1111/j.1420-9101.2009.01708.x

Isaac RE, Li C, Leedale AE, Shirras AD (2010) Drosophila male sex peptide inhibits siesta sleep and promotes locomotor activity in the post-mated female. *Proc Biol Sci* 277: 65–70. doi:10.1098/rspb.2009.1236

Kapelnikov A, Zelinger E, Gottlieb Y, Rhrissorrakrai K, Gunsalus KC, Heifetz Y (2008a) Mating induces an immune response and developmental switch in the Drosophila oviduct. *Proc Natl Acad Sci U S A* 105: 13912–13917. doi:10.1073/pnas.0710997105

Kapelnikov A, Rivlin PK, Hoy RR, Heifetz Y (2008b) Tissue remodeling: A mating-induced differentiation program for the drosophila oviduct. *BMC Dev Biol* 8: 114. doi:10.1186/1471-213X-8-114

Kocher SD, Tarpy DR, Grozinger CM (2010) The effects of mating and instrumental insemination on queen honey bee flight behaviour and gene expression. *Insect Mol Biol* 19: 153–162. doi:10.1111/j.1365-2583.2009.00965.x

Kocher SD, Richard FJ, Tarpy DR, Grozinger CM (2009) Queen reproductive state modulates pheromone production and queen-worker interactions in honeybees. *Behav Ecol* 20: 1007–1014. doi:10.1093/beheco/arp090

Kocher SD, Richard FJ, Tarpy DR, Grozinger CM (2008) Genomic analysis of post-mating changes in the honey bee queen (Apis mellifera). *BMC genomics* 9: 232. doi:10.1186/1471-2164-9-232

Kubli E (2003) Sex-peptides: Seminal peptides of the Drosophila male. *Cell Mol Life Sci* 60: 1689–1704. doi:10.1007/s00018-003-3052

Lawniczak MK, Begun DJ (2004) A genome-wide analysis of courting and mating responses in Drosophila melanogaster females. *Genome* 47: 900–910. doi:10.1139/g04-050

Lebreton S, Grabe V, Omondi AB, Ignell R, Becher PG, Hansson BS, Sachse S, Witzgall P (2014) Love makes smell blind: Mating suppresses pheromone attraction in Drosophila females via Or65a olfactory neurons. *Sci Rep* 4: 7119–7126. doi:10.1038/srep07119

Leiserson MD, Vandin F, Wu HT, Dobson JR, Eldridge JV, Thomas JL, Papoutsaki A, Kim Y, Niu B, McLellan M, et al (2015) Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. *Nat Genet* 47: 106–114. doi:10.1038/ng.3168

Liu H, Kubli E (2003) Sex-peptide is the molecular basis of the sperm effect in Drosophila melanogaster. *Proc Natl Acad Sci U S A* 100: 9929–9933. doi:10.1073/pnas.1631700100

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15: 550–621. doi:10.1186/s13059-014-0550-8

Ma H, Schadt EE, Kaplan LM, Zhao H (2011) COSINE: COndition-SpecIfic sub-NEtwork identification using a global optimization method. *Bioinformatics* 27: 1290–1298. doi:10.1093/bioinformatics/btr136

Mack PD, Kapelnikov A, Heifetz Y, Bender M (2006) Mating-responsive genes in reproductive tissues of female drosophila melanogaster. *Proc Natl Acad Sci U S A* 103: 10358–10363. doi:10.1073/pnas.0604046103

McDonough-Goldstein CE, Borziak K, Pitnick S, Dorus S (2021) Drosophila female reproductive tract gene expression reveals coordinated mating responses and rapidly evolving tissue-specific genes. *G3 (Bethesda)* 11: jkab020. doi:10.1093/g3journal/jkab020

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. doi:10.1016/j.cub.2004.08.028

McGraw LA, Clark AG, Wolfner MF (2008) Post-mating gene expression profiles of female Drosophila melanogaster in response to time and to four male accessory gland proteins. *Genetics* 179: 1395–1408. doi:10.1534/genetics.108.086934

Manfredini F, Brown MJ, Vergoz V, Oldroyd BP (2015) RNA-sequencing elucidates the regulation of behavioural transitions associated with the mating process in honey bee queens. *BMC Genomics* 16: 563. doi:10.1186/s12864-015-1750-7

Miller CA, Settle SH, Sulman EP, Aldape KD, Milosavljevic A (2011) Discovering functional modules by identifying recurrent and mutually exclusive mutational patterns in tumors. *BMC Med Genomics* 4: 34. doi:10.1186/1755-8794-4-34

Nacu S, Critchley-Thorne R, Lee P, Holmes S (2007) Gene expression network analysis and applications to immunology. *Bioinformatics* 23: 850–858. doi:10.1093/bioinformatics/btm019

Newell NR, Ray S, Dalton JE, Fortier JC, Kao JY, Chang PL, Nuzhdin SV, Arbeitman MN (2020) The *Drosophila* post-mating response: Gene expression and behavioral changes reveal perdurance and variation in cross-tissue interactions. *G3 (Bethesda)* 10: 967–983. doi:10.1534/g3.119.400963

Nguyen H, Shrestha S, Tran D, Shafi A, Draghici S, Nguyen T (2019) A comprehensive survey of tools and software for active subnetwork identification. *Front Genet* 10: 155. doi:10.3389/fgene.2019.00155

Page RE, Peng CY (2001) Aging and development in social insects with emphasis on the honey bee, Apis mellifera L. *Exp Gerontol* 36: 695–711. doi:10.1016/s0531-5565(00)00236-9

Pal S, Wu J, Wu LP (2008) Microarray analyses reveal distinct roles for Rel proteins in the *Drosophila* immune response. *Dev Comp Immunol* 32: 50–60. doi:10.1016/j.dci.2007.04.001

Parkash R, Sharma V, Chahal J, Lambhod C, Kajla B (2011) Impact of body melanization on mating success in Drosophila melanogaster. *Drosophila Melanogaster Entomologia Experimentalis Applicata* 139: 47–59. doi:10.1111/j.1570-7458.2011.01102.x

Pasquier C, Guerlais V, Pallez D, Rapetti-Mauss R, Soriani O (2021) Identification of active modules in interaction networks using node2vec network embedding. *BioRxiv*: 2021.09.22.461345. doi:10.1101/2021.09.22.461345. (Preprint posted September 24, 2021).

Patil A, Nakai K (2014) TimeXNet: Identifying active gene sub-networks using time-course gene expression profiles. *BMC Syst Biol* 8: S2–S8. doi:10.1186/1752-0509-8-S4-S2

Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C (2017) Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* 14: 417–419. doi:10.1038/nmeth.4197

Peng J, Zipperlen P, Kubli E (2005) Drosophila sex-peptide stimulates female innate immune system after mating via the toll and imd pathways. *Curr Biol* 15: 1690–1694. doi:10.1016/j.cub.2005.08.048

Pérez-Silva JG, Araujo-Voces M, Quesada V (2018) nVenn: generalized, quasi-proportional Venn and Euler diagrams. *Bioinformatics* 34: 2322–2324. doi:10.1093/bioinformatics/bty109

Petrochilos D, Shojaie A, Gennari J, Abernethy N (2013) Using random walks to identify cancer-associated modules in expression data. *BioData Min* 6: 17–25. doi:10.1186/1756-0381-6-17

Prokupek AM, Kachman SD, Ladunga I, Harshman LG (2009) Transcriptional profiling of the sperm storage organs of drosophila melanogaster. *Insect Mol Biol* 18: 465–475. doi:10.1111/j.1365-2583.2009.00887.x

Qiu YQ, Zhang S, Zhang XS, Chen L (2010) Detecting disease associated modules and prioritizing active genes based on high throughput data. *BMC Bioinformatics* 11: 26. doi:10.1186/1471-2105-11-26

Rajan A, Perrimon N (2011) Drosophila as a model for interorgan communication: Lessons from studies on energy homeostasis. *Dev Cell* 21: 29–31. doi:10.1016/j.devcel.2011.06.034

Rapaport F, Zinovyev A, Dutreix M, Barillot E, Vert JP (2007) Classification of microarray data using gene networks. *BMC Bioinformatics* 8: 35. doi:10.1186/1471-2105-8-35

Rassoulzadegan M, Grandjean V, Gounon P, Vincent S, Gillot I, Cuzin F (2006) RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. *Nature* 441: 469–474. doi:10.1038/nature04674

Rezával C, Pavlou HJ, Dornan AJ, Chan YB, Kravitz EA, Goodwin SF (2012) Neural circuitry underlying Drosophila female postmating behavioral responses. *Curr Biol* 22: 1155–1165. doi:10.1016/j.cub.2012.04.062

Richard FJ, Tarpy DR, Grozinger CM (2007) Effects of insemination quantity on honey bee queen physiology. *PLoS One* 2: e980. doi:10.1371/journal.pone.0000980

Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: A bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26: 139–140. doi:10.1093/bioinformatics/btp616

Robinson S, Nevalainen J, Pinna G, Campalans A, Radicella JP, Guyon L (2017) Incorporating interaction networks into the determination of functionally related hit genes in genomic experiments with Markov random fields. *Bioinformatics* 33: i170–i179. doi:10.1093/bioinformatics/btx244

Rogers DW, Whitten MM, Thailayil J, Soichot J, Levashina EA, Catteruccia F (2008) Molecular and cellular components of the mating machinery in Anopheles gambiae females. *Proc Natl Acad Sci U S A* 105: 19390–19395. doi:10.1073/pnas.0809723105

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res* 13: 2498–2504. doi:10.1101/gr.1239303

Short SM, Lazzaro BP (2010) Female and male genetic contributions to post-mating immune defence in female Drosophila melanogaster. *Proc Biol Sci* 277: 3649–3657. doi:10.1098/rspb.2010.0937

Short SM, Lazzaro BP (2013) Reproductive status alters transcriptomic response to infection in female Drosophila melanogaster. *G3 (Bethesda)* 3: 827–840. doi:10.1534/g3.112.005306

Silberberg Y, Kupiec M, Sharan R (2017) GLADIATOR: A global approach for elucidating disease modules. *Genome Med* 9: 48. doi:10.1186/s13073-017-0435-z

Simola DF, Graham RJ, Brady CM, Enzmann BL, Desplan C, Ray A, Zwiebel LJ, Bonasio R, Reinberg D, Liebig J, et al (2016) Epigenetic (re)programming of caste-specific behavior in the ant Camponotus floridanus. *Science* 351: aac6633. doi:10.1126/science.aac6633

Skinner MK, Guerrero-Bosagna C, Haque MM (2015) Environmentally induced epigenetic transgenerational inheritance of sperm epimutations promote genetic mutations. *Epigenetics* 10: 762–771. doi:10.1080/15592294.2015.1062207

Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, et al (2017) The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res* 45: D362–D368. doi:10.1093/nar/gkw937

Swanson WJ (2003) Sex peptide and the sperm effect in Drosophila melanogaster. *Proc Natl Acad Sci U S A* 100: 9643–9644. doi:10.1073/pnas.1834127100

Tanay A, Sharan R, Kupiec M, Shamir R (2004) Revealing modularity and organization in the yeast molecular network by integrated analysis of highly heterogeneous genomewide data. *Proc Natl Acad Sci U S A* 101: 2981–2986. doi:10.1073/pnas.0308661100

Tarpy DR, Page RE (2000) No behavioral control over mating frequency in queen honey bees (Apis mellifera L.): Implications for the evolution of extreme polyandry. *Am Nat* 155: 820–827. doi:10.1086/303358

Toivonen JM, Partridge L (2009) Endocrine regulation of aging and reproduction in Drosophila. *Mol Cell Endocrinol* 299: 39–50. doi:10.1016/j.mce.2008.07.005

Ulitsky I, Shamir R (2007) Identification of functional modules using network topology and high-throughput data. *BMC Syst Biol* 1: 8–17. doi:10.1186/1752-0509-1-8

Ulitsky I, Shamir R (2009) Identifying functional modules using expression profiles and confidence-scored protein interactions. *Bioinformatics* 25: 1158–1164. doi:10.1093/bioinformatics/btp118

Vlaic S, Conrad T, Tokarski-Schnelle C, Gustafsson M, Dahmen U, Guthke R, Schuster S (2018) ModuleDiscoverer: Identification of regulatory modules in protein-protein interaction networks. *Sci Rep* 8: 433–511. doi:10.1038/s41598-017-18370-2

Walker SJ, Corrales-Carvajal VM, Ribeiro C (2015) Postmating circuitry modulates salt taste processing to increase reproductive output in Drosophila. *Curr Biol* 25: 2621–2630. doi:10.1016/j.cub.2015.08.043

Wang L, Tu Z, Sun F (2009) A network-based integrative approach to prioritize reliable hits from multiple genome-wide RNAi screens in Drosophila. *BMC Genomics* 10: 220–318. doi:10.1186/1471-2164-10-220

White MA, Bonfini A, Wolfner MF, Buchon N (2021) Drosophila melanogaster sex peptide regulates mated female midgut morphology and physiology. *Proc Natl Acad Sci U S A* 118: e2018112118. doi:10.1073/pnas.2018112118

Whitfield CW, Cziko AM, Robinson GE (2003) Gene expression profiles in the brain predict behavior in individual honey bees. *Science* 302: 296–299. doi:10.1126/science.1086807

Wigby S, Slack C, Grönke S, Martinez P, Calboli FC, Chapman T, Partridge L (2011) Insulin signalling regulates remating in female Drosophila. *Proc Biol Sci* 278: 424–431. doi:10.1098/rspb.2010.1390

Wigby S, Chapman T (2005) Sex peptide causes mating costs in female Drosophila melanogaster. *Curr Biol* 15: 316–321. doi:10.1016/j.cub.2005.01.051

Woyke J (1964) Causes of repeated mating flights by queen honeybees. *J Apicultural Res* 3: 17–23. doi:10.1080/00218839.1964.11100077

Yapici N, Kim YJ, Ribeiro C, Dickson BJ (2008) A receptor that mediates the post-mating switch in Drosophila reproductive behaviour. *Nature* 451: 33–37. doi:10.1038/nature06483

Zhou S, Mackay T, Anholt RR (2014) Transcriptional and epigenetic responses to mating and aging in *Drosophila melanogaster*. *BMC Genomics* 15: 927. doi:10.1186/1471-2164-15-927