Neurogenomic Signatures of Successes and Failures in Life-History Transitions in a Key Insect Pollinator

Fabio Manfredini[1,2,*], Alfonso E. Romero[2], Inti Pedroso[3], Alberto Paccanaro[2], Seirian Sumner[4,5,†], and Mark J.F. Brown[1,†]

[1]School of Biological Sciences, Royal Holloway University of London, Egham, United Kingdom

[2]Department of Computer Science, and Centre for Systems and Synthetic Biology, Royal Holloway University of London, Egham, United Kingdom

[3]Center for Systems Biotechnology, Fraunhofer Chile Research Foundation, Santiago, Chile

[4]School of Biological Sciences, University of Bristol, United Kingdom

[5]Present address: Centre for Biodiversity & Environment Research, Department of Genetics, Evolution & Environment, University College London, London, United Kingdom

[†]These authors contributed equally to this work.

*Corresponding author: E-mail: fmanfredini79@gmail.com.

Abstract

Life-history transitions require major reprogramming at the behavioral and physiological level. Mating and reproductive maturation are known to trigger changes in gene transcription in reproductive tissues in a wide range of organisms, but we understand little about the molecular consequences of a *failure* to mate or become reproductively mature, and it is not clear to what extent these processes trigger neural as well as physiological changes. In this study, we examined the molecular processes underpinning the behavioral changes that accompany the major life-history transitions in a key pollinator, the bumblebee *Bombus terrestris*. We compared neuro-transcription in queens that succeeded or failed in switching from virgin and immature states, to mated and reproductively mature states. Both successes and failures were associated with distinct molecular profiles, illustrating how development during adulthood triggers distinct molecular profiles within a single caste of a eusocial insect. Failures in both mating and reproductive maturation were explained by a general up-regulation of brain gene transcription. We identified 21 genes that were highly connected in a gene coexpression network analysis: nine genes are involved in neural processes and four are regulators of gene expression. This suggests that negotiating life-history transitions involves significant neural processing and reprogramming, and not just changes in physiology. These findings provide novel insights into basic life-history transitions of an insect. Failure to mate or to become reproductively mature is an overlooked component of variation in natural systems, despite its prevalence in many sexually reproducing organisms, and deserves deeper investigation in the future.

Key words: bumblebee, *Bombus terrestris*, mating, reproductive maturation, brain, gene network.

Introduction

Successful negotiation of key life-history transitions is essential for individuals to pass on genes to the next generation. Life-history transitions (e.g., reproductive maturation, initiation of foraging, and seasonal migrations, Lutterschmidt and Maine 2014) are characterized by distinct switches in the behavioral and physiological traits of an individual in response to ontogenetic and/or environmental cues. Neurogenomic analyses (whole genome analyses of brain gene expression) have revealed the molecular changes that occur when individuals successfully mate or attain reproductive maturity, typically by analyzing successful phenotypes at different time points or by comparing the successful phenotype of interest to the previous stage of development, like mated individuals versus virgin

(Kocher et al. 2008; Dalton et al. 2010). However, we do not know what molecular processes are triggered at the genomic level when organisms are given the chance to achieve the same important transitions but fail to do so. Only a few studies have specifically characterized the global consequences of a failed biological transition (Almansa et al. 2015; Engelstädter 2015; Li et al. 2015). This is an important step toward understanding the molecular implications of life-history failures, which may include, for example, determining the role of environmental or developmental perturbations in failure (Wilburn and Swanson 2016). Achieving a mechanistic understanding of life-history failure is therefore fundamental, but also timely, especially for species that provide important ecosystem services, such as pollinating insects.

Neurogenomic methods allow the characterization and comparisons of pathways of gene activity in neural tissues across phenotypic states, offering a functional understanding of how ontogenetic switches are regulated at the molecular level (Harris and Hofmann 2014). Important behavioral transitions for more stable phenotypes (e.g., behavioral maturation or foraging in honey bees, see Zayed and Robinson 2012) are typically associated with significant changes in neurogenomic signatures (i.e., numbers of differentially expressed genes and overrepresentation of biological functions), but we do not know whether failing to accomplish the same transitions results in equally large effects. Studies in a range of organisms have shown that failing to win social interactions is associated with distinct neurogenomic states that contrast with the outcome of a successful interaction, like, for example, cichlid fish (Maruska 2014), zebrafish (Oliveira et al. 2016), and social wasps (Toth et al. 2014). Similar analyses applied within the context of fundamental life-history transitions like mating and reproductive activation can address questions such as whether failed outcomes trigger specific molecular processes, to what extent they do so (i.e., magnitude of the effect) and what processes are activated or suppressed.

Bumblebees provide tractable models for investigating the molecular mechanisms that regulate both successes and failures in life-history transitions. *Bombus terrestris* is well characterized at the molecular level thanks to the development of genomic resources (Colgan et al. 2011; Barribeau et al. 2015; Sadd et al. 2015), and this insect displays a complex social life (see Amsalem et al. 2015) that can be easily observed in the field or successfully reproduced in the lab. However, *B. terrestris* is subjected to many possibilities of failure both in the wild and in artificial rearing conditions. *Bombus terrestris* queens mate once when they are a few days old. Shortly after mating, queens enter diapause for the winter; in the spring they emerge from diapause, found a nest, lay eggs, and rear the first generation of workers (reviewed in Goulson 2010). Bumblebee queens often fail to accomplish these key transitions, making them biologically relevant models for understanding failures in life history. For example, in populations

of *B. pratorum*, 38% of queens fail to become reproductives (Rutrecht and Brown 2008); in captive *B. terrestris* up to 60% of queens fail to mate (Imran et al. 2015), and up to 65% fail to become mature reproductives (Karsli and Gurel 2013). Understanding the molecular mechanisms associated with these failures will provide valuable knowledge on the general biology of an organism that provides important pollination services (Kleijn et al. 2015).

Here, we sequence brain transcriptomes from queens of *B. terrestris* that failed to complete two key life-history transitions, that is, mating and reproductive activation ("Failed Mated" and "Failed Reproductive," respectively, fig. 1). To investigate how failure to mate or to become reproductively mature shapes the neurogenomic profile of bumblebee queens, we compare failed phenotypes to their successful counterparts (i.e., "Successfully Mated" and "Successfully Reproductive," respectively, see Materials and Methods): these represent the currently available and appropriate control groups, as successful queens shared the same age, the same social environment, the same previous life-history experiences and the same rearing conditions as failed queens. We focus on brain tissue because we are interested in behavioral transitions: by restricting our investigation to the organ that is the major regulator of behavior in animals, we increase the chances of detecting even subtle differences in the expression of genes that play a major role in behavioral performance. Furthermore, the neurogenomic approach is extremely powerful: a high proportion of genes in the genome are expressed in the brain (Lein et al. 2007) and the brain usually has the most diverse population of RNA compared with other tissues (Naumova et al. 2013). First, we characterize neurogenomic processes associated with the four different phenotypes at multiple molecular levels, that is, gene expression patterns, enrichment of molecular functionality, and gene coexpression network (AIM 1). Then, we analyze the molecular processes associated with the successful transition from mated to reproductive queens (AIM 2), which represents the baseline successful transition across the two life-history stages, and we compare our results to previously published studies on other insects addressing a similar question. Finally, we characterize the molecular patterns associated with queens that have failed to mate (AIM 3) or become reproductively mature (AIM 4). These data allow us to test three fundamental hypotheses on the molecular basis of life-history transitions. Firstly, each bumblebee queen phenotype has a unique neurogenomic profile (Hypothesis A); secondly, the successful transition from mating to reproductive maturation is associated with specific neurogenomic signatures that are conserved across organisms (Hypothesis B); and thirdly, both failed groups of queens have distinct neurogenomic profiles compared with their successful counterparts (Hypothesis C). The genes and molecular pathways associated with failed, rather than successful, mating and reproductive maturation will provide
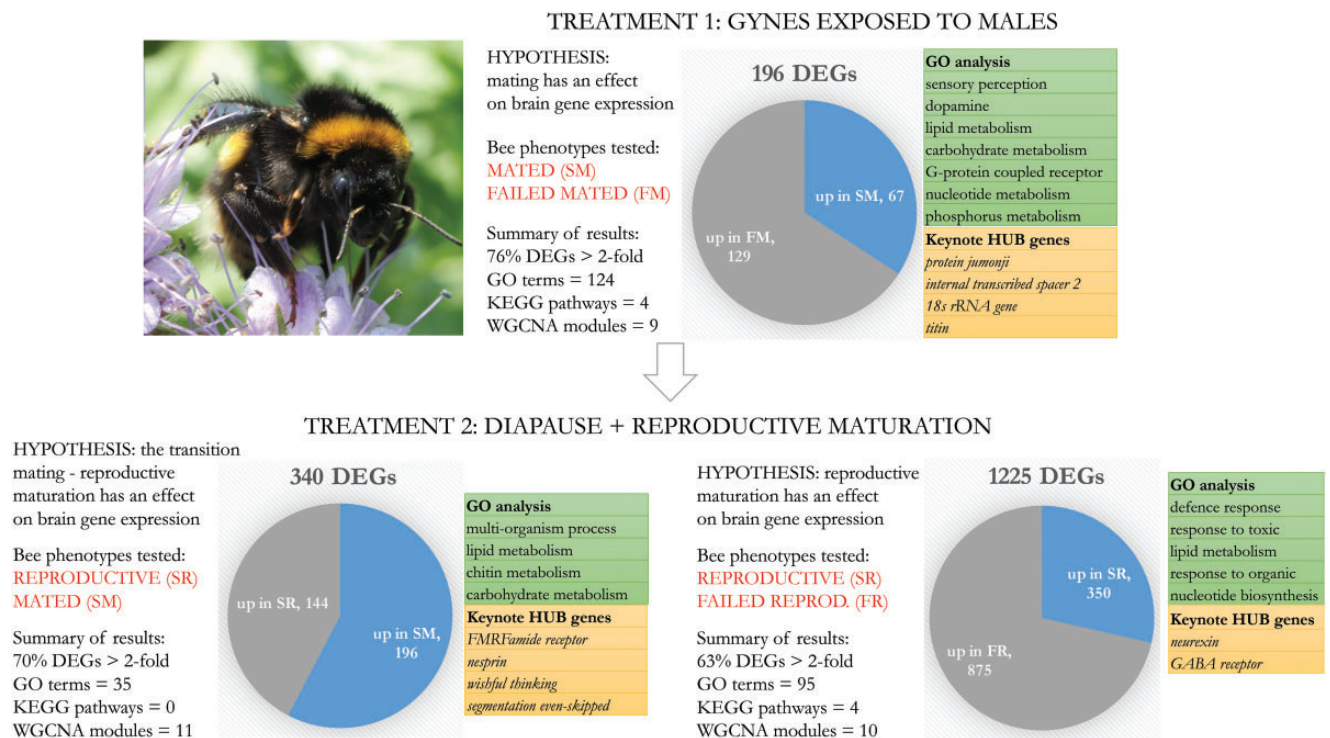
Fig. 1.—Neurogenomics of bumblebee life-history transitions. This diagram shows the experimental design for this study and the main findings. For each pairwise comparison of interest, we report the numbers of differentially expressed genes (DEGs), the proportion of DEGs that were above the 2-fold expression threshold, and the GO terms, the KEGG pathways and the WGCNA subnetworks (modules) that were significantly associated with that comparison. Inset (top left): queen of the buff-tailed bumblebee *Bombus terrestris* foraging on *Phacelia* blossoms (by Holger Casselmann, Wikimedia Commons).

essential insights into the mechanisms limiting (and promoting) these two key life-history transitions.

Materials and Methods

Preparation of Samples

Gynes (virgin queens) of *Bombus terrestris* were obtained from four commercial colonies of similar genetic background and reared following standardized conditions (Koppert BV, The Netherlands). These gynes were allowed to mate when they were 5 days old and were sampled 3 days after mating ("Successfully Mated" treatment or SM). Mating sessions lasted for 30–40 min and happened within a large cage with full visibility from each side. Bees were monitored for the full duration of the session and mating couples were removed as soon as we noticed them. Males and females remain attached with their bodies during sperm transfer for between 15 and 75 min (the majority 30–40 min), with sperm transfer taking place in the first few minutes (Duvoisin et al. 1999), hence we are 100% confident that all queens were correctly allocated to the SM group. The Failed Mated queen group (FM), instead, was obtained by handling the queens in the same way as in SM and sampling those who failed to mate. Hence, FM and SM groups were of comparable age and they were exposed to the same conditions of physical and social environments, making SM a suitable control group to

investigate failure during mating. A subset of mated queens was hibernated for a period of 8 weeks and sampled 1 month after emerging from hibernation. Queens showing fully developed ovaries with visible mature eggs were defined as "Successfully Reproductive" (SR, 45% of all queens successfully emerged from hibernation), whereas queens showing undeveloped ovaries were defined as "Failed Reproductive" (FR, 55%). As with the two previous groups, FR and SR were directly comparable for age and the environment they experienced, hence SR was a well-suited control group to investigate failure during reproductive maturation. For a detailed description of queen rearing, see sections "a–b–c" in supplementary methods in Supplementary Material online.

We dissected brains from focal bees and isolated total RNA from individual brains (see section "d" in supplementary methods in Supplementary Material online for a full description of how these steps were achieved). RNA from brain samples was used to perform an RNAseq experiment. Samples that provided the highest amount of RNA of good quality were used for RNAseq (between 5.83 and 19.5 µg of total RNA, RIN scores between 5 and 9.3, median 7.7). We included the four treatment groups described earlier in our sequencing experiment, with queen samples for each treatment coming from three different colonies, so that colony of origin was a random factor in the experimental design (see "bee_samples" in supplementary table S2, Supplementary Material online). The four

groups are: A) Successfully Mated ($N = 8$); B) Failed Mated ($N = 9$); C) Successfully Reproductive ($N = 8$); D) Failed Reproductive ($N = 8$). These samples were arranged in five lanes of an Illumina HiSeq 2000 Sequencing System to produce 90-bp paired-reads by means of TruSeq mRNA sequencing (Beijing Genomics Institute, China). Raw reads were preprocessed at BGI: this included removing adapters, quality control, and filtering out low-quality sequences—Q20% >98 and GC(%) >39. Clean reads were aligned with TopHat for Illumina using default settings (Trapnell et al. 2012) on the Galaxy web-based platform (https://usegalaxy.org/; last accessed October 27, 2017) to the latest version of the bumblebee genome including 10,673 predicted genes (*Bter 1.1*, Sadd et al. 2015). Mapped reads were converted into raw read counts with SAMtools idxstats (Li et al. 2009) and these were used to quantify differential gene expression. Only genes with at least ten reads per sample were kept for the analysis of gene expression (7,724 genes, 72% of the total).

Analysis of Gene Expression

To analyze global patterns of brain gene expression, we used hierarchical clustering (Ward method) and principal component analysis in JMP Pro 10.0 (SAS, Cary, NC). For more detailed analyses of gene expression, we imported raw sequence data into R and processed them with the edgeR package (v3.6.0) from Bioconductor (Robinson et al. 2010), following two separate approaches. First, we applied a glmLRT (Genewise Negative Binomial Generalized Linear Model) to the count data and identified differentially expressed genes using planned linear contrasts (table 1), as described in Mikheyev and Linksvayer 2015. Second, we performed pairwise comparisons to identify genes that differed between groups. Here, we focused on the three comparisons that were more useful to understand the difference between behavioral states or transitions: SM versus FM, SR versus FR, and SR versus SM (fig. 1). For this analysis, we used a modified Fisher's exact test that takes into account both dispersion and multiple samples, as described in Manfredini et al. (2015). Results of gene expression analyses were corrected for multiple testing (FDR, threshold = 0.05) using the Benjamini–Hochberg method (Benjamini and Hochberg 1995). For both analyses, raw sequence data were normalized using the default method for edgeR that produces trimmed mean of M values (TMM) between each pair of samples.

We used the output of the second set of gene expression analyses to identify enriched biological processes (GO terms, see "GO" in supplementary table S8, Supplementary Material online) and metabolic pathways (KEGG pathways, see "KEGG" in supplementary table S10, Supplementary Material online) by means of overrepresentation analyses ($P < 0.05$, Benjamini–Hochberg correction for multiple testing). For this set of analyses, we matched *B. terrestris* sequences with sequences from three reference organisms: the fruit

Table 1

Summary Table for glmLRT Analysis of Traits of Interest

Trait of Interest	Contrast	Genes	Up	Down
Mating: success	SM versus (FM+SR+FR)	33	21	12
Mating: failure	FM versus (SM+SR+FR)	68	58	10
Reproduction: success	SR versus (FR+SM+FM)	266	105	161
Reproduction: failure	FR versus (SR+SM+FM)	1,578	943	635
Success/Failure	(SM+SR) versus (FM+FR)	409	94	315
Age	(SM+FM) versus (SR+FR)	204	104	100

Note.—Numbers of up and down-regulated genes that are significantly different at $P < 0.05$ for each contrast are given (after Benjamini–Hochberg correction for multiple testing). The difference in the number of genes regulated in successful versus failed queens was statistical significant for both mating and reproductive maturation: mating (Fisher's exact test with odds ratio: 0.30 [0.10–0.89], $P = 0.02$); reproductive maturation (Fisher's exact test with odds ratio: 0.43 [0.33–0.57], $P = 8.5e$-10).

fly *Drosophila melanogaster*, as this is the globally recognized model for all insect transcriptomic studies, and two social insects, the honey bee *Apis mellifera* and the carpenter ant *Camponotus floridanus*. Honey bees and carpenter ants were chosen as they display similar behaviors to *B. terrestris* and therefore could provide a higher coverage for gene function prediction of behavior-related genes. In particular, queens of these social insects perform mating, founding, and reproductive activation in a similar fashion to bumblebee queens (but they lack diapause, as they overwinter with other colony members in an active state). Full details on the protocol that we used for overrepresentation analyses can be found in section "e" in supplementary methods and "BLAST" in supplementary table S7, Supplementary Material online.

We used Venny (http://bioinfogp.cnb.csic.es/tools/venny/index.html; last accessed October 27, 2017) to overlap lists of genes or GO terms and identify elements that were in common between pairwise comparisons or unique, and we used REVIGO to classify lists of GO terms in a hierarchical fashion (Supek et al. 2011). We also used Venny to compare our lists of GO terms to the lists identified in other studies that analyzed the transcriptomic basis for mating, reproductive maturation, and ageing in bumblebees or other insects (see "Comparative_studies" in supplementary table S11, Supplementary Material online, for details). We adopted a Hypergeometric test to identify significant overlaps between lists of GO terms (threshold = 0.05).

Network Analyses

To elucidate the transcriptomic organization in the brain of bumblebee queens, we performed weighted gene coexpression network analysis (WGCNA). This approach identifies sets of coregulated genes that share similar expression profiles and groups them in clusters or modules (Langfelder and Horvath 2008). Genes in the same modules (here called subnetworks for simplicity) show similar responses to analogous changes in behavioral or physiological conditions and therefore are assumed to play a similar role in a particular biological function. The WGCNA approach has been used to

describe the functional arrangement of gene networks in different organisms, from humans to social insects (Oldham et al. 2008; Mikheyev and Linksvayer 2015; Patalano et al. 2015; Morandin et al. 2016), thus complementing the quantification of gene expression. The bumblebee gene coexpression network was built using the WGCNA standard protocol, with a few minor modifications (see section "g" in supplementary methods and supplementary figures, Supplementary Material online). We used VisANT (Hu et al. 2013) to visualize subnetworks, reveal their structure, and also to identify one or two "hub" genes within each subnetwork (see section "g" in supplementary methods, Supplementary Material online).

We conducted an analysis of the proportions of differentially expressed genes (DEGs hereafter) within individual subnetworks, as this can provide useful insights about the subnetwork's structure. In fact, DEGs that are highly connected with a subnetwork (positively correlated) potentially drive the patterns of expression for that subnetwork and, if the subnetwork is significantly associated with a phenotypic trait, those DEGs are very likely to be highly relevant genes regulating the expression of the trait. We tested for nonrandom distribution of DEGs across subnetworks by determining whether the proportion of DEGs is drawn from the same binomial distribution or whether DEGs are clustered within subnetworks. To do this test, we fitted two general linear models in R following the same approach as in Patalano et al. 2015: one model with a single parameter (the global proportion of DEGs across all subnetworks) and the other (the saturated model) with a separate parameter for each of the 33 subnetworks, that is, one parameter per subnetwork indicating the proportions of DEGs within the individual subnetwork. We used the output of the GLM analyses to compare the best fit of the two models with an Analysis of Deviance for Generalized Linear Model Fits. Furthermore, we performed an analysis on the whole set of genes to measure the correlation between expression levels quantified as normalized read counts (the output of the edgeR analysis) and connectivity by performing a Spearman's rank correlation test. Finally, we performed a series of analyses to better characterize individual subnetworks: 1) DEGs enrichment analysis (threshold = 2e-3 after Bonferroni correction for multiple testing), 2) overrepresentation analysis (threshold = 0.01 after Benjamini–Hochberg correction), and 3) subnetwork–trait association analysis (threshold for significant correlation = 0.05 after Benjamini–Hochberg correction). The output of these analyses is summarized in table 2, whereas a full description of how these analyses were achieved is contained in section "g" in supplementary methods in Supplementary Material online. Different thresholds for statistical significance were chosen across this set of analyses in order to obtain meaningful significant elements: when the output of significant elements was too large, we chose a more conservative significance threshold in order to minimize false positives.

Results

Global Patterns of Gene Expression Reveal Distinct Molecular Differentiation Associated with Phenotype (Aim 1)

On an average, 53 million reads were generated per sample (min. 45,703,860 and max. 65,410,272 clean reads after filtering), achieving 10× coverage. Between 92% and 93% of clean reads per sample were aligned to single locations in the bumblebee genome ("RNAseq_stats" in supplementary table S3, Supplementary Material online).

The four bee phenotypes (i.e., Successfully Mated, Failed Mated, Successfully Reproductive, and Failed Reproductive) exhibited distinct molecular signatures in brain gene expression, gene functionality, and gene network. This is apparent from three different analyses of shared levels of gene expression among phenotypes: 1) Genewise Negative Binomial Generalized Linear Model (glmLRT) to examine which of the four queen phenotypes had distinct expression profiles versus all of the others, and which phenotype explained most of the differences in global gene expression; 2) hierarchical clustering (HC); and 3) principal component analysis (PCA). Furthermore, we performed pairwise comparisons across phenotypes (analysis with edgeR) and we detected 1,441 DEGs between any two queen phenotypes: this represents 18.66% of the total genes analyzed.

"Failed Reproductives" were the most distinct phenotype among the set of phenotypes that we compared in this study: this phenotype explained the highest amount of variance in gene expression, with 1,578 genes differentially expressed in these bees compared with the other phenotypes (glmLRT, $P < 0.05$, table 1). Failed reproductives were the outgroup in the HC analysis (fig. 2) and explained 39.7% of the differences in the PCA (fig. 3). This pattern was confirmed at the gene coexpression network level. Failed Reproductives were associated with two network modules or subnetworks (brown $r = -0.54$, $P = 0.04$; and turquoise $r = 0.55$, $P = 0.04$), that showed significant enrichment for DEGs (brown $R = -2$, $P < 2.8e-13$; and turquoise $R = 1.2$, $P = 1.7e-13$) and were both overrepresented in two key pairwise comparisons (table 2). "Failed Mated" and "Successfully Mated" were the most similar phenotypes (fig. 2): these phenotypes had the smallest numbers of differentially expressed genes (68 and 33, respectively, table 1), and explained the least variation (29% in the PCA, fig. 3).

Each queen phenotype was characterized by a unique set of gene coexpression subnetworks. We detected 33 subnetworks with more than 10 coexpressed genes (fig. 4 and "WGCNA" in supplementary table S13, Supplementary Material online). The average number of genes per subnetwork was 203 (range 14–3,835, SD = 666). Out of the 6,706 genes that were included in the network analysis, 1,329 were DEGs in at least one of the pairwise comparisons analyzed with edgeR. DEGs were nonrandomly distributed across

Table 2

Characterization of Gene Subnetworks Obtained with the WGCNA Approach

Subnetwork (gene number)	i) DEGs Enrichment		ii) Overrepresentation Analysis			iii) Subnetwork–trait Association		
	R-factor	Significance	SR/SM	SM/FM	SR/FR	Trait	*R*	Significance
black (99)	0.9	0.3		2.2e-03				
blue (1,002)	0.2	9.8e-58*	1.0e-06	7.7e-04	3.6e-26			
brown (224)	2	2.8e-13*	3.0e-05		3.5e-28	FR	−0.54	0.04
cyan (40)	3.9	5.0e-15*	1.5e-10	6.0e-03	2.6e-19			
green (123)	2.1	7.9e-09*	6.1e-22	2.2e-18	4.7e-05			
greenyellow (47)	1.8	7.0e-03	7.7e-12					
grey (389)	0.6	5.9e-06*			0.01	Age	0.54	0.02
grey60 (32)	2.2	2.0e-03	5.0e-05	2.1e-05	8.1e-03			
lightcyan (34)	1.6	0.05	4.3e-09			SM	−0.55	0.04
						SR	0.59	0.04
						Age	0.7	4.0e-04
magenta (67)	1.7	8.0e-03		4.4e-14				
paleturquoise (15)	1	0.4	0.01					
purple (52)	0.1	1.3e-04*			3.4e-03	SM	−0.53	0.04
red (111)	1	0.4		8.9e-15	2.5e-04			
royalblue (27)	1.5	0.14		1.3e-03				
saddlebrown (17)	1.8	0.1	1.7e-03					
tan (47)	1.1	0.45	3.5e-03					
turquoise (3,835)	1.2	1.7e-13*	1.5e-03		9.7e-128	FR	0.55	0.04
yellow (133)	1.3	0.04		9.4e-07	3.5e-03			

Note.—The numbers of genes within each subnetwork are in brackets. i) Enrichment of DEGs within the subnetwork: significantly enriched subnetworks (threshold = 2e-3 after Bonferroni correction for multiple testing) are indicated with an asterisk; the representation factor "R-factor" indicates positive enrichment (more genes than expected by chance) when > 1. ii) Overrepresentation analysis indicating the subnetworks that were significantly associated with each pairwise comparison (threshold = 0.01, after Benjamini–Hochberg correction). iii) Association between subnetwork and phenotypic trait: the correlation factor "*r*" shows the direction of the expression for the subnetwork (a positive value indicates higher expression while a negative value indicates lower expression); only significant associations are reported (threshold = 0.05, after Benjamini–Hochberg correction).

subnetworks (GLM × 2[32] = 171.59; *P* < 2.2e-16, "WGCNA" in supplementary table S13, Supplementary Material online); they were clustered in four subnetworks that contained more DEGs than expected by chance (Hypergeometric test, *P* < 0.05, see table 2).

We identified four sets of highly connected genes ("hub" genes) in the subnetworks that were significantly associated with traits of interest ("hub-genes" in supplementary table S14 and supplementary networks, Supplementary Material online). Hub genes were also likely to be highly expressed (Spearman's rank correlation, *r* = 0.57, *P* < 2.2e-16). Being the most connected genes and also highly expressed, "hub" genes are likely to play a key role in the regulation of biological functions and therefore deserve special attention. The first set of hub genes that we identified are important regulators of neurogenesis: *Peroxidasin*, part of a family of genes characterized by leucine-rich repeats and usually involved in protein–protein interactions across many different processes, including neuronal development (Soudi et al. 2012), *Nesprin*, a regulator of motor neuron innervation (Morel et al. 2014), and *segmentation even-skipped*, controlling neuronal fate (Doe et al. 1988). A second set of hub genes are involved in synapses and synaptogenesis: *neurexin*, previously characterized in the honey bee brain (Biswas et al. 2010), *BAI1* (Cork and Van Meir 2011) and *wishful thinking*, a regulator of neuromuscular synaptic transmission (Marqués et al.

2003). Additional genes associated with neural processes were *GABA receptor*, a neurotransmitter involved in learning processes in *Drosophila* and *Apis mellifera* (Liu et al. 2007; Liang et al. 2014) and the neuropeptide *FMRFamide receptor* (Walker et al. 2009). A third set of hub genes are likely to be involved in core biological functions of relevance to insect life-history: *Fatty acyl-CoA reductase* for pheromone synthesis (Teerawanichpan et al. 2010), *Jumonji* for olfactory learning (Walkinshaw et al. 2015), *GATA zinc finger domain-containing 7* for haematopoiesis (Waltzer et al. 2002), and *Titin* for muscle development (Ma et al. 2006). Finally, we identified hub genes that potentially play an important role as regulators of gene expression. These are the transposable element *botmar-15 transposon mariner*, previously characterized in *Bombus terrestris* (Rouleux-Bonnin et al. 2005), and the *histone-lysine N-methyltransferase Su(var)3-9*, involved in DNA methylation (Li et al. 2016).

Significant Shifts in Neurogenomic Signatures in the Transition from Mated to Reproductively Active Queens (Aim 2)
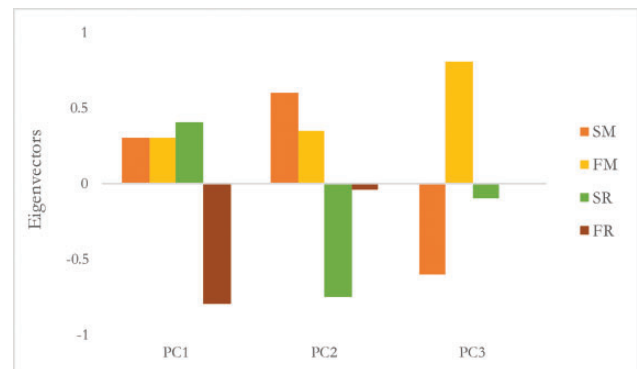
Marginally, more genes were down-regulated in Successfully Reproductive queens relative to Successfully Mated queens. We found 340 DEGs between the two phenotypes (4.4% of all genes analyzed): 196 were down-regulated in Successfully
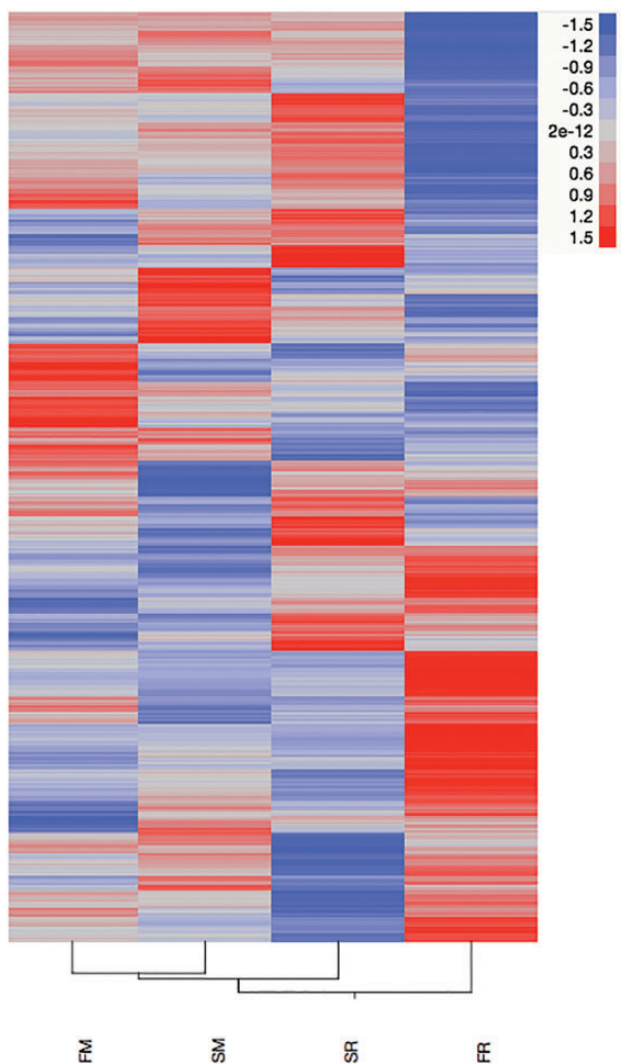
FIG. 2.—Global patterns of gene expression. The Hierarchical Clustering analysis (by treatment groups) shows the expression patterns of 10,673 genes that resulted from mapping RNAseq reads to the *Bombus terrestris* genome. Successfully and Failed Mated phenotypes (SM and FM, respectively) are the most similar groups; Failed Reproductive (FR) is the most distinct. The heatmap is color-coded, with genes that are highly expressed in red and genes that are expressed at lower levels in blue (see legend for conversion of the color intensity to normalized averaged read counts).

FIG. 3.—Components of global gene expression. Principal Component Analysis was performed on the 10,673 genes that resulted from mapping RNAseq reads to the *Bombus terrestris* genome. The three components represent the proportions of differentially regulated genes associated with Failed Reproductive queens (FR, PC1 = 39.7%), Successfully Reproductive queens (SR, PC2 = 31.3%), and Successfully versus Failed queens at mating (SM vs. FM, PC3 = 29%).

including functions associated with the response to stimulus and defence response; metabolism of lipids and hormones (6 elements); chitin metabolic process (7 elements); and the metabolism of carbohydrates (5 elements). However, we detected no significant overrepresentation of specific KEGG pathways among the DEGs.

Eleven networks were overrepresented in the comparison between Successfully Reproductive versus Successfully Mated queens, and five of these were uniquely overrepresented in this comparison (table 2). The whole set of overrepresented subnetworks included 13 hub genes, of which 7 were differentially expressed in this contrast (see, also, "hub_genes" in supplementary table S14, Supplementary Material online). These are: the transposable element *piggyBac*, and the previously described *wishful thinking*, *segmentation even-skipped*, *botmar-15 transposon mariner*, and *FMRFamide receptor*.

The design of this experimental comparison allowed us to simultaneously control for age effects (Successfully Reproductive queens were consistently older than Successfully Mated), hence we investigated the possibility that the difference between the two queen phenotypes was due to age more than to the transition of interest. The outcome of our ageing studies (see section "f" in supplementary methods, Supplementary Material online, for details) included 216 DEGs that were unique to the Successfully Reproductive/Successfully Mated comparison and therefore not age-related. This confirms that age was not the major factor influencing the expression patterns in this behavioral transition. Our comparative studies support these findings. Older mated bumblebee queens shared more GO terms with older queens of the ant *Cardiocondyla obscurior* that had mated (Von Wyschetzki et al. 2015), compared with old virgins (10 vs. 5, "Comparative_studies" in supplementary

Reproductive queens, whereas 144 were up-regulated in this group (χ^2 test from equal: $\chi^2 = 3.99$, df = 1, $P = 0.04$; fig. 1). Furthermore, 238 genes showed >2-fold changes in expression levels, and 71 of these showed >4-fold changes ("edgeR_SR-SM" in supplementary table S4, Supplementary Material online). Across all 340 DEGs, 35 Gene Ontology (GO) terms were significantly enriched ($P < 0.05$), and could be clustered into four major groups of related GO terms ("REVIGO" in supplementary table S9, Supplementary Material online): multiorganism processes (9 elements)
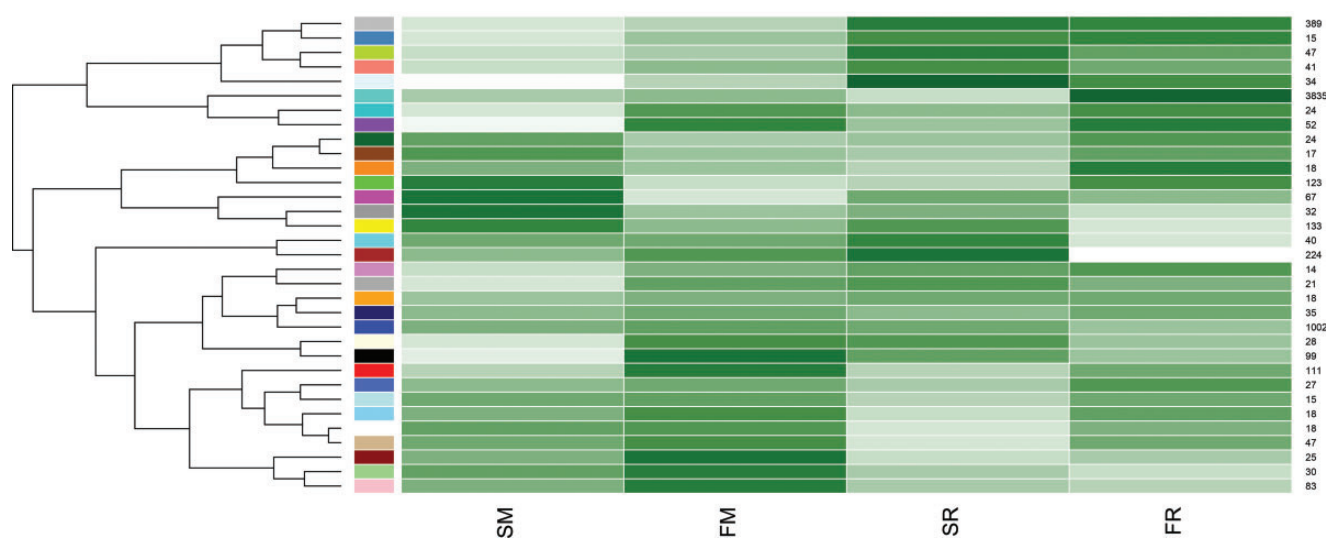
Fig. 4.—Weighted gene coexpression network analysis. WGCNA identified 33 subnetworks (color coded) of coexpressed genes that clustered according to expression profiles in the four queen phenotypes (from low expression = white, to high expression = dark green). Subnetwork size is indicated by the number of genes reported on the right-hand side of the figure.

table S11, Supplementary Material online), suggesting that mating was a more important factor in our analysis than age. Nevertheless, we were able to detect a set of 34 DEGs that were consistently age-related across comparisons ("age_studies" in supplementary table S12, Supplementary Material online), 2 subnetworks positively correlated with age (see table 2 and fig. 5), and 3 hub genes that significantly associated with age (*neurexin*, *FMRFamide receptor*, and *nesprin*; "hub_genes" in supplementary table S14, Supplementary Material online).

Neurogenomic Signatures of Success and Failure in Mating (Aim 3)

We found relatively small differences in gene expression between Successfully Mated and Failed Mated phenotypes (fig. 1). Both phenotypes were exposed to males and experienced courtship, hence our analysis enabled us to isolate those responses specifically linked to postmating changes in the brain or to the lack of it. There were 196 DEGs (2.5% of all genes analyzed FDR < 0.05): 149 DEGs showed >2-fold changes in expression levels, and out of these 24 showed >4-fold changes ("edgeR_SM-FM" in supplementary table S5, Supplementary Material online). Interestingly, Failed Mated queens were characterized by a general up-regulation of transcription: they had 1.9 times the number of up-regulated genes compared with down-regulated genes (129 vs. 67, χ^2 test from equal: $\chi^2 = 10.05$, df = 1, $P = 1.5e$-3).

We found significant differences in biological functionality between Successfully Mated and Failed Mated phenotypes. Of the 196 DEGs, 124 GO terms were enriched. The GO terms clustered into seven related groups ("REVIGO" in supplementary table S9, Supplementary Material online): sensory

perception (30 elements) including several associated with learning, visual behavior, response to light and chemical stimulus, and taxis; dopamine receptor signaling pathway (4 elements); metabolism of lipids (9 elements) including several related to hormone metabolism; metabolism of carbohydrates (21 elements); G-protein coupled receptor signaling pathway (6 elements); regulation of nucleotide metabolic process (20 elements); regulation of phosphorous metabolic process (4 elements). Analysis of KEGG pathways identified four pathways that were significantly overrepresented (table 3). Two of these pathways, *galactose* and *starch and sucrose*, are linked to the metabolism of carbohydrates (as identified in the GO analysis above). A third pathway was *neuroactive ligand–receptor interaction*.

Two subnetworks showed negative correlations with Successfully Mated queens (table 2 and fig. 5) and three hub genes were associated with Successfully Mated queens: *FMRFamide receptor*, *nesprin*, and *GABA receptor* (see "hub_genes" in supplementary table S14, Supplementary Material online). No subnetworks were linked to Failed Mated queens while three subnetworks were overrepresented in the comparison Successfully Mated versus Failed Mated queens. These included 11 hub genes of which 3 were differentially expressed (one of these is *titin*).

Some of the biological functions associated with mating in bumblebees appear to be conserved across other insect taxa. We compared the lists of biological functions putatively involved in successful mating events with similar lists for other insects and found some similarities with two species of fruit fly—*Drosophila melanogaster* (Dalton et al. 2010) and *Ceratitis capitata* (Gomulski et al. 2012)—and the honey bee *Apis mellifera* (Manfredini et al. 2015). We found 22 shared biological functions overall ("Comparative_studies"
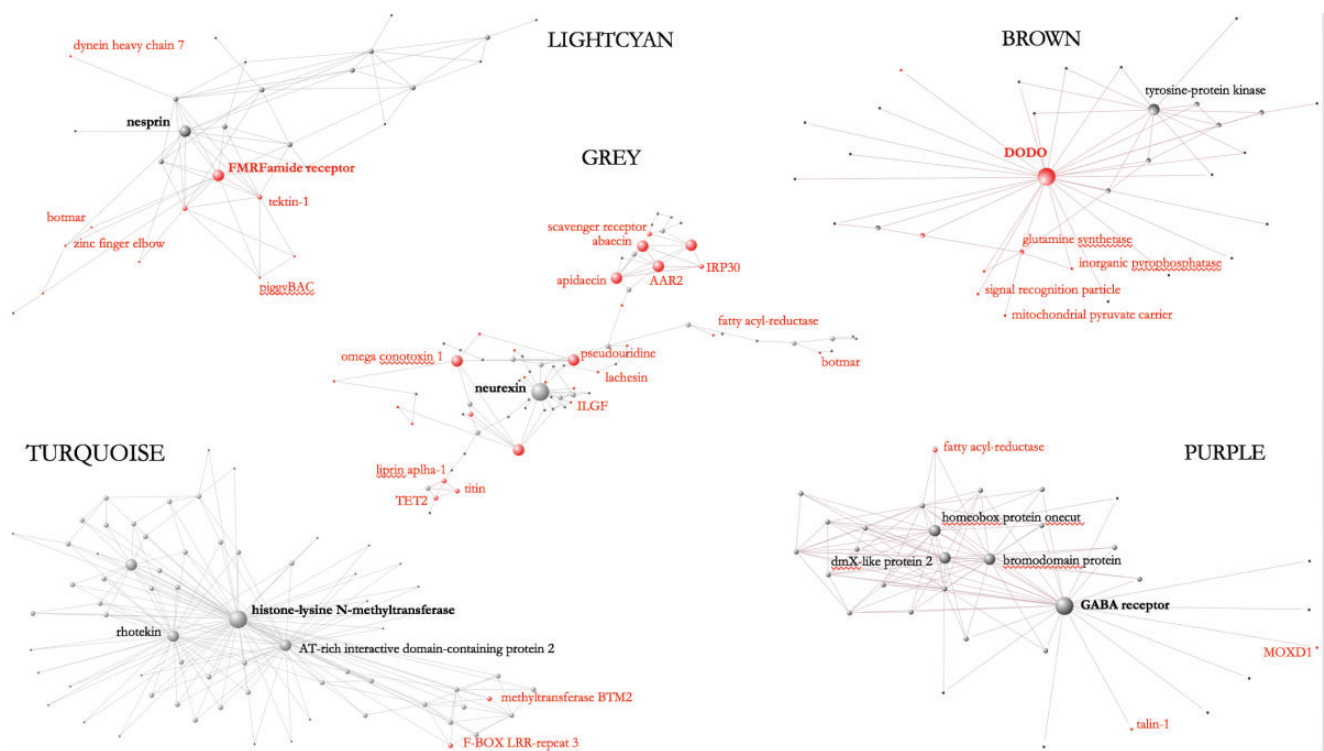
FIG. 5.—Visualization of gene subnetworks in bumblebee brains. The five WGCNA subnetworks that were significantly associated with queen phenotypes (subnetwork–trait association analysis) are displayed here. Larger nodes indicate hub genes (i.e., genes that are highly connected): names for these genes are in bold. In red are genes that were differentially expressed in at least one pairwise comparison. Names are provided for all genes where annotations could be retrieved. Names are also provided for hub genes irrespective of their expression patterns.

Table 3
KEGG Analysis of Metabolic Pathways

KEGG id	Description	SR/SM	SM/FM	SR/FR
4145	*Phagosome*			X
4391	*Hippo signaling pathway*			X
4745	*Phototransduction*			X
4080	*Neuroactive ligand–receptor interaction*	X		X
52	*Galactose metabolism*			X
500	*Starch and sucrose metabolism*			X
1100	*Metabolic pathways*			X

Note.—Reported here are the KEGG pathways that were significantly overrepresented among the genes that were significantly differentially regulated in one of the three focal pairwise comparisons ($P < 0.05$, after Benjamini–Hochberg correction for multiple testing).

in supplementary table S11, Supplementary Material online): this overlap is significantly larger than expected by chance (Hypergeometric test: Representation factor: 5.2, $P = 1.2e$-10). The majority of shared GO terms (59%; $n = 13$) is associated with sensory perception and response to stimuli. Another set of three shared GO terms are related to the metabolism of carbohydrates (carbohydrate metabolic process, cellular carbohydrate metabolic process and cellular ketone metabolic process).

Neurogenomic Signatures of Success and Failure in Reproductive Maturation (Aim 4)

The two most contrasting phenotypes in our study were the Successfully Reproductive and Failed Reproductive queens (fig. 1): 1,225 genes were differentially expressed (15.9% of all genes analyzed, FDR < 0.05), 769 showing >2-fold changes in expression levels, of which 181 had >4-fold changes ("edgeR_SR-FR" in supplementary table S6, Supplementary Material online). Interestingly, more than two-thirds of DEGs were up-regulated in Failed Reproductive queens (875, vs. 350 that were down-regulated, χ^2 test from equal: $\chi^2 = 116.12$, df = 1, $P < 1e$-4). These differences are likely to reflect reproductive physiology but also behavior (e.g., nest building, producing and modeling wax, foraging and caring for the brood, Goulson 2010). The differences are not likely to be due to age, as both groups of queens were ∼3 months old ("bee_samples" in supplementary table S2, Supplementary Material online).

Successfully Reproductive and Failed Reproductive queens differed for a variety of biological functions. Across the 1,225 DEGs, 95 GO terms were significantly enriched ($P < 0.05$), clustering into five major clusters of related terms ("REVIGO" in supplementary table S9, Supplementary

Material online): defence response to other organism (3 elements); response to toxic substance (6 elements); lipid and hormone metabolism (7 elements); response to organic substance (22 elements) including several functions related to the metabolism of carbohydrates; and regulation of nucleotide biosynthetic process (22 elements).

KEGG analysis on the same set of DEGs identified four pathways that were significantly enriched (*P* < 0.05, table 3) one of which, *neuroactive ligand–receptor interaction*, was identified as a molecular signature in the Successfully Mated/Failed Mated comparison. The three other pathways are particularly interesting: 1) *Hippo signaling pathway* is an evolutionarily conserved pathway, from flies to humans, and controls organ size during development by regulating cell-to-cell signaling and cell proliferation (Halder and Johnson 2011); 2) the *phototransduction* pathway is part of the process of visual signaling and involves the conversion of light signals (photons) into a change of membrane potential in photoreceptor cells (Katz and Minke 2009); 3) the *phagosome* pathway is linked to phagocytosis, that is, the process of particle engulfment by cells that operates during tissue remodeling, inflammation, and defence against infectious agents (Stuart and Ezekowitz 2005).

Two subnetworks were significantly associated with Failed Reproductive queens and they showed opposite directions of expression (table 2 and fig. 5). One subnetwork was correlated (positively) with Successfully Reproductive queens. If we look at the comparison between Successful versus Failed Reproductive queens, ten subnetworks were overrepresented and two of these were uniquely overrepresented in this comparison. About 11 hub genes were associated with overrepresented subnetworks and 5 of these were differentially expressed between the two queen phenotypes ("hub_genes" in supplementary table S14, Supplementary Material online). These genes are: *flocculation*, the *peptidyl-prolyl cis-trans isomerase dodo*, a signal transducer member of the MAP kinase pathway (Hsu et al. 2001), and the previously described *Peroxidasin*, *botmar-15 transposon mariner*, and *piggyBac*.

Discussion

In this study, we applied the neurogenomic approach to the key pollinator *Bombus terrestris* to explore the relationship between brain gene expression and important life-history transitions in queens. The multi-level analysis of RNAseq data, combining detailed characterization of gene expression, gene functions, and gene network, produced three major results that address our initial hypotheses: 1) each queen phenotype displays a unique neurogenomic profile defined by subtle differences at the levels of gene expression, biological functions, KEGG pathways, and gene networks (Hypothesis A); 2) the key transition from successful mating to reproductive maturation has a distinct neurogenomic signature and presents both similarities and striking differences compared

with closely related organisms such as honey bees or fruit flies (Hypothesis B); 3) queens that failed to mate or become reproductively mature are characterized by distinct neurogenomic profiles compared with their successful counterpart (Hypothesis C).

Failure during Life-History Transitions

Failure has a broad impact on brain gene regulation for the biological processes that we considered. In fact, for both mating and reproductive maturation, failure was associated with more DEGs than success, and with increased up-regulation of gene expression. Furthermore, Failed Reproductive queens represented the most distinct phenotype among the set of phenotypes that we analyzed in this study.

As a proxy for failed reproductive maturation, we used ovary dissection and lack of egg development. Ovary development and egg production in queens and workers of social insects is regulated by Juvenile Hormone, which is synthesized by the *corpora allata*, paired glands associated with the brain (Page et al. 2012). Such reproductive maturation is known to be positively correlated with the levels of biogenic amines in the brain (Harris and Woodring 1995; Boulay et al. 2001; Sasaki et al. 2007). These compounds are neuroendocrine modulators that act as major drivers of behavior. The distinct neurogenomic state of Failed Reproductive queens is therefore likely to be due to the absence of reproductive behaviors, including the behavioral patterns associated with colony founding (Goulson 2003).

Our KEGG analyses showed that several key metabolic pathways differ between Failed and Successfully reproductive queens, indicating possible mechanisms that could mediate failure in this important biological transition. The fact that *Hippo* is differentially regulated could indicate that neural cells undergo different paths of restructuring in Failed versus Successfully Reproductive queens. Interestingly, work on harvester ants has shown that behavioral changes associated with mating and ovary activation in ant queens are linked to a reduction in the size of the brain (Julian and Gronenberg 2002). Alternatively, regulation of *Hippo* in the brain could mirror the failed development of ovaries/eggs in the abdomen of Failed Reproductive queens. The differential regulation of the *phototransduction* pathway could be associated with the transition from photophilic (attracted by light) to photophobic (repulsed by light) behavior: in nature, newly reproductive queens progressively reduce their foraging activity and display more nest-bound behavior after ovary development (Goulson 2003). Failed Reproductive queens lack this transition and this might be why they differ from Successfully Reproductive for the regulation of this pathway. A similar pattern of regulation of genes associated with visual perception has been observed in honey bee queens (Manfredini et al. 2015) while they transition from virgin and photophilic (in preparation for mating flights) to mated and photophobic

(as they return to the colony where they live in total darkness). Finally, *phagosome* also differed between Failed and Successfully Reproductive queens. This result could be linked to the different stress or health conditions that the two groups of queens experience.

Another set of analyses at the gene and gene network levels indicate significant differences between Failed and Successfully Reproductive queens, and provide possible explanations for the mechanisms underpinning failure to become reproductively mature. Firstly, at the gene level there is the overall up-regulation of genes associated with "defence response," "response to toxic substance," and "phagosome" in Failed Reproductive queens, which could indicate a suboptimal state of health. We investigated the possibility that Failed Reproductive queens could be affected by parasitic infections that could undermine their health status. For this purpose, we screened the RNAseq data to find sequences matching the most common bee viruses (the only bee parasites that have been detected in the brain tissue so far) but obtained no matches, indicating that these viruses were not present. Secondly, at the network level, we identified two highly connected (hub) genes that are associated with Failed Reproductives: *histone-lysine N-methyltransferase Su(var)3-9* (turquoise, positively associated), implicated in histone residue methylation (Li et al. 2016), suggesting that there may be some epigenetic basis to the failure; and *dodo* (brown, negatively associated), involved in the process of oogenesis in *Drosophila* (Hsu et al. 2001), indicating that the failed activation of the ovaries could be mirrored by down-regulation of *dodo* in the brain. *FMRFamide receptor*, and *nesprin* (lightcyan, for details see "lightcyan" in supplementary table S15, Supplementary Material online) are instead positively associated with Successfully Reproductive queens. The *FMRFamide receptor* regulates key behaviors such as locomotion (Kiss et al. 2013) and response to environmental stress—for example, intense light exposure (Klose et al. 2010)— whereas *nesprin* is linked to muscle activity (Zhang et al. 2002): up-regulation of these genes in Successfully Reproductive queens is clearly associated with their transition to nest-bound behavior associated with egg-laying and brood rearing. These results show that the successful accomplishment of reproductive maturation and the maintenance of reproductive functions are not associated with a massive activation of reproductive genes (as one would expect) but with the subtle coordination of reproductive behavior and physiology.

We cannot say whether failure in general is the cause or the consequence of the observed changes in gene expression. Failed mating/reproductive maturation can be triggered by multiple external factors (e.g., interaction with males, rearing environment, diapause, food regime) or internal physiological processes that impact brain gene expression. Alternatively, faulty patterns of gene expression might be the cause of unsuccessful mating/reproductive maturation in these queens. In

this scenario, for example, a lack of canalization of gene expression in a process that is typically canalized (as reproduction in insects typically is, e.g., Hatle et al. 2003) can result in a global up-regulation of gene expression in the brain. Only a time-course analysis of gene expression across the mating and reproductive processes can address this question. These observations are in line with the idea that failed systems are less stable as they contain less information and require more regulation. If we analyze our system in terms of the Shannon's Information Theory criteria (Gatenby and Frieden 2007; Mousavian et al. 2016), then failed mated/reproductive queens require more regulation as they contain less information (i.e., they have higher entropy). In fact, they are characterized by higher uncertainty as their fate is less predictable, whereas successful queens have more information in their transcriptome as they successfully accomplished an important life-history transition. In an evolutionary perspective, this could be explained as a lack of selection for mechanisms that control gene regulation after failure (as failure is a reproductively and evolutionary dead end) while success has selected for mechanisms that improve stability of biological systems *via* tighter control on the regulation of gene expression. Similar considerations have been used to explain looser control in regulation of gene expression (or the higher frequency of errors observed) in human and mice that have passed the reproductive age (Hong et al. 2008).

Successful Life-History Transitions in Comparison with Other Insect Systems

After having discussed the implications of failure during mating and reproductive maturation in bumblebee queens, it is important now to consider the neurogenomic characterization of success to achieve these important transitions, and to analyze how this relates to other systems where similar questions have been addressed in the past. Despite the small number of genes differentially expressed after mating (SM/FM comparison), this biological transition is associated with large numbers of GO terms and KEGG metabolic pathways, showing that even small numbers of regulated genes can be responsible for important changes at the behavioral and physiological levels.

The regulation of many carbohydrate-related functions indicates that carbohydrates are fundamental compounds needed for brain activity, but also that carbohydrates could be involved in other physiological processes related to mating and regulated at the brain level. Several functions associated with neuroactive compounds were also regulated after mating. Typical neuroactive compounds are, for example, dopamine, histamine, and serotonin, and the roles of these molecules in mediating behavioral responses are well documented in insect models (Kamhi and Traniello 2013). Our GO analyses reveal that genes associated with the dopamine receptor signaling pathway are differentially

regulated in Successfully Mated versus Failed Mated, stressing the importance of these molecules in mediating the behavioral outcome of the mating process. Finally, our network analyses indicate that many genes linked to neurogenesis and neural processes are important "hubs" in the regulation of gene expression, as one would expect from the analysis of brain activity. However, we also detected "hub" genes broadly associated with gene regulation, such as, for example, transposable elements and methylation-related genes. Both these groups of genes have been shown to play a role in the regulation of brain plasticity (Reilly et al. 2013). These findings suggest that key genes associated with brain activity in bumblebee queens that are experiencing important life-history transitions can have a double nature. They are either key players of relevant biological functions (e.g., neural processes) or they are general regulators of gene expression (methylation or transposable elements).

The lack of a substantial shift in gene expression of bumblebee queens a few days after mating is surprising, and contrasts with studies in other insects that looked at similar time windows (Kocher et al. 2008; Dalton et al. 2010; Manfredini et al. 2015). In honey bees, 829 DEGs (6.4% of all genes analyzed; RNAseq) were detected 2 days after mating (Manfredini et al. 2015), whereas in *Drosophila* 545 DEGs (3.9% of the genes in the microarrays) were detected 3 days after mating (Dalton et al. 2010). Bumblebee queens typically start diapause shortly after they mate and resume reproductive maturation only after emergence from diapause. Hence, global brain gene expression in mated queens may be indicative of reduced behavioral activity in preparation for diapause (Alford 1969), rather than regulation of reproductive pathways, and this might explain the contrasts that we observe with organisms that directly transition from mating to reproductive maturation. These comparisons highlight the importance of life history in explaining patterns of molecular processes of shared traits (i.e., mating) across species. However, it is important to highlight the fact that our experimental approach to investigate mating was significantly different from previous studies. Although we compared individuals that successfully mated versus individuals that failed despite the fact that they were given the same chance to do so, previous studies had either analyzed successfully mated individuals across different time points after mating (Dalton et al. 2010), or they compared successfully mated individuals versus virgin individuals that were not given the same chance to mate (Kocher et al. 2008; Manfredini et al. 2015). At a different level, it is interesting to notice that the comparison between successful and failed mated queens was associated with a large number of GO terms (124). This suggests that even a small change in the patterns of gene expression could potentially trigger the differential regulation of many biological functions.

Despite the difference in the global patterns of gene expression across organisms, our comparative analyses identify common functions that are shared after mating. Shared functions include sensory perception and response to stimuli, indicating that regulation at the brain level could be linked to sensorial activity at the physiological level (e.g., egg maturation, fertilization, and egg-laying activity) reflecting the cross-talk between different compartments in the insect body. In addition, functions related to the metabolism of carbohydrates are shared. These compounds are the major source for quickly available metabolic fuel and these results highlight the importance of the brain, a highly energetically demanding organ (Gallagher et al. 1998), as a major coordinator of post-mating changes in insects. This analysis strongly supports the existence of a genetic toolkit that regulates mating-induced changes in behavior across different organisms (Rittschof and Robinson 2016).

Conclusions

Success and failure are the opposite outcomes of many biological processes. Traditionally, successful phenotypes have been better characterized (in particular at the level of molecular organization), due to the importance of understanding how key biological processes are regulated. However, failure also deserves great attention, as it might explain the decline of many species in the wild and can provide better tools to understand success itself. Our study provides a first characterization of the molecular underpinnings of failure associated with mating and reproductive maturation in a key pollinator insect.

It will be interesting in the future to understand whether failure associated with other biological processes is characterized by similar molecular patterns, in bumblebees as well as in other organisms. This will be a key step toward defining the molecular underpinnings of failure per se—whereas, in this study, we limit our investigation to failure in two specific life-history transitions and in relation to success within the same transitions. For example, an interesting area of research would be neurogenomic studies on aggressive interactions among conspecifics, where global gene expression profiles of individuals that succeed (i.e., win the competition over a resource) or fail (i.e., lose) have been compared. Interestingly, studies like these on male zebrafish or fire ant founding queens have shown that losers are characterized by larger changes in patterns of gene expression compared with winners (Manfredini et al. 2013; Oliveira et al. 2016). These results mirror our findings that failed bumblebee queens are characterized by massive up-regulation of genes in the brain, and suggest that failing to complete a key life transition or losing a competition with a rival have equally important consequences for the neurogenomic profile of an organism. Future studies should explore whether similar molecular mechanisms are in place in both scenarios, that is, whether there is a conserved genetic toolkit for biological success and failure across different animal systems.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

Literature Cited

Alford DV. 1969. A study of the hibernation of bumblebees (Hymenoptera: Bombidae) in Southern England. J Anim Ecol. 38(1):149–170.

Almansa R, et al. 2015. Transcriptomic correlates of organ failure extent in sepsis. J Infect. 70(5):445–456.

Amsalem E, Grozinger CM, Padilla M, Hefetz A. 2015. Chapter two – the physiological and genomic bases of bumblebee social behaviour. Adv Insect Physiol. 48:37–93.

Barribeau SM, et al. 2015. A depauperate immune repertoire precedes evolution of sociality in bees. Genome Biol. 16:83.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B 57(1):289–300.

Biswas S, et al. 2010. Sensory regulation of neuroligins and *neurexin I* in the honeybee brain. PLoS One 5(2):e9133.

Boulay R, Hooper-Bui LM, Woodring J. 2001. Oviposition and oogenesis in virgin fire ant females *Solenopsis invicta* are associated with a high level of dopamine in the brain. Physiol Entomol. 26(4):294–299.

Colgan TJ, et al. 2011. Polyphenism in social insects: insights from a transcriptome-wide analysis of gene expression in the life stages of the key pollinator, *Bombus terrestris*. BMC Genomics 12:623.

Cork SM, Van Meir EG. 2011. Emerging roles for the BAI1 protein family in the regulation of phagocytosis, synaptogenesis, neurovasculature, and tumor development. J Mol Med. 89(8):743–752.

Dalton JE, et al. 2010. Dynamic, mating-induced gene expression changes in female head and brain tissues of *Drosophila melanogaster*. BMC Genomics 11(1):13.

Doe CQ, Smouse D, Goodman CS. 1988. Control of neuronal fate by the *Drosophila segmentation gene even-skipped*. Nature 333(6171):376–378.

Duvoisin N, Baer B, Schmid-Hempel P. 1999. Sperm transfer and male competition in a bumblebee. Anim Behav. 58(4):743–749.

Engelstädter J. 2015. Host-parasite coevolutionary dynamics with generalized success/failure infection genetics. Am Nat. 185(5):E117–E129.

Gallagher D, et al. 1998. Organ-tissue mass measurement allows modeling of REE and metabolically active tissue mass. Am J Physiol Endocrinol Metab. 275:E249–E258.

Gatenby RA, Frieden BR. 2007. Information theory in living systems: methods, applications, and challenges. Bull Math Biol. 69(2):635–657.

Gomulski LM, et al. 2012. Transcriptome profiling of sexual maturation and mating in the Mediterranean fruit fly, *Ceratitis capitata*. PLoS One 7(1):e30857.

Goulson D. 2003. Bumblebees: their behaviour and ecology. Oxford: Oxford University Press, 234pp.

Goulson D. 2010. Bumblebees: behaviour, ecology, and conservation. Oxford: Oxford University Press, 317pp.

Halder G, Johnson RL. 2011. Hippo signaling: growth control and beyond. Development 138(1):9–22.

Harris JW, Woodring J. 1995. Elevated brain dopamine levels associated with ovary development in queenless worker honey bees (*Apis mellifera* L.). Comp Biochem Physiol C Pharmaco Toxicol Endocrinol. 111(2):271–279.

Harris RM, Hofmann HA. 2014. Neurogenomics of behavioral plasticity. In: Landry C., Aubin-Horth N. (eds) Ecological Genomics. Advances in Experimental Medicine and Biology, vol 781. Dordrecht: Springer.

Hatle JD, Borst DW, Juliano SA. 2003. Plasticity and canalization in the control of reproduction in the lubber grasshopper. Integr Comp Biol. 43(5):635–645.

Hong M-G, Myers AJ, Magnusson PK, Prince JA. 2008. Transcriptome-wide assessment of human brain and lymphocyte senescence. PLoS One 3(8):e3024.

Hsu T, McRackan D, Vincent TS, Gert De Couet H. 2001. *Drosophila Pin1 prolyl isomerase Dodo* is a MAP kinase signal responder during oogenesis. Nat Cell Biol. 3(6):538–543.

Hu Z, et al. 2013. VisANT 4.0: integrative network platform to connect genes, drugs, diseases and therapies. Nucleic Acids Res. 41(W1):W225–W231.

Imran M, Ahmad M, Nasir MF, Saeed S. 2015. Effect of different nest box materials on the mating of European bumblebee, *Bombus terrestris* (Hymenoptera: Apidae) under controlled environmental conditions. Pak J Zool. 47:241–247.

Julian GE, Gronenberg W. 2002. Reduction of brain volume correlates with behavioral changes in queen ants. Brain Behav Evol. 60(3):152–164.

Kamhi JF, Traniello JF. 2013. Biogenic amines and collective organization in a superorganism: neuromodulation of social behavior in ants. Brain Behav Evol. 82(4):220–236.

Karsli BA, Gurel F. 2013. Effects of CO_2 narcosis on the onset of oviposition and colony founding success of post diapausing bumblebee (*Bombus terrestris*) queens. Kafkas Universitesi Veteriner Fakyltesi Dergisi 19:221–224.

Katz B, Minke B. 2009. *Drosophila* photoreceptors and signaling mechanisms. Front Cell Neurosci. 3:2.

Kiss B, et al. 2013. Selective elimination/RNAi silencing of FMRF-related peptides and their receptors decreases the locomotor activity in *Drosophila melanogaster*. Gen Comp Endocrinol. 191:137–145.

Kleijn D, et al. 2015. Delivery of crop pollination services is an insufficient argument for wild pollinator conservation. Nat Commun. 6:7414.

Klose MK, Dason JS, Atwood HL, Boulianne GL, Mercier AJ. 2010. Peptide-induced modulation of synaptic transmission and escape response in *Drosophila* requires two g-protein-coupled receptors. J Neurosci. 30(44):14724–14734.

Kocher SD, Richard F-J, Tarpy DR, Grozinger CM. 2008. Genomic analysis of post-mating changes in the honey bee queen (*Apis mellifera*). BMC Genomics 9:232.

Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9(1):559.

Lein ES, et al. 2007. Genome-wide atlas of gene expression in the adult mouse brain. Nature 445(7124):168–176.

Li H, et al. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25(16):2078–2079.

Li S, et al. 2016. *SUVH1*, a Su (var) 3–9 family member, promotes the expression of genes targeted by DNA methylation. Nucleic Acids Res. 44(2):608–620.

Li X, et al. 2015. Metabolic and transcriptomic signatures of rice floral organs reveal sugar starvation as a factor in reproductive failure under heat and drought stress. Plant Cell Environ. 38:2171–2192.

Liang ZS, et al. 2014. Comparative brain transcriptomic analyses of scouting across distinct behavioural and ecological contexts in honeybees. Proc R Soc Lond B Biol Sci. 281(1797):20141868.

Liu X, Krause WC, Davis RL. 2007. *GABA A receptor RDL* inhibits *Drosophila* olfactory associative learning. Neuron 56(6):1090–1102.

Lutterschmidt DI, Maine AR. 2014. Sex or candy? Neuroendocrine regulation of the seasonal transition from courtship to feeding behavior in male red-sided garter snakes (*Thamnophis sirtalis parietalis*). Horm Behav. 66(1):120–134.

Ma K, Forbes JG, Gutierrez-Cruz G, Wang K. 2006. *Titin* as a giant scaffold for integrating stress and Src homology domain 3-mediated signaling pathways. J Biol Chem. 281(37):27539–27556.

Manfredini F, Brown MJF, 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(1):563.

Manfredini F, et al. 2013. Sociogenomics of cooperation and conflict during colony founding in the fire ant *Solenopsis invicta*. PLoS Genet. 9(8):e1003633.

Marqués G, et al. 2003. Retrograde Gbb signaling through the Bmp type 2 receptor *wishful thinking* regulates systemic *FMRFa* expression in *Drosophila*. Development 130(22):5457–5470.

Maruska KP. 2014. Social regulation of reproduction in male cichlid fishes. Gen Comp Endocrinol. 207:2–12.

Mikheyev AS, Linksvayer TA. 2015. Genes associated with ant social behavior show distinct transcriptional and evolutionary patterns. Elife 4:e04775.

Morandin C, et al. 2016. Comparative transcriptomics reveals the conserved building blocks involved in parallel evolution of diverse phenotypic traits in ants. Genome Biol. 17:43.

Morel V, Lepicard S, Rey AN, Parmentier M-L, Schaeffer L. 2014. *Drosophila Nesprin-1* controls glutamate receptor density at neuromuscular junctions. Cell Mol Life Sci. 71(17):3363–3379.

Mousavian Z, Kavousi K, Masoudi-Nejad A. 2016. Information theory in systems biology. Part I: gene regulatory and metabolic networks. Semin Cell Dev Biol. 51:3–13.

Naumova OY, Lee M, Rychkov SY, Vlasova NV, Grigorenko EL. 2013. Gene expression in the human brain: the current state of the study of specificity and spatiotemporal dynamics. Child Dev. 84(1):76–88.

Oldham MC, et al. 2008. Functional organization of the transcriptome in human brain. Nat Neurosci. 11(11):1271–1282.

Oliveira RF, et al. 2016. Assessment of fight outcome is needed to activate socially driven transcriptional changes in the zebrafish brain. Proc Natl Acad Sci U S A. 113(5):E654–E661.

Page RE, Rueppell O Jr, Amdam GV. 2012. Genetics of reproduction and regulation of honey bee (*Apis mellifera* L.) social behavior. Annu Rev Genet. 46:97.

Patalano S, et al. 2015. Molecular signatures of plastic phenotypes in two eusocial insect species with simple societies. Proc Natl Acad Sci U S A. 112(45):13970–13975.

Reilly MT, Faulkner GJ, Dubnau J, Ponomarev I, Gage FH. 2013. The role of transposable elements in health and diseases of the central nervous system. J Neurosci. 33(45):17577–17586.

Rittschof C, Robinson G. 2016. Chapter five – behavioral genetic toolkits: toward the evolutionary origins of complex phenotypes. Curr Top Dev Biol. 119:157–204.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26(1):139–140.

Rouleux-Bonnin F, Petit A, Demattei M-V, Bigot Y. 2005. Evolution of full-length and deleted forms of the mariner-like element, *Botmar1*, in the genome of the bumble bee, *Bombus terrestris* (Hymenoptera: Apidae). J Mol Evol. 60(6):736–747.

Rutrecht ST, Brown MJF. 2008. The life-history impact and implications of multiple parasites for bumble bee queens. Int J Parasitol. 38(7):799–808.

Sadd BM, et al. 2015. The genomes of two key bumblebee species with primitive eusocial organization. Genome Biol. 16:1–32.

Sasaki K, Yamasaki K, Nagao T. 2007. Neuro-endocrine correlates of ovarian development and egg-laying behaviors in the primitively eusocial wasp (*Polistes chinensis*). J Insect Physiol. 53(9):940–949.

Soudi M, Zamocky M, Jakopitsch C, Furtmüller PG, Obinger C. 2012. Molecular evolution, structure, and function of peroxidasins. Chem Biodivers. 9(9):1776–1793.

Stuart LM, Ezekowitz RAB. 2005. Phagocytosis: elegant complexity. Immunity 22(5):539–550.

Supek F, Bošnjak M, Škunca N, Šmuc T. 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One 6(7):e21800.

Teerawanichpan P, Robertson AJ, Qiu X. 2010. A fatty acyl-CoA reductase highly expressed in the head of honey bee (*Apis mellifera*) involves biosynthesis of a wide range of aliphatic fatty alcohols. Insect Biochem Mol Biol. 40(9):641–649.

Toth AL, et al. 2014. Shared genes related to aggression, rather than chemical communication, are associated with reproductive dominance in paper wasps (*Polistes metricus*). BMC Genomics 15(1):75.

Trapnell C, et al. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 7(3):562–578.

Von Wyschetzki K, Rueppell O, Oettler J, Heinze J. 2015. Transcriptomic signatures mirror the lack of the fecundity/longevity trade-off in ant queens. Mol Biol Evol. 32(12):3173–3185.

Walker RJ, Papaioannou S, Holden-Dye L. 2009. A review of FMRFamide- and RFamide-like peptides in metazoa. Invert Neurosci. 9(3–4):111–153.

Walkinshaw E, et al. 2015. Identification of genes that promote or inhibit olfactory memory formation in *Drosophila*. Genetics 199(4):1173–1182.

Waltzer L, Bataillé L, Peyrefitte S, Haenlin M. 2002. Two isoforms of *Serpent* containing either one or two GATA zinc fingers have different roles in *Drosophila* haematopoiesis. EMBO J. 21(20):5477–5486.

Wilburn DB, Swanson WJ. 2016. From molecules to mating: rapid evolution and biochemical studies of reproductive proteins. J Proteomics 135:12–25.

Zayed A, Robinson GE. 2012. Understanding the relationship between brain gene expression and social behavior: lessons from the honey bee. Annu Rev Genet. 46:591–615.

Zhang Q, Ragnauth C, Greener MJ, Shanahan CM, Roberts RG. 2002. The nesprins are giant actin-binding proteins, orthologous to *Drosophila melanogaster* muscle protein MSP-300. Genomics 80(5):473–481.

Associate editor: Soojin Yi