BMC Genomics

RESEARCH

Open Access

Post-transcriptional regulation of Dufour's gland reproductive signals in bumble bees

Nathan Derstine[1*], Tatiana Laremore[2] and Etya Amsalem[1]

Abstract

Pheromone communication is a key mechanism by which the reproductive division of labor is maintained within insect communities. Understanding how pheromones evolved to regulate social behavior requires knowledge of the molecular regulation of their production. However, even in cases where pheromones were identified, our understanding of their biosynthesis and molecular regulation remains limited. Bumble bees provide a unique system to explore pheromone biosynthesis since workers produce ester sterility signals in their Dufour's gland that differ from gyne-specific esters and are not produced by queens. These esters are hypothesized to be produced in the exocrine gland where they are stored, and indeed queens, gynes and workers differ significantly in the expression of Dufour's gland genes coding to enzymes involved in the biosynthesis of esters. However, a previous transcriptome analysis revealed no gene expression differences in the Dufour's gland of workers despite differences in both ester production and ovarian activation, suggesting that ester production may be regulated lower down. Proteomics of the Dufour's gland of queens, gynes, and workers recovered over 2400 proteins and broadly matched the previous RNAseq data. However, more than 100 differentially expressed proteins were found between the worker groups, including key enzymes in fatty acid biosynthesis, indicating that the regulation of reproductive signal biosynthesis in workers is done post-transcription. Overall, our data provide evidence that pheromone biosynthesis in the Dufour's gland is caste specific, that gynes and workers are likely using different enzymes to make their respective wax esters, and that the regulation on pheromone production in queens, gynes and workers is likely done at different regulatory levels, with workers signals being subjected to regulation at the protein level.

Keywords Pheromone biosynthesis, Social insects, Dufour's gland, Proteomics

Background

Chemical signals mediate key aspects of insect life and guide fundamental behaviors such as finding food and mating. They are particularly important in social insects for regulating reproduction and reproduction-related behaviors in females, the hallmark of sociality. Signals regulating reproduction are produced in multiple exocrine glands [1, 2] and are an important factor in the evolution of eusociality. However, despite the immense progress in identifying their structure and function [3–5], their biosynthesis and genetic basis remain poorly understood.

*Correspondence:
Nathan Derstine
ntd34@psu.edu
[1]Department of Entomology, Center for Chemical Ecology, Center for Pollinator Research, Huck Institutes of the Life Sciences, Pennsylvania State University, University Park, PA 16802, USA
[2]Proteomics and Mass Spectrometry Core Facility, Huck Institutes of the Life Sciences, Pennsylvania State University, University Park, PA 16802, USA

One common source of reproductive signals in Hymenoptera is the female Dufour's gland.

This exocrine gland is associated with the sting complex which secretes a wide variety of chemical compounds with many functions [6, 7]. In solitary species, the primary functions of the secretion are waterproofing brood cells and provisioning larvae [8], but in many social taxa the secretion has gained a communicative function, often by signaling reproductive status. For example, in the honey bee, *Apis mellifera*, the secretion of fertile workers mimics the queen secretion by producing a larger amount of wax esters compared to sterile workers [9]. In two species of bumble bees, sterile workers produce wax esters in their Dufour's gland that are not produced by queens and may function to decrease aggression from nestmates [10–13]. While the production of wax esters in these cases is assumed to occur within the Dufour's gland [14, 15], the specific biosynthetic pathway is unknown. Here, we examine the biosynthetic regulation of Dufour's gland compounds at the protein level in the common eastern bumble bee *Bombus impatiens*.

B. impatiens is a social insect forming annual colonies that are founded by a single queen. After a several month period of worker production and colony growth, gynes (unmated, young queens) and males are produced [16, 17]. In the initial growth phase, very few workers activate their ovaries, but after a certain point in the colony cycle called the 'competition point', workers activate their ovaries and compete both with the queen and other workers by laying male-destined eggs [18]. The Dufour's gland secretion in this species contains mostly esters, terpenes, and hydrocarbons. Worker-specific esters are produced in higher quantities by workers with undeveloped ovaries compared to fertile workers, whereas queen-specific esters and terpenes are produced in higher quantities by gynes compared to active queens [13]. Worker-specific esters have shorter carbon chains than the long and saturated esters produced by gynes. The esters found in gynes further contain double bonds in either the fatty acid or the fatty alcohol component of the ester, or in both. Worker glands are therefore predicted to express enzymes involved in chain shortening, but not desaturation. In addition, workers do not produce any of the terpenes or terpene esters found in gynes and queens. Terpenes are synthesized via the isoprenoid pathway [19] and therefore gynes and queens are predicted to over-express enzymes involved in this pathway relative to workers. The dynamic reproduction of females and the selective production of reproductive signals in the Dufour's gland make *B. impatiens* a great model to study the mechanisms regulating the biosynthesis of reproductive signals. Furthermore, fatty acid esters and terpenes are frequently used as semiochemicals in insect communication. They function as brood, alarm, aggregation, and sex pheromones in taxa from aphids [20], to beetles [21], to true bugs [22], and bees [23–25]. Thus, understanding their biosynthesis has broad implications for understanding the chemical communication systems of many species.

In a previous study, we sequenced the transcriptome of the Dufour's gland tissue and identified thousands of differentially expressed genes between queens, gynes, and queenless (fertile) and queenright (sterile) workers [14]. The gene expression we previously described largely matched the production of queen- and worker-specific esters, and the terpenes and terpene-esters only found in gynes and queens (e.g. higher expression of desaturases in groups producing more unsaturated compounds). While we highlighted many candidate genes from the fatty acid (e.g., fatty acyl reductases that are involved in producing esters) and isoprenoid biosynthetic pathways (e.g., hydroxymethylglutaryl-CoA reductase, responsible for producing terpenes) that likely contribute to the unique chemical phenotypes, there were two areas where the data did not match the Dufour's gland chemistry. First, we found no differentially-expressed genes between queenright and queenless workers, even though queenright workers are sterile and produce a significantly higher percentage of wax esters than queenless workers, and second, workers overexpressed several genes in the isoprenoid biosynthetic pathway, even though workers produce none of the diterpenes and terpene esters present in the queen caste. These patterns could be explained by compound synthesis being regulated in workers at the protein level.

While transcriptomic and proteomic measurements of gene expression are often highly correlated, discrepancies arise because the rates of transcription and translation and the half-lives of mRNA and protein can differ between cell types and/or condition [26], often due to numerous post-translational modifications that can affect protein abundance, function, and localization [27–29]. In addition, several lines of evidence point towards protein-level buffering of transcriptional variation. For instance, the co-variation in mRNA expression observed in neighboring genes does not follow through to their proteins [30], and mRNA, but not protein abundances have been found to scale with alterations to DNA copy number [31–33]. These factors, combined with proteomics measuring a functional level more directly related to phenotype, make it an excellent complement to transcriptomic data.

Here, we use shotgun proteomics to further examine Dufour's gland biosynthesis across caste and reproductive state in *B. impatiens*. We used female groups identical to our previous study [14] to examine whether the chemical variation between females is more accurately reflected at the protein level, and to compare the underlying genetics

responsible for pheromone biosynthesis at the mRNA and protein levels. Based on chemical phenotype data, and retaining our initial hypotheses from the RNAseq experiment, we hypothesize that fatty acyl reductases and acyl transferase enzymes (both involved in ester production) would be differentially expressed between worker groups, and enzymes in the isoprenoid biosynthesis pathway (involved in terpene production) would be upregulated in the queen caste, particularly in gynes. We further use the data to understand the relationship between the mRNA and the protein levels with regard to the biosynthesis of chemical signals within a well conserved exocrine gland across female bees.

Methods

Insect rearing and sampling

Bees were collected from *B. impatiens* colonies that were obtained from Koppert Biological Systems (Howell, Michigan, USA) and colonies reared in-house. Colonies were maintained in nest boxes at a constant temperature of 28–30 °C and 60% relative humidity, constant darkness and supplied *ad libitum* with a 60% sugar solution and honeybee-collected pollen [13, 14, 34]. These colonies were used as a source for mated, egg-laying queens (hereafter "queens"), newly-emerged, unmated queens (hereafter "gynes"), and newly-emerged workers. Workers were assigned to one of two treatments after emergence: queenright workers (QRW) and queenless and broodless workers (QLW). QRW were labeled and placed back in their natal colony in the presence of the queen, brood, and nestmates until collection, while QLW were housed in plastic cages (11 cm diameter × 7 cm height) in groups of 3–5 workers until collection. Both groups of workers were sampled on day 7 after emergence. Since the presence of larvae can inhibit worker reproduction [35, 36], QLW cages were monitored daily for egg laying, and newly-laid eggs were removed. Gynes were collected upon emergence from late-season colonies and housed individually in small cages until sampling on day 7. To generate egg-laying queens, gynes were mated with non-sibling males between 7 and 10 days of age, treated with CO_2 to initiate reproduction [37], and reared until their colonies contained ~ 50–100 workers before sampling. At this point queens were 2–3 months old. The time points of the treatment groups were selected to capture differences in reproductive and chemical phenotypes based on previous work [13, 14].

Ovarian activation and Dufour's gland dissection

Freeze-killed bees were dissected by first pinning them on their back and making cuts with a micro-scissors laterally near the overlap of the sternites and tergites. A drop of HPLC grade water was applied to the sting complex to help float the Dufour's gland, which was pinched and severed at the base by fine forceps and placed directly into T-PER (Tissue Protein Extraction Reagent, Thermo Scientific). Oocyte size was measured for each bee in all samples. After removing the Dufour's gland, ovaries were removed and placed in a drop of distilled water, and the largest three terminal oocytes across both ovaries (at least one from each ovary) were measured with an eyepiece micrometer [12]. The average terminal oocyte was used as a measure of ovary activation for each bee. Additionally, the head width of each bee was measured with a digital caliper as a proxy for size [12].

Protein extraction

Dufour's glands from queens, gynes, QRW, and QLW were dissected to create 3 biological replicates of each group (total of 12 samples). Each replicate consisted of 3 pooled glands (total of 36 bees) of bees of similar body size, the same age, and same natal colony, except for mated queens where this was not possible. For a complete description of how many individuals were sampled from each colony, and a size comparison of the bees, please see Table S1 and Figure S1. For each sample, glands were homogenized in 120 µL T-PER (queens and gynes) or 60 µL T-PER (workers) with a sterile plastic pestle in a 2 mL tube, centrifuged, and were measured for total protein with a Bradford Assay. Samples were subsequently analyzed by SDS-PAGE to ensure successful extraction.

Protein digestion

Samples were loaded onto centrifugal filters (Pall 10,000 MWCO) for on-column trypsin digestion following a standard protocol [38]. Briefly, 100 µg or the total amount of the sample (if lower than 100 µg), was used. Samples were reduced with dithiothreitol, alkylated with iodoacetamide, and exchanged to a volatile digestion buffer (50 mM ammonium bicarbonate). A 1:50 weight: weight ratio of trypsin: protein was added and incubated for 8 h at 37° C. Peptides were collected by centrifugation, dried in a speed-vac, and stored at -80° C until analysis.

HPLC-MS

Tryptic digests were desalted using reverse-phase spin columns (Thermo Pierce p/n 89870) according to the product instructions and dried down in a SpeedVac vacuum concentrator. Each sample was reconstituted in a volume of 4% acetonitrile, 0.1% formic acid to achieve 2 mg/mL concentration (1 mg/mL QR1 and QR2). To 5 µL of each sample, a 1 µL of BSA digest (Thermo p/n) was added at 0.1% of the protein concentration in the sample, and approximately 67 pmol of each sample was injected onto a column for the nano-flow LC-MS analysis.

Peptides were separated on an Acclaim PepMap RSLC reverse-phase column (250 mm x 0.05 mm, Thermo p/n 164944) with a 20 mm x 0.075 mm precolumn (Thermo

p/n 164946). The gradient was delivered by an Easy-nLC 1200 pump (Thermo) and consisted of 3 min of 5% B, 5–35% B over 90 min, 35–60% B over 20 min, followed by a 7-min flush with 90% B. Mobile phase A was 0.1% formic acid in water and mobile phase B was 80% acetonitrile, containing 0.1% formic acid. The flow rate was 300 nL/min.

The mass spectra were acquired on a Thermo Orbitrap Eclipse instrument using a modification of the default Universal method. The MS resolution was 120 K and the data-dependent MS^2 scans were acquired in the ion trap. Collision-induced dissociation was used as the ion fragmentation technique. The ion isolation window was 0.7 m/z, the scan range was 150–1500 m/z, and the normalized AGC target was 250%. Dynamic exclusion duration was 60 s, and the precursor tolerances were set at ±10 ppm. Only charge states 2–6 were included and the single charge state per precursor filter was not used. The minimum intensity threshold was 1.5 e^4.

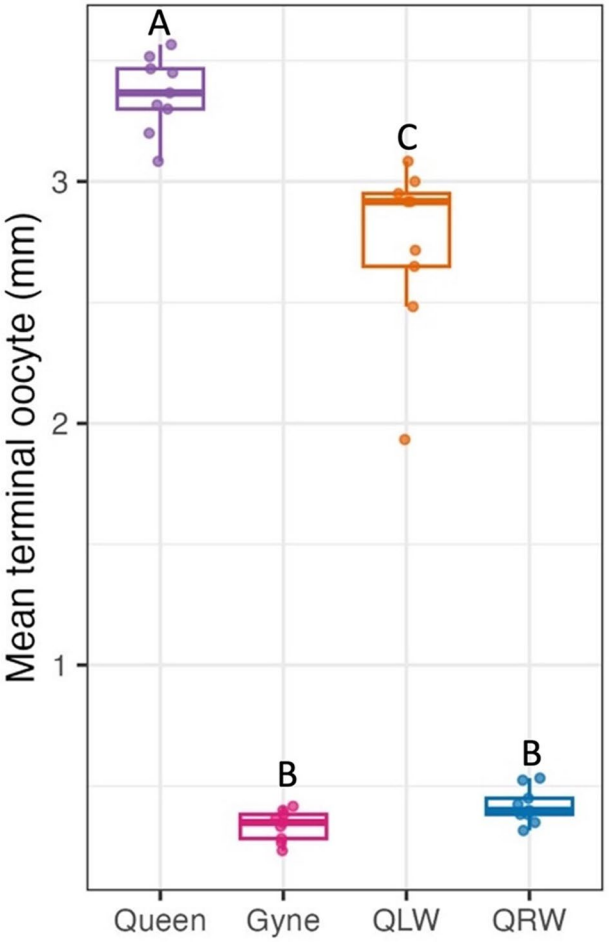

Fig. 1 Mean terminal oocyte length of bees used for proteomic analysis in four groups of *B. impatiens* females: queens, gynes, queenless workers (QLW) and queenright workers (QRW), $n=9$ for each treatment group. Each treatment group contained 3 samples of 3 pooled glands. Both groups of workers were 7 days old. Letters above boxplots indicate significant differences at $p<0.05$ (calculated from a linear mixed effects model with colony as a random factor, using Tukey's HSD post-hoc test for pairwise comparisons)

Data processing and analysis

The data was analyzed with Proteome Discoverer 2.5 (Thermo). The study definition included four sample groups: named G, M, QL, and QR referring to gynes, (mated) queens, queenless workers, and queenright workers; and each group contained three biological replicates, B1 through B12. A preconfigured processing step workflow for precursor quantification and LFQ was modified as follows: The *Bombus impatiens* (TaxID 132113) database from Uniprot was selected in the Spectrum Files RC and Sequest HT processing nodes; Trypsin was selected as the enzyme. All other processing parameters were left at their default values.

A preconfigured consensus step workflow for comprehensive enhanced annotation, LFQ, and precursor quantification was left unchanged except in two nodes. The *Bombus impatiens* (TaxID 132113) database was specified in the Protein Annotation node. The Precursor Ions Quantifier node had the following values selected - in the general quantification settings, Unique+Razor peptides were used; protein groups were considered for peptide uniqueness; shared quan results were used; quan results with missing channels were not rejected. Precursor quantification was based on intensity. The intensities were normalized on the specific protein amount and scaled on all average; a FASTA file containing BSA sequence was specified as the protein for normalization. For the quan rollup and hypothesis testing, protein abundances were calculated based on summed abundances. The protein ratios were calculated as the median of all possible pairwise peptide ratios calculated between replicates of all connected peptides. Missing values were imputed through random sampling of the lower 5% of detected values. The p-values for the protein ratios were obtained from the background-based t-test and adjusted using the Benjamini and Hochberg procedure to control for false discovery rate. Subcellular localizations and signaling sequences of proteins were predicted with Deeploc 2.0 [39]. Fisher's Exact test was used to assess whether any of the categories of localizations or signals were significantly enriched in the subset of proteins which were differentially expressed between any two groups.

Results

Social condition and reproductive phenotype

Ovarian activation of females followed previously reported patterns (Fig. 1) using the same treatment groups [14]. Queens had significantly larger terminal oocytes than gynes (3.36±0.05 and 0.34±0.02, mean mm±SE, respectively, linear mixed effects model with

colony as a random factor, followed by Tukey HSD post-hoc test, $p<0.001$), and QLW had significantly larger terminal oocytes than QRW (2.74±0.11 and 0.42±0.02, mean mm±SE, respectively, $p<0.001$). In groups with activated ovaries, queens' terminal oocyte size was also larger than of QLW ($p<0.001$). Gynes and QRW ovaries were not significantly different ($p=0.825$).

Shotgun proteomics results

We identified 2477 proteins across all Dufour's gland samples, of which 188 were differentially expressed (DE) between at least two of the groups. Details of the numbers of up and down-regulated proteins in each comparison are given in Table 1. The full data set, including FDR corrected p-values and all pairwise comparisons, is available as a supplementary data file. A t-SNE plot (t-distributed stochastic neighbor embedding) shows samples clustering primarily according to caste, with some overlap within worker and queen groups (Fig. 2). V1 separates queens and workers by caste, whereas V2 separates by reproductive status, with notable exceptions. A heatmap based on the scaled, log2 abundance values of all DE proteins shows clear clustering by the treatment with some overlap between queenright and queenless workers (Fig. 3). The same queenright worker sample (QRW1) that clustered with queenless workers in the t-SNE plot also clustered with queenless workers in the heatmap. This pattern could not be explained by either the maternal colony or the ovary status of this pool of bees which did not stand out compared to other samples. The highest number of DE and uniquely DE proteins (i.e., not DE in any other comparisons) was 82 proteins DE between queens and queenright workers of which 39 were uniquely DE. These groups vary not only by caste and ovarian activation but also by mating status. This and other comparisons are visualized in an Upset plot which shows the overlap of DE proteins between all pairwise

Table 1 Summary table of differentially-expressed proteins between pairwise comparisons of the treatment groups. QLW – queenless workers. QRW – queenright workers. "Up" refers to proteins which were significantly more abundant in the first group of each pairwise comparison, while "down" refers to those more abundant in the second group. For instance, there were 28 proteins more highly expressed in the queens of the "Queen vs Gyne" comparison, whereas there were 12 higher in the gynes

Comparison	Down	NotSig	Up	Groups differ in the status of
Queen vs. Gyne	12	2,437	28	Reproduction, Mating
QLW vs. Gyne	26	2,407	44	Reproduction, Caste
QRW vs. Gyne	11	2,424	42	Caste
QLW vs. Queen	10	2,439	28	Caste, Mating
QRW vs. Queen	26	2,395	56	Reproduction, Caste, Mating
QRW vs. QLW	16	2,446	15	Reproduction

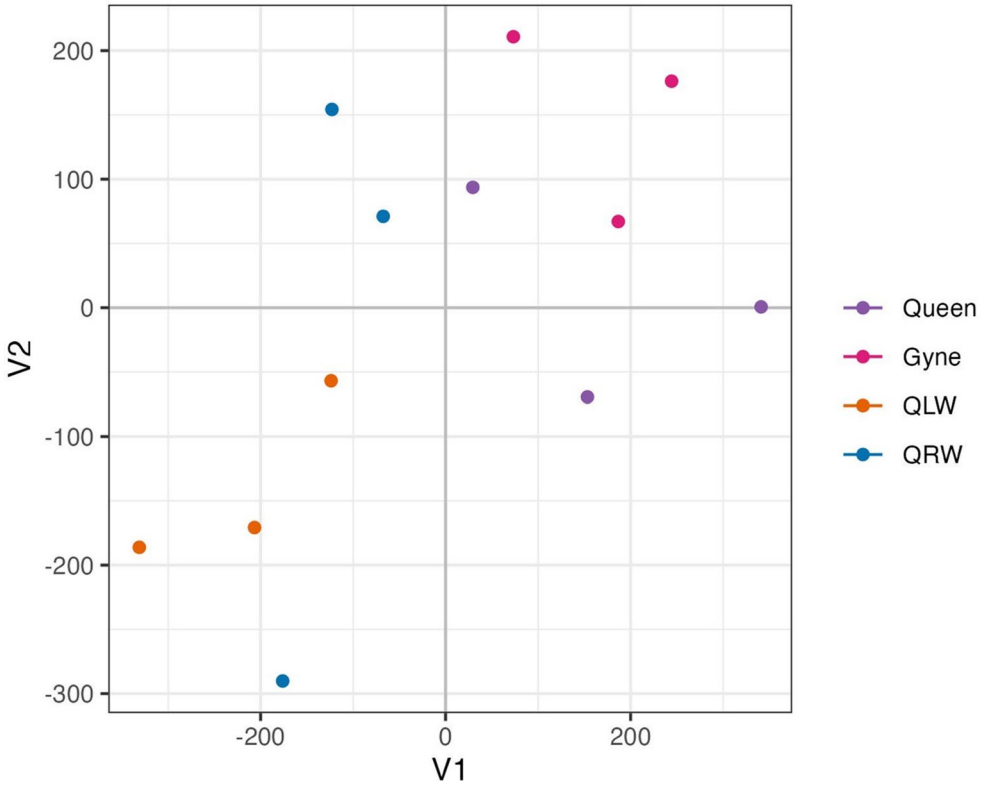

Fig. 2 t-SNE (t-distributed stochastic neighbor embedding) plot showing clustering of Dufour's gland proteomic samples based on treatment group (queen, gyne, QLW (queenless worker), QRW (queenright worker) using normalized abundance data from all identified proteins

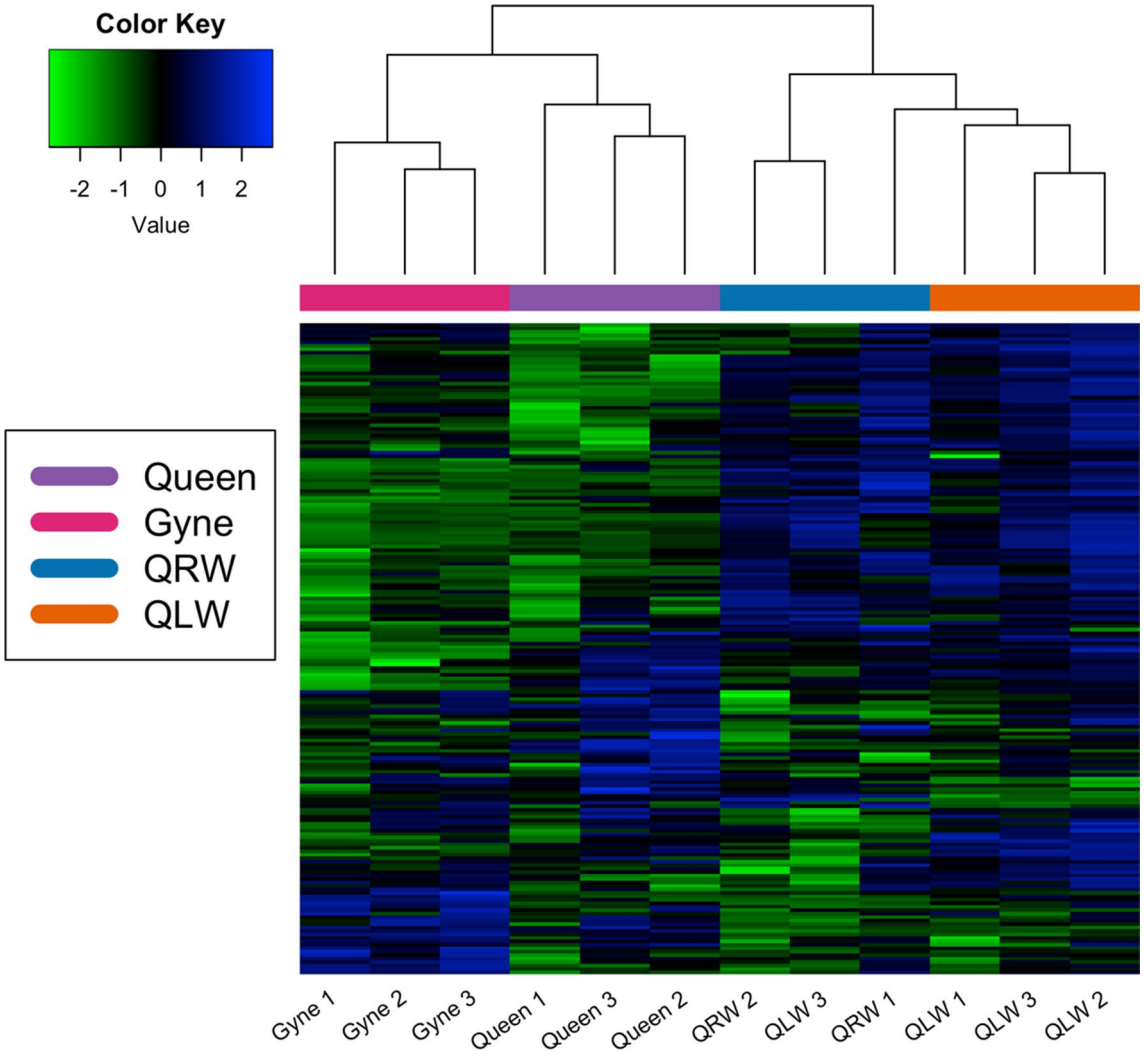

Fig. 3 Heatmap showing the scaled (mean = 0, sd = 1), log2 abundance values for the 188 proteins which were differentially expressed between the examined groups (3 samples per treatment). Rows are proteins and columns are samples. Colors at the tips of the dendrogram represent the treatment group (queen, gyne, QLW (queenless worker), QRW (queenright worker) of the samples

comparisons of the treatment groups (Fig. 4). Notably, 31 proteins were differentially expressed between worker groups (Fig. 5). Among these were proteins involved in fatty acid biosynthesis: four fatty acyl reductases upregulated in the QRW group (A0A6P3UUI7, A0A6P6FKS8, A0A6P8L4×3, A0A6P8M112) which convert fatty acids to fatty alcohols, a precursor of wax esters; a long chain fatty acid transport protein upregulated in QRW (A0A6P3URM9), and two proteins involved in reproduction which were up regulated in QLW: IRP30 (G7H7V6), an immune-responsive protein which promotes egg-laying in bumble bees [40], and vitellogenin (A0A6P-3DVD4), a transporter and precursor of egg proteins. Volcano plots of each pairwise comparison are shown in Fig. S5.

Fatty acid biosynthesis

One of the primary compounds of interest found in the Dufour's gland are wax esters, which are thought to be synthesized by a combination of fatty acyl reductases (FAR) and acyltransferase type wax synthases. We found 5 DE FAR proteins that showed variable expression patterns with respect to treatment group. One, A0A6P3UUI7, was most upregulated in gynes and queenright workers, the groups which produce the largest amount of wax esters. All five FARs were higher in QRW compared to QLW, but only four were significantly different (Fig. 6). Two of these (A0A6P8L4×3 and A0A6P8M112), had a log2 foldchange between 3 and 7 compared to gyne and queen groups, indicating remarkably caste-specific expression. In addition, a protein involved in chain shortening by B-oxidation (EHT, A0A6P3UTI4) was upregulated in queenright workers

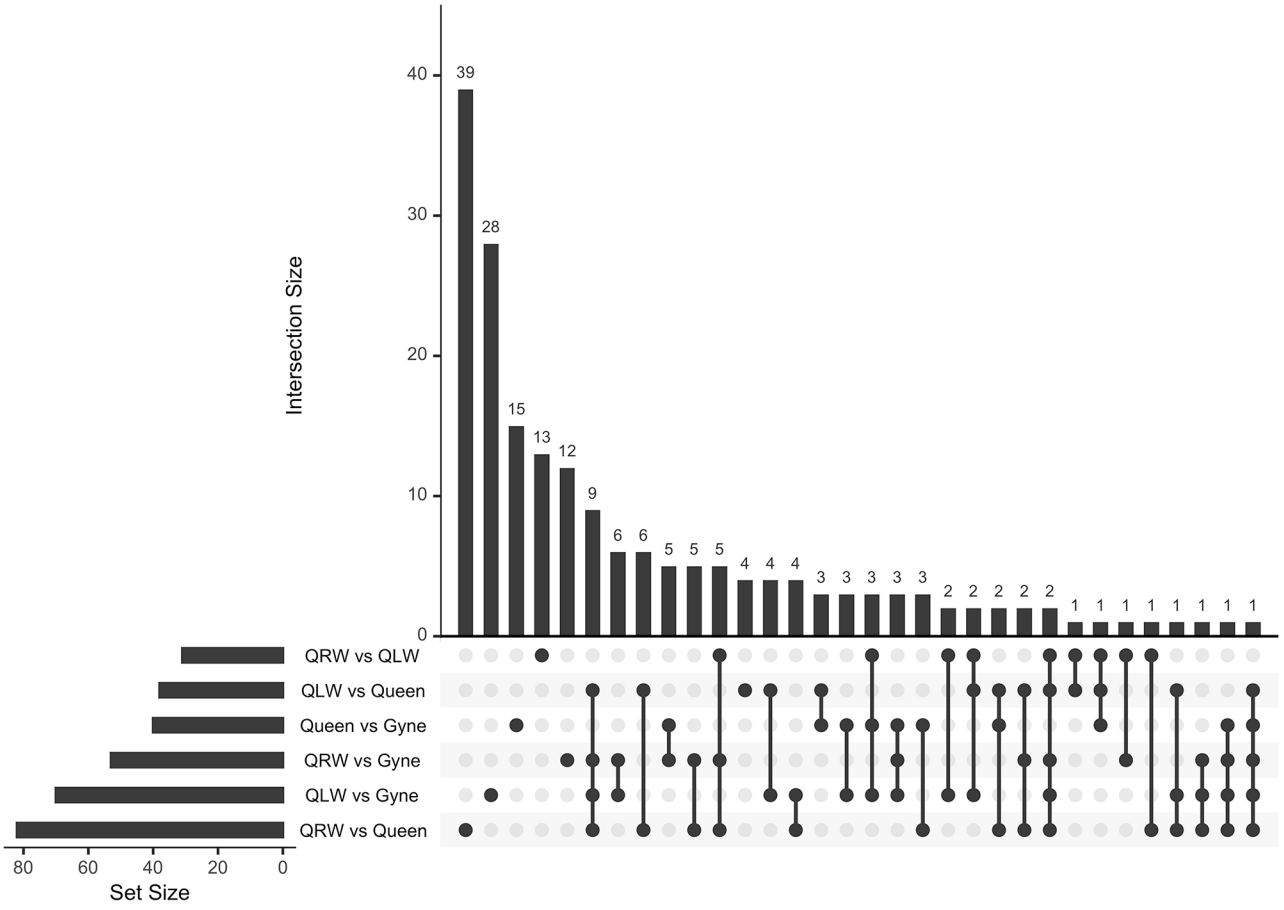

Fig. 4 Upset plot showing the intersection mode of overlap of differentially expressed proteins in the Dufour's gland between all pairwise comparisons. Pairwise comparisons were between the treatment groups (queen, gyne, QLW - queenless worker, QRW - queenright worker)

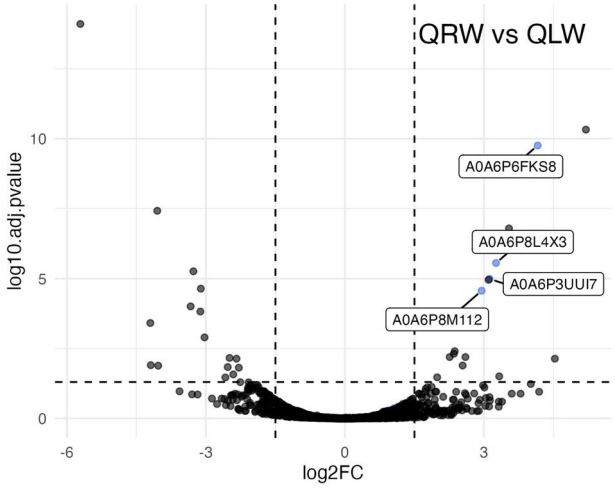

Fig. 5 Volcano plot showing comparison of protein expression between QRW and QLW. Vertical-dotted lines show a log2 fold change with a cutoff of ±1.5, and the horizontal-dotted line shows an FDR adjusted p-value with a cutoff of 0.05. Points colored blue are fatty acyl-CoA reductases (FAR). Four of seven identified FAR proteins are upregulated in QRW (which produce more wax esters than QLW), matching the chemical phenotype of these bees

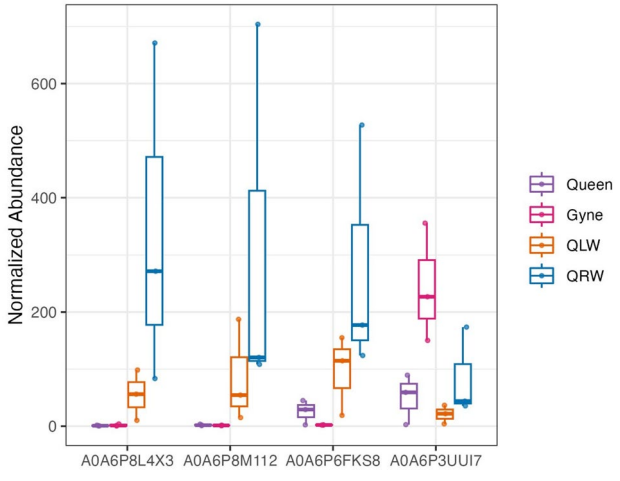

Fig. 6 Boxplots showing the normalized abundances of the four fatty acyl-CoA reductases (FARs) more highly expressed by queenright (QRW) than queenless (QLW) workers. Note that the first two FARs are highly worker specific

compared to queens, which matches the chemistry of workers who have shorter wax esters than queens in their Dufour's gland.

Isoprenoid biosynthesis

Hydroxymethylglutaryl-CoA synthase (HMGS, A0A6P8LCZ2) and Hydroxymethylglutaryl-CoA reductase (HMGR, A0A6P3V078) are key enzymes early in the isoprenoid biosynthetic pathway that is responsible for synthesizing terpenes and were most highly expressed in gynes. This matches the chemical phenotype of gynes who produce the most terpenoid compounds across all treatment groups including diterpenes and terpene esters. Isopentyl diphosphate isomerase (IPPI, A0A6P3E1J1) which catalyzes a later step in the pathway, was significantly upregulated in queens compared to gynes and queenright workers, and also in both worker groups compared to gynes (Fig. S2).

Comparison with previous RNA-seq study

After filtering to include proteins identified by a minimum of two peptides, shotgun proteomics identified proteins from 2469 genes or 25.8% of the 9569 genes in the previous RNA-seq analysis [14]. These proteins tended to correspond to transcripts that were found to be more highly expressed by RNA-seq, but still captured variation among the least abundant transcripts (Fig. 7). Of the 54 DE genes identified for their likely role in Dufour gland pheromone biosynthesis in [14], 24 were found in the current study, and 9 of these were DE as proteins. These 9 proteins included 4 FARs, a desaturase, a chain shortening enzyme, a component of an elongase, and two key proteins in the early isoprenoid biosynthesis pathway HMGS and HMGR (hydroxymethylglutaryl-CoA synthase and reductase). A comparison of the expression of these 9 genes between the transcript and protein levels is given in Table 2. We also recovered 29 proteins not found in the RNA-seq experiment, of which 8 were differentially expressed. Among them is an uncharacterized protein with hemolymph juvenile hormone binding protein domain (A0A6P8LBT1, *PFAM* - PF06585), indicating the transport of an important insect hormone involved in reproduction.

Deeploc 2.0 protein localizations

The majority of identified proteins were predicted to be localized in the cytoplasm, nucleus, or mitochondrion, however the differentially expressed proteins were significantly enriched in the "extracellular", "endoplasmic reticulum", and "cell membrane" categories (Fig. S3).

Discussion

Shotgun proteomics of the Dufour's gland in *B. impatiens* revealed protein expression specific to caste and reproductive state. Many of the nearly 200 DE proteins followed expression patterns that matched predictions

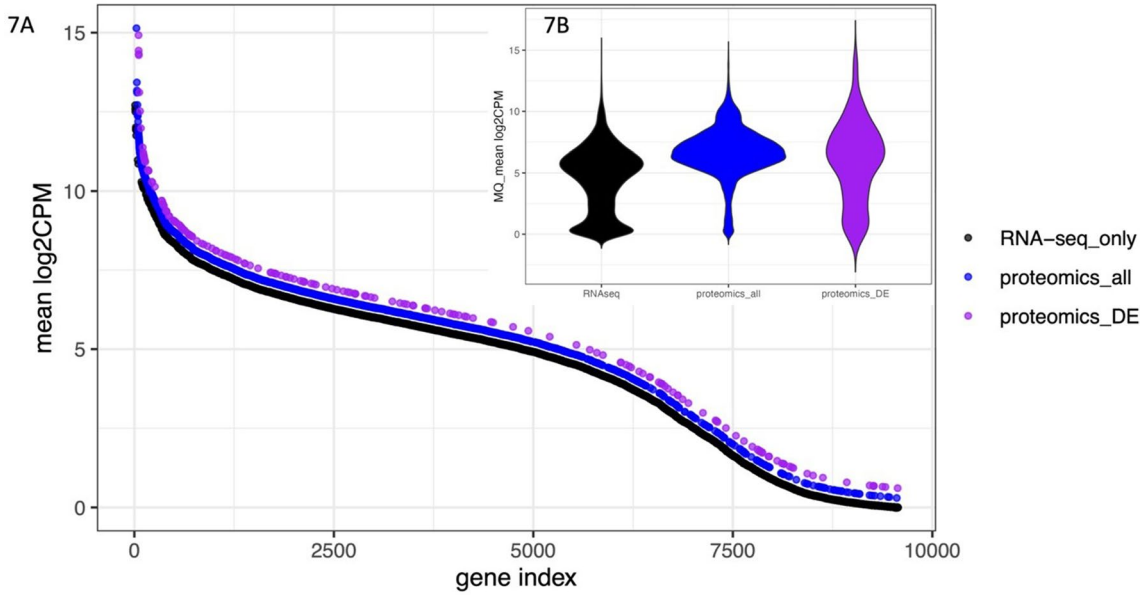

Fig. 7 Panel A-Comparison between identified proteins (current study) and transcripts (previous study) in the Dufour's gland. The y-axis shows mean log2CPM (TMM normalized counts per million) expression values from a previous RNA-seq study [14] and orders genes by descending value. Genes for which we also identified proteins with shotgun proteomics (this study) are shifted up and shown in blue, and those proteins which are differentially expressed are shifted up and shown in purple. Panel B - The inset violin plot shows the distributions of the genes that were found in the RNA-seq study (black), found in both proteomics and RNAseq studies (blue), and those that were differentially expressed in the proteomics study (purple). As expected, the set of identified proteins was biased towards more abundantly-expressed genes, but still captured variation in some of the least abundantly-expressed genes. For purpose of visualization, only the queen values are shown but the pattern is similar for all groups

Table 2 Comparison of differentially-expressed genes involved in the biosynthesis of Dufour's gland ester signals identified in a previous RNA-seq study [14] with protein expression in the current study. Pairwise comparisons were conducted between queens, gynes, queenless workers (QLW), and queenright workers (QRW). For each pairwise comparison, we provide the log2 fold change (logFC) and FDR adjusted p-value (adjP) in both RNAseq and proteomic studies. Cells shaded in teal show comparisons with concordance between proteomics and RNA-seq studies (i.e., the transcripts and proteins were significantly up or down regulated in the same direction). Orange cells show comparisons where the protein was differentially expressed but the transcript was not. Unshaded cells contain values which were either only differentially expressed in the RNAseq study or in neither study

		Queen vs. Gyne				**Queen vs. QLW**				**Queen vs. QRW**				**Gyne vs. QRW**				**Gyne vs. QLW**				**QLW vs. QRW**			
		Proteomics		**RNA-seq**		**Proteomics**		**RNA-seq**		**Proteomics**		**RNA-seq**		**Proteomics**		**RNA-seq**		**Proteomics**		**RNA-seq**		**Proteomics**		**RNA-seq**	
Accession	Description	logFC	adjP	logFC	adjP	logFC	adjP	logFC	adjP	logFC	adjP	logFC	adjP	logFC	adjP	logFC	adjP	logFC	adjP	logFC	adjP	logFC	adjP	logFC	adjP
A0A6P6F860	Fatty acyl-CoA reductase	2.98	**0.002**	9.83	**0.000**	2.51	**0.002**	7.43	**0.000**	2.40	0.051	8.67	**0.000**	-0.28	0.970	-1.16	0.659	0.00	0.966	-2.40	0.264	-0.26	0.434	1.23	0.958
A0A6P8M112	Fatty acyl-CoA reductase	0.04	0.999	3.06	**0.038**	-3.09	0.622	-3.08	0.181	-5.99	**0.000**	-3.57	0.111	-6.89	**0.000**	-6.63	**0.003**	-3.39	0.051	-6.15	**0.003**	-2.15	**0.000**	-0.49	0.991
A0A6P3DW33	Very-long-chain enoyl-CoA reductase	-2.16	0.633	0.28	0.565	-4.07	**0.025**	-0.60	0.292	-4.75	**0.000**	-0.77	0.145	-3.15	**0.016**	-1.05	**0.037**	-2.89	0.081	-0.88	0.072	-0.04	0.626	-0.16	0.981
A0A6P3V0×0	acyl-CoA Delta(11) desaturase isoform X2	-3.63	**0.032**	-1.21	0.075	-0.91	0.989	1.20	0.088	-1.51	0.565	0.70	0.319	1.49	0.925	1.91	**0.006**	2.72	**0.014**	2.41	**0.001**	-0.80	0.131	-0.50	0.958
A0A6P3V078	3-hydroxy-3-methyl-glutaryl coenzyme A reductase	-0.74	0.999	-0.06	0.969	1.50	0.191	3.49	**0.001**	2.05	0.196	3.13	**0.003**	3.32	0.097	3.19	**0.002**	2.43	**0.027**	3.54	**0.001**	0.67	1.000	-0.35	0.975
A0A6P3UUI7	Fatty acyl-CoA reductase	-2.45	0.506	-4.69	0.130	1.20	0.134	3.50	0.119	-0.62	1.000	1.97	0.412	2.39	0.466	6.66	**0.027**	4.92	**0.000**	8.19	**0.006**	-2.32	**0.000**	-1.53	0.958
A0A6P3UTI4	Very-long-chain (3R)-3-hydroxyacyl-CoA dehydratase	-1.27	0.982	-2.42	**0.000**	-2.83	0.723	-2.08	**0.000**	-3.62	**0.002**	-1.72	**0.000**	-1.86	0.201	0.70	0.126	-1.26	0.966	0.34	0.495	-0.60	0.253	0.36	0.958
A0A6P8LCZ2	Hydroxymethylglutaryl-CoA synthase	-3.13	0.136	-2.59	**0.002**	-0.10	0.865	0.55	0.527	0.29	1.000	-0.94	0.241	3.59	0.062	1.65	0.060	3.04	**0.005**	3.14	**0.000**	0.42	0.940	-1.49	0.925
A0A6P8L4×3	Fatty acyl-CoA reductase	-0.96	0.999	2.92	0.091	-4.76	**0.009**	-2.99	**0.045**	-7.21	**0.000**	-3.40	**0.021**	-5.88	**0.000**	-6.32	**0.002**	-4.79	**0.000**	-5.92	**0.003**	-2.46	**0.000**	-0.40	0.986

based on chemical phenotype and refine the understanding of the Dufour's gland as a site of signal biosynthesis for both fatty acid and isoprenoid derived signals. For instance, there was a clear differentiation of FAR enzymes that were most expressed in workers and gynes, and between QRW and QLW that differ in the amount of wax esters they produce. This indicates that the FAR enzymes catalyze different substrates to produce the different chemical phenotypes observed in gynes and workers, rather than a common FAR being supplied with different substrates. The de-coupling of these enzymes by caste may have facilitated the evolution of caste-specific chemical communication. Deciphering which of the 23 FAR genes in the *B. impatiens* genome [41] catalyze which reactions is important to understanding these evolutionary processes. Combining RNAseq and shotgun proteomic approaches helped narrow down the number of potential candidates and will facilitate future functional genomics work using gene silencing or heterologous expression techniques to map the biosynthesis pathways involved in pheromone production.

Dufour's gland proteomics has been examined in two other social insects, the honey bee *Apis mellifera* and the social wasp *Polybia paulista*. In each, enzymes related to pheromone biosynthesis and venom have been found [42, 43], underscoring the versatile and complex role played by the Dufour's gland across Hymenopterans. In *P. paulista*, the Dufour's gland produces a combination of aliphatic alkenes, alcohols, aldehydes, carboxylic acids, and methyl esters which are also found in the venom reservoir [42]. Of these, the methyl esters elicit the most alarm behavior from conspecifics. Of the 59 proteins that were identified from the gland, 15 were related to fatty acid biosynthesis, including chain elongation and shortening, desaturation, dehydrogenation, and an O-methyltransferase domain containing protein likely involved in the production of methyl esters. In *(A) mellifera*, because the Dufour's gland produces fatty-acid derived pheromones involved in the division of labor, the Dufour's gland protein content of nurse and forager bees were compared using 2D- gel electrophoresis followed by MALDI-TOF mass spectrometry [43]. They identified 131 spots that were differentially expressed and subsequently excised, 28 of which were identified. The identified proteins primarily related to protein metabolism and stress responses, heat shock proteins comprising the most upregulated proteins in foragers, and not directly fatty acid biosynthesis. However, this could be explained by the methodology, which only sought to identify differentially expressed proteins. The results reported here for *(B) impatiens* greatly expand the number of identified proteins in the Dufour's gland of any hymenopteran (from ~30–60 to ~2400) and presents the first caste-based comparison of an exocrine gland at the protein level.

One unique aspect of this study is the ability to compare patterns of gene expression gathered from proteomic and RNA-seq data. Shotgun proteomics identified ~25% of the genes found by RNA-seq [14]. The genes that were identified followed broadly similar patterns, but transcript level changes were often not observed at the protein level, indicating RNA-seq's higher sensitivity. Shotgun proteomics experiments generally sample the more abundant subset of a proteome for several reasons. First, proteomics is not amenable to the amplification technologies that lend incredibly high sensitivity to modern transcriptomics. Second, mass spectrometry-based proteomics typically uses abundance-based ion selection. The resulting expectation is that proteomics will be biased towards highly expressed genes [44]. In general, this was true for our data, but the approach taken here also recovered some of the most lowly-expressed transcripts in [14] (Fig. 6), indicating that this does not function as a hard cutoff. While differences between the datasets are in part due to methodological differences, another explanation is the many levels of regulation after translation that may affect protein expression and activity.

For instance, many post-translation modifications (PTMs) such as N-terminal acetylation, phosphorylation, and ubiquitylation, affect where proteins are localized within cells, how they bind to other proteins and substrates, and how quickly they are degraded [29, 45]. However, even before such modifications occur, an emerging body of evidence suggests that translation itself can be regulated through heterogeneous ribosomes which are selective for different mRNA transcripts [46]. Furthermore, there is evidence that ribosomes themselves can be modified (e.g., phosphorylated or ubiquitylated) with effects on translation [47]. These findings add additional layers of regulatory control to a key player in gene expression, the ribosome, which was previously thought to be highly conserved and invariant, processing whatever transcripts were produced. Determining whether such ribosomal or PTM regulatory mechanisms explain protein expression differences between workers of differing ovarian activation and chemical phenotype would be a fascinating area for future research. It is plausible that these mechanisms also help regulate gene expression between the queens and gynes, which show much greater transcriptional change.

There is also evidence from yeast that fatty acid biosynthetic processes are subject to posttranslational regulation. In yeast, the first and rate-limiting enzyme of fatty acid metabolism, acetyl-CoA carboxylase, is post-translationally regulated by Snf1 protein kinase, which plays roles in regulating lipid metabolism and response to nutritional stress [48, 49]. Phosphorylation with Snf1 inactivates Acc1 (a cytosolic yeast enzyme), which is also

involved in other processes, such as the B-oxidation of fatty acids.

The location of wax synthase activity seems to be highly conserved to the endoplasmic reticulum (ER) across the tree of life [50–52]. Based on this, we screened the protein localizations of the differentially expressed proteins between QR and QL workers and found four fatty acyl reductases, which are known to be active in the ER, and several other acyltransferase enzymes. Based on homology, the acyltransferase enzymes are predicted to produce sphingolipids, not wax esters, by catalyzing the incorporation of oleoyl-CoA to lysophosphatidylserine (A0A6P3DKV4, lysophospholipid aceyltransferase 1) and arachidonoyl-CoA to lysophosphatidylinositol (A0A6P-3DPR8, lysophospholipid aceyltransferase 7). However, the functions of these enzymes have not been directly investigated, and it is possible that these are candidates for wax synthase activity.

Two interesting questions that arise from the results presented here are: what is the advantage of post-transcriptional regulation of pheromone production, and why would workers exhibit it but gynes and queens would not? One possibility is that post-translational regulation allows a faster [53] or less energetically expensive option to alter pheromone production in line with the flexibility workers will have to exhibit in their reproductively dynamic social environment. This could include PTMs or changes to the subcellular location of mRNA binding proteins [54], which play a role, for instance, in the dynamic response to environmental stress [55]. This could make sense for social insect workers whose fitness depends on assessing the reproductive potential of the queen and either cooperating in raising her offspring or competing by laying their own eggs. Because *Bombus* workers compete to establish dominance hierarchies in the event of queen's death or after gyne production, and these contests involve chemical signaling, workers which lag in their ability to signal may be selected against. In contrast, the reproductive incentives never change with social context for queens, or gynes, reducing the necessity of this mechanism.

Such a mechanism is more likely to evolve in seasonal social insect species which have not yet passed the "point of no return" of extreme caste specialization [56], and where worker reproduction makes a significant contribution to overall male production. This is the case in several bumble bee species [57, 58], where workers produce a large proportion of the males (~20–85%, depending on the species). While some males are produced by workers in queenright conditions, the majority of males are queen-born and are produced after the queen is killed [59, 60]. At this point, the speed of reproduction is important, because if winter comes and there are no floral resources to feed the brood or there are no gynes in the population to mate with, then the potential fitness benefit for worker-born males is lost. Future studies could examine specific post-translational modifications between queens and workers of species with different levels of sociality, potentially focusing on enzymes involved in pheromone production, where known.

In species with more rigid social hierarchies and stronger reproductive division of labor, it is not known if post-transcriptional regulation is a factor, but current evidence suggests transcriptional changes are sufficient to explain pheromone production. For instance, in the biosynthesis of the queen mandibular pheromone in *A. mellifera*, there is a caste-based bifurcation point where stearic acid is hydroxylated in different positions to produce either the queen or worker versions of hydroxy acids [61]. Different *CYP* enzymes catalyze hydroxylation in different positions, and their gene expression matches the pheromonal phenotype found in queenright workers, queenless workers, and queens [62, 63].

Overall, our data provide evidence that pheromone biosynthesis in the Dufour's gland is caste specific, that gynes and workers are likely using different FAR enzymes to make their respective wax esters, and that for workers, the presence of the queen and the resulting changes to ovarian activation combine to alter the expression of fatty acid biosynthetic enzymes at the protein level, leading to differences in chemical phenotype. These data highlight the potentially different selective pressures operating on queens and workers in social insects and to our knowledge the first to highlight differences in regulatory level of signal production within members of the same species. We further show that shotgun proteomics is able to identify proteins from even lowly-expressed transcripts, demonstrating the utility of integrating transcriptional and proteomic analyses of pheromone biosynthesis.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-10873-3.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Author contributions

N.D. designed and carried out the experiments, analyzed the data and prepared the figures, and wrote the main manuscript text.E.A. designed experiments, helped write and revise the manuscript, and provided funding. T.L. carried out the proteomic mass spectrometry, assisted in data analysis, and contributed writing to the methods.All authors reviewed the manuscript.

Funding

This work was funded by NSF-CAREER IOS-1942127 to EA.

Data availability

Data presented in this manuscript is available in supplementary information files. The raw proteomic datasets generated and analyzed are available in

ProteomeXchange via the MassIVE repository, with the ProteomeXchange accession number PXD056402 and MassIVE ID MSV000095994.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 4 April 2024 / Accepted: 7 October 2024
Published online: 17 October 2024

References

1. Amsalem E. One problem, many solutions: female reproduction is regulated by chemically diverse pheromones across insects. Adv Insect Physiol. 2020;59:131.
2. Billen J. Exocrine glands and their key function in the communication system of social insects. Formos Entomol. 2011;31(2):75–84.
3. Keeling CI, Plettner E, Slessor KN. Hymenopteran Semiochemicals Top Curr Chem. 2004;239:133–77.
4. Hefetz A. The critical role of primer pheromones in maintaining insect sociality. Z Naturforsch C J Biosci. 2019;74(9–10):221–31.
5. Le Conte Y, Hefetz A. Primer pheromones in social hymenoptera. Annu Rev Entomol. 2008;53:523–42.
6. Mitra A. Function of the Dufour's gland in solitary and social Hymenoptera. J Hymenoptera Res, 2013. 35.
7. Abdalla FC, Cruz-Landim C. Dufour gland in the hymenopterans (Apidae, Formicidae, Vespidae); a review. Brazilian J Biology. 2001;6(1):95–106.
8. Cane JH. Dufour's gland secretion in the cell linings of bees (Hymenoptera: Apoidea). J Chem Ecol. 1981;7(2):403–10.
9. Dor R, Katzav-Gozansky T, Hefetz A. Dufour's gland pheromone as a reliable fertility signal among honeybee (Apis mellifera) workers. Behav Ecol Sociobiol. 2005;58(3):270–6.
10. Amsalem E, et al. Reproductive competition in the bumble-bee Bombus terrestris: do workers advertise sterility? Proc Biol Sci. 2009;276(1660):1295–304.
11. Amsalem E, et al. Dufour's gland secretion, sterility and foraging behavior: correlated behavior traits in bumblebee workers. J Insect Physiol. 2013;59(12):1250–5.
12. Amsalem E, Hefetz A. The appeasement effect of sterility signaling in dominance contests among Bombus terrestris workers. Behav Ecol Sociobiol. 2010;64(10):1685–94.
13. Derstine N, et al. Dufour's gland analysis reveals caste and physiology specific signals in Bombus impatiens. Sci Rep. 2021;11(1):2821.
14. Derstine N, et al. Differential gene expression underlying the biosynthesis of Dufour's gland signals in Bombus impatiens. Curr Res Insect Sci. 2023;3:100056.
15. Katzav-Gozansky T, Soroker V, Hefetz A. The biosynthesis of Dufour's gland constituents in queens of the honeybee (Apis mellifera). Invertebr Neurosci. 1997;3(2–3):239–43.
16. Michener CD. The social behavior of the bees: a comparative study. Harvard University Press; 1974.
17. Williams PH, et al. *Bumble bees of North America*, in *Bumble Bees of North America*. Princeton University Press; 2014.
18. Sramkova A, et al. Fertility signals in the bumblebee Bombus terrestris (Hymenoptera: Apidae). Naturwissenschaften. 2008;95(6):515–22.
19. Holstein SA, Hohl RJ. Isoprenoids: remarkable diversity of form and function. Lipids. 2004;39(4):293–309.
20. Vandermoten S, et al. Aphid alarm pheromone: an overview of current knowledge on biosynthesis and functions. Insect Biochem Mol Biol. 2012;42(3):155–63.
21. Gilg AB, Tittiger C, Blomquist GJ. Unique animal prenyltransferase with monoterpene synthase activity. Naturwissenschaften. 2009;96(6):731–5.
22. Tholl D. Biosynthesis of terpene pheromones in hemiptera/stink bugs. Insect pheromone Biochemistry and Molecular Biology. Elsevier; 2021. pp. 269–84.
23. Brabcova J, et al. Characterisation of Acetyl-CoA thiolase: the first enzyme in the biosynthesis of Terpenic Sex Pheromone Components in the labial gland of Bombus terrestris. ChemBioChem. 2015;16(7):1047–51.
24. Maisonnasse A, et al. E-beta-ocimene, a volatile brood pheromone involved in social regulation in the honey bee colony (Apis mellifera). PLoS ONE. 2010;5(10):e13531.
25. Valterová I, et al. Sexual attraction: a review of bumblebee male pheromones. Z Naturforsch C J Biosci. 2019;74(9–10):233–50.
26. Buccitelli C, Selbach M. mRNAs, proteins and the emerging principles of gene expression control. Nat Rev Genet. 2020;21(10):630–44.
27. Pieroni L, et al. Enrichments of post-translational modifications in proteomic studies. J Sep Sci. 2020;43(1):313–36.
28. Ree R, Varland S, Arnesen T. Spotlight on protein N-terminal acetylation. Exp Mol Med. 2018;50(7):1–13.
29. Deng S, Marmorstein R. Protein N-Terminal acetylation: structural basis, mechanism, versatility, and Regulation. Trends Biochem Sci. 2021;46(1):15–27.
30. Kustatscher G, Grabowski P, Rappsilber J. Pervasive coexpression of spatially proximal genes is buffered at the protein level. Mol Syst Biol. 2017;13(8):937.
31. Geiger T, Cox J, Mann M. Proteomic changes resulting from gene copy number variations in cancer cells. PLoS Genet. 2010;6(9):e1001090.
32. Stingele S, et al. Global analysis of genome, transcriptome and proteome reveals the response to aneuploidy in human cells. Mol Syst Biol. 2012;8(1):608.
33. Dephoure N, et al. Quantitative proteomic analysis reveals posttranslational responses to aneuploidy in yeast. Elife. 2014;3:e03023.
34. Treanore E, et al. Optimizing Laboratory Rearing of a key Pollinator, Bombus impatiens. Insects. 2021;12(8):673.
35. Orlova M, Starkey J, Amsalem E. A small family business: synergistic and additive effects of the queen and the brood on worker reproduction in a primitively eusocial bee. J Exp Biol, 2020. 223(Pt 3).
36. Starkey J, Brown A, Amsalem E. The road to sociality: brood regulation of worker reproduction in the simple eusocial bee Bombus impatiens. Anim Behav. 2019;154:57–65.
37. Röseler P-F. A technique for year-round rearing of Bombus terrestris (Apidae, Bombini) colonies in captivity. Apidologie. 1985;16(2):165–70.
38. Distler U, et al. Label-free quantification in ion mobility–enhanced data-independent acquisition proteomics. Nat Protoc. 2016;11(4):795–812.
39. Thumuluri V, et al. DeepLoc 2.0: multi-label subcellular localization prediction using protein language models. Nucleic Acids Res. 2022;50(W1):W228–34.
40. Tian Y, et al. IRP30 promotes worker egg-laying in bumblebee, Bombus terrestris (Hymenoptera: Apidae). Gene. 2021;776:145446.
41. Tupec M et al. Expansion of the fatty acyl reductase gene family shaped pheromone communication in Hymenoptera. Elife, 2019. 8.
42. Esteves FG, et al. Using a proteometabolomic approach to investigate the role of Dufour's gland in pheromone biosynthesis in the social wasp Polybia paulista. J Proteom. 2017;151:122–30.
43. Teixeira ADD, et al. Proteomic analysis in the Dufour's gland of Africanized Apis mellifera workers (Hymenoptera: Apidae). PLoS ONE. 2017;12(5):e0177415.
44. Ankney JA, Muneer A, Chen X. Relative and absolute quantitation in mass spectrometry–based proteomics. Annual Rev Anal Chem. 2018;11:49–77.
45. Walsh CT, Garneau-Tsodikova S, Gatto GJ Jr. Protein posttranslational modifications: the chemistry of proteome diversifications. Angew Chem Int Ed. 2005;44(45):7342–72.
46. Xue S, Barna M. Specialized ribosomes: a new frontier in gene regulation and organismal biology. Nat Rev Mol Cell Biol. 2012;13(6):355–69.
47. Simsek D, Barna M. An emerging role for the ribosome as a nexus for post-translational modifications. Curr Opin Cell Biol. 2017;45:92–101.
48. Shi S, et al. Improving production of malonyl coenzyme A-derived metabolites by abolishing Snf1-dependent regulation of Acc1. MBio. 2014;5(3):pmbio101128–01130.
49. Hedbacker K, Carlson M. SNF1/AMPK pathways in yeast. Front Bioscience: J Virtual Libr. 2008;13:2408.
50. Cheng JB, Russell DW. Mammalian wax biosynthesis. II. Expression cloning of wax synthase cDNAs encoding a member of the acyltransferase enzyme family. J Biol Chem. 2004;279(36):37798–807.
51. Li F, et al. Identification of the wax ester synthase/acyl-coenzyme A: diacylglycerol acyltransferase WSD1 required for stem wax ester biosynthesis in Arabidopsis. Plant Physiol. 2008;148(1):97–107.
52. Wenning L, et al. Establishing very long-chain fatty alcohol and wax ester biosynthesis in Saccharomyces cerevisiae. Biotechnol Bioeng. 2017;114(5):1025–35.

53. Shimoni Y, et al. Regulation of gene expression by small non-coding RNAs: a quantitative view. Mol Syst Biol. 2007;3(1):138.
54. Cammas A, et al. Post-transcriptional control of gene expression through subcellular relocalization of mRNA binding proteins. Biochem Pharmacol. 2008;76(11):1395–403.
55. Holcik M, Sonenberg N. Translational control in stress and apoptosis. Nat Rev Mol Cell Biol. 2005;6(4):318–27.
56. Wilson EO. The insect societies. Cambridge, Massachusetts, USA: Harvard University Press; 1971.
57. Huth-Schwarz A, et al. Workers dominate male production in the neotropical bumblebee Bombus wilmattae (Hymenoptera: Apidae). Front Zool. 2011;8:1–6.
58. Owen RE, Plowright R. Worker-queen conflict and male parentage in bumble bees. Behav Ecol Sociobiol. 1982;11:91–9.
59. Takahashi J-i, et al. Male parentage and queen mating frequency in the bumblebee Bombus ignitus (Hymenoptera: bombinae). Ecol Res. 2008;23(6):937–42.
60. Alaux C, et al. Does the queen win it all? Queen-worker conflict over male production in the bumblebee, Bombus terrestris. Naturwissenschaften. 2004;91(8):400–3.
61. Plettner E, Slessor KN, Winston ML. Biosynthesis of mandibular acids in honey bees (Apis mellifera): De novo synthesis, route of fatty acid hydroxylation and caste selective β-oxidation. Insect Biochem Mol Biol. 1998;28(1):31–42.
62. Malka O, et al. The gene road to royalty–differential expression of hydroxylating genes in the mandibular glands of the honeybee. FEBS J. 2009;276(19):5481–90.
63. Wu Y, et al. Comparative transcriptome analysis on the synthesis pathway of honey bee (Apis mellifera) mandibular gland secretions. Sci Rep. 2017;7(1):4530.

Publisher's note