Insect Molecular Biology

Insect Molecular Biology (2016) **25**(3), 295–314 doi: 10.1111/imb.12221

Exploring complex pheromone biosynthetic processes in the bumblebee male labial gland by RNA sequencing

A. Buček*[1], J. Brabcová*[1], H. Vogel†, D. Prchalová*, J. Kindl*, I. Valterová* and I. Pichová*

**Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Prague, Czech Republic; and †Max Planck Institute for Chemical Ecology, Jena, Germany*

Abstract

Male marking pheromones (MPs) are used by the majority of bumblebee species (Hymenoptera: Apidae), including a commercially important greenhouse pollinator, the buff-tailed bumblebee (*Bombus terrestris*), to attract conspecific females. MP biosynthetic processes in the cephalic part of the bumblebee male labial gland (LG) are of extraordinary complexity, involving enzymes of fatty acid and isoprenoid biosynthesis, which jointly produce more than 50 compounds. We employed a differential transcriptomic approach to identify candidate genes involved in MP biosynthesis by sequencing *Bombus terrestris* LG and fat body (FB) transcriptomes. We identified 12 454 abundantly expressed gene products (reads per kilobase of exon model per million mapped reads value $>$ 1) that had significant hits in the GenBank nonredundant database. Of these, 876 were upregulated in the LG ($>$ 4-fold difference). We identified more than 140 candidate genes potentially involved in MP biosynthesis, including esterases, fatty acid reductases, lipases, enzymes involved in limited fatty acid chain shortening, neuropeptide receptors and enzymes involved in biosynthesis of triacylglycerols, isoprenoids and fatty acids. For selected candidates, we confirmed their abundant expression in LG using quantitative real-time reverse transcription-PCR (qRT-PCR). Our study shows that the *Bombus terrestris* LG transcriptome reflects both fatty acid and isoprenoid MP biosynthetic processes and identifies rational gene targets for future studies to disentangle the molecular basis of MP biosynthesis. Additionally, LG and FB transcriptomes enrich the available transcriptomic resources for *Bombus terrestris*.

Keywords: RNA-seq, transcriptome, *Bombus terrestris*, labial gland, marking pheromone biosynthesis, apoptosis.

First published online 4 March 2016.

Correspondence: Iva Pichová or Irena Valterová, Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, V.V.I. Flemingovo N. 2, 166 10 Prague 6, Czech Republic. Tel.: + 420 220 183 251; e-mails: iva.pichova@uochb.cas.cz (Iva Pichová); irena@uochb.cas.cz (Irena Valterová)

[1]These authors contributed equally to this work.

Introduction

In the majority of bumblebee species (Apidae: *Bombus*) with described mate-finding strategies, males produce marking pheromones (MPs), which they deposit on prominent objects along their flight routes to attract conspecific females (Kullenberg *et al.*, 1973; Ayasse & Jarau, 2014). In bumblebee species that exhibit marking behaviour, the MP-producing cephalic part of the labial gland (LG) fills a large part of the male's head, whereas the females have relatively small LGs (Terzo & Coppens, 2007; Albert *et al.*, 2014). By contrast, honey bee *Apis mellifera* males do not accumulate cephalic LG secretions (Poiani & Da Cruz-Landim, 2010).

MP biosynthesis has traditionally been studied in the buff-tailed bumblebee, *Bombus terrestris* (Ayasse & Jarau, 2014), an important crop pollinator in commercial greenhouses (Velthuis & Van Doorn, 2006). The major compounds in *Bombus terrestris* LG extracts that are active in electroantennography are terpenoid compounds (2,3-dihydrofarnesol, geranylcitronellol, 2,3-dihydrofarnesal) and fatty acid (FA) derivatives, including fatty alcohols (hexadecan-1-ol, octadeca-9,12,15-trien-1-ol) and fatty acyl esters (ethyl dodecanoate) (Žáček *et al.*, 2009). In total, more than 50 compounds have been identified in *Bombus terrestris* LG extracts, suggesting complex biosynthetic processes underlying MP biosynthesis (Coppée *et al.*, 2008). The large quantity of MP, which culminates at approximately 6 mg 7 days after

hatching (Žáček *et al*., 2009), facilitates comprehensive identification and quantification of MP components of diverse biochemical origins (Kullenberg *et al*., 1970; Žáček *et al*., 2009). Established protocols for year-round breeding of *Bombus terrestris* provides a continuous source of specimens of defined ages and physiological condition. This allows for the investigation of temporal changes in MP composition, along with other aspects of *Bombus terrestris* physiology and behaviour (Šobotník *et al*., 2008; Žáček *et al*., 2009, 2013, 2015; Jiroš *et al*., 2011; Jarau *et al*., 2012; Brabcová *et al*., 2013a,b, 2015; Buček *et al*., 2013; Prchalová *et al*., 2016).

Sesquiterpene and diterpene derived compounds constituting the majority of MPs in *Bombus terrestris* and also in *Bombus hortorum* and *Bombus pratorum* (Kullenberg *et al*., 1970) are produced *de novo* in the LG (Žáček *et al*., 2013) via enzymes of the isoprenoid pathway (Prchalová *et al*., 2016). *de novo* terpenoid pheromone biosynthesis has been extensively studied in male bark beetles (Coleoptera: *Ips* spp.), which produce monoterpene aggregation pheromones via the classical isoprenoid pathway (Tillman *et al*., 1999; Gilg *et al*., 2005). Pheromone biosynthesis in bark beetles is controlled by juvenile hormone (JH), which increases the transcript abundance of 3-hydroxy-3-methylglutaryl–coenzyme A reductase (HMGR) (Tittiger *et al*., 1999) – an enzyme that also regulates isoprenoid biosynthesis in vertebrates (Goldstein & Brown, 1990). Enzymes catalysing subsequent sesquiterpene formation are well described in insect JH biosynthesis (Bellés *et al*., 2005) and alarm pheromone biosynthesis (Lewis *et al*., 2008). Diterpene formation has been described and biosynthetic enzymes tentatively identified in termites (Hojo, 2007; Hojo *et al*., 2012) and terpene-derivatizing enzymes have been identified in various insects and other organisms (eg Addlesee *et al*., 1996; Nishimura & Eguchi, 2006; Cao *et al*., 2009; Mayoral *et al*., 2009). Together, homologues of these enzymes provide candidates for the biosynthetic pathway of terpene-derived MPs in bumblebees (Fig. 1A).

As a model for FA-derived MP biosynthesis in bumblebees we propose to employ female moth sex pheromone biosynthesis because (1) both female moth sex pheromones and bumblebee FA-derived MPs are similar in structure (ie saturated or unsaturated FAs with derivatized carboxyl group) and therefore presumably produced via similar biosynthetic pathways; (2) fatty acid desaturases, key sex pheromone biosynthetic enzymes in moths (Tillman *et al*., 1999), are also abundantly expressed in male LGs of several bumblebee species (Matoušková *et al*., 2008; Buček *et al*., 2013); and (3) pheromone biosynthesis in both moths and bumblebee males is associated with lipase activity (Du *et al*., 2012a; Brabcová *et al*., 2013b; Zhang *et al*., 2014). The proposed MP biosynthetic pathway is depicted in Fig. 1B. In moths, a multitude of proteins involved in FA-derived pheromone biosynthesis has been identified, including acyl-coenzyme A-binding protein (ACBP; Matsumoto *et al*., 2001), fatty acid reductase (FAR; Moto *et al*., 2003), pheromone biosynthesis activating neuropeptide receptor (PBANr; Choi *et al*., 2003), fatty acid desaturase (Moto *et al*., 2004), acetyl coenzyme A carboxylase (ACC; Tsfadia *et al*., 2008), fatty acid transport protein (FATP; Ohnishi *et al*., 2009), lipid storage droplet protein (LSD; Ohnishi *et al*., 2011), diacylglycerol acyltransferase 2 (DGAT2; Du *et al*., 2012b), lipases (Du *et al*., 2012a,b; Zhang *et al*., 2014) and glycerol-3-phosphate O-acyltransferase (GPAT; Du *et al*., 2015). Generally, the transcripts coding for these proteins are specifically and abundantly expressed in the moth pheromone gland (PG). In contrast to *de novo* biosynthesis of pheromones in the moth PG, MP biosynthesis in bumblebees is supplied partially by transport of MP precursors from the FB, presumably in the form of diacylglycerols (DAGs; Žáček *et al*., 2013, 2015). In particular, the highly similar composition of FA pools in the FB and LG in some bumblebee species points to a major role for FA transport from the FB to the LG (Cvačka *et al*., 2006; Kofroňová *et al*., 2014). However, the relative contribution of alternative fatty-acyl-supplying pathways to MP biosynthesis is not known (Žáček *et al*., 2013).

In general, regulation of insect pheromone biosynthesis is not conserved across insects and involves a variety of hormones (Tillman *et al*., 1999). In moths, female sex pheromone biosynthesis is regulated by pheromone biosynthesis activating neuropeptide (PBAN; Tillman *et al*., 1999). PBAN activates a pyrokinin/PBAN G-protein-coupled receptor, which subsequently mobilizes a cascade of second messengers such as calcium ions and cyclic adenosine monophosphate (Choi *et al*., 2003). The second messengers activate a range of enzymes including protein kinases, which in turn phosphorylate lipid droplet associated proteins such as LSD and thus activate the triacylglycerols (TAGs) in lipid droplets for lipolytic cleavage and release of pheromone precursors in the form of fatty acyls (Ohnishi *et al*., 2011). PBAN can control distinct biosynthetic steps depending on the moth species, for example, the ACC-catalysed step of FA biosynthesis (Tsfadia *et al*., 2008), reduction of FAs to fatty alcohols (Ozawa *et al*., 1993) or aldehydes (Fang *et al*., 1995), and lipolytic release of the FA pheromone precursors stored in TAGs (Fang *et al*., 1996; Ohnishi *et al*., 2011). The pyrokinin/PBANr has also been identified in *Drosophila melanogaster* (Park *et al*., 2002) and mosquitoes (Olsen *et al*., 2007), and is predicted to be ubiquitous in insects (Jurenka & Nusawardani, 2011). In hymenopteran insects, pyrokinin/PBAN-like peptides have been identified in the fire ant *Solenopsis invicta* (Choi *et al*., 2009) and in *Ap. mellifera* (Hummon *et al*., 2006). However, the role

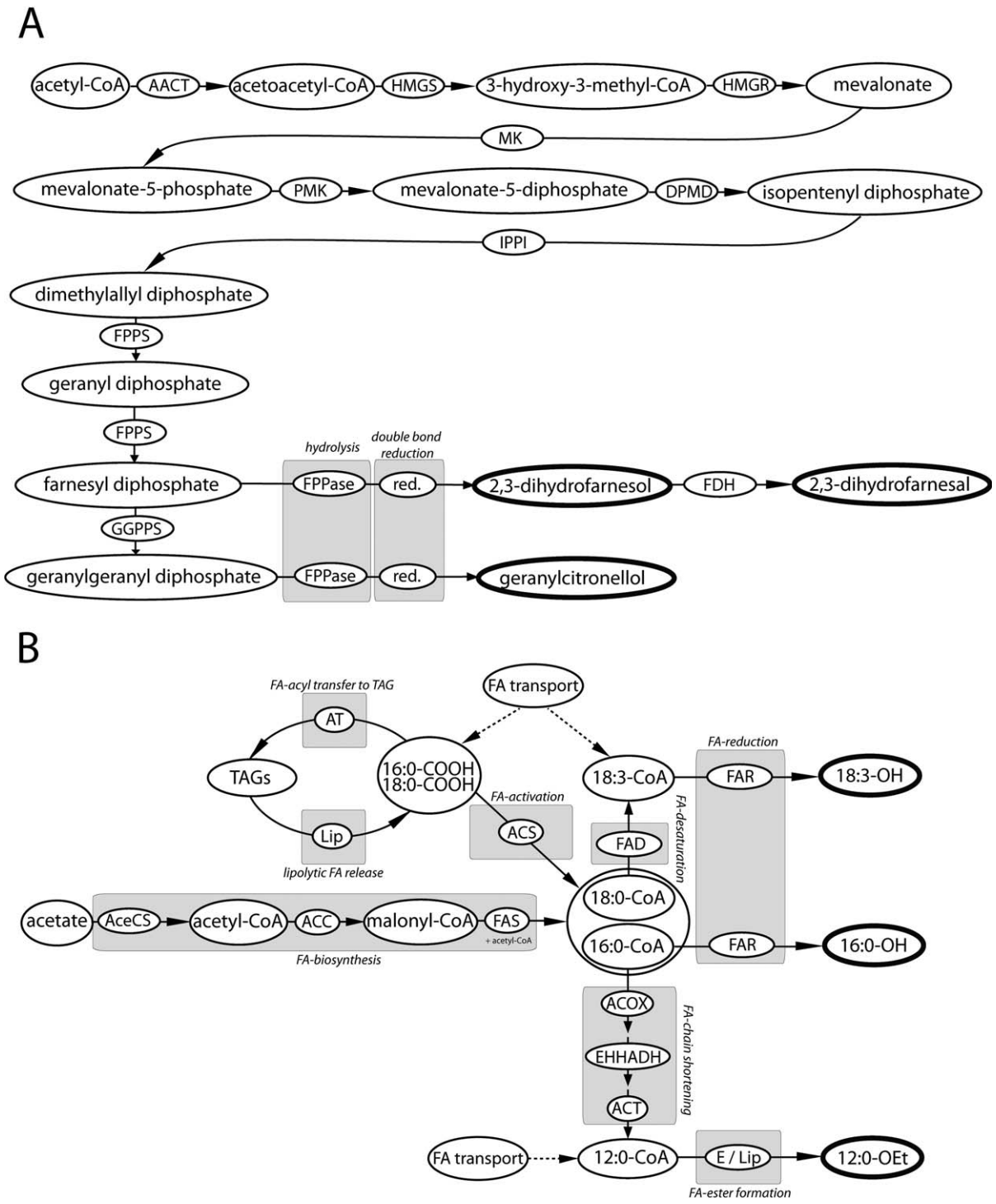

Figure 1. Proposed *Bombus terrestris* marking pheromone (MP)-biosynthetic scheme. Putative enzymes and intermediate compounds of (A) isoprenoid MP biosynthesis and (B) fatty acid (FA)-derived MP biosynthesis are displayed. Major *Bombus terrestris* MP components representing the final products of the biosynthetic pathways are highlighted in bold frames and possible FA-transport routes are indicated by dashed lines. The proposed MP-biosynthetic steps are based on the biosynthetic enzymes and pathways described in other insect species as cited in the main text. AACT, acetoacetyl-CoA thiolase; HMGS, 3-hydroxy-3-methylglutaryl coenzyme A synthase; HMGR, 3-hydroxy-3-methylglutaryl coenzyme A reductase; MK, mevalonate kinase; PMK, phosphomevalonate kinase; DPMD, diphosphomevalonate decarboxylase; IPPI, isopentenyl diphosphate isomerase; FPPS, farnesyl diphosphate synthase; GGPPS, geranylgeranyl diphosphate synthase; FPPase, farnesyl phosphatase; red., hypothetical reductase; FDH, farnesyl dehydrogenase; AceCS, acetyl-coenzyme A (acetyl-CoA) synthase; ACC, acetyl-CoA carboxylase; FAS, fatty acid synthase; FAD, fatty acid desaturase; AT, acyl transferase; TAGs, triacylglycerols; Lip, lipase; ACS, acyl-CoA synthases; FAR, fatty acid reductase; ACOX, acyl-CoA oxidases; EHHADH, peroxisomal enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase; ACT, 3-ketoacyl-CoA thiolase; E, esterase.

of PBAN in pheromone biosynthesis regulation in nonlepidopteran species is largely unknown. Experiments suggesting the role of PBAN in pheromone biosynthesis regulation were performed in the fly *Mayetiola destructor* (Foster *et al*., 1991), and recently, the role of PBAN and PBANr in terpenic trail pheromone biosynthesis was discovered in *S. invicta* (Choi & Vander Meer, 2012).

The complexity of processes connected to MP production in the *Bombus terrestris* LG is further increased by the occurrence of secretory cell apoptosis starting at day 5 in the life cycle of the bumblebee male imago (Šobotník *et al*., 2008). The biological role of secretory cell apoptosis in the LG is unclear, as the phenomenon is unparalleled in any other insect pheromone gland.

The advent of next-generation sequencing, particularly sequencing of cDNA libraries [RNA sequencing (RNA-seq)], has enabled transcriptomic studies of nonmodel organisms and provided access to the molecular basis of diverse insect biology (reviewed by Oppenheim *et al*., 2015). Although our study of individual MP-biosynthetic steps was boosted by the recently available *Bombus terrestris* genomic and transcriptomic data (Buček *et al*., 2013; Sadd *et al*., 2015), knowledge of the interplay of the metabolic processes underlying the complex *Bombus terrestris* MP biosynthesis remains limited.

To provide a framework for integration of the accumulated knowledge about the physiology and biochemistry of male bumblebee LGs, we present results from *Bombus terrestris* LG and FB transcriptome sequencing and compare them with available male head, queen head and queen ovary transcriptomes (Sadd *et al*., 2015). We describe over 140 MP-biosynthetic candidate genes, such as lipases and enzymes involved in TAG biosynthesis, proteins involved in FA transport and binding, FA biosynthetic enzymes, enzymes involved in derivatization of FAs to final MP components, biosynthetic enzymes involved in production of isoprenoid MP components, genes involved in PBAN-mediated regulation of pheromone biosynthesis and genes related to programmed cell death of LG secretory cells. For selected MP-biosynthetic gene candidates, we confirmed their upregulation in LG as compared to FB and whole queen head using quantitative real-time reverse transcription-PCR (qRT-PCR).

Results and discussion

RNA-seq, transcriptome assembly and gene product annotation

Sequenced cDNA libraries were prepared from RNA isolated from male LGs and FBs. Tissues from laboratory-bred specimens of defined ages (0-day-old pharate imago and 1-day-, 3-day- and 5-day-old imagoes) were pooled to cover transcripts involved in different stages of LG development and pheromonogenesis. HiSeq2500 Illumina sequencing yielded 40 000 000 reads for the LG and 30 000 000 reads for the FB, which were assembled into a total of 38 564 contigs (Buček *et al*., 2013; Supporting Information Data set S1).

The average contig length was 864 bp and N50 was 1351 bp. Nearly half of contigs (16 779; 43.5%) had a significant hit (E-value $< 10^{-5}$) in the GenBank nonredundant (NR) protein database. The complete RNA-seq data were submitted to the European Nucleotide Archive (http://www.ebi.ac.uk/ena) under study accession number PRJEB9937 (http://www.ebi.ac.uk/ena/data/view/PRJEB9937). Each contig was annotated with the highest-scoring BLAST hits retrieved in BLAST search of the National Center for Biotechnology Information (NCBI) NR database. We annotated 12 775 contigs (33.1% of all contigs and 76.1% of contigs with a significant BLAST hit) with gene ontology (GO) terms (Data set S1) using BLAST2GO software.

Analysis of GO terms enriched in the LG and FB of Bombus terrestris *males*

To uncover the biological processes enriched in the LG and FB, we identified GO terms significantly enriched (False discovery rate-corrected *P*-value < 0.05) in FB and LG. We selected a reads per kilobase of exon model per million mapped reads (RPKM) cut-off value of > 1 in at least one of the sequenced libraries to include a contig in the GO-enrichment analysis. By employing this threshold, contigs with overall low relative abundances were excluded. For a contig to be regarded as differentially expressed, the minimal fold change (ie ratio of RPKM values in the LG and FB) was set with moderately high stringency to a value of 4. By combining these cut-off values, we identified 9816 abundantly expressed contigs with GO annotation, of which 656 contigs were upregulated in the LG and 1421 were upregulated in the FB (Fig. 2). A large percentage (56.5%) of contigs from the combined LG and FB libraries did not have a significant BLAST hit (E-value $> 10^{-5}$). Contigs without a significant BLAST hit presumably represent noncoding transcripts, untranslated transcript regions, short transcript fragments and putative orphan genes. Notably, 1362 of the contigs without a significant BLAST hit were enriched in the LG as compared to the FB (fold change > 4) and also abundantly expressed in the LG (RPKM > 1). Some of these contigs might represent genes associated with bumblebee LG metabolism that do not have homologues in related insect species and that are not annotated in the available *Bombus terrestris* genome (Data set S1).

GO terms that were significantly over-represented in the LG were primarily related to biosynthesis of FAs and isoprenoids, namely the FA biosynthetic process, stearoyl-coenzyme A-9-desaturase activity, the terpenoid biosynthetic process and the pheromone biosynthetic process (Fig. 3). These data emphasize the LG as a

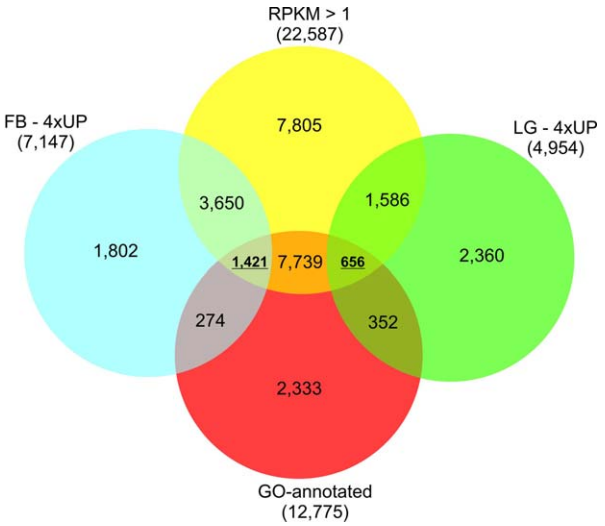

Figure 2. Venn diagram showing numbers of contigs with assigned gene ontology (GO) annotations, abundantly expressed contigs [reads per kilobase of exon model per million mapped reads (RPKM) value > 1], and contigs substantially upregulated (fold change > 4) in either the labial gland (LG-4xUP) or fat body (FB-4xUP). The numbers of GO-annotated contigs that are abundantly and differentially expressed in the FB and LG and were used as a test-set in the GO-enrichment analysis are underlined.

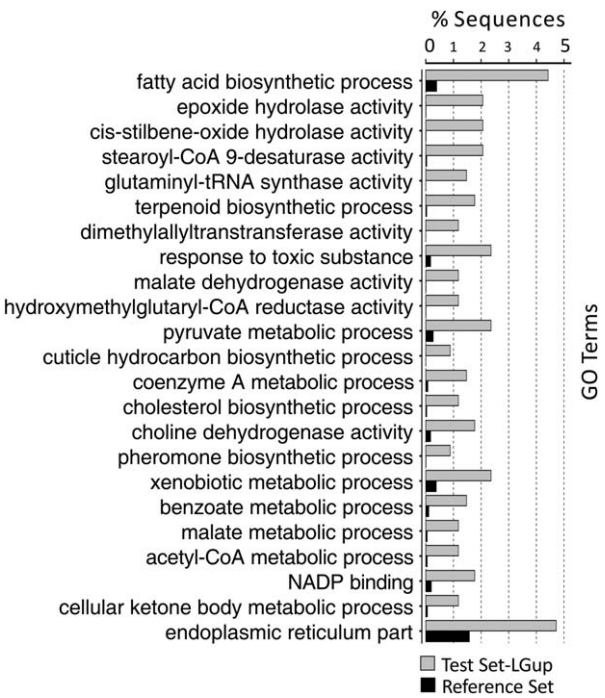

Figure 3. Gene ontology (GO) terms enriched in the *Bombus terrestris* labial gland (LG). Bar charts show the GO terms that were significantly ($P < 0.05$) enriched in the *Bombus terrestris* LG library. The GO terms are sorted in descending order according to their *P*-values. Only the most specific GO terms are displayed. For more general GO terms, see Fig. S1. CoA, coenzyme A; NADP, Nicotinamide adenine dinucleotide phosphate.

gland highly active in the biosynthesis of both terpenic and FA-derived MP components. The significantly enriched terms connected to FA biosynthesis strongly support the hypothesis that a substantial proportion of FA-derived MP components are biosynthesized *de novo* in the LG rather than supplied exclusively by transport from other tissues, such as the FB (Fig. 3).

The FB is a highly metabolically and biosynthetically active insect organ of general importance for energy storage and management, and has been extensively studied (reviewed by Arrese & Soulages, 2010). GO categories connected to metabolism of FAs, saccharides, nucleotides and amino acids were over-represented in the FB ($P < 0.05$), as expected because the FB serves as a major site of primary metabolism (Arrese & Soulages, 2010). All GO terms over-represented in the FB and LG ($P < 0.05$) are displayed in Figs S1 and S2, respectively. Besides GO terms connected to metabolism, GO terms connected to viral infection (RNA-directed RNA polymerase activity, viral genome replication and RNA helicase activity) were enriched in the FB (Fig. S2). We identified two sources of contigs annotated with GO terms related to viral infection, ie (1) contigs encoded by *Bombus terrestris*, categorized as 'retrovirus receptor', 'adenovirus-interacting', 'retrovirus-related transposon', etc. and (2) sequences from a virus very closely related or identical to the acute bee paralysis virus (ABPV, Data set S1). In all of the RNA-seq samples, including the male heads, queen heads and ovaries obtained from public databases, a low percentage of reads mapped to these viral sequences. The number of reads mapped to a total of 34 viral contigs was approx. 42 000 (0.14% of the total number of LG reads) for the LG and 150 000 (0.6% of the total number of FB reads) for the FB (Table S1). These data indicate that the *Bombus terrestris* specimens sequenced were infected by an ABPV-related virus, which is presumed to cause common inapparent infection in adult bees (reviewed by de Miranda *et al.*, 2010, Azzami *et al.*, 2012). In agreement with the frequent absence of ABPV symptoms at an individual or colony level, we did not observe any symptoms of viral infection in the bumblebees used for this study (data not shown). Although a low-level occurrence of this virus apparently exists in independent *Bombus terrestris* populations, ie this study and *Bombus terrestris* specimens sequenced by Sadd *et al.* (2015), the overall amount of total reads mappable to this virus is small (Table S1).

Gene groups potentially involved in pheromone biosynthesis and its regulation

In a gene-targeted analysis of the *Bombus terrestris* transcriptomes, we focused on individual gene products

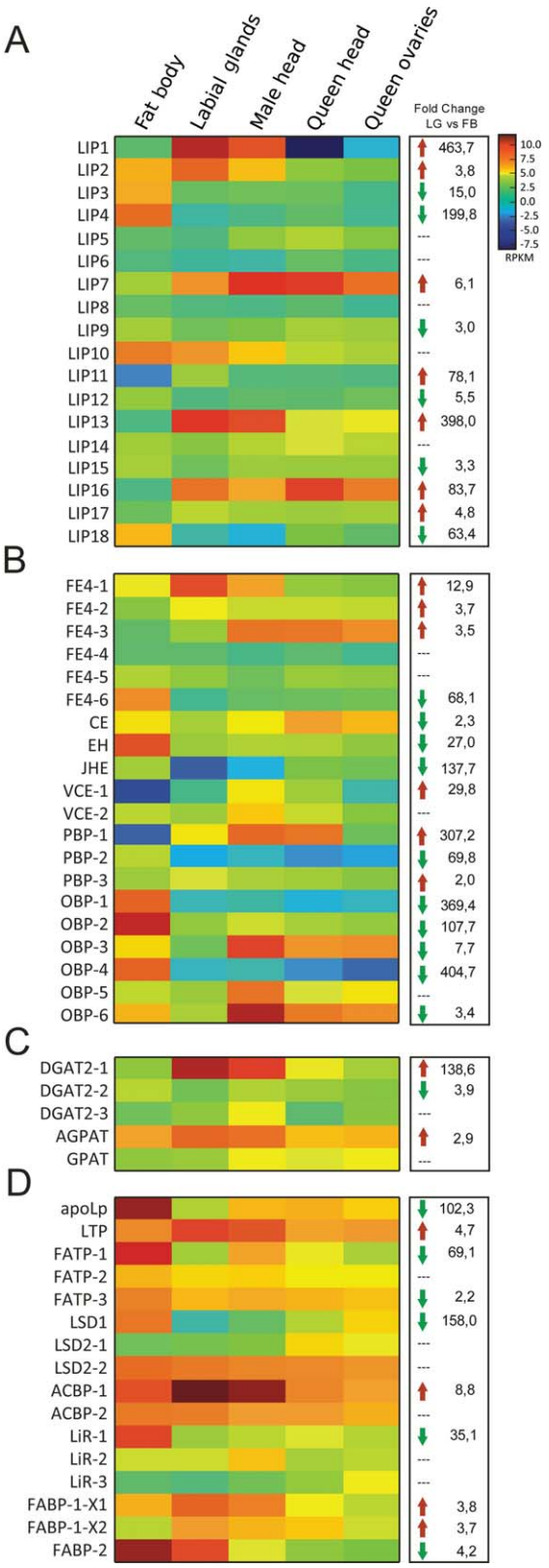

Figure 4. Relative expression of marking pheromone (MP)-biosynthetic candidate genes related to (A, B) lipolysis and fatty acid (FA) ester biosynthesis, (C) triacylglycerol biosynthesis and (D) lipid binding. The colour-coded expression values across tissues are based on $\log_2$-transformed reads per kilobase of exon model per million mapped reads (RPKM) values. FB, fat body; LG, labial gland; LIP1 - LIP18, lipase LIP; FE4-1 - FE4-6, esterase FE4; CE, carboxylesterase 5a-like; EH, ester hydrolase c11orf54 homolog; JHE, juvenile hormone esterase/venom carboxylesterase-3; VCE-1 - VCE-2, venom carboxylesterases; PBP-1 - PBP-3, pheromone binding proteins; OBP-1 - OBP-6, odorant binding proteins; DGAT2-1 - DGAT2-3, diacylglycerol acyltransferases 2; AGPAT, 1-acylglycerol-3-phosphate O-acyltransferase; GPAT, glycerol-3-phosphate O-acyltransferase; apoLP, apolipophorin; LTP, lipid transfer protein; FATP-1 - FATP-3, fatty acid (FA) transport-related proteins; LSD1 - LSD2-2, lipid storage droplet proteins; ACBP-1 - ACBP-2, acyl-coenzyme A-binding proteins; LiR-1 - LiR-3, lipophorin receptors; FABP-1-X1 - FABP-2, FA-binding proteins.

potentially involved in MP biosynthesis. For the majority of contigs, we identified the *Bombus terrestris* protein sequences by BLAST searches of the *Bombus terrestris* RefSeq database publicly available via NCBI GenBank (Sadd *et al.*, 2015). Although for some of the transcripts we had identified in our assembly we did not find the respective *Bombus terrestris* gene predictions, in all these cases we identified highly similar *Bombus impatiens* genomic sequence-derived predicted proteins, which we used for reconstruction of phylogenetic gene trees. We divided the candidate genes involved in MP biosynthesis into groups according to their enzymatic activity and proposed role in MP biosynthesis. To increase the reliability of differential gene expression prediction, we also used published RNA-seq data, including data from *Bombus terrestris* male head (MH), queen head (QH) and queen ovaries (QOs) (Sadd *et al.*, 2015) in our mappings.

Hydrolysis and esterification of FAs

Lipolysis

Lipases, members of a catalytically diverse α/β-hydrolase fold protein superfamily (Holmquist, 2000), are defined as lipolytic carboxyester hydrolases (Chahinian & Sarda, 2009; Ali *et al.*, 2012). The role of lipases in FA-derived bumblebee MP biosynthesis could be analogous to the role of lipases in the *Bombyx mori* PG, which catalyse lipolytic release of pheromone precursors from TAGs (Matsumoto *et al.*, 2002). Although we detected only traces of TAGs in the *Bombus terrestris* LG (data not shown), our previous experiments indicate that lipases presumably play a role in *Bombus terrestris* MP biosynthesis (Brabcová *et al.*, 2013b). The temporal increase in lipase activity of crude LG protein extract correlates well with MP biosynthesis onset in young *Bombus terrestris* males (Brabcová *et al.*, 2013b). Alternatively, lipases might play a role in biosynthesis of FA-ethyl ester MP components as discussed further below.

We identified 18 lipase genes (LIPs; Fig. 4A). The diversification of lipase genes after separation of insect orders (Horne *et al.*, 2009) and limited functional

information on insect lipases (Horne *et al*., 2009; Brabcová *et al*., 2013b) make it difficult to infer LIPs substrate specificity based solely on sequence. However, limited sequence–function relationships can be deduced from clustering LIPs into traditional lipase groups defined by protein sequence similarity, the presence of conserved domains and substrate preference (Horne *et al*., 2009). We reconstructed a lipase gene tree that includes LIPs and predicted *Ap. mellifera* lipases, which enabled us to extend the sequence-based predictions from *Ap. mellifera* lipases to their putative orthologues from *Bombus terrestris* (Fig. S3).

The lipase gene tree displays several well-supported clades that generally overlap with traditional lipase groups: a large group of neutral lipases (11 lipases), acidic lipases (three lipases), hormone-sensitive lipases (one lipase), class 3 lipases (two lipases) and pancreatic-like lipases (nine lipases) encompassed within the neutral lipases (Fig. S3).

The predicted lipases LIP1, LIP7, LIP11, LIP13, LIP16 and LIP17 were more abundantly expressed in the LG than in the FB (fold change > 4). LIP2 and LIP10 were abundantly expressed in both the LG and FB (Fig. 4A). Notably, LIP1 and LIP13 were more abundantly expressed in the LG and MH than in FB, QH, or QOs. Recently, we isolated LIP1 protein from the LG of *Bombus terrestris* males (Brabcová *et al*., 2013b). LIP1 exhibited (1) hydrolase activity with TAGs, fatty acyl-storage lipids, and (2) weak hydrolase activity with DAGs, a major form in which fatty acids are transported in insects, indicating its possible contribution to FA precursor release for pheromone biosynthesis. LIP13, a class 3 lipase presumably capable of hydrolysis of TAGs and a range of FA carboxy esters (Horne *et al*., 2009), represents a second candidate MP-biosynthetic lipase. The multitude of abundantly expressed lipase genes in the *Bombus terrestris* LG explains the lipase activity measured in this gland (Brabcová *et al*., 2013b) and supports the involvement of lipases in MP biosynthesis.

FA-ester production

The only functionally characterized insect enzymes capable of FA ethyl ester biosynthesis are two α/β-hydrolases involved in biosynthesis of ethyl octadecenoate, which serves as a primer pheromone in *Ap. mellifera* (Castillo *et al*., 2012). Although the equilibrium of hydrolase-catalysed reactions is generally shifted toward ester hydrolysis rather than ester formation, carboxyl esterases can exhibit substantial esterification activity in hydrophobic organic solvents (Yahya *et al*., 1998; Alvarez-Macarie & Baratti, 2000). Esterases, together with lipases, are members of the α/β-hydrolase fold protein superfamily capable of ester formation. In contrast to lipases, esterases are defined as carboxylesterases that prefer soluble carboxylic esters (ie simple esters consisting of short aliphatic chains) and are unable to hydrolyse water-insoluble lipids (Ali *et al*., 2012). The involvement of hydrolases in LG-localized ester formation is consistent with the formation of FA-ethyl esters in *Bombus lucorum* LGs incubated in the hydrophobic solvent hexane (Luxová *et al*., 2003). In the *Bombus terrestris* LG, the lipid droplet surface or interior might provide a natural environment favouring ester formation (Šobotník *et al*., 2008).

Although we did not find a *Bombus terrestris* orthologue of *Ap. mellifera* esterase/lipase EsLi (herein called AmelLIP19) capable of FA ethyl ester biosynthesis (Castillo *et al*., 2012; Fig. S3), the multitude of abundantly expressed lipases and esterases in the *Bombus terrestris* LG provides substantial genetic potential for ester biosynthetic activity (Fig. 4A, B). We identified 11 transcripts coding for enzymes with putative esterase activity, of which an esterase annotated as FE4−1 was abundantly and specifically expressed in the LG and MH (Fig. 4B). *Bombus terrestris* esterase VCE-2, a putative orthologue of *Ap. mellifera* α/β-3 hydrolase (XP_391943) capable of FA ethyl ester formation (Castillo *et al*., 2012), was not specifically expressed in the LG (Figs 4B, S4). A predicted *Bombus terrestris* gene product that exhibits high sequence similarity to hymenopteran venom carboxylesterases and juvenile hormone esterases (JHE) (Fig. 4B) exhibits homology to a previously functionally characterized *Ap. mellifera* JHE (Mackert *et al*., 2008; Fig. S4). This suggests that JHE might be involved in control of JH titre rather than FA ester metabolism in *Bombus terrestris*. The FB-specific expression of *Bombus terrestris* JHE (Fig. 4B) would be consistent with high JHE activity in insect FBs (Hammock, 1985).

Other proteins suggested to be involved in FA ethyl ester biosynthesis include odorant binding proteins (OBPs), or more generally, proteins with the capacity to bind aliphatic compounds. OBPs may be involved in biosynthesis of the honeybee forager prime pheromone, ethyl hexadecenoate, by binding it and preventing its hydrolysis (Castillo *et al*., 2012). Several OBPs and pheromone binding protein (PBP) homologues were also expressed in the LG transcriptome. Only PBP-1 was substantially more abundantly expressed in the LG than FB; however, it was also abundantly expressed in MH and QH (Fig. 4B), indicating a role in antennal pheromone detection rather than MP biosynthesis.

TAG biosynthesis

We identified three homologues of *Bombyx mori* DGAT2, which catalyses the rate-limiting step of TAG synthesis by covalently joining acyl-coenzyme A (acyl-CoA) and DAG (Du *et al*., 2012a,b). Notably, DGAT2-1 was

abundantly and specifically expressed in the LG and MH (Fig. 4C). Additionally, we identified a homologue of *Bombyx mori* GPAT involved in pheromone biosynthesis (Du *et al*., 2015), and 1-acylglycerol-3-phosphate O-acyltransferase (AGPAT), which was previously not considered to be involved in pheromone biosynthesis. AGPAT was more abundant in the LG and MH than in other tissues examined, indicating its possible role in MP biosynthesis (Fig. 4C).

Together, these findings suggest that despite only trace TAG content in the *Bombus terrestris* LG, TAGs might be involved in LG metabolism. Presumably, the turnover of TAGs in the LG is rapid, and therefore, TAGs are present only in small quantities at any given time point.

Lipid transport and lipid binding

Lipid transport is one of the presumed mechanisms that supplies fatty acyl precursors for biosynthesis of FA-derived MPs (Žáček *et al*., 2013). In insects, the transport form of lipids comprise DAG and apolipophorins apoLp-I, apoLp-II and apoLP-III, which together form low-density lipophorin (LDLp). LDLp can selectively unload lipid cargo in target tissues. Lipophorin receptors are presumed to further contribute to this process by receptor-mediated lipophorin endocytosis (Van der Horst *et al*., 2009).

Amongst the candidate genes involved in lipid transport, we identified (1) apoLP, which is abundantly and specifically expressed in the *Bombus terrestris* FB (Fig. 4D) and shares 46% protein sequence similarity with the *Anopheles gambiae* precursor protein apoLp-I/apoLp-II (XP_321226.5), and (2) three lipophorin receptors (LiRs) similar to the insect LiR that presumably mediates uptake of high-density lipophorin (HDLp) by FB cells and oocytes (Dantuma *et al*., 1999). However, none of the LiRs was specifically expressed in the LG (Fig. 4D), indicating that LiR-mediated uptake of lipids does not play a specific role in MP biosynthesis.

Additionally, we identified a range of transcripts coding for putative FA transport-related proteins, including FATPs (FATP-1–FATP-3) homologous to *Bombyx mori* FATPs that mediate uptake of extracellular FAs in the PG (Ohnishi *et al*., 2009) (Fig. 4D). FATP-1 was abundantly and specifically expressed in the FB, whereas FATP-2 and FATP-3 were expressed at comparable levels across the tissues, indicating their role in primary FA metabolism rather than LG-localized FA-derived MP biosynthesis (Fig. 4D). Notably, some FATPs exhibit acyl-CoA synthase activity in various organisms (DiRusso *et al*., 2008), suggesting a potential alternative role for *Bombus terrestris* FATPs in activation of FAs for downstream metabolic processes. We also identified lipid storage droplet proteins (LSDs; Fig. 4D) homologous to *Bombyx mori* LSDs that play a role in activation of lipid droplets for lipolysis (Ohnishi *et al*., 2011). *Bombus terrestris* LSD1 was highly expressed in the FB, whereas LSD2-2 was abundantly expressed at comparable levels across all the tissues. Lipid transfer protein was enriched in the LG and MH (Fig. 4D), and short predicted proteins ($<$ 140 amino acid residues) similar to the so-called FA binding proteins (Glatz & van der Vusse, 1996) were abundantly expressed in the FB and LG, indicating active FA transport in both tissues (Fig. 4D).

Additionally, we found two ACBPs that were highly abundantly expressed in the *Bombus terrestris* LG. In particular, ACBP-1 was specifically enriched in the *Bombus terrestris* LG and MH and was the second most abundant transcript sequenced in the LG (Fig. 4D, Data set S1). We presume that the role of ACBPs in *Bombus terrestris* MP biosynthesis is protection of acyl-CoA pheromone precursors against hydrolysis, analogous to the role of an ACBP characterized in the PG of *Bombyx mori* (Matsumoto *et al*., 2001; Ohnishi *et al*., 2006).

Regulation of MP biosynthesis

We searched the *Bombus terrestris* FB and LG transcriptomes for homologues of PBAN and PBANr, which are involved in regulation of FA-derived pheromone biosynthesis in moths (Tillman *et al*., 1999) and terpenic trail pheromone biosynthesis in the fire ant *S. invicta* (Choi & Vander Meer, 2012).

To identify a candidate *Bombus terrestris* PBANr, we searched for all *Bombus terrestris* homologues of previously predicted *Ap. mellifera* neuropeptide receptors (Jurenka & Nusawardani, 2011; Caers *et al*., 2012). Amongst the 26 putative *Bombus terrestris* neuropeptide receptors, NMUr2-2 (named after its homology to mammalian neuromedin U receptor 2) shared the highest protein sequence identity (74%) with *Ap. mellifera* pyrokinin 2/PBANr. Contrary to its presumed role in MP biosynthesis regulation, it was less abundant in the *Bombus terrestris* LG than FB and generally transcribed at low levels across all tissues (Figs 5A, S5).

We also identified the PBAN precursor, which shares high protein sequence identity with PBAN precursors from *Ap. mellifera* (79%) and *S. invicta* (68%). The *Bombus terrestris* PBAN gene presumably codes for four neuropeptides homologous to diuretic hormone, β-neuropeptide, PBAN and γ-neuropeptide (Fig. 6). Analogous to other hymenopteran PBAN precursors, the *Bombus terrestris* PBAN precursor lacks the α-neuropeptide coding region present in moth PBANs (Choi & Vander Meer, 2009). By sequence comparison, we predict that the *Bombus terrestris* PBAN is either a 10-amino acid peptide (IFPPLFAPRLamide) similar to a neuropeptide detected in *Ap. mellifera* (Hummon *et al*., 2006) or a 33-amino acid peptide (ESGEEYFSYGFPKDQEELYAEEQIFPPLFAPRLamide)

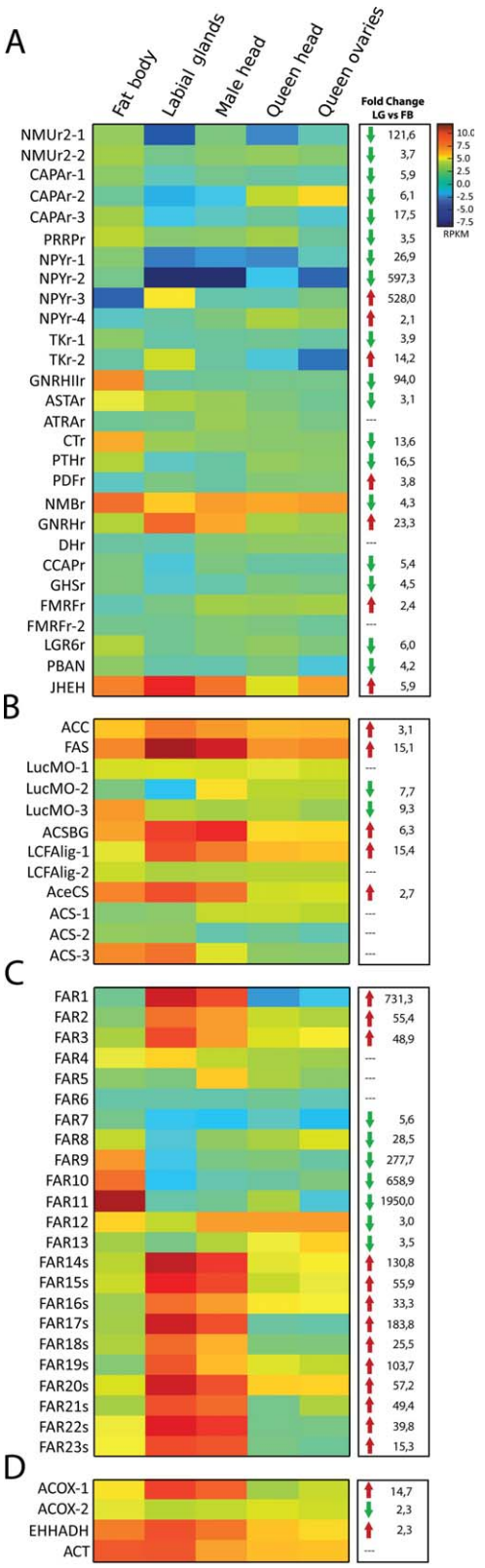

Figure 5. Relative expression of candidate (A) neuropeptide receptor genes and genes related to hormone regulation, (B) fatty acid (FA)-biosynthetic genes, (C) fatty acid reductases and (D) FA-chain shortening genes. The colour-coded expression values across tissues are based on $\log_2$-transformed reads per kilobase of exon model per million mapped reads (RPKM) values. FB, fat body; LG, labial gland; NMUr2-1 - NMUr2-2, neuromedin-U receptors 2-like; CAPAr-1 - CAPAr-3, neuropeptides capa receptors-like; PRRPr, prolactin-releasing peptide receptor-like; NPYr-1 - NPYr-4, neuropeptide Y receptors-like; TKr-1, tachykinin-like peptides receptor 99D-like; TKr-2, tachykinin-like peptides receptor 86C-like; GNRHIIr, gonadotropin-releasing hormone II receptor-like; ASTAr, allatostatin-A receptor-like; ATRAr, allatotropin receptor precursor; CTr, calcitonin receptor-like; PTHr, parathyroid hormone/parathyroid hormone-related peptide receptor-like; PDFr, pigment-dispersing factor receptor-like; NMBr, neuromedin-B receptor-like; gonadotropin-releasing hormone receptor-like; DHr, diuretic hormone receptor-like isoform 1; CCAPr, cardioacceleratory peptide receptor-like; GHSr, growth hormone secretagogue receptor type 1-like; FMRFr - FMRFr-2, FMRFamide receptors-like; LGR6r, leucine-rich repeat-containing GPC receptor 6-like; PBAN, pheromone biosynthesis activating neuropeptide; JHEH, juvenile hormone epoxide hydrolase; ACC, acetyl-CoA carboxylase; FAS, FA synthase; LucMO-1 - LucMO-3, luciferin monooxygenases; ACSBG, acyl-CoA synthetase "bubblegum"; LCFAlig-1 - LCFAlig-2, long-chain fatty acyl-CoA ligases; AceCS, acetyl-CoA synthase; ACS-1 - ACS-3, acyl-CoA synthases; FAR1 - FAR13, fatty acid reductases, FAR14s - FAR23s, short FAR-like proteins; ACOX-1 - ACOX-2, peroxisomal acyl-CoA oxidases; EHHADH, peroxisomal enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase; ACT, mitochondrial 3-ketoacyl-CoA thiolase.

homologous to PBAN from *S. invicta* (Choi & Vander Meer, 2009) resulting from the cleavage of a precursor PBAN protein at neuropeptide endoproteolytic cleavage sites (Southey *et al.*, 2008; Choi & Vander Meer, 2009). The two alternative *Bombus terrestris* PBANs share 85 and 70% identity with *Ap. mellifera* PBAN, respectively (Hummon *et al.*, 2006; Choi & Vander Meer, 2009).

Amongst the other neuropeptide receptors, we identified several putative receptor genes expressed more in the LG than in other examined tissues: neuropeptide Y receptor homologue (NPYr-3), tachykinin-like peptides receptor homologue (TKr-2) and gonadotropin-releasing hormone receptor (GNRHr) homologue (Fig. 5A).

FA biosynthesis

FA derivatives, such as ethyl dodecanoate, hexadecane-1-ol and octadeca-9,12,15-trien-1-ol, constitute over 15% of the *Bombus terrestris* MP (Kullenberg *et al.*, 1970; Žáček *et al.*, 2009), suggesting that genes involved in FA biosynthesis and derivatization are abundantly expressed in the *Bombus terrestris* LG. Previously, we confirmed high expression of fatty acid synthase (FAS) in the *Bombus terrestris* LG by qRT-PCR (Žáček *et al.*, 2013).

The GO-enrichment analysis of the LG transcriptome indicates significant enrichment ($P < 0.05$) of GO terms related to FA biosynthetic processes, providing further evidence for a highly active FA-biosynthetic enzymatic apparatus (Fig. 3). In addition to FAS, we identified other FA-biosynthesis-related genes, including acetyl-CoA synthetase (AceCS). AceCS activates acetate and supplies various metabolic processes with acetyl-CoA (Ikeda *et al.*, 2001), which is highly abundant in both the LG and FB

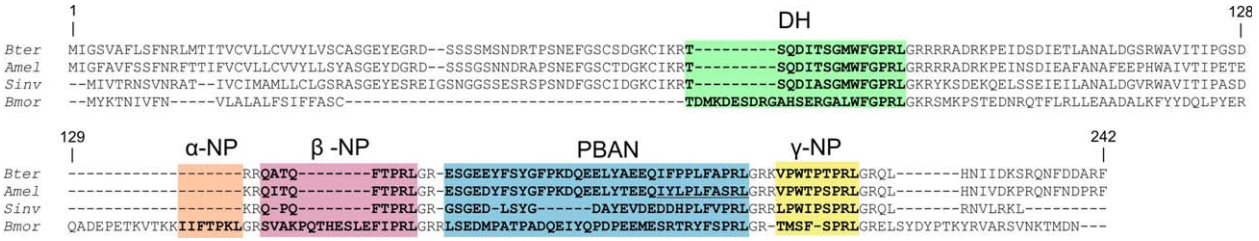

Figure 6. Sequence alignment of selected insect pheromone biosynthesis activating neuropeptide (PBAN) precursors. Predicted diuretic hormone (DH), α-neuropeptide (α-NP), β-neuropeptide (β-NP), PBAN and γ-neuropeptide (γ-NP) sequences are shaded in colour. The alternative shorter PBAN detected in *Apis mellifera* (Hummon *et al.*, 2006) is underlined. The sequence alignment was calculated with the MUSCLE algorithm and manually corrected. *Bter*, *Bombus terrestris* PBAN precursor (XP_003396136); *Amel*, *Apis mellifera* PBAN precursor (NP_001104182); *Sinv*, *Solenopsis invicta* PBAN precursor (NP_001291527); *Bmor*, *Bombyx mori* PBAN precursor (NP_001037321).

(Fig. 5B). *Bombus terrestris* AceCS shares high protein sequence identity with human AceCS1 (59%), which probably supplies FA biosynthesis in the liver (AAF75064, Luong *et al.*, 2000), and lower protein sequence identity (39%) with human mitochondrial AceCS2, which is involved in energy production (NP_001239605, Ikeda *et al.*, 2001), pointing toward Bombus terrestris AceCS playing a role in FA biosynthesis. We also identified an ACC homologue that catalyses the first committed step and one of the rate-limiting steps of FA biosynthesis: synthesis of malonyl-CoA from acetyl-CoA (Fig. 5B).

FA activation

Virtually all cellular FA-modifying processes require activation of FAs to the respective fatty acyl-CoA, a process that is catalysed by acyl-CoA synthetases (ACSs). In our search for *Bombus terrestris* ACSs, we benefited from the work on human ACSs by Watkins *et al.* (2007), who categorized 26 ACSs into subfamilies based on their FA-chain-length preference and sequence similarity. BLAST searches of *Bombus terrestris* LG and FB transcriptomes using representatives of human ACS subfamilies as queries yielded 10 putative ACS candidates with homology to human short-chain ACSs, including *Bombus terrestris* AceCS (Figs 5B, S6). Amongst the ACS candidates, we identified *Bombus terrestris* genes previously automatically annotated as luciferin monooxygenases (LucMOs). These candidates share sequence homology with the bifunctional luciferase from the firefly *Photinus pyralis* (AAA29795), which exhibits both LucMO and ACS activity (Oba *et al.*, 2003), and with a *Drosophila melanogaster* homologue (NP_651221) that possesses ACS activity exclusively (Oba *et al.*, 2004). The firefly luciferase and its *D. melanogaster* homologue prefer dodecanoic acid as a substrate (Oba *et al.*, 2005), making the *Bombus terrestris* LucMOs promising candidates for enzymes involved in biosynthesis of dodecanoic-acid-derived MP components, although the *Bombus terrestris* LucMOs are not overexpressed in the LG (Fig. 5B). Amongst the highly abundantly expressed and LG-specific *Bombus terrestris* ACS candidates, we found: (1) a gene product tentatively categorized into the 'bubblegum' enzyme subfamily (ACSBG) based on its homology to human ACSBGs (Steinberg *et al.*, 2000; Pei *et al.*, 2006) and (2) an ACS homologue of human long-chain ACSs [*Bombus terrestris* long-chain fatty acyl-CoA ligase (LCFAlig-1)] (Figs 5B, S6). In addition to these *Bombus terrestris* ACS candidates, *Bombus terrestris* FATPs (Figs 4D, S6) were found as the highest-scoring BLAST hits when *Bombus terrestris* transcriptomes were searched with representatives of human very long-chain ACSs (Watkins *et al.*, 2007), and they represent additional candidate genes involved in activation of FAs.

FA derivatization

FAs do not generally serve as long-range insect pheromones owing to their low volatility. To increase volatility, FAs undergo a variety of enzymatically catalysed modifications (Tillman *et al.*, 1999). The high number of possible pheromone structures produced by FA derivatization plays a pivotal role in generating the variability and species specificity of pheromone signals (Smadja & Butlin, 2009).

FARs and fatty acid desaturases

In our initial analysis of the *Bombus terrestris* LG transcriptome, we identified highly abundant and LG-specific expression of *Bombus terrestris* fatty acid desaturases that introduce double bonds at the Δ9- and Δ4-positions of fatty acyl chains. As *Bombus terrestris* does not utilize substantial amounts of unsaturated FA-derived MPs, we proposed that the activity of fatty acid desaturases in *Bombus terrestris* is controlled at the post-transcriptional level (Buček *et al.*, 2013). In this study, we identified another large multigene family encoding FA-modifying enzymes, FARs, which are presumably involved in reduction of FAs to fatty alcohol MP components. Amongst the putative *Bombus terrestris* FARs, FAR1, FAR2 and FAR3 were differentially and highly abundantly expressed in the LG and MH (Fig. 5C), indicating

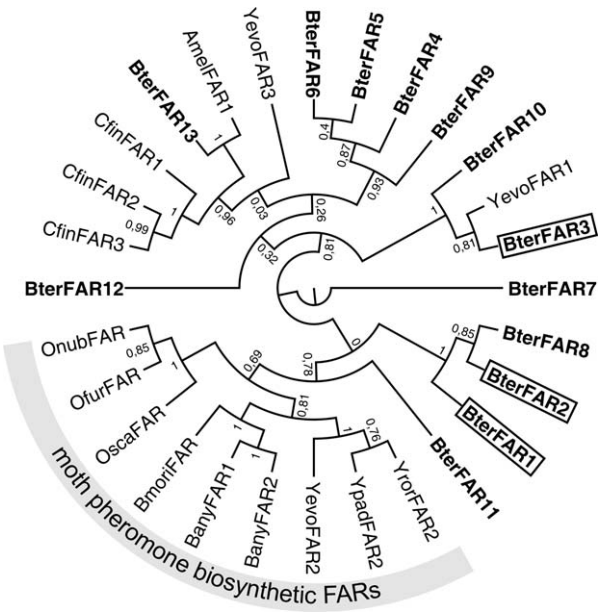

Figure 7. Arthropod fatty acid reductase (FAR) gene tree. Framed *Bombus terrestris* FARs were upregulated in the labial gland and male head. Numbers along branches indicate branch support calculated by an approximate likelihood ratio test (minimum of SH-like and Chi2-based values). Cfin, *Calanus finmarchicus*; Amel, *Apis mellifera*; Yevo, *Yponomeuta evonymellus*; Yror, *Yponomeuta rovellus*; Ypad, *Yponomeuta padellus*; Bany, *Bicyclus anynana*; Bmori, *Bombyx mori*; Osca, *Ostrinia scapulalis*; Ofur, *Ostrinia furnacalis*; Onub, *Ostrinia nubilalis*. See Data set S2 for GenBank sequence accession numbers.

their potential role in hexadecane-1-ol and octadeca-9,12,15-trien-1-ol biosynthesis. By contrast, FAR9, FAR10 and FAR11 were abundantly and specifically expressed in the FB, indicating a role in primary metabolism (Fig. 5C). We additionally identified numerous highly abundantly expressed, LG-specific contigs coding for short or partial FAR proteins (FAR11s–FAR20s, 47–234 amino acid residues; Fig. 5C). The short predicted proteins exhibited highest homology to FAR2 (Fig. S7). Notably, they also exhibited high sequence identity to short *Bombus terrestris* FAR-like proteins derived from genomic sequences available via GenBank. The role of these FAR-like open reading frames abundantly transcribed in the LG is unclear, but they might represent transcribed FAR pseudogenes.

To explore the homology of *Bombus terrestris* FARs to previously isolated insect and crustacean FARs, we reconstructed a FAR gene tree (Fig. 7). In the FAR tree, the moth pheromone biosynthetic FARs group together. Only YevoFAR1 and YevoFAR3 from the small ermine moth (*Yponomeuta evonymellus*) cluster outside the 'moth pheromone biosynthetic FAR' group, with YevoFAR1 being a putative orthologue of FAR3 (Fig. 5C). FAR13 was most closely related to AmelFAR1, which is abundantly expressed in the *Ap. mellifera* worker bee head and is presumed to be involved in pheromone or wax biosynthesis (Teerawanichpan *et al*., 2010; Figs 5C, 7). However, FAR13 was not overexpressed in LG or head tissues (MH, QH). Both AmelFAR1 and FAR13 group together with the CfinFARs from the marine crustacean *Calanus finmarchicus* (Fig. 7), which are proposed to be involved in biosynthesis of wax precursors (Teerawanichpan & Qiu, 2012). FAR1 and FAR2, which are highly specifically and abundantly expressed in the LG, do not exhibit orthology to any previously functionally characterized insect FAR.

FA chain-shortening

A plausible biosynthetic route to C12-derivatives involves peroxisomal β-oxidation, which is – in contrast to complete mitochondrial β-oxidation – capable of partial β-oxidation, resulting in chain-shortened FAs (Schulz, 1991). In moths, the specificity of FA chain-shortening enzymes in pheromone biosynthesis has been studied, but the enzymes involved have not yet been identified (Rosell *et al*., 1992; Jurenka *et al*., 1994).

We identified two putative peroxisomal acyl-CoA oxidases (ACOX-1 and ACOX-2; Fig. 5D) that catalyse the first step in the peroxisomal β-oxidation pathway (Schulz, 1991; Osumi, 1993). *Bombus terrestris* ACOX-2 exhibits high sequence identity with predicted insect ACOX3s, which are homologues of functionally characterized ACOX3s specific for branched acyl-CoA chains (eg human ACOX3, NP_003492; Baumgart *et al*., 1996). By contrast, the high protein sequence identity of *Bombus terrestris* ACOX-1 with other ACOX1s [45% with functionally characterized human ACOX1 (NP_004026), which prefers C12 to C18 substrates (Chu *et al*., 1995)] is consistent with the role of *Bombus terrestris* ACOX-1 in FA chain-shortening involved in the MP biosynthesis.

We also identified the multifunctional peroxisomal enzyme enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase (EHHADH), which catalyses two successive steps of peroxisomal β-oxidation: hydration of a β-double bond to 3-hydroxy acyl and dehydrogenation of 3-hydroxyacyl-CoA to 3-ketoacyl-CoA (Fig. 5D) (Osumi, 1993). Notably, we did not find in *Bombus terrestris* transcriptomes homologues of peroxisomal 3-ketoacyl-CoA thiolase (peroxisomal ACT), which catalyses the final step of peroxisomal β-oxidation – the release of a C2-shortened acyl-CoA (Hijikata *et al*., 1990). Furthermore, we found a homologue of mitochondrial 3-ketoacyl-CoA thiolase (mitochondrial ACT; Fig. 5D), which shares 57% protein sequence identity with human mitochondrial ACT (NP_006102; Abe *et al*., 1993) and only 34% with human peroxisomal ACT (NP_001598; Hijikata *et al*., 1990). Notably, we did not find the peroxisomal ACT in either the *Bombus terrestris* or *Ap. mellifera* NCBI RefSeq genome sequence, indicating that peroxisomal

ACT might be functionally replaced by an as-yet-unidentified enzyme.

Although chain-shortening was not observed in previous experiments employing isotopically labelled FAs of various chain lengths fed to *Bombus terrestris* males (Žáček *et al*., 2015), the results of the present study suggest that the identified genes represent rational targets for clarifying the origin of C12-derived *Bombus terrestris* MPs.

Isoprenoid biosynthesis

Terpenoid compounds comprise major proportion of the *Bombus terrestris* MP (Žáček *et al*., 2009), and isoprenoid biosynthesis is therefore expected to be highly active in the *Bombus terrestris* LG. Employing the RNA-seq data discussed in this paper, Prchalová *et al*. (2016) identified a complete set of isoprenoid biosynthetic candidate genes: acetoacetyl-CoA thiolase, 3-hydroxy-3-methylglutaryl coenzyme A synthase, HMGR, mevalonate kinase, phosphomevalonate kinase, diphosphomevalonate decarboxylase, isopentenyl diphospate isomerase, bifunctional farnesyl diphosphate synthase capable of sequential biosynthesis of geranyl diphosphate and farnesyl diphosphate, and geranylgeranyl diphosphate synthase. In addition, the researchers confirmed the high LG-expression of these enzymes by qRT-PCR (Prchalová *et al*., 2016).

After their backbone is built, isoprenoids undergo hydrolysis and various redox reactions, resulting in the final *Bombus terrestris* isoprenoid MP components. The major isoprenoid MP components have their first double bond reduced, which presumably originates from enzymatic reduction of isoprenoid precursors. The biochemical processes and enzymes involved in reduction of isoprenoid double bonds in insects remain unknown. Geranylgeranyl reductases (GGRs) capable of reducing geranylgeranyl double bonds have been characterized only in archaea (Nishimura & Eguchi, 2006; Murakami *et al*., 2007; Sato *et al*., 2008), plants (Keller *et al*., 1998; Tanaka *et al*., 1999), cyanobacteria (Addlesee *et al*., 1996) and bacteria (Bollivar *et al*., 1994). BLAST searches of FB and LG transcriptomes using these functionally characterized GGRs resulted only in low-quality BLAST hits (E-value $> 10^{-5}$) predominantly annotated as short-chain dehydrogenases/reductases (SDRs; Fig. S8). Despite their low sequence similarity with GGRs, these contigs represent possible candidates for genes involved in hydrogenation of the isoprenoid chain, as the SDR gene family encodes highly divergent enzymes with a broad substrate range, including steroids (Kallberg *et al*., 2002).

The isoprenoid compounds in the *Bombus terrestris* MP contain: (1) hydroxy groups presumably produced by hydrolysis of farnesyl diphosphate or geranylgeranyl diphosphate

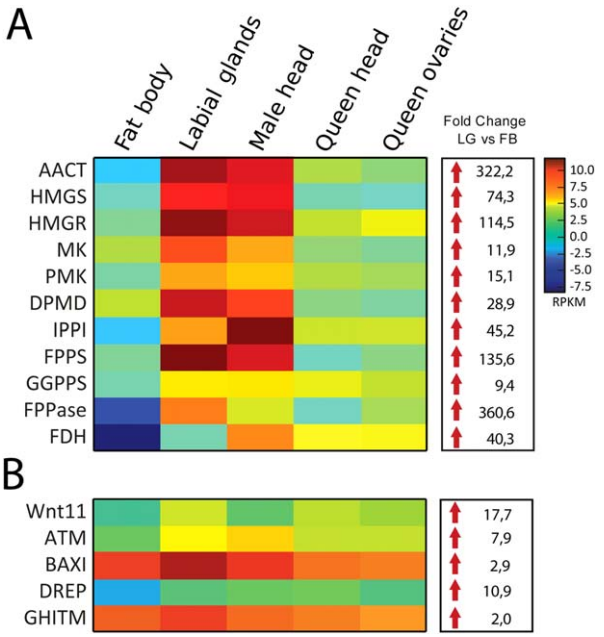

Figure 8. Relative expression of candidate (A) isoprenoid biosynthetic genes and (B) genes connected to apoptosis. The genes involved in biosynthesis of the isoprenoid backbone were identified in the *Bombus terrestris* labial gland (LG) and fat body (FB) by Prchalová *et al*. (2016). The colour-coded expression values across tissues are based on $\log_2$-transformed reads per kilobase of exon model per million mapped reads (RPKM) values. AACT, acetoacetyl-CoA thiolase; HMGS, 3-hydroxy-3-methylglutaryl coenzyme A synthase; HMGR, 3-hydroxy-3-methylglutaryl coenzyme A reductase; MK, mevalonate kinase; PMK, phosphomevalonate kinase ; DPMD, diphosphomevalonate decarboxylase; IPPI, isopentenyl diphospate isomerase; FPPS, bifunctional farnesyl diphosphate synthase; GGPPS, geranylgeranyl diphosphate synthase; FPPase, farnesyl diphosphate phosphatase; FDH, farnesol dehydrogenase; Wnt11, Wingless-Type MMTV Integration Site Family - Member 11; ATM, Ataxia Telangiectasia Mutated Protein; BAXI, BCL2-Associated X Protein inhibitor-1; DREP, DNA fragmentation factor (DFF) related protein; GHITM, Growth hormone inducible transmembrane protein.

to their respective alcohols and (2) aldehydic groups putatively produced by subsequent oxidation of these alcohol groups. The enzymes that catalyse these steps, farnesyl diphosphate phosphatase (FPPase; Cao *et al*., 2009; Nyati *et al*., 2013) and farnesol dehydrogenase (FDH; Mayoral *et al*., 2009), have previously been identified in insects. FDH and FPPase are presumably highly expressed in insect heads given their role in biosynthesis of JH, which is localized to the *corpora allata*, part of the insect brain (Tillman *et al*., 1999). We found *Bombus terrestris* FDH to be highly abundantly expressed in the MH and substantially less expressed in the LG and QH (Fig. 8A), indicating its rather marginal role in MP biosynthesis. The low abundances of FDH transcripts in the LG is consistent with only a small percentage (1%) of its presumed product, 2,3-dihydrofarnesal, being present in the MP (Žáček *et al*., 2009). By contrast, the high expression of FPPase in LG suggests its role in biosynthesis of terpene-derived MPS (Fig. 8A).

LG secretory cell apoptosis

The *Bombus terrestris* LG undergoes apoptosis of secretory cells after the fifth day of adult male life, a phenomenon that is unparalleled in any other pheromone-producing gland (Šobotník *et al.*, 2008). In general, apoptotic processes play a key role in organ development, but apoptosis also represents a cellular response to cell damage (Elmore, 2007). Because apoptosis starts at day 5 of the adult male lifespan (Šobotník *et al.*, 2008), apoptosis is presumably not vital for LG development. We hypothesize that, in the *Bombus terrestris* LG, terpenic MP components that accumulate to high, potentially toxic levels over time might trigger apoptosis. This hypothesis is supported by: (1) the previously described pro-apoptotic effect of farnesol in a range of organisms (Joo & Jetten, 2010); (2) the absence of secretory cell apoptosis in *Bombus lucorum* (Žáček *et al.*, 2009) and our preliminary data indicating the absence of apoptosis in *Bombus lapidarius*, both of which do not produce significant amounts of terpenic compounds; and (3) the temporal correlation between apoptosis onset and the peak of MP accumulation (Šobotník *et al.*, 2008; Žáček *et al.*, 2009).

To uncover candidate genes involved in LG apoptosis, we performed targeted gene exploration of the *Bombus terrestris* LG transcriptome and found 173 contigs annotated with cell death- or apoptosis-related GO terms (Data set S3). Three candidate genes were abundantly expressed in the LG (RPKM >1) and also substantially enriched in the LG compared with the FB (fold change- >4). However, they were not in all cases enriched in the LG compared with queen tissues (Fig. 8B). Homology with human proteins offers preliminary insight into the possible roles of *Bombus terrestris* proteins in LG apoptosis: (1) *Bombus terrestris* Wnt11 (Wingless-Type MMTV Integration Site Family, Member 11) is a homologue of a poorly described Wnt gene family member that encodes secreted signalling molecules involved in a variety of biological processes such as cell growth, cell differentiation, and apoptosis (Uysal-Onganer & Kypta, 2012); (2) *Bombus terrestris* Ataxia Telangiectasia Mutated Protein (ATM) is a very large protein kinase (2873 amino acids) homologous to human ATM, which is involved in response to DNA damage and a variety of other stress signalling networks and acts as a tumour suppressor via induction of cell-cycle arrest and apoptosis (Cremona & Behrens, 2014); and (3) *Bombus terrestris* DNA fragmentation factor (DFF) related protein (DREP) is homologous to *D. melanogaster* DREP-2, which is proposed to be involved in apoptosis based on its homology to human DFF (Inohara & Nuñez, 1999).

The most abundant cell-death-related gene products were Bax (BCL2-Associated X Protein) inhibitor-1 BAXI and a growth hormone inducible transmembrane protein GHITM. However, these proteins were expressed at comparable levels across all the investigated tissues. Both belong to the Bax inhibitory protein-like family, which has anti-apoptotic activity (Reimers *et al.*, 2007; Fig. 8B). None of the gene products annotated as caspases (endoproteases involved in regulation of cell death and inflammation; McIlwain *et al.*, 2013) was upregulated in the *Bombus terrestris* LG (Data set S3).

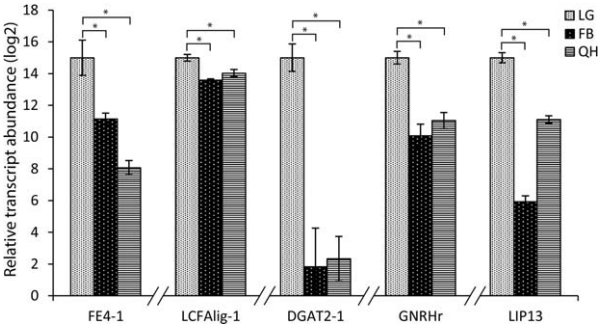

Figure 9. Quantitative real-time reverse transcription-PCR analysis of selected marking pheromone (MP)-biosynthetic gene candidates in the male labial gland (LG), male fat body (FB) and queen head (QH). The means $\pm$ SD of relative expression are shown in $\log_2$ scale. The average relative LG expression of each gene was arbitrarily set as 15. The expression of individual genes is not to scale. The significant differences between relative expression of genes in LG vs. FB and in LG vs. QH, are marked with an asterisk (*) and were calculated using a paired two-tailed *t*-test ($P<0.05$). FE4-1, esterase FE4-1; LCFAlig-1, long-chain fatty acyl-CoA ligase; DGAT2-1, diacylglycerol acyltransferase DGAT2-1; GNRHr, gonadotropin-releasing hormone receptor; LIP13, lipase LIP13.

The set of apoptosis-related gene products presented here represents the first insight into the molecular processes underlying the onset of LG secretory cell apoptosis. In the future, transcriptome analysis of LGs from *Bombus terrestris* specimens of different ages may highlight temporal trends in expression of apoptosis-related genes and thus provide further insight into the molecular basis of the LG secretory cell apoptosis.

qRT-PCR analysis of selected MP-biosynthetic gene candidates

The mRNA enrichment of several *Bombus terrestris* MP-biosynthetic gene candidates in the LG in comparison to the FB, ie fatty acid desaturases (Buček *et al.*, 2013), lipase (Brabcová *et al.*, 2013b), fatty acid synthase (Žáček *et al.*, 2013) and enzymes of terpene biosynthesis (Prchalová *et al.*, 2016), which was previously demonstrated by qRT-PCR, is in good agreement with the herein analysed *Bombus terrestris* LG and FB RNA-seq data sets. To further increase plausibility of the *Bombus terrestris* FB and LG transcriptomic data sets and to validate their comparison to publicly available *Bombus terrestris* transcriptomic data (Sadd *et al.*, 2015), we selected five genes representing various MP-biosynthetic pathways and protein classes that we found

were considerably enriched in the LG and measured their mRNA abundances in male FB, LG and QH using qRT-PCR. All of these candidate MP-biosynthetic genes, ie the putative esterase FE4-1, the acyl-CoA synthetase LCFAlig-1, the diacylglycerol acyltransferase DGAT2-1, the gonadotropin-releasing hormone receptor GNRHr and the lipase Lip13, were significantly more expressed in the LG (Fig. 9) than in the FB or QH, providing additional evidence for their involvement in male LG-specific processes and supporting the suitability of the transcriptomic approach for identification of *Bombus terrestris* LG-enriched genes.

Conclusion

The *Bombus terrestris* LG and FB transcriptomes enrich the transcriptomic resources available for this species (Sadd *et al*., 2010, 2015; Colgan *et al*., 2011; Woodard *et al*., 2011). Together with genomic resources (Munoz-Torres *et al*., 2011; Stolle *et al*., 2011; Amarasinghe *et al*., 2014; Sadd *et al*., 2015), these transcriptomic resources contribute to the establishment of *Bombus terrestris* as a model organism to provide insights into various aspects of primitively eusocial hymenopteran lives. The wealth of knowledge about genes involved in pheromone biosynthesis in moths and honeybees, along with the available knowledge about mammalian FA-metabolic enzymes, facilitated gene-targeted exploration of the *Bombus terrestris* LG transcriptome, which yielded MP biosynthetic gene candidates such as lipases, esterases, fatty acid reductases and enzymes involved in peroxisomal FA chain-shortening. Over 140 candidate genes identified in this work represent rational targets for future studies to disentangle the molecular basis of MP biosynthesis using methods such as (1) RNA interference knockdown of candidate genes; (2) detailed analysis of temporal gene expression changes across various tissues and developmental phases; and (3) cloning and functional characterization of genes encoding biosynthetic enzymes.

Experimental procedures

Tissue collection and RNA isolation

Bombus terrestris males were obtained from laboratory colonies established as previously described (Šobotník *et al*., 2008). LG and FB samples used for RNA-seq were prepared from 0-, 1-, 3- and 5-day-old bumblebee males by pooling tissues from two specimens of the same age for each tissue. The whole cephalic part of the LG and a section of the abdominal peripheral FB were dissected and pooled immediately after dissection in TRIzol (Invitrogen, Carlsbad, CA, USA) and flash-frozen by transferring the TRIzol-preserved samples to −80 °C, at which they were stored prior to RNA isolation. Total RNA was isolated using TRIzol according to the manufacturer's instructions. RNA quantity was assessed with a NanoDrop ND-100 UV/Vis spectrophotometer (Thermo Scientific, Waltham, MA, USA). RNA integrity was confirmed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). RNA for qRT-PCR was isolated from tissues from 3-day-old males (FB and LG) and 3-day-old queens (QH) using the above-described protocol.

qRT-PCR analysis

Total RNA (400 ng) served as templates for cDNA synthesis using SuperScript III Reverse Transcriptase (Invitrogen) and random hexamer primers according to the manufacturer's instructions. qRT-PCR was performed using a LightCycler 480 Real-Time PCR System (Roche, Indianapolis, IN, USA). PCR reactions were carried out using LightCycler 480 SYBR Green I Master (Roche), 0.625 mM of each gene-specific primer, and 2 μl of 6× diluted cDNA template. All measurements were performed in two technical and three biological replicates, and data were exported from the LightCycler 480 Software (version 1.5) into Microsoft Excel (Microsoft, Redmond, WA, USA) for statistical analysis. Relative gene expression was normalized to phospholipase A2 and elongation factor 1a (Horňáková *et al*., 2010). qRT-PCR primers were designed using Primer-BLAST as implemented on the NCBI website (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) such that for each gene one primer spanned the exon–exon junction. For detailed parameters of qRT-PCR analysis and qRT-PCR primer sequences see Supplementary Materials and Methods in Supporting Information.

RNA-sequencing, annotation, and GO-enrichment analysis

For each of the samples for RNA-seq, 5 μg of total RNA was used as starting material. Random primed cDNA library generation was performed using poly(A)+ enriched mRNA and standard Illumina TrueSeq protocols (Illumina, San Diego, CA, USA). The resulting cDNA was fragmented to an average of 150 bp. RNA-seq was carried out by Fasteris (Fasteris, Geneva, Switzerland) and was performed using the HiSeq 2500 Sequencing System from Illumina as previously described (Buček *et al*., 2013). Quality control, including filtering high-quality reads based on the fastq score and trimming the read lengths, was carried out using CLC Genomics Workbench software v. 7.0.1 (http://www.clcbio.com). After these filtering steps, the complete transcriptome library (70 000 000 reads) was assembled *de novo* using CLC Genomics Workbench software. Data assembly and annotation were performed as described previously (Buček *et al*., 2013).

Illumina reads mapped to the obtained reference backbone sequences were used to estimate expression levels and fold-change differences between the FB and LG. In addition to our RNA-seq data, we included previously published Illumina transcriptome sequence data from *Bombus terrestris* BioProject ID PRJNA45869, namely queen heads, queen ovaries and male heads available via NCBI Sequence Read Archive (SRA) under the IDs SRX090531, SRX090532, and SRX090533, respectively. Correction for biases in the sequence data sets and different transcript sizes was addressed by using the RPKM algorithm (Mortazavi *et al*., 2008). The RPKM values of all contigs with no reads mapped were set to a value smaller than the

value of the contig with the lowest RPKM value to enable fold-change calculations.

Complete short read (Illumina HiSeq2500) data were deposited in the European Nucleotide Archive with the following accession number: PRJEB9937 (SRA). The complete study can be directly accessed here: http://www.ebi.ac.uk/ena/data/view/PRJEB9937. The assembly of the *Bombus terrestris* data with contig consensus sequences, Blast2GO software v. 2.4.1 (Götz *et al*., 2008) hits against the NR database, hit accessions, putative annotations and relative expression levels across the five RNA-seq samples can be found in Data set S1.

Venn diagrams were calculated using software freely available at the University of Ghent web pages (http://bioinformatics.psb.ugent.be/webtools/Venn/).

GO term annotation of *Bombus terrestris* contigs, which was based on the sequence similarity of *Bombus terrestris* contigs to genes with available GO annotation, was performed in Blast2GO software v. 2.4.1. The annotations were subsequently utilized for GO-term enrichment analysis using a web-based version of Blast2GO software v. 2.7.2 (www.blast2go.de). The GO-enriched bar charts were reduced to display only the most specific GO-terms by removing parent terms of existing child terms using the function 'Reduce to most specific terms' implemented in Blast2GO. A GO term was considered significantly enriched if the *P*-value corrected by false discovery rate control was lower than 0.05.

Targeted gene discovery

A local BLAST database encompassing all *Bombus terrestris* contig sequences obtained in this study was set up in BioEdit software version 7.0.9.0 (Hall, 1999). Known insect protein sequences that had been previously demonstrated or proposed to be involved in pheromone biosynthesis were retrieved from NCBI GenBank and used as queries in BLAST searches of the local *Bombus terrestris* database. Additionally, the Blast2GO contig annotations were searched using keywords related to the presumed pheromone biosynthetic processes. Contigs that were annotated with GO terms related to presumed pheromone biosynthetic processes were also visually inspected. The resulting *Bombus terrestris* hits were used in subsequent rounds of local *Bombus terrestris* database BLAST searches to expand the initial search.

The predicted *Bombus terrestris* protein sequences were initially retrieved by BLAST searches of the *Bombus terrestris* RefSeq database publicly available via NCBI (accessed in 2014). The BLAST results were updated (May 2015) to reflect the most recent version of the *Bombus terrestris* genome annotation, which preceded publication of work on the *Bombus terrestris* and *Bombus impatiens* genomes (Sadd *et al*., 2015). The older versions of Bombus terrestris protein predictions (described in NCBI GenBank as obsolete as of May 2015) were used for reconstruction of gene phylogenies in cases for which (1) the obsolete prediction of *Bombus terrestris* protein was in greater agreement with its predicted or functionally characterized homologue from other insect species than the updated annotation (May 2015) or (2) the obsolete annotations of *Bombus terrestris* genes were in greater agreement with the gene models predicted from our RNA-seq data.

Phylogenetic reconstruction

Maximum-likelihood phylogenetic analysis was performed using the web-based pipeline Phylogeny.fr (Dereeper *et al*., 2008), consisting of multiple sequence alignment with the MUSCLE algorithm (Edgar, 2004) and reconstruction of a phylogenetic tree with the PhyML package (Guindon *et al*., 2009). The protein sequences used in phylogenetic reconstruction were downloaded from the GenBank database. For *Bombus terrestris* coding regions that did not have a genome-derived protein coding sequence available in GenBank, the translated protein-coding region of contig sequences obtained through the sequencing of the *Bombus terrestris* LG and FB libraries was used. Alternatively, *Bombus terrestris* contigs encompassing only a fragment of a predicted protein-coding region were replaced with *Bombus impatiens* protein-coding regions downloaded from GenBank if they exhibited clear signs of orthology with the *Bombus terrestris* contigs (over 90% protein sequence identity). The statistical support of branches was calculated by an approximate maximum likelihood-ratio test using Chi2-based parametric and SH-like branch supports. The phylogenetic trees were visualized with FigTree v. 1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/).

Acknowledgements

We are thankful to A. Votavová (Agricultural Research, Troubsko, Czech Republic) for *Bombus terrestris* breeding and colony maintenance, H. Hoffman for proof-reading of the manuscript, P. Žáček (Institute of Organic Chemistry and Biochemistry, AS CR, Prague, Czech Republic) for his help with interpretation of *Bombus terrestris* LG lipid analyses, two anonymous reviewers for comments on an earlier version of the manuscript and Wiley production team for editing the manuscript. This work was supported by research project RVO 61388963, NPU project no.: LO 1302 from the Ministry of Education, Czech Republic, and by grant no. 15-06569S from the Czech Science Foundation. The work was conducted within the framework of the COST Action FA1307 (Super-B).

References

Abe, H., Ohtake, A., Yamamoto, S., Satoh, Y., Takayanagi, M., Amaya, Y. *et al*. (1993) Cloning and sequence analysis of a full length cDNA encoding human mitochondrial 3-oxoacyl-CoA thiolase. *Biochim Biophys Acta* **1216**: 304–306.

Addlesee, H.A., Gibson, L.C.D., Jensen, P.E. and Hunter, C.N. (1996) Cloning, sequencing and functional assignment of the chlorophyll biosynthesis gene, chlP, of *Synechocystis* sp. PCC 6803. *FEBS Lett* **389**: 126–130.

Albert, S., Spaethe, J., Grübel, K. and Rössler, W. (2014) Royal jelly-like protein localization reveals differences in hypopharyngeal glands buildup and conserved expression pattern in brains of bumblebees and honeybees. *Biol Open* **3**: 281–288.

Ali, Y.B., Verger, R. and Abousalham, A. (2012) Lipases or esterases: does it really matter? Toward a new bio-physicochemical classification. *Methods Mol Biol* **861**: 31–51.

Alvarez-Macarie, E. and Baratti, J. (2000) Short chain flavour ester synthesis by a new esterase from *Bacillus licheniformis*. *J Mol Catal B Enzym* **10**: 377–383.

Amarasinghe, H.E., Clayton, C.I. and Mallon, E.B. (2014) Methylation and worker reproduction in the bumble-bee (*Bombus terrestris*). *Proc R Soc B Biol Sci* **281**: 20132502.

Arrese, E.L. and Soulages, J.L. (2010) Insect fat body: energy, metabolism, and regulation. *Annu Rev Entomol* **55**: 207–225.

Ayasse, M. and Jarau, S. (2014) Chemical ecology of bumble bees. *Annu Rev Entomol* **59**: 299–319.

Azzami, K., Ritter, W., Tautz, J. and Beier, H. (2012) Infection of honey bees with acute bee paralysis virus does not trigger humoral or cellular immune responses. *Arch Virol* **157**: 689–702.

Baumgart, E., Vanhooren, J.C.T., Fransen, M., Mannaerts, G.P. and Van Veldhoven, P.P. (1996) Mammalian peroxisomal acyl-CoA oxidases. III. Molecular characterization of human branched chain fatty acyl-CoA oxidase. *Ann N Y Acad Sci* **804**: 678–679.

Bellés, X., Martín, D. and Piulachs, M.D. (2005) The mevalonate pathway and the synthesis of juvenile hormone in insects. *Annu Rev Entomol* **50**: 181–199.

Bollivar, D.W., Suzuki, J.Y., Beatty, J.T., Dobrowolski, J.M. and Bauer, C.E. (1994) Directed mutational analysis of bacteriochlorophyll a biosynthesis in *Rhodobacter capsulatus*. *J Mol Biol* **237**: 622–640.

Brabcová, J., Kindl, J., Valterová, I., Pichová, I., Zarevúcka, M., Brabcová, J. *et al*. (2013a) Serine protease from midgut of *Bombus terrestris* males. *Arch Insect Biochem Physiol* **82**: 117–128.

Brabcová, J., Prchalová, D., Demianová, Z., Bučánková, A., Vogel, H., Valterová, I. *et al*. (2013b) Characterization of neutral lipase BT-1 isolated from the labial gland of *Bombus terrestris* males. *PLoS ONE* **8**: e80066.

Brabcová, J., Demianová, Z., Kindl, J., Pichová, I., Valterová, I. and Zarevúcka, M. (2015) Characterisation of Acetyl-CoA thiolase: the first enzyme in the biosynthesis of terpenic sex pheromone components in the labial gland of *Bombus terrestris*. *ChemBioChem* **16**: 1047–1051.

Buček, A., Vogel, H., Matoušková, P., Prchalová, D., Záček, P., Vrkoslav, V. *et al*. (2013) The role of desaturases in the biosynthesis of marking pheromones in bumblebee males. *Insect Biochem Mol Biol* **43**: 724–731.

Caers, J., Verlinden, H., Zels, S., Vandersmissen, H.P., Vuerinckx, K. and Schoofs, L. (2012) More than two decades of research on insect neuropeptide GPCRs: an overview. *Front Endocrinol* **3**: 151.

Cao, L., Zhang, P. and Grant, D.F. (2009) An insect farnesyl phosphatase homologous to the N-terminal domain of soluble epoxide hydrolase. *Biochem Biophys Res Commun* **380**: 188–192.

Castillo, C., Chen, H., Graves, C., Maisonnasse, A., Conte, Y.L. and Plettner, E. (2012) Biosynthesis of ethyl oleate, a primer pheromone, in the honey bee (*Apis mellifera* L.). *Insect Biochem Mol Biol* **42**: 404–416.

Chahinian, H. and Sarda, L. (2009) Distinction between esterases and lipases: comparative biochemical properties of sequence-related carboxylesterases. *Protein Pept Lett* **16**: 1149–1161.

Choi, M.Y. and Vander Meer, R.K. (2009) Identification of a new member of the PBAN family of neuropeptides from the fire ant, *Solenopsis invicta*. *Insect Mol Biol* **18**: 161–169.

Choi, M.Y. and Vander Meer, R.K. (2012) Ant trail pheromone biosynthesis is triggered by a neuropeptide hormone. *PLoS ONE* **7**: e50400.

Choi, M.Y., Fuerst, E.J., Rafaeli, A. and Jurenka, R. (2003) Identification of a G protein-coupled receptor for pheromone biosynthesis activating neuropeptide from pheromone glands of the moth *Helicoverpa zea*. *Proc Natl Acad Sci USA* **100**: 9721–9726.

Choi, M.Y., Raina, A. and Vander Meer, R.K. (2009) PBAN/pyrokinin peptides in the central nervous system of the fire ant, *Solenopsis invicta*. *Cell Tissue Res* **335**: 431–439.

Chu, R., Varanasi, U., Chu, S., Lin, Y., Usuda, N., Rao, M.S. *et al*. (1995) Overexpression and characterization of the human peroxisomal acyl-CoA oxidase in insect cells. *J Biol Chem* **270**: 4908–4915.

Colgan, T.J., Carolan, J.C., Bridgett, S.J., Sumner, S., Blaxter, M.L. and Brown, M.J. (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.

Coppée, A., Terzo, M., Valterova, I. and Rasmont, P. (2008) Intraspecific variation of the cephalic labial gland secretions in *Bombus terrestris* (L.) (Hymenoptera: Apidae). *Chem Biodivers* **5**: 2654–2661.

Cremona, C. and Behrens, A. (2014) ATM signalling and cancer. *Oncogene* **33**: 3351–3360.

Cvačka, J., Hovorka, O., Jiroš, P., Kindl, J., Stránský, K. and Valterová, I. (2006) Analysis of triacylglycerols in fat body of bumblebees by chromatographic methods. *J Chromatogr A* **1101**: 226–237.

Dantuma, N.P., Potters, M., De Winther, M.P., Tensen, C.P., Kooiman, F.P., Bogerd, J. *et al*. (1999) An insect homolog of the vertebrate very low density lipoprotein receptor mediates endocytosis of lipophorins. *J Lipid Res* **40**: 973–978.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F. *et al*. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* **36**: W465–W469.

DiRusso, C.C., Darwis, D., Obermeyer, T. and Black, P.N. (2008) Functional domains of the fatty acid transport proteins: studies using protein chimeras. *Biochim Biophys Acta* **1781**: 135–143.

Du, M., Yin, X., Zhang, S., Zhu, B., Song, Q. and An, S. (2012a) Identification of lipases involved in PBAN stimulated pheromone production in *Bombyx mori* using the DGE and RNAi approaches. *PLoS ONE* **7**: e31045.

Du, M., Zhang, S., Zhu, B., Yin, X. and An, S. (2012b) Identification of a diacylglycerol acyltransferase 2 gene involved in pheromone biosynthesis activating neuropeptide stimulated pheromone production in *Bombyx mori*. *J Insect Physiol* **58**: 699–703.

Du, M., Liu, X., Liu, X., Yin, X., Han, S., Song, Q. *et al*. (2015) Glycerol-3-phosphate O-acyltransferase is required for PBAN-induced sex pheromone biosynthesis in *Bombyx mori*. *Sci Rep* **5**: 8110.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792–1797.

Elmore, S. (2007) Apoptosis: a review of programmed cell death. *Toxicol Pathol* **35**: 495–516.

Fang, N., Teal, P.E.A. and Tumlinson, J.H. (1995) PBAN regulation of pheromone biosynthesis in female tobacco hornworm moths, *Manduca sexta* (L.). *Arch Insect Biochem Physiol* **29**: 35–44.

Fang, N., Teal, P.E.A. and Tumlinson, J.H. (1996) Effects of decapitation and PBAN injection on amounts of triacylglycerols in the sex pheromone gland of *Manduca sexta* (L.). *Arch Insect Biochem Physiol* **32**: 249–260.

Foster, S.P., Bergh, J.C., Rose, S. and Harris, M.O. (1991) Aspects of pheromone biosynthesis in the Hessian fly, *Mayetiola destructor* (Say). *J Insect Physiol* **37**: 899–906.

Gilg, A.B., Bearfield, J.C., Tittiger, C., Welch, W.H. and Blomquist, G.J. (2005) Isolation and functional expression of an animal geranyl diphosphate synthase and its role in bark beetle pheromone biosynthesis. *Proc Natl Acad Sci USA* **102**: 9760–9765.

Glatz, J.F. and van der Vusse, G.J. (1996) Cellular fatty acid-binding proteins: their function and physiological significance. *Prog Lipid Res* **35**: 243–282.

Goldstein, J.L. and Brown, M.S. (1990) Regulation of the mevalonate pathway. *Nature* **343**: 425–430.

Götz, S., García-Gómez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J. *et al.* (2008) High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res* **36**: 3420–3435.

Guindon, S., Delsuc, F., Dufayard, J.F. and Gascuel, O. (2009) Estimating maximum likelihood phylogenies with PhyML. *Methods Mol Biol* **537**: 113–137.

Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser.* **41**: 95–98.

Hammock, B.D. (1985) Regulation of juvenile hormone titre: degradation. In *Comprehensive Insect Physiology, Biochemistry & Pharmacology* (Kerkut, G.A. and Gilbert, L.I., eds), pp. 431–472. Pergamon Press, New York.

Hijikata, M., Wen, J.K., Osumi, T. and Hashimoto, T. (1990) Rat peroxisomal 3-ketoacyl-CoA thiolase gene: occurrence of two closely related but differentially regulated genes. *J Biol Chem* **265**: 4600–4606.

Hojo, M. (2007) Cloning and expression of a geranylgeranyl diphosphate synthase gene: insights into the synthesis of termite. *Insect Mol Biol* **16**: 121–131.

Hojo, M., Maekawa, K., Saitoh, S., Shigenobu, S., Miura, T., Hayashi, Y. *et al.* (2012) Exploration and characterization of genes involved in the synthesis of diterpene defence secretion in nasute termite soldiers. *Insect Mol Biol* **21**: 545–557.

Holmquist, M. (2000) Alpha/Beta-hydrolase fold enzymes: structures, functions and mechanisms. *Curr Protein Pept Sci* **1**: 209–235.

Horňáková, D., Matoušková, P., Kindl, J., Valterová, I. and Pichová, I. (2010) Selection of reference genes for real-time polymerase chain reaction analysis in tissues from *Bombus terrestris* and *Bombus lucorum* of different ages. *Anal Biochem* **397**: 118–120.

Horne, I., Haritos, V.S. and Oakeshott, J.G. (2009) Comparative and functional genomics of lipases in holometabolous insects. *Insect Biochem Mol Biol* **39**: 547–567.

Hummon, A.B., Richmond, T.A., Verleyen, P., Baggerman, G., Huybrechts, J., Ewing, M.A. *et al.* (2006) From the genome to the proteome: uncovering peptides in the *Apis* brain. *Science* **314**: 647–649.

Ikeda, Y., Yamamoto, J., Okamura, M., Fujino, T., Takahashi, S., Takeuchi, K. *et al.* (2001) Transcriptional regulation of the murine acetyl-CoA synthetase 1 gene through multiple clustered binding sites for sterol regulatory element-binding proteins and a single neighboring site for Sp1. *J Biol Chem* **276**: 34259–34269.

Inohara, N. and Nuñez, G. (1999) Genes with homology to DFF/CIDEs found in *Drosophila melanogaster*. *Cell Death Differ* **6**: 823–824.

Jarau, S., Žáček, P., Šobotník, J., Vrkoslav, V., Hadravová, R., Coppée, A. *et al.* (2012) Leg tendon glands in male bumblebees (*Bombus terrestris*): structure, secretion chemistry, and possible functions. *Naturwissenschaften* **99**: 1039–1049.

Jiroš, P., Cvačka, J., Hanus, R., Kindl, J., Kofroňová, E. and Valterová, I. (2011) Changes in the composition of triacylglycerols in the fat bodies of bumblebee males during their lifetime. *Lipids* **46**: 863–871.

Joo, J.H. and Jetten, A.M. (2010) Molecular mechanisms involved in farnesol-induced apoptosis. *Cancer Lett* **287**: 123–135.

Jurenka, R. and Nusawardani, T. (2011) The pyrokinin/pheromone biosynthesis-activating neuropeptide (PBAN) family of peptides and their receptors in Insecta: evolutionary trace indicates potential receptor ligand-binding domains. *Insect Mol Biol* **20**: 323–334.

Jurenka, R.A., Haynes, K.F., Adlof, R.O., Bengtsson, M. and Roelofs, W.L. (1994) Sex pheromone component ratio in the cabbage looper moth altered by a mutation affecting the fatty acid chain-shortening reactions in the pheromone biosynthetic pathway. *Insect Biochem Mol Biol* **24**: 373–381.

Kallberg, Y., Oppermann, U., Jörnvall, H. and Persson, B. (2002) Short-chain dehydrogenase/reductase (SDR) relationships: a large family with eight clusters common to human, animal, and plant genomes. *Protein Sci* **11**: 636–641.

Keller, Y., Bouvier, F., D'Harlingue, A. and Camara, B. (1998) Metabolic compartmentation of plastid prenyllipid biosynthesis – evidence for the involvement of a multifunctional geranylgeranyl reductase. *Eur J Biochem* **251**: 413–417.

Kofroňová, E., Nekola, A., Cvačka, J., Kindl, J. and Valterová, I. (2014) Fatty acids from pool lipids as possible precursors of the male marking pheromone in bumblebees. *Molecules* **19**: 2330–2343.

Kullenberg, B., Bergström, G. and Ställberg-Stenhagen, S. (1970) Volatile components of the cephalic marking secretion of male bumble bees. *Acta Chem Scand* **24**: 1481–1485.

Kullenberg, B., Bergstrom, G., Bringer, B., Carlberg, B. and Cederberg, B. (1973) Observations on scent marking by *Bombus* Latr. and *Psithyrus* Lep. males (Hym. Apidae) and localization of site of production of the secretion. *Zoon* **1**: 23–30.

Lewis, M.J., Prosser, I.M., Mohib, A. and Field, L.M. (2008) Cloning and characterisation of a prenyltransferase from the aphid *Myzus persicae* with potential involvement in alarm pheromone biosynthesis. *Insect Mol Biol* **17**: 437–443.

Luong, A., Hannah, V.C., Brown, M.S. and Goldstein, J.L. (2000) Molecular characterization of human acetyl-CoA synthetase, an enzyme regulated by sterol regulatory element-binding proteins. *J Biol Chem* **275**: 26458–26466.

Luxová, A., Valterová, I., Stránský, K., Hovorka, O. and Svatoš, A. (2003) Biosynthetic studies on marking pheromones of bumblebee males. *Chemoecology* **87**: 81–87.

Mackert, A., do Nascimento, A.M., Bitondi, M.M., Hartfelder, K. and Simões, Z.L (2008) Identification of a juvenile hormone esterase-like gene in the honey bee, *Apis mellifera* L. – expression analysis and functional assays. *Comp Biochem Physiol B Biochem Mol Biol* **150**: 33–44.

Matoušková, P., Luxová, A., Matoušková, J., Jiroš, P., Svatoš, A., Valterová, I. *et al*. (2008) A delta9 desaturase from *Bombus lucorum* males: investigation of the biosynthetic pathway of marking pheromones. *ChemBioChem* **9**: 2534–2541.

Matsumoto, S., Yoshiga, T., Yokoyama, N., Iwanaga, M., Koshiba, S., Kigawa, T. *et al*. (2001) Characterization of acyl-CoA-binding protein (ACBP) in the pheromone gland of the silkworm, *Bombyx mori*. *Insect Biochem Mol Biol* **31**: 603–609.

Matsumoto, S., Fónagy, A., Yamamoto, M., Wang, F., Yokoyama, N., Esumi, Y. *et al*. (2002) Chemical characterization of cytoplasmic lipid droplets in the pheromone-producing cells of the silkmoth, *Bombyx mori*. *Insect Biochem Mol Biol* **32**: 1447–1455.

Mayoral, J.G., Nouzova, M., Navare, A. and Noriega, F.G. (2009) NADP+-dependent farnesol dehydrogenase, a corpora allata enzyme involved in juvenile hormone synthesis. *Proc Natl Acad Sci USA* **106**: 21091–21096.

McIlwain, D.R., Berger, T. and Mak, T.W. (2013) Caspase functions in cell death and disease. *Cold Spring Harb Perspect Biol* **5**: 1–28.

de Miranda, J.R., Cordoni, G. and Budge, G. (2010) The Acute bee paralysis virus-Kashmir bee virus-Israeli acute paralysis virus complex. *J Invertebr Pathol* **103**: S30–S47.

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L. and Wold, B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* **5**: 621–628.

Moto, K., Yoshiga, T., Yamamoto, M., Takahashi, S., Okano, K., Ando, T. *et al*. (2003) Pheromone gland-specific fatty-acyl reductase of the silkmoth, *Bombyx mori*. *Proc Natl Acad Sci USA* **100**: 9156–9161.

Moto, K., Suzuki, M.G., Hull, J.J., Kurata, R., Takahashi, S., Yamamoto, M. *et al*. (2004) Involvement of a bifunctional fatty-acyl desaturase in the biosynthesis of the silkmoth, *Bombyx mori*, sex pheromone. *Proc Natl Acad Sci USA* **101**: 8631–8636.

Munoz-Torres, M.C., Reese, J.T., Childers, C.P., Bennett, A.K., Sundaram, J.P., Childs, K.L. *et al*. (2011) Hymenoptera Genome Database: integrated community resources for insect species of the order Hymenoptera. *Nucleic Acids Res* **39**: D658–D662.

Murakami, M., Shibuya, K., Nakayama, T., Nishino, T., Yoshimura, T. and Hemmi, H. (2007) Geranylgeranyl reductase involved in the biosynthesis of archaeal membrane lipids in the hyperthermophilic archaeon *Archaeoglobus fulgidus*. *FEBS J* **274**: 805–814.

Nishimura, Y. and Eguchi, T. (2006) Biosynthesis of archaeal membrane lipids: digeranylgeranylglycerophospholipid reductase of the thermoacidophilic archaeon *Thermoplasma acidophilum*. *J Biochem* **139**: 1073–1081.

Nyati, P., Nouzova, M., Rivera-Perez, C., Clifton, M.E., Mayoral, J.G. and Noriega, F.G. (2013) Farnesyl phosphatase, a Corpora allata enzyme involved in juvenile hormone biosynthesis in *Aedes aegypti*. *PLoS One* **8**: e71967.

Oba, Y., Ojika, M. and Inouye, S. (2003) Firefly luciferase is a bifunctional enzyme: ATP-dependent monooxygenase and a long chain fatty acyl-CoA synthetase. *FEBS Lett* **540**: 251–254.

Oba, Y., Ojika, M. and Inouye, S. (2004) Characterization of CG6178 gene product with high sequence similarity to firefly luciferase in *Drosophila melanogaster*. *Gene* **329**: 137–145.

Oba, Y., Sato, M., Ojika, M. and Inouye, S. (2005) Enzymatic and genetic characterization of firefly luciferase and *Drosophila* CG6178 as a fatty acyl-CoA synthetase. *Biosci Biotechnol Biochem* **69**: 819–828.

Ohnishi, A., Hull, J.J. and Matsumoto, S. (2006) Targeted disruption of genes in the *Bombyx mori* sex pheromone biosynthetic pathway. *Proc Natl Acad Sci USA* **103**: 4398–4403.

Ohnishi, A., Hashimoto, K., Imai, K. and Matsumoto, S. (2009) Functional characterization of the *Bombyx mori* fatty acid transport protein (BmFATP) within the silkmoth pheromone gland. *J Biol Chem* **284**: 5128–5136.

Ohnishi, A., Hull, J.J., Kaji, M., Hashimoto, K., Lee, J.M., Tsuneizumi, K. *et al*. (2011) Hormone signaling linked to silkmoth sex pheromone biosynthesis involves Ca2+/calmodulin-dependent protein kinase II-mediated phosphorylation of the insect PAT family protein *Bombyx mori* lipid storage droplet protein-1 (BmLsd1). *J Biol Chem* **286**: 24101–24112.

Olsen, S.S., Cazzamali, G., Williamson, M., Grimmelikhuijzen, C.J.P. and Hauser, F. (2007) Identification of one capa and two pyrokinin receptors from the malaria mosquito *Anopheles gambiae*. *Biochem Biophys Res Commun* **362**: 245–251.

Oppenheim, S.J., Baker, R.H., Simon, S. and DeSalle, R. (2015) We can't all be supermodels: the value of comparative transcriptomics to the study of non-model insects. *Insect Mol Biol* **24**: 139–154.

Osumi, T. (1993) Structure and expression of the genes encoding peroxisomal beta-oxidation enzymes. *Biochimie* **75**: 243–250.

Ozawa, R.A., Ando, T., Nagasawa, H., Kataoka, H. and Suzuki, A. (1993) Reduction of the acyl group: the critical step in bombykol biosynthesis that is regulated in vitro by the neuropeptide hormone in the pheromone gland of *Bombyx mori*. *Biosci Biotechnol Biochem* **57**: 2144–2147.

Park, Y., Kim, Y.J. and Adams, M.E. (2002) Identification of G protein-coupled receptors for *Drosophila* PRXamide peptides, CCAP, corazonin, and AKH supports a theory of ligand-receptor coevolution. *Proc Natl Acad Sci USA* **99**: 11423–11428.

Pei, Z., Jia, Z. and Watkins, P.A. (2006) The second member of the human and murine 'bubblegum' family is a testis- and brainstem-specific Acyl-CoA synthetase. *J Biol Chem* **281**: 6632–6641.

Poiani, S.B. and Da Cruz-Landim, C. (2010) Morphological changes in the cephalic salivary glands of females and males of *Apis mellifera* and *Scaptotrigona postica* (Hymenoptera, Apidae). *J Biosci* **35**: 249–255.

Prchalová, D., Buček, A., Brabcová, J., Žáček, P., Kindl, J., Valterová, I. *et al*. (2016) Regulation of isoprenoid pheromone biosynthesis in bumblebee males. *ChemBioChem* **17**: 260–267.

Reimers, K., Choi, C.Y., Bucan, V. and Vogt, P.M. (2007) The Growth-hormone inducible transmembrane protein (Ghitm) belongs to the Bax inhibitory protein-like family. *Int J Biol Sci* **3**: 471–476.

Rosell, G., Hospital, S., Camps, F. and Guerrero, A. (1992) Inhibition of a chain shortening step in the biosynthesis of the sex pheromone of the Egyptian armyworm *Spodoptera littoralis*. *Insect Biochem Mol Biol* **22**: 679–685.

Sadd, B.M., Kube, M., Klages, S., Reinhardt, R. and Schmid-Hempel, P. (2010) Analysis of a normalised expressed

sequence tag (EST) library from a key pollinator, the bumblebee *Bombus terrestris*. *BMC Genomics* **11**: 110.

Sadd, B.M., Barribeau, S.M., Bloch, G., de Graaf, D.C., Dearden, P., Elsik, C.G. *et al*. (2015) The genomes of two key bumblebee species with primitive eusocial organization. *Genome Biol* **16**: 76.

Sato, S., Murakami, M., Yoshimura, T. and Hemmi, H. (2008) Specific partial reduction of geranylgeranyl diphosphate by an enzyme from the thermoacidophilic archaeon *Sulfolobus acidocaldarius* yields a reactive prenyl donor, not a dead-end product. *J Bacteriol* **190**: 3923–3929.

Schulz, H. (1991) Beta oxidation of fatty acids. *Biochim Biophys Acta Lipids Lipid Metab* **1081**: 109–120.

Smadja, C. and Butlin, R.K. (2009) On the scent of speciation: the chemosensory system and its role in premating isolation. *Heredity (Edinb)* **102**: 77–97.

Southey, B.R., Sweedler, J.V. and Rodriguez-Zas, S.L. (2008) Prediction of neuropeptide cleavage sites in insects. *Bioinformatics* **24**: 815–825.

Steinberg, S.J., Morgenthaler, J., Heinzer, A.K., Smith, K.D. and Watkins, P.A. (2000) Very long-chain acyl-CoA synthetases: human 'bubblegum' represents a new family of proteins capable of activating very long-chain fatty acids. *J Biol Chem* **275**: 35162–35169.

Stolle, E., Wilfert, L., Schmid-Hempel, R., Schmid-Hempel, P., Kube, M., Reinhardt, R. *et al*. (2011) A second generation genetic map of the bumblebee *Bombus terrestris* (Linnaeus, 1758) reveals slow genome and chromosome evolution in the Apidae. *BMC Genomics* **12**: 48.

Šobotník, J., Kalinová, B., Cahlíková, L., Weyda, F., Ptáček, V. and Valterová, I. (2008) Age-dependent changes in structure and function of the male labial gland in *Bombus terrestris*. *J Insect Physiol* **54**: 204–214.

Tanaka, R., Oster, U., Kruse, E., Rudiger, W. and Grimm, B. (1999) Reduced activity of geranylgeranyl reductase leads to loss of chlorophyll and tocopherol and to partially geranylgeranylated chlorophyll in transgenic tobacco plants expressing antisense RNA for geranylgeranyl reductase. *Plant Physiol* **120**: 695–704.

Teerawanichpan, P. and Qiu, X. (2012) Molecular and functional analysis of three fatty acyl-CoA reductases with distinct substrate specificities in copepod *Calanus finmarchicus*. *Mar Biotechnol (NY)* **14**: 227–236.

Teerawanichpan, P., Robertson, A.J. and 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**: 641.

Terzo, M. and Coppens, P. (2007) Reduced cephalic labial glands in the male bumblebees of the subgenus *Rhodobombus* Dalla Torre (Hymenoptera: Apidae: *Bombus* Latreille). *Ann La Société Entomol Fr* **43**: 497–503.

Tillman, J., Seybold, S.J., Jurenka, R. and Blomquist, G.J. (1999) Insect pheromones–an overview of biosynthesis and endocrine regulation. *Insect Biochem Mol Biol* **29**: 481–514.

Tittiger, C., Blomquist, G.J., Ivarsson, P., Borgeson, C.E. and Seybold, S.J. (1999) Juvenile hormone regulation of HMG-R gene expression in the bark beetle *Ips paraconfusus* (Coleoptera: Scolytidae): implications for male aggregation pheromone biosynthesis. *Cell Mol Life Sci* **55**: 121–127.

Tsfadia, O., Azrielli, A., Falach, L., Zada, A., Roelofs, W. and Rafaeli, A. (2008) Pheromone biosynthetic pathways: PBAN-regulated rate-limiting steps and differential expression of desaturase genes in moth species. *Insect Biochem Mol Biol* **38**: 552–567.

Uysal-Onganer, P. and Kypta, R.M. (2012) Wnt11 in 2011 - the regulation and function of a non-canonical Wnt. *Acta Physiol* **204**: 52–64.

Van der Horst, D.J., Roosendaal, S.D. and Rodenburg, K.W. (2009) Circulatory lipid transport: lipoprotein assembly and function from an evolutionary perspective. *Mol Cell Biochem* **326**: 105–119.

Velthuis, H.H.W. and Van Doorn, A. (2006) A century of advances in bumblebee domestication and the economic and environmental aspects of its commercialization for pollination. *Apidologie* **37**: 421–451.

Watkins, P.A., Maiguel, D., Jia, Z. and Pevsner, J. (2007) Evidence for 26 distinct acyl-coenzyme A synthetase genes in the human genome. *J Lipid Res* **48**: 2736–2750.

Woodard, S.H., Fischman, B.J., Venkat, A., Hudson, M.E., Varala, K., Cameron, S.A. *et al*. (2011) Genes involved in convergent evolution of eusociality in bees. *Proc Natl Acad Sci USA* **108**: 7472–7477.

Yahya, A.R.M., Anderson, W. and Moo-Young, M. (1998) Ester synthesis in lipase-catalyzed reactions. *Enzyme Microb Technol* **23**: 438–450.

Žáček, P., Prchalová-Horňáková, D., Tykva, R., Kindl, J., Vogel, H., Svatoš, A. *et al*. (2013) De novo biosynthesis of sexual pheromone in the labial gland of bumblebee males. *Chembiochem* **14**: 361–371.

Žáček, P., Kalinová, B., Šobotník, J., Hovorka, O., Ptáček, V., Coppée, A. *et al*. (2009) Comparison of age-dependent quantitative changes in the male labial gland secretion of *Bombus terrestris* and *Bombus lucorum*. *J Chem Ecol* **35**: 698–705.

Žáček, P., Kindl, J., Frišonsová, K., Průchová, M., Votavová, A., Hovorka, O. *et al*. (2015) Biosynthetic studies of the male marking pheromone in bumblebees by using labelled fatty acids and two-dimensional gas chromatography with mass detection. *Chempluschem* **80**: 839–850.

Zhang, S.D., Li, X., Bin, Z., Du, M.F., Yin, X.M. and An, S.H. (2014) Molecular identification of a pancreatic lipase-like gene involved in sex pheromone biosynthesis of *Bombyx mori*. *Insect Sci* **21**: 459–468.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Data set S1. The assembly of the *Bombus terrestris* data with contig consensus sequences, Blast2GO hits against the nonredundant database, hit accessions, putative annotations, and relative expression levels across labial gland, fat body, male head, queen head and queen ovary samples.

Data set S2. GenBank sequence accession numbers of genes and proteins used here for the reconstruction of gene trees.

Data set S3. List of *Bombus terrestris* contigs annotated with cell death or apoptosis related gene ontology (GO) terms.

Figure S1. Analysis of enriched gene ontology (GO) terms in the *Bombus terrestris* labial gland (LG). The nonreduced bar chart shows GO terms at various levels of specificity that were significantly ($P < 0.05$) enriched in the LG. The GO terms are sorted in descending order according to their *P*-values.

Figure S2. Analysis of enriched gene ontology (GO) terms in the *Bombus terrestris* fat body (FB). The nonreduced bar chart shows GO

terms at various levels of specificity that were significantly ($P < 0.05$) enriched in the FB. The GO terms are sorted in descending order according to their *P*-values.

Figure S3. Lipase gene tree. *Bombus terrestris* lipases are highlighted in red and *Apis mellifera* lipases in blue. The functionally characterized *Ap. mellifera* lipase is framed in blue; *Bombus terrestris* lipases specifically expressed in the labial gland and male head are framed in red. Numbers along branches indicate branch support calculated by an approximate likelihood ratio test (minimum of SH-like and Chi2-based values). See Data set S2 for GenBank sequence accession numbers.

Figure S4. Esterase gene tree. Predicted *Bombus terrestris* esterases are highlighted in red and predicted *Apis mellifera* esterases in blue. The functionally characterized *Ap. mellifera* esterases are framed in blue; the FE4-1 esterase specifically expressed in the *Bombus terrestris* labial gland and male head is framed in red. Numbers along branches indicate branch support calculated by an approximate likelihood ratio test (minimum of SH-like and Chi2-based values). Amel_JHE, *Ap. mellifera* juvenile hormone esterase. See Data set S2 for GenBank sequence accession numbers and other abbreviations.

Figure S5. Neuropeptide receptor gene tree. *Bombus terrestris* neuropeptide receptors (or *Bombus impatiens* orthologues) are highlighted in red and *Apis mellifera* receptors in blue. Framed *Bombus terrestris* neuropeptide receptors were upregulated in the labial gland compared with the fat body, queen head and queen ovaries. Numbers along branches indicate branch support calculated by an approximate likelihood ratio test (minimum of SH-like and Chi2-based values). See Data set S2 for GenBank sequence accession numbers and abbreviations.

Figure S6. Gene tree of enzymes with predicted acyl-coenzyme A synthetase (ACS) activity. Framed genes are most abundant in the *Bombus terrestris* labial gland and male head. *Bombus terrestris* genes (or *Bombus impatiens* orthologues) are displayed in blue and *Apis mellifera* genes in red. Numbers along branches indicate branch support calculated by an approximate likelihood ratio test (minimum of SH-like and Chi2-based values). Homo, *Homo sapiens* genes; FATP, fatty acid transport protein; ACSVL, very long-chain ACS; ACSL, long-chain ACS; ACSM, medium chain ACS; ACSS, short-chain ACS; ACSF, ACS-Family member; LucMO, luciferine monooxygenase; LCFAlig, long chain fatty acid ligase; AceCS, acetyl-CoA synthase; ACSBG, 'bubblegum' type ACS. See Data set S2 for GenBank sequence accession numbers.

Figure S7. Fatty acid reductase (FAR) gene tree including the short predicted *Bombus terrestris* FAR-like proteins. The short FAR-like proteins were deduced from the FAR-like contigs (BterFAR14s-BterFAR23s, see Fig. 5C). Numbers along branches indicate branch support calculated by an approximate likelihood ratio test (minimum of SH-like and Chi2-based values). Cfin, *Calanus finmarchicus*; Amel, *Apis mellifera*; Yevo, *Yponomeuta evonymellus*; Yror, *Yponomeuta rovellus*; Ypad, *Yponomeuta padellus*; Bany, *Bicyclus anynana*; Bmori, *Bombyx mori*; Osca, *Ostrinia scapulalis*; Ofur, *Ostrinia furnacalis*; Onub, *Ostrinia nubilalis*. See Data set S2 for GenBank sequence accession numbers.

Figure S8. Relative expression of contigs retrieved as highest-scoring BLAST hits using functionally characterized geranylgeranyl reductases as BLAST queries. The colour-coded expression values across tissues are based on $\log_2$-transformed reads per kilobase of exon model per million mapped reads (RPKM) values. See Data set S1 for contig sequences and annotation.

Table S1. Number and percentage of reads mapped to the acute bee paralysis virus (ABPV) contigs. The number of reads mapping to ABPV contigs was deduced from the *Bombus terrestris* male fat body and labial gland RNA sequencing (RNA-seq) data (this study) and male head, queen head and queen ovary RNA-seq data (Sadd *et al.*, 2015).