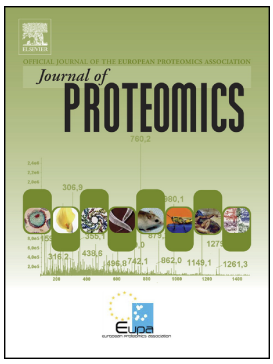


Combination of label-free quantitative proteomics and transcriptomics reveals intraspecific venom variation between the two strains of Tetrastichus brontispae, a parasitoid of two invasive beetles

Bao-Zhen Tang, E. Meng, Hua-Jian Zhang, Xiao-Mei Zhang, Sassan Asgari, Ya-Ping Lin, Yun-Ying Lin, Zheng-Qiang Peng, Ting Qiao, Xia-Fang Zhang, You-Ming Hou

PII: S1874-3919(18)30307-5
DOI: doi:10.1016/j.jprot.2018.08.003
Reference: JPROT 3189

To appear in: *Journal of Proteomics*

Received date: 26 February 2018
Revised date: 25 July 2018
Accepted date: 3 August 2018

Please cite this article as: Bao-Zhen Tang, E. Meng, Hua-Jian Zhang, Xiao-Mei Zhang, Sassan Asgari, Ya-Ping Lin, Yun-Ying Lin, Zheng-Qiang Peng, Ting Qiao, Xia-Fang Zhang, You-Ming Hou , Combination of label-free quantitative proteomics and transcriptomics reveals intraspecific venom variation between the two strains of Tetrastichus brontispae, a parasitoid of two invasive beetles. Jprot (2018), doi:10.1016/j.jprot.2018.08.003

Combination of Label-free Quantitative Proteomics and Transcriptomics Reveals Intraspecific Venom Variation between the Two Strains of *Tetrastichus brontispae*, a Parasitoid of Two Invasive Beetles

Bao-Zhen Tang[a,b,1], E Meng[a,b,1], Hua-Jian Zhang[a,b], Xiao-Mei Zhang [a, b], Sassan Asgari [c], Ya-Ping Lin [a, b], Yun-Ying Lin [a, b], Zheng-Qiang Peng [d], Ting Qiao [a, b], Xia-Fang Zhang [a, b], and You-Ming Hou[a,b,*]

[a]State Key Laboratory of Ecological Pest Control for Fujian and Taiwan Crops, Fujian Agriculture and Forestry University, Fuzhou 350002, China

[b]Fujian Provincial Key Laboratory of Insect Ecology, Department of Plant Protection, Fujian Agriculture and Forestry University, Fuzhou, Fujian, China

[c]School of Biological Sciences, the University of Queensland, Brisbane, QLD 4067, Australia

[d]Environment and Plant Protection Institute, Chinese Academy of Tropical Agricultural Sciences, Haikou 571101, China

[1] These authors contributed equally to this work.

* Corresponding author at: Fujian Provincial Key Laboratory of Insect Ecology, Department of Plant Protection, Fujian Agriculture and Forestry University, No.15 Shangxiadian Road, Cangshan District, Fuzhou City, Fujian Province, China.

E-mail address: ymhou@fafu.edu.cn (You-Ming Hou).

This article contains supplemental Figs. S1 to S5 and Tables S1 to S11.

Abstract:

The venom apparatus is a conserved organ in parasitoids that shows adaptations correlated with life-style diversification. Combining transcriptomics and label-free quantitative proteomics, here we explored the venom apparatus components of the endoparasitoid *Tetrastichus brontispae* (Eulophidae), and provide a comparison of the venom apparatus proteomes between its two closely related strains, *T. brontispae-Octodonta nipae* (Tb-On) and *T. brontispae-Brontispa longissima* (Tb-Bl). Tb-Bl targets the *B. longissima* pupa as its habitual host. However, Tb-On is an experimental derivative of Tb-Bl, which has been exposed to the *O. nipae* pupa as host consecutively for over 40 generation. Results showed that approximately 1505 venom proteins were identified in the *T. brontispae* venom apparatus. The extracts contained novel venom proteins, such as 4-coumarate-CoA ligase 4. A comparative venom proteome analysis revealed that significant quantitative and qualitative differences in venom composition exist between the two strains; although the most abundant venom proteins were shared between them. The differentially produced proteins were mainly enriched in fatty acid biosynthesis and melanotic encapsulation response. Six of these enriched proteins presented increased levels in Tb-On, and this result was validated by parallel reaction monitoring (PRM) analysis. Overall, our data reveal that venom composition can evolve quickly and respond to host selection.

Keywords: *Tetrastichus brontispae*; *Octodonta nipae*; *Brontispa longissima*; parallel reaction monitoring (PRM); label-free; parasitoid venom.

Abbreviations

Tb-On, *Tetrastichus brontispae-Octodonta niap*e; Tb-Bl, *Tetrastichus brontispae-Brontispa longissima*; LFQ, label-free quantitation; TEM, transmission electron microscopy; DIC, differential interference contrast; NCBI, National Center for Biotechnology Information; NR, NCBI non-redundant; KEGG; Kyoto Encyclopedia of Genes and Genome; KAAS, KEGG Automatic Annotation Server; COG, Cluster of Orthologous Groups; FASP, filter-aid sample preparation; SDS, sodium dodecyl sulfate; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; DTT, dithiothreitol; UA buffer, 8 M Urea, 150mM Tris-HCl pH 8.0; IAA, iodoacetamide; ABC, ammonium bicarbonate; HCD, higher energy collisional dissociation; FDR, false discovery rate; iBAQ, intensity-based absolute quantification; GO, gene ontology; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; AMPK, AMP-activated

protein kinase; PPAR, peroxisome proliferator-activated receptors; NEP, neprilysin; KA, kynurenic acid; KAT, kynurenine aminotransferase; Try, tryptophan; PDC, phenylpropanoid-derived compound; LGT, lateral gene transfer; GH19, glucoside hydrolase family 19; CRT, calreticulin; RCL, reactive center loop; Serpin, serine protease inhibitor; SPH, serine protease homologs; PAP, proPO-activating protease; PRM, parallel reaction monitoring.

1. Introduction

Parasitoid wasps are insects that achieve their development at the expense of other arthropods, ultimately leading to host death [1]. They are thus invaluable in suppressing host populations and are extensively used for biological control of various insect pests. To ensure successful parasitism, endoparasitoids, which develop inside their hosts, must evade or counteract the specific physiological and immunological defenses of their hosts [2]. During the antagonistic evolution between endoparasitoids and hosts, endoparasitoids have evolved original strategies to alter host immunity and/or manipulate its development and physiology ranging from displaying egg surface features preventing recognition by the host [3, 4], to injection of various virulence factors such as venoms [5], virus-like particles [6] and polydnaviruses [7, 8] together with the egg. In many cases, such as endoparasitoids devoid of polydnaviruses, venoms play a key role in virulence towards the host [5]. Typically, parasitoid venoms consist of a complex cocktail of proteinaceous (proteins and peptides, commonly referred to as toxins) and non-proteinaceous (salt and organic components) compounds. The proteinaceous components are generally the most abundant [2, 9].

Through advances in "omic" technologies, parasitoid venom composition data from selected species have recently become available, revealing considerable complexity and diversity in these arsenals, along with a high degree of functional redundancy [9-11]. Moreover, it is intriguing that venom components can differ between closely related species that parasitize the same host, and even between strains and individuals [10, 12], yet the most abundant venom proteins in a given species can be quite similar to those in a phylogenetically distant species [13]. Despite the richness and diversity of parasitoid venom composition, during the past 50 years, venoms have so far been analyzed in only 18 species belonging to the Ichneumonidae, Braconidae, Pteromalidae, Figitidae, Eucoilidae or Bethylidae using different approaches [2, 11, 14], with no available data from the remaining more than 20,000 species [15]. It is worth noting that among these selected species studied to date, venomic studies have mainly focused on the identification of primary venom components while overlooking secondary venom factors that could possibly potentiate or regulate their virulence in parasitism. Moreover, sparse information is available yet regarding the accurate quantity of venom composition. The only quantitative analyses available are those from *Leptopilina* and *Aphidius ervi* which relied on the numbers of ESTs (expressed sequence tags) at the transcriptional level rather than the exact abundance at the protein level [12, 13]. Hence, it is obvious that analyses of venom components are far from being complete. Knowledge of broader genomic and transcriptomic data together with an accurate proteomic analysis from a great

diversity of parasitoids will provide a foundation for unravelling the mechanisms of evolution under strong selection pressures.

Tetrastichus brontispae Ferriere (Hymenoptera: Eulophidae) is a gregarious and koinobiont endoparasitoid native to Java, Indonesia, and has been introduced into many countries to efficiently control the coconut hispine beetle *Brontispa longissima* (Gestro) (Coleoptera: Chrysomelidae), a currently serious invasive pest of palm plants in southern China [16]. The potential of *T. brontispae* to control another invasive beetle of palms in southern China, *Octodonta nipae* (Maulik) (Coleoptera: Chrysomelidae), has also been validated [17]. *T. brontispae* mostly prefers one-day-old host pupae and has high fecundity (*per capita* 22 wasps emerging from each host pupa) [16, 17]. The offspring are usually female-biased [16, 17]. In addition, it takes only approximately 20 days for *T. brontispae* to fulfil its lifecycle [16, 17]. These biological characteristics enhance the success of *T. brontispae* as a biological control agent for the management of these beetles. *T. brontispae* is devoid of polydnavirus (unpublished data), and thus, its venom is the main source of factors regulating the host. However, to date, its venom composition at the molecular level is not available, and only sparse information is accessible regarding how *T. brontispae* interacts with host physiology and how its venom is involved in regulating host genes post-parasitization [18-20].

Next-generation sequencing technologies combined with bioinformatics tools can generate extensive data to facilitate the construction of a reference protein database for proteomics analyses, which is invaluable, particularly in the absence of a sequenced genome. The large-scale label-free quantitative proteomic approach is a mass spectrometry-based proteomics method, which enables the identification and quantification of low-abundance proteins within samples in a comprehensive and unbiased manner with high confidence [21]. Moreover, because of the replacement of stable-isotope labeling steps with advanced informatics, it allows the comparison of any samples of any biological origin [22, 23]. Parallel reaction monitoring (PRM) is a recently introduced targeted method of quantification performed on high-resolution quadrupole-Orbitrap mass spectrometers [24, 25], and generates improved sensitivity, reproducibility, and quantitative dynamic range compared to selected reaction monitoring. PRM has been successfully applied in the validation of relative abundance of target proteins of interest [26]. Therefore, here, we utilized these current state-of-the-art, label-free and PRM-based quantitative proteomic tools in combination with a transcriptomic approach to investigate and compare venom composition between two strains of *T. brontispae*, *T. brontispae*-Bl (Tb-Bl) and *T. brontispae*-On (Tb-On). Tb-On is a sub-population of Tb-Bl that has been experimentally

adapted to a new host, *O. nipae*, continuously for over two years. This study thus yields a valuable dataset, which enables us to address the questions of whether venom apparatus composition responds to selection associated with host availability/switching using experimental evolution, and how and which venom proteins evolve, e.g., quantitative and/or qualitative intraspecific variation. In addition, this dataset contributes to further functional analysis of parasitoid venoms.

2. Materials and Methods

2.1. Wasp strains

The origin of *T. brontispae* has been previously described [17]. Two strains of *T. brontispae* were used in this study. *T. brontispae*-Bl (Tb-Bl) was cultured with one-day-old (newly exuviated) *B. longissima* pupae as hosts. *T. brontispae*-On (Tb-On), a derived sub-population of Tb-Bl, was reared with one-day-old *O. nipae* pupae as hosts consecutively for over two years. Parasitoid strains were maintained at 25±1°C, 85±5% RH, and a photoperiod of 12:12 L:D. After emergence, adult wasps were fed on a streak of 10% sucrose solution in plastic containers, as previously described [17]. Because one-day-old females from both strains have the highest parasitism rate, all the following experiments were performed on one-day-old females without parasitism experience unless otherwise specified.

2.2. Preparation of venom apparatus and transmission electron microscopy (TEM)

Venom apparatus were collected from the *T. brontispae*-On strain. Females were anaesthetized on ice and then dissected in physiological saline (150 mM NaCl, 2.7 mM KCl, 1.8 mM KH_2PO_4 and 10.1 mM Na_2HPO_4, pH 7.2) on a slide glass under a stereomicroscope with the assistance of pins and forceps. The morphology of the venom apparatus was observed through differential interference contrast (DIC) microscopy (Nikon Ni-U, Japan). The excised venom apparatus (gland and reservoir, dissected separately) were fixed with 2.5% glutaraldehyde in phosphate-buffered saline, pH 7.2, at 4 °C overnight and then post-fixed with 1% osmium tetroxide in PBS at room temperature for 2 h. The samples were dehydrated in an ethanol series and embedded in Epon. Ultrathin sections cut with a Leica EM UC7 Ultramicrotome were contrasted with uranyl acetate and lead citrate for 5 min and examined under a Hitachi H-7650 transmission electron microscope.

2.3. Tetrastichus brontispae *reference transcriptome*

A comprehensive *T. brontispae* library was constructed using pooled samples with adequate female wasps from the Tb-On and Tb-Bl strains, i.e., each strain containing approximately 575 abdomens of female wasps of mixed ages (from one to five days of age). Total RNA was isolated using Trizol Reagent (Invitrogen) according to the manufacturer's instructions and treated with DNase I. RNA sample concentration and integrity were determined using a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). Poly(A)-containing mRNAs were enriched using oligo (dT) magnetic beads (Qiagen), fragmented with RNA fragmentation buffer, and subjected to the following procedure: first- and second-strand cDNA synthesis, purification, end repair, single-nucleotide A addition, ligation of adapters, purification of ligated products, and PCR amplification for cDNA template enrichment. The cDNA library was qualified and quantified with an Agilent 2100 Bioanalyzer and ABI StepOnePlus Real-time PCR system, respectively, and then, it was sequenced for 125 bp paired-end reads using the Illumina HiSeq ™ 2500 platform at Genedenovo (Guangzhou, China).

After filtering out the sequencing adapters, reads with more than 5% unknown nucleotides (N), and low-quality reads, the resulting clean reads were assembled using Trinity 2.1.1 [27] and output as unigenes. The clean datasets containing the sequences and the quality scores are available at the NCBI Short Read Archive (SRA) with accession number SRP148751. Unigenes were then annotated by BLASTx with an E-value cut-off of $1e^{-5}$ against the NCBI non-redundant (NR), Swiss-Prot (http://www.expasy.ch/sprot), Kyoto Encyclopedia of Genes and Genome (KEGG, http://www.genome.jp/kegg/), and Cluster of Orthologous Groups (COG, www.ncbi.nlm.nih.gov/COG) protein databases. The open reading frames (ORFs) of unigenes without homology to these databases were predicted using the ESTScan software 1.6 [28]. A search for protein domains was carried out using the PfamScan (ftp://ftp.sanger.ac.uk/pub/databases/Pfam/Tools/) and annotated with the Pfam 26.0 database (November 2011, 13672 families) [29].

2.4. Preparation of T. brontispae *venom apparatus proteins*

Two samples, namely, Tb-On and Tb-Bl, were prepared, and each sample consisted of the harvested venom apparatus (gland plus reservoir) from 2500 female wasps (one day old) of the corresponding strain. The samples were

suspended on ice in 200 μL lysis buffer containing 4% sodium dodecyl sulfate (SDS), 1 mM dithiothreitol (DTT), 150 mM Tris-HCl pH 8.0 by agitating with a homogenizer (FastPrep-24®, MP Biomedical). After boiling for 15 min, they were ultrasonicated (80 W, ten times, 10 s each, 15 s intervals) and centrifuged at 13,400 rpm for 30 min at 25 °C. The resulting supernatant was collected, and the protein concentration was quantified using the BCA Protein Assay Kit (Bio-Rad). For each sample approximately 30 μg of protein was subjected to 12.5% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) to assess protein integrity.

2.5. In-solution trypsin digestion of venom apparatus proteins

The pooled venom apparatus proteins from 2500 female wasps mentioned above were then divided into three replicates for each sample (Tb-On or Tb-Bl), i.e., each replicate contained the extract of approximately 830 venom apparatus. An aliquot of 100 μg of each replicate was subjected to in-solution digestion according to the filter-aided sample preparation (FASP) procedure described previously [30]. In brief, proteins were reduced with DTT (final concentration 100 mM) and boiled for 5 min. The detergent, DTT and other low-molecular-weight components were removed using 200 μL of UA buffer (8 M urea, 150 mM Tris-HCl pH 8.0) by repeated ultrafiltration (Microcon units, 30 kD) facilitated by centrifugation and then alkylated with 100 μL iodoacetamide (IAA, 0.05 M in UA buffer) in darkness for 45 min at room temperature. The filter was washed twice with UA buffer and then with twice 25 mM ammonium bicarbonate (ABC). Next, the proteins were digested with 40 μL of 100 mM ABC containing 50 ng/μL of sequencing-grade trypsin (Promega) at 37 °C for 16-18 h. The digested peptides were collected by centrifugation, and ABC (25 mM, 40 μL) was added to the filter (30 kD) before centrifuging again at 13,400 rpm for 30 min. The resulting peptides were collected as the pooled filtrate, and the peptide content was determined by UV light spectral density at 280 nm using an extinction coefficient of 1.1 of 0.1% (g/l) solution.

2.6. Liquid chromatography (LC) – electrospray ionization (ESI) tandem mass spectrometry (MS/MS) analysis

Approximately 4 μg peptides from each replicate was desalted using C18 Cartridges (standard density, Empore™ SPE, bed I.D. 7 mm, volume 3 ml, Sigma, St. Louis, MO), concentrated by vacuum centrifugation, and resuspended in 40 μL of 0.1% trifluoroacetic acid. Identification of proteins was performed by reversed-phase LC-MS/MS on an Easy nLC 1000 (ThermoFisher Scientific) coupled to a Q Exactive mass spectrometer (ThermoFisher Scientific). The

peptides were loaded on a trap column (Thermo Scientific Acclaim PepMap 100, 3 cm long 100 μm i.d., nano Viper C18) by an autosampler, followed by separation on an analysis column (Thermo Scientific EASY column, 10 cm long 75 μm i.d., 3 μm, C18-A2) in buffer A (2% acetonitrile and 0.1% formic acid) at a flow rate of 300 nL/min using the following gradients: from 0 to 45% buffer B (84% acetonitrile and 0.1% formic acid) in 100 min, from 45 to 100% buffer B in 8 min and 100% buffer B in 12 min. Mass spectrometric data were acquired in a data-dependent manner on the Q Exactive mass spectrometer. Full MS scans (range from 300-1800 m/z with a resolution of 70,000 at m/z 200) were followed by intensity-dependent MS/MS of the top 10 most abundant precursor ions. Higher energy collisional dissociation (HCD) with normalized collision energy of 30 eV was used to fragment the ions, and the fragment ion spectra were acquired in the analyser with a resolution of 17,500 at m/z 200. The dynamic exclusion duration was 40 s, and target values were determined based on predictive automatic gain control with the underfill ratio, which specifies the minimum percentage of the target value likely to be reached at maximum fill time, set to 0.1%. Peptide recognition mode was enabled. Each of the Tb-On and Tb-Bl venom apparatus samples was subjected to three independent LC-MS/MS runs.

2.7. Protein identification

Mass spectra were processed, and protein identification was performed using MaxQuant software version 1.3.0.5 (Max Planck Institute, Germany) [31]. The Q Exactive raw files were searched against the above *T. brontispae* transcriptome database with 16,068 protein entries (including 14503 annotated and 1565 unknown protein sequences) derived from 23,053 assembled unigenes. The search parameters were set as follows: first and main search peptide tolerances of 20 ppm and 6 ppm, respectively; precursor ion and fragment mass tolerances of 20 ppm and 0.1 Da, respectively; a maximum of two missed trypsin cleavage sites; fixed cysteine carbamidomethylation; and variable N-terminal acetylation and methionine oxidation. A target-decoy database approach was used with a cut-off value for the global false discovery rate (FDR) of 1% for both peptide and protein identifications. If peptides matched multiple members of a protein family, a leading razor protein (best scoring protein) was selected. The MS proteomic data were deposited in the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository with the dataset identifier PXD004653.

2.8. Label-free quantitation (LFQ) of protein abundance

Intensity-based absolute quantification (iBAQ) calculated by MaxQuant was adopted for protein abundance as previously described [32]. In iBAQ, the sum of all peptide peak intensities for each protein is divided by the number of theoretically observable tryptic peptides, which provides an accurate determination of relative protein amounts. The expression level of each protein was calculated as its average iBAQ from three independent replicates. Proteins were considered to have a significant difference in abundance between the two strains (Tb-On and Tb-Bl) with a difference of greater than twofold and a corrected *p* value significance *A* (calculated by the MaxQuant software) < 0.05 using a two-tailed Student's *t* test [31]. Statistical analysis was performed with Perseus software version 1.3.0.4 (Max Plank Institute, Germany).

2.9. Bioinformatic analyses

To determine the putative biological and functional properties of *T. brontispae* venom proteins, all identified venom proteins were retrieved and locally searched against the NCBI nr database using the NCBI BLAST-client software (ncbi-blast-2.2.28+-win32.exe) to find homologous sequences with the top 10 blast hits and a cut-off E-value of $1e^{-3}$ for each query sequence, followed by Gene Ontology (GO) terms using Blast2GO (version 2.7.1) for GO mapping and annotation. After annotation and annotation augmentation, the proteins were also blasted against KEGG GENES (insects) to retrieve the corresponding KOs, followed by mapping to pathways in KEGG by the KEGG Automatic Annotation Server (KAAS). Signal peptide prediction was performed online (http://www.cbs.dtu.dk/services/SignalP/). Specialized similarity searches were performed against the Tox-Prot database (http://expasy.org/sprot/tox-prot/) [33] and the demonstrated or putative factors of parasitoids described in the literature, manually.

Moreover, the resulting differentially produced proteins between the two samples were subjected to GO functional and KEGG pathway enrichment analyses using Fisher's exact test. Only GO terms and biological pathways with *p* values < 0.01 and 0.05, respectively, was considered as statistically significantly enriched. In addition, expression profiles of the differentially produced proteins were created by hierarchical clustering analysis using the pheatmap package in R version 3.2.4.

2.10. Quantitative real-time PCR

Total RNA samples were prepared in three replicates for each strain, and each RNA sample was extracted from 250 venom apparatus using TRIzol Reagent (Invitrogen) according to the manufacturer's instructions and subjected to cDNA synthesis using the ThermoFisher Scientific Verso cDNA Kit, where the RT enhancer can remove contaminating DNA and eliminate the need for DNase I treatment. Seven differentially produced proteins and eight abundant proteins with similar iBAQ levels between strains were selected for quantitative real time PCR (qRT-PCR) analysis. qRT-PCR was performed according to our previous protocol [18], with three technical replicates for each biological replicate using the Power SYBR Green Master Mix Kit (Invitrogen) in an ABI 7500 System. Transcripts were normalized to glyceraldehyde-3-phosphate dehydrogenase (GAPDH). Data analysis was performed by an unpaired two-tailed Student's *t* test using GraphPad InStat (GraphPad Software Inc.). Primer sequences are provided in Supplemental Table 1.

2.11. Immunoblotting

The same protein samples as those in the label-free profiles described above were used. Equal venom apparatus protein aliquots (6 μg in each lane) for the Tb-On and Tb-Bl strains were separated by (10%) SDS-PAGE gel and then transferred to nitrocellulose membranes using Trans-Blot SD (Bio-Rad). The membranes were blocked with 5% (w/v) skim milk powder in Tris-buffered saline containing 0.1% (v/v) Tween-20 for 1 h at room temperature before incubating with primary antibodies at a dilution of 1:1000. The antibodies for unigene0022364 (serpin 3/4) and unigene0018080 (calreticulin) were generated in rabbits using peptides (Supplemental Table 2) produced by the pGEM-4T-1 expression plasmid. GAPDH, as a reference control, was purchased from Abcam. The secondary antibody for detection of the primary rabbit antibody was horseradish peroxidase-conjugated goat anti-rabbit IgG at a dilution of 1:5000, and that for GAPDH was rabbit anti-goat at a dilution of 1:3000. The immunoreactive signal was detected using Immobilon Western Chemiluminescent HRP Substrate (Millipore) on an Amersham Imager 600 QC (GE Healthcare).

2.12. Targeted analyses by PRM

Seven significantly differentially produced proteins of interest (Supplemental Table 10) between the two strains

were selected for further targeted quantification by PRM in Shanghai Applied Protein Technology Co., Ltd, China. Briefly, proteins were prepared from 800 newly dissected venom apparatus for each sample. Peptides were prepared as described above for the label-free profiles. An AQUA stable-isotope peptide as an internal standard reference (Supplemental Table 11) was spiked in with each sample. Digested peptides were desalted on C18 stage tips prior to reversed-phase chromatography on an Easy nLC-1200 system (ThermoFisher Scientific). One hour liquid chromatography at a flow rate of 250 nL/min was used with the following gradients: 5 to 23% buffer B in 42 min, 23 to 40% buffer B in 8 min and 40 to 100% buffer B in 8 min. PRM analysis was performed using a Q Extractive HF mass spectrometer (ThermoFisher Scientific). Optimal collision energy, charge state, and retention time for the most significantly regulated peptides were generated experimentally using unique peptides (Supplemental Table 11) of high intensity and confidence for each target protein. The mass spectrometer was operated in position ion mode with the following parameters: the full scan was collected with a resolution of 60,000 at 200 m/z, the automatic gain control target was 3×10^6 and the maximum injection time was at 200 ms; the following 20 PRM scans were performed at a resolution of 30,000 at 200 m/z, AGC target of 3×10^6, maximum injection time of 120 ms and individual isolation window of 1.6 Th. A normalized collision energy of 27 in an HCD collision cell was employed for fragmentation. All PRM data analyses and data integration were performed using Skyline software version 3.5.0 (MacCoss Laboratory, University of Washington) [34]. Three replicates were included for each sample in the PRM-MS analysis. Relative peptide quantification was calculated by dividing the peptide peak area by the labelled reference peptide peak area. A two-tailed Student's *t* test was used to estimate the significance of the difference in relative peptide abundance between the two strains of *T. brontispae*. The MS proteomics data for PRM were deposited in the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository with the dataset identifier PXD008566.

3. Results

3.1. Structure of the T. brontispae *venom apparatus*

The venom apparatus of *T. brontispae* consists of a large, transparent, spherical or ovoid venom reservoir and a venom gland, associated with an elongated and tubular Dufour's gland, which is closely appressed to the base of the venom reservoir (Fig. 1A). The venom gland is located at the anterior end of the reservoir and appears as an enteroid

structure consisting of secretory units in a compact form and a central collecting duct (Fig. 1A). The venom duct, linking the reservoir to the ovipositor, is invisible due to its small size on the photomicrograph. Cross sections reveal that the enteroid venom gland possesses a central lumen lined with an intimal layer highly folded into a pronounced basal labyrinth, narrow and tightly packed duct cells and peripheral concentrically arranged secretory cells with large nuclei (Fig. 1B). As described in *Ampulex compressa* [35], a cuticle-lined small ductule named the smooth canal in *Leptopilina* spp. [36] was also found (Fig. 1B, 1C and 1E). This histological organization corresponds to the "type III" insect epidermal glands described previously [37], in which the ductule drains secretory products into the central collecting duct. The venom gland is separated from the hemocoel space by a thin basal membrane. The secretory cells contain numerous cellular organelles, such as mitochondria, extensive rough endoplasmic reticulum and vacuoles; these cells typically contain secretory vesicles and have end-apparatus surrounded by well-developed microvilli (Fig. 1C and 1D). The secretory vesicles are present throughout the cytoplasm but occur more frequently proximal to the extracellular portion of the end-apparatus (Fig. 1D). These features indicate that the secretory units produce large quantities of protein.

The venom reservoir is surrounded by a thin longitudinal muscle fibres layer of uniform thickness (Fig. 1F), similar to that of the "type II" venom apparatus described previously in braconid wasps [38, 39]. Therefore, the muscles of *T. brontispae* are not presumed to be innervated and are likely of minor importance in the injection of venom into hosts. Instead, the ejection of venom into the ovipositor may be aided in part by general metasomal movement and may be brought about gradually [40]. Inside the muscle layer, an epithelial non-secretory cell layer with distinct, relatively lengthened nucleus is closely attached to a thin, highly folded, cuticle-lined intima (Fig. 1F), resembling epithelial cells described in *Pteromalus puparum* [41]. Very few organelles were observed in the cytoplasm of the epithelial cells. The lumen of the reservoir is filled with putative aggregates of venom components (Fig. 1F). Thus, the reservoir mainly serves to store venom secretions and is filled passively, as has also been suggested for some other wasp species [42].

No remarkable age-related changes in the fine structure of the venom apparatus were observed for adult females from one to five days of age (Table 1), which is indicative of the "type A" characteristic according to Van Merle and Piek [43]. Thus, there are no generative changes of the venom gland in the whole lifespan of *T. brontispae*.

3.2. General overview of the T. brontispae *reference transcriptome*

Since no reference genome was available, the recently described sequencing/bioinformatic approach was adopted to identify the venom proteins [44]. RNA-Seq deep sequencing analysis generated approximately 22.7 and 20.3 million paired-end reads from the Tb-On and Tb-Bl libraries, respectively. Using *de novo* assembly, the pool of all reads from the two libraries yielded 23,054 unigenes with an N50 length and an average size of 1791 and 1030 bp (Table 2), respectively, including 7791 unigenes (33.79%) over 1000 bp in length (Supplemental Fig. 1). Using BLASTx, more than 60% (14,592) of unigenes had significant similarities in the NR, Swiss-Prot, KEGG and COG databases with a cut-off E-value of $1e^{-5}$ (Table 2). As expected, most of the unigenes had best matches and first hit against proteins of the Hymenoptera, with the greatest number of matches to *Nasonia vitripennis* (Table 2). To construct a database against which our proteomics data could be searched, each unigene was translated, and the amino acid sequences of the predicted protein-coding regions of a total of 16,068 unigenes, including 1565 unknown sequences with ESTScan predictions, were produced, more than 60% (9887) of which could be assigned a Pfam-A annotation (Table 2).

3.3. The T. brontispae venom apparatus proteome

The protein content of the *T. brontispae* venom apparatus was resolved in bands ranging from 15 kDa to more than 250 kDa on a 12.5% SDS-PAGE gel, and most of the proteins migrated between 30 and 100 kDa (Fig. 2). A comparison of these electrophoretic profiles showed that most bands migrated at the same apparent mass between the two strains of *T. brontispae*, but quantitative variation certainly occurred in some of the proteins (Fig. 2, one example is shown by red arrows). For an exhaustive proteomic analysis, the proteins of the venom apparatus extract were subjected to in-solution trypsin digestion, followed by data-dependent LC-MS/MS analysis. Using MaxQuant software, the mass spectrometric data from the pools of the two samples corresponding to 8638 peptide sequences (Supplemental Table 9) were processed to search against the *T. brontispae* transcriptome, which resulted in the identification of 1505 proteins (ranging from 3 to 450 kDa, Supplemental Table 3) with a minimum of one unique peptide, mainly via hymenopteran proteins (92%) (Supplemental Table 3). It is worth noting that while lower molecular weight proteins could still be recovered, due to the enrichment method we used in the study, the resulting venom extract would be skewed towards high molecular weight proteins. Of these proteins, 186 were predicted to be

secreted with the SignalP program (Supplemental Table 3). Notably, the true number of proteins predicted to be secreted is probably higher than 186, because of the incompleteness of some assembled unigenes.

To obtain a global functional level view of the proteome, these proteins were sorted into 12 functional level 2 GO terms on the basis of their molecular function. Among these terms, the over-represented categories were binding and catalytic activity, followed by structural molecule activity, transporter activity, enzyme regulator activity, and electron carrier activity (Fig. 3A). To examine the functions in more detail, the proteins were assigned at a more detailed level. At level 3, the venom apparatus extract contained a large number of heterocyclic and organic cyclic compound binding proteins, enzymes (particularly hydrolases, oxidoreductases and transferases), and transmembrane and substrate-specific transporters (Fig. 3B and Supplemental Table 4). Further, to understand the key metabolic pathways, the proteins in the venom apparatus were annotated with the KEGG database, with a total of 671 proteins assigned to 275 KEGG pathways (Supplemental Table 5). Among these, the major represented functional groupings were proteins related to ribosome, protein processing in endoplasmic reticulum, carbon metabolism and oxidative phosphorylation (Fig. 4). These results support the largely secretory function of the venom gland and its contribution to venom production.

3.4. Comparison of proteomic analysis in the venom apparatus of Tb-On versus Tb-Bl

An overview of protein abundance was obtained using iBAQ analysis with three replicates of each strain (Supplemental Table 3). The iBAQ data were compared with a limit of 1081 proteins identified at least in two of the three replicates for both strains, which represented a good correlation between replicates (Pearson's correlation coefficient of r>0.95) (Supplemental Fig. 2). To visualize the distribution of proteins between the two strains, the iBAQ ratios ($\log_2$(Tb-On/Tb-Bl)) were calculated, and a volcano plot was generated for the 1081 proteins (Fig. 5). Using the criteria of fold change > 2 ($\log_2$ ratio of ± 1.0) and significance $A < 0.05$ (Student's t test), approximately 7.6% of the 1081 proteins were significantly differentially produced (Supplemental Table 6). The majority of proteins were shared and present at similar iBAQ levels in the two strains (Fig. 5), e.g., the top eight abundant proteins (Table 3). Considering that virulence in parasitism mainly relies on the abundant venom factors, the four proteins annotated with putative functions among these top eight abundant proteins are likely implicated in the virulence activity of the venom. Among the proteins exhibiting significant differences, 63 proteins showed dramatically increased levels and

19 proteins showed significantly reduced levels in Tb-On *versus* Tb-Bl (Supplemental Table 6). To better visualize the differences in protein production between the two strains, the 82 significantly differentially produced proteins were subjected to hierarchical clustering analysis, and the dendrogram showed two unambiguous clusters of proteins, characterized by down-regulation or up-regulation (Fig. 6). Each cluster could be divided into several subclusters. A difference occurred between one replicate and the other two replicates in the Tb-On strain in subcluster B1, which is possibly due to the deviation among the replicates. In addition, 49 proteins (identified in at least two of the three replicates) were identified in only one of the strains: 39 proteins were exclusively identified in Tb-On, and ten proteins were identified in only Tb-Bl (Supplemental Table 6). Proteins exclusively identified in one of the strains may be present in the other strain but at levels that are beyond our limit of detection; alternatively, they could indeed be absent in the other strain. Proteins selected for independent validation are marked on the volcano plot (Fig. 5).

To further interpret the biological significance of the complex cohort data, the significantly differentially produced proteins (including those identified in only one of the two strains) were plotted against GO categories. As with all the 1505 identified venom apparatus proteins, binding and catalytic activity were among the largest functional categories in the molecular function cluster for these significantly differentially produced proteins (Supplemental Table 7). GO enrichment analysis revealed that 12 GO terms in the biological process category were significantly enriched ($p < 0.01$) and mainly related to melanotic encapsulation of foreign target and melanization defense response (Fig. 7A). In the KEGG enrichment analysis, eight pathways, e.g., fatty acid biosynthesis, adipocytokine, AMPK (AMP-activated protein kinase), and PPAR (peroxisome proliferator-activated receptors) signaling pathways, were significantly enriched (Fig. 7B). Furthermore, almost all the proteins enriched, except two proteins, mitochondrial fission 1 protein-like (unigene0019473) and rRNA 2-O-methyltransferase fibrillarin-like (unigene0019200), showed dramatically increased levels (Fig. 6, marked in red) or were uniquely identified in the Tb-On strain (excluding unigene0016298, Supplemental Table 8).

3.5. Validation of proteins at the levels of mRNA and protein

Abundant venom proteins are likely associated with parasitism success. To assess the transcriptional levels of the abundant proteins identified in the venom apparatus, we performed quantitative real-time PCR on six of the top eight most abundant proteins (the assembled lengths of the remaining two proteins were relatively short, therefore, they

were not taken into consideration) (Table 3). For all the six proteins, the transcripts indeed presented dramatically higher transcript levels (Fig. 8A), which confirmed the proteomic data. But two proteins (5-unigene0018515, 4-coumarate-CoA ligase-like 4; and 6-unigene0000125, unknown protein) that showed similar abundance between the two strains in the proteomic analysis showed differential expression at the mRNA level (Fig. 8A). This deviation may be because protein level in some cases may not be synchronized with its corresponding transcript level. We also used qRT-PCR to validate the differential expression of venom apparatus proteins between the two strains by randomly selecting seven proteins. Differential expression of six out of the seven genes, with the exception of unigene0019736 (succinyl-CoA ligase), was validated (Fig. 8B).

Calreticulin is a multifunctional Ca^{2+}-binding chaperone protein, which has been demonstrated to function in preventing host encapsulation and hemocyte aggregation [45, 46]. Serine protease inhibitors (serpins) are also key venom proteins, which may interact with serine proteases to inhibit phenoloxidase (PO) activation (see discussion for details) [47]. As these two proteins (calreticulin, unigene0018080; and serpin 3/4, unigene0022364) showed significant differential expression between the two strains at the mRNA level but not in the proteomic data (Fig. 8A and 8B), their protein production in the wasps was also assessed by western blot analysis with anti-Calreticulin and anti-Serpin antibodies. The results supported the proteomic data that no difference in the protein production levels exists between the two strains (Fig. 8C).

3.6. *Validation of significantly differentially produced proteins by PRM*

The label-free analysis revealed that the significantly differentially produced proteins between the two strains were mainly enriched in the melanotic encapsulation of foreign target GO term, and in the fatty acid biosynthesis, adipocytokine and AMPK signaling pathways. To confirm these observations, seven significantly differentially produced proteins involved in these GO terms and pathways were selected for PRM analysis. Two unique peptides with anticipated chemical stability were chosen for each protein (Supplemental Table 11). Relative protein abundance was expressed as the average of the two normalized peptide peak areas. Based on this, all the selected venom proteins, except unigene0019710 (fatty acid synthase-like), displayed elevated expression levels (> 1.3-fold, $p < 0.05$, Supplemental Table 11) in the Tb-On strain compared with those of the Tb-Bl strain, which showed identical trends to the label-free data (Fig. 9). However, PRM analysis enabled the detection of unigene0021082 (carnitine o-

palmitoyltransferase), which was not detected by label-free analysis in the Tb-Bl strain (Supplemental Table 10). Thus, for the majority of differentially produced proteins, the label-free proteomics data and PRM data were broadly consistent, showing the same trends.

4. Discussion

Venom apparatus is virtually ubiquitous in Hymenoptera, and its anatomies have been well described in numerous ichneumonid and braconid wasps [40], and more extensively, in some pteromalid species, e.g., *N. vitripennis* [48], *Anisopteromalus calandrae*, *Pteromalus cerealellae* [49], and *P. puparum* [41]. In the present study, the gross morphology of the venom apparatus of *T. brontispae* belonging to Eulophidae was analyzed. The stacked enteroid unbranched venom gland and the associated single elongated tubular Dufour's gland with approximately an equal length to that of the venom gland constitute a different combination of features from those of other documented parasitoids. This thereby is also consistent with the previously stated view that venom apparatus are considerably divergent, and their variety seems to be correlated to some degree with the life strategies of parasitoids as well as the evolutionary paths followed by different taxa [5].

In species of parasitic wasps devoid of polydnaviruses or other virus-like particles, venom plays essential roles in the virulence of parasitoids towards their hosts [42]. Thus, the identification of venom components is of great value for understanding interactions between venoms and hosts. Here, we provide an in-depth venom apparatus proteome of *T. brontispae* and identify a total of 1505 individual proteins. The reasons for the detection of a large number of proteins, far more than those proteins identified in the venoms of other parasitoids, may be due to 1) the large number of female wasps we used (2500 female wasps from each strain), which allowed the detection of proteins present in low abundance in individual parasitoids, 2) our digestion of all proteins rather than of only the major bands observed by electrophoretic analysis, and/or 3) collection of the venom reservoir and gland together, rather than from the venom reservoir only. Investigators focus mainly on the abundant components in venoms, which can potentially lead to overlooking factors that also have key functions in parasitism even though they are at low concentrations. For example, an extracellular superoxide dismutase in low abundance in venom may protect stored venom proteins from oxidative stress rather than acting as a virulence factor [50]. Indeed, the proteins identified here from the venom apparatus (venom gland plus reservoir) contain many cellular proteins, e.g., a large number of proteins associated with ribosome

and processing in the endoplasmic reticulum (Fig. 4). However, a number of analyses can be performed to increase the chance of identification of the "pure" venom proteins, e.g., assignment of putative functions on the basis of bioinformatics such as KEGG analysis, and extensive functional characterization of a given protein, such as enzymatic functions, together with accurate analyses using a purified recombinant protein and/or RNAi analysis as performed for the venoms of *L. boulardi* [51] and *N. vitripennis* [52].

The identified *T. brontispae* venom apparatus proteins are mainly involved in the binding and catalytic activity terms according to the GO system at level 2; these terms are also the most common functional categories assigned to venom gland ESTs from *Chelonus inanitus* [53]. It is now established that most parasitoids tend to use a common set of venom proteins, though in different abundances, e.g., enzymes, protease inhibitors, binding proteins and some immune related proteins [5, 11, 54]. Our GO classifications at level 3 are also consistent with this view.

Comparisons of parasitoid venom proteins between different hymenopteran species, such as those of more closely related species, have attracted the attention of investigators interested in the composition of venoms in Hymenoptera. However, variability of venom components in intraspecific strains or populations has rarely been documented except in *Leptopilina boulardi*, where some of the main venom factors identified in the two different geographical strains of *L. boulardi*, ISm and ISy, display quantitative and qualitative variations [12]. For instance, the major virulence factor RhoGAP shows a quantitative difference between the two *L. boulardi* lines [55], and their serpin proteins (LbSPNm and LbSPNy) differ in SDS-PAGE migration despite their similar abundance [12]. Moreover, one of the novel parasitoid venom proteins, tissue inhibitor of metalloproteinase (TIMP), is secreted in high amounts in the venom of the ISm strain only [12]. These significant quantitative and qualitative variations in venom components are responsible for the observed difference in virulence properties [10, 56]: ISm, highly virulent against only *D. melanogaster*, induces permanent immunosuppression and targets immune cellular components with morphological changes of lamellocytes in *D. melanogaster* [57, 58]; whereas ISy, able to suppress immune defenses of both *D. melanogaster* and *D. yakuba* depending on the host resistance genotype, transiently suppresses the encapsulation of *D. yakuba*, and targets the humoral components of encapsulation via inhibition of the phenoloxidase cascade [47, 59, 60].

In the present study, in addition to specialized similarity searches against the Tox-Prot database and the demonstrated or putative factors of parasitoids described in the literature in the last few years, a comparison of venom apparatus proteins between two strains of *T. brontispae* were investigated. The Tb-Bl strain parasitizes the pupae of *B.*

longissima as its habitual host, whereas the Tb-On strain is an experimental derivative of Tb-Bl, which has been consecutively exposed to *O. nipae* pupae as host for over two years. Our combined analyses of venom composition revealed the presence of two categories of proteins, which we examined in more detail: (i) a set of proteins in high abundance in both strains and (ii) a set of differentially produced proteins with similarity to host regulatory functions described in other parasitoid species.

4.1. The most abundant proteins shared by the two strains

Since the abundant venom proteins are most likely involved in parasitism success, here, we mainly focused on the eight most abundant proteins (Table 3), whose corresponding transcript levels in the venom apparatus tissue were validated by qRT-PCR (Fig. 8A).

Neprilysin-like (NEP-like) proteins, which belong to the M13 peptidase family, are zinc-metalloendopeptidases and integral plasma membrane ectopeptidases with their active catalytic site on the extracellular surface [61]. In addition to the membrane-bound forms, soluble isoforms of neprilysin can be found in body fluids [62]. In the present study, a NEP venom protein (unigene0000904) was identified with a signal peptide, as expected, and without transmembrane helices, when predicted using TMHMM Server v.2.0, which suggested that it is a soluble protein. An amino acid sequence alignment showed that, similar to *Drosophila melanogaster* NEP2 [63], a soluble member of the neprilysin family, *T. brontispae* NEP also contained a zinc binding motif ($HE_{XX}H$) and two catalytic motifs (GENID and VNAFY) but with substitutions for some residues (Fig. 10A). NEP-like proteins have key roles in inactivating signaling peptides involved in modulating neuronal activity, blood pressure and the immune system in mammals and insects [64, 65]. Until now, NEP-like proteins have been described in the venom of many parasitoid species. For example, similar to the venom of *T. brontispae* in the present study, NER-like protein was also one of the most abundant proteins in the venom of *Microctonus hyperodae* [66], whereas it was present in low abundances in the venoms of *A. ervi* [13], *L. boulardi* and *L. heterotoma* [12] on the basis of number of ESTs. The difference may be due to the specificity of each parasitoid and/or their evolutionary adaptations to regulate host physiology. In addition, a NEP-like protein was found in association with virus-like particles produced in the calyx region of *Venturia canescens* and was hypothesized to result in systemic induction of specific pathways of host immune regulation by releasing immune-specific peptides [67]. Notably, another protein with 66 amino acid residues (unigene0018250) in

low abundance, also coding for neprilysin-2-like, was identified in only the Tb-On strain based on the iBAQ results (Supplemental Table 6). According to an NCBI BLAST analysis, this new protein is membrane-bound and shares 63% similarity with unigene0000904 in the known region (Supplemental Fig. 3A). Further analysis is required to elucidate the role of these proteins in virulence in parasitism or whether unigene0018250 is a neutral variant, and also the evolutionary relationship between these two NEP-like proteins.

Kynurenine--oxoglutarate transaminase (unigene0004125), also named kynurenine 2-oxoglutarate transaminase, L-kynurenine aminotransferase or kynurenine aminotransferase (KAT), is responsible for the production of kynurenic acid (KA) using kynurenine as the direct precursor and plays a key role in maintaining physiological levels of KA [68, 69]. KA is a metabolite in the tryptophan (Try) oxidation pathway and is recognized as a broad-spectrum antagonist of ionotropic excitatory amino acid receptors in vertebrates, protecting the central nervous system from over-stimulation by excitatory cytotoxins [70]. In insects, kynurenines are implicated in a variety of biological functions such as eye development, tissue remodeling during metamorphosis, and modulation of complex behaviours [70-74]. This venom protein is possibly associated with neuronal dysfunction and developmental manipulation of the hosts. To the best of our knowledge, there is no report yet of KAT's involvement in host-parasitoid interactions either as a virulence factor or a venom protein.

The 4-coumarate-CoA ligase 4 (4CL4, unigene0018515), belonging to the adenylate forming domain (AFD) Class I superfamily, is an enzyme that catalyzes the activation of 4-coumarate and related substrates to the respective CoA esters, playing a pivotal role in the phenylpropanoid-derived compound (PDC) pathway [75]. The PDC pathway, as well as its branch pathways, generates various classes of natural compounds with essential functions in plant growth, development and environmental interactions, such as structural support, flower color, pollination and pathogen defense [75, 76]. It is thus interesting to investigate how 4CL4 has been recruited and evolved to perform venom functions or be involved in venom metabolism. An amino acid sequence alignment indicated that the 4CL4 from *T. brontispae* also harbors three conserved catalytic residues, Lys-Gln-Lys, and the residues essential for substrate binding, which are located between the putative AMP binding domain (Box I) and the GEICIRG domain (Box II) (Fig. 10B). It is worth noting that most of the conserved residues for substrate binding in parasitoids were replaced compared with those in plants. Maximum likelihood phylogenetic analysis showed that the 4CLs from parasitoids clustered within a single clade, whereas firefly luciferases, which are also members of the AFD Class I superfamily from the

Lampyridae, clustered together with the 4CLs from plants (Supplemental Fig. 4). Recently, lateral gene transfer (LGT) of glucoside hydrolase family 19 (GH19) chitinase from the unicellular microsporidia/Rozella clade into parasitoid wasps of the superfamily Chalcidoidea was described [77]. LGT is a recognized source of eukaryotic genome innovation and can provide eukaryotes with novel biochemical and other functions. GH19 chitinases are widespread in plants, bacteria and microsporidia and are used in nutrient acquisition or defense, while in Chalcidoidea, it has been recruited as a venom protein. Therefore, further studies are needed to determine the exact function of the laterally transferred 4-coumarate-CoA ligase 4 in the parasitoid wasp or its interaction with the host. In medicinal plants, certain PDCs are known to have antioxidative, anti-inflammatory, immunomodulatory, antibacterial and antiviral effects [76, 78]. These suggest a possible role for this ligase, upon envenomation, in defending the host from opportunistic microbe infection or in depressing the host immunity.

Venom protein r-like protein (unigene0009814) was the most abundant according to the iBAQ analysis in *T. brontispae*. This protein has not been described yet in any parasitoid venom but was identified in *Streptomyces clavuligerus* (fungus), *Coptotermes formosanus* (Formosan subterranean termite) and *Rhynchophorus ferrugineus* (red palm weevil) based on the UniProtKB database. Surprisingly, another venom protein r-like protein (unigene0001674) in low abundance was detected in only the Tb-On strain. These two proteins presented only 31% identity, despite having equivalent known sequence lengths (Supplemental Fig. 3B). Although these r-like proteins may not be actual venom proteins, but part of the proteome of the venom apparatus, they have been annotated as "salivary secreted peptide" and also show similarity to venom protein R from *Nasonia vitripennis* venom reservoir [79], though with low identify (30%) and high E value (2e-08). To determine whether "venom protein r-like" proteins from *T. brontispae* are "pure" venom proteins needs further investigations.

The remaining four abundant proteins had no similarity with other known proteins in generalist databases, suggesting that they are newly discovered venom proteins, thus constituting venom proteins specific to *T. brontispae*. Further investigation is needed to understand the functions of these proteins.

Notably, another venom protein calreticulin (CRT, unigene0018080), which was similar in transcript abundance to the most abundant venom proteins by qRT-PCR (Fig. 8A), was also shared between the two strains by the iBAQ analysis and western blot conformation using the calreticulin-specific antibody (Fig. 8C). Calreticulin is a multifunctional Ca^{2+}-binding chaperone protein identified in the venom fluids of many parasitoids [11]. The virulence

of *Cotesia rubecula* venom CRT against its host in preventing encapsulation *in vitro* by inhibiting hemocyte spreading behavior was evidenced previously without deciphering the mechanism [45]. In addition, a hemocyte anti-aggregation effect of venom CRT was observed in *P. hypochondriaca* [46]. We are now aiming at elucidating the underlying mechanism of CRT function in *T. brontispae* parasitism of the host *O. nipae*.

4.2. *Differences in venom apparatus proteins between* the *Tb-On and Tb-Bl* strains

Our enrichment analyses demonstrated that the significantly differentially produced proteins were mainly enriched in fatty acid biosynthesis and melanotic encapsulation response by KEGG and GO enrichment analyses (Fig. 7 and Supplemental Table 10), respectively. Further, these significantly enriched proteins presented strikingly increased levels in the Tb-On strain, which was confirmed by PRM analysis as well (Fig. 9), except the fatty acid synthase-like protein (unigene0019710). These enrichments are most likely due to adaptations that are related to host specificity. Although the pupae of *O. nipae* are morphologically and biologically similar to those of *B. longissima* [80], *O. nipae* pupae generally have a smaller body size and lower fat body abundance than those of *B. longissima* (Supplemental Fig. 5). *T. brontispae* typically prefers to embed its eggs within the host fat body (our observation), which seems to be more beneficial for immunoevasion. From another perspective, lipid is an essential nutrient for the development and growth of parasitoids [81]. In *N. vitripennis*, it has been demonstrated that, although this species can develop on over 60 different host species, its venom induces lipid metabolism alterations in only the host *Sarcophaga*'s [82]. Parasitoid species with superficially similar developmental strategies can differ considerably in physiological dependence on their host, and their adaptations to different hosts can result in highly specific mechanisms to increase lipid levels and optimize host exploitation [83]. After injection of *T. brontispae* venom, freely floating lipid particles were observed in the hemolymph of the host *O. nipae*, and some of the released lipid particles were phagocytized by hemocytes (unpublished data). These suggest that the host *O. nipae* is probably forced to alter its lipid levels to the benefit of the parasitoid in the Tb-On strain.

In addition to the proteins enriched in specific GO terms and pathways described above, some differentially produced venom proteins common to other wasp venoms were examined. Two proteins, unigene0020436 and unigene0020437, both corresponding to serpin 5, presented significantly reduced abundance in the Tb-On strain at the protein level (Fig. 8B). Serpins belong to the serine protease inhibitor family, and their expression specialized in the

venom has been presumed to function in the PO cascade by interacting with one or several serine proteases when injected into hosts [5]. This inhibition of activation of the PO cascade was experimentally shown in *L. boulardi* [47]. In addition, serpins have been reported in the venoms of several other parasitoids, such as *Hyposoter didymator* [84], *L. boulardi*, *L. heterotoma* [12, 44], and *Microplitis demolitor* [85]. Interestingly, another serpin (unigene0022364) with homology to serpin 3/4 was found to be present in greater abundance than those of unigene0020436 and 0020437 at the transcript and protein levels (Fig. 8B). But this new serpin (unigene0022364) was found in similar abundance in the two strains at the protein level by iBAQ analysis (Fig. 8B), which was confirmed by western blot using a unigene0022364-specific antibody (Fig. 8C). Moreover, an alignment of amino acid sequences at key residues in the reactive center loop (RCL) (Supplemental Fig. 3C), which might be responsible for protease target specificities [86], implied that these serpins may have distinct biological characteristics, e.g., some function during parasitism, while others function in stabilization of venom components, as described previously [11]. Because of the incompleteness of sequence, RCL analysis could not be performed for unigene0020437. It is worth noting that unigene0022364 resembles a serpin from the host *O. nipae* at the positions P1, P1' and P2' of RCL (unpublished data). Its possible venomous function and mechanism in parasitism are now under investigation.

Another venom protein with expected signal peptide was exclusively found in the Tb-On strain, though in low abundance based on the iBAQ analysis; This protein is homologous to serine protease homologs (SPHs) (Supplemental Table 6). SPH venom proteins have also been identified in other parasitic wasps, such as *C. inanitus* [53] and *C. rubecula* [87]. SPHs present in the hemolymph of insects normally function as cofactors for proPO-activating protease (PAP) [88]. In contrast, an SPH in the venom of *C. rubecula* inhibits melanization of the host *Pieris rapae* hemolymph [87]. It may therefore be informative to confirm the function of this protein in the Tb-On strain and assess its effect on the host *O. nipae* melanization.

In our present study, the combined analyses of two datasets corresponding to two different strains originating from the same geographical region but adapted to two different hosts demonstrated that these two strains differed in their venom composition (quantity or quality), and thus, the strains varied in virulence factors when targeting their different hosts, *O. nipae* and *B. longissima*, even though the major venom proteins were shared between them. It has been noted that venom variability can occur not only between closely related species or populations but also among individuals of a given population [2, 10]. The occurrence of such variability is a prerequisite for selection on parasitoid

venom, and a driver in parasitoid adaptation in response to selection pressure imposed by the host species. Thus, it is not surprising to observe the presence of variation in venom composition between the Tb-On and Tb-Bl strains despite their coming from the same geographical region. The presence or abundance of specific venom proteins in one of these strains suggests a large potential for rapid adaptation of *T. brontispae* to changes in host species, since the Tb-On strain was experimentally derived from a sub-population of Tb-Bl using *O. nipae* pupae as hosts consecutively for less than three years. Therefore, this reflects a short-term adaptive process. Although the involved molecular machinery responsible for this process remains to be elucidated, it could likely be due to differences in cis-regulation of transcription or to quick and reversible regulation by microRNAs [12, 89]. Moreover, the emergence of a multigenic family for the abundant venom proteins (e.g., neprilysin-like protein and venom protein r-like protein) was detected in the Tb-On strain. Similar events were also identified for the major venom protein RhoGAP in *L. boulardi* [12] and for reprolysin metalloprotease-like genes in the *M. demolitor* venom gland [85]. It thus suggests that duplication of genes with other functions, as commonly reported in snakes or spiders [9, 90], occurs in *T. brontispae* in response to host selection, though this hypothesis remains to be tested.

5. Conclusion

Thanks to the state-of-the-art approaches in proteomics and transcriptomics, our present study provides data from the Eulophidae for comparative investigation of the diversity in the composition of parasitoids' venom, as well as short-term adaptations to host selection in an experimental evolution setting. Our data suggests the presence of a number of conserved as well as novel proteins in the *T. brontispae* venom apparatus, and that this composition may change when the parasitoid is experimentally adapted to a new host. Interesting questions that remain to be investigated include whether these changes translate into functional shits at the physiological level, and what is the rate of adaptive changes and intra-population variability following host change. Further, because our venom proteomics data are from venom reservoir plus the gland, and the abundant proteins based on iBAQ value do not necessarily represent venom components, the roles of essential proteins in parasitoid virulence should be investigated by functional approaches, such as RNA interference techniques. In addition, the key findings that fatty acid biosynthesis and melanotic encapsulation response in the host may be responsible for the intraspecific venom variation between the two strains of *T. brontispae* should be validated with additional experiments, such as quantitative and

qualitative analyses of host fat body cells.

Funding

This work was supported by grants from the National Natural Science Foundation of China (31471829, 31301727 and 31672086), the FAFU's Science Foundation for Distinguished Young Scholars (XJQ201502), the Natural Science Foundation of Fujian Province, China (2014J01088), and the National Key R&D Program of China (2017YFC1200600).

Biological significance

Recent venom analyses from selected species have revealed considerable complexity and diversity in venom composition existing even between strains and individuals, which may partly determine the potential for parasitoid adaptation. However, the data in regard to full and accurate quantity of venom compositions at the protein level is not available. The current study provides an accurate proteomic data for comparative investigation of the diversity in the composition of the parasitoid venom apparatus, thus informing us of the nature of venom evolution in parasitic wasps, mainly through rapid changes in regulation of protein abundance and/or the emergence of multigenic families by gene duplication. Our results additionally provide invaluable data for further functional venom analyses.

Conflicts of interest

There are no conflicts of interest.

References

[1] H.C.J. Godfray, Parasitoids: behavioural and evolutionary ecology, Princeton University Press, Princeton 1994.

[2] S.J. Moreau, S. Asgari, Venom proteins from parasitoid wasps and their biological functions, Toxins 7 (2015) 2385-2412.

[3] L.S. Corley, M.R. Strand, Evasion of encapsulation by the polyembryonic parasitoid *Copidosoma floridanum* is mediated by a polar body-derived extraembryonic membrane, J. Invertebr. Pathol. 83 (2003) 86-89.

[4] J. Hu, Q. Xu, S. Hu, X. Yu, Z. Liang, W. Zhang, Hemomucin, an O-glycosylated protein on embryos of the wasp *Macrocentrus cingulum* that protects it against encapsulation by hemocytes of the host *Ostrinia furnacalis*, J. Innate Immun. 6 (2014) 663-675.

[5] S. Asgari, Venoms from endoparasitoids, in: N.E. Beckage, J.M. Drezen (Eds.), Parasitoid Viruses, Symbionts and Pathogens, Academic Press, London, UK, 2012, pp. 217-231.

[6] J.l. Gatti, A. Schmitz, D. Colinet, M. Poirié, Diversity of virus-like particles in parasitoids' venom, in: N.E. Beckage, J.M. Drezen (Eds.), Parasitoid Viruses, Symbionts and Pathogens, Academic Press, London, UK, 2012, pp. 181-192.

[7] N.E. Beckage, Polydnaviruses as endocrine regulators, in: N.E. Beckage, J.M. Drezen (Eds.), Parasitoid Viruses, Symbionts and Pathogens, Academic Press, London, UK, 2012, pp. 163-168.

[8] M.R. Strand, Polydnavirus gene products that interact with the host immune system, in: N.E. Beckage, J.M. Drezen (Eds.), Parasitoid Viruses, Symbionts and Pathogens, Academic Press, London, 2012, pp. 149-161.

[9] N.R. Casewell, W. Wuster, F.J. Vonk, R.A. Harrison, B.G. Fry, Complex cocktails: the evolutionary novelty of venoms, Trends Ecol. Evol. 28 (2013) 219-229.

[10] D. Colinet, H. Mathe-Hubert, R. Allemand, J.L. Gatti, M. Poirie, Variability of venom components in immune suppressive parasitoid wasps: from a phylogenetic to a population approach, J. Insect Physiol. 59 (2013) 205-212.

[11] M. Poirié, D. Colinet, J.L. Gatti, Insights into function and evolution of parasitoid wasp venoms, Curr. Opin. Insect Sci. 6 (2014) 52-60.

[12] D. Colinet, E. Deleury, C. Anselme, D. Cazes, J. Poulain, C. Azema-Dossat, M. Belghazi, J.L. Gatti, M. Poirie, Extensive inter- and intraspecific venom variation in closely related parasites targeting the same host: the case

of *Leptopilina* parasitoids of *Drosophila*, Insect Biochem. Mol. Biol. 43 (2013) 601-611.

[13] D. Colinet, C. Anselme, E. Deleury, D. Mancini, J. Poulain, C. Azema-Dossat, M. Belghazi, S. Tares, F. Pennacchio, M. Poirie, J.L. Gatti, Identification of the main venom protein components of *Aphidius ervi*, a parasitoid wasp of the aphid model *Acyrthosiphon pisum*, BMC Genomics 15 (2014) 342.

[14] J.Y. Zhu, Deciphering the main venom components of the ectoparasitic ant-like bethylid wasp, *Scleroderma guani*, Toxicon 113 (2016) 32-40.

[15] F. Pennacchio, M.R. Strand, Evolution of developmental strategies in parasitic hymenoptera, Annu. Rev. Entomol. 51 (2006) 233-258.

[16] Q. Chen, Z. Peng, C. Xu, C. Tang, B. Lu, Q. Jin, H. Wen, F. Wan, Biological assessment of *Tetrastichus brontispae*, a pupal parasitoid of coconut leaf beetle *Brontispa longissima*, Biocontrol Sci. Technol. 20 (2010) 283-295.

[17] B.Z. Tang, L. Xu, Y.M. Hou, Effects of rearing conditions on the parasitism of *Tetrastichus brontispae* on its pupal host *Octodonta nipae*, BioControl 59 (2014) 647-657.

[18] B.Z. Tang, J. Chen, Y.M. Hou, E. Meng, Transcriptome immune analysis of the invasive beetle *Octodonta nipae* (Maulik) (Coleoptera: Chrysomelidae) parasitized by *Tetrastichus brontispae* Ferrière (Hymenoptera: Eulophidae), PLoS ONE 9 (2014) e91482.

[19] E. Meng, B.Z. Tang, Y.M. Hou, X.X. Chen, J.T. Chen, X.Q. Yu, Altered immune function of *Octodonta nipae* (Maulik) to its pupal endoparasitoid, *Tetrastichus brontispae* Ferriere, Comp. Biochem. Physiol. B, Biochem. Mol. Biol. 198 (2016) 100-109.

[20] K. Liu, J. Lin, Y. Fu, Z. Peng, Q. Jin, Effects of parasitization by *Tetrastichus brontispae* (Hymenoptera: Eulophidae) on immunoreaction of the coconut hispine beetle, *Brontispa longissima* (Coleoptera: Chrysomelidae), Acta Entomol. Sin. 51 (2008) 1011-1016.

[21] R.M. Branca, L.M. Orre, H.J. Johansson, V. Granholm, M. Huss, A. Perez-Bercoff, J. Forshed, L. Kall, J. Lehtio, HiRIEF LC-MS enables deep proteome coverage and unbiased proteogenomics, Nat. Methods 11 (2014) 59-62.

[22] S. Tate, B. Larsen, R. Bonner, A.C. Gingras, Label-free quantitative proteomics trends for protein-protein interactions, J. Proteomics 81 (2013) 91-101.

[23] B. Domon, R. Aebersold, Options and considerations when selecting a quantitative proteomics strategy, Nat. Biotechnol. 28 (2010) 710-721.

[24] S. Gallen, E. Duriez, C. Crone, M. Kellmann, T. Moehring, B. Domon, Targeted proteomic quantification on quadrupole-orbitrap mass spectrometer, Mol. Cell Protoemics 11 (2012) 1709-1723.

[25] A.C. Peterson, J.D. Russell, D.J. Bailey, M.S. Westphall, J.J. Coon, Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics, Mol. Cell Protoemics 11 (2012) 1475-1488.

[26] N. Rauniyar, Parallel reaction monitoring: a targeted experiment performed using high resolution and high mass accuracy mass spectrometry, Int. J. Mol. Sci. 16 (2015) 28566-28581.

[27] M.G. Grabherr, B.J. Haas, M. Yassour, J.Z. Levin, D.A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B.W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman, A. Regev, Full-length transcriptome assembly from RNA-Seq data without a reference genome, Nat. Biotechnol. 29 (2011) 644-652.

[28] C. Iseli, C.V. Jongeneel, P. Bucher. ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. Proceedings International Conference on Intelligent Systems for Molecular Biology 1999, pp. 138-148.

[29] M. Punta, P.C. Coggill, R.Y. Eberhardt, J. Mistry, J. Tate, C. Boursnell, N. Pang, K. Forslund, G. Ceric, J. Clements, A. Heger, L. Holm, E.L. Sonnhammer, S.R. Eddy, A. Bateman, R.D. Finn, The Pfam protein families database, Nucleic Acids Res. 40 (2012) D290-301.

[30] J.R. Wisniewski, A. Zougman, N. Nagaraj, M. Mann, Universal sample preparation method for proteome analysis, Nat. Methods 6 (2009) 359-362.

[31] J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification, Nat. Biotechnol. 26 (2008) 1367-1372.

[32] B. Schwanhausser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, M. Selbach, Global quantification of mammalian gene expression control, Nature 473 (2011) 337-342.

[33] F. Jungo, L. Bougueleret, I. Xenarios, S. Poux, The UniProtKB/Swiss-Prot Tox-Prot program: a central hub of integrated venom protein data, Toxicon 60 (2012) 551-557.

[34] B. MacLean, D.M. Tomazela, N. Shulman, M. Chambers, G.L. Finney, B. Frewen, R. Kern, D.L. Tabb, D.C. Liebler, M.J. MacCoss, Skyline: an open source document editor for creating and analyzing targeted proteomics experiments, Bioinformatics 26 (2010) 966-968.

[35] W. Gnatzy, J. Michels, W. Volknandt, S. Goller, S. Schulz, Venom and Dufour's glands of the emerald cockroach wasp *Ampulex compressa* (Insecta, Hymenoptera, Sphecidae): structural and biochemical aspects, Arthropod Struct. Dev. 44 (2015) 491-507.

[36] R. Ferrarese, J. Morales, D. Fimiarz, B.A. Webb, S. Govind, A supracellular system of actin-lined canals controls biogenesis and release of virulence factors in parasitoid venom glands, J. Exp. Biol. 212 (2009) 2261-2268.

[37] C. Noirot, A. Quennedey, Fine structure of insect epidermal glands, Annu. Rev. Entomol. 19 (1974) 61-80.

[38] K.M. Edson, S.B. Vinson, A comparative morphology of the venom apparatus of female braconids (Hymenoptera: Braconidae), Can. Entomol. 111 (1979) 1013-1024.

[39] K.M. Edson, M.R. Barlin, S.B. Vinson, Venom apparatus of braconid wasps comparative ultrastructure of reservoirs and gland filaments, Toxicon 20 (1982) 553-562.

[40] D.L.J. Quicke, Internal and reproductive anatomy, in: D.L.J. Quicke (Eds.) The Braconid and Ichneumonid parasitoid wasps: biology, systematics, evolution and ecology, Wiley Blackwell, Oxford, 2015, pp. 57-70.

[41] J.Y. Zhu, G.Y. Ye, C. Hu, Morphology and ultrastructure of the venom apparatus in the endoparasitic wasp *Pteromalus puparum* (Hymenoptera: Pteromalidae), Micron 39 (2008) 926-933.

[42] S.J.M. Moreau, S. Vinchon, A. Cherqui, G. Prévost, Components of *Asobara* venoms and their effects on hosts, Adv. Parasitol. 70 (2009) 217-232.

[43] J. Van Merle, T. Piek, Morphology of venom apparatus, in: T. Piek (Eds.) Venoms of the Hymenoptera, Academic Press, New York, 1986, pp. 17-44.

[44] J. Goecks, N.T. Mortimer, J.A. Mobley, G.J. Bowersock, J. Taylor, T.A. Schlenke, Integrative approach reveals composition of endoparasitoid wasp venoms, PLoS One 8 (2013) e64125.

[45] G. Zhang, O. Schmidt, S. Asgari, A calreticulin-like protein from endoparasitoid venom fluid is involved in host hemocyte inactivation, Dev. Comp. Immunol. 30 (2006) 756-764.

[46] D.B. Rivers, M.P. Dani, E.H. Richards, The mode of action of venom from the endoparasitic wasp *Pimpla hypochondriaca* (Hymenoptera: Ichneumonidae) involves Ca^{+2}-dependent cell death pathways, Arch. Insect Biochem. Physiol. 71 (2009) 173-190.

[47] D. Colinet, A. Dubuffet, D. Cazes, S. Moreau, J.M. Drezen, M. Poirie, A serpin from the parasitoid wasp *Leptopilina boulardi* targets the *Drosophila* phenoloxidase cascade, Dev. Comp. Immunol. 33 (2009) 681-689.

[48] N.A. Ratcliffe, P.E. King, Morphological, ultrastructural, histochemical and electrophoretic studies on the venom system of *Nasonia vitripennis* Walker (Hymenoptera: Pteromalidae), J. Morphol. 127 (1969) 177-204.

[49] R.W. Howard, J.E. Baker, Morphology and chemistry of Dufour glands in four ectoparasitoids: *Cephalonomia tarsalis, C. waterstoni* (Hymenoptera: Bethylidae), *Anisopteromalus calandrae*, and *Pteromalus cerealellae* (Hymenoptera: Pteromalidae), Comp. Biochem. Physiol. B: Biochem. Mol. Biol. 135 (2003) 153-167.

[50] D. Colinet, D. Cazes, M. Belghazi, J.L. Gatti, M. Poirie, Extracellular superoxide dismutase in insects: characterization, function, and interspecific variation in parasitoid wasp venom, J. Biol. Chem. 286 (2011) 40110-40121.

[51] D. Colinet, L. Kremmer, S. Lemauf, C. Rebuf, J.L. Gatti, M. Poirie, Development of RNAi in a *Drosophila* endoparasitoid wasp and demonstration of its efficiency in impairing venom protein production, J. Insect Physiol. 63 (2014) 56-61.

[52] A.L. Siebert, D. Wheeler, J.H. Werren, A new approach for investigating venom function applied to venom calreticulin in a parasitoid wasp, Toxicon 107 (2015) 304-316.

[53] B. Vincent, M. Kaeslin, T. Roth, M. Heller, J. Poulain, F. Cousserans, J. Schaller, M. Poirie, B. Lanzrein, J.M. Drezen, S.J. Moreau, The venom composition of the parasitic wasp *Chelonus inanitus* resolved by combined expressed sequence tags analysis and proteomic approach, BMC Genomics 11 (2010) 693.

[54] S. Asgari, D.B. Rivers, Venom proteins from endoparasitoid wasps and their role in host-parasite interactions, Annu. Rev. Entomol. 56 (2011) 313-335.

[55] D. Colinet, A. Schmitz, D. Cazes, J.L. Gatti, M. Poirie, The origin of intraspecific variation of virulence in an eukaryotic immune suppressive parasite, PLoS Pathog. 6 (2010) e1001206.

[56] A. Dubuffet, D. Colinet, C. Anselme, S. Dupas, Y. Carton, M. Poirié, Variation of *Leptopilina boulardi* success in *Drosophila* hosts, Adv. Parasitol. 70 (2009) 147-188.

[57] S. Dupas, M. Boscaro, Geographic variation and evolution of immunosuppressive genes in a *Drosophila* parasitoid, Ecography 22 (1999) 284-291.

[58] J. Russo, M. Brehelin, Y. Carton, Haemocyte changes in resistant and susceptible strains of *D. melanogaster* caused by virulent and avirulent strains of the parasitic wasp *Leptopilina boulardi*, J. Insect Physiol. 47 (2001) 167-172.

[59] A. Dubuffet, S. Dupas, F. Frey, J.M. Drezen, M. Poirie, Y. Carton, Genetic interactions between the parasitoid wasp Leptopilina boulardi and its *Drosophila* hosts, Heredity (Edinb) 98 (2007) 21-27.

[60] S. Dupas, Y. Carton, M. Poirie, Genetic dimension of the coevolution of virulence-resistance in *Drosophila*-parasitoid wasp relationships, Heredity (Edinb) 90 (2003) 84-89.

[61] A.J. Turner, R.E. Isaac, D. Coaates, The neprilysin (NEP) family of zinc metalloendopeptidases: genomics and function, Bioessays 23 (2001) 261-269.

[62] C. Antczak, I. De Meester, B. Bauvois, Ectopeptidases in pathophysiology, Bioessays 23 (2001) 251-260.

[63] J.E. Thomas, C.M. Rylett, A. Carhan, N.D. Bland, R.J. Bingham, A.D. Shirras, A.J. Turner, R.E. Isaac, *Drosophila melanogaster* nep2 is a new soluble member of the neprilysin family of endopepdidases with implications for reproduction and renal function, Biochem. J. 386 (2005) 357-366.

[64] N.D. Bland, J.W. Pinney, J.E. Thomas, A.J. Turner, R.E. Isaac, Bioinformatic analysis of the neprilysin (M13) family of peptidases reveals complex evolutionary and functional relationships, BMC Evol. Biol. 8 (2008) 16.

[65] R.E. Isaac, N.D. Bland, A.D. Shirras, Neuropeptidases and the metabolic inactivation of insect neuropeptides, Gen. Comp. Endocrinol. 162 (2009) 8-17.

[66] A.M. Crawford, R. Brauning, G. Smolenski, C. Ferguson, D. Barton, T.T. Wheeler, A. Mcculloch, The constituents of Microctonus sp. parasitoid venoms, Insect Mol. Biol. 17 (2008) 313-324.

[67] S. Asgari, A. Reineke, M. Beck, O. Schmidt, Isolation and characterization of a neprilysin-like protein from *Venturia canescens* virus-like particles, Insect Mol. Biol. 11 (2002) 477-485.

[68] G. Allegri, C.V.L. Costa, A. Bertazzo, M. Biasiolo, E. Ragazzi, Enzyme activities of tryptophan metabolism along the kynurenine pathway in various species of animals, Il Farmaco 58 (2003) 829-836.

[69] J. Fang, Q. Han, J. Li, Isolation, characterization, and functional expression of kynurenine aminotransferase cDNA from the yellow fever mosquito, *Aedes aegypti*, Insect Biochem. Mol. Biol. 32 (2002) 943-950.

[70] A. Cerstiaens, J. Huybrechts, S. Kotanen, I. Lebeau, K. Meylaers, A. De Loof, L. Schoofs, Neurotoxic and neurobehavioral effects of kynurenines in adult insects, Biochem. Biophys. Res. Commun. 312 (2003) 1171-1177.

[71] S.J. Chiou, S. Kotanen, A. Cerstiaens, D. Daloze, J.M. Pasteels, A. Lesage, J.W. Drijfhout, V. P., L. Dillen, M. Claeys, H. De Meulemeester, B. Nuttin, A. De Loof, L. Schoofs, Purification of toxic compounds from larvae of

the gray fleshfly: the identification of paralysins, Biochem. Biophys. Res. Commun. 246 (1998) 457-462.

[72] J. Li, B.T. Beerntsen, A.A. James, Oxidation of 3-hydroxykynurenine to produce xanthommatin for eye pigmentation: a major branch pathway of tryptophan catabolism during pupal development in the yellow fever mosquito, *Aedes aegypti*, Insect Biochem. Mol. Biol. 29 (1999) 329-338.

[73] E. Savvateeva, A. Popov, N. Kamyshev, J. Bragina, M. Heisenberg, D. Senitz, J. Kornhuber, P. Riederer, Age-dependent memory loss, synaptic pathology and altered brain plasticity in the *Drosophila* mutant cardinal accumulating 3-hydroxykynurenin, J. Neural. Transm. 107 (2000) 581-601.

[74] V.B. Smirnov, E.G. Chesknokova, N.G. Lopatina, E. Voike, Electrophysiological characteristics of a population of mushroom body neurons in *Apis mellifera* honeybees in kynurenine deficiency, Neurosci. Behav. Physiol. 37 (2007) 799-802.

[75] B. Hamberger, K. Hahlbrock, The 4-coumarate:CoA ligase gene family in *Arabidopsis thaliana* comprises one rare, sinapate-activating and three commonly occurring isoenzymes, Proc. Natl. Acad. Sci. 101 (2004) 2209-2214.

[76] F. Ververidis, E. Trantas, C. Douglas, G. Vollmer, G. Kretzschmar, N. Panopoulos, Biotechnology of flavonoids and other phenylpropanoid-derived natural products. Part I: Chemical diversity, impacts on plant biology and human health, Biotechnol. J. 2 (2007) 1214-1234.

[77] E.O. Martinson, V.G. Martinson, R. Edwards, Mrinalini, J.H. Werren, Laterally transferred gene recruited as a venom in parasitoid wasps, Mol. Biol. Evol. 33 (2016) 1042-1052.

[78] Z. Blach-Olszewska, B. Jatczak, A. Rak, M. Lorenc, B. Gulanowski, A. Drobna, E. Lamer-Zarawska, Production of cytokines and stimulation of resistance to viral infection in human leukocytes by *Scutellaria baicalensis* flavones, J. Interferon Cytokine Res. 28 (2008) 571-581.

[79] D.C. de Graaf, M. Aerts, M. Brunain, C.A. Desjardins, F.J. Jacobs, J.H. Werren, B. Devreese, Insights into the venom composition of the ectoparasitoid wasp *Nasonia vitripennis* from bioinformatic and proteomic studies, Insect Mol. Biol. 19 (2010) 11-26.

[80] X. Zhang, B.Z. Tang, Y.M. Hou, A rapid diagnostic technique to discriminate between two pests of palms, *Brontispa longissima* and *Octodonta nipae* (coleoptera: chrysomelidae), for quarantine applications, J. Econ. Entomol. 108 (2015) 95-99.

[81] Y. Nakahara, T. Hiraoka, K. Iwabuch, Effects of lipophorin and 20-hydroxyecdysone on in vitro development of the larval endoparasitoid *Venturia canescens*, J. Insect Physiol. 45 (1999) 453-460.

[82] D.B. Rivers, D.L. Denlinger, Venom-induced alterations in fly lipid metabolism and its impact on larval development of the ectoparasitoid *Nasonia vitripennis* (Walker) (Hymenoptera: Pteromalidae), J. Invertebr. Pathol. 66 (1995) 104-110.

[83] B. Visser, C. Le Lann, F.J. den Blanken, J.A. Harvey, J.J. van Alphen, J. Ellers, Loss of lipid synthesis as an evolutionary consequence of a parasitic lifestyle, Proc. Natl. Acad. Sci. 107 (2010) 8677-8682.

[84] T. Doremus, S. Urbach, V. Jouan, F. Cousserans, M. Ravallec, E. Demettre, E. Wajnberg, J. Poulain, C. Azema-Dossat, I. Darboux, J.M. Escoubas, D. Colinet, J.L. Gatti, M. Poirie, A.N. Volkoff, Venom gland extract is not required for successful parasitism in the polydnavirus-associated endoparasitoid *Hyposoter didymator* (Hym. Ichneumonidae) despite the presence of numerous novel and conserved venom proteins, Insect Biochem. Mol. Biol. 43 (2013) 292-307.

[85] G.R. Burke, M.R. Strand, Systematic analysis of a wasp parasitism arsenal, Mol. Ecol. 23 (2014) 890-901.

[86] M.R. Kanost, Serine proteinase inhibitors in arthropod immunity, Dev. Comp. Immunol. 23 (1999) 291-301

[87] S. Asgari, G. Zhang, R. Zareie, O. Schmidt, A serine proteinase homolog venom protein from an endoparasitoid wasp inhibits melanization of the host hemolymph, Insect Biochem. Mol. Biol. 33 (2003) 1017-1024.

[88] X.Q. Yu, H.B. Jiang, Y. Wang, M.R. Kanost, Nonproteolytic serine proteinase homologs are involved in prophenoloxidase activation in the tobacco hornworm, *Manduca sexta*, Insect Biochem. Mol. Biol. 33 (2003) 197-208.

[89] O. Hobert, Gene Regulation by transcription factors and microRNAs, Science 319 (2008) 1785-1786.

[90] E.S. Wong, K. Belov, Venom evolution through gene duplications, Gene 496 (2012) 1-7.

Figure Legends

FIG. 1. **Structure of *T. brontispae* venom apparatus.** *A*, Overview of the venom apparatus of one-day-old female *T. brontispae*: venom gland (VG), venom reservoir (VR), Dufour's gland (DG) and ovipositor (OP) using differential interference contrast (DIC) microscopy. *B*, *C*, *D*, *E*, Ultrastructure of venom gland observed by transmission electronic microscopy (TEM): *B* cross-section of venom gland showing a central lumen lined with an intimal layer, the narrow tightly packed duct cells and the peripheral concentrically arranged secretory cells with large nuclei; *C* details of a secretory cell, which contains numerous cellular organelles, such as extensive rough endoplasmic reticulum and vacuoles, and are typical of end-apparatus surrounded by well-developed microvilli; *D* details of the end apparatus bordered by abundant microvilli, into which secreted products are discharged, and around which mitochondria and secretory vesicles are spread; *E* details of nucleus of the duct cell, smooth canals lined with cuticle, and lumen filled with the discharged venom aggregates and bordered by cells of the intimal layer; *F*, TEM view of the venom reservoir showing the muscle fibres layer of uniform thickness, the nuclei of epithelial cell, the cuticle-lined intima lining the reservoir wall, and putative venom aggregates in the reservoir lumen.

EA, end apparatus; Ec, epithelial cell; ER, extensive rough endoplasmic reticulum; Hsp, hemocoel space; I, intima; IL, intimal layer; LuVr, lumen of venom reservoir; M, mitochondria; Mf, muscle fibre; MV, microvilli; NDC, nucleus of the duct cell; NSC, nucleus of the secretory cell; Nu, nucleus; SC, smooth canals lined with cuticle; SV, secretory vesicles; V, vacuoles; VA, venom aggregates.

FIG. 2. **SDS-PAGE profiles of the venom apparatus extracts from two strains of *T. brontispae*, Tb-On and Tb-Bl.** Equalized venom apparatus extracts of Tb-On and Tb-Bl were separated on a 12.5% SDS-PAGE and stained with Coomassie Brilliant blue R-250. Molecular mass markers (M) are indicated to the left of the gel in kilodaltons. One

example of quantitative variation is shown by red arrows.

FIG. 3. **Main distribution of identified proteins from *T. brontispae* venom apparatus by their functional classification.** Venom proteins (ranging from 3 to 450 kDa, although more skewed towards high molecular weight proteins) are classified according to the molecular functional categories at (*A*) level 2 and (*B*) level 3 of the gene ontology system (GO) using Blast2GO (version 2.7.1).

FIG. 4. **Main biological pathways of identified proteins from *T. brontispae* venom apparatus.** Venom apparatus proteins were blasted against KEGG GENES (insects) and analyzed by KEGG Automatic Annotation Server (KAAS).

FIG. 5. **Relative quantitation of venom apparatus proteins from the two stains of *T. brontispae*, Tb-On and Tb-Bl.** Ratios for the 1081 venom apparatus proteins having iBAQ values at least in two of three replicates are shown plotted by $\log_2$(Tb-On/Tb-Bl) and -$\log_{10}$(significance *A*) < 0.05 (*significance A* is a corrected *p* value calculated by the MaxQuant software using a two-tailed Student's *t* test). The shaded area of the graph indicates venom apparatus proteins displaying two-fold differences and significance *A* < 0.05. Fifteen venom apparatus proteins selected for qRT-PCR validation are marked on the plot, including eight abundant proteins presenting in similar amounts between the strains (blue square), four proteins in more abundance (red square) and three proteins in less abundance (green square) in the Tb-On strain.

FIG. 6. **Unsupervised hierarchical clustering of the differentially produced venom apparatus proteins (fold change > 2 and significance *A* < 0.05) between two strains of *T. brontispae*, Tb-On and Tb-Bl.** The columns

represent replicates of each strain, and the rows represent the individual venom apparatus proteins. The up- or down-regulated venom apparatus proteins are indicated by red and green color code, respectively. The color intensity changes with the protein abundance as commented on the key bar on the top right. Cluster divided into several subclusters was marked with A or B (B1, B2, B3 and B4). The significantly enriched venom apparatus proteins performed by KEGG Pathway and/or GO enrichment analysis are marked in red in the rows.

FIG. 7. **GO and KEGG pathway enrichment of the significantly differentially produced venom apparatus proteins between two strains of *T. brontispae*, Tb-On and Tb-Bl.** Enrichment analyses were conducted by Fisher's exact test, and only the *p* values < 0.01 and 0.05 were considered as statistically significantly enriched GO terms and pathways, respectively. Rich factor is the ratio of the number of differentially produced venom apparatus proteins mapped to a certain GO term or pathway to the total number of venom apparatus proteins mapped to this GO term or pathway. Higher rich factor indicates greater intensiveness. Gene number represents the number of venom apparatus proteins mapped to a certain GO term or KEGG pathway.

FIG. 8. **Validation of the selected venom apparatus proteins at the level of mRNA and protein by quantitative real time PCR and western blot analysis, respectively, between the two strains of *T. brontispae*, Tb-On and Tb-Bl.** *A*, Relative expression levels of highly produced venom apparatus proteins and a lowly produced venom apparatus protein (unigene0019736, succinyl-CoA ligase) as a control validating the proteomics results. Arabic number (one to six) in front of the protein names in X-axis represents six out of the top eight abundant venom apparatus proteins. *B*, qRT-PCR validation of differentially produced venom apparatus proteins between the two strains. The mRNA expression was normalized with the reference gene glyceraldehyde-3-phosphate dehydrogenase (GAPDH). Error bars

indicate standard deviations of the mean from three independent biological replications. "*" and "**" denotes the significant differential expression levels between strains at α=0.05 and α=0.01 level, respectively. *C*, Western blot analysis of two selected proteins, calreticulin (unigene0018080) and serpin 3/4 (unigene0022364). Equalized total venom apparatus extracts of Tb-On and Tb-Bl (the same samples as those in label-free profiles) were analyzed by immunoblot using anti-Calreticulin and anti-Serpin 3/4 antibodies to monitor differences in protein levels. Two non-specific bands (NSB1 and NSB2) are shown as loading control.

Protein types for the selected unigenes are listed as follows: Unigene0019736, succinyl-CoA ligase; Unigene0020620, vitellogenin; Unigene0022824, serine/threonine-protein phosphatase; Unigene0022142, cytochrome P450; Unigene0010415, unknown protein; Unigene0020436, serpin 5 precursor; Unigene0020437, serpin 5 precursor; Unigene0022364, serpin 3/4; Unigene0018080, calreticulin.

FIG. 9. **Confirmation of differentially produced proteins detected in label-free analysis using PRM assay.** Fold change denotes the ratio of protein level in the Tb-On strain to that in the Tb-Bl strain. For unigene0021082, its protein abundance in the Tb-Bl strain was not detected in label-free analysis, thus no fold change is displayed.

FIG. 10. **A section of sequence alignment of (*A*) neprilysin (NEP) and (*B*) 4-coumarate CoA ligase 4 (4CL4).** *A*, Conserved residues including the $HE_{XX}H$ zinc binding motif and the catalytically important GENIAD and VNAFY motifs are highlighted with Box (*Drosophila melanogaster* NEP2, FlyBase number CG5894; *Nasonia vitripennis* NEP, Genbank accession number XP_001600059). *B*, Box I and II represent conserved motifs. Residues involved in hydroxycinnamate binding and enzymatic function are indicated by circles and triangles, respectively. Accession numbers of the sequences are listed in Supplemental Fig. S4.

Table 1 Biological characteristics of venom apparatus of *T. brontispae* at different developmental stages.

Biological parameters of venom apparatus	**Developmental stages**			***F***	***P***
	1 d[a]	**3 d**	**5 d**		
Area[b] of venom gland（μm²）	14794.283 ± 536.270	12338.937± 024.710	12961.999± 343.184	0.499	0.609
Area of venom reservoir（μm²）	12148.089 ± 671.082	13202.209 ± 816.443	15207.684 ± 1995.160	1.647	0.200
Length of Dufour's gland (μm）	198.936 ± 15.073	200.448 ± 13.129	185.599 ± 25.422	0.196	0.822
Φ[c] of nucleus of secretory cell (μm)	60.824 ± 14.940	59.821 ± 8.018	44.4175 ± 12.97995	0.230	0.797
Φ of secretory vesicles (μm）	0.524 ± 0.086	0.412 ± 0.031	0.3171 ± 0.016	2.737	0.082
Φ range of secretory vesicles (μm)	0.130-2.039	0.107-1.406	0.139-0.696	--	--

Data are expressed as mean ± SE. Significant difference was analyzed by one-way ANOVA and Tukey's (homogeneity of variances) or Tamhane's T2 test (heterogenicity of variances).

[a] days after eclosion of *T. brontispae*.

[b] length * diameter.

[c] *Φ*.

Table 2 General features of the pooled *T. brontispae* transcriptome and results of similarity searches.

Sequencing Parameters	
Total reads	43,023,274
Total nucleotides (bp)	5,377,909,250
Q20 percentage (%)[a]	96.12
N percentage (%)[b]	0.00
Number of unigenes	23,054
GC percentage (%)	43.95
N50 of unigene set (bp)[c]	1791
Max length (bp)	13,739
Min length (bp)	201
Mean length of unigenes (bp)	1030
Similarity Searches	
With public databases	
NCBI NR	14473
Swiss-Prot	10847
COG	5170
KEGG	6100
With insect proteomes	
Nasonia vitripennis	8818
Camponotus floridanus	741
Harpegnathos saltator	695
Acromyrmex echinatior	622
Megachile rotundata	403
Tribolium castaneum	281
Bombus impatiens	277
Apis florea	235
Apis mellifera	234
Bombus terrestris	233
Translation	
Unisequences with CDS prediction	14,503
Unisequences with ESTScan prediction	1565
Unisequences with Pfam A annotation	9887

[a] Percentage of nucleotide error rate under 0.01.

[b] Uncertain base in the output sequencing data.

[c] Median length of all unigenes.

Table 3 The eight most abundant venom apparatus proteins shared by the two strains of *T. brontispae*, Tb-On and Tb-Bl.

Unigene ID	Sequence description	Assembled length (bp)	Average iBAQ		log2(Tb-On/Tb-Bl)	*Significance A*
			Tb-On	Tb-Bl		
Unigene0009814	venom protein r-like protein	641	28957333333	28600000000	0.0179	0.459
Unigene0004953	--[a]	771	27727000000	28134000000	-0.0210	0.402
Unigene0021495	--	267	13269333333	10031700000	0.404	0.817
Unigene0000904	neprilysin-2-like[b]	2377	13247666667	12629333333	0.0690	0.540
Unigene0000192	--	203	13215666667	15739000000	-0.252	0.156
Unigene0004125	kynurenine--oxoglutarate transaminase 3-like	1633	12709666667	12434333333	0.0316	0.480
Unigene0018515	4-coumarate--CoA ligase-like	1850	10968133333	10662400000	0.0408	0.494
Unigene0000125	--	1281	10389133333	11418533333	-0.136	0.260

Intensity-based absolute quantification (iBAQ) calculated by MaxQuant was adopted for protein abundance. Differential expression in protein abundance between the two strains was analyzed using a two-tailed Student's *t* test and a *Significance A* value (a corrected *p* value calculated by the MaxQuant software).

[a] No similarity with other known proteins in generalist databases.

[b] Another neprilysin-2-like protein (unigene0018250) was identified only in Tb-On strain.

Fig. 1

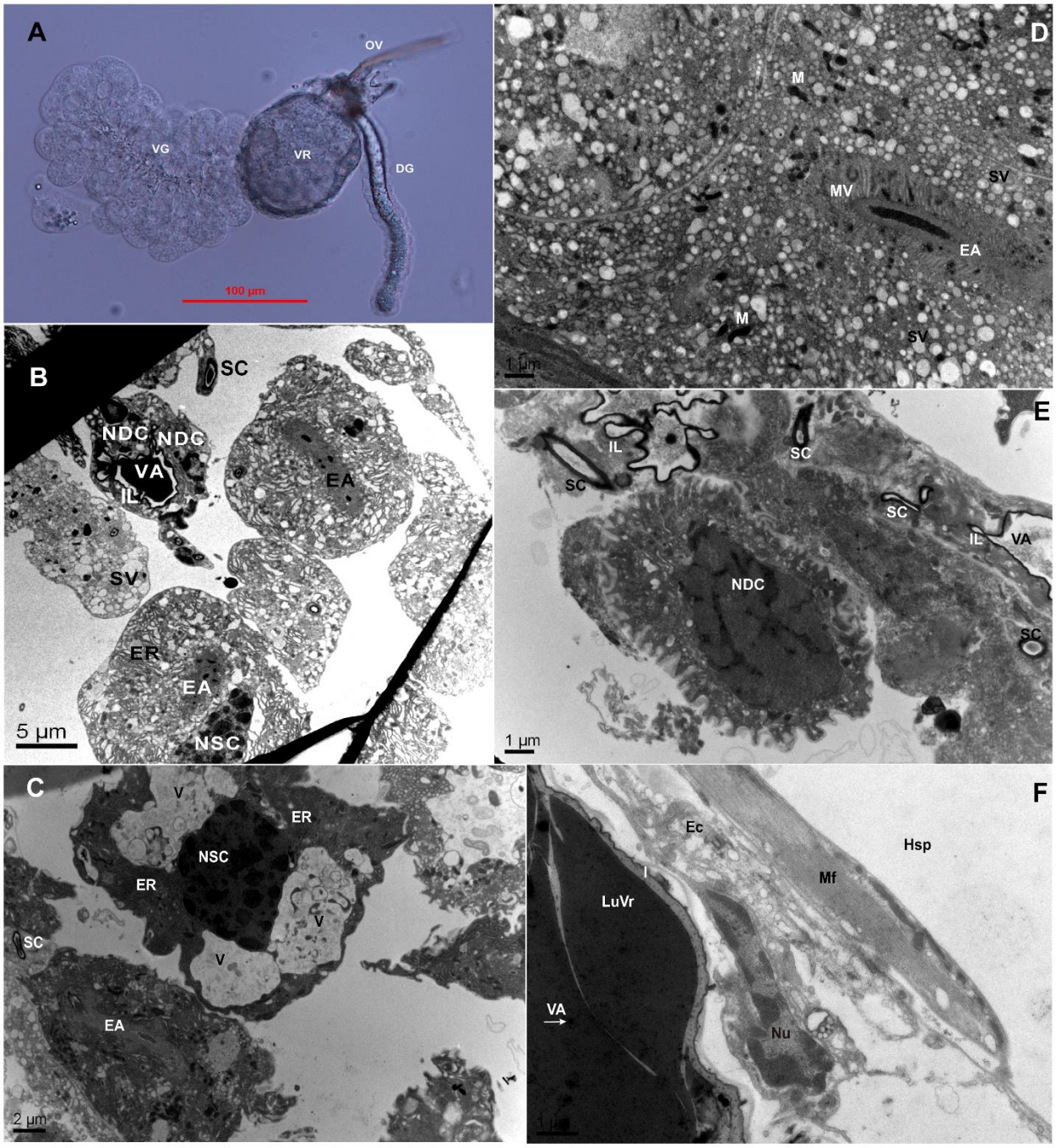


Fig. 2

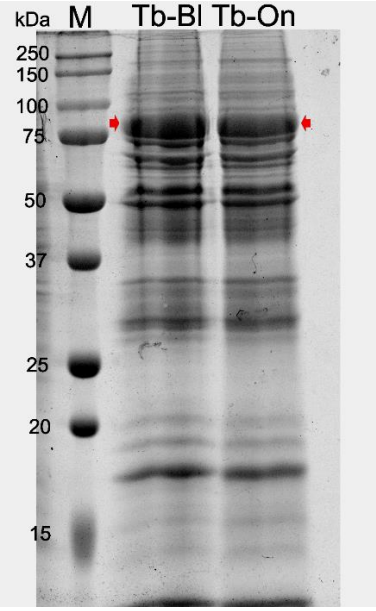


Fig. 3

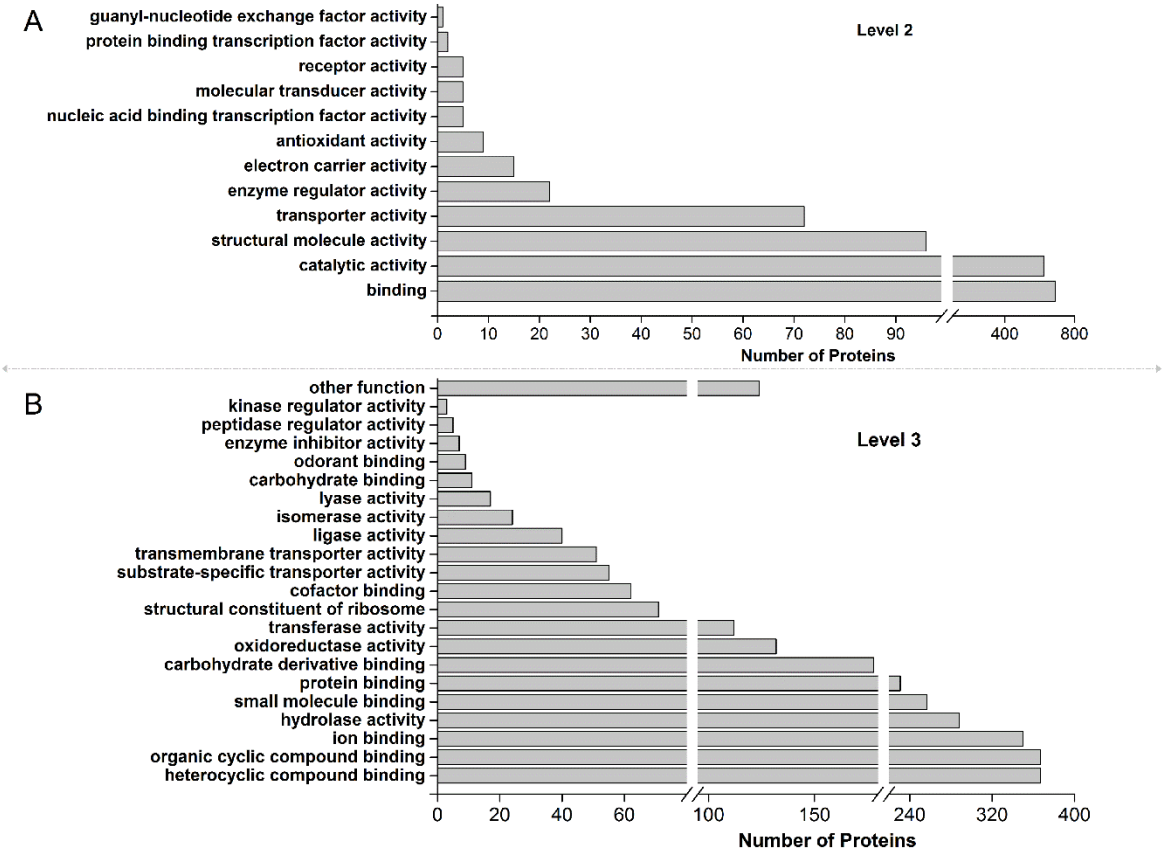


Fig. 4

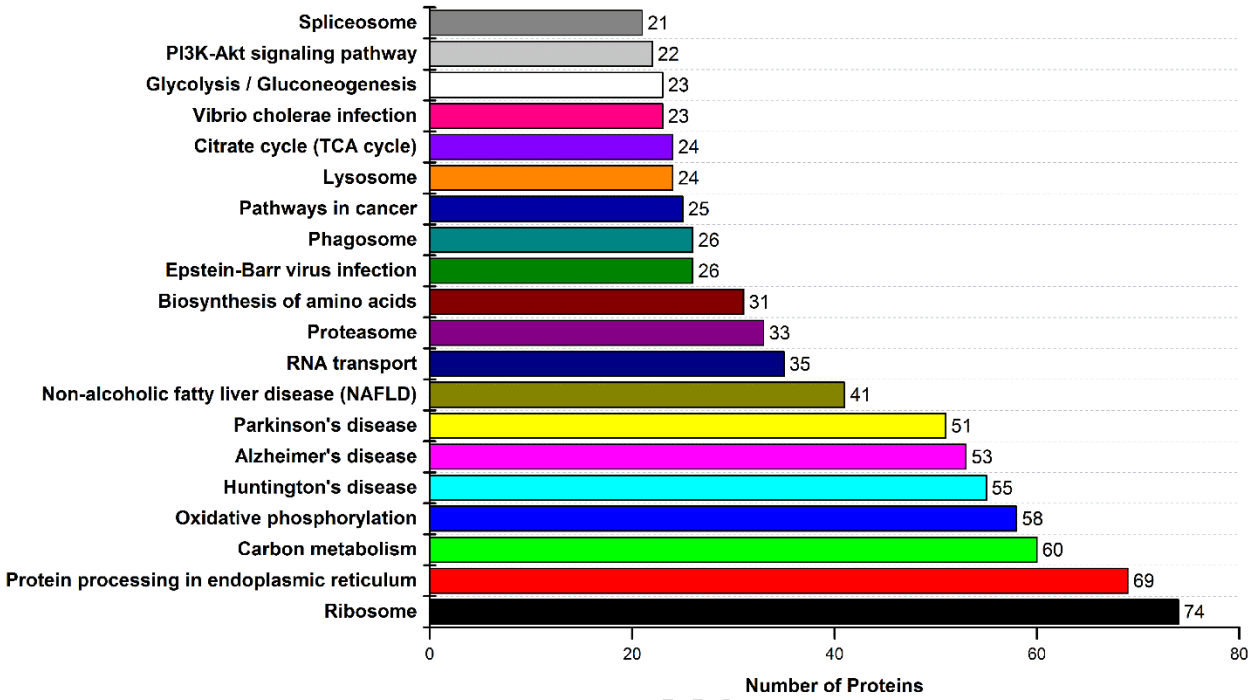


Fig. 5

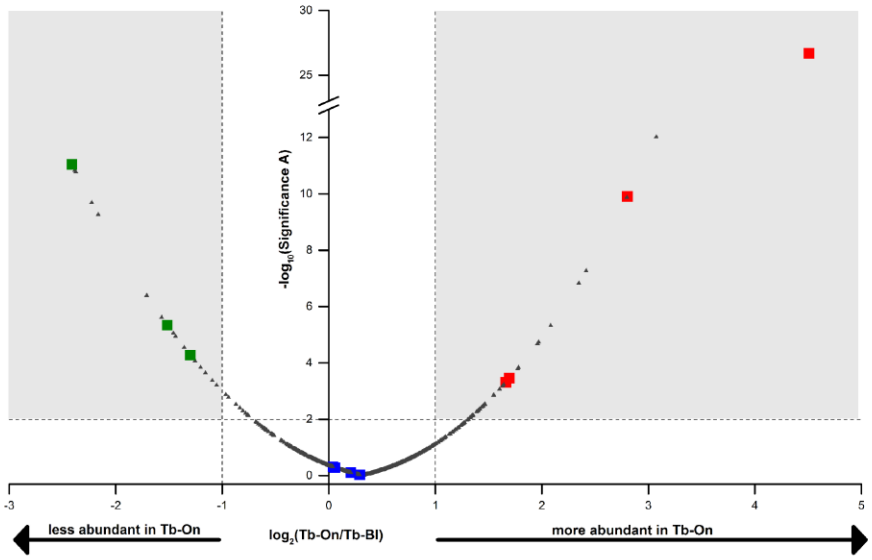


Fig. 6

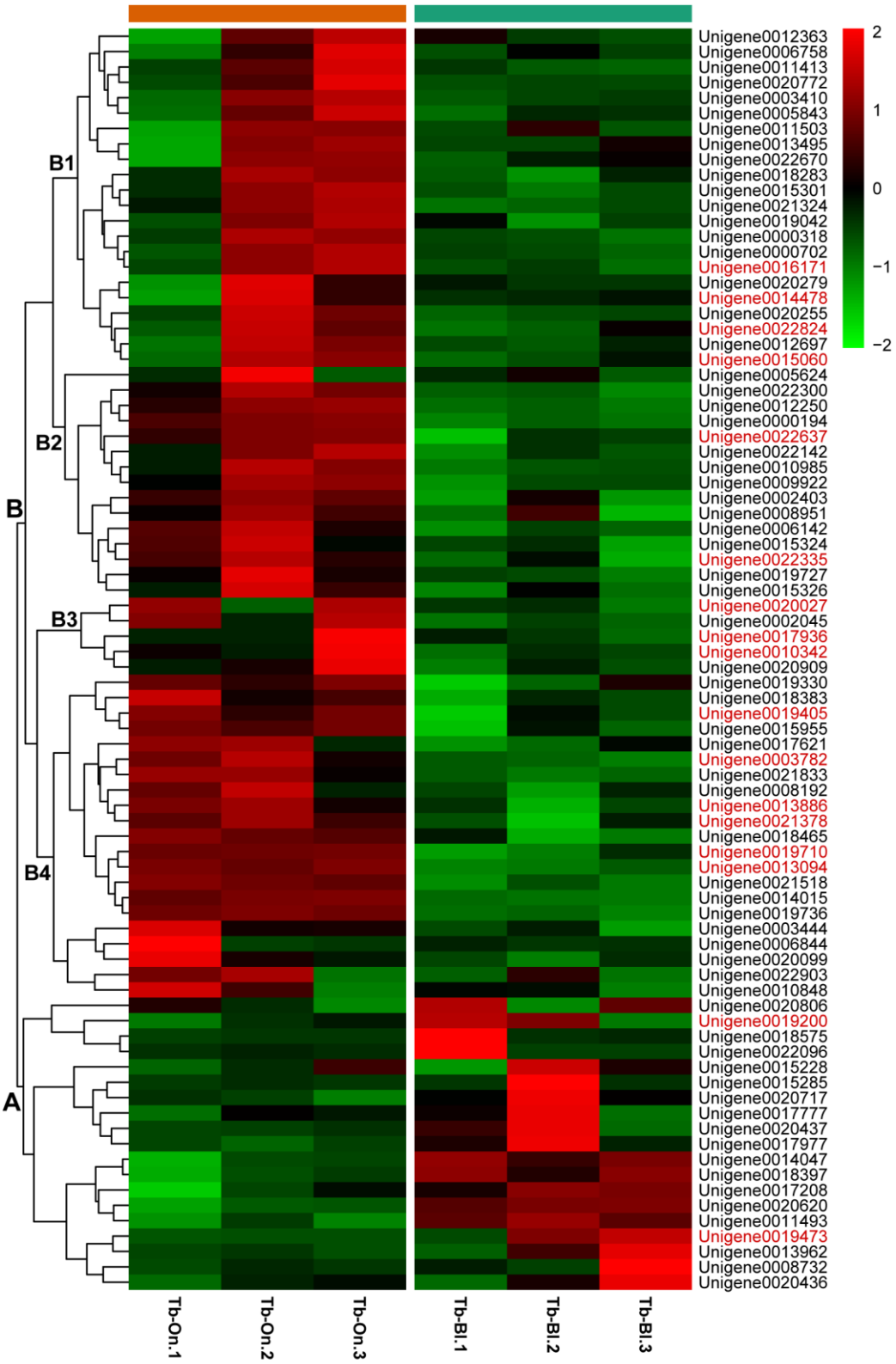


Fig. 7

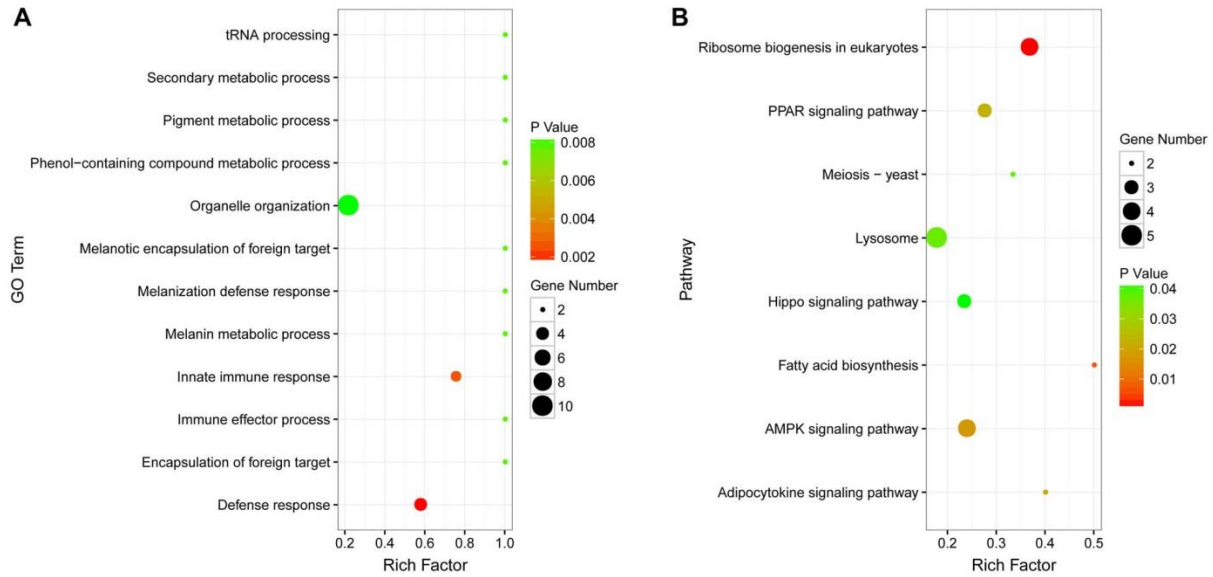


Fig. 8

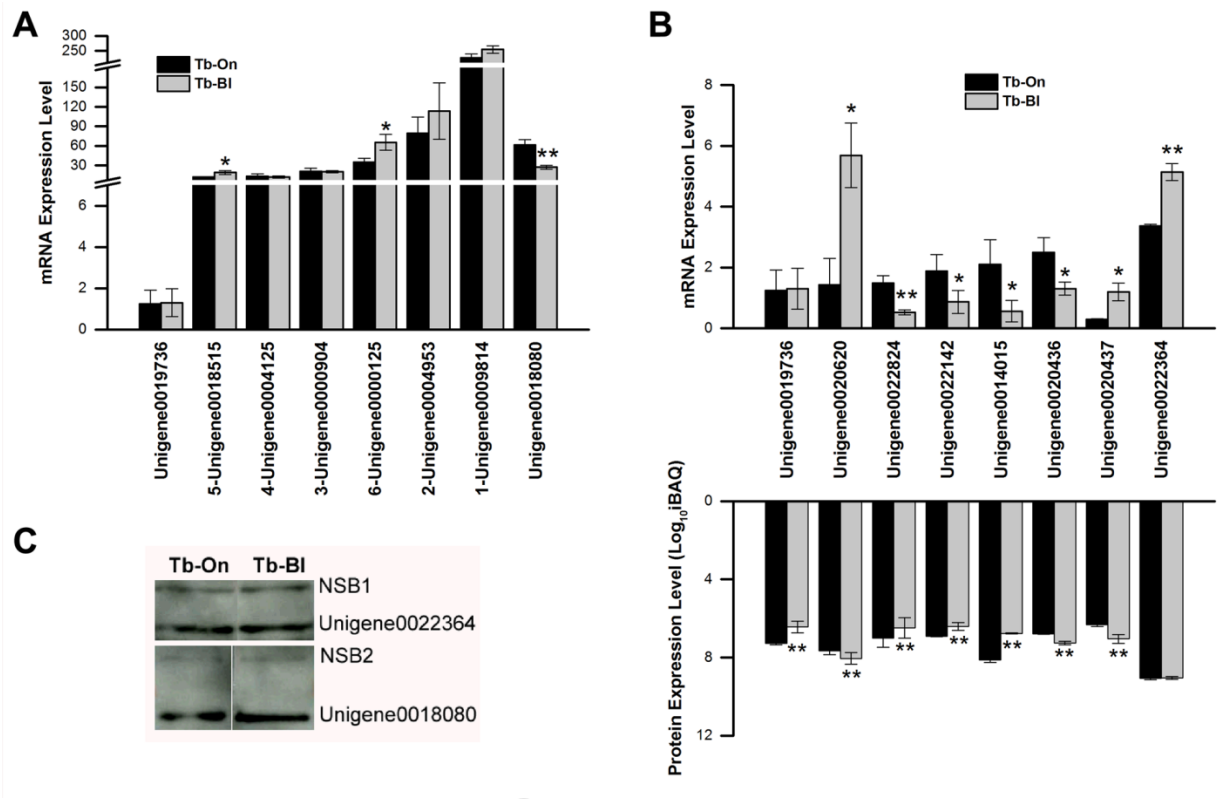


Fig. 9

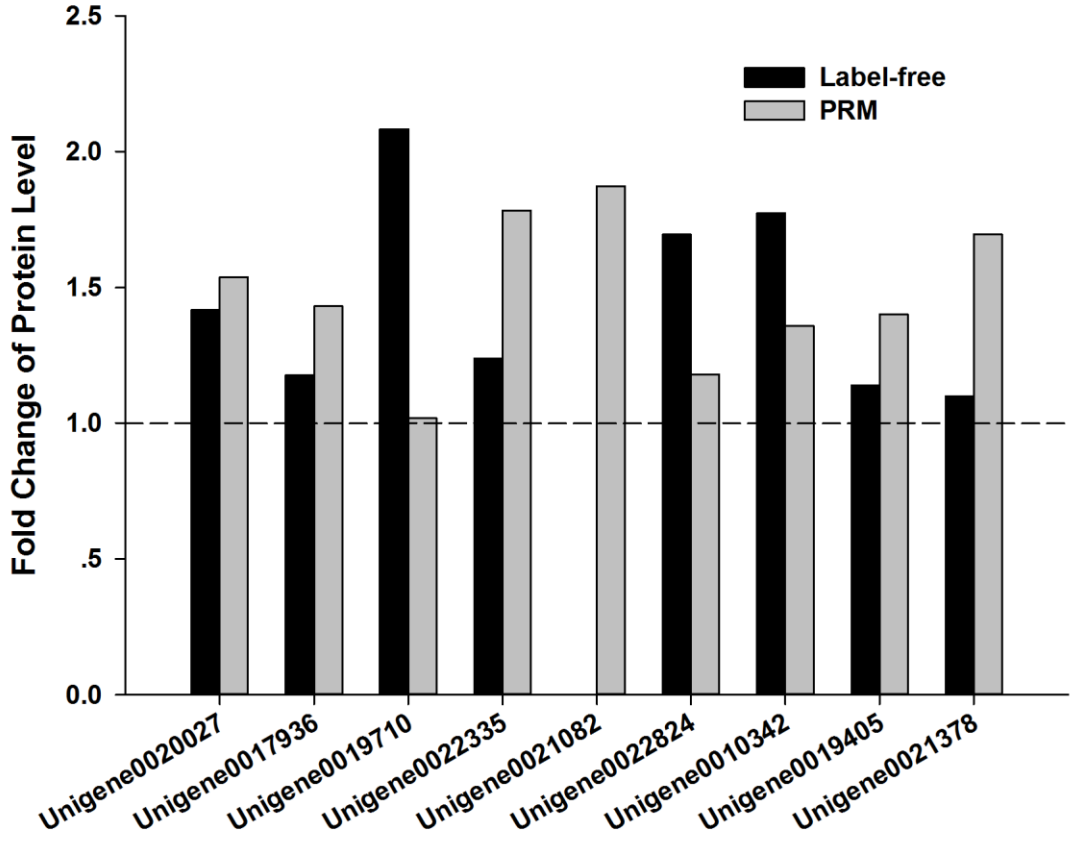
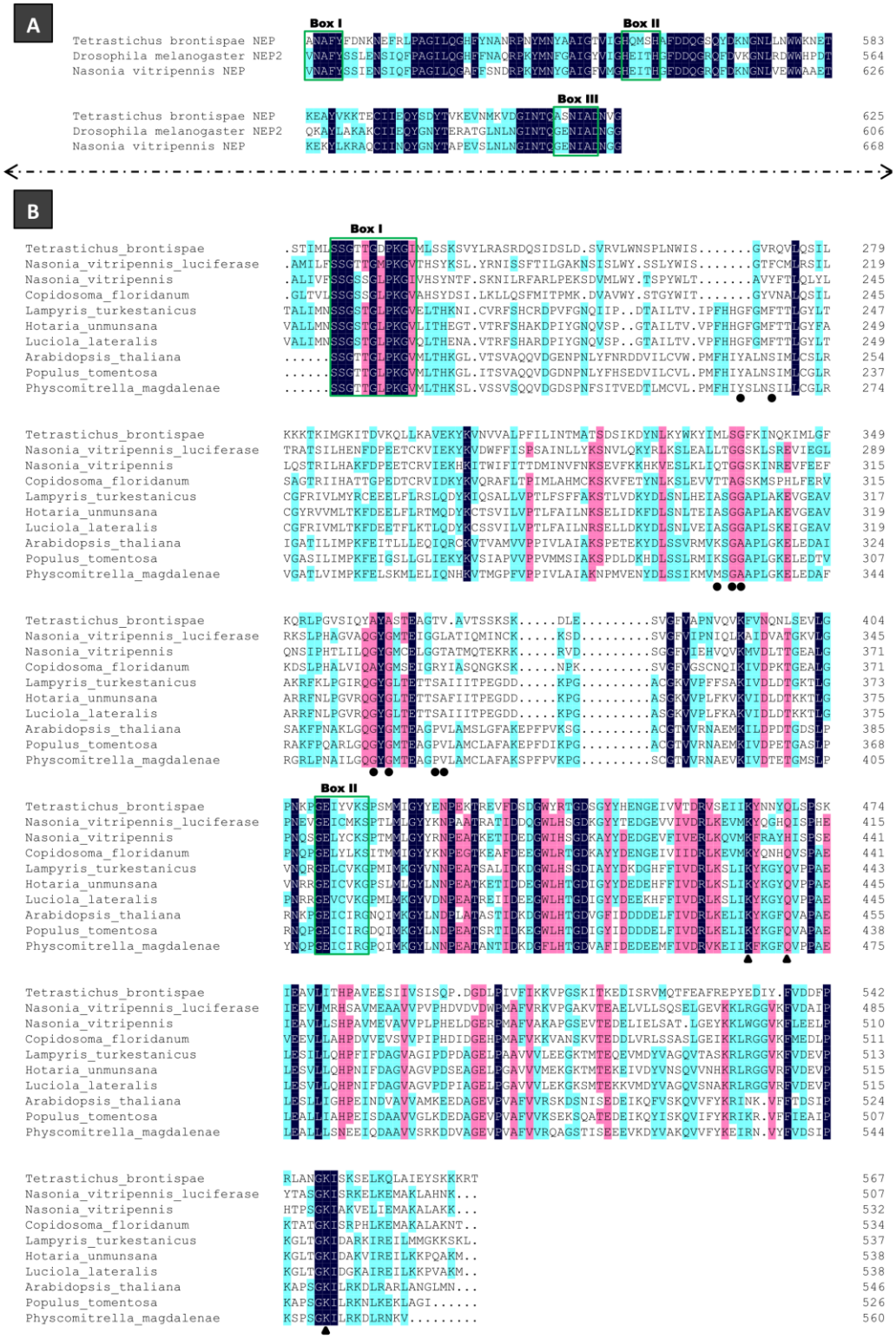


Fig. 10



Highlights:

- Venom apparatus of *Tetrastichus brontispae* contains novel venom proteins, such as 4-coumarate CoA ligase 4.
- Venom apparatus proteins vary between the two strains of *T. brontispae*.
- Venom variation is mainly enriched in fatty acid biosynthesis and melanotic encapsulation response.
- Venom composition can evolve quickly and respond to host selection.