

Expression profiles of venom components in some social hymenopteran species over different post-capture periods

Kyungjae Andrew Yoon, Woo-Jin Kim, Si Hyeock Lee



PII: S1532-0456(21)00274-X

DOI: <https://doi.org/10.1016/j.cbpc.2021.109247>

Reference: CBC 109247

To appear in: *Comparative Biochemistry and Physiology, Part C*

Received date: 14 October 2021

Revised date: 14 November 2021

Accepted date: 18 November 2021

Please cite this article as: K.A. Yoon, W.-J. Kim and S.H. Lee, Expression profiles of venom components in some social hymenopteran species over different post-capture periods, *Comparative Biochemistry and Physiology, Part C* (2021), <https://doi.org/10.1016/j.cbpc.2021.109247>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Expression profiles of venom components in some social hymenopteran species over different post-capture periods

Kyungjae Andrew Yoon^{a,*}, Woo-Jin Kim^b and Si Hyeock Lee^{a,c,*}

^aResearch Institute of Agriculture and Life Sciences, Seoul National University, Seoul 08826, Korea

^bEntoCode Co., Seoul 06028, Korea

^cDepartment of Agricultural Biology, Seoul National University, Seoul 08826, Korea

*Authors to whom correspondence should be addressed:

E-mail: kongbob89@snu.ac.kr (KAY), shlee22@snu.ac.kr (SHL)

Tel.: +82-2-880-4694 (KAY); +82-2-880-4704 (SHL)

Abstract

To explore and compare the expression patterns of venom components depending on post-capture periods, venom gland-specific transcriptome and proteome analyses were conducted for five model hymenopteran species at a series of time points after capture. Venom gland-specific genes with signal sequences were considered as putative venom component genes. Expression patterns of venom gland-specific genes in all the social wasps and bees examined varied considerably depending on the post-capture period. Higher numbers of venom genes exhibited a decreasing expression pattern than an increasing pattern as the capture period increased. For example, genes encoding most of the allergens (dipeptidyl peptidase 4, endocuticle structural glycoprotein, odorant-binding protein, phospholipase A1, A2, B1, serine protease, serine protease inhibitor and venom allergen 5), pain-producing factor (mast cell degranulating peptide), and paralyzing factor (neprilysin) commonly exhibited decreasing expression patterns in all of the hymenopteran species tested, except for some of the major venom genes in *Apis mellifera* and *Bombus ignitus*, which showed an increasing pattern. These findings indicate species- or group-specific variations in the expression patterns of major venom genes. Taken together, flash freezing in liquid nitrogen immediately after capture was determined to be the best way to obtain the most natural expression profiles of venom components in social wasp species, thus, enabling a better understanding of the toxic potential of venom in wasp sting accidents. This study provides guidance for establishing optimal protocols for venom gland isolation and venom extraction from wasps and bees that can ensure the most naturally represented venom composition.

Keywords: Stress-dependent toxicity; Transcriptome analysis; Proteome analysis; Social

wasp; Honeybee; Bumblebee

Journal Pre-proof

1. Introduction

Hymenopteran venoms are composed of various amines and molecules that mainly produce pain and allergic reactions. Stings from Apidae and Vespidae species (Hymenoptera) cause a wide range of reactions, from mild irritation to fatal systemic complications such as anaphylactic shock. Since more than 40 deaths are reported in the United States per year and approximately 4% of the population is susceptible to systematic hymenopteran sting allergic reactions (Warrell, 2012), it is necessary to analyze the venom components to understand their toxic potential and intoxication symptoms, thus, preventing the possible injury.

Venom components are known to vary significantly between different hymenopteran species (Leluk et al., 1989). Species-specific variations in the expression levels of venom components were found in *Vespa crabro* and *V. analis*, which belong to the same genus (Yoon et al., 2015). Overall, venom components in *V. crabro* and *V. analis* were similar, but *V. crabro* possessed venom with highly enriched in major venom components, thus exhibiting potentially higher toxicity than *V. analis*. Intraspecific variations were also found among individuals of the same population and the same age in the parasitoid wasp *Leptopilina boulardi* (Colinet et al., 2013). The giant ant *Dinoponera quadriciceps* also showed differences in venom proteins depending on collected regions, suggesting that environmental factors affect venom composition (Cologna et al., 2013). Seasonal variation of venom components in queen and worker bees implies their needs for venom at different times in their lives (Danneels et al., 2015). Taken together, the venom components of hymenopteran species can vary between individuals from the same population and of the same age, depending on various environmental factors. However, it is unknown yet whether the expression patterns of venom components change significantly over the post-capture periods.

Omics technologies have recently been employed in venom research. Rapid identification

of venom genes has been achieved by cost-effective deep sequencing of venom glands dissected from captured target organisms, and proteomic and peptidomic components of venoms can be efficiently determined using highly sensitive mass spectrometry analyses of isolated venoms. However, it remains to be elucidated at what post-capture time point venom gland-specific transcriptome and proteome analyses should be conducted to obtain the most representative natural venom compositions. Few studies have analyzed variations in the hymenopteran venom system and there are no reports on venom expression patterns depending on the post-capture period, which is likely proportional to the capture stress. Environmental stress such as a dynamic change in temperature was found to regulate the venom composition of the land snake *Bungarus multicinctus* (Yin et al., 2020) and to affect the developmental rate, sex ratio, and offspring density of the endoparasitoid wasp *Pteromalus puparum* (Hu, 1986; Hu and War, 1985; Wang et al., 2012). The capture stress, which could be similar to the situation of being threatened by intruders, likely affects the expression patterns of venom components in hymenopteran species. If the post-capture period extends, thereby increasing the capture stress, the overall venom expression patterns would likely change, resulting in altered venom compositions at the time of sampling and analysis. Therefore, to understand the toxic potential of hymenopteran venoms, it is essential to determine the effects of capture stress on alterations in venom composition and the expression level of each component at the normal state prior to capture stress. With this in mind, the goal of our study is to provide insights into the dynamic changes in venom components and their expression levels depending on different post-capture periods in some hymenopteran species.

2. Materials and methods

2.1. Insect collection, total RNA preparation, and venom extraction

Female social wasps (*Vespa crabro*, *Vespula flaviceps*, and *Polistes rothneyi*), bumblebee (*Bombus ignitus*), and honeybee (*Apis mellifera*) were collected from several regions of Korea. Some of the collected hymenopteran species were directly frozen in liquid nitrogen. The rest were frozen after being captured in a 50 mL tube for 2 and 6 h at room temperature. Since honeybees were much sensitive to capture stress than other social wasps and bumblebee, time points were modified to 0, 2, and 3 h. Total RNA was extracted from 10 specifically dissected venom glands of the hymenopteran species using 200 μ L of TRI reagent (Molecular Research Center, Cincinnati, OH, USA) according to the manufacturer's protocol. A total of 10 venom sacs were dissected out from each species, punctured using forceps, and transferred into GeneAll® SV Column type D tubes (GeneAll, Seoul, Korea) containing 100 μ L of 1x phosphate-buffered saline (PBS). Venom was extracted by centrifugation at $12,000 \times g$ for 2 min at 4 °C and stored at -70°C until use.

2.2. Construction of cDNA libraries and orthologous analysis

cDNA libraries were prepared for 151 bp paired-end sequencing using the TruSeq stranded mRNA sample preparation kit (Illumina, San Diego, CA, USA). The mRNA molecules were purified and fragmented from 1 μ g of total RNA using oligo (dT) magnetic beads. Single-stranded cDNAs were synthesized from the fragmented mRNAs through random hexamer priming, and second-strand synthesized double-stranded cDNA was prepared for downstream analysis. cDNA libraries were amplified after the sequential process of end repair, A-tailing, and adapter ligation. The quality of the cDNA libraries was evaluated using the Agilent 2100 BioAnalyzer (Agilent, CA, USA) and the libraries were quantified using the KAPA library quantification kit (Kapa Biosystems, MA, USA) according to the manufacturer's library

quantification protocol. Following cluster amplification of denatured templates, paired-end (2×151 bp) sequencing was performed using Illumina NovaSeq6000 (Illumina). Gene annotation and calculation of expression rates in transcripts per kilobase million (TPM) were conducted as previously described methods (Yoon et al., 2020). The raw transcriptome sequencing data from the following species have been uploaded to NCBI's Short Read Archive (SRA); *V. crabro* 0 h (SRR13249510), *V. crabro* 2 h (SRR13249450), *V. crabro* 6 h (SRR13249596), *V. flaviceps* 0 h (SRR13249600), *V. flaviceps* 2 h (SRR13249688), *V. flaviceps* 6 h (SRR13249601), *P. rothneyi* 0 h (SRR13235787), *P. rothneyi* 2 h (SRR13235789), *P. rothneyi* 6 h (SRR13235788), *B. ignitus* 0 h (SRR13222047), *B. ignitus* 2 h (SRR13235781), *B. ignitus* 6 h (SRR13235782), *A. mellifera* 0 h (SRR13213116), *A. mellifera* 2 h (SRR13221833), and *A. mellifera* 3 h (SRR13221838).

2.3. Short time-series expression miner (STEM) and hierarchical clustering of venom genes

Short time-series expression miner (STEM) software (version 1.3.13) was used to analyze the expression profiles of total venom genes over different post-capture periods (Ernst and Bar-Joseph, 2006). Each gene was assigned to the closest model profile using a Pearson correlation-based distance metric and profiles with p-values < 0.05 were considered significantly enriched for the genes (Ernst et al., 2005).

Sequence similarities across 26 venom gland-specific major venom genes were analyzed using a local BLAST search in BioEdit software version 7.2 (Hall et al., 2011) (bioedit.software.informer.com/) and false positives that covered less than 50% of the queries were filtered. Hierarchical clustering was conducted for the 26 orthologous major venom genes by using the unweighted-pair group method with arithmetic mean (UPGMA) using

POPTREE2 software (Takezaki et al., 2010) (med.kagawa-u.ac.jp/~genomelb/takezaki/poptreew/index.html) and genetic distance was calculated using Nei's D_A distance measurement. A heat map was generated using Heml version 1.0 (Deng et al., 2014) (hemi.biocuckoo.org/index.php).

2.4. Proteomic analysis

The venom extracts from *V. crabro* and *B. ignitus* were separated using sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) (4%–12% Bis-Tris gel). Once the dye front had moved approximately 1 cm from the wells, the gel above the dye front was diced and rinsed with 200 μ L water, 200 μ L of 25 mM ammonium bicarbonate in 50% (v/v) acetonitrile, and 100 μ L of acetonitrile to dehydrate the gel plugs. The dry gel plugs were rehydrated in 150 μ L of 25 mM ammonium bicarbonate containing 25 ng/ μ L trypsin. An additional 100 μ L of 25 mM ammonium bicarbonate was added and the gel plugs were incubated overnight at 37°C. After adjusting to approximately pH 2-3, the samples were desalted in C_{18} spin columns (Harvard Apparatus, Holliston, MA, USA) and completely dried in a centrifugal concentrator. The pooled extracts were reduced to dryness and reconstituted in 80% acetonitrile/0.1% formic acid for trapped ion mobility spectrometry-time-of-flight mass spectrometry (TIMS-TOF MS). The NanoElute and TIMS-TOF MS settings were followed by previously described methods (Yoon et al., 2020).

Data files were uploaded to PEAKS X (Bioinformatics Solutions Inc., Waterloo, ON, Canada) for processing, de novo sequencing and database searching. The sequences were searched against the deduced protein sequence database obtained from the venom gland-specific transcriptome data of all hymenopteran species tested with mass error tolerances of 20 ppm and 0.05 Da for parent and fragment, respectively. False discovery rate (FDR)

estimation was enabled and peptides were filtered at a 1% FDR at the peptide spectrum match level.

The analyzed results from peptide/protein identification and quantification software PEAKS X (Bioinformatics Solutions Inc.) were described as coverage, average mass, and peak area. The average mass was calculated by summing the masses of the isotopes of the elements, each of which was multiplied by its natural abundance. The peak area is the area under the curve of the peptide feature found at the same mass-to-charge ratio (m/z) and retention time as the MS/MS scan, which can be used to quantitatively identify the abundance of the peptide.

The *V. crabro* and *B. ignitus* proteome data were uploaded to Zenodo (European Organization for Nuclear Research) with the corresponding DOI number (10.5281/zenodo.5149772) and link address (<https://zenodo.org/record/5149772>).paralog groups

2.5. Quantitative real-time PCR

DNase I (Takara, Kyoto, Japan) was treated to eliminate DNA contamination of total RNA extracted from 10 venom glands of three social wasps (*V. crabro*, *V. flaviceps*, and *P. rothneyi*), one bumblebee (*B. ignitus*), and one honeybee (*A. mellifera*), and cDNA was synthesized using SuperScript IV reverse transcriptase (Invitrogen, Carlsbad, CA, USA). Nine venom genes (apidaecin, arginine kinase, defensin 1, endocuticle structural glycoprotein, hymenoptaecin, venom allergen 5, hyaluronidase, phospholipase A1, and mastoparan) and a reference gene (dimethyladenosine transferase) were selected to verify the transcriptional abundance of transcriptome data (Table S1). Gene-specific primers were designed with

similar conditions, including length and %GC content, and qPCR was conducted as previously described methods (Yoon et al., 2020).

3. Results

3.1. Transcriptional profiles of the venom glands of some social hymenopteran species

The total number of bases (bp) and total trimmed reads obtained from RNA sequencing of *V. crabro*, *V. flaviceps*, *P. rothneyi*, *B. ignitus*, and *A. mellifera* are shown in Table 1. A total of 11,097; 23,426; 25,438; 29,138, and 15,314 assembled transcripts from each hymenopteran species were identified using a BLAST search (Table 1).

Quantitative real-time PCR (qPCR) was conducted for seven genes in Apidae (*B. ignitus*) and Vespidae (*V. crabro*, *V. flaviceps*, and *P. rothneyi*) species to verify the transcriptional abundance of venom genes (Table S1). As determined by TPM values, all the genes tested in Vespidae showed a continuously decreasing pattern, whereas those in Apidae exhibited a continuously increasing pattern in their relative transcription levels as post-capture periods increased. Similar patterns of expression level between relative transcription levels and TPM values suggest that transcriptional abundances of venom genes estimated from TPM values of transcriptome analysis are reliable (Fig. S1).

3.2. Expression patterns of total venom genes in some hymenopteran species over different post-capture periods

A total of 7.4%, 8.4%, 8.7%, 6.8%, and 9.5% of assembled transcripts from venom glands of *V. crabro*, *V. flaviceps*, *P. rothneyi*, *B. ignitus*, and *A. mellifera*, respectively, were predicted

to possess signal peptides and were thus considered to encode venom components (Table S2). To investigate and analyze the overall expression profiles of entire venom genes over different post-capture periods, we conducted a STEM analysis. More than 44% of venom genes in *V. crabro* (44.9%), *V. flaviceps* (65.2%), *P. rothneyi* (52.8%), *B. ignitus* (73.3%), and *A. mellifera* (52.5%) showed highly significant changes in their expression over the different post-capture periods ($p < 0.001$; Table 2 and Fig. 1). A total of 16 profiles were observed and categorized based on expression patterns, and the genes showing similar expression patterns were clustered under the same profile numbers (Table 2). Overall, genes with decreasing expression patterns over the post-capture period were more predominant than genes with increasing expression patterns in most of the hymenopteran species examined, except *A. mellifera*. Profiles 1, 3, 4, 5, 8, and 10, which were mainly composed of genes with decreasing expression patterns, were commonly observed in at least three of the hymenopteran species examined (profiles 1, 5, and 4 in *V. flaviceps*, *B. ignitus*, and *A. mellifera*, and profiles 8, 3, and 10 in *V. crabro*, *P. rothneyi*, and *B. ignitus*). Profiles 11, 12, and 15, composed of genes showing slightly increasing expression patterns, were observed with high significance only in *V. crabro*, *P. rothneyi*, and *A. mellifera*. Interestingly, genes clustered in profiles 12 and 15 with constant or gradually increasing expression patterns were only observed in *A. mellifera* (Fig. 1). Genes clustered in profiles 2 and 6 with ‘decreasing and increasing’ expression patterns were only observed in *V. flaviceps* (Fig. 1).

3.3. Stress-dependent expression patterns of major venom genes in some social hymenopteran species

The expression levels of 26 major venom genes in the hymenopteran species examined exhibited changes with the following six different patterns (Fig. S2): continuously decreasing

as the capture period increased (profiles 1, 3, and 4), decreasing with the highest expression level at 2 h of capture stress (profiles 8 and 10), decreasing with the lowest expression level at 2 h of capture stress (profiles 2, 5, and 6), continuously increasing as capture period increased (profiles 13, 14, and 16), increasing with the highest expression level at 2 h of capture stress (profiles 11, 12, and 15) and increasing with the lowest expression level at 2 h of capture stress (profiles 7 and 9). The number of major venom genes with a pattern of ‘decreasing with the highest expression level at 2 h of capture’ showed the highest percentages in *V. crabro* and *P. rothneyi*; the number with a continuously decreasing pattern exhibited the highest percentages in *V. flaviceps* and *A. mellifera* and the highest number of genes with a continuously increasing pattern was found in *B. ignitus* (Fig. S2). The major venom genes in social wasps (*V. crabro*, *V. flaviceps*, and *P. rothneyi*) and honeybee (*A. mellifera*) were mainly composed of genes with decreasing pattern of expression levels (76.9%, 88.5%, 61.5%, and 53.8%, respectively). However, the bumblebee (*B. ignitus*) had a significantly higher percentage of genes with the increasing pattern of expression levels (84.6%), indicating that the different patterns of dynamic changes in expression levels in *B. ignitus* depending on post-capture periods could be caused by various factors, including ecological traits or immune systems.

At 6 h post-capture in *V. crabro*, *V. flaviceps*, *P. rothneyi*, and *B. ignitus* and 3 h post-capture in *A. mellifera*, antimicrobial peptides such as apidaecin, defensin 1, and hymenoptaecin exhibited decreasing expression patterns compared with the rest of the genes represented and were clustered in a clade (Fig. 2). In particular, apidaecin, defensin 1, and hymenoptaecin from 6 h-captured *V. flaviceps* showed 7589-, 1139- and 3022-fold lower expression levels, respectively, than those from directly frozen species, and most of the social wasps (*V. crabro*, *V. flaviceps*, and *P. rothneyi*) and *A. mellifera* exhibited decreasing pattern in their expression levels. In contrast, apidaecin, defensin 1, and hymenoptaecin from 6 h-

captured *B. ignitus* showed 17-, 13- and 121-fold higher expression levels, respectively, compared with the same genes from directly frozen *B. ignitus*.

Most of the allergens (dipeptidyl peptidase 4, endocuticle structural glycoprotein, odorant-binding protein, phospholipase A1/A2/B1, serine protease, serine protease inhibitor, and venom allergen 5), pain-producing factor (mast cell degranulating peptide), and paralyzing factor (neprilysin), which are highly expressed in social wasp venoms as colony-defending substances (Yoon et al., 2020) commonly exhibited decreasing pattern of expression levels in all of the social wasps examined (*V. crabro*, *V. flaviceps*, and *P. rothneyi*) (Fig. 2). In particular, hyaluronidase, mast cell degranulating peptide, phospholipase A1, and venom allergen 5 showed 3- to 295-fold lower expression levels at 6 h post-capture compared with directly frozen species.

The mast cell degranulating peptide and acid phosphatase in *A. mellifera* showed significantly high fold-changes (85- and 265 fold, respectively) of expression levels between 3 h-captured and directly frozen samples (Fig. 2). Although the expression levels of mast cell degranulating peptide and acid phosphatase in *A. mellifera* exhibited decreasing expression patterns with significantly high fold-changes as the capture period increased, the proportion of genes with increasing expression patterns was higher than those of social wasps. Social wasps (*V. crabro*, *V. flaviceps*, and *P. rothneyi*) belonging to the Vespidae family showed higher percentages of major venom genes with decreasing expression levels (77%, 88%, and 62%, respectively) compared with genes of increasing expression patterns. Both *B. ignitus* and *A. mellifera*, which belong to the Apidae family, exhibited slightly and significantly lower percentages of genes with decreasing expression pattern (12% and 54%, respectively) compared with genes in social wasps, indicating species-specific variations in expression patterns of major venom genes. In *B. ignitus* and *A. mellifera*, 85% and 38% of the major

venom genes, respectively, exhibited increasing expression patterns as the capture period increased. In contrast, phospholipase A2, one of the major hymenopteran venom allergens, commonly showed decreasing expression in the Vespidae and Apidae species examined.

3.4. Stress-dependent proteome analysis of venom proteins in *V. crabro* and *B. ignitus*

The fold-changes in peak areas of major venom proteins between 6 h-captured and directly frozen samples were similar to those of TPM values of major venom genes from *V. crabro* and *B. ignitus*. In addition, the increasing and decreasing patterns of peak areas of major venom proteins matched well with those of TPM values from major venom genes (Table 3). Most of the major venom proteins from *V. crabro* and *B. ignitus* exhibited the highest peak areas at 2 h post-capture, whereas only defensin 1 of *B. ignitus* showed continuously increasing peak areas as the capture period increased. The total percentages of venom proteins with decreasing peak areas were 79% and 60% of the total identified proteins in *V. crabro* and *B. ignitus*, respectively. Among the proteins with decreasing peak areas, 71% and 68% of proteins showed decreasing transcription patterns of genes encoding them in *V. crabro* and *B. ignitus*, respectively (Table S3). These increasing and decreasing patterns of peak areas from proteome data analysis showed trends similar to expression levels identified from venom gland-specific transcriptome data analysis. The total identified venom proteins from *V. crabro* and *B. ignitus* were composed of 92% and 82% of the total proteins exhibiting the highest or lowest peak areas at 2 h post-capture, respectively, indicating that *V. crabro* and *B. ignitus* venom proteins might also be sensitive at 2 h post-capture, as shown by the results of venom gland-specific transcriptome analysis.

3.5. Identification of uncharacterized soluble proteins in some social hymenopteran species

Genes encoding uncharacterized proteins that possess signal peptides with high expression levels (TPM > 500) from directly frozen hymenopterans were identified in the food allergy research resource program-based allergen protein database (allergenonline.org). Most of the uncharacterized soluble proteins (USPs) were predicted to be allergens and muscle-regulating proteins. Most of the allergens in the hymenopteran species examined, except *V. flaviceps* and *P. rothneyi*, showed increasing expression patterns as the capture period increased and the number of allergens with the highest or lowest expression levels at 2 h of capture stress was higher than that of allergens with continuously increasing or decreasing expression levels, implying that abundantly expressed soluble allergens also exhibit dynamic changes in expression levels, depending on the post-capture period (Table 4).

4. Discussion

Invertebrates possess innate immune system to defense themselves against environmental stresses (Qasim et al., 2020; Qasim et al., 2021). Especially antimicrobial peptides are known to play a crucial role in the immune system of bumblebees. Defensin 1 and abaecin from *B. terrestris* reduced parasitemia of the bumblebee parasite *Crithidia bombi*, demonstrating their functional role in anti-*Crithidia* defense (Deshwal and Mallon, 2014). The expression levels of abaecin, apidaecin, and hymenoptaecin in *B. impatiens* were significantly increased, and defensin 1 also increased after imidacloprid treatment. The expression levels of these antimicrobial peptides were dependent on concentration and time of exposure, indicating that neonicotinoids might be modulators of immune gene expression in *B. impatiens* (Simmons and Angelini, 2017). This implies the possibility that environmental stress may enhance the

immune system of *B. ignitus*, with dynamic changes in the expression levels of antimicrobial peptides.

Decreasing expression pattern of most of the allergens, pain-producing factor, and paralyzing factor in all of the social wasps examined indicates that the potential for expressing these venom components is maintained at a maximal level at the unthreatened normal state but decreases gradually upon exposure to external threats such as intruding vertebrates. This finding suggests that the allergenic potential of hymenopteran venom should be assessed using transcriptome profiles of flash-frozen venom glands or proteome profiles obtained from venom isolated immediately after capture. *A. mellifera* showed a high ratio of increasing expression pattern in total venom genes with high significance and exhibited 38% of the major venom genes with an increasing expression levels as capture periods increased. Profiles 12 and 16 in the STEM analysis of *A. mellifera* were clustered with major venom genes such as dipeptidyl peptidase 4, exocuticle structural glycoprotein, hyaluronidase, neprilysin, odorant-binding protein and serine protease showing 2.1-to 10.3-fold higher expression levels at 3 h post-capture (Fig. 1D and Fig. 2). Increasing expression levels in these genes were only observed in the Apidae family. Especially, hyaluronidase, which is known as a venom spreading factor (Chain and Duthie, 1940), was clustered in profile 16 and only showed an increasing expression pattern in *A. mellifera*. This unique expression pattern of some venom genes, which is highly composed of increasing expression pattern as the capture period increased in *A. mellifera*, might be explained by the domestication of honeybees during which bees get used to frequent interaction with the potential intruder; beekeepers. Through the domestication process, honeybees may have evolved to desensitize their venom response system, responding with less sensitive to human activity and enabling resource trade-off for stress management.

Species-specific changes in the venom gene expression patterns were observed over the different post-capture periods. The number of soluble venom genes showing decreasing expression pattern was higher than that of increasing expression pattern in the social hymenopteran species tested, indicating that the overall venom expression pattern changes as the capture period increases. Therefore, freezing the samples in liquid nitrogen immediately after capture could be the best way to obtain venom components with the most natural expression profiles in some social hymenopteran species. This is the first study to provide guidance for establishing optimal protocols for the extraction of naturally represented venom components.

5. Conclusion

The major venom genes encoding most of the allergens (dipeptidyl peptidase 4, endocuticle structural glycoprotein, odorant-binding protein, phospholipase A1, A2, B1, serine protease, serine protease inhibitor and venom allergen 5), pain-producing factor (mast cell degranulating peptide), and paralyzing factor (neprilysin) exhibited a higher ratio of decreasing expression pattern than increasing pattern as the capture period increased in all of the hymenopteran species tested. These findings indicate species- or group-specific variations in the expression patterns of major venom genes and suggest that flash freezing in liquid nitrogen after capture is determined to be the best way to obtain the most natural expression profiles of venom components in social wasp species. Taken together, this study provides guidance for establishing optimal protocols for venom gland isolation and venom extraction from wasps and bees that can ensure the most naturally represented venom composition, enables a better understanding of the toxic potential of venom in wasp sting accidents, and provides insights into the dynamic changes in venom components and their expression levels

depending on different post-capture periods in some hymenopteran species.

Journal Pre-proof

References

- Chain, E., Duthie, E.S., 1940. Identity of hyaluronidase and spreading factor. *Brit J Exp Pathol* 21, 324-338.
- Colinet, D., Deleury, E., Anselme, C., Cazes, D., Poulain, J., Azema-Dossat, C., Belghazi, M., Gatti, J.L., Poirie, M., 2013. Extensive inter- and intraspecific venom variation in closely related parasites targeting the same host: The case of *Leptopilina* parasitoids of *Drosophila*. *Insect Biochem Molec* 43, 601-611.
- Cologna, C.T., Cardoso, J.D., Jourdan, E., Degueldre, M., Upert, G., Gilles, N., Uetanabaro, A.P.T., Neto, E.M.C., Thonart, P., de Pauw, E., Quinton, L., 2013. Peptidomic comparison and characterization of the major components of the venom of the giant ant *Dinoponera quadricaps* collected in four different areas of Brazil. *J Proteomics* 94, 413-422.
- Danneels, E.L., Van Vaerenbergh, M., Debyser, G., Devreese, F., de Graaf, D.C., 2015. Honeybee venom proteome profile of queens and winter bees as determined by a mass spectrometric approach. *Toxins* 7, 4468-4483.
- Deng, W.K., Wang, Y.B., Liu, Z.X., Cheng, H., Xue, Y., 2014. HeatM: a toolkit for illustrating heatmaps. *Plos One* 9.
- Deshwal, S., Mallon, E.B., 2014. Antimicrobial peptides play a functional role in bumblebee anti-trypanosome defense. *Dev Comp Immunol* 42, 240-243.
- Ernst, J., Bar-Joseph, Z., 2006. STEM: a tool for the analysis of short time series gene expression data. *Bmc Bioinformatics* 7.
- Ernst, J., Nau, G.J., Bar-Joseph, Z., 2005. Clustering short time series gene expression data. *Bioinformatics* 21, 1159-1168.
- Hall, T., Biosciences, I., Carlsbad, C., 2011. BioEdit: an important software for molecular biology. *GERF Bull Biosci* 2, 60-61.
- Hu, C., 1986. The Developmental Rate of *Pteromalus puparum* in relation to temperature. *Acta Entomol Sinica* 29, 101-103.
- Hu, C., Wan, X.S., 1988. Studies on the sex-ratio of *Pteromalus puparum* L. *Acta Entomol Sinica* 31, 332-335.
- Leluk, J., Schmidt, J., Jones, D., 1989. Comparative studies on the protein-composition of hymenopteran venom reservoirs. *Toxicon* 27, 105-114.
- Qasim, M., Ul Islam, S., Islam, W., Noman, A., Khan, K.A., Hafeez, M., Hussain, D., Dash, C.K., Bamsile, B.S., Akatse, K.S., Rizwan, M., Nisar, M.S., Jan, S., Wang, L.D., 2020. Characterization of mycotoxins from entomopathogenic fungi (*Cordyceps fumosorosea*) and their toxic effects to the development of asian citrus psyllid reared on healthy and diseased citrus plants. *Toxicon* 188, 39-47.
- Qasim, M., Xiao, H.M., He, K., Omar, M.A.A., Hussain, D., Noman, A., Rizwan, M., Khan, K.A., Al-zoubi, O.M., Alharbi, S.A., Wang, L.D., Li, F., 2021. Host-pathogen interaction between asian citrus psyllid and entomopathogenic fungus (*Cordyceps fumosorosea*) is regulated by modulations in gene expression, enzymatic activity and HLB-bacterial population of the host. *Comp Biochem Phys C* 248.
- Simmons, W.R., Angelini, D.R., 2017. Chronic exposure to a neonicotinoid increases expression of antimicrobial peptide genes in the bumblebee *Bombus impatiens*. *Sci Rep-Uk* 7.
- Takezaki, N., Nei, M., Tamura, K., 2010. POPTREE2: Software for constructing population trees from allele frequency data and computing other population statistics with windows interface. *Mol Biol Evol* 27, 747-752.
- Wang, H., Li, K., Zhu, J.Y., Fang, Q., Ye, G.Y., Wang, H., Li, K., Zhu, J.Y., 2012. Cloning and expression pattern of heat shock protein genes from the endoparasitoid wasp, *Pteromalus puparum* in response to environmental stresses. *Arch Insect Biochem* 79, 247-263.

- Warrell, D.A., 2012. Venomous bites, stings, and poisoning. *Infect Dis Clin N Am* 26, 207-+.
- Yin, X.M., Guo, S., Gao, J.H., Luo, L., Liao, X.J., Li, M.Q., Su, H., Huang, Z.H., Xu, J., Pei, J., Chen, S.L., 2020. Kinetic analysis of effects of temperature and time on the regulation of venom expression in *Bungarus multicinctus*. *Sci Rep-Uk* 10.
- Yoon, K.A., Kim, K., Kim, W.-J., Bang, W.Y., Ahn, N.-H., Bae, C.-H., Yeo, J.-H., Lee, S.H., 2020. Characterization of venom components and their phylogenetic properties in some aculeate bumblebees and wasps. *Toxins* 12, 47.
- Yoon, K.A., Kim, K., Nguyen, P., Seo, J.B., Park, Y.H., Kim, K.G., Seo, H.Y., Koh, Y.H., Lee, S.H., 2015. Comparative functional venomomics of social hornets *Vespa crabro* and *Vespa analis*. *J Asia-Pac Entomol* 18, 815-823.

Acknowledgments

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The authors thank Sung Jin Kim for participation in collecting bees and providing technical assistance to extraction of total RNA.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Corresponding authors

Correspondence to Kyungjae Andrew Yoon and Si Hyeock Lee.

Figures & Tables

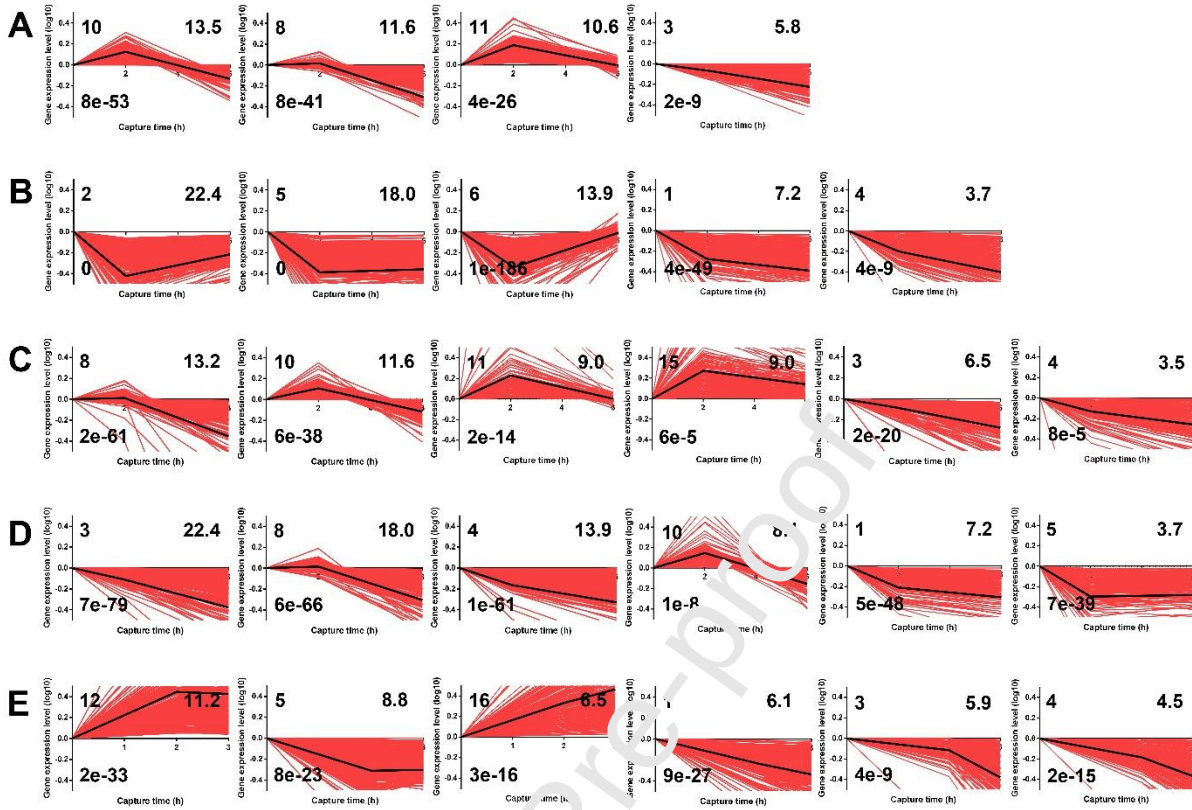


Fig. 1. Venom gene expression profiles clustered by Short Time-series Expression Miner (STEM) analysis in *V. crabro* (A), *V. flaviceps* (B), *P. rothneyi* (C), *B. ignitus* (D), and *A. mellifera* (E). The profiles are in order of *p*-value significance. The x- and y-axis represent capture period and log₁₀ values from fold change of gene expression level, respectively. Black solid line in each cluster indicates the mean values of expression levels. The number of profiles, *p*-value, and the proportion (%) in total genes are at the top-left, bottom-left, and top-right corner in each cluster, respectively.

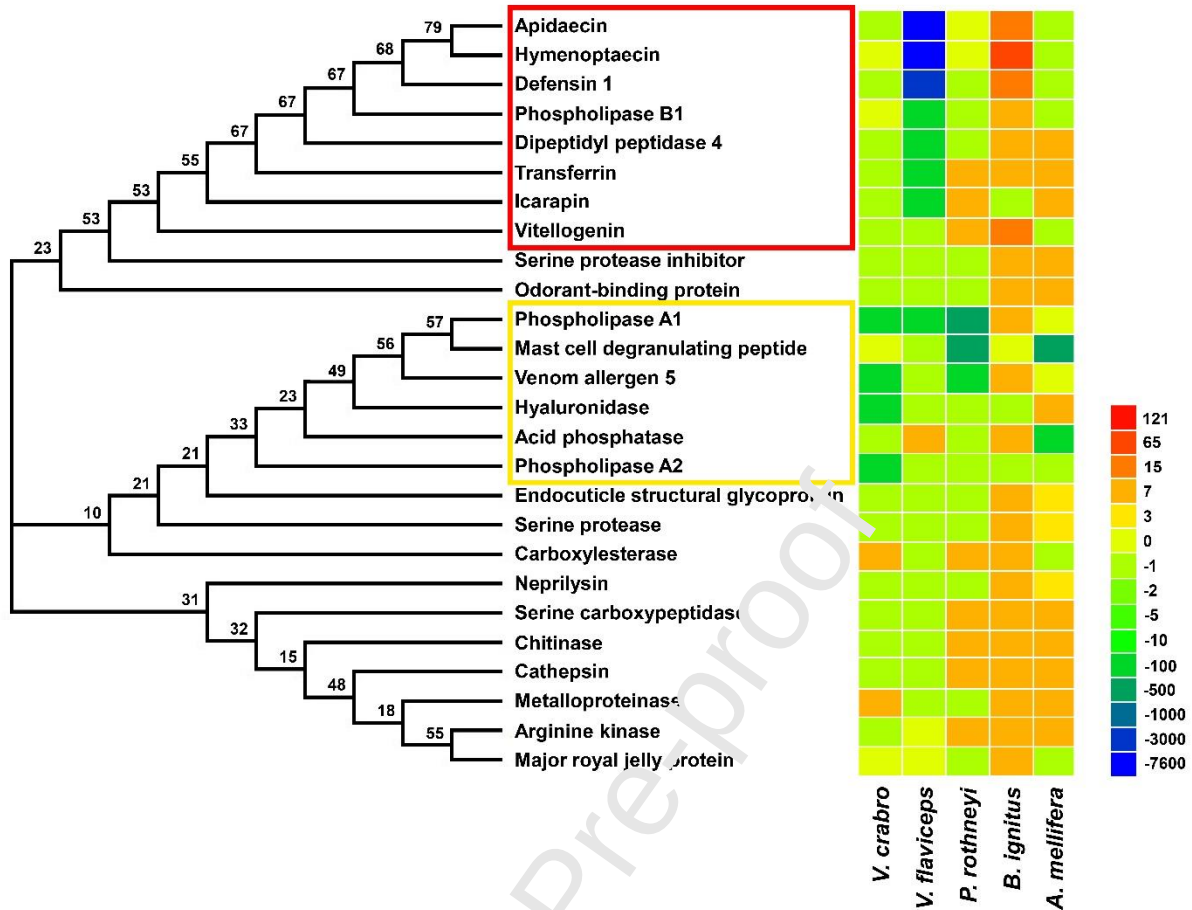


Fig. 2. Hierarchical clustering of 25 orthologous major venom genes in five vespid and apidae species (left panel), and fold-change of their differential expression levels between directly frozen and 6 h of capture period (directly frozen and 3 h of capture period for *A. mellifera*) represented in a heat map (right panel). Each row represents the venom genes analyzed and each column represents the species tested. The percentages of bootstrap values based on 1,000 pseudoreplicates are shown on each dendrogram node. Red and yellow rectangles indicate the clades clustered with proteins showing extremely and significantly increased or decreased expression levels, respectively.

Table1. Summary of the venom gland transcriptome libraries of five hymenopteran insects.

Species	Post-capture period (h)	Total number of bases (bp)	Trimmed reads	Total number of assembled transcripts
<i>V. crabro</i>	0	3,436,794,126	16,443,746	11,097
	2	1,526,913,208	6,599,372	
	6	1,089,076,024	4,577,244	
<i>V. flaviceps</i>	0	6,298,583,574	41,712,474	23,426
	2	6,452,282,548	42,730,348	
	6	5,729,652,720	37,944,720	
<i>P. rothneyi</i>	0	4,246,724,302	20,453,384	25,438
	2	10,198,681,336	45,712,116	
	6	4,638,207,506	18,927,256	
<i>B. ignitus</i>	0	6,390,211,582	42,319,282	29,138
	2	6,409,123,426	42,444,516	
	6	5,833,025,206	38,629,306	
<i>A. mellifera</i>	0	6,048,028,670	40,053,170	15,314
	2	5,836,378,010	38,651,510	
	3	5,680,854,050	37,121,510	

Table 2. Patterns of venom gene expression profiles over post-capture periods in some social hymenopteran species as determined by Short Time-series Expression Miner (STEM) analysis.

Profiles	Expression pattern	Pattern	Profiles	Expression pattern	Pattern
1	Decreasing and gradually decreasing		9	Unchanged and increasing	
2	Decreasing and gradually increasing		10	Increasing and sharply decreasing	
3	Decreasing and sharply decreasing		11	Increasing and gradually decreasing	
4	Constantly decreasing		12	Increasing and unchanged	
5	Decreasing and unchanged		13	Constantly increasing	
6	Decreasing and increasing		14	Increasing and sharply increasing	
7	Decreasing and sharply increasing		15	Increasing and gradually decreasing	
8	Unchanged and decreasing		16	Increasing and gradually increasing	

Table 3. Major venom proteins of *V. crabro* and *B. ignitus* identified by mass spectrometry analysis and expression level of encoding genes.

Species	Description	Coverage ¹ (%)	Avg. mass ² (Da)	Area ³ (0 h)	Area (2 h)	Area (6 h)	Fold-Change ⁴ (TPM)
<i>V. crabro</i>	Arginine kinase [<i>Vespa mandarinia</i>]	79	39892	2285100	3604700	1495000	-1.22
	Transferrin [<i>Vespa mandarinia</i>]	54	78511	500200	740330	373540	-1.03
	Apolipophorins [<i>Vespa mandarinia</i>]	46	378343	439880	1639400	294990	-1.32
	Venom allergen 5 [<i>Vespa mandarinia</i>]	61	25209	303100	654370	60557	-62.6
	Vitellogenin [<i>Vespa mandarinia</i>]	39	202992	111440	562870	68337	-1.01
	Hyaluronidase [<i>Vespa mandarinia</i>]	29	42443	88813	41710	77417	-47.1
	Dipeptidyl peptidase 4 [<i>Vespa mandarinia</i>]	30	88475	57356	213700	19666	-6.91
	Phospholipase A1 [<i>Vespa affinis</i>]	55	37144	33201	280730	6394.7	-46.2
	Odorant binding protein [<i>Polistes canadensis</i>]	49	18959	24774	168840	6138.2	-2.34
<i>B. ignitus</i>	Serine carboxypeptidase [<i>Vespa mandarinia</i>]	21	53881	2519.3	11633	311.23	-1.25
	Apolipophorins [<i>Bombus vosnesenskii</i>]	48	375331	13670	742780	23781	-1.56
	Hyaluronidase [<i>Bombus vosnesenskii</i>]	35	44405	3539.8	26208	1995.5	-3.33
	Odorant binding protein [<i>Bombus vosnesenskii</i>]	51	14822	2029	74303	9669.5	1.63
	Defensin 1 [<i>Bombus impatiens</i>]	20	19739	0	1876.3	11769	12.86

¹The percentage of the protein sequence that is covered by supporting peptides.

²The sum of the abundance-weighted masses of all of the stable isotopes from elements.

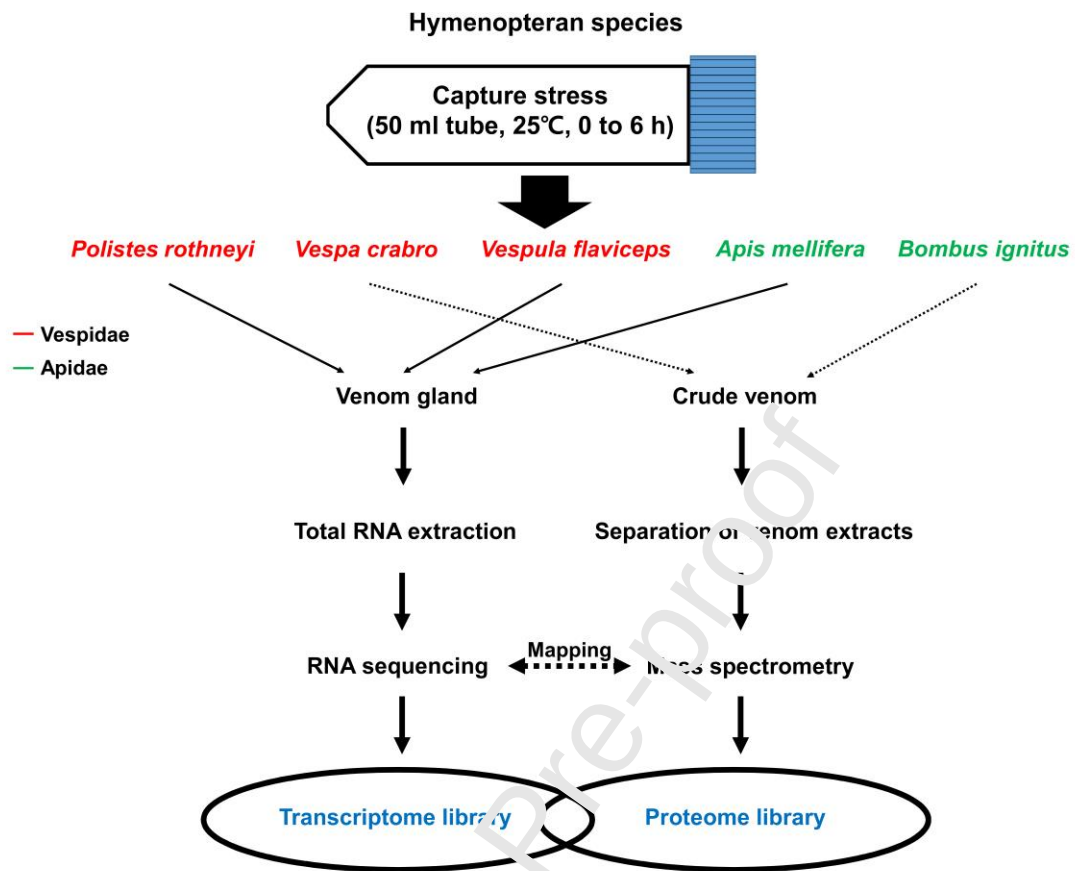
³The area under the curve of the peptide feature found at the same mass-to-charge ratio (m/z) and retention time.

⁴Fold change of expression levels between directly frozen and catch capture period.

Table 4. Variation of expression levels of secretory uncharacterized proteins in some social hymenopteran species depending on stress.

<i>V. crabro</i>					
No.	Description	TPM (0 h)	TPM (2 h)	TPM (6 h)	Similarity (%)
1	Allergen [<i>Sarcoptes scabiei</i>]	24009	21438	27712	54.8
2	Major allergen [<i>Parietaria judaica</i>]	5395	5457	6864	65.4
3	Allergen [<i>Lupinus angustifolius</i>]	1438	1872	2282	79.7
4	Tropomyosin [<i>Macrobrachium rosenbergii</i>]	702	1038	429	61.7
<i>V. flaviceps</i>					
No.	Description	TPM (0 h)	TPM (2 h)	TPM (6 h)	Similarity (%)
1	Conarachin [<i>Arachis hypogaea</i>]	16497	9337	12032	58.2
2	Allergen [<i>Sarcoptes scabiei</i>]	2446	2086	3566	55.1
3	Tropomyosin [<i>Octopus vulgaris</i>]	1126	862	785	60.9
4	Ovomucoid [<i>Gallus gallus</i>]	565	10	8	63.2
<i>P. rothneyi</i>					
No.	Description	TPM (0 h)	TPM (2 h)	TPM (6 h)	Similarity (%)
1	Allergen [<i>Dermatophagoides farinae</i>]	1166	1105	1025	53.2
2	Myosin light chain 1 [<i>Procambarus clarkii</i>]	948	996	788	55.9
3	Allergen [<i>Penaeus chinensis</i>]	688	724	423	57
<i>B. ignitus</i>					
No.	Description	TPM (0 h)	TPM (2 h)	TPM (3 h)	Similarity (%)
1	Tropomyosin [<i>Procambarus clarkii</i>]	6835	10320	6560	64.6
2	Kappa-casein [<i>Bos taurus</i>]	2973	2732	3205	51.1
3	Allergen [<i>Triticum aestivum</i>]	1436	1184	1473	57.1
4	Major allergen [<i>Aspergillus fumigatus</i>]	1297	3439	2394	46.7
5	Pollen allergen [<i>Holcus lanatus</i>]	1264	2079	5925	58.7
6	Venom protease [<i>Bombus pensylvanicus</i>]	644	842	822	60.5
7	Vitellogenin [<i>Apis mellifera</i>]	571	1266	552	54.8
<i>A. mellifera</i>					
No.	Description	TPM (0 h)	TPM (2 h)	TPM (6 h)	Similarity (%)
1	Allergen [<i>Euroglyphus maynei</i>]	8340	24236	40942	50.7

Graphical abstract



Highlights

Venom expression pattern changes as capture period increases in hymenopteran species

Venom genes of decreasing expression pattern are predominant in hymenopteran species

Honeybee possesses high ratio of major venom genes of increasing expression pattern

This study provides guidance for extraction of naturally represented venom components

Journal Pre-proof