Journal of Proteomics 329 (2026) 105676

Contents lists available at ScienceDirect

Journal of Proteomics

journal homepage: www.elsevier.com/locate/jprot

Comparative peptidomics of four wasp venoms reveals extensive peptide diversity, proteolytic patterns, and predicted bioactivities

Kai Wang [a,b], Jiangtao Qiao [c], Loïc Quinton [d], Gauthier Eppe [d], Xianbin Meng [e], Guozhong Huang [f], Eric Haubruge [b,*], Hongcheng Zhang [a,*]

[a] State Key Laboratory of Resource Insects, Institute of Apicultural Research, Chinese Academy of Agricultural Sciences, Beijing 100093, China
[b] Terra Research Center, Gembloux Agro-Bio Tech, University of Liege, 5030 Gembloux, Belgium
[c] School of Life Sciences and Food Engineering, Hebei University of Engineering, Handan 056038, China
[d] Laboratory of Mass Spectrometry, MolSys Research Unit, University of Liege, 4000, Liege, Belgium
[e] Equipment R&D Center, Shenzhen Medical Academy of Research and Translation, Shenzhen 518107, Guangdong, China
[f] Longling Huangshi Bee Industry Co., Ltd, Longling County, 678300, Yunnan, China.

ARTICLE INFO

Keywords:
Wasp venom
Peptidomics
Peptide diversity
Bioactive peptides
Cryptides

ABSTRACT

Wasp venoms possess complex compositions and diverse bioactivities, making them potential pharmacological sources. In this study, venoms from four wasp species (*Vespa mandarinia*, *V. velutina*, *V. basalis*, and *Provespa barthelemyi*) were collected by electrical stimulation and analyzed using liquid chromatography–tandem mass spectrometry. A total of 681 peptides were identified, nearly 90% of which had not been previously reported. Comparative analyses revealed pronounced species-specific signatures at both peptide and peptide family levels. Sequence-based analyses indicated that peptide release is consistent with targeted proteolytic cleavage patterns, exhibiting features resembling known substrate preferences of metalloproteases and serine proteases, rather than stochastic degradation. Bioinformatic predictions identified 291 peptides with potential bioactive properties spanning multiple functional categories, with angiotensin-converting enzyme (ACE) and dipeptidyl peptidase IV (DPP4) inhibitory activities being the most prominently represented. Among these, eight ACE inhibitory peptides and seventeen DPP4 inhibitory peptides were prioritized as candidates based on predicted safety profiles, and sequence-based analysis further identified ten putative cryptides. Overall, this study establishes the first comparative peptidomic dataset across four wasp venoms, providing insights into peptide diversity, inferred generation patterns, and predicted activities.

Significance: Venoms are rich sources of biologically active molecules and have historically provided templates for clinically used therapeutics. Although major protein toxins have been extensively characterized, the endogenous low-molecular-weight peptide fraction remains comparatively underexplored, particularly in social wasps, and systematic comparative resources remain limited. Venom peptidomic datasets inherently contain multiple layers of biological information, including diversification patterns, peptide origin, and potential bioactivity, yet these aspects are often interpreted independently. By integrating these dimensions, this study establishes a multi-level analytical framework for extracting biological insights and function-related information from venom peptidomes. The identification of consistent cleavage patterns suggests a degree of regulation in peptide generation, shifting the interpretation of venom peptides from degradation by-products toward biologically organized repertoires. Moreover, candidate prioritization illustrates how peptidomic datasets can generate experimentally testable functional hypotheses rather than serving solely as descriptive catalogs. The resulting dataset serves as a reference resource for cumulative comparative analyses. The analytical framework presented here also provides a transferable strategy for functional peptide discovery in other complex secretions.

* Corresponding authors at: Institute of Apicultural Research, Chinese Academy of Agricultural Sciences, Beijing 100093, China.
E-mail addresses: e.haubruge@uliege.be (E. Haubruge), 460414874@qq.com (H. Zhang).

https://doi.org/10.1016/j.jprot.2026.105676
Received 7 February 2026; Received in revised form 22 April 2026; Accepted 3 May 2026
Available online 5 May 2026

1. Introduction

Animal venoms are finely tuned biochemical arsenals that have evolved over millions of years, primarily serving as predation, defense, and ecological competition [1,2]. Beyond their ecological functions, venoms represent valuable sources of pharmacologically active compounds, and several venom-derived agents have been successfully translated into clinical applications, including captopril, developed from a peptide in *Bothrops jararaca* snake venom, and ziconotide, derived from the marine cone snail *Conus magus* [3–5]. Hymenopterans (sawflies, wasps, ants, and bees) represent one of the most species-rich insect lineages, comprising more than 153,000 described species [6]. Within this order, venoms from wasps, ants, and bees have been reported to exert a wide range of biological functions, including neuroregulatory, immunomodulatory, antimicrobial, anticoagulant, and metabolic effects [7–11]. Bee venom has been extensively investigated, particularly in the context of apitherapy and bee venom-based interventions, owing to its long-standing therapeutic use and reported clinical applications [12,13]. In contrast, wasp venoms are often associated with more intense pain and immune responses, remain comparatively less explored [14].

Given that naturally occurring bioactive peptides often exhibit high target specificity, wasp venoms represent a largely underexplored yet pharmacologically promising resource for therapeutic discovery [3]. Wasp venoms are complex biochemical mixtures comprising high-molecular-weight proteins (such as enzymes and allergens), biogenic amines, and diverse small peptides, including mastoparan, kinins, and cytolytic peptides [14]. To date, research on wasp venoms has predominantly focused on a limited number of high-abundance proteins and a small subset of well-characterized functional peptides, leaving the overall composition and diversity of venom peptides insufficiently explored [14,15]. In addition, a hallmark of animal venoms is their remarkable diversity, as even closely related species can exhibit substantial variation in venom composition and biological activity [16]. To date, studies focusing on venom peptides in wasps have largely been restricted to individual species, and multi-species comparative datasets remain scarce [14]. At the peptide level, fundamental features, including peptide diversity, sequence variability, and species-specific distribution patterns, remain poorly characterized across different wasp venoms [14,15]. Considering the remarkable species richness of wasps, their broad ecological distribution, and their pronounced biological effects, comparative analyses across multiple species are therefore essential.

Recent advances in trace-level analytical techniques, together with the maturation of high-resolution liquid chromatography coupled to mass spectrometry (LC–MS/MS) platforms, have enabled high-throughput peptidomic analyses, facilitating to identify known peptides and discover these previously unreported molecules [17,18]. Peptidomics has therefore been widely applied to elucidate the peptide composition of animal venoms across diverse taxa, including snakes, scorpions, and ants, and has proven particularly valuable for organisms with limited venom availability [19–21]. Moreover, the direct characterization of naturally occurring peptides provides unique insights into their generation mechanisms and potential biological activities [19]. Accumulating evidence indicates that animal venom peptides are not only produced through canonical precursor processing but may also arise from regulated proteolytic cleavage events, highlighting non-random mechanisms underlying venom peptide generation [22–25]. For example, cryptides, sequence-embedded bioactive peptides, have been reported in venom systems and may represent an additional source of functional peptides [25]. Thus, systematic analyses of venom peptides may offer a powerful framework to gain deeper insights into species-specific venom diversification and to uncover overlooked bioactive peptides.

Based on the above considerations, the present study analyzed peptidomics of venoms from four wasp species, *V. mandarinia*, *V. velutina*, *V. basalis*, and *Provespa barthelemyi*. Venoms were collected by electrical stimulation and analyzed using LC–MS/MS to comprehensively characterize the repertoire of naturally occurring peptides. On this basis, comparative analyses were performed to evaluate similarities and differences in peptide composition among species, with the aim of elucidating species-specific distribution patterns of wasp venom peptides. Sequence-based analyses were further employed to explore potential proteolytic cleavage features underlying peptide generation. Bioinformatic approaches were applied to predict the putative pharmacological activities of the identified peptides, providing an initial assessment of their functional potential. This study also reported the first discovery of potential cryptide candidates from wasp venoms, offering new perspectives for exploring the diversity and origins of bioactive peptides in wasp venoms and their prospective relevance to drug discovery.

2. Material and methods

2.1. Materials

Venoms from *Vespa mandarinia*, *V. velutina*, *V. basalis*, and *Provespa barthelemyi* were collected in Yunnan Province, China, in July 2022. The venoms were obtained using specialized collection devices placed at the entrance of the hives. Adult wasps were electrically stimulated to trigger venom release onto sterile glass plates. The collected venom was immediately lyophilized and stored at −80 °C in the dark prior to analysis.

2.2. Preparation of crude wasp venom (WV) extracts

Lyophilized wasp venom (4 mg) was reconstituted in 1 mL of ultrapure water and mixed by gentle vortexing for 1 min. The suspension was subsequently sonicated (100 W, 3 min) and clarified by centrifugation at 11,180 ×*g* for 10 min at 4 °C. The resulting supernatant was passed through a 0.22 μm membrane filter to obtain crude WV extracts.

2.3. Purification of proteins and peptides from crude WV extracts

Proteins and peptides in the crude WV extracts were separated by cation-exchange chromatography using a Toyopearl GigaCap S—650S column (Tosoh Bioscience, Tokyo, Japan). The system was operated with 0.2 M citrate buffer (pH 5.0) as solvent A and the same buffer supplemented with 1 M NaCl as solvent B. Samples (2 mL per injection) were loaded, then eluted at 1 mL/min, while UV absorbance was recorded at 280 nm. Based on venom diversity, species-specific multi-step gradient programs were applied. For *V. mandarinia*, the gradient was as follows: 0–37 min, 0% B; 37–220 min, linear 0 → 100% B; 220–250 min, hold at 100% B. For *V. velutina*, *V. basalis*, and *Provespa barthelemyi*, the gradient program was: 0–20 min, 0% B; 20–470 min, linear 0 → 50% B; 470–480 min, linear 50 → 100% B; 480–530 min, hold at 100% B. Fractions corresponding to individual peaks were collected, and peak areas were calculated to estimate their relative contributions to the total WV protein contents.

2.4. SDS-PAGE analysis of individual fractions

Protein profiles of the collected fractions were examined by reducing SDS-PAGE in a Tris–glycine buffer system. A pre-stained molecular weight marker (Beyotime Biotechnology, Shanghai, China; 10–180 kDa) was included for size estimation. Samples were mixed with 5× loading buffer, heat-denatured at 95 °C for 5 min, and briefly centrifuged to remove insoluble material. Electrophoresis was carried out at 120 V using a 4–20% Tris–glycine gel (Beyotime Biotechnology, Shanghai, China). Gels were subsequently fixed and stained with Coomassie Brilliant Blue R-250 for protein visualization.

2.5. Protein identification of electrophoresis bands

Protein bands were excised from SDS-PAGE gels and processed by in-gel digestion. Briefly, gel pieces were rinsed with ultrapure water, cut into 1 mm^3 cubes, and destained using 50% acetonitrile in 25 mM ammonium bicarbonate. Disulfide bonds were reduced with 10 mM dithiothreitol (56 °C, 30 min) followed by alkylation with 55 mM iodoacetamide. The proteins were then digested with sequencing-grade trypsin (12.5 ng/μL) overnight at 37 °C. Peptides were recovered by two consecutive extractions with 1% trifluoroacetic acid in 50% aqueous acetonitrile (30 min each) and subsequently concentrated to near dryness using a SpeedVac centrifugal vacuum concentrator.

Protein identification was performed on an Orbitrap Fusion mass spectrometer coupled to an Easy-nLC 1000 system (Thermo Fisher Scientific). Peptide mixtures were resolved on a C18 trap column (Acclaim PepMap, 100 μm i.d. × 2 cm, 3 μm; Thermo Fisher Scientific) using a 60-min linear gradient at a flow rate of 300 nL/min. Solvent A was 0.1% formic acid in water and solvent B was 0.1% formic acid in acetonitrile. Data were acquired in a data-dependent acquisition (DDA) mode, with full MS scans collected in the Orbitrap over an *m/z* range of 350–1500 at a resolution of 120,000. The most intense precursors were subsequently selected for top-speed MS/MS analysis in the ion trap using a 1.6 m/z isolation window and 35% normalized collision energy. Raw spectra were processed in Proteome Discoverer (v2.4, Thermo Fisher Scientific, USA) and searched against the *Vespinae* FASTA database downloaded from NCBI (taxonomy ID: 7439; 92,332 entries; accessed May 6, 2023).

2.6. Peptidomics analysis

Peptide-containing fractions obtained from the 2.3 section were first processed using 3 kDa ultrafiltration centrifugal devices (Millipore, MA, USA) to enrich low-molecular-weight peptides (<3 kDa). The resulting filtrates were desalted and concentrated on a C18 solid-phase extraction cartridge (Waters, USA). After drying under vacuum, peptides were reconstituted in 20 μL of 0.1% (*v*/v) formic acid prepared in ultrapure water and subjected to LC–MS/MS analysis.

Peptidomics data were acquired under the same LC–MS/MS conditions described in Section 2.5. Raw MS/MS files were processed in PEAKS X Pro (Peaks Studio 10.6; Bioinformatics Solutions Inc., Waterloo, Canada) using the *de novo* sequencing workflow with the following settings: precursor mass tolerance, 20 ppm; fragment ion tolerance, 0.02 Da; enzyme specificity, none; and variable modifications including oxidation, deamidation, and protein *N-terminal* acetylation, allowing up to three variable post-translational modifications (PTMs) per peptide. As tandem mass spectrometry cannot reliably distinguish isoleucine from leucine, isoleucine residues were reported as leucines in *de novo* sequences. Peptides with ALC ≥ 80% and non-zero peak areas were retained as high-confidence endogenous venom peptides and compiled into an in-house venom peptide database consisting of four species-specific libraries (Vm, Vv, Vb, and Pb) for downstream comparative and bioactivity analyses (see Supplemental Peptide Database_Vm, _Vv, _Vb, and _Pb.txt). Homology-based annotation was performed using the UniProt BLAST tool against UniProtKB entries restricted to *Vespa* (taxonomy ID: 7443) and *Provespa* (taxonomy ID: 76993) (accessed October 10, 2024). Peptide novelty was further evaluated using the UniProt Peptide Search tool (https://www.uniprot.org/peptide-search) against UniProtKB (accessed October 18, 2024).

2.7. Sequence-level specificity and conservation analysis of peptides across four wasp venoms

To assess the sequence-level specificity and conservation of venom peptides among the four wasp species, we firstly performed an UpSet analysis based on exact sequence matching, enabling a direct comparison of fully identical peptides across species. To further explore sequence conservation, peptides were clustered at 90% sequence identity using MMseqs2 implemented on the Galaxy server (https://usegalaxy.eu/), and the resulting clusters were visualized as an additional UpSet plot. All analyses were performed in RStudio (version 2025.051; R version 4.5.1) using the UpSetR package [26] to quantify unique and shared peptides (or peptide family clusters) among species.

2.8. Analysis of P4–P4' cleavage site motifs

To define the sequence motifs governing venom peptide proteolysis, we generated P4–P4' motifs from all mature peptides (≥8 aa). This was achieved by concatenating the residues immediately flanking the predicted cleavage sites (P4–P1 and P1'–P4'). For example, the peptide AAPPEFLK yields the motif EFLKAAPP. This analytical strategy follows established approaches used in venom peptidomics [24,27]. Amino acid frequency patterns of the motifs were visualized as sequence logos using the ggseqlogo R package implemented in RStudio (version 2025.051; R version 4.5.1).

2.9. In silico prediction of peptide bioactivities

Peptides from the in-house venom databases were evaluated for *in silico* bioactivity. General bioactivity was first predicted using PeptideRanker (https://peptide.ucd.ie/peptideranker/), and sequences with scores ≥0.5 were retained as putative bioactive candidates. To further investigate specific biological functions, these candidates were analyzed using the BIOPEP-UWM database of bioactive peptides (https://biochemia.uwm.edu.pl/biopep-uwm/) [28], which contains literature-reported bioactive peptides (n = 4765, as of December 2023). This platform was selected to enable a broad and integrative initial overview of potential bioactivities within a single analytical framework. The frequency of bioactive fragments within each peptide (A-value; A = a/N, where a is the number of activity-related fragments and N is the total number of residues) was calculated, and peptides with A-values ≥0.8 were considered to possess potential bioactivity. These bioactive candidates were further evaluated for safety, including toxicity, hemolytic potential, and allergenicity. Toxicity was predicted using the ToxinPred server (https://webs.iiitd.edu.in/raghava/toxinpred/index.html; accessed August 16, 2025). Hemolytic potential was assessed with HemoPI (https://webs.iiitd.edu.in/raghava/hemopi/; accessed August 16, 2025), where the PROB score represents the normalized SVM output ranging from 0 (very unlikely to be hemolytic) to 1 (very likely to be hemolytic). Allergenic potential was predicted using the AlgPred online tool (https://webs.iiitd.edu.in/raghava/algpred/submission.html; accessed August 18, 2025), with MEME/MAST motif matching applied as a primary filter to identify non-allergenic peptides.

2.10. Criteria for identifying putative cryptide candidates

Peptides predicted to possess bioactivity were further evaluated to assess their suitability as putative cryptide candidates based on parent protein origin and sequence reliability. Each peptide sequence was mapped to its corresponding precursor protein, and amino acid assignments were carefully examined for both physiological plausibility and mass spectrometry detectability. Particular attention was given to residue pairs that cannot be unambiguously distinguished by tandem mass spectrometry, such as isoleucine and leucine. Peptides lacking identifiable precursor proteins or exhibiting ambiguous *de novo* sequencing assignments were excluded. The novelty of the retained peptide candidates was evaluated as described in Section 2.6 (*Peptidomics analysis*).

2.11. Data visualization and data availability

Graphical representations were generated using RStudio (version 2025.051; R version 4.5.1) and GraphPad Prism 8.0 (GraphPad Software, San Diego, CA, USA). The mass spectrometry data have been deposited to the ProteomeXchange Consortium *via* the PRIDE partner

repository with the dataset identifier PXD074345.

3. Results

3.1. Purification of protein and peptide fractions from crude extracts

3.1.1. V. mandarinia (Vm) venom

The cation exchange chromatographic elution profile and

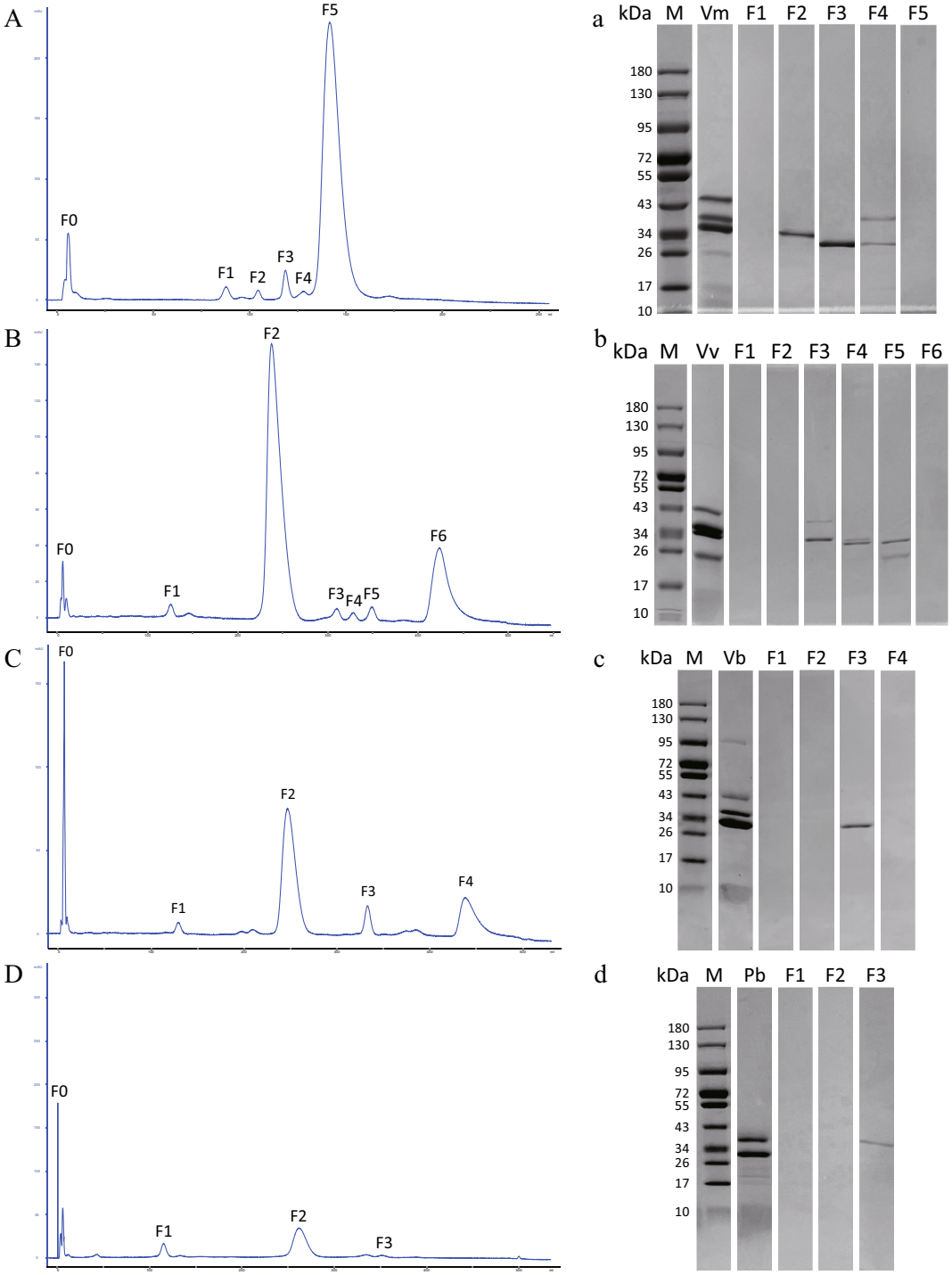

Fig. 1. Cation exchange chromatography elution profiles (A–D) and corresponding SDS-PAGE analyses (a–d) of four wasp venom samples. Venom samples A–D (a–d) represent: A, a – *Vespa mandarinia; B, b – V. velutina; C, c – V. basalis; D, d – Provespa barthelemyi. (A–D) Chromatographic profiles of crude venom extracts separated by cation exchange chromatography using a Toyopearl GigaCap S-650S column. Absorbance at 280 nm was monitored, and bound proteins were eluted using a stepwise NaCl gradient (0–1 M) in 50 mM acetate buffer (pH 5.0) at a flow rate of 1.0 mL/min. A total of five (A), six (B), four (C), and three (D) main protein fractions were collected from V. mandarinia, V. velutina, V. basalis, and P. barthelemyi, respectively. (a–d) SDS-PAGE analysis (10% gel, non-reducing conditions) of crude venom extracts and their corresponding eluted fractions. Lanes: M, molecular weight markers; Vm, crude venom extract from V. mandarinia; Vv, crude venom extract from V. velutina; Vb, crude venom extract from V. basalis; Pb, crude venom extract from P. barthelemyi; F1–F6, eluted fractions F1 to F6 (number of fractions varies by sample).—*

corresponding SDS-PAGE analysis of the crude *V. mandarinia* venom are presented in Fig. 1 A. Five distinct fractions (F1–F5) were obtained (Fig. 1A), respectively accounting for 1.65%, 0.91%, 2.62%, 0.51%, and 87.41% of the total protein content. These fractions were analyzed by 10% SDS-PAGE under non-reducing conditions (Fig. 1a). Mass spectrometry analysis (Table 1) identified probable phospholipase A1 (PLA1) magnifin isoform X2 and PLA1 1 as the predominant components in fractions F2 and F3, respectively. Fraction F4 mainly contained hyaluronidase A and PLA1 1 (Table 1). The poor SDS-PAGE resolution of peptides in fractions F1 and F5, particularly those below 10 kDa, is attributed to diffusion during electrophoresis or staining. This indicates that these fractions are enriched in peptides that are poorly retained or visualized by conventional SDS-PAGE. Given the high abundance of F5, this fraction was selected for further processing, including sequential membrane filtration, desalting, and additional purification for peptidomics.

3.1.2. *V. velutina (Vv) venom*

The crude *V. velutina* venom yielded six fractions (F1–F6) following cation exchange chromatography (Fig. 1B), representing 1.23%, 69.08%, 1.48%, 0.83%, 1.58%, and 21.30% of the total protein content, respectively. These fractions were separated and analyzed by 10% SDS-PAGE under non-reducing conditions (Fig. 1b). Mass spectrometry analysis (Table 1) revealed that fraction F3 primarily contained hyaluronidase A and PLA1, while F4 was predominantly composed of PLA1 1. Fraction F5 included both PLA1 1 and venom allergen 5. As with Vm venom, several peptides smaller than 10 kDa were poorly resolved in SDS-PAGE. F1, F2, and F6 are presumed to consist mainly of small peptides undetectable by gel electrophoresis. Fractions F2 and F6, given their high abundance, were selected for further peptidomics analysis following membrane filtration and desalting.

3.1.3. *V. basalis (Vb) venom*

The elution profile of crude *V. basalis* venom revealed four major fractions (F1–F4) (Fig. 1C), contributing 2.07%, 55.42%, 6.00%, and 18.32% of the total protein content, respectively. SDS-PAGE analysis was performed under non-reducing conditions (Fig. 1c). Fraction F3 was found to be enriched in PLA1 based on mass spectrometry data (Table 1). Peptides with molecular weights below 10 kDa were not clearly resolved in SDS-PAGE. Therefore, F1, F2, and F4 are likely to contain predominantly small peptides. Due to their relative abundance, fractions F2 and F4 were further purified and processed for downstream peptidomics studies.

3.1.4. *Provespa barthelemyi (Pb) venom*

Cation exchange chromatography of crude *Provespa barthelemyi* venom yielded three major fractions (F1–F3), as shown in Fig. 1D. These fractions accounted for 12.09%, 55.19%, and 1.79% of the total protein content, respectively. The fractions were analyzed by 10% SDS-PAGE under non-reducing conditions (Fig. 1d). Mass spectrometry identified PLA1 1 as the major protein in fraction F3 (Table 1). As in previous samples, fractions F1 and F2 exhibited poor SDS-PAGE resolution for low-molecular-weight peptides, suggesting the predominance of small peptide components. Given its high abundance, F2 was selected for further peptide-focused purification and analysis.

All four wasp venoms were successfully fractionated using a one-step AKTA purification strategy. Distinct elution patterns were observed among species, reflecting interspecific variation in venom composition. Pure PLA1-related proteins were electrophoretically isolated from the venoms of *V. mandarinia*, *V. basalis*, and *Provespa barthelemyi*. Two proteins, Probable PLA1 magnifin isoform X2 and PLA1 1, were purified from *V. mandarinia* venom. Mass spectrometry and database searches indicated that both correspond to entries annotated as "predicted from genomic sequence," suggesting that these results represent the first experimental confirmation of these proteins in *V. mandarinia* venom. From *V. basalis* venom, a single electrophoretical pure PLA1 protein was identified. The protein isolated from *Provespa barthelemyi* venom shared high peptide homology with *V. mandarinia* PLA1 1, although no gene or protein record for this enzyme exists in public databases for *Provespa barthelemyi*. Collectively, these results demonstrate the presence of multiple PLA1 isoforms across different wasp venoms and provide a biochemical foundation for subsequent peptidomics analyses.

3.2. *Peptidomics analysis of peptide-containing fractions*

The peptidomics analysis of purified fractions are presented in Fig. 2. A total of 91, 182, 303, and 105 high-confidence peptides (Fig. 2A, Table S1) were respectively identified from the fraction F5 of *V. mandarinia* venom (in Fig. 1A), fractions F2 and F6 of *V. velutina* (in Fig. 1B), fractions F2 and F4 of *V. basalis* (in Fig. 1C), and fraction F2 of *Provespa barthelemyi* (in Fig. 1D). Among these, 76, 161, 274, and 97 peptides respectively from *V. mandarinia* (Vm), *V. velutina* (Vv), *V. basalis* (Vb), and *Provespa barthelemyi* (Pb) venoms, were not recorded in the UniProt peptide database, suggesting that they may represent novel peptides. All identified peptides across the four wasp venoms had molecular weights below 2 kDa (Fig. 2B). For peptides derived from Vm, Vb, and Pb venoms, the 1.5–2.0 kDa range represented the most abundant molecular weight interval, respectively accounting for 45.05%, 43.89%, and 48.57% of peptides. In contrast, approximately 58.24% of peptides from Vv venom fell within the 1.0–1.5 kDa range, indicating a species-specific distribution pattern. Peptide length distributions across the four wasp venoms are shown in Fig. 2C. The identified peptides ranged from 6 to 19 amino acids in Vm, 5 to 18 in Vv, 5 to 20 in Vb, and 6 to 21 in Pb. Despite differences in length ranges, the median peptide

Table 1
Protein identification of SDS–PAGE bands by LC–MS/MS*.

Fractions	Accession NO.	Coverage (%)	Peptides	MW (kDa) /PI	Protein name
Vm F2	XP_035732009.1	48	15	37.6/8.5	Probable phospholipase A1 magnifin isoform X2, OS = *V. mandarinia*
Vm F3	XP_035732026.1	59	23	37/8.53	Phospholipase A1 1, OS = *V. mandarinia*
Vm F4–1	XP_035735901.1	21	10	42.1/9.44	Hyaluronidase A, OS = *V. mandarinia*
Vm F4–2	XP_035732026.1	46	14	37/8.53	Phospholipase A1 1, OS = *V. mandarinia*
Vv F3–1	C0HLL4.1	23	9	40/9.35	Hyaluronidase A, OS = *V. velutina*
Vv F3–2	C0HLL3.1	55	15	33.9/8.84	Phospholipase A1, OS = *V. velutina*
Vv F4–1	XP_047357581.1	41	12	37.1/8.6	Phospholipase A1 1, OS = *V. velutina*
Vv F4–2	XP_047357581.1	48	15	37.1/8.6	Phospholipase A1 1, OS = *V. velutina*
Vv F5–1	XP_047357581.1	43	15	37.1/8.6	Phospholipase A1 1, OS = *V. velutina*
Vv F5–2	P0DMB9.2	50	8	22.7/9.01	Venom allergen 5, OS = *V. velutina*
Vb F3	4QNN_B	46	17	33.2/8.63	Phospholipase A1, OS = *V. basalis*
Pb F3	XP_035732026.1	32	10	37/8.53	Phospholipase A1 1, OS = *V. mandarinia*

Mw, molecular weight; pI, isoelectric point; *Protein identification was performed using a Thermo Orbitrap Fusion mass spectrometer. Vm, *Vespa mandarinia*; Vv, *V. velutina*; Vb, *V. basalis*; Pb, *Provespa barthelemyi*. Protein band labels such as "F4–1" and "F4–2" indicate multiple visible bands observed within a single SDS-PAGE lane corresponding to elution fraction F4. The numbering reflects relative molecular weight in descending order, *i.e.*, F4–1 represents the band with higher molecular weight than F4–2 within the same lane.

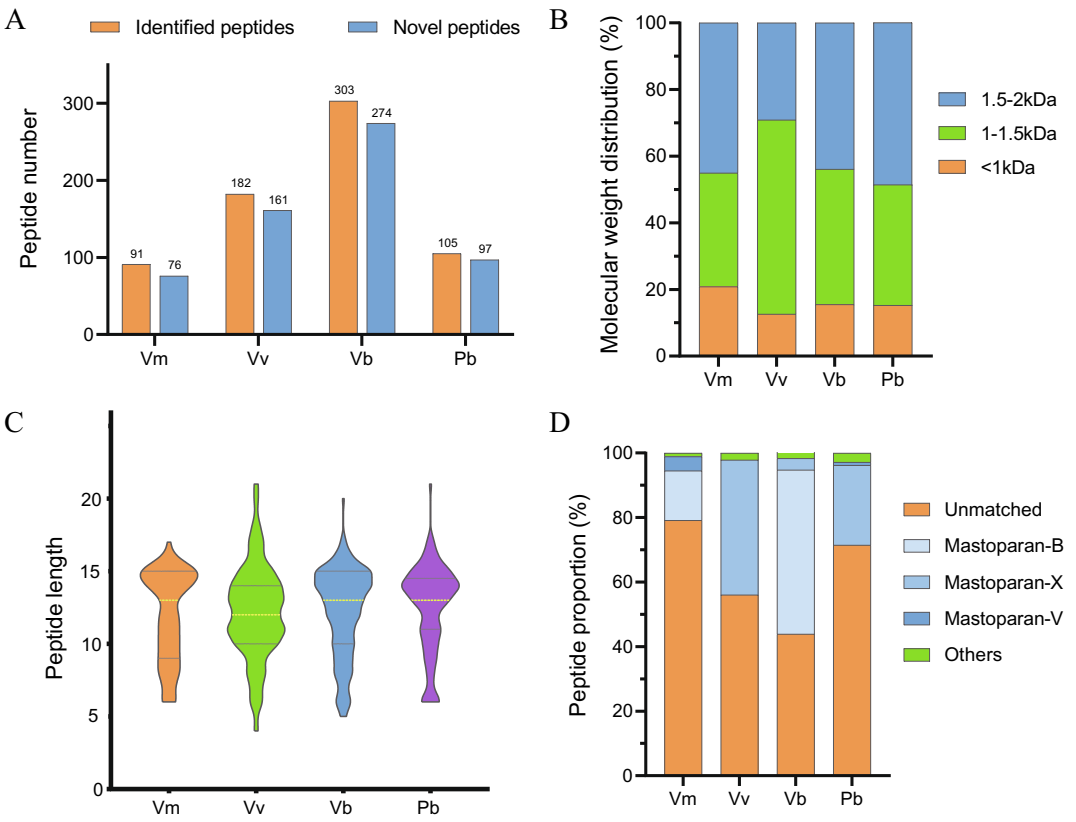

Fig. 2. Peptidomic characterization of venom extracts from four wasp species. (A) Number of peptides identified in each wasp venom. (B) Molecular weight distribution of peptides. All peptides fell below 2 kDa in molecular weight; thus, only three intervals were used for comparative distribution analysis. (C) Peptide length distribution in each venom sample. (D) Protein source classification of the identified peptides based on BLASTP annotation against the UniProt wasp protein database.

lengths were relatively consistent among species, with values of 13, 12, 13, and 13 amino acids for Vm, Vv, Vb, and Pb, respectively. *V. mandarinia*, *V. basalis*, and *Provespa barthelemyi* venom exhibited unimodal and relatively concentrated peptide length distributions, whereas *V. velutina* venom showed a broader profile with potential bimodal characteristics, suggesting greater variability in peptide processing or composition. The protein origins of the identified peptides are presented in Fig. 2D and detailed in Table S1. A large proportion of peptides in all four venoms did not match any known proteins in the UniProt wasp protein database, accounting for 79.12%, 56.04%, 43.89%, and 71.43% in Vm, Vv, Vb, and Pb, respectively. Among the peptides with identifiable protein origins, most were derived from mastoparan-related precursors, including Mastoparan-B, Mastoparan-X, and Mastoparan-V, whereas only a small fraction (1–3%) originated from other protein sources. These results indicate the coexistence of conserved and species-specific components in the wasp venom peptidomes.

3.3. Sequence-level specificity and conservation of venom peptides and peptide families

To evaluate the sequence-level specificity and conservation of venom peptides among four wasp species (*V. mandarinia*, *V. velutina*, *V. basalis*, and *Provespa barthelemyi*), we firstly performed an UpSet analysis based on exact sequence matching (Fig. 3A), enabling direct comparison of fully identical peptides across species. The number of peptides identified in each species was consistent with the overall peptidomic profiling results (Fig. 3A), with 303 in *V. basalis*, 182 in *V. velutina*, 105 in *Provespa barthelemyi*, and 91 in *V. mandarinia*. The vast majority of peptides were species-specific: 279 (92.08%) in *V. basalis*, 162 (89.01%) in *V. velutina*, 89 (84.76%) in *Provespa barthelemyi*, and 74 (81.32%) in *V. mandarinia*. Only two peptides were conserved across all four species, indicating an extremely limited core venom repertoire. A small number of peptides were shared between two or three species, such as 13 between *V. basalis* and *V. mandarinia*, 9 between *V. velutina* and *Provespa barthelemyi*, and 4 between *V. basalis* and *V. velutina*. These findings highlight the high degree of species specificity and compositional diversity in wasp venom peptidomes.

To further investigate whether peptides with similar sequences might be conserved across species, we performed peptide clustering at 90% sequence identity using MMseqs2, and generated an additional UpSet plot (Fig. 3B). This approach grouped highly similar peptides into putative families for comparative analysis, potentially reflecting shared structural or functional properties. As a result, 300, 175, 105, and 91 peptide families were detected in *V. basalis*, *V. velutina*, *Provespa barthelemyi*, and *V. mandarinia*, respectively. Despite this relaxed threshold, the majority of families remained species-specific: 276 (92.00%), 155 (88.57%), 89 (84.76%), and 74 (81.32%) in each species, respectively. These values exhibited negligible deviation from the exact-match analysis, supporting the conclusion that cross-species peptide family overlap is constrained, even when sequence similarity is high. Notably, *V. mandarinia* showed the lowest proportion of species-specific clusters, suggesting slightly greater interspecies similarity. Only two peptide families were shared among all four species, underscoring the narrow extent of conserved venom features. These two shared families are represented by the peptides FVFGCAEALYK (Vm_013) and LNWKGLAAMAKKLL (Vm_067), respectively (Table S1). Taken together, these results demonstrate a strong species-specific signature in wasp venom peptides at both the individual and family level, while revealing a small but noteworthy set of conserved peptide families that may reflect shared evolutionary pressures or ecological adaptations.

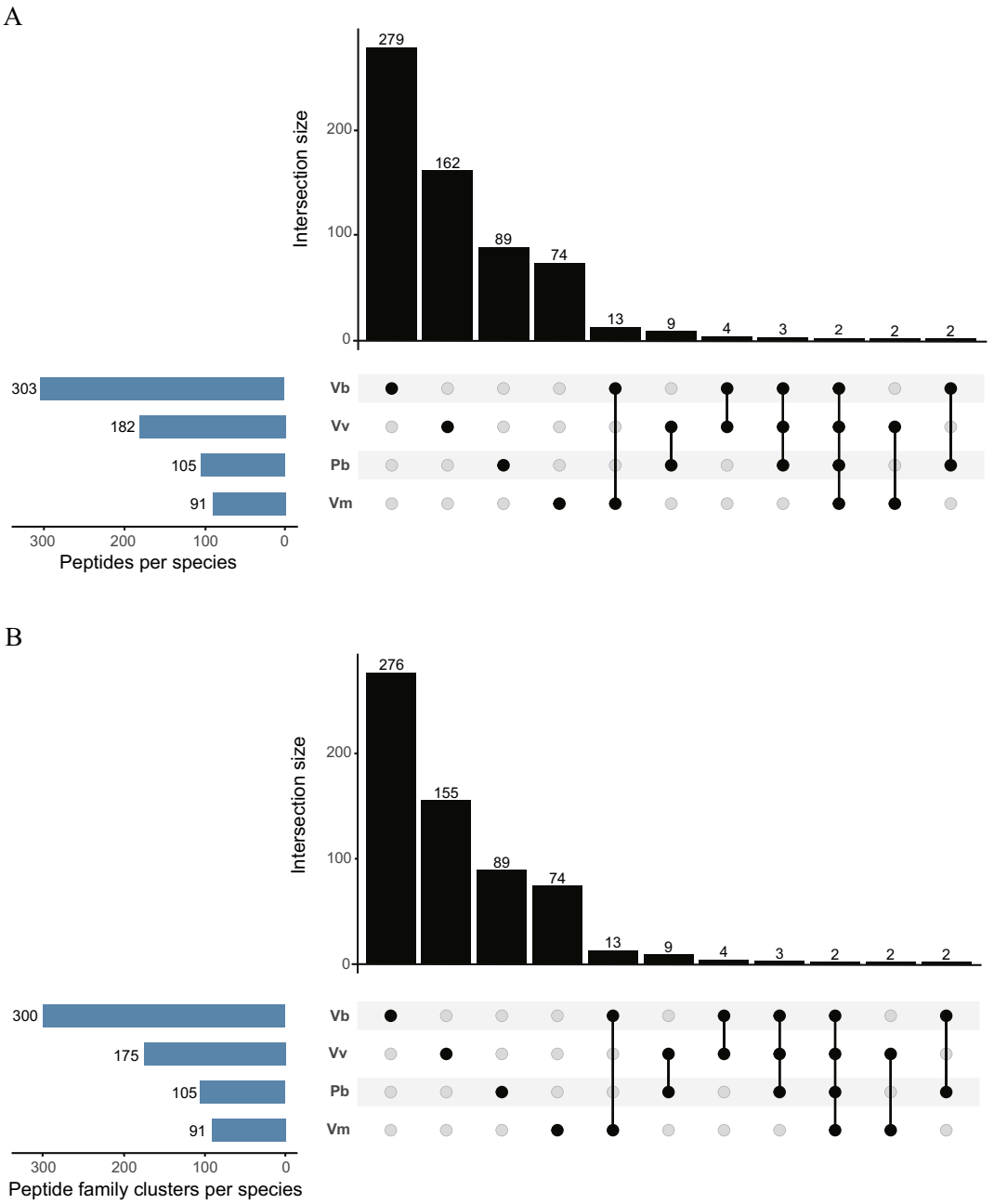

Fig. 3. Comparative analysis of unique and shared peptides (A) and peptide family clusters (B) in the venoms of four wasp species. (A) UpSet plot showing the distribution of unique and shared peptides identified from the venom peptidomes of *V. mandarinia (Vm), V. velutina (Vv), V. basalis (Vb), and Provespa barthelemyi (Pb), based on exact sequence matches. The bar chart on the bottom left indicates the set size, representing the total number of peptides identified per species. The bar chart on the top right shows the intersection size, indicating the number of peptides shared by the species highlighted by black dots in the matrix panel below. (B) UpSet plot showing peptide family clusters identified based on 90% sequence similarity using MMseqs2. The bar chart on the bottom left indicates the set size, representing the total number of peptide family clusters per species. The top right bar chart shows the intersection size, indicating the number of clusters shared by the species highlighted by black dots in the matrix panel. For example, two peptide family clusters are present in all four species.*

3.4. Characterization of P4–P4' proteolytic cleavage motif

To examine proteolytic cleavage preferences, we analyzed amino acid frequencies at the P4–P4' positions for peptides (≥8 amino acids) identified from four wasp venoms (Fig. 4, Supplementary Fig. S1, Supplementary Table S2).

3.4.1. V. mandarinia

Cleavage sites in *V. mandarinia* venom peptides exhibited a distinctly non-random distribution (Fig. 4, Supplementary Table S2). For P1, leucine (L, 30%), arginine (R, 20%), and lysine (K, 18%) were predominant. P1' was similarly enriched in L (39%) and K (26%), with alanine (A) and phenylalanine (F) more frequent than at P1. Upstream positions (P4–P2) were mainly hydrophobic, whereas downstream sites (P2'–P4') included both hydrophobic residues and small polar residues such as A, glycine (G), and serine (S).

3.4.2. V. velutina, V. basalis, and Provespa barthelemyi

In these three species (Fig. 4, Supplementary Table S2), P1/P1' positions were enriched in basic (K, R) and hydrophobic (L, valine (V)) residues, with species-specific patterns. In *V. velutina* venom peptides, L (35%) dominated P1', accompanied by elevated G and asparagine (N)

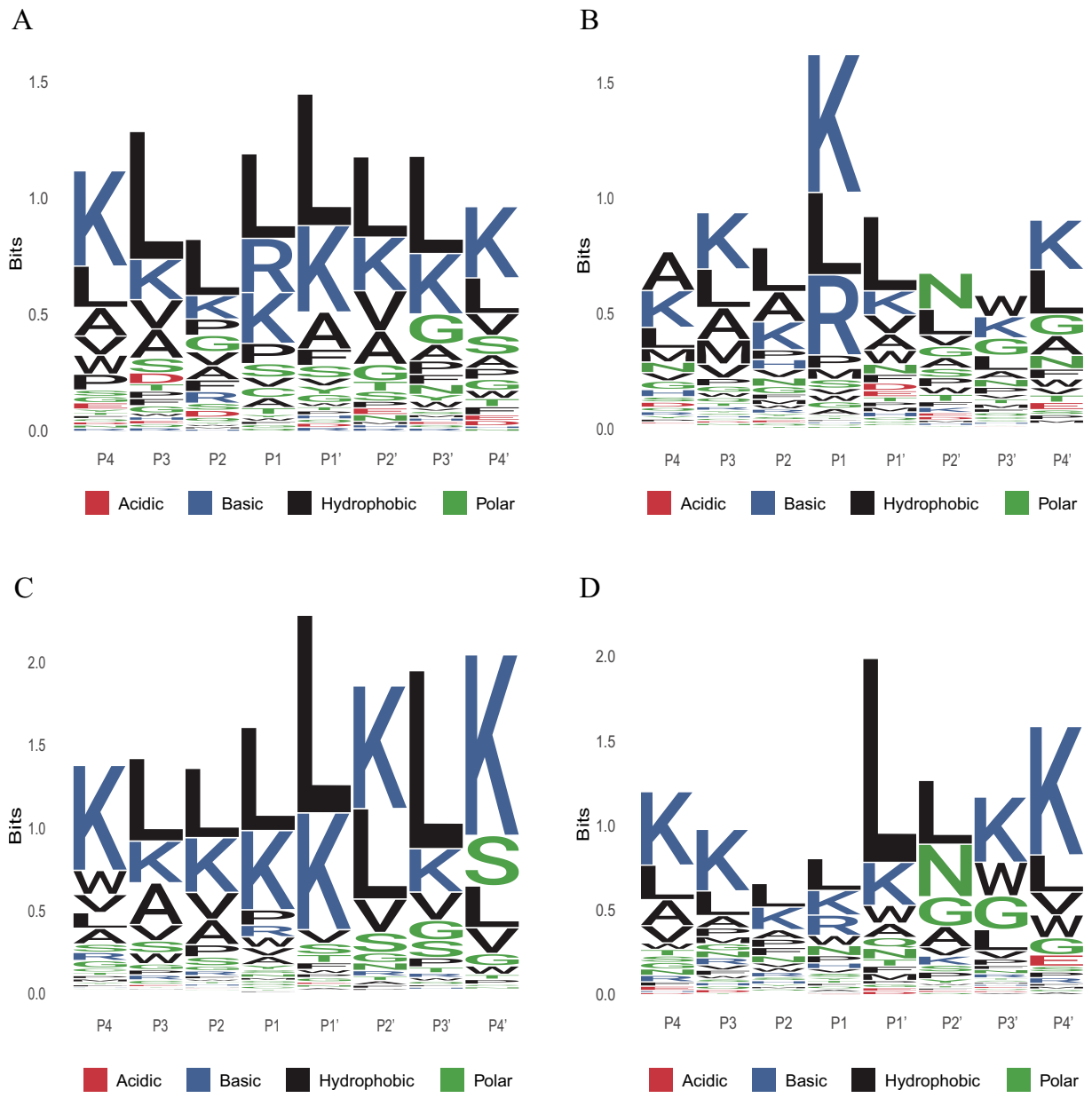

Fig. 4. Sequence logos showing amino acid frequency motifs from positions P4 to P4' in venom-derived peptides of *V. mandarinia* (A), *V. velutina* (B), *V. basalis* (C), and *Provespa barthelemyi* (D). Letter height reflects the relative frequency of each residue at the given position, scaled by information content (bits). The x-axis indicates motif positions (P4–P4'), and the y-axis represents information content. Due to the inability of mass spectrometry to discriminate between isoleucine and leucine, all isoleucine residues are represented as leucine (L) in the sequence logos.

downstream. *V. basalis* venom peptides showed a pronounced L enrichment at P1' (52%) and a higher occurrence of S at downstream sites. *Provespa barthelemyi* venom peptides contained abundant L at P1' (61%), with downstream increases in G and N.

3.4.3. Combined analysis

Across all 619 peptides (≥8 amino acids) from the four venoms, the P1 position was occupied primarily by L (31%) and K (28%), while R was also a prevalent residue (13%) (Supplementary Fig. S1, Supplementary Table S2). P1' was enriched in L (47%) and K (22%). Upstream positions favored hydrophobic residues (L, A, V, tryptophan (W)), whereas downstream sites contained more small polar residues (G, S, N). These results reveal a conserved cleavage motif featuring basic and hydrophobic residues at core positions, alongside modest interspecies variation in downstream polar residue usage.

3.5. Bioactivity in silico prediction for venom-derived peptides

A total of 91, 182, 303, and 105 peptides from *V. mandarinia*, *V. velutina*, *V. basalis*, and *Provespa barthelemyi* venoms were subjected to *in silico* bioactivity assessment. General bioactivity was evaluated using PeptideRanker as an initial screening step. In total, 291 peptides with PeptideRanker scores ≥0.5 were identified as putative bioactive candidates, distributed as follows: 30 from *V. mandarinia*, 88 from *V. velutina*, 108 from *V. basalis*, and 65 from *Provespa barthelemyi* venoms (Fig. 5A; Supplementary Table S3). To further characterize their specific biological activities, the candidate peptides were analyzed using the BIOPEP-UWM database. Bioactivity prediction suggested that the identified peptides were associated with multiple functional categories, such as antibacterial, anti-inflammatory, antioxidative, and neuropeptide-related activities (Fig. 5B–E). Based on the A-value threshold (A ≥ 0.8), the predicted activities retained at this cutoff were angiotensin-

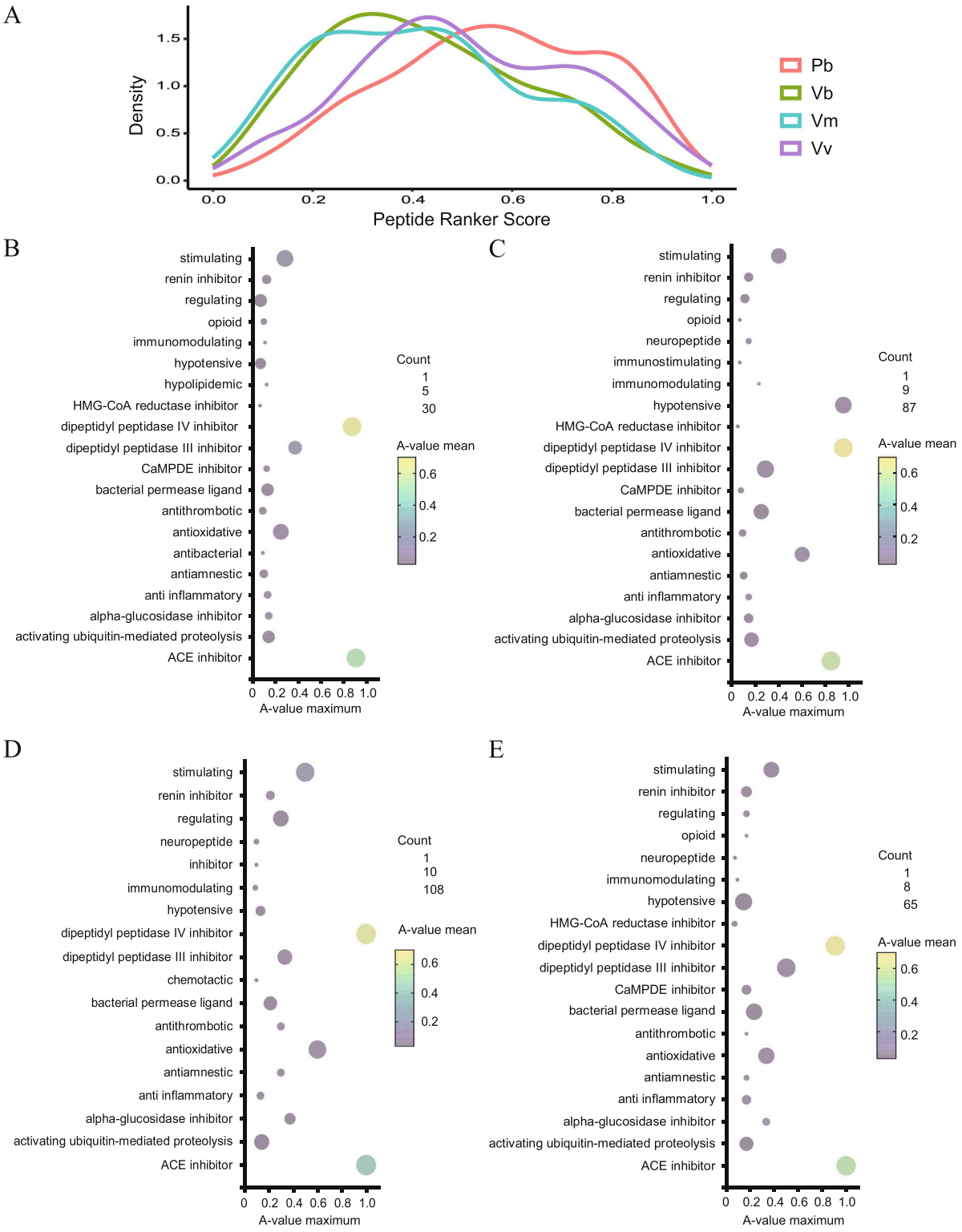

Fig. 5. Predicted bioactivity landscape of wasp venom peptides. (A) Density distributions of PeptideRanker scores for venom peptides from *V. mandarinia (Vm), V. velutina (Vv), V. basalis (Vb), and Provespa barthelemyi (Pb). (B–E) Bubble plots of predicted bioactivity categories for Vm (B), Vv (C), Vb (D), and Pb (E). Bubble size indicates the number of peptides assigned to each category, and bubble color represents the mean A-value of peptides within that category.*

converting enzyme (ACE) inhibition, dipeptidyl peptidase IV (DPP4) inhibition, and hypotensive effects (Supplementary Table S4). Peptides with predicted bioactivity (A-value ≥0.8) were further evaluated for safety, including toxicity, hemolytic potential, and allergenicity (Supplementary Table S5). Eight peptides with ACE-inhibitory potential were identified as predicted to be non-toxic, non-hemolytic, and non-allergenic (Table 2). Similarly, 17 peptides with DPP4-inhibitory potential fulfilled the predicted non-toxic, non-hemolytic, and non-allergenic criteria (Table 3). Together, these predicted non-toxic, non-hemolytic, and non-allergenic peptides with ACE- and/or DPP4-inhibitory potential represent candidates for future functional validation and potential drug discovery applications.

Table 2
Predicted non-toxic, non-hemolytic, and non-allergenic ACE-inhibitory peptides from wasp venoms.

Peptide ID	Peptide sequence	Species	Length	PeptideRanker score	A-value	Toxicity	Hemolysis	Allergenicity
Vb_014	GSGGGGGGLGSGEKR	Vb	15	0.65899	1	Non	0.49	Non
Vm_082	SSYVLGVPGRP	Vm	11	0.554323	0.9091	Non	0.49	Non
Vv_070	LGGWKGLAPGAKK	Vv	13	0.700802	0.8462	Non	0.49	Non
Vb_099	KRPPGFTPFRLW	Vb	12	0.956529	0.8333	Non	0.49	Non
Vb_098	KRPPGFTLPFR	Vb	11	0.89885	0.8182	Non	0.49	Non
Pb_093	QKRPEGAFPFR	Pb	11	0.768945	0.8182	Non	0.49	Non
Vb_017	KAVKLKGWSLKLKLPA	Vb	16	0.745363	0.8125	Non	0.47	Non
Vb_114, Vm_048	LGGWKGLAAFKAKKP	Vb, Vm	15	0.820258	0.8	Non	0.46	Non

Species are abbreviated as follows: *V. mandarinia* (Vm), *V. velutina* (Vv), *V. basalis* (Vb), and *Provespa barthelemyi* (Pb). PeptideRanker score reflects the likelihood of a peptide being bioactive. The A-value denotes the frequency of ACE-inhibitory fragments within a peptide sequence, as calculated using the BIOPEP-UWM database. Toxicity, hemolytic potential, and allergenicity were predicted using ToxinPred2, HemoPred, and AllerTOP, respectively. Hemolysis values are presented as prediction scores (0–1); scores <0.5 are interpreted as non-hemolytic.

Table 3
Predicted non-toxic, non-hemolytic, and non-allergenic DPP4-inhibitory peptides from wasp venoms.

Peptide ID	Peptide sequence	Species	Length	PeptideRanker score	A-value	Toxicity	Hemolysis	Allergenicity
Vb_098	KRPPGFTLPFR	Vb	11	0.89885	1	Non	0.49	Non
Vb_099	KRPPGFTPFRLW	Vb	12	0.956529	0.9167	Non	0.49	Non
Pb_018	KAPVPLAKKFL	Pb	11	0.815208	0.9091	Non	0.49	Non
Vv_170	WGPAGLAAMAK	Vv	11	0.789403	0.9091	Non	0.48	Non
Vv_080	LLGGEPTAAALAHRPWK	Vv	17	0.563777	0.8824	Non	0.49	Non
Vb_261	PPLVTLKKVKGPLLKAS	Vb	17	0.540995	0.8824	Non	0.48	Non
Vb_240	LVGLKSLVSWAKKVL	Vb	15	0.56363	0.8667	Non	0.48	Non
Pb_068	LNWKGLAAFAKKPK	Pb	14	0.622539	0.8571	Non	0.49	Non
Pb_069	LNWKGLAAFKKAPK	Pb	14	0.620752	0.8571	Non	0.49	Non
Vv_070	LGGWKGLAPGAKK	Vv	13	0.700802	0.8462	Non	0.49	Non
Vb_022	KKPKALPSWAKK	Vb	12	0.708863	0.8333	Non	0.48	Non
Vv_033	FPPTPLESLVVAKASLPT	Vv	18	0.69374	0.8333	Non	0.49	Non
Vb_232	LSKKGWLVKKAL	Vb	12	0.537241	0.8333	Non	0.44	Non
Vb_023	KKPKSLVSWAKK	Vb	12	0.504754	0.8333	Non	0.49	Non
Vb_248	LVKKAVVLVSWAKLLL	Vb	16	0.561746	0.8125	Non	0.49	Non
Vb_115	LGGWKGLAAFKKAPK	Vb	15	0.808432	0.8	Non	0.46	Non
Vb_113	LGGWKGLAAAFKKPK	Vb	15	0.800076	0.8	Non	0.46	Non

Species are abbreviated as follows: *V. mandarinia* (Vm), *V. velutina* (Vv), *V. basalis* (Vb), and *Provespa barthelemyi* (Pb). PeptideRanker Score reflects the likelihood of a peptide being bioactive. The A-value denotes the frequency of DPP4-inhibitory fragments within a peptide sequence, as calculated using the BIOPEP-UWM database. Toxicity, hemolytic potential, and allergenicity were predicted using ToxinPred2, HemoPred, and AllerTOP, respectively. Hemolysis values are presented as prediction scores (0–1); scores <0.5 are interpreted as non-hemolytic.

3.6. Cryptide candidates with predicted bioactivities

According to the classification by Autelitano et al. (2007), cryptides, functional sequences encrypted within proteins, comprise three categories [29,30]: **Type I**, natural proteolytic fragments with novel bioactivities unrelated to the parent protein; **Type II**, natural proteolytic fragments with similar bioactivity to the parent protein, though potentially modified; and **Type III**, protein–peptide fragments generated *in vitro* with novel bioactivities, regardless of whether they are produced *in vivo*.

Following the prediction of bioactivity and safety, peptides were screened for identifiable parent proteins and reliable sequences, leading to the exclusion of those with ambiguous *de novo* sequences or no clear precursor. From the initial 25 predicted safe peptides (8 with ACE-inhibitory potential and 17 with DPP4-inhibitory potential), 3 peptides with dual predicted activities were counted once, leaving 22 unique sequences for cryptide assessment. Ten of these peptides met the criteria for reliable cryptides, exhibiting both a confirmed parent protein and a reliable sequence (Supplementary Table S6). A search of the UniProtKB database using the Peptide search tool revealed no prior annotation of these sequences as mature venom peptides, corroborating their novel status. Notably, nine peptides embedded within mastoparan prepropeptide sequences were predicted to have inhibitory activity against ACE or DPP4. According to Autelitano's classification, these peptides represent type I cryptides, subject to future experimental confirmation of their biogenesis and bioactivity. For example, the peptide LVGLKSLVSWAKKVL (Vb_240) from *V. basalis* corresponds to a segment within the mastoparan-B precursor (UniProt: P21654; residues 48–59 in the precursor sequence). While mastoparans are classically associated with antimicrobial, mast cell degranulating, and membranolytic activities, *in silico* prediction suggested that Vb_240 may function as a DPP4 inhibitor, indicating a potential functional difference from its parent (mastoparan) activity. In addition, the peptide KRPPGFTPFRLW (Vb_099), which is embedded within the Vespakinin-X sequences (residues 2–11), a bradykinin-like peptide with vasodilatory and hypotensive effects, was predicted to have inhibitory activities against ACE and DPP4. This predicted function aligns with, but is not identical to, the activity of the precursor, leading to its classification as a candidate type II cryptide under Autelitano's framework. For these candidates, their natural proteolytic liberation and predicted bioactivities represent testable hypotheses that await experimental verification.

4. Discussion

PLA1 is widely recognized as a major allergenic component across diverse wasp venoms [31], and in species such as *V. affinis*, it is reported to be highly abundant [32]. In our study, one-step AKTA purification of electrophoretically pure PLA1-like proteins, followed by proteomic confirmation, provides the first protein-level evidence of two PLA1-like isoforms in *V. mandarinia* venom. The occurrence of multiple PLA1 isoforms aligns with findings in other wasps; for instance, at least three isoforms (VesT1.01a, VesT1.01b, and VesT1.02) have been purified from *V. tropica* venom [33], indicating that PLA1 family expansion is a

recurrent feature in wasps. Although functional characterization was not conducted here, previous reports suggest that isoform diversification may enable venom PLA1-like proteins to adopt differentiated functional roles, including variations in enzymatic activity, toxicity, or other biological effects [31,32,34]. Furthermore, the PLA1-like protein purified and identified from *Provespa barthelemyi* venom exhibits high peptide homology to *V. mandarinia* PLA1, providing the first protein-level evidence for a PLA1-like homolog in this species. This cross-genus conservation is consistent with previous studies indicating that venom PLA1 enzymes are evolutionarily conserved across wasp lineages [34,35]. Taken together, our results demonstrate that even the nocturnal and phylogenetically distinct genus *Provespa* retains PLA1 homologs comparable to those of typical *Vespa* species, highlighting the evolutionary stability of this enzyme family and supporting PLA1 as a core component of wasp venoms. The identification of PLA1-like proteins across four wasp species significantly expands the known phylogenetic distribution of venom phospholipases.

Beyond the conserved PLA1-like proteins, our peptidomics analyses revealed numerous previously unreported peptides across the four wasp venoms, highlighting their substantial molecular diversity. Notably, the four venoms exhibited pronounced species-specific peptidomics signatures, with only a small subset of peptides or peptide families shared among them. Such high divergence and limited conservation are consistent with observations in other venomous taxa. In ants, for example, *Myrmicines* exhibit highly heterogeneous venom peptide repertoires and, with the exception of EGF-like peptides, share very few peptide families with toxins from other ant subfamilies, such as *Ponerinae* and *Myrmeciinae* [36]. Similar interspecific divergence has also been documented in snakes, where closely related *Bothrops* species (*e.g.*, *B. cotiara*, *B. fonsecai*, *B. jararaca*) display marked differences in their venom peptide profiles [24]. Comparable trends are observed in funnel-web spiders, with species such as *Hadronyche valida* and *H. infensa* exhibiting substantial heterogeneity in toxin components despite close phylogenetic relatedness [37]. Conversely, several spider lineages, including *Heteropoda davidbowie*, *Poecilotheria formosa*, *Viridasius fasciatus*, and *Latrodectus mactans*, retain highly conserved cysteine-rich peptide families [38], indicating that peptide diversification is not universal across all venom systems. Overall, these comparative data demonstrate that venom peptide peptidomes can vary widely in their degree of species specificity or conservation. Consistent with this broader pattern, the four wasp venoms examined here showed strongly species-specific signatures at both the peptide and peptide family levels. Beyond this species-level divergence, evidence from other venomous taxa further demonstrates that peptide repertoires are highly dynamic and responsive to ecological pressures. In ants, for example, *Neoponera villosa* shows pronounced seasonal and habitat-dependent shifts in venom peptide [21], reflecting changes in prey availability and predator pressure. Similar patterns are observed in mygalomorph spiders, where only ~20% of venom peptides are shared among conspecific individuals, and substantial additional variation can occur even within a single individual over time [37]. Such plasticity has been attributed to genetic background, age, geographic origin, behaviour, and predator–prey interactions [37]. Collectively, these observations illustrate that short, structurally flexible peptides tend to evolve more rapidly than larger, functionally constrained venom proteins. The marked species-specificity in wasp venom peptidomes is primarily driven by rapid peptide evolution and ecological adaptation. Accordingly, our findings suggest that elucidating venom evolution, particularly within the rapidly evolving peptidome, requires consideration of peptide sequence diversity, functional diversification, and the ecological contexts that shape these repertoires.

The cleavage motifs of peptides across the four wasp venoms exhibit distinctly non-random P4–P4' distributions, suggesting that peptide maturation is unlikely to arise from stochastic degradation and may instead reflect regulated proteolytic processes. Despite interspecies variation in peptide repertoires, all venoms consistently showed strong enrichment of leucine (L), lysine (K), and arginine (R) at P1, along with over-representation of L and K at P1', a pattern consistent with preferential cleavage after hydrophobic or basic residues. The pronounced enrichment of L at the P1' position resembles cleavage preferences of hydrophobic-residue–preferring proteases, commonly observed in metalloproteases, including matrix metalloproteinases (MMPs) [39]. In venom systems, most prominently in snake venoms, related metalloproteases such as snake venom metalloproteinases (SVMPs) [40] and hemorrhagic SVMPs [41] have been reported to cleave adjacent to hydrophobic residues. Large proteome-derived substrate screens have consistently shown that L is the dominant P1' determinant in SVMP-mediated peptide processing [40,41]. Outside snake venoms, metalloproteases are also present in other venom systems, including spiders [42], scorpions [43], and parasitoid wasps [44]. However, their functions in these taxa have not yet been experimentally linked to the generation of endogenous venom peptides. The strong enrichment of K and R at P1 is consistent with cleavage patterns of basic-residue–directed serine proteases, particularly trypsin-like serine proteases, which canonically cleave after these residues [45]. In snake venoms, trypsin-like serine proteases are well characterized, and proteome-derived substrate screens have revealed their strong P1 preference for K and R, thereby generating defined peptide products [46]. Within this context, the pronounced R enrichment resembles cleavage preferences of Arg-selective proteases, such as thrombin-like serine proteases, which have been reported to exhibit strict P1-Arg specificity in snake venoms [47]. Beyond snake venoms, serine proteases with trypsin- or thrombin-like activities have also been reported in bee [48] and parasitoid wasp [49] venoms, suggesting that basic-residue–directed proteases are widespread across diverse venom systems. Nevertheless, in these non-snake venoms, their substrate specificities and potential roles in generating endogenous venom peptides remain far less well defined. Taken together, the cleavage profiles across P4–P4' reveal a characteristic architecture: hydrophobic or basic residues (L, K, R) dominate P1/P1', hydrophobic/basic residues are favored upstream (P4–P2), and small residues such as G, S, and N occur more frequently downstream (P2'–P4'). This motif aligns with precursor-processing patterns described in snake venoms, where coordinated metalloprotease–serine protease networks preferentially cleave at hydrophobic or basic P1/P1' residues followed by small downstream residues [22,50,51]. Therefore, we propose that wasp venoms are likely to operate within a comparable proteolytic framework, a view supported by the broad distribution of metalloproteases and serine proteases across *Hymenoptera* venoms [44,49,52–54]. The modest interspecies differences observed here may reflect variation in protease repertoires, catalytic efficiencies, or venom-gland microenvironments. Collectively, these findings support the hypothesis that wasp venoms may employ a conserved yet adaptable proteolytic system to generate diverse, species-specific repertoires of endogenous peptides. Further experimental validation will be required to elucidate the roles of specific proteases in peptide generation.

The predicted bioactivity profiling and cryptide mining analyses collectively suggest that wasp venom peptidomes may contain an additional layer of sequence-embedded functional features, extending their functional scope beyond classical neurotoxic and membrane-interactive activities. Rather than representing nonspecific degradation products, venom-derived peptides may constitute a versatile biochemical repertoire with broader regulatory potential, including enzyme-modulating activities. It should be noted that the bioactivity predictions were derived from the BIOPEP-UWM database, which integrates literature-reported bioactive peptides across multiple functional categories. Due to differences in data availability and annotation depth, certain activity classes, particularly ACE and DPP4 inhibition, are more extensively represented in this platform, which may contribute to their prominence in the prediction results. Accordingly, the present analysis should be interpreted as an initial, integrative screening rather than a comprehensive assessment of all possible bioactivities. It provides a framework for prioritizing candidate peptides and generating testable

hypotheses. Future work will extend this framework by incorporating additional prediction strategies and experimental validation to enable more comprehensive and systematic investigation of peptide functional properties. Notably, the predicted ACE-inhibitory potential is in line with previous preclinical reports showing that apitherapy formulations incorporating bee venom ameliorate hypertension-associated phenotypes in animal models, including improvements in blood pressure and modulation of renin–angiotensin–aldosterone system (RAAS)–associated markers [55]. Similarly, the predicted DPP4-inhibitory potential is in line with experimental studies suggesting that bee venom administration can improve diabetic phenotypes, including reductions in blood glucose levels and modulation of metabolic parameters [56]. Although these effects have thus far been demonstrated primarily in animal models, they collectively suggest a broader association between venom-derived peptides and the regulation of blood pressure and glucose homeostasis. Beyond predicted bioactivities, cryptides represent a particularly underappreciated dimension of venom peptide diversity. While cryptides have been increasingly documented in snake, scorpion, and spider venoms [25,57,58], their systematic investigation in *Hymenoptera* venoms has been limited. Our study therefore provides, to our knowledge, the first integrative, hypothesis-driven peptidomic framework for exploring sequence-embedded cryptide candidates in wasp venoms. Although experimental validation is required to confirm their biogenesis and biological activities, this approach offers accessible starting points for future biochemical, structural, and pharmacological investigations, and further highlights the potential of wasp venoms as reservoirs of bioactive molecules.

5. Conclusion

This study presents a comparative peptidomics framework for the systematic investigation of wasp venoms across multiple species. Our work indicates that wasp venom peptidomes display pronounced, species-specific composition patterns consistent with potentially regulated proteolytic processing, rather than arising from nonspecific protein degradation. These findings suggest that proteolysis may be an important contributor to venom peptide diversification in wasps. Beyond compositional insights, predictive bioactivity profiling suggests that wasp venoms may represent a source of peptides with diverse functional potential, including enzyme-modulating activities relevant to cardiometabolic regulation. The identification of putative cryptide candidates further expands current views of venom complexity by implicating sequence-embedded functional peptides. Collectively, this work positions comparative peptidomics as a powerful strategy for exploring venom-derived peptide diversity and provides a foundation for future studies on peptide biogenesis, activity, and evolutionary significance in *Hymenoptera* venoms.

CRediT authorship contribution statement

Kai Wang: Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization, Writing – original draft. **Jiangtao Qiao:** Validation, Software, Methodology, Data curation. **Loïc Quinton:** Resources, Methodology, Writing – review & editing. **Gauthier Eppe:** Resources, Methodology. **Xianbin Meng:** Software, Data curation. **Guozhong Huang:** Resources, Methodology. **Eric Haubruge:** Software, Resources, Methodology, Conceptualization, Writing – review & editing. **Hongcheng Zhang:** Project administration, Funding acquisition, Conceptualization, Writing – review & editing.

Declaration of competing interest

The authors declare that there is no conflict of interest.

Acknowledgments

This study was funded by the Modern Agro-Industry Technology Research System of the Ministry of Agriculture of P.R. China (CARS-45-KXJ19), the Agricultural Science and Technology Innovation Program of the Ministry of Agriculture of P.R. China (CAAS-ASTIP-2019-IAR), the Major Science and Technology Special Project of Yunnan Province, P.R. China (2018ZF01204), and the China Scholarship Council (CSC; Student ID: 202403250005).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jprot.2026.105676.

Data availability

The mass spectrometry data have been deposited to the ProteomeXchange Consortium *via* the PRIDE partner repository with the dataset identifier PXD074345.

References

[1] M.V. Modica, K. Sunagar, M. Holford, S. Dutertre, Diversity and evolution of animal venoms: neglected targets, ecological interactions, future perspectives, Front. Ecol. Evol. 8 (2020) 65, https://doi.org/10.3389/fevo.2020.00065.
[2] V. Schendel, L.D. Rash, R.A. Jenner, E.A. Undheim, The diversity of venom: the importance of behavior and venom system morphology in understanding its ecology and evolution, Toxins 11 (11) (2019) 666, https://doi.org/10.3390/toxins11110666.
[3] L. Freuville, C. Matthys, L. Quinton, J.-P. Gillet, Venom-derived peptides for breaking through the glass ceiling of drug development, Front. Chem. 12 (2024) 1465459, https://doi.org/10.3389/fchem.2024.1465459.
[4] E. Messadi, Snake venom components as therapeutic drugs in ischemic heart disease, Biomolecules 13 (10) (2023) 1539, https://doi.org/10.3390/biom13101539.
[5] H. Safavi-Hemami, S.E. Brogan, B.M. Olivera, Pain therapeutics from cone snail venoms: from Ziconotide to novel non-opioid pathways, J. Proteome 190 (2019) 12–20, https://doi.org/10.1016/j.jprot.2018.05.009.
[6] Z.-Q. Zhang, Animal Biodiversity: An Outline of Higher-Level Classification and Survey of Taxonomic Richness, Magnolia press, Auckland, New Zealand, 2011.
[7] P.G. Czaikoski, D.L. Menaldo, S. Marcussi, A.L. Baseggio, A.L. Fuly, R.C. Paula, A. U. Quadros, P.R. Romao, M.L. Buschini, F.Q. Cunha, Anticoagulant and fibrinogenolytic properties of the venom of Polybia occidentalis social wasp, Blood Coagul. Fibrinolysis 21 (7) (2010) 653–659, https://doi.org/10.1097/MBC.0b013e32833cea7a.
[8] V. Isidorov, A. Zalewski, G. Zambrowski, I. Swiecicka, Chemical composition and antimicrobial properties of honey bee venom, Molecules 28 (10) (2023) 4135, https://doi.org/10.3390/molecules28104135.
[9] H.R. El-Seedi, N. Eid, A.A. Abd El-Wahed, M.E. Rateb, H.S. Afifi, A.F. Algethami, C. Zhao, Y. Al Naggar, S.M. Alsharif, H.E. Tahir, Honey bee products: preclinical and clinical studies of their anti-inflammatory and immunomodulatory properties, Front. Nutr. 8 (2022) 761267, https://doi.org/10.3389/fnut.2021.761267.
[10] H.S. Kachel, R.N. Patel, H. Franzyk, I.R. Mellor, Block of nicotinic acetylcholine receptors by philanthotoxins is strongly dependent on their subunit composition, Sci. Rep. 6 (1) (2016) 38116, https://doi.org/10.1038/srep38116.
[11] M.Y. Hanafi, E.L. Zaher, S.E. El-Adely, A. Sakr, A.H. Ghobashi, M.H. Hemly, A. H. Kazem, M.A. Kamel, The therapeutic effects of bee venom on some metabolic and antioxidant parameters associated with HFD-induced non-alcoholic fatty liver in rats, Exp. Ther. Med. 15 (6) (2018) 5091–5099, https://doi.org/10.3892/etm.2018.6028.
[12] T.-Y. Lin, C.-L. Hsieh, Clinical applications of bee venom acupoint injection, Toxins 12 (10) (2020) 618, https://doi.org/10.3390/toxins12100618.
[13] S. Jang, K.H. Kim, Clinical effectiveness and adverse events of bee venom therapy: a systematic review of randomized controlled trials, Toxins 12 (9) (2020) 558, https://doi.org/10.3390/toxins12090558.
[14] A. Abd El-Wahed, N. Yosri, H.H. Sakr, M. Du, A.F. Algethami, C. Zhao, A. H. Abdelazeem, H.E. Tahir, S.H. Masry, M.M. Abdel-Daim, Wasp venom biochemical components and their potential in biological applications and nanotechnological interventions, Toxins 13 (3) (2021) 206, https://doi.org/10.3390/toxins13030206.
[15] S. Sharma, R.K. Upadhyay, Hymenoptera toxins: biological activity, pharmaceutical and therapeutic uses, Int J Pharm Pharm Sci 13 (2021). 10.22159/ijpps.2021v13i4.40473.
[16] L. Hemajha, S. Singh, C.A. Biji, A. Balde, S. Benjakul, R.A. Nazeer, A review on inflammation modulating venom proteins/peptide therapeutics and their delivery strategies: a review, Int. Immunopharmacol. 142 (2024) 113130, https://doi.org/10.1016/j.intimp.2024.113130.

[17] D.C. Dallas, A. Guerrero, E.A. Parker, R.C. Robinson, J. Gan, J.B. German, D. Barile, C.B. Lebrilla, Current peptidomics: applications, purification, identification, quantification, and functional analysis, Proteomics 15 (5–6) (2015) 1026–1038, https://doi.org/10.1002/pmic.201400310.
[18] E. Clynen, G. Baggerman, S. Husson, B. Landuyt, L. Schoofs, Peptidomics in drug research, Expert Opin. Drug Discov. 3 (4) (2008) 425–440, https://doi.org/10.1517/17460441.3.4.425.
[19] L.T. Ito, J.G. Miyamoto, S.S. Sant'Anna, K.F. Grego, A.M. Tanaka-Azevedo, A. K. Tashima, Unveiling the peptidome diversity of *Lachesis muta* snake venom: discovery of novel fragments of metalloproteinase, L-amino acid oxidase, and bradykinin potentiating peptides, Biochem. Biophys. Res. Commun. 683 (2023) 149090, https://doi.org/10.1016/j.bbrc.2023.10.022.
[20] L. Zeng, C. Zhang, M. Yang, J. Sun, J. Lu, H. Zhang, J. Qin, W. Zhang, Z. Jiang, Unveiling the diversity and modifications of short peptides in Buthus martensii scorpion venom through liquid chromatography-high resolution mass spectrometry, Toxins 16 (3) (2024) 155, https://doi.org/10.3390/toxins16030155.
[21] C.T. Cologna, R.S. Rodrigues, J. Santos, E. De Pauw, E.C. Arantes, L. Quinton, Peptidomic investigation of Neoponera villosa venom by high-resolution mass spectrometry: seasonal and nesting habitat variations, J. Venom. Anim. Toxins Incl. Trop. Dis. 24 (1) (2018) 6, https://doi.org/10.1186/s40409-018-0141-3.
[22] J.A. Portes-Junior, N. Yamanouye, S.M. Carneiro, P.S. Knittel, S.v.S. Sant'Anna, F. C. Nogueira, M. Junqueira, G.S. Magalhães, G.B. Domont, A.M. Moura-da-Silva, Unraveling the processing and activation of snake venom metalloproteinases, J. Proteome Res. 13 (7) (2014) 3338–3348. https://doi.org/10.1021/pr500185a.
[23] C.A. Nicolau, P.C. Carvalho, I.L. Junqueira-de-Azevedo, A. Teixeira-Ferreira, M. Junqueira, J. Perales, A.G.C. Neves-Ferreira, R.H. Valente, An in-depth snake venom proteopeptidome characterization: benchmarking *Bothrops jararaca*, J. Proteome 151 (2017) 214–231. https://doi.org/10.1016/j.jprot.2016.06.029.
[24] A.K. Tashima, A. Zelanis, E.S. Kitano, D. Ianzer, R.L. Melo, V. Rioli, S.S. Sant'anna, A.C. Schenberg, A.C. Camargo, S.M. Serrano, Peptidomics of three Bothrops snake venoms: insights into the molecular diversification of proteomes and peptidomes, Mol. Cell. Proteomics 11 (11) (2012) 1245–1262, https://doi.org/10.1074/mcp.M112.019331.
[25] J.G. Miyamoto, E.S. Kitano, A. Zelanis, P.G. Nachtigall, I. Junqueira-de-Azevedo, S. S. Sant'Anna, R.L. da Silva, P.A. Bersanetti, A.K. Carmona, P.J.B. Pereira, A novel metalloproteinase-derived cryptide from *Bothrops cotiara* venom inhibits angiotensin-converting enzyme activity, Biochimie 216 (2024) 90–98, https://doi.org/10.1016/j.biochi.2023.10.010.
[26] A. Lex, N. Gehlenborg, H. Strobelt, R. Vuillemot, H. Pfister, UpSet: visualization of intersecting sets, IEEE Trans. Vis. Comput. Graph. 20 (12) (2014) 1983–1992, https://doi.org/10.1109/TVCG.2014.2346248.
[27] O. Schilling, C.M. Overall, Proteome-derived, database-searchable peptide libraries for identifying protease cleavage sites, Nat. Biotechnol. 26 (6) (2008) 685–694, https://doi.org/10.1038/nbt1408.
[28] P. Minkiewicz, A. Iwaniak, M. Darewicz, BIOPEP-UWM database of bioactive peptides: current opportunities, Int. J. Mol. Sci. 20 (23) (2019) 5978, https://doi.org/10.3390/ijms20235978.
[29] N. Ueki, K. Someya, Y. Matsuo, K. Wakamatsu, H. Mukai, Cryptides: functional cryptic peptides hidden in protein structures, Pept. Sci. 88 (2) (2007) 190–198, https://doi.org/10.1002/bip.20687.
[30] H. Mukai, T. Seki, H. Nakano, Y. Hokari, T. Takao, M. Kawanami, H. Tsukagoshi, H. Kimura, Y. Kiso, Y. Shimonishi, Mitocryptide-2: purification, identification, and characterization of a novel cryptide that activates neutrophils, J. Immunol. 182 (8) (2009) 5072–5080, https://doi.org/10.4049/jimmunol.0802965.
[31] A. Perez-Riverol, A.M. Lasa, J.R.A. dos Santos-Pinto, M.S. Palma, Insect venom phospholipases A1 and A2: roles in the envenoming process and allergy, Insect Biochem. Mol. Biol. 105 (2019) 10–24, https://doi.org/10.1016/j.ibmb.2018.12.011.
[32] S. Sukprasert, P. Rungsa, N. Uawonggul, P. Incamnoi, S. Thammasirirak, J. Daduang, S. Daduang, Purification and structural characterisation of phospholipase A1 (Vespapase, Ves a 1) from Thai banded tiger wasp (Vespa affinis) venom, Toxicon 61 (2013) 151–164. https://doi.org/10.1016/j.toxicon.2012.10.024.
[33] P. Rungsa, S. Peigneur, S. Daduang, J. Tytgat, Purification and biochemical characterization of VesT1s, a novel phospholipase A1 isoform isolated from the venom of the greater banded wasp Vespa tropica, Toxicon 148 (2018) 74–84, https://doi.org/10.1016/j.toxicon.2018.03.015.
[34] H. Yang, X. Xu, D. Ma, K. Zhang, R. Lai, A phospholipase A1 platelet activator from the wasp venom of Vespa magnifica (smith), Toxicon 51 (2) (2008) 289–296, https://doi.org/10.1016/j.toxicon.2007.10.003.
[35] L. Luo, P.M. Kamau, R. Lai, Bioactive peptides and proteins from wasp venoms, Biomolecules 12 (4) (2022) 527, https://doi.org/10.3390/biom12040527.
[36] V. Barassé, N. Téné, C. Klopp, F. Paquet, N. Tysklind, V. Troispoux, H. Lalägue, J. Orivel, B. Lefranc, J. Leprince, Venomics survey of six myrmicine ants provides insights into the molecular and structural diversity of their peptide toxins, Insect Biochem. Mol. Biol. 151 (2022) 103876, https://doi.org/10.1016/j.ibmb.2022.103876.
[37] L.H. Duran, T.L. Rymer, D.T. Wilson, Variation in venom composition in the Australian funnel-web spiders *Hadronyche valida*, Toxicon X 8 (2020) 100063, https://doi.org/10.1016/j.toxcx.2020.100063.
[38] V. Oldrati, D. Koua, P.-M. Allard, N. Hulo, M. Arrell, W. Nentwig, F. Lisacek, J.-L. Wolfender, L. Kuhn-Nentwig, R. Stöcklin, Peptidomic and transcriptomic profiling of four distinct spider venoms, PLoS One 12 (3) (2017) e0172966, https://doi.org/10.1371/journal.pone.0172966.
[39] G. Murphy, H. Nagase, Progress in matrix metalloproteinase research, Mol. Asp. Med. 29 (5) (2008) 290–308, https://doi.org/10.1016/j.mam.2008.05.002.
[40] A.F.P. Leme, T. Escalante, J.G. Pereira, A.K. Oliveira, E.F. Sanchez, J.M. Gutiérrez, S.M. Serrano, J.W. Fox, High resolution analysis of snake venom metalloproteinase (SVMP) peptide bond cleavage specificity using proteome based peptide libraries and mass spectrometry, J. Proteome 74 (4) (2011) 401–410, https://doi.org/10.1016/j.jprot.2010.12.002.
[41] L. Bertholim, A. Zelanis, A.K. Oliveira, S.M. Serrano, Proteome-derived peptide library for the elucidation of the cleavage specificity of HF3, a snake venom metalloproteinase, Amino Acids 48 (5) (2016) 1331–1335, https://doi.org/10.1007/s00726-016-2218-z.
[42] D. Trevisan-Silva, A.V. Bednaski, L.H. Gremski, O.M. Chaim, S.S. Veiga, A. Senff-Ribeiro, Differential metalloprotease content and activity of three Loxosceles spider venoms revealed using two-dimensional electrophoresis approaches, Toxicon 76 (2013) 11–22, https://doi.org/10.1016/j.toxicon.2013.08.065.
[43] K.C. Bordon, G.C. Santos, J.G. Martins, G.A. Wiezel, F.G. Amorim, T. Crasset, D. Redureau, L. Quinton, R.E. Procópio, E.C. Arantes, Pioneering comparative proteomic and enzymatic profiling of Amazonian scorpion venoms enables the isolation of their first α-Ktx, metalloprotease, and phospholipase A2, Toxins 17 (8) (2025) 411, https://doi.org/10.3390/toxins17080411.
[44] S.J. Moreau, S. Asgari, Venom proteins from parasitoid wasps and their biological functions, Toxins 7 (7) (2015) 2385–2412, https://doi.org/10.3390/toxins7072385.
[45] L. Hedstrom, Serine protease mechanism and specificity, Chem. Rev. 102 (12) (2002) 4501–4524, https://doi.org/10.1021/cr000033x.
[46] A. Zelanis, P.F. Huesgen, A.K. Oliveira, A.K. Tashima, S.M. Serrano, C.M. Overall, Snake venom serine proteinases specificity mapping by proteomic identification of cleavage sites, J. Proteome 113 (2015) 260–267, https://doi.org/10.1016/j.jprot.2014.10.002.
[47] A. Ullah, R. Masood, I. Ali, K. Ullah, H. Ali, H. Akbar, C. Betzel, Thrombin-like enzymes from snake venom: structural characterization and mechanism of action, Int. J. Biol. Macromol. 114 (2018) 788–811, https://doi.org/10.1016/j.ijbiomac.2018.03.164.
[48] Y.M. Choo, K.S. Lee, H.J. Yoon, B.Y. Kim, M.R. Sohn, J.Y. Roh, Y.H. Je, N.J. Kim, I. Kim, S.D. Woo, Dual function of a bee venom serine protease: prophenoloxidase-activating factor in arthropods and fibrin (ogen) olytic enzyme in mammals, PLoS One 5 (5) (2010) e10393, https://doi.org/10.1371/journal.pone.0010393.
[49] A. Becchimanzi, M. Avolio, H. Bostan, C. Colantuono, F. Cozzolino, D. Mancini, M. L. Chiusano, P. Pucci, S. Caccia, F. Pennacchio, Venomics of the ectoparasitoid wasp Bracon nigricans, BMC Genomics 21 (1) (2020) 34, https://doi.org/10.1186/s12864-019-6396-4.
[50] K. Kalogeropoulos, A.F. Treschow, U. auf dem Keller, T. Escalante, A. Rucavado, J. M. Gutiérrez, A.H. Laustsen, C.T. Workman, Protease activity profiling of snake venoms using high-throughput peptide screening, Toxins 11 (3) (2019) 170, https://doi.org/10.3390/toxins11030170.
[51] A.M. Moura-da-Silva, M.T. Almeida, J.A. Portes-Junior, C.A. Nicolau, F. Gomes-Neto, R.H. Valente, Processing of snake venom metalloproteinases: generation of toxin diversity and enzyme inactivation, Toxins 8 (6) (2016) 183, https://doi.org/10.3390/toxins8060183.
[52] D. Dashevsky, K. Baumann, E.A. Undheim, A. Nouwens, M.P. Ikonomopoulou, J. O. Schmidt, L. Ge, H.F. Kwok, J. Rodriguez, B.G. Fry, Functional and proteomic insights into aculeata venoms, Toxins 15 (3) (2023) 224, https://doi.org/10.3390/toxins15030224.
[53] J.-L. Gatti, M. Belghazi, F. Legeai, M. Ravallec, M. Frayssinet, S. Robin, D. Aboubakar-Souna, R. Srinivasan, M. Tamo, M. Poirié, Proteo-trancriptomic analyses reveal a large expansion of metalloprotease-like proteins in atypical venom vesicles of the wasp Meteorus pulchricornis (Braconidae), Toxins 13 (7) (2021) 502, https://doi.org/10.3390/toxins13070502.
[54] L.D. dos Santos, K.S. Santos, J.R.A. Pinto, N.B. Dias, B.M.d. Souza, M.F. dos Santos, J. Perales, G.B. Domont, F.M. Castro, J.E. Kalil, Profiling the proteome of the venom from the social wasp *Polybia paulista*: a clue to understand the envenoming mechanism, J. Proteome Res. 9 (8) (2010) 3867–3877, https://doi.org/10.1021/pr1000829.
[55] F.E.-Z. Abd El-Hakam, G. Abo Laban, S. Badr El-Din, H. Abd El-Hamid, M. H. Farouk, Apitherapy combination improvement of blood pressure, cardiovascular protection, and antioxidant and anti-inflammatory responses in dexamethasone model hypertensive rats, Sci. Rep. 12 (1) (2022) 20765, https://doi.org/10.1038/s41598-022-24727-z.
[56] A.K. Hassan, D.A. El-kotby, M.M. Tawfik, R.E. Badr, I.M. Bahgat, Antidiabetic effect of the Egyptian honey bee (*Apis mellifera*) venom in alloxan-induced diabetic rats, J. Basic Appl. Zool. 80 (1) (2019) 58, https://doi.org/10.1186/s41936-019-0127-x.
[57] C. Rocha-Resende, N.M. Leão, M.E. de Lima, R.A. Santos, A.M. de Castro Pimenta, T. Verano-Braga, Moving pieces in a cryptomic puzzle: Cryptide from Tityus serrulatus Ts3 Nav toxin as potential agonist of muscarinic receptors, Peptides 98 (2017) 70–77, https://doi.org/10.1016/j.peptides.2016.12.018.
[58] R. Guo, G. Guo, A. Wang, G. Xu, R. Lai, H. Jin, Spider-venom peptides: structure, bioactivity, strategy, and research applications, Molecules 29 (1) (2023) 35, https://doi.org/10.3390/molecules29010035.