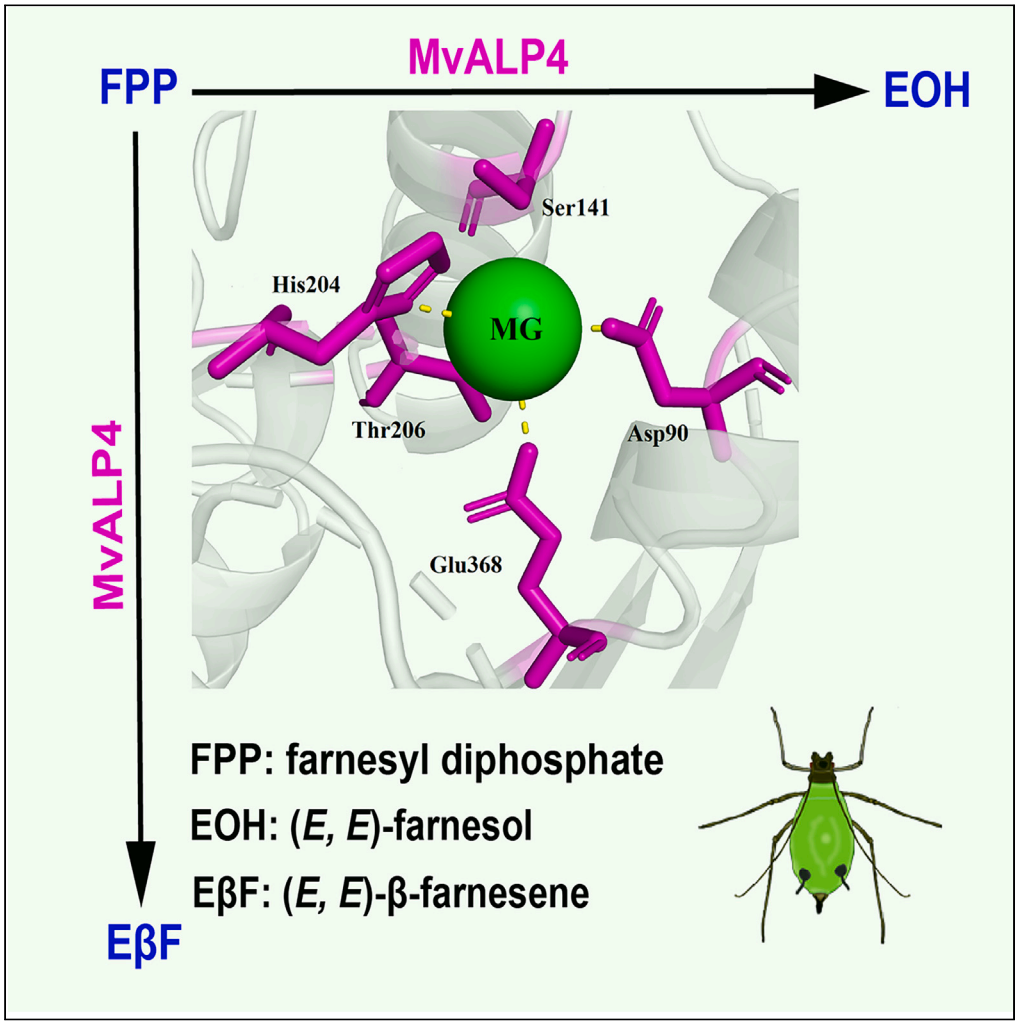
iScience

CellPress
OPEN ACCESS

Article

Functional characterization of alkaline phosphatases involved alarm pheromone in the vetch aphid *Megoura viciae*



Xuan Song, Yao-Guo Qin, Yi-Han Zhang, Yu-Bei Zhou, Dan Chen, Dong-Hai Xie, Zheng-Xi Li

zxli@cau.edu.cn

Highlights

Four ALP genes are identified in the vetch aphid

The binding residues of (*E*, *E*)-FPP are varied in the *MvALP1-4*

The expression patterns of *MvALP1-4* indicated promiscuous functions

One alkaline phosphatase has terpene synthase activity

Song et al., iScience *26*, 108115
November 17, 2023
https://doi.org/10.1016/j.isci.2023.108115

iScience

CellPress
OPEN ACCESS

Article

Functional characterization of alkaline phosphatases involved alarm pheromone in the vetch aphid *Megoura viciae*

Xuan Song,[1] Yao-Guo Qin,[1] Yi-Han Zhang,[1] Yu-Bei Zhou,[1] Dan Chen,[1] Dong-Hai Xie,[1] and Zheng-Xi Li[1,2,*]

SUMMARY

The alkaline phosphatases (ALPs) are highly promiscuous enzymes and have been extensively investigated in mammals for their medical significance, but their functional promiscuity is relatively poorly understood in insects. Here, we first identified four ALP genes (designated as *MvALP1-4*) in the vetch aphid *Megoura viciae* that contained one alkaline phosphatase site, three metal-binding sites, and varied other functional sites. Phylogenetic analysis, molecular docking and the spatiotemporal expression profiling of *MvALP1-4* were very different, indicating a promiscuous functionality. We also found that MvALP4 involved the biosynthesis of aphid alarm pheromones (EβF) *in vitro* and *in vivo*. Finally, transcriptome analysis in the stimulated and unstimulated aphids supported the involvement of MvALPs in the biosynthesis of aphid alarm pheromones. Our study identified a multifunctional ALP involved terpene synthase enzyme activity in the aphid, which contributes to the understanding of the functional plasticity of ALPs in insects.

INTRODUCTION

Alkaline phosphatases (ALPs) are a large family of ubiquitous membrane-bound glycoproteins that catalyzes the hydrolysis of phosphate monoesters at basic pH values.[1] They are present in microorganisms and animal cells but absent from some higher plants.[2–4] Mammalian ALPs are zinc-containing metalloenzymes encoded by a multigene family, which can be divided into four types according to the site of tissue expression, *i.e.*, intestinal ALPs, placental ALPs, germ cell ALPs, and tissue nonspecific ALPs or liver/bone/kidney ALPs.[4] Alignment of amino acid sequences showed that the tissue-specific ALP isoenzymes shared 86–98% identity to one another,[5] but only 52–56% identity with tissue nonspecific ALPs.[6] In the past years, ALPs have been continuously investigated for their diagnostic and therapeutic values. The bone isoenzymes have long been recognized to play a role in skeletal mineralization.[7,8] The activity of liver and bone ALPs in serum has been extensively applied in routine diagnosis of human diseases. Many studies regarding the prophylactic and therapeutic potential of intestinal alkaline phosphatases have confirmed their potent role in protecting the gut and the entire body against inflammation.[9] Nevertheless, up to now, our knowledge regarding the physiological functions of ALPs is still relatively limited.

The ALP superfamily consists of a large group of highly promiscuous metallohydrolases, sharing a similar active site architecture and substrate preference, but exhibiting limited sequence homology.[10] The members of ALPs have served extensively as both experimental and computational model systems for deciphering the mechanisms of catalytic promiscuity.[11] The classical perception of enzyme catalysis is that enzymes are highly specific, with each enzyme for the turnover of a single substrate. Nonetheless, mounting evidence suggests that many enzymes are capable of facilitate multiple reactions within the same active site.[12] The ALP superfamily provides an excellent model system for both *in vitro* and *in silico* studies of enzyme promiscuity, as members of ALP superfamily not only exhibit catalytic promiscuity, but also show crosswise promiscuity.[10]

The first cDNA encoding an insect ALP was cloned and sequenced in the silkworm, *Bombyx mori*.[13] It was a membrane-bound ALP (m-ALP) with a highly hydrophobic domain presumed to be a membrane anchoring region at the C-terminus. Based on sequence analysis, the similarities between *Bombyx* m-ALP and mammalian ALPs were very low (41.7 and 45.6%), but higher homologies existed between m-ALP and tissue-nonspecific ALPs. Nonetheless, the functionally important substrate-binding sites and metal-binding sites of various ALP isozymes were highly conserved, suggesting a common ancestral gene for ALPs, although the silkworm ALP genes might have diverged long before mammalian ALP genes.[14] Earlier studies observed two midgut ALP isozymes in *Bombyx* larvae.[15] Further biochemical analyses revealed that the slow migrating band (membrane-bound alkaline phosphatase, m-ALP) and the fast moving one (soluble alkaline phosphatase, s-ALP),[16] but both forms were shown to be a monomer, different from the functional dimeric architecture of mammalian ALPs.[16–18] In addition, the activity of silkworm ALPs was strongly inhibited by L-cysteine, although mammalian ALPs are inhibited by L-phenylalanine, L-homoarginine and L-leucine.[14] More impressively, ATPase activity was detected in the purified s-ALP of silkworm.[19] Despite all these studies, the

[1]Department of Entomology and MOA Key Laboratory for Monitoring and Environment-Friendly Control of Crop Pests, College of Plant Protection, China Agricultural University, Beijing 100193, China
[2]Lead contact
*Correspondence: zxli@cau.edu.cn
https://doi.org/10.1016/j.isci.2023.108115

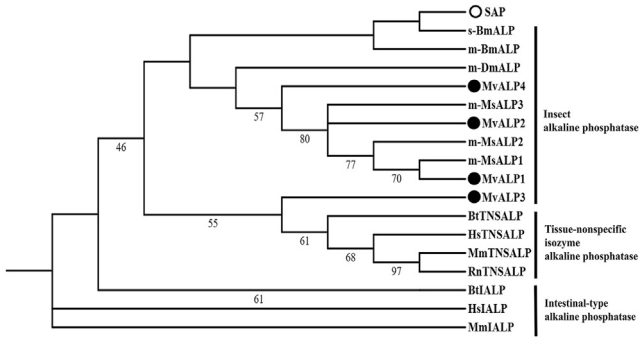


Figure 1. An a maximum-likelihood tree based on ALPs amino acid sequences

The numbers represent the confidence calculated by bootstrapping. The ALPs of *M. viciae* and SAP are indicated by dark and hollow dots, respectively.

physiological roles of insect ALPs are poorly understood. Till today, only limited progress has been made in the involvement of insect ALPs as toxin-binding proteins or functional receptors in the resistance of lepidopteran insects to Cry1Ac, an insecticidal crystal toxin from *Bacillus thuringiensis* (Bt).[20–23]

The vetch aphid *Megoura viciae* as an important agricultural pest causes serious damage to the broad bean *Vicia faba* and the pea *Pisum sativum*.[24,25] Gas chromatography/mass spectrometry combined with behavioral assays identified the major components of alarm pheromones in *M. viciae*.[26] In addition, the molecular mechanism underlying the biosynthesis of alarm pheromones in the aphid is still partially understood.[27–30] Here we first identified four ALPs in *M. viciae* based on transcriptome sequencing and molecular cloning. We next focused on the functional characterization of these ALPs by *in vitro* enzymatic activity assay, RNA interference and differential transcriptome sequencing. Our results confirmed that one of these ALPs is a multifunctional enzyme. This is the first report of an insect ALP involved terpene synthase enzyme activity.

RESULTS

Sequences of *M. viciae* alkaline phosphatases

The full-length cDNAs encoding ALPs in *M. viciae* were sequenced (designated as *MvALP1-4*) (GenBank acc. nos. UXP86334.1, UXP86335.1, UXP86336.1, and UXP86337.1). The open reading frames of *MvALP1-4* consisted of 1665-bp, 1542-bp, 1605-bp and 1425-bp nucleotides, encoding 557, 514, 535 and 475 amino acids, with a predicted molecular weight and theoretical pI of 60.64 kDa/6.04, 56.7 kDa/6.18, 59.01 kDa/6.61 and 52.19 kDa/6.84, respectively. Sequence alignment showed that these ALPs shared an amino acid identity of 49.81% with their homologs from *Drosophila melanogaster*, *Bombyx mori,* and shrimp alkaline phosphatase (SAP) (Figure S4). All of these ALPs contained one alkaline phosphatase site and three metal binding sites (Figure S4; Table S5).[13,31] A hydrophobic signal peptide existed in *MvALP1-4* and all membrane-bound ALPs but not in soluble silkworm ALP and SAP (Figure S4).[32] Prediction of functional sites identified *N*-glycosylation site,[33] amidation,[34,35] Tyr-phosphatase site,[36] cell attachment site,[37] P-II protein family signatures,[38] and hydrophobic domain[39–41] in these ALPs (Tables S3–S6). It is noteworthy that only MvALP4 and *m*-DmALP contained both amidation site and Tyr-phosphatase site (Tables S5 and S6).

Phylogenetic analysis

Phylogenetic analysis showed that MvALP1, MvALP2, MvALP4 and other insect ALPs were clustered in a subclade with bootstrap score (46%), while MvALP3 was clustered in a subclade belonging to tissue-nonspecific isozyme ALPs with a bootstrap score (55%) (Figure 1). The alkaline phosphatases from *Homo sapiens*, *Bos taurus*, and *Mus musculus* formed a distinct group (intestinal-type ALPs). In addition, the silkworm soluble ALP (s-BmALP) showed a close phylogenetic relationship to SAP from *Pandalus borealis*.

Spatiotemporal expression profiling of *MvALP1-4*

Temporal dynamics analysis showed that *MvALP1* was expressed significantly higher during the adult stage compared to the four nymphal stages, followed by the 4th-instar nymphal stage ($p < 0.001$) (Figure 2A); there was no significant difference in the expression of *MvALP2* for all developmental stages ($p = 0.164$) (Figure 2B); a significantly higher expression of *MvALP3* was observed during the 2nd- and 3rd-instar stages ($p = 0.004$ and 0.048, respectively) (Figure 2C), and a significantly higher expression of *MvALP4* was detected during the 4th-instar nymphal stage ($p = 0.039$). Tissue-specific expression analysis showed that *MvALP1* was expressed highest in the head ($p < 0.001$), followed by in the cornicle and embryo ($p < 0.001$) (Figure 2E); *MvALP2* was expressed highest in the embryo, followed by in the cornicle and midgut ($p < 0.001$) (Figure 2F). In contrast, both *MvALP3* and *MvALP4* showed an extremely higher expression in the midgut ($p < 0.001$) (Figures 2G and 2H).

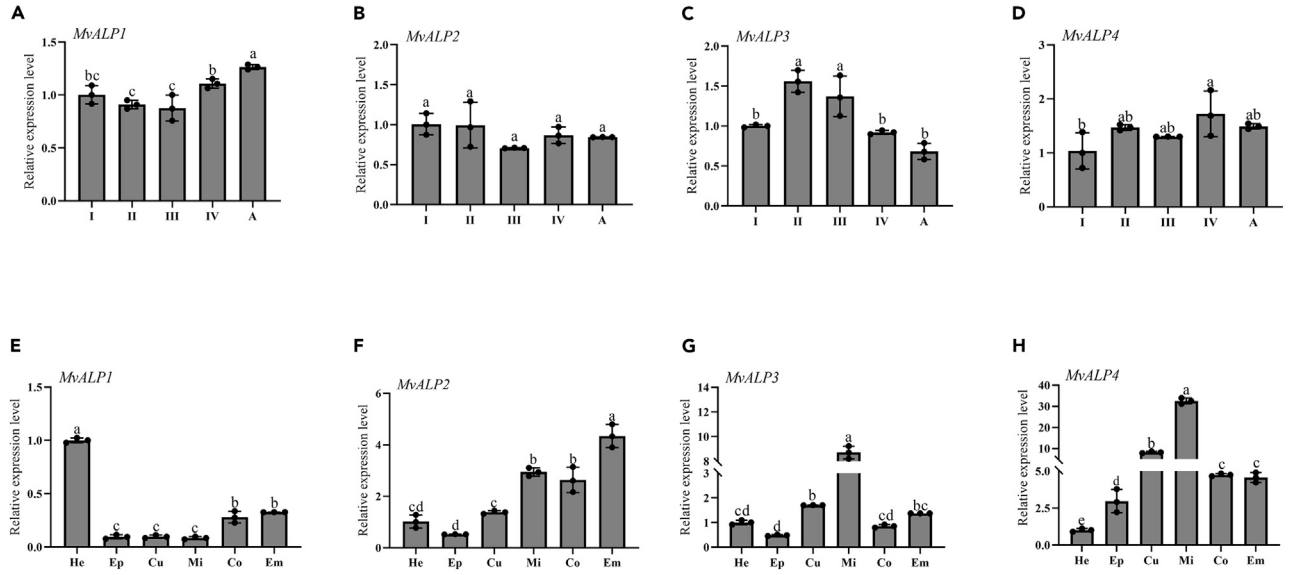


Figure 2. Spatiotemporal expression profiling of *MvALP1-4* by real-time quantitative PCR

(A–D) Stage-specific expression analysis. I-IV: 1st-4th instar nymph; A: 2-day-old adult.

(E–H) Tissue-specific expression analysis. He: head; Ep: epicuticle; Cu: cuticle; Mi: midgut; Co: cornicle; Em: embryo. Data are means ± SD of three biological replicates. The significance of differences were analyzed by ANOVA and Tukey's test. Different lowercase letters indicate significant difference at $p < 0.05$.

Molecular modeling and docking

The best 3D models (rank 1) of MvALP1-4 were built (Figures 3A, 3D, 3G, and 3J), with an overall quality factor of 91.829%, 87.344%, 94.336%, and 94.643%, respectively (Figure S5). A Ramachandran plot indicated that 88.0%, 87.1%, 91.9%, and 91.7% of MvALP1-4 residues were in the most favored regions, respectively (Figure S6). The binding residues for divalent metal ions (MG) in MvALP1 active site were Asp78 (2.14 Å), Ser128, His191, Ser193 (2.39 Å) and Glu350 (3.22 Å) (Figures 3B and 4A). Molecular docking indicated that (*E*, *E*)-FPP was tightly bound to MvALP1 with a G value of −5.9 kcal/mol, via hydrogen bonds with Arg188 (3.31 Å), His191 (2.96 Å), His312 (3.19 Å) and His359 (3.18 Å) (Figures 3C and 4B; Table S7). These four key residues and additional residues including Asp127, Tyr143, Lys144, Arg204, Glu205, Val356, Phe471, Glu472, Thr473, and His474 formed the binding site (Figure 4B; Table S8). Similarly, divalent metal ions (MG) tightly bound to different binding residues in MvALP2 (Asp90, Ser142, His205, Ser207 and Glu352 (2.99 Å)), MvALP3 (Asp59 (2.21 Å), Ser109, Asp172, Ser174 (2.21 Å) and Glu332 (3.29 Å)) and MvALP4 (Asp90 (2.51 Å), Ser141, His206 (2.21 Å), Thr207 and Glu368 (3.03 Å)), respectively (Figures 3E, 3H, 3K and 4C, 4E, 4G). The glide G value of MvALP2, MvALP3, and MvALP4 were −6.2 kcal/mol, −6.2 kcal/mol and −6.7 kcal/mol, respectively (Table S7). (*E*, *E*)-FPP was bound to MvALP2 via hydrogen bonds with Ser142 (2.91 Å), Arg218 (2.78 Å) and His361 (2.80 Å), which formed the binding site along with the other 13 residues (Ser34, Met36, Asp37, Phe40, Arg202, His205, Ser312, His313, Met314, Arg355, Asp357, His358 and Gln362) (Figures 3F and 4D; Table S8). (*E*, *E*)-FPP was bound to MvALP3 via hydrogen bonds to Arg169 (3.11 Å), Asp172 (2.91 Å and 3.00 Å), Arg185 (3.04 Å and 3.21 Å) and Lys338 (2.93 Å and 2.97 Å), forming a binding pocket with additional 4 residues (His294, Met295, Asp296, and Tyr297) (Figures 3I and 4E; Table S8). (*E*, *E*)-FPP was bound to MvALP4 via hydrogen bonds to His204 (3.11 Å and 3.20 Å), Arg217 (2.92 Å and 3.06 Å) and Trp374 (3.06 Å), forming a binding pocket with additional 3 residues (Tyr333, His377 and Ser378) (Figures 3L and 4G; Table S8).

In vitro enzymatic activity assay

The *in vitro* reaction using the commercial SAP enzyme (positive control) and (*E*, *E*)-FPP generated (*E*, *E*)-farnesol (RT = 33.666 min) and a small amount of (*E*, *E*)-β-farnesene (RT = 26.071 min) (Figures S7 and S8). The reactions using the recombinant proteins of MvALP1-4 demonstrated that only MvALP4 could catalyze the formation of (*E*, *E*)-farnesol (Figure 5D). Moreover, MvALP4 could generate a small amount of (*E*, *E*)-β-farnesene based on GC-MS analysis with extracted ion chromatogram (EIC) (Figures 6, S9, and S11).

RNAi-mediated functional analysis of *MvALP4*

The expression of *MvALP4* was successfully knocked down by treatment with *dsMvAP4* for 36 h or 48 h (Figure 7A). Compared to the control (treatment with *dsGFP*), *MvALP4* was significantly down-regulated by 62.62% at 48 h ($p < 0.001$). After RNAi against *MvALP4*, the contents of four major alarm pheromone components per aphid (ng/μL) were significantly decreased in *M. viciae* (Figure 7B). Specifically, the amount of (−)-β-pinene, (*E*, *E*)-β-farnesene, (−)-α-pinene and (+)-limonene was decreased by 29.94%, 14.01%, 54.40% and 36.31%, respectively, after treatment with *dsMvALP4* ($p = 0.000294$, 0.003818, 0.000128 and 0.003158, respectively).

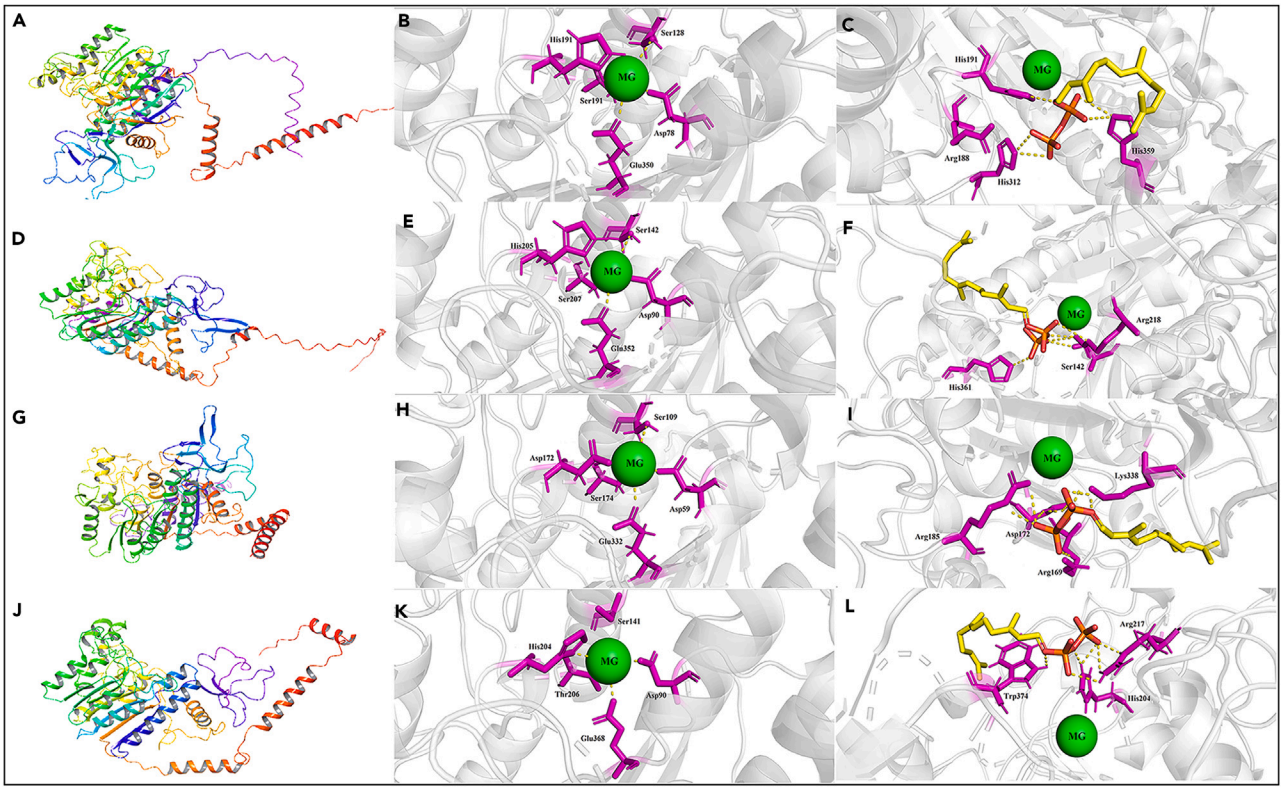


Figure 3. Three-dimensional model and molecular docking of MvALP1-4

(A, D, G, J) Three-dimensional the active site architectures structures of MvALP1, MvALP2, MvALP3, and MvALP4, respectively.

(B, E, H, K) The active site architectures structures and binding residues for Mg^{2+} (green) in MvALP1-4, respectively.

(C, F, I, L) Interactions between MvALP1-4 and (*E*, *E*)-FPP via key residues (purple).

Alkaline phosphatase activity of crude proteins

The enzymatic activity assay using crude protein extracts from the 3rd-instar nymphs with (*E*, *E*)-FPP as the substrate identified (*E*, *E*)-farnesol (RT = 33.719 min) and (*E*, *E*)-β-farnesene (RT = 26.074 min) based on GC-MS product analysis (Figure 8A). No expected products were detected in the negative control group (Figure 8B).

Quantitative analysis of alkaline phosphatases and (*E*, *E*)-β-farnesene in *M. viciae* under simulated stimulation

Quantitative analysis showed that the amount of (*E*, *E*)-β-farnesene was significantly decreased at different time points after droplet release from the cornicle by simulated stimulation ($p < 0.01$) (Figure 9A), but the amount of (*E*, *E*)-β-farnesene rebound sharply at 12 h after droplet release, although the level was still lower than the control group (Figure 9B). In contrast, the total content of alkaline phosphatases gradually increased and reached a peak at 24 h after simulated stimulation (Figures 9C and 9D).

Transcriptome sequencing in response to simulated stimulation

A total of 47,244 unigenes were annotated based on transcriptome sequencing. The KEGG enrichment analysis classified these unigenes into six functional categories: Organismal systems, Metabolism, Human diseases, Genetic information processing, Environmental information processing, and Cellular processes (Figure S8). Transcriptome sequencing in response to simulated stimulation identified 895 differentially expressed genes (DEGs), including 555 up-regulated and 340 down-regulated DEGs (Figures 9A and 9B). Based on the transcriptome data, the expressions of *MvALP1-4* were up-regulated by 0.06, 0.49, 1.02, and 0.17-fold changes transformed by Log2, respectively, whereas the expressions of *FPPS* and *GGPPS* were not significantly changed (Figure 10).

DISCUSSION

In the present study, we identified four ALP genes in *M. viciae* based on transcriptome sequencing, molecular cloning, sequence analysis, and phylogenetic analysis. The four MvALPs shared a relatively low amino acid identity with their homologs from other insect species. Phylogenetic analysis showed that all insect ALPs are clustered in a group, except MvALP3 grouped with tissue-nonspecific isozyme ALPs.

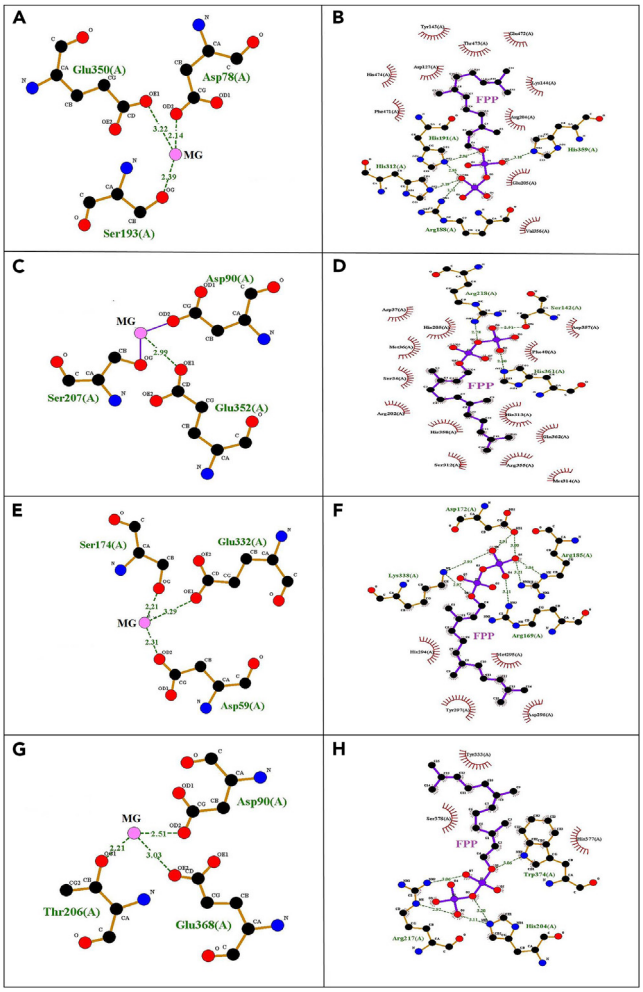


Figure 4. Molecular docking of MvALP1-4 with ligands visualized in 2D

(A, C, E, G) The binding residues for Mg^{2+} (pink) in MvALP1-4, respectively.

(B, E, H, K) Interactions between MvALP1-4 and (*E*, *E*)-FPP via binding residues. The numbers indicate the positions of amino acid residues. The green dashes and corresponding values represent hydrogen bonds and distances, respectively.

Interestingly, the shrimp SAP is clustered with the silkworm soluble s-BmALP, which are closely related to a neighboring subclade containing *m*-BmALP, *m*-DmALP, and MvALP4. Prediction of functional sites showed that all of these ALPs contained one alkaline phosphatase site and three metal-binding sites, but varied in other functional sites such as *N*-glycosylation site, amidation, Tyr-phosphatase site, cell attachment site, P-II protein family signatures, and hydrophobic domain, indicating functional promiscuity of this family of enzymes.

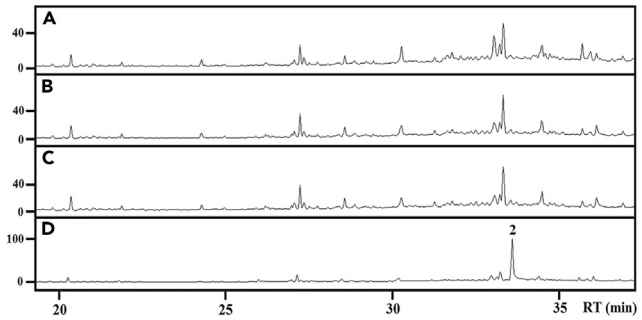


Figure 5. Identification of the products generated by recombinant MvALP1-4 with (*E*, *E*)-FPP as the substrate by using GC-MS

(A) MvALP1; (B) MvALP2; (C) MvALP3, and (D) MvALP4. Peak no.2 indicate (*E*, *E*)-farnesol (RT = 33.618 min).

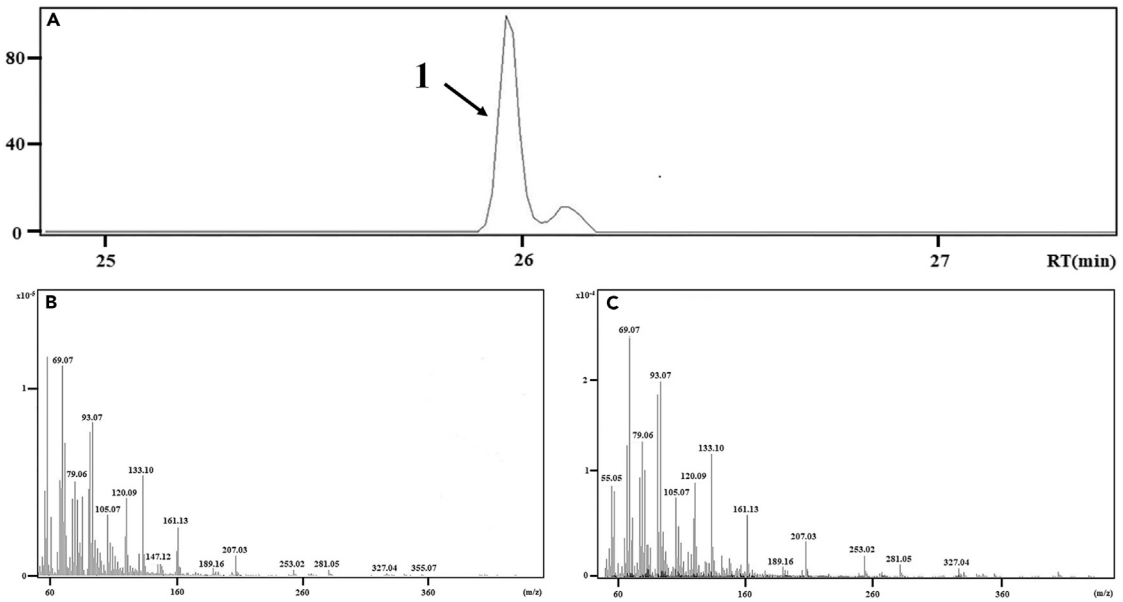


Figure 6. Identification of the products generated by recombinant MvALP4 with (*E, E*)-FPP as the substrate by using GC-MS

(A) Gas chromatogram of (*E, E*)-β-farnesene with EIC (RT = 26.012 min), Peak no.1.

(B) Mass spectra of the product (m/z: 69.07, 79.05, 93.07, 105.07, 120.09, 133.10, 147.12, 161.13, 189.16, 207.03, 253.02, 281.05, 327.04, 355.07).

(C) Mass spectra of the reference standard (*E, E*)-β-farnesene (m/z: 55.05, 69.07, 79.05, 93.07, 105.07, 120.09, 133.10, 161.13, 189.16, 207.03, 253.02, 281.05, 327.04).

To predict the potential action sites of MvALPs, we conducted a spatiotemporal expression analysis of *ALPs* in *M. viciae*. As expected, *MvALPs* were expressed during all developmental stages of the aphid. Nonetheless, tissue-specific analysis uncovered distinct expression patterns of different *MvALP* genes: *MvALP1* was expressed highest in the head; *MvALP2* was expressed highest in the embryo, but both *MvALP3* and *MvALP4* had an extremely higher expression in the midgut. In insects, pheromones are usually synthesized in specialized cells associated with the epidermis,[42] and midgut cells as the site for pheromone biosynthesis have been reported in males of the bark beetles *Ips pini* and *Dendroctonus jeffreyi*.[43,44] Although there is not a pheromone gland in the midgut, cells appear to be involved in isoprenoid pheromone production throughout the tissue.[45] In aphids, evidences from biochemical and molecular studies have shown that the terpene alarm pheromone components are synthesized *de nova*.[28,30] Thus, the possibility of MvALPs being involved in the biosynthesis of aphid alarm pheromone should not be excluded.

Molecular modeling and docking help us to better understand the relationships between structures and functions of enzymes. The active site of silkworm *m*-BmALP contains three metal binding sites, including two Zn^{2+} (the substrate binding arginine) and one Mg^{2+} (coordinating water molecule hydrogen-bonding to the substrate). In the magnesium site, the residues Asp51, Asp153, Thr155 and Glu322 were presumed to be important for constituting the active pocket in *m*-BmALP.[13] Based on 3-D modeling in this study, the active pocket (magnesium site) in MvALP1-4 were formed by the conserved residues Asp, Ser, His/Asp/Thr, Ser and Glu, respectively, and the first Ser is the key amino acid residue in alkaline phosphatase sites. Nonetheless, molecular docking simulation revealed that, although all four MvALPs can tightly bind to the substrate (*E, E*)-FPP, the residues forming the active pocket of different MvALPs varied. These results suggested a structure-based functional promiscuity in MvALPs.

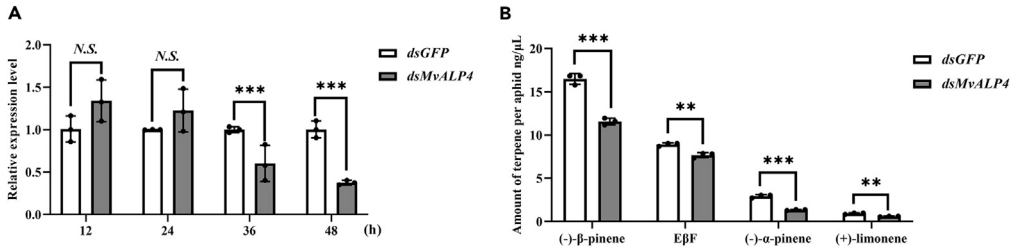


Figure 7. RNAi-mediated knockdown of *MvALP4* and its effects on the amount of aphid alarm pheromones in *M. viciae*

(A) RNAi efficiency of *dsMvALP4* at different time points (12 h, 24 h, 36 h and 48 h) after treatment; (B) The amounts of aphid alarm pheromones after treatment with *dsMvALP4* compared to the negative control group (*dsGFP*). Data are means ± SD of three biological replicates. The significance of differences is analyzed by using Student's *t* test, **, $p < 0.01$; ***, $p < 0.001$; *N.S.*, $p > 0.05$.

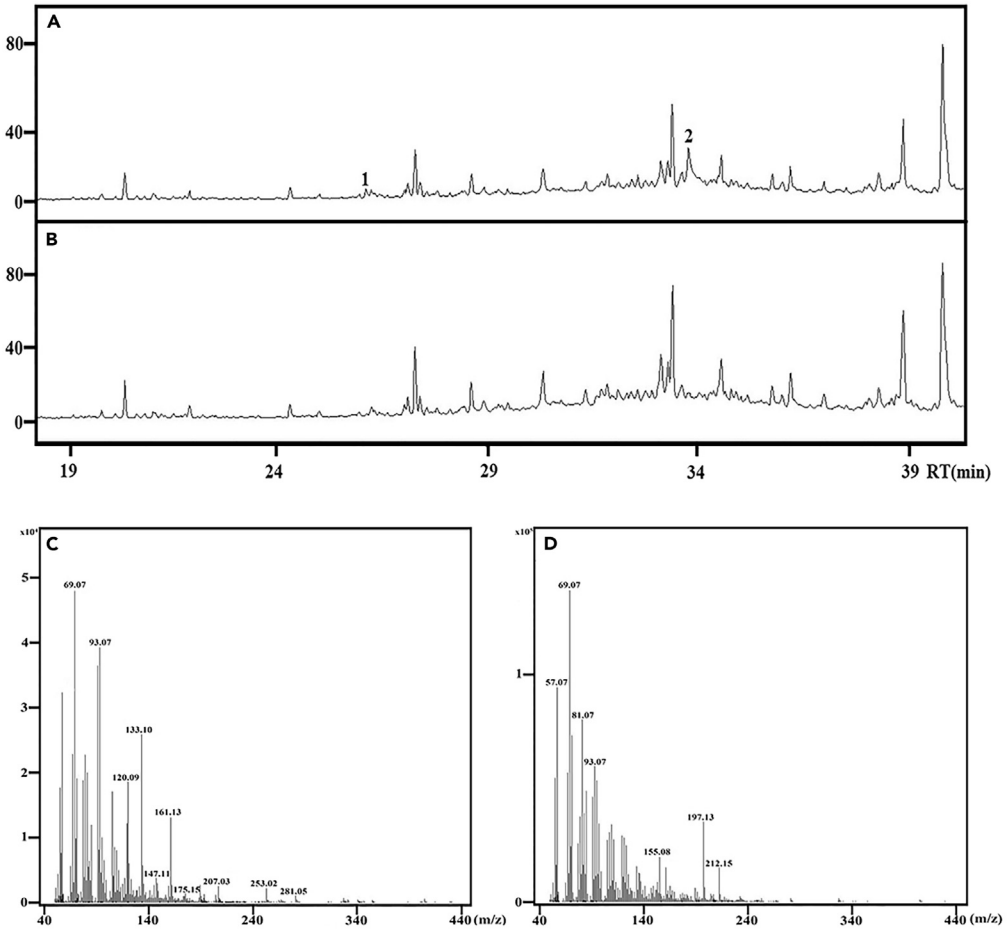


Figure 8. Identification of the product generated by crude proteins extracted from *M. vicia* with (*E*, *E*)-FPP as the substrate by using GC-MS

(A) Crude proteins with (*E*, *E*)-FPP.

(B) Crude proteins without (*E*, *E*)-FPP (negative control). Peaks nos. 1 and 2 indicate (*E*, *E*)-β-farnesene (RT = 26.074 min) and (*E*, *E*)-farnesol (RT = 33.719 min), respectively.

(C) Mass spectra of (*E*, *E*)-β-farnesene (m/z: 69.07, 93.07, 120.09, 133.10, 147.11, 161.13, 175.15, 207.03, 253.02, 281.05).

(D) Mass spectra of (*E*, *E*)-farnesol (m/z: 57.07, 69.07, 81.07, 93.07, 155.08, 197.13, 212.15).

In vitro enzymatic activity assay using (*E*, *E*)-FPP as the substrate found that MvALP4 can catalyze the formation of (*E*, *E*)-farnesol and generate a small amount of EβF. This result is interesting because it not only corroborated the instinct ALP activity of MvALP4 but also uncovered a novel sesquiterpene synthase activity in MvALP4. Moreover, this finding is surprising as all four MvALPs are predicted to be functional ALPs based on structural and molecular studies. One of the possible explanations is that the reaction conditions for the commercial

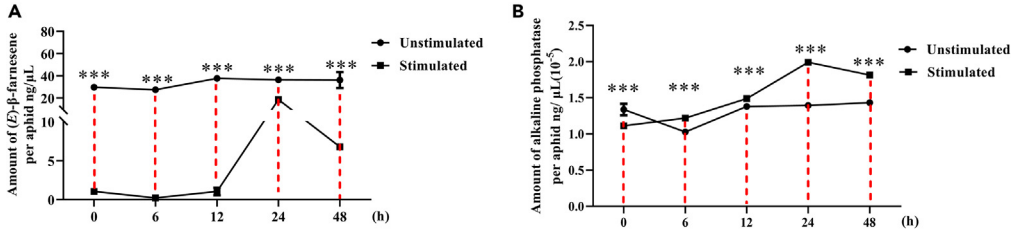


Figure 9. Quantitative analysis of (*E*, *E*)-β-farnesene and ALPs in *M. viciae* under stimulated and unstimulated conditions

(A) Amounts of (*E*, *E*)-β-farnesene; (B) dynamics of (*E*, *E*)-β-farnesene represented by line chart.

(C) Amounts of ALPs.

(D) dynamics of ALPs represented by the line chart. Data are means ± SD of three biological replicates. The significance of differences is analyzed by using Student's *t* test, ***, $p < 0.001$; *N.S.*, $p > 0.05$.

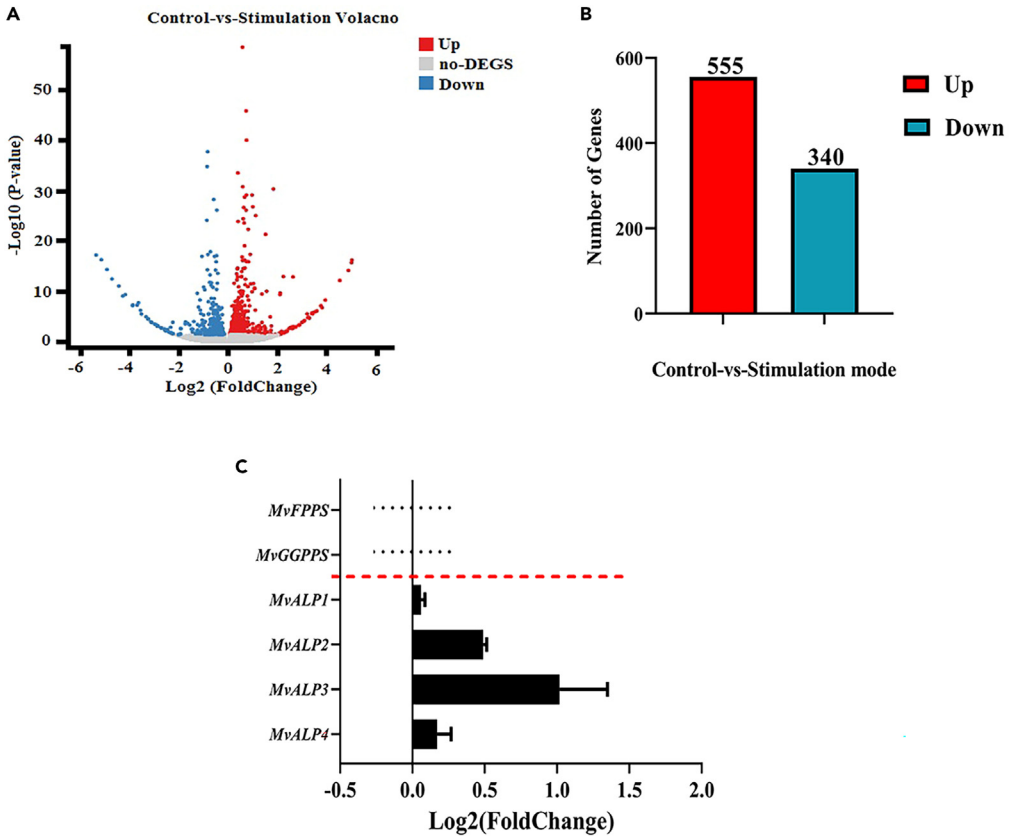


Figure 10. Comparative transcriptome analysis in response to simulated stimulation

(A) Volcano plot.

(B) Numbers of DEGs in the 3rd-instar nymphs at 24 h after simulated stimulation.

(C) Fold changes in the expression levels of *MvFPPS*, *MvGGPPS* and *MvALP1-4* transformed by Log2. The gray splashes indicate genes without significantly differential expression. The red splashes indicate significantly up-regulated genes, and the blue splashes indicate significantly down-regulated genes.

alkaline phosphatase SAP (pH 9.0, 37°C, Mg^{2+}) are only optimal for MvALP4 but not for other three MvALPs. Most known common catalytic scaffolds employed by alkaline phosphatases include one or more divalent metal ions (Zn^{2+}, Ca^{2+} or Mn^{2+}) with different pH values that played important roles in catalytic function.[11] The pH stability of m-ALP and s-ALP were 7–12 (optimum: 10.9) and 7–9 (optimum: 9.8), respectively.[14] If this is the case, more optimizations are needed for the reaction conditions for each of MvALPs, including temperature, pH values, and metal ions. Thus, it is still a challenge for us to identify the optimal combination of the catalytic promiscuity in the alkaline phosphatase.

The function of MvALP4 is involved terpene synthase activity that based on *in vitro* enzymatic assay, which is supported by our RNAi-mediated functional analysis. After expression knockdown of *MvALP4*, the four major alarm pheromone components were significantly reduced in the aphid, including EβF, (−)-β-pinene, (−)-α-pinene and (+)-limonene. These results suggested that MvALP4 was not only involved in the biosynthesis of EβF but of the other three components, although the reaction conditions for the other pheromone components need further investigation. Moreover, the terpene synthase activity of MvALP4 was also corroborated by enzymatic activity assay using crude proteins from aphids, in which (*E*, *E*)-farnesol and EβF were generated with (*E*, *E*)-FPP as the substrate. This confirmed that EβF was biosynthesized *de nova* in the aphid.

We also performed a quantitative analysis of EβF and ALPs after droplet release by the aphid in response to simulated stimulation. It revealed that EβF was significantly reduced after droplet release, with a sharp rebound after 12 h, while ALPs were enhanced by stimulation, with a gradual rebound. These results showed that ALPs are negatively correlated with the amount of EβF, suggesting a role of ALPs in the biosynthesis of EβF.

Finally, we performed a transcriptome analysis in the aphid in response to simulated stimulation, which identified 895 DEGs (555 up and 340 down). The expressions of *MvALP1-4* were up-regulated to different extents. Impressively, *MvALP3* exhibited the highest extent of upregulation, followed by *MvALP2*, while *MvALP1* and *MvALP4* displayed the least upregulation. It is thus interesting to examine the novel functions of *MvALP3* and *MvALP2* in the future. In addition, the expression of farnesyl diphosphate synthase *FPPS* and geranylgeranyl diphosphate synthase *GGPPS* (the key genes of the terpenoid backbone biosynthesis pathway) showed no significant change after stimulation, which was consistent with our previous finding that the immediate compensation of alarm pheromones was not required after droplet release.[46]

In conclusion, we identified four ALPs in *M. viciae*. We further showed that MvALP4 is a multifunctional enzyme with ALP and sesquiterpene synthase activities based on molecular docking simulation, *in vitro* enzymatic assay, RNAi-mediated gene knockdown, crude protein assay and quantitative analysis. Our results suggest that at least one of the four ALPs in *M. viciae* is involved in the biosynthesis of alarm pheromones. This is the first report of an insect ALP having terpene synthase activity. Our study contributes to the understanding of the functional plasticity of ALPs in insects.

Limitations of the study

Our results revealed a multifunctional ALP involved the biosynthesis of aphid alarm pheromones. However, our enzymatic activity assay is based recombinant ALP enzymes in *Escherichia coli*. It is still necessary to verify whether we can get purified ALPs in insects. Our results only represent an insect ALP having terpene synthase activity, while the MvALPs contained similar functional sites that indicated they could have same function. We also need know about more reaction conditions for each of MvALPs. In summary, promiscuous functions of ALPs need more power evidence in the future research.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Insects
- METHOD DETAILS
  - RNA extraction
  - Sequencing of ALP genes in *M. viciae*
  - Phylogenetic analysis
  - Spatiotemporal expression profiling of ALP genes in *M. viciae*
  - Molecular modeling and docking
  - Recombinant expression and purification of ALPs
  - *In vitro* enzymatic activity assay of ALPs
  - Product analysis by GC-MS
  - Functional analysis of MvALP4 by RNA interference (RNAi)
  - Enzymatic activity assay of crude proteins
  - Quantitative analysis of ALPs and alarm pheromone component in *M. viciae* under simulated stimulation
  - Transcriptome sequencing in response to simulated stimulation
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Statistical analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.108115.

ACKNOWLEDGMENTS

This work is supported by the National Natural Science Foundation of China (Grant nos. 31972267) and Chinese Universities Scientific Fund (Grant nos. 2023TC109).

AUTHOR CONTRIBUTIONS

X.S.: conceptualization, methodology, investigation, data curation, formal analysis, writing - original draft, and writing-review & editing. Y.-G.Q.: investigation, data curation, and formal analysis. Y.-H.Z.: investigation, data curation, and formal analysis. Y.-B.Z.: investigation, data curation, and formal analysis. D.C.: data curation and formal analysis. D.-H.X.: data curation and formal analysis. Z.-X.L.: conceptualization, writing - original draft, writing - review & editing, validation, and supervision.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: July 20, 2023
Revised: September 1, 2023
Accepted: September 29, 2023
Published: October 1, 2023

REFERENCES

1. Trowsdale, J., Martin, D., Bicknell, D., and Campbell, I. (1990). Alkaline phosphatases. Biochem. Soc. Trans. *18*, 178–180. https://doi.org/10.1042/bst0180178.
2. Shigenari, A., Ando, A., Baba, T., Yamamoto, T., Katsuoka, Y., and Inoko, H. (1998). Characterization of alkaline phosphatase genes expressed in seminoma by cDNA cloning. Cancer Res. *58*, 5079–5082. https://doi.org/10.1201/b16830-6.
3. Sadeghirizi, A., and Yazdanparast, R. (2007). Plasma membrane homing of tissue nonspecific alkaline phosphatase under the influence of 3-hydrogenkwadaphnin, an anti-proliferative agent from *Dendrostellera lessertii*. Acta Biochim. Pol. *54*, 323–329. https://doi.org/10.18388/abp.2007_3253.
4. Sharma, U., Pal, D., and Prasad, R. (2014). Alkaline phosphatase: an overview. Indian J. Clin. Biochem. *29*, 269–278. https://doi.org/10.1007/s12291-013-0408-y.
5. Mornet, E., Stura, E., Lia-Baldini, A.S., Stigbrand, T., Ménez, A., and Le Du, M.H. (2001). Structural evidence for a functional role of human tissue nonspecific alkaline phosphatase in bone mineralization. J. Biol. Chem. *276*, 31171–31178. https://doi.org/10.1074/jbc.m102788200.
6. Whyte, M.P., Landt, M., Ryan, L.M., Mulivor, R.A., Henthorn, P.S., Fedde, K.N., Mahuren, J.D., and Coburn, S.P. (1995). Alkaline phosphatase: placental and tissue-nonspecific isoenzymes hydrolyze phosphoethanolamine, inorganic pyrophosphate, and pyridoxal 5′-phosphate. Substrate accumulation in carriers of hypophosphatasia corrects during pregnancy. J. Clin. Invest. *95*, 1440–1445. https://doi.org/10.1172/jci117814.
7. Orimo, H. (2010). The mechanism of mineralization and the role of alkaline phosphatase in health and disease. J. Nippon Med. Sch. *77*, 4–12. https://doi.org/10.1272/jnms.77.4.
8. Vimalraj, S. (2020). Alkaline phosphatase: Structure, expression and its function in bone mineralization. Gene *754*, 144855. https://doi.org/10.1016/j.gene.2020.144855.
9. Lallès, J.P. (2019). Recent advances in intestinal alkaline phosphatase, inflammation, and nutrition. Nutr. Rev. *77*, 710–724. https://doi.org/10.1093/nutrit/nuz015.
10. Jonas, S., and Hollfelder, F. (2009). Mapping catalytic promiscuity in the alkaline phosphatase superfamily. Pure Appl. Chem. *81*, 731–742. https://doi.org/10.1351/pac-con-08-10-20.
11. Pabis, A., and Kamerlin, S.C.L. (2016). Promiscuity and electrostatic flexibility in the alkaline phosphatase superfamily. Curr. Opin. Struct. Biol. *37*, 14–21. https://doi.org/10.1016/j.sbi.2015.11.008.
12. Khersonsky, O., and Tawfik, D.S. (2010). Enzyme promiscuity: a mechanistic and evolutionary perspective. Annu. Rev. Biochem. *79*, 471–505. https://doi.org/10.1016/b978-008045382-8.00155-6.
13. Itoh, M., Takeda, S., Yamamoto, H., Izumi, S., Tomino, S., and Eguchi, M. (1991). Cloning and sequence analysis of membrane-bound alkaline phosphatase cDNA of the silkworm, *Bombyx mori*. Biochim. Biophys. Acta *1129*, 135–138. https://doi.org/10.1016/0167-4781(91)90229-f.
14. Eguchi, M. (1995). Alkaline phosphatase isozymes in insects and comparison with mammalian enzyme. Comp. Biochem. Physiol. B Biochem. Mol. Biol. *111*, 151–162. https://doi.org/10.1016/0305-0491(94)00248-s.
15. Yoshitake, N. (1964). Genetical studies on the alkaline phosphatase in the midgut of the silkworm, *Bombyx mori* L. J Sericult Sci *33*, 28–33. https://doi.org/10.1046/j.1439-0418.2002.00625.x.
16. Itoh, M., Kanamori, Y., Takao, M., and Eguchi, M. (1999). Cloning of soluble alkaline phosphatase cDNA and molecular basis of the polymorphic nature in alkaline phosphatase isozymes of *Bombyx mori* midgut. Insect Biochem. Mol. Biol. *29*, 121–129. https://doi.org/10.1016/s0965-1748(98)00115-5.
17. Okada, N., Azuma, M., and Eguchi, M. (1989). Alkaline phosphatase isozymes in the midgut of silkworm: Purification of high pH-stable microvillus and labile cytosolic enzymes. J. Comp. Physiol. B *159*, 123–130. https://doi.org/10.1007/bf00691732.
18. Terra, W.R., and Ferreira, C. (1994). Insect digestive enzymes: properties, compartmentalization and function. Comp. Biochem. Physiol. Part B Comp. Biochem. *109*, 1–62. https://doi.org/10.1016/0305-0491(94)90141-4.
19. Azuma, M., Takeda, S., Yamamoto, H., Endo, Y., and Eguchi, M. (1991). Goblet cell alkaline phosphatase in silkworm midgut epithelium: its entity and role as an ATPase. J. Exp. Zool. *258*, 294–302. https://doi.org/10.1002/jez.1402580304.
20. Jurat-Fuentes, J.L., and Adang, M.J. (2004). Characterization of a Cry1Ac-receptor alkaline phosphatase in susceptible and resistant *Heliothis virescens* larvae. Eur. J. Biochem. *271*, 3127–3135. https://doi.org/10.1111/j.1432-1033.2004.04238.x.
21. Jurat-Fuentes, J.L., and Ferré, J. (2007). A proteomic approach to study Cry1Ac binding proteins and their alterations in resistant *Heliothis virescens* larvae. J. Invertebr. Pathol. *95*, 187–191. https://doi.org/10.1016/j.jip.2007.01.008.
22. Jurat-Fuentes, J.L., Karumbaiah, L., Jakka, S.R.K., Ning, C., Liu, C., Wu, K., Jackson, J., Gould, F., Blanco, C., Portilla, M., et al. (2011). Reduced levels of membrane-bound alkaline phosphatase are common to lepidopteran strains resistant to Cry toxins from *Bacillus thuringiensis*. PLoS One *6*, e17606. https://doi.org/10.1371/journal.pone.0017606.
23. Wei, J., Zhang, M., Liang, G., and Li, X. (2019). Alkaline phosphatase 2 is a functional receptor of Cry1Ac but not Cry2Ab in *Helicoverpa zea*. Insect Mol. Biol. *28*, 372–379. https://doi.org/10.1111/imb.12556.
24. Kunert, G., Schmoock-Ortlepp, K., Reissmann, U., Creutzburg, S., and Weisser, W.W. (2008). The infuence of natural enemies on wing induction in *Aphis fabae* and *Megoura viciae* (Hemiptera: Aphididae). Bull. Entomol. Res. *98*, 59–62. https://doi.org/10.1017/s0007485307005391.
25. Leroy, P.D., Wathelet, B., Sabri, A., Francis, F., Verheggen, F.J., Capella, Q., Thonart, P., and Haubruge, E. (2011). Aphid-host plant interactions: does aphid honeydew exactly refect the host plant amino acid composition? Arthropod-Plant Inte. *5*, 193–199. https://doi.org/10.1007/s11829-011-9128-5.
26. Song, X., Qin, Y.G., Yin, Y., and Li, Z.X. (2021). Identification and behavioral assays of alarm pheromone in the vetch aphid *Megoura viciae*. J. Chem. Ecol. *47*, 740–746. https://doi.org/10.21203/rs.3.rs-467323/v1.
27. Vandermoten, S., Mescher, M.C., Francis, F., Haubruge, E., and Verheggen, F.J. (2012). Aphid alarm pheromone: an overview of current knowledge on biosynthesis and functions. Insect Biochem. Mol. Biol. *42*, 155–163. https://doi.org/10.1016/j.ibmb.2011.11.008.
28. Sun, Z.J., and Li, Z.X. (2017). Host plants and obligate endosymbionts are not the sources for biosynthesis of the aphid alarm pheromone. Sci. Rep. *7*, 6041. https://doi.org/10.1038/s41598-017-06465-9.
29. Sun, C.X., and Li, Z.X. (2021). Biosynthesis of aphid alarm pheromone is modulated in responseto starvation stress under regulation by the insulin, glycolysis and isoprenoid pathways. J. Insect Physiol. *128*, 104174. https://doi.org/10.1016/j.jinsphys.2020.104174.
30. Song, X., Qin, Y.G., Zhang, Y.H., Zhou, Y.B., and Li, Z.X. (2023). Farnesyl/geranylgeranyl diphosphate synthases regulate the biosynthesis of alarm pheromone in a unique manner in the vetch aphid *Megoura viciae*. Insect Mol. Biol. *32*, 229–239. https://doi.org/10.1111/imb.12826.
31. Manes, T., Glade, K., Ziomek, C.A., and Millán, J.L. (1990). Genomic structure and comparison of mouse tissue-specific alkaline phosphatase genes. Genomics *8*, 541–554. https://doi.org/10.1016/0888-7543(90)90042-s.
32. Almagro Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O., Brunak, S., von Heijne, G., and Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat.

Biotechnol. *37*, 420–423. https://doi.org/10.1038/s41587-019-0036-z.
33. Hirata, T., and Kizuka, Y. (2021). *N*-Glycosylation. Adv. Exp. Med. Biol. *1325*, 3–24. https://doi.org/10.1007/978-3-030-70115-4_1.
34. Kreil, G. (1984). Occurrence, detection, and biosynthesis of carboxy-terminal amides. Methods Enzymol. *106*, 218–223. https://doi.org/10.1016/0076-6879(84)06023-7.
35. Bradbury, A.F., and Smyth, D.G. (1987). Biosynthesis of the C-terminal amide in peptide hormones. Biosci. Rep. *7*, 907–916. https://doi.org/10.1007/bf01122123.
36. Cooper, J.A., Esch, F.S., Taylor, S.S., and Hunter, T. (1984). Phosphorylation sites in enolase and lactate dehydrogenase utilized by tyrosine protein kinases *in vivo* and *in vitro*. J. Biol. Chem. *259*, 7835–7841. https://doi.org/10.1016/s0021-9258(17)42869-9.
37. Ruoslahti, E., and Pierschbacher, M.D. (1986). Arg-Gly-Asp: a versatile cell recognition signal. Cell *44*, 517–518. https://doi.org/10.1016/0092-8674(86)90259-x.
38. Cheah, E., Carr, P.D., Suffolk, P.M., Vasudevan, S.G., Dixon, N.E., and Ollis, D.L. (1994). Structure of the *Escherichia coli* signal transducing protein PII. Structure *2*, 981–990. https://doi.org/10.2210/pdb1pil/pdb.
39. Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. *157*, 105–132. https://doi.org/10.1016/0022-2836(82)90515-0.
40. Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. *305*, 567–580. https://doi.org/10.1093/bioinformatics/bti303.
41. von Heijne, G. (2006). Membrane-protein topology. Nat. Rev. Mol. Cell Biol. *7*, 909–918. https://doi.org/10.1002/pro.5560050221.
42. Tillman, J.A., Seybold, S.J., Jurenka, R.A., and Blomquist, G.J. (1999). Insect pheromones—an overview of biosynthesis and endocrine regulation. Insect Biochem. Mol. Biol. *29*, 481–514. https://doi.org/10.1016/s0965-1748(99)00016-8.
43. Gilg, A.B., Bearfield, J.C., Tittiger, C., Welch, W.H., and Blomquist, G.J. (2005). Isolation and functional expression of an animal geranyl diphosphate synthase and its role in bark beetle pheromone biosynthesis. Proc. Natl. Acad. Sci. USA *102*, 9760–9765. https://doi.org/10.1073/pnas.0503277102.
44. Hall, G.M., Tittiger, C., Blomquist, G.J., Andrews, G.L., Mastick, G.S., Barkawi, L.S., Bengoa, C., and Seybold, S.J. (2002). Male Jeffrey pine beetle, *Dendroctonus jeffreyi*, synthesizes the pheromone component frontalin in anterior midgut tissue. Insect Biochem. Mol. Biol. *32*, 1525–1532. https://doi.org/10.1016/s0965-1748(02)00073-5.
45. Nardi, J.B., Young, A.G., Ujhelyi, E., Tittiger, C., Lehane, M.J., and Blomquist, G.J. (2002). Specialization of midgut cells for synthesis of male isoprenoid pheromone components in two scolytid beetles, *Dendroctonus jeffreyi* and *Ips pini*. Tissue Cell *34*, 221–231. https://doi.org/10.1016/s0040-8166(02)00004-6.
46. Sun, Z.J., and Li, Z.X. (2018). The terpenoid backbone biosynthesis pathway directly affects the biosynthesis of alarm pheromone in the aphid. Insect Mol. Biol. *27*, 824–834. https://doi.org/10.1111/imb.12521.
47. Song, X., and Li, Z.X. (2022). Functional characterization of two different decaprenyl diphosphate synthases in the vetch aphid *Megoura viciae*. Arch. Insect Biochem. Physiol. *110*, e21900. https://doi.org/10.1002/arch.21900.
48. Pirovano, W., Feenstra, K.A., and Heringa, J. (2008). PRALINE™: a strategy for improved multiple alignment of transmembrane proteins. Bioinformatics *24*, 492–497. https://doi.org/10.1093/bioinformatics/btm636.
49. Trifinopoulos, J., Nguyen, L.T., von Haeseler, A., and Minh, B.Q. (2016). W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res. *44*, W232–W235. https://doi.org/10.1093/nar/gkw256.
50. Nguyen, L.T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. Mol. Biol. Evol. *32*, 268–274. https://doi.org/10.1093/molbev/msu300.
51. Letunic, I., and Bork, P. (2021). Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. *49*, W293–W296. https://doi.org/10.1093/bioinformatics/btl529.
52. Hulo, N., Bairoch, A., Bulliard, V., Cerutti, L., De Castro, E., Langendijk-Genevaux, P.S., Pagni, M., and Sigrist, C.J.A. (2006). The PROSITE database. Nucleic Acids Res. *34*, 227–230. https://doi.org/10.1093/nar/gkj063.
53. Singh, A.K., Kumar, R., Tripathi, A.K., Gupta, B.K., Pareek, A., and Singla-Pareek, S.L. (2015). Genome-wide investigation and expression analysis of Sodium/Calcium exchanger gene family in rice and Arabidopsis. Rice *8*, 54. https://doi.org/10.1186/s12284-015-0054-5.
54. Cristiano, G., Grossi, G., Scala, A., Fanti, P., and Falabella, P. (2016). Validation of reference genes for qRT-PCR analysis in *Megoura viciae* (Hemiptera: Aphididae). B Insectol *69*, 229–238. https://doi.org/10.7717/peerj.10395/fig-5.
55. Bruno, D., Grossi, G., Salvia, R., Scala, A., Farina, D., Grimaldi, A., Zhou, J.J., Bufo, S.A., Vogel, H., Grosse-Wilde, E., et al. (2018). Sensilla morphology and complex expression pattern of odorant binding proteins in the vetch aphid *Megoura viciae* (Hemiptera: Aphididae). Front. Physiol. *9*, 777. https://doi.org/10.3389/fphys.2018.00777.
56. Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. *29*, e45–e2007. https://doi.org/10.1093/nar/29.9.e45.
57. Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using realtime quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods *25*, 402–408. https://doi.org/10.1006/meth.2001.1262.
58. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589. https://doi.org/10.3410/f.740477161.793588825.
59. Colovos, C., and Yeates, T.O. (1993). Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. *2*, 1511–1519. https://doi.org/10.1002/pro.5560020916.
60. Yang, J., Roy, A., and Zhang, Y. (2013). Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. Bioinformatics *29*, 2588–2595. https://doi.org/10.1093/bioinformatics/btt447.
61. O'Boyle, N.M., Banck, M., James, C.A., Morley, C., Vandermeersch, T., and Hutchison, G.R. (2011). Open Babel: An open chemical toolbox. J. Cheminf. *3*, 33. https://doi.org/10.1186/1758-2946-3-33.
62. Trott, O., and Olson, A.J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. J. Comput. Chem. *31*, 455–461. https://doi.org/10.1002/jcc.21334.
63. Aguilar, O.A., Berry, R., Rahim, M.M.A., Reichel, J.J., Popović, B., Tanaka, M., Fu, Z., Balaji, G.R., Lau, T.N.H., Tu, M.M., et al. (2017). A viral immunoevasin controls innate immunity by targeting the prototypical natural killer cell receptor family. Cell *169*, 58–71.e14. https://doi.org/10.3410/f.727432488.793530003.
64. Laskowski, R.A., and Swindells, M.B. (2011). LigPlot^{+}: multiple ligand-protein interaction diagrams for drug discovery. J. Chem. Inf. Model. *51*, 2778–2786. https://doi.org/10.1021/ci200227u.
65. Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. *72*, 248–254. https://doi.org/10.1016/0003-2697(76)90527-3.
66. Yan, S., Qian, J., Cai, C., Ma, Z., Li, J., Yin, M., Ren, B., and Shen, J. (2020). Spray method application of transdermal dsRNA delivery system for efficient gene silencing and pest control on soybean aphid *Aphis glycines*. J. Pest. Sci. *93*, 449–459. https://doi.org/10.1007/s10340-019-01157-x.
67. Yuan, X., Zhang, M., Wang, Y., Zhao, H., and Sun, D. (2019). Using co-axial electrospray deposition to eliminate burst release of simvastatin from microparticles and to enhance induced osteogenesis. J. Biomater. Sci. Polym. Ed. *30*, 355–375. https://doi.org/10.1080/09205063.2018.1559978.
68. Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. *29*, 644–652. https://doi.org/10.3410/f.13296969.14657090.
69. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009). BLAST^{+}: architecture and applications. BMC Bioinf. *10*, 421. https://doi.org/10.1186/1471-2105-10-421.
70. Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian

transcriptomes by RNASeq. Nat. Methods *5*, 621–628. https://doi.org/10.1038/nmeth.1226.

71. Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for dierential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140. https://doi.org/10.1093/bioinformatics/btp616.

72. Ma, X., Wang, P., Zhou, S., Sun, Y., Liu, N., Li, X., and Hou, Y. (2015). De novo transcriptome sequencing and comprehensive analysis of the drought-responsive genes in the desert plant *Cynanchum komarovii*. BMC Genom. *16*, 753. https://doi.org/10.1186/s12864-015-1873-x.

73. Benjamini, Y., and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. Ann. Stat. *29*, 1165–1188. https://doi.org/10.1214/aos/1013699998.

74. Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinf. *12*, 323. https://doi.org/10.1186/1471-2105-12-323.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
M. viciae feed on *Pisum sativum*	This paper	N/A
Bacterial and virus strains		
Escherichia coli Trans1-T1	TransGen Biotech	CD501
Escherichia coli BL21(DE3)	TransGen Biotech	CD601
Chemicals, peptides, and recombinant proteins		
TRIzol reagent	Invitrogen	15596026CN
AG RNAex Pro reagent	Accurate Biotechnology	AG21101
(*E*, *E*)-FPP	Sigma-Aldrich	44270
(*E*, *E*)-β-farnesene	Sigma-Aldrich	73492
(*E*, *E*)-farnesol	Sigma-Aldrich	277541
SAP	TaKaRa Biotechnology	2660A
nanocarrier SPc	Research Group of Insect Development in the Department of Entomology, China Agricultural University	N/A
Critical commercial assays		
2×KODmix-plus	Sino-US Taihe Biotechnology	PC203-01
pCloneEZ-NRS-Blunt-Kan/HC	Sino-US Taihe Biotechnology	C5852-50
AidQuick Gel Extraction kit	Aidlab	DR0401
SYBR Green Pro Taq HS Mix	Accurate Biotechnology	AG11701
GoldTM His-tag purification resin	Beyotime	P2210FT
Cell lysis buffer for Western and IP without inhibitors	Beyotime	P0013J
Alkaline Phosphatase Assay Kit	Beyotime	P0321S
T7 RiboMAX™ Express RNAi System	Promega	P1700
Oligonucleotides		
Primers are listed in Table S1	This paper	N/A
Software and algorithms		
Prism 8	Graphpad	https://www.graphpad.com/scientific-software/prism/
OpenBabel v. 3.1.1	N/A	https://openbabel.org/
AutoDocktools v 1.5.7	Vina	https://autodocksuite.scripps.edu/adt
Pymol	N/A	https://www.pymol.org/pymol
LigPlot^{+}	Java	https://www.ebi.ac.uk/thornton-srv/software/LigPlus/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Zheng-Xi Li (zxli@cau.edu.cn).

Materials availability

All unique reagents or aphid generated in this study are available from the lead contact without restriction.

Data and code availability

All data reported in this paper will be shared by the lead contact upon request.
This paper does not report original code.
Any additional information required to re-analyse the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Insects

M. viciae (Hemiptera: Aphididae) (wild type) was obtained from Hebei Agricultural University and reared on *Pisum sativum* L. in the Laboratory of Insect Molecular Ecology in China Agricultural University, Beijing. The aphids were maintained under the conditions of 70±5% R.H. and 16L: 8D photoperiod at 19±1 °C in a growth chamber (RXZ-300B, Ningbo, China).[47] These aphids had agamogenesis mode. The 1st-instar nymphs of aphids (n = 30) were moved to *P. sativum* and placed in 1.5% (w/v) agar in a box (13.5 cm × 11.8 cm × 6.5 cm). The 3rd-instar nymphs were used in this study.

METHOD DETAILS

RNA extraction

Total RNA was extracted from different tissues of 2-day-old adult aphids (head, epicuticle, cuticle, midgut, cornicle and embryo) and from whole aphid bodies at different developmental stages (1st-4th instar nymphs and 2-day-old adult), respectively, using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. The integrity of RNA templates was examined using 1.2% agarose and checked using a NanoDrop 2000 spectrophotometer (NanoDrop, Wilmington, DE, USA). First-strand cDNAs were synthesized from 500 ng RNAs by using the HiScript III qRT superMix with gDNase (Vazyme, Nanjing, China) and stored at -80 °C for further use.[47]

Sequencing of ALP genes in *M. viciae*

The primers for amplifying the ALP genes of *M. viciae* were designed according to its transcriptome data previously sequenced by our group (Table S1).[47] The polymerase chain reaction (PCR) procedures were performed as followed: predenaturation at 95 °C for 2 min, followed by 35 cycles of amplification (95°C for 10 s, 55°C for 30 s, and 68°C for 1 min), and a final extension at 68 °C for 7 min. The reaction system consisted of 25 μL 2 × KODmix-plus (Sino-US Taihe Biotechnology, Beijing), 1 μL cDNA template, 1 μL of each forward and reverse primer (10 μmol/L), and 22 μL RNA-free H_2O. PCR products were checked by 1.2% agarose gel electrophoresis and recovered by AidQuick Gel Extraction kit. The purified product was ligated to the vector pCloneEZ-NRS-Blunt-Kan/HC (Sino-US Taihe Biotechnology, Beijing) and then transferred to *Escherichia coli* Trans1-T1 (TransGen Biotech, Beijing, China) for sequencing by Sino-US Taihe Biotechnology Company (Beijing, China).

Phylogenetic analysis

Alignment of the amino acid sequences of ALPs was carried out on the PRALINE online server (http://www.ibi.vu.nl/programs/pralinewww/).[48] To construct a maximum-likelihood tree, aligned sequences were analyzed using IQ-TREE web server (http://iqtree.cibiv.univie.ac.at/).[49] The reliability of nodes were assessed by bootstrap resampling (1000 times).[50] The resulting tree was further edited using the Interactive Tree of Life (iTOL, https://itol.embl.de).[51] The reference sequences included SAP, membrane-bound and soluble, intestinal-type and tissue-nonspecific isozyme alkaline phosphatases (Table S2). Analysis of ALP motifs was performed on the Novopro online servers (https://www.novopro.cn/tools/motifscan.html) and Protter (http://wlab.ethz.ch/protter/).[52,53]

Spatiotemporal expression profiling of ALP genes in *M. viciae*

The spatiotemporal expression patterns of *M. viciae* ALP genes were analyzed by real-time quantitative PCR (RT-qPCR) on the ABI QuantStudio 6 Flex Detection system. The reactions were performed in a volume of 20 μL containing 10 μL of SYBR Green Pro Taq HS Mix, 0.4 μL ROX Reference Dye (4 μM), 6.6 μL RNA-free H_2O, 1.0 μL forward and reverse primers (10 μmol/L), and 1.0 μL cDNA template (500 ng). The thermal cycling program consisted of denaturation at 95°C for 20 s, followed by 40 cycles of 95°C for 5 s and 60°C for 30 s. The melt curve stage was set up as followed: 95 °C for 15 s, 60 °C for 1 min, and 95 °C for 15 s. The ribosomal protein L32 of *A. pisum* (GenBank acc. no. NM001126210.2) was used as the internal reference gene.[54,55] The primers used in this study were listed in Table S1. The specificity and amplification efficiency of the primers were assessed using regression coefficient (R^2) (Figure S1).[56] The relative expression levels of target genes were quantitatively analyzed in different tissues and during different developmental stages using the $2^{-\Delta\Delta CT}$ method.[57] The head and the 1st-instar nymph were used as the calibrator samples for tissue- and stage-specific RT-qPCR analysis, respectively. Three biological replicates were conducted for each sample.

Molecular modeling and docking

The three-dimension (3D) structures of ALP proteins were modeled via the ColabFold: AlphaFold2 using MMseqs2 (https://alphafold.com/).[58] The stereochemical quality of ALP models were validated (ERRAT and PROCHECK) using the SAVES v6.0 structure server (https://saves.mbi.ucla.edu/).[59] The active site architectures and binding residues for divalent metal ions (MG) were evaluated by COACH online server (https://seq2fun.dcmb.med.umich.edu/COACH/).[60] The structure data file (SDF) of the target compound was downloaded from the PubChem

database (https://pubchem.ncbi.nlm.nih.gov/). The SDF was then converted to Mol2 format using OpenBabel v. 3.1.1.[61] Then the modeled structures and ligands were converted to PDBQT format on AutoDockTools.[62] Docking was carried out using Vina integrated with AutoDocktools v 1.5.7 and the docking scores showed the interaction energies (kcal/mol). The top 10 poses of the ligand (*E, E*)-FPP were ranked based on the lowest binding affinity. The interaction between receptor and ligand was analyzed and visualized using Pymol and LigPlot^{+} to draw 3D and 2D pictures, respectively.[63,64]

Recombinant expression and purification of ALPs

ALP genes were fused with the vector pET28a(+) cut with *Bam*H I and *Hind* III (TaKaRa Biotechnology, Beijing, China). The fused expression vectors were transformed into *Escherichia coli* BL21(DE3) (TransGen Biotech, Beijing, China). Single positive colonies were cultured in 200 mL Luriae-Bertani media (LB, pH7.3) with 50 mg/mL kanamycin at 37°C (200 rpm), and then 0.1 mM isopropylthio-b-D-galactoside (IPTG) was added to the culture medium, which was allowed to grow under 18°C and 150 rpm until the OD600 reached 0.1. The cells were collected from liquid culture by centrifugation at 12,000 g for 30 min at 4°C and resuspended in 50 mM Tris-HCl buffer (pH7.3). After sonication and centrifugation (16,000 g for 30 min at 4°C), the protein was purified by GoldTM His-tag purification resin (Beyotime, Shanghai, China) with 50 mM imidazole (Solarbio, Beijing, China) buffer (50 mM Tris-HCl buffer with imidazole added to a final concentration of 50 mM, pH7.3). Fractions were analyzed by SDS-PAGE (Figure S2). Purified ALP proteins were resuspended in the storage buffer (25 mM Tris-HCl, pH 7.0, 25 mM $MgCl_2$, 1 mM DL-1,4-dithiothreitol and 50% glycerol).[30] Their concentrations were determined at OD595 on the SpectraMax I3x using the Bradford method.[30,65]

In vitro enzymatic activity assay of ALPs

The enzymatic reaction system contained recombinant ALP enzymes (5.0 μM) in the assay buffer (25 mM Tris-HCl, pH9.0, 5 mM $MgCl_2$ and 50% glycerol) and the substrate (*E, E*)-FPP (Sigma-Aldrich, Oakville, Canada) in a 1.5-mL centrifuge tube. The reaction mixture was incubated at 37°C (200 rpm) for 21 h. The non-polar solution *n*-hexane (100 μL) was then added to extract the products. The supernatant after centrifuge at 4°C for 30 min was analyzed by Gas chromatography-mass spectrometry (GC-MS). The enzymatic reaction with SAP (TaKaRa Biotechnology, Beijing, China) was used as the positive control.

Product analysis by GC-MS

The supernatant (50 μL) from *in vitro* enzymatic reaction was transferred into a 2-mL vial for product analysis by GC-MS on an Agilent 7890B/7200 (Agilent Technologies Inc., California, USA). The samples were analysed on an Agilent, equipped with an HP-5 capillary column (300 mm × 0.25 mm × 0.25 μm). The program of GC-MS was the same as described earlier.[30] Briefly, the sample (2 μL) was injected with helium as carrier gas at a split flow of 2:1. The oven temperature program consisted of 40°C (hold for 1 min), increased to 220°C at a rate of 4°C/min and then to 280°C at a rate of 11°C/min (hold for 10 min). The injector and ion source temperatures were set to 250°C. The MS was operated with electron impact ionization (EI, 60 eV) and a scan range of *m/z* 20-1020. The pure external standards of (*E, E*)-β-farnesene and (*E, E*)-farnesol (Sigma-Aldrich, Oakville, Canada) were analyzed under the same conditions.[30] The amount of pheromone component was estimated by comparing the peak area ratio of the sample to the standard. Three biological replicates were performed for each reaction.

Functional analysis of MvALP4 by RNA interference (RNAi)

The functional MvALP4 gene verified by *in vitro* enzymatic activity assay was further analyzed by RNAi. The double-stranded RNAs (*dsMvALP4*) of ALP gene and green fluorescent protein gene (KX247384) were synthesized by using the T7 RiboMAX™ Express RNAi System (Promega, Madison, WI, USA) and the primers listed in Table S1. The 3rd-instar nymphs were used for RNAi analysis at four different time points (12 h, 24 h, 36 h and 48 h). The relative gene expression level was analyzed by RT-qPCR, and the amounts of aphid alarm pheromone components were measured by GC-MS. Briefly, dsRNAs (working concentration: 500 ng/μL) were mixed with the nanocarrier SPc (provided by the Research Group of Insect Development in the Department of Entomology, China Agricultural University) at a mass ratio of 1:1, and added to a formulation containing 0.3% (v/v) detergent.[30,66] The mixture (40 nL) was then applied on the notum of aphid by using a Nanoliter 2000 microinjector (World Precision Instruments, Sarasota, USA) on ice and kept for 15 min. The treated nymphs were moved to plants after the solution was assimilated. The protocol for GC-MS after RNAi was the same as described above. Three biological replicates were performed for each group.

Enzymatic activity assay of crude proteins

The 3rd-instar nymphs (n = 100) were fully milled in the Cell lysis buffer for Western and IP without inhibitors P0013J (Beyotime, Shanghai, China) (100 μL) in a 1.5-mL centrifuge tube on ice, which was then centrifuged at 4°C for 30 min. The supernatant (crude proteins) was moved into a clean tube for setting up an enzymatic reaction system with (*E*)-FPP as the substrate as described above. The supernatant without adding (*E*)-FPP was used as the negative control.

Quantitative analysis of ALPs and alarm pheromone component in *M. viciae* under simulated stimulation

Under a stereomicroscope (SMZ-850; Olympus; Japan), the 3rd-instar nymphs were gently pressed on its back with a needle to artificially stimulate them to emit droplets from the cornicle (n = 50).[46] Subsequently, ALPs and aphid alarm pheromone components were quantitatively

analyzed at different time points (0 h, 6 h, 12 h, 24 h and 48 h) under stimulated and unstimulated conditions, respectively. The unstimulated groups were used the negative control. The amount of (*E*, *E*)-β-farnesene (alarm pheromone component) was for GC-MS analysis as described above. The Alkaline Phosphatase Assay Kit P0321S (Beyotime, Shanghai, China) was used to quantify ALPs. The fully milled aphids in the Cell lysis buffer (200 μL) was centrifuged (12000 rpm) at 4°C for 30 min. The supernatant (50 μL) and chromogenic substrate P0321S-2 (50 μL) were incubated at 37°C for 30 min, and then 100 μL stop-reaction solution (P0321S-4) was added. The mixture turned into yellow, and the quantity of ALPs was determined at OD595 nm on the SpectraMax I3x (Molecular devices, California, USA). ALPs were also quantified during different developmental stages (1^{st}-4^{th} instars and 2-day-old adult) (Figure S3A). P0321S-3 (*p*-nitrophenol) was used as the standard (Figure S3B).[67] Six biological replicates were carried out for each group.

Transcriptome sequencing in response to simulated stimulation

Total RNA (150 μg) was extracted from the 3^{rd}-instar nymphs at 24 h after simulated stimulation by using *AG RNAex Pro* reagent (Accurate Biotechnology, Changsha, China) according to the manufacturer instructions. The unstimulated aphid was used as the negative control. Transcriptome sequencing was completed by BGI Company (Beijing, China).[68] All clean reads data were used to predict unigenes by Blastx search against the Nr, Nt, Swiss-Prot, Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases.[69] Functional annotation and classification for unigenes were performed by using KEGG. The read counts were further used to quantify transcript abundance into FPKM (Fragments per kilobase of exon model per million mapped reads) values in two groups.[70] DEGseq provided statistical routines for R package in differential gene expression analysis.[71,72] The *P*-value set the threshold for the difierential gene expression test. The fold change (FC>1.0) and the false discovery rate (FDR $\leq$0.001) were the cut-off conditions for DEGseq.[73,74] For KEGG classification of differentially expressed genes (DEGs), the DEGs were mapped to terms in the KEGG database.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis

All statistical analyses and preliminary graphs were made using the software GraphPad Prism 8.0 (GraphPad Software, San Diego, CA). Statistical analyses on the differences in the relative gene expression levels of different developmental stages and different tissues were analyzed using ANOVA and Tukey's B multiple range test (Figure 2). Molecular modeling and docking were carried out using Vina integrated with AutoDocktools v 1.5.7 (Figure 3). The relative expression levels of knockdown genes were statistically compared by using the Student's *t*-test (Figure 7). The differences in the amounts of ALPs and aphid alarm pheromone components between different groups also were analyzed using Student's *t*-test (Figure 9). The statistical details and the replicate number of experiments can be found in the figure legends.