

RESEARCH ARTICLE

Open Access



# Molecular diversity and function of jasmintides from *Jasminum sambac*

Geeta Kumari<sup>†</sup>, Ka Ho Wong<sup>†</sup>, Aida Serra, Joon Shin, Ho Sup Yoon, Siu Kwan Sze and James P. Tam<sup>\* id</sup>

## Abstract

**Background:** Jasmintides jS1 and jS2 from *Jasminum sambac* were previously identified as a novel family of cysteine-rich peptides (CRPs) with an unusual disulfide connectivity. However, very little else is known about jasmintides, particularly their molecular diversity and functions. Here, we report the discovery and characterization of a novel suite of jasmintides from *J. sambac* using transcriptomic, peptidomic, structural and functional tools.

**Results:** Transcriptomic analysis of leaves, flowers and roots revealed 14 unique jasmintide precursors, all of which possess a three-domain architecture comprising a signal peptide, a pro-domain and a mature jasmintide domain. Peptidomic analysis, using fractionated mixtures of jasmintides and chemical derivatization of cysteine to pseudolysine, trypsin digestion and MS/MS sequencing, revealed an additional 86 jasmintides, some of which were post-translationally modified. NMR analysis showed that jasmintide jS3 has three anti-parallel  $\beta$ -strands with a three-disulfide connectivity of CysI–CysV, CysII–CysIV and CysIII–CysVI, which is similar to jasmintide jS1. Jasmintide jS3 was able to withstand thermal, acidic and enzymatic degradation and, importantly, exhibited antifeedant activity against mealworm *Tenebrio molitor*.

**Conclusion:** Together, this study expands the existing library of jasmintides and furthers our understanding of the molecular diversity and cystine framework of CRPs as scaffolds and tools for engineering peptides targeting pests.

**Keywords:** Jasmintides, Cysteine-rich peptides, Antifeedant, Peptide biosynthesis, *Jasminum sambac*

## Background

*Jasminum sambac*, which belongs to the Oleaceae family, is a small shrub with fragrant white flowers native to India. Jasmine flowers, regarded as the “queen of flowers”, have traditionally been used for relieving symptoms caused by fever, sunburn, stomach ulcers and anxiety [1]. Its leaves are known to exhibit anti-inflammatory and analgesic properties in rat models [2]. A diverse array of secondary metabolites has been isolated from the flowers and leaves of *J. sambac*, including flavonoids, polyphenols and glycosides [3].

Plants have been a rich source of bioactive compounds and inspiring structures for drug design and discovery [4]. Historically, a major focus of plant-based medicinal research has been directed towards small-molecule metabolites. Recently, peptides have attracted interest as potential therapeutic leads from plants [5–16]. Peptides, however, are generally susceptible to proteolytic degradation, leading to

poor metabolic stability or oral bioavailability [6, 9, 11]. In contrast, cysteine-rich peptides (CRPs), particularly those constrained by three to five disulfide bonds, are known to overcome this limitation [6, 9, 14–16]. Consequently, our laboratory has focused on this superfamily of metabolically stable plant-derived peptides, particularly CRPs with molecular weights of 2 to 6 kDa. CRPs in this chemical space generally contain a well-defined tertiary structure due to the cross-embraced multiple disulfide bonds that provide them unprecedented stability against thermal and proteolytic degradation [14–17]. Of particular interest are plant CRPs, which also display a wide range of biological and pharmacological activities [4, 13–16]. As potential therapeutics, plant CRPs with a larger molecular size than small-molecule metabolites have the advantages of increased on-target specificity and decreased off-target toxicity. An additional appeal in discovering novel plant CRPs is their highly stable and evolutionarily conserved structures, which are known to serve as scaffolds that are useful for grafting labile bioactive peptides to increase their metabolic stability in drug development [14–16].

\* Correspondence: [jptam@ntu.edu.sg](mailto:jptam@ntu.edu.sg)

<sup>†</sup>Geeta Kumari and Ka Ho Wong contributed equally to this work.  
School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551, Singapore



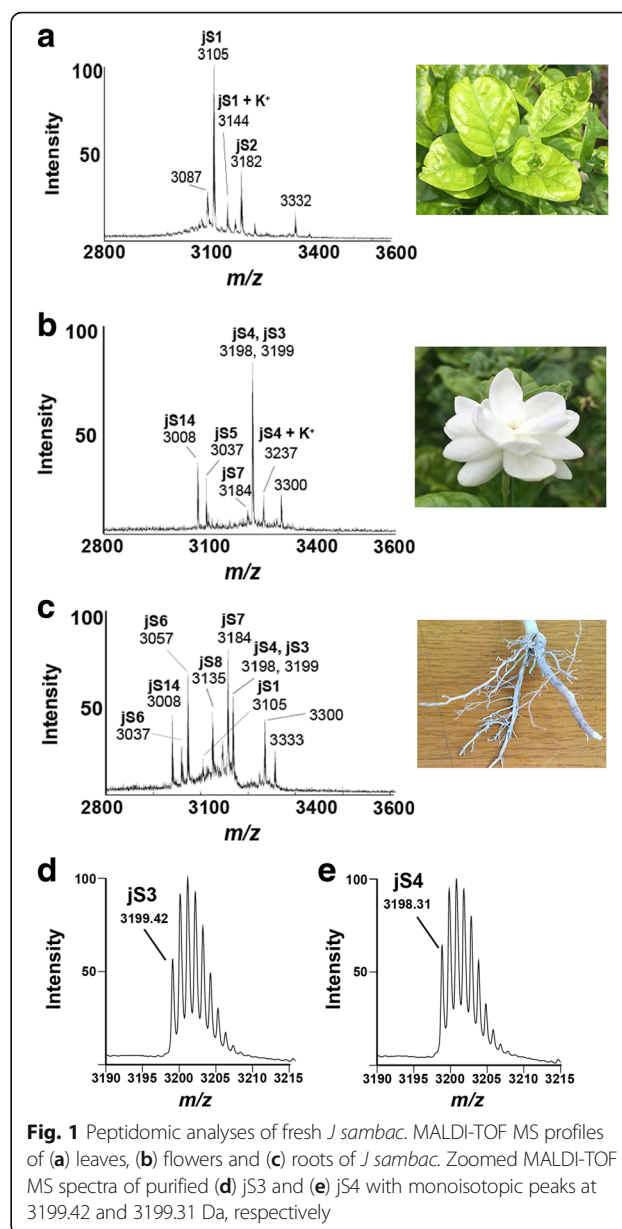
CRPs are classified into different families according to the number of cysteine residues and their disulfide connectivity [4]. Cystine knot peptides, well represented by knottins and hevein-like peptides, are the most commonly encountered six-cysteine CRPs (6C-CRPs) with a knotted disulfide connectivity between CysI–CysIV, CysII–CysV, and CysIII–CysVI [18]. A second family is represented by thionins, comprising a symmetric disulfide pattern or an onion-like topology between CysI–CysVI, CysII–CysV and CysIII–CysIV [19]. Recently, our laboratory reported the discovery of three novel patterns of disulfide threads, including jasmintides,  $\beta$ -ginkgotides and lybatides [7, 13, 16]. The (*Jasminum sambac* peptides) jasmintides, which contain disulfide connectivity between CysI–CysV, CysII–CysIV and CysIII–CysVI, are 6C-CRPs with a 27 amino acid length and a molecular weight of 3.1 kDa, isolated from *J. sambac* leaves [7]. The  $\beta$ -ginkgotides, which contain disulfide connectivity between CysI–CysIV, CysII–CysVI and CysIII–CysV, are hyperdisulfide-constrained peptides with a 20 amino acid length and a molecular weight of 2 to 3 kDa, isolated from *Ginkgo biloba* nuts [16]. The lybatides, which contain disulfide connectivity between CysI–CysVI, CysII–CysVIII, CysIII–CysVII and CysIV–CysV, are cystine-stapled helical peptides with a 33 amino acid length and a molecular weight of 3.6 kDa, isolated from *Lycium barbarum* root bark [13]. Thus far, only two jasmintides (jS1 and jS2) have been reported [7], and little is known about their molecular diversity, precursor sequences, distribution and functions.

Here, we report the molecular diversity, structure and function of jasmintides. We identified 14 non-redundant jasmintide precursors at transcriptomic level which encoding jasmintides jS1–jS15 and 86 novel jasmintides at peptidomic level. NMR structural determination showed that jasmintide jS3 is composed of three anti-parallel sheets constrained by three disulfide bonds with a connectivity similar to that of the previously reported jasmintide jS1 [7]. We also showed that jS3 displays a potent antifeedant effect on *Tenebrio molitor* mealworms. Our results expand the library of jasmintides and demonstrate that jasmintides are much more diverse than we previously thought. The highly stable jasmintide structure can potentially be used as a scaffold for grafting bioactive peptides.

## Results

### Isolation of jasmintides jS3 and jS4

Fresh tissues from the roots, flowers and leaves of *J. sambac* were extracted using 50% (v/v) ethanol, purified with C<sub>18</sub> Ziptip and profiled by MALDI-MS. Distinct MS profiles were observed from different parts of *J. sambac* extracts (Fig. 1). A cluster of peaks



**Fig. 1** Peptidomic analyses of fresh *J. sambac*. MALDI-TOF MS profiles of (a) leaves, (b) flowers and (c) roots of *J. sambac*. Zoomed MALDI-TOF MS spectra of purified (d) jS3 and (e) jS4 with monoisotopic peaks at 3199.42 and 3198.31 Da, respectively

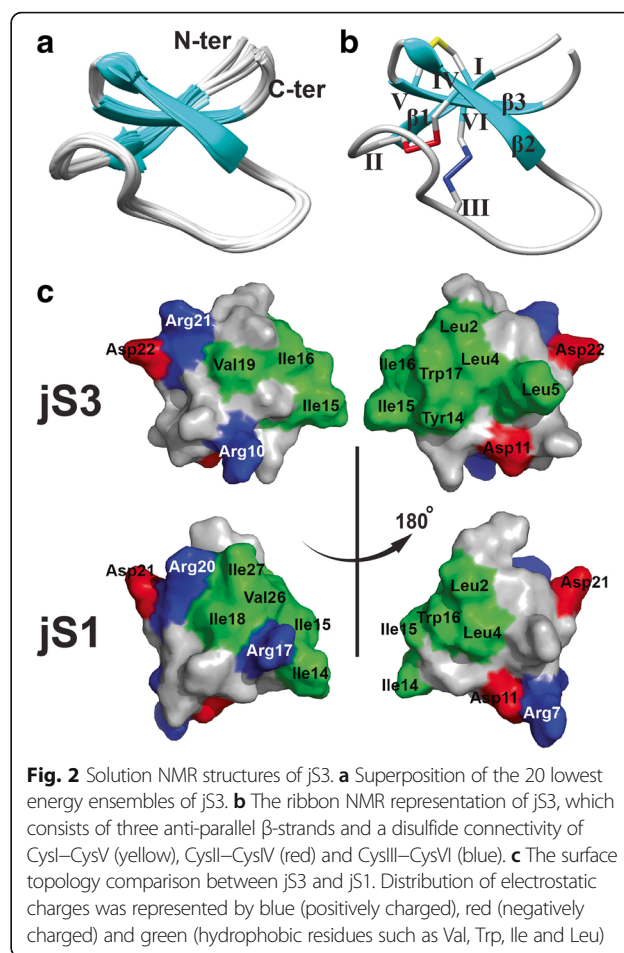
was observed between 3000 and 3400 Da of *J. sambac* flowers. To further characterize these compounds, a scale-up extraction of jasmintides was performed using fresh flowers of *J. sambac*, which was then fractionated by reversed-phase flash chromatography and purified by repeated reversed-phase and strong cation-exchange high-performance liquid chromatography. Jasmintides jS3 and jS4, with relative monoisotopic molecular masses  $[M + H]^+$  of 3199.42 and 3198.31 Da, were isolated, yielding approximately 40 and 20 mg per kg of fresh material, respectively. Jasmintide jS3 was chosen as the representative for further characterization due to its higher abundance than jS4.

### Solution NMR structure of jasmintide jS3

Two-dimensional nuclear Overhauser effect spectroscopy (NOESY) and total correlation spectroscopy (TOCSY) spectra of jasmintide jS3 with ~98% proton resonance were unambiguously assigned. Additional file 1: Table S1 and Additional file 2: Figure S1 illustrate the complete amide proton assignment of the NOESY spectrum. The solution structure of jS3 was determined based on a total of 503 NMR-derived distance restraints and 23 dihedral angle restraints (Table 1). PROCHECK analysis indicated that all residues were distributed in the allowed region of the Ramachandran map. Figure 2a shows the 20 ensembles of jS3 with root-mean-square deviation (RMSD) values of 0.29 Å for backbone atoms and 1.00 Å for all heavy atoms. On the basis of the strong  $\alpha\text{N}$  ( $i, i+1$ ) NOE connectivity, coupling constants measured by DQF-COSY and 1H-1D experiments, and hydrogen bond patterns determined by amide-hydrogen exchange experiments, jS3 was determined to be composed of a compact disulfide-constrained structure. The overall structure of jS3 consisted of three short antiparallel  $\beta$ -strands ( $\beta$ 1: C3 – L5,  $\beta$ 2: T18 – R21, and  $\beta$ 3: C24 – N26) and two loop segments (L1: C6 – W17 between  $\beta$ 1 and  $\beta$ 2 and L2: D22 – G23 between  $\beta$ 2 and  $\beta$ 3), as shown in Fig. 2b and Additional file 3: Figure S2. The disulfide connectivity of jS3 was confirmed by  $d\beta\beta$  ( $i, j$ ) and  $d\alpha\beta$  ( $i, j$ ) NOE cross-peaks, as illustrated in Additional file 4: Figure S3. NOE signals between  $\text{H}\beta - \text{H}\beta$  from Cys3/Cys24 and Cys6/Cys20 were observed, which corresponded to the CysI–CysV and CysII–CysIV cysteine pairs, respectively.

**Table 1** Structural statistics of NMR structures of jasmintide jS3

Number of NOE constraints	
All	503
Intra residues $ i-j  = 0$	139
Sequential, $ i-j  = 1$	141
Medium-range, $1 <  i-j  < 5$	48
Long-range, $ i-j  > 5$	175
Number of hydrogen bond constraints	14
Number of dihedral angle constraints	23
Number of constraint violations ( $> 0.5$ Å)	0
Number of angle violations ( $> 5^\circ$ )	0
Energies (kcal/mol)	
$E_{\text{NOE}}$	$0.74 \pm 0.33$
$E_{\text{cdih}}$	$0.02 \pm 0.01$
$E_{\text{bond}} + E_{\text{angle}} + E_{\text{improper}}$	$18.53 \pm 0.78$
$E_{\text{VDW}}$	$7.27 \pm 1.20$
RMS Deviation of the structural segment(Cys3-Ser28) for final 20 structures to REM structure	
Backbone(N,C $\alpha$ ,C $\prime$ )	$0.29 \pm 0.06$ Å
Heavy atoms	$1.00 \pm 0.11$ Å

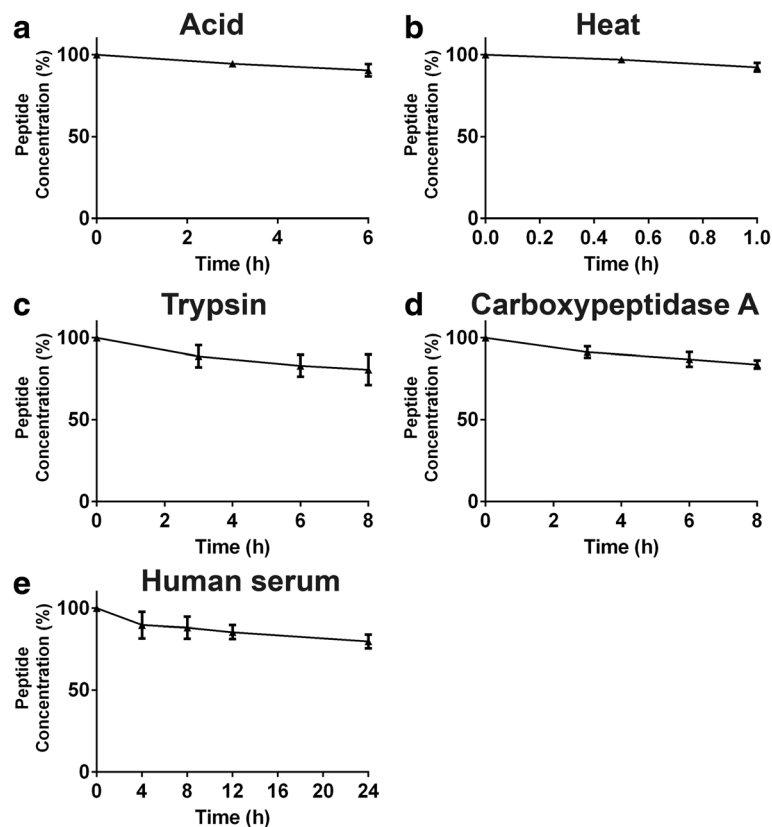


**Fig. 2** Solution NMR structures of jS3. **a** Superposition of the 20 lowest energy ensembles of jS3. **b** The ribbon NMR representation of jS3, which consists of three anti-parallel  $\beta$ -strands and a disulfide connectivity of CysI–CysV (yellow), CysII–CysIV (red) and CysIII–CysVI (blue). **c** The surface topology comparison between jS3 and jS1. Distribution of electrostatic charges was represented by blue (positively charged), red (negatively charged) and green (hydrophobic residues such as Val, Trp, Ile and Leu)

No NOE cross-peak was observed for Cys12/Cys25 due to spectral overlap. However, several long-range NOE correlations were observed between the  $\text{HN}/\text{H}\alpha$  atoms of Cys3 and  $\text{H}\alpha/\text{H}\beta$  of Cys23, and therefore, the disulfide connectivity of jS3 was concluded to be arranged as CysI–CysV, CysII–CysIV and CysIII–CysVI, which agrees with the disulfide arrangement determined by partial alkylation reported for jasmintide jS1 [7]. Figure 2c displays the surface topology of jS3. The positively charged residues (Arg10 and Arg21) and negatively charged residues (Asp11 and Asp22) are displayed sporadically, whereas the hydrophobic residues (Tyr14, Ile15, Ile16, and Trp17) in the loop between  $\beta$ -strand 1 and  $\beta$ -strand 2 formed a continuous hydrophobic surface combined with the N-terminal hydrophobic residues (Leu2, Leu4, and Leu5).

### Thermal and enzymatic stability of jasmintide

To show that jasmintide could be a potentially active component in medicinal plants and a highly stable scaffold for grafting bioactive peptides, we performed stability assays mimicking the decoction process and digestive environment (Fig. 3). Chromatographic and mass spectroscopic data revealed that jS3 was resistant



**Fig. 3** Stability of jS3 against various simulated conditions. **a** Acidic stability was performed at 37 °C and pH 2.0 for six hours. **b** Thermal stability was performed at 100 °C for one hour. **c** Trypsin and **d** carboxypeptidase A stability assays were performed at 37 °C for eight hours. **e** Human serum stability was performed at 37 °C for 24 h. The experiments were performed in triplicated ( $n = 3$ ) and were presented as mean  $\pm$  standard deviation

in boiling water for 1 h and under acidic conditions (pH 2.0) for up to 6 h with > 92 and > 90% of jasmintide remaining, respectively. Jasmintide jS3 was resistant to the endopeptidase trypsin and the exopeptidase carboxypeptidase A for up to 6 h with > 80% remaining intact. Furthermore, more than 79% of jS3 remained intact after incubation with human serum for up to 24 h at 37 °C.

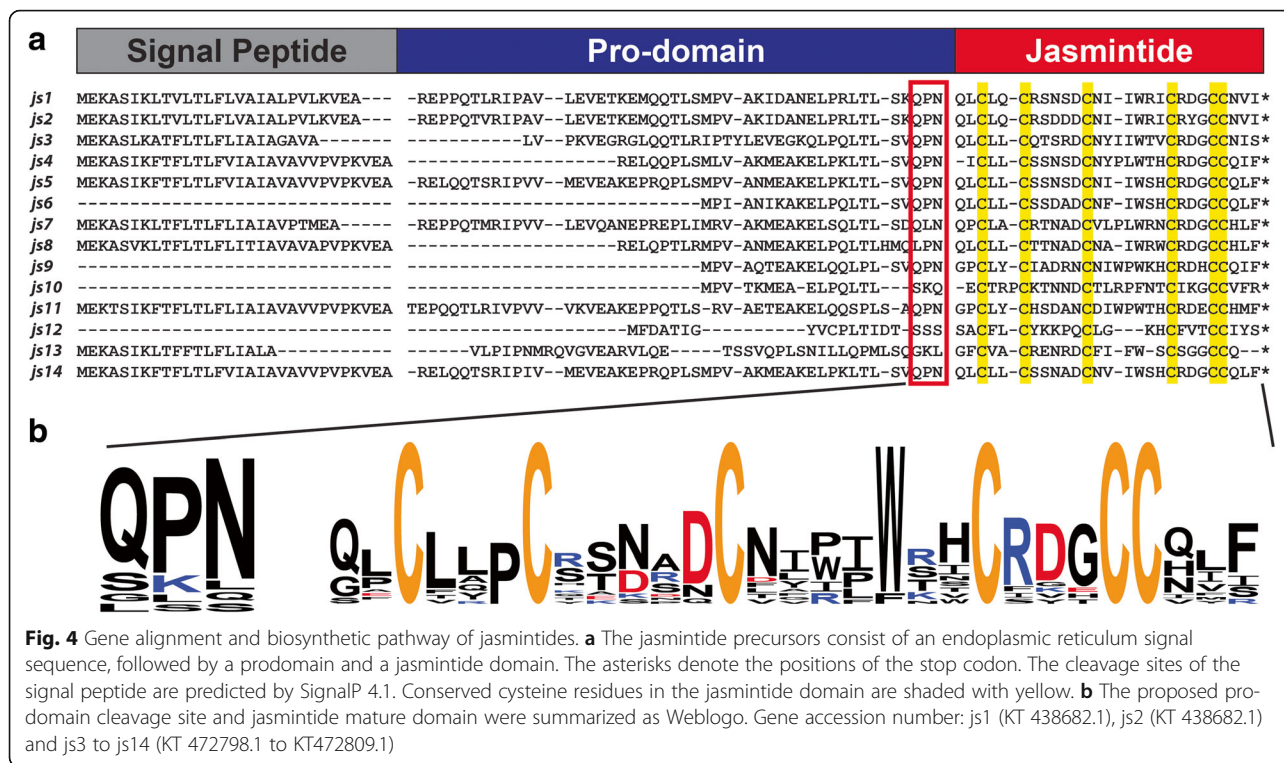
#### Biosynthesis of jasmintides

Total RNA extraction was performed from the roots, flowers and leaves of *J. sambac*, and RNA was sequenced using an Illumina Hiseq 2000. A total of 284.2 million clean reads were generated, with 41.7% GC content and an average length of 829 bp. The raw cDNA reads were assembled by Trinity into 296,026 contigs with a minimum length of 201 nucleotides. A tBLASTn search using the reported jasmintides jS1 and jS2 as queries led to the identification of 14 unique precursor sequences (Fig. 4a). In addition, we found two precursors, *js2* and *js14*, which were identical to those previously reported by our laboratory via transcriptomic analysis [7]. All 14 full-length precursors, each about 100 amino acids in length, shared a three-domain architecture, consisting of

an endoplasmic reticulum signal peptide (27–33 aa), a pro-domain (27–47 aa) and a mature jasmintide domain (25–28 aa). The three-domain precursor architecture has been found in other CRP families such as cystine knot  $\alpha$ -amylase inhibitors and carboxypeptidase inhibitors. It differs from the two-domain precursor arrangements of thionins and 6C-hevein-like CRPs, which do not have a pro-domain [5, 6, 9, 14, 16]. In addition, a conserved cleavage site, QXN, was found in 11 out of 14 jasmintide precursors at the C-terminus of the pro-domain. Thus, it is likely that an asparaginyl endopeptidase is involved in the bioprocessing of the jasmintides [9, 20]. The remaining three jasmintide precursors (*js10*, *js12* and *js13*) did not contain a QXN processing motif (X represents any amino acid), suggesting that a different processing enzyme could be involved (Fig. 4b). Interestingly, apart from the six highly conserved cysteine residues, some jasmintides were rich in hydrophobic leucine, isoleucine and valine residues.

#### Derivatization of cysteine to pseudo-lysine

A mass-spectrometry-driven peptidomic analysis of CRPs using mixtures would be useful to study the molecular



**Fig. 4** Gene alignment and biosynthetic pathway of jasmintides. **a** The jasmintide precursors consist of an endoplasmic reticulum signal sequence, followed by a prodomain and a jasmintide domain. The asterisks denote the positions of the stop codon. The cleavage sites of the signal peptide are predicted by SignalP 4.1. Conserved cysteine residues in the jasmintide domain are shaded with yellow. **b** The proposed pro-domain cleavage site and jasmintide mature domain were summarized as Weblogo. Gene accession number: *js1* (KT 438682.1), *js2* (KT 438682.1) and *js3* to *js14* (KT 472798.1 to KT472809.1)

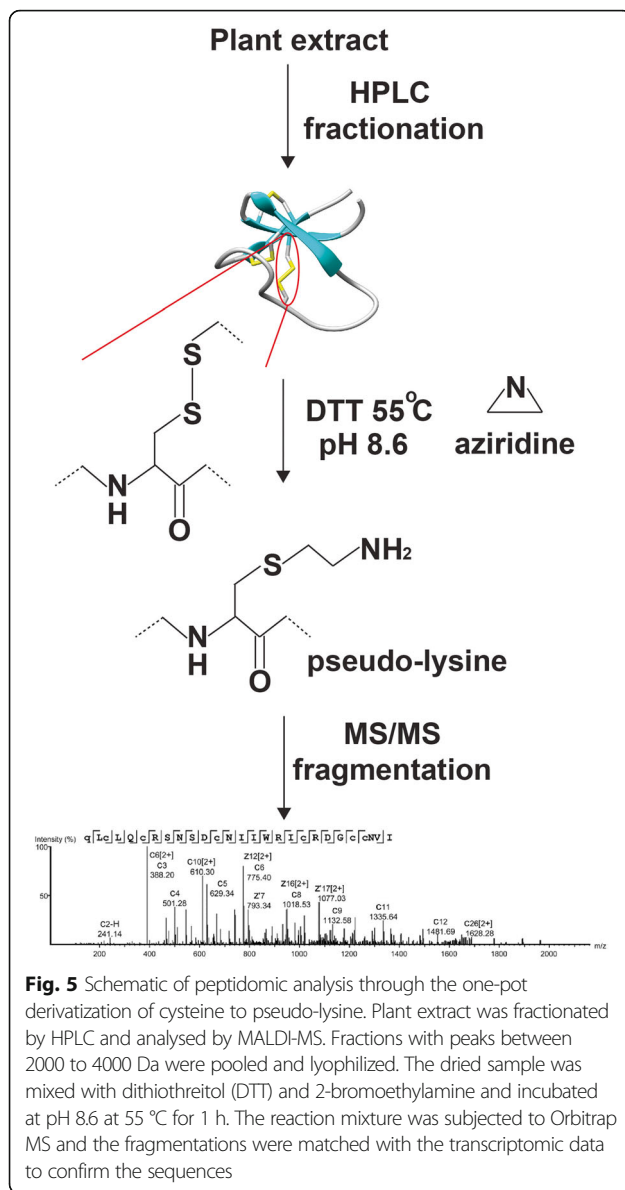
diversity of jasmintides. Our laboratory recently developed a one-pot derivatization method to convert cystine into pseudo-lysine, as shown in Fig. 4 [21]. In this one-pot reaction, dithiothreitol (DTT) and 2-bromoethylamine were mixed with fractionated CRPs. DTT reduced cystine to cysteine, and 2-bromoethylamine spontaneously cyclized to form an aziridine. Ring opening of aziridine by a cysteinyl thiol formed from the DTT-reduced cystine resulted in S-alkylation, converting cysteine into pseudo-lysine. The S-alkylated mixture was trypsinized, fractionated and analyzed by MS/MS, similar to conventional proteomic analysis. This one-pot approach enables us to analyze a complex mixture of CRPs and minimizes the labor-intensive purification steps [21].

**Peptidomic analysis of jasmintide expression**

A mass-spectrometry-driven peptidomic analysis of CRPs using mixtures would be useful to study the molecular diversity of jasmintides and the tissue-specific expression of jasmintides. Individual samples of fresh roots, flowers and leaves of *J. sambac* were extracted with water, fractionated using high-performance liquid chromatography and analyzed using MS as previously described (Fig. 5) [21]. The jasmintide-containing fractions were S-reduced by 30 mM dithiothreitol (DTT) and S-alkylated by 60 mM bromoethylamine at pH 8.6 and 55 °C for one hour. The derivatized mixtures were sequenced by electron-transfer dissociation (ETD) Orbitrap MS (Additional file 5: Figure S4). MS/MS fragmentation data of the alkylated extracts

were recorded and queried against the customized transcriptomic database using the PEAKS studio. Identification of unique full or partial sequences with -10LogP values ≥30 was considered valid.

Table 2 summarizes 86 jasmintides identified by MS/MS and nine jasmintide contigs (*js1 – js8* and *js14*) confirmed at the peptidomic level, representing 64.3% coverage of the total number of jasmintide genes present in the *J. sambac* transcriptome. The phylogenetic tree of the 14 jasmintide precursors were illustrated in (Additional file 6: Figure S5). Based on the highly conserved cysteine motif of C-C-C-C-C and high sequence identity (27.3 to 88.9%) and similarity (36.4 to 92.6%) to the previously reported *js1* (Additional file 7: Table S2), jasmintides *js2* to *js15* were classified as new members of the jasmintide family. Five jasmintide contigs (*js9 – js13*) were not detected at the peptidomic level. Our dataset contained 32 full sequences, including jasmintides *js1* and *js2* previously reported by our group [7]. In addition, 12 truncated jasmintide sequences with six cysteine residues and 54 truncated sequences with five or less cysteine residues were identified. A total of 71 sequences were identified in all three parts of *J. sambac*, while three sequences were expressed only in the roots. In addition, eleven sequences were detected in the flowers and leaves of *J. sambac* but were absent in the roots (Table 2). We identified 39 jasmintides with post-translation modifications (PTMs) using peptidomic analysis. They included 34 pyroglutamations at the N-terminus of Gln and



six deamidations (five at Asn and one at Gln). Together, we identified 84 novel jasmintides using a combination of transcriptomic and peptidomic approaches.

#### Antifeedant activity of jS3 against *Tenebrio molitor* larvae

To determine the effect of jS3 on *Tenebrio molitor* mealworm larvae, a no-choice assay was performed where the larvae were fed diets containing different amounts of jasmintide (0.5 to 30 mg/g of diet). After 24 h, the remaining food in the control and treatment groups were recorded, and the feeding deterrence index (FDI) was calculated. Figure 6a shows that the FDI increased from 42.9% at 1 mg/g of jasmintide in the diet to the highest value of 93.6% at 30 mg/g of jasmintide in the diet, suggesting a dose-dependent antifeedant effect of jS3 against *T. molitor*

larvae. Figure 6b shows the long-term effect of jasmintide on larvae growth, determined using a diet containing 1 mg of jS3 per gram of oat flakes. After a nine-day treatment, the average body weight of the treatment group ( $n = 30$ ) increased from 44.3 to 62.2 mg, which was significantly less than that of the control group without supplementation with jS3, for which the mean body weight increased from 47.4 to 71.2 mg. These results suggested that the jasmintide-containing diet was not preferred by the *T. molitor* larvae; hence, diet consumption in the treatment group was reduced and their mean body weight declined. No fatality was observed during the period of treatment. More importantly, co-incubation of insect Sf9 cells with jasmintide jS3 at concentrations up to 100  $\mu$ M showed insignificant effects on the change in morphology and mortality rate (Additional file 8: Figure S6).

## Discussion

### Unique cystine framework in jasmintides

Plant CRPs are grouped by their sequence homology and cystine framework. For our purpose, cystine framework is defined as a framework containing both a conserved cysteine motif (one-dimensional) and a conserved disulfide connectivity (two-dimensional). Cysteine spacing refers to the number of amino acids between two cysteine residues and is often referred to as a loop. Thus, jasmintides are a four-looped 6C-CRP family because they contain the cysteine motif of C-C-C-C-CC with a tandem CC motif [7]. Additionally, cystine framework is not equivalent to cysteine motif because CRP families can share the same disulfide connectivity with different cysteine motifs. For example, the knottin family comprises a cystine-knot disulfide connectivity between CysI–CysIV, CysII–CysV and CysIII–CysVI, which is a cystine connectivity found in three CRP families: 6C-hevein-like peptides [5], cystine-knot  $\alpha$ -amylase inhibitors [10, 11] and carboxypeptidase inhibitors [22]. However, the cysteine motifs differ between the families, with a C-C-CC-C-C motif found in cysteine knot  $\alpha$ -amylase inhibitors and a C-C-C-C-C-C motif found in carboxypeptidase inhibitors.

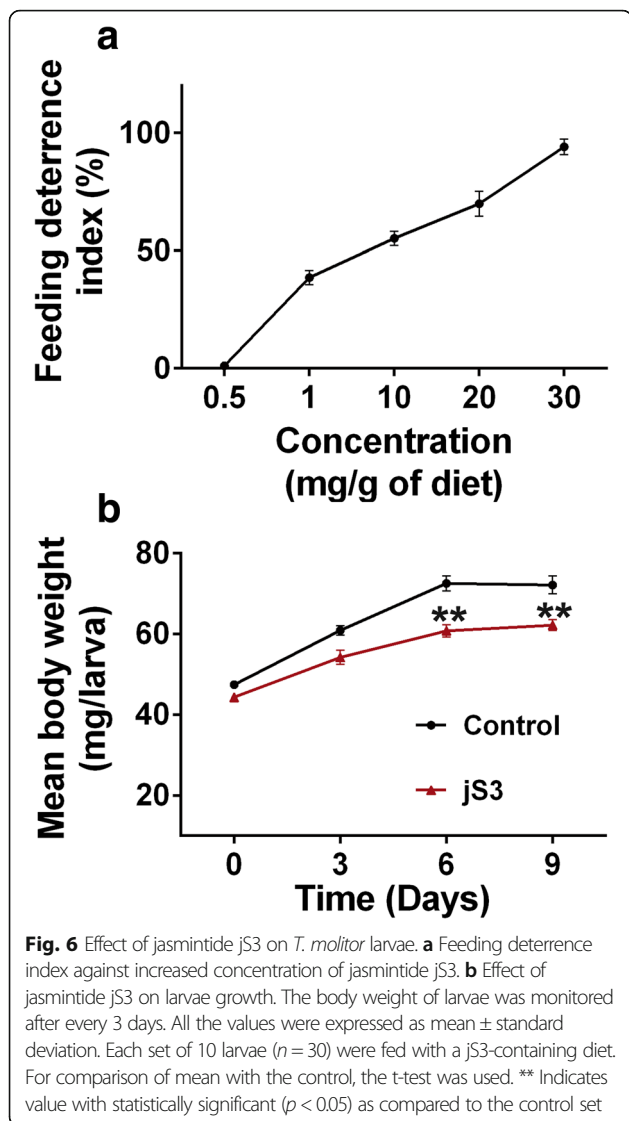
The presence of consecutive cysteine residues in a plant 6C-CRP can be used as a clue to identify its family. For example, the CC motif in jasmintides is located at the C-terminus as CysV and CysVI. This C-terminal CC-motif is a mirror image of the cysteine motif found in thionins, whose CC motif is located at the N-terminus as CysI and CysII (CC-C-C-C-C) [4, 7, 19]. On the other hand, the CC motif in the 6C-hevein-like peptides and cystine knot  $\alpha$ -amylase inhibitors is located at the middle of the C-C-CC-C-C framework as CysIII and CysIV [10, 14, 15]. In contrast,  $\beta$ -ginkgotides contain two CC motifs with a cystine framework of C-CC-C-CC [16].

Interestingly, the jasmintide-like cysteine motif with a CC motif at the C-terminus has also been found in

**Table 2** Identification and discovery of jasmitides from transcriptomic and peptidomic approaches

Name	Sequence	Mass	Transcriptome	LC-MS	Tissue	-10logP	Error	PTM
j81-1	<u><b>Z</b></u> LCIQCRNSDCNI IWRICRDGCCNVI	3104.38	x	x	R, L, F	172.71	1.2	Pyr (Q)
j81-2	<u><b>Z</b></u> LCIQCRNSDCNI IWRICRDGCCNVI	2991.29	x	x	R, L, F	165.43	0.6	Pyr (Q)
j81-3	<u><b>Z</b></u> LCIQCRNSDCNI IWRICRDGCCN	2909.25	x	x	R, L, F	186.11	0.7	Pyr (Q)
j81-4	<u><b>Z</b></u> LCIQCRNSDCNI IWRICRDGCCNVI	2993.35	x	x	R, L, F	200	0.4	
j81-5	<u><b>Z</b></u> LQCRNSDCNI IWRICRDGCCNVI	2880.26	x	x	R, L, F	117.6	0.2	
j81-6	<u><b>Z</b></u> LQCRNSDCNI IWRICRDGCCNVI	2778.25	x	x	R, L, F	200	0.3	
j81-7	<u><b>Z</b></u> CRNSDCNI IWRICRDGCCNVI	2537.11	x	x	R, L, F	150.64	-1.8	
j81-8	<u><b>Z</b></u> DCNI IWRICRDGCCNVI	2077.93	x	x	R, L, F	110.5	-0.3	
j81-9	<u><b>Z</b></u> CNI IWRICRDGCCNVI	1875.87	x	x	R, L, F	108.48	-0.3	
j81-10	<u><b>Z</b></u> IWRICRDGCCNVI	1659.82	x	x	R, L, F	200	-0.2	
j81-11	<u><b>Z</b></u> IWRICRDGCCNVI	1546.73	x	x	R, L, F	153.12	-0.4	
j82-1	<u><b>Z</b></u> LCIQCRSDDCNI IWRICRYGCCNVI	3180.41	x	x	R, L, F	193.97	0.8	Pyr (Q), Dea (Q)
j82-2	<u><b>Z</b></u> LCIQCRSDDCNI IWRICRYGCCNVI	3181.39	x	x	R, L, F	29.84	5	Pyr (Q)
j82-3	<u><b>Z</b></u> LCIQCRSDDCNI IWRICRYGCCNVI	3068.31	x	x	R, L, F	164.75	0.1	Pyr (Q)
j82-4	<u><b>Z</b></u> LQCRSDDCNI IWRICRYGCCNVI	2855.27	x	x	R, L, F	200	0.8	
j82-5	<u><b>Z</b></u> LQCRSDDCNI IWRICRYGCCNVI	2742.19	x	x	R, L, F	200	-0.5	
j82-6	<u><b>Z</b></u> LQCRSDDCNI IWRICRYGCCN	2643.12	x	x	R, L, F	200	-0.3	
j83-1	<u><b>Z</b></u> LCLLCQTSRDCNY IWTVCRDGCCNIS	3199.39	x	x	R, L, F	153.61	4.7	Pyr (Q)
j83-2	<u><b>Z</b></u> LCLLCQTSRDCNY IWTVCRDGCCNIS	3216.42	x	x	R, L, F	41.89	-8	
j83-3	<u><b>Z</b></u> LCLLCQTSRDCNY IWT	2050.94	x	x	R, L, F	163.06	-0.8	Pyr (Q)
j83-4	<u><b>Z</b></u> LCLLCQTSRDCNY	1638.69	x	x	R, L, F	74.91	0.3	Pyr (Q)
j83-5	<u><b>Z</b></u> LCLLCQTSRDCN	1475.63	x	x	R, L, F	133.47	0.6	Pyr (Q)
j83-6	<u><b>Z</b></u> LCLLCQTSRD	1259.58	x	x	R, L, F	111.19	-0.2	Pyr (Q)
j83-7	<u><b>Z</b></u> LCQTSRDCNY IWTVCRDGCCNIS	2760.19	x	x	R, L, F	42.73	1.4	
j83-8	<u><b>Z</b></u> CQTSRDCNY IWTVCRDGCCNIS	2647.10	x	x	R, L, F	120.33	-0.3	
j83-9	<u><b>Z</b></u> QTSRDCNY IWTVCRDGCCNIS	2545.09	x	x	R, L, F	64.09	0.1	
j83-10	<u><b>Z</b></u> CNY IWTVCRDGCCNIS	1957.83	x	x	R, L, F	117.04	-0.4	
j83-11	<u><b>Z</b></u> YI IWTVCRDG	1427.61	x	x	R, L, F	76.12	1	
j83-12	<u><b>Z</b></u> I IWTVCRDGCCNIS	1578.71	x	x	R, L, F	167.9	-0.6	
j83-13	<u><b>Z</b></u> IWTVCRDGCCNIS	1465.63	x	x	R, L, F	44.85	-0.4	
j84-1	<u><b>Z</b></u> LCILCSSNSDCNY PLWTHCRDGGCCQIF	3198.30	x	x	R, L, F	153.62	2	Pyr (Q)
j84-2	<u><b>Z</b></u> QCILCSSNSDCNY PLWTHCRDGGCCQIF	3215.33	x	x	R, L, F	58.47	-8.9	DF (C)
j84-3	<u><b>Z</b></u> LCILCSSNSDCNY PLWTHCRDGGCCQI	3051.23	x	x	R, L, F	79.59	0	Pyr (Q)
j84-4	<u><b>Z</b></u> LCILCSSNSDCNY PL	1751.73	x	x	R, L, F	33.38	0.6	Pyr (Q)
j84-5	<u><b>Z</b></u> LCILCSSNSDCN	1378.53	x	x	R, L, F	77.57	0.8	Pyr (Q)
j84-6	<u><b>Z</b></u> LCILCSSNSDC	1264.48	x	x	R, L, F	97.07	-0.9	Pyr (Q)
j84-7	<u><b>Z</b></u> LCSSNSDCNY PLWTHCRDGGCCQIF	2759.10	x	x	R, L, F	109.64	-0.7	
j84-8	<u><b>Z</b></u> CSNSDCNY PLWTHCRDGGCCQIF	2646.01	x	x	R, L, F	154.64	0	
j84-9	<u><b>Z</b></u> SSNSDCNY PLWTHCRDGGCCQIF	2544.00	x	x	R, L, F	133.19	-0.3	
j84-10	<u><b>Z</b></u> SDCNY PLWTHCRDGGCCQIF	2456.97	x	x	R, L, F	90.58	-0.7	
j84-11	<u><b>Z</b></u> CNY PLWTHCRDGGCCQIF	2053.84	x	x	R, L, F	151.69	-0.5	
j85-1	<u><b>Z</b></u> QCILCSSNSDCNY IWSHCRDGGCCQLF	3053.30	x	x	R, L, F	92.81	-7.1	Dea (N)
j85-2	<u><b>Z</b></u> QCILCSSNSDCNY IWSHCRDGGCCQLF	3037.25	x	x	R, L, F	200	0.2	Pyr (Q)
j85-3	<u><b>Z</b></u> QCILCSSNSDCNY IWSHCRDGGCCQLF	3053.30	x	x	R, L, F	78.17	-5.7	Dea (N)
j85-4	<u><b>Z</b></u> QCILCSSNSDCNY IWSHCRDGGCCQLF	3052.31	x	x	R, L, F	102.64	1.4	Dea (N)
j85-5	<u><b>Z</b></u> QCILCSSNSDCNY IWSHCRDGGCCQLF	3053.30	x	x	R, L, F	85.21	-6.5	Dea (N)
j85-6	<u><b>Z</b></u> QCILCSSNSDCNY IWSHCRDGGCCQLF	2907.22	x	x	R, L, F	144	-0.5	Pyr (Q)
j85-7	<u><b>Z</b></u> QCILCSSNSDCNY IWSHCRDGGCCQ	2777.10	x	x	R, L, F	137.26	0.2	Pyr (Q)
j85-8	<u><b>Z</b></u> QCILCSSNSDCNY IWSHCRDGGCCQLF	2813.14	x	x	R, L, F	114.15	1.8	
j85-9	<u><b>Z</b></u> LCSSNSDCNY IWSHCRDGGCCQLF	2598.05	x	x	R, L, F	45.77	0	
j85-10	<u><b>Z</b></u> CSNSDCNY IWSHCRDGGCCQLF	2484.96	x	x	R, L, F	116.53	-0.5	
j85-11	<u><b>Z</b></u> SSNSDCNY IWSHCRDGGCCQLF	2382.96	x	x	R, L, F	51.01	0.4	
j85-12	<u><b>Z</b></u> SDCNY IWSHCRDGGCCQLF	2094.85	x	x	R, L, F	91.96	-4.1	
j85-13	<u><b>Z</b></u> CNY IWSHCRDGGCCQLF	1892.79	x	x	R, L, F	138.02	-0.4	
j85-14	<u><b>Z</b></u> IWSHCRDGGCCQLF	1563.65	x	x	R, L, F	116.95	0.6	
j86-1	<u><b>Z</b></u> QCILCSSNSDADCNFY IWSHCRDGGCCQLF	3056.23	x	x	R, L, F	200	0.3	Pyr (Q)
j86-2	<u><b>Z</b></u> QCILCSSNSDADCNFY IWSHCRDGGCCQLF	3073.26	x	x	R, L, F	120.44	-4.3	
j86-3	<u><b>Z</b></u> QCILCSSNSDADCN	1264.13	x	x	R, L, F	90.03	-0.1	Pyr (Q)
j86-4	<u><b>Z</b></u> CNFY IWSHCRDGGCCQLF	1926.77	x	x	R, L, F	108.67	-0.5	
j87-1	<u><b>Z</b></u> QPCLACRNTADCVPLWRNCRDGGCHLF	3183.40	x	x	R, L, F	190.75	0.7	Pyr (Q)
j87-2	<u><b>Z</b></u> QPCLACRNTADCVPLWRNCRDGGCHLF	3200.43	x	x	R, L, F	89.43	-5.1	
j87-3	<u><b>Z</b></u> QPCLACRNTADCVPLWRNCRDGGCHLF	3036.33	x	x	R, L, F	200	0.3	Pyr (Q)
j87-4	<u><b>Z</b></u> QPCLACRNTADCVPLWRNCRDGGCH	2923.25	x	x	R, L, F	150.62	-0.1	Pyr (Q)
j87-5	<u><b>Z</b></u> QPCLACRNTADCVPLWRNCRDGGCC	2786.19	x	x	R, L, F	104.52	0.1	Pyr (Q)
j87-6	<u><b>Z</b></u> QPCLACRNTADCVPL	1582.70	x	x	R, L, F	110.95	-0.5	Pyr (Q)
j87-7	<u><b>Z</b></u> QPCLACRNTADCVL	1485.65	x	x	R, L, F	159.69	0.3	Pyr (Q)
j87-8	<u><b>Z</b></u> CRNTADCVPLWRNCRDGGCC	2291.97	x	x	R, L, F	47.17	-1.3	
j87-9	<u><b>Z</b></u> RTNADCVPLWRNCRDGGCHLF	2587.18	x	x	R, L, F	66.88	0.4	
j87-10	<u><b>Z</b></u> RTNADCVPLWRNCRDGGCC	2189.97	x	x	R, L, F	59.1	-0.7	
j87-11	<u><b>Z</b></u> TNADCVPLWRNCRDGGCH	2170.92	x	x	R, L, F	91.74	-0.1	
j87-12	<u><b>Z</b></u> TNADCVPLWRNCRDGGCC	2033.86	x	x	R, L, F	135.94	-0.4	
j87-13	<u><b>Z</b></u> CVLPLWRNCRDGGCHLF	2029.92	x	x	R, L, F	158.47	1.6	
j87-14	<u><b>Z</b></u> CVLPLWRNCRDGGCHLF	1882.85	x	x	R, L, F	70.17	0.7	
j87-15	<u><b>Z</b></u> LWRNCRDGGCHLF	1618.71	x	x	R, L, F	49.74	0.7	
j87-16	<u><b>Z</b></u> WRNCRDGGCHLF	1505.62	x	x	R, L, F	96.7	0	
j88-1	<u><b>Z</b></u> LCLLCTTNADCN IWRNCRDGGCHLF	3134.33	x	x	R	182.32	0.6	DF (C), Pyr (Q)
j88-2	<u><b>Z</b></u> LCLLCTTNADCN	1390.56	x	x	R	94.24	-0.5	DF (C), Pyr (Q)
j88-3	<u><b>Z</b></u> CTTNADCN IWRNCRDGGCHLF	2582.04	x	x	R	167.41	-1.8	
j89-1	<u><b>Z</b></u> GPCFLYCIADBNCFIWFWRNCRDGGCCQIF	3387.47	x	x				
j810-1	<u><b>Z</b></u> ECTFPKRTNDCILRFPNFCIKGGCVFV	3215.45	x	x				
j811-1	<u><b>Z</b></u> GPCFLYCHSDAMCDIWFWRNCRDGGCCQIF	3335.25	x	x				
j812-1	<u><b>Z</b></u> SACFLCYKRPQCLGRKCFVTCIYS	2838.28	x	x				
j813-1	<u><b>Z</b></u> GPCVACRENRCDFIWFSCGGCCQ	2684.04	x	x				
j814-1	<u><b>Z</b></u> LCILCSSNADCNFY IWSHCRDGGCCQLF	3007.24	x	x	R, L, F	179.32	0.1	DF (C), Pyr (Q)
j814-2	<u><b>Z</b></u> LCILCSSNADCNFY IWSHCRDGGCCQLF	2860.18	x	x	R, L, F	200	-0.3	DF (C), Pyr (Q)
j814-3	<u><b>Z</b></u> LCILCSSNADCNVI	1574.69	x	x	R, L, F	58.49	0.6	DF (C), Pyr (Q)
j814-4	<u><b>Z</b></u> LCILCSSNADCN	1362.53	x	x	R, L, F	78.98	0	DF (C), Pyr (Q)
j814-5	<u><b>Z</b></u> LCILCSSNADCV	1248.49	x	x	R, L, F	82.23	0.4	DF (C), Pyr (Q)
j814-6	<u><b>Z</b></u> CSNSNADCNFY IWSHCRDGGCCQLF	2454.95	x	x	R, L, F	107.86	1	
j814-7	<u><b>Z</b></u> V IWSHCRDGGCCQLF	1662.72	x	x	R, L, F	88.73	-1.3	
j815-1	<u><b>Z</b></u> QPCLACRNTFDCVPLWRNCRDGGCHLF	3226.44	x	x	F	145.47	0.7	

Mass: calculated mass expressed in Dalton. The tissue with the presence of jasmitides was detected by LC-Orbitrap MS/MS, where R, L and F represent roots, leaves and flowers, respectively. A higher -10logP value indicates a more confident sequencing result. Sequences with a -10logP value  $\geq 30$  is equivalent to false discovery rate (FDR) < 0.03%. Error represent the error in parts-per million (ppm). PTM post-translational modifications, where Pyr and Dea represent pyroglutamation and deamidation, respectively. The bracket represents the post-translational modified amino acid residues. Residues with modifications are bolded and underlined



$\beta$ -defensin hBD1 isolated from *Homo sapiens* [23] and conotoxin as24a isolated from the murine cone snail *Conus cancellatus* [24] (Additional file 9: Table S3).  $\beta$ -Defensin hBD1 has a similar disulfide connectivity to that of jS3, even though it comprises significantly longer loops 1 and 4 than jasmintides [23]. The disulfide connectivity of as24a has not yet been determined, and it would be of interest to investigate whether it shares similar disulfide connectivity with jasmintides. Molecular diversity also arises in each family with a conserved cystine framework by varying the size and amino acid composition of the displayed loops to diversify their biological function. Compared with jasmintides, hBD1 and as24a are relatively hydrophilic and have been shown to possess antimicrobial and paralytic effects, respectively [23, 24]. This observation suggests that the jasmintide-like framework could have been divergently evolved for several different purposes in plants and animals.

#### Unique three-dimensional structure in jasmintides

To determine whether jasmintide jS3 shared a similar peptide fold to other 6C-CRPs, pairwise structure alignment using the TM-align algorithm was performed and is summarized in Additional file 10: Table S4. A TM-score  $> 0.5$  indicates that the two structures shared the same peptide fold, whereas a value  $< 0.3$  represents that they have random structural similarity. Jasmintide jS3 had an average TM-score of 0.21, 0.26, 0.42, 0.38, 0.27 and 0.32 when compared to  $\beta$ -ginkgotide, cystine knot  $\alpha$ -amylase inhibitors, carboxypeptidase inhibitors, 6C-hevein-like peptides, thionins and human  $\beta$ -defensin, respectively, whereas its score was 0.56 compared to jS1. Taken together, these results show that jS3 displays a significantly different fold from those of other 6C-CRPs but a high similarity to that of jS1, a result consistent with the notion that they share a similar disulfide connectivity and cysteine spacing [6, 7, 9, 13–16].

#### Truncation and PTM increase sequence diversity

The presence of truncated peptides could be contributed by two mechanisms: diversity and turnovers. First, the same gene generates a diverse pool of candidates to respond to various environmental threats for their survival. This is particularly true in small-molecule natural products such as metabolites in which processing enzymes and post-translational modifications as well as degradation provide diversity. Finn and his colleagues proposed the ‘screening model’ to explain how molecular diversity would confer an organism with superior advantages [25]. A pool of diversified molecules “provides a better chance to possess rare components that are useful for changing environments” [21, 25]. Instead of producing a pool of DNA encodes for an array of peptides, plants use one gene to encode multiple peptides via truncation, mutation and other post-translation approaches. In turn, our work shows that such molecular diversity created via biosynthesis and cysteine frameworks can be achieved biosynthetically. Truncation of peptide sequences is a means for diversification and has been studied extensively in cyclotides [21, 26]. In a recent study conducted by Serra et al., a total of 38 cyclotide precursors were identified from *Clitoria ternatea*, which encoded for  $> 270$  ciotides [21]. To show that the presence of truncated peptides are neither methodological artefacts nor generated during sample preparation, the raw plant extract without chemical derivatization was analyzed using LC-MS. The results revealed a wide range of low-abundance shortened peptides with molecular weights  $\leq 3$  kDa, suggesting that truncated peptides occurred naturally. In another study, a sum of  $> 160$  cyclotides were identified from the plant extract of *Viola tricolor*, which were encoded by 108 precursor sequences. This “one-gene-multiple-products” approach is more



energy conserved than producing multiple genes to adopt the changing environment and has been observed in species other than the plant kingdom. For example, > 100 truncated conotoxins were generated by variable peptide processing from a single gene (MrIA) isolated from the cone snail *Conus marmoreus* [27, 28].

Second, plant peptides and proteins are regularly degraded. Cellular protein degradation is achieved by autophagy of cytosolic complex and organelles through vacuole-based protein degradation and the coordinated action of the proteasome on ubiquitinated proteins in the cytosol [29]. For example, leaf proteins in *Arabidopsis thaliana* rosettes have degradation rates of 0.03–0.04 day<sup>-1</sup> and a half-life of 3.5 days [30], whereas the total proteins in *Dactylis glomerata* have degradation rates of 0.12–0.16 day<sup>-1</sup> and a half-life of 4–6 days [31]. This is also observed in conotoxins cysteine-rich peptides by Dutertre et al. [27].

Glutaminy cyclization of the side chain of the N-terminal glutamic acid is one of the most common PTMs found in plant CRPs and has been shown to contribute to stability against enzymatic degradation by exopeptidases such as aminopeptidase I [32]. However, pyroglutamation at the N-terminus could be a spontaneous event. Previous studies demonstrated that deamidation of asparagine is the most likely spontaneous non-enzymatic reaction and is directly related to pH and temperature [33].

#### Antifeedant activity of jasmintides

*T. molitor* is not a common pest on *J. sambac*. However, this model has widely been used in determining the insecticidal effect of triterpenes [34], diterpenes [35] and benzoylhydrazine [36], molting disruption effect of long chain n-alkanes [37], and the effect of growth development exerted by pyriproxyfen, a juvenile hormone analog in insect [38]. The use of a similar insect strain generates results that can be compared from various laboratories. The large quantity of larvae required in carrying out various experiments limits the use of common pests. In our study, more than 210 larvae were used. In contrast, *T. molitor* larvae are commercially available in local aquarium store, and hence the results can be easily reproduced in other laboratories. Therefore, instead of a common pest on jasmine, commercially available *T. molitor* larvae were used herein.

It is important to note that the yield of jS3 obtained from the extraction is not the physiological concentration of jasmintide in *J. sambac* plant. Since the extraction involved six steps, our estimation is that the physiological concentration would likely be tenfold higher than 40 mg/kg of fresh material. The antifeedant activity could be explained by three major mechanisms. An olfactory repellent effect is the most common mechanism exerted

by volatile aromatic compounds [39]. The second mechanism contributes to the inhibition of digestive enzymes, which disrupts digestion and causes gut discomfort. This mechanism has been observed in 6C-CRPs such as the carboxypeptidase inhibitor from potato *Solanum tuberosum* [40]. The last mechanism relates to true taste deterrence, which stimulates sensory receptors, generates uncomfortable taste and, hence, decreases appetite and food consumption [41]. Certain hydrophobic amino acids have been shown to elicit a bitter taste, including Trp, Tyr, Leu, Ile, Phe and Val [42, 43]. Jasmintide jS3 contains nine such amino acid residues, accounting for 32% of its sequence. Importantly, they are located in close proximity, forming a hydrophobic patch revealed by NMR analysis. As a result, we speculated that the antifeedant effect exerted by jS3 may be related to this hydrophobic patch, which binds to the sensory receptors and evokes a bitter sensation, leading to a decrease in consumption of the jasmintide-containing diet. However, a detailed mechanism of the antifeedant activity exerted by jasmintide remains to be studied.

#### Conclusion

In this study, 14 precursor sequences encoding > 80 jasmintides were identified from the roots, flowers and leaves of *J. sambac* using transcriptomic and peptidomic methods. Jasmintides, compact in structure and rich in hydrophobic amino acids such as Leu and Ile, belong to a new family of CRPs with disulfide connectivity that is different from that of knottins and thionins. Jasmintides are resistant to degradation by heat, acid and human serum. Importantly, an in vivo study showed that jasmintide jS3 exhibits an antifeedant effect against *T. molitor* larvae in a dose-dependent manner, suggesting jasmintide jS3 could play a role in host defense. The combination of a compact structure and hyperstability together with in vivo antifeedant effects suggests that jasmintides could be a useful template for developing bioactive peptides. In addition, our results showed that molecular diversity exists on several levels, including amino acid composition, cystine framework, disulfide connectivity and variable peptide processing. Such molecular diversity could be exploited for protein engineering to design feeding deterrents to prevent crop damage.

#### Methods

##### General materials

*J. sambac* was cultivated at the community herb garden in Nanyang Technological University, Singapore (1° 21' 22.878" N, 103° 41' 23.388" E). It was authenticated by Mr. Paul Leong from the Singapore Botanic Garden, Singapore, and a voucher specimen was stored at the Singapore Botanic Garden Herbarium Singapore with accession number SING 2015–202. NMR experiments were

performed on a 600 or 700 MHz Bruker Avance spectrometer equipped with a 5-mm triple-resonance cryogenic probe head and a z-axis gradient coil at 25 °C. Mass spectrometry analysis was monitored using a MALDI-TOF/TOF ABI 4800 (Applied Biosystems, Framingham, MA, USA). Data acquisition was performed in reflection mode using a mass range of 1000–6000 Da and a focusing mass of 3000 Da. A saturated solution of  $\alpha$ -cyano-4-hydroxycinnamic acid prepared in 75% acetonitrile and 0.1% trifluoroacetic acid was used as the MALDI matrix. Desalting of samples was performed as needed using C<sub>18</sub> ZipTip. HPLC and ultra-performance liquid chromatography (UPLC) were performed on a Shimadzu system. Grace Vydac C<sub>18</sub> columns (particle size, 5  $\mu$ m; pore size, 300 Å; Hesperia, CA, USA) with dimensions of 250 × 22 mm, 250 × 10 mm, and 250 × 4.6 mm were used for preparative, semi-preparative, and analytical RP-HPLC, respectively. A PolyLC polysulfoethyl A column (250 × 4.6 mm) was used for SCX-HPLC. LC-MS/MS was performed in an Orbitrap Elite mass spectrometer (Thermo Scientific Inc., Bremen, Germany) coupled with a Dionex UltiMate 3000 UHPLC system (Thermo Scientific Inc., Bremen, Germany) using an Acclaim PepMap RSL column (75  $\mu$ m × 15 cm; 2  $\mu$ m particles; Thermo Scientific Inc., Bremen, Germany). All chemical reagents used in this study were of analytical grade and were purchased from Sigma-Aldrich (MO, USA) unless otherwise stated.

#### RNA extraction, sequencing and assembly

Total RNA was extracted from the fresh leaves, flowers and roots of *J. sambac* by Trizol according to the manufacturer's protocol (Life Technologies, California, USA). Sequencing and de novo assembly were performed by Macrogen Inn (Seoul, Korea). Poly-A mRNA was isolated from total RNA using magnetic oligo (dT) beads. Following purification, the mRNA was fragmented into small pieces using divalent cations under elevated temperature. The cleaved RNA fragments were copied into first-strand cDNA using reverse transcriptase and random primers. This was followed by second-strand cDNA synthesis using DNA Polymerase I and RNase H. These cDNA fragments then went through an end-repair process, the addition of a single 'A' base, and ligation of the adapters. The products were then purified and enriched by PCR to create the final cDNA library. The cDNA library was examined using an Agilent Bioanalyzer High Sensitivity DNA chip to assess the library quality. Finally, the library was sequenced using an Illumina HiSeq™ 2000 and assembled using Trinity [44].

#### Transcriptome analysis of jasmintide precursor sequences

The identification of jasmintide sequences was performed from the raw data using tBLASTn with jasmintides jS1 and jS2 as the query. The e-value cut-off was set at 0.01. The hits were then translated into their respective amino

acid sequences via EMBOSS Transeq [45]. Sequences with redundancy and absence of a stop codon were removed. The cleavage sites of precursor sequences were predicted by SignalP 4.1. The jasmintide precursors were aligned using Clustal Omega and visualized by BioEdit. Novel precursor sequences are available in the GenBank database under the accession numbers KY472798.1 to KY472809.1.

#### Characterization of jasmintides by mass spectrometry

Fresh *J. sambac* leaves, flowers and roots (800 mg) were washed and individually blended with 50% ethanol (1:3 w/v). After centrifugation at 1000 rpm for 10 min, the supernatant was collected, diluted to 20% (v/v) ethanol and injected into SCX-HPLC. Peaks with intensities >100 absorption units were pooled and desalted using a C<sub>18</sub> Sep-Pack. The eluent was concentrated to 200  $\mu$ L and mixed with 30 mM DTT and 60 mM bromoethylamine in 0.2 M Tris buffer (pH 8.6). The mixture was incubated at 55 °C for 1 h. Subsequently, the alkylated jasmintides were analyzed by HPLC tandem mass spectrometry as previously described, with minor modifications [21]. LC-MS/MS analyses were performed with an Orbitrap Elite mass spectrometer (Thermo Scientific Inc., Bremen, Germany) coupled with a Dionex UltiMate 3000 UHPLC system (Thermo Scientific Inc., Bremen, Germany). Separation of peptides was conducted using a reverse-phase Acclaim PepMap RSL column (75  $\mu$ m ID × 15 cm, 2  $\mu$ m particles, Thermo Scientific). The mobile phase was 0.1% formic acid (FA) as eluent A and 90% ACN 0.1% FA as eluent B, with a flow rate of 0.3  $\mu$ L/min. A 60-min gradient was used for the elution as follows: 3% B for 1 min, 3–35% B over 47 min, 35–50% B over 4 min, 50–80% B over 6 s, and 80% for 78 s and then reverted to the initial state over 6 s and maintained for 6.5 min. Samples were sprayed with a Michrom Thermo Captive Spray nano-electrospray ion source (Bruker-Michrom Inc., Auburn, CA, USA) set at a source voltage of 1.5 kV and a capillary temperature of 250 °C. Data were collected using Xcalibur 2.2 software (Thermo Scientific Inc., Bremen, Germany) in positive ion mode alternating between full scan-MS (350–1600  $m/z$ , 60,000 resolution at 400  $m/z$ , 1 microscan per spectrum) and MS/MS (150–2000  $m/z$ , 30,000 resolution at 400  $m/z$ , 1 microscan averaged per spectrum). High-energy collisional dissociation fragmentation was performed for the 3 most intense ions with a 500-count threshold using 27, 30 and 34% normalized collision energy per spectrum and a 3-Da isolation window. Automatic gain control for full scan-MS and full scan-MS/MS was set to  $1 \times 10^6$  ions.

The database search and post-translational modification analysis were performed in PEAKS studio (version 7.0, Bioinformatics Solutions, Waterloo, Canada) using 10 ppm MS and 0.05 Da MS/MS tolerances. The PEAKS PTM function [46] was used to search our data for > 650 PTMs and mutations listed in the Unimod database

[47]. The SPIDER algorithm was employed to identify peptide mutations and homology [48]. All identified sequences were validated manually.

#### Isolation of jasmintides from *J. sambac* flowers

Fresh *J. sambac* flowers (1 kg) were homogenized and extracted with 5 L of 50% (v/v) ethanol. After centrifugation (8500 g, 10 min, 4 °C), the supernatant was filtered, diluted to 20% (v/v) ethanol and purified using a C<sub>18</sub> flash column (Grace Davison, Columbia, MD, USA). The unbound compounds were washed with Milli-Q water and eluted using 1 L of 80% (v/v) ethanol. The eluent was lyophilized, re-dissolved in 200 mL of 20% acetonitrile and purified using several rounds of SCX-HPLC and RP-HPLC, as previously described [9, 10, 49]. Eluents were monitored by MALDI-MS, and fractions with the same molecular weight were pooled and lyophilized. The dried samples were stored at -20 °C until further analysis.

#### NMR spectroscopy

Lyophilized peptides (3 mg) were dissolved in 500 µL of 90% H<sub>2</sub>O/10% D<sub>2</sub>O or 99.99% D<sub>2</sub>O at pH 7.0 in 20 mM Na<sub>3</sub>PO<sub>4</sub> buffer containing 50 mM NaCl and 0.01% NaN<sub>3</sub>. Homonuclear 2D NOESY experiments were performed with mixing times of 200 and 300 ms. COSY data was recorded in both H<sub>2</sub>O and D<sub>2</sub>O solutions with a mixing time of 78 ms using MLEV17 spin-lock pulses. Vicinal coupling constants were determined using DQF COSY and 1H NMR experiments. All 2D-NMR data were recorded in the phase-sensitive mode using the time-proportional phase increment method, with 2048 data points in the t<sub>2</sub> domain and 512 points in the t<sub>1</sub> domain. Slowly exchanging amide protons were identified by lyophilization of a fully protonated sample in H<sub>2</sub>O solution to dryness, resuspension in 99.99% D<sub>2</sub>O solution and immediate acquisition of a series of 1D spectra. All NMR data were processed using Bruker TOPSPIN 2.1 (Bruker Instruments) or NMRPipe [50] and were analyzed using Sparky 3.12. Solution structures of jS3 were calculated by hybrid distance geometry and simulated annealing protocols in torsion angle space with CNS 1.2 [51]. A total of 656 distance and 22 torsion angle constraints were used for the structure calculations. NOE distance restraints were classified as strong (1.8–3.0 Å), medium (1.8–3.5 Å), weak (1.8–5.0 Å), or very weak (1.8–6.0 Å). Corrections for pseudo atom representations were used for non-stereo specifically assigned methylenes, methyl groups, and aromatic ring protons. Backbone dihedral angle restraints were derived from the 3JHN-Hα coupling constants in the DQF-COSY or 1H NMR spectra in H<sub>2</sub>O solution. MOLMOL [52] and PyMOL [53] were used for structure visualization, and PROCHECK-NMR [54] was used for structure validation.

The jasmintide jS3 NMR structure was uploaded to the Protein Data Bank with accession number 5WXE.

#### Stability assays

For the thermal stability assay, jS3 were heated in boiling water for 1 h. At each time interval (1.0 h), a 30-µL sample was aliquoted and quenched in an ice bath for 15 min. For the acid stability assay, jS3 and jS4 were incubated in hydrochloric acid (pH 2.0) for 6 h. At each time interval (2, 4 and 6 h), a 30-µL sample was aliquoted and quenched by adding 1 M sodium hydroxide. For the enzymatic stability assay, jS3 and jS4 were incubated for 6 h at the optimal temperature with a buffer solution, as recommended by the manufacturer. At each time interval (2, 4 and 6 h), a 30-µL sample was aliquoted and quenched by adding 1 M hydrochloric acid. For the serum stability assay, jS3 were incubated with 25% human AB-type male serum in phenol red-free Dulbecco's modified Eagle's medium at 37 °C for 24 h. At each time interval (4, 8, 12 and 24 h), a 30-µL sample was aliquoted and quenched with 96% ethanol. Samples were kept at 4 °C for 15 min and then centrifuged at 18,000 rpm for 10 min. For all stability assays, the reaction mixture was analyzed by UPLC and monitored by MALDI-MS. The areas under the peak prior to and after treatment were integrated to determine stability. For each stability assay, three independent experiments were performed. At each time point, the sample was quenched by the specific reagent aforementioned and analyzed by UPLC in triplicates.

#### Feeding deterrent assay

The antifeedant activity was examined using a no-choice assay [55]. Oat flakes were soaked in different concentrations of jS3 (0.5 to 30 mg/g of oat) and lyophilized. The control diet was prepared in the same manner without addition of jS3. Ten *T. molitor* larvae of similar weights (41 ± 3 mg) were placed in a tissue culture dish with 500 mg of oats. The dishes were transferred into a rearing chamber and kept at 29 ± 1 °C in the dark. After 24 h, the remaining oat flakes were weighed, and the amount consumed was calculated. In total, 30 larvae were used for each concentration, and three independent biological experiments were performed. The feeding deterrence activity was determined as follows [56]:

$$\text{Feeding Deterrence Index (FDI)} = (C - T) * 100 / (C + T)$$

where C represents the average weight of consumed control diet and T represents the average weight of consumed treatment diet within the groups. The overall differences in feeding amounts were compared to the control and were compared to each other using one-way ANOVA.

### Effect of jasmintides on the growth of larvae

At day 0, 10 *T. molitor* larvae of similar weights ( $43 \pm 2$  mg and  $47 \pm 2$  mg in the treatment and control sets, respectively) were fed 200 mg of a diet containing 1 mg of jS3/g of oats. Every three days, the weights of individual larva were recorded, and larvae were transferred to a new dish with the same amount of supplemented diet to maintain a continuous supply. Three independent biological experiments were carried out over a period of 9 days. The control diet was prepared in the same manner without the addition of jS3. Student's t-test was used to compare the mean with the control.

### Additional files

**Additional file 1: Table S1.** NMR restraints and statistics of the solution structure of jS3. (DOCX 14 kb)

**Additional file 2: Figure S1.** NH/COH fingerprint region of 2D-NOESY spectrum of jS3 recorded in 90% H<sub>2</sub>O/10% D<sub>2</sub>O at 298 K. Sequential connectivity of each amino acid residue are shown by solid lines. (DOCX 4392 kb)

**Additional file 3: Figure S2.** Summary of sequential and medium NOE connectivity. Slowly exchanging amide protons (filled box) and backbone vicinal coupling constants (open circle: <sup>3</sup>JHNa < 6 Hz, filled circle: <sup>3</sup>JHNa > 8 Hz) are indicated. (DOCX 4739 kb)

**Additional file 4: Figure S3.** NOE cross peak between the Hβs of the each disulfide bond in jS3. (DOCX 171 kb)

**Additional file 5: Figure S4.** Annotated MS/MS spectra of jasmintides identified using proteomics approach. Fragments are labeled with c-, z-, z + 1 (z'), z + 2 (z(+ 2)), b- and y- ions. (DOCX 825 kb)

**Additional file 6: Figure S5.** Phylogenetic tree of the 14 jasmintide precursors. The precursors were aligned using Clustal Omega and the tree was constructed with 1000 bootstraps using neighbor-joining algorithm. (DOCX 17 kb)

**Additional file 7: Table S2.** Comparison of the 14 jasmintides against jS1. (DOCX 12 kb)

**Additional file 8: Figure S6.** The effect of jasmintide jS3 on the cell viability of Sf9 cells. (DOCX 111 kb)

**Additional file 9: Table S3.** TM align score between jasmintide jS3 and other CRPs with 6 cysteine residues. (DOCX 12 kb)

**Additional file 10: Table S4.** The amino acid sequences of jasmintide jS3, human β-defensin hBD1 and conotoxin as24a. (DOCX 12 kb)

### Abbreviations

6C-CRPs: Six-cysteine cysteine-rich peptides; CRPs: Cysteine-rich peptides; D2O: Deuterium oxide; DTT: Dithiothreitol; ETD: Electron-transfer dissociation; FA: Formic acid; FDI: Feeding deterrence index; HPLC: High-performance liquid chromatography; LD<sub>50</sub>: Average lethal dose; MALDI-TOF MS: Matrix-assisted laser desorption/ionization-time of flight mass spectrometry; MS: Mass spectrometry; MS/ MS: Tandem mass spectrometry; NOESY: Nuclear Overhauser effect spectroscopy; PTMs: Post-translation modifications; RMSD: Root-mean-square deviation; RP: Reversed-phase; SCX: Strong cationic exchange; TOCSY: Total correlation spectroscopy; UPLC: Ultra-performance liquid chromatography

### Acknowledgments

We would like to thank Mr. Paul Leong from the Singapore Botanic Garden, Singapore for the authentication of plant sample *Jasmintide sambac*.

### Funding

This research was supported by a Nanyang Technological University internal funding – Synzyme and Natural Products Center (SYNC) and the AcRF Tier 3 funding (MOE2016-T3–1-003). The funding bodies were not contributed to

the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

### Availability of data and materials

All the data pertaining to the present study has been included in the tables and figures of the manuscript and the data has been uploaded to the public databases as follows: The NMR structure of jS3 can be viewed with Protein Data Bank accession code 5WXE (<https://www.rcsb.org/structure/5WXE>). The transcriptome data was uploaded to NCBI GenBank SRA database with a BioProject, BioSample and SRA accession number of PRJNA470789, SAMN09103990 and SRR7151744, respectively. The complete cDNA coding sequence of jasmintide can be found via GenBank with the accession number as follows: js1 (KT438682.1; <https://www.ncbi.nlm.nih.gov/nucore/KT438682.1>), js2 (KT438683.1; <https://www.ncbi.nlm.nih.gov/nucore/KT438683.1>), js3 (KY472798.1; <https://www.ncbi.nlm.nih.gov/nucore/KY472798.1>), js4 (KY472799.1; <https://www.ncbi.nlm.nih.gov/nucore/KY472799.1>), js5 (KY472800.1; <https://www.ncbi.nlm.nih.gov/nucore/KY472800.1>), js6 (KY472801.1; <https://www.ncbi.nlm.nih.gov/nucore/KY472801.1>), js7 (KY472802.1; <https://www.ncbi.nlm.nih.gov/nucore/KY472802.1>), js8 (KY472803.1; <https://www.ncbi.nlm.nih.gov/nucore/KY472803.1>), js9 (KY472804.1; <https://www.ncbi.nlm.nih.gov/nucore/KY472804.1>), js10 (KY472805.1; <https://www.ncbi.nlm.nih.gov/nucore/KY472805.1>), js11 (KY472806.1; <https://www.ncbi.nlm.nih.gov/nucore/KY472806.1>), js12 (KY472807.1; <https://www.ncbi.nlm.nih.gov/nucore/KY472807.1>), js13 (KY472808.1; <https://www.ncbi.nlm.nih.gov/nucore/KY472808.1>) and js14 (KY472809.1; <https://www.ncbi.nlm.nih.gov/nucore/KY472809.1>). The authors are pleased to share all the data and plant materials upon reasonable request.

### Authors' contributions

JPT, GK and KHW conceived and designed the experiments. GK, KHW, AS and SJ performed the experiments, analyzed the data and wrote the manuscript. JPT, HSY and SKS revised the manuscript. All authors read and approved the final version of the manuscript.

### Ethics approval and consent to participate

*Jasmintide sambac* was cultivated at the community herb garden in Nanyang Technological University, Singapore (1° 21' 22.878" N, 103° 41' 23.388" E). It was authenticated by Mr. Paul Leong from the Singapore Botanic Garden, Singapore, and a voucher specimen was stored at the Singapore Botanic Garden Herbarium Singapore with accession number SING 2015–202.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 3 April 2018 Accepted: 2 July 2018

Published online: 11 July 2018

### References

- Kiritikar KR, Basu BD. Indian medicinal plants with illustrations. 2nd edition. Oriental Enterprises, Vol 7; 2003. p. 2093–6.
- Sengar N, Joshi A, Prasad SK, Hemalatha S. Anti-inflammatory, analgesic and anti-pyretic activities of standardized root extract of *Jasminum sambac*. *J Ethnopharmacol.* 2015;160:140–8.
- Ito Y, Sugimoto A, Kakuda T, Kubota K. Identification of potent odorants in Chinese jasmine green tea scented with flowers of *Jasminum sambac*. *J Agric Food Chem.* 2002;50(17):4878–84.
- Tam JP, Wang S, Wong KH, Tan WL. Antimicrobial peptides from plants. *Pharmaceuticals.* 2015;8(4):711–57.
- Kini SG, Nguyen PQT, Weissbach S, Mallagaray A, Shin J, Yoon HS, Tam JP. Studies on the chitin-binding property of novel cysteine-rich peptides from *Alternanthera sessilis*. *Biochemistry.* 2015;3(54):6639–49.
- Kini SG, Wong KH, Tan WL, Xiao T, Tam JP. Morintides: cargo-free chitin-binding peptides from *Moringa oleifera*. *BMC Plant Biol.* 2017;17(1):68.

7. Kumari G, Serra A, Shin J, Nguyen PQT, Sze SK, Yoon HS, Tam JP. Cysteine-rich peptide family with unusual disulfide connectivity from *Jasminum sambac*. *J Nat Prod*. 2015;78(11):2791–9.
8. Nguyen GK, Wang S, Qiu Y, Hemu X, Lian Y, Tam JP. Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis. *Nat Chem Biol*. 2014;10(9):732–8.
9. Nguyen GKT, Zhang S, Nguyen NT, Nguyen PQT, Chiu MS, Hardjojo A, Tam JP. Discovery and characterization of novel cyclotides originated from chimeric precursors consisting of albumin-1 chain a and cyclotide domains in the Fabaceae family. *J Biol Chem*. 2011;286(27):24275–87.
10. Nguyen PQT, Luu TT, Bai Y, Nguyen GK, Pervushin K, Tam JP. Allotides: proline-rich Cystine knot alpha-amylase inhibitors from *Allamanda cathartica*. *J Nat Prod*. 2015;78(4):695–704.
11. Nguyen PQT, Wang S, Kumar A, Yap LJ, Luu TT, Lescar J, Tam JP. Discovery and characterization of pseudocyclic cystine-knot  $\alpha$ -amylase inhibitors with high resistance to heat and proteolytic degradation. *FEBS J*. 2014;281(19):4351–66.
12. Nguyen PQT, Ooi JSG, Nguyen NTK, Wang SJ, Huang M, Liu DX, Tam JP. Antiviral Cystine knot  $\alpha$ -amylase inhibitors from *Alstonia scholaris*. *J Biol Chem*. 2015;290(52):31138–50.
13. Tan WL, Wong KH, Jian L, Sakai N, Tan HW, Rolf H, Tam JP. Lybatides from *Lycium barbarum* contain an unusual Cystine-stapled helical peptide scaffold. *Sci Rep*. 2017;7:5194.
14. Wong KH, Tan WL, Kini SG, Xiao TS, Serra A, Sze SK, Tam JP. Vaccatides: antifungal glutamine-rich Hevein-like peptides from *Vaccaria hispanica*. *Front Plant Sci*. 2017;8:1100.
15. Wong KH, Tan WL, Serra A, Xiao T, Sze SK, Yang D, Tam JP. Ginkgotides: proline-rich Hevein-like peptides from gymnosperm *Ginkgo biloba*. *Front Plant Sci*. 2016;7:1639.
16. Wong KH, Tan WL, Xiao TS, Tam JP. Beta-Ginkgotides: Hyperdisulfide-constrained peptides from *Ginkgo biloba*. *Sci Rep*. 2017;7:6140.
17. Bhardwaj G, Mulligan VK, Bahl CD, Gilmore JM, Harvey PJ, Cheneval O, Buchko GW, Pulavarti SV, Kaas Q, Eletsky A, et al. Accurate de novo design of hyperstable constrained peptides. *Nature*. 2016;538(7625):329–35.
18. Gracy J, Le-Nguyen D, Gelly JC, Kaas Q, Heitz A, Chiche L. KNOTTIN: the knottin or inhibitor cystine knot scaffold in 2007. *Nucleic Acids Res*. 2008;36:D314–9.
19. Stec B. Plant thionins—the structural perspective. *Cellular and molecular life sciences : CMLS*. 2006;63(12):1370–85.
20. Saska I, Gillon AD, Hatsugai N, Dietzgen RG, Hara-Nishimura I, Anderson MA, Craik DJ. An asparaginyl endopeptidase mediates in vivo protein backbone cyclization. *J Biol Chem*. 2007;282(40):29721–8.
21. Aida Serra XH, Nguyen GK, Nguyen NT, Sze SK, Tam JP. A high-throughput peptidomic strategy to decipher the molecular diversity of cyclic cysteine-rich peptides. *Sci Rep*. 2016;6:23005.
22. Chiche L, Heitz A, Padilla A, Le-Nguyen D, Castro B. Solution conformation of a synthetic bis-headed inhibitor of trypsin and carboxypeptidase a: new structural alignment between the squash inhibitors and the potato carboxypeptidase inhibitor. *Protein Eng*. 1993;6(7):675–82.
23. Bensch KW, Raida M, Magert HJ, Schulz-Knappe P, Forssmann WG. hBD-1: a novel beta-defensin from human plasma. *FEBS Lett*. 1995;368(2):331–5.
24. Aguilar MB, Zugasti-Cruz A, Falcon A, Batista CV, Olivera BM, de la Cotera EP. A novel arrangement of Cys residues in a paralytic peptide of *Conus cancellatus* (jr. syn.: *Conus austini*), a worm-hunting snail from the Gulf of Mexico. *Peptides*. 2013;41:38–44.
25. Finn RD, Jones CG. Natural products—a simple model to explain chemical diversity. *Nat Prod Rep*. 2003;20(4):382–91.
26. Hellinger R, Koehbach J, Soltis DE, Carpenter EJ, Wong GK, Gruber CW. Peptidomics of circular cysteine-rich plant peptides: analysis of the diversity of cyclotides from *Viola tricolor* by transcriptome and proteome mining. *J Proteome Res*. 2015;14(11):4851–62.
27. Dutertre S, Jin A-h, Kaas Q, Jones A, Alewood PF, Lewis RJ. Deep venomomics reveals the mechanism for expanded peptide diversity in cone snail venom. *Mol Cell Proteomics*. 2013;12(2):312–29.
28. Han TS, Teichert RW, Olivera BM, Bulaj G. *Conus* venoms—a rich source of peptide-based therapeutics. *Curr Pharm Des*. 2008;14(24):2462–79.
29. Nelson CJ, Millar AH. Protein turnover in plant biology. *Nat Plants*. 2015;1:15017.
30. Hachiya T, Terashima I, Noguchi K. Increase in respiratory cost at high growth temperature is attributed to high protein turnover cost in *Petunia x hybrida* petals. *Plant Cell Environ*. 2007;30(10):1269–83.
31. Ishihara H, Obata T, Sulpice R, Fernie AR, Stitt M. Quantifying protein synthesis and degradation in *Arabidopsis* by dynamic <sup>13</sup>C<sub>2</sub>O<sub>2</sub> labeling and analysis of enrichment in individual amino acids in their free pools and in protein. *Plant Physiol*. 2015;168(1):74–93.
32. Tayo LL, Lu B, Cruz LJ, Yates JR 3rd. Proteomic analysis provides insights on venom processing in *Conus textile*. *J Proteome Res*. 2010;9(5):2292–301.
33. Yang H, Zubarev RA. Mass spectrometric analysis of asparagine deamidation and aspartate isomerization in polypeptides. *Electrophoresis*. 2010;31(11):1764–72.
34. Quiroz S, Cespedes CL, Alderete JB, Alarcon J. Ceanothane and oleanane-type triterpenes from *Talguenea quinquenervia* have insecticidal activity against *Cydia pomonella*, *Tenebrio molitor* and *Drosophila melanogaster*. *Ind Crop Prod*. 2015;74:759–66.
35. Gallardo O, Tonn CE, Nieto M, Morales G, Giordano OS. Bioactive neoclerodane diterpenoids toward *Tenebrio molitor* larvae from *Teucrium nudicaule* H and *Baccharis spicata* (lam) Beill. *Nat Prod Lett*. 1996;8(3):189–97.
36. Soltani N, Aribi N, Berghiche H, Lakbar S, Smaghe G. Activity of RH-0345 on ecdysteroid production and cuticle secretion in *Tenebrio molitor* pupae in vivo and in vitro. *Pestic Biochem Phys*. 2002;72(2):83–90.
37. Cespedes CL, Molina SC, Munoz E, Lamilla C, Alarcon J, Palacios SM, Carpinella MC, Avila JG. The insecticidal, molting disruption and insect growth inhibitory activity of extracts from *Condalia microphylla* Cav. (Rhamnaceae). *Ind Crop Prod*. 2013;42:78–86.
38. Aribi N, Smaghe G, Lakbar S, Soltani-Mazouni N, Soltani N. Effects of pyriproxyfen, a juvenile hormone analog, on development of the mealworm, *Tenebrio molitor*. *Pestic Biochem Phys*. 2006;84(1):55–62.
39. Isman MB. Plant essential oils as green pesticides for pest and disease management. *ACS Sym Ser*. 2004;887:41–51.
40. Ryan CA, Hass GM, Kuhn RW. Purification and properties of a carboxypeptidase inhibitor from potatoes. *J Biol Chem*. 1974;249(17):5495–9.
41. Aljibory Z, Chen MS. Indirect plant defense against insect herbivores: a review. *Insect Sci*. 2018;25(1):2–23.
42. Kawai M, Sekine-Hayakawa Y, Okiyama A, Ninomiya Y. Gustatory sensation of L- and D-amino acids in humans. *Amino Acids*. 2012;43(6):2349–58.
43. Solms J, Vuataz L, Egli R. The taste of L- and D-amino acids. *Experientia*. 1965;21(12):692–4.
44. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, et al. De novo transcript sequence reconstruction from RNA-seq using the trinity platform for reference generation and analysis. *Nat Protoc*. 2013;8(8):1494–512.
45. Li W, Cowley A, Uludag M, Gur T, McWilliam H, Squizzato S, Park YM, Buso N, Lopez R. The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res*. 2015;43(W1):W580–4.
46. Han X, He L, Xin L, Shan B, Ma B. PeaksPTM: mass spectrometry-based identification of peptides with unspecified modifications. *J Proteome Res*. 2011;10(7):2930–6.
47. Creasy DM, Cottrell JS. Unimod: protein modifications for mass spectrometry. *Proteomics*. 2004;4(6):1534–6.
48. Han Y, Ma B, Zhang K. SPIDER: software for protein identification from sequence tags with de novo sequencing error. *J Bioinform Comput Biol*. 2005;3(03):697–716.
49. Wong KH, Li GQ, Li KM, Razmovski-Naumovski V, Chan K. Optimisation of Pueraria isoflavonoids by response surface methodology using ultrasonic-assisted extraction. *Food Chem*. 2017;231:231–7.
50. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR*. 1995;6(3):277–93.
51. Brunger AT. Version 1.2 of the crystallography and NMR system. *Nat Protoc*. 2007;2(11):2728–33.
52. Koradi R, Billeter M, Wüthrich K. MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph*. 1996;14(1):51–5.
53. DeLano WL. The PyMOL molecular graphics system Version 1.7.4 Schrödinger, LLC. 2002.
54. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr*. 1993;26(2):283–91.
55. Huang SH, Xian JD, Kong SZ, Li YC, Xie JH, Lin J, Chen JN, Wang HF, Su ZR. Insecticidal activity of pogostone against *Spodoptera litura* and *Spodoptera exigua* (Lepidoptera: Noctuidae). *Pest Manag Sci*. 2014;70(3):510–6.
56. Li W, Hu J, Yang J, Yuan G, Guo X, Luo M. Feeding deterrence of common spices against *Helicoverpa armigera* larvae. *Adv Biosci Biotechnol*. 2014;5(13):1025.