The thionin family of antimicrobial peptides

Katharina Höng*, Tina Austerlitz*, Timo Bohlmann*, Holger Bohlmann

Department of Crop Sciences, Institute of Plant Protection, University of Natural Resources and Life Sciences, Vienna, Austria

* These authors contributed equally to this work.

* holger.bohlmann@boku.ac.at

Abstract

Thionins are antimicrobial peptides found only in plants. They are first produced as prepro-proteins and then processed to yield the usually 5 kDa, basic thionin peptide with three or four disulfide bridges. So far, thionins had only been found in some plant families of angiosperms. The One Thousand Plant Transcriptomes Initiative (1KP project) has sequenced the transcriptomes of more than 1000 plant species. We have used these data to search for new thionin sequences which gave 225 hits. After removing doublets these resulted in 133 new thionins. No sequences were found in algae and mosses. The phylogenetically earliest hits were from Selaginella species and from conifers. Many hits were from angiosperm plant families which were previously not known to contain thionins. A large gene family for thionins was found in Papaver. We isolated a genomic clone from Papaver somniferum which confirmed the general genomic structure with two small introns within the acidic domain. We also expressed the thionin encoded by the genomic clone and found that it had antimicrobial activity in vitro, especially against fungi. Previously, we had grouped thionins into four classes. The new data reported here led us to revise this classification. We now recognize only class 1 thionins with eight cysteine residues and class 2 thionins with six cysteine residues. The different variants that we found (and also previously known variants) can all be traced back to one of these two classes. Some of the variants had an uneven number of cysteine residues and it is not clear at the moment what that means for their threedimensional structure.

Introduction

A large variety of relatively small, basic, and often cysteine-rich polypeptides have been isolated from different organisms and shown to have antimicrobial activities in vitro [1–3]. Some prominent examples from animals include magainins from the skin of Xenopus laevis and mammalian defensins and β-defensins. Antimicrobial peptides (AMPs) seem to be distributed ubiquitously in multicellular organisms. Plants have also been shown to contain a variety of different AMPs, including thionins [4,5] and plant defensins [6,7]. It is thought that the molecular targets of the majority of the usually basic AMPs are acidic phospholipids in biomembranes [8]. However, at least for some AMPs, biomembranes may just be a barrier hindering access to their primary target inside the cell [9–11].
Thionins (Fig 1) are a group of AMPs whose toxic activity was apparently first described by Jago and Jago in 1885 [12] who reported that wheat flour contains a substance that was toxic to yeast. Balls and Hale [13] found that the polypeptide responsible for the toxicity could be extracted with petroleum ether and named it purothionin. It was later shown to consist of three different isoforms [14–16] whose sequences have been determined. Purothionins, as the majority of thionins, are basic and cysteine-rich polypeptides with a molecular weight of about 5 kDa. Related polypeptides, all with eight conserved cysteine residues, have been isolated and sequenced from the endosperm of oat and barley and are also present in other species of the Poaceae [17].

Cultivated barley (Hordeum vulgare L.) and also many wild Hordeum species contain a large multigene family of closely related genes for leaf thionins [20,28,29] in addition to the small gene family for endosperm specific hordothionins [19]. Mature leaf thionins have been found in the vacuole [30] and in the cell wall [28] of barley leaves. A prominent feature of leaf thionin genes in barley is their light responsiveness. Etiolated barley seedlings contain a very high level of leaf thionin transcripts which drops drastically after illumination, probably mediated by two photoreceptors, phytochrome and a blue-light-absorbing photoreceptor [31]. Barley leaf thionins and hordothionins have repeatedly been shown to have antimicrobial activities in vitro against different phytopathogenic bacteria and fungi [28,32,33]. This in vitro toxicity and the cellular distribution of barley leaf thionins suggested a role in plant defense.

Thionins are also present in dicot species. One large group of thionins (Fig 1) has been found in different species of mistletoes [17]. The toxic effects of viscotoxins towards mammals have been investigated in detail, for instance by Rosell and Samuelsson [34] because of the medical use of mistletoe extracts in the treatment of cancer [35]. Several thionins with six cysteine residues have been described from species of the family Brassicaceae (Fig 1). These include crambin with six cysteine residues from the seeds of Crambe abyssinica [36]. Crambin has a pI of 5.73 and has no known toxic activities [37]. Schrader-Fischer and Apel found in addition to crambin several novel and highly variable thionins in Crambe abyssinica using a PCR based approach [38]. Arabidopsis contains four thionin genes, of which two have been studied in detail [26,39–42], giving strong support for a role in plant defense.

Thionins are characterized by two consecutive cysteine residues at position 3 and 4. Some variants are known that are missing one of these cysteine residues. Accordingly, thionins have been classified [18] as class 1 with eight cysteine residues, class two with six cysteine residues and variants of class 3 and 4 (Fig 1).

The seed specific crambin has been used intensively for three-dimensional structure analysis [43]. Structures have also been determined for several other thionins. The general three-dimensional structure of thionins shows a compact molecule consisting of two alpha helices and two beta sheets (reviewed by [44]). The three-dimensional structure is related to the toxicity and antimicrobial activity of thionins. Many studies have shown that thionins can induce cell membrane permeability [5]. A phospholipid binding site, which includes the tyrosine at position 13 (Fig 1) which has been shown to be indispensable for toxicity, is present in all known structures of thionins with antimicrobial activity and it has been proposed that thionins interact with negatively charged phospholipids present in the cell membrane. This binding and the withdrawal of phospholipids disturbs the fluidity of the membrane and finally results in membrane lysis [44]. Another possible mechanism has been reported, claiming that thionins which are inserted into the membrane act as water channels. Water is delivered through these channels to the lipophilic center of the biomembrane, leading to local membrane disruption [45].

Thionins have been isolated as the 5 kDa mature peptides (Fig 1) from different plant species but are synthesized as much larger precursors [46]. Cloning of different thionin cDNAs and genomic clones confirmed the original observation and revealed that these mature
Thionins are first synthesized as preproproteins (S1 and S2 Figs). All precursors have a typical N-terminal signal peptide with a conserved signal peptide cleavage site that directs the proprotein into the endoplasmic reticulum. The proprotein consists of the thionin itself and a C-terminal acidic domain. In those cases where genomic DNA sequences are known [19,42,47], there are two small introns in the acidic domain (S2 Fig).

The function of the acidic domain is still not clear. Acidic domains have six cysteine residues (S1 Fig) which are highly conserved as are the cysteine residues in the thionin and are usually acidic (as the name says). An exception are the acidic domains deduced from the Tulipa thionin precursors that have only three cysteine residues (sequences submitted to GenBank by Luyten et al. 1994). In the case of the acidic crambin, the acidic domain too is slightly acidic (S1 Table). The acidic domain of the Arabidopsis THI2.3 proprotein is actually basic (S1 Table). A protein corresponding to the acidic domain has, to our knowledge, never been isolated from plants. To date, there is no experimental information available about possible functions of the acidic domain. But it is clearly not dispensable as shown by the high conservation of the cysteine residues, even in the case of viscotoxin precursors which have several deletions in the acidic domain [23]. Furthermore, Florack et al. [48] found that expression of α-hordeothionin in transgenic tobacco plants without the acidic domain resulted in significantly lower levels of the mature thionin.

Fig 1. Examples of thionins including previous classification of thionins [18]. 1HTH1, Hordeothionin α [19]; 2Barley leaf-thionin DB4 [20]; 3Pyrularia thionin [21]; 4Viscotoxin A3 [22,23]; 5Crambin 2 [24,25]; 6Arabidopsis thaliana thionin 2.1 [26]; 7pTTH20 Neutral wheat thionin [27]; 8Tulipa gesneriana thionin4.1 (Genbank, submitted by Luyten et al. 1994).
One possible function might be that the acidic domain contains information to guide the thionin through the secretory pathway to its final destination (vacuoles, cell walls, protein bodies). Another function of the acidic domain might be to neutralise the basic thionin and thereby protect the cell against its own toxin. A toxic effect of barley leaf thionins has been demonstrated against tobacco protoplasts and against barley protoplasts [49].

We have recently purified a thionin precursor processing enzyme (TPPE) from etiolated barley seedlings which contain large amounts of leaf-specific thionins [31]. We used a fluorescently labelled peptide that incorporated the flanking sequence between barley leaf-specific thionins and their acidic domains to identify a TPPE protein as the subtilase BAJ93208. The barley TPPE was produced as a strep-tagged protein in *E. coli* and used to study the processing of a recombinant leaf-thionin precursor in vitro. It could be shown that this protease produced thionins with a correct C-terminus while the acidic domain was cleaved several times. This explains why a peptide corresponding to the acidic domain has never been isolated from plants. Furthermore, it could be shown that the thionin domain was protected by its threedimensional structure against cleavage by the TPPE [50].

While plant defensins are ubiquitously distributed in plants, thionins were only found in a limited number of plant families. The One Thousand Plant Transcriptomes Initiative (1KP project) which sequenced the transcriptomes of more than 1000 plant species [51] was therefore a good opportunity to study the distribution of thionins in the plant kingdom. We searched the data obtained in that project for thionin sequences. We found thionin sequences from several families which had previously no reported thionins, including a large gene family in the genus *Papaver*. However, a large number of plants was without thionin sequences. We also found a number of sequences that would code for unknown thionin variants which prompted us to propose a new classification system for thionins. A genomic clone was isolated from *P. somniferum* and the encoded thionin was expressed in *E. coli* and shown to have antimicrobial activity.

**Materials and methods**

**Database searches**

We used BLAST to search the translated 1KP data with thionin amino acid sequences. As bait sequences we used a range of well-known thionins, for instance THI2.1 from Arabidopsis, VaTHI2.1 (Viscotoxin A3), CaTHI2.1 (Crambin), HvTHI1.3 (leaf thionin DB4 from Barley) and PpTHI1.1 (Pyrularia thionin) as well as some sequences that were found using the former baits (S2 Table). Searching was done with each of these bait sequences against all samples separately. From these results we removed the sequences which were repeatedly found by the different bait sequences. This resulted in a list of 2130 primary hits. All hits were then curated manually to remove the non-thionin sequences. Typically, thionins would have a double cysteine motif at position 3 and 4 as well as an additional four or six cysteines and an aromatic amino acid at position 13. Furthermore, we also looked for a possible acidic domain extension with its typical three and three cysteines. Manual inspection also allowed us to discover several thionin variants with untypical cysteine motifs. This produced a final list of 225 hits of which some resulted in the same sequences because the 1KP project often sequenced several different tissues for one species. Sometimes these different tissues were in addition pooled and sequenced, also resulting in redundancy. Signal peptides and acidic domain sequences were removed because they were often truncated. In some cases where the end of the signal peptide was not obvious (in case of typical thionin sequences with a basic amino acid at position 1) SignalP 5.0 (http://www.cbs.dtu.dk/services/SignalP/) was used to predict the signal peptide
cleavage site [52]. The cleavage site between thionin domain and acidic domain was predicted according to known thionin sequences being six amino acids after the last cysteine.

Mw and pI were computed at https://web.expasy.org/compute_pi/. Pairwise alignment was performed by using emboss needle (https://www.ebi.ac.uk/Tools/psa/emboss_needle/) [53] and clustal omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) was used for multiple sequence alignments [53]. Sequence logos were created at http://weblogo.threeplusone.com/create.cgi [54]. For creation of the logo two truncated class 1 sequences were removed. For the class 2 logo we removed a truncated sequence and two sequences (ChTHI2.1 and MfTHI2-1) with insertions in the middle for easier comparison with class 1.

**Cloning**

All primers used for cloning can be found in S3 Table. Genomic DNA was prepared from *P. somniferum* seedlings according to [55]. A genomic clone was amplified using primers PsoThi1.5forNco and PsoThi1.5revBam. A fragment of approximately 730 bp was obtained. It was digested with NcoI and BamHI and ligated to the vector fragment of pMAAred [56] digested with the same enzymes, thus replacing the GUS sequence. The insert was sequenced and shown to be a thionin preproprotein with two introns.

For expression of the thionin encoded by the genomic clone we constructed an expression vector for *E. coli* which was based on a modified pETtrx1a vector [57] with an NdeI site at the start codon and a BamHI site behind the stop codon. The vector contained a His-tag in front of the thioredoxin sequence followed by a TEV site. We amplified the thioredoxin part of the vector by PCR using primers pETtrxfor1 and pETtrxTEVrev. The thionin coding sequence of the PsoTHI1.7 genomic clone was also amplified by PCR using the primers TEV1PsoTH1.7for and PsoTH1.7Bamrev. The thioredoxin fusion fragment and the thionin fragment were combined and amplified with primers pETtrxfor1 and PsoTH1.7Bamrev. The final PCR fragment was digested with NdeI and BamHI and cloned into the modified pETtrx1a vector previously digested with the same enzymes. The final expression vector was verified by sequencing.

**Expression and purification of PsoTHI1.7**

The recombinant pETtrx::PsoTHI1.7 vector was transformed into the *E. coli* strain C3030 Shuffle [58]. A 100 ml TB medium pre-culture (containing 50 mg/ml kanamycin) was inoculated with a single colony and incubated overnight at 37°C and 180 rpm. The next day, 2 L TB medium (containing 50 mg/ml kanamycin) were inoculated with 3% v/v pre-culture and grown at 30°C and 140 rpm until the OD$_{600}$ reached 1. Expression was induced by addition of 0.75 mM IPTG and incubated overnight at 20°C and 140 rpm. Cells were harvested by centrifugation for 30 min at 4°C (Sorvall RC6+ centrifuge, Thermo Scientific) and resuspended in 30 ml of ice-cold buffer (20 mM NaH$_2$PO$_4$, 0.5 M NaCl, 50 mM imidazole, pH 7.4). Cells were lysed by sonication (Sonifier W-250D, Branson Ultrasonics) in an ice-water bath. The lysate was clarified by centrifugation at 18,000 rpm for 30 min at 4°C.

The supernatant was used to isolate the fusion protein by immobilized metal ion affinity chromatography on an AKTA purifier10 FPLC system (GE healthcare) with a HisTrap HP 5ml column (GE Healthcare), equilibrated with 10 CV binding buffer (20 mM NaH$_2$PO$_4$, 0.5 M NaCl, 50 mM imidazole, pH 7.4). The sample was loaded with the sample pump at a flow rate of 5 ml/min and the column was washed with binding buffer until the 280 nm absorbance reached a steady baseline. The fusion protein was eluted with 2 CV of Elution buffer (20mM NaH$_2$PO$_4$, 0.5 M NaCl, 500 mM imidazole, pH 7.4) and collected with a fraction collector. Flow-through and the eluted fractions were analyzed by SDS-PAGE [59] using a Mini PRO-TEAN Tetra Cell 1mm (BioRAD). Fractions containing fusion protein were pooled and
dialysed (SpectraPor RC tubing, MWCO 3.5 kDa Spectrumlabs) to a buffer containing 25 mM Tris-HCl pH 8, 0.5 mM EDTA, and 0.25 mM DTT was added directly to the sample before proteolytic digest. The fusion protein was cleaved using TEV protease at a ratio of 1:10 and cleavage was confirmed by PAGE (S3A Fig).

The cleaved fusion protein was dialysed to a binding buffer (20 mM NaH$_2$PO$_4$, 0.5 M NaCl, 5 mM imidazole, pH 7.4) for negative His-tag purification. Binding and loading to a column were performed as described before. The flow through, which contained PsoTHI1.7, was collected and the washing step was omitted. (His)6-TRX, remaining undigested fusion protein and the TEV protease were eluted (20 mM NaH$_2$PO$_4$, 0.5 M NaCl, 500 mM imidazole, pH 7.4) and discarded. The volume of the flowthrough was reduced to 500 μl using an Amicon Ultra centrifugal filter unit (MWCO 3 kDa). The sample was applied to a Superdex 75 10/300 GL column on an ÄKTA purifier10 (GE Healthcare) using the 100 μl sample loop with a flow of 0.5 ml/min. The running buffer was 100 mM TRIS-HCl pH 8, 500 mM NaCl. The purified peptide was run on a PAGE gel (S3B Fig) and silver stained according to [60]. For protein quantification, Thermo Scientific™ Pierce™ BCA Protein Assay kit was used. Page Ruler Unstained (Thermofisher 26614) and Page Ruler Unstained Low range (Thermofisher 26632) were used as PAGE marker.

Antimicrobial assays

Antimicrobial assays were done according to [61]. Fungal test organisms were Botrytis cinerea 05.10 and Fusarium oxysporum f.sp. matthiolae. Both fungi were grown on PDA for two weeks at room temperature. B. cinerea was grown in the dark while F. oxysporum was grown in the light on the lab bench. Spores were harvested by flooding the plate with 10 ml of sterile water and filtering the spore suspension through two layers of cheesecloth. The spores were counted in a Thoma hemocytometer and the number of spores per μl was calculated. Spore concentration was adjusted to 2 x 10$^4$ spores/ml in ½ concentrated PDB. 25 μl of the test peptide were pipetted into the wells of a clear 96 well cell culture microplate (PS,TC, F-bottom with a lid, Greiner Bio-One). Then 75 μl of the spore suspension was added. Final peptide concentrations were 100 μg/ml, 50 μg/ml, 25 μg/ml, 12.5 μg/ml, 6.25 μg/ml, 3.125 μg/ml, 1.56 μg/ml and 0 μg/ml. Plates were incubated in the dark at room temperature. The OD$_{600}$ was measured on a Fluostar Omega Plate reader (BMG Labtech) using the well scan setting with a diameter of 3 mm and nine measuring points and double orbital shaking for 30 sec. before the measurement. OD was measured at the beginning of the experiment and again after 24 h and 48 h. The percentage of growth inhibition was calculated, using the median of all measuring points per well, as follows:

\[
\% \text{growth inhibition} = 100 \times \frac{\text{absorbance of control} - \text{absorbance of test}}{\text{absorbance of control}}
\]

Bacteria were tested in a similar way. Test organisms were Agrobacterium tumefaciens GV3101 and Pseudomonas syringae pv. tomato DC3000. Bacteria were grown in 5 ml LB medium over night at 30˚C. The cultures were centrifuged for 10 min at 7000 rpm and resuspended in 1% tryptone. Bacteria were diluted to an OD$_{600}$ of 0.05 in 1% tryptone before using them for the test. The peptide was pipetted into a 96 well plate as described above and 75 μl of bacterial culture was added. OD$_{600}$ was measured in a Fluostar Omega plate reader as described above. Plates were incubated in the dark at 30˚C and measured again after 24h. Percentage of growth inhibition was calculated as described above.
Results

New thionins discovered by the 1KP project

The 1KP project sequenced the transcriptomes of more than 1000 plant species including algae, mosses, ferns, gymnosperms and angiosperms [51,62,63]. We used these data to search for thionin sequences in the translated protein sequences. Our search produced 225 hits (Table 1) of which some resulted in the same sequences because the 1KP project often sequenced different tissues from the same species. We thus ended up in a total of 133 thionin sequences (Table 2).

No thionin sequences were found in the transcriptomes of algae and mosses. The phylogenetically oldest hits were a class 1 thionin sequence from Selaginella wallacei and from Selaginella stauntoniana (phylum Lycophyta) and two class 2 thionin sequences from the fern Cystopteris fragilis, indicating that thionins have perhaps evolved in the first vascular plants. Furthermore, class 1 thionin sequences were found in the conifers Manoao colensoi, Sequoia-dendron giganteum and Athrotaxis cupressoides (two hits). Already these phylogenetically oldest thionins were produced as preproproteins including a typical pro-domain (“acidic” domain).

The 1KP project had reported that some data have to be treated with caution (https://cyverse.atlassian.net/wiki/spaces/iptol/pages/242170988/Sample+source+and+purity), including the Cystopteris fragilis XXHP sample that was flagged as “Worrisome Contamination”. The two hits for Cystopteris fragilis were from fronds (sample XXHP). A separate tissue sample from young Cystopteris fragilis leaves (not reported to be problematic) contained no hits. The 1KP project had also sequenced the transcriptomes of three other Cystopteris species and these contained also no hits for thionins. We tried to confirm the thionin sequences in the genome of Cystopteris fragilis by PCR but were unable to find these sequences. We must therefore conclude that the thionin sequences from Cystopteris fragilis were probably a contamination. In case of the Selaginella thionins, the hits were from two different species (in total, transcriptomes of 8 species were sequenced). In this case we had no plant material available to confirm the sequence by PCR.

In total, 53 different class 1 sequences and 62 class 2 sequences were found, many of them in plant families of angiosperms from which previously no thionins had been reported (Fig 2). These class 1 and class 2 thionins had an average molecular weight of 4939.41 and 4959.49 Da, respectively. In general, there was more variability within the class 2 thionins than within the class 1 thionins (Table 2 and Fig 3). Of the 50 class 1 sequences that were not truncated at the N-terminus, four started with arginin as the first amino acid and one with isoleucine. All others had lysin as the first amino acid. In case of the class 2 thionins only 40 of 59 N-terminally untruncated sequences had lysin as first amino acid. In addition, seven sequences had an additional amino acid at the N-terminus, thus the double cysteine was at position 4/5 instead of 3/4 as determined by SignalP. One thionin had a N-terminal sequence of KSKKACC.

The average pl was 9.18 and 8.47 for class 1 and class 2 thionins, respectively. Thus, the class 1 thionins are more basic with no member having a pl below 7 while there were eight class two members with an acidic pl. The lowest pl of 4.04 was found for AaTHI2.3 from Arabis alpina. There were also two more thionins from this species with an acidic pl while three had a basic pl. All A. alpina hits were from young leaves. The most basic thionin with a pl of 11.17 was also a class 2 thionin (DsTHI2.1 from Draba sachalinensis) with nine basic residues and no acidic residue. In addition to class 1 and class 2 thionins, variants were discovered that were previously unknown and which could have been evolved from class 1 (seven) or class 2 [11] thionins by deletion or insertion of cysteine residues. Some of these variants had an uneven number of cysteine residues (Table 2).
| species/genus        | family         | order       | Class I/II hits |
|---------------------|----------------|-------------|-----------------|
| Selaginella wallacei| Selaginellaceae | Selaginellales | 1 1             |
| Selaginella stauntoniana | Selaginellaceae | Selaginellales | 1 1             |
| Cystopteris fragilis | Cystopteridaceae | Polypodiales | 2 2             |
| Manoao colensoi     | Podocarpaceae  | Coniferales  | 1               |
| Sequoiadendron giganteum | Cupressaceae  | Coniferales  | 1               |
| Athrotaxis_cupressoides | Cupressaceae   | Coniferales  | 2 2             |
| Papaver             | Papaveraceae   | Ranunculales | 72 21            | 93   |
| Thalictrum_thalictroides | Ranunculaceae  | Ranunculales | 1 2 3           |
| Hakea prostrata     | Proteaceae     | Proteales    | 2 2             |
| Oresitrophe_rupifraga | Saxifragaceae  | Saxifragales | 3 3             |
| Ribes_off_giraldii  | Grossulariaceae | Saxifragales | 1               |
| Sesuvium_           | Aizoaceae      | Caryophyllales | 10 10           |
| Cypsealetum_humifusum | Aizoaceae     | Caryophyllales | 3 1            |
| Trianthema_portulacastra | Aizoaceae    | Caryophyllales | 6 6            |
| Zaleya_pentandra    | Aizoaceae      | Caryophyllales | 1 1            |
| Aerva               | Amaranthaceae  | Caryophyllales | 8 8            |
| Alternanthera_caracassana | Amaranthaceae | Caryophyllales | 2 2            |
| Polygonum_convolvulus | Polygonaceae  | Caryophyllales | 1 1            |
| Polycarpacea_repens | Caryophyllaceae | Caryophyllales | 3 3            |
| Portulaca_mauii      | Portulacaceae  | Caryophyllales | 2 2            |
| Mirabilis_jalapa    | Nyctaginaceae  | Caryophyllales | 1 1            |
| Sassafras_albidum   | Lauraceae      | Laurales     | 1               |
| Lindera_benzoin     | Lauraceae      | Laurales     | 1               |
| Cannabis_sativa     | Cannabaceae    | Rosales      | 7 7             |
| Humulus_lupulus     | Cannabaceae    | Rosales      | 4 4             |
| Urtica_dioica       | Urticaceae     | Rosales      | 2 2             |
| Linum_perenne       | Linaceae       | Malpighiales | 2 2             |
| Chrysobalanus_icaco | Chrysobalanaceae | Malpighiales | 1 1             |
| Viola_tricolor      | Violaceae      | Malpighiales | 1 1             |
| Brassica_nigra      | Brassicaceae   | Brassicales  | 3 3             |
| Arabis_alpina       | Brassicaceae   | Brassicales  | 6 6             |
| Sinapis_alba        | Brassicaceae   | Brassicales  | 3 3             |
| Draba               | Brassicaceae   | Brassicales  | 3 6 9           |
| Cochlearia_officinalis | Brassicaceae | Brassicales | 1 1             |
| Daphne_geraldii     | Thymelaeaceae  | Malvales     | 2 2             |
| Sarcodes_sanguinea  | Ericaceae      | Ericales     | 1 1             |
| Eleusine_coracana   | Poaceae        | Poales       | 4 4             |
| Panicum_miliaceum   | Poaceae        | Poales       | 1               |
| Neurachne           | Poaceae        | Poales       | 3 3 6           |
| Thyridolepis        | Poaceae        | Poales       | 2 2             |
| Lepidosperma_gibsonii | Cyperaceae    | Poales       | 1 1             |
| Polyscias_fruticosa | Araliaceae     | Apiaceae     | 1               |
| Psychotria_ipecacuha | Rubiaceae      | Gentianales  | 2 2             |
| Myristica_fragrans  | Myristicaceae  | Magnoliaceae | 3 3             |
| Hydrocotyle_umbellata | Araliaceae    | Apiaceae     | 1 1             |
| Prunella_vulgaris    | Lamiaceae      | Lamiaceae    | 1 1             |

(Continued)
A large gene family for thionins in *Papaver* species

Our results revealed a large gene family for thionins in *Papaver* species. Different tissues of several *Papaver* species resulted in a total of 93 hits. We found 11 class 1 thionins for *P. rhoesas*, six for *P. somniferum*, three for *P. setigerum* and seven for *P. bracteatum* (Table 2). In addition to class 1 thionins, all four *Papaver* species contained a few class 2 thionins, two in *P. rhoesas* and one each in the other three species. While the majority of *Papaver* thionins had the usual acidic domain, there were a few class 1 thionins from *P. somniferum* and *P. setigerum* which had an unusual pro domain (Fig 4). We found three hits each in both species. These were from developing fruits, leaves and stems in case of *P. setigerum* while the *P. somniferum* hits were from flower buds, leaves and “five samples combined”. Following the thionin domain (pI 8.63) was a sequence with six cysteine residues and a pI of 8.89. Up to there all sequences were identical, which holds also for the DNA sequences. Following this were extensions with several cysteine residues which were rich in proline and serine residues. However, they were all truncated with the exception of EPRK-20112 from *P. setigerum* with a total of 496 amino acids. This extension contained 21 cysteine residues and was also basic with a pI of 8.53.

All *Papaver* species contained in addition variants (Table 2) of class 1 thionins with most of them having an uneven number of cysteine residues. We found variants with seven cysteines in *P. rhoesas*, *P. somniferum* and *P. bracteatum*. *P. setigerum* had one variant with nine cysteines. One *P. setigerum* thionin contained six cysteine residues, thus, seemingly being a class 2 member. However, closer inspection revealed a deletion of two cysteines from a class 1 thionin.

From other thionin genes it is known that the acidic domain contains two small introns (S3 Fig). The transcriptome data from the 1KP project give of course no indication about introns. We have therefore cloned a genomic sequence from *Papaver somniferum* as described in Materials and methods. We derived the primer sequences from the DNA sequence of the *P. somniferum* clone BMRX-2017616 which encodes PsoTHI1.5. The genomic clone that we obtained coded for a typical thionin preprothionin and contained two small introns within the region of the acidic domain (Fig 5). The thionin sequence had two mismatches with PsoTHI1.5 and one with PsoTHI1.6 (Table 2) and was named PsoTHI1.7. It has a pl of 9.54 and a Mw of 4859.75 while the acidic domain has a pl of 4.26 and a Mw of 7724.47. The PsoTHI1.7 pl is similar to that of the other class 1 *Papaver* thionins (Table 2).

Basic thionins have repeatedly been reported to have antimicrobial activity. Having cloned the PsoTHI1.7 we used this sequence to produce an expression vector for *E. coli*. We produced PsoTHI1.7 as a fusion protein with thioredoxin, purified the fusion protein and cleaved it with TEV protease. Finally, PsoTHI1.7 was purified as described in Methods. MALDI-TOF mass spectroscopy showed that the size of the purified peptide corresponded to an oxidized peptide

### Table 1. (Continued)

| species/genus | family | order | Class I/II hits |
|---------------|--------|-------|-----------------|
| *Teucrium chamaedrys* | Lamiaceae | Lamiales | I II both |
| *Orobanchus fasciculata* | Orobanchaceae | Lamiales | 3 3 |
| *Schlegelia parasitica* | Schlegeliales | Lamiales | 2 2 |
| *Heliotropium texanum* | Boraginaceae | Boraginales (Lamiales) | 2 2 |
| *Phoradendron serotinum* | Santalaceae | Santalales | 2 2 |
| *Brachypodium pinnatum* | Poaceae | Poales | 110 115 |
| Total hits | | | 225 |

https://doi.org/10.1371/journal.pone.0254549.t001
### Table 2. All thionin sequences from the 1 KP project.

| Name          | Species/Class | Sequence | pI   | Mw (Da) |
|---------------|---------------|----------|------|---------|
| HvTHI1.1      | Hordeum vulgare | 1 KSCCRSTLGRNCYNLCRVRGAQKLCAGVCRCKLTSSGKCPTGFPK | 9.75 | 4855.81 |
| CaTHI2.1      | Crambe abyssinica | 2 TTCCPSIVARSNFNVCRLPGTSEAICATYTGCIIIPGATCPGDYAN | 5.73 | 4726.42 |
| VaTHI2.1      | Viscum album | 3 KSCCPNTTGRNIYNACRLTGAPRPTCAKLSGCKIISGSTCPSDYPK | 9.30 | 4835.60 |
| AtTHI2.1      | Arabidopsis thaliana | 4 KICCPSNQARNGYSVCRIRFSKGRCMQVSGCQNSDTCPRGWVN | 9.43 | 4811.52 |
| TaTHI3.1      | Triticum aestivum | 4 VDCGANPFKVACFNSCLLGPSTVFQCADFCACRLPAG | 5.90 | 3826.47 |
| TgTHI4.1      | Tulipa gesneriana | 4 KSCFPSTAAKYCYNACRLPGCRPETICAARCGCKIISSGNCPPGYDY | 8.71 | 5040.85 |
| SwTHI1.1      | Selaginella wallacei | 9 KSCCPSTAARNCYNACRLVGTSQTTCASLCGCIHVDGNTCPPNYPS | 8.21 | 4769.40 |
| SsTHI1.1      | Selaginella stauntoniana | 1 KSCCPSTAARNCYNACRLVGTSQTTCASLCGCIHVDGNTCPPNYP | - | - |
| McTHI1.1      | Manoao colensoi | 10 KSCCPSTAARNCYNACRLVGTSQTTCASLCGCIHVDGNTCPPNYPK | 8.51 | 4810.50 |
| SgTHI1.1      | Sequoiadendron giganteum | 1 KSCCPSTAARNCYNACRLVGTSQTTCASLCGCIHVDGNTCPPNYPK | 8.21 | 4857.60 |
| AcTHI1.1      | Athrotaxis cupressoides | 10 KSCCPSTAARNCYNACRLVGTSQTTCASLCGCIHVDGNTCPPNYPK | 8.51 | 4810.50 |
| OrTHI1.1      | Oresitrophe rupifraga | 11 KSCCVNTTARNCYNVCRLTGTQAFCANLCGCIHIDGTTCPPDYPK | 8.21 | 4857.60 |
| OrTHI1.2      | Oresitrophe rupifraga | 6 KSCCPSTAARNCYNACRLVGTSQTTCASLCGCIHVDGNTCPPNYPK | - | - |
| OrTHI1.3      | Oresitrophe rupifraga | 6 ARNCYNVCRLSGASRATCAKLCGCIHINGSTCPSNYPK | - | - |
| VtTHI1.1      | Viola tricolor | 9 KSCCPSTAARNCYNACRLVGTSQTTCASLCGCIHVDGNTCPPNYPS | 8.21 | 4769.40 |
| SpTHI1.1      | Schlegelia parasitica | 6,9 KSCCPSTAARNCYNACRLVGTSQTTCASLCGCIHVDGNTCPPNYP | 2 | - |
| RgTHI1.1      | Ribes giraldii | 6 TTARNCYNTCRLTGTSQARCASLCGCIHITGTTCPPN | - | - |
| DgTHI1.1      | Daphne geraldii | 1 KSCCRNTLGRNCYNTCRFGGAPRPVCASLCDCINIDGTRCPNTHPS | 8.76 | 4956.68 |
| SasTHI1.1     | Sarcodes sanguinea | 1 KSCCKNTTGRNCYNACRFAGGSRPVCATACGCKIISGPTCPRDYPK | 9.25 | 4860.65 |
| PrThi1.1      | Papaver rhoeas | 1 RSCCDSKAGRNCYQACIRRTGATQLCARSCGCRFTRENRCPSSHPW | 9.37 | 5197.91 |
| PrThi1.2      | Papaver rhoeas | 1 KSCCNTTIKRNCYNICRFKFSQETCAKTCGCTLIHGTKCPSRNDN | 9.08 | 5003.87 |
| PrThi1.3      | Papaver rhoeas | 1 KSCCKTTIKRNCYNICRLKFSQETCAKTCGCTLIQGKKCPSRNDN | 9.35 | 5091.00 |
| PrThi1.4      | Papaver rhoeas | 1 KSCCRNTTARNCYNLCRVPGTPREVCAKACDCKIISGKKCPSDYPS | 9.06 | 5041.88 |
| PrThi1.5      | Papaver rhoeas | 1 RICCKDSVARSCFNSCQPGTPRSVCATTCRCRTISGLCPSSYPS | 8.97 | 4703.43 |
| PrThi1.6      | Papaver rhoeas | 1 KICCKSTTARSCFKACRIRLSRETCASTCSCKILTGNCPSDYPK | 9.25 | 4835.72 |
| PrThi1.7      | Papaver rhoeas | 1 KPCCTTYNGKKCNNRCRWDGGNKDSCADMCGCKSCPRIEVG | 8.72 | 4503.17 |
| PrThi1.8      | Papaver rhoeas | 1 KSCCKSTVGRNCYNACRLKFARQVCASTCSCKIVGGNRCPPGYPK | 9.74 | 4905.86 |
| PrThi1.9      | Papaver rhoeas | 1 KSCCKSTLGRNCYNACRLKFPRKTCSTCSCKILKGNRCPSGYPK | 9.80 | 4970.94 |
| PrThi1.10     | Papaver rhoeas | 1 KSCCKSTLGRNCYNACRLRLPRKTCASTCSCKILKGNRCPSGYPK | 9.80 | 4970.94 |
| PrThi1.11     | Papaver rhoeas | 3 KSCCKSTVGRNCYNACRLRFARQVCASTCSCKIVGGNRCPRGYPK | 9.70 | 4918.82 |
| PsoThi1.1     | Papaver somniferum | 1 RSCCGSKAGRNCYQACVRRTGATQLCARSCGCRFTRENRCPSSHPW | 9.60 | 5125.85 |
| PsoThi1.2     | Papaver somniferum | 1 KSCCKTTAARNCYNVCRLTGTSRQVCAATCGCKIISGNKCPRGYDK | 9.38 | 4934.77 |
| PsoThi1.3     | Papaver somniferum | 1 KNCCKTAFGRHCYNLCRLTSPRQNCDAICNCIRWGFSRCPRTYPH | 9.30 | 5295.15 |
| PsoThi1.4     | Papaver somniferum | 1 KSCCKSTVGRNCYNACRLRFARQVCASTCSCKIVGGNRCPRGYPK | 9.75 | 4946.83 |
| PsoThi1.5     | Papaver somniferum | 1 KSCCKSTVGRNCYNACRLRFARQVCASTCSCKIVGGNRCPRGYPK | 9.58 | 4887.76 |
| PsoThi1.6     | Papaver somniferum | 12 KSCCKSTVGRNCYNACRLRFARQVCASTCSCKIVGGNRCPRGYPK | 9.58 | 4918.82 |
| PsoThi1.7     | Papaver somniferum | 13 KSCCKSTVGRNCYNACRLKFARQVCASTCSCKIVGGNRCPRGYPK | 9.54 | 4946.83 |
| PseThi1.1     | Papaver setigerum | 12 KSCCKSTVGRNCYNACRLRFARQVCASTCSCKIVGGNRCPRGYPK | 9.58 | 4887.76 |
| PseThi1.2     | Papaver setigerum | 1 KSCCKSTVGRNCYNACRLRFARQVCASTCSCKIVGGNRCPPGYPK | 9.58 | 4887.76 |
| Name       | Species                          | Class   | Sequence                | Hits | PI    | Mw (Da) |
|------------|----------------------------------|---------|-------------------------|------|-------|---------|
| PbThi1.2   | Papaver bracteatum               |         | KSCCQNTLARNCYNVCRFAGGSREACAKACNCKIITETDCPSDYPK | 2    | 8.50  | 5037.76 |
| PbThi1.3   | Papaver bracteatum               |         | KSCCKSASGRRCYNVCRLRFPRQNCGAICNCIAWWASNTCPYFAPY | 1    | 9.27  | 5270.14 |
| PbThi1.4   | Papaver bracteatum               |         | KSCCKSTFGRNCYNACRLKFPRKTCASTCSCKIVGGNRCPPGYPK | 2    | 9.65  | 4935.85 |
| PbThi1.5   | Papaver bracteatum               |         | KSCCKSTLGRNCYNACRLKFPRKTCASTCNCKILKGNKCPSGYPK | 1    | 9.70  | 4975.95 |
| PbThi1.6   | Papaver bracteatum               |         | KSCCKSTCYNACRLKFPRKTCASTCNCKILKGNKCPSGYPK     | 1    | 9.57  | 4535.45 |
| PbThi1.7   | Papaver bracteatum               |         | KSCCKSTAARECYNACHSAGAPRYVCPYICKCLIISGTKCPPAYRY | 7    | 9.03  | 5051.96 |
| SaTHI1.1   | Sassafras albidum                |         | KSCCRSTTARNCYNVCRLSGSSRPTCASLCDCKIITGTTCPSDYPK | 1    | 8.92  | 4952.69 |
| LbTHI1.1   | Lindera benzoin                  |         | KSCCRSTTARNCYNVCRLAGTPRETCAKLCDCIIITGTTCPSGYPK | 1    | 8.92  | 4960.80 |
| CiTHI1.1   | Chrysobalanus icaco              |         | KSCCVCNTTARNCYNVCRLTGTQAFCANLCGCIHIDGTTCPPDYPK | 1    | 8.21  | 4857.60 |
| PmTHI1.1   | Panicum miliaceum                |         | KSCCKSTLARNCYNVCRLRGARSVCATTCGCKIIKGTKCPPGYPK | 1    | 9.65  | 4840.83 |
| PfTHI1.1   | Polyscias fruticosa              |         | KSCCVNTTARNCYNVCRLTGTQAFCANLCGCIHIDGTTCPPDYPK | 1    | 8.21  | 4857.60 |
| TtTHI1.1   | Thalictrum thalictroides         |         | KSCCPGTLQRNCYNLCRVGGKITSETCAKTCGCKHVVGRVCPPGWQS | 1    | 9.10  | 5030.90 |
| DsTHI1.1   | Draba sachalinensis              |         | KSCCPSTSARNCYNVCRVTGTSQRTCASLCGCKIISGNTCPPGFPS | 3    | 8.95  | 4760.48 |
| EcTHI1.1   | Eleusine coracana                |         | KSCCPDTTKRNCYNVCRHSMKKEICANVCGCKLVSGVKCPRDYPK  | 4    | 9.05  | 5053.01 |
| NaTHI1.1   | Neurachne alopecuroidea          |         | KSCCRSTTARNCYNLCRLRRPQATCASLCGCKIIKGNTCPRDFPK  | 1    | 9.63  | 5038.97 |
| NmTHI1.1   | Neurachne munroi                 |         | KSCCKNTAGRNCYNICRRAGGSQQVCARRCGCIIITGNRCPPNYPK | 1    | 9.63  | 5035.89 |
| TmuTHI1.1  | Thyridolepis multiculmis         |         | KSCCKSTMARNCYNICRFKGPRLVCAQMCGCKIIGGQKCPSDFPK  | 1    | 9.38  | 4976.03 |
| TmiTHI1.1  | Thyridolepis mitchelliana        |         | KSCCKSTMARNCYNICRFKGPRLVCAQMCGCKIIGGQKCPSDFPK  | 1    | 9.38  | 4976.03 |
| OfTHI2.1   | Orobanche fasciculata            |         | KSCCEDTTARYCYNVCRLPGTPRQTCAKICGCIITTSTTCPSNYPK | 3    | 8.72  | 5052.85 |
| PiTHI2.1   | Psychotria ipecacuanha           |         | KISCCCPSTYARSTYNLCSLYKSQIICARLSGCILIDGTSCPSNYPK | 7    | 8.62  | 4954.71 |
| MfTHI2.1   | Myristica fragrans               |         | ESCCPSAKAKNLYNVCRNQYSDPHYFTKSFCANLAGCKLADGKKCEPPYDH | 2    | 8.31  | 5717.46 |
| CfTHI2.1   | Cystopteris fragilis             |         | KSCCPTIVARNQYSVCRFAGASRPECAKLSGCKIVDGECPGGYNR  | 1    | 8.90  | 4796.53 |
| CfTHI2.2   | Cystopteris fragilis             |         | KSCCPSSTARSIYRTCRFGGSTQTCAQISGCKIVSGECPGGYNK  | 1    | 9.13  | 4608.25 |
| HtTHI2.1   | Heliotropium texanum             |         | KSCCPSTTARNTYNVCRLAGTPRPVCASISGCKIITGTKCPKGY  | 2    | -     | -       |
| MjTHI2.1   | Mirabilis jalapa                 |         | KSCCPSTTARNIYNTCRFGGGSRPMCASISGCKIISGTKCPKGYEK | 1    | 9.46  | 4863.67 |
| HpTHI2.1   | Hakea prostrata                 |         | KVACCPSIAASNYYSICRLYGASGPKCAKIEDCKIVDGEECPGSTYP | 2    | 6.24  | 4962.69 |
| LgTHI2.1   | Lepidosperma gibsonii            |         | KVACCPSIAASNYYSICRLYGASGPKCAKIEDCKIVDGEECPGSTYP | 1    | 6.24  | 4962.69 |
| TcTHI2.1   | Teucrium chamaedrys              |         | KSCCPSTSARNIYNTCRLAGGTRPFCASISGCKIVDGKCPTGWDK  | 1    | 9.13  | 4754.49 |
| PrThi2.1   | Papaver rhoeas                   |         | TSCCPSAYARSTYNLCSLYKSQIICARLSGCILIDGTSCPSNYPK  | 7    | 8.62  | 4853.61 |
| PrThi2.2   | Papaver rhoeas                   |         | KICCMNDTRRNRYKDCLNTGASVTSCAGVSGCLIVSGSLCPPNYPY  | 4    | 8.65  | 4932.69 |
| PsoTHi2.1  | Papaver somniferum               |         | TSCCESTKARNSYSVCRLRLGASKNCAKLTGCIIIDGTSCPSDYPI | 3    | 8.63  | 4887.63 |
| PseTHi2.1  | Papaver setigerum                |         | TSCCESTKARNSYSVCRLRLGASKNCAKLTGCIIIDGTSCPSDYPI | 3    | 8.63  | 4887.63 |
| PbThi2.1   | Papaver bracteatum               |         | KICCMNDTRRNRYEVCLNTGASVASCAGVSGCLIISGSVCPPNYPH | 4    | 8.34  | 4894.18 |
| CoTHI2.1   | Cochlearea officinalis           |         | TLCCPNKKTADIYATCRTSGVSKYMCERLSGCKNVSGTCPDILQN  | 1    | 8.62  | 4948.37 |
| HuTHI2.1   | Hydrocotyle umbellata            |         | KSCCPNTTARNIYNTCRITGASRSVCASLSGCIIQSSSTCLPPNTH | 1    | 8.96  | 4818.48 |
| PvTHI2.1   | Prunella vulgaris                |         | KSCCPSTSARNIYNVCRLPGTARETCAKLSGCKIQDPPCVPPFDH  | 1    | 8.66  | 4865.63 |
| TtTHI2.1   | Thalictrum thalictroides         |         | KSCCPSTLKRNIYNACRLKFSQETCAKTSGCKLEDKTCPEGWQK  | 1    | 9.06  | 4985.78 |
| TtTHI2.2   | Thalictrum thalictroides         |         | YVVCCKNIQARNYFNACLSLGSGTSDCLRHSNGNCIRKTGATCPANFPR | 1    | 9.15  | 5326.08 |
| SvTHI2.1   | Sesuvium ventricosum             |         | KSCCPSTTARNTYNVCRLAGTPRPMCASISGCKIITGTKCPKGYEK | 1    | 9.46  | 4897.77 |
| ChTHI2.1   | Cypselea humifusum               |         | KSCCPNTTARNIYNTCRSVQTPAEIYKACRITGGTRSFCAQLSGCKIKKLT | 1    | 9.64  | 5587.53 |
| name          | Species                  | class | sequence                                                                 | hits | pI   | Mw (Da) |
|---------------|--------------------------|-------|---------------------------------------------------------------------------|------|------|---------|
| ChTHI2.2⁵     | Cypsela humifusa         | 2     | LSCCPSTTSVRHFSRCSCRLGGAGRVSGLSCGKGVSGQTCPDRAK                            | 1    | 9.18 | 4810.55 |
| TpTHI2.1      | Trianthema portulacastrum| 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 1    | 9.46 | 4865.71 |
| TpTHI2.2      | Trianthema portulacastrum| 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 1    | 9.38 | 4792.45 |
| TpTHI2.3      | Trianthema portulacastrum| 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | 9.30 | 4863.53 |
| TpTHI2.4      | Trianthema portulacastrum| 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 1    | 9.30 | 4819.47 |
| ApTHI2.1      | Aerva persica            | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | 9.27 | 4935.74 |
| AcTHI2.1      | Alternanthera caracasana | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| PrTHI2.1      | Polycarpacea repens      | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| PrTHI2.2      | Polycarpacea repens      | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| CsTHI2.1      | Cannabis sativa          | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| CsTHI2.2      | Cannabis sativa          | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| CsTHI2.3      | Cannabis sativa          | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| UdTHI2.1      | Urtica dioica            | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| UdTHI2.2      | Urtica dioica            | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| HlTHI2.1      | Humulus lupulus          | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| HlTHI2.2      | Humulus lupulus          | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| LpTHI2.1      | Linum perenne            | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| BnTHI2.1      | Brassica nigra           | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| BnTHI2.2      | Brassica nigra           | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| AsTHI2.1      | Arabis alpina            | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| AsTHI2.2      | Arabis alpina            | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| AuTHI2.1      | Aralia spinosa           | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| AuTHI2.2      | Aralia spinosa           | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| NaTHI2.1      | Neurachne alopecuroidea  | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| NaTHI2.2      | Neurachne alopecuroidea  | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| DoTHI2.1      | Draba sachalinensis      | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| DhTHI2.1      | Draba hispida            | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| DoTHI2.2      | Draba oligosperma        | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| DaTHI2.1      | Draba aizooides          | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| DmTHI2.2      | Draba magellanica        | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| NlTHI2.1      | Neurachne lanigera       | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| NtTHI2.1      | Neurachne lanigera       | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| TmTHI2.1      | Thyridolepis matthielliana| 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| TmTHI2.2      | Thyridolepis multicus    | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| MfTHI2.2      | Myristica fragrans       | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| PsTHI2.1      | Phoradendron serotinum   | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| PsTHI2.2      | Phoradendron serotinum   | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| PrTHI2.3⁶     | Polycarpacea repens      | 2     | KSCCPSTTSNRNYVCRAGSGRSPQASGLSCHIHSNGKCPSTH                              | 2    | -    | -       |
| name       | Species          | class | sequence                                      | hits | pI    | Mw (Da) |
|------------|------------------|-------|-----------------------------------------------|------|-------|---------|
| PrTHI-BERN-18329* | *Papaver rhoeas* | 1X    | KSCCKSTVGRNCYNACRLFARCTSCSCKIVGGNCPFRGYPK    | 62   | 8.47  | 4959.49 |
| PsoTHI-RQNK-5714* | *Papaver somniferum* | 1X    | KSCCYNACRLRFARKVCASTSCSCKIVGGNRCPRGYPK     | 1    | 9.78  | 4402.25 |
| PseTHI-JSCV-2507* | *Papaver setigerum* | 1X    | KSCCKSTVGRNYARQVCASTSCSCKIVGGNRCPRGYPK     | 1    | 9.73  | 4100.91 |
| PseTHI-QOCU-26075* | *Papaver setigerum* | 1X    | KSCCKSTVGRNCYNACRLRFARKVCASTSCSCKIVGGNRCPRGYPK | 1    | 9.70  | 4138.85 |
| PrTHI-SSDU-32671* | *Papaver bracteatum* | 1X    | KSCCKSTACRLKFPRKTCASTNCCKILGNKCPFRGYPK    | 1    | 9.88  | 5774.83 |
| PcTHI-FYSJ-17384* | *Polygonum convolvulus* | 1X    | TVCCPSTARNAYTVCSGGLSSDCALVCGCTQFSTCSESGYD | 7    | 8.32  | 4608.0  |
| NaTHI-ZENX-85534* | *Neurachne alopecuroidea* | 1X    | KVCKTEIIDMYCYDVCHLCRPEHELAPLNTSIDICPDHPK | 7    | 8.32  | 4608.0  |
| SvTHI-EDIT-887* | *Sesuvium ventricosum* | 2X    | RICCRTPTAKTIYTNCRNAG5RQMACAKLSCGIVGGPCPVGCNHSQL | 3    | 9.27  | 5206.14 |
| SvTHI-HIT3-5985* | *Sesuvium portulacastrum* | 2X    | KSCCPFXGRSMCSLSCSCKIVGGNKPFRGYPK    | 1    | 9.04  | 3468.08 |
| SvTHI-OPZ-76* | *Sesuvium ventricosum* | 4X    | SISCPSRNAYTVCSGGLSSDCALVCGCTQFSTCSESGYD | 1    | 9.30  | 3892.41 |
| SvTHI-2JK-458* | *Sesuvium ventricosum* | 1X    | KSCCPITARNYIVCRVPPTFVPVCASLYEYCKLGVSSVCGAMTL | 1    | 8.82  | 5052.00 |
| ChTHI-GJNK-209* | *Cypselea humifusum* | 2X    | VLSLMSQFGVVGICPSASARNIYICRVAEGCSESACRSLGTLNLK | 1    | 8.50  | 5183.13 |
| SpTHI-BERS-225* | *Saleya pentandra* | 2X    | QIVQVEANITYNCRFAGGSRDQCALKLSCHMDTCVPPYTK  | 1    | 8.49  | 4435.06 |
| ALTHI-EMIG-63* | *Aerva lanata* | 6X    | RTKKCCPDYIATWYKVRFAGASRAQACASLRTGCICHTSCPNSNI | 6    | 8.57  | 4913.69 |
| BYNZ-2059306* | *Portulaca mauii* | 2X    | RTKKCCPDYIATWYKVRFAGASRAQACASLRTGCICHTSCPNSNI | 2    | 8.81  | 5170.98 |
| NaTHI-ZENX-85558* | *Neurachne alopecuroidea* | 2X    | KSCCSSSTARNYIVCRPSRESEKCVASCSCKIVGGNKPFRGYPK | 1    | 8.91  | 5225.99 |
| BnTHI-IPWB-14968* | *Brassica nigra* | 2X    | KVCPSLISNVFNVCQFQFRACSLVSQCIQVGNVCPFRGYK | 1    | 8.28  | 4675.46 |
| TpTHI-QNIK-96290* | *Trianthema portulacastrum* | 2X    | KSCCPSTARNYINCRFGGSSRDMCSLSCSCKIVGGNKPFRGYK  | 1    | 8.92  | 5236.08 |

For comparison, some typical previously known thionins have been included. Shown are the thionin name, plant species, class affiliation, sequence, number of hits, pI and molecular weight (Mw).

Cysteine yellow, basic amino acids blue, acidic amino acids red, tyrosine pink, phenylalanine green.

1HTH1, 2Crambin, 3Viscotoxin A3, 4pTTH20, 5Same sequence, unusual pro-domains!, 6Not complete, 7Same sequence, 8Signal peptide cleavage site unclear, 9Same sequence, 10Same sequence, 11Same sequence, 12Same sequence, 13Genomic sequence, 14Same sequence, 15Same sequence.
with all four disulfide bonds. We tested the antimicrobial activity against *Fusarium oxysporum* and *Botrytis cinerea* (Figs 6 and 7). The IC₅₀ was approximately 1 μg/ml for *F. oxysporum* and 6 μg/ml for *B. cinerea* after 48 hours. Microscopic inspection of the hyphae revealed no morphological alterations such as hyperbranching or swelling of hyphal tips (Fig 7). The activity
Fig 4. Alignment of *Papaver* thionin preproproteins with unusual pro domains. RQNK-2047842, BMRX-2023477 and FPYZ-2131139 are from *P. somniferum* (PsoTHI2.1) while EPRK-2020112, FNXH-2029216 and MLPX-2033082 are from *P. setigerum* (PseTHI2.1). Cysteine yellow, basic amino acids blue, acidic amino acids red, tyrosine pink, phenylalanine green, proline grey, serine dark grey.

https://doi.org/10.1371/journal.pone.0254549.g004
against bacteria was much lower. PsoTHI1.7 was not active against Agrobacterium tumefaciens but had some activity against Pseudomonas syringae with an IC$_{50}$ of approximately 25 μg/ml (Fig 6).

**Variants**

In addition to the thionin variants found in Papaver species, variants were in addition found in a number of species from the order Caryophyllales. Sesuvium ventricosum (Fam. Aizoaceae,

![Fig 5. PsoTHI1.7 encoded by the genomic clone from Papaver somniferum. Capital letters are exons while lowercase letters indicate introns. Signal peptide marked grey, thionin domain marked yellow and acidic domain marked cyan. NcoI site marked green and BamHI site marked red.](https://doi.org/10.1371/journal.pone.0254549.g005)

against bacteria was much lower. PsoTHI1.7 was not active against Agrobacterium tumefaciens but had some activity against Pseudomonas syringae with an IC$_{50}$ of approximately 25 μg/ml (Fig 6).

**Variants**

In addition to the thionin variants found in Papaver species, variants were in addition found in a number of species from the order Caryophyllales. Sesuvium ventricosum (Fam. Aizoaceae,

![Fig 6. PsoTHI1.7 antimicrobial activity. A, Fusarium oxysporum; B, Botrytis cinerea. Growth inhibition was calculated after growth for 48h. Values are from 3 biological replicates with standard deviation. C, Pseudomonas syringae; D, Agrobacterium tumefaciens. Growth inhibition was calculated after growth for 24h. Values are from 3 biological replicates with standard deviation.](https://doi.org/10.1371/journal.pone.0254549.g006)
order Caryophyllales) had a typical class 2 thionin and in addition three different class 2 variants with a typical acidic domain. Two of them had an uneven number of cysteines resulting from an additional residue. The other one had only four cysteine residues, missing cysteine number one plus another cysteine residue which was hard to define. This thionin was rather small with only 28 amino acids and was basic with a typical acidic domain. A second Sesuvium species, S. portulacastrum, contained another class 2 variant with a total of only 34 amino acids, this time with five cysteine residues but also with a typical acidic domain. The class 2 variant from Cypsela humifusum (Fam. Aizoaceae) had a typical acidic domain but was probably truncated at the N-terminus as SignalP did not find a signal peptide.

Several other species of the Caryophyllales contained also class 2 variants. Aerva lanata and Portulaca mauii both contained a variant with three consecutive cysteines at position 4–6.
resulting in an uneven number of cysteines. Another class 2 variant with three consecutive cysteines, this time at position 3–5, was found in Neurachne alopecuroidea, a perennial grass. This variant has another additional cysteine residue, resulting in an even number of eight cysteines. *N. alopecuroidea* is a species with class 1 and 2 thionins and which contains in addition variants of class 1 (with an uneven number of cysteines due to an additional residue) and the already mentioned variant of class 2.

Discussion

Thionins have originally been isolated as peptides from a number of plant species, especially mistletoes and cereals. Later on, DNA sequences have revealed thionin sequences in various other plant species. However, compared to defensins, there were many plant families for which no thionins had been reported. The 1KP project was thus an opportunity to extend the knowledge about the distribution of thionins within plants. Since the 1KP project also sequenced a number of algae, mosses, ferns and conifers it was also possible to test if thionins might be found in plant species other than angiosperms. Using various bait sequences we found 225 thionins. The 1KP project sequenced cDNAs from specific plant organs but also pooled samples which resulted in multiple hits for some thionins. In addition, some thionins were found in different samples. This reduced the number to a total of 133 different thionins.

We did not find thionin sequences in algae and mosses. Shelenkov et al. [64] also recently mined the 1KP data for AMP sequences, including thionins. They reported that they found thionin sequences in algae with 36 thionins in *Chlamydomonadaceae* and 10 in *Desmidiaceae*. We had 31 primary hits in each of these families. However, thorough inspection showed that all these hits did not contain thionins. Since Shelenkov et al. [64] reported only numbers without sequences it is at the moment not possible to resolve these differences.

The phylogenetically earliest hits were from *Selaginella wallacei* and *S. stauntoniana* which belonged to class 1. We had no plant material available to confirm the sequences by PCR. We found four class 1 sequences from conifers, one from *Manoa colensoi* (Fam. Podocarpaceae) and the same sequence with two hits from *Athrotaxis cupressoides* (Fam. Cupressaceae). Furthermore, another species from the family Cupressaceae, *Sequoiadendron giganteum*, also had one hit. All other sequences were from angiosperms. Up to the 1KP project there were eight plant families of the angiosperma which were known to contain thionins. The 1KP project identified 25 plant families which contained previously no known thionins. However, there are still many plant families which do not seem to contain thionins, although the negative results from the 1KP project have to be treated with care. All these data are from transcriptomes, sometimes only from one tissue. It is therefore possible that genes which are not expressed in that tissue or only expressed at a low level are missed. In those cases the genome sequence is needed.

According to the data from the 1KP project, thionins emerged first in lycophytes as class 1 thionins and class 1 thionins are also found in conifers. Class 2 thionins then came up within the angiosperms which now contain both classes. Most plant families contain only one class of thionins. However, there are also plant species which contain both classes of thionins. *Poaceae* have mainly class 1 thionins while *Brassicaceae* have mainly class 2 thionins but there are some exceptions. *N. alopecuroidea*, family *Poaceae*, had a class 1 and a class 2 thionin. The majority of thionins in the *Papaver* species are class 1 but all four species that were included in the 1KP project also contain one class 2 thionin (two in the case of *P. rhoeas*).

There was a slightly higher number of class 2 thionins than class 1 and half of the class 1 thionins were from *Papaver* species. This might indicate that class 2 thionins are perhaps more
widely distributed than class 1 thionins. In line with this, the variability among class 2 thionins was higher and acidic thionins were only found within class 2.

**Large gene family for thionins in *Papaver* species**

The 1KP project revealed a large number of thionin sequences in all four *Papaver* species that were sequenced. It was previously not known that *Papaver* species contain thionins, supporting the importance of large scale transcriptome and genome projects. The *Papaver* thionins included typical class 1 and class 2 thionins as well as several variants (see below). In addition, class 1 thionins from *P. somniferum* and *P. setigerum* were discovered which had an unusual pro domain. Most of the clones were truncated but one seemed to be complete with a total of 496 amino acids. It is unlikely that these clones were artifacts because they were found in 3 samples each of 2 different *Papaver* species. However, it would still be important to confirm these with genomic sequences. Furthermore, this unusual structure also raises the question how the pro domain is processed. Could the processing give rise to not only a thionin but also additional protein/s with specific functions? This should be investigated experimentally.

We isolated a genomic clone from *P. somniferum* which allowed us to check for introns because this information is of course not evident from transcriptomics. We found two small introns within the region encoding the acidic domain. These two small introns seem to be generally conserved in thionin genes [19,42,47]. It would be interesting to see if these are already found in the phylogenetically earliest thionin genes.

We expressed the *Papaver* thionin PsoTHI1.7 in *E. coli*. It had significant antimicrobial activity *in vitro* against the fungi we tested but only low activity against *P. syringae*. This is in line with other thionins that have been tested *in vitro*, for instance TaTHI1.3, α-purothionin [65]. Molina et al. [33] also tested wheat purothionins against different pathogens. For three different *B. cinerea* strains they found IC$_{50}$ values between 5 and 12 μg/ml which is comparable to the IC$_{50}$ value of 6 μg/ml for the *B. cinerea* strain that we tested. All *Papaver* class 1 thionins have a basic pI and contain lysin as first amino acid. Most of them also have tyrosin at position 13, indicating that these have most likely antimicrobial or toxic activities [4]. If we compare PsoTHI1.7 with the other *Papaver* class 1 thionins it is safe to conclude that also the other class 1 thionins will have antimicrobial activity, maybe with the exception of those that do not have tyrosin at position 13. With all the data available, it is very likely that the *Papaver* thionin gene family is involved in resistance against pathogens.

**Variants and a new classification for thionins**

Thionins were previously grouped into four classes [18]. With the data obtained by the 1KP project we will now recognize only two classes and variants thereof. The reason for this is that all the variants that we found can be traced back to either class 1 (eight cysteines) or class 2 (six cysteines). With all the different variants (a total of 18 variants) uncovered by the 1KP project we would have to extend the number of classes substantially, making the system confusing. Accordingly, TaTHI3.1 and TgTHI4.1, formerly class 3 and class 4 thionins [18], respectively, are both variants of class 1. Many of the variants that we found here, six of the seven class 1 variants and six of the 11 class 2 variants, have an uneven number of cysteine residues. This raises the question about the structure of the peptides. We can assume that most of these cysteines form disulfide bridges, as has repeatedly been shown for many class 1 and class 2 thionins, leaving probably one single cysteine. It could be possible that the single cysteines engage in forming dimers. It has been shown before that thionins without single cysteines form dimers [44]. However, formation of intermolecular disulfide bridges has to be determined experimentally by isolating the peptides from the plants or by expressing them in expression hosts.
Supporting information
S1 Fig. Sequences of different thionin proproteins. (DOCX)
S2 Fig. General structure of thionin preproproteins and position of the introns. (DOCX)
S3 Fig. Expression and purification of PsoTH1.7. (DOCX)
S1 Table. pI of some thionins, acidic domains and proproteins. (DOCX)
S2 Table. Bait sequences that were used to search the 1KP data. (DOCX)
S3 Table. Primers used in this work. (DOCX)

Acknowledgments
We thank Gane Ka-Shu Wong and Michael K. Deyholos for access to data from the 1KP project. We appreciate Clemens Grünwald-Gruber for performing the mass spectroscopy.

Author Contributions
Conceptualization: Holger Bohlmann.
Data curation: Holger Bohlmann.
Investigation: Katharina Höng, Tina Austerlitz, Timo Bohlmann.
Supervision: Holger Bohlmann.
Visualization: Holger Bohlmann.
Writing – original draft: Katharina Höng, Tina Austerlitz, Holger Bohlmann.
Writing – review & editing: Holger Bohlmann.

References
1. Broekaert WF, Cammue BPA, DeBolle MFC, Thevissen K, DeSamblanx GW, Osborn RW. Antimicrobial peptides from plants. Critical Reviews in Plant Sciences. 1997; 16(3):297–323.
2. Jønsson H, Hamill P, Hancock RE. Peptide antimicrobial agents. Clin Microbiol Rev. 2006; 19(3):491–511. https://doi.org/10.1128/CMR.00056-05 PMID: 16847082
3. Yeaman MR, Yount NY. Mechanisms of antimicrobial peptide action and resistance. Pharmacol Rev. 2003; 55(1):27–55. https://doi.org/10.1124/pr.55.1.2 PMID: 12615953
4. Bohlmann H. The Role of Thionins in Plant-Protection. Critical Reviews in Plant Sciences. 1994; 13(1):1–16.
5. Bohlmann H, Apel K. Thionins. Annu Rev Physiol Plant Mol Biol. 1991; 42:227–40.
6. Parisi K, Shafee TMA, Quinbar P, van der Weerden NL, Bleackley MR, Anderson MA. The evolution, function and mechanisms of action for plant defensins. Semin Cell Dev Biol. 2019; 88:107–18. https://doi.org/10.1016/j.semcdb.2018.02.004 PMID: 29432955
7. Thomma BP, Cammue BP, Thevissen K. Plant defensins. Planta. 2002; 216(2):193–202. https://doi.org/10.1007/s00425-002-0902-6 PMID: 12475332
8. Zasloff M. Antimicrobial peptides of multicellular organisms. Nature. 2002; 415(6870):389–95. https://doi.org/10.1038/415389a PMID: 11807545

PLOS ONE | https://doi.org/10.1371/journal.pone.0254549  July 14, 2021 20 / 23
9. Cardoso MH, Menegueti BT, Costa BO, Buccini DF, Oshiro KGN, Preza SLE, et al. Non-Lytic Antibacterial Peptides That Translocate Through Bacterial Membranes to Act on Intracellular Targets. Int J Mol Sci. 2019; 20(19). https://doi.org/10.3390/ijms20194877 PMID: 31581426

10. Hale JD, Hancock RE. Alternative mechanisms of action of cationic antimicrobial peptides on bacteria. Expert Rev Anti Infect Ther. 2007; 5(6):951–9. https://doi.org/10.1586/14787210.5.6.951 PMID: 18039080

11. Otvos L Jr. Antibacterial peptides and proteins with multiple cellular targets. J Pept Sci. 2005; 11(11):697–706. https://doi.org/10.1002/psc.698 PMID: 16059966

12. Ohtani S, Okada T, Yoshizumi H, Kagayama H. Complete primary structures of two subunits of purothionin A, a lethal protein for brewer’s yeast from wheat flour. J Biochem. 1977; 82(3):753–67. https://doi.org/10.1093/oxfordjournals.jbchem.a131752 PMID: 914810

13. Balls AK, Hale WS, Harris TH. A crystalline protein obtained from a lipoprotein of wheat flour. Cereal Chemistry. 1942; 19:279–88.

14. Nimmo CC, Kasarda DD, Lew EJ. Physical characterization of the wheat protein purothionin. J Sci Food Agric. 1974; 25(6):951–9. https://doi.org/10.1002/jsfa.2740250603 PMID: 18039080

15. Mak AS, Jones BL. The amino acid sequence of wheat beta-purothionin. Can J Biochem. 1976; 54(10):835–42. https://doi.org/10.1139/o76-120 PMID: 990986

16. Fisher N, Redman DG, Elton GAH. Fractionation and characterization of purothionin. Cereal Chemistry. 1968; 45:48–57.

17. Bohlman H, Apel K. Thionins. Annual Review of Plant Physiology and Plant Molecular Biology. 1991; 42:227–40.

18. Bohlmann H, Apel K, Garcia-Olmedo F. Thionins. Plant Molecular Biology Reporter. 1994; 12(2):S75–S.

19. Rodriguez-Palenzuela P, Pintor-Toro JA, Carbone P, Garcia-Olmedo F. Nucleotide sequence and endosperm-specific expression of the structural gene for the toxin alpha-hordothionin in barley (Hordeum vulgare L.). Gene. 1988; 70(2):271–81. https://doi.org/10.1016/0378-1119(88)90199-0 PMID: 2850969

20. Bohlmann H, Apel K. Isolation and characterization of cDNAs coding for leaf-specific thionins closely related to the endosperm-specific hordothionin of barley (Hordeum vulgare L.). Molecular General Genetics. 1987; 207:446–54.

21. Vernon LP, Evett GE, Zeikus RD, Gray WR. A toxic thionin from Pyrularia pubera: purification, properties, and amino acid sequence. Arch Biochem Biophys. 1985; 238(1):18–29. https://doi.org/10.1016/0003-9861(85)90136-5 PMID: 3985614

22. Samuelsson G, Seger L, Olson T. The amino acid sequence of oxidized viscotoxin A3 from the European mistletoe (Viscum album L, Loranthaceae). Acta Chem Scand. 1968; 22(8):2624–42. https://doi.org/10.3891/acta.chem.scand.22-2624 PMID: 5719166

23. Schrader G, Apel K. Isolation and characterization of cDNAs encoding viscotoxins of mistletoe (Viscum album). Eur J Biochem. 1991; 198(3):549–53. https://doi.org/10.1111/j.1432-1033.1991.tb16049.x PMID: 1710983

24. Teeter MM, Mazer JA, L’Italien JJ. Primary structure of the hydrophobic plant protein crambin. Biochemistry. 1981; 20(19):5437–43. https://doi.org/10.1021/bi00052a013 PMID: 6895315

25. Vermeulen JAWH, Lamerichs RMJN, Berliner LJ, De Marco A, Llinas M, Boelens R, et al. 1H-NMR characterization of two crambin species. FEBS Lett. 1987; 219:426–30.

26. Epple P, Apel K, Bohlmann H. An Arabidopsis thaliana thionin gene is inducible via a signal transduction pathway different from that for pathogenesis-related proteins. Plant Physiol. 1995; 109:813–20. https://doi.org/10.1104/pp.109.3.813 PMID: 8552715

27. Castagnaro A, Marana C, Carbonero P, Garcia-Olmedo F. Extreme divergence of a novel wheat thionin generated by a mutational burst specifically affecting the mature protein domain of the precursor. J Mol Biol. 1992; 224(4):1003–9. https://doi.org/10.1006/jmbi.1992.1021 PMID: 1569564

28. Bohlmann H, Clausen S, Behnke S, Giese H, Hiller C, Reimann-Philipp U, et al. Leaf-specific thionins of barley-a novel class of cell wall proteins toxic to plant-pathogenic fungi and possibly involved in the defence mechanism of plants. EMBO J. 1988; 7(6):1559–65. PMID: 2722812

29. Gausing K. Thionin genes specifically expressed in barley leaves. Planta. 1987; 171(2):241–6. https://doi.org/10.1007/BF00391100 PMID: 24227332

30. Reimann-Philipp U, Schrader G, Martinova E, Barkholt V, Apel K. Intracellular thionins of barley. A second group of leaf thionins closely related to but distinct from cell wall-bound thionins. J Biol Chem. 1989; 264(15):8978–84. PMID: 2722812

31. Reimann-Philipp U, Behnke S, Batschauer A, Schafer E, Apel K. The effect of light on the biosynthesis of leaf-specific thionins in barley, Hordeum vulgare. Eur J Biochem. 1989; 182(2):283–9. https://doi.org/10.1111/j.1432-1033.1989.tb14828.x PMID: 2737201
32. Florack DE, Visser B, De Vries PM, Van Vuurde JWL, Stekema W. Analysis of the toxicity of purothionins and hordeothionins for plant pathogenic bacteria. Netherland Journal Plant Pathology. 1993; 99:259–68.

33. Molina A, Ahi-Goy P, Fraile A, Sanchez-Monge R, Garcia-Olmedo F. Inhibition of bacterial and fungal plant pathogens by thionins of types I and II. Plant Science. 1993; 92:169–77.

34. Rosell S, Samuelsson G. Effect of mistletoe viscotoxin and phoratoxin on blood circulation. Toxicon. 1966; 4(2):107–10. https://doi.org/10.1016/0041-0101(66)90005-5 PMID: 6005043

35. Selawry OS, Vester F, Mai W, Schwartz MR. On the identification of the constituents of Viscum album. Phytochemistry. 1966; 4(2):107–10. https://doi.org/10.1016/0041-0101(66)90005-5 PMID: 6005043

36. Vanetten CH, Nielsen HC, Peters JE. A crystalline polypeptide from the seed of Crambe abyssinica. Phytochemistry. 1965; 4(3):467–73.

37. Van Etten CH, Gagne WE, Robbins DJ, Booth AN, Daxenbichler ME, Wolff IA. Biological evaluation of Crambe seed meals and derived products by rat feeding. Cereal Chemistry. 1969; 46:145–55.

38. Schrader-Fischer G, Apel K. Organ-specific expression of highly divergent thionin variants that are distinct from the seed-specific crambin in the crucifer Crambe abyssinica. Mol Gen Genet. 1994; 245(3):380–9. https://doi.org/10.1007/BF00290119 PMID: 7816048

39. Bohlmann H, Vignutelli A, Hilpert B, Miersch O, Wasternack C, Apel K. Wounding and chemicals induce expression of the Arabidopsis thaliana gene Th2.1, encoding a fungal defense thionin, via the octadecanoid pathway. FEBS Lett. 1998; 437(3):281–6. https://doi.org/10.1016/s0014-5793(98)01251-4 PMID: 9824308

40. Eppl P, Apel K, Bohlmann H. Overexpression of an endogenous thionin enhances resistance of Arabidopsis against Fusarium oxysporum The Plant Cell. 1997; 9:509–20. https://doi.org/10.1105/tpc.9.4.509 PMID: 9144959

41. Eppl P, Vignutelli A, Apel K, Bohlmann H. Differential induction of the Arabidopsis thaliana Th2.1 gene by Fusarium oxysporum f. sp. matthiolae. Mol Plant Microbe Interact. 1998; 11(6):523–9. https://doi.org/10.1094/MPMI.1998.11.6.523 PMID: 9612950

42. Vignutelli A, Wasternack C, Apel K, Bohlmann H. Systemic and local induction of an Arabidopsis thionin gene by wounding and pathogens. The Plant Journal. 1998; 14(3):285–95. https://doi.org/10.1046/j.1365-313x.1998.00117.x PMID: 9628023

43. Stec B. Plant thionins—the structural perspective. Cell Mol Life Sci. 2006; 63(12):1370–85. https://doi.org/10.1007/s00018-005-5574-5 PMID: 16715411

44. Stec B, Markman O, Rao U, Helfron G, Henderson S, Vernon LP, et al. Proposal for molecular mechanism of thionins deduced from physico-chemical studies of plant toxins. J Pept Res. 2004; 64(6):210–24. https://doi.org/10.1011/j.1399-3011.2004.00187.x PMID: 15613085

45. Oard SV. Deciphering a mechanism of membrane permeabilization by alpha-hordeothionin peptide. Biochim Biophys Acta. 2011; 1808(6):1737–45. https://doi.org/10.1016/j.bbamem.2011.02.003 PMID: 21315063

46. Ponz F, Paz-Ares J, Hernandez-Lucas C, Garcia-Olmedo F, Carbonero P. Cloning and nucleotide sequence of a cDNA encoding the precursor of the barley toxin alpha-hordeothionin. Eur J Biochem. 1986; 156(1):131–5. https://doi.org/10.1111/j.1432-1033.1986.tb09557.x PMID: 3082629

47. Holtof S, Apel K, Bohlmann H. Specific and different expression patterns of two members of the leaf thionin multigene family of barley in transgenic tobacco. Plant Science. 1995; 111(1):27–37.

48. Florack DE, Dirksse WG, Visser B, Heidekamp F, Stekema WJ. Expression of biologically active hordeothionins in tobacco. Effects of pre- and pro-sequences at the amino and carboxyl termini of the hordeothionin precursor on mature protein expression and sorting. Plant Mol Biol. 1994; 24(1):83–96. https://doi.org/10.1007/BF00405776 PMID: 8111029

49. Lee D-W, Apel K, Binding H. The Formation of Leaf Thionins and their Effect on the Viability of Isolated Protoplasts of Barley (Hordeum vulgare L.). Journal of Plant Physiology. 1995; 145(1–2):131–6.

50. Plattner S, Gruber C, Stadlmann J, Widmann S, Gruber CW, Altmann F, et al. Isolation and Characterization of a Thionin Proprotein-processing Enzyme from Barley. J Biol Chem. 2015; 290(29):18056–67. https://doi.org/10.1074/jbc.M115.647859 PMID: 26013828

51. Leebens-Mack JH. One thousand plant transcriptomes and the phylogenomics of green plants. Nature. 2019; 574(7780):679–88. https://doi.org/10.1038/s41586-019-1693-z PMID: 31645766

52. Almagro Armenteros JJ, Tsigris KD, Sanderby CK, Petersen TN, Winther O, Brunak S, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol. 2019; 37(4):420–3. https://doi.org/10.1038/s41598-019-0036-2 PMID: 30778233

53. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. 2019; 47(W1):W636–W41. https://doi.org/10.1093/nar/gkz268 PMID: 30876793
54. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome Res. 2004; 14(6):1188–90. https://doi.org/10.1101/gr.849004 PMID: 15173120

55. Edwards K, Johnstone C, Thompson C. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucleic Acids Research. 1991; 19(6):1349. https://doi.org/10.1093/nar/19.6.1349 PMID: 2030957

56. Ali MA, Shah KH, Bohlmann H. pMAA-Red: a new pPZP-derived vector for fast visual screening of transgenic Arabidopsis plants at the seed stage. BMC Biotechnol. 2012; 12:37. https://doi.org/10.1186/1472-6750-12-37 PMID: 22747516

57. Bogomolovas J, Simon B, Sattler M, Stier G. Screening of fusion partners for high yield expression and purification of bioactive viscosotoxins. Protein Expr Purif. 2009; 64(1):16–23. https://doi.org/10.1016/j.pep.2008.10.003 PMID: 18983922

58. Lobstein J, Emrich CA, Jeans C, Faulkner M, Riggs P, Berkmen M. SIFuffle, a novel Escherichia coli protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. Microb Cell Fact. 2012; 11:56. https://doi.org/10.1186/1475-2859-11-56 PMID: 22569138

59. Schägger H. Tricine-SDS-PAGE. Nat Protoc. 2006; 1(1):16–22. https://doi.org/10.1038/nprot.2006.4 PMID: 17406207

60. Gharahdaghi F, Weinberg CR, Meagher DA, Imai BS, Mische SM. Mass spectrometric identification of proteins from silver-stained polyacrylamide gel: a method for the removal of silver ions to enhance sensitivity. Electrophoresis. 1999; 20(3):601–5. PMID: 10217175

61. Broekaert WF, Terras FRG, Cammue BPA, Vanderleyden J. An automated quantitative assay for fungal growth inhibition. FEMS Microbiol Lett. 1990; 69:55–60.

62. Wong GK, Soltis DE, Leebens-Mack J, Wickett NJ, Barker MS, Van de Peer Y, et al. Sequencing and Analyzing the Transcriptomes of a Thousand Species Across the Tree of Life for Green Plants. Annu Rev Plant Biol. 2020; 71:741–65. https://doi.org/10.1146/annurev-arplant-042916-041040 PMID: 31851546

63. Carpenter EJ, Matasci N, Ayyampalayam S, Wu S, Sun J, Yu J, et al. Access to RNA-sequencing data from 1,173 plant species: The 1000 Plant transcriptomes initiative (1KP). Gigascience. 2019; 8(10).

64. Shelenkov A, Slavokhotova A, Odintsova T. Predicting Antimicrobial and Other Cysteine-Rich Peptides in 1267 Plant Transcriptomes. Antibiotics (Basel, Switzerland). 2020; 9(2). https://doi.org/10.3390/antibiotics9020060 PMID: 32032999

65. Terras F, Schoofs H, Thevissen K, Osborn RW, Vanderleyden J, Cammue B, et al. Synergistic Enhancement of the Antifungal Activity of Wheat and Barley Thionins by Radish and Oilseed Rape 2S Albumins and by Barley Trypsin Inhibitors. Plant Physiol. 1993; 103(4):1311–9. https://doi.org/10.1104/pp.103.4.1311 PMID: 12232024