bioRxiv preprint doi: https://doi.org/10.1101/2020.09.12.294728; this version posted September 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

# Arginine citrullination of proteins as a specific response mechanism in *Arabidopsis thaliana*

Claudius Marondedze[1,2,*], Giuliano Elia[3], Ludivine Thomas[1], Aloysius Wong[4] and Chris Gehring[1,5,*]

1: Division of Biological and Chemical Science and Engineering, 4700 King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia

2: Rijk Zwaan, De Lier 2678 ZG, Netherlands

3: Philochem AG, Libernstrasse, CH-8112 Otelfingen, Switzerland

4: Department of Biology, College of Science and Technology, Wenzhou-Kean University, 88 Daxue Road, Ouhai, Wenzhou, Zhejiang Province 325060, China

5: Department of Chemistry, Biology & Biotechnology, University of Perugia, I-06121 Perugia, Italy

*Corresponding authors: christophandreas.gehring@UniPG.it; cmarondedze@gmail.com

## Abstract

Arginine citrullination, also referred to as arginine deimination, is a post-translational modification involved in an increasing number of physiological processes in animals, including histone modifications and transcriptional regulation, and in severe diseases such as rheumatoid arthritis and neurodegenerative conditions. It occurs when arginine side chains are deiminated and converted into side chains of the amino acid citrulline, a process catalysed by a family of $Ca^{2+}$-dependent peptidyl arginine deiminases (PADs). PADs have been discovered in several mammalian species and in other vertebrates, like birds and fish, but have not been observed in bacteria, lower eukaryotes or higher plants.

Here we show, firstly, that the *Arabidopsis thaliana* proteome does contain citrullinated proteins; secondly and importantly, that the citrullination signature changes in response to cold stress. Among the citrullinated proteins are DNA- or RNA-binding proteins thus implying a role for it the control of the transcriptional programming in plant cells. Thirdly, through sequence and structural analysis, we identify one arabidopsis protein, currently annotated as agmatine deiminase (At5g08170), as a candidate protein arginine deiminase. Finally, we show biochemical evidence that AT5G08170 can citrullinate peptides from LHP1-interacting factor 2 (AT4G00830) an RNA-binding protein that has been identified

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.12.294728; this version posted September 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

as citrullinated in cell suspension cultures of *Arabidopsis thaliana* roots. In addition, we show that, *in vitro*, agmatine deiminase can undergo auto-citrullination. In conclusion, our work established the presence of protein arginine citrullination in higher plants and assigns it a role in post-translational modifications during abiotic stress responses.

## Introduction

Citrulline, an intermediate of the urea cycle, is not comprised among the 20 amino acids that constitute the building blocks of the proteins, However, it can be formed in proteins and peptides as a result of arginine deimination (Fig. 1). The presence of citrulline in protein extracts was first reported well over half a century ago in the red alga *Chondrus crispus* [1]. It was later discovered in hair follicles and established unequivocally in peptide linkages [2, 3].

The enzymatic conversion of an arginine residue into a citrulline in a protein results in the loss of a positive charge that perturbs a number of intra- and inter-molecular electrostatic interactions causing partial unfolding of the citrullinated protein [4]. Consequently, these may become more prone to protease attack and may become immunogenic. A disturbance of the cellular citrullination rates in proteins has been shown to play a central role in the pathogenesis of a number of autoimmune and neurological disorders such as rheumatoid arthritis [5, 6]. Citrullination is also important in physiology. Examples include terminal differentiation of the epidermis, apoptosis or the modulation of brain plasticity during postnatal life, as well as the regulation of pluripotency [7] and cancer [8].

Importantly, citrullination has been reported to have a role in the epigenetic regulation of gene transcription [9]. In several species, histones H2A, H3 and H4 have been shown to be citrullinated by the peptidyl arginine deiminase type IV (PAD4) in the nucleus of granulocytes where the reaction uses both arginine and methylarginine (but not dimethylarginine) as substrates. It has been hypothesized that arginine demethylation by PAD4 might act to “reverse” the transcriptional co-activator function of the protein arginine methyltransferases (PRMTs) by hydrolyzing methylated arginines to citrulline [9, 10]. PAD-catalyzed hydrolysis of methylated arginines would not truly reverse the modification but instead leave peptidylcitrulline as a final product.

To-date, reversibility by an arginine demethylase has not been reported. However, the PAD4 enzyme, which can remove the methyl groups introduced by the PRMTs, might serve an analogous function, possibly together with as yet undiscovered transaminases,

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.12.294728; this version posted September 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

thereby restoring arginine groups in histones. The structural basis of PAD4 activation by $Ca^{2+}$ in humans has also been reported [11].

PADs have been discovered in several mammalian species and in other vertebrates, like birds and fish, but have not been observed in bacteria, lower eukaryotes or higher plants [12]. A notable exception is the anaerobic human pathogen *Porphyromonas gingivalis* that harbours an enzyme termed PPAD that shows little sequence similarity to the human PAD enzymes, but can convert both peptidyl arginine and free arginine to peptidylcitrulline and free citrulline, respectively. However, PPAD appears to be $Ca^{2+}$-independent [12].

In plants, the presence of the non-proteinogenic amino acid citrulline was known for over a century (review see [13]). More recently many aspects of its metabolism and localization, transport and accumulation have been elucidated [13]. However, citrullination as a post-translational modification catalyzed by PADs has not been seriously considered and in fact it has been suggested that in the absence of PAD function *in planta*, the regulation of protein function through citrullination was an unlikely proposition (for review see [14]).

Here we asked the following questions: (1) Is there evidence for citrullinated proteins in arabidopsis, and if so, (2) are citrullination signatures stimulus-specific and (3), what are plausible *Arabidopsis thaliana* deiminase candidates that can citrullinate proteins.

## Results and discussion

**Identification of citrullinated peptides and proteins in *Arabidopsis thaliana*.** Cell suspension cultures derived from roots of *Arabidopsis thaliana* (ecotype Columbia-0) were used in this study to examine the presence of citrullinated arginines in higher plants. Cells (control) were subjected to nuclear fraction enrichment and citrullinated proteins were enriched by immunoprecipitation followed by western blotting (Supplementary Fig. 1). Given the novel exciting results from western blot analysis and the relatively uncharacterized nature of the anti-ctrullination antibody in plant systems, we validated the results by tandem mass spectrometry (LC-MS/MS) following in-blot digestion or direct in-solution digestion of the immunoreactant proteins. The LC-MS/MS analysis shows that citrulline was detected on proteins from the immunoreactant samples including the

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.12.294728; this version posted September 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

detection of citrullinated peptide of chromatin remodelling 34 (CHR34; AT2G21450; Fig. 2). Here we noted that citrullination was rather targeted and to specific arginine residues, for example, only two of the 52 arginine residues were citrullinated in CHR34. Besides, a total of 14 citrullinated peptides, corresponding to 14 individual proteins were identified from the control samples. Ten of these did not show any changes in response to stimuli (described below) and are considered a baseline citrullination signature of the nuclear enriched proteome in our sample (Table 1). One of the citrullinated proteins, target of Rapamycin (TOR) related protein (AT1G50030), belongs to the family of phosphatidylinositol 3-kinase. They are targets of the anti-proliferative drug rapamycin and are annotated as having a role in many processes, including metabolic re-programming [15], transcription and rRNA processing and protein self-association. In addition, it may also contribute to ending seed dormancy. TOR is also a DNA-binding protein (Table 1). Other notable DNA-binding proteins identified include far-red impaired responsive protein (FAR1; AT4G12850) and methyltransferase MT-A70 (AT4G09980). The FAR1 is a transcription factor co-opted from ancient mutator-like transposase, which is located in the nucleus and is a component of phytochrome A mediated far-red light signalling [16]. The extensive similarity of FAR1 with the mutator-like transposases points to a multitude of roles for FAR1 in developmental and physiological processes including stress responses, programmed cell death, reactive oxygen species (ROS) homeostasis and abscisic acid (ABA) signalling and branching. These diverse roles of FAR1 may implicate citrullination in modulating DNA-binding, light signaling and stress responses. However, with methyltransferase MT-A70, citrullination may be linked to seed dormancy and embryogenesis. Methyltransferase MT-A70 is a S-adenosylmethionine-dependent methyltransferase with a role in seed dormancy release, which in turn is controlled by the balance between ABA, GA and ethylene [17]. Methyltransferase MT-A70 gene is also strongly up-regulated in suspension culture cells exposed to cold (4°C) for 24 h and then to cordycepin (3'-deoxyadenosine, an inhibitor of transcription) 1 h prior to harvesting cells. Furthermore, a ubiquitin associated protein, RING/FYVE/PHD zinc finger protein (AT3G02890) is also involved in DNA methylation, RNA interference and the biological process of methylation-dependent chromatin silencing. However, the role of citrullination on all these biological processes is yet to be established.

Protein decapping 5 (DCP5) (At1G26110), one of the three RNA-binding proteins identified with a citrullinated peptide is a positive regulator of cytoplasmic mRNA and a negative regulator of translation [18]. In addition to having protein binding and protein

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.12.294728; this version posted September 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

homodimerization activities, it is transcriptionally up-regulated by the herbicides primisulfuron-methyl. DCP5 has been identified as a candidate protein interacting with the arabidopsis Cold Shock Domain Protein 3 (AtCSP3) that shares an RNA chaperone function with *E. coli* cold shock proteins and regulates freezing tolerance during cold acclimation [19]. Like DCP5, LHP1-INTERACTING FACTOR 2 (LIF2, AT4G00830) is an RNA-binding protein that encodes a heterogeneous nuclear ribonucleoprotein that is involved in the plant innate immune response and may function as a suppressor of cell-autonomous immunity [20]. LIF2 is involved in cell identity and cell fate decision and has been postulated to modulate the activity of LHP1, a polycomb repressive complex 1 subunit, at specific loci, during specific developmental windows or in response to environmental cues that control cell fate determination [21]. These key functional roles of LIF2 correlate well with the role of citrullination as described in the animal systems [7].

Three proteins were associated with protein binding activities and these include armadillo (ARM)-repeat protein (AT4G31890), YCF4 (ATCG00520) and thioredoxin family protein (AT4G10000). Armadillo protein is part of a large family with diverse functions in many eukaryotes. It has been proposed that the structure of ARM-repeats enables functional versatility [22]. The tandem ARM-repeat can fold itself to form a superhelix that can enable protein-protein interactions and in turn activate or modulate signalling networks. It is therefore conceivable that citrullination could affect cellular signalling events via armadillo citrullination. On the other hand, thioredoxins are small disulfide-containing proteins that have electron carrier activity and participate in redox homeostasis [23, 24]. Plant thioredoxins can act as antioxidants by activating ROS scavenging enzymes and regulate activity of key enzymes in the processes such as photosynthetic carbon fixation and anther dehiscence [24]. In addition, thioredoxins have been shown to activate their target enzymes by reducing Cys disulfides [25]. Considering that most of the proteins detected in this study are linked to stress events, it is conceivable that the role of thioredoxins in ROS activity may be modulated by citrullination. In addition, calcium exchanger 7 (AT5G17860) is also linked to stress responses. Calcium exchanger 7 is a $Ca^{2+}$:$Na^{+}$ antiporter that is highly expressed during senescence and in response to fungal infection [26].

**Stimulus-specific changes in the citrullination signature**. In order to detect possible stimulus-specific changes in citrullination signatures, we decided to test the effects of exposure of cells to low temperature and this for two reasons. Firstly, because the

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.12.294728; this version posted September 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

arabidopsis agmatine iminohydrolase is itself transcriptionally regulated by cold (see https://genevestigator.com/gv/ [27]) and secondly, temperature changes can be applied easily and reproducibly to suspension culture cells. In response to cold, four proteins showed differential expression and abundance in their citrullination signature over time (0, 1 h and 24 h) (Table 2). These proteins include the chromatin remodelling 34 (CHR34; AT2G21450) (see spectrum in Fig. 2) that shows an increase in citrullination after cold stress at both 1 h and 24 h. CHR34 is strongly transcriptionally up-regulated by several herbicides including primisulfuron-methyl and cloransulam-methyl and sulfometuron-methyl (see https://genevestigator.com/gv/ [27]). In addition, citrullination of three proteins, F-box family protein (AT5G27920), GDA1/CD39 nucleoside phosphatase (AT4G19180) and hypothetical exostosin family protein (AT2G29040), increased in response to cold treatment, and these proteins are expressed in the male gametophyte, pollen tube and flower. The F-box protein is annotated as having ubiquitin-protein transferase activity. F-box proteins contain a conserved motif (40 - 50 amino acids) at their N-terminal that functions as a site of protein-protein interactions. F-box protein components of the Skp I-Cullin-F-box complex are critical for regulating cell signaling as well as controlling the degradation of cellular proteins [28]. The F-box protein-encoding genes are influenced by light in addition to other abiotic stress conditions as shown by differential expression of at least 43 F-box protein-encoding genes in rice (*Oryza sativa*) seedlings [28]. The GDA1/CD39 nucleoside phosphatase (AT4G19180) is annotated as expressed in cultured plant cell and the flower, and in particular the gynoecium, pollen and sepal.

Furthermore, five proteins undergo *de novo* citrullination in response to cold stress. For the AAA-type ATPase (AT1G24290) citrullination is observed at both 1 h and 24 h after the onset of cold stress conditions (Table 2). The protein is annotated as playing a role in DNA replication and is most highly expressed in the carpel and the seed, and during senescence. It is noteworthy that the gene is transcriptionally up-regulated in response to cold, while it is repressed in response to ABA (see https://genevestigator.com/gv/ [27]). The YCF4 chloroplast protein (ATCG00250) and the hypothetical protein (AT1G16950) are citrullinated only after 1 h of cold treatment (Table 2). YCF4 is annotated as having a role in transcription, DNA-templated and elongation as well as “unfolded protein-binding” and the assembly of photosystem II. It may also act in translational elongation and triplet codon-amino acid adaptor activity [29]. The hypothetical protein (AT1G16950) is citrullinated only 24 h after cold treatment and is

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.12.294728; this version posted September 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

also expressed during the flowering stage and in particular in the stamen. The most pronounced transcriptional increase (>8-fold) occurs in double mutants of the transcription factors Auxin Response Factor 6 and 8 (ARF 6 and 8) that activate jasmonate biosynthesis. It is also up-regulated in severely gibberellic acid (GA) deficient mutants (*ga1-3*) that are dwarfed and male sterile (see https://genevestigator.com/gv/ [27]). The unknown DUF607 (AT2G23790) is citrullinated after only 1 h cold treatment (Table 2) and is again highly expressed in flowers and seedlings. This unknown protein is also referred to as the mitochondrial inner membrane calcium uniporter and it mediates calcium uptake into mitochondria [30]. Finally, the UDP-xylose synthase 4 (AT2G47650) is citrullinated only after prolonged exposure to cold (24 h) and cold stress causes marked (>2-fold) transcript accumulation. Overall, we noted that the majority of the proteins that are differentially citrullinated in response to low temperature stress are linked to general stress responses and importantly reproductive stages of plant development including key organs such as flower and pollen. This is indicative of the vital role citrullination may have in stress adaptation during this critical stage in the plant development.

Using Weblogo, a software used to predict specific signatures in a set of proteins, two main sequence specificity signatures or consensus were predicted from the citrullinated peptides both highlighting arginine residue flanked by four amino acids on the N-terminus and eight amino acid residues on the C-terminus (Supplementary Fig. 2).

**In search of candidate deiminases in *Arabidopsis thaliana*.** We first checked if the arabidopsis proteome contains any orthologues of annotated deiminases and noted that an arabidopsis agmatine iminohydrolase (AT5G08170) has also been annotated as porphyromonas-type peptidyl-arginine deiminase family protein containing agmatine deiminase (InterPro:IPR017754) and peptidyl-arginine deiminase of the Porphyromonas-type (IPR007466) sites. Crystal structures representing the Michaelis complex and the thiouronium reaction intermediate of *Pseudomonas aeruginosa* arginine deiminase [31] as well as an arabidopsis agmatine iminohydrolase [32] have been elucidated. The predicted highly conserved catalytic residues of the peptidylarginine deiminase (Q9RQJ2) from *Porphyromonas gingivalis* include Asp130, His236, Asp238 and Cys351, and an Asp187 that is absent in two family members of iminohydrolase families [33, 34]. The arabidopsis iminohydrolase in turn does contain these conserved residues, the Asp169, His224, Asp226 and Cys366 residues diagnostic for peptidyl-arginine deiminases and notably the aspartates that engage in ionic interactions with the guanidinium group in the cysteine.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.12.294728; this version posted September 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We also aligned key amino acid residues in the catalytic center of annotated bacterial agmatin deiminases and created a search motif: W-X-R-D-[TS]-G-X(100,140)-H-[VIL]-D (Figure 3a). This motif identified AT5G08170 as a sole in arabidopsis. Based on the crystal structure of AT5G08170 (PDB ID: 3H7K), the key amino acids appearing in the motif were located. Notably, these conserved amino acids constitute a clear cavity (highlighted in yellow) that is solvent exposed (Figure 3b) and that is also spatially able to accommodate an arginine residue. This cavity is thus assigned the catalytic center for arginine deiminase (Figure 3b and 3c).

Next, we evaluated if arginines in two selected citrullinated proteins (At2g21450 and At4g00830) identified from cell cultures of arabidopsis roots, can dock at the assigned catalytic center of AT5G08170. In order for docking simulations, models of both At2g21450 and At4g00830 citrullinated proteins were first generated by homolog modeling against the chain K of an ATPase domain of a chromatin remodeling factor (PDB ID: 6PWF) and the chain D of decaheme c-type cytochrome (PDB ID: 6R2K) respectively. Based on the 3D models, the citrullinated arginines (colored according to surface charges) were determined to be solvent exposed and protruding into space that is visibly large enough for the arginines to have unimpeded interactions with the catalytic center of AT5G08170 (Supplementary Figure 3 a and b). The individual citrullinated residues: R427 and R431 of AT2G21450, and R477 and R487 of At4g00830, were respectively docked at the catalytic center cavity of At5g08170, keeping all bonds in the R ligand non-rotatable so that their poses from the generated 3D models, are preserved. All citrullinated arginines docked at the catalytic cavity in a binding pose deemed suitable for catalysis i.e., with the amine rich region pointing into the cavity, as shown in the surface models (right panels) (Supplementary Figure 3 b and c).

### *In vitro* citrullination activity of AT5G08170

*In vitro* citrullination involving recombinant agmatine deiminase were performed using various substrates including fibrinogen (Sigma-Aldrich, St. Louis, MO) and peptides generated from LIF2 and chromatin remodeling 34 (Table 2). Fibrinogen has been shown to be a target of citrullination by agmatine deiminase from *Porphyromonas gingivalis* [35]. Interestingly, fibrinogen was also citrullinated by the recombinant agmatine deiminase generated from the arabidopsis sequence (Supplementary Table 1). We then tested whether the recombinant enzyme could citrullinate peptides from proteins identified by mass spectrometry as potential citrullination candidates. Using LIF2 and chromatin

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.12.294728; this version posted September 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

remodeling 34, we could show that agmatine deiminase can citrullinate the LIF2 peptide (Supplementary Table 2). However, citrullination of the chromatin remodeling 34 peptide had a low confidence score and we postulated that this could be attributed to the peptide failing to form the needed structural conformation. Since, it has been shown that the orthologs of agmatine deiminase can undergo auto-citrullination [36], we tested if this is the case for the arabidopsis agmatine deiminase. When we tested auto-citrullination in the presence or absence of $Ca^{2+}$, we observed that agmatine deiminase auto-citrullination does occur and is more pronounced in the presence of $Ca^{2+}$ (Table 3).

## Concluding remarks

We conclude from this, that citrullination events do occur in higher plants, particularly in arabidopsis and that the arabidopsis proteome does contain proteins that incur citrullination. In addition, low temperature stress causes stimulus-specific protein citrullination of a set of proteins many of which have annotated roles in stress responses. Furthermore, our data also suggest that citrullination may be part of the general regulation of pluripotency. Finally, we also predict that specific citrullination could occur in response to biotic and abiotic stimuli and that stress specific citrullination signatures and their down-stream effects influence plant growth, development and stress responses.

## Methods

**Arabidopsis cell suspension culture growth.** Cells derived from roots of *Arabidopsis thaliana* (ecotype Columbia-0) were grown in 100 mL of Gamborg's B-5 [37] basal salt mixture (Sigma-Aldrich, St Louis, MO, USA) with 2,4-dichlorophenoxyacetic acid (0.5 µg $mL^{-1}$) and kinetin (0.05 µg $mL^{-1}$) in 250 mL sterile flasks in a growth chamber (Innova® 43, New Brunswick Scientific Co., NJ, USA) at 120 rpm, with photosynthetic light set for 12 h light/12 h dark cycles at 23°C, and sub-cultured every seven days [38-40]. Cells were cold (4°C) treated and collected at 0, 1 hour and 24 h post-treatment. Media was drained off using Stericup® filter unit (Millipore, Billerica, MA), and cells were immediately frozen in liquid nitrogen and stored at -140°C until further use.

**Nuclear protein extraction.** Nucleus enrichment was performed using a nucleus enrichment kit for tissue (Pierce, Rockford, IL, USA). Briefly, protease inhibitors were added to the appropriate volume of nucleus enrichment reagents A and B. A total of 500

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.12.294728; this version posted September 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

mg of frozen callus cells was transferred to a 50 mL Falcon tube and 800 µL of reagent A and 800 µL of reagent B were added. The tissue was homogenized for 4 sec twice on ice using a PowerGen 125 grinder (Fisher Scientific, Rockford, IL, USA). The homogenate was transferred to a fresh 2 mL microcentrifuge tube, mixed by inverting the tube five times and centrifuged at 500 × g for 10 min at 4°C to collect the supernatant. A discontinuous density gradient was prepared by carefully overlaying three OptiPrep™ gradients (first layer consisted of 1.5 mL of 27.5% (v/v), the second of 1.5 mL of 23% (v/v) and the third of 1 mL of 20% (v/v)). Gradients were prepared from the OptiPrep media and gradient dilution buffer. The tissue extract was mixed with 60% (v/v) OptiPrep™ cell separation media and diluted to a final density gradient of 7.5% (v/v) with gradient dilution buffer and overlaid on top of the density gradients. Ultracentrifugation was performed at 40,000 × *g* for 90 min at 4°C. The top 2 mL, containing the nuclei, was pipetted out and precipitated with ice-cold acetone overnight at −20°C. Proteins were pelleted at 3,900 × *g* for 15 min at 4°C, washed three times in 80% (v/v) ice-cold acetone, solubilized in urea lysis buffer (7 M urea, 2 M thiourea) as described in [40]. Protein concentration was estimated by Bradford [41]. Approximately 200 µg of total nucleus protein extract was reduced with 5 mM dithiothreitol, alkylated, and used for protein digestion with sequencing-grade modified trypsin (Promega, Madison, WI, USA). Resulting peptides were purified using Sep-Pak Vac tC18 100 mg cartridge (Waters, Milford, MA, USA), as described previously [42] completely dried in a Speed Vac concentrator (Thermo Scientific, Bremen, Germany) and analysed by tandem mass spectrometry (LC-MS/MS).

**Immunoprecipitation, western blot and quantitative analysis of citrullinated proteins.** Immunoprecipitation was performed using two protocols. In the first experiment immunoprecipitation was done using immunoprecipitation kit Dynabeads® Protein A (Life technologies) and according to manufacturer's instruction. Here, Dynabeads® were resuspended by vortexing >30 sec and 50 µL of Dynabeads® were pipetted to a 1.5 mL microcentrifuge tube. The supernatant was removed using a magnetic Eppendorf tube rack prior to adding 2 µg of anti-citrulline antibody (ACAb, Abcam, Cambridge, UK) diluted in 200 µL phosphate buffered saline (PBS) containing Tween-20. The beads-Ab mix was incubated with rotation for 10 min at room temperature. The supernatant was removed by placing the tube on the magnetic rack and the beads-Ab complex was resuspended in 200

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.12.294728; this version posted September 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

µL PBS/Tween-20 buffer, resuspended by gentle pipetting and the supernatant was removed as stated previously. The protein sample (200 µg) was then added to the beads-Ab and gently pipetted to mix, incubated with rotation for 20 min at room temperature and supernatant decanted. Beads-Ab-protein mix was washed three times using 200 µL of washing buffer for each wash and resuspended in 100 µL washing buffer by gentle pipetting and transferred to a fresh tube. Supernatant was removed by means of a magnetic rack. A 20 µL of elution buffer was gently pipetted to resuspend the Dynabeads®-Ab-protein complex and incubated with rotation for 2 min at room temperature to dissociate the complex.

In the second experiment, 200 µg of protein sample and 2 µL of anti-citrulline antibody in TBS were mixed and incubated overnight at 4ºC with gentle rocking. Dynabeads® Protein A (Novex®, Life technologies) were then added and incubated for 3 h at 4ºC with gentle rocking. The protein-antibody-beads complex was washed five times with 1 × lysis buffer diluted in TBS, resuspended in 20 µL of elution buffer, mixed gently and incubated with end-over-end mixing for 2 min at room temperature to dissociate the complex. Supernatant was collected using a magnetic rack.

For both experiments, the protein solution collected was split into two-10 µL aliquots, one was used for in-solution digestion and tandem mass spectrometry and the remaining 10 µL was mixed with equal volume of 2 × reducing buffer for preparation prior to one-dimensional gel electrophoresis and western blot analysis. Proteins were transferred from polyacrylamide gel to polyvinylidine difluoride (PVDF) membranes as described by [43] using a Trans-blot® electrophoretic transfer cell (Bio-Rad, city, country). The PVDF membrane with transferred proteins was blocked overnight at 4°C in blocking solution (5% (w/v) bovine serum albumin in TBS), washed three times with TBS buffer containing 0.05% (v/v) Tween 20 (TBST) for 5 min and incubated with anti-citrulline (primary antibody) diluted in TBST for 1 h at 37°C. Membranes were washed three times in PBST for 5 min, incubated with the secondary antibody, Alexa Fluor® 488 goat anti-rabbit IgG (Molecular Probes, Eugene, OR, USA) diluted in TBS for 2 h at room temperature, thoroughly washed in TBST and imaged with a Typhoon™ 9410 scanner (GE Healthcare, city, country). Quantitative analyses and the relative-fold changes for protein bands were computed using ImageQuant-TL 7.0 software (GE Healthcare). Membranes were stained using the MemCode™reversible stain according to the manufacturer's instructions (Pierce,

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.12.294728; this version posted September 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

city, country). All visibly stained bands were cut, digested with trypsin and analyzed by LC-MS/MS.

**Tandem mass spectrometry and database search.** Peptides were analyzed by mass spectrometry as described previously [44]. Briefly, peptides were re-suspended in 5% (v/v) acetonitrile and 0.1% (v/v) formic acid prior to identification by LTQ Orbitrap Velos mass spectrometer (Thermo Electron, Bremen, Germany) coupled with an Ultimate 3000 UPLC (Dionex-Thermo-Scientific) for nano-LC-MS/MS analyses. A volume of 4 µL of peptide mixtures was injected onto a 5 mm long × 300 µm C18 PepMap 100 µ-precolumn, 5 µm, 100 Å (Thermo-Scientific) and then to a 150 mm × 75 µm Acclaim PepMap 100 NanoViper C18, 3 µm, 100 Å (Thermo-scientific) column. The MS scan range was *m/z* 350 to 1600 and spray voltage 1500 V. Top 10 precursor ions were selected in the MS scan by Orbitrap with a resolution r = 60,000 for fragmentation in the linear ion trap using collision induced dissociation. Normalized collision-induced dissociation was set at 35.0. Spectra were submitted to a local MASCOT (Matrix Science, London, UK) server and searched against arabidopsis in the TAIR (release 10), with a precursor mass tolerance of 10 ppm, a fragment ion mass tolerance of 0.5 Da, and strict trypsin specificity allowing up to one missed cleavage, carbamidomethyl modification on cysteine residues as fixed modification and oxidation of methionine residues, citrullination (also known as deamination) of arginine residues as variable modifications. Proteins were considered positively identified if the Mascot score was over the 95% confidence limit corresponding to a score ≤32 and peptide was considered citrullinated if the Mascot ion score was ≤30.

**Data analysis.** WebLogo3 (http://weblogo.threeplusone.com, 2015) was used to predict specificity of citrullination on the identified citrullination candidates. Gene ontology (GO) analysis toolkit and database for agricultural community, AgriGO (http://bioinfo.cau.edu.cn/agriGO/) was used for GO based analysis to detect enriched cellular components, biological processes and molecular functions from the set of citrullinated candidate proteins.

**Structural analysis.** Amino acid consensus motifs of catalytic centers were obtained as first employed in the search for plant nucleotide cyclases [45] and the motif searches of Swiss-Prot were done on-line (https://www.genome.jp/tools/motif/MOTIF2.html).The

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.12.294728; this version posted September 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

crystal structure of *Arabidopsis thaliana* agmatine deiminase At5g08170 was obtained from protein data bank (PDB ID: 3H7K) and the location of the arginine citrullination catalytic center was identified and visualized using UCSF Chimera (ver. 1.10.1) [46]. Two selected citrullinated proteins At2g21450 and At4g00830 were modeled against the chain K of a ATPase domain of a chromatin remodeling factor (PDB ID: 6PWF) and the chain D of decaheme c-type cytochrome (PDB ID: 6R2K) respectively using the Modeller (ver. 9.14) software [47]. The citrullinated arginines in the generated models were visualized and assessed for their ability to spatially fit the catalytic center of At5g08170. Docking simulations of the citrullinated arginines from the two peptides to the catalytic center of At5g08170 was performed using AutoDock Vina (ver. 1.1.2) [47, 48]. All bonds in the ariginines were set as non-rotatable to keep their poses similar to that in the peptides. The arginine docking poses were analyzed, and all images created by UCSF Chimera (ver. 1.10.1) [46]. Chimera was developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311).

**Validation**

**Cloning and Expression of At5g08170**

Total RNA from arabidopsis (Col-0) leaves was extracted using RNeasy Plant Mini Kit (Qiagen). Followed by synthesis of cDNA using SuperScript™ II Reverse Transcriptase system (Invitrogen). Full length AT5G08170 was then detected and amplified by polymerase chain reaction (PCR) using the forward primer (5'-ATGGAGGAGTCACGAGAATCG- 3') and reverse primer (5'-TCAGTGGCCATTTTCGGC-3'). PCR amplified gene was cloned into pCR8™/GW/TOPO® plasmid prior to clone it into pDEST™ 17 destination vector using GATWAY cloning technology (Invitrogen). Plasmid contained the full length gene was then transformed into *E. coli* BL-21 A1 cells (Invitrogen). The expression of recombinant protein was induced by 0.2% (w/v) L-Arabinose (Sigma Aldrich). The expressed protein was purified as His-tagged fusion protein using Ni-NTA agarose beads (Qiagen) and analyzed by SDS-PAGE.

**Enzyme activity analysis via mass spectrometry**

Two synthetic peptides from two of the identified citrullinated proteins, LIF2 and chromatin remodelling 34 were designed and purchased (Thermo Fisher Scientific). The

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.12.294728; this version posted September 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

selected peptides were designed to ensure that all peptides had at least one arginine or lysine to confirm the success of digestion. A tryptic digestion was performed following a citrullination assay that was performed in 50 mM CHES, pH 9.5, 5mM $CaCl_2$, 10 mM DTT, 50 µg of different substrates (including CHES buffer as control, the two synthetic peptides and fibrinogen) and 50 µg of purified AT5G01870 in a total volume of 200µl in microcentrifuge tube. Notably, autocitrullination of the agmatine deiminase in the presence and absence of calcium was also assessed by tandem mass spectrometry as described earlier.

ACKNOWLEDGMENTS

G.E. was the recipient of an Erasmus Mundus Action 2, Strand 2, lot 4 (Gulf Countries) award. A.W. was supported by the National Natural Science Foundation of China (31850410470) and the Zhejiang Provincial Natural Science Foundation of China (LQ19C130001).

AUTHOR CONTRIBUTIONS

G.E. conceived of the project, C.M. and C.G. planned the experiments, C:M and L.T. did the experiments and A.W. did the structural modeling. All authors contributed to the writing of the manuscript.

Competing financial interests: The authors declare no competing financial interests.

REFERENCES

1. Smith DG, Young EG. The combined amino acids in several species of marine algae. The Journal of Biological Chemistry. 1955; 217: 845-53.
2. Rogers GE, Simmonds DH. Content of citrulline and other amino-acids in a protein of hair follicles. Nature. 1958; 182: 186-7.
3. Rogers GE. Occurrence of citrulline in proteins. Nature. 1962; 194: 1149-51.
4. Tarcsa E, Marekov LN, Mei G, Melino G, Lee SC, Steinert PM. Protein unfolding by peptidylarginine deiminase. Substrate specificity and structural relationships of the natural substrates trichohyalin and filaggrin. The Journal of Biological Chemistry. 1996; 271: 30709-16.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.12.294728; this version posted September 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

5. van Venrooij WJ, Pruijn GJ. Citrullination: a small change for a protein with great consequences for rheumatoid arthritis. Arthritis Research. 2000; 2: 249-51.
6. Klareskog L, Lundberg K, Malmstrom V. Autoimmunity in rheumatoid arthritis: citrulline immunity and beyond. Advances in Immunology. 2013; 118: 129-58.
7. Christophorou MA, Castelo-Branco G, Halley-Stott RP, Oliveira CS, Loos R, Radzisheuskaya A, Mowen KA, Bertone P, Silva JC, Zernicka-Goetz M, Nielsen ML, Gurdon JB, Kouzarides T. Citrullination regulates pluripotency and histone H1 binding to chromatin. Nature. 2014; 507: 104-8.
8. Yuzhalin AE. Citrullination in Cancer. Cancer research. 2019; 79: 1274-1284.
9. Thompson PR, Fast W. Histone citrullination by protein arginine deiminase: is arginine methylation a green light or a roadblock? ACS Chemical Biology. 2006; 1: 433-41.
10. Arita K, Shimizu T, Hashimoto H, Hidaka Y, Yamada M, Sato M. Structural basis for histone N-terminal recognition by human peptidylarginine deiminase 4. Proceedings of the National Academy of Sciences of the United States of America. 2006; 103: 5291-6.
11. Arita K, Hashimoto H, Shimizu T, Nakashima K, Yamada M, Sato M. Structural basis for $Ca^{2+}$-induced activation of human PAD4. Nature Structural & Molecular Biology. 2004; 11: 777-83.
12. Gyorgy B, Toth E, Tarcsa E, Falus A, Buzas EI. Citrullination: a posttranslational modification in health and disease. The International Journal of Biochemistry & Cell Biology. 2006; 38: 1662-77.
13. Joshi V, Fernie AR. Citrulline metabolism in plants. Amino Acids. 2017; 49: 1543-1559.
14. Gudmann NS, Hansen NU, Jensen AC, Karsdal MA, Siebuhr AS. Biological relevance of citrullinations: diagnostic, prognostic and therapeutic options. Autoimmunity. 2015; 48: 73-9.
15. Masui K, Cavenee WK, Mischel PS. mTORC2 in the center of cancer metabolic reprogramming. Trends Endocrinololy and Metabolism 2014; 25: 364-73.
16. Wang H, Wang H. Multifaceted roles of FHY3 and FAR1 in light signaling and beyond. Trends Plant Science. 2015.
17. Linkies A, Muller K, Morris K, Tureckova V, Wenk M, Cadman CS, Corbineau F, Strnad M, Lynn JR, Finch-Savage WE, Leubner-Metzger G. Ethylene interacts with abscisic acid to regulate endosperm rupture during germination: a comparative approach using *Lepidium sativum* and *Arabidopsis thaliana*. Plant Cell. 2009; 21: 3803-22.
18. Xu J, Chua NH. Arabidopsis decapping 5 is required for mRNA decapping, P-body formation, and translational repression during postembryonic development. Plant Cell. 2009; 21: 3270-9.
19. Kim MH, Sonoda Y, Sasaki K, Kaminaka H, Imai R. Interactome analysis reveals versatile functions of Arabidopsis COLD SHOCK DOMAIN PROTEIN 3 in RNA processing within the nucleus and cytoplasm. Cell Stress Chaperones. 2013; 18: 517-25.
20. Molitor AM, Latrasse D, Zytnicki M, Andrey P, Houba-Herin N, Hachet M, Battail C, Del Prete S, Alberti A, Quesneville H, Gaudin V. The Arabidopsis hnRNP-Q Protein LIF2 and the PRC1 Subunit LHP1 Function in Concert to Regulate the Transcription of Stress-Responsive Genes. Plant Cell. 2016; 28: 2197-2211.
21. Latrasse D, Germann S, Houba-Herin N, Dubois E, Bui-Prodhomme D, Hourcade D, Juul-Jensen T, Le Roux C, Majira A, Simoncello N, Granier F, Taconnat L,

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.12.294728; this version posted September 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Renou JP, Gaudin V. Control of flowering and cell fate by LIF2, an RNA binding partner of the polycomb complex component LHP1. PloS one. 2011; 6: e16592.
22. Tewari R, Bailes E, Bunting KA, Coates JC. Armadillo-repeat protein functions: questions for little creatures. Trends Cell Biology. 2010; 20: 470-81.
23. Bertoni G. CBS domain proteins regulate redox homeostasis. Plant Cell. 2011; 23: 3562.
24. Meyer Y, Siala W, Bashandy T, Riondet C, Vignols F, Reichheld JP. Glutaredoxins and thioredoxins in plants. Biochimica et Biophysica Acta. 2008; 1783: 589-600.
25. Montrichard F, Alkhalfioui F, Yano H, Vensel WH, Hurkman WJ, Buchanan BB. Thioredoxin targets in plants: the first 30 years. J Proteomics. 2009; 72: 452-74.
26. Shigaki T, Hirschi KD. Diverse functions and molecular properties emerging for CAX cation/H+ exchangers in plants. Plant Biology (Stuttg). 2006; 8: 419-29.
27. Zimmermann P, Hirsch-Hoffmann M, Hennig L, Gruissem W. GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. Plant Physiology. 2004; 136: 2621-32.
28. Jain M, Nijhawan A, Arora R, Agarwal P, Ray S, Sharma P, Kapoor S, Tyagi AK, Khurana JP. F-box proteins in rice. Genome-wide analysis, classification, temporal and spatial gene expression during panicle and seed development, and regulation by light and abiotic stress. Plant Physiology. 2007; 143: 1467-83.
29. Heyndrickx KS, Vandepoele K. Systematic identification of functional plant modules through the integration of complementary data sources. Plant Physiology. 2012; 159: 884-901.
30. Stael S, Wurzinger B, Mair A, Mehlmer N, Vothknecht UC, Teige M. Plant organellar calcium signalling: an emerging field. Jornal of Experimental Botany. 2012; 63: 1525-42.
31. Galkin A, Lu X, Dunaway-Mariano D, Herzberg O. Crystal structures representing the Michaelis complex and the thiouronium reaction intermediate of Pseudomonas aeruginosa arginine deiminase. The Journal of Biological Chemistry. 2005; 280: 34080-7.
32. Levin EJ, Kondrashov DA, Wesenberg GE, Phillips GN, Jr. Ensemble refinement of protein crystal structures: validation and application. Structure. 2007; 15: 1040-52.
33. Shirai H, Blundell TL, Mizuguchi K. A novel superfamily of enzymes that catalyze the modification of guanidino groups. Trends Biochemical Science. 2001; 26: 465-8.
34. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH. CDD: NCBI's conserved domain database. Nucleic Acids Research. 2015; 43: D222-6.
35. Lundberg K, Wegner N, Yucel-Lindberg T, Venables PJ. Periodontitis in RA-the citrullinated enolase connection. Nature Reviews - Rheumatology. 2010; 6: 727-30.
36. Andrade F, Darrah E, Gucek M, Cole RN, Rosen A, Zhu X. Autocitrullination of human peptidyl arginine deiminase type 4 regulates protein citrullination during cell activation. Arthritis and Rheumatism. 2010; 62: 1630-40.
37. Gamborg OL, Eveleigh DE. Culture methods and detection of glucanases in suspension cultures of wheat and barley. Canadian Journal of Biochemistry. 1968; 46: 417-21.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.12.294728; this version posted September 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

38. Turek I, Marondedze C, Wheeler JI, Gehring C, Irving HR. Plant natriuretic peptides induce proteins diagnostic for an adaptive response to stress. Frontiers in Plant Science. 2014; 5: 661.
39. Piazza A, Zimaro T, Garavaglia BS, Ficarra FA, Thomas L, Marondedze C, Feil R, Lunn JE, Gehring C, Ottado J, Gottig N. The dual nature of trehalose in citrus canker disease: a virulence factor for Xanthomonas citri subsp. citri and a trigger for plant defence responses. Jornal of Experimental Botany. 2015.
40. Marondedze C, Wong A, Groen A, Serrano N, Jankovic B, Lilley K, Gehring C, Thomas L. Exploring the Arabidopsis proteome: influence of protein solubilization buffers on proteome coverage. International Journal of Molecular Science. 2015; 16: 857-70.
41. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical Biochemistry. 1976; 72: 248-54.
42. Groen A, Thomas L, Lilley K, Marondedze C. Identification and quantitation of signal molecule-dependent protein phosphorylation. Methods in Molecular Biology. 2013; 1016: 121-37.
43. Towbin H, Staehelin T, Gordon J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proceedings of the National Academy of Sciences of the United States of America. 1979; 76: 4350-4.
44. Marondedze C, Turek I, Parrott B, Thomas L, Jankovic B, Lilley KS, Gehring C. Structural and functional characteristics of cGMP-dependent methionine oxidation in *Arabidopsis thaliana* proteins. Cellular Communication Signaling. 2013; 11: 1.
45. Ludidi N, Gehring C. Identification of a novel protein with guanylyl cyclase activity in *Arabidopsis thaliana.* The Journal of Biological Chemistry. 2003; 278: 6490-4.
46. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera--a visualization system for exploratory research and analysis. Journal of Computational Chemistry. 2004; 25: 1605-12.
47. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. Journal of Molecular Biology. 1993; 234: 779-815.
48. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of Computational Chemistry. 2010; 31: 455-61.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.12.294728; this version posted September 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figures and legends:

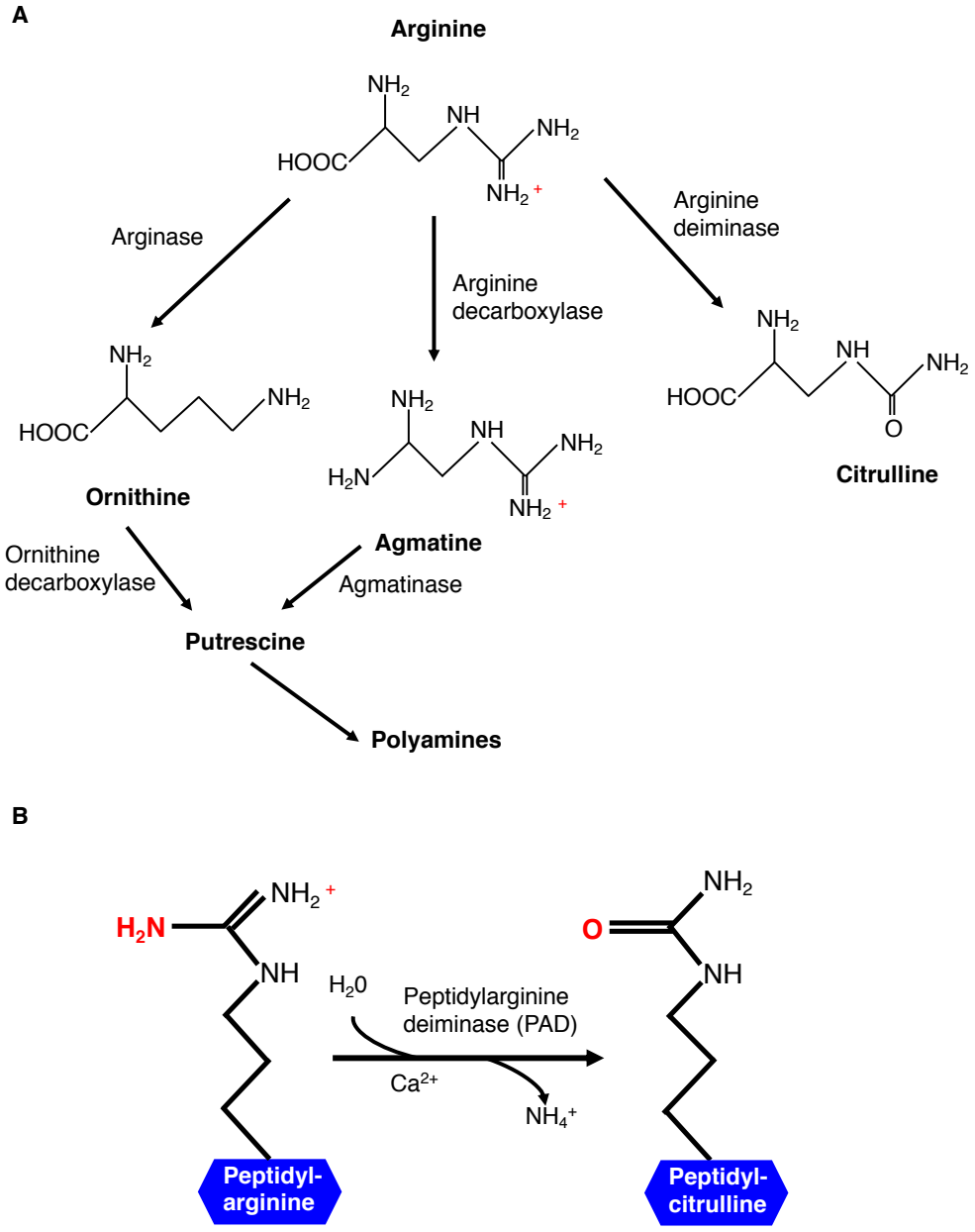



**Figure 1** (A) Overview of arginine modifications. (B) Enzymatic formation of a citrulline. Enzymatic hydrolysis of peptidyl-arginine to peptidyl-citrulline by peptidylarginine deiminase (PAD).

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.12.294728; this version posted September 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

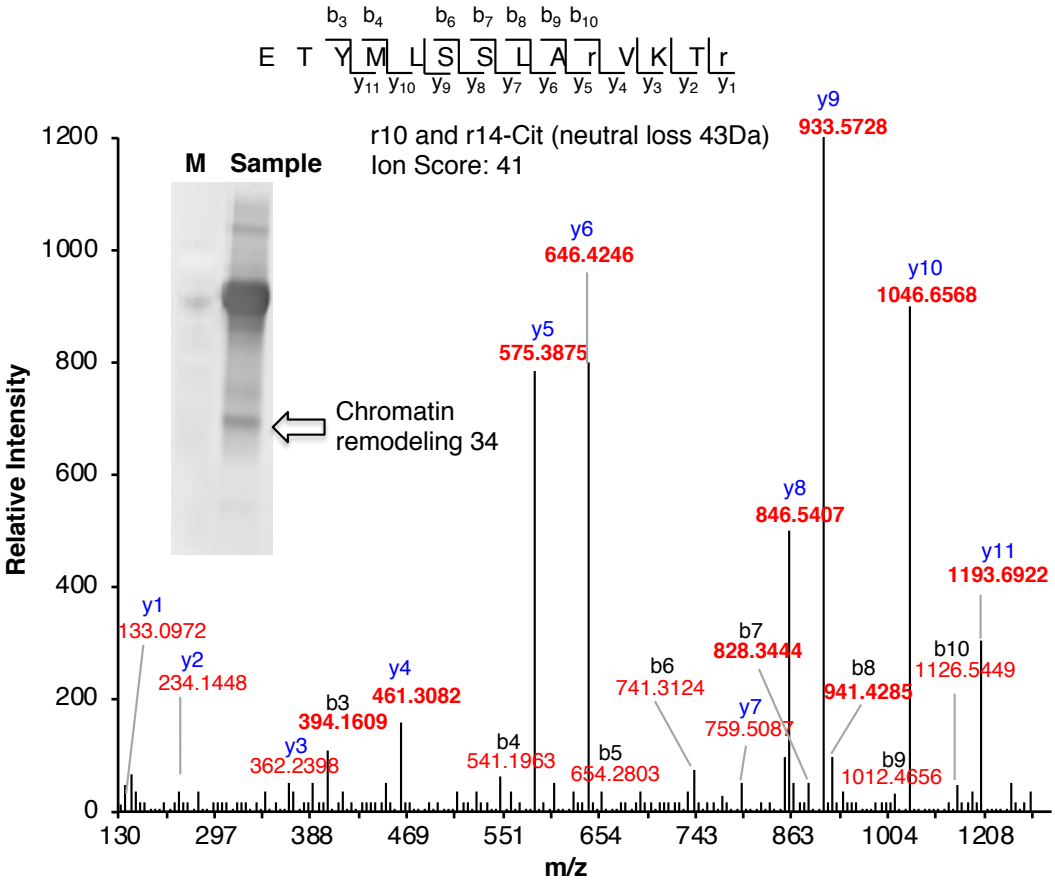



**Figure 2** An MS/MS spectra of a citrullinated chromatin remodeling 34 (CHR34). The MS/MS spectrum of chromatin remodeling 34 (AT2G21450) contains citrullinated residues on position 10 and 14. Indicated on the peptide sequence are the detected fragment ions, as indicated by the b- and y-ions and their relative intensities. Ion score of the peptide is 41. The insert shows a western blot of anti-citrulline immunoprecipitated proteins. The arrow shows the protein band identified as "Chromatin remodeling 34".

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.12.294728; this version posted September 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

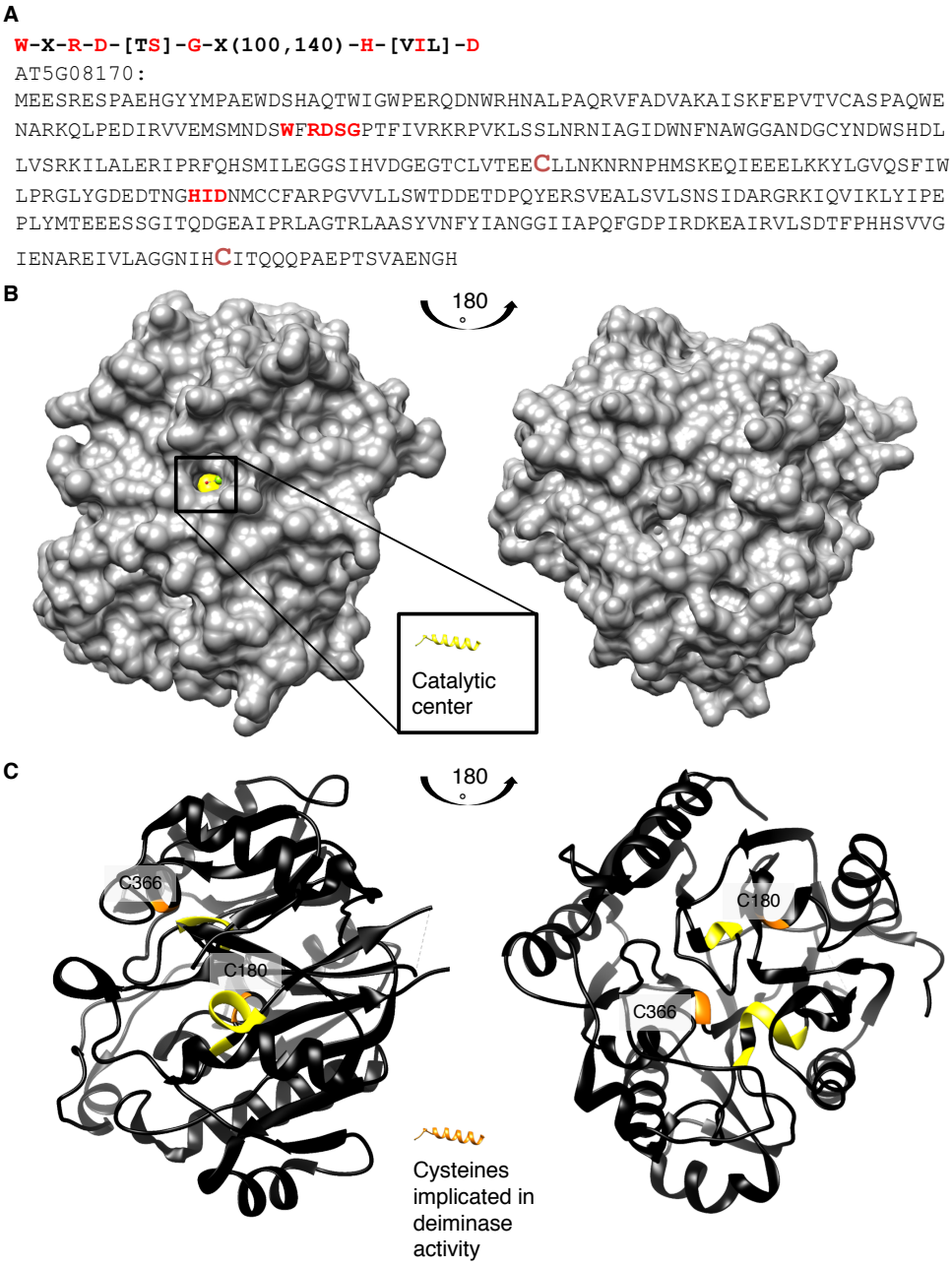
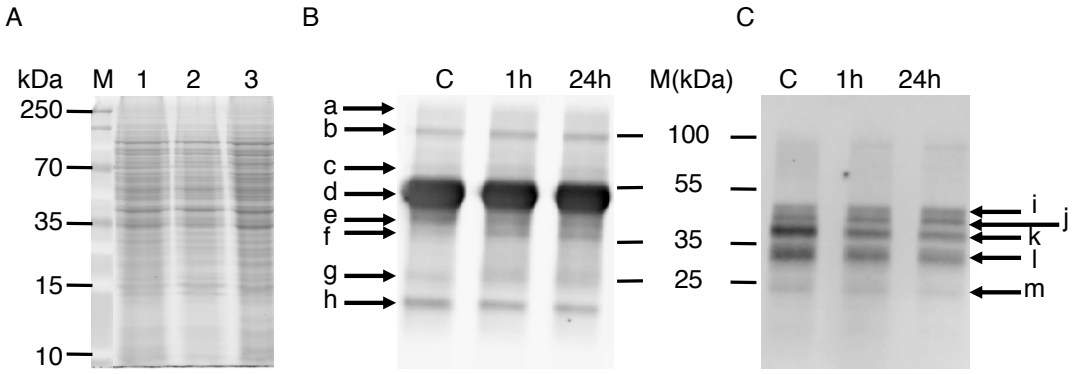



**Figure 3** (A) Amino acid sequence, (B) surface and (c) ribbon structure of At5g08170 (PDB ID: 3H7K). Amino acids appearing in the motif: W-X-R-D-[TS]-G-X(100,140)-H-[VIL]-D are highlighted in red in (A) the full-length At5g08170 amino acid sequence, and colored yellow in (B) the surface and (C) ribbon At5g08170 structures. The cysteines (C180 and C366) implicated in catalytic activity of agmatine deiminase are highlighted in green in (a) the amino acid sequence and orange in (B) the surface and (C) the ribbon structures. Amino acids R93 – G96 from the motif, occupy a clear cavity that can spatially accommodate an R residue (b), thus is assigned the catalytic center for arginine deiminase (highlighted in yellow). All analysis and images were created using UCSF Chimera (ver. 1.10.1) [46]. Chimera was developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311).

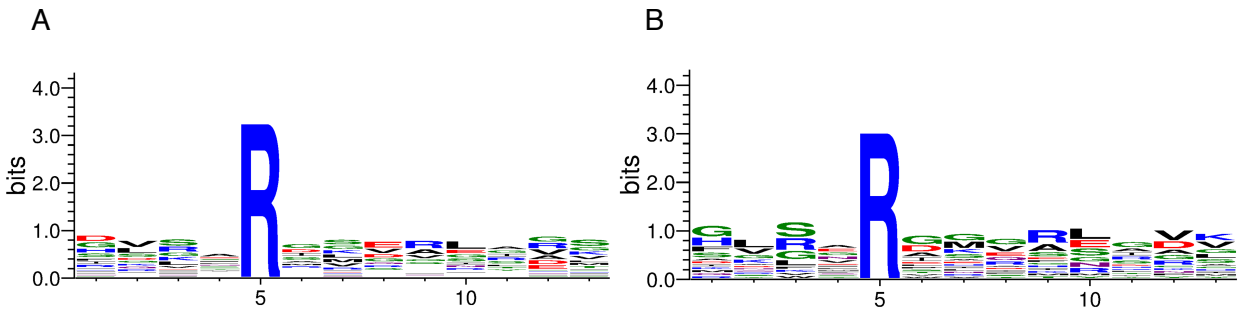
bioRxiv preprint doi: https://doi.org/10.1101/2020.09.12.294728; this version posted September 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Figures and Legends:



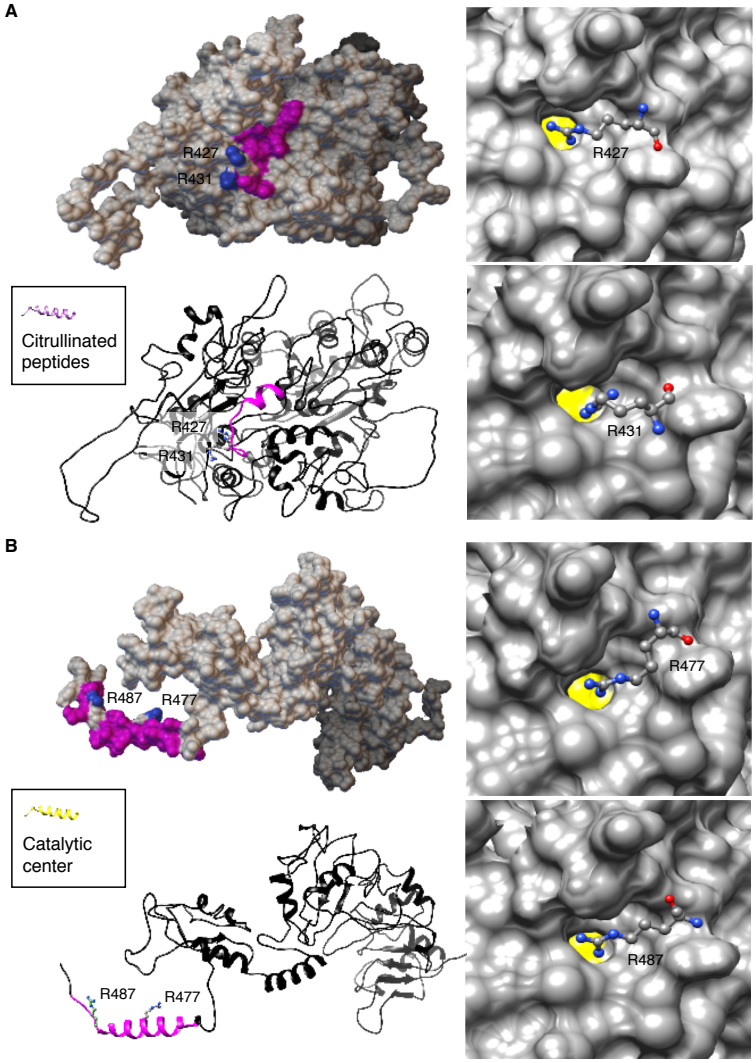

**Supplementary Figure 1** (A) One-dimensional gel electrophoresis scans of nuclear enriched proteins. Lane 1: protein ladder (M); lane 2: nuclear extract from control samples of arabidopsis cell suspension culture; lane 3: nuclear extract from 1 h-cold treated arabidopsis cell suspension culture sample; lane 4: nuclear extract from 24 h-cold treated arabidopsis cell suspension culture sample. (B) Western blot of nuclear extracts treated for 1 h or 24 h with cold and incubated overnight with the anti-citrulline antibody following incubation IgA protein coated beads. (C) Western blot with anti-citrulline antibody conjugated to IgA protein beads and then incubated with the nuclear extract for 10 min. Identified proteins in (B) and (c) are: I = AAA-type ATPase family protein, Target of rapamycin (TOR), Methyltransferase MT-A70, Far-red impaired responsive (FAR1), Calcium exchanger 7 (CAX7); II = Chromatin remodeling 34, Unknown protein, Armadillo (ARM) repeat superfamily protein; III = no positive identification (ND); IV = RNA-binding (RRM/RBD/RNP motifs) family protein; V = Unknown protein; VI = AAA-type ATPase family protein; VII = Chromatin remodeling 34, GDA1/CD39 nucleoside phosphatase; VIII = RING/FYVE/PHD zinc finger superfamily protein; IX = F-box family protein; X = Decapping 5, Thioredoxin family protein; XI = AAA-type ATPase family protein and RNA-binding (RRM/RBD/RNP motifs) family protein; XII = Peroxidase superfamily protein; XIII = Chromatin remodeling 34.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.12.294728; this version posted September 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.




**Supplementary Figure 2** (A) Sequence logo of citrullinated peptides. (b) Sequence logo of peptides with the citrullinated residue not located on the N- or C-terminus. Proteins are presented as WebLogo sequences using WebLogo 3.4 (http://weblogo.threeplusone.com/create.cgi). In the plot, the citrullinated residue was centralized and flanked with four amino acids on N- terminus and eight amino acids on C-terminus.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.12.294728; this version posted September 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.




**Supplementary Figure 3.** Computational assessment of the citrullinated arginines of two selected proteins (A) At2g21450 and (B) At4g00830. At2g21450 and At4g00830 were modeled against the chain K of a ATPase domain of a chromatin remodeling factor (PDB ID: 6PWF) and the chain D of decaheme c-type cytochrome (PDB ID: 6R2K) respectively using the Modeller (ver. 9.14) software [47]. The citrullinated arginines (colored according to surface charges) in the generated models were visualized and assessed for their ability to spatially fit the catalytic center of At5g08170. Citrullinated peptides were colored magenta and citrullinated arginine residues are all solvent exposed as shown in the ribbon and surface models of At2g21450 and At4g00830 respectively (left panels). Individual citrullinated residue: R427 and R431 of At2g21450, and R477 and R487 of At4g00830, were respectively docked at the catalytic center cavity of At5g08170, keeping all bonds in the R ligand non-rotatable so that their poses in the generated 3D models are retained. All citrullinated arginines docked at the catalytic cavity in a binding pose deemed suitable for catalysis i.e., with the amine rich region pointing into the cavity, as shown in the surface models (right panels). All docking simulations were performed using AutoDock Vina (ver. 1.1.2) [48]. Docking poses were analyzed, and all images created using UCSF Chimera (ver. 1.10.1) [46]. Chimera was developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311).

**Table 1: Citrullinated *Arabidopsis thaliana* protein identified by mass spectrometry**

| Accession # | Description | Mascot score | Mol. Mass | Peptide (Pep.) | Pep. Score (≥30) | Expect |
|---|---|---|---|---|---|---|
| **DNA-binding** | | | | | | |
| AT1G50030 | Target of rapamycin (TOR) | 59 | 281237 | R.ERAVEALR.A | 59 | 0.00031 |
| AT2G21450 | Chromatin remodeling 34 (CHR34) | 68 | 94037 | K.ETYmLSSLARVKTR.R | 41 | 0.0018 |
| AT4G12850 | Far-red impaired response (FAR1) | 37 | 21522 | K.VRDIVENVKK.L | 36 | 0.035 |
| AT1G24290# | AAA-type ATPase | 37 | 57702 | K.SMRGGDANAAIYWLAR.M | 35 | 0.013 |
| AT4G09980 | Methyltransferase MT-A70 protein | 32 | 86022 | R.ERTHGSSSDSSK.R | 32 | 0.024 |
| AT3G02890^ | RING/FYVE/PHD zinc finger protein | 33 | 110280 | R.RVGNRPMGRR.G | 33 | 0.022 |
| **RNA-binding** | | | | | | |
| AT1G26110 | Decapping 5 (DCP5) | 47 | 64331 | R.GYGGyGGRGGGGGGYGYGGRGQGR.G | 47 | 0.007 |
| AT4G00830 | LHP1-INTERACTING FACTOR 2 | 33 | 55273 | R.NNGSSGGSGRDNSHEHDGNRGGR.R | 33 | 0.027 |
| **ATP-/GTP-binding** | | | | | | |
| AT4G19180 | GDA1/CD39 nucleoside phosphatase | 37 | 82544 | R.FQRWSPMSTGVK.T | 37 | 0.0077 |
| AT1G29040 | 50S ribosomal protein L34 | 33 | 20111 | K.FLSSRSK.Q | 33 | 0.013 |
| **Protein-binding** | | | | | | |
| AT4G31890 | Armadillo (ARM) repeat protein | 35 | 56650 | R.VtLAMLGAIPPLVSmIDDsR.I | 35 | 0.0097 |
| ATCG00520 | YCF4 (unfolded protein binding) | 35 | 21564 | K.DIQsIRIEVK.E | 35 | 0.014 |
| AT4G10000 | Thioredoxin family protein | 30 | 37375 | R.ISGNGNWVRER.R | 30 | 0.018 |
| **Ubiquitination** | | | | | | |
| AT5G27920 | F-box family protein | 41 | 72636 | R.ARGLETLAR.M | 47 | 0.0029 |
| AT3G02890^ | RING/FYVE/PHD zinc finger protein | 33 | 110280 | R.RVGNRPMGRR.G | 33 | 0.022 |
| **Others** | | | | | | |
| AT5G58390 | Peroxidase superfamily protein | 67 | 35142 | K.AQGLSTRDMVALSGAHTIGR.A | 56 | 0.00074 |
| AT5G17860 | Calcium exchanger 7 | 40 | 62679 | K.RMSDQILR.S | 39 | 0.033 |
| AT1G16950 | Unknown transmembrane protein | 34 | 10044 | .MARPRIsIsmICLLILIVGFVLQssQAR.K | 34 | 0.0071 |
| AT4G27810 | Hypothetical protein | 34 | 22156 | R.RSLSVIR.R | 34 | 0.031 |
| AT2G23790 | Calcium uniporter (DUF607) | 32 | 38438 | K.LLRAAQIEIVK.S | 32 | 0.02 |
| AT2G47650 | UDP-xylose synthase 4 | 32 | 50085 | R.VVVtGGAGFVGSHLVDRLmAR.G | 32 | 0.0087 |

# -ATP-binding, ^-DNA and RNA binding

**Table 2. Low stress induced citrullination events in *Arabidopsis thaliana* proteins.** Citrulline candidate proteins indicating their accession numbers, description, Mascot scores, citrullinated peptide(s) with the citrullinated residue highlighted in red, peptide ion scores, expect and fold changes (FC) following cold treatment.

| Accession number | Description | Mascot score | Mass | Peptide (Pep.) | Pep. Score (≥30) | Expect | FC C vs 1h | FC C vs 24h |
|---|---|---|---|---|---|---|---|---|
| **Citrullinated peptides in the control** | | | | | | | | |
| AT1G50030[#] | Target of rapamycin (TOR) | 59 | 281237 | R.ERAVEALR.A | 59 | 0.00031 | ns | ns |
| AT1G26110[#] | Decapping 5 (DCP5) | 47 | 64331 | R.GYGGyGGRGGGGGGYGYGGRGQGR.G | 47 | 0.007 | ns | ns |
| AT5G17860[#] | Calcium exchanger 7 (CAX7) | 40 | 62679 | K.RMSDQILR.S | 39 | 0.033 | ns | ns |
| AT4G12850[#] | Far-red impaired responsive (FAR1) | 37 | 21522 | K.VRDIVENVKK.L | 36 | 0.035 | ns | ns |
| AT4G31890[#] | Armadillo (ARM) repeat protein | 35 | 56650 | R.VtLAMLGAIPPLVSmIDDsR.I | 35 | 0.0097 | ns | ns |
| AT4G27810[#] | Hypothetical DUF688 protein | 34 | 22156 | R.RSLSVIR.R | 34 | 0.031 | ns | ns |
| AT3G02890[#] | RING/FYVE/PHD zinc finger | 33 | 110280 | R.RVGNRPMGRR.G | 33 | 0.022 | ns | ns |
| AT4G00830[#] | RNA-binding family protein | 33 | 55273 | R.NNGSSGGSGRDNSHEHDGNRGGR.R | 33 | 0.027 | ns | ns |
| AT4G09980[#] | Methyltransferase MT-A70 | 32 | 86022 | R.ERTHGSSSDSSK.R | 32 | 0.024 | ns | ns |
| AT4G10000[#] | Thioredoxin family protein | 30 | 37375 | R.ISGNGNWVRER.R | 30 | 0.018 | ns | ns |
| **Peptides differentially citrullinated in response to cold treatment** | | | | | | | | |
| AT2G21450[*] | Chromatin remodeling 34 (CHR34) | 68 | 94037 | K.ETYmLSSLARVKTR.R | 41 | 0.0018 | 1.90 | 1.84 |
| AT5G27920[#,] | F-box family protein | 41 | 72636 | R.ARGLETLAR.M | 47 | 0.0029 | 1.86 | 1.53 |
| AT4G19180[*] | GDA1/CD39 nucleoside phosphatase | 37 | 82544 | R.FQRWSPMSTGVK.T | 37 | 0.0077 | 1.9 | 1.8 |
| AT1G29040[#] | Hypothetical exostosin family prot. | 33 | 20111 | K.FLSSRSK.Q | 33 | 0.013 | ns | 0.67 |
| ***De novo* cold-induced citrullinated peptides** | | | | | | | | |
| AT1G24290[#] | AAA-type ATPase family protein | 37 | 57702 | K.SMRGGDANAAIYWLAR.M | 35 | 0.013 | at 1h | at 24h |
| ATCG00520 | YCF4 (unfolded protein binding) | 35 | 21564 | K.DIQsIRIEVK.E | 35 | 0.014 | at 1h | nd |
| AT1G16950 | Hypothetical protein | 34 | 10044 | .MARPRIsIsmICLLILIVGFVLQssQAR.K | 34 | 0.0071 | nd | at 24h |
| AT2G23790 | Unknown DUF607 protein | 32 | 38438 | K.LLRAAQIEIVK.S | 32 | 0.02 | at 1h | nd |
| AT2G47650 | UDP-xylose synthase 4 | 32 | 50085 | R.VVVtGGAGFVGSHLVDRLmAR.G | 32 | 0.0087 | nd | at 24h |

* - Identified by gel-based and gel-free techniques, # - identified by gel-based technique, ns - not significant at $p \leq 0.05$ and FC = $0.68 < x > 1.5$; nd - not identified.

**Table 3: Autocitrullination of agmatine deiminase in the presence or absence of calcium**

| Agmatine peptide | Citrullinated site(s) | Stoichiometry (%; A,B) |
|---|---|---|
| ESPAEHGYYMPAEWDSHAQTWIGWPErQDNWr | 32, 37 | 13, 25 |
| ESPAEHGYYMPAEWDSHAQTWIGWPErQDNWR | 32 | 30, 37 |
| FEPVTVCASPAQWENArK | 73 | 11, 17 |
| GLYGDEDTNGHIDNMCCFArPGVVLLSWTDDETDPQYEr | 233, 252 | 14, 18 |
| GLYGDEDTNGHIDNMCCFArPGVVLLSWTDDETDPQYER | 233 | 49, 35 |
| LAASYVNFYIANGGIIAPQFGDPIrDK | 331 | 20, 43 |
| LYIPEPLYMTEEESSGITQDGEAIPrLAGTr | 301, 306 | 11, 33 |
| LYIPEPLYMTEEESSGITQDGEAIPrLAGTR | 301 | 33 |
| NIAGIDWNFNAWGGANDGCYNDWSHDLLVSrK | 144 | 10, 18 |
| QDNWrHNALPAQR | 37 | 14, 19 |
| QLPEDIrVVEMSMNDSWFR | 81 | 11, 2 |
| QLPEDIrVVEmSMNDSWFR | 81 | 2 |
| VVEMSMNDSWFrDSGPTFIVR | 93 | 44, 50 |

A= Agmatine - CaCl, B= Agmatine + CaCl

Supplementary Table 1:

**Commonly citrullinated peptides following plant Agmatine and animal rmPAD2 on fibrinogen**

| Protein Acc. | Protein Description | Peptide Sequence | Identified citrullinated p R219) |
|---|---|---|---|
| gi\|237823914 | chain A, crystal structure of human fibrinogen | ADSGEGDFLAEGGGVrGPR | X |
| gi\|182439 | fibrinogen gamma chain | ANQQFLVYCEIDGSGNGWTVFQKr | X |
| gi\|237823915 | chain B, crystal structure of human fibrinogen | EEAPSLrPAPPPISGGGYR | X |
| gi\|4503689 | fibrinogen alpha chain isoform alpha-E preproprotein | GDFSSANNrDNTYNR | X |
| gi\|237823915 | chain B, crystal structure of human fibrinogen | GGETSEMYLIQPDSSVKPYrVYCDMNTENGGWTVIQNr | X(C-Term.) |
| gi\|4503689 | fibrinogen alpha chain isoform alpha-E preproprotein | GGSTSYGTGSETESPrNPSSAGSWNSGSSGPGSTGNr | X |
| gi\|4503689 | fibrinogen alpha chain isoform alpha-E preproprotein | GGSTSYGTGSETESPrNPSSAGSWNSGSSGPGSTGNR | X |
| gi\|4503689 | fibrinogen alpha chain isoform alpha-E preproprotein | HrHPDEAAFFDTASTGK | X |
| gi\|237823916 | chain C, crystal structure of human fibrinogen | IHLISTQSAIPYALrVELEDWNGR | |
| gi\|182439 | fibrinogen gamma chain | IHLISTQSAIPYALrVELEDWNGR | |
| gi\|182439 | fibrinogen gamma chain | IIPFNrLTIGEGQQHHLGGAK | X |
| gi\|4503689 | fibrinogen alpha chain isoform alpha-E preproprotein | QFTSSTSYNrGDSTFESK | X |
| gi\|4503689 | fibrinogen alpha chain isoform alpha-E preproprotein | TFPGFFSPMLGEFVSETESrGSESGIFTNTK | X |
| gi\|182439 | fibrinogen gamma chain | VELEDWNGrTSTADYAMFK | X |
| gi\|237823915 | chain B, crystal structure of human fibrinogen | VYCDMNTENGGWTVIQNrQDGSVDFGR | X |
| gi\|237823915 | chain B, crystal structure of human fibrinogen | VYCDMNTENGGWTVIQNrQDGSVDFGr | |

Supplementary Table 2:

**Citrullinated peptides ≥ 30 ion scores**

**LHP1-INTERACTING FACTOR 2 RNA binding protein peptides that are citrullinated**

| prot_acc | prot_desc | prot_scor | prot_mas | prot_mat | prot_mat | prot_sequ | prot_sequ | pep_quer | pep_rank | pep_isbol | pep_isuni | pep_exp_ | pep_exp_ | pep_exp_ | pep_calc_ | pep_delta | pep_miss | pep_score | pep_expe | pep_res_b | pep_seq | pep_res_af | pep_va |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AT4G0083 | RNA-binding (RRM/RBD | 50 | 55273 | 25 | 6 | 4 | 2 | 16632 | 1 | 1 | 1 | 661.2514 | 2640.977 | 4 | 2640.944 | 0.0328 | 2 | 31.62 | 0.032 | R | NDrNNGSSGGSGRDNSHEHDGNR | G | Citrul ( |
| AT4G0083 | RNA-binding (RRM/RBD | 50 | 55273 | 25 | 6 | 4 | 2 | 16776 | 1 | 1 | 1 | 671.9999 | 2683.971 | 4 | 2683.95 | 0.0212 | 2 | 30.78 | 0.023 | R | NDrNNGSSGGSGrDNSHEHDGNR | G | 2 Citru |
| | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | R.NNGSSGGSGrDNSHEHDGNrGGR.R | original peptide afte | |